J.C.A. van Peet
Royal Netherlands Meteorological Institute
PO-box 201
3730 AE, De Bilt
The Netherlands
peet@knmi.nl

To: Editor of Atmospheric Measurement Techniques
Subject: submission of manuscript AMT-2013-185 after full referee reports

Dear editor,

Please find enclosed our detailed answers to the comments made by the referees of our article "Ozone ProfilE Retrieval Algorithm for nadir-looking satellite instruments in the UV-VIS" (AMT-2013-185). We would like to thank the reviewers for their valuable time and useful comments. On the following pages, we give detailed answers to each of their comments. Please note that some of the changes are directly included in the attached supplement, and they will not be fully repeated here. Major changes are indicated in green, minor textual changes have not been highlighted.

With kind regards,

J. van Peet

Answers to the comments made by Anonymous Referee #2 in RC C3468: 'AMTD Reviewer Comment', 29 Nov 2013

General comments:
We are glad that the referee is satisfied with the changes after the quick access review.

Specific Comments:

I understand that any erroneous reflectances are not used as reference points in the filtering algorithm. If so, I would clarify this by saying "previous accepted wavelength" instead of "previous wavelength" on lines 6 and 8.
Text has been updated

I suggest to add a small paragraph on wavelength calibration. Are there any residual wavelength shifts between the radiance and solar irradiance to account for in the fit? If the level 1 of both instruments is good enough in this respect, then mention it.
It is assumed that the wavelength calibration is good enough, since both instruments have a daily wavelength calibration based on solar measurements. Section 3.2.7 (see the supplement) has been updated accordingly.

please specify that the bias is wavelength independent.
Text has been updated

The discussion refers to convergence and the number of iterations in many places. Therefore, I suggest to add here the formula for the next step of the iteration to give the reader better understanding on the contributions of the weighting function and covariance matrices in the iteration process.
The formula for the next iteration step and references to the relevant equations in Rodgers (2000) have been added to section 3.2.8 (see the supplement).

Equation 5 is just a repeat of equation 1 with $x_t = x$. Please remove eq. 5 and refer to eq.1 here.
Equation 5 has been removed and the text has been updated.

Too generic statement about all UV-VIS retrievals. At the end of section 3.2.3, you mention that only a cloud-free retrieval is done if snow or ice is detected. Therefore you should be more specific here, saying something like: "Since the current algorithm only does a cloud-free retrieval over snow and ice, this ..."
Text has been updated to explain the working of the algorithm over snow and ice.

Is it ambiguous what the “relative means of the differences” means. I believe you take the relative difference of each (satellite,sonde) pair first, and then the latitude band mean of these relative differences. I would call this “means of the relative differences”. Please clarify what you mean and indicate it (i.e. where the average is taken) in the formula given in the x-axis label.

Text now reads “Mean of the relative differences”, and the axis labels now include angular brackets <…> to indicate where the average is taken.

… I suggest to add plots similar to figs. 3 and 6 but for the errors, plotting for example 100*(σsat − σpri)/σpri to see how the prior error (σpri) at different pressure is reduced by the retrieval (σsat) on average for the different latitude bands and for both instruments. Please add a short discussion on the reduction of the uncertainty for both instruments.

Plots with new numbers 4 and 10 have been added to the manuscript (see supplement) and a short discussion has been added to section 4.2 (for GOME-1) and to section 4.3 (for GOME-2)

“Clouds are modelled” should refer to the old version as “Clouds were modelled”

Text has been updated

“That version does an performs well”, should read as “That version performs well”

Text has been updated

Answers to the comments made by L. Flynn in RC C3889: 'Review with revisions', 06 Jan 2014

The abstract should provide the specific performance as compared to the limits that are given in the CCI
The CCI limits have been added to the abstract.

Specific comments:

_>_>_>_ A set of averaging kernel figures should be provided for both layer resolution retrievals, preferably in the paper but certainly in supplementary material for some standard retrievals. _<_<_<_
Plots 5 (averaging kernels for GOME-1) and 11 (averaging kernels for GOME-2) have been added to the manuscript (see supplement) and a short discussion regarding the averaging kernels has been added to the text.

Figure 1 could easily accommodate DFS values for GOME-2.
Figure 1 has been updated with DFS values for GOME-2 and DFS values for a GOME-2 retrieval without additive offset.

_>_>_>_ It would be useful to have a clear mathematical description of the "additive offset" term (S 3.7) and the temporal and spatial variation in this retrieval quantity. _<_<_<_ How does this term affect the averaging kernels?
A mathematical description of the additive offset has been added to section 3.2.7. (equation 5). The temporal and spatial variation of the additive offset for GOME-2 has been illustrated by plots 12, 13 and 14. A discussion on the results shown in these plots can be found in section 4.3 (see supplement).

Additional information on the failure of the RTM to model what is primarily a single scattering phenomenon below 290 nm should be provided. (Are there other studies to cite for this behavior?)
The original sentence ("In addition to the degradation of the detector…") is confusing and has been changed to "As a result of the degradation of the detector, the modeled radiance…". We believe that the observed difference in modeled and measured radiance is not due to the RTM, but that it is only a consequence of the degradation of the detector.

What are the retrieval algorithm behaviors for a theoretical ozone profile variation such as is expected during the last 20 years? That is, how large are the expected measurements changes, what part of these changes would be incorrectly fit with the additive term, and with what fidelity will this algorithm retrieve these profile changes both for GOME-1 and GOME-2.
Based on two studies by Kyrola et al. (2013) and Gebhardt et al. (2014) (references added to the manuscript), a short discussion has been added at the end of section 4.3 (see supplement).

What are the retrieval algorithm behaviors for measurement errors of the sizes and relations that are expected given the calibration uncertainty of the two instruments? That is, what retrieved profile uncertainties would one expect as computed by positing a measurement uncertainty pattern and multiplying by Dy? Since the GOME-2 makes daily solar measurements, why doesn't the instrument degradation cancel in the radiance/irradiance ratios? Are the instrument characterization deficiencies correlated in wavelength, e.g., imprecision in the etalon corrections?

The lightpaths for Earth and Solar radiation differ in GOME-2 and therefore the instrument degradation does not cancel in the radiance/irradiance ratios. This is reflected in the additive offset term in equation 5, which is only added to the earth radiance, and not the solar irradiance. The sensitivity of the retrieved profile on the additive offset is shown in figure 14.

While there is a discussion of the effects of different FOVs and the measurement noises of the GOME-1 and GOME-2, I did not see specifics on the assumed measurement noise (Table 1 would be a good location), and number of wavelengths (all of them in the window?). How dependent are the DFS on the measurement noise (Se) and number of channels? With regard to the measurement noise, there is a danger in optimal estimation in assuming that the measurement errors are uncorrelated. If this is not the case, then the DFS calculations will be biased higher they actually are. This could be explored with the approach in the previous paragraph where a pattern of measurement biases is proscribed (with sizes related to the instrument radiance/irradiance calibration uncertainty) with a persistent relationship across wavelength intervals.

Values for the measurement noise for some selected wavelengths has been added to sections 4.2 (GOME-1) and 4.2 (GOME-2). The effect of correlated measurement noise has been investigated by setting the sub-diagonal elements of the covariance matrix to 0.01 and 0.10 of the main diagonal. This lowered the DFS by 0.3 and 3.0 percent respectively. The text in section 4.3 has been updated accordingly.

Is the increased coverage at the edges of the SAA worth the difficulties and potential biases of the accepted data? Are the "successful" retrievals identified/flagged as coming from the region where we expect this signal contamination?

Using the filter adds successful retrievals in a region where otherwise no successful retrievals would be done. No special flags are raised to indicate if the retrieval comes from the SAA region. See section 3.2.6.

_>_>_>_ What portions of the ozone profile variability (relative to the a priori) observed by the sondes (both before and after application of the averaging kernel) is captured by the retrievals? _<_<_<_

Figure 9 has been added (see supplement) to the manuscript and is based on GOME-2 data for the Northern Hemisphere. Figure 9a gives the mean of the relative differences and gives an indication of what part of the bias comes from the a priori. Figure 9b gives the RMS of the absolute differences as an indication of the "true" variation that is captured by the retrieval.

From the preliminary review:

The critical requirements for producing consistent time series include stable algorithms and stable measurements. The degradation correction and additive offset would seem to be key items in this case. What sensitivity experiments have been carried out? For example, one could construct forward model data for an expected trend in the ozone profile (before and after) and see how the retrievals from the algorithm change if it is allowed to include the additive offset term in one or both retrievals.
These issues are addressed in figures 1, 5, 11 and 14 and corresponding discussions in the manuscript.

What information is coming from the A Priori and what is coming from the measurements? Figure 3: Are the results for averaging kernel processed sonde data or simple layer sums? How would they differ? Figures 3, 6, 7, & 8: How do the A Pioris for the retrievals compare to the "Truth" data in the mean differences?
These issue are addressed in figures 4, 9 and 10 and corresponding discussions in the manuscript.