Anonymous Referee #1
Atmos. Meas. Tech. Discuss., 8, C1595–C1601, 2015
Received and published: 18 June 2015

We would like to thank reviewer #1 for the constructive comments, which helped to improve the quality of our manuscript. In this document we provide our replies to the reviewer's comments. The original comments of the reviewer are numbered and set in italic font. Line numbers, page numbers and figure numbers in the reviewer's comment refer to the original manuscript. Following every comment we provide our reply.
We provide a revised version of the manuscript, in which all changes are highlighted, New sections and added text are provided in red. In our replies to the comment we provide line numbers, page numbers and figure numbers of the revised version of the manuscript, if not stated differently.

Major comments:

M-1.) The authors tend to put in opposition total O3 products generated using given a priori profiles and their effective column approach. To my opinion, such a distinction doesn't make sense. Indeed, using a single (scaled) O3 profile is somehow a very simple climatology. I think that the important message is that a smoothing error is associated to every O3 column estimate and that this error is directly related to the quality of the a priori profile data base. Providing the averaging kernels is important in any case to allow users to apply them in case they have more reliable a priori information. Of course, when using a single profile for all ozone retrievals, having the averaging kernels is even more important as the smoothing error associated to the "effective" column will be generally much larger. Without those kernels and realistic information on the true profile, the effective columns are probably useless in numerous conditions on contrary to a total O3 product generated with a more reliable a priori profile data base. In summary, I think the best option is to use the best a priori profile database for the retrievals and also to provide users with averaging kernels. Since the authors used a better climatology in the validation section, I think that they will agree with that statement but they should provide a clearer message through the whole manuscript.

Adjusted: With restructuring the paper and rephrasing the abstract, we clarified that there are two different ways of interpreting the retrieval product, the total ozone column and the effective ozone column. To our knowledge for the first time, a validation of the retrieved effective ozone column is presented, which requires the use of the column averaging kernel. We thus put the focus of the validation on the ozone information that can be extracted from the measurements thereby diminishing the impact of the particular choice of the reference profile. This is clearly shown in the manuscript (see Fig. 13, in which the red bars show similar retrieval errors for the three different scaling profiles used in the validation).
To validate the effective column method, however, we require measurements of the vertical ozone distribution which limits the validation in its temporospatial coverage. This is contrary to the retrieved total ozone columns which are interpreted as an estimate of the true total ozone column and hence can be directly compared to ground-based total ozone measurements, which are available in a much larger quantity. This approach relies on accurate a priori knowledge of the reference scaling profile in the direct fitting method to retrieve total ozone columns as accurately as possible and is applied in the literature by Lerot et al., 2010 and 2014.
We do not intend to put both products in opposition, since both are valid and useful. We want to stress the different and complementary ways of interpreting these ozone column products. For this reason, we think it does make sense to distinguish between both interpretations and consequently, we keep the discussion of the different views on the data product in the revised version of the manuscript.

M-2.) It is not exact to say that Lerot et al. (2014) rely on the scaling of a reference profile. They use a total O3 column classified climatology, which provides O3 profiles as a function of the month, latitude and the O3 column itself. The stratospheric profile shape depends therefore on the O3 column. At each iteration of the fit, the appropriate O3 profile is interpolated through the database (see also Van Roozendael et al., JGR, doi:10.1029/2005JD006375). Because of this, the product quality is expected to be better than using a simple scaling approach. In addition, the O3 products presented in their work also contain the averaging kernels as well as a priori profiles corresponding to the retrieved columns. Please adapt the manuscript wherever it is necessary.

Changed: The text has been adjusted in the revised version of the manuscript on page 3, lines 80 – 86 and page 8, lines 247 – 250. When referencing to the work of Lerot et al.(2010) and Lerot et al. (2014) we extended the terminology 'profile scaling' by mentioning that the a priori profile is being updated during the iteration through a total ozone column classified climatology.

M-3.) Page 4920 - lines 28-29: Have you estimated how a real trend in the ozone columns would impact this assumption? I agree that for short time series, it is not really relevant, but that might be the case for longer periods.

Changed: On page 10, lines 316 – 319, we provide the trend of the total ozone column of 0.08% ±0.13% per year between 2000 and 2013 as reported in the WMO 2014 ozone assessment report (Pawson et al., 2014) and conclude that this trend would not introduce significant errors if not accounted for in the degradation correction in the observed period 2007 – 2010. Furthermore, we like to stress that we perform a degradation correction relative to 2007.

M-4.) Page 4923 - lines 16-19: Have you verified that the spectral structures of the ratio of the tabulated reflectances do not depend significantly on the viewing zenith angle, azimuth angle, albedo and altitude height?

Changed: On page 5, lines 158 – 164, we state the dependence of the lookup table on solar zenith angle as well as the ozone abundance in the atmosphere to avoid interference of the filling-in due to Raman scattering and the retrieved ozone column. Adding extra parameters to the lookup table does not improve the spectral residuals any further than for the combination of the effective amplitude 'a' with the lookup table.
However, the reviewer is right that Raman scattering depends on a suite of atmospheric parameters including viewing zenith angle, azimuthal angle albedo and scattering height but also the amount of ozone in the atmosphere (see e.g. Joiner et al. 1995, Vountas et al. 1998, Landgraf et al. 2004). A proper tabulation of the reflectance ratios as function of all relevant parameters would result easily in a huge lookup table, which next to storage issues also hampers any practical implementation in the retrieval. Due to this, we followed a slightly different approach to account for atmospheric Raman scattering in the retrieval, using tabulated values of the reflectance ratio combined with the fit of an effective amplitude 'a'. For the lookup table, we have tabulated the reflectance ratio as function of solar zenith angle and the total amount of ozone. Solar zenith angle is one of the main parameters which determines the scattering geometry, the fraction of multiple scattering and so the effect of inelastic Raman scattering (see e.g. Joiner et al. 1995). Although the effect of ozone on corresponding Ring spectra is significantly smaller (see e.g. Landgraf et al, Fig. 6), this parameter is chosen as a relevant parameter of the lookup table to avoid the interference of the filling-in of ozone absorption features in the Huggins band with the retrieved total column of ozone.

M-5.i) Page 4923 - lines 23-29: Could you specify the instrumental spectral response function you

use in your algorithm?
M-5.ii) Could you explain the physical origin of the two spectral calibration parameters? Does s account for the Doppler shift and $\Delta\lambda ISRF$ for possible wavelength registration issues in the measured solar spectrum? Are those two parameters not cross-correlated in the fit?

Changed:
M-5.i) On page 5, lines 134 – 135, we implemented that we use a Gaussian instrumental spectral response function has been implemented with a FWHM of 0.3nm.
M-5.ii) On page 6, lines 170 – 174, we explain that the spectral shift $\Delta\lambda_S$ mainly accounts for the Doppler shift of the solar spectrum, but in fact is also used to mitigate issues with respect to the spectral calibration of the solar spectrum. In contrast, the spectral shift $\Delta\lambda_{ISRF}$ allows for adjusting the spectral position of atmospheric absorption features in the simulated earth radiance measurement with respect to the satellite measurement. Thus, both parameters are of different physical origin and thus can be retrieved independently.

M-6.i) Pages 4927-4928 and Fig. 1: Please provide additional information on the parameters used for the simulations (solar and viewing zenith angle, cloudiness, albedo,. . .). The impact of the a priori information may be significantly dependent on those parameters. In particular, at extreme solar zenith angles (>80), the error due to the a priori profile may increase up to several percent. So stating that the error due to the profile is on the order of 1% is overoptimistic for those conditions.
M-6.ii) Could clarify this in the text and provide specific estimates at those large SZAs?

Changed:
M-6.i) In the figure caption of Fig. 1, we provide information on the addiational parameter used for the simulation: a satellite solar zenith angle of 70 degrees, a satellite viewing angle of 30 degrees have been used. The simulations were done for a clear sky atmosphere using an albedo of 0.1. The values are representative for the chosen day.
M-6.ii) On page 9, lines 293 – 294, we provide the result of a sensitivity study for a high-latitude station with large solar zenith angles. In that case the null space error mounts up to 5%.

M-7.) Removing the smoothing error from the satellite-ground-based comparisons with the application of the averaging kernels relies on the linearization of the forward model. I wonder how efficient is the process when the a priori profile used for the inversion is very different than the real profile (e.g. the US standard profile is used for retrieving columns in ozone hole conditions). Have you tested this? I'd recommend to perform a few sensitivity tests to better assess what are the possible limitations of the approach.

Changed: See also M-6.) For a southern hemisphere station we performed retrievals with simulated measurements for austral winter (low ozone conditions) and several solar zenith angles and found null space errors up to 5%. Moreover, we would like to stress here that the validation accuracy of the effective column approach does not depend significantly on the choice of the a priori to be scaled which we show in Sect. 5. This is what we expect from theory in the linear domain of the forward model.

M-8.) Page 4929 - lines 26-27: Do you have any idea why the statistics is poorer for pixel 1? Systematic clouds? Does it mean that pixel 1 is systematically discarded in your O3 column product?

Changed: Fig. 2 has been re-done and includes pixel 1 and the text on page 10, lines 334 – 335 is adjusted. Pixel 1 had been systematically excluded from the degradation analysis due to an index error. The column product was not affected by that bug.

M-9.) Page 4931 - lines 18-20: It is also important to mention that the systematic error in the Dobson measurements due to their temperature dependence leads to significant seasonal/SZA dependences at mid- and high-latitudes. This is most likely the main cause for the larger SZA dependence in the satellite-Dobson differences in the validation section. You should adapt the text accordingly.

Changed: On page 11, line 381 – page 12, line 384 we included a reference to Bernhard et al. (2005) who report that a fixed stratospheric temperature in total ozone retrievals from Dobson spectrometers, as well as the assumption of a delta ozone layer at a fixed height cause both seasonal and solar zenith angle dependencies in the ground-based measurements.

M-10.) Page 4935 - line 12: Again for high latitude stations, the error may be much larger, especially in local winter and during the ozone hole season.

Adjusted. See also bullets M-6.) and M-7.) of this reply.

M-11.) In the paragraph dealing with the effect of cloudiness, could you clarify how you model the effects of clouds within your algorithm (IPA?)? Do you add a ghost column to account for the partial O3 column between the cloud and the ground or is the systematic underestimation in case of strong cloud contamination could come from the shielding effect?

Adjusted: In Sect. 4.3.1, we clarify that the presented algorithm assumes a clear sky atmosphere, thus no cloud model is applied. Hence, we rely on cloud filtering to exclude "heavily contaminated" scenes. The dependence of the retrieval error on a misleadingly assumed clear sky atmosphere is addressed in the performance study of the key parameter cloudiness. Indeed, the correlation between the systematic underestimation of the total ozone column and "strong cloud contamination" comes from the shielding effect.

M-12.) Page 4939 - line 14 (and abstract): I wouldn't say that polarization can be completely ignored as the current goal in the development of total ozone algorithms is to reach a level of 1% accuracy in most geophysical conditions. To reach that level of accuracy, I think that the forward model should be as accurate as possible, and including polarization contributes to this. Although I agree that the effect is relatively small, some seasonal-dependent errors may be reduced when accounting for polarization.

Changed: On page 1, lines 9 – 10 (the abstract), the strong statement that polarization can be ignored in the abstract has been weakened to 'pseudo spherical scalar radiative transfer is fully sufficient for the purpose of this retrieval' in the revised version of the manuscript. Moreover, in Sect. 4.3.3 (page 17, lines 590 – 591) a statement to the benefit of higher accuracy accounting for polarization in radiative transfer has been implemented.

M-13.) About the analysis of the instrumental degradation, is it not possible to extend the analysis to the full time series? We have now more than eight years of GOME-2A data, while only the first *four years are presented here. That might give an even more convincing demonstration of the importance of correcting for this degradation to have accurate retrievals.*

We agree that analyzing the whole GOME-2/A data set would demonstrate the need of degradation correction. However, the main goal to compare the two different ways of validation (direct approach and effective approach) and to study forward model errors have been achieved with the dataset at hand. Extending the dataset would be a major effort and is not feasible within the constraints of the project.

M-14.) Page 4943 - lines 13-15: This has already been reported in Loyola et al. (JGR, 2011, doi:10.1029/2010JD014675). You might refer to this study.

Changed: On page 14, lines 468 – 470 and page 20, line 692, a reference to this study has been implemented.

M-15.) Fig. 16: The Lerwick Dobson station is not really appropriate for this illustration as the Dobson measurements suffer from a SZA dependence due to the temperature seasonal variation. The SZA is unfortunately directly correlated to the scattering angle. Could consider to use a Brewer station instead?

Changed: We acknowledge the critics because we were not aware of the results when implementing Fig. 16 in the original version of the manuscript. We considered the Brewer data of Goose Bay instead. This has no consequences for the conclusions of the manuscript. In order to reduce the number of figures in the revised version of the manuscript, this figure has been removed, but the message is brought by Fig. 10 as well.

Editorial comments:
E-1.) Page 4919 - line 27:
The latest version of the operational O3MSAF/EUMETSAT total ozone product is GDP v4.7 and as been presented in Hao et al., AMT, 2014 (doi:10.5194/amt-7-2937-2014). You should use this reference.

Changed: On page 2, line 47 we implemented a reference to this study.

E-2.) Page 4920 - line 2:
"measurments" to be replaced by "measurements"

Corrected: On page 2, line 51, the typing error has been corrected.

E-3.) Page 4920 - line 3-4:
please specify which data products and versions have been used in that study for those different sensors.

Changed: Page 2, lines 47 – 57. The data products and processor versions used in the study by Koukouli et al. (2012) are mentioned. For GOME-2/MetOp-A processor version GDPv4.4 has been used for level 1b data version 4. For GOME/ERS-2 retrievals have been performed with processor version GDPv4.1 and for SCIAMACHY with processor version SGP5.0. OMI/Aura level 1b data version collection 3 have been processed with processor versions TOMSv8.5 and DOAO3v1.0.1.

E-4.) Page 4923 - line 10: I would specify that several O3 cross-sections at different temperatures are used.

Adapted: On page 5, line 151, we adapted the text to "..., which are calculated for the given temperature profile, ..." to address the used of temperature dependent ozone cross sections.

E-5.) Eq. (9):
please define 'i'

Adjusted: On page 7, line 210 the index 'i' has been defined. It is the index for a wavelength.

E-6.) Page 4930 - line 20:
one extra comma.

Changed: On page 18, lines 624 – 625 the extra comma after both has been removed.

E-7.) Page 4932 - line 16:
I think 2 is not defined.

Changed: In Eq. (18) we define a χ^2 criterion of 2 and $\chi^2 = 1/(N-1)\ \Sigma_{i=1}^{N}\ [(r_{meas,i} - r_{mod,i})/sqrt(S_{e\ i,i})]^2$ where r_{meas} and r_{mod} are the measured and simulated reflectance, respectively, and S_e is the measurement error covariance.

E-8.) Section 4.2:
I'd recommend to split this section into several subsections to facilitate the reading.

Done: The whole validation section has been split into sub sections and furthermore, we separated the study addressing the retrieval performance with respect to the key parameters and moved it into its own section:
4. Performance analysis
Here, we describe the preparation of the datasets for the performance analysis and validation and introduce accuracies of the ground-based instruments.
4.1. Data filtering
This section introduces the quality filtering applied to the retrieved total ozone column.
4.2. Topography correction and instrument degradation correction
The effect of a topography correction and instrument degradation are discussed.
4.3. Forward model errors
Applying the data filtering and data correction we discuss three forward model errors in the following sub sections.
4.3.1. Cloudiness
4.3.2. Earth's sphericity
4.3.3. The scalar radiative transfer approximation
5. The effect of regularization
In this section we discuss the effective column validation using the total column averaging kernel.

E-9.) Page 4934 - line 27:
‘coregistation’ to be replaced by ‘coregistration’

Corrected: The sentence changed completely in the course of restructuring.

E-10.) Page 4935 - line 8:
minute?

Changed: On page 18, line 598 'minute' has been replaced by 'small'.

E-11.) Page 4935 - line 10:
‘SOAZ’ to be replaced by ‘SAOZ’

Corrected: On page 18, line 599, the typing error has been corrected.

E-12.) Page 4935 - line 27:
one extra comma.

Corrected: On page 15, lines 496 – 497 the comma after “Applying this correction” has been removed.

E-13.) Page 4937 - line 5:
“to” instead of “for”. The Dobson is more susceptible to solar zenith angle dependencies than what?the Brewer?

Changed: On page 16, lines 533 – 534 we changed 'for' into 'to' and added Brewer at the end of the Sentence.

E-14.) Page 4937 - line 8:
“Ushuaha” to be replaced by “Ushuahia”.

Corrected: On page 16, line 537 the typing error has been corrected.

E-15.) Page 4943 - line 10:
“which are larger than the retrieval errors. . .” instead of “which are more strongly affected compared to the retrieval errors. . .”

Changed: The whole sentence has been rephrased in the course of restructuring.

E-16.) Page 4943 - line 26:
this sentence is confusing. The sensitivity of nadir-viewing UV-Visible instruments is never

maximum in the lowermost troposphere. However, it is clearly larger there when there is no cloud contamination. Is it what you mean? Please rephrase.

Changed: The sentence has been removed from the revised version of the manuscript.

E-17.) Figs. 4 and 5:
Please homogenize the legends (effective or with nullspace, and direct or without nullspace).

Changed: In Fig. 12, the figure caption has been adapted to match the legend of the figure. Fig. 5 (original version) has been removed from the revised version of the manuscript.

E-18.) Fig 9.: top-right panel:
legend not consistent with caption.

Changed: In Fig. 3, the caption of the top panels has been changed accordingly to retrieved and ground-based. Furthermore, the caption describing the lower panels of the figure has been adapted to clearly relate the text of the figure caption to the legend in the figure.

E-19.) Fig. 10:
"percent " instead of "per cent"

Corrected: In Fig. 6, the typing error has been corrected.

E-20.) Fig. 14:
Please specify the values of the other parameters used for these simulations (e.g. VZA, O3 column,. . .).

Changed: In Fig. 9, the figure caption has been adjusted. The measurements are simulated using satellite solar zenith angles and viewing angles adapted from the Lerwick validation data set and a clear sky model atmosphere using a Lambertian surface albedo of 0.1. The ozone profile has been taken from the US standard atmosphere.

E-21.) Fig. 15:
"include" instead of "including".

Changed: In Fig. 10, a redraft has been implemented

E-22.) Fig. 19:
"west pixel bins, which include pixels numbers 1-12 and 13-24, respectively" instead of "west pixel bins, separated between pixel numbers 12 and 13"

Changed: In Fig. 5, the redraft has been implemented.