Reply on RC1
Gérard Ancellet et al.

Author comment on "Homogenization of the Observatoire de Haute Provence electrochemical concentration cell (ECC) ozonesonde data record: comparison with lidar and satellite observations" by Gérard Ancellet et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-7-AC1, 2022

Response to Reviewer 1

This document is the list of our responses to the reviewer's comments and a revised version of the text is also attached to this response to show the changes in red and the deleted sentences using strikethrough text

This is one of the first papers describing the results of homogenizing a historic ozone sonde record by applying the corrections suggested by the OzoneSonde Data Quality Activity (Smit et al., 2012; Smit and Thompson, 2021). The paper is generally well written and scientifically solid. The English could benefit significantly from copy editing. My major suggestions are to improve the Figures, and add additional important information. After that, the paper is well suited for publication in AMT.

We warmly thank the reviewer for his/her suggestions and comments. We have modified figures 6 and 9 to 11 to add new data requested by the reviewer or to facilitate the comparison of the different information initially provided. A careful copy editing of English writing has been made. The revised version of the text attached to this response shows the changes in red and the deleted sentences using strikethrough text

Major suggestions:

The tight < 6 hours coincidence criterium for matching sonde ascents and lidar data results in only about 40 matches ( sondes launched at night, ~4% of all sondes), out of more than thousand sondes (most of which are launched during daytime). I suggest to also allow < 12 hours coincidence, which will match many more sondes with the nighttime lidar measurements. It would be very interesting to see if this changes the results and
statistics presented in Figs. 5 and 6.

We agree that above 20 km the comparison of lidar and ECC measurements is still relevant with a time shift of 12h instead of the 6 h initially used. Such a comparison has been added in Fig. 6 for the period corresponding to the comparison with the MLS measurements. The number of coincidences increases from 40 to 366. Although the difference between uncorrected and homogenized ECC is now very significant, the bias between LIO3St lidar measurements and homogenized ECC remains of the same order of magnitude at altitudes greater than 20 km. The following sentence has been also added in section 3. Line 137: « For the sake of a more complete discussion of the two types of comparisons made in the stratosphere, we also considered a lidar data set of 366 profiles from 2005 to 2021 with less restrictive measurement time difference with the ECC launches (<12h). Such a criterion is valid as long as the rapid O3 variations typically encountered below 18 km are not included. »

Are geopotential or geometric altitudes used in the sonde vs lidar comparison? This needs to be clarified. If geopotential altitudes are now used for the sondes, switching to geometric altitudes (as used by the lidar) should improve the comparison above about 25 km.

We thank the reviewer for this insightful suggestion. This decreases differences between the ECC and LIO3St or MLS above 26 km (from -10% to -8%). All comparisons shown in Figure 6 have been recalculated with geometric altitude. The following sentence has been added in section 4.2 line 208: "The means of the relative difference between ECC and LIO3St are then calculated for 8 vertical layers between 14 and 30 km using the geometric altitude for the ECC sondes as geopotential altitudes become significantly greater than lidar geometrical altitudes above 25 km.”

In addition to showing the mean differences sonde minus lidars in Figs. 5 and 6, the authors should also show the mean difference sonde minus MLS (R. Stauffer is one of the co-authors, and should be able to provide that quite easily). This would be important to compare with the stratospheric sonde-lidar difference. It might help to elucidate the significant ~5% difference seen in the stratosphere. From Fig. 7 it looks like the ECChomogenization is moving the sondes in the right direction, and that there is no significant difference between homogenized sondes and MLS around 20 hPa (~26 km). This is different from Fig. 6, where the homogenization seems to move the sondes in the wrong direction.

We agree and ECC-MLS comparisons are now also shown in Fig. 6b. Indeed the ECC-MLS and ECC-LIO3St do not have the same sign after the homogenization. The following paragraphs are now included in section 4.2 line 217:

« For the period 2005-2021 and using a time difference less than 12 hours, the negative bias between the homogenized ECC and the lidar decreases down to -2% between 22 and 24 km, but remains as large as -7% above 28 km (Fig.6b). Note also that the mean uncorrected ECC and lidar difference is now slightly positive (+1%) for the 2005-2021 period in good agreement with the N2 negative trend shown in Fig.4. Below 18 km, the ~4% negative bias between homogenized ECC and lidar (Fig.6b) should be interpreted by
possible significant concentration changes within 12 hours in this altitude range. »

and section 4.3 line 258

« The fact that the average ECC-MLS difference shown in Fig. 6b is slightly positive (+2%) in the 22-26 km altitude range, while the average ECC-LiO3St difference is slightly negative (-2%) means that homogenization is a good compromise for intercomparability with other techniques measuring O3 in the stratosphere below 26 km. Above 26 km, both comparisons indicate a negative bias in homogenized ECC O3 concentrations of less than -6% »

Figures 5 and 6 should be combined into one Figure. The sonde - MLS differences could be included in that Figure as well

The instruments used to plot figures 5 and 6 respectively in the troposphere and stratosphere are very different and the impact of ECC homogenization is not the same for these two regions. We therefore prefer to keep 2 separate figures for these 2 discussions.

Figure 9 needs to be improved. The different axes make comparison of ECC and surface data difficult. It might be better to include the surface data directly in the ECC

Figure 9 (now 10) is modified in order to be able to superimpose on the same temporal evolution of the annual averages both the measurements of the ECC and those of the OHP surface O3 analyzer.

Figure 10 should also show the MLS time series. That would be very helpful.

The MLS data record only starts in 2005 and is not as relevant as the lidar data record starting in 1990 for the evaluation of homogenization on the sign of the O3 trends or its magnitude in the stratosphere. The differences between uncorrected and homogenized ECC are not large enough after 2005 to significantly alter the long-term trend in O3. We believe that only the analysis of the temporal evolution of the difference between ECC and MLS, shown in Figure 8, is relevant to this paper.

Figure10 it is not necessary to plot all the error bars, since they are all very similar. Instead, I think it would be much better to show all respective time series (ECC old, ECC homogenized, Lidar, MLS) in one plot.

We agree with the reviewer. The annual standard deviation does not change a lot and all the time series can be shown in a single plot for a given altitude range. Fig. 10 and 11 (now 11 and 12) have been modified accordingly (not including MLS as explained above).
While the overall trends and their comparison is useful, it would also be quite important to look at time series of ECC minus surface, ECC minus lidar(s), and ECC minus MLS. Are there significant trends in these difference time series? Is there a significant annual cycle in these differences? How do trends (and possibly remaining annual cycles) change with the homogenized ECC data? These difference time series probably do not require subtraction of an annual cycle. Since common variations largely cancel out, trend uncertainty should be smaller than when comparing trends of the individual monthly or annual anomalies.

We understand the reviewer’s point of view. Such an analysis is complementary to, but different from, the trend analysis presented in Figures 11 and 12. The purpose of our trend analysis was to use all weekly ECC observations to assess how sensitive is a linear ozone trend to homogenization of the ECC. For our study the temporal evolution of the differences between two instruments is only relevant when using observations made on the same day. Such a discussion was already present in section 4.3 for the MLS measurements shown in Fig. 7 (now 8) because the time lag with the ECC measurements is less than 1 day and there are many MLS occurrences. This is not always the case for comparisons with lidar (only 366 profiles out of 1412). Nevertheless, the temporal evolution of the difference between ECC and lidar between 2005 and 2021 is now included in Figs. 6c,d and this temporal evolution is compared with the MLS time series already discussed in section 4.3. The following paragraphs are now included in Section 4.2 line 222:

« The time evolution of the relative difference of ECC and LiO3st ozone concentrations is shown in Fig 6c and Fig. 6d for uncorrected and homogenized ECC, respectively. Many of the differences between uncorrected ECC and LiO3St are greater than +6% between 2007 and 2016 while there are some negative differences approaching -6% in 2006. Homogenization improves the relative differences now remaining between -5% and +5\%, except in 2006 when the negative bias decreases down to values smaller than -6%. »

and in section 4.3 line 255

« An interesting feature of this MLS/ECC comparison is the interannual variability of the differences. It can be observed that differences using homogenized ECC data are more evenly distributed around zero. The same conclusion could be drawn from the time evolution of the relative differences between homogenized ECC and LiO3St presented in Fig. 6c,d. »

For all trend uncertainties: Is autocorrelation of the residuals accounted for? Please state that, and preferably account for it.

As explained in section 3, only basic linear trends of the ozone concentrations corrected for the mean seasonal variation at OHP is considered in this study for the assessment of the homogenization. The following sentence as been added in section 3, line 153 : « The trend uncertainties are calculated using the 95% confidence limit of the slope of the linear regression assuming that the residuals are not correlated for weekly (ECC) or 2/3 per week (lidar) observations. »
satellite total ozone, as well as MLS ozone profiles: It would be important to compare the OHP results / biases with those seen at other sonde stations. Since R. Stauffer is a coauthor, and has most of these data, an additional paragraph, or even additional Figures would be very important. This is needed to put the OHP results into the necessary wider context.

**The purpose of this work is to focus on the reanalysis of OHP because we can use both lidar and satellite observations for this site. Furthermore, the corrections are very site-dependent and the conclusions drawn from the use of O3S-DQA at OHP are not easily applicable to other sites.**

Nevertheless, the following paragraph has been added in the conclusion line 359 to compare our results with the homogenization performed in the SHADOZ network and for the Uccle/De Bilt sites in Europe: « While the objective of this paper is to discuss the impact of homogenization on the OHP dataset using lidar and satellite measurements, it is worth checking how such corrections have improved data quality at other sites. The impact of the homogenization is dependent on the site, because different homogenization steps have to be applied at different stations. In general, the additional corrections for the pump temperature will give higher ozone partial pressure amounts in the stratosphere. On the other hand, applying a constant background current subtraction instead of a pressure dependent background current and applying the transfer functions from EnSci-SST 1% will lead to lower ozone partial pressure values above 10 km. Witte et al. (2017) performed an extensive analysis of 7 SHADOZ network stations in the tropics, showing that the mean differences between ECC and MLS are reduced from -11.2±13.6% to -3.0 ±10% at 40 hPa (22 km) and from -3.2%±4% to -0.7±3.1% at 17 hPa (28 km). In Europe, Van Malderen et al. (2016) observed that the O3S-DQA corrections actually give higher (+1%) and lower (-2%) ozone concentrations in the stratosphere with respect to standard processing for the Uccle 1997-2014 and De Bilt 1993-2014 ECC observations, respectively. This is mainly due to the fact that the pump temperature correction was a major correction for Uccle, while changing the background current correction has a major effect for De Bilt. O3S-DQA corrections reduce the relative $O_3$ difference between Uccle and De Bilt in the lower stratosphere. The analysis of homogenized ECC at OHP using LiO3St or MLS show similar improvements in the stratosphere below 26 km. The remaining bias of -2% to -3.7% between homogenized ECC and other techniques measuring $O_3$ in the stratosphere at OHP is also in the range of the remaining negative differences between homogenized ECC and MLS observed in the 22 to 28 km altitude range by 4 stations of the SHADOZ network (Witte et al., 2017).»

Please also note the supplement to this comment: [https://amt.copernicus.org/preprints/amt-2022-7/amt-2022-7-AC1-supplement.pdf](https://amt.copernicus.org/preprints/amt-2022-7/amt-2022-7-AC1-supplement.pdf)