Review 2.

We thank the referee for the valuable comments that helped us to improve the manuscript. To address some of the referees' comments we decided to attach a Supplement, where some additional results and discussion are presented.

Please, find our responses to the referee's comments below.

General remarks:

Referee: - Some descriptions in the text are vague, and it is not clear if they refer to a specific figure or to something that is not shown in a figure. Some of these unclear references/descriptions are pointed out below in the "Specific comments" section. But I would also recommend that you check the rest of the manuscript to clear up other possible short or vague descriptions.
The SBUV ozone data set is one of the longest continuous records available, and is therefore used often for trend analyses. However, the vertical resolution of the measurements is coarse, and the lower levels have high uncertainties connected to the values. With combining several different layers, as suggested in the manuscript, it is possible to reduce the uncertainty on the ozone values, but this also clearly decreases the number of layers in the upper troposphere and stratosphere significantly. I am not sure if I understand correctly what the main message of this manuscript is: It is possible to decrease the uncertainty (and get the ozone variability right) on the SBUV layer data by combining the layers to be able to describe trends with less uncertainty? Or: It is possible to decrease the uncertainty on the SBUV layer data by combining the layers, but now layers are too broad to be of use to distinguish the differences in trends in specific atmospheric regions? I think it would be helpful if you would state more clearly what you think the data could be used for (maybe add some examples) after combining the layers and therefore reducing the smoothing errors.

Authors: The SBUV data are reported at 21 layers, though the actual vertical resolution is coarser. In this manuscript we analyzed the SBUV retrieval characteristics (Averaging Kernels) and smoothing error and ultimately recommended merging several lauers together to reduce the smoothing error and best represent the actual vertical resolution of the SBUV measuring system. These layer combinations are recommended for trend analysis and for comparisons with other instruments and/or model output. The SBUV data record is the longest consistent satellite ozone data record available for long-term time series analysis of total ozone, of profile ozone in the middle stratosphere, and of the broad tropospheric/lower stratospheric layer and as such represents an invaluable resource to the ozone community. However, the SBUV data cannot be used to resolve trends in the lower stratosphere, or as demonstrated in this paper, the QBO, for example. These limitations are widely understood in a general sense, but for the first time in this analysis a thorough, quantitative study of the errors has been completed to help guide the user.

Referee: p.6, lines 14-26: The Aura MLS climatological profiles calculated from measurement from January 2005 to December 2010, are merged with ozone sonde climatologies that are calculated from January 2000 to December 2005. So far there is no discussion about the possible problems for the total covariance calculation due to the different covered time periods. Do you expect that to have an impact? If no, why not?

Authors: Thank you for the comment. Aura MLS and sonde data are obtained over the two different 6-year time periods and are used to construct the covariance matrices **C**. The diagonal elements of **C** represent the squared standard deviation of the ozone variability, and we don't anticipate the variability to be significantly different from one 6-year period to another. In turn, the off-diagonal elements of **C** represent inter-level correlations and might be affected by the time mismatch. However, the sensitivity run demonstrated that for the SBUV retrievals the calculated smoothing errors are almost insensitive to the off diagonal elements of **C**. We included this discussion in the text, section 2.4, last paragraph: " Off-diagonal elements of **C** reveal correlations among the layers (see Fig. S2 in Supplement). If the correlation between any two layers is high, the corresponding off-diagonal elements will also be large, and vice versa. We do not analyze off-diagonal elements of **C** here, because a simple sensitivity test in which the smoothing error was calculated with the off-diagonal elements of covariance matrices set to zero showed that in our case the off-diagonal elements had a very small effect on the smoothing error values. For this reason, the mismatch of the time periods between MLS (2005-2010) and sonde measurements (2000-2005) used to construct the covariance matrices **C** should not affect the smoothing error calculations, because only the off-digonal elements of **C**, that indicate the inter-level correlation, are sensitive to the time mismatch. We expect the ozone variances (presented by the diagonal elements of **C**) over the two considered 6-year time periods to be very similar. "

Specific comments:

Referee: p.4, line 17/18: Why is the top layer not included?

Authors: The SBUV measurement system has almost no sensitivity to the ozone changes above 0.1 hPa (the top layer). However, the forward model used in the retrieval algorithm is required to specify ozone profiles from the ground to the top of the atmosphere. We also report the ozone amounts for this layer, but we don't report any correlative information for this layer, such as Averaging Kernels, smoothing error, Jacobians etc. We described this in new section 2.3.1 "SBUV version 8.6 algorithm".

Referee: p.4, line 19/20/21: What does the sign "/" stand for?

Authors: This is the letter '*l*' and refers to the index of the layer. We specified that: " A column of the **A** matrix at a given layer *l* (where *l* is a layer index from 1 to 20) gives the response of ..."

Referee: p.4, line 22: "AK" is already defined as "SBUV averaging kernel matrix". It would be better to use different abbreviations for the "averaging kernel matrix" and the "averaging kernels"

Authors: Thank you for the note. We corrected that. Throughout the text we follow the terminology described by Rodgers (2000) and use a term "**A** matrix" whenever we are talking about the entire matrix, and we use term "Averaging Kernels (AK)" to refer to the rows of the **A** matrix.

Referee: p.4, line 30: the lower boundary for the sharp SBUV AK is given as 25 hPa, but later (p.5 line 2) it is mentioned as 10 hPa. Are these the boundaries for different things? If yes, please explain this in more detail, if no, then please be consistent in mentioning the hPa value for the lower boundary.

Authors: Thank you for the comment. The lower boundary is 16 hPa outside of the tropics and 10 hPa in the tropics. We now keep it consistent throughout the manuscript.

Referee: p.5, line 2: "shifted upward (downward)," with respect to what?

Authors: We added: "... which are shifted upward (downward) from the nominal altitude, showing..."

Referee: p.5, line 11/12: I am not sure what is meant here by "fractional ozone changes".

Authors: Thank you for the note. We corrected: "... is applicable to the ozone profiles expressed as a fraction from the a priori ".

Referee: p.5, line 16: "i and j are layer indexes." Are these indexes of neighboring layers, or just any layers?

Authors: We added: "... and *i* and *j* are layer indices from 1 to 20".

Referee: p.5, line 27: How is the total DFS calculated?

Authors: The total DFS is equal to the sum of all diagonal elements of **A** matrix. We explained that in the text, sec. 2.3 par 5: " The sum of all diagonal elements of the **A** matrix - the total DFS -varies from 3.7 to 6.9 out of the 6-9 wavelengths used ...."

Referee: p.5, line 29: "higher SZAs and resulting increased vertical resolution in the upper layers." It is not clear why the higher SZA should result in an increased vertical resolution. It is explained on p.6, lines 4-7, so maybe just a rearrangement of the text would be helpful to clarify this.

Authors: We re-arrange the text in this paragraph to make it more clear.

Referee: p.5, line 30/31: The sentence "The layer DFS " is somehow vague. Could you be more specific about which upper layers you mean? In Figure 3 the uppermost layers (above 1 hPa) show smaller DFS than the layers directly below (between 10 hPa and 1 hPa)

Authors: We added: " The layer DFS also increases in the upper layers above 1 hPa when .."

Referee: p.6, line 3: Why does the DFS of the tropical profile below 40 hPa decrease abruptly?

Authors: We added: " ...due to fewer wavelengths used to retrieve ozone at small SZAs and the higher altitude of the ozone peak in the tropics.. "

Referee: p.6, line 13/21: Could you be more specific about the definition of the ozone sonde latitude bins?

Authors: Sonde data were distributed in 10-degree latitude bins and then interpolated to the 5-degree scale. We added: " We distribute sonde data by 10° latitude bins following the guidance provided by McPeters and Labow (2012) to account for limited sampling in some latitude bins. The 10° zonal sonde data for each month and altitude level have been interpolated to the 5° latitude scale. "

Referee: p.6, line 25: What does it mean if you convert "MLS and sonde profiles preliminary"?

Authors: We re-phrased: "Both MLS and sonde profiles were converted into ozone partial columns at SBUV pressure layers."

Referee: p.6, line 25/26: Did you have to use any auxiliary data to convert the MLS and sonde data to partial columns? If yes, could you please mention them?

Authors: Both MLS and sonde data are reported as mixing ratios, and we didn't use any auxiliary data (e.g. temperature) to convert them into the partial ozone columns.

Referee: p.6, line 27-29: Could you explain in more detail how the merging was done exactly? It is not clear to me what "proportional 75/50/25% weighting for lower/mid/upper parts of the range." means.

Authors: We re-phrased: " For each month and latitude bin we merge sonde and Aura MLS mzm profiles in layers 5, 6 and 7 (between 160 and 40 hPa) using a proportional 75%/50%/25% weighting for the sonde data in layers 5/6/7. "

Referee: p.7, line 6: What do you mean with a priori here? Is this the a priori profile used in the SBUV retrieval? And does that change with different latitude? Please be more specific in your description.

Authors: We added: " Figure 4 shows the square roots of the diagonal elements of the covariance matrices for three different latitude bins as a percent of the mean annual SBUV a priori at each latitude.".

Referee: p.7, line 6/7: the range given in the sentence "Standard deviations vary between 2-15

Authors: We changed: "Standard deviations vary between 2% and 15%, ..."

Referee: p.7, line 8/9: There is a reference to a high latitude standard deviation profile, but no high latitude profile is shown in Figure 4.

Authors: We added some plots in the Supplement that show standard deviations for all latitude bins.

Referee: p.7, line 22/23: Where does the total column ozone come from that is mentioned here? Is this just the integration over the different SBUV layers? Be more specific.

Authors: We added new section 2.3.1 that briefly describes the main features of the retrieval algorithm. In this section we added: " The total ozone columns are calculated as sums of ozone columns at all 21 layers."

Referee: p.7, line 30: "(as large as 15-20%)" These numbers cannot be seen in Figure 5. Do you refer to a different figure here?

Referee: p.7, line 30/31: it is not clear to which profile (latitude band and/or season) the sentence "Errors also increase up to 5%..." refers. Please be more specific.

Authors: Figure 5 shows the errors for the NOAA 17 instrument for two months in 2007 to demonstrate typical error profiles. In the text we provide a range of errors that we found analyzing all SBUV instruments. We changed the Figure 5 caption and specified the instrument, year and month: " **Fig. 5.** Typical profiles of SBUV smoothing error (% from the mean a priori profile) for the northern mid-latitudes in summer (green line), in winter (blue line) and for the tropics (red line). Profiles for NOAA 17, January and July 2007."

We also added several more figures in the Supplement that show error profiles for some other latitude bins.

Referee: p.8, line 2/3: It was not totally clear to me at this stage in the manuscript that the TC matrices are only functions of latitude. Does that the sonde data was averaged over the whole period (Jan 2000 to Dec 2005) within the different latitude bins? And the same was done for MLS? Could you explain this with a little more detail somewhere?

Authors: The covariance matrices **C** was calculated from a 6-year time series of the mzm merged data (MLS+sonde). We wrote in Sec. 2.4 par. 4: " The covariance matrices **C** for each 5-degree latitude bin have been calculated by employing Eq. (2), and are included in the SBUV mzm data files. The resulting covariance matrix **C** for each latitude bin is a matrix with dimensions of number of layers by number of layers (20 by 20; top layer not included), with the diagonal elements equal to the squares of the standard deviations of mzm merged MLS/sonde profiles."

Referee: p.8, line 4: "correspondence" might not be the best possible word choice here (made me stumble over the meaning of the sentence).

Authors: We replaced "correspondence" with "consistency".

Referee: p.8, line 10/11: Where does the QBO signal come from? The averaging of MLS and sonde data should average the QBO data out, so the signal should not be present in the TC. Does the signal come from the AK?

Authors: The diagonal elements of the covariance matrix **C** represent the square roots of standard deviations of ozone variability for this layer relative to the mean values. These deviations were calculated using a 6-year time series of mzm merged data (sonde + MLS) as described in section 2.4. Both sonde and MLS time series contain the QBO signal, which is responsible for major portion of inter-annual ozone variability in the tropical lower stratosphere. As a result the standard deviations (diagonal elements of **C**) increase.

Referee: p.8, line 17/18: Could you explain (again?) why there is an increased sensitivity in the upper layers?

Authors: We use a different number of wavelengths to retrieve ozone depending on SZA. We added explanation in the text: ", since the total DFS increase with increasing SZA due to the larger number of wavelengths used to retrieve ozone for high SZA, implying that we have ... "

Referee: p.9, line 3: What does Xa stand for? Still the a-priori profile?

Authors: Yes, $\boldsymbol{x_a}$ corresponds to the SBUV a priori profile. We use the same terms throughout the paper.

Referee: p.9, line 25/26: Could you explain in a little more detail why the physical interpretation would be a challenge?

Authors: We decided to remove this sentence from the text. We believe that we provided the detailed explanation of the practical difficulties of applying Averaging Kernels in this paragraph.

Referee: p.9, line 27-30: How exactly are the different SBUV layers combined? And does the method of combining them affect the resulting reduction in smoothing errors?

Authors: Since the SBUV data reported as partial ozone columns at 21 pressure layers it makes it easier to calculate ozone amount for the merged layer just simply adding ozone amounts for the parent layers. To achieve a reduction of the smoothing error the right layer combinations should be chosen. We advise using the recommended layer combinations in the manuscript.

Referee: p.10, line 3: You state here that the "AK are significantly reduced below 250 hPa", but on page 5 you state that the AK are significantly reduced below 60 hPa (in the tropics). Is there a threshold you apply to decide where the AK is significantly reduced? And why didn't you test combining the layers from 60 hPa to 16 hPa in the tropics (since the AK is reduced below 60 hPa in the tropics already)?

Authors: Thank you for the comment. We changed the text to avoid any confusion: " For comparison with the limb instruments (e.g. Aura MLS), that do not measure ozone below 250 hPa, the layer combination from 250 hPa to 25 hPa (or 16 hPa) can be used outside of the tropics (in the tropics). Even though SBUV measures ozone down to the ground, the signal significantly reduces for layers below 250 hPa and the amplitude of the AK below 250 hPa is negligibly small (see Fig. 2). Thus, we also calculate the smoothing error for the layer combinations from 250 hPa to 25 hPa (or 16 hPa) outside of the tropics (in the tropics)".

Referee: p.10, line 14: "The smoothing errors are larger ", larger than what?

Authors: We added: " The smoothing errors are larger in the tropics compared to the mid- and high- latitudes. At high latitudes errors are larger in winter and smaller in summer as the SZA changes. "

Referee: p.11, line 1-15: I assume the MLS data was also combined to make a comparison between the combined SBUV layers possible. Could you make that clearer in the text? As well as the method of combining the MLS layers?

Authors: We clarified that in the caption to Fig. 8: " The 254-16 hPa ozone columns (for both SBUV and MLS) were calculated by simply summing partial ozone columns in 6 individual layers (layers 4-9)." We also added some explanation in the caption to Fig. 1:" The MLS mixing ratio profiles were converted first to the partial ozone columns at the SBUV layers."

Referee: p.11, line 29: You mention two parameters on which the smoothing error depends on. However, on page 2 line 29-31 you mention three things on which the smoothing error depends on: the vertical resolution of the observing system, the accuracy of the a priori data, and the magnitude of the natural ozone variability. Is it two or three parameters? And if it is three, then please discuss the third parameter somewhere in the manuscript as well.

Authors: The smoothing error depends on two parameters: Averaging Kernels and Covariance matrix, that we fully described in section 2. At the same time Averaging Kernels contain information from the measurements and from the retrieval. As a result AK represents a measure of the vertical resolution of the measuring system and the accuracy of the a priori data.

Referee: p.15: Add somewhere the time period for which the ozone profiles are used.

Authors: We added in the caption to table 1: "... calculations over the time period from 1 January 2000 to 31 December 2005."

Referee: p.16: Why are there no columns for the latitude bands 90S-80S and 80N-90N?

Authors: We added in caption to table 2: "The orbit of the SBUV satellites and the geometry of the measurement does not allow for sampling atmospheric ozone poleward of 82 degrees."

Referee: p.19, line 2: Do the mid-latitude profiles represent a specific hemisphere? If yes, please specify in the figure caption.

Authors: We changed the caption to fig. 3: "**Fig. 3.** Typical profile of layer DFS for the northern mid-latitudes in summer (green line), in winter (blue line) and for the tropics (red line). Profiles from NOAA 17, January and July 2007."

Referee: p.20: What do the vertical profiles of the square roots of the diagonal elements of the TC for latitude bins in the SH look like? Does the fact that there are clearly less ozone sonde measurements available in the SH have an effect on the profiles? It would also be helpful to see a high-latitude profile here.

Authors: Fig. S1 in Supplement shows vertical profiles for all latitude bins in NH and SH.

Referee: p.21, line 2: Do the mid-latitude profiles represent a specific hemisphere? If yes, please specify in the figure caption.

Authors: We added in the caption of Fig. 5: "Typical profiles of SBUV smoothing error (% from the mean a priori profile) for the northern mid-latitudes in summer (green line), in winter (blue line) and for the tropics (red line). Profiles from NOAA 17, January and July 2007.'

Referee: p.22: Why are the smoothing errors so small for Nimbus 4 and Nimbus 7 (both figures)? And could you add “a)” and “b)” to the respective figures?

Authors: These two instrument were in stable near-noon orbits. Thus the SZA for these instruments were changing only with season, and the number of wavelength used for the retrieval didn't change from year to year.

Referee: p.23: Would it be possible to have the same y-axis range for all four plots? That would show the differences of the smoothing errors for the different layer combination clearer

Authors: We made the range of y-axis different to demonstrate the latitudinal structures. When the Y-axis range was set the same for all plots the latitudinal structures are not well seen for figures b and d. However, we made the same range for (a) and (c) and for (b) and (d).