REPLY TO REFEREE #2

We thank Referee #2 for his/her thorough review of the article and useful comments.

General comments:
This is an interesting and generally well written manuscript dealing with the investigation of possible long-term drifts between different sets of stratospheric ozone time series. Still, overall this is a relevant and timely paper that will make an important contribution to the field. I recommend publication of this manuscript after the following specific comments were considered and addressed.

 # We thank referee #2 for recommending the paper to publish in AMT.

Specific comments:

1. Page 473, line 11: "and drifts are less than ±0.3%yr-1 at all stations". Looking at Fig. 4 this statement is not correct. There are larger drifts between 20 and 40 km, e.g. for MLS/Aura relative to several of the lidar stations. See also comments below.

 # Corrected. Please find the revised Abstract, Line 11—15, i.e., "All measurement techniques show their best agreement with respect to the lidar at 20—40 km, where the differences and drifts are generally within +/-5% and +/-0.5%/yr, respectively, at most stations."

2. Page 473, Line 22: "of THE Antarctic ozone hole" ?
Page 473, Line 26: "The analysis of stratospheric ozone trends .. are currently" -> "The analysis of stratospheric ozone trends .. is currently"

 # All are corrected.

3. Page 474, Line 15: "In a recent study .." The exact meaning of the sentence is not clear (at least to me). What irritates me is the phrase "to a lesser extent". Do you mean to suggest that Dhomse et al. found that the TOC increase is to a lesser extent (compared to what?) caused by an enhanced residual circulation? What is the main effect then? What exactly is confirmed by Harris et al. (2008)?

 # The sentence has been reformulated. Please find the revised Introduction, Page 2, Line 3—7.

4. Page 476, line 15: I believe the expression "radiations" does not exist. Suggest to change to: " .. emission of lidar radiation at two wavelengths"
Page 477, Line 4: "in the use of reference wavelength" -> "in the choice of the reference wavelength"?
Page 477, line 16: Remove comma after "and"
Page 481, line 1: "and are" ->"that are"
Page 482, line 1: "in the work by Froidevaux et al. (2008); Jiang et al. (2007); Livesey et al. (2008)." -> "in the works by Froidevaux et al. (2008), Jiang et al. (2007) and Livesey et al. (2008)."

 # Corrected.

5. Table 1: Why is SBUV/2 data really only used until 2007?

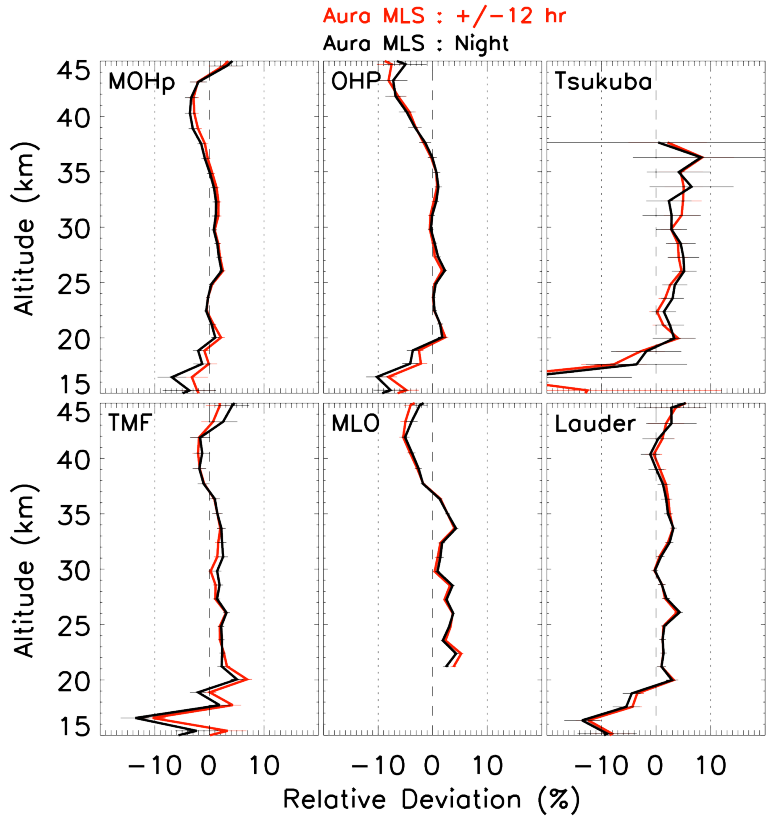
 # SBUV(/2) NOAA series is still measuring ozone. But the data were not publicly available at the time of this study. Hence, SBUV(/2) data are used until 2007 only.

6. Fig. 1, panels d – f: These panels suggest that there are no co-locations between ozone sondes and SBUV, SAGE II and HALOE, which can certainly not be true.

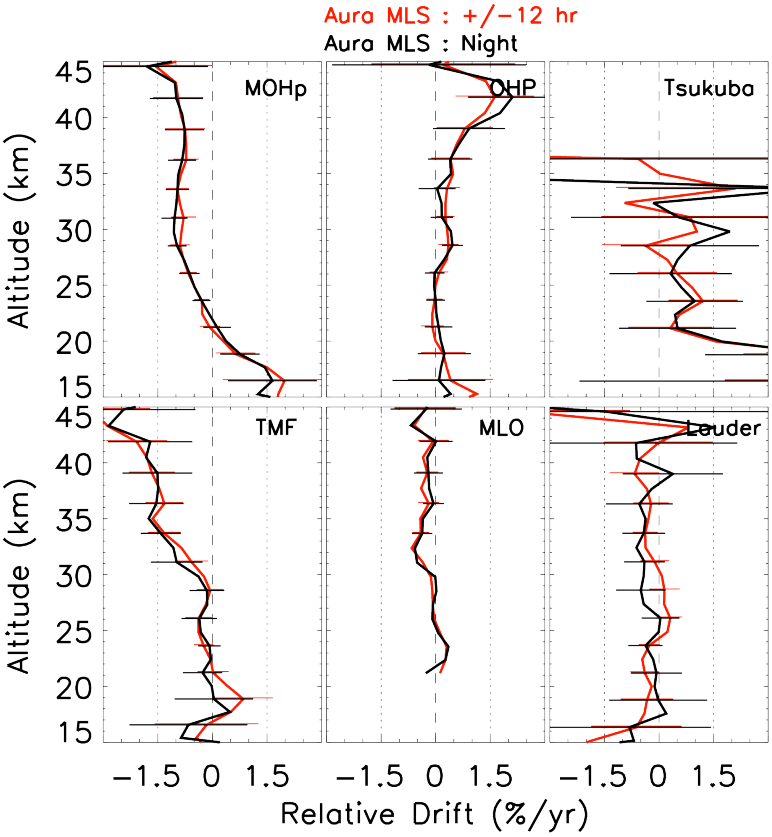
 # The analysis includes the cross-comparison of long-term data sets such as lidar, SBUV(/2), SAGE II and HALOE only. Ozonesondes are not used for the comparison with satellite observations. Therefore, panels d, e and f of Figure 1, do not include ozonesondes.

7. Page 483, line 26: "with a time difference maximum of ±12 h." Diurnal variations of stratospheric ozone may also be important below 40 km altitude, and certainly between 40 and 50 km.

 # In order to check this, we have performed an analysis using the night-time Aura MLS profiles (profiles measured between 8.00 PM and 3.00 AM) and then, average bias and relative drifts are calculated at all stations. Please see Figures : SUPP_FIG_03 and SUPP_FIG_04 for the results (figures are shown below). The average bias and relative drifts do not vary significantly between these two analyses (Night-time and +/-12 hr analyses). So we have kept the temporal criterion as +/-12h.



SUPP_FIG_03: *Vertical distribution of the average relative differences of Aura MLS observations with various lidar measurements for two different temporal criteria.*



SUPP_FIG_04: *Vertical distribution of the relative drifts of Aura MLS observations with various lidar measurements for two different temporal criteria.*

8. Page 484, equation (2): Why is the sum in the numerator on the right hand side also over j, i.e. altitude or pressure? I think this should just be a sum over i.

 # This has been corrected. Please find the revised Equation 2.

9. Page 486, line 10: "Hence, the analysis excludes the outliers such that the profile consisting of very low and high ozone values at all altitudes are removed from the analysis." What are the exact quantitative criteria for this selection process? What are very low and high values?

 # "The comparison of lidar with SAGE II and HALOE shows relative differences higher than 200% below 17 km and at 45 km for some profiles. These altitudes have been removed from those profiles for the analysis. But, they are very few in number, less than 5 in total for a station. Similarly, ozonesondes at OHP show very small ozone values of the order of 1*10^10 molecules/cm^3 in two days over the period. Those profiles are also removed from the analysis." Please find this text in the revised Section 3.2, Paragraph 2, Line 7—14.

10. Page 486, lines 23 – 27: Are the MLS averaging kernels not available for this purpose?

 # The vertical resolution of the MLS profiles are higher than that of SBUV(/2). The comparison of SBUV(/2) with the convolved lidar does not show significant differences compared to that between SBUV(/2) and non-convolved lidar, as shown in Nair et al., (2011). In addition, ozone lidar measurements have a vertical resolution of about 3 km above 30 km, which is comparable to that of MLS. Therefore, we opted not to use the Averaging kernel information for the comparison between lidar and MLS.

11. Page 487, line 15: What is a "slightly large drift"? Can you provide a quantitative statement?

 # The comparison between SBUV(/2) and lidar for which the lidar ozone number density is converted to DU using NCEP data at OHP shows a drift of about 0.5%/yr (Nair et al., 2011). This is termed as the slightly larger drift because when Arletty pressure-temperature (P/T) data are used instead of NCEP P/T data, the drifts are reduced to 0.2%/yr. This has been cleared in the revised version. Please find the revised Section 3.3, Paragraph 3, Line 3—5.

12. Page 487, line 19, Arletty data: If I understand correctly, then this data set consists of the MSIS-90 climatology for altitudes above 30 km. MSIS probably does not include any possible temperature trends caused by dynamics changes (possible enhancement of Brewer-Dobson Circulation), changing radiative balance due to increasing green- house gases etc. Please provide more detailed information on MSIS 90 and its limitations. I'm not an expert here, but I'd be quite skeptical whether MSIS 90 really performs better than ECMWF in this context and at these altitudes.

 # Our analysis reveals a temperature trend of less than –0.5K/dec above 30 km from MSIS-90 temperature data in 1987—1998. This is smaller than that estimated using the ECMWF data. The MSIS-90 model data are based on the Middle Atmospheric Program (MAP) Handbook (Labitzke et al., 1985) tabulation of zonal average P/T data below 72.5 km and the NCEP P/T data below 20 km. These are stated in the revised Section 3.3, Paragraph 3, Line 11—34.

13. Page 490, line 26: I suggest changing "slightly large" to "slightly larger" or "an increasing negative"

 # This has been changed.

14. Page 491, line 15: "However, some significant drifts .. for SAGE II .., for HALOE .." MLS should also be mentioned here, because it exhibits really large drifts at MOHp, OHP, and TMF.

 # The sentence has been modified.

15. Page 492, lines 1-6: This paragraph suggests that anomalous drifts in the lidar measurements at MOHp and TMF are responsible for the relative drift between MLS/Aura and the lidars. Is this the intended meaning of this paragraph? Ins't there good reason to believe that the lidar measurements are more accurate? I also don't fully understand the statement "and TMF lidar shows high ozone values in 2008 and 2009 above 30 km compared to all other measurements". What does "other measurements" refer to? Out of the satellite instruments discussed here, only MLS/Aura was operational in 2008 and 2009, and perhaps the drift is actually mainly caused by a drift in the MLS data set. Please provide more detailed information on the possible issues with the TMF measurements in 2008 and 2009.

 # The other measurements are SBUV(/2), SAGE II, HALOE, UARS MLS and ozonesondes. Because ozone concentration and estimated ozone anomalies from lidar measurements at TMF agree well with those evaluated from these data sets until 2007. Conversely, a clear increase in ozone anomaly is found in 2008 and 2009 for TMF lidar compared to the ozone anomalies of TMF lidar prior to 2008 above 30 km. Similarly, MOHp lidar shows an increase in ozone anomaly after 2007 compared to those evaluated from MOHp lidar before 2007 above 30 km. However, the ozone anomalies of Aura MLS do not show any discontinuity and exhibit similar pattern over the analysis period at all altitudes. These are stated in the revised Section 4.2.1, Paragraph 2.

16. Page 492: "It implies that the lidars can be taken as a reliable reference for drift evaluation of satellite and other ground-based measurements." Given the apparent drift between MLS/Aura and the lidars at MOHp and TMF this statement is somewhat contentious, because it is not generally applicable. It is not generally applicable, because SBUV/2 data were only used up to 2007, and the differences

between MLS/Aura and the lidars at MOHp and MTF start in 2007 and 2008/2009, respectively. I suggest changing the statement, otherwise it directly suggests that there is an anomalous drift in MLS/Aura (see previous point).

 # The sentence has been removed.

17. Page 493, line 4: "drifts maximum" -> "maximum drifts"
Same sentence: "and Lauder and ±0.4%yr-1 at OHP". Looking at the Fig. 6a) this doesn't seem to be correct, because differences well exceed 0.5%/yr between about 33 km and 43 km and also below 18 km.

 # The sentence has been modified.

18. Page 493, line 10: "At MLO, the coincidences are available in 1999–2003 only. This is the reason for the estimated large drifts at MLO." These two sentences already appear 3 lines above.

 # The sentences have been removed.

19. Page 494, line 13: "relatively fewer" -> "relatively few"

 # Corrected.

20. Page 494, section 2.4.2: "Average of the drifts of long-term measurements" Are the averaged drifts simply arithmetic mean values of the 3 individual drift values, or are the weighted by the number of coincidences with each of the other instruments?

 # This is the mean of the drifts of one measurement technique (e.g., SBUV(/2)) at one station, obtained from the comparisons with other measurement techniques (e.g., lidar, SAGE II and HALOE) over that station.

21. Page 495, line 1: "Generally, as found in the previous comparisons, all data sets show small drifts of around ±0.2%yr-1 at 18–45 km.. " Does this statement refer to drifts averaged over the 18-45 km altitude range, or to typical values at individual altitudes within this range? Looking at the panels, the first is/should be the case, and I suggest to mention this explicitly in the paper.

 # No, this is a general statement considering each altitude, not the mean of the drifts over the 18—45 km altitude range.

22. Page 495, line 3: "slightly large" -> "slightly larger"
23. Page 495, line 4: "play" -> "plays"

 # These are corrected.

24. Page 495, line 8: "It is obvious (from Fig. 4) that the 8 yr data record of Aura MLS yields comparable drifts as of the long-term measurements at all regions." Looking at Fig. 4 this statement is not correct. Fig. 4 shows that at MOHp there is a -1 % / year negative drift between MLS and the lidar over an extended altitude range. At OHP the drift reaches -1.5 % / year near 40 km, and at TMF a large negative drift is observed. 1% / year corresponds to 10% / decade, and if MLS is the origin of these anomalous drifts, then MLS is probably not suitable to extend the SAGE/HALOE time series.

 # The sentence has been rephrased in Section 4.3, Paragraph 1, Line 1—4 as, "In general, the 8 years data record of Aura MLS yields comparable drifts to the long-term measurements with respect to most of the ozone lidar measurements, except with MOHp and TMF ozone lidars above 30 km."

25. Page 496, line 6: "because of the increase in ozone lidar measurements after 2007, as discussed in Sect. 4.2.1." Again, this suggests that the MOHp lidar measurements are erroneous. From the analysis presented here this conclusion cannot be drawn, in my opinion. How do we knw that the increase in the lidar measurements has not actually occurred? Looking at Fig. 8, and at altitudes between 23 and 42 km my visual impression is that there is a drift between MLS/Aura and MOHp lidar not only after 2007, but already from 2004 onwards.

 # The MOHp ozone lidar measurements show a clear increase in ozone anomaly after 2007. However, the Aura MLS ozone does not show an increase. From this, we reached into a conclusion that the negative difference of the Aura MLS data at MOHp can be due to the increase in ozone lidar measurements. But more data sets and more years are required for a general conclusion, as we have only the Aura MLS measurements now. Therefore, the text has been revised. Please find the revised Section 4.2.1, Paragraph 2 and Section 4.3.1, Paragraph 1, Line 11—14.

26. Page 496, line 18: "which can be due to higher lidar ozone during the period as compared to other measurements". Again, what "other measurements" are you referring to? Can this be backed up with other published results?

 # Please find answer to the comment No. 15.

27. Page 497, line 3: "It indicates that the combination of these satellite observations provides a potential long-term data set for the evaluation of long-term ozone trends in the stratosphere." I agree that the drifts of the combined time series are very small, and this is a very good result. However, given the inconclusive drift analysis between MLS/Aura and MOHp, OHP and TMF, I'm not convinced that the MLS data set

is suitable for long-term trend studies at the current stage. I believe the remaining inconsistencies have to be resolved (not necessarily as part of this study), before such a general statement can be made.

 # The sentence has been restated in Section 4.3.2, Paragraph 1, Line 16—20 as, "It indicates that the combination of these satellite observations can be a potential long-term data set for the evaluation of long-term ozone trends in the stratosphere even though Aura MLS shows significant drifts with lidars at some stations above 30 km."

28. *Page 497, line 20: "This is attained first by comparing all measurements with respect to lidars, which yields drifts of less than ±0.3%yr-1 at 20–40km for all observations. Aura MLS with 8 yr of observation also shows drifts that are comparable to those from the long-term data sets at all stations." I disagree with these statements. Looking at Fig. 4 there are larger drifts of MLS/Aura with several lidar stations. They may of course – as you suggest – be due to issues with the lidar measurements, but this needs to be demonstrated.*

 # The sentence has been changed. Also, please find answer to the comment No. 15.

29. *Page 511, caption Fig. 5: "The error bars correspond TO the .."*
30. *Page 514, caption Fig. 8, line 2: remove extra space after "(left panel)"*

 # All are done.

31. *Page 499, line 13: "[Data]." ?*

 # This is the citation method, suggested by WOUDC.