GOFAP-KNMI-ASD-01

GOME Ozone Fast Delivery and value-Added Products Algorithm Specification Document

Version 1.0
March 4, 1999

Authors: A.J.M. Piters, P.J.M. Valks (KNMI), D.M. Stam (KNMI/VU)

Reviewed by: C. Zehner, ESRIN

Contents

Chapter 1

Introduction

The aim of the project "GOME Ozone Fast Delivery and value-Added Products" (GOFAP) is to develop a Fast Delivery (FD) processor retrieving total ozone columns from the Extracted GOME Instrument header data (EGOIs), which are used at ESRIN for the instrument performance and quality assurance monitoring. The project is performed within the framework of ESA's Data User Programme (DUP).

The EGOIs consist of selected windows of the raw GOME data, among which is an ozone retrieval window (325–335 nm). The FD-processor performs a simple calibration on this raw data, and uses the differential spectrum (Earth spectrum divided by the Sun spectrum) to retrieve a slant ozone column density, i.e. the amount of ozone molecules per square centimeter along the light path. A look-up table of pre-calculated Air Mass Factors (AMFs) is used to derive the vertical column density.

The retrieval of total ozone columns from the EGOI data takes in the order of a few seconds per GOME orbit on a workstation. The collection on board, downloading to the ground station, extraction of the EGOI data, and transport to ESRIN and from there to KNMI takes considerably more time. However, the total time delay before the total ozone column is available on the web is well within 3 hours after observation (hence near-real time). Therefore it is possible to use this level 2 data into on-line numerical weather prediction systems, e.g. into the ECMWF weather prediction model.

Furthermore, the near-real time ozone columns are continuously assimilated into a simple off-line transport model, which makes use of the ECMWF analysed and forecast wind, to construct up-to-date ozone maps (level 4 data). These maps can be used for, e.g., measurement campaign preparations.

This document describes the algorithms used for calibration (level 0 →'1' processing) and for the retrieval of total ozone columns (level '1'→ 2 processing). Chapter 4 gives a description of the level 2 → 4 processor, and Chapter 5 describes the operational environment. The calibration contains only those steps which influence the retrieved ozone column value, therefore the result of this calibration is called level '1' (not completely calibrated) instead of level 1.

Chapter 2

Level 0→'1' Algorithms

2.1 Introduction

The only product retrieved from the raw GOME data is total column of ozone. Therefore all algorithms are optimised for the retrieval of this product. The calibration of the raw data only needs those steps which influence the retrieved ozone column values.

From the EGOI data a level '1' product is derived which, after dividing the Solar and Earth level '1' products (differential spectrum), can be used to derive an accurate slant column, using the DOAS method. After the quality checking of each data record (see Section 2.3), the following calibration steps are performed:

- The correction for fixed pattern noise and leakage current (Section 2.4). This is especially important for low intensity measurements.

- The spectral calibration (Section 2.6). This should be performed, because the refence spectra are given as a function of wavelength.

The following calibration steps are neglected:

- The pixel-to-pixel gain correction. This correction is expected not to be important for the slant column retrieval, because it is a high frequency effect (small wavelength scales). No large effect on the scale of the ozone absorption features is expected.

- The stray-light correction. This correction, especially for uniform stray light, might be important, because it is a subtraction, like the leakage current and fixed pattern noise correction. It therefore will effect the ratio of radiance and irradiance. However, the expected influence on the slant column retrieval is expected to be in the order of 0.2%, so this correction is not performed.

- The radiometric calibration. This calibration is expected not to have large impact on the slant column. This is because the slant column directly depends on the ratio of radiance and irradiance, which is equal to the ratio of the uncorrected spectra multiplied by the BSDF of the diffuser. The radiance response function vanishes in the calculation of the ratio. The BSDF of the diffuser is a slowly varying function of wavelength and neglecting this factor will be corrected by the polynomial in the DOAS fit.

- The polarization correction. Leaving out this correction results in an expected error less than 0.1%.

2.2 EGOI description

A GOME science packet is generated by ERS-2 every 1.5 seconds. On-ground each GOME science packet is extracted in the form of a single EGOC product, consisting of a main product header (176 bytes), a single product header (304 bytes) and a data set record (8004 bytes). An EGOI (Extracted GOME Instrument header) product is generated by extracting from a variable number of EGOC products a maximum of 10 specific data windows from the data set records, and adding one specific product header (40 bytes) and one main product header (176 bytes) for one EGOI. Typically three EGOIs per orbit are generated. The EGOIs are transmitted from the ERS ground stations at Gatineau, Kiruna and Maspalomas via land lines.

The main product header consists of product identifier information (which satellite, which instrument, which ground station, etc), information on the location of the satellite, product confidence flags, and information on the data record lengths.

The specific product header gives the byte positions and lengths of the specific data windows extracted from the original EGOC products. The maximum number of windows are 10, and the maximum total number of bytes in these window are 1026.

At the time of writing the EGOI data set records consist of the following windows:

1. Auxiliary Data: These data form a contiguous block at the beginning of each GOME science packet. They consist of information on instrument status, integration times, temperatures, etc. The auxiliary data also contain samples of, on average, three pixels per detector band, plus all PMD measurements.
2. Stray Light: These are six pixels used for measuring stray light (three in Channel 1, and three in Channel 2)
3. Lamp Lines 1: This window consists of the data from pixels 411 to 446 in Channel 2 ($\sim$ 337–341 nm), which includes some clear calibration lamp lines at 337.919 nm, 339.3775 nm and 340.9113 nm.
4. Lamp Lines 2: This window consists of the data from pixels 233 to 279 in Channel 2 ($\sim$ 317–322 nm), which includes a relatively broad calibration lamp line at 321.9122 nm.
5. DOAS ozone window: This is the main window for the DOAS fit, it consists of pixels 302 to 390 in Channel 2 ($\sim$ 325–335 nm).
6. Chlorophyll: This window consists of pixels 398 to 400 of Channel 4 ($\sim$ 665 nm), it is used together with the next window to derive a normalized difference vegetation index.
7. 780.0 nm window: This window consists of pixels 955 to 957 of Channel 4 ($\sim$ 780 nm).
8. ICFA window: This is the largest window from $\sim$ 758 nm to $\sim$ 778 nm, containing an oxygen line. This window can be used to obtain cloud information.
9. Channel 1: This window contains five pixels from Channel 1: 656–660 ($\sim$ 283 nm).
10. Channel 3: This window contains three pixels from Channel 3: 764–766 ($\sim$ 555 nm).

In total the length of each EGOI data set record is 974 bytes.

For the Fast Delivery processor described in this document, currently only the EGOI windows 1, 3, 4, and 5 are used.

2.3 Quality checking / identification of records

The EGOI products are checked for transport and instrumental problems. First, the instrument and integration status of each EGOI product Data Set Record (DSR) is checked for the following:

- word 12, bit 0, Operating mode should be 'normal'
- word 12, bit 2, Last DEU command should be 'valid'
- word 12, bit 3&4, Last S.U.-DDHU protocol message should be 'ACK' (command OK)
- word 12, bit 5, Auxiliary data external stimulus should be 'disabled'
- word 12, bit 6&7, FPA and RTM latch-up protection should be 'OFF'
- word 12, bit 8&9 and 11-13: LED's drivers, Scan Mirror heaters and ADC calibration should all be 'OFF'
- word 13, bit 0–7, Coolers 1-4 should be 'ON at 235 K'
- word 13, bit 8–15, FPA temperatures 1-2 channel 1-4 should be 'enabled'
- word 14, bit 8–15, Integration times of bands 2A to 4 should be completed nominally in the considered ground pixel

If one of these criteria is not fulfilled, the DSR will not be processed. Additionally the following checks are performed

- FPA temperatures should be 235 ± 1 K
- Predisperser temperature should be 282 ± 10 K
- Calibration lamp voltage should be between 180 and 210 V
- Calibration lamp current should be between 8 and 11 mA

Each data record is identified as one of the following measurement types:

- Earth-shine measurement: if the cover is closed (word 12, bit 14&15), the calibration unit is off (word 12, bit 10), and the first and last scan mirror recordings (word 30&90) are less than 40° off nadir direction.
- Sun measurement: if the cover is open, the calibration unit is off, and the first and last scan mirror recordings are less than 5° off the 180° direction.
- Dark measurement: if the cover is closed, the calibration unit is off, and the first and last scan mirror recordings are less than 5° off the 100° direction.
- Calibration lamp measurement: if the cover is closed, the calibration unit is on and the first and last scan mirror recordings are less than 5° off the 165° direction.
- Other measurement: in all other cases, these measurements are not processed.

If one of the 20 frames of a GOME science packet is not completed during downloading the end of the frame will be padded with 'FFFF Hex'. If the complete frame is missing it will be padded with 'BBBB Hex'. Frame 6, which includes the DOAS ozone window and one of the lamp line windows, is checked for this padding and for saturation (signal larger than 55500 for one or more detector pixels). Frame 5 is checked for padding and saturation for Calibration lamp measurements only. In the case of padding or saturation the DSR is not processed.

If the coadding mode is on (word 174, bit 5,9,13&15 are 'ON', the other bits are 'OFF', and the integration times of band 2B to 4 are all 0.375 s), the signals of the Earth-shine measurement, the Sun measurement, and the Dark measurement are divided by 4. If, in addition, a division by 2 in band 2B was performed on board to prevent numerical saturation (word 14, bit 3 is 'ON'), the signals in band 2B (which includes the ozone window) are divided by 2 instead of 4.

2.4 Correction of leakage current and fixed pattern noise

The dark measurements consist of both the fixed pattern noise and the integration time dependent leakage current. The correction is performed in one algorithm step. The dark measurements are collected and averaged for the same band 2B integration time. The averaging is done when there are 2 or more dark measurements with this integration time in one EGOI file (one third of an orbit) as follows:

$$D_t(i) = \frac{1}{n_t} \sum_{j=1}^{n_t} d_t(i,j)\,, \tag{2.1}$$

where $D_t(i)$ is the average dark signal in detector pixel i for integration time t, n_t is the number of dark measurements with integration time t, and $d_t(i,j)$ is the signal of the j^{th} dark measurement with integration time t in detector pixel i. The standard deviation is calculated:

$$\sigma_t(i) = \sqrt{\frac{1}{n_t - 1} \sum_{j=1}^{n_t} (d_t(i,j) - D_t(i))^2}\,. \tag{2.2}$$

If individual dark measurements $d_t(i,j)$ exist which deviate more than $3\sigma_t(i)$ from the average value $D_t(i)$, the calculation of the average and standard deviation is done again without these deviating measurements. The average dark signals and standard deviations are stored so that they can be used for processing other EGOI files.

The correction is performed to Earth-shine and Sun measurements, by subtracting the average dark signal with the same integration time as the Earth-shine or Sun measurement. If there is no dark signal with this integration time in the same EGOI file, a stored dark signal with the same integration time t is used (closest in observation time to the Earth-shine or Sun measurement):

$$I(i) = S(i) - D_t(i)\,; \tag{2.3}$$

$$\Delta I(i) = \sqrt{\frac{cI(i)}{N_e} + \sigma_t^2(i)}\,, \tag{2.4}$$

where $S(i)$ is the measured signal, $I(i)$ is the dark signal corrected signal, $\Delta I(i)$ is the uncertainty in $I(i)$, and N_e is the number of electrons per binary unit ($N_e = 937$). $c = 0.25$ for co-adding mode, and $c = 1$ for non-coadding mode. If $I(i)$ is negative, it will be made 0. For proper calculation of the surface reflectivity, the signal $I(i)$ and its uncertainty $\Delta I(i)$ are divided by the integration time. If the co-adding mode is on, they are also divided by 4.

2.5 Calculating the differential spectrum

For each EGOI file all calibrated Sun measurements are collected. The sun spectra are added, and divided by the total integration time. This last division has no effect on the DOAS fit, its main purpose is to limit the numerical values in order of magnitude, and for proper calculation of the surface reflection. The average Sun spectrum is stored with the estimated errors so that they can be used for processing other EGOI files.

Each Earth-shine measurement is divided by the average Sun measurement closest in time. This division is done pixel-by pixel, so differences in wavelength between the two spectra are neglected.

2.6 Wavelength calibration

Three lamp lines are used for the wavelength calibration: $\lambda_1 = 321.9122$ nm, $\lambda_2 = 332.4692$ nm, and $\lambda_3 = 337.9190$ nm, which are approximately at detector pixel positions $p_1 = 276$, $p_2 = 368$, and $p_3 = 416$.

The Calibration lamp measurements are collected and added per EGOI file. First, it is checked whether the lamp lines are approximately at the detector pixel position they are expected to be: in the detector pixel ranges $[p_k - 2, p_k + 2]$ the maximum signal should be at p_k. If this is not the case, these calibration measurements are not used. Then the three lamp lines are fitted with a parabola, using only three detector pixels per line $(p_k - 1, p_k, p_k + 1)$. The pixel position of line k then is (using $y_1 = I(p_k) - I(p_k + 1)$; $y_2 = I(p_k + 1) - I(p_k - 1)$; and $y_3 = I(p_k) - I(p_k - 1)$):

$$x_k = p_k + \frac{y_2}{y_1 + y_3} \,. \tag{2.5}$$

The error in this position is approximated by:

$$\Delta x_k = \frac{\sqrt{\sum_{k=1}^{3} y_k \Delta I^2(p_k)}}{y_1 + y_3} \,. \tag{2.6}$$

A straight line $x = a + b\lambda$ is fitted through these three points (λ_k, x_k), by minimizing

$$\chi^2(a,b) = \sum_{k=1}^{3} \left(\frac{x_k - a - b\lambda_k}{\Delta x_k} \right)^2 \,. \tag{2.7}$$

The wavelength calibration then consists of attaching to all detector pixel positions p of the ozone window a wavelength λ by:

$$\lambda = \frac{p - a}{b} \,. \tag{2.8}$$

Chapter 3

Level '1'→2 Algorithms

3.1 DOAS slant fit

A linear fit is performed on the negative natural logarithm of the differential spectrum, $y_i = -\ln(I_e(\lambda_i)/I_s(\lambda_i))$, by minimizing

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - P_3(\lambda_i) - a\sigma_R(\lambda_i) - \mathrm{SCD}\sigma_{O_3,T_{\mathrm{eff}}}(\lambda_i)}{\Delta y_i} \right)^2 , \tag{3.1}$$

where SCD is the fitted slant column density and $P_3(\lambda_i)$ is a third order polynomial. $\sigma_R(\lambda_i)$ is a theoretical Ring spectrum calculated by Chance and Spurr [1997], and $\sigma_{O_3,T_{\mathrm{eff}}}(\lambda_i)$ is the GOME-FM Ozone absorption spectrum at the effective temperature T_{eff}. The spectra are interpolated to the detector pixel wavelengths λ_i; the determination of the effective temperature is described at the end of this section. The minimization of χ^2 is done using singular value decomposition. χ^2 above can be written as:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{N} \left[\frac{y_i - \sum_{k=1}^{6} a_k X_k(\lambda_i)}{\Delta y_i} \right]^2 \\ &= |\mathbf{A} \cdot \mathbf{a} - \mathbf{b}| , \end{aligned} \tag{3.2}$$

where $A_{ij} = X_j(\lambda_i)/\Delta y_i$, and $b_i = y_i/\Delta y_i$, and a_1 to a_4 represent the polynomial parameters ($X_1(\lambda_i) = 1, \cdots, X_4(\lambda_i) = \lambda_i^3$), a_5 represents the Ring parameter ($X_5 = \sigma_R$), and a_6 represent the slant column densities SCD ($X_6 = \sigma_{O_3,T_{\mathrm{eff}}}$). The matrix $\mathbf{A}$ is written as a product of a column-orthogonal matrix $\mathbf{U}$, a diagonal matrix $\mathbf{W}$ with poitive or zero elements, and the transpose of an orthogonal matrix $\mathbf{V}$:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^T . \tag{3.3}$$

The variance in the estimate of a_6 = SCD then is:

$$\sigma^2(a_j) = \sum_{i=1}^{M} \left(\frac{V_{ji}}{w_i} \right)^2 , \tag{3.4}$$

with $w_i = W_{ii}$. The error in the slant column is estimated to be:

$$\Delta\mathrm{SCD} = \sigma(a_6)\sqrt{\frac{\chi^2}{N-6}} \tag{3.5}$$

The linear fit is improved by shifting the Ring and ozone absorption spectra, to allow for uncertainties in the wavelength calibration. The optimum shift values are found by non-linear fitting (steepest descent method). This is an iterative procedure by which the linear fit for the slant column density SCD is followed by the non-linear fitting for the shift values.

3.1.1 *Determination of the effective temperature*

A temperature climatology from ECMWF as a function of month, height (up to 50 hPa), latitude and longitude, is zonally averaged and extended to 0.3 hPa by a smooth transition to the MODTRAN climatology between 50 and 0.3 hPa. With this climatology, an effective temperature climatology is determined for the DOAS slant column fit. This effective temperature is a function of month, latitude and longitude. To derive the effective temperature we use simulations of the differential spectrum.

The differential spectrum will be

$$\ln I_{\text{out}}(\lambda) - \ln I_{\text{in}}(\lambda) = -\int \rho(z)\,\sigma(T(z),\lambda)\,dz + P(\lambda)$$

where ρ is the ozone density, and σ is the ozone cross-section and $P(\lambda)$ is a slowly varying function of λ.

The DOAS method assumes:

$$\ln I_{\text{out}}(\lambda) - \ln I_{\text{in}}(\lambda) = -\sigma(T_{\text{eff}},\lambda)\int \rho(z)\,dz + P(\lambda)$$

where $\int \rho(z)dz$ is the slant column. We define T_{eff} as the temperature for which $\sigma(T_{\text{eff}},\lambda)$ fits best the function

$$\frac{\int \rho(z)\,\sigma(T(z),\lambda)\,dz}{\int \rho(z)\,dz}$$

using a least squares fit over the DOAS spectral window.

3.2 Viewing geometry

ESA's ERS orbit propagator is used to find the geolocation for each ground pixel. The following information is extracted from the orbit propagator at the start time of integration of each ground pixel:

- λ_0: geographic space craft and nadir longitude
- ϕ_0: geodetic space craft and nadir latitude
- $r_{\|}$: nadir radius of curvature parallel to meridian
- $r_{\perp}$: nadir radius of curvature perpendicular to meridian
- T_{lst}: true local solar time
- δ: true sun's declination
- z: geodetic altitude of space craft above nadir
- v_{ϕ}: Northward ground velocity component with respect to Earth
- v_{λ}: Eastward ground velocity component with respect to Earth

To calculate the longitude and latitude of the two corners of a ground pixel, at a specific viewing angle γ and a viewing direction ψ (with respect to North), the angle is projected along and perpendicular to the meridian:

$$\begin{aligned} \gamma_x &= \arctan(\sin\psi\tan\gamma) && (3.6)\\ \gamma_y &= \arctan(\cos\psi\tan\gamma) && (3.7) \end{aligned}$$

The corresponding angles along the Earth's surface are:

$$\delta_x = \arcsin\left(\left(1+\frac{z}{r_\perp}\right)\sin\gamma_x\right) - \gamma_x \quad (3.8)$$

$$\delta_y = \arcsin\left(\left(1+\frac{z}{r_\parallel}\right)\sin\gamma_y\right) - \gamma_y \quad (3.9)$$

The new latitude and longitude are:

$$\phi = \arcsin\left(\sin\left(\phi_0+\delta_y\right)\cos\delta_x\right) \quad (3.10)$$

$$\lambda = \lambda_0 + \arcsin\left(\frac{\sin\delta_x}{\cos\phi}\right) \quad (3.11)$$

The viewing direction ψ is calculated from v_ϕ and v_λ: $\sin\psi = v_\phi/|v|$ and $\cos\psi = -v_\lambda/|v|$, the line-of-sight zenith angle is given by the first scan-mirror position. The latitude and longitude calculated here give the geolocation of the centre of the field of view at the beginning of the measurement. To account for the extent of the ground pixel, the instantaneous field of view is taken to be $\epsilon = \pm 1.5°$ in the along-track direction. The calculation above is repeated for

$$\gamma_x = \arctan(\cos\psi\tan\epsilon) \quad (3.12)$$

$$\gamma_y = -\arctan(\sin\psi\tan\epsilon) \quad (3.13)$$

$$\phi_0 = \phi \quad (3.14)$$

$$\lambda_0 = \lambda \quad (3.15)$$

to get the latitude and longitude of the two corner points. The corner points at the end of the ground pixel are taken to be equal to the corner points at the beginning of the next ground pixel. If the next ground pixel is not available, the geolocation is calculated for an extrapolated scan-mirror position γ:

$$\gamma = \gamma_1 + \frac{16}{15}(\gamma_2 - \gamma_1) \quad (3.16)$$

where γ_1 is the first scan-mirror position and γ_2 is the 16^{th} scan-mirror position.

3.3 Surface reflectivity

The surface reflectivity R is calculated from the 7 detector pixels at the long wavelength end of the second lamp-line window ($\sim$ 340 nm). It is defined as

$$R = \frac{I}{\mu_0 F} \quad (3.17)$$

where I is the radiance at 340 nm, πF is the irradiance, $\mu_0 \equiv \cos\theta$, θ is the solar zenith angle. The surface reflectivity will be used to derive cloud cover information. The reflectivity is approximated by the ratio of the level '1' earth spectrum and the level '1' solar spectrum integrated over these 7 detector pixels, and multiplied by a constant, which accounts for the sensitivity of the diffuser effecting the solar radiation. The constant is approximated by BSDF_0 (≈ 0.14) as determined by TPD.

3.4 Cloud information

The cloud cover fraction is retrieved from the measured surface reflectivity R by using climatological values of the surface albedo and calculated surface reflectivities for a cloud free and cloudy situation. The surface albedo values are taken from the monthly reflectivity climatology derived from 340-380 nm TOMS data in which sea-ice and snow cover is taken into account (for a description of the TOMS reflectivity climatology see Herman and Celarier [1997]). The surface reflectivities for a cloud free and cloudy situation are retrieved from a look-up table calculated with the Doubling Adding Model KNMI (DAK). The look-up table contains the reflectivities for 340 nm, calculated for different viewing scenarios and surface albedo values.

The cloud cover fraction f is given by:

$$f = \frac{R - R_s}{R_c - R_s} \tag{3.18}$$

where R is the measured surface reflectivity, R_s is the calculated surface reflectivity for a cloud free situation using the TOMS surface reflectivity climatology and R_c the calculated reflectivity for a cloudy situation assuming a cloud albedo of 0.8. (The approximation of reflection by clouds by that of a Lambertian surface works best when a high cloud albedo is used, because in that case the retrieval assumption that no radiation penetrates the cloud is most accurate, see Koelemeijer and Stammes [1999]).

The cloud top height is assumed to be known a-priori, and is taken from the International Satellite Cloud Climatology Project (ISCCP) (Rossow and Garder, 1993).

3.5 Air Mass Factor calculation

The Air Mass Factors are calculated off-line. A table is generated with different Air Mass Factors for different viewing scenarios. The AMFs are calculated at 325 nm. One set of air mass factors, for 15 different solar zenith angles, 9 surface heights, 4 surface albedos, 12 months, and 17 latitude bands, is calculated for each pixel-type/swath-width combination (see Table 3.1).

The pixel types considered are: East, Centre and West with swath widths of 960 km, 240 km, and 0 km (i.e. no scanning). The effective viewing angle for each pixel-type is calculated as follows: ten AMFs are calculated for an 960 km swath pixel-type, varying only the viewing angle (ten equally spaced angles over the ground pixel) and the azimuth angle (when the viewing angle passes the nadir viewing direction), the average AMF is calculated, and the viewing angle which corresponds to this average AMF (determined by interpolation) is taken as the effective viewing angle for this pixel-type. All AMF calculations are then performed using only this single effective viewing angle. The resulting viewing angles for the East, Centre and West pixel in a 960 km swath are: 21.5°, 3.3°, and −21.8°. In the current algorithm, the effective viewing angles of the 960 km swath Centre pixel are used for the 240 and 0 km swath-width pixels. It is planned that in the future, the AMF table will be extended with effective viewing angles for the Backscan pixels and for the 240 and 0 km pixel-type/swath width combinations.

Since the orbit of the ERS-2 is fixed with respect to the Sun, the azimuth angle $\phi - \phi_0$ (plus or minus 180°)can be approximated by:

$$\phi - \phi_0 = \arccos\left(\frac{\sin\alpha}{sin\theta}\right), \tag{3.19}$$

where θ is the solar zenith angle and α is the angle between the direction to the Sun and the plane of the satellite orbit.

The AMFs are calculated with the Doubling Adding model KNMI (DAK). In order to be able to calculate AMFs at high solar zenith angles, the original plane-parallel approximation is changed into a quasi-spherical approximation.

3.6 Vertical column calculation

The ozone vertical column is calculated by dividing the slant column by the appropriate Air Mass Factor. However, the presence of clouds complicates the calculation of the vertical column. Clouds are approximated as reflecting surfaces; no radiation is penetrating through the cloud. Therefore, the calculated airmass factor for the cloudy sky pertains to the atmosphere above the cloud and it is necessary to calculate two Air Mass Factors: one down to the ground surface (AMF_{clear}) and one down to the cloud top

Table 3.1: All parameters for which AMFs are calculated. Note that in the different pixel types/swath widths the AMFs are averaged over 10 viewing angles. The azimuth angle is approximated by assuming a simple relationship with solar zenith angle (see text).

pixel type/swath width (km)	East, Center, West (for 960 km swath width)
solar zenith angle	15,20,25,30,35,40,45,50,55,60,65,70,75,80,85
surface height (km)	0,1,2,3,4,6,8,10,12
surface albedo	0,0.2,0.8,1
latitude	-80,-70,-60,-50,-40,-30,-20,-10,0,10,20,30,40,50,60,70,80
month	Jan,Feb,Mar,Apr,May,Jun,Jul,Aug,Sep,Oct,Nov,Dec

(AMF_{cloud}). The total Air Mass Factor is given by:

$$AMF_{tot} = fAMF_{cloud} + (1-f)AMF_{clear} \tag{3.20}$$

A correction is made to account for the ozone amount below the cloud top (the so-called ghost vertical column, GVC). The ghost column is estimated using the climatological ozone profiles from Fortuin and Kelder [1998], with the cloud top height taken from the ISCCP database. The total ozone vertical column is given by:

$$VCD_{tot} = \frac{SCD + f \cdot GVC \cdot AMF_{cloud}}{AMF_{tot}} \tag{3.21}$$

The error in the ozone vertical column is estimated from the error in the slant column ΔSCD (neglecting the error in the Air Mass Factors):

$$\Delta VCD_{tot} = \frac{\Delta SCD}{AMF_{tot}} \tag{3.22}$$

Since the error in the AMF is neglected in this estimate, the real error in the vertical column is generally larger. However, the error in the AMF is not easily quantified because it depends on many variables, such as the presence of aerosols, the uncertainty in the climatological ozone profiles and the errors in the surface albedo and cloud parameters.

Chapter 4

Level 2→4 Algorithms

For the production of level 4 assimilated ozone maps we use the technique of 4D variational data assimilation. (An overview of the data assimilation technique can be found in Eskes et al., 1998; a detailed description of the 4D-VAR program and the error estimation approach can be found in Eskes et al., 1999). ECMWF forecast windfields are used for the description of the ozone transport. The Variational Assimilation Model KNMI (VAMK) processes one day of data in approximately 20 minutes (depending on the computer). The level 4 product (synoptic total ozone fields for 12 GMT) is updated every time a new orbit of level 2 data is available.

Data assimilation is a statistical technique to combine measurements of variables, such as for instance the atmospheric temperature, pressure or ozone concentration, with models that describe the time evolution of these variables. The error statistics of both the measurements and the model is used to find the best guess of the actual field. When the measurements are very accurate the analysed field will reproduce the observations, and when the model produces a good forecast in comparison to the accuracy of the measurement, the field variables will be adjusted only slightly. Both the model and the measurements can profit from the assimilation procedure. Modelling problems can be identified, as they will manifest themselves in the form of large forecast errors. Similarly the assimilation may provide information about the quality of the instruments and retrieval code (Piters et al., 1996).

Most assimilation techniques analyse the measurements sequentially. When a measurement (a dot in the figure) is encountered a new analysed field value is computed (the field value, the solid line, shows a jump). This will in general lie somewhere in between the observation and the forecast. After the analysis the field is propagated forward in time until the next measurement, and the procedure is repeated.

Here a different approach is used, namely variational data assimilation, also known as "4D-VAR" (see Figure 4.1) The procedure starts by choosing a time interval [0,T]. In our case T=24 h. First a first-guess model trajectory (dashed line) is computed by running the model from time 0 to time T. During this run the mismatches (arrows) between the model and all measurements in the interval are recorded. Then a second model run is performed (the "adjoint" model run) which transports the mismatches (jumps in the lower solid line) backwards in time. This results in a field of errors at time t=0 ("gradient") which

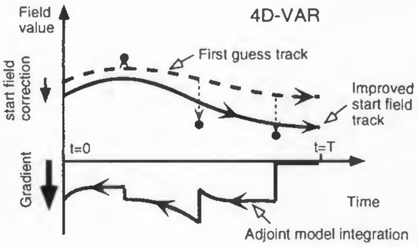

Figure 4.1: Schematic explanation of the variational data assimilation approach, 4DVar.

is then used to improve the ozone field at time t=0. From this new start field an improved trajectory is computed by the model (solid line). Typically the above procedure is repeated several times to converge to the most likely actual ozone distribution. The evolution of the ozone field in the interval is described purely by the model, and no jumps occur. Note that all measurements in the interval [0,T] are taken into account in a single analysis. This results in a smoother behaviour of the fields. For instance the first measurement would suggest a positive correction, but the actual correction made is negative, because the measurements at later times indicate a low ozone value. An important advantage of the variational approach is the ability to incorporate future measurements in the analysis. As interval we use t=0 GMT to t=24 GMT, but we will show the field at t=12 GMT. This field reflects 12 hours of past and 12 hours of future measurements. In this sense there are twice as many measurements available compared to the sequential technique, improving the accuracy of the assimilated field.

Chapter 5

System Design Overview

5.1 Introduction

In this chapter the operational environment of the FD system is described. The operational environment takes care of the acquisition of the EGOI data, the processing to level 2 and 4 data and the distribution of the data to the users and the FD WWW-site. This chapter is organized as follows: first the operational system and the technical environment are described, then the product distribution is presented. An overview of the FD WWW-site and the user service is given. Finally, the near-real time performance of the system is discussed.

5.2 System description

The FD system performs the following four functions: acquisition, processing, distribution and monitoring. The input and output for these functions and their place in the system are shown in figure 5.1.

The acquisition function takes care of the transport of EGOI files from an ESRIN ftp server to the processing environment at KNMI, and of the acquisition of the ECMWF wind fields. The processing function consists of two independent parts (not indicated separately in the figure): a level $0 \rightarrow 2$ processor and a level $2 \rightarrow 4$ processor. The EGOI files are input for the level $0 \rightarrow 1$ processor, the wind fields and the level 2 data are input for the level $2 \rightarrow 4$ processor. Auxiliary data (like an AMF look-up table) is stored in the processing environment. The output of the processor function are level 2 files (ASCII format, containing geo-locations, observation times and ozone vertical column densities with uncertainties), level 4 files (ASCII format, containing gridded (latitude/longitude) ozone column densities for one specific time), and images (GIF format, showing the ozone distribution over the globe). The distribution function takes care of the distribution of the data to the users and is described in Section 5.4. The monitoring function monitors the performance of the three previously described functions.

5.3 Operational Environment

The development, integration, testing and operation of the FD system is done on a UNIX Silicon Graphics workstation at KNMI. The system is controlled by several UNIX scripts, called deamons, which are dealing with specific functions in the FD system. The main deamons for the FD system are described below.

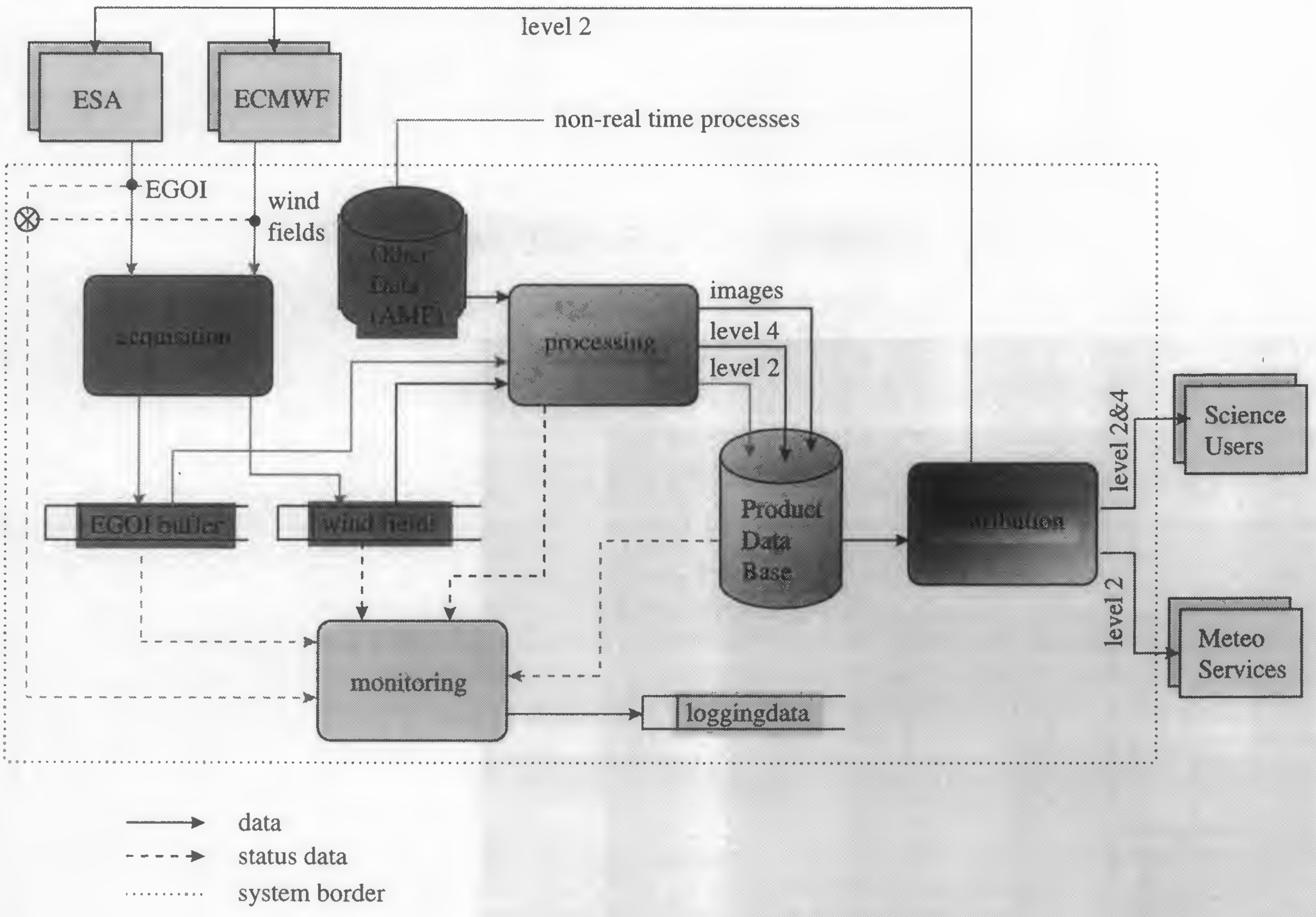

Figure 5.1: System concept of the Fast Delivery system

5.3.1 *ftpfd*

This is the ftp acquisition deamon. It checks every three minutes if there are new EGOI-files present on the ESRIN ftp-server. The new EGOI-files are transferred via FTP directly to the KNMI workstation. The ftp-transfer time of the EGOI files is logged and if there were no transfer errors, the EGOI-files on the ESRIN ftp-server are deleted.

5.3.2 *process2fd*

This is the level 0 → 2 processing deamon. It runs the level 0 → 2 processor for each new EGOI-file. The processor creates level 2 ASCII data files for further distribution and processing, and data files for validation purposes. The retrieval time delay is logged for the near-real time statistics.

5.3.3 *plotlv2fd*

This deamon creates near real time GIF images of the level 2 data for the past 24 hours.

5.3.4 *process4fd*

This deamon runs the data assimilation model VAMK with the level 2 data. The model creates an assimilated level 4 ozone fields for 12^h of the present day. GIF images and ASCII data files of the assimilated ozone fields are created for World, North Pole and South Pole projections.

Apart from the main deamons listed above, additional deamons are run for the transfer to KNMI of the ECMWF wind fields, to update the near-real time statistics and to monitor the FD system.

5.4 Product distribution

The distribution deamon *distributefd* takes care of the distribution of the level 2 and 4 products to the users. The distribution is organized via a product list and a distribution list. In the product list, the products that have to be distributed are given. The level 2 and 4 ASCII data files and the level 4 GIF files are the products that are currently distributed. In the distribution list, the locations to which the products have to be distributed are given. The distribution list also includes the protocol by which the products have to be distributed, for instance local file copy or ftp transfer.

5.5 The Fast Delivery WWW-site

The FD WWW-site contains level 2 and 4 images and ASCII data. A level 2 near-real time image of the GOME orbits for the past 24 hours is shown and an image of the latest level 4 assimilated ozone field for 12^h of the present day is shown for World, North Pole and South Pole projections.

Level 2 and 4 data and level 4 images for the present and previous month are available. The level 2 data is given in hourly ASCII files and the assimilated level 4 ozone fields are given in daily ASCII files and GIF images.

In addition to the level 2 and 4 data, the WWW-site also contains on-line documentation on the FD system. General information on the FD system is available as well as a description of the data formats used. There is also documentation on the level $0 \rightarrow 1$, $1 \rightarrow 2$ and $2 \rightarrow 4$ processors.

As a user-service, there are e-mail links for questions or special requests for additional data or documentation.

The address of the Fast Delivery WWW-site is:

`http://www.knmi.nl/neonet/atmo_chem/gome/fd/index.html`

Chapter 6

References

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

ER-TN-ESA-GO-0096 GOME Science Packet Description

ER-TN-DPE-PA-1137 New EGOI Windows

Eskes H.J., A.J.M. Piters, P.F. Levelt, M.A.F. Allaart, and H. Kelder, Variational data assimilation: How to extract more information from GOME total ozone data, Earth Observation Quarterly - GOME special, Vol.58, p.35-38, 1998.

Eskes H.J., A.J.M. Piters, P.F. Levelt, M.A.F. Allaart, and H. Kelder, Variational assimilation of total-column ozone satellite data in a 2D lat-lon tracer-transport model, submitted to Journal of Atmospheric Sciences, 1999.

Fortuin, J.P.F. and H. Kelder, An ozone climatology based on ozonesonde and satellite measurements, Journal of Geophysical Research, Vol. 103, p. 31,709-31,731, 1998

Herman, J.R., E.A. Celarier, Earth surface reflectivity climatology at 340-380 nm from TOMS data, Journal of Geophysical Research, Vol. 102, p. 28,003-28,011, 1997.

Koelemeijer, R.B.A. and P. Stammes, Effects of clouds on ozone column retrieval from GOME UV measurements, to appear in the Journal of Geophysical Research, 1999.

Piters, A. J. M., P. F. Levelt, M. A. F. Allaart, and H. M. Kelder, Validation of GOME total ozone column with the assimilation model KNMI, in "GOME Geophysical Validation Campaign", ESA WPP-108, p209, 1996.

Rossow, W.B. and L.C. Garder, Cloud detection using satellite measurements of infrared and visible radiances for ISCCP, J. Climate, 6, 2341-2369, 1993.

Chapter 7

Acronyms

AMF Air Mass Factor
BU Binary Unit
BSDF Bi-directional Scattering Distribution Function
DAK Doubling Adding model KNMI
DOAS Differential Optical Absorption Spectroscopy
DSR Product Data Set Record
ECMWF European Centre for Medium Range Weather Forecast
EGOI Extracted GOME Instrument header data
ERS European Remote Sensing Satellite
ESA European Space Agency
FTP File Transfer Protocol
GOFAP GOME Ozone Fast delivery and value-Added Products
GOME Global Ozone Monitoring Experiment
ICFA Initial Cloud Fitting Algorithm
ISCCP International Satellite Cloud Climatology Project
KNMI Koninklijk Nederlands Meteorologisch Instituut
PMD Polarisation Measurement Device
TOMS Total Ozone Mapping Spectrometer
VAMK Variational Assimilation Model KNMI
VU Vrije Universiteit
WWW World Wide Web

GOFAP-KNMI-PSD-01

GOME Ozone Fast Delivery and value-Added Products Product Specification Document

March 4, 1999

Authors: A.J.M. Piters and P.J.M. Valks, KNMI

Reviewed by: C. Zehner, ESRIN

Contents

Chapter 1

Introduction

The aim of the project "GOME Ozone Fast Delivery and value-Added Products" (GOFAP) is to develop a Fast Delivery (FD) processor retrieving total ozone columns from the Extracted GOME Instrument header data (EGOIs), which are used at ESRIN for the instrument performance and quality assurance monitoring. The project is performed within the framework of ESA's Data User Programme (DUP).

The EGOIs consist of selected windows of the raw GOME data, among which is an ozone retrieval window (325–335 nm). The FD-processor performs a simple calibration on this raw data, and uses the differential spectrum (Earth spectrum divided by the Sun spectrum) to retrieve a slant ozone column density, i.e. the amount of ozone molecules per square centimeter along the light path. A look-up table of pre-calculated Air Mass Factors (AMFs) is used to derive the vertical column density.

The retrieval of total ozone columns from the EGOI data takes in the order of a few seconds per GOME orbit on a workstation. The collection on board, downloading to the ground station, extraction of the EGOI data, and transport to ESRIN and from there to KNMI takes considerably more time. However, the total time delay before the total ozone column is available on the web is well within 3 hours after observation (hence near-real time). Therefore it is possible to use this level 2 data into on-line numerical weather prediction systems, e.g. into the ECMWF weather prediction model.

Furthermore, the near-real time ozone columns are continuously assimilated into a simple off-line transport model, which makes use of the ECMWF analysed and forecast wind, to construct up-to-date ozone maps (level 4 data). These maps can be used for, e.g., measurement campaign preparations.

This document describes the products resulting from the fast-delivery processor.

Chapter 2

Total ozone columns: Level 2

The GOME level 2 archive files contains the measured total ozone columns and the corresponding error estimate for one hour intervals. Each row in an ASCII-archive files contains the data of one GOME pixel measurement. First the GOME pixel date and time (GMT) is given, then the longitudes and latitudes of the four corner points of the GOME pixel are given in 1/100 degree. The longitude runs from 0° to 360° and the latitude from 90° (North Pole) to −90° (South Pole). The pixel scan position (East, Nadir or West) is given and the total ozone column and the corresponding error estimate in Dobson Units. The format of the ASCII data files is:

Field	Number of characters	Position	Description
1	14	1–14	GOME pixel time (YYYYMMDDhhmmss)
2	4 * 6	15–38	Longitude of the four corner points of the pixel in 1/100 degree (_XXXXX)
3	4 * 6	39–62	Latitude of the four corner points of the pixel in 1/100 degree (_XXXXX)
4	3	63–67	Total ozone column in Dobson Units (_XXX_)
5	2	68–69	Pixel scan position (_X, 0=East, 1=Nadir, 2=West)
6	3	70–72	Error estimate of the total ozone column in Dobson Units (_XX)

As an example, the first few lines of the data file for 10^h, May 19, 1998 look like the following:

```
19980519100000  6936  7131  6126  6304  7891  7909  8112  8131 480  1  9
19980519100002  6126  6304  4613  4744  8112  8131  8330  8348 461  2  8
19980519100005  7391  7588  6772  6961  7602  7621  7874  7893 478  0  9
19980519100006  6772  6961  5940  6111  7874  7893  8091  8110 483  1  9
19980519100008  5940  6111  4425  4550  8091  8110  8302  8321 468  2  9
19980519100011  7256  7447  6613  6795  7588  7608  7856  7876 474  0  9
19980519100012  6613  6795  5763  5927  7856  7876  8068  8089 483  1  9
19980519100014  5763  5927  4252  4370  8068  8089  8273  8293 479  2  9
19980519100017  7123  7308  6459  6635  7573  7593  7837  7858 470  0  9
19980519100018  6459  6635  5594  5752  7837  7858  8045  8067 481  1  9
19980519100020  5594  5752  4092  4203  8045  8067  8243  8264 482  2  9
```

Chapter 3

Assimilated total ozone maps: Level 4

The GOME level 4 archive files contain the assimilated total ozone columns at 12^h for each day. The data are gridded into 1 degree latitude zones by 1.25 degree longitude zones. Latitudes go from -90 degrees (South Pole) to 0 degrees (Equator) to $+90$ degrees (North Pole) in 1 degree steps, so there are 180 latitude zones. The first zone extends from -90 to -89, so the grid cell is considered to be centered on -89.5. Similarly, longitudes go from -180 (West longitude) to 0 (Greenwich, England) to $+180$ (East longitude) in 1.25 degree steps, so there are 288 longitude zones (360/1.25). The first longitude zone extends from -180 to -178.75, so the grid cell is considered to be centered on -179.375.

The first three lines are header information which includes the date the data were taken, the type of data and the generation date of the file.

Then, the 288 longitude values for one latitude zone, centered at -89.5, are given. The next 288 longitude values are given for the latitude zone centered at -88.5, and so on. All measurements are given in Dobson units and are integers with 3 significant figures.

As an example, the first few lines of the data file for May 19, 1998 look like the following:

```
 Day: 139 May 19, 1998 Assimilated GOME total O3 (KNMI/ESA) 9805194; T=12:00
Longitudes:  288 bins centered on 179.375 W to 179.375 E  (1.25 degree steps)
Latitudes :  180 bins centered on  89.5   S to  89.5   N  (1.00 degree steps)
300300300300300300300300300300300300300300300300300300300300300300301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301301301301301301301301301301301301301301301301301301301301301301301301301
301300300300300300300300300300300300300300300300300300300300300300300300300
300300300300300300300300300300300300300   lat =  -89.5
```

Chapter 4

Data distribution

The distribution part of the Fast Delivery System takes care of the delivery of level 2 and 4 products to the users. This distribution process is described in Section 5.4 of the Algorithm Specification Document. At the moment, ASCII data files of the level 2 total ozone columns and images and ASCII data files of the level 4 assimilated total ozone fields are distributed to the Fast Delivery WWW-site.

The WWW-site contains the level 2 and 4 ASCII data and level 4 images for the present and previous month. New data created by the FD system are copied to the WWW-site every 15 minutes. The level 2 data is given in hourly ASCII files and the assimilated level 4 ozone fields are given in daily ASCII files and GIF images in World, North Pole and South Pole projections. The WWW-site also contains two compressed tar files with all the level 2 data for the present and previous month. The format of the level 2 and level 4 ASCII files is described in Chapters 2 and 3 of this document.

Special user services, such as automatic FTP transfer of level 2 and 4 data to the users' FTP-server, can be implemented in the FD-system. Requests for such a data transfer service or for additional data can be made by using the e-mail link on the WWW site.

The address of the Fast Delivery WWW-site is:

http://www.knmi.nl/neonet/atmo_chem/gome/fd

Chapter 5

Documentation

Documentation on the Fast Delivery system can be found in the Algorithm Specification Document. This document describes the level 0→'1', level '1'→2 and level 2→4 algorithms and gives an overview of the operational environment of the Fast Delivery system. Documentation on the Fast Delivery products can be found in this document (the Product Specification Document).

Likewise, the Fast Delivery WWW-site contains on-line documentation on the Fast Delivery system. There is general information on the FD system and descriptions of the level 0→'1', level '1'→2 and level 2→4 processors. In addition, the formats of the available level 2 and level 4 ASCII data files on the WWW-site are described.