



GOMOS One-step Algorithm Theoretical Basis Document

Date: 6 June 2016

Version: 1.02

Written by: Janne Hakkarainen (janne.hakkarainen@fmi.fi)

Purpose of this document

The purpose of this ATBD is to give an outline of the GOMOS one-step retrieval algorithm. In addition, this document discusses the differences between the one-step and the two-step approaches, and the role of priors in the algorithm. This document is created in the framework of the ALGOM (GOMOS Level 2 algorithm evolution studies) project funded by the European Space Agency. This document will be distributed along with GOMOS One-step UTLS ozone dataset.

Contents

1	Introduction to the GOMOS measurements	2
2	Operational two-step retrieval algorithm	3
3	One-step retrieval algorithm	4
4	Main differences between two-step and one-step approaches	5
5	Limitations of the Algorithm	6

1 Introduction to the GOMOS measurements

Global Ozone Monitoring by Occultation of Stars (GOMOS) is a satellite instrument onboard ENVISAT spacecraft that was launched in March 2002. It was operative until the end of ENVISAT mission in 2012. During the mission, GOMOS observed about 880 000 vertical profiles of ozone, NO_2 , NO_3 , aerosol extinctions and other species. About half of these measurements were made during nighttime. In this section, we recall the GOMOS measurement principle and the operative retrieval algorithm. For a more comprehensive introduction to GOMOS, see the GOMOS Algorithm Theoretical Basis Document (ATBD)¹ and ACP's GOMOS special issue² and the papers therein; in particular [1, 3, 4, 8].

The GOMOS measurement principle is relatively simple and it is based on the stellar-occultation technique. The stellar spectrums at different tangent altitudes are obtained with the sampling resolution of 0.3 – 1.7 km. One occultation contains roughly 100 measurement spectrums at 1 416 different wavelengths in the UV–Visible wavelength region. Hence, roughly 150 000 individual measurements per occultation are obtained. To obtain the so-called transmission spectrum, the stellar spectrum is divided by the reference spectrum that is measured above the atmosphere. This can be written as

$$T_{\text{ext}}^{\text{meas}}(\lambda, l) = \frac{I(\lambda, l)}{I^{\text{ref}}(\lambda)}, \quad (1)$$

where $I(\lambda, l)$ is the stellar spectrum measured along the line-of-sight l and at wavelength λ . $I^{\text{ref}}(\lambda)$ is the reference spectrum measured above the atmosphere. The error characteristics of the measured transmission spectrum—used later to obtain the atmospheric profiles—vary strongly from star to star with the stellar brightness and temperature. In the lower atmosphere, in particular in the UTLS region, the signal-to-noise ratio is low due to increasing amount of aerosols. See [8] for discussion and examples.

The transmission can be modeled using the Beer-Lambert law

$$T_{\text{ext}}^{\text{mod}}(\lambda, l) = \exp(-\tau(\lambda, l)) \quad (2)$$

where the optical depth $\tau(\lambda, l)$ is given as

$$\tau(\lambda, l) = \sum_{j=1}^{\text{nconst}} \int_l \alpha_j(\lambda, T(s)) \rho_j(s) ds, \quad (3)$$

where ρ_j is a local density profile and $\alpha_j(\lambda, T(s))$ is the constituent, wavelength and temperature dependent cross-section. The integration is done along the line-of-sight. This model can be reduced using the so-called effective cross-sections that are assumed to be constant along the integration. This reduced model can be written as

$$\tau(\lambda, l) = \sum_{j=1}^{\text{nconst}} \alpha_j^{\text{eff}}(\lambda) N_j(l), \quad (4)$$

where $N_j(l) = \int_l \rho_j(s) ds$ is the so-called line density of the constituent j and $\alpha_j^{\text{eff}}(\lambda)$ is the constituent and wavelength dependent effective cross-section. When the problem has been discretized, we can use the geometry kernel matrix K to obtain the line densities N_j from the actual local density profiles ρ_j . This can be written as

$$N_j = K \rho_j. \quad (5)$$

¹Available on the Internet at https://earth.esa.int/documents/10174/384988/GOMOS_ATBD_V3.pdf

²Available on the Internet at http://www.atmos-chem-phys.net/special_issue153.html

Using the vectorized notation, the full transmission model can be written as

$$T_{\text{ext}}^{\text{mod}}(\rho) = \exp(-(K \otimes A)\rho), \quad (6)$$

where $K \otimes A$ is the Kronecker product between the kernel operator K and the cross-sections A . If three gases (O_3 , NO_2 and NO_3), three aerosol parameters and 100 transmission spectrums at 1500 wavelengths are considered, the sizes of the elements of the model (6) are the following. The length of the transmission model output $T_{\text{ext}}^{\text{mod}}$ is $1500 \times 100 = 150\,000$ and the length of the input ρ is $6 \times 100 = 600$. Hence, the linear operator $K \otimes A$ has to be a $150\,000 \times 600$ matrix. Here the altitudes of the transmission model input and output are assumed to be the same. In addition to O_3 , NO_2 , NO_3 and aerosol parameters, neutral density and some minor trace gases like OCIO and BrO can be included in the model, too.

The aim of the GOMOS retrieval algorithms is to find the atmospheric profiles ρ_j that model the measured transmission in statistical sense given the measurement error and some *a priori* knowledge. Next, we present an outline of the operational two-step approach.

2 Operational two-step retrieval algorithm

The operational GOMOS retrieval algorithm is based on a two-step approach, where the *spectral* and the *vertical* inversion are conducted separately. This procedure can be iterated for improving accuracy. The use of the reduced model with the effective cross-sections enables the splitting.

In the spectral inversion, the line densities N are obtained by minimizing the following cost function

$$J(N) = \frac{1}{2}(T_{\text{ext}}^{\text{mod}}(N) - T_{\text{ext}}^{\text{meas}})^T C^{-1} (T_{\text{ext}}^{\text{mod}}(N) - T_{\text{ext}}^{\text{meas}}). \quad (7)$$

using the Levenberg-Marquardt (LM) optimizing algorithm. The covariance matrix C includes the measurement noise and possible modeling uncertainties. Note that this part of the retrieval uses uninformative flat prior and only the likelihood part of the Bayesian cost function is considered.

When the spectral inversion part has been done and the different line densities N_j obtained, the vertical inversion can be started. In the vertical inversion, the aim is to find the vertical profiles ρ_j given the line densities N_j separately for every constituent. In principle, this could be done solving the linear Eq. (5). However, because of the low signal-to-noise ratio and other factors, the solution is based on Tikhonov regularization. Using the matrix calculus, the solution can be written as

$$\rho_{\gamma,j} = (K^T K + \gamma L^T L)^{-1} K^T N_j, \quad (8)$$

where γ is the Tikhonov regularization parameter. In the retrieval, the smoothness requirements are considered. Hence, L is selected to be the tridiagonal matrix that approximate the second derivative. It can be written as

$$L = \frac{1}{h^2} \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -2 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (9)$$

where h is the layer thickness (in practice different for every layer). In the operative algorithm, the regularization parameter γ has been selected so that the so-called target resolution—defined

as a spread of the averaging kernel computed using (5) and (8), see, e.g., [6] for details—for ozone is 2 km below 30 km and 3 km above 40 km. For NO₂, NO₃ and aerosol parameters the selected target resolution is 4 km at all altitudes.

Two features about the algorithm can be noted. First, the (smoothness) priors have only effect in the vertical inversion part and are given for each constituent separately. This means that the prior given, e.g., to aerosol profiles does not affect the retrievals of other species. Second, the averaging kernel of the Tikhonov solution (8) does not depend of the measurement noise, and it depends of the actual measurements only via the kernel matrix K . This allows the setting of the target resolution, since the measurement geometry is relatively similar from one occultation to another and the layer thickness is included in (9).

3 One-step retrieval algorithm

The basic idea of the GOMOS one-step retrieval algorithm is to conduct the spectral and the vertical inversion of the operative GOMOS algorithm simultaneously using the full model (6). From the methodological point of view this makes sense, since the assumption of the effective cross-sections, see formula (4), can be dropped. In addition, the splitting of the inversion in two parts is somewhat artificial. Historically, the use of the two-step strategy in the operative algorithm is mainly related to the enormous size of the problem. The possibility of the one-step inversion was studied already in the early days of the GOMOS mission [9, 2]. Here we present our own version of the GOMOS one-step retrieval algorithm. The algorithm is written in MATLAB programming language.

The one-step retrieval of the GOMOS measurements can be seen as a minimization of the following cost function consisting of the likelihood and the prior:

$$J(\rho) = \frac{1}{2}(T_{\text{ext}}^{\text{mod}}(\rho) - T_{\text{ext}}^{\text{meas}})^T C^{-1}(T_{\text{ext}}^{\text{mod}}(\rho) + T_{\text{ext}}^{\text{meas}}) + \frac{1}{2}(\rho_{\text{prior}} - H\rho)C_{\text{prior}}^{-1}(\rho_{\text{prior}} - H\rho), \quad (10)$$

where H is a linear operator based on how the prior given. If, e.g., the smoothness prior is considered, the matrix H is selected like the matrix L in (9) and ρ_{prior} is set to zero. This cost function could, in principle, be minimized using the non-linear optimizing algorithms like the LM algorithm used in the spectral inversion part of the operative algorithm. However, in practice, we solve this problem using generalized linear model (GLIM) approach with the iterative re-weighted least squares algorithm (IRLS), where the special exponential structure of the problem can be exploited for numerical efficiency [5]. The IRLS algorithm is a special case of a more general approach to optimization called the Newton-Raphson algorithm.

Let us next show how the GLIM approach works in practice. Let us consider the following model

$$y = \exp(Ax) + \varepsilon, \quad (11)$$

where y are the measurements, A is the linear operator, $\varepsilon \sim N(0, \sigma I)$ is the measurement noise and x is the unknown state vector. Readers familiar with the GLIM vocabulary, can note that the *link function* between the measurements and the linear operator is logarithmic and—as the measurement error is assumed Gaussian—the *variance function* is simply identity.

Using the IRLS algorithm, starting from the initial guess of x , the unknown x can be solved iteratively. In the iteration steps, the auxiliary problem based on the linear equation

$$z = Ax, \quad (12)$$

is solved using the standard-deviation weights w , where

$$\eta = Ax, \quad (13)$$

$$\mu = \exp(\eta), \quad (14)$$

$$z = \eta + (y - \mu)/\mu, \quad (15)$$

$$w = \sigma/\mu. \quad (16)$$

For more information, see [5] for details. The linear step is operating on the original unknowns x and if we assign prior regularization to each step then the final solution uses the same prior constraints.

When prior information x_{prior} is considered, the iterative solution x can be seen as the argument that minimizes the following quadratic cost function

$$J(x) = \frac{1}{2}(Ax - z)^T C_w^{-1} (Ax - z) + \frac{1}{2}(x_{\text{prior}} - Hx)^T C_{\text{prior}}^{-1} (x_{\text{prior}} - Hx), \quad (17)$$

where C_w is the diagonal covariance matrix induced by the weights w . As the operators A and H in the cost function (17) are linear, we can obtain the solution x using basic linear algebra

$$x = (A^T C_w^{-1} A + H^T C_{\text{prior}}^{-1} H)^{-1} (A^T C_w^{-1} z + H^T C_{\text{prior}}^{-1} x_{\text{prior}}). \quad (18)$$

In practice, we solve this linear equation using Cholesky factorization of $(A^T C_w^{-1} A + H^T C_{\text{prior}}^{-1} H)$ and MATLAB's backslash³ operator, although other options exist too. We note that in full case, the presented GLIM approach is computationally much lighter than the computationally rather expensive LM algorithm.

Prior information. In the one-step algorithm, there are currently three ways of giving the prior information: one can give the prior in absolute units or give the prior for the first or the second derivative of the profiles. Naturally, all three types of priors can be considered at the same time. The non-trivial part in the one-step algorithm is the selection of the (diagonal) prior covariance matrix C_{prior} for the derivatives. From the statistical point of view, the covariance matrix should reflect our prior knowledge, e.g., from the theory, about how smooth the profiles are or how big steps can the profiles make between the altitudes. See [2, 7] for details. From the practical point of view, the prior covariance matrix can be seen as a similar regularization tuning handle of the inversion as the Tikhonov regularization parameter γ is in the operative algorithm.

4 Main differences between two-step and one-step approaches

Although the operative two-step and the presented one-step algorithms consist of exactly the same elements, they are still fundamentally different. In one-step algorithm some of the approximation can be avoided and the prior and the measurement errors are correctly treated together. Some of the algorithmic differences are discussed next.

The main difference of the retrieval algorithms comes from the use of the prior information. In the one-step algorithm, the prior given to one constituent affects the other constituent too. This can be clearly seen from the formulation of the cost function (10). The opposite is true for the operative algorithm, where the prior takes place only in the vertical inversion and is given for every constituent separately.

³See <http://www.mathworks.se/help/matlab/ref/mldivide.html>

On the other hand, the resolution of the operative algorithm depends very little of the actual occultations. In particular, it is independent of the measurement noise that varies strongly from one star type to another. This allows the setting of the so-called target resolution, which makes the operative dataset user-friendly and easy to use in, e.g., time-series analysis and validation studies [4].

In the one-step algorithm, similar target resolution cannot be set, since the resolution depends on the occultation depended measurement noise. In particular, the target resolution cannot be set, if the prior covariance matrix is kept fixed for all star types.

5 Limitations of the Algorithm

Although the one-step retrieval algorithm has proven its usefulness in various case studies and in creation of the full ALGOM dataset, it is good to list some limitations of the algorithm: 1) Like the operational two-step algorithm, it is very sensitive to the aerosols; 2) Target resolution cannot be set; 3) Application of the so-called full-covariance matrix is challenging (i.e., too time consuming when creating full dataset); 4) Error-bars are expected to be too small; 5) Current best setup for ozone does not allow good reliable retrievals for NO_2 and NO_3 .

References

- [1] J. L. Bertaux, E. Kyrölä, D. Fussen, A. Hauchecorne, F. Dalaudier, V. Sofieva, J. Tamminen, F. Vanhellemont, O. Fanton d’Andon, G. Barrot, A. Mangin, L. Blanot, J. C. Lebrun, K. Pérot, T. Fehr, L. Saavedra, G. W. Leppelmeier, and R. Fraisse. Global ozone monitoring by occultation of stars: an overview of GOMOS measurements on ENVISAT. *Atmos. Chem. Phys.*, 10(24):12091–12148, 2010.
- [2] H. Haario, M. Laine, M. Lehtinen, E. Saksman, and J. Tamminen. Markov chain Monte Carlo methods for high dimensional inversion in remote sensing. *J. R. Stat. Soc. (B)*, 66(3):591–607, 2004.
- [3] E. Kyrölä, J. Tamminen, V. Sofieva, J. L. Bertaux, A. Hauchecorne, F. Dalaudier, D. Fussen, F. Vanhellemont, O. Fanton D’Andon, G. Barrot, M. Guirlet, T. Fehr, and L. Saavedra de Miguel. GOMOS O_3 , NO_2 , and NO_3 observations in 2002-2008. *Atmos. Chem. Phys.*, 10(16):7723–7738, 2010.
- [4] E. Kyrölä, J. Tamminen, V. Sofieva, J. L. Bertaux, A. Hauchecorne, F. Dalaudier, D. Fussen, F. Vanhellemont, O. Fanton d’Andon, G. Barrot, M. Guirlet, A. Mangin, L. Blanot, T. Fehr, L. Saavedra de Miguel, and R. Fraisse. Retrieval of atmospheric parameters from GOMOS data. *Atmos. Chem. Phys.*, 10(23):11881–11903, 2010.
- [5] P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Chapman & Hall, London, second edition, 1989.
- [6] Clive D. Rodgers. *Inverse methods for atmospheric sounding: theory and practice*. World Scientific, 2000.
- [7] V. F. Sofieva, J. Tamminen, H. Haario, E. Kyrölä, and M. Lehtinen. Ozone profile smoothness as a priori information in the inversion of limb measurements. *Ann. Geophys.*, 22(10):3411–3420, 2004.

- [8] J. Tamminen, E. Kyrölä, V. F. Sofieva, M. Laine, J.-L. Bertaux, A. Hauchecorne, F. Dalaudier, D. Fussen, F. Vanhellemont, O. Fanton-d'Andon, G. Barrot, A. Mangin, M. Guirlet, L. Blanot, T. Fehr, L. Saavedra de Miguel, and R. Fraisse. GOMOS data characterisation and error estimation. *Atmos. Chem. Phys.*, 10(19):9505–9519, 2010.
- [9] Filip Vanhellemont, Didier Fussen, and Christine Bingen. Global one-step inversion of satellite occultation measurements: A practical method. *J. Geophys. Res.*, 109(D9), 2004.