Anonymous Referee #1

Received and published: 17 February 2016

In the manuscript by Leventidou and co-authors, Tropical tropospheric ozone columns from nadir retrievals of GOME-1/ERS-2, SCIAMACHY/Envisat, and GOME-2/MetOpA (1996-2012), monthly averaged tropospheric ozone columns from 1996 – 2012 for GOME/ERS-2, SCIAMACHY/ENVISAT and GOME-2/MetopA data are presented using a new CCD algorithm. Unfortunately, the manuscript as written fails to convince firstly of the actual necessity of a new algorithm, then of the validation of said algorithm and surely of the fact that this new data may be "used in climate models and tropospheric ozone trend studies" as the authors conclude. Summarizing the comments made inline the text, I found the text lacking many important points as to how the three different satellite data were homogenized in order to produce a stable long term record; how the method was applied to all three datasets; how crucial choices with respect to statistical analysis were made [for e.g. why is the data in 2.5x5deg bins but are then compared to 5x5deg bins to the ozonesondes.] The error analysis section is not an actual error analysis, but a statistical consideration. As such, it should be changed, renamed and its focus re-established. Finally, the validation to the ozonesondes leaves a lot to be desired, whereas the validation with the other version of the SCIAMACHY tropical troposphere ozone columns is inadequate. I suggest that the team re-shapes the manuscript by including sections, such as the section of presenting the new dataset. One Figure out of 11 Figures, is not what one would call "present the data". And also by excluding, to shortening sections, and re-establishing the focus.
Further to my comments in the text, the following major points may be made on tables and figures: Tables 2, 3 and 4 require re-thinking and re-writing, as per my reservations about the relevant section and its findings. Figure 2, bottom plots: the legend is extremely difficult to read. Figure 3, no explanation given for the bottom plots, why are some histograms pink and others not? Figure 4, extremely difficult to understand and to extract relevant information from. I suggest a complete overhauling. Figure 6. Why not make similar maps from the LNM method on SCIAMACHY and compare these? Figure 10. I strongly disagree with this figure, with putting all ozonesondes, all satellites, all years, all seasons, all averaging techniques, in one pot and calculating a correlation as such. Remove this plot, which does not add any new information to what you already described in the text for each ozonesonde separately anyway. Figure 11. As already discussed, this does not show comparisons between the two methods but rather comparisons of each of the methods with the ozonesondes. I suggest you re-think this Figure and its accompanying section.

Most of the answers to the points raised above are given in the specific answers to the following comments.

1. **There is no reason to have this in parenthesis, I would suggest to exclude them.**

   Done

2. **What is SHADOZ?**

   SHADOZ is the Southern Hemisphere ADditional OZonesondes which has been included in the text.

3. **... from the CCD technique....**

   Changed

4. **This part of the sentence appears to be out of sequence, please rephrase.**

   Rephrased

5. **I am assuming you should delete this, since you give NMVOCs as acronym.**

   Sentence rephrased as:

   "volatile organic compounds (VOCs) and non-methane volatile organic compounds (NMVOCs)"

6. **Substitute by: ... as well as non-human activities....**

   Done

7. **Which convective system is that?**

   Convective systems such as tropical cyclones

8. **Where what builds up?**

   The sentence has been rephrased as:

   "Second, convective systems such as tropical cyclones, can transport ozone precursors many kilometers away from their source, resulting in ozone production at remote areas where it builds up"

9. **In which altitude within the troposphere? also in the PBL? also in the UTLS?**

   Regarding the tropospheric ozone lifetime, the following text has been included in the manuscript:

   "The globally averaged tropospheric ozone lifetime is 22±2 days (Stevenson et al., 2006). On the other hand, ozone's lifetime in the boundary layer is much shorter (a few hours) because it is more probable to get destroyed by surface deposition and chemical reactions, whereas in the middle and upper troposphere it's lifetime is on the order of weeks to months (Cooper et al., 2014)."

10. **I suggest you start the new paragraph here and move the previous sentence to the previous paragraph.**

    The paragraph has been divided into two paragraphs. The first, describes the role of ozone as an oxidizing agent and a greenhouse gas and the second presents the influence of convection on tropospheric ozone abundance and the different ozone lifetimes at different tropospheric altitudes.

11. **Why does it increase? a sentence should be given on this issue, it is not enough to reference other works. This reason is basically why this new CCD retrieval is necessary, isn't this so?**

    A new paragraph has been created discussing the reasons for tropospheric ozone increase and highlighting the need for a reliable tropospheric ozone dataset:

    "Ozone is removed from the troposphere by several chemical reactions (3470±520 $Tgyr^{-1}$) but it is also dry deposited (770±180 $Tgyr^{-1}$) at the surface (IPCC, 2007). Nevertheless, the world population growth and the industrialization have led to a strong increase in anthropogenic emissions, resulting in an increase in the tropospheric ozone burden (300±30 Tg (IPCC, 2007)) by 1-7% per decade in the tropics (Beig and Singh, 2007; Cooper et al., 2014). Undoubtedly, the need to control the tropospheric ozone increase is crucial. Every potential monitor and study of long-term tropospheric O3 changes as well as the quantification of associated radiative forcing using chemical transport or climate models have to rely on the availability of reliable tropospheric ozone

data."

12. Using which satellite instrument?

As refered in the manuscript the satellite istrument used is :

"Total Ozone Mapping Spectrometer (TOMS) for ozone and temperature-humidity and infrared radiometer (THIR) for cloud data."

13. Is this the work by Fishman et a., 1990? this phrase is then out of sequence.

It is work of Fishman et al., 1990. The citation has been moved.

14. ...using...

Corrected

15. ... (LNM) technique....

"technique" has been added

16. I suggest changing to "In a subsequent step, the monthly mean..."

The word "Afterwards" has been replaced with the phrase "In a subsequent step"

17. Replace by Coldewey-Egbers here and in all references to Melanie's 2005 work.

Done

18. Sections

Changed to *"sections"*

19. ... for ozone below the cloud layer that cannot be measured.

Corrected as proposed.

20. This phrase feels out of sequence here, what is one to see below?

The phrase *"(see below)"* has been replaced with the phrase "*as it will be discussed later* "

21. Hence, you are stating that the method starts with an inherent 5 D.U. natural variability which cannot be excluded from the result?

As discussed in section 3 that describes the method, the assumption made in the original CCD method by Ziemke et al. (1998) requires that the (stratospheric) ozone column above 200hPa is independent of longitude in the tropics (Valks et al., 2003). However, a zonal variability in the ACCO of less than ~5DU (Fig. 2a) exists on monthly time scales in the tropical region due to episodic tropical waves (Kelvin and Rossby waves) and the natural variability in the cloud top height.

In order to adjust the ACCO of each orbit to 200 hPa, we use the climatology of Fortuin et al. (1998) to extend or subtract partial columns in order to determine the ACCO down to the fixed 200hPa level.

Afterwards, we restrict the zonal variability of the ACCO by removing outliers. We consider as outliers, the daily averaged ACCO values per 2.5°x5° gridbox having a 1σ standard deviation of the mean greater than 10 DU and being greater than the daily averaged total column in the same gridbox. Additionally, in order to homogenise the ACCO, we keep only the values that differ less than 5 DU from their neighbouring grid boxes. Finally, the average for each 2.5° latitude band ACCO is calculated and the standard deviation and the number of measurements are computed.

22. It's a bit peculiar to mention Fig. 2a before Fig. 1, especially since it is not discussed further below in the text. I suggest that you move the phrase where you mention and discuss Fig.1 before the zonal variability discussion.

The order of the figures has been changed and the whole section has been reshaped.

23. What is this acronym?

The Inter-Tropical Convergence Zone (ITCZ) has been added

24. This Figure shows only two months of one year. A comment should be made as to the distribution in other months/years.

The citation of Sassen et al. (2009) has been added referring to the distribution of deep convective clouds (DCC) from 2 years averaging using CALIPSO/Cloudsat. Also the following text has been included:

"Due to the immigration of the ITCZ, these clouds are located south over the western and central Pacific Ocean, northern South America, and equatorial Africa in boreal winter and spring, whereas in boreal summer, the highest DCC occurrences are located over the Indonesian region and the Bay of Bengal. Fig 3a shows the distribution of the DCCs in January and August 2008 for SCIAMACHY ((cf>0.8 and cth>9km) SACURA) and GOME-2 ((cf>0.8 and cth>7km) FRESCO), indicating the location of the ITCZ. Both instruments and cloud algorithms agree in the location of the DCCs but not on the number of the DCCs per grid box, mainly due to differences in the cloud algorithms used and the spatial resolution of the instruments. "

25. Is this due to the SCIAMACHY pixel size? the different cloud algorithm used? another factor? is this difference expected?

This is mainly due to the different cloud algorithms used. FRESCO assumes that clouds behave as opaque Lambertian surfaces resulting usually in the effective (optical centroid) cloud top height which lies below the physical/geometrical cloud top height. SACURA cloud top height retrieval algorithm takes into account radiative transfer inside, above and below the clouds, therefore, it provides more realistic cloud top heights. The part referring to the different cloud algorithms approaches has been shifted before discussing the fact that for SACURA (SCIAMACHY), 25% of the cloud top heights in the western Pacific are higher than 9km, whereas for Fresco (GOME-2), the same frequency is only met for clouds above 7 km.

26. I do not think it is very polite to state this, to be honest, even though it might be true in some respects. Please remove.

The phrase has been removed and replaced by the following discussion later in the text:

"Another correction approach for the difference between the cloud pressure level and the 200 hPa level was also used by Valks et al. (2014) assuming a constant (small) ozone volume mixing ratio of 5 ppbv. Valks et al. (2014) concluded that the correction term is small (less than 2 DU) and therefore the difference with the climatology considered negligible. "

As seen in Figure 4b, this can be true for the case of SCIAMACHY ACCO in August 2008 (yellow dots before and green dots after climatological correction). On the other hand, in the case of GOME-2this can be true only for northern latitude bands where there are enough cloudy data and the deviation within the month is small.

27. What are these numbers? where does this formula originate? if the authors wish to have their technique better documented, all relevant details should be added here.

The formula has been explained as follows:

"In order to convert the volume mixing ratios (ppm) at the i-th level to Dobson units (DU), the following formula was used, taking into account the ideal gas law and the horizontal surface density (Ziemke et al., 2001):

$$VC(i)= c \times 0.5\times[vmr(i)+vmr(i+1)]\times[p(i)-p(i+1),] \quad (1)$$

where vmr is the volume mixing ratio (ppmv), p, the pressure (hPa), and

$$c=k_B\cdot T_s\cdot N_A/(\mu\cdot p_s\cdot g)= 0.7889 \ DU \cdot Pa^{-1} \cdot ppmv^{-1}.$$

T_S is the standard temperature (273.16 K), p_s, the standard pressure (101325 Pa), k_B, Boltzmann's constant ($1.3806\times10^{23} J\cdot K^{-1}$), N_A, Avagadro's number (6.022×10^{26} molecules$\cdot$kmol^{-1}), μ, the mean molecular weight of the atmosphere (approximately 29), and g, the mean acceleration of gravity ($9.76\ m\cdot s^{-2}$)."

28. replace by ... above 200hPa....

Replaced

29. Which database are the ozonesonde data from? how many ozonesonde data have been averaged for that month? did any burst below the tropopause? since the ozonesonde data are used here to validate the choice of ACCO values after screening, more detail is required as to this dataset.

The SHADOZ network dataset was used and has been included in the text. Also the following sentences give more information about the ozonesonde data used in August 2008:

"The comparison of the ozone column above 200 hPa with six ozonesonde stations from SHADOZ network (Ascension, Natal, Nairobi, Kulala Lumpur, Paramaribo, and Hilo) is presented in Fig. 4b. The number of ozonesonde data varies between 1 and 4 ozonesonde launches per station in August 2008. The ozonesonde burst altitude resides within the stratosphere (~30 km), therefore, the above 200 hPa ozone column from the ozonesondes had to be indirectly calculated for these stations. The ozonesonde measurements from the surface up to 200 hPa were integrated and monthly averaged and then they where subtracted from the GOME-2 monthly averaged total ozone measurements, deriving the ozone column above 200 hPa. The difference between the ozonesonde's ACCO and the corrected CCD ACCO is less than 3 DU for these six stations.

"

30.Replace by ... from the ozonesondes....

Replaced

31.You mean, it is better than 2 D.U., right?

The following sentence has been added:

"The difference between the ozonesonde's ACCO and the corrected CCD ACCO is less than 3 DU for these six stations."

32.Remove.

The figure and the discussion about the difference in ACCO between the western Pacific and the Atlantic basin have been removed.

33.Rephrase as: In Fig. 4b the difference .... is plotted.

The figure and the discussion about the difference in ACCO between the western Pacific and the Atlantic basin have been removed.

34.Put these two references in parentheses.

The figure and the discussion about the difference in ACCO between the western Pacific and the Atlantic basin have been removed.

35.In general, I found Figure 4 hard to read and to extract information from. Another point that a reader might be interested in is how much the ozonesonde TOCs agree with the original SCIAMACHY, GOME, GOME2 TOCs, without any "cloud" post-corrections. In general, I suggest a careful read-through this point and adding more information to strengthen the case in hand.

The figure has been replaced. The new figure 4 is divided in 2 sections. a) presents the different ACCOs retrieved for GOME-2 and SCIAMACHY using different cloud top heights and cloud fractions. b) shows the GOME-2 ACCO without any correction applied (blue dashed line) and after the screening out the outliers and correction to 200 hPa level (red line). Additionally, SCIAMACHY data before (yellow dashed line) and after the corrections (green line) are shown. We can see that the ACCOs agree better with each other and with the six ozonesonde stations that have been selected for the validation of the ACCO after adjusting them to 200 hPa and removing the outliers. Additionally, the 1σ standard deviations of the means are becoming smaller.

36.Why is this year used when previously in the text year 2008 was shown?

This year was selected to estimate the uncertainty on tropospheric ozone column because it was the same year selected by Ebojie et al. (2014) for the Limb-Ndir Marching (LNM) technique error estimation. The approach of the uncertainties estimation has been changed in the revised manuscript and the comparison with LNM technique as well. For this reason the years presented for the uncertainty estimation now are 2002 for GOME and 2008 for SCIAMACHY and GOME-2.

37. Valid points. Did you consider the case that the errors are log-normally distributed? why was this case not considered, when it is widely used in other branches of the atmospheric sciences, for e.g. emission inventories?

The Section 4 for the uncertainty estimation section has been re-written.

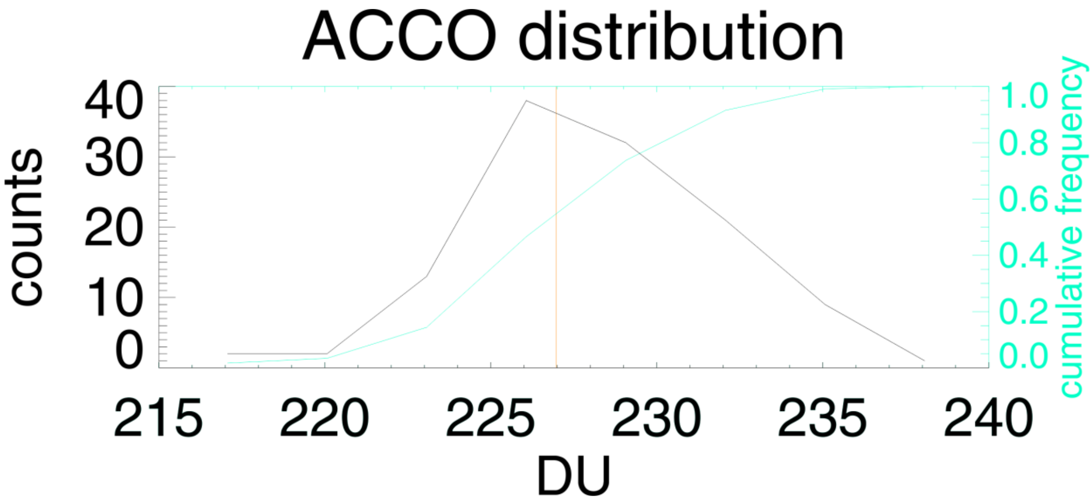

Figure 1. ACCO distribution for the latitude band -3.75 in January 2008

We concluded that the above cloud column (ACCO) in a given month and the total ozone in each grid box follows more or less a Gaussian distribution and therefore the uncertainties are supposed to follow a Gaussian distribution and are uncorrelated (see also above figure).

38. I.e. only systematic errors? and only systematic errors that originate from the algorithm? how about instrumental errors? and how about random errors?

The WFDOAS total ozone column retrieval uncertainty reported by Coldewey-Egbers et al. (2005) in page 1023, is a combination of several error sources, systematic and random. The following text in the manuscript describes the WFDOAS total ozone column retrieval uncertainty:

"The largest contribution in the WFDOAS total ozone column retrieval uncertainty originates from the a-priori errors associated with the use of the ozone climatology and simplifying assumptions made in the derivation of effective parameters (e.g. look up tables for albedo, altitude, and solar zenith angle, other errors like the absorbing aerosol load, the ghost vertical column and the Ring ozone filling) (Coldewey-Egbers et al., 2005)."

39. Is this an actual calculated error? or a statistically extracted value from the comparisons to the ground based data? I skimmed through the ACP Weber et al.,2004, and didn't not find an error propagation section. Please be careful when applying the same word to different meanings: error as in error propagation and calculation is difference from the "monthly mean difference between satellite and ground". Please rephrase this part accordingly.

The uncertainty in the total ozone column retrieval is a statistically extracted value. The overall uncertainty in the total and the above cloud column includes the uncertainty on the retrieved ozone as well as the statistical uncertainty from averaging in a certain grid-box. The uncertainty analysis has been re-written and the following text has been added:

"The individually retrieved total ozone column has an uncertainty of 3% (Coldewey-Egbers et al., 2005). The comparison with ground data show an RMS difference of about 1.5% in the tropics (Weber et al., 2005). Assuming that the precision of satellite and ground data equally contribute to the RMS difference, it results in a precision of WFDOAS total ozone, of about 1% ($X_{TCO\_retrieval}$ ~3 DU is the uncertainty for the WFDOAS TCO and $X_{ACCO\_retrieval}$ ~2.5 DU is the uncertainty for the WFDOAS ACCO retrieval)."

40.Is the zonal variability of 5 DU you show in Fig 2a included in this analysis?

This sentence has been removed from the text since the natural zonal variability of the ACCO is included in the ACCO uncertainty (see earlier answer). The uncertainty in the ACCO combines the retrieval uncertainty and the uncertainty from taking the zonal mean.

41. I am truly concerned with the assumptions made in this calculation. You have mixed up 1-sgma variabilities, variances, uncertainties and so on. All these are not "errors" in the true sense of the term and do not follow the same mathematical behavior as true errors do. You are in effect taking a monthly mean variability from a previous version of the algorithm against ground-based data [and you can imagine how many different error/variances are included in this!] and adding the 1sigma of the TOCs and claiming that this is the error? I am sure you realise that this cannot be so. I suggest a complete overhauling and re-writing of this section.

Section 4 has been completely re-written.

42.Again, these are statistics extracted from spatiotemporal averaging and I fail to see how these are simply called error later on.

The term error is not used in the manuscript, we now use the term uncertainty.

43.I believe that more detail of how exactly the differences between the sensors have been smoothed out is needed. Since the paper aims to provide a continuous 16 year records based on three instruments with different footprints, different instrumental degradation, different global coverage, data screening, to name but a few, a lot more detail on these parameters is needed.

The three datasets have not been harmonized yet. The scope of this paper is to present the CCD tropospheric ozone results for the individual instruments (1996-2002 for GOME, 2003-2007 for SCIAMACHY and 2008-2012 from GOME-2A) and their comparison with the SHADOZ ozone data. The harmonization of the three separate datasets with the overlapping periods into one consistent, will be the subject of another paper, along with the trend analysis of the harmonized time-series. Nevertheless, a comparison between the CCD results for the common years of GOME/SCIAMACHY (2003) and SCIAMACHY/GOME-2 (2008) has been included here.

44.You are showing one year for GOME and another for SCIA/GOME2. Shouldn't this "small detail" be discussed? if you assume that these features are so constant throughout the years, then you can simply show a multi-annual average for each month and instrument. Then you might be able to compare them. As is, I fail to see the point of Fig.6. You also do not discuss at all the rather obvious differences between these plots.

Three new plots have been included in the results section, presenting the TTCOs for the overlapping years of GOME, SCIAMACHY and GOME-2 and the differences and similarities between them are discussed. Additionally, a figure of the seasonal TTCOs is added before them and the comparison with other results from literature is discussed.

45. And how are you planning to deal with the overlaying years, i.e. years that there are data from two instruments?

We are planning to harmonize the three separate datasets into one consistent in the future using correction factors for GOME and GOME-2 having as reference the SCIAMACHY CCD results.

46. This bit reads like general discussion and general knowledge of tropospheric ozone behavior and NOT discussion of results. Move to the introduction and use the space to discuss your results.

This part has been shifted to the Introduction. The presentation of results now contains the discussion of the seasonal variation of the TTCOs retrieved with the CCD_IUP algorithm and the comparison for the common years (2002 and 2008) between the instruments.

47. So approximately, 500 by 500 km in spatial search radius. How about a temporal constraint? especially considering your convective cell discussion below?

After the comments and from other reviewers, the comparison of CCD TTCOs with ozonesondes using a 5^ox5^o box around the sonde station has been replaced by the comparison with the closest gridbox that the sonde station belongs in the regular 2.5^ox5^o TTCO gridded data.

In order to have sufficient sample of ACCO the algorithm is based upon monthly mean ACCO (as monthly mean total ozone in each gridbox). For this reason, only monthly mean comparisons between TTCOs and ozonesondes were made.

48. Considering the limitations of these comparisons, as you well note, how do these values change if you pick say a different spatiotemporal criterion for averaging per month? that would give a clear indication as to the robustness of your method.

We made the statistics (mean, standard deviation, RMS, relative differences and R) for the case of a 5^ox5^o box around the sonde station and for the case of the nearest gridbox of the 2.5^ox5^o TTCO gridded data that the sonde belongs and we concluded that the statistics do not change dramatically.

49. Are you implying that when no seasonal cycle is present one cannot have good correlations between two different datasets?

The correlation requires some variation. When there is no seasonal cycle present, the variation is random and results in low correlation. Therefore we have used other statistical tools like bias, relative difference and RMS to conclude if two time series agree well or not with each other.

50. I fail to see why are you making this point. Are you saying that when El Nino is strong, your technique cannot capture the trop ozone correctly?

No, this is just an observation that we see an ozone enhancement at Java station during El Nino.

51. Again, are you claiming that when stratospheric intrusions occur the ozonesondes and or/ the satellite cannot retrieve Trop Ozone correctly?

This part has been changed. The stratospheric intrusions are mentioned in order to explain the enhanced ozone concentrations at winter and not to justify any failure of the retrieval. Due to the movement of the ITCZ at southern latitudes in NH winter months, it is difficult to retrieve ACCO and consequently the TTCO.

52. What do you mean by errors? is this value an outcome of error analysis within the Ebojie et al., 2014, algorithm or is this the 1-sigma of the monthly mean values? if so, what are the bins used? what are the collocations used? to be correct in this work, you should actually use 1-sigma/sqrt(number of points)

This is the outcome from the Ebojie et al. (2014) error analysis (page 2090). For the comparison between the 2 methods, the CCD data have been gridded with the same grid as the one used in LNM.

53. Is this 10% for all years, all sensors, all seasons? it seems like such an ad hoc choice....

This part has been changed and the comparison with sondes have been removed. The two datasets are now compared directly with each other in Figs. 15 and 16.

In order to make the LNM columns comparable to the CCD TTCOs, we adjust the LNM columns to the 200 hPa level using climatological values from the Fortuin et al. (1998) climatology. Therefore, we subtract the ozone between the tropopause and the 200 hPa.

54. How much noisier? Numbers?

The RMS is less than 10 DU and in the LNM dataset we notice elevated ozone columns (40 DU) appearing e.g. over the Pacific Ocean where the neighboring grid-boxes are around 15 DU.

55. In this section, based on the title, we were expecting to see comparisons and correlations between the LNM and CCD methods. Here are are shown comparisons between LNM and ozonesondes and CCD and ozonesondes. This section has a lot of information missing. Are the LNM and CCD data binned in the same way? are the monthly mean variabilities the same order of magnitude? why is CCD a better method than LNM?

As answered before, the data are binned in the same 2.5x5 bins and have been adjusted using climatology to the 200 hPa level for a better comparison. The monthly values now are in better agreement but still there is a consistent underestimation of LNM due to the overestimation of limb ozone v2.9 in the tropics. The LMN data seem to have a consistent lack of data over the Pacific Ocean due to cloud contamination.

56. Shouldn't this be updated with the actual ACP paper?

Done

57. This is not the error in the ozone retrieval. This is the statistical uncertainty due to spatiotemporal averaging.

Everywhere in the text only uncertainties are mentioned.

58. This is typically given in %, I recommend you do the same.

Changed to %

59. As above.

Changed to %

60. Replace by: The final ACCO values are plotted with red lines....

The figure has changed

61. Replace by: The stratospheric columns are shown in black dots.....

Done

62. How about adding the standard deviation of the mean in the plot. That would give an indication of the variability of each datasets.

The standard deviation has been added on the plots.