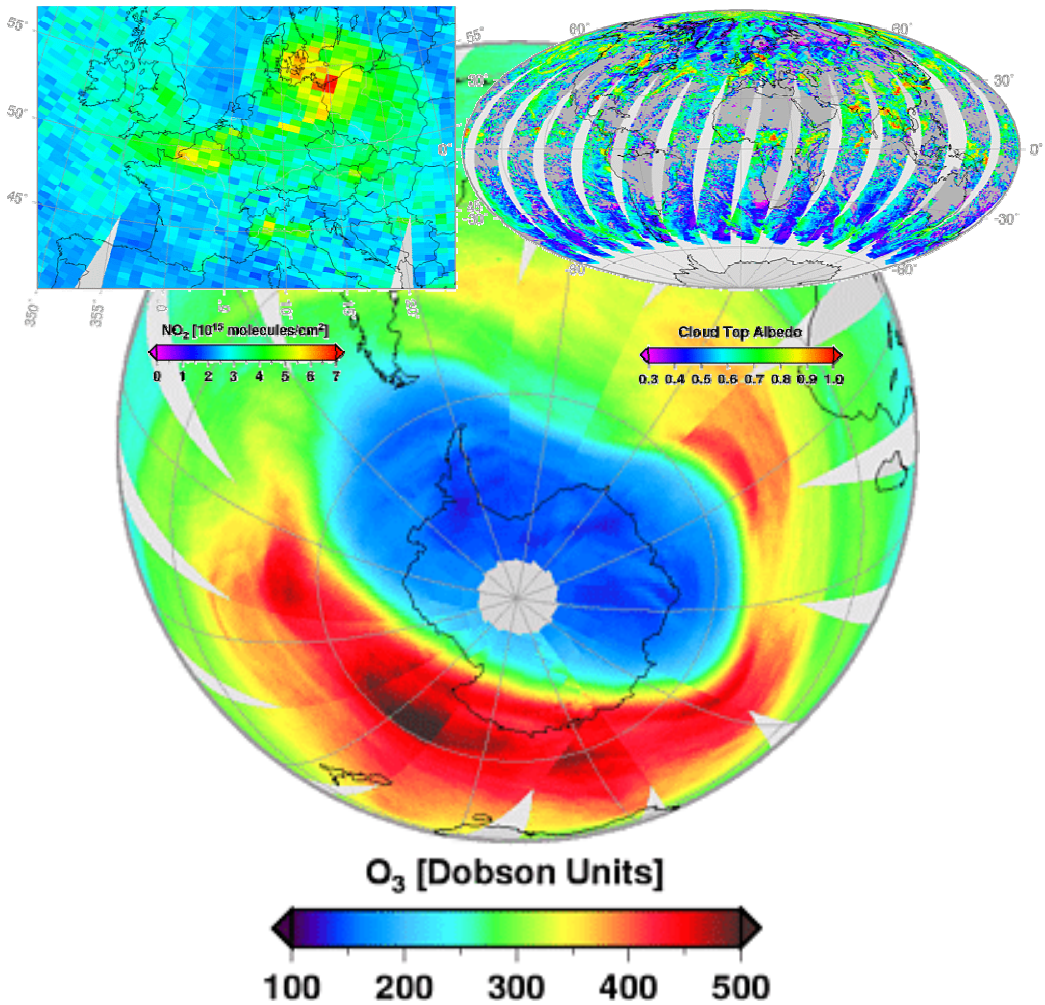
# Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, Minor Trace Gases, and Cloud Properties

## (GDP 4.2 for O3M-SAF OTO and NTO)





Doc.No.: DLR/GOME-2/ATBD/01
Iss./Rev.: 2/A
Date: 28 January 2009



Deutsches Zentrum für Luft und Raumfahrt e.V. – DLR
Oberpfaffenhofen, Germany

**Signatures**

| *Action: Name* | *Affiliation* | *Function* | *Date* | *Signature* |
|---|---|---|---|---|
| prepared by:<br>P. Valks<br>D. Loyola<br>N. Hao<br>M. Rix | <br>DLR-MF<br>DLR-MF<br>DLR-MF<br>DLR-MF | <br>GOME Project Scientist<br>GOME Project Manager<br>GOME Project Scientist<br>GOME Project Scientist | 28 Jan. 2009 | |
| released by:<br>D. Loyola | DLR-MF | GOME Project Manager | 28 Jan. 2009 | |

**Distribution List**

| *Function* | *Organization* |
|---|---|
| UPAS Team | DLR-MF, DLR-DFD |
| GOME Team | ESA, BIRA, RTS, AUTH, various |
| O3M-SAF Team | EUMETSAT, FMI, KNMI, DMI, various |

**Document Change Log**

| *Issue* | *Rev.* | *Date* | *Section* | *Description of Change* |
|---|---|---|---|---|
| 1 | A | 4 May 2007 | All | Completely new |
| 1 | B | 12 Oct. 2007 | All | Revised following ORR-A |
| 1 | C | 14 April 2008 | 3, 4 | Update tropospheric $NO_2$ and BrO algorithms |
| 2 | A | 17 Nov. 2008<br>28 Jan. 2009 | All | Revisions following ORR-A3<br>Harmonize with PUM |

1 INTRODUCTION ..... 5

1.1 Purpose and scope ..... 5

1.2 GOME-2 instrument ..... 5

1.3 Overview of the GDP 4.2 algorithm ..... 5

1.4 Abbreviations and acronyms ..... 7

2 THE OZONE COLUMN ALGORITHM ..... 9

2.1 Introduction ..... 9

2.2 DOAS slant column fitting ..... 9

2.3 Air Mass Factor and vertical column computations ..... 10

2.3.2 The ozone profile-column map ..... 13

2.3.3 Intra-Cloud correction ..... 13

2.3.4 Radiative Transfer Model for the AMF calculation ..... 14

2.3.5 Atmospheric and surface setups for the RT model ..... 15

2.4 Molecular Ring correction ..... 16

2.4.1 DOAS implementation ..... 17

2.5 Error budgets and sensitivity studies ..... 19

2.5.1 Error budgets for the total ozone algorithm ..... 19

2.5.2 Sensitivity issues for GDP 4.2 algorithm ..... 20

3 THE NO2 COLUMN ALGORITHM ..... 24

3.1 DOAS slant column fitting ..... 24

3.2 AMF and VCD determination ..... 24

3.3 Tropospheric NO2 column calculation for polluted conditions ..... 25

3.4 Error budget for the total and tropospheric NO2 column ..... 26

4 THE BRO COLUMN ALGORITHM .....27

5 THE SO2 COLUMN ALGORITHM .....28

5.1 DOAS slant column fitting.....28

5.2 $SO_2$ background correction .....28

5.3 AMF and VCD determination.....28

6 OTHER TRACE GAS COLUMN PRODUCTS .....30

6.1 Formaldehyde column algorithm.....30

6.2 OClO column algorithm .....30

7 CLOUD ALGORITHMS .....31

7.1 OCRA cloud fraction algorithm .....31

7.2 ROCINN cloud-top height and albedo algorithm .....31

7.3 Cloud-top pressure and cloud optical thickness calculation.....32

8 EXPECTED ACCURACY .....33

REFERENCES .....34

# 1 INTRODUCTION

## 1.1 Purpose and scope

This document describes the GOME-2 Data Processor Version 4.2 (GDP 4.2), the operational algorithm for the retrieval of total columns of trace gases from the GOME-2/MetOp instrument, as part of the O3M-SAF. GDP 4.2 is based on the DOAS-style algorithm being used operational for GOME/ERS-2 and its corresponding ATBD [Spurr et al., 2004].

This document contains descriptions of the different trace gases and cloud retrieval algorithms, including a last chapter that summarizes the expected accuracies. The product format and dissemination information are given in the corresponding product user manual [Loyola et al., 2009]. Preliminary validation results of the GOME-2 total ozone, $NO_2$, BrO and $SO_2$ column products with ground-based measurements are described in the O3M-SAF Validation Reports [Balis et al., 2007, 2008; Lambert et al., 2007, 2008; Van Roozendael et al., 2008; Van Geffen et al., 2008].

In this document, the terms GOME/ERS-2 and GOME-2/MetOp are used to reference the specific instruments. The general term GOME applies to both sensors.

## 1.2 GOME-2 instrument

On 30 January 1998, the ESA Earth Observation Programme Board gave its final go-ahead for the MetOp Programme. The instruments on the MetOp satellites will produce high-resolution images of the Earth's surface, vertical temperature and humidity profiles, and temperatures of the land and ocean surface on a global basis. In addition, there will be instruments for monitoring trace gases and wind flow over the oceans. This instrument payload will be of significant value to meteorologists and other scientists, particularly to those studying the global climate.

Given the need for global-scale routine monitoring of the abundance and distribution of ozone and associated trace gas species, a proposal was put forward for the inclusion of GOME-2 on the MetOp satellites. MetOp-A was launched on 19 October 2006 as part of the Initial Joint Polar System (IJPS) in co-operation with NOAA in the USA.

The GOME-2 field of view of each step may be varied in size from 5 km x 40 km to 80 km x 40 km. The mode with the largest footprint (twenty four steps with a total coverage of 1920 km x 40 km) provides daily near global coverage at the equator.

Based on the successful work with the GOME data processors, the German Aerospace Centre (DLR) plays a major role in the design, implementation and operation of the GOME-2 ground segment for total column products. DLR is a partner in the Satellite Application Facility on Ozone and Atmospheric Chemistry Monitoring (O3M-SAF), which is part of the EUMETSAT Polar System (EPS) ground segment, and is responsible in this project for the generation of total column amounts of the various trace gases and cloud properties which may be retrieved from GOME-2 level 1b products.

## 1.3 Overview of the GDP 4.2 algorithm

The GOME Data Processor (GDP) operational algorithm is the baseline algorithm for the trace gas column retrievals from GOME-2/MetOp. The GDP 4.2 is a classical DOAS-AMF fitting algorithm for the generation of total column amounts of ozone, $NO_2$, BrO, $SO_2$, $H_2O$, HCHO, and OClO [Van Roozendael et al., 2006]. The algorithm has two major steps: a DOAS least-squares fitting for the

trace gas slant column, followed by the computation of a suitable Air Mass Factor to make the conversion to the vertical column density. Figure 1 is a schematic flowchart for the GDP 4.2 trace-gas column algorithm. In a pre-processing step, cloud information (fractional cover, cloud-top height and cloud albedo) is derived before the above two major algorithm components are executed. In GDP 4.2 cloud algorithm products are computed directly by calls to the OCRA/ROCINN algorithms [Loyola, 2004, 2007]. Table 1 lists the wavelength regions used for the retrieval of the trace gas column and cloud products.
In the next chapter, the ozone column algorithm is described. A general description of the DOAS slant column algorithm is given, and the vertical ozone column calculation using the Air Mass Factor is described. The ozone column error budgets and sensitivity studies are addressed as well. In the following chapters, specific retrieval algorithm aspects for the $NO_2$, BrO, $SO_2$ and other trace gas column products are described. Finally, a description of the GOME cloud algorithms OCRA and ROCINN is provided.

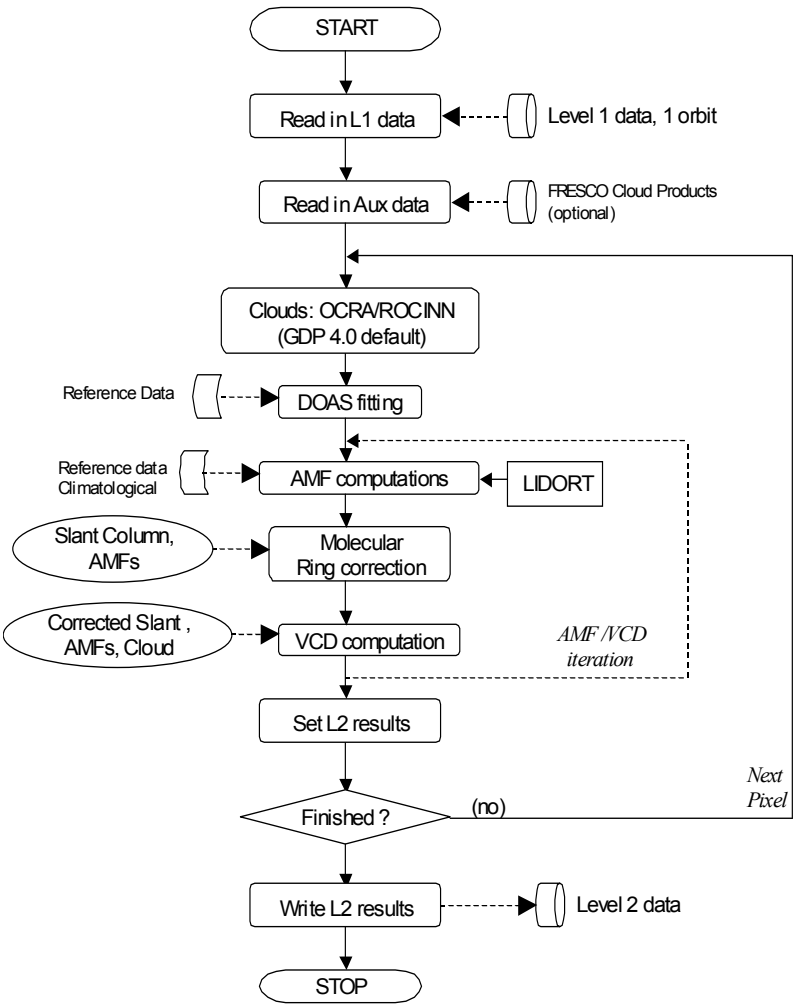



**Figure 1** Flow diagram of the GDP 4.2 algorithm for GOME-2/MetOp (from Van Roozendael et al. [2006]).

**Table 1** GOME-2/MetOp trace gas column and cloud products generated by the O3M-SAF, with the corresponding wavelength regions used for the retrieval.

| Product | Wavelength region |
|---|---|
| Ozone column | 325.0-335.0 nm |
| $NO_2$ column | 425.0-450.0 nm |
| BrO column | 336.0-351.5 nm |
| $SO_2$ column | 315.0-326.0 nm |
| $H_2O$ column | 611.0-700 nm |
| HCHO column | 337.5-359.0 nm |
| OClO column | 365.0-389.0 nm |
| cloud fraction | 300-800 nm (PMD-p) |
| cloud-top height (pressure) & albedo (optical thickness) | 758-771 nm |

## 1.4 Abbreviations and acronyms

A list of abbreviations and acronyms which are used throughout this document is given below:

| | |
|---|---|
| AMF | Air Mass Factor |
| BIRA-IASB | Belgian Institute for Space Aeronomy |
| DLR | Deutsches Zentrum für Luft- und Raumfahrt e.V. (German Aerospace Centre) |
| DOAS | Differential Optical Absorption Spectroscopy |
| DU | Dobson Unit |
| EPS | EUMETSAT Polar System |
| ERS-2 | European Remote Sensing Satellite-2 |
| ESA | European Space Agency |
| ESC | Effective Slant Column |
| EUMETSAT | European Organisation for the Exploitation of Meteorological Satellites |
| GDOAS | GODFIT-DOAS |
| GDP | GOME Data Processor |
| GOME | Global Ozone Monitoring Experiment |
| IMF | Remote Sensing Technology Institute |
| LER | Lambertian Equivalent Reflectivity |
| LIDORT | Linearized Discrete Ordinate Radiative Transfer Forward Modeling |
| MetOp | Operational Meteorological Satellite |
| NRT | Near Real Time |
| NTO | Identifier used for near-real-time total column trace gas products |
| O3M-SAF | SAF on Ozone and Atmospheric Chemistry Monitoring |
| OCRA | Optical Cloud Recognition Algorithm |
| OL | Off-line |
| OTO | Identifier used for offline total column trace gas products |
| P-S | Pseudo-Spherical |
| PMD | Polarisation Measurement Device |
| RMS | Root Mean Square |
| ROCINN | Retrieval of Cloud Information using Neural Networks |

| | |
|---|---|
| RRS | Rotational Raman Scattering |
| RT | Radiative Transfer |
| SAF | Satellite Application Facility |
| SZA | Solar Zenith Angle |
| TOA | Top of Atmosphere |
| TOMS | Total Ozone Mapping Spectrometer |
| UMARF | Unified Meteorological Archiving and Retrieval Facility |
| UV | Ultra Violet |
| UPAS | Universal Processor for UV/VIS Atmospheric Spectrometers |
| UTC | Universal Time Coordinate |
| VCD | Vertical Column Density |
| VIS | Visible |

# 2 THE OZONE COLUMN ALGORITHM

## 2.1 Introduction

The first major algorithm component is the DOAS fitting. This is a straightforward least-squares inversion to deliver the effective slant column of total ozone, plus a number of auxiliary fitted parameters and error diagnostics. The latter include an effective temperature for the ozone absorption, a slant column for $NO_2$ (regarded as an interfering species in the ozone UV window), wavelength registration parameters for re-sampling the earthshine spectrum, scaling factors for interference due to undersampling and Ring effects, and low-pass filter closure parameters.

The second major component is the iterative AMF/VCD (Air Mass Factor, Vertical Column Density) computation to generate the final vertical column. An initial guess is made for the VCD. At each iteration step, ozone air mass factors (to ground level and to cloud-top) are computed for the current guess of the vertical column. This radiative transfer calculation uses a column-classified ozone profile climatology. Then the DOAS slant column is adjusted using the molecular Ring correction (to compensate for interference effects in ozone absorption features due to inelastic rotational Raman scattering). This adjusted slant column is then used in conjunction with pre-processed cloud information and the AMF values to update the VCD guess. The pixel processing is completed with an assignation of the Level 2 output (total column, errors and retrieval diagnostics, and auxiliary output such as surface pressure and selected Level 1 geolocation information) for one orbit of data.

The ozone column algorithm components are described in the next sections, starting with the DOAS fitting (section 2.2), moving on to the iterative AMF/VCD computation (section 2.3), and the molecular Ring correction (section 2.4). In section 2.5 we present an error budget for the total ozone algorithm, and discuss a number of sensitivity tests.

## 2.2 DOAS slant column fitting

In DOAS fitting, the basic model is the Beer-Lambert extinction law for trace gas absorbers [Spurr et al., 2005]. An external polynomial closure term accounts for broadband effects: molecular scattering, aerosol scattering and absorption and reflection from the Earth's surface. We also include additive spectra for Ring effect interference. The fitting model is then:

$$Y(\lambda) \equiv \ln\left[\frac{I_\lambda(\Theta)}{I_\lambda^0(\Theta)}\right] = -\sum_g E_g(\Theta)\sigma_g(\lambda) - \sum_{j=0}^{3}\alpha_j\left(\lambda-\lambda^*\right)^j - \alpha_R R(\lambda) \tag{1}$$

Here, $I_\lambda$ is the earthshine spectrum at wavelength $\lambda$, $I_\lambda^0$ the solar spectrum, $E_g(\Theta)$ the effective slant column density of gas $g$ along geometrical path $\Theta$, $\sigma_g(\lambda)$ is the associated trace gas absorption cross section. The second term in Eq.1 is the closure polynomial (a cubic filter has been assumed), with $\lambda^*$ a reference wavelength for this polynomial. The last term on the right hand side of Eq.(1) is the additive terms for the Ring reference spectrum R($\lambda$). The fitting minimizes the weighted least squares difference between measured and simulated optical densities $Y_{meas}(\lambda)$ and $Y_{sim}(\lambda)$ respectively. The model in Eq. (1) is linear in the slant columns $E_g(\Omega)$, the polynomial coefficients $\{\alpha_k\}$ and the Ring scaling parameters $\alpha_R$.

Shift and squeeze parameters may be applied to cross-section wavelength grids to improve wavelength registration against Level 1 spectra. Experience with DOAS in the operational GDP processor has shown that fitting of such non-linear parameters on a pixel-by-pixel basis can sometimes leads to numerical instability, and an optimized pre-shift value needs to be applied.

Furthermore, it was found that DOAS fitting for GOME total columns achieves greater accuracy when two ozone cross-sections at different temperatures are used as reference spectra [Richter and Burrows, 2002].

At the pre-operational phase, the use of re-convolved GOME-FM98 ozone cross-sections [Burrows et al., 1999b] in the DOAS ozone slant column retrieval provides the most consistent and stable results for GOME-2 (the de-convolved GOME-FM98 ozone cross-sections have been convolved with the latest GOME-2 slit function v1.1 data [Siddans et al., 2006]). GOME-2 FM3 cross-sections for ozone and $NO_2$ (version 2.1) are also available [Gür et al., 2005], as well as other established laboratory ozone cross-sections [Bass and Paur, 1985; Malicet et al., 1995]. Here, it should be noted that the final version of the GOME-2 FM3 cross-section data has not been released yet.

Preshifting of the ozone and $NO_2$ cross-sections is required to compensate for inaccuracies in the wavelength calibration of the cross-section data. The re-convolved GOME-FM98 ozone cross-sections require a preshift of + 0.016 nm and are corrected for the so-called I0 effect [Aliwell et al., 2002].

In the GDP, the solar spectrum is used as the wavelength reference. Shift and squeeze parameters are applied to each Earthshine wavelength grid in order to re-sample the Earthshine spectrum. If necessary, the wavelength calibration of the GOME-2 level-1 spectra can be improved by applying window-dependent pre-shifts to parts of the solar spectrum before each orbit of data is processed. These pre-shifts are established by cross-correlation with a high-resolution solar spectrum [Chance and Spurr, 1997] over limited wavelength ranges covering the fitting window (325-335 nm for $O_3$, 425-450 nm for $NO_2$ in the visible, and 758-772 nm covering the $O_2$ *A* band as used in the ROCINN algorithm). For GOME-2 a relatively small pre-shift of ~0.0015 nm is found for the ozone fitting window and ~0.015 nm for the $NO_2$ fitting window

The Ring effect (filling-in of well-modulated solar and absorption features in earthshine spectra) is due to inelastic rotational Raman scattering (RRS). In DOAS fitting, it is treated as an additional absorber, by means of an additive Ring reference spectrum and associated scaling parameter, as in Eq. (1) above. The simplest 'Fraunhofer' Ring spectrum is obtained by folding rotational Raman cross-sections at a fixed temperature with a high-resolution Fraunhofer spectrum taken from the Kitt Peak Observatory [Chance and Spurr, 1997], but this does not include a telluric contribution. In the UV window 325-335 nm, Ring effect distortion of $O_3$ Huggins bands absorption features is large enough to seriously compromise total ozone fitting accuracy. As noted already, a new molecular Ring effect correction was developed for GOME total ozone in GDP 4.1. This correction is an *ex post facto* scaling of the DOAS slant column result, and it is performed at each iteration step in the AMF/VCD calculations (see section 2.3). A description of this molecular Ring correction algorithm is presented in Section 2.4.

The DOAS state vector for linear fitting in GDP 4.2 has 9 parameters: 2 effective slant columns of $O_3$ and $NO_2$, 1 fitting parameter for a second $O_3$ cross-section (to derive the effective temperature $T_{eff}$), 4 closure coefficients, and 2 additive scaling factors (corresponding to Fraunhofer Ring and undersampling reference spectra). There are 2 parameters in the nonlinear least-squares fitting: a wavelength shift and squeeze for re-sampling the earthshine spectrum on to the solar spectrum reference wavelength grid.

## 2.3 Air Mass Factor and vertical column computations

### 2.3.1 Iterative AMF/VCD method

The Air Mass Factor definition that is used in the GDP is the traditional one:

$$A = \frac{\log\left(I_{nog} / I_g\right)}{\tau_{vert}}, \tag{2}$$

where $I_g$ is the radiance for an atmosphere including the particular trace gas as an absorber, $I_{nog}$ is the radiance for an atmosphere without this trace gas and $\tau_{vert}$ is the vertical optical thickness of the trace gas.

To simulate the backscatter radiances $I_g$ and $I_{nog}$ in the AMF definition (Eq. 2), the LIDORT radiative transfer model is used [Spurr et al., 2001]. LIDORT is a multiple scatter multi-layer discrete ordinate radiative transfer code. The atmosphere is assumed stratified into a number of optically uniform layers. The LIDORT code used here neglects light polarization. Although polarization in RT simulations is an important consideration for ozone profile algorithms, in DOAS retrievals with narrow fitting windows in the UV, the polarization signature is subsumed in the closure polynomial. We use the LIDORT Version 2.2+ [Spurr, 2003] which possesses corrections for beam attenuation along curved line-of-sight paths, needed for the wide viewing angles of GOME-2 (scan angles in the range 40-50°)

For GOME scenarios, computation of the vertical column density (VCD) proceeds via the relation:

$$V = \frac{E + \Phi G A_{cloud}}{(1-\Phi) A_{clear} + \Phi A_{cloud}}, \tag{3}$$

where $E$ is the DOAS-retrieved slant column, $A_{clear}$ the clear sky AMF, $A_{cloud}$ the AMF for the atmosphere down to the cloud-top level, and the "ghost column" $G$ is the quantity of ozone below the cloud-top height, which cannot be detected by GOME and is derived from an ozone profile climatology (see section 2.3.2). This formula assumes the independent pixel approximation for cloud treatment. In GDP 4.2, we use the "intensity-weighted cloud fraction" $\Phi$. defined as:

$$\Phi = \frac{c_f I_{cloud}}{(1-c_f) I_{clear} + c_f I_{cloud}}, \tag{4}$$

where $I_{clear}$ and $I_{cloud}$ are the backscattered radiances for cloud-free and cloud-covered scenes respectively. $I_{clear}$ and $I_{cloud}$ are calculated with the LIDORT radiative transfer model, and depend mainly on the surface and cloud albedos and on the GOME viewing geometry.

AMFs depend on ozone profiles through the radiative transfer model. In traditional DOAS retrievals, the ozone AMF depends on a fixed ozone profile taken from climatology; one application of Eq. (3) yields the VCD. In the iterative approach to AMF calculation, we use a column-classified ozone profile climatology to establish a unique relationship between the ozone profile and its corresponding total column amount. The AMF values are now considered to be functions of the VCD through this profile-column relation, and the above formula in Eq. (3) is used to update the VCD value according to:

$$V^{(n+1)} = \frac{E + \Phi G^{(n)} A_{cloud}^{(n)}}{(1-\Phi) A_{clear}^{(n)} + \Phi A_{cloud}^{(n)}} \tag{5}$$

Here, the ($n$) superscript indicates the iteration number. The AMFs $A_{clear}^{(n)}$ and $A_{cloud}^{(n)}$, and the ghost column $G^{(n)}$, depend on the value of VCD $V^{(n)}$ at the $n^{th}$ iteration step. In this iteration, the slant column $E$ reflects the true state of the atmosphere and acts as a constraint on the iteration. Equation (5) is applied repeatedly until the relative change in $V^{(n)}$ is less than a prescribed small number $\varepsilon$. In other words, convergence is reached when $\left|(V_{n+1} / V_n) - 1\right| < \varepsilon$. For a value of $\varepsilon$ set at $10^{-4}$ (the GDP 4.2

operational baseline), convergence is rapid and 3-5 iterations are usually sufficient. The first guess choice $V_0$ comes from a zonally-averaged total column climatology derived from many years of TOMS data.
In GDP 4.2, there is a molecular Ring correction $M$ applied to the slant column $E$, and we must therefore use a corrected slant column $E_{corr} = E/M$ in the iteration. As we will see in section 2.4, $M$ depends on the total AMF, defined to be $A_{total} = (1-\Phi)A_{clear} + \Phi A_{cloud}$ . Clearly $M$ will need to be updated at each AMF/VCD iteration step, and our iteration formula now reads:

$$V^{(n+1)} = \frac{\dfrac{E}{M^{(n)}} + \Phi G^{(n)} A_{cloud}^{(n)}}{(1-\Phi)A_{clear}^{(n)} + \Phi A_{cloud}^{(n)}} . \tag{6}$$

The iterative AMF/VCD algorithm is straightforward to implement, and a flow diagram of the GDP 4.2 application is shown in Figure 2.

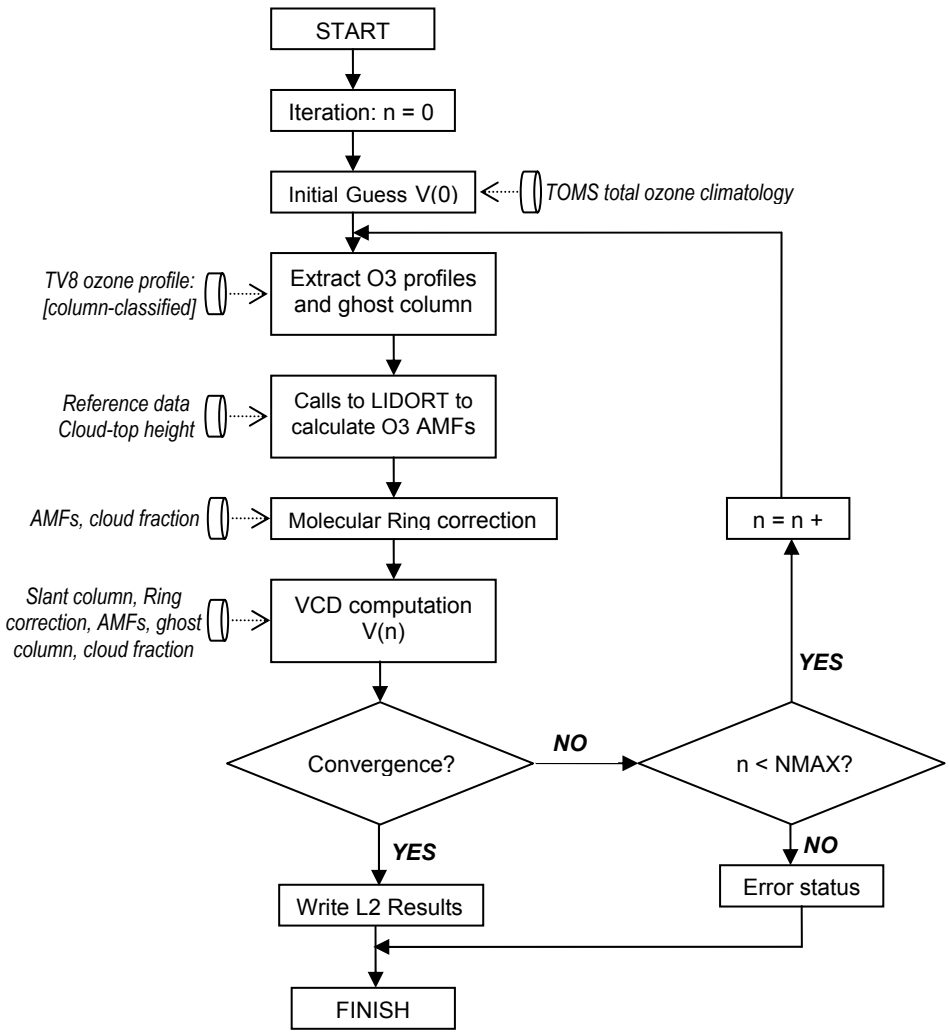



**Figure 2** Functional diagram of the iterative solution scheme for ozone air mass factors and vertical column densities (from Van Roozendael et al. [2006]).

### 2.3.2 The ozone profile-column map

A column-classified ozone profile climatology has recently been released for TOMS Version 8 [*Bhartia*, 2003] and this is used for GDP 4.2. This has a more sophisticated classification scheme than its predecessor, with 12 monthly profiles in 18 latitude zones at 10º intervals. The TV8 data has a variable column classification, from 3-5 columns at tropical latitudes and as much as 11 columns for polar regions. Column amounts vary from 125 DU to 575 DU and are separated at 50 DU intervals. Profile partial column amounts are also given in Dobson units.

The total ozone column $V$ is the sum of the partial columns $\{U_j\}$ that make up a given ozone profile, where $j$ is an index for the atmospheric layering. In the TV8 climatology, we are given a number of partial column profiles corresponding to fixed total column amounts. The profile-column mapping establishes the profile to be used for arbitrary values of the total column. For the linear profile-column map, the desired profile is expressed as a linear combination of two adjacent profiles $\{U_j^{(1)}\}$ and $\{U_j^{(2)}\}$ with corresponding total columns $V^{(1)}$ and $V^{(2)}$ bracketing $V$:

$$U_j(V) = \left(\frac{V - V^{(1)}}{V^{(2)} - V^{(1)}}\right) U_j^{(2)} + \left(\frac{V^{(2)} - V}{V^{(2)} - V^{(1)}}\right) U_j^{(1)} . \tag{7}$$

If the vertical column lies outside the range of values classifying the climatology, the profile is determined using a stable spline extrapolation scheme. This situation may occur in extreme ozone-hole scenarios ($V$ < 125 DU). Latitude and time of GOME-2 measurements are specified from Level 1 geolocation information. In order to avoid jump artefacts associated with discrete latitude and time classifications, the climatological profiles are interpolated between latitude bands using a linear weighting scheme based on the cosine of the latitude, and over time using a linear weighting based on the day of the month.

In GDP 4.2, we use the pressure grid of the ozone profile climatology for calculating layer optical properties required for the LIDORT computations. The TV8 climatology uses 11 partial columns with layer pressure differences based on atmospheric scale heights (pressures are halved for each successive atmospheric boundary). For each GOME pixel, it is necessary to adjust the lowest-layer partial column to account for the actual surface pressure (this depends for the most part on the assigned topographical height). This adjustment is done by scaling the partial column with the logarithm of the layer pressure difference. For the computation of AMFs to cloud-top, the lowest layer is bounded by the cloud-top pressure, and the corresponding partial column will also scale with the logarithmic pressure drop. The ghost column is the difference between clear and cloudy sky total columns, and it emerges directly from the profile-column mapping.

### 2.3.3 Intra-Cloud correction

GDP 4.x uses the Lambertian Equivalent Reflectivity cloud model (LER), also called clouds as reflecting boundaries model. The intra-cloud ozone column is improperly modeled in the LER approach, it may have a significant effect on the backscatter signal and total column errors could be large [*Liu et al*., 2004].
The total column below cloud-top is actually the sum of the intra-cloud ozone column ($V_{ic}$) plus the column below the cloud itself. In reality, backscatter measurements are sensitive to $V_{ic}$, and the traditional LER methods will overestimate the total atmospheric column by ignoring $V_{ic}$. GDP 4.2 uses a simple correction called Semi-transparent Lambertian cloud (STLC) model [*Loyola*, 2007]. It provides an initial empirical characterization of $V_{ic}$ as function of the climatological ozone column below cloud-top (ghost column), the cloud albedo, and the solar zenith angle.

### 2.3.4 Radiative Transfer Model for the AMF calculation

In GDP 4.2, the AMFs are computed directly using a fast radiative transfer model that is able to deliver all necessary AMF results well within the data turn-over rate. The LIDORT radiative transfer model [Spurr et al., 2001] is used to simulate backscatter radiances $I_g$ and $I_{nog}$ in the AMF definition in Eq. (2). LIDORT is a multiple scatter multi-layer discrete ordinate radiative transfer code. The atmosphere is assumed stratified into a number of optically uniform layers (in the ozone AMF computations, the layering scheme follows the TV8 pressure grid). The LIDORT code uses the pseudo-spherical (P-S) approximation: all scattering takes place in a plane-parallel medium, but attenuation of the solar beam before scatter is determined by ray-tracing through a spherical-shell atmosphere. The LIDORT code used here neglects light polarization. Although polarization in RT simulations is an important consideration for ozone profile algorithms, in DOAS retrievals with narrow fitting windows in the UV, the polarization signature is subsumed in the closure polynomial.
The P-S approximation is sufficiently accurate for AMF computations with solar zenith angle (SZA) up to 90° and for line-of-sight viewing angles up to 30-35° from the nadir. However, the P-S implementation is not accurate enough for the large viewing angles of GOME-2. This requires additional corrections for beam attenuation along curved line-of-sight paths, and for this we use the LIDORT Version 2.2+ [Spurr, 2003] which possesses this line-of-sight correction. LIDORT V2.2+ is used for all viewing modes in order to maintain consistency.

For DOAS applications with optically thin absorbers, the trace gas AMF wavelength dependence is weak and it is customary to choose the mid-point wavelength of the fitting window. This does not apply to ozone in the 325-335 nm DOAS fitting window, and for GDP versions up to and including 3.0, the $O_3$ AMF was always calculated at 325.0 nm. The motivation and explanation for this choice of wavelength may be found in [Burrows et al., 1999a]. Further testing of the AMF wavelength choice was done using simulated Level 1 GOME/ERS-2 radiances in [Van Roozendael et al., 2002], and it was shown that with this choice of 325.0 nm, total ozone errors of up to 5% are possible for solar zenith angles in excess of 80°, and generally, errors at the 0.5-1% level are found for sun angles < 80°. In the same study, it was shown that these errors are reduced (to the 1-2% level for SZA> 80°) when 325.5 nm is used as the representative AMF wavelength. The impact of the change in wavelength for the computation of the ozone AMFs is illustrated in Figure 3. The ozone vertical column error displayed in Figure 3 (lower panel) includes all basic aspects of the DOAS retrieval approach adopted for GDP 4.1 (except for cloud effects), and can be regarded as the “best-case “accuracy that can be expected from actual GOME retrievals. Errors below 1% are obtained in all typical GOME observation conditions, which is compliant with requirements on GOME total ozone accuracy, given the size of error sources in actual measuring conditions.

LIDORT is pure scattering code, and requires as input the following optical properties in each layer: (1) total extinction optical thickness, (2) total single scatter albedo, and (3) total phase function scattering coefficients. LIDORT also requires knowledge of the surface reflection (assumed Lambertian). In the GDP 4.2 application, there is an “atmospheric/surface setup module” which deals with detailed radiative transfer physics of molecules, trace gases, aerosols, clouds and surface reflection as needed to create the necessary LIDORT inputs. This setup function is completely decoupled from LIDORT, and this gives the AMF computation great flexibility. It is straightforward to change input climatology and other reference atmospheric and surface datasets. The setup function is described in the next section.

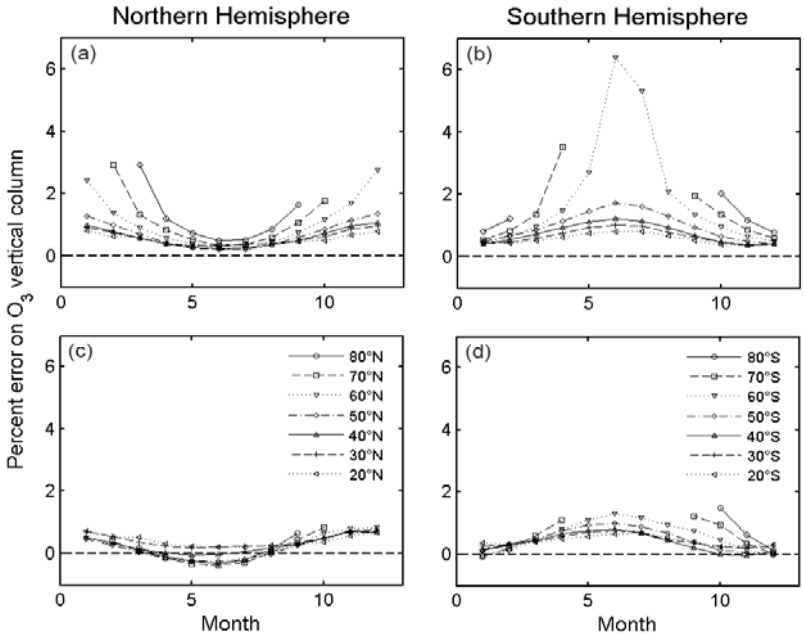



**Figure 3** Impact on the total ozone accuracy of the choice of single wavelength for ozone AMF computations. Retrievals were made using synthetic radiance data based on the ozone profile climatology of [Fortuin and Kelder, 1998] (12 months, 7 latitude bands, both hemispheres). Panels (a) and (b): percentage error on total ozone columns for AMFs calculated at 325.0 nm. Panels (c) and (d): percentage error on total ozone with AMFs at 325.5 nm (from Van Roozendael et al. [2006]).

### 2.3.5 Atmospheric and surface setups for the RT model

As noted above, GDP 4.2 uses pressure levels from the TV8 ozone profile climatology. Top of the atmosphere (TOA) is set at 0.03 hPa. Temperature profiles are required for hydrostatic balance and the determination of ozone cross sections. GDP 4.2 uses a zonal mean (18 latitude bands) and monthly mean temperature climatology that is supplied with the TV8 ozone profiles. Altitudes are determined by hydrostatic balance, with the acceleration due to gravity varying with latitude and height according to the specification in Bodhaine et al. [1999]. For surface topography, GDP 4.2 uses the GTOP30 topographical database (http://lpdaac.usgs.gov/gtopo30/gtopo30.asp). In the calculation of ozone absorption optical thickness, the pre-shifted GOME flight-model $O_3$ cross sections (as used in the DOAS fitting) are interpolated quadratically to account for the temperature dependence.

Rayleigh scattering is determined from a standard formula, but using the latest parameterizations as given in [Bodhaine et al., 1999]. The Rayleigh phase function depolarization ratio is taken from [Chance and Spurr, 1997]. In GDP 4.2 total ozone retrievals, aerosols are neglected in the AMF computations, since AMF and VCD values are insensitive to aerosols to first order. For sensitivity testing, we have used the MODTRAN aerosol data sets [Kneizys et al., 1988] to provide aerosol loading and optical properties. We return to the aerosol sensitivity issue in section 2.5 below.

In GDP 4.2, a dynamic albedo data set derived from accumulated satellite reflectance data is used: a combination of the GOME Lambertian equivalent reflectivity (LER) data set of albedos prepared from 5.5 years of reflectivity data [Koelemeijer et al., 2003], and the Nimbus-7 TOMS LER data set prepared from 14.5 years of data from 1978 [Herman and Celarier, 1997], and valid for 340 and 380

nm. The GOME LER data has monthly and yearly entries on a 1°x1° latitude/longitude grid, at 12 different wavelengths spanning the GOME range; the TOMS data is also monthly. We use GOME LER data at 335 and 380 nm, and TOMS LER data at 380 nm; the desired combination albedo is $a(\lambda) = s(\lambda)a_{TOMS}(380)$, where the scaling is $s(\lambda) = a_{GOME}(\lambda)/a_{GOME}(380)$, and $\lambda$ = 335 nm for total ozone fitting [Boersma et al., 2004]. In this way, the strengths of both data sets are combined: the long duration of the TOMS record (1978-1992) and the spectral information (11 wavelengths) of the shorter GOME record (1995-2001).

Changes in surface albedo values will chiefly affect the clear-sky AMF $A_{clear}$ and the intensity-weighted cloud fraction $\Phi$. The effect on the total ozone column is largest for cloud-free and partly cloudy scenes; for completely cloud-covered scenes the effect is generally small, since the clear-sky AMF plays no part in the total ozone column calculations (see Eq. (3) with $\Phi$ = 1).

In the independent pixel approximation, cloud information is reduced to the specification of 3 parameters (cloud fraction, cloud-top albedo and cloud-top pressure). Clouds are regarded as highly reflecting Lambertian surfaces. GDP 4.2 employs the OCRA and ROCINN cloud pre-processing steps before the total column retrieval. OCRA uses the GOME-2 sub-pixel PMD output and it delivers the geometric cloud fraction [Loyola, 1998]. ROCINN [Loyola, 2004] is a fitting algorithm using $O_2$ *A* band reflectivities from GOME-2, and it retrieves cloud-top pressure and cloud-top albedo. Cloud fraction in the ROCINN algorithm is constrained to take the OCRA value when the algorithms are used in tandem. The algorithms are summarized in Chapter 7. The GDP 4.2 algorithm can ingest cloud results derived from other algorithms, e.g. the FRESCO cloud parameters provided in the GOME-2 Level 1b data.

## 2.4 Molecular Ring correction

The smoothing ("filling-in") of Fraunhofer features in zenith sky spectra was reported in [Grainger and Ring, 1962] and has become known as the Ring effect. It is also present in satellite instruments measuring in the UV and visible. It is now known to be caused in large part by inelastic rotational Raman scattering (RRS) from air molecules. The Ring reference spectrum is defined as the change in optical depth between intensities calculated with and without RRS. The Ring effect is generally small, as RRS contributes only 4% of all scattering by air molecules. The Ring effect shows up best in spectral regions of significant intensity modulation such as the well known Fraunhofer Ca II lines around 394-398 nm. However, modulations of backscattered light in the ozone Huggins bands are also large enough for inelastic RRS effects to appear as the filling-in of ozone absorption features (the molecular or telluric Ring effect). Spectral dependence in this molecular Ring effect correlates quite strongly with the behavior of the ozone absorption.

As noted in section 2.2, the Ring effect is treated as "pseudo-absorber" interference in the DOAS algorithm using a Ring reference spectrum and additive fitting parameter. It was found that neglect of the telluric Ring effect in GDP 3.0 leads to systematic underestimation of ozone total columns (up to 10%) [Van Roozendael et al., 2002]. From this study, a correction for the molecular Ring effect in ozone retrieval was developed during the GOME geophysical validation campaign in 2002, as explained below.

Considering only $O_3$ absorption, the correction is based on a simplified forward model of the intensity at satellite $I(\lambda)$ which includes an explicit contribution due to inelastic RRS:

$$I(\lambda) = I^0(\lambda)\cdot\exp\left[-\sigma_{O3}(\lambda)\cdot E_{O3} - P_1^{\lambda}\right] + E_{Ring}\cdot I_0^{RRS}(\lambda)\cdot\exp\left[-\sigma_{O3}(\lambda)\cdot E_{O3}^{RRS} - P_2^{\lambda}\right]. \quad (8)$$

The first term on the right-hand follows the Lambert-Beer law for ozone absorption, with $I^0(\lambda)$ the solar intensity, and $\sigma_{O3}$ and $E_{O3}$ the ozone absorption cross-section and effective slant column

respectively. Elastic scattering effects are subsumed by means of the low band pass polynomial $P_1^{\lambda}$. The Ring effect is modeled by the second term in Eq. (8), in which there are several approximations. First, it is assumed that Raman-scattered light is generated close to the surface of the atmosphere, with the spectral shape given by a source spectrum for Raman scattering $I_0^{RRS}(\lambda)$. This source spectrum only treats the spectral smoothing effect of RRS on the solar intensity. In practice it is calculated by the convolution of a GOME irradiance spectrum using Raman cross sections appropriate to inelastic scattering into the wavelength of interest. The fractional intensity of Raman light (the $E_{Ring}$ parameter) is freely adjustable. This may vary considerably and will depend on parameters such as cloud coverage, cloud altitude and surface albedo. Ozone absorption (the term $\sigma_{O3}(\lambda).E_{O3}^{RRS}$) is then treated consistently, assuming that Raman photons produced at the surface and/or above clouds travel upward to the satellite. Ozone absorption taking place in the incoming light is assumed to be fully smeared out in the inelastic process, so that it can be neglected in the first approximation.

Raman scattered light smoothes out structured information in incident solar radiation. It can be seen as a source of atmospheric straylight which produces a low-side bias on any retrieved trace gas total column. This bias will nevertheless be modulated by atmospheric absorption in light paths above the region of RRS generation in the lower troposphere. For ozone, the bulk of the column is located in the stratosphere and upper troposphere, mostly above the source of RRS. Hence, ozone absorption that takes place in RRS light can be easily estimated. This is not necessarily the case for other trace gases, which may have significant partial columns in the lower troposphere. In summary, Raman scattering has a similar impact on all atmospheric absorbers, but it can only be accounted for accurately in a simple way for stratospheric trace gases such as $O_3$.

### 2.4.1 DOAS implementation

Equation (8) can be rewritten (after a Taylor expansion, discarding higher-order terms) in the following way:

$$\ln\left[\frac{I(\lambda)}{I^0(\lambda)}\right] = -\sigma_{O3}(\lambda).E'_{O3} - \sigma_{Ring}(\lambda).E_{Ring} - P(\lambda), \tag{9}$$

with the Ring cross-section $\sigma_{Ring}(\lambda)$ defined as:

$$\sigma_{Ring}(\lambda) = -\frac{I_0^{RRS}(\lambda)}{I^0(\lambda)}. \tag{10}$$

Equation (9) is the familiar DOAS fitting model, from which $E'_{O3}$, $E_{Ring}$ and the $P(\lambda)$ polynomial coefficients can be derived in the usual manner. The major difference with Ring correction methods used in previous studies comes in the definition of the modified $O_3$ effective slant column $E'_{O3}$, which is related to the effective slant column for *elastic* scattering ($E_{O3}$) by the following formula:

$$E'_{O3} \cong E_{O3}.\left\{1 + E_{Ring}\cdot\overline{\sigma}_{Ring}\cdot\left(1 - \frac{\sec(\theta_0)}{A_{total}}\right)\right\} = E_{O3}.M_{Ring}, \tag{11}$$

where $A_{total}$ is the ozone AMF, $\theta_0$ the solar zenith angle, and $\overline{\sigma}_{Ring}$ an average Ring cross-section calculated over the spectral fitting interval. Equation (11) defines the molecular Ring correction $M_{Ring}$. From section 2.3, we have $A_{total} = (1-\Phi)A_{clear} + \Phi A_{cloud}$ in the independent pixel approximation, where $\Phi$ is the intensity-weighted fractional cloud cover.

In this formulation, the DOAS fitting is essentially unchanged, and it gives fitted parameters $E'_{O3}$ and $E_{Ring}$. The effective slant column for ozone is then adjusted *after* the fit through the relation $E_{O3} = M_{Ring}\, E'_{O3}$. Note that the molecular Ring term $M_{Ring}$ can also be used to quantify the error due to an incorrect estimation of the Ring effect in previous GDP versions. Studies have shown that for moderate SZA, the geometrical AMF is sufficiently accurate to approximate $A_{total}$ in Eq. (11). For high SZA with a long path through absorbing ozone layers, a more precise calculation is needed. However in GDP 4.2, we use the LIDORT-calculated total AMF already computed at each AMF/VCD iteration step to obtain $M_{Ring}$ and the corrected slant column $E'_{O3} = E_{O3} / M_{Ring}$ as required for the VCD update (Eq.(6)).

Figure 4 shows values of the molecular Ring correction term $M_{Ring}$ for four seasonally representative GOME/ERS-2 orbits. In GDP 4.1, ozone slant columns are clearly scaled up by 2 to 9% and this is more than enough to compensate for the negative bias observed in several GOME validation campaigns. The general shape of the correction factor is due to the variation of SZA across the GOME orbit. Pronounced peaks and high-frequency oscillations are mainly due to clouds, but changes of surface albedo and surface height can influence the correction. The cloud impact is especially visible for orbit 18248 (orange) at latitudes of 10°N and 30°S where the GOME measurements were affected by high clouds and the high cloud fractional cover typically found in tropical regions. With RRS dominant in the lower troposphere, high cloud cover implies an immediately noticeable reduction in the RRS contribution to the measured radiance, and a consequent reduction in the Ring correction factor (closer to unity). The influence of the surface albedo is obvious at high Southern latitudes where a sharp increase of the albedo around 60°S due to sea ice and the Antarctic ice shield is associated with a corresponding decrease of the Ring correction term.

The accuracy of our simplified approach has been tested under various conditions using the same SCIATRAN model as a reference. We describe here some verification tests performed using simulations of the earthshine radiances provided by the SCIATRAN code. The atmosphere was set up using ozone profiles from a seasonally classified and latitude resolved climatology [Fortuin and Kelder, 1998]. Simulations of synthetic radiances in the range 320-340 nm were performed at SZAs representative of GOME observations at latitudes and seasons sampled by the climatology. DOAS retrievals were then performed using configuration settings as they appear in operational GOME retrievals. This systematic underestimation is largely compensated by the new correction, for all conditions applied in the tests.

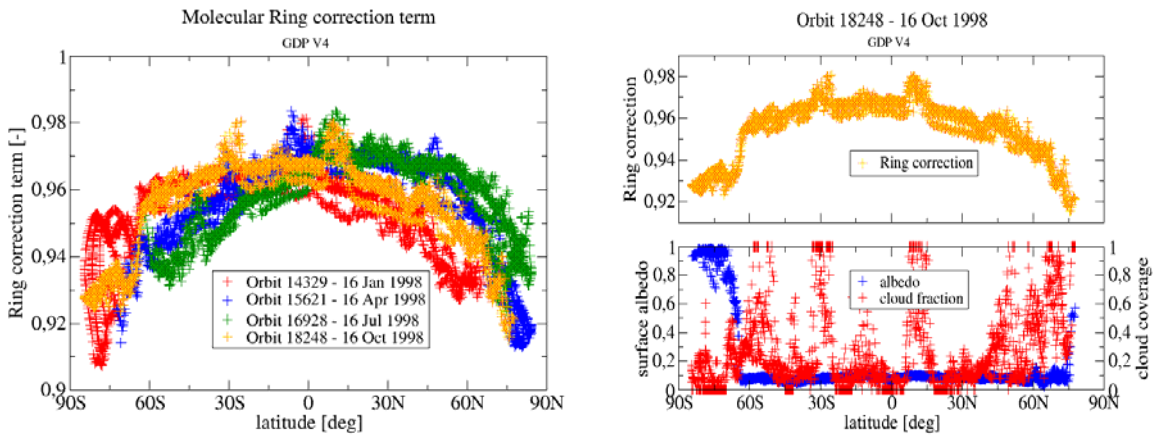



**Figure 4** Molecular Ring correction factors for four GOME/ERS-2 orbits in 1998 (left panel) and Ring correction, surface albedo and fractional cloud cover for one GOME Orbit (right panel). See text for more information (from Van Roozendael et al. [2006]).

## 2.5 Error budgets and sensitivity studies

### 2.5.1 Error budgets for the total ozone algorithm

Referring to Eq. (3) in section 2.3, the error on vertical column $V$ (denoted as $s_V$) can be expressed as a function of the error on component parameters $E$ (ozone slant column), $G$ (ghost vertical column), $\Phi$ (radiance-weighted cloud fraction), $A_{clear}$ (AMF for a clear sky scene), $A_{cloud}$ (AMF to cloud-top). A complete definition can be derived from error propagation rule:

$$s_V^2 = \left(\frac{\partial V}{\partial E}\right)^2 \cdot s_E^2 + \left(\frac{\partial V}{\partial A_{clear}}\right)^2 \cdot s_{A_{clear}}^2 + \left(\frac{\partial V}{\partial A_{cloud}}\right)^2 \cdot s_{A_{cloud}}^2 + \left(\frac{\partial V}{\partial \Phi}\right)^2 \cdot s_\Phi^2 + \left(\frac{\partial V}{\partial G}\right)^2 \cdot s_G^2 \cdot \tag{12}$$

This error propagation formula is strictly valid under the assumption that error sources are mutually uncorrelated. In general we would expect some correlations (for example between the cloud fraction, and cloud-top height and cloud-top albedo), but the derivation of a complete error covariance for all sources is beyond the scope of the present work. With this in mind, we may use the definition of $V$ in Eq. (3) to obtain:

$$\frac{\partial V}{\partial E} = \frac{1}{A_T}; \qquad \frac{\partial V}{\partial G} = \Phi \cdot \frac{A_{cloud}}{A_T}; \quad \frac{\partial V}{\partial \Phi} = \frac{1}{A_T}\left[V \cdot A_{clear} - (V - G) \cdot A_{cloud}\right];$$

$$\frac{\partial V}{\partial A_{clear}} = -\frac{V}{A_T}(1-\Phi); \qquad \frac{\partial V}{\partial A_{cloud}} = -\frac{\Phi}{A_T}(V - G) \cdot$$

$$A_T = (1-\Phi)A_{clear} + \Phi A_{cloud}$$

Error component $s_E$ comes from the DOAS slant column fitting, and $s_\Phi$ from the OCRA cloud pre-processing. In GDP 4.2, an AMF error is assumed that is dependent on the solar zenith angle, and the ghost column error is taken as $s_G$ = 30%. As discussed below, the solar zenith angle dependency of the AMF error has been determined empirically from an examination of the variability of the $O_3$ AMFs over a wide range of ozone profiles. It should be noted that this simplified error formulation is introduced for the calculation of the errors on a pixel-by-pixel basis, and it only includes the largest contributors to the total error budget.

A more comprehensive estimation of the error budget for the GDP 4.2 ozone columns is provided in Table 2. This includes typical errors on ozone slant columns, ozone AMFs, cloud fractions and ozone ghost column, and are for the most part derived from the GDOAS delta-validation report for GOME/ERS-2 [Van Roozendael et al., 2004]. The error budget in Table 2, which has been derived for GOME/ERS-2, can serve as an initial (theoretical) error assessment for GOME-2/MetOp.

The error budget has been separated into two parts: errors affecting the retrieval of slant columns (DOAS-related errors) and errors affecting the conversion of slant columns into vertical columns (AMF-related errors). Since several AMF-related error sources are significantly enhanced at large SZA, the AMF-related part of the error budget has been divided into two regimes (SZA < 80°, and SZA ≥ 80°).

The DOAS-related (slant column) uncertainties quoted in Table 2 are for the most part extracted from the study associated with the GDP 3.0 Delta-validation for GOME/ERS-2 [Van Roozendael et al., 2002]. Error values are determined from a number of sensitivity tests dealing with the impact of uncertainties on absorption cross-sections and their temperature dependence, as well as wavelength calibration and convolution issues. We include the molecular Ring effect error under the DOAS

heading. Errors due to the molecular Ring effect are derived from retrieval tests using synthetic radiance data, as presented in the GODFIT validation report [Van Roozendael et al., 2003].

Errors relating to $O_3$ AMF values are determined from a series of sensitivity tests carried out using different settings for the AMF calculations (e.g. different $O_3$ profile climatologies, or the error from the assumption of a single wavelength choice for the AMF calculation). In addition, the impacts of surface albedo errors as well as cloud and aerosol uncertainties have been considered explicitly. Several error sources are significantly enhanced at large solar zenith angles (typical of polar spring and autumn observations), and this justifies the division in the error budget in Table 2 between values representative of solar zenith angles lower than and greater than 80°. Independently of albedo and cloud/aerosol effects, errors on AMFs will depend significantly on the shape of the ozone profile as well as its column content. Hence an upper limit of the AMF error (and its SZA dependence) can be obtained from consideration of the variability of $O_3$ AMFs calculated using a wide range of climatological ozone profiles. The AMF variability is a strong function of the SZA, especially above 80°.

In an attempt to parameterize the main dependency of the AMF error, we have assumed that the AMF uncertainty can be linked to atmospheric profile shape errors, which will have a larger impact at high SZA values. For operational implementation in GDP 4.2, this curve has been used to derive an empirical relationship between AMF uncertainty and solar zenith angle. A simple scaling (by a factor of 2) has been applied to the variability curve in such a way that the resulting error curve matches up with the error estimates shown in Table 2 for both SZA ranges. Although it is not the result of a rigorous error analysis, this empirical parameterization has the advantage of providing realistic uncertainties on the GDP 4.2 total ozone product both at low and at high SZA. The ghost column estimate of 30% used in GDP is a composite value based on error contributions from a number of sources (in particular, the ROCINN estimate of cloud-top height error and the uncertainty on the tropospheric part of the ozone profile).

### 2.5.2 Sensitivity issues for GDP 4.2 algorithm

In GDP 4.2, the largest impact of atmospheric temperature is through the temperature-dependence of the ozone absorption cross-sections. Two ozone spectra at two different temperatures are used in the DOAS fitting; the accuracy of this approach is limited (1) at large SZA, due to the breakdown of the optically thin approximation, (2) at extreme stratospheric temperatures (due to non-linearity in the temperature dependence of the ozone cross-sections), and (3) by the intrinsic accuracy of the laboratory cross-sections. It is possible that instrument degradation also has an impact on the accuracy of the effective temperature determination. This has not been tested explicitly, but results from overpass processing over Hohenpeissenberg and Lauder, extending from 1996 until 2003 and retrieved with no particular attempt to compensate for known GOME degradation problems, suggest that the DOAS algorithm is stable and not strongly influenced by the degradation of the instrument (see Figure 5 for the GOME/ERS-2 time-series).

As noted already, the long-term stability of the GOME total ozone record is a key consideration for trend analysis. In Figure 5, monthly mean ozone differences between GDP 4.1 and Brewer measurements at Hohenpeissenberg are shown for a 10-year period from July 1995 through April 2005. A sine function has been fitted to the time series in order to highlight seasonal variations in the differences. The amplitude of these variations is about 0.5% and the mean bias is 0.3%. The long-term stability of GOME and the absence of any significant time-dependent bias are clear. It is worth noting that the stability is still evident after more than 8 years, despite some loss of ozone accuracy from June 2003 to December 2004 caused by the absence of daily solar calibration measurements in the GOME Level 1 product during that period (this problem has been solved in the updated GOME/ERS-2 Level 1 processor).

**Table 2** Estimation of error sources of the GDP 4.2 total ozone retrievals as derived for GOME/ERS-2 [Van Roozendael et al., 2004].

| Error source | Percent error | |
|---|---|---|
| | SZA < 80° | SZA > 80° |
| **Ozone slant column** | | |
| $O_3$ absorption cross-sections | <2 | <2 |
| Atmospheric (effective) temperature determination | <1.5 | <3 |
| Instrument signal-to-noise | 0.5 | <2 |
| Instrument spectral stability (wavelength registration) | 0.5 | 0.5 |
| Solar $I_0$-effect | 0.2 | 0.2 |
| Ring and molecular Ring effect | <2 | <2 |
| **Ozone Air Mass Factor** | | |
| Single wavelength calculation (325.5 nm) | <1 | <2 |
| $O_3$ profile | <1 | <4 |
| Surface albedo | 0.3 | 0.3 |
| Cloud fraction | 0.8 | 0.8 |
| Cloud top pressure (height) | 1 | 1 |
| Cloud top albedo (optical thickness) | 0.8 | 0.8 |
| Ghost column | <2 | <3 |
| Tropospheric aerosols (background conditions) | 0.2 | 0.2 |
| **Ozone vertical column (accuracy)** | | |
| Clear | <3.6 | <6.4 |
| Cloudy | <4.3 | <7.2 |
| **Ozone vertical column (precision)** | | |
| Clear | <2.4 | <4.9 |
| Cloudy | <3.3 | <5.9 |

Since the iterative AMF/VCD algorithm relies on an ensemble of ozone profiles to define the profile-column map needed for the iteration, the choice of ozone profile climatology is important. Ozone profile shape is a key factor controlling the accuracy of the total ozone retrieval, especially at high latitudes where the ozone profile-shape sensitivity of the AMFs is enhanced by the extreme variations in the ozone field (e.g. ozone hole) combined with large solar zenith angles. The GDP 4.1 code has been tested using both the TOMS Version 7 and Version 8 ozone profile climatologies. Differences in retrieved total ozone columns using the two climatologies are shown in Figure 6 for a sample data set consisting of 465 orbits from 1997. Largest differences are found in polar regions (especially in the southern hemisphere) close to the terminator where GOME SZAs are at their maximum. In [Spurr et al., 2005], it was noted that the fixed ozone burden in the troposphere was a significant error source for ozone AMFs in GDP 3.0, particularly at low SZA (maximum photon penetration). In the TV7 data set, ozone partial columns are fixed at 9 DU and 15 DU in the lowest two layers. There is much more tropospheric variation in ozone content with the Version 8 profile data, but it remains the case that errors of 10-15 DU in the tropospheric boundary layer ozone burden can induce AMF errors of 3-5% for low SZA values (~25°). This may explain the surprisingly large sensitivity in Figure 6 for the Northern sub-tropics during summer when the GOME SZA is at minimum.

It is difficult to extract any information about aerosols from a DOAS fitting of ozone in the UV Huggins bands. Aerosol scattering and extinction are subsumed in the DOAS slant column fit through the closure polynomial, and the introduction of parameterized aerosol information in the AMF RT simulations is an additional source of error. For a 10-year operational reprocessing of the GOME/ERS-2 record, it is impossible to account for aerosol variability in anything but the simplest terms, and the policy in GDP 4.1 has been to avoid the use of aerosols altogether, and to use a Rayleigh atmosphere for the baseline AMF calculations. For scattering aerosols in the troposphere, the AMF is relatively insensitive to aerosol content. For background aerosol conditions the error is small: ~0.2%; for more optically thick aerosol regimes, the error generally remains below the 1% level. It is known however that for scenarios with absorbing aerosols present (in particular biomass burning, industrial pollution, desert dust outbreaks and volcanic plumes), ozone AMFs may be significantly in error if the aerosol presence is ignored or not treated accordingly. These effects are again largest for low SZA. Aerosols are not treated explicitly in the GDP 4.2 AMF calculations. However, a significantly scattering aerosol layer will be detected by OCRA/ROCINN as a thin cloud layer, and the aerosol affect will thus be included indirectly in the vertical column calculation. To first-order, aerosol uncertainties in the GDP 4.2 total algorithm will be picked up in the cloud parameter error budget estimates. Although cloud fractions are in general weakly influenced by the presence of aerosols, cloud algorithms such as FRESCO and OCRA/ROCINN are sensitive to strong aerosol pollution episodes.

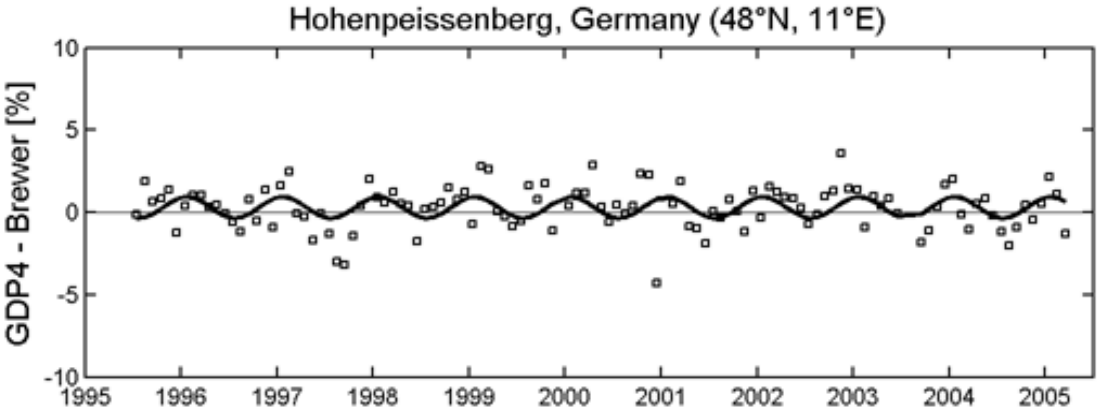



**Figure 5** GDP v4 of GOME/ERS-2 – Hohenpeissenberg Brewer monthly mean ozone differences from July 1995 until April 2005. A sinusoidal fit to the time series (thick black line) highlights the size of seasonal variations in the differences (amplitude: 0.5%). The mean bias over the 10-year time period is 0.3% (from Van Roozendael et al. [2006]).

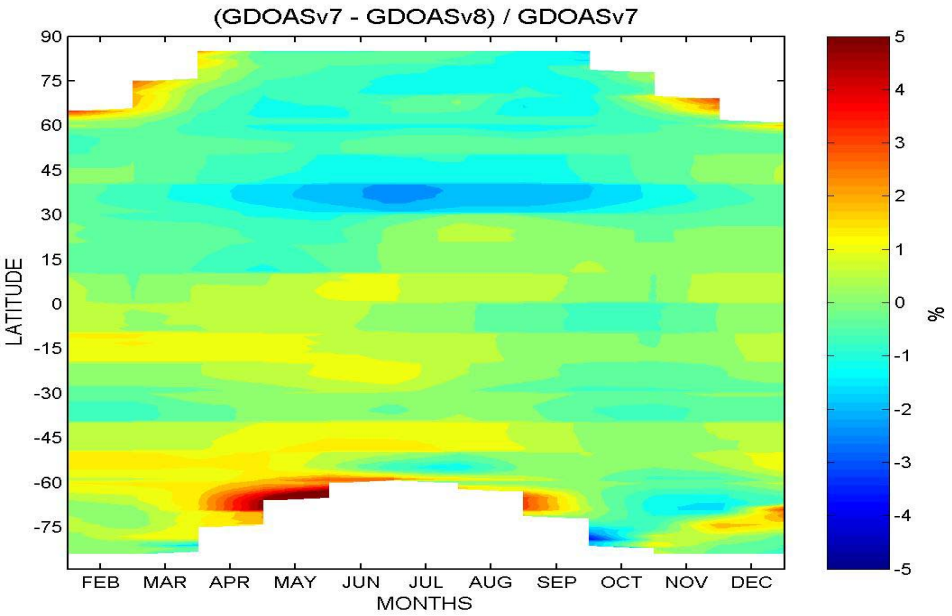



**Figure 6** Relative differences in GDP v4 GOME/ERS-2 total ozone retrieved using the two TOMS version 7 and version 8 ozone profile climatologies. Differences are mostly significant in polar regions, close to the terminator, as well as in northern tropical regions around the place of minimum GOME solar zenith angle (from Van Roozendael et al. [2006]).

# 3 THE NO2 COLUMN ALGORITHM

## 3.1 DOAS slant column fitting

The GDP 4.2 $NO_2$ DOAS algorithm is very similar to that for total ozone, and uses the same least squares fitting package; the description in Section 2.2 is relevant here, with the following differences:

- The fitting window is 425-450 nm in GOME-2 Channel 3. $NO_2$ absorption features are prominent, and GOME measurements have high signal-to-noise and manageable interference effects.
- A single $NO_2$ cross-section reference spectrum is used. For GOME-2, the GOME-2 FM3/CATGAS cross-sections for Channel 3 at 243 K are used [Gür et al., 2005]. There is no retrieval of an effective temperature; temperature dependence of the cross-sections is accounted for on the AMF level (see below).
- There is one additive Fraunhofer Ring spectrum for this region of Channel 3; An updated Fraunhofer spectrum for GOME-2 FM3 have been prepared by BIRA-IASB.
- Intensity offset effects that may be induced by residual stray-light or remaining calibration issues in the level-1 product are known to be sources of bias in DOAS retrievals of minor trace species; to correct for offset the inverse of the sun spectrum is fitted as another effective cross-section.
- $O_3$ is an interfering species and the slant column amplitude for it is included in the fit. However, $O_3$ absorption in this part of the Chappuis bands is weak (one reason for the fitting window choice). In this wavelength region, the GOME-2 FM3/CATGAS cross-sections data at 221 K can be used [Gür et al., 2005].
- $O_2$-$O_2$ and $H_2O$ are interfering species and slant column amplitudes for them are included in the fit. Sources are [Greenblatt et al., 1990] for $O_2$-$O_2$ (recalibrated) and HITRAN [Rothman et al., 2003] for $H_2O$ (the latter as input to line-by-line computations which are followed by GOME-2 FM3 slit function convolution).
- There is no molecular Ring correction implemented for the pre-operation phase. For $NO_2$, the error in the retrieved total column due to the molecular Ring effect is small (1-2%) as compared to the other error sources, see also Section 3.4.
- The broadband filtering polynomial is cubic (4 coefficients).

The total number of fitting parameters is 10, comprising 4 trace gas slant columns, 4 polynomial coefficients, and 2 amplitudes for additive reference spectra. Wavelength registration is done as for total ozone DOAS: the solar spectrum is the wavelength standard, with a shift-and-squeeze fitting performed for each footprint for resampling the earthshine spectrum. “Post Level 1” wavelength registration for the solar spectrum is improved at the orbit start by an additional cross-correlation covering the 425-450 nm fitting window.

The $NO_2$ absorption cross-section has a marked temperature depedence, which has to be taken into account to improve the accuracy of the retrieved columns. In the GDP 4.2, a single $NO_2$ cross-section reference spectrum at 243 K is used, and the temperature dependence of the cross-sections is accounted for on the AMF level using the correction scheme developed by Boersma et al. [2004]. This method uses a correction factor as a function of temperature, and then apply it to the slant column using a temperature and $NO_2$ profile, and the altitude dependent AMF.

## 3.2 AMF and VCD determination

The AMF is calculated with the LIDORT 2.2+ model for the window mid-point (437.5 nm), since $NO_2$ is an optically thin absorber in this wavelength region. To incorporate the seasonal and latitudinal variation in stratospheric $NO_2$ in the AMF calculations, a composite climatology of stratospheric $NO_2$ profiles is used [Lambert and Granville, 2004]. The computation of the $NO_2$ vertical column density

proceeds via Eq. (3). An AMF/VCD iteration (as implemented in the total ozone algorithm) is not needed given the small optical thickness of $NO_2$.

With this choice of profiles, the vertical resolution need not be too fine, and it will be sufficient to use the 13-layer grid based on TOMS pressure levels that was used for the ozone AMF computations. Molecular scattering and aerosol optical properties will again be drawn from the sources mentioned in section 2.3. Ozone profiles will be taken from the TOMS climatology (this is not a critical consideration). Cloud information will be used in the same way as before. The choice of surface albedo will again be combined from the GOME LER (values at 380 nm and 440 nm) and TOMS LER (values at 380 nm) databases.

### 3.3 Tropospheric NO2 column calculation for polluted conditions

The $NO_2$ retrieval method described above uses a stratospheric AMF to compute the $NO_2$ total column density. This method is valid over much of the Earth, but it underestimates the total column density in polluted areas with significant $NO_2$ in the troposphere. For polluted areas, a more accurate $NO_2$ column retrieval is achieved by subtraction of the estimated stratospheric $NO_2$ column before evaluation of the tropospheric component. This correction procedure consists of three steps: 1) estimate the stratospheric component of the $NO_2$ column using spatial filtering, 2) recognition of geographic regions that contain significant tropospheric pollution, and 3) determine the tropospheric $NO_2$ column using an accurate tropospheric AMF for these polluted regions, and correct the initial total $NO_2$ column for this tropospheric component. This approach is similar to the one used for the $NO_2$ product of the Ozone Monitoring Instrument (OMI) on EOS-Aura [Bucsela et al., 2006].

The spatial filtering approach to determine the stratospheric $NO_2$ component is based on the assumption that the gradients in stratospheric $NO_2$ are much larger in the latitude direction than in the longitude direction, and that the spatial variability of tropospheric $NO_2$ occurs on smaller scales than that of stratospheric $NO_2$. First a global map from the initial $NO_2$ columns is constructed by binning 24 hours of GOME-2 data on a high resolution spatial grid. To minimize tropospheric bias in the stratospheric field, an *a priori* global mask is used to eliminate large areas with potentially high amounts of tropospheric $NO_2$. The stratospheric $NO_2$ column $V_s$ is then determined by low-pass filtering the initial $NO_2$ columns in the zonal direction.

A tropospheric correction is applied to all GOME-2 observation with an initial total $NO_2$ column that is significantly larger than the estimated stratospheric component. In those cases, the tropospheric $NO_2$ column $V_t$ is determined, and a corrected total column $V_c$ is calculated:

$$V_t = \frac{E - A_s V_s}{A_t} \tag{13}$$

$$V_c = V_s + V_t \tag{14}$$

where $E$ is the slant column density calculated in the DOAS fit and $V_s$ is the stratospheric component, as calculated with the spatial filtering method. $A_s$ is the stratospheric air mass factor, and is calculated using the composite stratospheric $NO_2$ profile climatology (as described above). $A_t$ is a tropospheric air mass factor based an *a priori* tropospheric $NO_2$ profile. For this, climatological monthly tropospheric $NO_2$ profiles from the MOZART-2 chemistry transport model in 1.875° longitude-latitude bins are used [Nüß et al., 2005]. For GOME-2 observations with a tropospheric correction applied, both the corrected total vertical column density $V_c$ and the tropospheric column density $V_t$ are reported in the data product, as well as the initial vertical column density. The tropospheric correction is complicated in case of (partly) cloudy conditions. For most measurements over cloudy scenes, the cloud-top is well

above the $NO_2$ pollution in the boundary layer. In those cases, the enhanced tropospheric $NO_2$ concentrations can not be detected by GOME. Therefore, a tropospheric correction will only be applied to GOME observations with an “intensity-weighted” cloud fraction smaller than 50%.

## 3.4 Error budget for the total and tropospheric NO2 column

A preliminary estimation of the error budget for the total and tropospheric $NO_2$ column is provided in Table 3. This includes typical errors on $NO_2$ slant columns and the AMF for the total $NO_2$ column for unpolluted conditions and the tropospheric $NO_2$ column (for polluted conditions). The preliminary error-estimates are mainly based on initial DOAS analyses using GOME-2 data (see Section B in Lambert et al. [2007]), and the $NO_2$ error analysis of Boersma et al. [2004].

An initial validation of the $NO_2$ total column product with ground-based NDACC/UVVIS spectrometers generally show a good agreement, except for the southern mid- and high latitudes, where GOME-2 reports systematically smaller $NO_2$ verical columns than the ground-based measurements [Lambert et al., 2007].

**Table 3** Initial estimation of error sources for the total $NO_2$ column for unpolluted conditions and the tropospheric $NO_2$ column (for polluted conditions).

| Error source | Percent error | |
|---|---|---|
| | Total column (unpolluted) | Tropospheric column |
| **$NO_2$ slant column** | | |
| $NO_2$ absorption cross-sections | 2-5 | 2-5 |
| Instrument signal-to-noise | 5 | 5 |
| Instrument spectral stability (wavelength registration) | 0.5 | 0.5 |
| Ring and molecular Ring effect | <2 | <2 |
| Stratospheric NO2 column | n.a. | 10-20 |
| **$NO_2$ Air Mass Factor** | | |
| $NO_2$ profile shape | <1 | 10 |
| Surface albedo | <1 | 8 |
| Cloud fraction | <1 | 8 |
| Cloud top pressure (height) | <0.5 | 3 |
| Tropospheric aerosols (background conditions) | <0.5 | <10 |
| **$NO_2$ vertical column (accuracy)** | 5-10 | 50-100 |

# 4 THE BRO COLUMN ALGORITHM

The original baseline algorithm uses the 344.6-359 nm wavelength range for the DOAS slant column fit of BrO (so-called "GOME" fitting-window). A BrO cross-section is included in the fit, as well as the cross-sections of the interfering trace gases: ozone, $NO_2$, $O_2$-$O_2$, and HCHO. The BrO cross-sections are from Wilmouth et al. [1999] and convolved with the latest GOME-2 slit function data [Siddans et al., 2006]; the re-convolved GOME-FM98 ozone cross-sections [Burrows et al*.,* 1999] and the GOME-2 Flight Model/CATGAS $NO_2$ cross-sections are used [Gür et al., 2005]. Ozone cross-sections at two temperatures (221K and 241K) are included, and the $NO_2$ cross-sections at 241K. The $O_2$-$O_2$ cross-sections are from Greenblatt et al. [1990], the HCHO cross-sections are from Cantrell et al. [1990].Two Ring reference spectrums are included as an additive fitting parameter.

Initial DOAS fit analyses with GOME-2 data show unexpectedly large scatter in the retrieved BrO slant columns. This is a major issue that is currently under investigation. A more stable retrieval is obtained by an alternative, UV-shifted, fitting window: 336–351.5 nm (so-called "SCIAMACHY" fitting-window). In this fitting window, the most consistent and stable results are obtained using the BrO cross-section data from Fleishmann et al. [2004] and the re-convolved GOME-FM98 ozone cross-sections at two temperatures (221K and 241K) [Burrows et al*.,* 1999b]. The $NO_2$ and $O_2$-$O_2$ cross-sections are the same as in "GOME" fitting-window. Because of the strong interference between the BrO and HCHO cross-sections in this wavelength region, HCHO is not included in the DOAS-fit.

The AMF is calculated with the LIDORT 2.2+ model for the fitting window mid-point, since BrO is an optically thin absorber in this wavelength region. To incorporate the seasonal and latitudinal variation in stratospheric BrO in the AMF calculations, a stratospheric BrO profile climatology is used [Bruns et al., 2003]. This climatology contains monthly mean BrO profiles as a function of latitude, based on the chemistry transport model SLIMCAT. The computation of the BrO vertical column density proceeds via Eq. 3. An AMF/VCD iteration (as implemented in the total ozone algorithm) is not needed given the small optical thickness of BrO. Activities on further improvements of the BrO column algorithm are ongoing [Van Roozendael and Theys, 2005]. This work focuses on optimizing the accuracy of global total BrO columns, as well as polar tropospheric BrO columns.

Because of the issues in the BrO DOAS fit described above, only a preliminary error estimate for the BrO column can be given at this stage (see Table 4).

**Table 4** Initial estimation of error sources for the total BrO column.

| Error source | Percent error |
|---|---|
| **BrO slant column** | 15-30 |
| BrO absorption cross-sections | 5-10 |
| Instrument signal-to-noise | 10-20 |
| **BrO Air Mass Factor** | 5-20 |
| **BrO vertical column (accuracy)** | 20-50 |

# 5 THE SO2 COLUMN ALGORITHM

## 5.1 DOAS slant column fitting

The DOAS algorithm for $SO_2$ is based on the algorithm for ozone, as described in Section 2.2. The DOAS algorithm settings for $SO_2$ are listed below:

- The DOAS slant column fit of $SO_2$ is performed in the UV wavelength range 315-326 nm [Thomas et al., 2005].
- A single $SO_2$ cross-section is included in the fit, the $SO_2$ cross-sections are the SCIA Flight Model cross-sections from Bogumil et al. [1999], reconvolved with the GOME-2 slit function data. To account for the temperature dependence of the $SO_2$ cross-sections based on the assumed height of the $SO_2$ plume, cross-sections at three different temperatures are used: 203K for an assumed plume height of 15km, 243K for a plume height of 6km and 273K for 2.5km plume height (see also next section on the AMF determination).
- Cross-sections of the interfering trace gases ozone and $NO_2$ are included. The best results in this wavelength region are obtained using the Malicet et al. [1995] ozone cross-sections at two temperatures (218K and 243K) with a pre-shift of -0.01 nm; for $NO_2$, the GOME-2 Flight Model/CATGAS cross-sections is used at 241K [Gür et al., 2005].
- Furthermore two Ring reference spectra calculated with the SCIATRAN model are included as additive fitting parameters to account for the molecular ring effect.
- Intensity offset effects that may be induced by residual stray-light or remaining calibration issues in the level-1 product are known to be sources of bias in DOAS retrievals of minor trace species; to correct for possible offsets, the inverse of the sun spectrum is fitted as another effective cross-section.
- The broadband filtering polynomial is cubic.

## 5.2 $SO_2$ background correction

In the wavelength range 315-326 nm, there is a strong interference of the $SO_2$ and ozone absorption signals resulting in “negative” $SO_2$ slant columns for higher solar zenith angles. Therefore, a three step offset correction is applied to the $SO_2$ slant column values. In the first step, an equatorial offset is calculated that accounts for any systematic bias in the $SO_2$ column. The offset is calculated on a daily basis from GOME-2 measurements in the equatorial region, where no $SO_2$ sources are present. This offset is then subtracted from the original $SO_2$ slant column densities. In the second step, a correction is made for the dependency of the $SO_2$ slant column on the ozone column. This ozone correction factor is determined from one year of GOME-2 ozone and $SO_2$ data and applied to the individual $SO_2$ slant columns. A possible remaining dependence of the $SO_2$ slant column on the solar zenith angle is corrected in a third step. Correction values are calculated as a function of solar zenith angle from one year GOME $SO_2$ data, and also applied to the individual $SO_2$ slant columns.

## 5.3 AMF and VCD determination

For $SO_2$, the conversion from the slant column to a vertical column is complicated by the strong dependence of the Air Mass Factor on clouds, aerosols, and most importantly, on the *a priori* vertical profile of $SO_2$ in the atmosphere (see Figure 7). Especially the different emission sources of $SO_2$ (volcanic emissions at different altitudes, as well as anthropogenic pollution), should be taken into account in the AMF calculations. For the AMF calculations, a volcanic $SO_2$ profile is assumed with a predefined central plume height and a Gaussian $SO_2$ distribution around that central height. The $SO_2$ column for volcanic eruptions is computed for three different assumed $SO_2$ plume heights: 2.5 km above ground level, 6 km and 15 km. The first one represents passive degassing of low volcanoes, the

second one effusive volcanic eruptions or passive degassing of high volcanoes and the third one explosive eruptions. The influence of clouds on the AMF is treated as explained in Section 2.3.1, but in the case of $SO_2$, no ghost column $SO_2$ is derived meaning that just the "visible" $SO_2$ amount is retrieved.

A preliminary error estimate for the retrieved $SO_2$ column is given in Table 5.

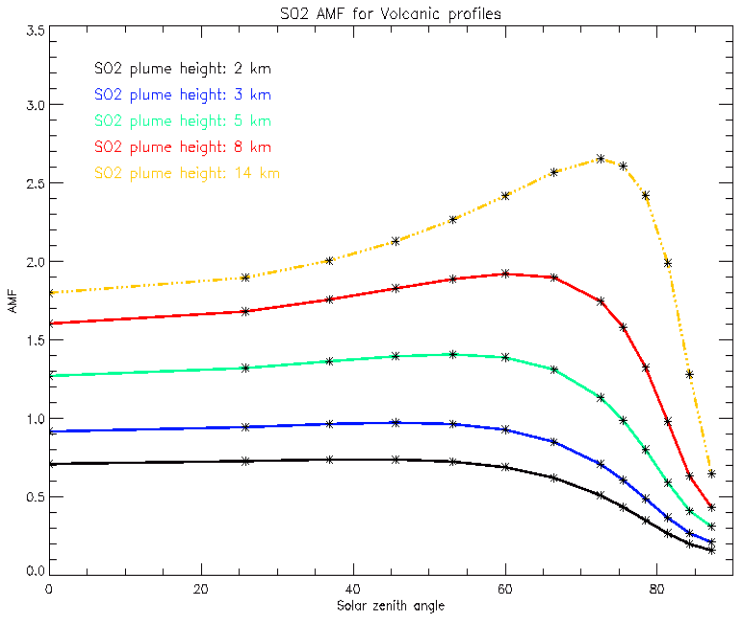



**Figure 7** Dependence of the $SO_2$ Air Mass Factor on the assumed volcanic plume height (2, 3, 5, 8 and 14 km). The AMF has been calculated as a function of solar zenith angle for clear-sky nadir viewing conditions and a surface albedo of 0.05. The assumed total $SO_2$ column is 3 DU.

**Table 5** Initial estimation of error sources for the total $SO_2$ column (volcanic $SO_2$).

| Error source | Percent error |
|---|---|
| **$SO_2$ slant column** | 30 – 50 |
| $SO_2$ absorption cross-sections | 5 – 10 |
| Atmospheric (effective) temperature | 5 – 10 |
| Other (Instrument signal-to-noise, Ozone abs. interference, Ring effect) | 20 – 30 |
| **$SO_2$ Air Mass Factor** | 20 – 50 |
| **$SO_2$ vertical column (accuracy)** | 50 – 100 |

# 6 OTHER TRACE GAS COLUMN PRODUCTS

The other trace gas column products include $H_2O$ total column product, the formaldehyde (HCHO) and OClO total column products, and the tropospheric ozone column product. A short description of the baseline retrieval algorithm for the formaldehyde and OClO column products are given below. Further algorithm development for the other trace gas column products will be done during the Continuous Development and Operations Phase (CDOP) of the O3M-SAF.

## 6.1 Formaldehyde column algorithm

The DOAS slant column fit of HCHO is performed in the 337.5-359.0 nm wavelength range [De Smedt et al., 2007]. The HCHO cross-section applied in the DOAS fit are from Cantrell et al. [1990], as well as the cross-sections of the interfering trace gases: ozone, $NO_2$, BrO, and $O_2$-$O_2$. A Ring reference spectrum is included as an additive fitting parameter, as well as a polynomial closure term of order 5.

The AMF is calculated with the LIDORT 2.2+ model for the fitting window mid-point (~348 nm), since HCHO is an optically thin absorber in this wavelength region. The AMF depends strongly on the vertical profile shape of HCHO in the troposphere, the surface albedo and the presence of clouds. *A priori* HCHO vertical profiles are provided by a tropospheric chemistry transport model (e.g. the MOZART-2 or IMAGES model). These CTMs can provide best-guess HCHO profiles on a monthly basis. The surface albedo is determined by combing the GOME LER (values at 335 nm) and TOMS LER databases.

## 6.2 OClO column algorithm

The DOAS slant column fit of OClO is performed in the 365-389.0 nm wavelength range. The OClO cross-section applied in the DOAS fit are from Kromminga et al. [1999], as well as the cross-sections of the interfering trace gases: ozone, $NO_2$, and $O_2$-$O_2$. A Ring reference spectrum is included as an additive fitting parameter, as well as a polynomial closure term.

OClO is determined by rapid photochemistry and is an indicator of chlorine activation during ozone hole conditions. Calculation of an AMF for OClO requires modelling of the radiative transfer through an inhomogeneous atmosphere due to the fast twilight photolysis; this will be a subject of study in the Continuous Development and Operations Phase (CDOP) of the O3M-SAF. In the pre-operational phase, only OClO slant columns at twilight conditions during ozone hole conditions will be provided.

# 7 CLOUD ALGORITHMS

GOME-2 footprints are comparably large and the retrieval is often affected by partially cloudy scenes. In such cases, the tropospheric contribution of trace species below clouds to the total content must be taken from climatological trace gas databases. Furthermore, clouds are usually opaque in the GOME spectral range and the cloud-top albedo is then taken as the lower reflecting boundary of the earth-atmosphere system, relative to the top of atmosphere. It is therefore vital to know the cloud fraction, the cloud-top height and cloud-top albedo parameters for providing reliable trace gas columns. These three parameters are needed for the computation of the different terms of Eq. (5).

Two algorithms OCRA and ROCINN [Loyola et al., 2007] are used for generating GOME-2 cloud information inputs for the trace gas column retrievals: OCRA for cloud fraction, and ROCINN for cloud-top height (pressure) and cloud-top albedo (optical thickness).

## 7.1 OCRA cloud fraction algorithm

The basic idea in OCRA (Optical Cloud Recognition Algorithm [Loyola and Ruppert, 1998]) is to break down each optical sensor measurement into two components: a cloud-free background and a residual contribution expressing the influence of clouds. The key to the algorithm is the construction of a cloud-free composite that is invariant with respect to the atmosphere, to topography and to solar and viewing angles. For a given location ($x$,$y$), we define a reflectance factor $\rho(x,y,\lambda)$ measured by the PMDs of GOME at wavelength $\lambda$ for the ground cover projection of the image. This reflectance is translated into normalized *rg*-color space via the relation:

$$r = \frac{\rho(x,y,\lambda_R)}{\sum_{i=R,G,B}\rho(x,y,\lambda_i)}, g = \frac{\rho(x,y,\lambda_G)}{\sum_{i=R,G,B}\rho(x,y,\lambda_i)} \cdot \tag{15}$$

with $R$ in [570-800 nm], $G$ in [400-570 nm] and $B$ in [300-400 nm]. If M is the set of $n$ normalized multi-temporal measurements over the same location ($x$,$y$), then a cloud-free (or minimum cloudiness) pixel $rg_{CF}$ in M is selected with the brightness criterion $\|rg_{CF} - w\| \geq \|rg_k - w\|$ for $k$ = 1,..,$n$, where $w$ = (1/3,1/3) is the *white point* in the *rg* chromaticity diagram. A global cloud-free composite is constructed by merging cloud-free reflectances $\rho_{CF}(\lambda)$ (corresponding to $rg_{CF}$) at all locations. The effective cloud fraction is determined by examining separations between measured reflectances and their cloud-free composite values:

$$c_f = \sqrt{\sum_i \alpha(\lambda_i)\max(0,[\rho(\lambda_i) - \rho_{CF}(\lambda_i)]^2 - \beta(\lambda_i))} \cdot \tag{16}$$

Scaling factors $\alpha$ ensure that the cloud fraction is mapped to [0,1], while offsets β account for aerosol and other radiative effects.

OCRA has been given an additional algorithm for the proper discrimination between clouds and Sun-glint - most of the GOME-2 orbits are affected by this phenomenon.

## 7.2 ROCINN cloud-top height and albedo algorithm

ROCINN [Loyola, 2004] is an algorithm based on $O_2$ *A* band reflectances from GOME: it delivers cloud-top height and cloud-top albedo. The independent pixel approximation is used; the cloud fraction $c_f$ derived from the OCRA algorithm is taken as a fixed input to the ROCINN algorithm. In the simulations, only attenuation through oxygen absorption of the direct solar beam and its reflection from

ground or cloud-top is considered. Molecular scattering, scattering and absorption by aerosols and diffuse surface reflection are neglected, as is absorption by oxygen within and below any clouds. Surfaces are assumed to be Lambertian reflectors. In this approximation, we need only consider reflactances along two photon paths through the atmosphere, and the forward model reflectivity is then:

$$R_{sim}(\lambda) = c_f \left\langle R(\lambda, \Theta, c_a, c_z) \right\rangle + (1 - c_f) \left\langle R(\lambda, \Theta, s_a, s_z) \right\rangle \tag{17}$$

Here, $\langle R \rangle$ denotes the convoluted reflectance to cloud-top or surface for path geometry $\Theta$ (solar zenith angle and line-of-sight angle), wavelength $\lambda$, surface albedo $s_a$ and cloud-top albedo $c_a$, and lower boundary heights $s_z$ (surface) and $c_z$ (cloud-top). Line-by-line transmittances must first be calculated using line spectroscopic information for the $O_2$ *A* band (taken from the HITRAN database), before convolution with the GOME-2 slit function. Quantities $s_z$ and $s_a$ are the surface height and albedo, taken from a suitable database and assumed known. ROCINN aims to retrieve cloud-top height $c_z$ and the cloud-top albedo $c_a$. Reflectance calculations based on Eq. (17) are used to create a complete data set of simulated reflectances for all viewing geometries and geophysical scenarios, and for various combinations of cloud fraction, cloud-top height and cloud-top albedo. High-resolution reflectances are computed with VLIDORT for the range 758-772 nm at resolution 0.002 nm before convolution. The inversion of Eq. (17) is performed using neural network techniques.

## 7.3 Cloud-top pressure and cloud optical thickness calculation

The cloud-top pressure for GOME scenes is derived from the cloud-top height provided by ROCINN and an appropriated pressure profile.

Cloud reflectivity is calculated with the libRadtran radiative transfer package by Mayer and Kylling [Mayer, 2005], as a function of cloud optical thickness, surface albedo, solar zenith angle, and viewing zenith and azimuth. An effective radius of 10 micron is assumed and the cloud is placed between 1 and 10 km. The midlatitude summer atmosphere is assumed as background atmosphere to include Rayleigh scattering. Cloud single scattering properties for 760 nm are calculated with Mie theory and the radiative transfer is solved with the plane-parallel discrete ordinate solver DISORT [Stamnes et al., 1988].

The reflectivity dependency on the cloud-top height is very small, for that reason a look-up table is created running libRadtran for a fixed cloud-top height of 4 km. A neural network is trained with this look-up table and the inverse problem is solved using the technique described in [Loyola, 2006]. Cloud optical thickness $\tau$ is computed as a function of $c_a$ cloud-top albedo, $s_a$ the surface albedo, $\theta_0$ the solar zenith angle, $\theta$ the satellite zenith angle, and $\phi$ the relative azimuth angle:

$$\tau = INV_{NN}(c_a, s_a, \theta_0, \theta, \phi) \tag{18}$$

The cloud optical thickness is computed using (18) taking as input the cloud-top albedo retrieved with ROCINN. For more details see [Loyola et al., 2009].

# 8 EXPECTED ACCURACY

The following table lists the GOME-2 total column trace gases and cloud products and estimated accuracy and precision.

**Table 5** Expected accuracy and precision of the GOME-2 total column trace gases and cloud products generated by the O3M-SAF

| Error source | Expected Accuracy | Expected Precision |
|---|---|---|
| Total ozone column | 3.6 - 4.3% (SZA < 80°)<br>6.4 - 7.2% (80°< SZA < 90°) | 2.4 - 3.3% (SZA < 80°)<br>4.9 - 5.9% (80°< SZA < 90°) |
| Tropospheric ozone column | 20-40% | 20-40% |
| Total $NO_2$ column | 5-10% (unpolluted conditions) | 3-10% (unpolluted conditions) |
| Tropospheric $NO_2$ column | 50-100% (polluted conditions)<br>> 100% (unpolluted) | 50-100% (polluted conditions)<br>> 100% (unpolluted) |
| Total BrO column | 20-50% | 10-50% |
| Total $H_2O$ column | 5-20% | 10-25% |
| Total $SO_2$ column | 50-100% (SZA < 70°)<br>> 100% (SZA > 70°) | 20-50% (SZA < 70°)<br>> 50% (SZA > 70°) |
| Total HCHO column | 50-100% (polluted conditions)<br>> 100% (unpolluted) | 20-50% (polluted conditions)<br>> 100% (unpolluted) |
| Total OClO column | 50-100% (SZA > 75°) | 20-50% (SZA > 75°) |
| Cloud fraction | < 10% | < 10% |
| Cloud-top height (pressure) | < 10% | < 10% |
| Cloud-top albedo (optical thickness) | < 10% | < 10% |

## REFERENCES

Aliwell, S. R., M. Van Roozendael, P. V. Johnston, A. Richter, T. Wagner, D. W. Arlander, J. P. Burrows, D. J. Fish, R. L. Jones, K. K. Tørnkvist, J.-C. Lambert, K. Pfeilsticker, and I. Pundt (2002), Analysis for BrO in zenith-sky spectra: An intercomparison exercise for analysis improvement, J. Geophys. Res., 107, D14, doi: 10.1029/2001JD000329.

Balis, D., M. Koukouli, D. Loyola and P. Valks (2007), Validation of GOME-2 total ozone products (OTO/O3, NTO/O3) processed with GDP 4.2, SAF/O3M/AUTH/GOME-2VAL/RP/01.

Balis, D., M. Koukouli, D. Loyola, P. Valks and N. Hao (2008), Second validation report of GOME-2 total ozone products (OTO/O3, NTO/O3) processed with GDP 4.2, SAF/O3M/AUTH/GOME-2VAL/RP/02.

Bass, A.M., and R.J. Paur (1985), The ultraviolet cross-sections of ozone, I, The measurements, in: Atmospheric Ozone, edited by C.S. Zerefos and A. Ghazi, pp. 606-610, D. Reidel, Norwell, Mass.

Bhartia, P. K. (2003), Algorithm Theoretical Baseline Document, TOMS v8 Total ozone algorithm, NASA.

Bodhaine, B., N. Wood, E. Dutton, and J. Slusser (1999), On Rayleigh optical depth calculations, *J.* Atmos. Ocean. Tech., 16, 1854-186.

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

Bruns M., H. Bovensmann, A. Richter, and J.P. Burrows (2003), A Stratospheric BrO climatology for the GOME-2 instrument, O3M-SAF Visiting Scientist Report, IUP University of Bremen, Feb. 2003.

Bucsela, E., E. Celarier, M. Wenig, J. Gleason, P. Veefkind, F. Boersma and E. Brinksma (2006), Algorithm for NO2 vertical column retrieval from the Ozone Monitoring Instrument, IEEE Trans. Geosci. Remote Sensing, 44, 1245-1258.

Burrows J.P., M. Weber, M. Buchwitz, V.V. Rozanov, A. Ladstaetter-Weissenmeyer, A. Richter, R. de Beek, R. Hoogen, K. Bramstadt, K.-U. Eichmann, M. Eisinger and D. Perner (1999a), The Global Ozone Monitoring Experiment (GOME): mission concept and first scientific results, *J. Atmos. Sci*, 56, 151-175.

Burrows, J.P., A. Richter, A. Dehn, B. Deters, S. Himmelmann, S. Voigt, J. Orphal (1999b), Atmospheric remote-sensing reference data from GOME - 2. Temperature-dependent absorption cross-sections of $O_3$ in the 231-794 nm range. J. Quant. Spectrosc. Radiat. Transfer, 61, 509-517.

Cantrell, Davidson, McDaniel, Shetter and Calvert (1990), Temperature-dependent Formaldehyde cross-sections in the near-ultraviolet spectral region, J. Phys. Chem., 94, 3902-3908.

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

De Smedt, I., M. Van Roozendael, T. Stavrakou, J.-F. Müller, R. van der A, H. Eskes (2007), Global observation of formaldehyde in the troposphere by satellites:GOME and SCIAMACHY results, Proc. Envisat Symposium 2007, Montreux, Switzerland, 23-27 April 2007.

Fortuin, J.P.F., and H. Kelder (1998), An ozone climatology based on ozonesonde and satellite measurements, J. Geophys. Res., 103, 31709-31734.

Fleschmann, O. C., Hartmann, M., Burrows, J. P. , Orphal, J., (2004), New ultraviolet absorption cross-sections of BrO at atmospheric temperatures measured by time-windowing Fourier transform spectroscopy, J. photochem. photobiol., A Chem., 168, no1-2, 117-132.

Grainger, J.F., and J. Ring (1962), Anomalous Fraunhofer Line Profiles, Nature, 193, 762.

Greenblatt, G.D., J.J. Orlando, J.B. Burkholder, and A.R. Ravishankara (1990), Absorption measurements of oxygen between 330 and 1140 nm, *J. Geophys. Res.*, 95, 18577-18582.

Gür, B., P. Spietz, J. Orphal and J. Burrows (2005), Absorption Spectra Measurements with the GOME-2 FMs using the IUP/IFE-UB´s Calibration Apparatus for Trace Gas Absorption Spectroscopy CATGAS, Final Report, IUP University of Bremen, Oct. 2005.

Herman, J.R, and E.A. Celarier (1997), Earth surface reflectivity climatology at 340 nm to 380 nm from TOMS data, J. Geophys. Res., 102, 28003-28011.

Koelemeijer, R.B.A., J.F. de Haan, and P. Stammes (2003), A database of spectral surface reflectivity in the range 335--772 nm derived from 5.5 years of GOME observations, J. Geophys. Res., 108, 4070, doi:10.1029/2002JD0024.

Kneizys F.X., E.P. Shettle, L.W. Abreu, J.H. Chetwynd, G.P. Anderson, W.O. Gallery, J.E.A. Selby, and S.A. Clough (1988), Users Guide to LOWTRAN 7, Air Force Geophysics Laboratory, Environmental Research Papers, No. 1010, AFGL-TR-88-0177.

Kromminga, H., J. Orphal, S. Voigt, and J.P. Burrows (1999) Fourier-transform-spectroscopy of symmetric chlorine dioxide (OClO), Proc. EC Advanced Study Course, Bergen, Norway.

Lambert, J.-C., and J. Granville, Harmonic climatology of stratospheric $NO_2$, BIRA-IASB, Brussels, 2004.

Lambert, J.-C., I. De Smedt, J. Granville, and P. Valks (2007), Initial validation of GOME-2 Nitrogen Dioxide columns (GDP 4.2 OTO/NO2 and NTO/NO2): March – June 2007, TN-IASB-GOME2-O3MSAF-NO-01-1/A.

Lambert, J.-C., G. Pinardi, N. Hao, and P. Valks (2008), ORR B – GOME-2 GDP 4.2 total NO2 (NTO/OTO) validation update and tropospheric NO2 validation set-up, TN-IASB-GOME2-O3MSAF-NO2-02_ORR-B_1/A.

Liu, X., M. Newchurch, R. Loughman, and P.K. Bhartia (2004), Errors resulting from assuming opaque Lambertian clouds in TOMS ozone retrieval, Journal of Quantitative Spectroscopy and Radiative Transfer, 85, 337-365.

Loyola, D., and T. Ruppert (1998), A new PMD cloud-recognition algorithm for GOME, ESA Earth Observation Quarterly, 58, 45-47.

Loyola, D. (2004), Automatic Cloud Analysis from Polar-Orbiting Satellites using Neural Network and Data Fusion Techniques, IEEE International Geoscience and Remote Sensing Symposium, 4, 2530-2534, Alaska.

Loyola D. (2006), "Applications of Neural Network Methods to the Processing of Earth Observation Satellite Data", Neural Networks, vol. 19, no. 2, pp. 168-177.

Loyola D., Thomas W., Livschitz Y., Ruppert T., Albert P.,and Hollmann R. (2007), Cloud properties derived from GOME/ERS-2 backscatter data for trace gas retrieval, IEEE Transactions in Geoscience and Remote Sensing, vol. 45, no. 9, pp. 2747-2758.

Loyola, D. (2007), A semi-transparent Lambertian cloud model for ozone retrieval, DLR presentation, September 2007.

Loyola D., W. Zimmer, S. Kiemle, P. Valks, T. Ruppert (2009), Product User Manual for GOME Total Column Productss of Ozone, Minor Trace Gases, and Cloud Properties, DLR/GOME/PUM/01, Iss./Rev. 2/A.

Loyola D., Thomas W., Spurr, R., B. Mayer (2009), Global patterns in daytime cloud properties derived from GOME backscatter UV-VIS measurements, International Journal of Remote Sensing, in press.

Malicet, J., D. Daumont, J. Charbonnier, C. Parisse, A. Chakir, and J. Brion (1995), Ozone UV spectroscopy. II. Absorption cross-sections and temperature dependence, J. Atmos. Chem., 21, 263-273.

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

Nüß, H., A. Richter, P. Valks and J. Burrows (2005), Improvement of the NO2 total column retrieval for GOME-2, O3M SAF Visiting Scientist Activity, Draft Final Report, IUP University of Bremen, Dec. 2005.

Richter, A., and J. Burrows (2002), Tropospheric $NO_2$ from GOME measurements, Adv. Space Res., 29, 1673-1683.

Richter, A., F. Wittrock and J. Burrows (2006), SO2 measurements with SCIAMACHY, in Proceedings of the First Conference on Atmosphereic Science, SP-628, ESA, Frascati, Italy, May 2006.

Rothman, L., et al. (2003), The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001, , J. Quant. Spectrosc. Rad. Transfer, 82, 5-44.

Siddans, R., B.J. Kerridge, B.G. Latter, J. Smeets and G. Otter (2006), Analysis of GOME-2 Slit function measurements, Algorithm Theoretical Basis Document, EUM/CO/04/1298/RM.

Spurr, R. J. D., T. P. Kurosu, and K. V. Chance (2001), A Linearized discrete Ordinate Radiative Transfer Model for Atmospheric Remote Sensing Retrieval, J. Quant. Spectrosc. Radiat. Transfer, 68, 689-735.

Spurr, R. (2003), LIDORT V2PLUS: a comprehensive radiative transfer package for nadir viewing spectrometers, remote Sensing of clouds and atmosphere, Proceedings SPIE conference 5235, Barcelona, Spain.

Spurr R.J.D., Van Roozendael M., Loyola D.G. (2004), "Algorithm Theoretical Basis Document for GOME Total Column Densities of Ozone and Nitrogen Dioxide. P/GDOAS: GDP 4.0", ERSE-DTEX-EOPG-TN-04-0007, Iss./Rev.:1/A.

Spurr, R., D. Loyola, W. Thomas, W. Balzer, E. Mikusch, B. Aberle, S. Slijkhuis, T. Ruppert, M. Van Roozendael, J.-C. Lambert, and T. V. Soebijanta (2005), GOME Level 1-to-2 Data Processor Version 3.0: A Major Upgrade of the GOME/ERS-2 Total Ozone Retrieval Algorithm, Applied Optics, 44, 7196-7209.

Stamnes, K. and Tsay, S.C. and Wiscombe, W. and Jayaweera, K (1988), A numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, Applied Optics, vol. 27,nr. 12, 2502-2509.

Thomas, W., Erbertseder, T., Ruppert, T., van Roozendael, M., Verdebout, J., Meleti, C., Balis, D., Zerefos, C. (2005), On the retrieval of Volcanic Sulfur Dioxide Emissions from GOME backscatter measurements, *J. Atm. Chem.*, 50, 295-320.

Van Geffen, J., M. Van Roozendael, M. Rix and P. Valks (2008), Initial validation of GOME-2 GDP 4.2 SO2 total columns – ORR-B, TN-IASB-GOME2-O3MSAF-SO2-01.

Van Roozendael, M., V. Soebijanta, C. Fayt, and J.-C. Lambert (2002), Investigation of DOAS Issues Affecting the Accuracy of the GDP Version 3.0 Total Ozone Product, in ERS-2 GOME GDP 3.0 Implementation and Delta Validation, Ed. J.-C. Lambert, ERSE-DTEX-EOAD-TN-02-0006, ESA/ESRIN, Frascati, Italy, Chap.6, pp.97-129.

Van Roozendael, M., and R.J.D. Spurr (2003), GOME Direct Fitting (GODFIT) Validation Report, ERS Exploitation AO/1-4235/02/I-LG.

Van Roozendael, M., J.-C. Lambert, R. J. D. Spurr, and C. Fayt (2004), GOME Direct Fitting (GODFIT) GDOAS Delta Validation Report, ERS Exploitation AO/1-4235/02/I-LG.

Van Roozedael, M., D. Loyola, R. Spurr, D. Balis, J.-C. Lambert, Y. Livschitz, P. Valks, T. Ruppert, P. Kenter, C. Fayt, and C. Zehner (2006), Ten years of GOME/ERS-2 total ozone data - The new GOME Data Processor (GDP) Version 4.0: I. Algorithm Description, J. Geophys. Res. 111, D14311, 10.1029/2005JD006375.

Van Roozedael, M. and N. Theys (2005), BrO column retrieval algorithms for GOME-2, O3M SAF Visiting Scientist Activity Proposal, Oct. 2005.

Van Roozedael, M., F. Hendrick and N. Hao (2008), Initial validation of GOME-2 GDP 4.2 BrO total columns – ORR-A3, TN-IASB-GOME2-O3MSAF-BrO-01.

Wahner, A., et al. (1988), Chem. Phys. Lett., 152, 507.

Wilmouth, D., T. Hanisco, N. Donahue, J. Anderson (1999), Fourier Transform Ultraviolet Spectroscopy of the transition of BrO, J. Phys. Chem., 103, 8935-8944.