Assessment of the error budget for stratospheric ozone profiles retrieved from OMPS limb-scatter measurements

Carlo Arosio[1], Alexei Rozanov[1], Victor Gorshelev[1], Alexandra Laeng[2], and John P. Burrows[1]

[1]University of Bremen, Institute of Environmental Physics, Bremen, Germany
[2]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Karlsruhe, Germany

Correspondence: Carlo Arosio, carloarosio@iup.physik.uni-bremen.de

Abstract. This study presents an error budget assessment for the ozone profiles retrieved at the University of Bremen from limb observations of the Ozone Mapper and Profiler Suite - Limb Profiler Suomi National Polar-orbiting Partnership (OMPS-LP ~~NPP~~SNPP) satellite instrument. The ~~presented error characterization intends to be complaint~~ error characteristics are presented in a form that aims at being compliant with the recommendations and the standardizing effort of the Towards Unified Error Reporting (TUNER) project. Besides the retrieval noise, contributions from retrieval parameters are extensively discussed and quantified by using synthetic retrievals performed with the SCIATRAN radiative transfer model. For this investigation, a representative set of OMPS-LP measurements is selected to provide a reliable estimation of the uncertainties as a function of latitude and season. Errors originating from model approximations and spectroscopic data are also taken into account and found to be non-negligible. The choice of the ozone cross section is found to be relevant, as expected. Overall, we ~~put effort to classify different error contributions~~ classify the estimated errors as random or systematic, and investigate correlations between errors from different sources. After summing up the relevant error components, we present an estimate of the total random uncertainty on the retrieved ozone profiles, which is found to be in the ~~5-30~~5–30 % range in the lower stratosphere, 3–5 % in the middle stratosphere and 5–7 % at upper altitudes. The systematic uncertainty is mainly ~~related~~ due to cloud contamination and model errors in the lower stratosphere, and due to the retrieval bias at higher altitudes~~, with total absolute values of about~~ . The corresponding total bias exceeds 5 % only above 50 km and below 20 km. After computing the estimate of the overall random and systematic error components, we also provide an ex-post assessment of the uncertainties using self-collocated OMPS-LP observations and collocated Microwave Limb Sounder (MLS) data in a χ^2 fashion.

1 Introduction

The objective of this paper is to provide an extensive uncertainty estimate for ozone profiles retrieved at the University of Bremen from the Ozone Mapper and Profiler Suite - Limb Profiler (OMPS-LP) onboard the Suomi National Polar-orbiting Partnership (~~NPP~~SNPP) spacecraft, within the framework of the Towards Unified Error Reporting (TUNER) project, which aims at homogenizing the error reporting in the atmospheric remote sensing community (von Clarmann et al., 2020). An extensive error characterization of the retrieved ozone profiles is relevant for a correct usage of the ozone profiles, e.g. to avoid misinterpretation of retrieval results in atmospheric regions where the product is affected by a large uncertainty. Such estimation

25 is also valuable for the validation of the product: in order to understand and assess the statistical significance of the differences between data sets it is namely important to have an estimation of their random and systematic errors. ~~Such estimation is also valuable for a correct usage of the ozone profiles, e.g. to avoid misinterpretation of retrieval results in atmospheric regions where the product is affected by a large uncertainty.~~

A study assessing the error budget for OMPS-LP ozone retrieval was conducted by NASA at the beginning of the satellite
30 mission and reported by Rault and Loughman (2012). The authors described the retrieval algorithm, which ~~was in the meantime~~ has since been substantially changed (Kramarova et al., 2018), and presented the results of Monte Carlo simulations to estimate both bias and random components of the uncertainty on their profiles. The authors generated randomly perturbed forward simulations, by accounting for the errors in a series of parameters, which were then fed into their standard retrieval routine. They estimated a bias within 4% and a 4% precision from 15 to 50 km, with aerosol, surface albedo and height registration
35 affecting the retrieval up to ~~10-15~~10–15% in the lower stratosphere (Rault and Loughman, 2012). A similar investigation of the error budget for the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) ozone profiles V2.5 retrieved at the University of Bremen was presented by Rahpoe et al. (2013). The focus of that study was on parameter errors, and it reported an estimated random error of about ~~10-15~~10–15% and a systematic component of about 7%. Recently, several satellite groups have been working on improving the uncertainty characterization of their own retrieval products within
40 the TUNER initiative (e.g. ~~Sheese et al. (2020)~~(Sheese et al., 2020))

Regarding the used terminology, we are aware of the effort of the Joint Committee for Guides in Metrology (JCGM) and the Bureau International des Poids et Mesures (BIPM) to replace the concept of error analysis with the concept of uncertainty analysis (Fox, 2010). Referring to von Clarmann et al. (2020), we define the term "error" as the difference between a reference (truth) and the value from measurements or simulations, whereas "uncertainty" describes the distribution of the
45 error. However, in this work, as in von Clarmann et al. (2020), the two terms are not mutually exclusive, especially when used in combined terms, e.g. parameter error or error budget. We also define "ex-ante" and "ex-post" estimates according to von Clarmann et al. (2020). The terms random and systematic are used here to characterize error sources which mainly contribute to the variance and to the bias respectively. This means that, when considering the average of several profiles, random contributions are expected to be reduced (proportionally to the square root of their number), whereas the systematic terms
50 remain constant. Errors on Level 1 (L1) data are not explicitly taken into consideration in this study, but rather assumed to be included in the Signal-to-Noise Ratio (SNR) of the instrument, because it is currently not possible to attribute L1 errors to specific causes and recent efforts from the NASA team largely mitigated the contributions from stray light and pointing ~~(Kramarova et al., 2018, and private conversations). In~~ (Kramarova et al., 2018, and private communications). Another source of error which is not considered in this study but relevant for limb observations, is related to the horizontal smoothing, tackled
55 for example in (Von Clarmann et al., 2009; Cortesi et al., 2007). Namely, limb observations smooth the atmospheric variability over a 250–400 km region around the tangent point (TP) along the line of sight (Zawada et al., 2018). In addition, a 1D retrieval cannot account for gradients along the line of sight, leading to an additional uncertainty component, which is not expected to be relevant on average, but might play a role for atmospheric scenes characterized by strong gradients along the line of sight, e.g. in the presence of the ozone hole (Zawada et al., 2018) or sharp reflectivity gradients.

60  In this study, it was not possible to carry out an error budget for the entire data set, for obvious computing time limitations. As a consequence, we focused on case studies accounting for variations in the geometry and observing conditions, whereas for the parameter and spectroscopic uncertainties we provide a case sample that is considered to be representative of the entire data set.

 With this study we intend to discuss and characterize in more detail the various error contributions, following the TUNER

65 recommendations, and to complete the analysis with an ex-post validation of the obtained error budget. This paper is structured as follows: after the presentation of the OMPS-LP instrument ~~and~~ in Sect. 2.1 and of the retrieval algorithm with its characterization ~~given~~ in Sect. 2.2, an assessment of the systematic retrieval error is described in Sect. 3. Parameter errors are extensively discussed in Sect. 4 as a function of altitude, latitude and season. Uncertainties related to spectroscopic errors are discussed in Sect. 5, with a focus on the used ozone cross section, whereas uncertainties related to model approximations are presented in

70 Sect. 6. Finally, an overall uncertainty estimate is reported in Sect. 7, followed by an ex-post assessment of the reported errors using self-collocated OMPS-LP observations and a validation of the error budget using collocated MLS data in a χ^2 fashion.

2  Description of the instrument and of the retrieval scheme

2.1  OMPS-LP instrument

OMPS was launched by NASA at the end of 2011 on board the SUOMI-NPP satellite platform and started to collect data in

75 January 2012 (Flynn et al., 2014). The spacecraft flies in a sun-synchronous orbit with a nominal 13:30 local time ascending node. The main objective of the mission is the observation of the ozone vertical distribution within the Earth's middle atmosphere at high accuracy. The OMPS-LP instrument, of interest for this study, images the atmosphere by observing the Earth's limb backwards with respect to its flying direction (Jaross et al., 2014). It collects 180 images (i.e. atmospheric spectra) per orbit (around 160 with solar zenith angle less then 80°) and completes ~~14-15~~ 14–15 orbits per day. The instrument measures

80 limb scattered radiance in the spectral region between 280 and 1000 nm. It employs a prism spectrometer, resulting in a spectral resolution that ~~increases~~ degrades from 1 nm in the ultraviolet (UV) ~~up~~ to 40 nm in the near infrared. It observes the altitude range from ground to ∼100 km with an instantaneous field of view of about 1.5 km and a sampling of 1 km (Jaross et al., 2014).

2.2  Retrieval scheme

85 The ~~retrieval of ozone profiles~~ official ozone product for OMPS-LP observations is released by NASA and described elsewhere, see e.g. Kramarova et al. (2018). This study assesses the error budget of the ozone profiles retrieved from OMPS-LP observations ~~is performed~~ at the University of Bremen by using a non-liner inversion technique with 1-st order Tikhonov regularization. This study is based on the retrieval of ozone profiles from OMPS-LP observations developed at the University of Bremen ~~by employing~~, by implementing an iterative approach and the regularized inversion technique with first order Tikhonov con-

90 straints (Rodgers, 2000). The forward modeling accounts for the atmospheric multiple scattering in 1-D geometry, within the

framework of the approximate spherical solver of the SCIATRAN radiative transfer model (RTM) (Rozanov et al., 2014). The measurement vector $\boldsymbol{y}$ contains the logarithms of the normalized limb radiance at all selected tangent heights (TH) and spectral points; e.g., the elements referring to the TH j are defined as:

$$\boldsymbol{y_j} = \log\left(\frac{\boldsymbol{I_{TH_j}}}{\boldsymbol{I_{TH_{norm}}}}\right) - \boldsymbol{P_n} \tag{1}$$

95   where a normalization to a limb measurement at an upper TH is performed. For a detailed description of the retrieval algorithm we refer to Arosio et al. (2018). In this study we used very similar settings ~~for what concerns the chosen~~ to choose spectral segments and the altitude ranges: Table 1 lists the details of the used wavelengths and normalization altitudes. The order of the subtracted polynomial ($\boldsymbol{P_n}$ in Eq. 1) in each spectral segment is also reported in the last column of Table 1. The a priori profile is set to zero: this means that at each iteration SCIATRAN constrains the difference between zero and the ozone 100   profile itself.

Table 1. List of the spectral segments considered in the ozone retrieval with corresponding TH range, altitude used for the normalization and order of the subtracted polynomial ("-" means that no polynomial is subtracted).

TH range [km]	Spectral segment [nm]	TH norm. [km]	Poly. order
~~48-60~~ 48–60	~~290-302~~ 290–302	61.5	-
~~34-49~~ 34–49	~~305-313~~ 305–313	51.5	-
~~28-39~~ 28–39	~~321-330~~ 321–330	51.5	0
~~10-31~~ 10–31	~~508-670~~ 508–670	42.5	1

Ozone, NO_2 and O_4 have relevant spectral signatures in the selected spectral ranges. Their respective cross-sections are taken from Serdyuchenko et al. (2014) and Bogumil et al. (2000), and are ~~beforehand convolved~~ pre-convolved to the OMPS-LP spectral resolution. Ancillary information regarding background temperature and pressure for each OMPS-LP observation is provided in L1 data and comes from the Goddard Earth Observing System (GEOS)-5 model.

105   The OMPS-LP SNR is estimated from the residuals between the measured and the modeled spectrum. This is done at a pre-processing step, where, in addition, a "shift and squeeze" spectral correction is applied in the Chappuis band, located in the visible spectral range. This correction is applied independently for each TH between 10 and 31 km and consists, as the name ~~tells~~ implies, in the calculation of a shift and a squeeze ~~parameters~~ parameter, which define the wavelength correction $\Delta\lambda$ for each wavelength grid point of the modeled spectrum relative to the measured one. This correction is meant to account for 110   potential issues with the spectral calibration and a possible thermal expansion of the sensor. Due to the relatively low spectral resolution of OMPS-LP, the ~~differential~~ fine-scale absorption structure in the UV is mostly not resolved. As a consequence, the UV retrieval uses normalized radiances or their slopes rather than the differential structure, and the influence of spectral misalignments is expected to be negligible, so that the shift and squeeze algorithm is not applied.

   Simultaneously with the ozone retrieval, the surface albedo estimation is performed, by exploiting the sun normalized ra- 115   diance: the two spectral intervals 355 – 365 nm and 670 – 680 nm are used, where the ozone absorption has a minimum.

Beforehand, a cloud filter is also applied and the retrieval of aerosol extinction profiles is performed (see Arosio et al. (2018) for more details).

2.3 Retrieval characterization and retrieval noise

To characterize the retrieval algorithm, the averaging kernels (AK) and the retrieval noise covariance matrix $\mathbf{S_{x,noise}}$ are

120   analyzed. AK provide the information content of the measurements and the sensitivity of the retrieval as a function of altitude. Following (Rodgers, 2000), they are defined for our case as follows:

$$\mathbf{S_{x,noise}} = (\mathbf{K^T}\mathbf{S_{y,noise}}^{-1}\mathbf{K} + \mathbf{S_{x,reg}})^{-1}\mathbf{K^T}\mathbf{S_{y,noise}}^{-1}\mathbf{K}(\mathbf{K^T}\mathbf{S_{y,noise}}^{-1}\mathbf{K} + \mathbf{S_{x,reg}})^{-1} \tag{2}$$

$$\mathbf{A} = (\mathbf{K^T}\mathbf{S_{y,noise}}^{-1}\mathbf{K} + \mathbf{S_{x,reg}})^{-1}\mathbf{K^T}\mathbf{S_{y,noise}}^{-1}\mathbf{K} \tag{3}$$

where $\mathbf{K}$ is the weighting function matrix, $\mathbf{S_{y,noise}}$ is the measurement noise covariance matrix and $\mathbf{S_{x,reg}}$ is the used regu-

125   larization matrix. Typical ~~percentage~~ values of the retrieval noise $\sigma_{x,ret\ noise}$, i.e. the square root of the diagonal elements of $\mathbf{S_{x,noise}}$, averaged over one month, are reported, expressed as a percentage of the true profile, as a function of altitude and latitude in Fig.1.
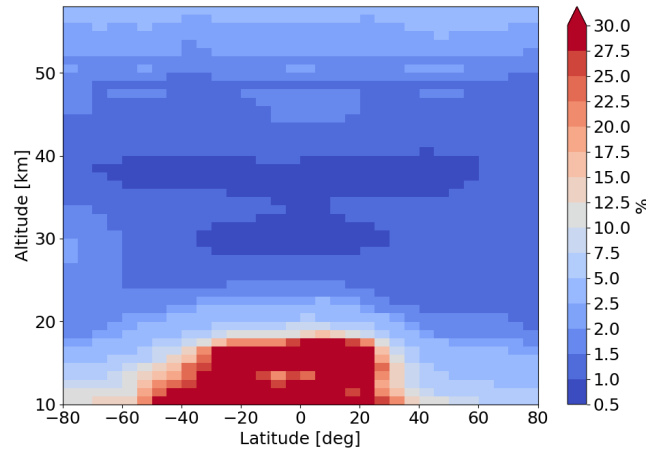
Figure 1. Monthly and zonally averaged retrieval noise for March 2016 as a function of altitude and latitude.

The retrieval noise associated to the ozone profiles does not show any significant latitudinal dependence above 25 km. It lies in the range of ~~1-4~~ 1–4 % up to 60 km and tends to increase up to ~~10-30~~ 10–30 % at lower altitudes, particularly in the tropical

130   Upper Troposphere-Lower Stratosphere (UTLS), where the ozone concentration drops significantly and the retrieval sensitivity degrades.

The vertical resolution of the retrieved profiles is computed as the inverse of the diagonal elements of the AK, multiplied by the altitude layer width, whereas the measurement response is defined as the area of the AK and gives an estimate of the a priori contribution (Rodgers, 2000) and it is dimensionless. The vertical resolution and the measurement response are reported

135 in Fig. 2, averaged over 1 day. Below 30 km, the ~~actual~~ vertical resolution of the retrieval scheme is typically about 2.5 ~~3~~ 3 km getting worst around 33 km, where the transition of the sensitivity between UV and Chappuis spectral ranges occurs. The best vertical resolution of the profiles is achieved around 45 km, whereas above 50 km it gets coarser, due to the increasing regularization. The measurement response is shown in the panel (b) of Fig. 2 and stays around 1.0 at most of the altitudes, except below 15 km in the tropics.
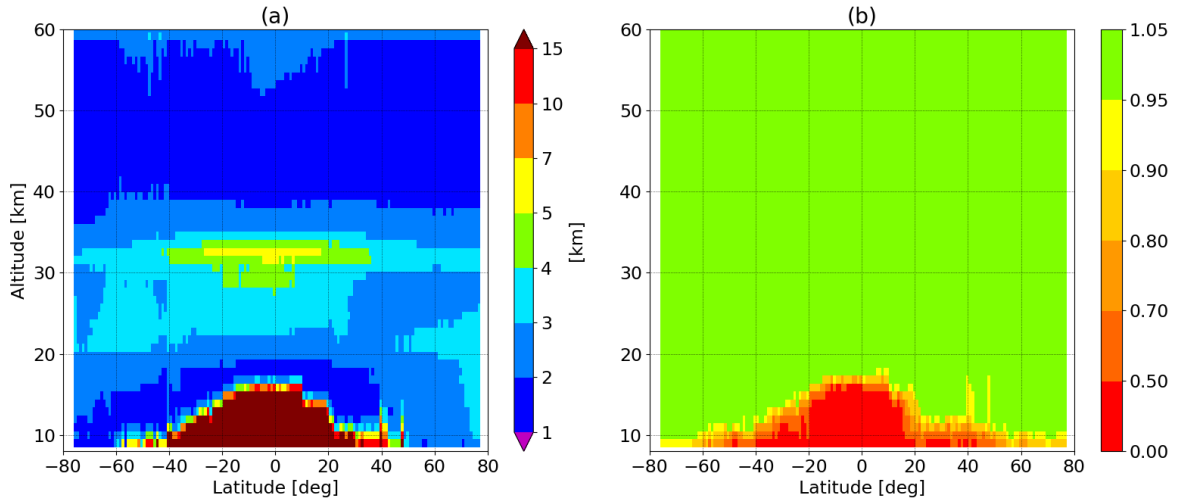
Figure 2. Vertical resolution (panel a) and measurement response (panel b) for the standard retrieval version over one day (15 March 2016)

140 3 Synthetic retrievals using the Monte Carlo approach

Firstly, by using synthetic retrievals with Monte Carlo approach, we evaluate the systematic component of the retrieval error and assess the linearity of the error propagation within the retrieval scheme.

The following procedure is implemented. The SCIATRAN v4.0 forward model is run in OMPS-LP observation geometry obtaining a synthetic OMPS-LP spectrum. This is then fed into the standard retrieval described in Sect. 2.2. Since in the
145 retrieval scheme the SNR of the measurements is estimated from the residual fit, a similar procedure is implemented for the synthetic cases. First, the root mean square (rms) of the fit residuals at every altitude and for each spectral segment is computed for a standard retrieval run of the same state, using OMPS-LP data. This rms is used as the estimated SNR of the measurement. Then, a Gaussian generator is implemented to randomly generate N = 50 noise sequences. Fig. 3 shows in the upper panel an example of residual fits, i.e. the difference between OMPS-LP and the ~~simulated~~ fitted model spectrum, together with the 50
150 generated noise samples, in the Chappuis band. The noise sequences are then multiplied by the SNR value estimated from the real measurements and added to the synthetic spectrum previously simulated. In this way, we obtained ~~N~~ 50 intensity matrices, which are fed into the retrieval scheme to get unperturbed ozone profiles, $O_3^{n,unpert}$. We use the ozone profile retrieved from the
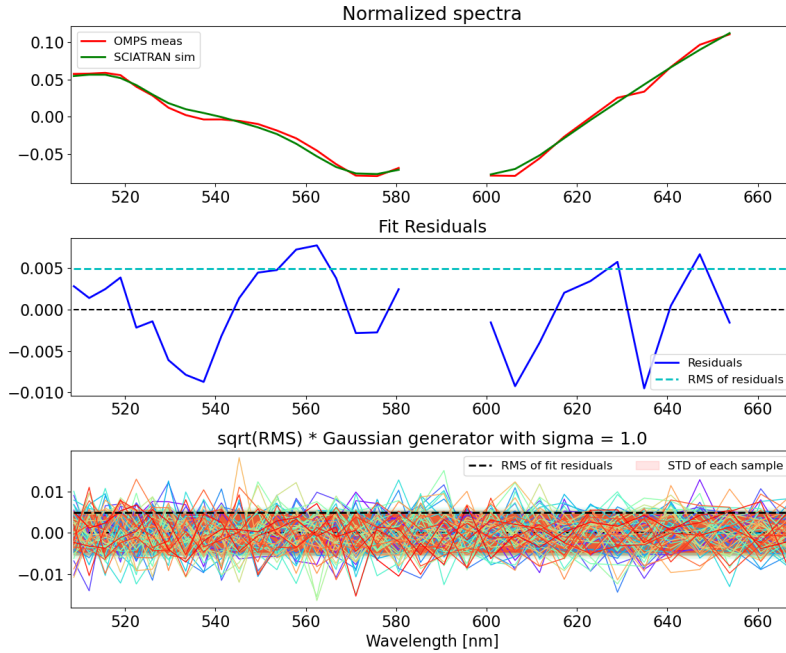
Figure 3. Example of spectra, fit residuals and randomly generated noise sequences (multiplied by $\sqrt{RMS}$) in the Chappuis band. The square root of the fit residuals (rms) is equal to the standard deviation (STD) of each generated noise sample.

real measurements as first guess ~~profile~~ profiles and we refer to it as the ~~"~~"true" ozone, which is then compared with the average of the retrieved unperturbed ozone profiles. Figure 4 shows, in panel (a), all 50 retrieved unperturbed profiles (with different noise sequences applied) and their average $O_3^{avg,unpert}$ in black, for one selected OMPS-LP state. Rigorous definitions of the average retrieved unperturbed ozone profile and the difference to the true profile $\delta_{x,unpert}$, as function of latitude (ϕ) and altitude (z), are given below:

$$O_3^{avg,unpert}(\phi,z) = \frac{1}{N}\sum_{n=1}^{N} O_3^{n,unpert}(\phi,z) \tag{4}$$

$$\delta_{x,unpert}(\phi,z) = \frac{O_3^{true}(\phi,z) - O_3^{avg,unpert}(\phi,z)}{O_3^{true}(\phi,z)} * 100 \tag{5}$$

Panel (b) of Fig. 4 displays the average relative difference $\delta_{x,unpert}$ for several latitudes. The relative discrepancy is significantly different from zero, contrary to what is ideally expected, in particular above 50 km, where the discrepancy increases with altitude up to ~~5-10~~5–10 %. Also in the lower stratosphere significant differences are observed, with peaks of 2.5–5% below 20 km, however their sign changes with the latitude and has an oscillating nature. These ~~features~~ discrepancies are related to the fact that we choose as a priori the null profile and SCIATRAN regularizes at each iteration the difference between zero and the resulting ozone profile. As a consequence, at altitudes where the Tikhonov regularization becomes stronger, the retrieved ozone profile tends to be underestimated. The large relative differences are also related to the very low ozone values
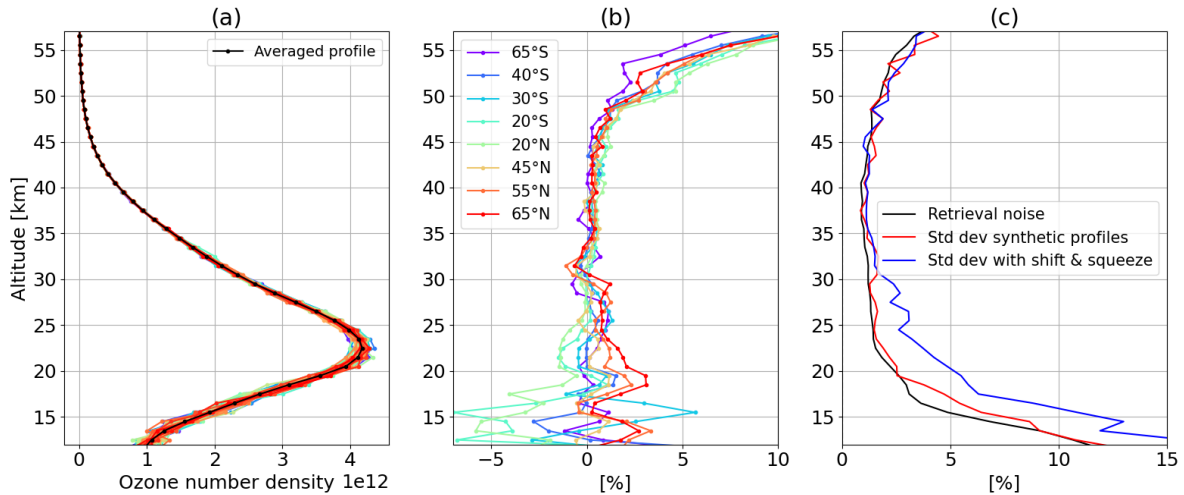
Figure 4. Panel (a): all 50 retrieved unperturbed ozone profiles with their average in black. Panel (b): relative differences $\delta_{x,unpert}$ at several latitudes. Panel (c): retrieval noise and standard deviation of the synthetic retrievals with and without shift and squeeze procedure.

at the upper altitudes. This feature is a component of the systematic error of the retrieved profiles and has to be taken into account when comparing the data set with other reference instruments. In panel (c) of Fig. 4 the retrieval noise and the standard deviation of the unperturbed synthetic retrievals are plotted together in relative values for a direct comparison. As soon as the shift and squeeze procedure is activated, the standard deviation of the synthetic profiles increases drastically. The maximum range of the shift and squeeze correction allowed in the retrieval is ± 1 nm, resulting in an increase of the noise error by about 50 % compared to the linearly propagated retrieval noise. This implies that the error propagation below 30 km has a non-linear component related to the shift and squeeze algorithm. This term has to be taken into account in the total error estimation and will be identified as $\sigma_{x,shift-squeeze}$ $\sigma_{x,\text{shift\&squeeze}}$. Otherwise the measurement noise is fairly linearly propagated into level 2 (L2) profiles.

By using synthetic simulations, we also checked the dependence of the retrieved ozone profile on the first guess profile and it was found to be negligible (not shown here). When the first guess profile was scaled by values between -50 % and +100 %, the retrieval scheme showed strong stability, with the resulting ozone profile being affected by less than 1 %, as expected by panel (b) of Fig. 4.

4 Parameter uncertainty

In this section we discuss the so called parameter errors, i.e. the uncertainty components originating from using certain information about the atmospheric state, which is either known or independently retrieved with a certain error, and which does not enter the state vector. The following parameters are considered: surface albedo, aerosol extinction profile, TH registration, pres-

sure and temperature profiles. For each parameter, a reasonable value of its error is estimated, corresponding to their standard deviation, and used to generate perturbed scenarios and synthetic retrievals. Table 2 indicates the five considered parameters and the values of their respective estimated errors.

Table 2. List of the parameters used for the sensitivity studies and respective values of the variations, i.e. their standard deviation, corresponding to their estimated errors.

	Albedo	Temperature	Pressure	Aerosol	TH
Variation	±20 %	±2 K	±2 %	±30 %	±100 m

The ~~final uncertainties on ozone profiles related to these parameter errors are indicated in the following as $\sigma_{x,albedo}$, $\sigma_{x,T}$, $\sigma_{x,P}$, $\sigma_{x,aerosol\ ext}$, and $\sigma_{x,TH}$ respectively. The~~ magnitudes of these errors are either estimated from the retrieval product, e.g. for aerosol extinction profiles, or are taken from literature. Temperature and pressure profiles are taken from the GEOS-5 global assimilation model and a reasonable estimation of their precision is 2 K and 2 % respectively, see for example Langland et al. (2008). These estimates were also reported, for example, by Nowlan et al. (2007), who studied the differences between MAESTRO and reanalysis from the National Centers for Environmental Prediction (NCEP) and the European Centre for Medium-range Weather Forecasts (ECMWF). The aerosol extinction retrieved at the University of Bremen (for reference Malinina et al. (2018)) shows on average an upper error value of about 30 % in the lower stratosphere. Similarly, an estimate of 20 % precision as an upper boundary value for the surface albedo comes form the analysis of the results of the surface albedo retrieval, performed at the same time with ozone. Finally, an important parameter for ozone limb retrievals is the pointing knowledge of the instrument. According to Moy et al. (2017), the results from NASA investigations about the pointing stability of the instrument, indicate that OMPS-LP pointing has an accuracy of about ±100 m. The final uncertainties on ozone profiles related to the errors in albedo, temperature, pressure, aerosol extinction and tangent height registration are indicated in following as $\sigma_{x,albedo}$, $\sigma_{x,T}$, $\sigma_{x,P}$, $\sigma_{x,aerosol\ ext}$, and $\sigma_{x,TH}$ respectively.

To have a representative sample of cases for the parameter uncertainty estimation, we selected cloud-clear OMPS-LP cases distributed over one year of observations, with the aim to have about $50 - 100$ cases distributed in 4 seasons and 5 latitude bands (southern and northern polar latitudes, southern and northern mid-latitudes, and tropics). All estimated errors are assumed to have a random nature, as we have no information about systematic deviations of the parameters. It has to be noted that for aerosol, temperature and pressure, the error on parameter was assumed to be constant in altitude, even though a random perturbation as a function of height might also be assumed. A comparison of the resulting error can be found in the Supplements.

4.1 Methodology

The Monte Carlo approach described in Sect. 3 has been also used to study the parameter uncertainty of the retrieval. In this case one parameter at a time is varied in the forward model, then the Gaussian noise is added to the simulated radiances in the same way as described above, and N intensity matrices obtained, which are fed into the retrieval scheme to get perturbed ozone

profiles, $O_3^{n,\text{pert\_fix}}$. We define then the averaged perturbed ozone profiles $O_3^{avg,\text{pert\_fix}}$ and the relative differences w.r.t. the averaged unperturbed retrievals ($O_3^{avg,\text{unpert}}$) as follows:

$$O_3^{avg,\text{pert\_fix}}(\text{param},\phi,z) = \frac{1}{N}\sum_{n=1}^{N} O_3^{n,\text{pert\_fix}}(\text{param},\phi,z) \tag{6}$$

$$\sigma_{x,\text{pert\_fix}}(\text{param},\phi,z) = \frac{O_3^{avg,\text{pert\_fix}}(\text{param},\phi,z) - O_3^{avg,\text{unpert}}(\phi,z)}{O_3^{avg,\text{pert\_fix}}(\text{param},\phi,z) + O_3^{avg,\text{unpert}}(\phi,z)} * 20 \tag{7}$$

where n is the running number of the noise sequence. $\sigma_{x,\text{pert\_fix}}$ provides an estimate of the effect of each parameter error on the retrieved ozone profile. In this case we assume a fixed error value on the parameters and evaluate its effect in the ozone space.

Another conceptually similar approach consists in perturbing a single parameter around its mean value by a normally distributed amount, which has a standard deviation equal to the estimated parameter error. For example, the aerosol extinction profile is perturbed by 30 %, so that 50 aerosol extinction profiles normally distributed around the original one are created. The standard retrieval version is then applied 50 times to the synthetic OMPS-LP spectra with perturbed parameters, using the residual fit from the corresponding retrieval of measured data as SNR. The quantity of interest, i.e. the random uncertainty on the retrieved ozone, is the standard deviation of the retrieved ozone profile distribution.

$$\sigma_{x,\text{pert\_rand}}(\text{param},\phi,z) = \frac{std(O_3^{n,\text{pert\_rand}}(\text{param},\phi,z))}{O_3^{avg,\text{unpert}}(\phi,z)} * 100 \tag{8}$$

4.2 Results

Synthetic simulations using the two methods described above are performed to assess the impact of errors of the forward model parameters on the retrieved ozone profiles. In this investigation the shift and squeeze in Chappuis band was turned on, as in the standard retrieval version, with the upper limit for the $\Delta\lambda$ equal to 1 nm. As defined by Eqs. 7 and 8, for the rest of the analysis the differences are computed relative to $O_3^{avg,\text{unpert}}$ rather than to the true ozone profile, in order to have an evaluation of the sensitivity of the retrieved ozone profile to forward model parameters, which is not affected by the sensitivity of the retrieval itself. Below 20 km, it is challenging to estimate the effect of each particular parameter because of a high variability found in the simulations, which is however reduced by averaging over several case studies.

In Fig. 5 an example of the comparison between the results of the two methodologies described above is presented for a single OMPS-LP observation, for pressure and aerosol extinction cases. The plot indicates that both approaches lead to very similar results. In both cases, in red and blue the perturbed profiles are shown in the left panels. The unperturbed one is shown in black, whereas the green shaded area represents the standard deviation of the N simulations from the second approach ($\sigma_{x,\text{pert\_rand}}$). For a better view, in the right panels $\sigma_{x,\text{pert\_fix}}$ and $\sigma_{x,\text{pert\_rand}}$ are plotted.

After verifying the agreement between the two methods, the final results will be reported using the first approach (fixed parameter perturbation), which is computationally less demanding and conveys information on
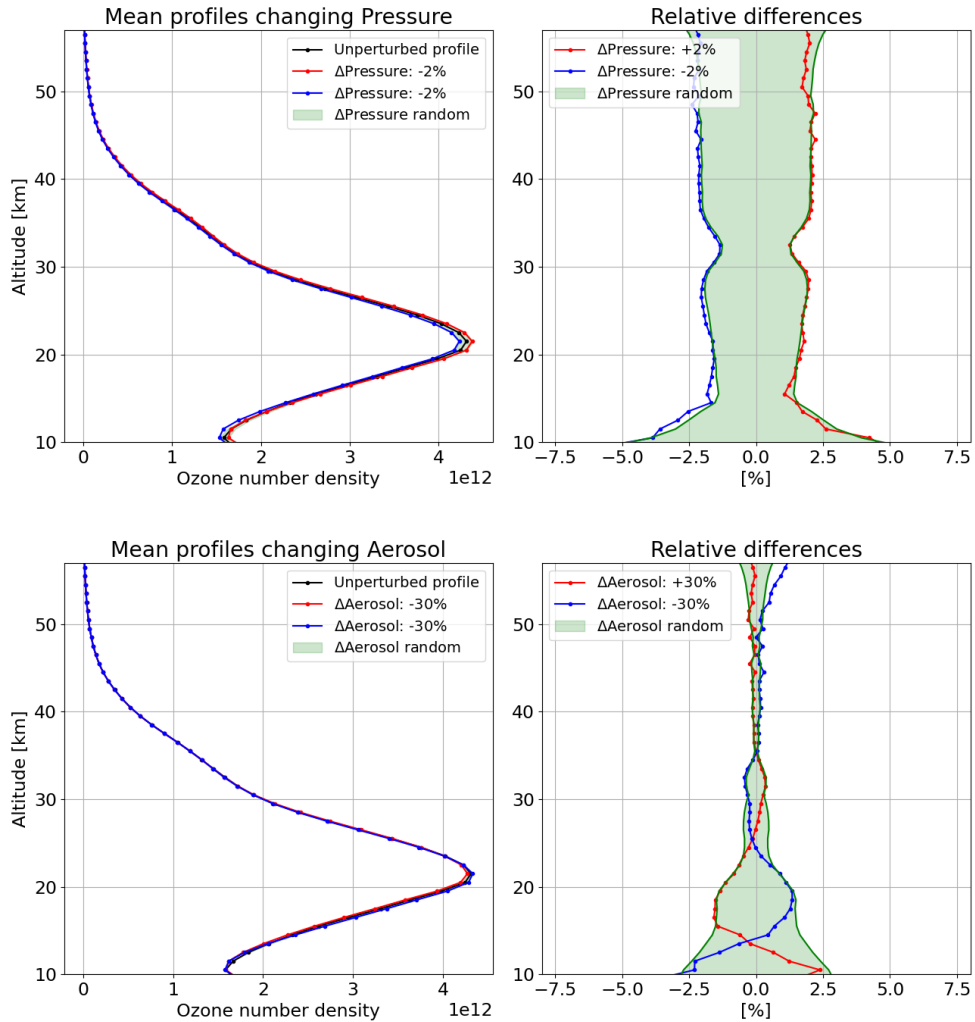
Figure 5. In the left panels the unperturbed ozone profile is shown together with the mean ozone profiles obtained by a Monte Carlo approach by perturbing in the forward simulation pressure (top row) and aerosol extinction (bottom row) by 2% and 30% ~~,~~respectively~~, in a Monte Carlo - like approach~~. The green shaded area shows (barely visible) the standard deviation of the 50 profiles obtained with the second approach. In the right panels, ~~$\sigma_{x,pert\_fix}$~~ $\sigma_{x,\text{pert\_fix}}$ and ~~$\sigma_{x,pert\_rand}$~~ $\sigma_{x,\text{pert\_rand}}$ are shown in relative values for the respective cases.

the relationship between the sign of the ~~ozone uncertainty according to~~ parameter perturbation and the sign of the ~~parameter consequent ozone~~ perturbation. In Fig. 6 the results of the parameter uncertainties are reported for the 5 identified parameters, as a function of latitude (as reported in the legends) and for both increment and decrement of the parameter. To give an indication of the variability in term of season, the shaded areas indicate the standard deviation of the uncertainties in different seasons.

From the upper left panel of Fig. 6 we notice that the effect of perturbing the pressure profile is fairly constant with height, with a deviation of ~~2-3~~2–3 % relative to the unperturbed case at most of the altitudes, without a significant dependence from latitude or season, though a larger scatter is seen at high latitudes. A reduction of the sensitivity of the retrieval at about ~~32-33~~ 32–33 km is the reason for the reduced effect of the perturbation, and it is also visible in other panels.

The effect of increasing the temperature profile by $\Delta = 2\ K$ has a different sign and a smaller amplitude w.r.t. the pressure perturbation, leading to an underestimation of the retrieved ozone, with the largest effects below 30 km and in the upper stratosphere. However, the deviations remain within ~~1-2~~1–2 % at all altitudes above 15 km and no evident variation as a function of latitude is detected, except above 50 km, where an underestimation of the temperature leads to a large overestimation of ozone.

The aerosol perturbation has also an inverse dependency with the ozone deviation and the largest effect is observed below 25 km. A positive variation of the aerosol extinction leads to an underestimation of the ozone concentration and it has the maximum impact at northern high latitudes, due to the scattering geometry of the satellite observations: in the northern hemisphere the forward scattering makes the effect of the aerosols stronger than the backward scattering observed in the southern hemisphere, due to the stronger forward peak in the aerosol phase function. The dispersion of the values as a function of season is also large in the northern polar regions.

As already mentioned, for limb retrievals the accurate knowledge of the instrument pointing is a highly relevant parameter. A perturbation of the TH of ±100 m leads to an uncertainty profile that changes sign at ~~20-25~~20–25 km, with values increasing with altitude up to ~~3-4~~3–4 % above 55 km.

In order to assess the impact of the retrieved surface albedo error on the ozone profile, we perturb the surface albedo value in the forward simulation and retrieve ozone switching off the albedo retrieval. We find that the deviation from the unperturbed case is evident only below 30 km and appears to be symmetric, in response to a positive or negative perturbation of the surface albedo: an increase of this parameter in the forward model leads to an underestimation of the ozone retrieved profile in the lower stratosphere and vice versa. The average unperturbed surface albedo values are reported in the legend for the displayed latitude bands.

We investigated whether parameter errors have an additive (i.e. their absolute values are independent from ozone concentration) or multiplicative (i.e. they scale with the ozone profile, have a constant relative values) nature. This is done by performing the simulations with a first guess ozone profile scaled by ±20 %, and comparing the results with Fig. 6. It was found that all parameter uncertainties are multiplicative: by doubling the amount of ozone, the relative error stays constant. Only for the aerosol extinction case we notice a stronger than linear dependence from the ozone amount, i.e. the relative error in ozone due the assumed background aerosol profile slightly increases with increasing ozone. In addition, since the errors on the parameters are random, we assume that the resulting ozone uncertainties are generally random as well. However, possible systematic
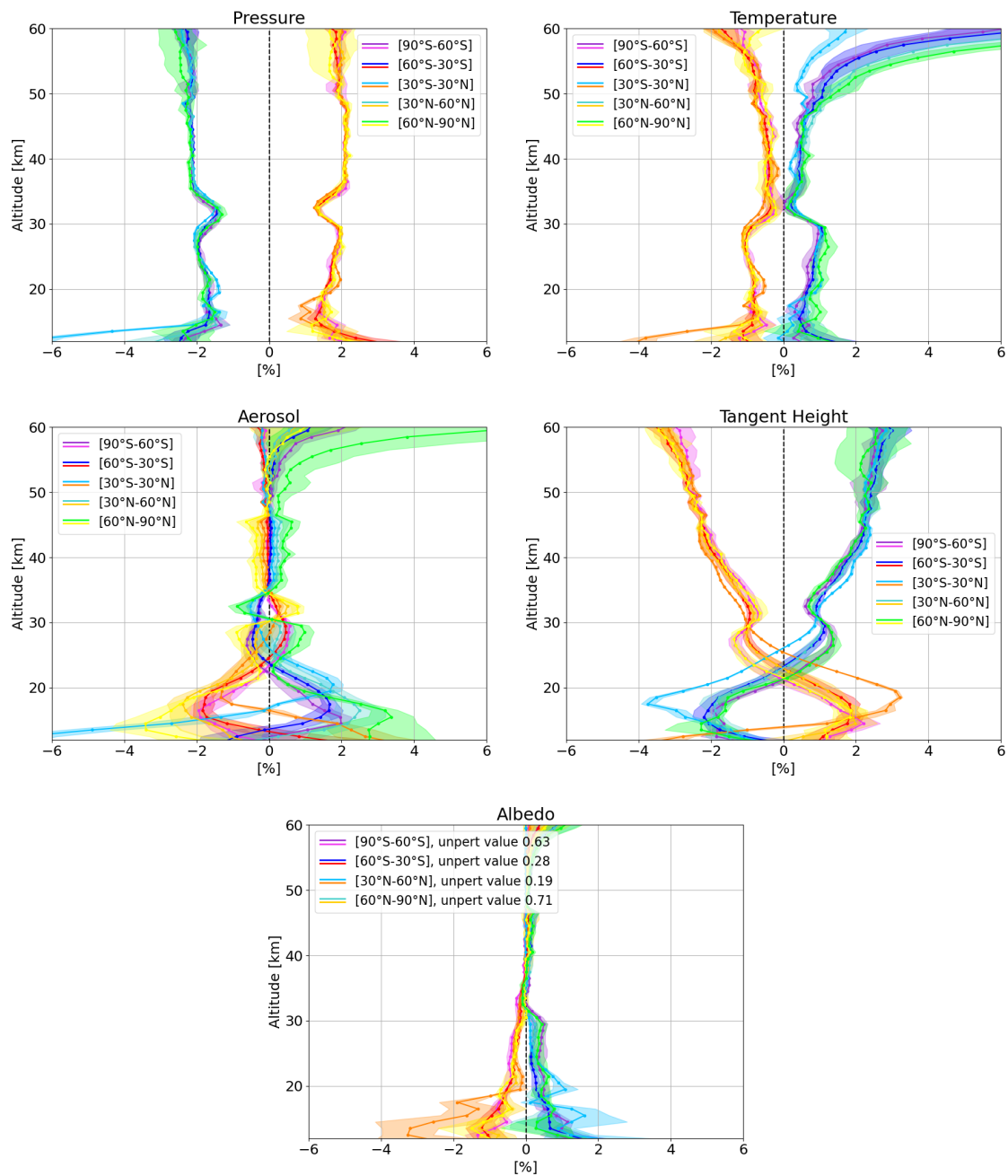
Figure 6. Estimation of parameter errors from synthetic retrievals, in terms of the relative uncertainty on the ozone profile. Each panel refers to the parameter specified in the title, which is perturbed according to Tab. 2. Warm colors indicate a positive perturbation of the parameter, cold colors a negative one. Latitude bands are indicated in the legend and the shaded areas around each profile indicate the scattering of the uncertainty as a function of season.

effects along the time series cannot be excluded, in particular for what concerns the aerosol extinction contribution, which depends on the aerosol loading. As shown, the magnitude of the aerosol error has also a systematic variation along the latitude.

4.3 Clouds

Synthetic simulations have been performed to assess the effect of undetected clouds on the ozone profile: forward simulations
280 are performed by adding a cloud layer in SCIATRAN using three scenarios: a cloud layer at ~~16-18 km, 10-12 km and 4-6~~ 16–18 km, 10–12 km and 4–6 km. Three optical depth (OD) values for each cloud altitude are also chosen: 0.1, 0.5 and 1.0. These are relatively small values, to simulate thin clouds that could remain undetected by the implemented cloud filter. However, by applying the cloud filtering to the simulated intensities, the cases with OD of 0.5 and 1.0 are always flagged as cloudy by our algorithm. Clouds are assumed to consist of hexagonal ice crystals for the first two high-clouds and of water droplets
285 for the lower-cloud case. The scenarios with an ice cloud of OD of 0.5 or 1.0 at 17 km refer to a typical thin cirrus in the tropics; the other 2 refer to a thin ice or low water cloud at mid-/high- latitudes. The cloud is added in the forward simulation and the retrieval is performed with the standard settings, using the SNR from an OMPS-LP retrieval. At the same time, a reference synthetic retrieval is run, i.e. a forward simulation without cloud followed by a standard retrieval in the same way as for cloud-perturbed cases. The difference is then computed w.r.t. this reference unperturbed scenario:

290 $$\delta_{x,\text{cloud}} = O_3^{\text{unpert}} - O_3^{\text{pert}} \tag{9}$$

Fig. 7 displays these relative differences for several ODs and cloud heights (for each panel). The shaded areas show the spread of the relative differences as a function of latitude within one orbit.

The discrepancies are generally positive, meaning that the presence of an undetected cloud leads to an underestimation of the ozone concentration, and increasing, as expected, as a function of optical depth. Above 20 km the differences are generally
295 negligible and the largest values are found in the case of a low-level water cloud, with deviations up to ~~15-20~~ 15–20 % below 15 km in the tropics. An undetected thin cirrus, affects the retrieved ozone with a deviation of ~~1-3~~ 1–3 % between 20 and 30 km. The nature of these uncertainties is rather systematic as thin clouds may not be filtered out by the employed cloud flag. As a consequence, this term $\delta_{x,\text{clouds}}$ will be added in the "systematic bias" $\delta_{x,\text{sys}}$ in Sect. 7.

5 Ozone cross section uncertainty

300 The uncertainty related to the ozone absorption cross section used in the retrieval has a relevant contribution to the error budget. A general problem for these considerations is the lack of precise information about the errors as a function of wavelength of the used cross sections (von Clarmann et al., 2020). In our case we decided first to have a look at the impact on the retrieved profile of changing the ~~used cross section~~ cross section source, and then to propagate an estimation of the cross section error into the ozone profile.
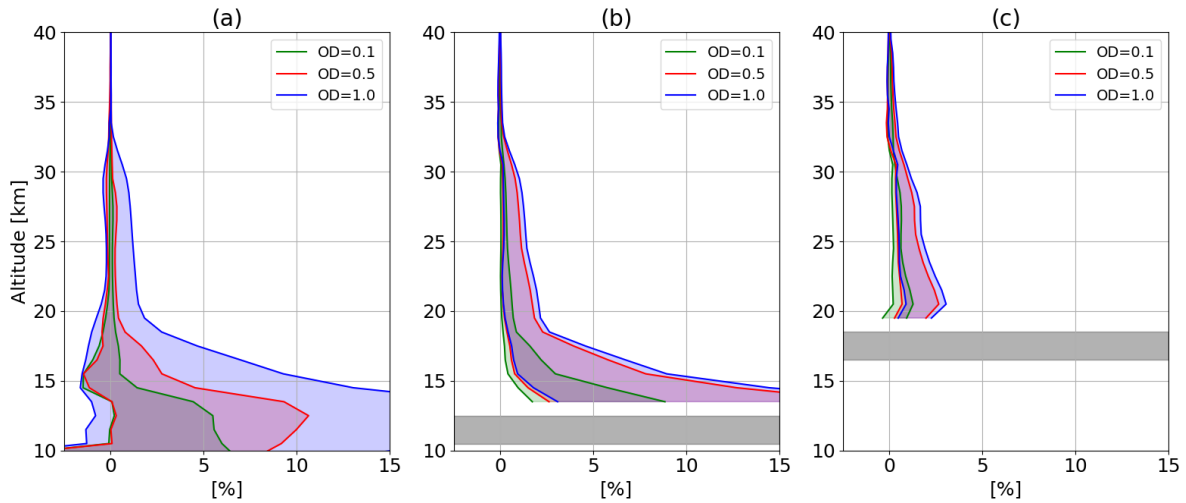
Figure 7. Relative difference between unperturbed and perturbed retrieved ozone profiles, with three different cloud scenarios ((a) water cloud at 7 km, (b) ice cloud at 10 km, (c) ice cloud at 18 km) represented by the gray horizontal bars, for three optical depths (in different colors). The shaded areas give an idea of the spread of the differences as a function of latitude.

5.1 Changing cross section

We consider the following four cross sections: Serdyuchenko (SER, Serdyuchenko et al. (2011); Gorshelev et al. (2014)), Bass and Paur (BP, Bass and Paur (1985)), Brion-Daumont-Malicet (BDM, Daumont et al. (1992); Malicet et al. (1995)) and Voigt (Voigt et al., 2001). It ~~has to be taken into consideration~~ should be borne in mind that different cross sections are reported at different resolutions, spectral ranges and temperatures. BP and BDM are considered for our case only in the UV band, as they do not cover the entire wavelength range needed in the visible spectral range for all temperatures. ~~On the contrary~~In contrast, the Voigt cross section covers ~~also~~ the entire Chappuis band but is available for ~~less~~ fewer temperatures than SER. All cross sections have been convolved to OMPS-LP resolution before use. Figure 8 reports the relative differences for the ozone profiles retrieved with the listed cross sections w.r.t. the standard SER case, averaged over one orbit. The shaded areas display the range of variations observed as a function of latitude/field of view within one orbit.

The obtained uncertainty is up to 6 % when considering the Voigt cross section, and within ± 2 % when comparing with the BP and BDM cases. The last two are of particular interest as they are used (or were used) in the UV retrievals from NASA and University of Saskatchewan for OMPS-LP (Kramarova et al., 2018; Zawada et al., 2018). ~~As a consequence, those values will be used for the final error estimation.~~
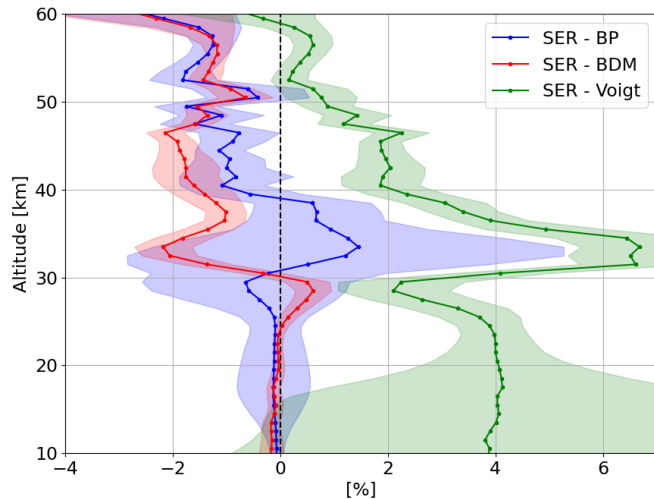
Figure 8. Relative differences between the ozone profiles retrieved considering different cross sections, with SER case used as a reference.

5.2 Uncertainty propagation

320   The SER ozone cross section was measured for 11 temperature values in a broad spectral range covering from 213 to 1100 nm, by using a Fourier transform and an Echelle grating spectrometers. According to Weber et al. (2016), in the UV spectral region, the wavelength-dependent total error on the cross section is in the range 1–3%. The light source stability and the detector noise play in this region a dominant role and, as a consequence, the error has a noise-like, random nature. In the Chappuis band, on the contrary, the light source stability improves whereas the optical densities used for the measurements are lower, leading to

325   an overall error of about 2%, with a predominant bias-like nature.

    As a consequence, the SER cross section is first perturbed by adding a normally distributed random bias with an amplitude of 2 % in the Chappuis band, obtaining 50 perturbed cross sections, used to retrieve a corresponding number of ozone profiles. Figure 9 shows the standard deviation of the retrieved profiles, which gives an estimation of the uncertainty on ozone due to the error on the cross section. Uncertainties for the UV are computed in a similar way: N=50 randomly generated error sequences

330   with an amplitude equal to 2 % of the cross section value are added to the SER cross section and used for the ozone retrieval. The standard deviation of the 50 retrieved ozone profiles is shown in the middle panel of the Fig. 9. For both cases, the reported values refer to an average of the results over several synthetic simulations at different latitudes. The cross section uncertainty in the visible spectral range leads to a larger contribution w.r.t. to the uncertainty in the UV, anyway within 2 %.

    This term $\delta_{\overline{x,cs}} \delta_{x,cs}$ is considered to be systematic, as the actual unknown noise sequence affecting the used cross sections

335   is the same for all retrievals. The uncertainty on ozone is then expected to be of the magnitude estimated via the synthetic retrievals, and constant for all observations, though with unknown sign. We also checked how the cross section error scales with the ozone amount and found that this error component is also multiplicative: by doubling the amount of ozone, the

relative error is found to be fairly constant. We finally performed the synthetic simulations for the same case samples used for the parameter uncertainties in order to better check the variation of this component as a function of latitude and season, without finding any relevant dependence, except for a general increase of the values at high solar zenith angles.
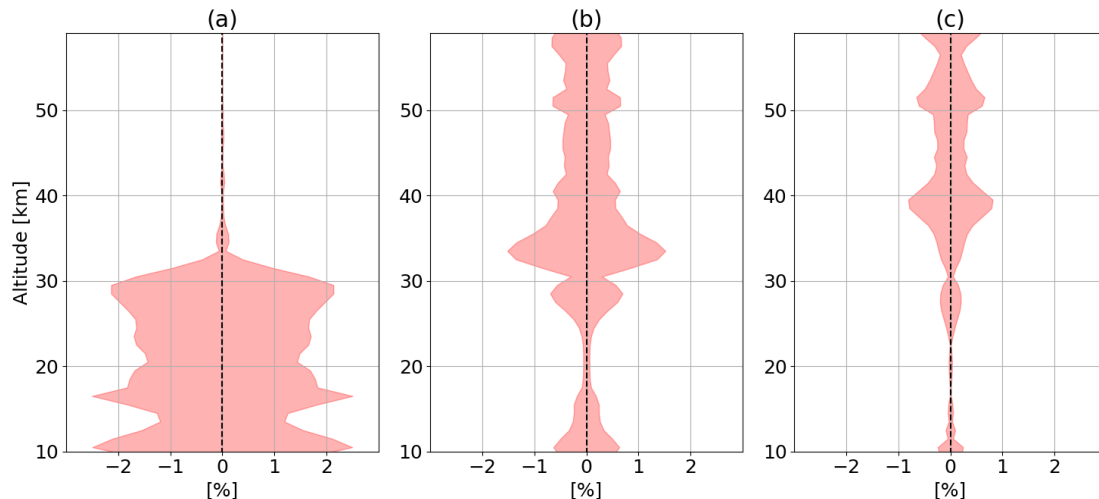
Figure 9. Standard deviation of 50 retrieved ozone profiles, using a perturbed cross section: in panel (a) adding a random bias of 2% (constant over wavelength) in the Chappuis band, and in panel (b) adding random noise with an amplitude of 2% in the UV. Panel (c): standard deviation of 50 ozone profiles after perturbing the temperature and pressure profiles to investigate the uncertainty from the cross section temperature dependence.

Errors from the temperature dependency of the ozone cross section have been also taken into consideration, similarly to Rahpoe et al. (2013): the temperature profile is perturbed by 2 K and the pressure is accordingly changed to keep the air density equal to the unperturbed case. 50 perturbed scenarios are generated and the standard retrieval is run. The standard deviation of the synthetic retrievals is shown in panel (c) of Fig. 9, averaged over one orbit, and this term is within ± 0.5 %, with largest value at about 40 km.

6 Model errors

Model errors refer to the approximations in the RTM used for the ozone retrieval. In order to keep the processing time reasonable, it is required to use a number of simplifications in the model simulations, with the goal to guarantee both computational efficiency and keep the error at the same level as the measurement uncertainties (von Clarmann et al., 2020). In this study we focus our analysis on the scattering approximations and the radiative solver used in SCIATRAN, by considering the geometries typically covered over an OMPS-LP orbit.

First, the radiative tranfer solver in SCIATRAN can be chosen between the scalar Discrete Ordinate Method (DOMS) and the Combined Differential-Integral (CDI) approach (Rozanov et al., 2001). The difference in the ozone profiles retrieved using

different solvers while keeping all other settings constant is shown in panel (a) of Fig. 10, averaged over one orbit, and it's
within 3 % at all altitudes. In addition, the approximation used in the RTM regarding the Earth's sphericity plays also a
non-negligible role. In the standard retrieval version, the CDI solver uses the so-called approximate spherical solution, which
estimates the multiple scattering contribution by integrating the pseudo-spherical radiative field along a line of sight crossing
a "spherical atmosphere" (Rozanov et al., 2001). To test its accuracy, a comparison is performed with a model run including
an additional iteration over the multiple scattering field, leading to a fully spherical solution. The difference shown in panel
(a) is 0.5 - 1 % below 35 km and negligible above. The third curve in panel (a) refers to the "convolution case", i.e. the
difference between the standard profile retrieved from a synthetic spectrum and a profile retrieved from a synthetic spectrum
that is simulated at high resolution, multiplied by a high-resolution solar spectrum, convolved at OMPS-LP resolution and
divided by the low-resolution solar spectrum. The difference is within 2 % at most altitudes, though it increases above 52 km.
The overall contribution of RTM errors $\delta_{\overline{x,RTM}}$ $\delta_{x,RTM}$ to the ozone uncertainty is considered to be systematic as the obtained
estimations are fairly constant regardless the chosen geometry and conditions. This, as shown below, is the largest contribution
to model errors and is taken into account in the total error estimation in Sect. 7.
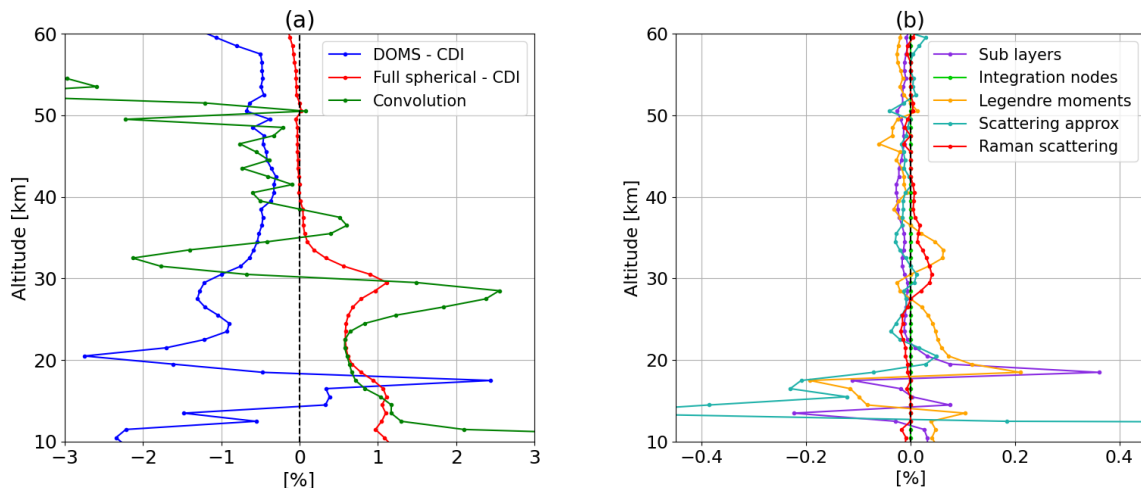
Figure 10. Left panel: relative difference in the ozone profiles for different RT solvers and approximations for atmospheric sphericity. Right panel: relative differences when changing scattering approximation and computation accuracy.

In addition to these main parameters, several other approximations are taken into account. Panel (b) of Fig. 10 displays the
relative differences between profiles retrieved using a different accuracy of the RTM calculations in terms of several parameters.
First, we modify the number of sub-layers (from 10 to 3), which the original altitude grid is divided into, to account for vertical
gradients. No significant differences are found. Negligible differences are seen also when changing the number of "integration
nodes", i.e. changing the nodes used to discretize the angular integrals in the RT equation (from standard 100 to 225), or
changing the number of "Legendre moments", i.e. the approximation used in the expansion of the scattering phase function

(from standard 35 to 64). We also check the relevance of the approximation used in the scattering treatment in SCIATRAN: panel (b) of Fig. 10 shows also the difference ("Scattering approx") between a run with the weighting functions which include
375 the multiple scattering contribution and the standard retrieval, where weighting functions are calculated considering the single scattering contribution only. Differences are also in this case negligible.

To investigate the importance of the Raman scattering, we perform a forward simulation at high resolution (0.01 nm) to simulate a Raman spectrum in SCIATRAN. We then convolve the obtained intensities to OMPS-LP resolution and calculate the difference between the simulated radiance with and without Raman scattering, which corresponds to the Ring contribution.
380 Finally, we add this matrix to the convolved and normalized intensity matrix, and perform a standard retrieval. Differences are shown in Fig.10 and are within 0.1 %, which is explained by a coarse spectral resolution of OMPS-LP.

To account for the impact of the polarization in the retrieved ozone profiles, forward simulations are performed in SCIA-TRAN using the vectorial DOM as radiative transfer solver. Then the synthetic radiance is fed into the standard version of the retrieval and the result is compared with the same scenario excluding polarization. In addition, we perform the following
385 multiplication:

$$\Delta \mathbf{x} = \mathbf{G} \Delta \mathbf{y} \qquad (10)$$

where $\mathbf{G}$ is the gain matrix and $\Delta \mathbf{y}$ is the difference between the intensity matrix from the forward simulations with and without polarization. The result is shown in Fig.11 for several latitudes. The uncertainty evaluation using the matrix multiplication, i.e. linear propagation of the polarization contribution within SCIATRAN (gray shaded area in Fig. 11), shows a very
390 similar behaviour to the analysis of the synthetic retrievals. The contribution is generally within 1% and is largest in the lower stratosphere, with a secondary peak at about ~~30-35~~ 30–35 km, where the retrieval sensitivity is low.
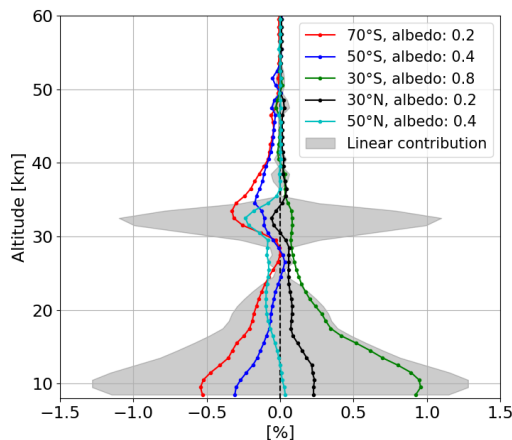
Figure 11. Relative difference between the results with and without polarization for several latitude/surface albedo combinations. The shaded area refers to contribution of the polarization from the linear error propagation, as given by Eq. 10.

7 Estimated total uncertainty

The most important sources of parameter uncertainty from the previous analysis are: pointing errors, particularly in the lower and upper stratosphere, errors on pressure, and, in the lower stratosphere, errors on aerosol extinction coefficient and surface albedo. In addition, the uncertainty originating from the usage of different ozone cross sections and from the propagation of the cross section errors have been shown to be non-negligible. Clouds have also been found to have a relevant impact in the lower stratosphere, as well as RTM approximations throughout the relevant altitude range.

In order to come up with an overall error estimation for the random component of the retrieved ozone profiles, we applied the following rms sum, which also includes a term related to the shift and squeeze procedure described in Sect. 2.2:

$$\sigma_{x,\text{random}}^2 = \sqrt{\sigma_{x,albedo}^2 + \sigma_{x,P}^2 + \sigma_{x,T}^2 + \sigma_{x,aerosol\ ext}^2 + \sigma_{x,TH}^2 + \sigma_{x,shift-squeeze}^2 + \sigma_{x,ret\ noise}^2} \tag{11}$$

This estimation of the random uncertainty implies that each component is independent from the others. As discussed in Fig. 13, there is however some correlation level between different error components and it is sensible to assume that they are, to some extent, interdependent between each other. However, not having reliable information, we assume them to be independent. The overall random uncertainty as a function of altitude and latitude is reported in Table 3 for three latitude bands in the lower, middle and upper stratosphere, in comparison with the retrieval noise component.

Table 3. Relative values of the estimated total random uncertainty due to parameter errors and retrieval noise for tropic, mid- and high latitudes, and for the lower, middle and upper stratosphere. The first value indicates the total estimation (Eq. 11) and the second value indicates the retrieval noise component (from Fig. 1).

Lat. band	Lower stratosphere	Middle stratosphere	Upper stratosphere
tropics	15–30 %, 14 %	4 %, 1 %	6 %, 3 %
mid-latitudes	12 %, 5 %	3 %, 1–2 %	5 %, 3 %
polar latitudes	9 %, 4 %	4 %, 2 %	6 %, 3 %

The systematic uncertainty includes the following terms:

$$\delta_{x,\text{systematic}} = \sqrt{(\delta_{x,\text{unpert}} + \delta_{x,\text{clouds}} + \delta_{x,RTM})^2 + \delta_{x,cs}^2} \tag{12}$$

where the terms with known signs are first summed up and finally the rms with the cross section term is performed. We do not include here the term related to the effects of changing the ozone cross section, as its choice is arbitrary, and shall be evaluated in a case by case base when comparing with other instruments or retrieved products.

Figure 12 displays relevant uncertainty contributions for tropics and high latitudes in panels (a) and (b) respectively, for the summer season, together with the estimated total random uncertainty from Eq. 11 and the systematic bias from Eq. 12.
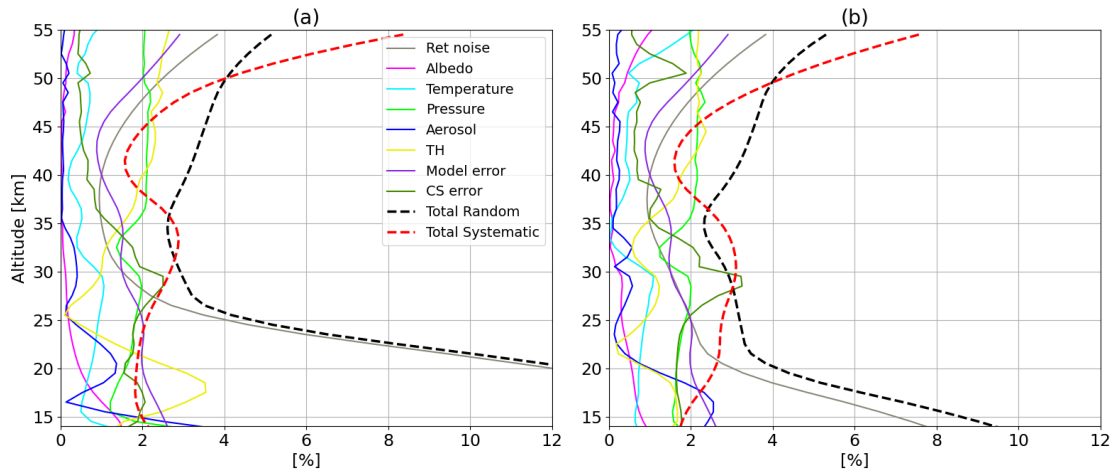
Figure 12. Contributions to the total uncertainty on ozone profiles in the tropics (panel (a)) and at high latitudes (panel (b)), in summer. The total random uncertainty is reported with dashed black line, whereas the estimated systematic bias is shown with red dashed line.

In compliance with the recommendations in von Clarmann et al. (2020) we also calculate correlations between relevant errors as a function of latitude, altitude and season (not shown). Figure 13 presents an example of correlation matrices, reporting correlation coefficients over latitude and altitude for autumn. As an ~~exemplification~~example, only the parameter uncertainties are here taken into consideration.
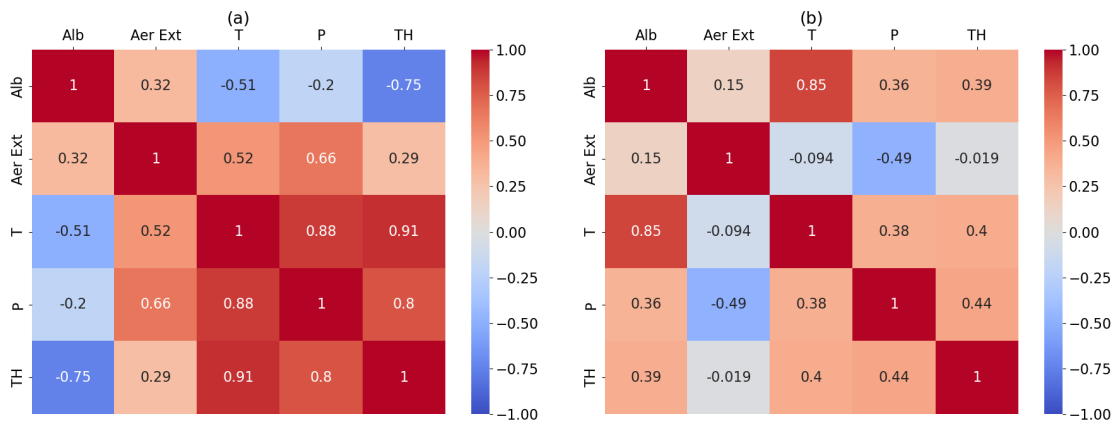
Figure 13. Correlation matrices between several error components, for autumn. Panel (a): Pearson correlation coefficient over latitudes at 30 km; panel (b): correlation coefficient for northern mid-latitudes over altitude.

Panel (a) of this picture shows the Pearson correlation coefficients at 30 km over latitudes, i.e. it indicates how the several error components are correlated in the middle stratosphere when changing the latitude. Panel (b) shows the correlation over

altitude, with low values indicating that the vertical structure of the uncertainties is different, e.g. the vertical profiles of pressure and aerosol extinction components.

One of the final goals of this study is to provide the user interested in uncertainty estimations with arrays giving the error budget as a function of altitude, in 5 latitude bands and 4 seasons. In addition to these values, correlation matrices as shown in Fig. 13 can be obtained upon request.

7.1 Verification of the uncertainties

A comparison between ex-ante and ex-post uncertainties (von Clarmann et al., 2020) is carried out to provide a first verification of the discussed errors. A simple test to validate the precision estimates, i.e. prove that the retrieval noise is not under- or overestimated, is the comparison between the variance of the differences of collocated observations of the same instrument s^2 with the estimated uncertainty σ^2. In general it is expected that:

$$s^2 = 2 * \sigma^2 + \sigma_{x,nat x,\text{nat}}{}^2 \tag{13}$$

Using self-collocated observations we can assume that the term $\sigma_{x,nat}\sigma_{x,\text{nat}}$, i.e. the natural atmospheric variability, approaches zero when collocation criteria become more tight, so that a direct proportionality between estimated precision and standard deviation of the collocated observations is predicted by Eq. 13. This method was followed for example by (Piccolo and Dudhia, 2007; Bourassa et al., 2012; Sofieva et al., 2021) and is referred to as "structure functions".

OMPS-LP is not designed to make repeated measurements of the same atmospheric scene, however an approximation is available from the pairs of measurements located at the intersections of OMPS orbits a few hours apart. If the collocations are sufficiently close in space and time so that atmospheric variations can be neglected, the actual precision of the retrievals can be estimated from the standard deviation of these observations.

Fig. 14 shows the results for OMPS-LP self-collocations with criteria Δ latitude 2°, Δ longitude 10° and Δ time 6 h, considering observations from different orbits, i.e. avoiding matching consecutive measurements when reducing the collocation distance. In different colours, collocations within progressively shorter distance are considered, from 400 down to 100 km. The number of available collocations per year is $\sim$ several thousands for the widest constraints and $\sim$100 for the strictest one.

As shown in the plot, the typical retrieval noise is smaller than the standard deviation of the collocated pairs. This is expected as the precision is obtained from a linear error propagation through the RTM, whereas the retrieval process is thought to be non-linear (Piccolo and Dudhia, 2007). In addition, the standard deviation of self-collocated profiles is expected to be larger because it includes the variability of the atmosphere and above 40 km diurnal variations might also play a role. On the other hand, the estimated total precision represents an upper boundary of the uncertainty for the ozone retrieval. At altitudes around 48-52 48–52 km we notice a peak of the s^2 term, indicating an underestimation of the uncertainties at these altitudes.

7.2 Validation of uncertainties

To validate the total uncertainty we used the χ^2 approach, applied for example by (Migliorini et al., 2004; Clarmann, 2006; Steck et al., 2007). In this approach, the bias between the current data set and collocated observations from an independent
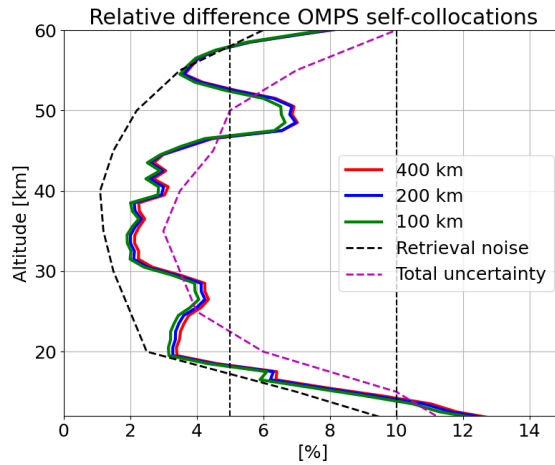
Figure 14. Relative difference between self-collocated OMPS-LP observations when changing the collocation criteria. Superimposed, typical retrieval noise and total random uncertainty for high latitudes.

instrument (MLS in our case) is calculated as follows:

$$b = \frac{\sum_{n=1}^{K}(O_{3_{OMPS,n}} - O_{3_{MLS,n}})}{K}$$ (14)

where K is the number of collocated OMPS-MLS profile pairs. The uncertainty at each altitude is then validated by comparing the de-biased mean square difference of the collocated measurements with the ex-ante estimate of the total uncertainty (σ_x^2), in a χ^2 sense.

$$\chi^2 = \frac{\sum_{n=1}^{K}(O_{3_{OMPS,n}} - O_{3_{MLS,n}} - b)^2/K}{\sigma_x^2}$$ (15)

where σ_x^2 is the sum of three squared terms:

$$\sigma_x^2 = \sigma_{x,\text{OMPS}}^2 + \sigma_{x,\text{MLS}}^2 + \sigma_{x,\text{mismatch}}^2$$ (16)

where $\sigma_{x,\text{OMPS}}$ corresponds to the total random uncertainty from OMPS-LP error budget ($\sigma_{x,\text{random}}$ in Eq. 11), $\sigma_{x,\text{MLS}}$ is the estimated precision for MLS ozone profiles from Livesey et al. (2020) and $\sigma_{x,\text{mismatch}}$ corresponds to expected atmospheric variability for the chosen collocation criteria. The latter has been assessed using model simulations, as described in the study by Laeng et al. (2021). The authors provided us with appropriate values of the expected atmospheric variability, as function of latitude, season and altitude. The χ^2 value, introduced in Eq. 15, gives an indication whether the two data sets agree with the estimated total random error.

Figure 15 shows in panel (a) the three contributions to σ_x^2 together with the numerator of Eq. 15, i.e. the de-biased rms difference between OMPS-LP and MLS, for mid-latitudes in autumn, as an ~~exemplification~~example. The atmospheric variability

profile is available only between 20 and 50 km. The right panel reports the obtained χ^2. Values above 1 indicate that the de-biased rms difference exceeds the estimated combined uncertainty. In our case the ratio stays generally around one, indicating a reasonable assessment of the combined uncertainty, at least between 20 and 45 km. However, at ~~42-45~~42–45 km as well as below 22 km σ_x is too small to explain the bias between the two data sets. Between 25 and 33 km, values about 0.75 indicate a slight overestimation of the combined uncertainty, possibly from OMPS (looking at panel (a)). The low χ^2 values above 45 km indicate that the combined random uncertainty is too large in the upper stratosphere, where MLS has an associated precision that exceed 100 %. A possible overestimation of MLS errors, due to the inclusion of the smoothing error into the reported total random uncertainty, was already reported by Laeng et al. (2015).
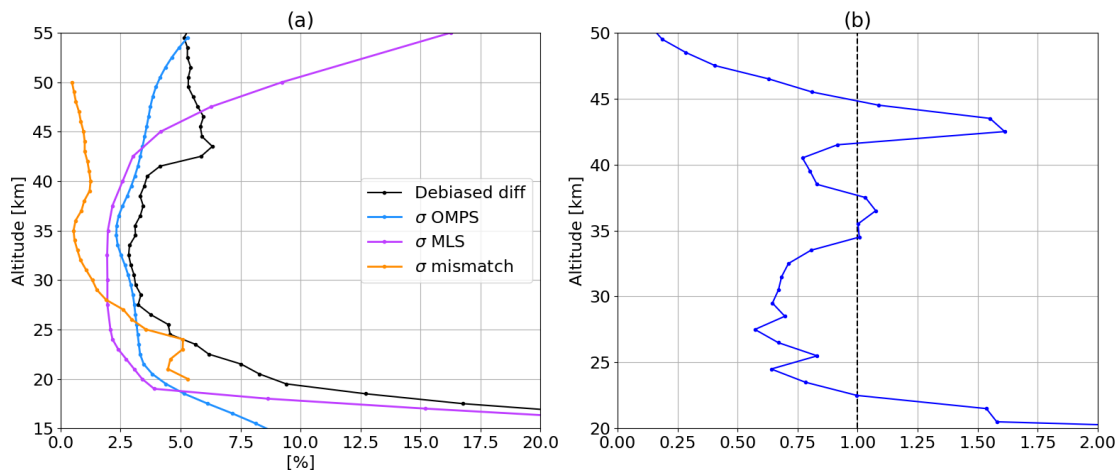
Figure 15. In panel (a), contributing terms from Eq. 16 and de-biased rms difference between OMPS-LP and MLS (numerator of Eq. 15). In panel (b), χ^2 values obtained as validation metric for the combined OMPS-LP and MLS uncertainty.

8 Conclusions

In this paper we presented an extensive error budget assessment for OMPS-LP ozone profiles retrieved at the University of Bremen. Synthetic simulations with the SCIATRAN RTM have been used to estimate a systematic and a random component of the uncertainty on ozone profiles. Effort has been put to characterize the uncertainty components and be compliant with the TUNER recommendations.

First, a retrieval bias was estimated by using unperturbed simulations and comparing them with the "true" ozone profile. This procedure highlighted a negative bias in the upper stratosphere, fairly constant with latitude, of about 5–8 % above 50 km. Then, parameter uncertainties have been extensively investigated and we found that the most relevant contribution is the pointing accuracy of the instrument, especially in the upper stratosphere, with values up to ~~3-4~~3–4 %, followed by pressure and temperature error contributions. The precision of the retrieved aerosol extinction profiles plays a relevant role below 25 km

with a parameter uncertainty up to 4 %. The effect of an undetected cloud was also discussed and a potential negative bias in the lower stratosphere highlighted. Another relevant source of error is related to the chosen ozone cross section. It was shown that changing the cross section in the UV leads to variations of up to 2 % above 30 km in ozone. The cross section error propagated into L2 profiles is estimated to be on the order of $\pm$ 2 % in the lower stratosphere and decreasing with altitude, this term has a systematic nature. In addition, the analysis of the error from model approximations revealed that the radiative transfer solver and polarisation effects may add to the total random ozone error a contribution up to ~~1-2~~1–2 %.

Overall, most error components have shown a multiplicative nature. By summing up the single random uncertainty components and the discussed systematic errors, we provided reference values for the total uncertainty as a function of altitude and latitude. Typically, the total random uncertainty on retrieved ozone profiles is expected to lie in the range ~~5-30~~5–30 % for the lower stratosphere, 3–5 % for the middle stratosphere and 5–7 % for the upper stratosphere. The systematic uncertainty is mainly related to cloud contamination below 20 km, producing an underestimation in the ozone profile, and to the retrieval bias above 50 km, also as underestimation of the true profile.

Uncertainties have been verified using self-collocated OMPS-LP measurements at high latitudes, finding that the standard deviation of the collocated profiles approaches the typical retrieval noise values, except at altitudes about 28 and 50 km. The validation of the uncertainties in a χ^2 approach showed that the combined OMPS-LP and MLS uncertainty is reasonable, but generally overestimated: a fair consistency between the combined uncertainty and the de-biased difference between the two satellites was found between 22 and 45 km, whereas in the upper stratosphere, MLS reported uncertainties are too large.

Author contributions. CA performed most of the simulations and data analysis for the uncertainty estimations and wrote the manuscript. AR provided the RT expertise needed in this study, supervised and guided the development of the error analysis and reviewed the paper. VG provided information about the CS uncertainty and reviewed the paper. AL provided atmospheric variability arrays and reviewed the paper. JPB, who leads the project, contributed to the review of the manuscript and the scientific outcome.

Competing interests. The authors declare that they have no conflict of interest

Acknowledgements. This work was partially funded by a DAAD PRIME scholarship and the Living Planet Fellowship SOLVE, and was supported by the University and State of Bremen. We are very grateful to Thomas von Clarmann, for his insights and suggestions about this work, especially regarding the characterization of the errors and the compliance with TUNER recommendations. We would like to acknowledge NASA team for the provided data and the support. Part of the data processing has been done on the German HLRN (High Performance Computer Center North). GALAHAD Fortran Library was used for calculations.

References

Arosio, C., Rozanov, A., Malinina, E., Eichmann, K.-U., von Clarmann, T., and Burrows, J. P.: Retrieval of ozone profiles from OMPS limb scattering observations, Atmospheric Measurement Techniques, 11, 2135–2149, https://doi.org/10.5194/amt-11-2135-2018, 2018.

Bass, A. and Paur, R.: The ultraviolet cross-sections of ozone: I. The measurements, in: Atmospheric ozone, pp. 606–610, Springer, 1985.

Bogumil, K., Orphal, J., Burrows, J. P., et al.: Temperature dependent absorption cross sections of O3, NO2, and other atmospheric trace gases measured with the SCIAMACHY spectrometer, in: Proceedings of the ERS-Envisat-Symposium, Goteborg, Sweden, 2000.

Bourassa, A., McLinden, C., Bathgate, A., Elash, B., and Degenstein, D.: Precision estimate for Odin-OSIRIS limb scatter retrievals, Journal of Geophysical Research: Atmospheres, 117, 2012.

Clarmann, T. v.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, Atmospheric Chemistry and Physics, 6, 4311–4320, 2006.

Cortesi, U., Lambert, J., De Clercq, C., Bianchini, G., Blumenstock, T., Bracher, A., Castelli, E., Catoire, V., Chance, K., De Maziere, M., et al.: Geophysical validation of MIPAS-ENVISAT operational ozone data, Atmospheric Chemistry and Physics, 7, 4807–4867, 2007.

Daumont, D., Brion, J., Charbonnier, J., and Malicet, J.: Ozone UV spectroscopy I: Absorption cross-sections at room temperature, Journal of Atmospheric Chemistry, 15, 145–155, 1992.

Flynn, L., Long, C., Wu, X., Evans, R., Beck, C., Petropavlovskikh, I., McConville, G., Yu, W., Zhang, Z., Niu, J., et al.: Performance of the ozone mapping and profiler suite (OMPS) products, Journal of Geophysical Research: Atmospheres, 119, 6181–6195, 2014.

Fox, N.: A Quality Assurance Framework for Earth Observation: A guide to expression of uncertainty of measurements, Tech. rep., Tech. Rep. QA4EO-QAEO-GEN-DQK-006, Group on Earth Observation., 2010.

Gorshelev, V., Serdyuchenko, A., Weber, M., Chehade, W., and Burrows, J.: High spectral resolution ozone absorption cross-sections–Part 1: Measurements, data analysis and comparison with previous measurements around 293 K, Atmospheric Measurement Techniques, 7, 609–624, 2014.

Jaross, G., Bhartia, P. K., Chen, G., Kowitt, M., Haken, M., Chen, Z., Xu, P., Warner, J., and Kelly, T.: OMPS Limb Profiler instrument performance assessment, Journal of Geophysical Research: Atmospheres, 119, 4399–4412, 2014.

Kramarova, N. A., Bhartia, P. K., Jaross, G., Moy, L., Xu, P., Chen, Z., DeLand, M., Froidevaux, L., Livesey, N., Degenstein, D., et al.: Validation of ozone profile retrievals derived from the OMPS LP version 2.5 algorithm against correlative satellite measurements, Atmospheric Measurement Techniques, 11, 2837–2861, 2018.

Laeng, A., Hubert, D., Verhoelst, T., Von Clarmann, T., Dinelli, B., Dudhia, A., Raspollini, P., Stiller, G., Grabowski, U., Keppens, A., et al.: The ozone climate change initiative: Comparison of four Level-2 processors for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), Remote Sensing of Environment, 162, 316–343, 2015.

Laeng, A., von Clarmann, T., Errera, Q., Grabowski, U., and Honomichl, S.: Satellite data validation: a parametrization of the natural variability of atmospheric mixing ratios, Atmospheric Measurement Techniques Discussions, 2021, 1–15, https://doi.org/10.5194/amt-2021-373, 2021.

Langland, R. H., Maue, R. N., and Bishop, C. H.: Uncertainty in atmospheric temperature analyses, Tellus A: Dynamic Meteorology and Oceanography, 60, 598–603, 2008.

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Lambert, A., Manney, G. L., Millán Valle, L. F., Pumphrey, H. C., and Santee, M. L.: Version 4.2x Level 2 and 3 data quality and description document, 2020.

Malicet, J., Daumont, D., Charbonnier, J., Parisse, C., Chakir, A., and Brion, J.: Ozone UV spectroscopy. II. Absorption cross-sections and temperature dependence, Journal of atmospheric chemistry, 21, 263–273, 1995.

550 Malinina, E., Rozanov, A., Rozanov, V., Liebing, P., Bovensmann, H., and Burrows, J. P.: Aerosol particle size distribution in the stratosphere retrieved from SCIAMACHY limb measurements, Atmospheric Measurement Techniques, 11, 2085–2100, 2018.

Migliorini, S., Piccolo, C., and Rodgers, C. D.: Intercomparison of direct and indirect measurements: Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) versus sonde ozone profiles, Journal of Geophysical Research: Atmospheres, 109, 2004.

Moy, L., Bhartia, P. K., Jaross, G., Loughman, R., Kramarova, N., Chen, Z., Taha, G., Chen, G., and Xu, P.: Altitude registration of limb-
555 scattered radiation, Atmospheric Measurement Techniques, 10, 167–178, 2017.

Nowlan, C., McElroy, C., and Drummond, J.: Measurements of the O2 A-and B-bands for determining temperature and pressure profiles from ACE–MAESTRO: Forward model and retrieval algorithm, Journal of Quantitative Spectroscopy and Radiative Transfer, 108, 371–388, 2007.

Piccolo, C. and Dudhia, A.: Precision validation of MIPAS-Envisat products, Atmospheric Chemistry and Physics, 7, 1915–1923, 2007.

560 Rahpoe, N., Savigny, C. v., Weber, M., Rozanov, A., Bovensmann, H., and Burrows, J.: Error budget analysis of SCIAMACHY limb ozone profile retrievals using the SCIATRAN model, Atmospheric Measurement Techniques, 6, 2825–2837, 2013.

Rault, D. F. and Loughman, R. P.: The OMPS Limb Profiler environmental data record algorithm theoretical basis document and expected performance, IEEE transactions on geoscience and remote sensing, 51, 2505–2527, 2012.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, vol. 2, World scientific, 2000.

565 Rozanov, A., Rozanov, V., and Burrows, J. P.: A numerical radiative transfer model for a spherical planetary atmosphere: Combined differential–integral approach involving the Picard iterative approximation, Journal of Quantitative Spectroscopy and Radiative Transfer, 69, 491–512, 2001.

Rozanov, V., Rozanov, A., Kokhanovsky, A., and Burrows, J.: Radiative transfer through terrestrial atmosphere and ocean: software package SCIATRAN, Journal of Quantitative Spectroscopy and Radiative Transfer, 133, 13–71, 2014.

570 Serdyuchenko, A., Gorshelev, V., Weber, M., and Burrow, J. P.: New broadband high-resolution ozone absorption cross-sections, Spectroscopy Europe, 23, 14, 2011.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections. Part 2: Temperature dependence , Atmospheric Measurement Techniques, 7, 625–636, https://doi.org/10.5194/amt-7-625-2014, 2014.

Sheese, P., Walker, K. A., and Boone, C. D.: Uncertainty budgets for ACE-FTS temperature, ozone, and water vapour profiles, in: AGU Fall
575 Meeting Abstracts, vol. 2020, pp. A044–0006, 2020.

Sofieva, V. F., Lee, H. S., Tamminen, J., Lerot, C., Romahn, F., and Loyola, D. G.: A method for random uncertainties validation and probing the natural variability with application to TROPOMI on board Sentinel-5P total ozone measurements, Atmospheric Measurement Techniques, 14, 2993–3002, 2021.

Steck, T., Clarmann, T. v., Fischer, H., Funke, B., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., et al.:
580 Bias determination and precision validation of ozone profiles from MIPAS-Envisat retrieved with the IMK-IAA processor, Atmospheric chemistry and physics, 7, 3639–3662, 2007.

Voigt, S., Orphal, J., Bogumil, K., and Burrows, J.: The temperature dependence (203–293 K) of the absorption cross sections of O3 in the 230–850 nm region measured by Fourier-transform spectroscopy, Journal of Photochemistry and Photobiology A: Chemistry, 143, 1–9, 2001.

585 Von Clarmann, T., De Clercq, C., Ridolfi, M., Höpfner, M., and Lambert, J.-C.: The horizontal resolution of MIPAS, Atmospheric Measurement Techniques, 2, 47–54, 2009.

von Clarmann, T., Degenstein, D. A., Livesey, N. J., Bender, S., Braverman, A., Butz, A., Compernolle, S., Damadeo, R., Dueck, S., Eriksson, P., et al.: Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature, Atmospheric Measurement Techniques, 13, 4393–4436, 2020.

590 Weber, M., Gorshelev, V., and Serdyuchenko, A.: Uncertainty budgets of major ozone absorption cross sections used in UV remote sensing applications, Atmospheric Measurement Techniques, 9, 4459–4470, 2016.

Zawada, D. J., Rieger, L. A., Bourassa, A. E., and Degenstein, D. A.: Tomographic retrievals of ozone with the OMPS Limb Profiler: algorithm description and preliminary results, Atmospheric Measurement Techniques, 11, 2375–2393, 2018.