

Atmos. Chem. Phys. Discuss., referee comment RC2
<https://doi.org/10.5194/acp-2021-122-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2021-122

Anonymous Referee #2

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-RC2>, 2021

On the Use of Satellite Observations to Fill Gaps in the Halley Station Total Ozone Record

The authors use observations from multiple satellite instruments to create an ozone column dataset above the historically and scientifically significant Halley station in Antarctica, using the Dobson instrument as a calibration anchor. The result is a dataset that can fill in gaps due to a recent ice crack, check for calibration issues affecting the new Dobson data, and fill in gaps caused by other future geophysical or social disruptions.

General comments:

Overall, I enjoyed reading this work. The manuscript is well written and organized. Gaps in scientifically significant long-term datasets – such as the Halley Station's Dobson – are an important problem. The paper contributes usefully to this topic. I have a few questions and requests for clarification.

The satellite instruments used in the study have a variety of measurement techniques, advantages and limitations. While it might not be necessary to do a deep dive into this, I think it is at least necessary to comment on the different spatial resolutions (and perhaps vertical sensitivities) and the spatial coincidence criteria used to define co-location with Halley Station.

Figure 1 is a key illustration of the datasets involved in this study – i.e., the adjusted average satellite and Halley Station Dobson. The figure is limited to 2013 – 2019, I assume because this allows features on the ~monthly timescale to be seen and because 2013 – 2015 was one of the time windows used to test the technique's ability to reproduce the Dobson. And 2017 – 2018 was the motivating gap in the Dobson's

timeseries. Nonetheless, I would like to see the full timeseries of the two datasets. Figure 6 does some of this, but I believe it only inserts the satellite dataset into the Dobson gaps (?). That's useful, but I'd also like to see both fully plotted to get a sense of how closely they agree.

I am not entirely comfortable with the conclusion that the adjusted satellite average reproduces the Halley Station Dobson to within about 2 DU. I don't think it was sufficiently stated that the satellite reproduction of the Halley Dobson data varies seasonally. In Figure 5, some months show very close agreement; others show much larger differences, e.g., ± 15 DU. 2 DU is the apparent result of averaging large positive differences and negative differences across the year. In addition, while there are similar annual patterns, there are also notable differences between the years shown.

In addition, the time periods chosen for the test have a particular combination of satellite instruments that are being used to compare with the Halley Station Dobson. What were the results of the comparison between the adjusted satellite average and the Dobson for wider time periods? How well do these test cases, 1998 – 2002 and 2013 – 2015, generally represent the physical conditions and satellite datasets available to fill other gaps and future gaps?

The approach described in the paper, and the resulting dataset, is useful, but I think care needs to be taken in asserting the accuracy in reproducing the Dobson shown here – especially since particular months of the year are often of more significant interest to the study of ozone chemistry than annual averages.

Early in the paper (section 2.2) it is stated that both absolute and relative differences are computed. The paper then focuses on the absolute differences. Given the annual pattern to the absolute differences, I'm curious to know what the results of the relative differences were? Do percent differences show the same seasonality as the absolute differences? Would constructing a delta adjustment on the basis of a relative difference or SZA remove some of the annual pattern in the difference?

Figures:

It would be helpful to expand the width of the figures to fill the width of the page. Figure 1, for example, would benefit from this since it can be difficult to see the structure in the data and the comparison between the Dobson and satellite measurements. This is important for understanding the work being described. Figure 3 as well.

Figure 4: why is there no monthly value for August (month 8) 2019? Caption title could add that 2019 is with the automated Dobson so that this context stands alone in the figure without the text. I'm assuming the error bars are the standard error average? Is the

August error bar larger because there are fewer days being used (polar night limitation)?

Figure 6: has tick labels that are too small to easily read; they are notably smaller than other figures.

Tables:

Table 1 is awkwardly split across page 5, which has the caption and titles, and page 6, which has all the values. Please note this to the typesetter and check that the proofs correct this.

Could be informative to add a column for the delta-corrected satellite average and/or the delta adjustment.

Abstract:

In a few cases, it might be helpful to the reader to add specifics. For example,

“by... adjusting overpass data” – adjusting how?

“Tests suggest that our method...” – what tests? Or rephrase to say “comparisons to ____ suggest that our method...”

“... our approach improves on the overall performance...” – what does overall performance refer to? Accuracy of the measurement? Comparison results? Completeness? Also want to be careful because the goal of the paper, as stated at P3L57-58, is not a high-performance dataset / “most accurate” dataset, but “to reproduce what the Dobson instrument would have observed”.

“... there was a significant difference between the two.” – I’d suggest being more quantitative. What was the difference? This also brings up a question of what the authors consider to be a threshold for good agreement?

Specific comments:

P2L1: A good paragraph starting sentence. But, suggest rephrasing "now", since the interruption being discussed was a few years ago.

P2L49: (warning: pedantic point) Most readers will understand what is meant by "With the advanced multi-satellite observing system now..." but the wording here might suggest that the current (and past) suite of satellites are part of a coherent and coordinated "system". We have a great scientific community and missions are chosen to provide complementary coverage. But I'm not entirely convinced the satellites used in this study, from different agencies, eras, and designs, are "an observing system".

P4L79: missing parenthesis before the Munroe citation.

P4L76: "...some also have spectral information at other wavelengths." – what instruments and what other wavelengths?

P4L80: Suggest re-wording "spin off of the" – it isn't clear what is meant by "spin off".

P4L81: "... of a somewhat improved..." – "somewhat" is hard to interpret. What was improved (or not) that is notable here?

P4L79:82: why put this information about GOME here when the details of this instrument are four paragraphs below? Doesn't serve the flow.

P4L83: SBUV acronym used before defined. Need to carefully define, given it is an instrument and set of instruments.

P5L111: It is important to be clear what has been done to define the spatial coincidence criteria for the satellites and Halley Station.

P5L117: "... due to unusually high differences..." – what were the differences.

P5: Section 2.3: You state that the delta value for each DOY is averaged across all years in each satellite series. Was there any trend or interannual variability in the differences before combining all years?

P6L149: Why is April different?

P8: The observed differences are typically larger than the average delta, which is reduced by large positive and negative values averaging out. If this pattern is due to SZA-dependence, would a bias correction based on (or including) SZA rather than DOY produce a useful result?

P8: If there was an unusually high or low level of ozone abundance, will the absolute delta correction sufficiently reproduce the Dobson? Would a relative (%) adjustment scale be better?

P9L190: suggest adding a comment about calculating averages for the months that border polar night, where there will be a reduced number of days contributing. I don't know offhand when Halley Station's latitude enters/exits polar night. That might be worth mentioning somewhere since it is relevant to data collection and chemistry.

P12L223: "Larger difference" – how large? Since this is an important finding, it would be good to state this. Hopefully this prompts investigation into why there might be differences between the measurements.

P13L227: were any of the satellites used in this study validated using the data collected at Halley by the Dobson?

P13: There was quite a bit of discussion about the DOY data but the conclusions focused on the monthly averages. Why not include a plot of the DOY comparison results as well and comment on their comparability to the monthly results?