We thank Referee #1 for his thorough review of the article and useful comments.

General comments
This paper is a valuable addition to the literature comparing measurements of the vertical distribution of ozone. I have a few concerns which should be addressed prior to publication.

 # We thank referee #1 for recommending the paper to publish in AMT.

1. I suggest that some degree of additional analysis is presented – for example, are there annual cycles in the differences? What about autocorrelation at different time-scales?

 # The relative differences of Aura MLS at MOHp shows seasonal differences with positive deviations in winter at 40 km. This is stated in the revised manuscript. Please see Section 4.3.1, Paragraph 1, Line 15—19.
The autocorrelation is calculated for all data sets and is found to be within +/-0.3 in the 20 – 40 km altitude range. The standard deviation of the slope or drift is also analysed using the equation, given by Frederick (1984), that makes use of the autocorrelation term. The standard deviations are very similar to those calculated using the equation from Press et al. (1989), as described in the manuscript . This is inserted in the revised Section 3.2, Paragraph 1, Line 6—11.

2. Second, while most of the paper is clear and well written, Section 4 (Results and Discussion) is less so. I make a few suggestions below, but the authors should look carefully through this section to make sure that the points they want made are clear. Part of this is making the relevance of some comments clear; part is clarifying the text. It is worth doing.

 # We have considered all these points. Please find answers to the specific comments below.

3. Third, the authors should consider putting some of the information in the instrument descriptions into tables. This needs careful consideration of how viable it is, but if it can be done it would be helpful to the reader.

 # Table 1 is modified with the instrument description. Please see the revised Table 1.

Specific comments

4. 474, 18-21: This was recognised before them, so if Weatherhead and Anderson are used, make it an example.

 # This has been done. Please find the revised Introduction, Page 2, Left column, Line 8.

5. 475, 7-11 A short summary of the main results from Nair et al 2011 would be helpful.

 # "The study showed that the considered data sets agree well with the ozone lidar observations at OHP, showing an average bias of less than +/- 0.5 % in the 20—40 km altitude range. All measurements are stable and their relative drifts are shown to be within +/- 0.5 %/yr in this altitude range". This is inserted in the revised text. Please find the revised Introduction, Paragraph 3, Line 21—25.

6. 476, 13 I would delete 'self-calibrating' here and leave it for section 2.4. It is probably worth pointing out in the summary that the lidars exhibit good stability against all other instruments.
7. 476, 21 '..range-corrected. . .'
8. 477, 16 '. . . and 8 channels at TMF. . .' (delete comma)

 # All are corrected.

9. 477, sxn 2.1 A comment on the effect of changes in the lidar setup on the quality and stability of the ozone records would be instructive.

 # Please find the revised Section 2.1, Paragraph 2, Line 18—20, where it is stated as "For instance, at OHP, the receiving system had 2 acquisition channels until 1993. It was then modified to accommodate 6 channels (4 at 308, 355 nm; 2 at 332, 387 nm) in 1994 (Godin-Beekmann et al., 2003), which improved the observational capacity of the lidar system."

10. 478, sxn 2.2 Similarly for the ozonesondes.

 # Please find the revised Section 2.2, Paragraph 1, Line 15—26, where it is written as, "Several studies (e.g. Johnson et al., 2002, Smit et al., 2007) revealed that the ENSCI sondes overestimate ozone by 5 % below 20 km and 5—10 % above 20 km as compared to SPC-6A sondes, when both sondes operate with 1% KI full buffer cathode solution. The BM sondes underestimate ozone by 10% while the ECC sondes with 1 % KI cathode solution overestimate ozone by 5 % compared to that with 0.5% KI (Stubi et al., 2008). Similarly, the KC sondes underestimate ozone by 10 % above 50 hPa (Deshler et al., 2008)."

11. 478, sxn 2.2 Given the on-going revision in the ozonesonde records, the authors should state exactly which ozonesonde data sets they

used.

 # This has been stated. Please find the revised Section 2.2, Paragraph 3, Line 9—10; Paragraph 4, Line 12—13; Paragraph 5, Line 11—12; Paragraph 6, Line 14—15; **and** Paragraph 7, Line 17—19.

12. 478, 28 '. . . sensor solution were used for ozonesonde measurements from. . .'

 # This has been changed.

13. 481, 9 Should it not be 14 years? I.e. from Sept 1991 to November 2005. And write 'year'.

**# Yes, It is 14 years. 'Year' is added. Thank you for commenting this.**

14. 482, 24/25 What is the sensitivity of the results to the spatial criteria used. What if either set of criteria had been used for all satellite overpasses?

 # An analysis is performed for all data sets at MOHp using both spatial criteria (please see Figures shown below : SUPP_FIG_01 **and** SUPP_FIG_02 **for the results). SAGE II and HALOE show comparatively larger biases and drifts for the comparison with lidar when the spatial criterion is tightened (i.e., Lat : +/- 2.5° and Long : +/-5°) for finding coincidences. In addition, the number of coincidences are fewer and that results in large uncertainty. For Aura MLS, the relaxed spatial criterion of Lat : +/- 5° and Long : +/- 10° results in small bias and drift below 20 km. Furthermore, the derived drifts are also insignificant. SBUV(/2) does not exhibit any difference in bias and drift by relaxing the spatial criterion (Lat : +/- 5° and Long : +/- 10°). However, these biases and drifts are still within the error bars of the measurements in each case.**

SUPP_FIG_01: *Vertical distribution of the average relative differences of different data sets with MOHp lidar for different spatial criteria.*
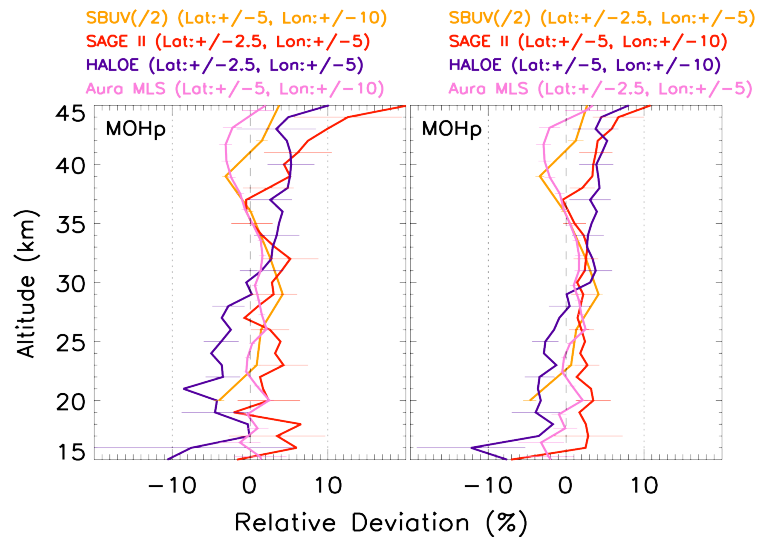
SUPP_FIG_02: *Vertical distribution of the relative drifts of different data sets with MOHp lidar for different spatial criteria.*
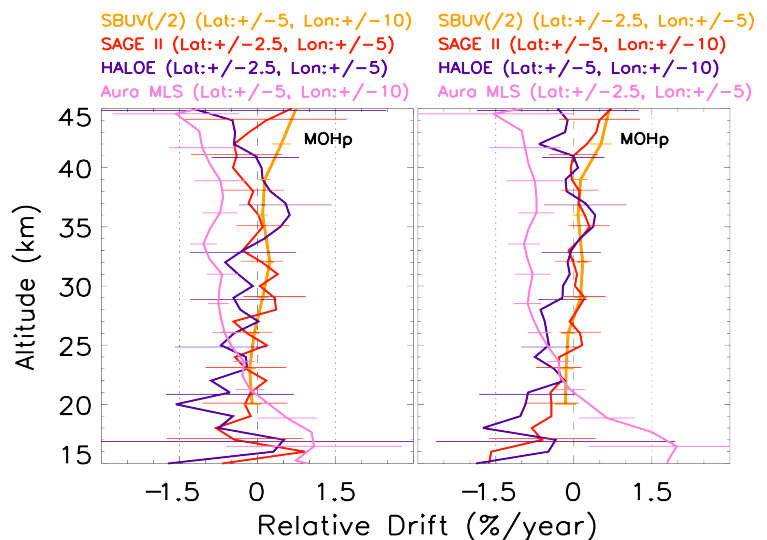
15. 486, sxn 3.2 See above – this is OK but needs some justification.

 # This has been done. Please find answer to the comment No. 1. Also, please see the revised Section 3.2, Line 6—11.

16. 486, 18 '. . .these units, and ozone. . .'

 # Changed.

17. 487, 19-27 Some comments are needed on the quality of the Arletty temperature data and on any trends found in them. This description is not sufficient given the problems that temperature errors, particularly drifts, can cause.

 # We have done some trend analysis using the temperature data and the results are given in the revised text. "Arletty temperature data are made using the ECMWF operational analyses at the studied altitude regions since 1999. Prior to 1999, Arletty makes use of the ECMWF data up to 30 km and the MSIS-90 above 30 km. In order to demonstrate which temperature data are useful for the analysis, trends in the NCEP and Arletty temperature data are calculated using a simple linear regression at MOHp, one of the stations considered in the study. The NCEP temperature data show insignificant trends of less than –1 K/decade below 30 km and about –1 to –2 K/decade in the 30—40 km altitude range over the period 1987—2009. However, temperature trends derived from the ECMWF data are within +/- 1 K/decade at 15—30 km in 1987—2009 and at 31—43 km in 1999—2009 and are also insignificant. On the other hand, the temperature data from the MSIS-90 model show insignificant trends of less than –0.5 K/decade above 30 km in 1987—1998." Please see the revised Section 3.3, Paragraph 3, Line 11—34.

18. 491, 2-4 Not sure what this sentence means – makes better sense without the final 4 words but I am not sure if that is what is meant.
19. 491, 6 '. . . for consistency, . . .'
20. 491, 7 '. . ..series and the drift is derived..
21. 491, 10 '. . . used as the reference . . .'
22. 491, 20 'Aura MLS exhibits comparable drifts to those of SAGE II and HALOE even though it has. . .'
23. 492, 16 'They are similar to those. . .'

 # All are corrected.

24. 492, 18 I am not sure what the point of this reference to 'certain altitudes' is. What should the reader understand from this?

 # This sentence has been modified. Please find the revised Section 4.2.2, Line 6—10, where the text is revised as "Generally, all lidars exhibit very small drifts (within +/- 0.2 %/yr) with SBUV(/2), but some of these are significant at MOHp (at 30, 32, 42 and 45 km), Tsukuba (at 26 and 32 km), TMF (at 32, 42 and 45 km) and MLO (at 30, 32 and 42 km)."

25. 493, 2-11 Same here. Greater clarity needed. If what is written is correct, what is the message?

 # This sentence has been restructured. Please find the revised Section 4.2.3, Paragraph 1.

26. 494, 4-14 I assume the point of this paragraph is to say that there are larger drifts between SAGE II and HALOE than between other instrument pairs, but that this needs to be interpreted in the light of the other comparisons, i.e. with other instruments effectively being used as transfer standards. If so, shorten the paragraph.

 # Done. Please find the revised Section 4.2.3, Paragraph 4.

27. 495, 9 '. . . drifts to the long-term. . .'
28. 495, 9-11 '. . . a strong candidate for extending the observations of SAGE II and HALOE. Here we assess. . .'

 # All are corrected.

29. 496, 5-7 I understand this point numerically, but what should I believe is right?

 # In order to test this, we have calculated ozone anomaly at MOHp and TMF from different measurement techniques (lidar, SBUV(/2), SAGE II, HALOE, UARS MLS, Aura MLS and ozonesondes). The ozone concentration and estimated ozone anomalies from lidar measurements at TMF agree well with those evaluated from these data sets until 2007. On the contrary, a clear increase in ozone anomaly is found in 2008 and 2009 compared to the ozone anomalies of TMF lidar prior to 2008 above 30 km. Similarly, the MOHp lidar shows an increase in ozone anomaly after 2007 compared to those evaluated before 2007. Nevertheless, the ozone anomalies from Aura MLS do not show any discontinuity and exhibit similar pattern over the analysis period. However, it is clear that we need more data sets to reach a general conclusion about these differences. Therefore, the sentence is revised. Please find the revised Section, 4.2.1, Paragraph 2 and Section, 4.3.1, Paragraph 1, Line 11—14.

30. 498, 20-22 Too strong. This study provides strong evidence to show that Aura MLS could be used to extend the SAGE II and HALOE and for the combined record to be suitable for trend studies. However more work is needed. . ..

 # This has been considered. Please find the revised Section 5, Paragraph 4, Line 9—11.

31. *498, 23 Is 'satellite calibration' really a goal for NDACC? Personally, I do not think it should be, but at least double-check that it is before saying so!!*

 # This has been removed.

32. *Figs 2&4 Use the layout for Fig 7, otherwise they are likely to be too small.*

 # The figures are modified. Please see the revised Figures 2 and 4.

33. *Fig 8, capt. '.. left panel), OHP. . .' (no space)*

 # Changed.