Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC1
https://doi.org/10.5194/acp-2021-122-RC1, 2021

Interesting, but needs clarification of uncertainty measures

Anonymous Referee #1

Referee comment on "On the use of satellite observations to fill gaps in the Halley station total ozone record" by Lily N. Zhang et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-122-RC1, 2021

Overall Comments

The manuscript describes the use of satellite data of total ozone to fill gaps in the ground-based Dobson total ozone record at Halley Bay, Antarctica.
As mentioned in the text, Halley Bay has one of the longest and most important total ozone records. This record was the key for the dectction of the Antarctic ozone hole.
It is a good idea, and scientifically sound, to fill gaps in this important record with satellite data, as well as check for consistency. Overall, the paper is well written and merits publication in ACP.

Before publication, however, I suggest a few important clarifications. Thoroughout the manuscript, I get confused about the use of individual total ozone measurements, daily averages and difference, monthly averages and differences, and the corresponding standard deviations. Sometimes standard deaviations appear to be mis-named "averages" as well. The description of the applied method to shift satellite data towards the Dobson data is also quite long-winded. It would benefit from shortening and clarification. There is no need to make a simple average bias correction appear much more complicated than it is.

Specific Comments

line 21, average of 2 Dobson Units: I don't think this is what is meant. My understanding is the each satellite record is shifted by the Δ from Fig. 3, so that it matches the Dobson data on average. Therefore the satellite average should reproduce the Dobson average exactly, by construction. What is probably meant here is "within a standard deviation of 2 Dobson units". Even more information is required here: are the (presumably) 2 DU standard deviation for monthly means or for daily means? Is the given value 1 or 2 standard deviations? Is it +-1 DU or +-2 DU? Is it even correct? In lines 184 and 185 the stated standard deviation of the differences is 6 to 7 DU. This is much larger than 2 and neesd to be checked.

From text and Table 1 it appears that the "root mean square difference" (which is the same as the standard deviation!) for daily average data is about 12 DU. So the 2 DU are probably for the monthly average data, but the 12 DU for the daily data should be mentioned here as well. (Assuming a Gaussian distribution, 67% of the data should be

within +-1 standard deviation of the mean (which should be zero here by construction), 95% of the data within +-2 standard deviations, ...

While I call this lack of clarity out here for line 21, it exists throughout the text, and needs to be fixed everywhere.

line 39: Here it says "throughout the year", line 37 said that no data are available for May to July. What is true now?

line 45: delete "the" before "satellite"?

line 50: replace "well tested" by "in place"?

Fig. 1: are those all measurements or daily averages? Please mention. Are the satellite data the original data from all satellites, or the adjusted data matching the Dobson?

Lines 79 to 104: Would be good to also give the size of the satellite ground pixels near Halley Bay for all the satellite instruments. In addition, I think it absolutely necessary to state which data version was used for each satellite, and where / from which URL the satellite data came from. For GOME2, for example, there are data from Uni-Bremen, from DLR / EUMETSAT, from RAL / ESA_CCI, ... A table of URLs and versions would help here.

Line 85: "cross-calibrated" My understanding is that the current SBUV 8.6 version is not cross-calibrated between satellites, but relies on improved calibration at the radiance level for each satellite. Please check. Natalya Kramarova will know.

Line 103: should be "polarization effects"

Line 107: How were overpasses defined? Satellite foot-point within what distance? Same for all satellites?

line 127: Would it not be better to have Figure 3 and lines 155 to 160 right here in section 2.3? After all, the Figure shows the Δ-s that are discussed in lines 120 to 127?

Line 128, Section 2.4: Would it not be clearer, to have section 2.2 here, after section 2.3. That way, you would have a more logical flow. a.) discuss Δ-s for individual satellites b.) discuss how you use all satellites to fill in, and how that looks for the different months.

Table 1: What is shown here? Differences between monthly averages, or differences between daily averages? From the numbers, around 12 DU, it looks like it was daily averageds. Were the satellite data Δ-adjusted or not? In April, that would make a large difference according to Fig. 3.

Line 143: Would the Δ-adjustment not take care of the Bass-Paur difference as well? Is it necessary to mention systematic biases here, since the filling-in method takes care of them anyways?

Figure 4: Having Figure 4 so close to Figure 3 confused me (Are they now using monthly Δ-s again? Or daily? Or what?). I guess the only point of Figure 4 is to show that 2019 was very different from the other years. This does not become very clear here. The stars for 2019 are easy to miss in the Figure, and they do not have error bars. It would be helpful to have a clearer Figure, that points out 2019 in a legend in the Figure, not just in the caption.

Also Figure 4: What are the error bars? Standard deviation of daily data or monthly data? Standard error of the mean? One or two standad deviations?

Lines 169 to 187, and Figure 5: I am confused. Does Fig. 5 show data, where 2003 to 2012 was the training period? Or which training periods were used to generate the data in the two panels of Fig. 5?

Lines 184, 185: I assume that the numbers are for the trained data? Please state the same numbers for the unadjusted satellite data. Only then you can conclude if the adjusted date are better, or not. Check consistency with numbers in abstract and conclusions!!

Figure 6: The differences between the Dobson monthly means and the Δ-adjusted satellite data look rather large in 2019 and 2020, 5 to 10 DU. Is that consistent with the numbers given in lines 184, 185? Figure 4 shows that the 2019 Dobson data are flawed. Are the 2020 Dobson data flawed as well? Flaws of the Dobson data should be stated, and maybe even marked with different sysmbols in the Figure. How do the Δ-adjusted satellite data look in the other years? It would be good to plot the entire red time series.

Line 230: Are the 2 Dobson Units the average difference? Is that really relevant? In principle, the average difference should be zero, due to the Δ-adjustment. Of course zero is not realized in every subset / realization of the data. Is not the standard deviation between Dobson and Δ-adjusted satellite data a much more meaningful quantity, to show how well the two data sets agree?

Also line 230: Check consistency with the numbers in abstract and in lines 184, 185. Please give (also) the standard deviations of Dobson minus Δ-adjusted satellite data on the basis of monthly and daily means.