Answers to Reviewer 2

We want to thank Reviewer 2 for his elaborate review of the paper. This criticism has considerably improved the paper. We apologize for many small errors and missing explanations. Building the real content of the paper has been quite a demanding undertaking and energy for polishing the text has been in short supply.

We have made an effort to take into account all the comments of the reviewers. All larger changes in the manuscript are shown in **bold**.

Changes by the authors

Independently from corrections suggested by the reviewers, authors have corrected typos and mistakes in the manuscript. The following corrections need to be mentioned:

1. The numbers of GOMOS and SAGE II measurements in text and in Figs. 1-2 are now consistent.

2. Figure reordering: 5 (old)->7(new), 6->8, 7->5, 8->6, 7->9. Note that in reviews and answers below old figure numbers are used.

3. Figure 5 in the discussion paper was erroneously based on pre-release of GOMOS version 6 data. The new figure (Fig. 7 after reordering the figures) is now based on the same data release as other results in this paper.

4. We have added references to a companion paper "Analyzing time varying trends in stratospheric ozone time series using state space approach" by M. Laine et al. that was submitted (11.7. 2013) to the same special issue as this paper. This new paper makes use a more flexible time series modeling than the present paper.

5. In order to handle autocorrelations and especially their impact on error estimates, we have applied the standard Cochrane-Orcutt transformation. This transformation causes only some minor changes to the final trend results, but shrinks somewhat the confidence regions. Text in the appropriate places has been modified. The cause of autocorrelations, modeling imperfections, are discussed in the companion paper mentioned above.

Below we have answers to general and specific comments from Reviewer 2.

Response to general comments (numbered by the authors)

1. One major question was the current state of GOMOS. In the description of GOMOS it is described in present tense. However, my understanding is that ESA lost contact with the ENVISAT satellite on 8 April 2012 and the mission was ended. Hence, should the description of GOMOS not be written in past tense as the mission is technically over? If so, I strongly suggest the authors correct for this.

Answer: The present tense was used because at the time of writing the manuscript we

thought optimistically that GOMOS and Envisat could still emerge from the darkness. Now we are sure that Envisat and its instruments are lost for science even if they are still up in then sky. We have changed the tense where appropriate.

2. There seems to be some inconsistency when describing the data sets. The vertical resolution is described for GOMOS, but not for SAGE II. Moreover, it is also stated that the data quality can be found at Kyrölä et al. (2010b) and Tamminen et al. (2010) for GOMOS, but there is nothing about data quality for SAGE II. There should be the same amount of detail for each instrument. The paper can be improved by just adding a few sentences about these subjects, whilst also maintaining a balance of the information between each instrument.

Answer: We have stated SAGE II vertical resolution in Table 1. We have added a fresh reference (Damadeo et al., 2013) to the new SAGE II version we have used in the paper.

3. Regarding Equation 2. Here, GOMOS is used as the denominator in the equation. While I understand that this is needed in the context of moving a time series so that it fits relative to another time series, it's not clear in your method yet that this is so. Here, you are simply comparing measurements. When doing so, it is common practice to use a combined average as the denominator using both data sets/measurements instead of using one as the standard, simply because it is not sure which one is best/correct. Is there a particular reason for choosing GOMOS as the standard? If not, then I suggest to use the combined average. Also, some explanation of what Fg and Fs is needed, although I'm assuming they are GOMOS and SAGE respectively?

Answer: There are two SAGE II populations (sunrise and sunset), so it is simpler to stick to GOMOS only. The differences between the profiles are small enough so that we are not doing any major deviation here. Furthermore, the ratio Δ written in Eq. (2) can be easily transformed into that normalized on combined average, $\Delta_{comb.av.}$. The relationship between these two parameters is $\Delta_{comb.av.} = \frac{\Delta}{1+\Delta/2}$, i.e., in case of GOMOS and SAGE-II, they are very close to each other.

Fg and Fs are now explained in text.

4. You clearly state that the comparison is made over 60S to 60N. However, there is no information about possible effects from coincidences occurring in different air masses, such as the polar vortex. Did you take this into account? Although 60N and 60S are "typically" too far south and north, respectively, of the vortex's location, it is not impossible that it would be situated around these locations. Would this impact your results? Some more clarification needs to be made in the text here about this possible problem.

Answer: We have not taken this into account. Polar air incursions to lower latitudes are natural events. To see if these events have any impact to our SAGE II-GOMOS comparison we imposed the limits -38<MPV (460K)<38. The number of collocations was changed from 251 to 230 i.e. about 10% in the latitude belt 40N-60N. But no change

in the bias figures. Because we have MPV calculations only for the GOMOS period 2002-2012 but not for 1984-2001, we do not impose any restriction to MPV-values.

5.There is no information about the common vertical grid that you use to make the comparison. What kind of grid do you use? Do you convolve profiles based on the averaging kernels?

Answer: The grid is 20-60 km with 1 km step. No averaging kernels as the vertical resolutions of these instruments are close to each other. This has now been mentioned after Eq. (2) and in the beginning of Sec. 6.

6.2 Removal of bias

6. I think Figures 8 and 9 are very good and valuable to the paper. I would however, like to see an extra subplot that presents the densities in relative units, this would give the reader a better overview of the asymmetries.

Answer: The upper subplot is now replaced by a plot where densities are scaled by the median density over the combined tine series.

6.3 The combined Data set

7. I suggest the use of a Median Absolute Deviation instead of mixing a median with the standard deviation when trying to identify outliers. MAD is far more robust and will provide a better overall representation of your data, without compromising the amount of data needed for a statistical analysis (especially when dealing with the number of coincidences you have).

Answer. The text was inaccurate. We are using (in Matlab)

```
find(abs(dd1-median(dd1)) < outlierlimit*1.4826*median(abs(dd1-
median(dd1)))=
Outlierlimit*1.4826* MAD.
Here outlierlimit=3 and the factor 1.4826 ensures agreement with
the standard outlier definition in the case of normal
distribution. We have clarified this in the text.
```

8. L25-29: I find this a little difficult to grasp. Are you making a weighted mean from the medians, based on the individual measurement uncertainties of the input profiles that make the median values? Hence, the larger the measurement uncertainty for a given profile, less weight is given to that profile in the overall median? You also state “medians of the individual instrument data sets”, but what does this mean exactly? The medians for a given lat-time cell? This paragraph needs a lot more explanation, maybe even an equation could alleviate some of the doubt.

Answer:

Our procedure is: In every latitude-month cell we calculate the median of GOMOS and SAGE measurements (separately). The errors of the two medians are calculated by the median error formula (see e.g., Kyrölä et al., 2010). This includes calculation using interquantile distance, which is a generalization of MAD. After this we combine these

two medians using weighting by the errors just calculated.

9. Sec. 7 Time series analysis of SAGE II and GOMOS combined data set

You use the term t in your analysis twice, but for different things, which makes it a little confusing (for time and also for t values). Are “t values” from a student t test? I think you need to separate this somehow and describe a little better that you are using a statistical technique to determine the inflexion point. Perhaps, time should be renamed to T?

Answer: Our t-value is simply a parameter value divided by its estimated error. This definition has been added to text. In our case we have a large number of degrees of freedom and, therefore, the Student distribution can be very closely approximated by the normal distribution.

10. Another important point is your choice of altitude/latitude analysis. Why are you looking between 30 and 50 km and in the mid-lats? I think the main point you are trying to get across here is that this is the region where we expect to see an ozone recovery first and our motivation is such that by looking at these regions we hope to be able to find the largest impact in a change in trend. Moreover, this point should probably be made earlier in the paper as a motivation, possibly as early as the introduction, also using references to back up your point.

Perhaps the most important thing in a trend analysis, is not just the trend value itself, but its uncertainty. Uncertainty depends greatly upon mainly 3 things; time series length, noise, and autocorrelation. There is nothing in this section about the trend uncertainties and how they have been calculated. I strongly suggest that some information about how the trend uncertainty is calculated is included in this article.

Answer: We are looking for all altitudes and latitudes where the results are significant. Trend estimate are coming directly from the least squares fit. And as you say, it is determined by data errors and the length of the data set. In addition, we have now included autocorrelation corrections in the analysis. Modeling imperfections, source of autocorrelations, will be scrutinized in the companion paper submitted to this same special issue on 12. 7. 2013 (Laine et al.). This is now mentioned in the text.

11. Sec. 8 Ozone trend

L17 and beyond. You have some very interesting results and your global trends also seem very reasonable compared to past analyses. In this section you compare your results only to Wang et al. However, in your introduction you mentioned many other results/references. How do your results compare to these? I think this section misses an important part and that is the discussion of the results. What have we learned from these results? Is there anything new to be taken away from it? Is the Montreal Protocol still working? What more can be done to try and improve future trend analyses (if anything)? Please try and make this section a bit more complete, after all, the results are always the most interesting part of a paper!

Answer: Some more text has been added. Because we are not doing any attribution study,

we cannot seriously comment Montreal Protocol and related issues. The change ozone we observe can originate also from other than CFC + bromine related issues.

Minor

Page 2

Title: The title would read better if it were to be : ". . . data set for 1984-2011.."

Answer: Done.

L7: comma needed: "Global coverage, good vertical resolution, and the self-calibrating measurement method.."

Answer: Done.

L12: comma needed: "Above 55 km, SAGE II.."

Answer: Done.

L14: comma needed: "Between 35–55 km, SAGE II.."

Answer: Done.

L19: comma and change needed: "Between 25–35 km, SAGE II sunrise.."

Answer: Done.

L22: change needed: "data set covers the time period of 1984–2011.."

Answer: Done.

L22: comma needed:" ..latitudes 60° S 60° N, and the.."

Answer: Done.

L24: regarding "(so-called "hockey stick" form)": This is not mentioned anywhere in the paper apart from the abstract. I would suggest to remove this or include a quick sentence or two covering this point in the methodology.

Answer: Added also to Sec. 7 and Conclusions.

Page 3

L5: There is nothing in the abstract describing the significance of the trend results, hence it is not clear how much trust we can put into them. Hence, I suggest you include a few words to help clarify this.

Answer: Added.

L9 : I would also include the WMO (2011) in your list of references as it provides a thorough review to ozone in the atmosphere.

Answer: Added.

L26-27 it is stated "Weak signs of the ozone recovery have already been detected." I suggest that you try to include some quantification to try to give the reader a better understanding of the "weak signs". This can be done by using the references above in the article. This will also be useful for later on in the paper for comparison of results.

Answer: Done.

Page 4

L1: "Past changes in ozone have taken place slowly and this pace is predicted to continue." This sentence is a little ambiguous. What do you mean by "past changes"? Ozone concentrations are constantly varying depending on season, dynamics etc.. do you mean the long term evolution of ozone?

Answer: Added a word "average".

L2: "Therefore, requirements for the stability of ozone observations are stringent" stability in terms of what? Precision, accuracy, mechanical failure? Maybe you can expand on this a little as it is not obvious what you mean.

Answer: This topic easily leads to deep waters. We try to avoid these. We have modified text to:

Past changes in the average ozone levels have taken place slowly and this pace is predicted to continue. Therefore, requirements for the stability of ozone observations are stringent. The stability means that a given ozone density in atmosphere produces the same retrieved ozone value (allowing variation by noise) independent of the measurement time. Instrumental factors or changes in retrieval parameters are obvious sources for affecting the stability.} From the instrumental point of view self calibrating occultation instruments are good candidates for long-term monitoring of ozone.

L9: "In this work we use only the GOMOS 410 000 night measurements." You rightly state the reason why you only use only night measurements later in the paper, but maybe you can just say in brackets "(reasoning to be described in methodology)" or something similar so the reader knows this is to come.

Answer: Added.

L10-12 “In this paper we combine these two ozone data sets to create a homogeneous time series from 1984 to 2011 and look for changes in the middle atmosphere ozone profiles”. I suggest change to “In this paper, we combine these two ozone data sets to create a homogeneous time series from 1984 to 2011 in order to look for changes in the middle atmospheric ozone profiles”

Answer: Changed.

L21: ERBS needs to be written out in full first – Earth Radiation Budget Satellite. L23: change : “.. from 1979–1981..”

Answer: Done.

L27: “(with the exception of 1993) but..” why was this? Due to Pinatubo? Please add a little more information here. Also, a comma is needed after the closed bracket.

Answer: The small number of sunset measurements is thought to be related to the battery problems of SAGE II. This is now mentioned in the text.

Page 5

L1: “The amount of measurements..” change to “The number of measurements..”

Answer: Done.

L5: comma needed : “..beginning of the SAGE II mission, but..”

Answer: Done.

L7; comma needed: 600, 940, and 1020 nm..

Answer: Done.

L8: “The ozone, NO2, H2O density and the”. Better to write ozone as O3 so as to stay consistent with the rest of the list.

Answer: Done.

L8: comma needed: H2O density, and the aerosol..

Answer: Done.

L10: comma needs changing: “channel and NO2, from..” to “channel, and NO2 from..”

Answer: Done.

L22: MERRA needs to be written in full first (Modern Era Retrospective-Analysis for

Research and Applications). A reference would also be advisable;

Answer: Done and reference added.

Rienecker, M.M., M.J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M.G. Bosilovich, S.D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, et al., 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Climate, 24, 3624-3648, doi:10.1175/JCLI-D-11-00015.1

L26 comma needed: "during day and night, but only nighttime..."

Answer: Done.

Page 6

L3-4 "The number of measurements peaked in 2004 and declined thereafter. An instrument problem strongly decreased the number of measurements in 2005" Try and combine these two sentences "The number of measurements peaked in 2004 and declined thereafter, due to an instrument problem that strongly decreased the number of measurements in 2005" .

Answer: Done.

Sentences in this paragraph tend to be a little short as a whole, so perhaps combining some of them will help produce a better flow to the text.

Answer: **Done**

L4: "The latitudinal distribution is more even than for SAGE II". I suggest to expand on this to "The latitudinal distribution of GOMOS is more even than for SAGE II, owing to the differences in the orbital parameters of the ENVISAT and ERBS satellites"

Answer: The flatness of the latitudinal distribution rather reflects the multitude of targets with GOMOS. Added some text.

L15-16 : change to: "within a few percent between 20–40 km with.."

Answer: Done.

L17: comma needed: "In this work, we use GOMOS"

Answer: Done.

Page 7

L5: ECMWF needs to be written out in full (European Centre for Medium-range Weather Forecasts)

Answer: Done.

L11: MSIS90 needs to be written out in full (Mass Spectrometer Incoherent Scatter 90). A reference should also be included. Possibly,

A. Hedin, Extension of the MSIS Thermospheric Model into the Middle and Lower Atmosphere, J. Geophys. Res. 96, 1159, 1991.

J.M. Picone, A.E. Hedin, D.P. Drob, and A.C. Aikin, NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, J. Geophys. Res., 107(A12), 1468, doi:10.1029/2002JA009430, 2002.

Answer: Done. Added Hedin as a reference.

L23-24: "The self-calibrating measurement principle diminishes possible reasons for drifts of data products" How so? A bit more information could be included here on how occultation can achieve this.

Answer: Some text added.

L24: change to "However, there are factors that are.."

Answer: Done.

L25: The Figures in this section need to be reordered. The authors start with figures 8 and 9, but then later introduce Figures 5 and 6 in the text on page 12.

Answer: Done.

Page 8

L1: change to: "However, due to the short measurement series.."

Answer: Done.

L6: change to "Both data sets are screened for erroneous ozone profiles"

Answer: Done.

L14: "of data eliminated is 5

Answer: This comment was not clear to authors.

L17: comma needed: "For GOMOS, we first eliminate measurements.."

Answer: Done.

L22: "..10 km and 110 km or the mixing ratios is outside.." change to "10 km and 110 km, or where the mixing ratios are outside the range of - 0.5 ppm to 20 ppm from.."

Answer: Done.

Page 9

L11: change to “where the brackets denote the median over all..”

Answer: Done.

L25: check for spelling consistencies between US and British English. Here, behavior is spelt the US way, but is spelt behaviour on page 10 line 20. Which version are you using?

Answer: Everything changed to US English.

Page 10

L15: Why does 01-05 show an offset in shape compared to the previous years in Fig. 4? Is this an impact of instrument degradation? Is it a possible sampling issue due to there being sparser measurements ? How have you made this calculation?

Answer: We have carefully inspected the sampling patterns but no evident reason found. The curves show the relative difference of the median populations

L23: “by chemical-transport models show”, but which chemical models? References and expansion of text is needed here.

Answer: The models are WACCM and MIROC3.2. As this information is not relevant to the present paper, we have modified this sentence by moving the relevant reference to a better position.

Page 11

L7: “(there are only a few GOMOS data in 2012).” And this is, I’m assuming, due to ENVISAT’s termination? Please clarify

Answer: Done.

L13: change to: “where at least two measurements are found”

Answer: Done.

L17: change to : “In order to obtain a good coverage”

Answer: Done.

L18: change to; “a monthly grid with five to twenty degree..”

Answer: Done.

L19: “In the following we use 10 latitude bands.” Why do you choose this? Use Table 2 for your motivation

Answer: Motivation added.

Page 12

L4: “This may explain the change seen in the ozone densities at 40km” Do you mean the variability here?

Answer: No, there seem to be a change in values.

L7: change to: “GOMOS is considerably larger (0.27).”

Answer: Done.

L8-10: change to “At other latitudes the asymmetry differences between the two instruments are smaller and probably do not cause jumps in the time series of ozone density.” Also, how small is small? Quantification is needed here.

Answer: The asymmetries are important only if the two instruments have very different asymmetries. In our grid (10 deg. in latitude and one month in time) 0.1 unit difference in asymmetry (typical in Fig. 9) means 0.5 deg. in latitude or 1.5 days in time. Our knowledge about ozone latitudinal and temporal scales would guarantee that these differences could have consequences only near the polar vortex. That is why we picked up the case for Fig. 8. Quantitative estimates would require atmospheric modeling. This kind of study has been submitted to JGR by M. Toohey in June 2013.

L10: “The temporal asymmetry differences are smaller and we believe that they have a smaller impact on densities.” Any grounds for this belief? Again, some more information is needed here.

Answer: See previous answer.

L22: “By comparing years and individual months the culprit for this outlier is found to be the GOMOS-SAGE II sunset comparisons in June 2003 around 48 S.” So how does the profile now compare if you were to remove the outlier? Would it not be better to remove this outlier and publish an improved plot? Perhaps the screening of data needs to be improved slightly?

Answer: Because we are using median, this outlier has no distinct effect on the average bias. The underlying reason should, however, be studied separately from this study.

Page 13

L1: comma needed: “Because this is not the case (see Figs. 1 and 2), we continue..”

Answer: Done.

L9: “in this work. In this work”. This is a bit awkward. I suggest rephrasing to “..in this work. Here, ..”

Answer: Done.

L11: “using the appropriate bias profile.” Might be wise to just state where the appropriate bias has come from.

Answer: Done.

L17: change to; “We have to decide what statistical estimators to use to..”

Answer: Done.

Page 14

L3: “dispersion correction” What is this exactly? Do you have a reference?

Answer: Added a reference.

L5: change to: “There are alternatives to calculate a representative value in..”

Answer: Done.

L11: change to: “In large latitude–time cells the neutrality..”

Answer: Done.

Page 15

L7: change to: “The observational basis for these proxies is discussed in Harris et al. (1999) and WMO (2007)”

Answer: Done.

L8: “The last terms represent harmonic variation up to the second order, i.e. annual and semi-annual terms” so these are the harmonics in terms of cos and sin, second order, n, and the constants a and b that need to be determined? You need to explain your variables, n , a, and b

Answer: n=summing index. Parameters are now explained in text.

L23: “In Fig. 12 we show an example of the fit in the latitude band 40–50 N at 30 km.”, but in the Figure caption it says 40 km, so which is it?

Answer: 40 km is correct.

L24: Your conclusion about using 1997 agrees with previous findings (Steinbrecht, Jones, Newchurch, Cunnold), hence, I suggest you add a few lines to just confirm this.

Answer: We have added text based on findings by Newchurch and Jones.

Page 16:

your values of the solar term seem reasonable and also agree with previous findings, hence, I suggest you add a few lines to register this.

Answer: Add some text.

L14: Remember that your trends are in

Answer: Corrected.

L17: “In addition to the difference in data version (comma here) the statistical method used differs from the method of this work (comma here) but overall the results show a reasonably good agreement” How does the statistical method differ? I think this point can be expanded on a little

Answer: Text added.

L18: “The statistically significant trend results” To what level are these significant, 90

Answer: Significance statements added.

L26: change “In the mesosphere the decrease of ozone is still ongoing at a rate of 2–6

Answer: Done.

Page 17

L4: change “rates are by 2–5

Answer: Done.

L6: change “Outside the tropics the decay trends have increased by 2–5

Answer: Done.

L7: Remove the word peculiar “The 30 to 35 km region in the tropics..”

Answer: Done.

L20: comma needed: “solar and QBO proxies, and a linear term with”

Answer: Done.

L22 change to “1984–1997 show clear statistically significant (*to what ever confidence level*) negative trends”

Answer: Done.

Fig 2. Caption change to “Latitude distributions of SAGE II (top) and GOMOS (bottom) measurements. Notation: Solid and dashed red: SAGE II sunrise for 1985 and 2004, solid and dashed blue: SAGE II sunset in 1985 and 2004, solid and dashed black: GOMOS for 2004 and 2010.”

Answer: Done.

Fig 3. What is in the brackets, please define in caption

Answer: Done.

Change: “The corresponding differences between SAGE II sunrise and sunset measurements are also shown”

Answer: Done.

Figures

Fig5. Remove the header or is it necessary?

Answer: Done.

Fig8. You state that the densities are monthly in the main text, maybe just state this in the caption just to be sure

Answer: Done.

Fig11. Do you need the header? Also labels on plots 1-7 could be a little larger

Answer: The header removed. Labels are changed to a larger font where possible without blocking details of graphs.

Fig12. Again, state that these are monthly values in the caption.

Answer: Done.

Fig16. Change to “..95

Answer: This comment was not clear to authors.