Anonymous Referee #2
Atmos. Meas. Tech. Discuss., 8, C1861–C1863, 2015
Received and published: 2 July 2015

We would like to thank reviewer #2 for the constructive comments that aided us to improve our manuscript. In this document we provide our replies to the reviewer's comments. The original comments made by the reviewer are numbered and typeset in italic font. Line, page and figure numbers in the reviewer's comments refer to the original manuscript. Following every comment we give our reply.

We provided a revised version of the manuscript in which all changes are indicated: Newly added sections are typeset in red. In our reply we give page and line numbers that refer to the revised manuscript, unless otherwise stated.

General comments:

G-1.) The first half of the paper is well written and flows nicely, while the second half [from the Validation Section onwards] leaves room for improvement. I believe that a restructuring of the Validation Section is in order, splitting it into clear sub-sections that will help the interested reader follow the steps taken.

Changed: We restructured Sect. 4 and split it into several subsections. The structure is as follows:
4. Performance analysis
4.1. Data filtering
This subsection 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.

G-2.) The abstract is not too clear either, as to what is the paper to present: a new, operational, scientific algorithm? A to-be operational scientific algorithm? A case-study type study? And if this algorithm is not to be evolved into one that can be used for creating TOCs, maybe the authors should consider giving recommendations based on their findings to the PIs of the operational algorithms, for e.g. the O3MSAF EUMETSAT consortium which hosts the operational GOME2/A and GOME2/B TOC retrievals.

Changed: The abstract has been rephrased to stress that the present study is a sensitivity study of a direct fitting approach by means of adjusting the scaling of a reference ozone profiles. Direct fitting requires accurate radiative transfer modeling and thus we address three forward model errors: (1) the use of a clear sky atmosphere in the radiative transfer, (2) the approximation of Earth's sphericity, and (3) the need of polarization in radiative transfer modeling. Finally, we address the effect of regularization which is inherent to the profile scaling approach. The different topics are supported by appropriate validation exercises.

G-3.) Furthermore, I believe that the amount of Figures is too large and can certainly be reduced substantially, if not by half then close to it. Relevant suggestions have been made in the text.

Changed: We agree with the reviewer that the number of figures can be reduced. We thus reduced the number of figures to 13 taking also the comments of other reviewers into account and removed the following figures from the original manuscript: Fig. 3, Fig. 5, Fig. 6, Fig. 8, Fig. 13, Fig. 16. However, most of the suggestions of the reviewer did not result in clear presentation of our results, to our opinion.

Specific comments:

S-1.) Page: 4918, Number 1 (Line 3): You mean GOME2/MetopA or /metopB? or both? this should be clear in the abstract. Also, which years are you referring to? which algorithm? this information should appear first.

Changed: on page 1, lines 2 – 3, we rephrased the abstract in the revised version of the manuscript and the first sentence now states that we use GOME-2/MetOP-A data in the period 2007 – 2010 to perform a sensitivity study of the direct fitting of total ozone columns..

S-2.) Page: 4918, Number 2 (Lines 4 – 8): This phrase is far too long and compact for the layman to follow. Split into two sentences at least.

Changed: On page 1, line 12 – page 2, line 24, we completely rephrased the sentence in the revised version of the manuscript.

S-3.) Page: 4918, Number 3 (Line 9): From this phrase I understand that you applied your technique to synthetic spectra? otherwise, how can you know the true column? again, this phrase is too compact for an abstract, re-phrase.

Changed: In the course of rephrasing the abstract, the sentence changed completely.

S-4.) Page: 4918, Number 4 (Line 25): A couple of general reading references might be a good idea here, as is common in introductory sections of papers.

Changed: On page 2, lines 28 – 29, we added the following general reading references:

- Fleming, E. L., Jackman, C. H., Stolarski, R. S., and Douglass, A. R.: A model study of the impact of source gas changes on the stratosphere for 1850–2100, Atmos. Chem. Phys., 11, 8515–8541, doi:10.5194/acp-11- 8515-2011, 2011.
- Fuhrer, J.: Ozone risk for crops and pastures in present and future climates, Naturwissenschaften, 96, 173–194, doi: 10.1007/s00114-008-0468-7, 2009.
- Guicherit, R. and Roemer, M.: Tropospheric ozone trends, Chemosphere – Global Change Science, 2, 167–183, 2000.
- World Health Organization (WHO) and others: Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide: report on a WHO working group, Bonn, Germany, 13-15 January 2003, Copenhagen: WHO Regional Office for Europe, 2003.
- World Meteorological Organization (WMO): Scientific assessment of ozone depletion, World Meteorological Organization, Global Ozone Research and Monitoring Project–Report No. 55, 416pp. Genova, Switzerland, 2014.

S-5.) Page: 4919, Number 1 (Line 1): Substitute with "in measuring".

Changed: On page 2, line 30, the redraft has been implemented.

S-6.) Page: 4919, Number 2 (Line 13): You also need a reference for the GOME2 instruments, as you referenced the rest of the satellites. Maybe Munro et al.,2006, will be enough.

Changed: on page 2, line 31, we implemented the following reference for the GOME-2 instrument: Munro, R., Eisinger, M., Anderson, C., Callies, J., Corpaccioli, E., Lang, R., Lefebvre, A., Livschitz, Y., and Albinana, A. P.: GOME-2 on MetOp, Proceedings of the EUMETSAT Satellite Conference, 12–16 June 2006, Helsinki, Finland, 2006.

S-7.) Page: 4920, Number 1 (Line 6): Either add a reference or explain where this degradation comes from, in one sentence.

Changed: On page 3, lines 58 – 59, we added the following references for GOME and SCIAMACHY instruments. For GOME:

- Snel, R.: In-orbit optical path degradation: GOME experience and SCIAMACHY prediction, ERS Envisat Symposium, SP-461, G¨oteborg, Sweden, 2000.
- Tanzi, C. P., Snel, R., Hasekamp, O., and Aben, I.: Degradation of UV earth albedo observations by GOME, ERS Envisat Symposium, SP-461, G¨oteborg, Sweden, 2000.
- van der A, R. J., van Oss, R. F., Piters, A. J. M., Fortuin, J. P. F., Meijer, Y. F., and Kelder, H. M.: Ozone profile retrieval from recalibrated Global Ozone Monitoring Experiment data, J. Geophys. Res., 107, 4239, 2002.

And for SCIAMACHY:

- Bramstedt, K., Noel, S., Bovensmann, H., Burrows, J. P., Lerot, C., Tilstra, L. G., Lichtenberg, G., Dehn, A., and Fehr, T.: SCIAMACHY monitoring factors: Observation and end-to-end correction of instrument performance degradation, Proceedings of the 2009 Atmospheric Science Conference, 2009.
- Noel, S., Bramstedt, K., Bovensmann, H., Burrows, J. P., Gottwald, M., and Krieg, E.: SCIAMACHY degradation monitoring results, Proceedings of the 2007 Envisat Symposium, 2007.

S-8.) Page: 4921, Number 1 (Line 21): I think the comma is not needed here.

Changed: On page 4, lines 112 – 113, the extra comma after both has been removed.

S-9.) Page: 4922, Number 1 (Line 3): Which one did you use? the one provided by EUMETSAT or did you correct/update/form that one?

Changed: On page 5, lines 134 – 135, we mention that we use a Gaussian instrument spectral response function with a full width at half maximum of 0.3nm in our study. This has not been mentioned earlier, since before we discussed the model setup more generally.

S-10.) Page: 4923, Number 1 (Line 16): Maybe a few more details on how extensive the LUT is, *how many nadir geometries for e.g., and SZA values?*

Changed: On page 5, lines 158 – 164, we provide a detailed description of the lookup table. The lookup table is generated for solar zenith angles between 10 and 80 degree in steps of 10 degrees and total ozone column between 280 and 400 DU in steps of 20 DU to address the dependence of the atmospheric Ring spectra on the solar zenith angle and to avoid interference between the filling-in of ozone absorption features due Raman scattering with the retrieved total ozone column.

S-11.) Page: 4923, Number 2 (Line 17): Also explain what the US standard atmosphere is, how detailed it is spatiotemporally

Changed: On page 5, lines 158, we added the following citation:
NOAA: US Standard Atmosphere, 1976, Technical Report NOAA-S/T76-1562, National Oceanic and Atmospheric Administration, US Gov. Printing Office, Washington, DC, 1976.
We think that the US standard atmosphere is a well-known reference and thus a reference should be sufficient here.

S-12.) Page: 4927, Number 1 (Line 11): Maybe spatiotemporal?

Not changed, because “temporospatial” is a synonym for “spatiotemporal”.

S-13.) Page: 4928, Number 1 (Line 2): Are you implying that the ozonesonde measurements are error-free? I am sure they are associated both with instrumetal errors as well as algorithm errors. Least of all, the error resulting from the choice of tropopause level in the ozonesonde calculations.

Changed: On page 7, line 237 to page 8, line 241, we added the sentence: “Borsdorff et al. (2014) discussed the meaning of e_n in terms of the profile scaling approach. Interpreting the effective column c_{eff} as an estimate of the true column, e_n represents the error made by the choice of the reference profile ρ_{ref} to be scaled in the inversion. Obviously, when the reference profile represents the correct relative vertical trace gas distribution, e_n vanishes.”
Also, we do not imply that ozonesonde measurement are free of errors, in fact that is why we apply a correction to the integrated amount of ozone in Sect. 5 of the revised manuscript, in which we address the effect of regularization. Addiationally, we are aware that errors due to ozonesonde profiles remain and impact the validation. Moreover, we also assume the ground-based measurement not be free of error (see the discussion on page 11, line 373 – page 12, line 386).

S-14.) Page: 4928, Number 3 (Line 4): I am rather weary that you base this important result of your article on one ozonesode profile, before you apply your method. Are you implying that you used all available ozonesode profiles in De Bilt for the years 2007-2010 and for all of them the error vanishes?

Changed: On page 9, lines 293 – 294, after discussing the inversion of the profile scaling approach, we illustrate the averaging kernel and the smoothing error using simulated measurements (for De Bilt) using three different reference ozone profiles. Furthermore, we performed additional simulations for a high-latitude polar ozone hole case and found that the null space error can mount up to 5%. This error is indeed avoided when either using a perfect a priori vertical profile or when interpreting the retrieved column as an effective column and subsequently applying the total column averaging kernel concept for validation. We think that this discussion is sufficient for the illustration purpose we are aiming for at this point in the manuscript.

Later in Sect. 5, we present a validation of the effective column approach using 647 collocated measurements (page 19, lines 630 – 631) and display the results in Fig. 13. Here, the lower panel clearly shows when ozonesondes are used as reference ozone profiles to be scaled by the inversion and in the validation, respectively, the null space error vanishes as expected from the theory.

S-15.) Page: 4929, Number 1 (Line 22): I assume you are accepting only forward viewing measurements of the descending node?

Changed: On page 10, lines 327 – 328, we state that we use only forward measurements of the descending node.

S-16.) Page: 4930, Number 1 (Line 5): How big a contribution?

Not Changed: That personal communication was merely about the kind of error contribution (additive of multiplicative) and this has been stated in the manuscript.

S-17.) Page: 4930, Number 2 (Line 21): ...and to be spatially co-aligned...

Changed: On page 18, line 625, the redraft has been implemented.

S-18.) Page: 4931, Number 1 (Line 3): And this number is?

Changed: On page 12, lines 421 – 422 and page 19, line 631, we give the numbers of collocations found in the observed period. The number of collocated GOME-2 retrievals with ozonesonde measurements is 647 for the validation of the effective column approach. The number is mentioned in the revised version of the manuscript. For the performance analysis and the direct validation we obtain 6861 validation measurements that fulfill the requirements defined in Eqs. (19) – (22) of the revised version of the manuscript.

S-19.) Page: 4931, Number 2 (Lines 5 – 6): You have to give the entire acronym of WOUDC and SHADOZ, as per norm.

Changed: On page 11, lines 365 – 366, the redraft has been implemented.

S-20.) Page: 4931, Number 4 (Line 25): Validation studies using Brewer, Dobson and ozonesonde simultaneously have also revealed differences between the ground measurements which reach higher values for "delicate" regions, such as the Poles, the tropics and so on. For e.g. Balis et al, 2007a;b, Antón et al., 2009, Loyola et al., 2011; Koukouli et al., 2012, Labow et al., 2013, Bak et al. 2015.

Changed: One page 11, line 381 – page 12, line 384, we added a reference to Bernhard et al. (2005) stating that seasonal and solar zenith angle dependencies can mount up to 4% due to the parameterization of the ozone layer, because here, we address the accuracy of the ground-based instruments, which is crucial to evaluate our validation. We do not discuss validation studies here. References to some of these studies have been given in the introduction (page 2, lines 48 – 57).

S-21.) Page: 4932, Number 1 (Line 8): It is not clear whether this section follows the cloud filtering section or whether it refers to new issues. If the latter, maybe you should start a new paragraph just in case.

Changed: On page 12, lines 390/391/399/407/409, for clarification, we specified the number of filters that we define and consequently apply. We use four filters and the text has been adjusted to clarify the point when a new filter is defined.

S-22.) Page: 4932, Number 2 (Lines 8 – 9): I do not follow the logic behind this. Why was this filter applied? I am sure this really reduces the amount of comparative datasets you have at your disposal. Please explain both here and in the text the reasoning.

Changed: On page 12, lines 399 – 400; 404 – 406, we introduce the filter δt to filter for strong ozone variation. We assume the scene to be ergodic, which means that a scene describes the same behavior averaged over space as it does averaged over time. This is an important assumption but does not need to hold for all stations, because ergodic statistics might not be given. However, for other stations we saw a clear improvement in the validation using this filter.

S-23.) Page: 4933, Number 1 (Line 13): I wonder what you happen if you assumed a more detailed and geographically appropriate climatology... please comment.

Not changed: A climatology superior in quality to the Fortuin & Kelder climatology but inferior in quality to ozonesonde measurements might improve the validation. However, this is not the goal of the paper. We want to stress the difference between the direct column approach and the effective column approach in the validation. We found that the accuracy of the effective column does not significantly depend on the choice of the a priori reference profile, whereas the accuracy of the true column estimate does, due to a smaller null space contribution for an improved quality of the reference profile. We therefore do not see the need to add an extra case for the reference profile which will provide an intermediate accuracy of the estimate of the true column between using the Fortuin & Kelder climatology or ozonesonde profile as a priori reference.

S-24.) Page: 4933, Number 2 (Lines 15 – 16): You hence assume that the total ozone column given by the Dobson is more correct that the integrated column given by the ozonesonde? previously you treated the ozonesonde profile as "truth' hence why the need for the scaling. Please explain both here and in the text.

Not changed: Yes, we assume the ground-based total ozone columns to be of better quality than the integrated ozonesonde profiles for reasons that are discussed in the original manuscript on page 4930, lines 14 – 20. We think the reviewer is referring to the paragraph page 4927, line 18 to page 4928 line 8 and Fig. 1. There indeed we assume the ozonesonde profile on 15 January 2009 to represent the truth. Since in that paragraph, we discuss the concept of the column averaging kernel more generally by means of synthetic measurements, where the scaling of the ozonesonde to a total ozone column of a ground-based measurement is not relevant.

S-25.) Page: 4933, Number 3 (Line 29): You mean 15 in total, 15 per annum, 15 per month?

Adapted: We defined a criterion that we used only stations in the validation for which there were a total of minimum 15 collocations with GOME-2 retrievals. However, that paragraph has been completely rephrased in the revised version of the manuscript to avoid any confusion.

S-26.) Page: 4934, Number 1 (Lines 10 – 11): Is there an actual physical reasoning for this, i.e. that something specific improves for the geophysical characteristics of ozone over those sites? I note that two of those sites are in the tropics. Please discuss.

Changed: On page 19, lines 663 – 664, we mention that a deficiency of the Fortuin & Kelder climatology might cause these effects.

S-27.) Page: 4935, Number 1 (Line 1): This is the difference between the GOME2 TOC and the ground-based TOC? of the same day, correct? are you in this way not ignoring the fact that the WOUDC TOCs are daily means whereas the GOME2 measures at a specific time? 30 D.U. difference is not unexplained for when comparing a daily mean which is derived from measurements throughout the daytime hours [but you cannot know which daytime hours those measurements have been taken] and the specific GOME2 overpass time.

Not changed. No. Here, we do not talk about the difference between the GOME-2 and the validation measurement. We rather mean differences between the consecutive days, which should not exceed 30 DU. See also the definition of the filter for scene heterogeneity and representation error. We like to refer the reviewer to point S-22.) and page 12, lines 399 – 406 for the description of the filter δt.

S-28.) Page: 4935, Number 2 (Line 4): Why wasn't this climatology used from the beginning of this study? practical reasons, I assume?

Adjusted: Section 4 is based on the Fortuin & Kelder climatology, which provides a reasonable estimate for the reference profile. In this section, we discuss forward model errors. For this purpose the Fortuin & Kelder climatology is sufficient. In Sect. 2.2, we consider different choices of the reference profile to illustrate the effect of regularization of the profile scaling approach for simulated measurements. In Sect. 5, we pick up the discussion from Sect. 2.2 and extend it with real measurements. We are convinced that this aspect becomes clearer in he revised version of the manuscript with the new structure of Sect. 4 and 5.

S-29.) Page: 4935, Number 4 (Line 7): Per annum? per the five year period?

Adapted: Here, we required a total of at least 30 collocations in the observed period. However, because of the restructuring of the manuscript and the reduction of figures in the revised version of the manuscript, this reference is not needed.

S-30.) Page: 4935, Number 5 (Line 13): Do these findings agree with other validation studies with other satellite products using the same ground-based stations?

Changed: On page 18, lines 602 – 603, we added a reference to Lerot et al. (2014) who present errors that are introduced by differences in the shape of the ozone profile < 1% for SZA < 80° and < 4% for SZA > 80°. This error is similar to the effective null space in our study and has not yet been addressed in validation studies. We base the value of 1% on the synthetic study from Sect. 2.2 but expect errors up to 5% for large solar angles in high-latitude ozone hole conditions.

S-31.) Page: 4935, Number 6 (Line 21): This is a very valid point. However, do you know off hand of recent/modern satellite algorithms that do not account for orography in their retrievals?

Changed: On page 13, lines 426 – 427, we added a reference to Fioletov et al.(2008). A common way is to exclude high altitude stations as e.g. Mauna Loa, Izana from the validation. This is done in the study by Fioletov et al.(2008). However, Balis et al.(2007a,b) do not explicitly exclude high altitude stations (only in case other filter criteria are not met). No correction is mentioned either.

S-32.i) Page: 4935, Number 7 (Line 24): First of all, I would start a new paragraph here, since this topic is truly quite different from what you showed before. Also, the entire analysis/discussion/algorithm presented before depends on there being no clouds in the GOME2 scene. Since from here on you also allow cloudy scenes, this should become a separate section.
S-32.ii) Furthermore, it is not entirely explained what you are showing here: you have analysed all GOME2 pixels assuming there are no clouds in the scene [and hence no cloud treatment in your algorithm]? or did you take the FRESCO algorithm into account during your retrievals for this section?
If the latter, then you should explain clearly what type of cloud algorithm this is, how it was applied, details, etc. As you are well aware these past five or ten years most of the algorithm community has been working on developing the appropriate cloud algorithm to treat cloudy satellite scene with various methods. If the former, then I truly do not understand the meaning of this section. If you assume no clouds and your scene has clouds, the resulting ozone will be erroneous of course. Please expand and explain.
S-32.iii) I also have serious concerns about this "correction factor" and how one justifies applying it, however, I reserve judgement for after I have read the explanations on this paragraph.

S-32.i) Changed: In the course of the restructuring, the paragraph dealing with the cloudiness parameter has been put into a separate section (Sect. 4.3.1).
S-32.ii) Changed: On page 14, lines 491 – 492, we rephrased the paragraph to clarify the procedure. We require the data set being nearly cloud free, i.e. $\eta_{cld} < 0.1$. In section 4.3.1 of the revised version of the manuscript, we investigate the influence of clouds on the validation. So Fig. 10 of the original version of the manuscript displays the retrieval error as function of the cloudiness parameter. A clear dependence of the retrieval error on the cloudiness is shown and from that we follow that we need strict filtering for clouds.
S-32.iii) Changed: On page 14, lines 493 – 495, we adjusted the text. To disentangle the retrieval error caused by clouds from other possible forward model error sources this correction has been obtained from nearly cloud-free scenes ($\eta_{cld} < 0.1$) for each validation station. In this way, it was possible to illustrate the retrieval error due to cloudiness. In the context of Fig. 6, it makes perfect sense to correct the retrieval for the other errors.

S-33.) Page: 4936, Number 1 (Line 3): I would also start a new sub-section here as well, since you are now talking about SZA, something new.

Changed: We introduced Sect. 4.3.2 in the course of restructuring Sect. 4 of the original manuscript. This section deals with dependence of the retrieval error on the scattering angle and the approximation of Earth's sphericity is discussed to mitigate this error.

S-34.) Page: 4936, Number 2 (Lines 9 – 11): This phrase is confusing. Which overall biases? between what and what? which dataset is corrected for its mean error? please expand.

Changed: On page 15, lines 509 – 512, we clarified that we correct each dataset for a bias caused by other errors than solar zenith angle dependent. This correction is obtained analogously to the correction for cloudiness in Set. 4.3.1. See also S-32.iii).

S-35.) Page: 4936, Number 3 (Lines 20 – 21): Very broad statement, needs more scientific backing up.

Changed: On page 15, lines 520 – page 16, line 523, we added the following sentences: “This may be caused by the combination of different measurement sites with different bias corrections, because the underlying dataset of this study consists of Brewer, Dobson, and SAOZ measurements with different solar zenith angle dependencies. Due to the sampling of different solar zenith angle ranges of the measurements at different stations, features can be introduced.” However, and attempt to use a more homogeneous dataset resulted in an unrepresentative sample for statistical analysis.

S-36.) Page: 4936, Number 4 (Line 22): Please make clear you mean the SZA of the satellite and not the ground since we do not have the SZA information of the daily mean groundbased TOCs.

Changed: On page 16, line 523, we clarified that we mean satellite solar zenith angles and satellite viewing angles.

S-37.) Page: 4936, Number 5 (Lines 24 – 25): Did you de-seasonalise the ground-based TOCs and the satellite TOCs before applying the linear regression? the Dobsons have a strong seasonality [for various reasons] which should be excluded before performing any such analysis. Ditto for their differences.

Not Changed. If the total ozone column time series were de-seasonalized, the solar zenith angle dependence of the measurements over the year will be lost for the particular stations.

S-38.) Page: 4937, Number 1 (Lines 3 – 4): On which scientific fact do you base this assessment?

Changed: On page 16, line 534, we added a reference to Bernhard et al. (2005).

S-39.i) Page: 4937, Number 2 (Line 5): The Dobsons in general as instruments have issues with high SZA, among others. In general, I have very strong reservations about this entire paragraph. The satellite SZA dependency on comparisons between satellite and ground is a well-documented fact in literature and it has already been thoroughly discussed "with stations comprising multiple instruments and sufficient data".
S-39.ii) There are numerous factors that affect the satellite-ground comparisons per SZA and one cannot simply extract a linear trend from differences.
S-39.iii) Also, you show on the same plot SAOZ instruments which is highly misleading. The SAOZ instruments only take measurements at very high SZAs and extreme conditions.
S-39.iv) The Brewer/Dobson daily mean TOCs that you show may have resulted from measurements during a day with different SZAs. I suggest entirely re-writing/re-thinking of this SZA section, including Fig. 13 and the numbers you show in the Table.

S-39.i) Changed: On page 16, line 534, we added a reference to Bernhard et al. (2005). See also S-38.)
S-39.ii) Not changed: We do not say that the dependence on the solar zenith angle is linear. We use linear regression to demonstrate a dependence on solar zenith angle, which might contain higher orders as well. The fact that by means of linear regression we can detect solar zenith angle dependencies of Dobson instruments, supports that this method is applicable. See also Table 2.
S-39.iii) Not changed: For that reason SAOZ measurements are used to include high large solar zenith angles in the analysis to show the limitations of the plane parallel assumption.
S-39.iv) Changed: On page 16, line 523, we discuss satellite solar zenith angles and not solar zenith angles of the ground-based instruments. The former are used in the forward model and this dependence in the validation is discussed in the performance analysis in Sect. 4.3.2.

S-40.) Page: 4937, Number 3 (Line 13): This section should definitely preceed all other validation sections.

Changed: Sect. 4.3.3 discusses the radiative transfer solver and we agree that this section should precede the validation sections.

S-41.) Page: 4937, Number 4 (Line 28): Is the SZA in the x-axis the satellite SZA or the ground SZA? if the latter, then I assume you are not using the daily mean WOUDC data for Lerwick. Please explain in detail which ground based data you are using.

Changed: On page 17, lines 558 – 559, we clarified that we mean satellite solar zenith angles and satellite viewing angles. For the validation of the data product we use daily mean values from the ground-based instruments.

S-42.) Page: 4938, Number 1 (Lines 1 – 2): These are errors between two different runs of your model? between satellite and ground? is it all simulations? please explain and place in a different section, maybe in the introductory/model discussion part.

Changed: On page 17, lines 559 – 561, we specify that the errors discussed here are retrieval errors for simulated measurements for satellite solar zenith angles and satellite viewing angles at one particular ground station caused by using scalar radiative transfer. The errors are shown in Fig. 10 and its figure caption holds more information on the underlying data used to generate the figure.
This figure is needed in this section to explain the error mitigation using a spectrally linear albedo.

S-43.) Page: 4938, Number 2 (Line 15): Which datapoints? which stations? which years/SZAs/etc?

Changed: The corresponding figure has been removed from the revised version of the manuscript in order to reduce the number of figures.

S-44.) Page: 4939, Number 1 (Line 16): References needed. Values also. How big a degradation is reported?how have other algorithms dealt with this issue?

Changed: On page 13, lines 433 – 435, the references mentioned in S-7.) are repeated here for GOME and SCIAMACHY. Additionally, in Sect. 3 in the original version of the manuscript, we give values for GOME-2 ranging from 0.6% to 2.2% from east to west pixels at 325nm (Cai et al., 2012) and report our findings of 3-4% (east pixels) to about 9.5% (west pixels) which are comparable to the findings of Tilstra et al. (2012).

S-45.) Page: 4940, Number 1 (Lines 18 – 19): Haven't other algorithms seen this? references and comments on such papers should be included.

Changed: On page 14, lines 468 – 470, references to Anton et al. (2009), Loyola et al. (2011), Koukouli et al. (2012), and Hao et al. (2014) have been included.

S-46.) Page: 4940, Number 2 (Lines 24 – 25): Above you give a value of 0.5%. What is it?

Changed: On page 14, lines 465 – 468, we removed 'Δ_{ret} ~ 0.5%' and kept Δ_{ret} = -1.3% and Δ_{ret} = -0.6% which are the differences between ground-based measurement and satellite retrieval without and with degradation correction applied, respectively.

Reviwer comments concerning Figures:

S-47.) Page: 4957, Number 1 (Figure 5): This Figures should be merged to FIgure 4, Naha, maybe using one colour for Hohenpeissenberg and one for Naha.

Not changed: We disagree with the reviewer. Merging Figs. 4 and 5 would result in an unclear figure. Hence, Fig. 5 has been removed from the revised version of the manuscript and Fig. 4 has become Fig. 12.

S-48.) Page: 4958, Number 1 (Figure 6): Figure not necessary.

Changed: We agree and consequently Fig. 6 has been removed from the revised version of the manuscript. The validation of the stations is summarized in Table 1 dataset 1.

S-49.) Page: 4959, Number 1 (Figure 7): Figure 7 and Figure 8 are hard to read and the information they contain can fit nicely into a table.

Not changed: We disagree on the reviewers opinion. Figure 7 nicely demonstrates the effect of direct and effective comparison, which cannot be visualized in similar way in a table. We leave Fig. 7 unchanged and it becomes Fig. 13 in the revised version of the manuscript.

S-50.) Page: 4960, Number 1 (Figure 8): See comment for Figure 7.

Changed: To reduce the number of figures, we removed Fig. 8 from the revised version of the manuscript and refer to Table 1 dataset 1 for the validation using the direct comparison method in the revised version of the manuscript.

S-51.) Page: 4962, Number 1 (Figure 10): This Figure may be substituted easily with a Table, or one single line plot, the histogram representation is not paramount for giving the message.

Not Changed: We disagree. The bar plot presents the message clearly and furthermore demonstrates that data has been aggregated into bins of $\eta = 0.05$. Moreover line plots suggest a linear dependence between data points which is not necessarily supported by the data. Figure 10 has become Fig. 6 in the revised version of the manuscript.

S-52.) Page: 4963, Number 1 (Figure 11): Both panels can be substituted with line plots which take a lot less space. Histogram representation is not necessary for giving the message.

Not changed: We do not understand why a line plot needs less space. The bar plot representation of the upper panel nicely puts the different approximations of Earth's sphericity into perspective regarding the solar zenith angle dependence of the retrieval error. Figure 11 has become Fig. 7 in the revised version of the manuscript.

S-53.) Page: 4964, Number 1 (Figure 12): These four panels may well be merged into two or even one using different colours per station, it is very common to do so.

Not changed: We disagree. Merging the four panels clearly deteriorates the clarity of the figure. To reduce space, we removed the panels displaying the data of Lerwick and Uccle from the figure. Figure 12 has become Fig. 8 in the revised version of the manuscript.

S-54.) Page: 4965, Number 1 (Figure 13): Not necessary, substitute with a table.

Changed: We agree and the data of the figure in presented in Table 2 in the revised version of the manuscript.

S-55.) Page: 4968, Number 1 (Figure 16): Merge into one panel.

Changed: In order to reduce the number of figures, we removed this figure from the revised version of the manuscript.

S-56.) Page: 4970, Number 1 (Figure 18): Merge into one line plot. See previous comments on Figures.

Not changed: We disagree. The bar representation of the top panel clearly put the effect of the degradation correction into perspective with retrievals from data that are not corrected for instrument degradation. Furthermore, this representation shows again that data has been aggregated into bins. Figure 18 becomes Fig. 4 in the revised version of the manuscript.

S-57.) Page: 4971, Number 1 (Figure 19): Substitute with a table.

Not changed: We think that the chosen representation shows the effect of scan angle dependent degradation much more clearly than a table. Figure 19 becomes Fig. 5 in the revised version of the manuscript.