

Technical note: Potential improvements of GOMOS retrievals in the UTLS using one-step approach

Date: 15 March 2016

Version: 1.20

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

Purpose of this document

The purpose of this technical note is to document, report and summarize the studies done with the one-step approach in the framework of the ALGOM (GOMOS Level 2 algorithm evolution studies) project funded by the European Space Agency. This technical note studies Task 1.3 of the ALGOM project. The scope of Task 1.3 is to search improvements to the existing GOMOS UTLS ozone dataset(s) and to optimize the one-step algorithm for UTLS ozone retrievals.

Motivation for using the one-step algorithm for UTLS ozone comes from the positive results of [5]. In addition, motivation comes from the previous validation studies and the UTLS validation study of this project. The latter shows that the operational GOMOS IPF v6 ozone product has a substantial bias up to 100% at the UTLS altitudes.

This document is divided in three parts. The first part gives an outline of the GOMOS one-step retrieval algorithm and discusses the differences between the one-step and the two-step approaches, and the role of priors in the algorithm. The second part searches the optimal setup for the UTLS ozone and gives validation and verification. The third part describes the processed sample dataset.

Contents

1	Description of the algorithm	2
1.1	Introduction to the GOMOS measurements	2
1.2	Operational two-step retrieval algorithm	3
1.3	One-step retrieval algorithm	4
1.4	Main differences between two-step and one-step approaches	6
2	Verification of UTLS ozone with different setups of the one-step retrieval algorithm	6
2.1	Validation of GOMOS one-step using NDACC station soundings	7
2.2	A search for geophysical features at UTLS altitudes	13
2.3	Validation of GOMOS one-step retrievals using OSIRIS dataset	16
2.4	Verification of the new setup	16
2.5	Analysis of error estimates, vertical resolution and effect of regularization	23
2.6	Conclusion from validation	30
2.7	Recommendations for the next IPF processing	30
3	One-step dataset of UTLS ozone	31
3.1	Disclaimer	31
3.2	Summary of variables in NetCDF files.	32
A	Supplementary material	42

1 Description of the algorithm

1.1 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, NO₂, NO₃, 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 [2, 6, 7, 12].

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

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

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 [12] 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

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

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.

1.2 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., [10] for details—for ozone is 2 km below 30 km and 3 km above 40 km. For NO_2 , NO_3 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).

1.3 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 [13, 4]. 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 [8]. 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 [8] 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.

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

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 [4, 11] 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.

1.4 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.

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 [7].

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.

2 Verification of UTLS ozone with different setups of the one-step retrieval algorithm

In this section, we test different setups of the one-step algorithm in order to improve the quality of the UTLS ozone. We search for the best setup of the algorithm in order to process the one-step version of the ALGOM UTLS dataset. In particular, we test the role of different aerosol models, which are expected to have the largest effect on the UTLS ozone.

We start our search by looking the same setup that was—before ALGOM project—used in our arctic ozone case study [5]. It includes the same three-aerosol-parameter model of the operational algorithm with very strong aerosol smoothness priors. In addition, ECMWF air is used as an absolute prior for air retrieval and the standard deviation is set to 5% of its value. This setup proved to be much better than GOMOS IPF v5 in the case study of “unprecedented stratospheric ozone loss in the arctic in spring of 2011” [5]. See Fig. 1 for an example. See also the animation at http://dl.dropboxusercontent.com/u/35228286/gomos_video20110330.gif.

N.B. Though out this section we use only the occultations of 49 brightest stars (i.e., GOMOS star index < 50) in full dark (SZA > 107 degrees) and have processed the measurements between 10 and 100 km altitude. A very little attention to the altitudes other that UTLS

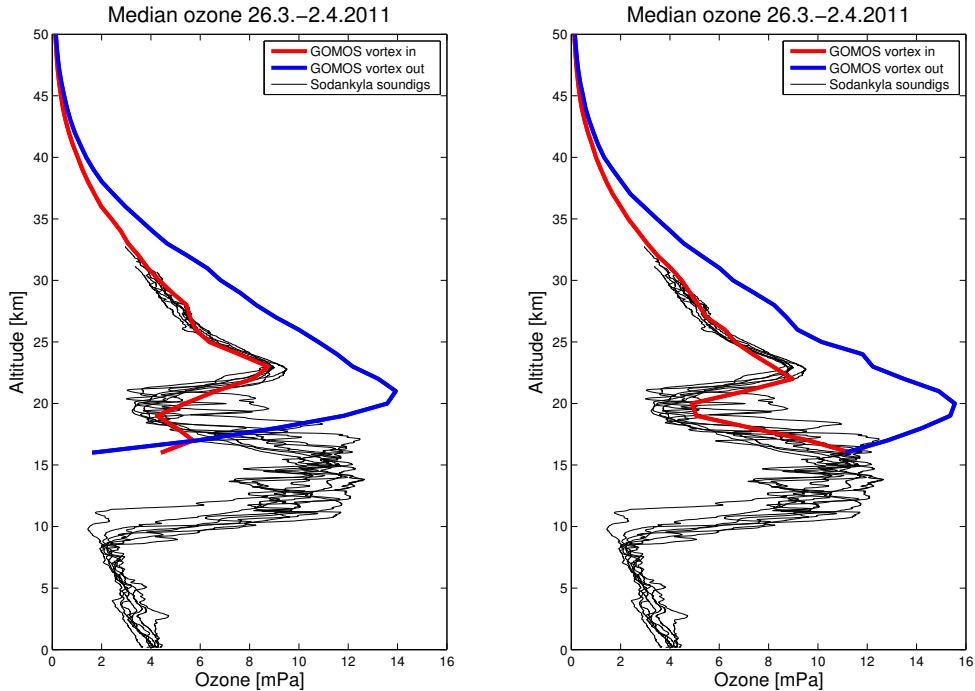


Figure 1: The median GOMOS ozone profiles retrieved using the operative (left panel) and the one-step (right panel) retrieval algorithms inside and outside the Arctic polar vortex. Sodankylä sounding profiles are illustrated too. Clear differences between the retrieval algorithms can be observed. Adopted from [5].

altitudes is given. No absolute priors are used for ozone (or relevant constituents)!

2.1 Validation of GOMOS one-step using NDACC station soundings

As in validation task of this project, we use the ozone soundings from different NDACC stations to validate GOMOS profiles. The stations used in Task 1.1. are illustrated in and Fig. 2 and summarized in Table 1. In this task, we concentrate only on the stations that contains most collocations in troposphere. These stations are highlighted in Fig. 2 and Table 1.

In Fig. 3 we show 104 profiles at Izaña processed with the setup from [5] (blue curve). We observe drastic improvement of UTLS ozone with respect to IPF v6 (black curve). The overall visual agreement between one-step ozone and soundings is “outstanding”. If we relax the strong aerosol priors (magenta curve), the agreement is lost. This proves that the strong smoothes has tremendous impact on the quality of the UTLS ozone.

Unfortunately, although very strong aerosol smoothness priors seem to work extremely well at Izaña and in [5], there is still a price to pay. In case of a clear aerosol layer—not just background aerosols—things can go “very wrong” with this setup. A failure of this kind is illustrated in Fig. 4. As aerosol profiles do not “bend” enough, and ozone starts to compensate its effect. This is the reason why this setup is not recommended for the processing of the whole mission.

Our next working hypothesis includes a change of the aerosol model to $1/\lambda$. Improvements are expected from the fact a) we get more degrees of freedom as this change reduces the number of parameters from 3 to 1, and b) this removes the second degree that competes with ozone.

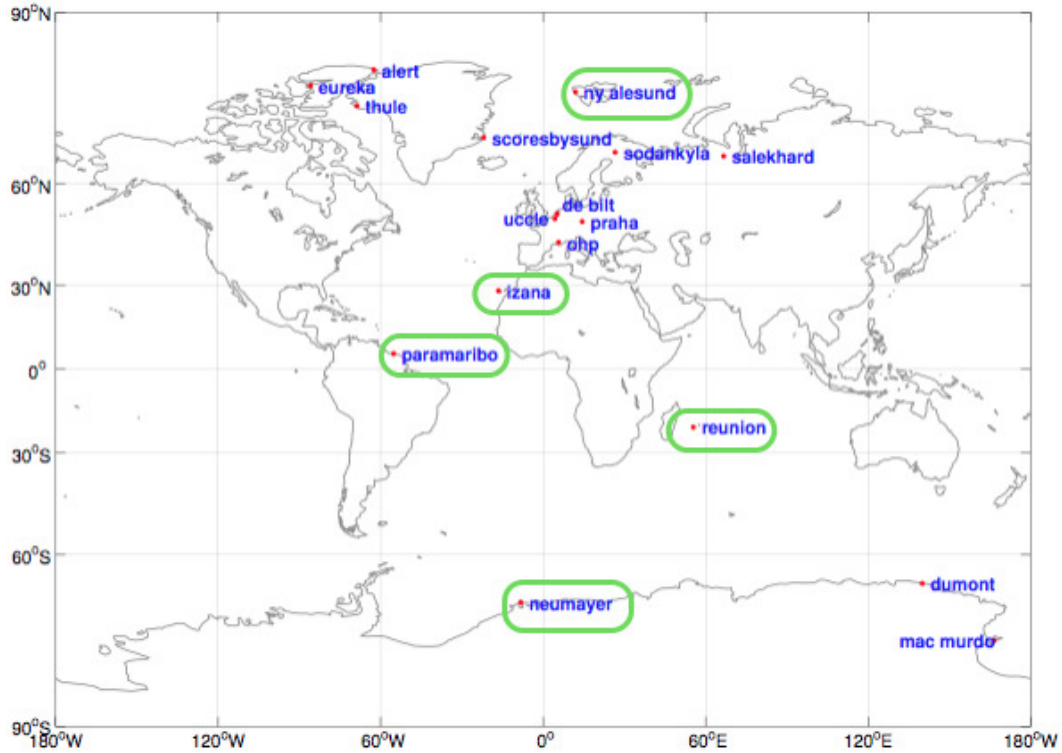


Figure 2: Map of the NDACC network stations used in Task 1.1. The stations analyzed in this task are highlighted in green.

In addition to $1/\lambda$ aerosol model, we tested also $1/\lambda^2$ aerosol model (as it only contains one variable), but this idea was later dropped as it produced inferior results with respect to $1/\lambda$. In addition, we also relaxed the strong aerosol priors.

Figures 5–9 show results—using $1/\lambda$ aerosol model—in five different NDACC stations. These stations are highlighted in Fig. 2. Right panel shows the standard deviations (STD). We can note that retrievals made with the one-step algorithm improve the quality of the UTLS ozone with respect to IPF v6 substantially. This is highlighted especially in the tropical regions. It can also be noted that GOMOS one-step retrievals have smaller STD than IPF v6 and it is very similar (only slightly higher) to the ozone STD of the soundings.

Table 1: Summary of collocations at different NDACC stations used in Task 1.1. The stations analyzed in this task are highlighted.

Station	(Lat ° N, Lon ° E)	#collocations in UTLS	#collocations in troposphere
Alert	(82.45, -62.51)	38	3
De Bilt	(52.10, 5.18)	53	3
Dumont	(-66.67, 140.02)	57	9
Eureka	(79.99, -85.93)	51	3
Mac Murdo	(-77.85, 166.63)	7	2
Izaña	(28.30, -16.50)	160	67
Neumayer	(-70.68, -8.26)	74	18
Ny Ålesund	(78.93, 11.93)	132	26
OHP	(43.94, 5.71)	74	3
Paramaribo	(5.8, -55.22)	75	15
Praha	(50.00, 14.44)	66	0
Reunion	(-21.06, 55.48)	82	34
Salekhard	(66.50, 66.70)	7	0
Scorebysund	(-70.68, -8.26)	112	20
Sodankylä	(78.93, 11.93)	58	0
Thule	(43.94, 5.71)	16	1
Uccle	(50.80, 4.35)	195	6

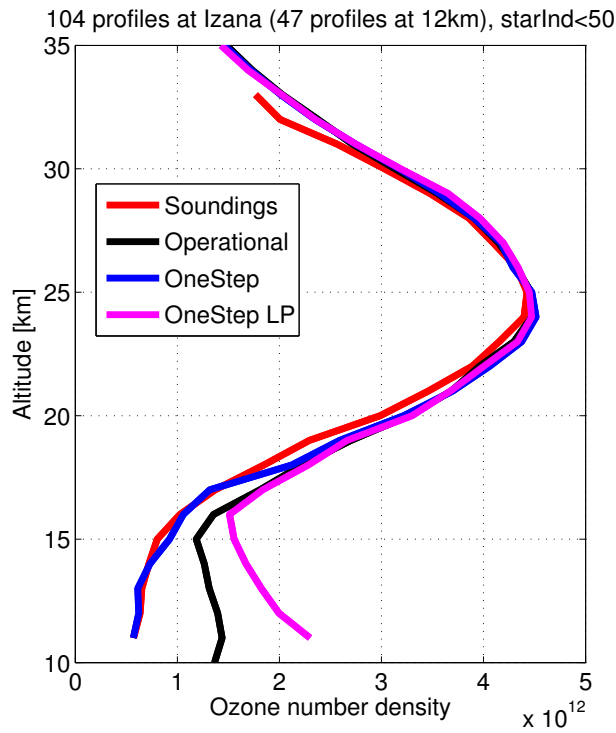


Figure 3: Illustration of 104 profiles at Izaña. See text for details.

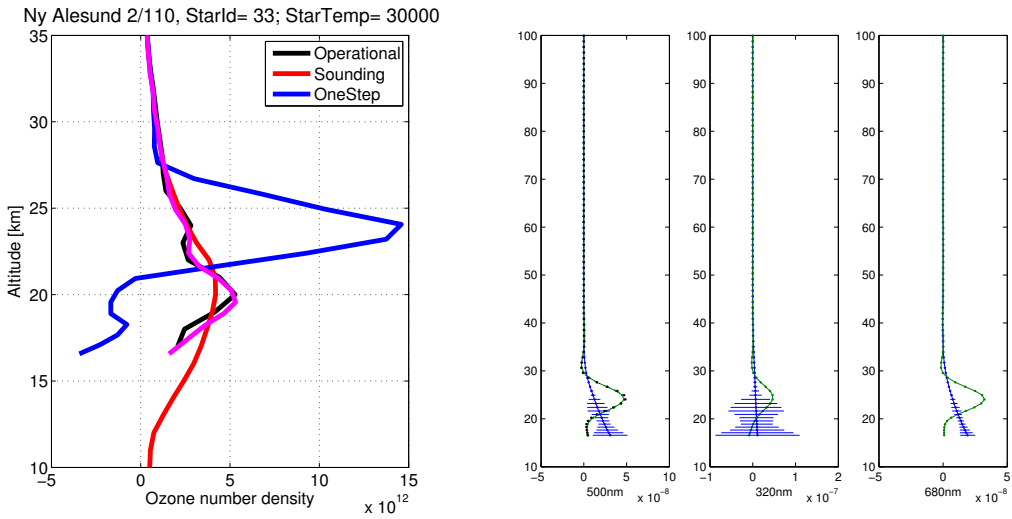


Figure 4: Example where the setup of [5] fails. Priors relaxed in magenta. See text for details.

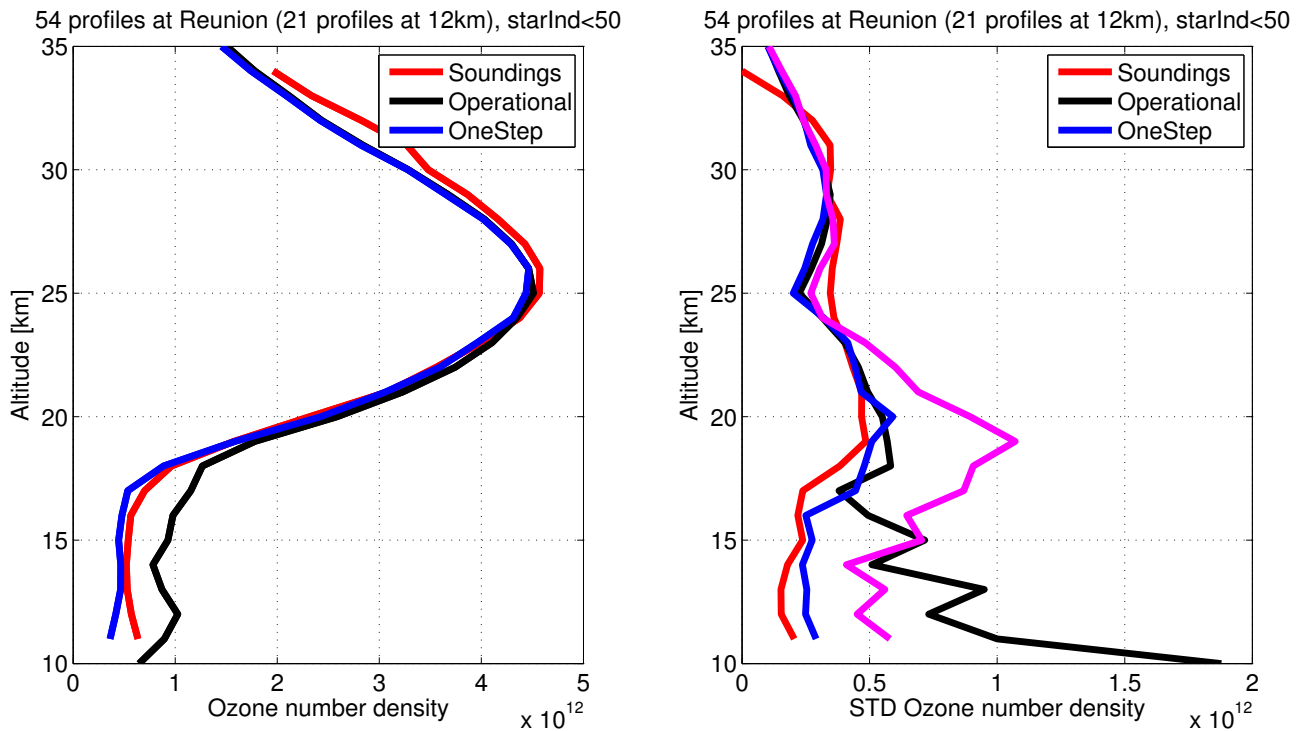


Figure 5: Reunion. Right: ozone profiles. One-step processed with the first version with $1/\lambda$ aerosol model. Left: standard deviation of ozone number densities. In magenta there is std of with the setup of [5] with relaxed aerosol priors.

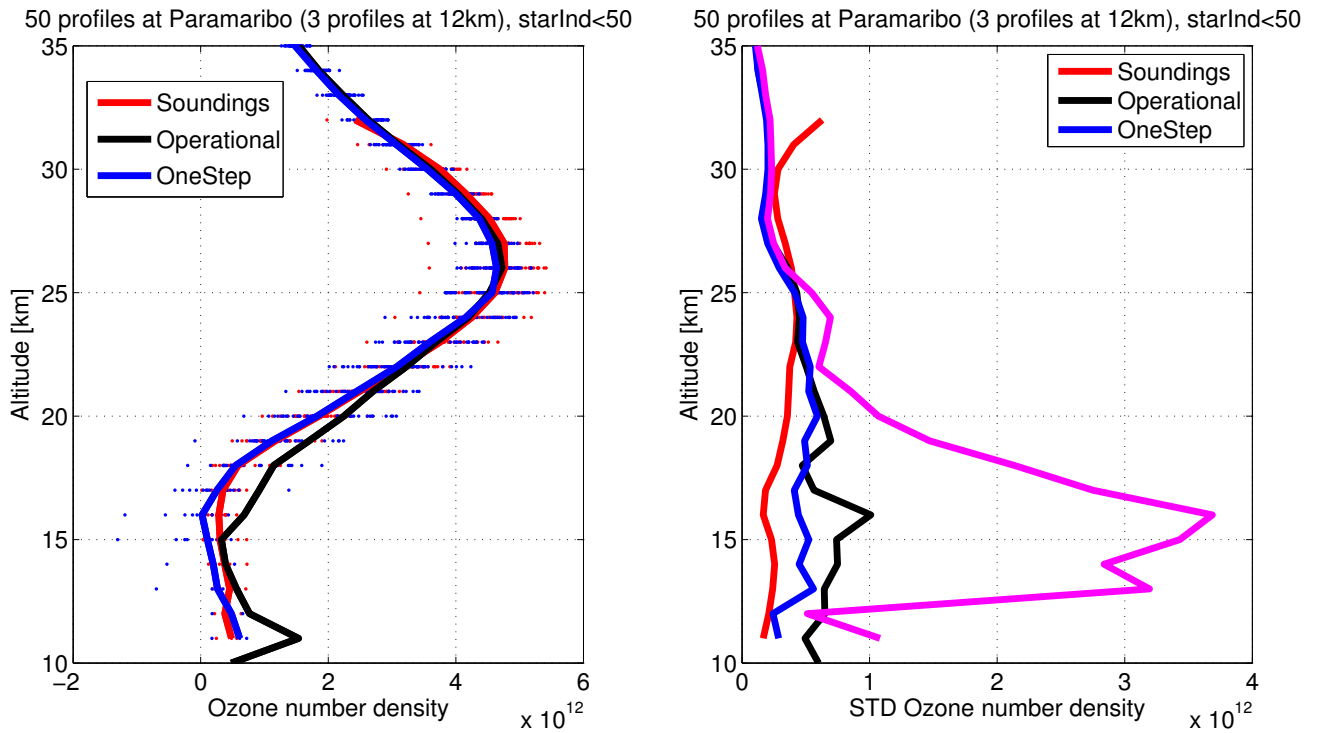


Figure 6: Paramaribo. Right: ozone profiles. One-step processed with the first version with $1/\lambda$ aerosol model. Left: standard deviation of ozone number densities. In magenta there is std of with the setup of [5] with relaxed aerosol priors.

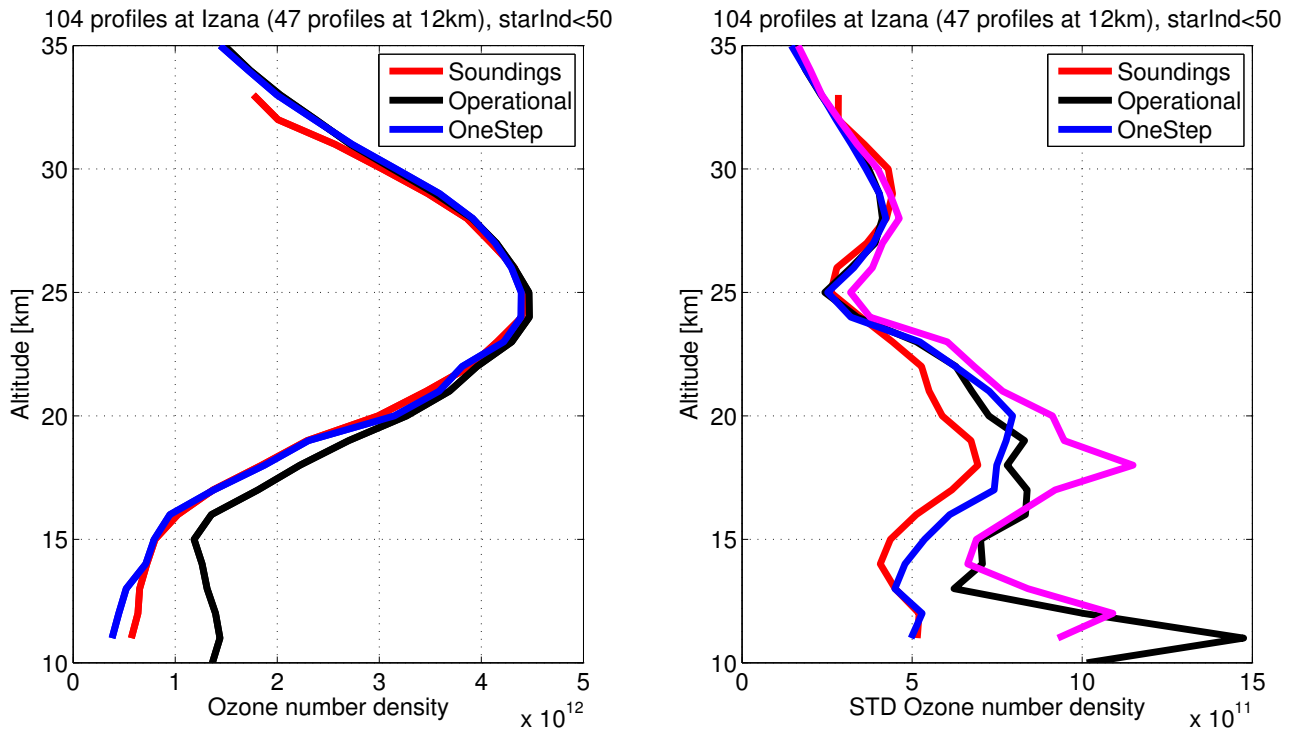


Figure 7: Izaña. Right: ozone profiles. One-step processed with the first version with $1/\lambda$ aerosol model. Left: standard deviation of ozone number densities. In magenta there is std of with the setup of [5] with relaxed aerosol priors.

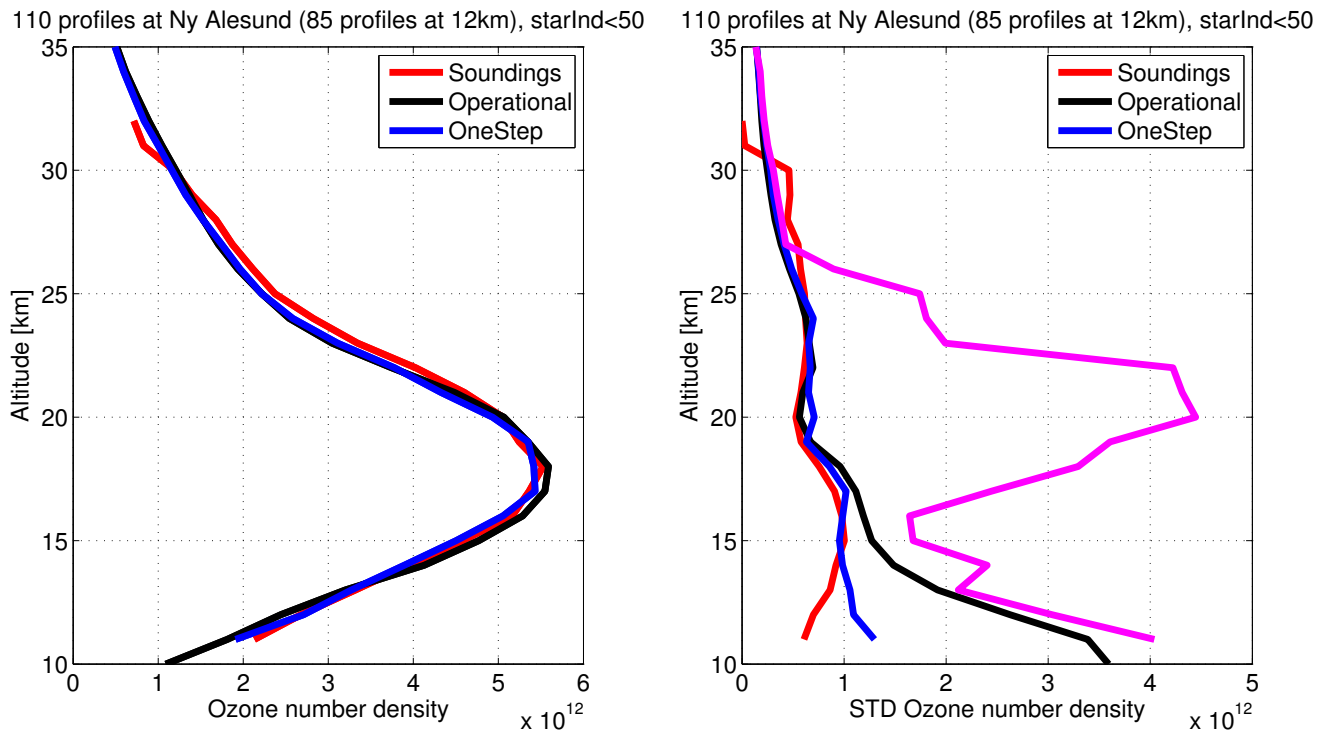


Figure 8: Ny Ålesund. Right: ozone profiles. One-step processed with the first version with $1/\lambda$ aerosol model. Left: standard deviation of ozone number densities. In magenta there is std of with the setup of [5] with relaxed aerosol priors.

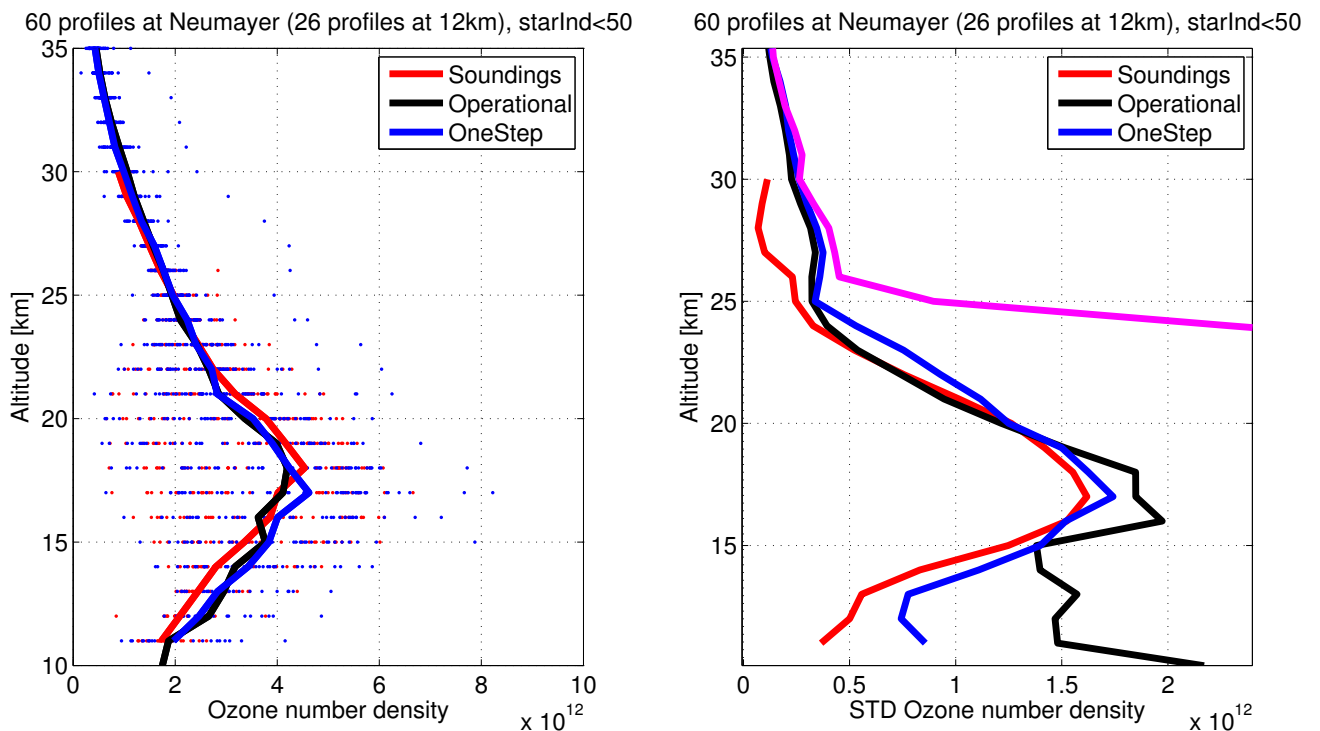


Figure 9: Neumayer. Right: ozone profiles. One-step processed with the first version with $1/\lambda$ aerosol model. Left: standard deviation of ozone number densities. In magenta there is std of with the setup of [5] with relaxed aerosol priors.

2.2 A search for geophysical features at ULTS altitudes

Soon after processing the results presented in the section before, it became clear that working only (with limited number of) collocations is not a good idea: Although, the NDACC stations give indeed very good reference. Just by looking at them, one can really “not see the forest for the trees”. Thus, in order to get the big picture a larger dataset needs to be created.

This is the reason why we decided to process with the best setup so far (presented in last section) of one-step algorithm all the occultations made with bright stars (star index < 50) in full dark ($SZA > 107$ degrees) from all latitudes.

During the mission 2002–2012 there are 178 313 such profiles. However, for technical reasons, we needed to set the maximum number of layers to 150. If this change was not made the processing would have stuck. In the end, 171 233 profiles (96%) got processed. The rejected profiles include the ones with number of layers higher than 150, but other processing failures existed too. This processing exercise was an important step, and it shows that the one-step algorithm can be used, if selected, in ALGOM Task 1.4.

In Fig. 14, we show the distribution of the GOMOS measurements in 2008 (black and green; note that the figure also illustrates OSIRIS data in 2008 in red). In Fig. 10 we compare these 2008 retrievals (right panel) to the McPeters and Labow (ML) climatology [9] (left panel). We observe that the overall features of these two climatologies are similar, but there is one latitude band where “things go wrong”. It is related to the well-known GOMOS bad star problem, and when new ALGOM bad star list was later applied, this problem went away. We observe negative bias (failed retrieval) at higher altitudes that also seem to cause positive bias at lower altitudes. This is further verified in Fig. 11, where relative difference between these two climatologies is presented.

Figure 12 studies the relative differences of two other GOMOS data products (IPF v6, AERGOM) with respect to the ML climatology in 2008. It can be seen that differences in IPF v6 and AERGOM data products are very high at UTLS altitudes. One-step ozone, on the other hand, shows negative relative differences at UTLS with respect to the ML climatology. Note that GOMOS one-step dataset was processed only for the 49 brightest stars.

A good check for any data product is to check if they can produce same (geophysical) features as other instruments. For example there might be a bias in the dataset, but it can still—qualitatively—give “good” results. This is a minimal requirement for a dataset to be useful in application studies. Fig. 13 show the relative differences between the 2008-datasets of MIPAS, SCIAMACHY, OSIRIS and ACE ozone with respect to the ML climatology. We note that they all show more or less same geophysical features. The main features include two “poles” at about 25 degrees North and South. Unfortunately, only faint hints of the same features can be found from the GOMOS one-step dataset, see Fig. 11.

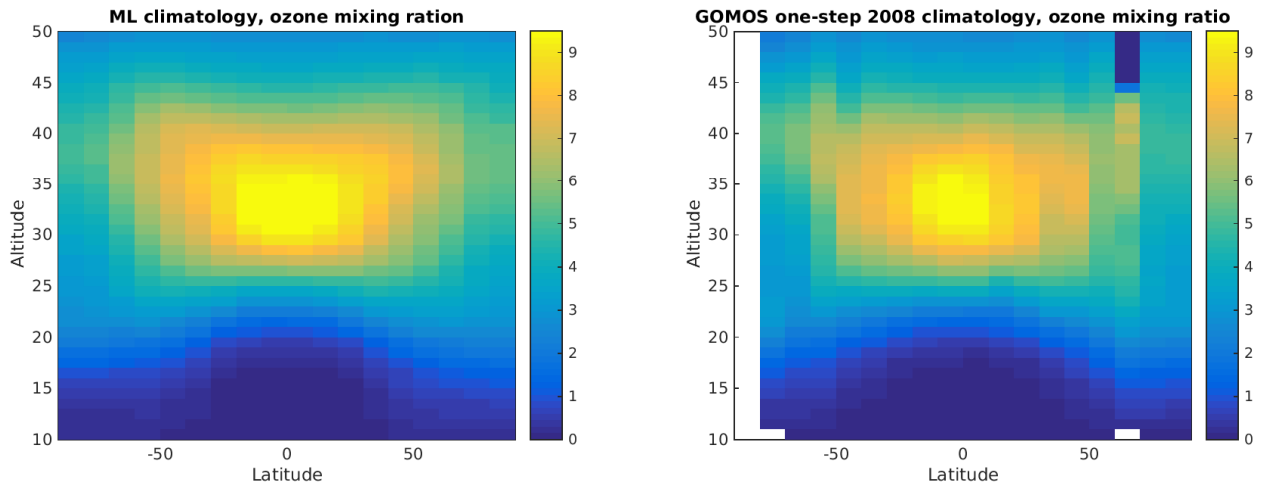


Figure 10: Right: ML climatology. Left: GOMOS one-step 2008 climatology.

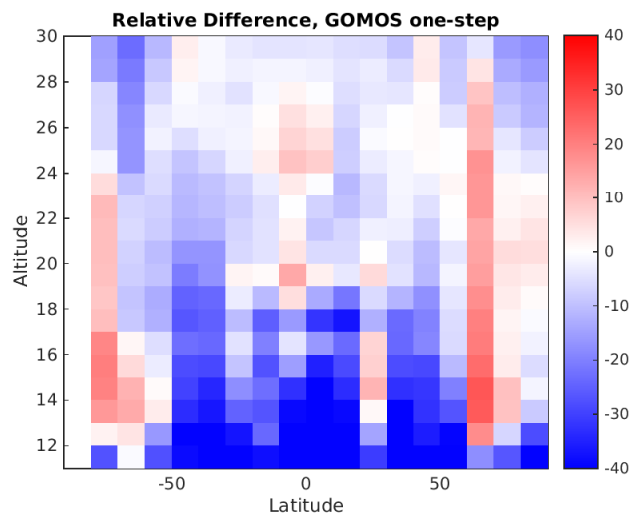


Figure 11: Relative differences [%] between the ML climatology and GOMOS 2008 one-step climatology. The two climatologies are presented in Fig. 10.

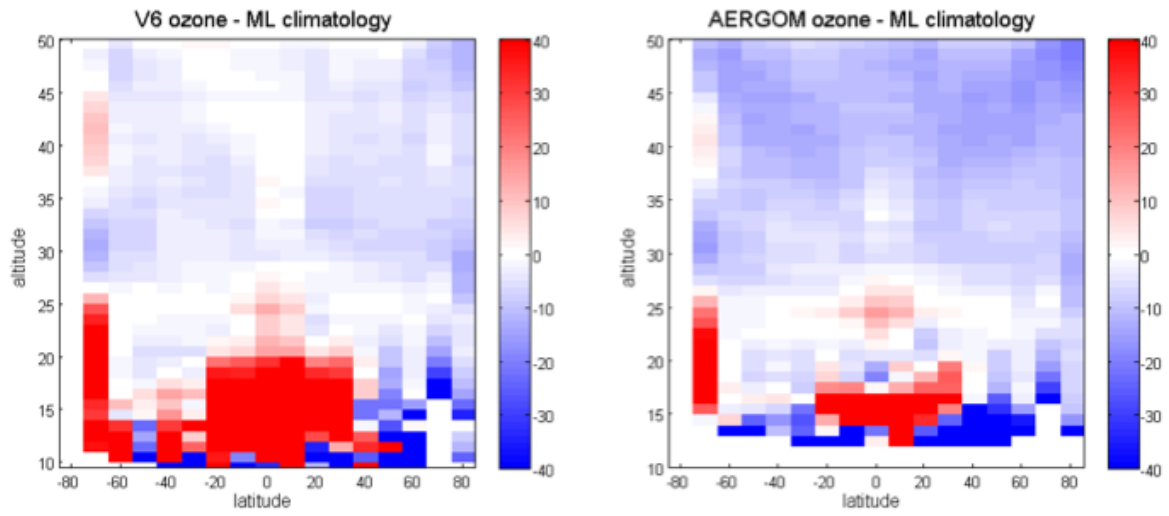


Figure 12: Relative differences [%] between the ML climatology and GOMOS 2008 climatologies of IPF v6 (left) and AERGOM (right).

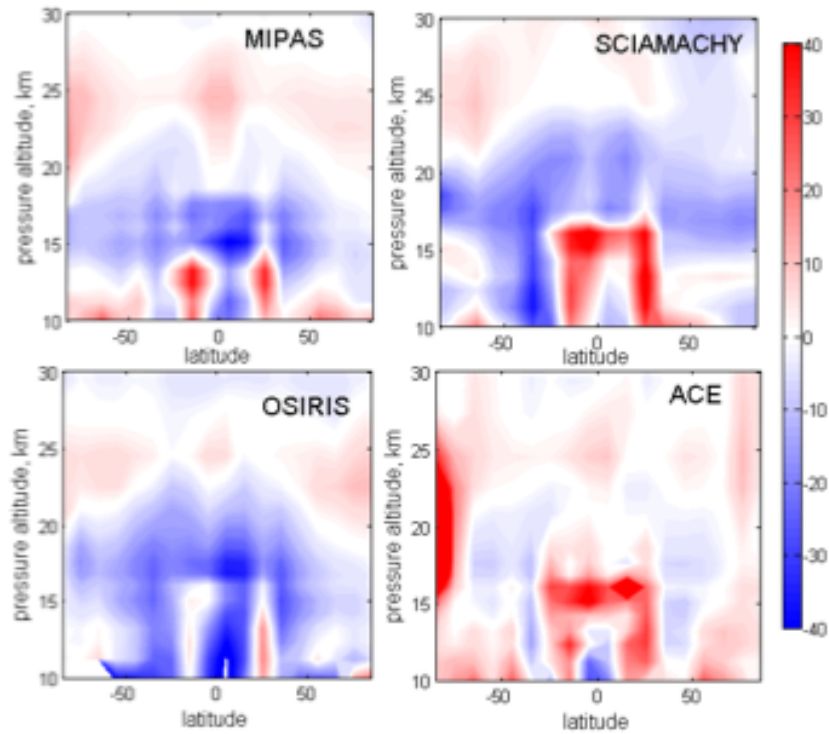


Figure 13: Relative differences [%] between the ML climatology and 2008 climatologies of MIPAS, SCIAMACHY, OSIRIS and ACE.

2.3 Validation of GOMOS one-step retrievals using OSIRIS dataset

One of the best ways to get good global idea of the quality a satellite-borne dataset is to compare it against another global satellite-borne dataset. One of the best global dataset of UTLS ozone is produced by OSIRIS (processed in the University of Saskatchewan) [3, 1].

Figure 14 shows the distribution of OSIRIS (red) and GOMOS (black) measurements in 2008. In order to find collocations between GOMOS and OSIRIS, we set the maximum longitude and latitude difference to 5 and 2 degrees, respectively. The time difference is required to be less than 12 hours. In Fig 14, the OSIRIS collocated GOMOS measurements are plotted in green. We note that as GOMOS provides useful measures in dark, and OSIRIS through sunrise/sunset. Hence, the datasets complements each other and not many collocations are found in polar regions.

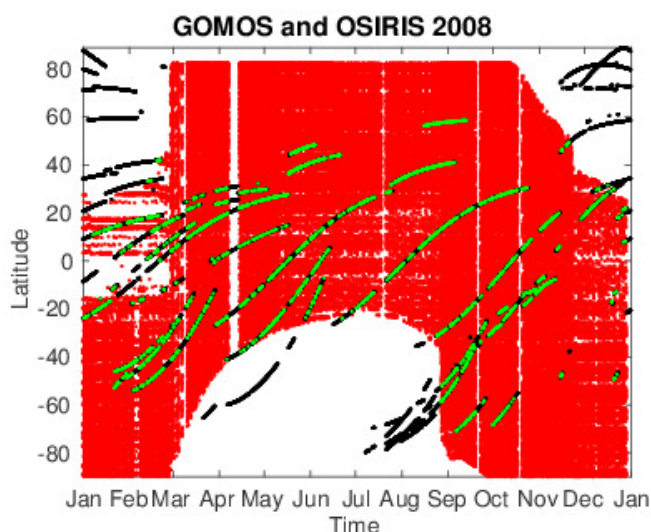


Figure 14: Distribution of OSIRIS (red) and GOMOS (black) data in 2008. OSIRIS collocated GOMOS points are in green.

In this verification study, we divide the collations in ten-degree latitude bands. Sample animations from the years 2002, 2004 and 2008 can be found from these links (respectively)

- <https://dl.dropboxusercontent.com/u/35228286/ALGOMTN/01d0ne2002.gif>
- <https://dl.dropboxusercontent.com/u/35228286/ALGOMTN/01d0ne2004.gif>
- <https://dl.dropboxusercontent.com/u/35228286/ALGOMTN/01d0ne2008.gif>

Other years produce similar results. Similar animations were also presented in ALGOM progress meeting number 2 and the presentation is available through GOMOS QWG FTP-site. It can be noted that “good” overall agreement (see animations, not appropriate to give exact numbers) between GOMOS and OSIRIS can be found at UTLS altitudes. However, still clear error structures can be observed. The improvement of the quality of the UTLS ozone with respect to the IPF v6 is drastic, in particular in the tropics.

2.4 Verification of the new setup

After the creation of the big dataset of about 171 000 profiles, we decided to study the following changes to the algorithm:

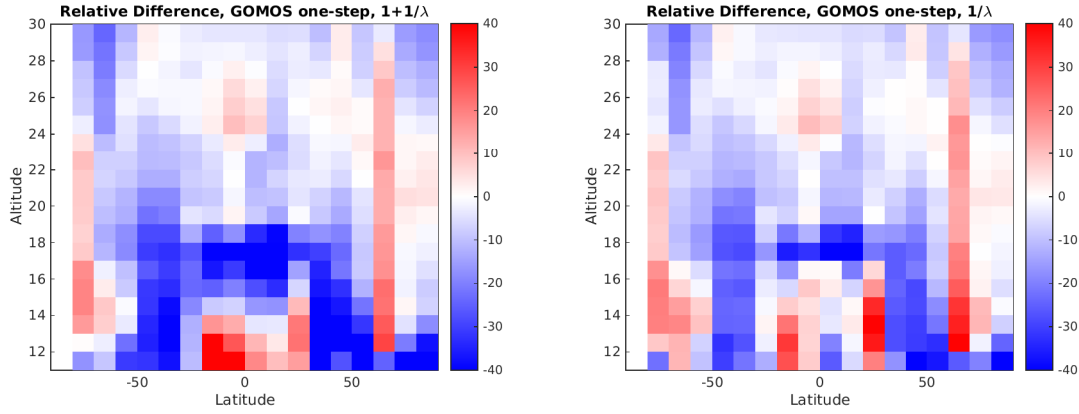


Figure 15: Relative differences [%] between the ML climatology and GOMOS 2008 one-step climatologies processed with the latest setup using “ $1/\lambda$ ” (right panel) and “ $1 + 1/\lambda$ ” aerosol models (left panel).

- We implemented “ $1 + 1/\lambda$ ” aerosol model. This model describes better the variety of possible aerosol extinction spectra. Hence, this model is expect to be slightly more “physical”.
- We “Fixed” neutral air retrieval. This change was made, because of non-orthogonality of aerosol and Rayleigh cross-sections. We also gain a degree of freedom.
- We removed the last layer from retrieval. This change was made in order to avoid outliers due to possible error in star tracking at the very end of the occultation.
- Implemented total variation (L1 norm) solution to two-step algorithm. This could help the one-step retrieval too as we get updated weights though the IRLS algorithm. This implementation however was a little bit unstable and did not help in all but trivial cases. This is why this addition was later dropped.

Using this setup we re-processed 23 000 measurements from the year 2008 (all latitudes, $SZA > 107^\circ$, star index < 49), see Fig. 14 for the distribution. The OSIRIS comparison results with this setup using “ $1 + 1/\lambda$ ” and “ $1/\lambda$ ” aerosol models are given in following animations (respectively):

- <https://dl.dropboxusercontent.com/u/35228286/ALGOMTN/OnePlusOneOverLambda.gif>
- <https://dl.dropboxusercontent.com/u/35228286/ALGOMTN/OneOverLambda.gif>

Note that the first animation includes also the operational dataset and the second animation show the comparisons with “ $1/\lambda$ ” and “ $1 + 1/\lambda$ ” aerosol-model dataset together.

We again note that the improvement of the quality of the UTLS ozone is drastic with respect to IPF v6. Both aerosol models, “ $1/\lambda$ ” and “ $1 + 1/\lambda$ ”, provide similar results. Although, some differences are observed. It cannot be concluded which one of them is “better” (i.e., closer to the OSIRIS UTLS ozone dataset).

In Fig. 15, we repeat the analysis of the previous section (comparison against ML climatology). This time, with this new setup, we can observe the same geophysical features that MIPAS, OSIRIS, SCIAMACHY and ACE observe, see Fig. 13. These features seem to be slightly better highlighted with “ $1/\lambda$ ” aerosol model (right panel) than with “ $1 + 1/\lambda$ ” (left panel).

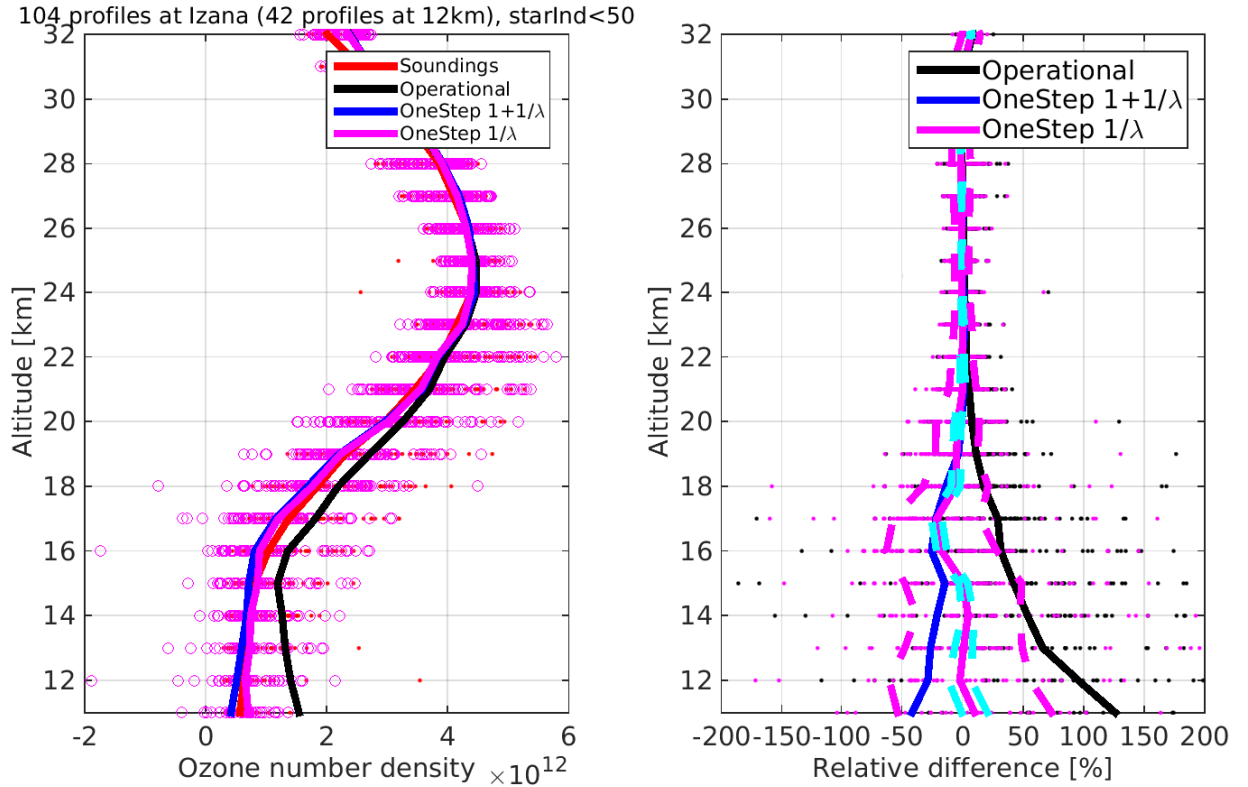


Figure 16: Izaña. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

We also repeat the NDACC validation study for five NDACC stations. Figures 16–20 present results using “ $1/\lambda$ ” (magenta) and “ $1 + 1/\lambda$ ” (blue). We observe that profiles processed with “ $1/\lambda$ ” aerosol model are closer to the sounding than the ones processed with “ $1 + 1/\lambda$ ”. In particular, this is highlighted at Izaña. In Appendix A we show these same figures, but with right panels zoomed to $\pm 50\%$. In addition, we show the relative errors with respect to tropopause. The appendix figures include also the results from Scoresbysund and Dumont.

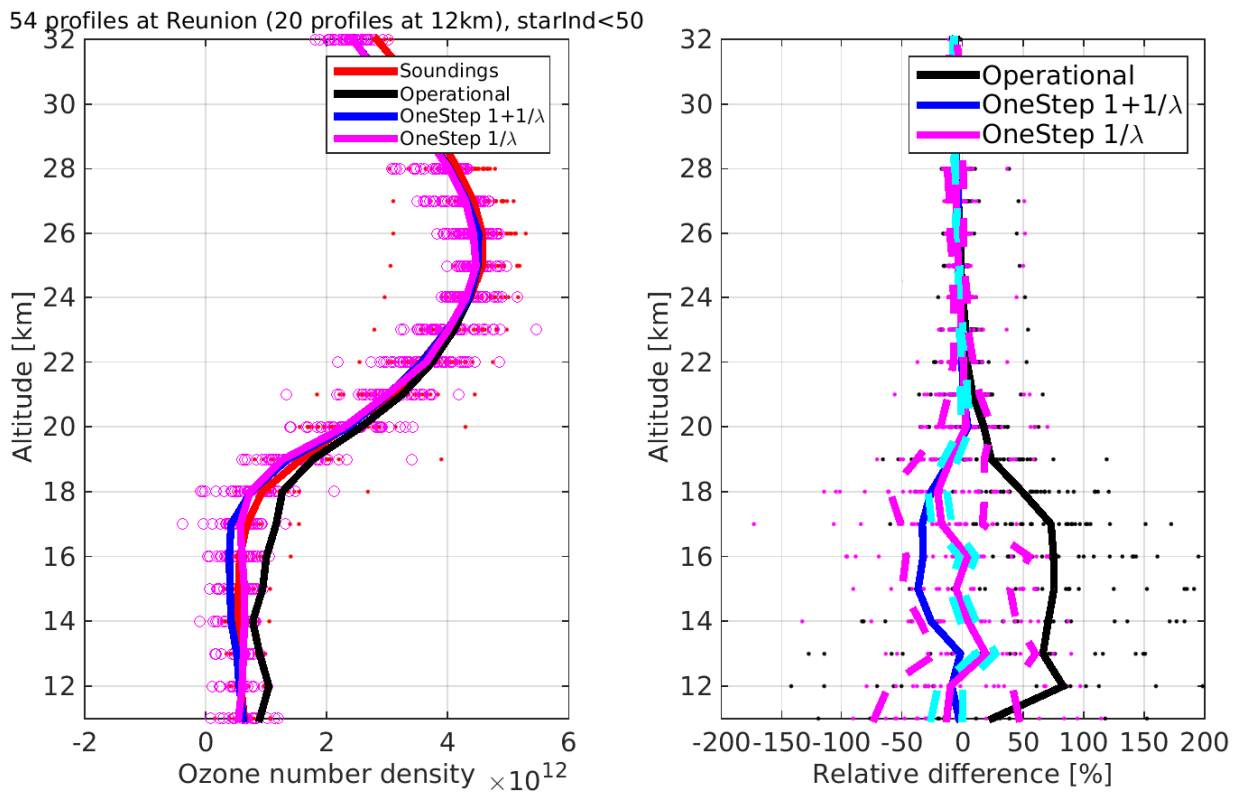


Figure 17: Reunion. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

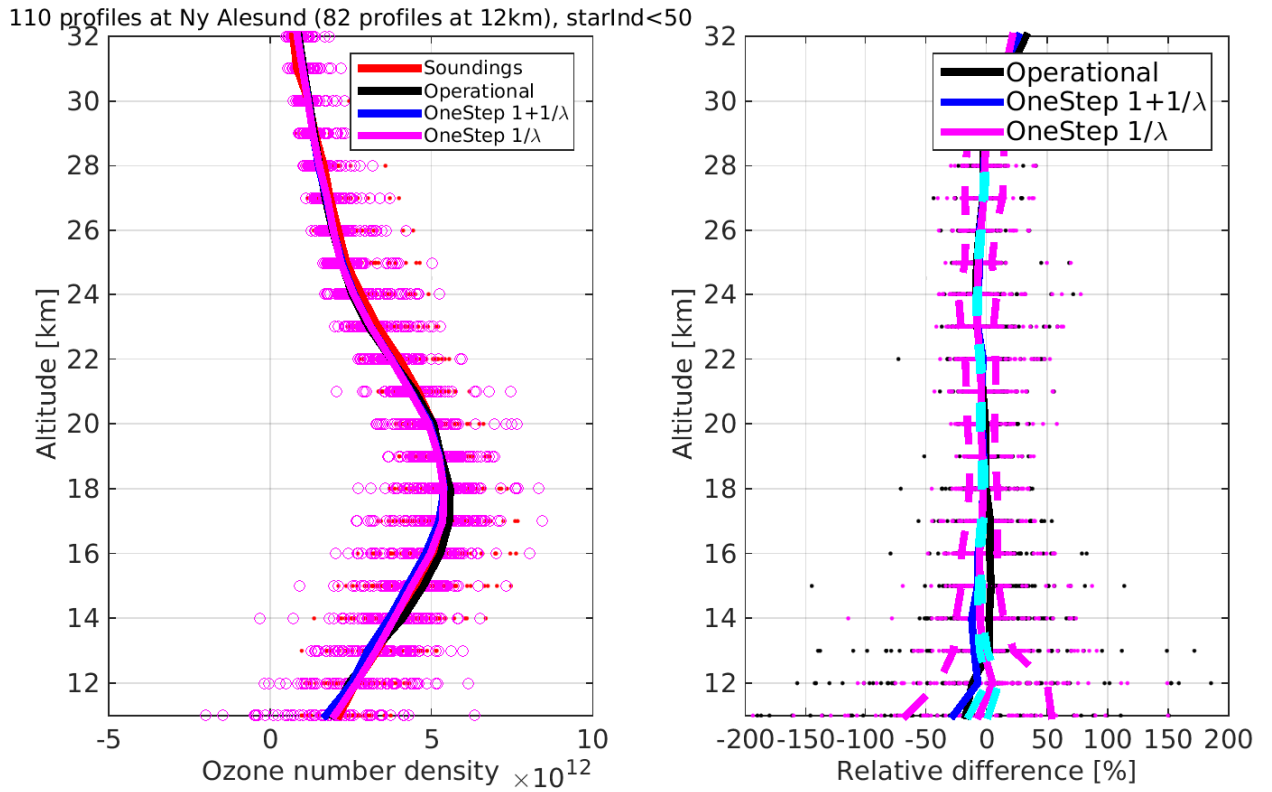


Figure 18: Ny Ålesund. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

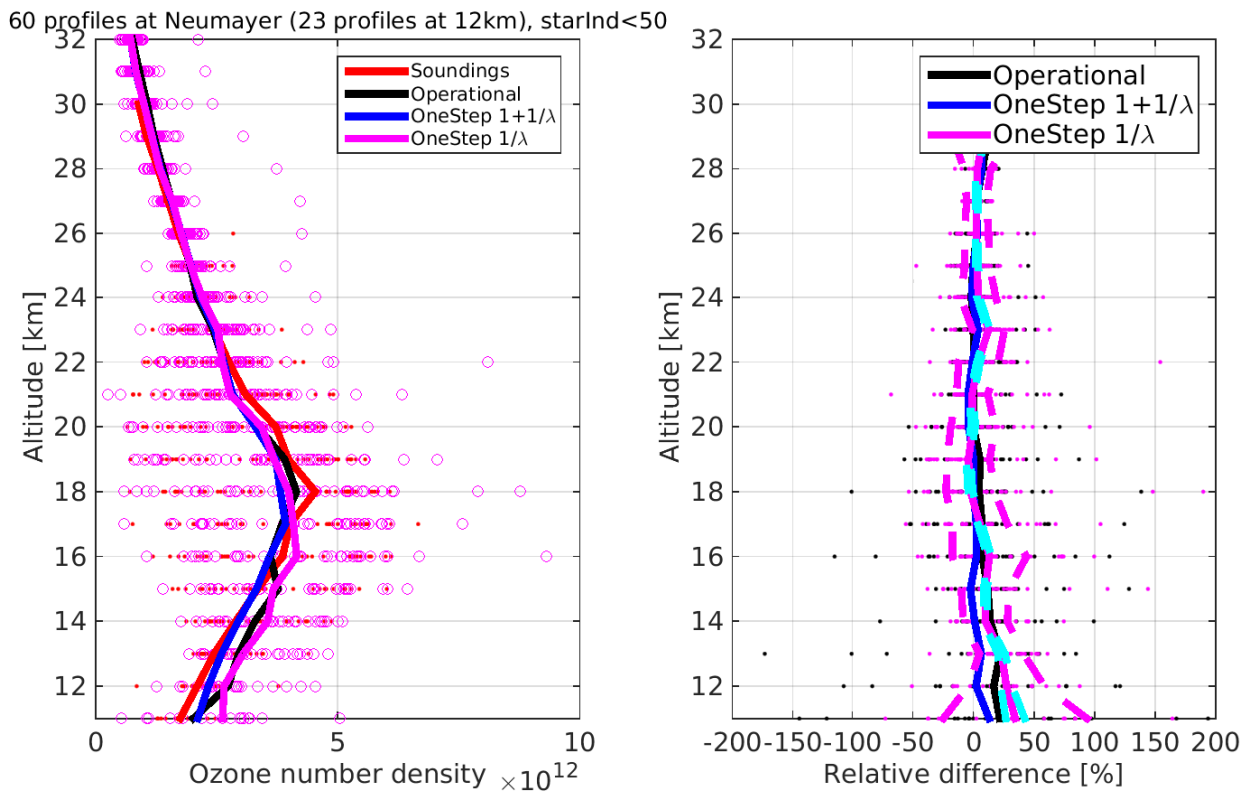


Figure 19: Neumayer. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

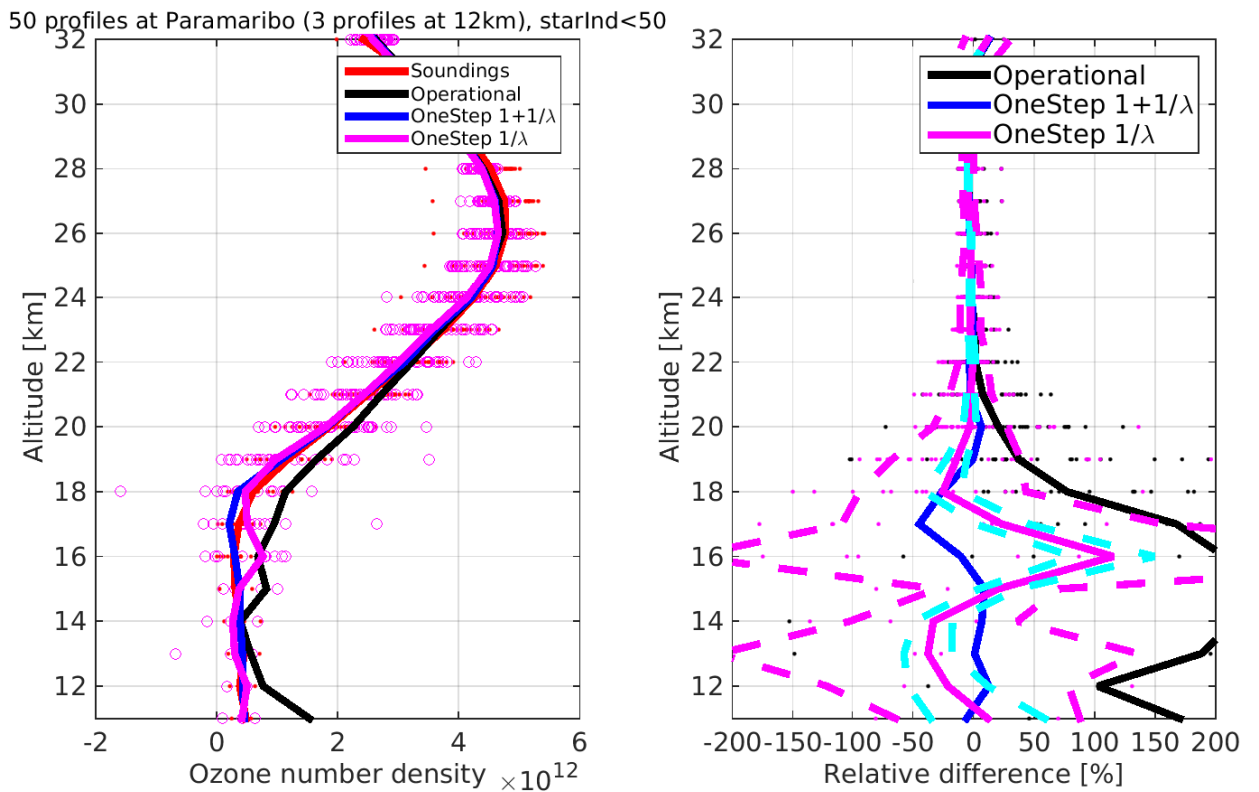


Figure 20: Paramaribo. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

2.5 Analysis of error estimates, vertical resolution and effect of regularization

Right panels of Figs. 16–20 present relative differences at five stations. See also Figs. 33–46 in Appendix A. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm. We note that the GOMOS one-step error estimates are too low. In Fig. 21, we plot for GOMOS, the sounding relative error estimates at Izaña for operational (black) and one-step (magenta) datasets. We observe that one-step error estimates are much smaller.

In order to analyze the error estimates in more detail, we test the effect of the so-called full-covariance matrix (FCM) and smoothes (2nd derivative). Figure 22 illustrates this analysis at Izaña. We note that at UTLS altitudes the largest effect to the error estimates comes from regularization—if we do not regularize error estimates come higher (and more realistic). In stratosphere, the effect of the FCM is proportionally larger (right panel). Figure 23 shows the effect of regularization and FCM—on average—at Izaña. On average, the effect to the relative difference is small, but not negligible.

As higher error values are desirable—it could be an option to process the data without any regularization at all. This would also give higher vertical resolution to the UTLS altitudes. In Figs. ??–26 we show the effect of regularization for three cases at Izaña (S002, S009 and S038, respectively). We note that the oscillation effect—caused by the measurement geometry—is clearly too large for the data to be used without averaging. Post retrieval smoothing—with created averaging kernels—is of course possible, but maybe not desirable.

As discussed in Sect. 1.4, in the one-step algorithm target resolution cannot be set, since the resolution depends on the occultation depended measurement noise. And in particular, the target resolution cannot be set, if the prior covariance matrix is kept fixed for all star types. In Figs. 27–29 we show the resolution for the same three cases using and not using the FCM. We note the following: resolution depends on star, and in particular on the noise of the signal. In addition, adding modeling error—with the FCM—worsens the resolution, as it should do. This is clearly visible at around 30 km. The reason for anomalous peaks at around 16 km is known and they are caused by strange signals. This effect will be shown in detail in ALGOM progress meeting number 3.

The vertical resolution presented in Figs. 27–29 was decided to be too high for the UTLS dataset and its applications. We decided to tune one parameter of Table 2. We set the ozone smoothness parameter at 15 km to 2×10^{11} . This yields the resolution comparable to the operational, see Figs. 30–32. The actual value still—of course—depends on the star type. We note however that the anomalous feature of the resolution at around 16 km is diminished. With this adjusted setup, we processed another sample dataset for the whole mission. It is available at <https://drive.google.com/file/d/0B96hSFTUTHmDZHhONTE2NOZKY1E/view?usp=sharing>.

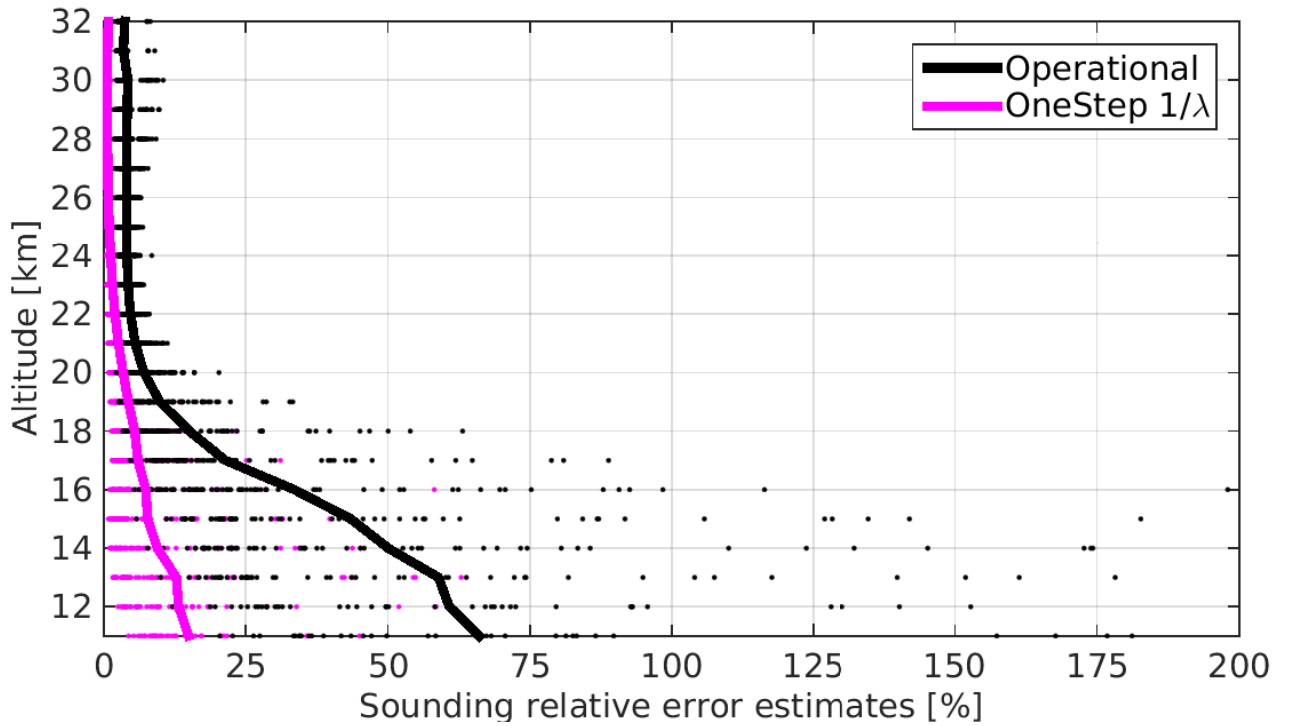


Figure 21: Sounding relative GOMOS error estimates at Izaña for operational (black) and one-step (magenta) dataset at Izaña.

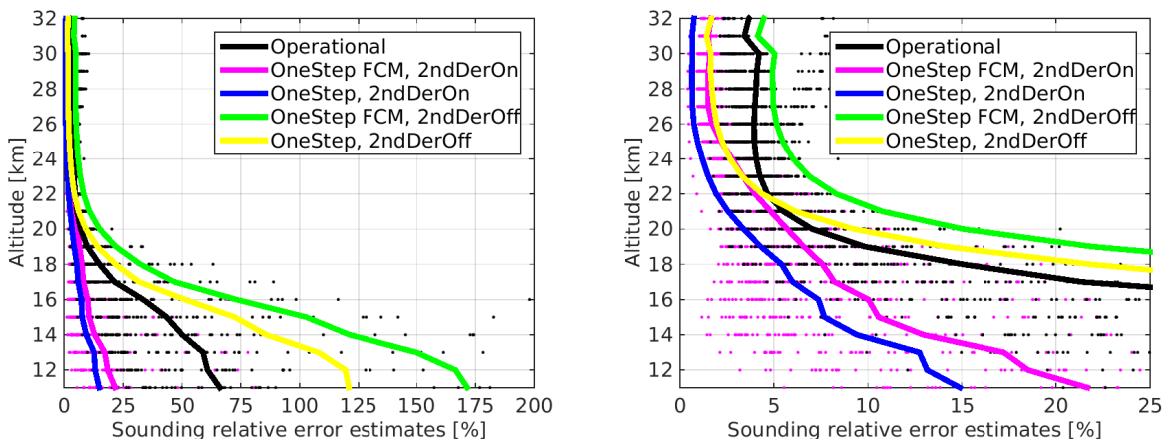


Figure 22: Sounding relative GOMOS error estimates at Izaña for several setups (see legend) at Izaña. Right panel is zoomed in order to see the stratospheric effect.

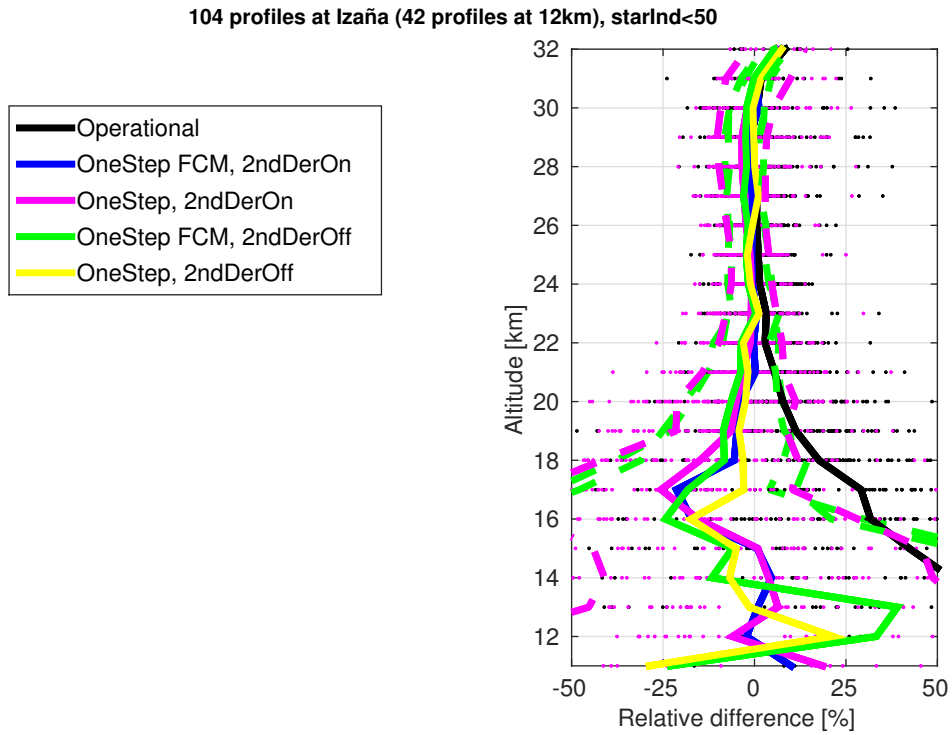


Figure 23: Effect of regularization and FCM to the validation results at Izaña.

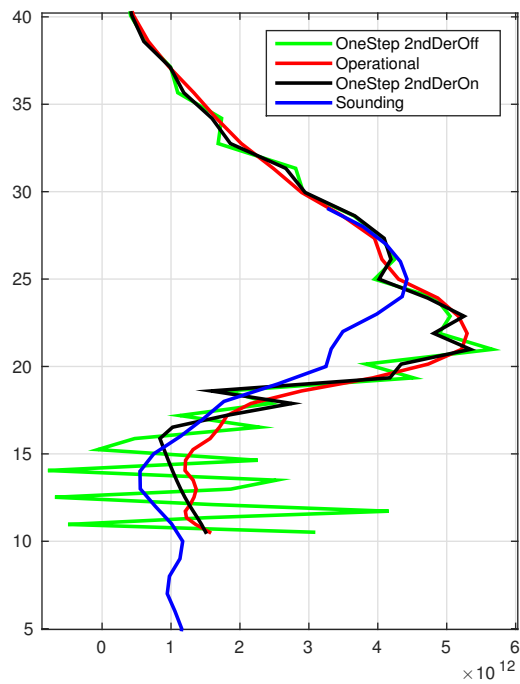


Figure 24: Effect of regularization to an individual occultation (S002). Collocation at Izaña.

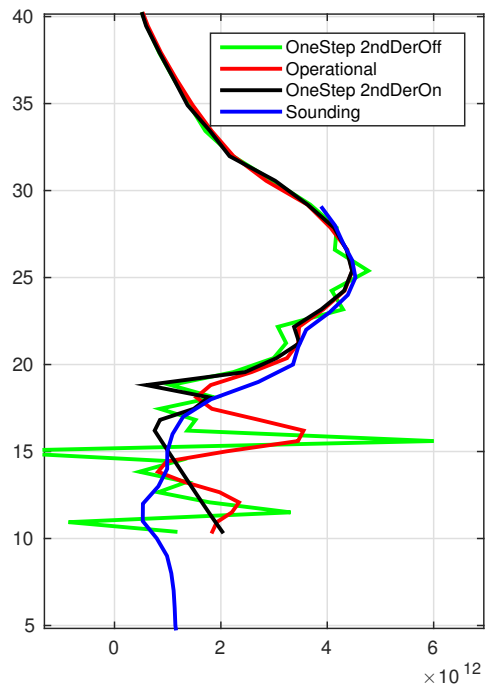


Figure 25: Effect of regularization to an individual occultation (S009). Collocation at Izaña.

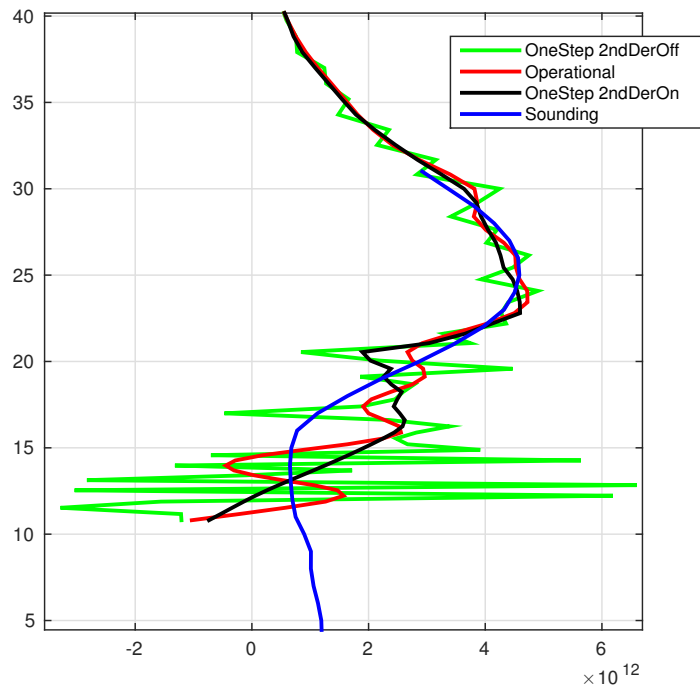


Figure 26: Effect of regularization to an individual occultation (S038). Collocation at Izaña.

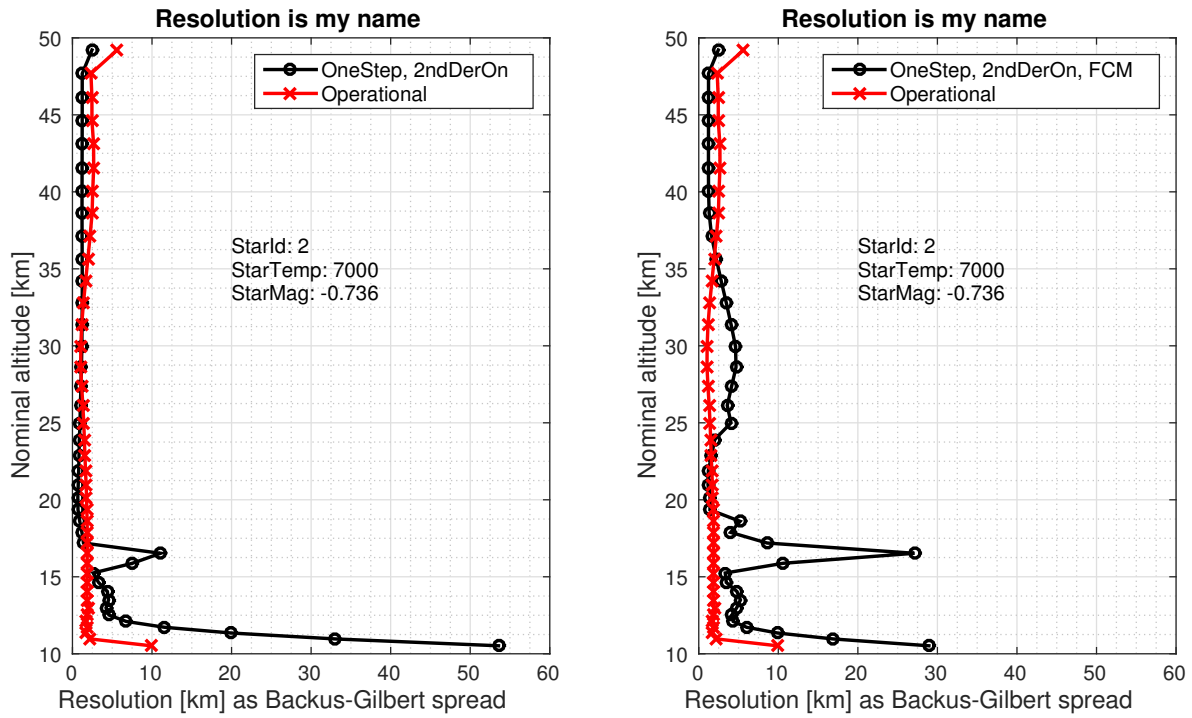


Figure 27: Resolution for S002 without (left) and with (right) full-covariance matrix.

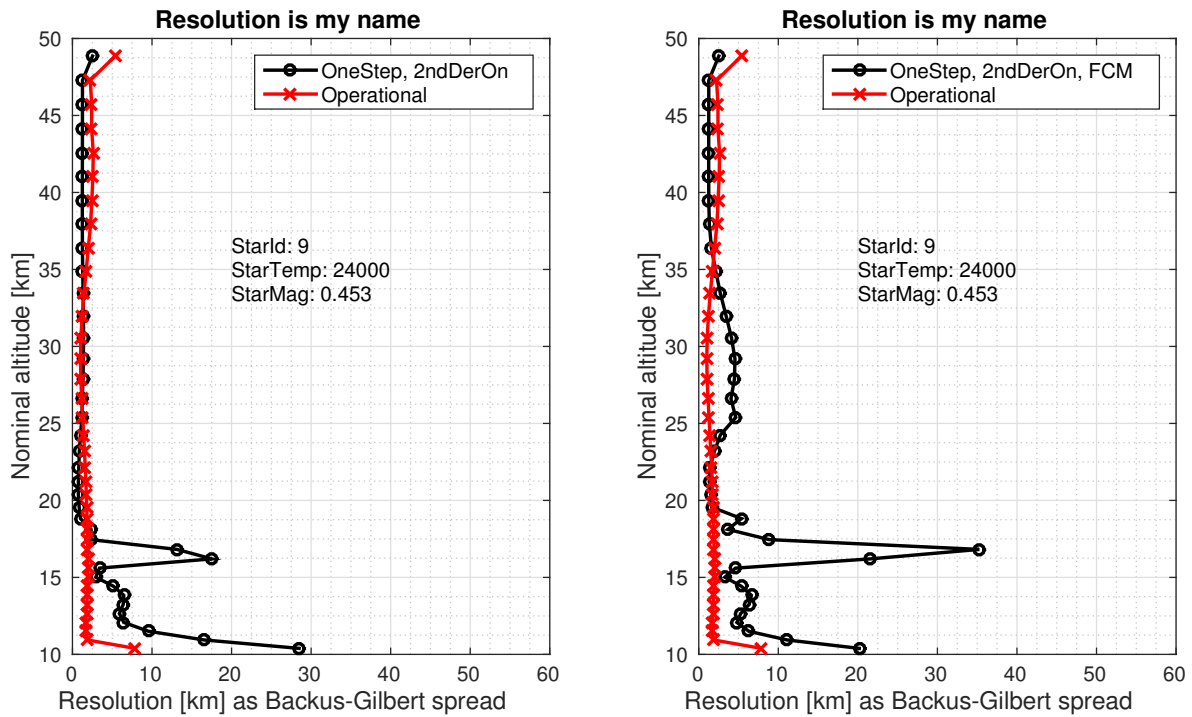


Figure 28: Resolution for S009 without (left) and with (right) full-covariance matrix.

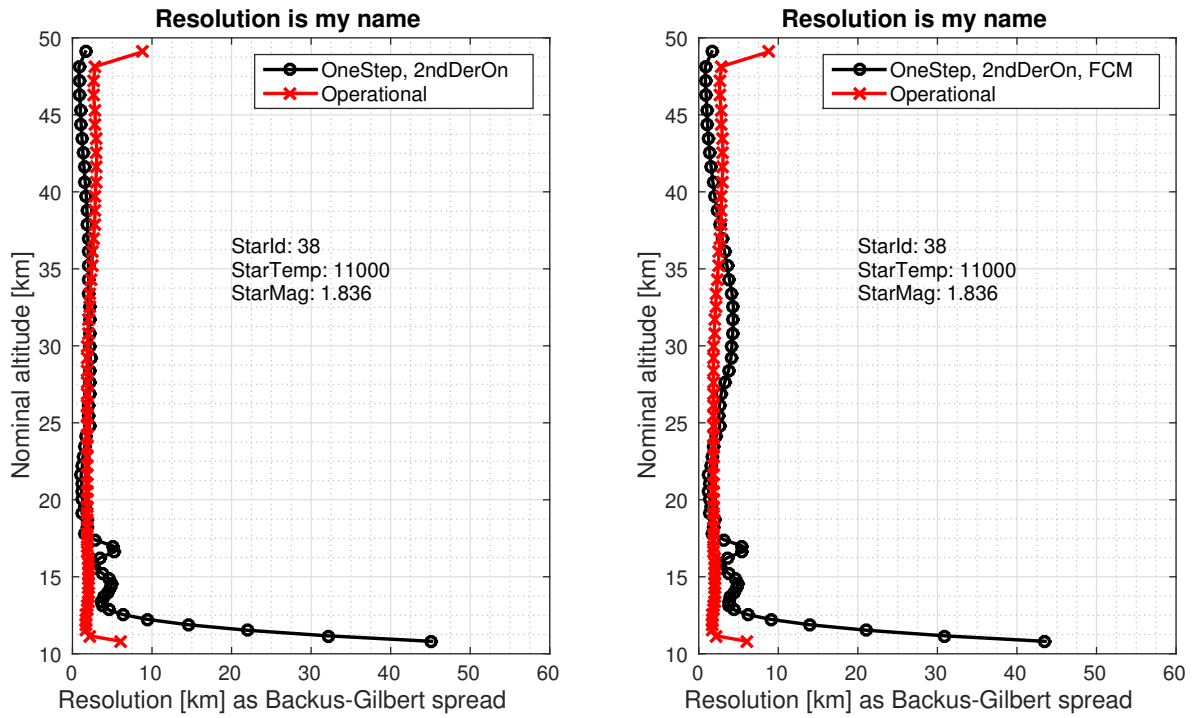


Figure 29: Resolution for S038 without (left) and with (right) full-covariance matrix.

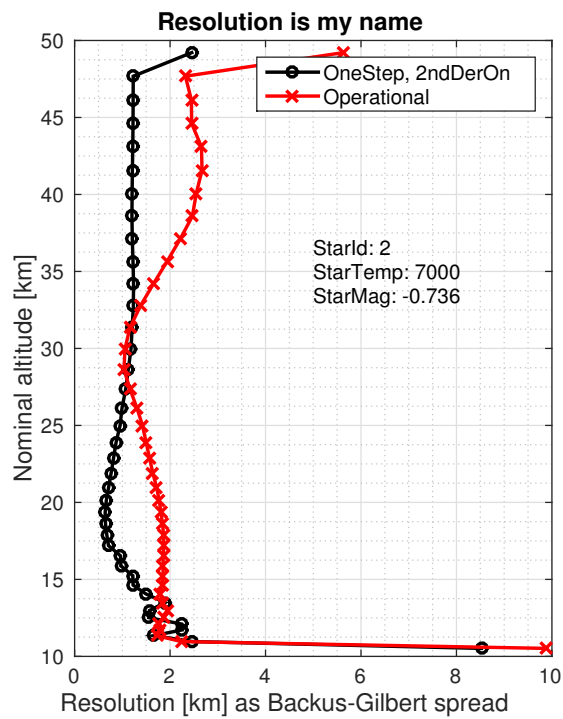


Figure 30: Resolution with adjusted setup (S002).

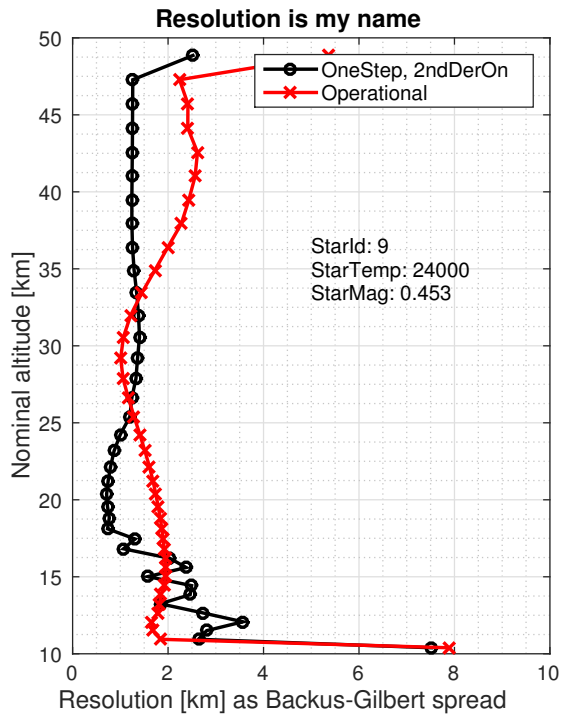


Figure 31: Resolution with adjusted setup (S009).

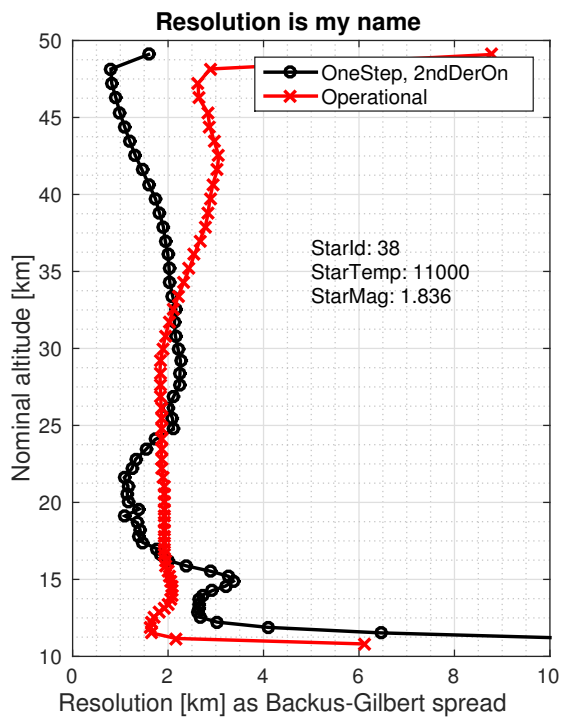


Figure 32: Resolution with adjusted setup (S038).

2.6 Conclusion from validation

From the analysis done earlier in this section, we can draw the following conclusions:

- Almost all one-step setups tested in the framework of the ALGOM project show—via geophysical validation—drastic improvement with respect to the IPF v6 UTLS dataset. The improvement is most pronounced in Tropics.
- The operational three-parameter aerosol model setup together with strong aerosol smoothes priors does improve the quality at some stations. However, in case of a clear aerosol layer, processing anomalies come. This is the reason why this setup is not recommended for the processing of the whole mission.
- We processed a big dataset of 171 000 profiles. This shows that processing all the approximately 400 000 full dark occultations of the whole mission is possible.
- Outlier still exist in the dataset. Also if the processing fails at upper layers of the atmosphere, it affects the UTLS layers too.
- With the latest setup, comparison against reliable OSIRIS dataset show massive improvement of the quality of the UTLS ozone with respect to IPF v6.
- With the latest setup we can see same geophysical features as MIPAS, SCIAMACHY, OSIRIS and ACE. This is related to the fact that we have fixed the air in the retrieval, since these features were not clearly visible in the earlier dataset. Also removing the lowest measurements help. The geophysical features can be seen better with “ $1/\lambda$ ” than with “ $1 + 1/\lambda$ ” aerosol model.
- NDACC validation of the latest version show that with the “ $1/\lambda$ ” we obtain closer results to soundings than with “ $1 + 1/\lambda$ ” aerosol model. Over all structure is close to zero and the median difference is always between $\pm 25\%$ (in all relevant cases, see Figs. 16–20 and the figures in Appendix A).
- Analysis of the GOMOS error estimates show that the errors given by the one-step algorithm are too low, and lower than operational ones.

Based on these conclusions, our recommendation is to produce the GOMOS one-step UTLS dataset with this latest setup using “ $1/\lambda$ ” aerosol model.

2.7 Recommendations for the next IPF processing

Based on our studies in this task, we make the following recommendations for the next IPF processing:

- Set aerosol model to $1/\lambda$. This the single most important update.
- Fix air in retrieval (if not already fixed).
- Remove the last occultation and the occultations below 10 km. This is based on “feeling” that strange values in the lower layers destroy the retrievals also on upper layers. This cannot be tested with the one-step algorithm and should be tested with the IPF algorithm.

3 One-step dataset of UTLS ozone

We have processed the entire 2002–2012 dataset for GOMOS with one-step algorithm. We have used the following limitations:

- Only the data in “full dark” – i.e., solar zenith angle $> 104^\circ$ – are processed.
- Processing is limited to the “good stars” only. Id est, all stars in new ALGOM “bad star list” are removed.
- In addition, also occultations that have no data below 50 km, or more than 150 lines are removed.

The whole dataset contains 243 883 profiles.

Key factors of this dataset are listed below:

- Gases: O_3 , (NO_2 , NO_3). Aerosol extinction at 500 nm.
- Measurements processed from 10 km to 100 km at occultation altitudes. Final layer is always removed from the retrieval.
- Aerosol model $1/\lambda$ is selected.
- Air is fixed to ECMWF value in retrieval.
- No absolute priors. No priors for first derivative.
- Smoother priors (second derivative) used are presented in Table 2. Modest smoothness requirements for ozone. Almost no regularization at all for aerosols.
- Variables are given in NetCDF files (one per occultation) are presented in Section 3.2.

Table 2: A priori standard deviation for second derivative (smoothness) per km.

Altitude [km]	O_3	NO_2	NO_3	Aerosol
1	2×10^8	10^8	10^6	10^{-2}
10	2×10^8	10^8	10^6	10^{-2}
15	2×10^{11}	10^8	10^6	10^{-2}
20	$2 \times 10^{11.5}$	10^8	10^6	10^{-2}
25	$2 \times 10^{10.5}$	10^8	10^6	10^{-2}
30	$10^{10.5}$	10^8	10^6	10^{-2}
40	$10^{10.5}$	10^8	10^6	10^{-2}
120	$10^{8.5}$	10^8	10^6	10^{-2}
160	$10^{8.5}$	10^8	10^6	10^{-2}

3.1 Disclaimer

- Although, NO_2 and NO_3 processed along ozone and aerosols they are not reported in data files, because we can not rigorously recommend using them for scientific analysis.
- We have validated only the “UTLS part” of the ozone profile. Other altitude regions have not been optimized, but are expected to have similar quality as IPF 6.01 ozone profiles. Same is true for aerosol extinctions.

3.2 Summary of variables in NetCDF files.

Output files:

- The NetCDF format follows the “user friendly” format created in ALGOM project.
- The filename syntax is “ESA_ALGOM-L2-GOMOS-FMI_onestep-,date,T,hourminsec-R,orbitnumber S,starnumber,-fv,fileversion,.nc”.
- Example filename: ESA_ALGOM-L2-GOMOS-FMI_onestep-20080820T013701-R33838-S001-fv001.nc.

Below there is a “ncdisp” print of a GOMS one-step NetCDF file.

Source:

ESA_ALGOM-L2-GOMOS-FMI_onestep-20080820T013701-R33838-S001-fv001.nc

Format:

netcdf4

Dimensions:

n_alt = 65

oneval = 1

Groups:

/geolocation_group/

Variables:

time

Size: 1x1

Dimensions: /oneval

Datatype: double

Attributes:

long_name = 'mean Modified Julian Date'

standard_name = 'time'

units = 'days since 1858-11-17 00:00:00'

description = 'Mean Modified
Julian Date between altitudes 20 and 50 km'

latitude

Size: 1x1

Dimensions: /oneval

Datatype: double

Attributes:

long_name = 'mean latitude'

standard_name = 'latitude'

units = 'degrees_north'

description = 'Mean latitude between
tangent altitudes 20 and 50 km'

longitude

Size: 1x1

Dimensions: /oneval

Datatype: double

Attributes:

long_name = 'mean longitude'

units = 'degrees_east'

standard_name = 'longitude'


```

description = 'Mean longitude between
              tangent altitudes 20 and 50 km'
time_start
  Size:      1x1
  Dimensions: /oneval
  Datatype:  double
  Attributes:
              long_name = 'first measurement Modified Julian Date'
              units     = 'days since 1858-11-17 00:00:00'
time_end
  Size:      1x1
  Dimensions: /oneval
  Datatype:  double
  Attributes:
              long_name = 'last measurement Modified Julian Date'
              units     = 'days since 1858-11-17 00:00:00'
latitude_start
  Size:      1x1
  Dimensions: /oneval
  Datatype:  double
  Attributes:
              long_name = 'first measurement latitude'
              units     = 'degrees'
latitude_end
  Size:      1x1
  Dimensions: /oneval
  Datatype:  double
  Attributes:
              units     = 'degrees'
              long_name = 'last measurement latitude'
longitude_start
  Size:      1x1
  Dimensions: /oneval
  Datatype:  double
  Attributes:
              long_name = 'first measurement longitude'
              units     = 'degrees'
longitude_end
  Size:      1x1
  Dimensions: /oneval
  Datatype:  double
  Attributes:
              long_name = 'last measurement longitude of occultation'
              units     = 'degrees'
altitude
  Size:      65x1
  Dimensions: /n_alt
  Datatype:  double
  Attributes:
              long_name = 'tangent height above mean sea level'

```

```

        standard_name = 'altitude'
        units          = 'km'
        description    = 'Tangent altitude'
altitude_min
    Size:            1x1
    Dimensions:      /oneval
    Datatype:        double
    Attributes:
        long_name    = 'minimum tangent altitude reached by the
                        occultation'
        units        = 'km'
altitude_parameters
    Size:            1x1
    Dimensions:      /oneval
    Datatype:        double
    Attributes:
        long_name    = 'parameter averaging altitude range'
        units        = 'km'
        description  = 'Mean tangent altitude for mean values of
                        parameters. Usually parameters are mean
                        over 20-50 km.'
duration
    Size:            1x1
    Dimensions:      /oneval
    Datatype:        double
    Attributes:
        long_name    = 'duration of the occultation'
        units        = 'sec'
obliquity
    Size:            1x1
    Dimensions:      /oneval
    Datatype:        double
    Attributes:
        long_name    = 'obliquity of the occultation'
        units        = 'degrees'
        description  = 'Obliquity of the occultation: Obliquity
                        is the angle between the vector velocity of the line
                        of sight in the atmosphere and the local vertical
                        at altitude 35 km.'

/radiation_group/
Variables:
    sza_tangentpoint
        Size:            1x1
        Dimensions:      /oneval
        Datatype:        double
        Attributes:
            long_name    = 'mean solar zenith angle
                            at tangent point'
            standard_name = 'solar_zenith_angle'

```

```

units          = 'degrees'
description    = 'Mean solar zenith angle
between tangent
altitudes 20 and 50 km at tangent point'
illumination_flag
  Size:        1x1
  Dimensions: /oneval
  Datatype:   int32
  Attributes:
    long_name   = 'illumination flag'
    description = 'Illumination conditions flag:
0=dark, 1=bright, 2=twilight,
3=stray light, 4=stray+twilight'
sza_satellite
  Size:        1x1
  Dimensions: /oneval
  Datatype:   double
  Attributes:
    long_name   = 'mean solar zenith angle
at satellite'
    units      = 'degrees'
    description = 'Mean solar zenith angle
at satellite location
between tangent altitudes 20 and 50 km'
saa_flag
  Size:        1x1
  Dimensions: /oneval
  Datatype:   int32
  Attributes:
    long:name   = 'south Atlantic Anomaly-flag'
    description = 'South Atlantic Anomaly-flag:
0=outside SAA, 1=inside SAA'

```

/startarget_group/

Variables:

```

star_id
  Size:        1x1
  Dimensions: /oneval
  Datatype:   int32
  Attributes:
    long_name   = 'star identification number'
    description = 'Star number in the GOMOS
star catalogue'
star_temperature
  Size:        1x1
  Dimensions: /oneval
  Datatype:   double
  Attributes:
    long_name   = 'star effective temperature'
    units      = 'K'

```

```

description = 'Target star effective
temperature in
the GOMOS star catalogue'
star_magnitude
Size: 1x1
Dimensions: /oneval
Datatype: double
Attributes:
long_name = 'star visual magnitude'
description = 'Star visual magnitude in
the GOMOS star catalogue'

/o3_density_group/
Variables:
o3_density
Size: 65x1
Dimensions: /n_alt
Datatype: double
Attributes:
long_name = 'ozone number density
at tangent altitude'
units = 'cm-3'

o3_density_std
Size: 65x1
Dimensions: /n_alt
Datatype: double
Attributes:
long_name = 'ozone number density
error estimate
at tangent altitude'
units = 'cm-3'

o3_vertres
Size: 65x1
Dimensions: /n_alt
Datatype: double
Attributes:
long_name = 'o3 vertical resolution'
units = 'km'
description = 'O3 vertical resolution'

/no2_density_group/
Variables:
no2_density
Size: 65x1
Dimensions: /n_alt
Datatype: double
Attributes:
long_name = 'nitrogen dioxide number density
at tangent altitude'
units = 'cm-3'

```

```

no2_density_std
    Size:          65x1
    Dimensions:    /n_alt
    Datatype:      double
    Attributes:
        long_name = 'nitrogen dioxide number density
                    error estimate at tangent altitude'
        units     = 'cm-3'

no2_vertres
    Size:          65x1
    Dimensions:    /n_alt
    Datatype:      double
    Attributes:
        long_name  = 'no2 vertical resolution'
        units      = 'km'
        description = 'NO2 vertical resolution'

/no3_density_group/
  Variables:
    no3_density
        Size:          65x1
        Dimensions:    /n_alt
        Datatype:      double
        Attributes:
            long_name = 'nitrogen trioxide number density
                        at tangent altitude'
            units     = 'cm-3'

    no3_density_std
        Size:          65x1
        Dimensions:    /n_alt
        Datatype:      double
        Attributes:
            long_name = 'nitrogen trioxide
                        number density
                        error estimate at tangent altitude'
            units     = 'cm-3'

    no3_vertres
        Size:          65x1
        Dimensions:    /n_alt
        Datatype:      double
        Attributes:
            long_name  = 'no3 vertical resolution'
            units      = 'km'
            description = 'NO3 vertical resolution'

/aerosol_group/
  Variables:
    aerext_500
        Size:          65x1
        Dimensions:    /n_alt

```

```

    Datatype: double
    Attributes:
        lon_name = 'aerosol extinction at 500 nm
        at tangent altitude'
        units    = '1/km'
aerext_500_std
    Size:      65x1
    Dimensions: /n_alt
    Datatype:  double
    Attributes:
        long_name = 'aerosol extinction at 500 nm
        error estimate at tangent altitude'
        units     = '%'
aerext_500_vertres
    Size:      65x1
    Dimensions: /n_alt
    Datatype:  double
    Attributes:
        long_name  = 'aerosol extinction
        at 500 nm vertical resolution'
        units      = 'km'
        description = 'Aerosol extinction
        at 500 nm vertical resolution'

```

/retrieval_quality_group/

Variables:

```

    chi2
        Size:      65x1
        Dimensions: /n_alt
        Datatype:  double
        Attributes:
            long_name  = 'chi2 '
            description = 'Chi2 from fit (normalised by
            the degrees of freedom)'

```

/aprior_data_group/

Variables:

```

    air_density_ecmwf
        Size:      65x1
        Dimensions: /n_alt
        Datatype:  double
        Attributes:
            long_name  = 'a priori neutral density profile'
            units      = 'cm-3'
            description = 'Neutral density
            from ECMWF (below
            altitudes 1 hPa) and MSIS90
            (above altitudes 1 hPa)
            at tangent altitude'
    air_pressure_ecmwf

```

```

Size:          65x1
Dimensions:    /n_alt
Datatype:     double
Attributes:
    long_name   = 'a priori pressure profile'
    units       = 'hPa'
    description = 'Pressure from ECMWF
                  (below altitudes 1 hPa)
                  and MSIS90 (above altitudes 1 hPa)
                  at tangent altitude'
air_temperature_ecmwf
Size:          65x1
Dimensions:    /n_alt
Datatype:     double
Attributes:
    long_name   = 'a priori temperature profile'
    units       = 'K'
    description = 'Temperature from ECMWF
                  (below altitudes 1 hPa)
                  and MSIS90 (above altitudes 1 hPa) at tangent altitude'

```

/satellite_geolocation_group/

Variables:

```

orbit_number
Size:          1x1
Dimensions:    /oneval
Datatype:     int32
Attributes:
    long_name   = 'ENVISAT orbit number'
latitude_satellite
Size:          1x1
Dimensions:    /oneval
Datatype:     double
Attributes:
    long_name   = 'mean satellite latitude'
    units       = 'degrees'
    description = 'Mean latitude of satellite between
                  tangent altitudes 20 and 50 km'
longitude_satellite
Size:          1x1
Dimensions:    /oneval
Datatype:     double
Attributes:
    long_name   = 'mean satellite longitude'
    units       = 'degrees'
    description = 'Mean longitude of satellite between
                  tangent altitudes 20 and 50 km'
latitude_satellite_start
Size:          1x1
Dimensions:    /oneval

```

```

Datatype: double
Attributes:
    long_name = 'first measurement satellite latitude'
    units     = 'degrees'
latitude_satellite_end
Size:        1x1
Dimensions: /oneval
Datatype:   double
Attributes:
    long_name = 'last measurement satellite latitude'
    units     = 'degrees'
longitude_satellite_start
Size:        1x1
Dimensions: /oneval
Datatype:   double
Attributes:
    long_name = 'first measurement satellite longitude'
    units     = 'degrees'
longitude_satellite_end
Size:        1x1
Dimensions: /oneval
Datatype:   double
Attributes:
    long_name = 'last measurement satellite longitude'
    units     = 'degrees'

```

/metadata_group/

```

Attributes:
    title = 'GOMOS One-step UTLS data'
    source_file_ext = 'GOM_EXT_2PRFIN20080820_013701_000000442071_00217_33838_7657.N1'
    source_file_nl = 'GOM_NL__2PRFIN20080820_013701_000000442071_00217_33838_7657.N1'
    GOMOS_OnestepUTLS_dataversion = '1.02'
    GOMOS_IPF_dataversion = 'GOMOS/6.01'
    file_creation_date = '12-Jan-2016 04:23:16'
    file_created_by = 'Janne Hakkarainen,
    janne.hakkarainen@fmi.fi'
    project = 'ESA-ALGOM'
    institute = 'Finnish Meteorological Institute'
    value_for_nodata = 'NaN'
    orbit_number = 33838
    platform = 'ENVISAT'
    instrument = 'GOMOS'

```


References

- [1] C. Adams, A. E. Bourassa, V. Sofieva, L. Froidevaux, C. A. McLinden, D. Hubert, J.-C. Lambert, C. E. Sioris, and D. A. Degenstein. Assessment of odin-osiris ozone measurements from 2001 to the present using mls, gomos, and ozonesondes. *Atmos. Meas. Tech.*, 7(1):49–64, 2014.
- [2] 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.
- [3] Matthew Cooper, Randall V. Martin, Bastien Sauvage, Chris D. Boone, Kaley A. Walker, Peter F. Bernath, Chris A. McLinden, Doug A. Degenstein, Andreas Volz-Thomas, and Catherine Wespes. Evaluation of ace-fits and osiris satellite retrievals of ozone and nitric acid in the tropical upper troposphere: Application to ozone production efficiency. *J. Geophys. Res.*, 116(D12), 2011. D12306.
- [4] 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.
- [5] Janne Hakkarainen. *On state and parameter estimation in chaotic systems*. PhD thesis, Lappeenranta University of Technology, 2013.
- [6] 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 O₃, NO₂, and NO₃ observations in 2002-2008. *Atmos. Chem. Phys.*, 10(16):7723–7738, 2010.
- [7] 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.
- [8] P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Chapman & Hall, London, second edition, 1989.
- [9] Richard D. McPeters and Gordon J. Labow. Climatology 2011: An MLS and sonde derived ozone climatology for satellite retrieval algorithms. *J. Geophys. Res.*, 117(D10), 2012. D10303.
- [10] Clive D. Rodgers. *Inverse methods for atmospheric sounding: theory and practice*. World Scientific, 2000.
- [11] 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.
- [12] 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.

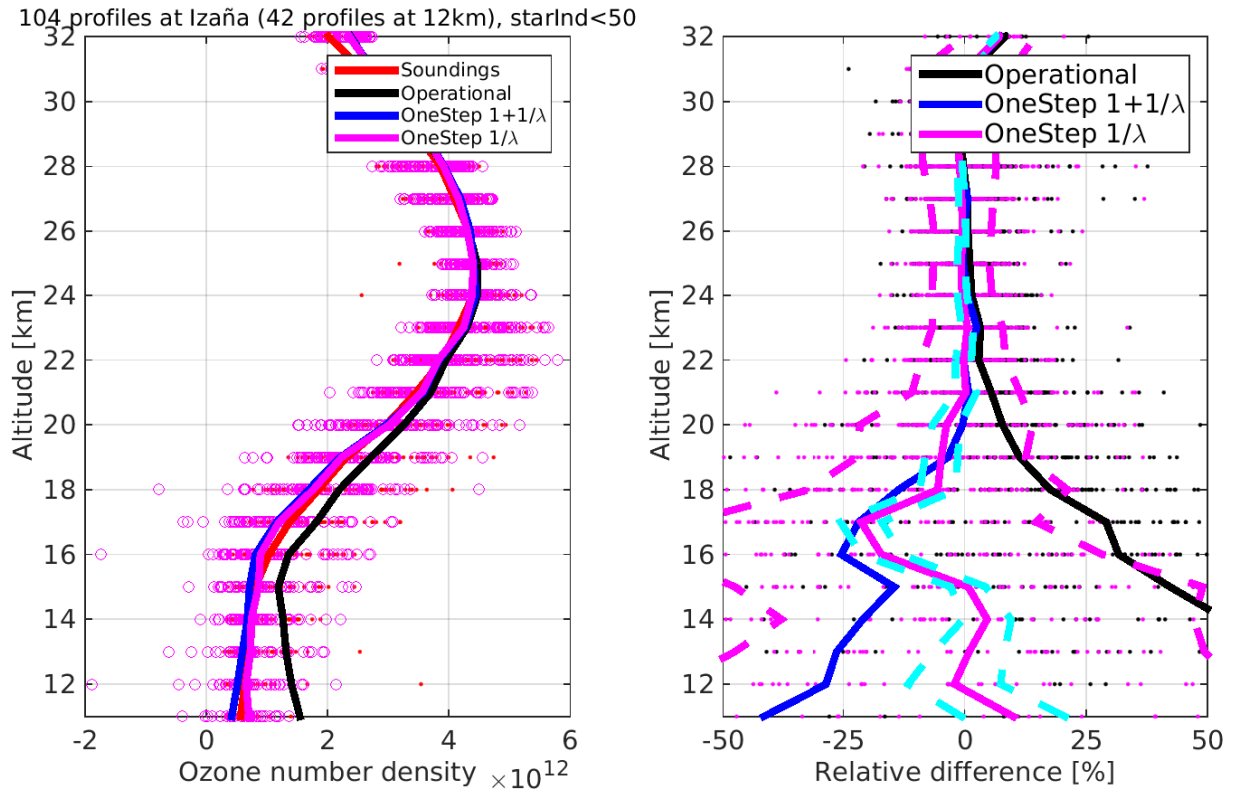


Figure 33: Izaña. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1+1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

[13] Filip Vanhellemont, Didier Fussen, and Christine Bingen. Global one-step inversion of satellite occultation measurements: A practical method. *J. Geophys. Res.*, 109(D9), 2004.

A Supplementary material

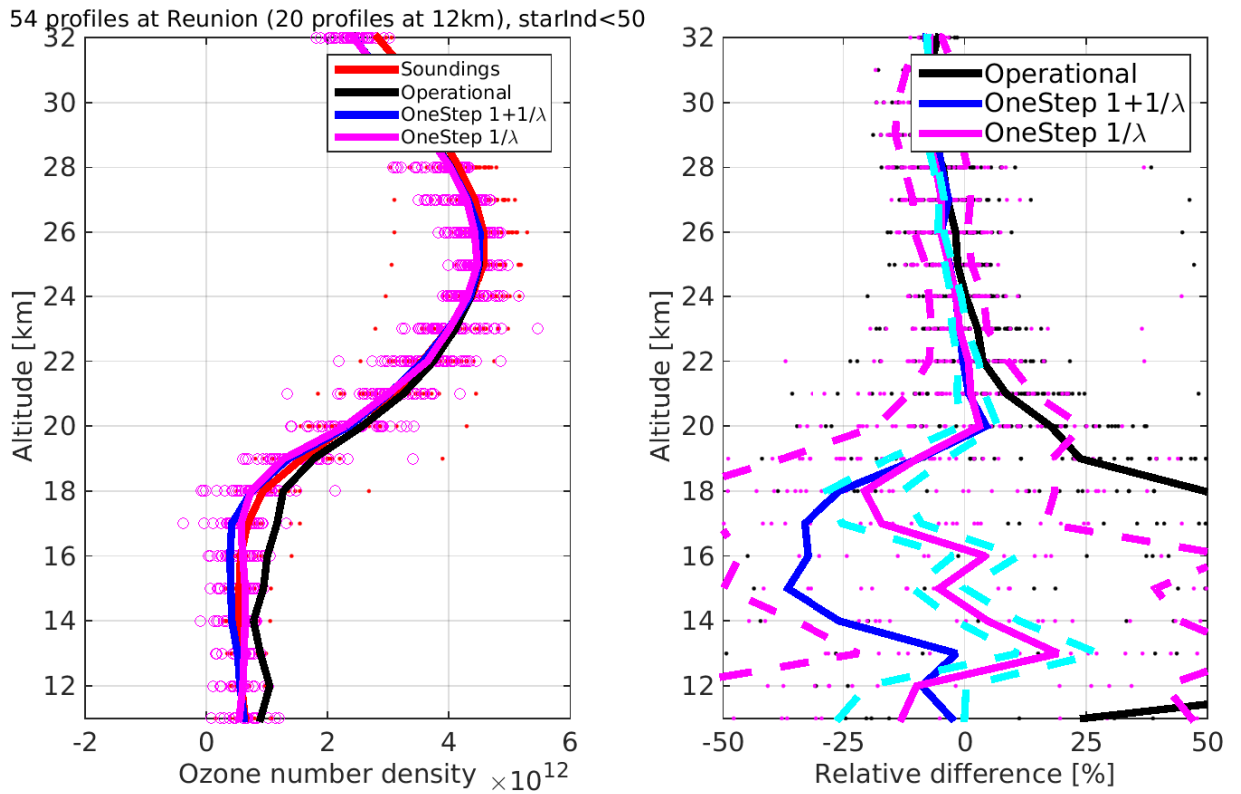


Figure 34: Reunion. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

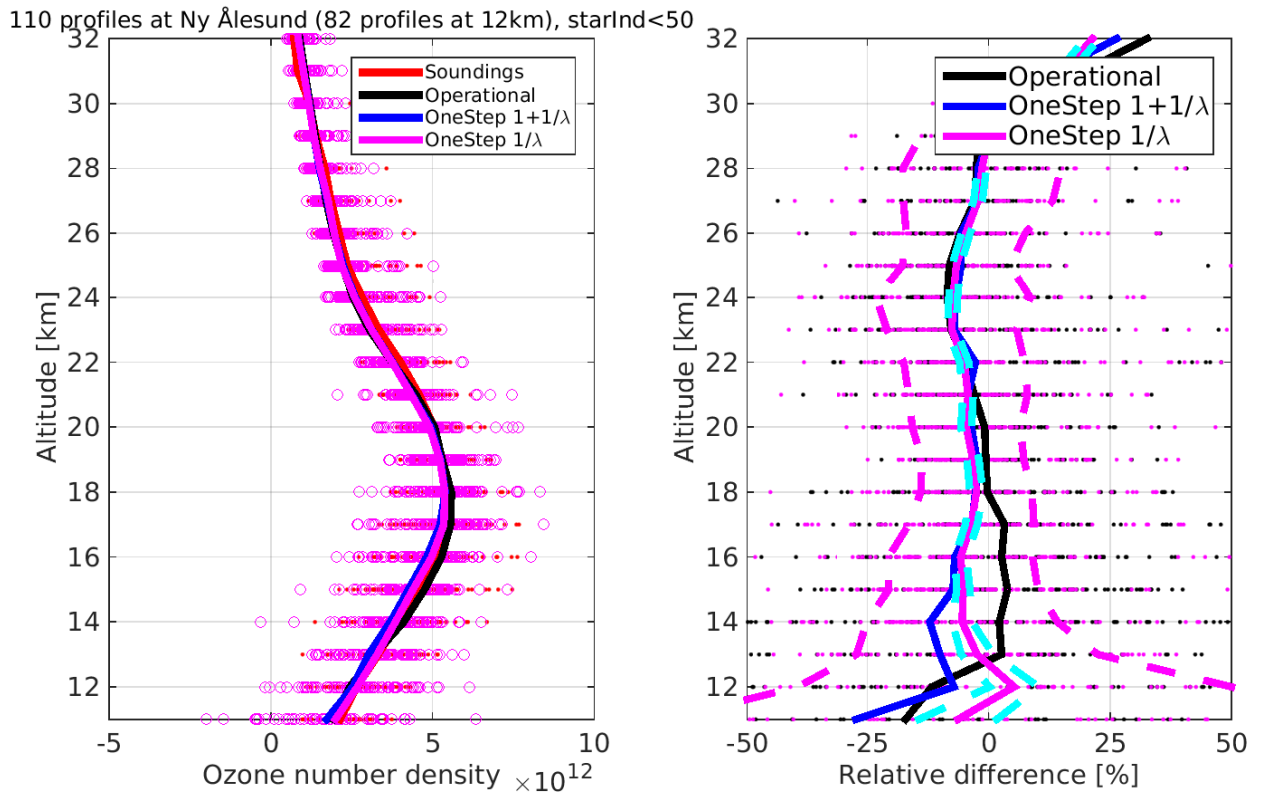


Figure 35: Ny Ålesund. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

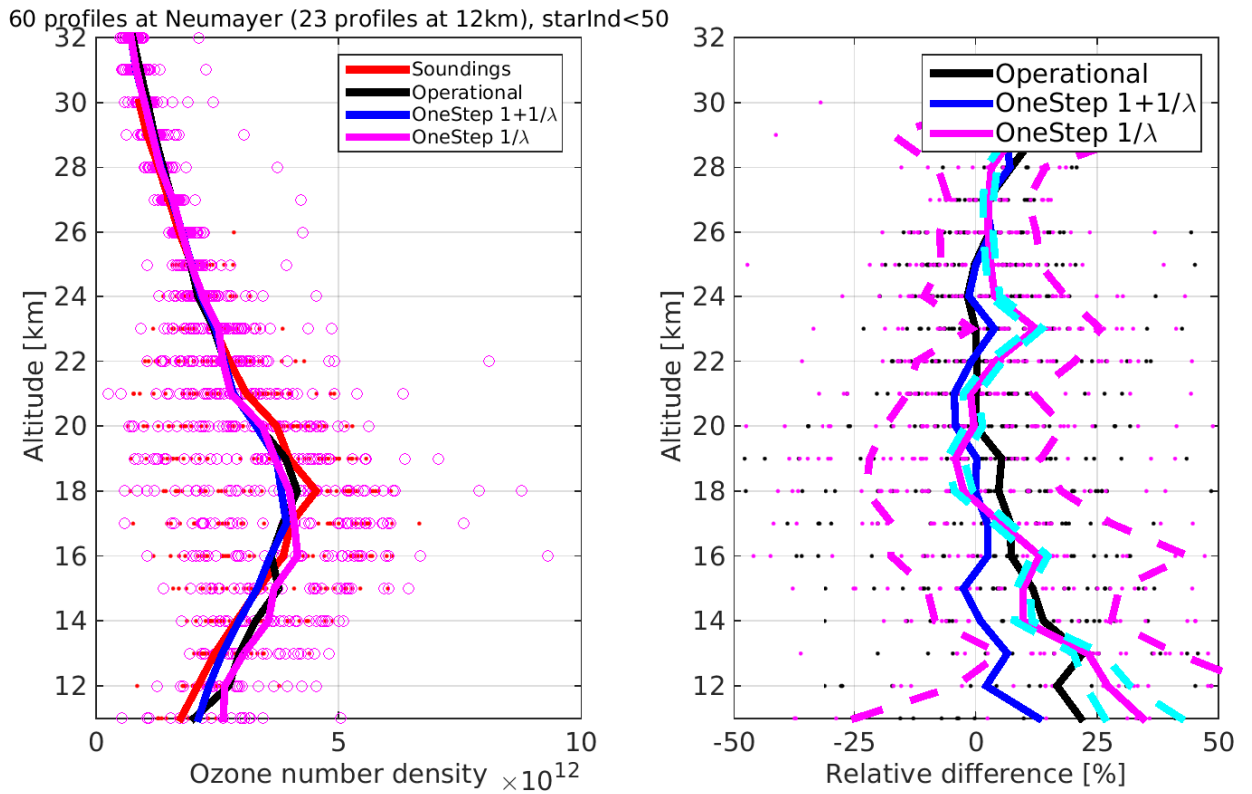


Figure 36: Neumayer. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

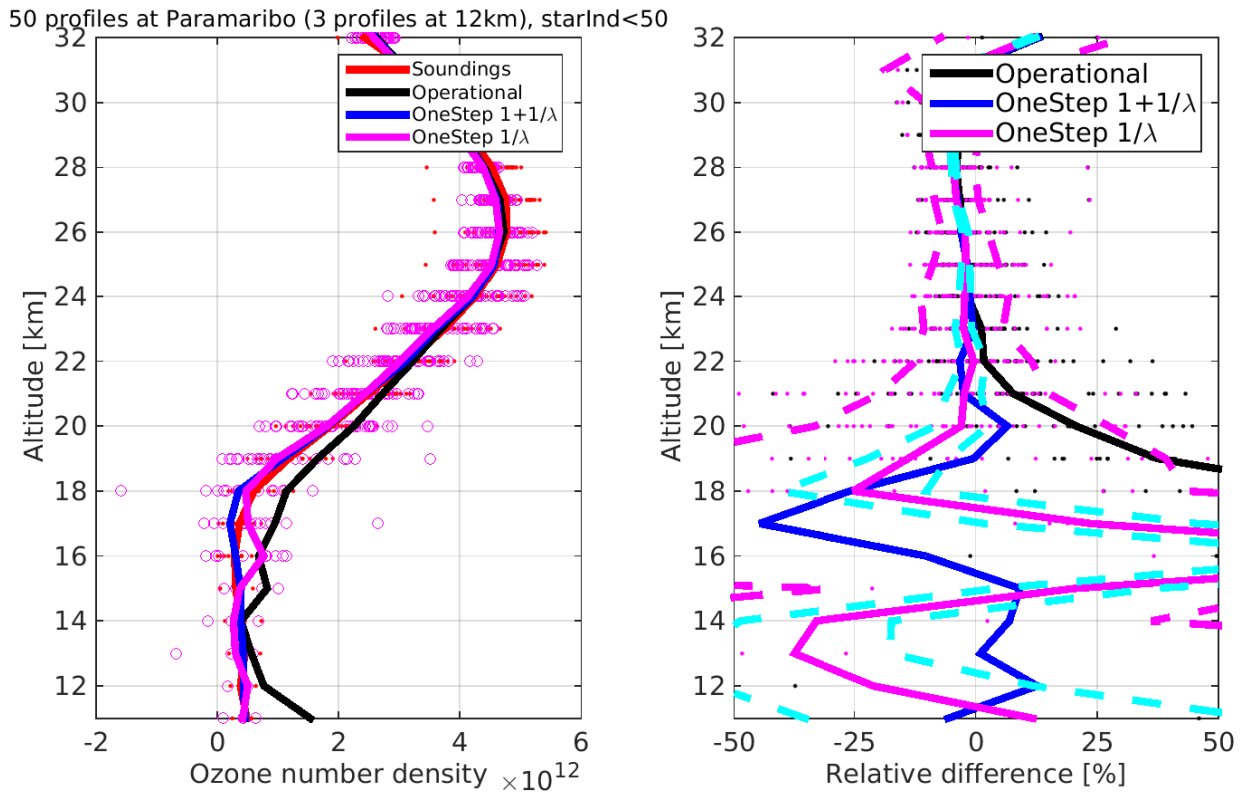


Figure 37: Paramaribo. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

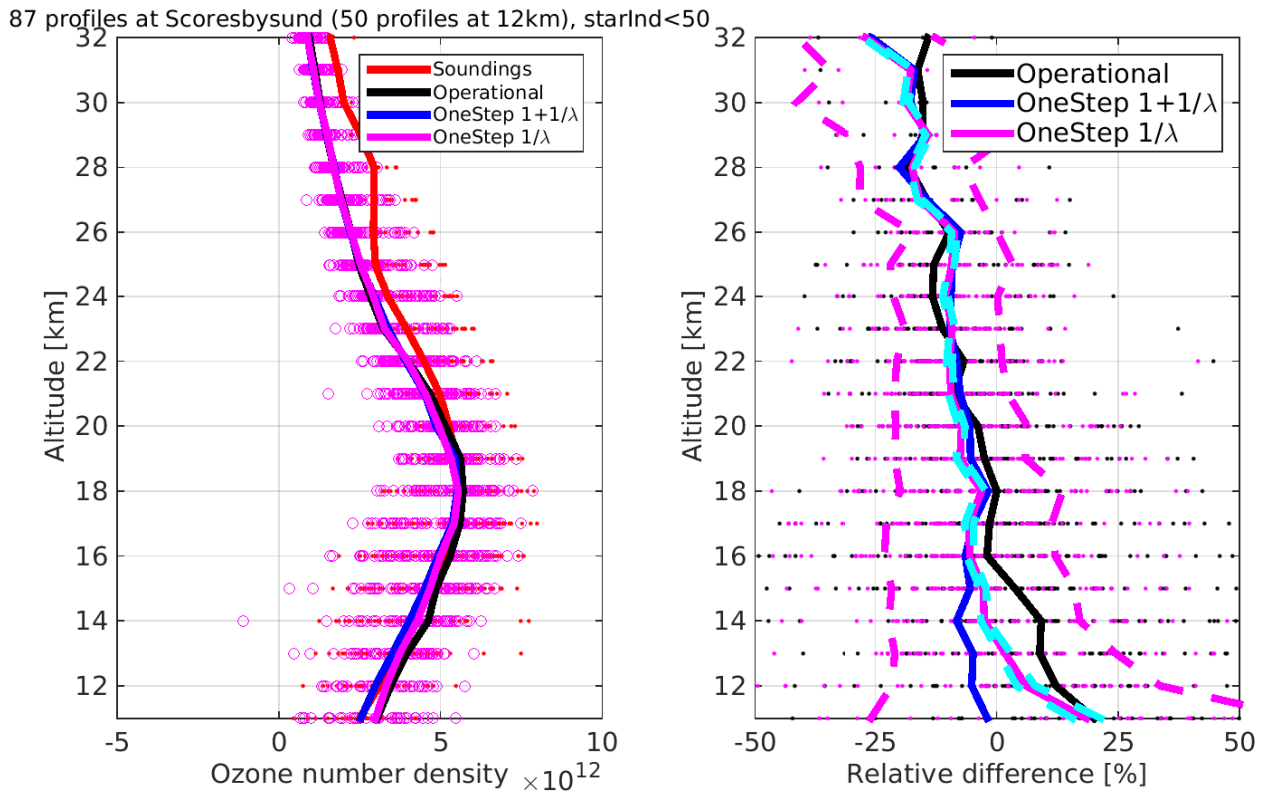


Figure 38: Scoresbysund. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

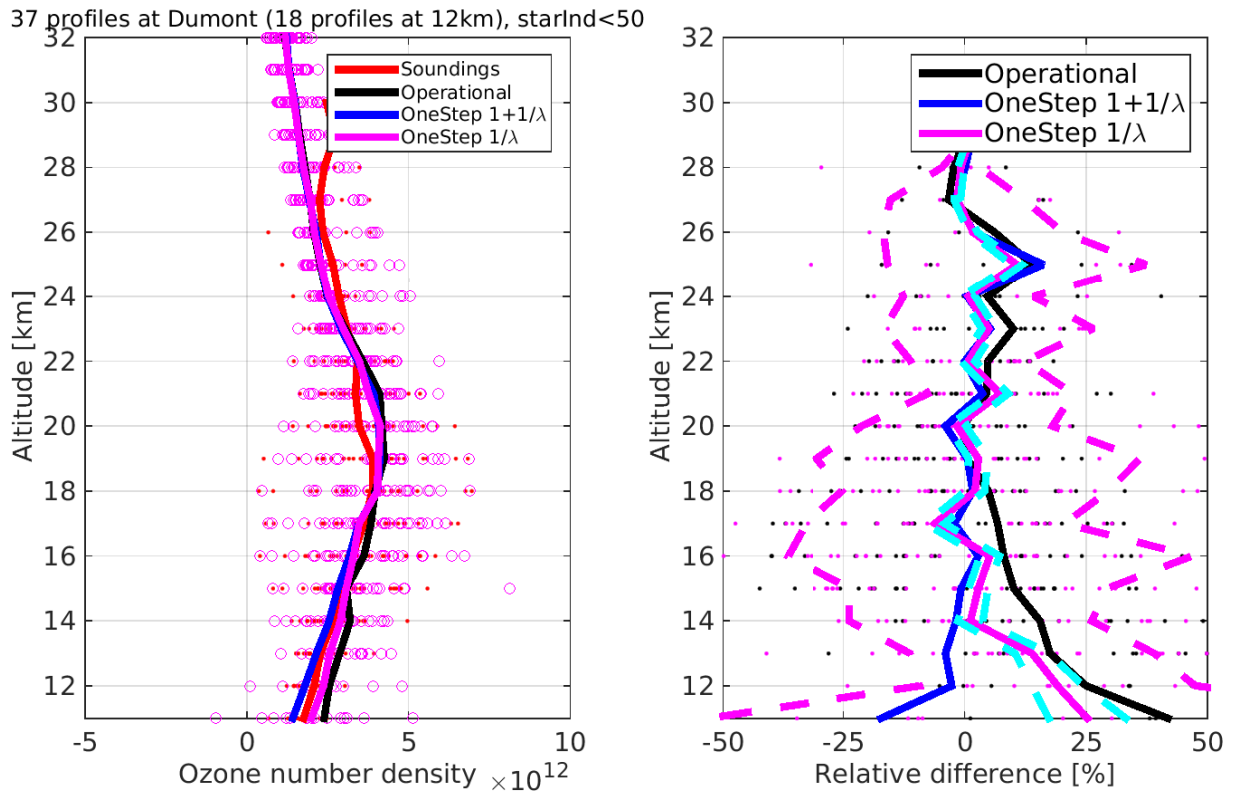


Figure 39: Dumont. Right: ozone profiles. One-step processed with the latest version using $1/\lambda$ (magenta) and $1 + 1/\lambda$ (blue) aerosol models. Left: Relative differences. Magenta dashed lines show the median \pm the standard deviation of the relative differences (for “ $1/\lambda$ ”). Cyan dashed lines show the median \pm median GOMOS error estimates, reported by the one-step algorithm.

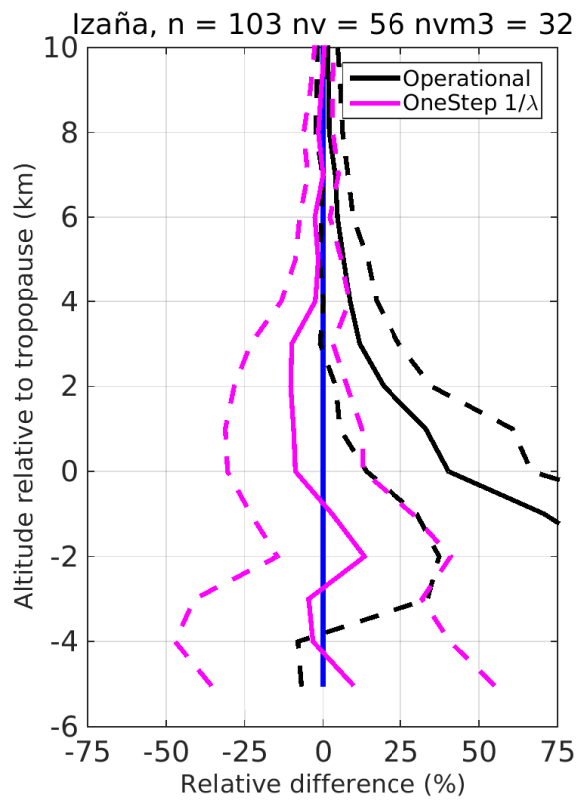


Figure 40: Izaña. Relative difference with respect to soundings. Altitude is relative to tropopause.

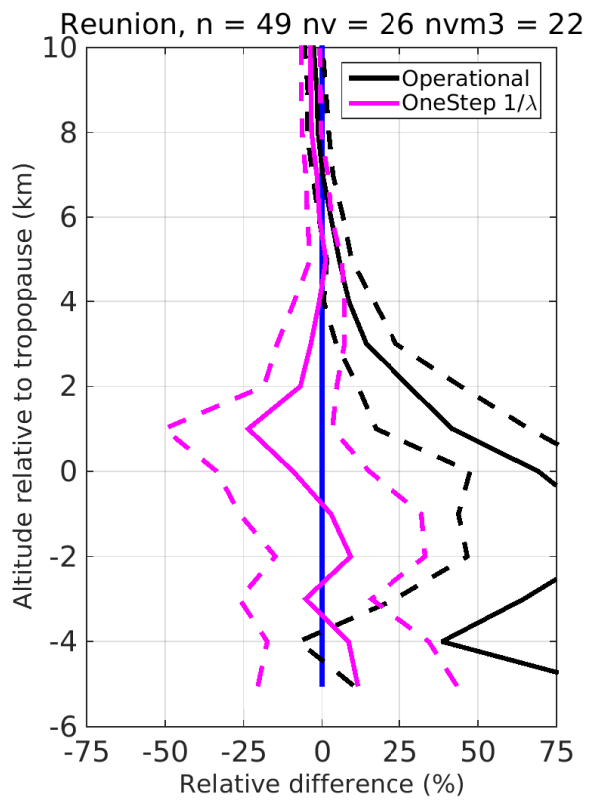


Figure 41: Reunion. Relative difference with respect to soundings. Altitude is relative to tropopause.

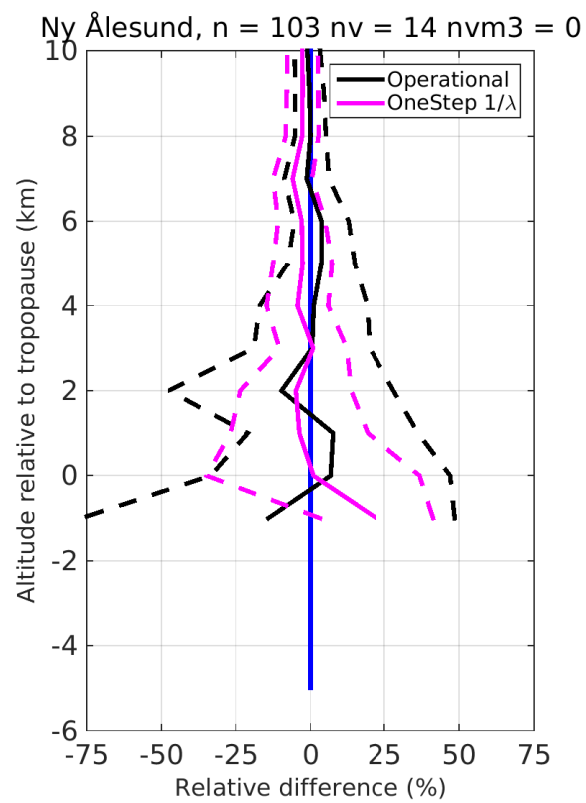


Figure 42: Ny Ålesund. Relative difference with respect to soundings. Altitude is relative to tropopause.

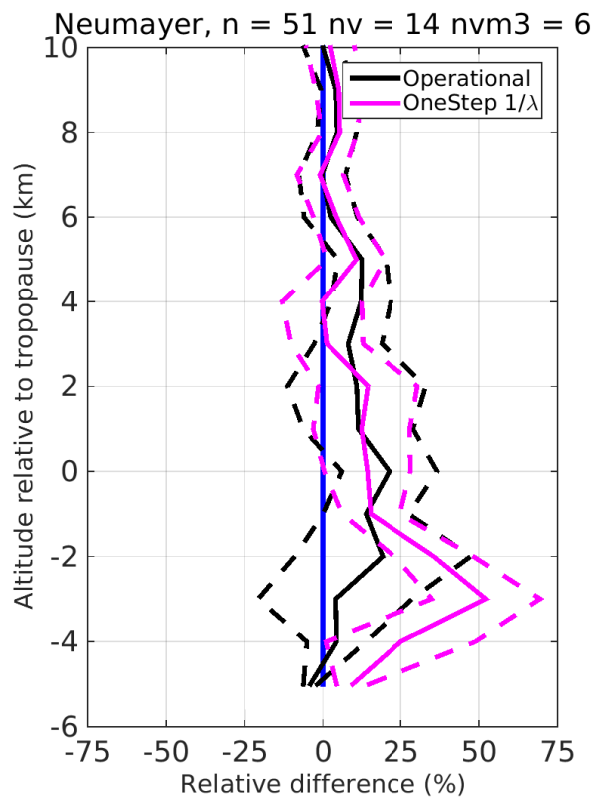


Figure 43: Neumayer. Relative difference with respect to soundings. Altitude is relative to tropopause.

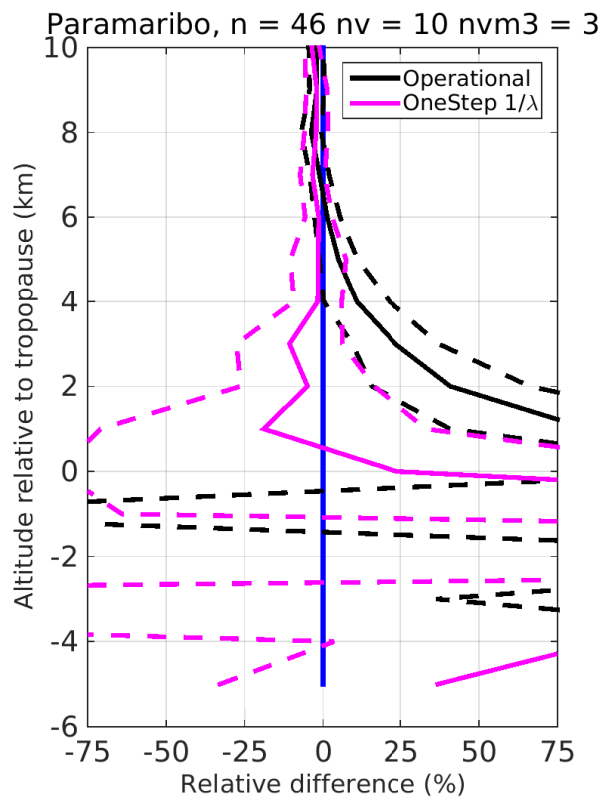


Figure 44: Paramaribo. Relative difference with respect to soundings. Altitude is relative to tropopause.

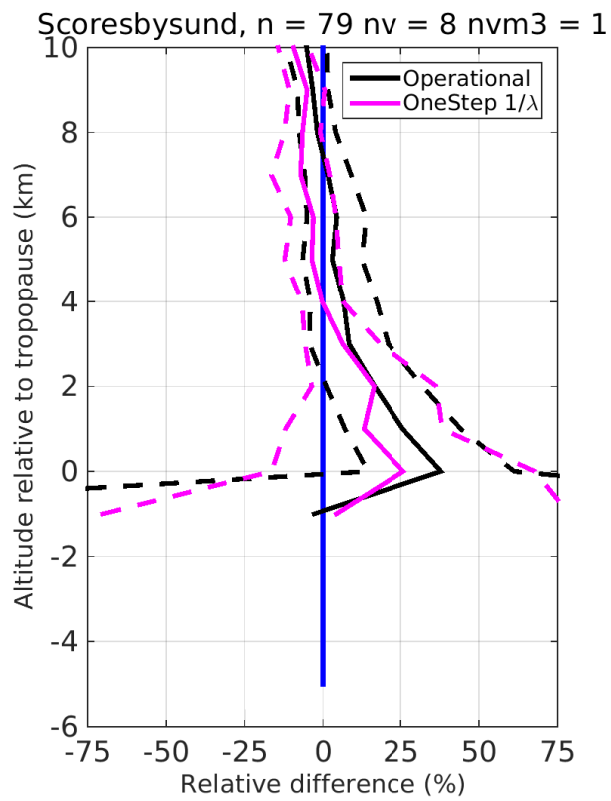


Figure 45: Scoresbysund. Relative difference with respect to soundings. Altitude is relative to tropopause.

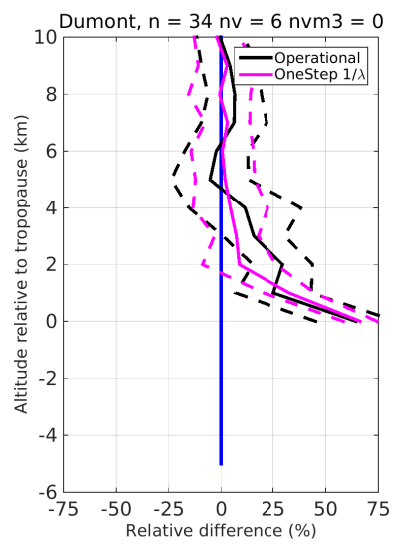


Figure 46: Dumont. Relative difference with respect to soundings. Altitude is relative to tropopause.