Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2020-324-RC1, 2020

Atmospheric Measurement Techniques Discussions

Interactive comment on "The world Brewer reference triad – updated performance assessment and new double triad" *by* Xiaoyi Zhao et al.

Anonymous Referee #1

Received and published: 5 October 2020

This is the first review of the paper "The world Brewer reference triad – updated performance assessment and new double triad by Xiaoyi Zhao et al. This paper is of great importance for the WMO Brewer network as it discusses the stability of the world Brewer triad maintained by the ECCC, Canada. Comparisons between the single Brewer triad (BrT) and the double Brewer triad (BrT-D) are reported for the 1999-2019 period. The previous assessment of the BrT performance (Fioletov et al., 2005) is used to verify the stability of the reference instruments over an extended period (1984-2019). Four statistical methods to evaluate the uncertainty of each instrument relative to the BrT and BrT-D baseline, to the independent reference observations (Pandora and eleven satellite records), and to the reanalyses (MERRA-2) are presented and summarized in

plots and tables.

The paper is well written, the figures used to demonstrate the analyses are clear. The summary tables support the discussion and allow us to evaluate stability and random uncertainties of the total ozone observation originating from uncertainties in the extra-terrestrial constant (ETC) and the effective absorption cross-section coefficients specific to each instrument in the triad. There are a couple of inconsistencies in the analyses, including grouping of the data in either monthly, 3-months, or 6-months averages. It is not clear why the time periods for averages are changing depending on the analyses. It would make sense to present all data as monthly averages. The 2005 paper analyzed data starting in 1984. Why does this paper exclude the 1984-1988 period? Since the triad is independently calibrated at Mauna-Loa observatory, where the station Dobson (since 1957) is located, why not to perform comparisons for data collected by triad at MLO? The traveling Brewer reference is used to calibrate station instruments. It would be good to include its record with respect to BrT in this paper. Here are specific comments: 1) line 68. The text "230 Brewer instruments deployed" is in contradiction with the abstract where 230 instruments are referred to as "produced". Were all produced instruments deployed? 2) Line 70-71. The paper states that 123 instruments are currently in operations and are located at 88 stations. How many countries use Brewer instruments for ozone monitoring? Are there Brewers that are not part of the WMO GAW network and do not submit data to WOUDC for archiving? 3) Lines 73-74. Do I understand correctly that effective ozone cross-section is determined once after the instruments are produced? Are there in-field instrument adjustments that can change the instrument-specific absorption cross-section, overtime, or abruptly? Is there a method to check the stability of the ozone cross-sections? Is it done when the instrument is calibrated at MLO? 4) Line 80. "reference instrument . . .is independently calibrated every 2-6 years". What are the WMO GAW requirements for the frequency of calibrations of the triad? Is it consistent with the requirement for the in-field instrument calibrations? Table 1 shows that some instruments were not calibrated for 6 years. Would this affect the triad stability? What is the requirement for the

traveling standard calibration? 5) Line 109, Table 1, right column, row 6 – “Significantly less instrumental stray light than in single instrument” – please quantify what it means, include information about the level of rejection of the stray light, i.e. 10ˆ-4, 10ˆ-5 in the wings? Is stray light here attributed to the out-of-of band light? How much does it contribute to the total column ozone error at representative air mass over Toronto? 6) Line 145. The period of evaluation includes 2019 which is after the BrT was moved to a different location in 2018. Why not exclude 2018-2019? 7) Line 170 – “seasonal mean” – is it the same 3-month averages that are discussed later (line 432)? 8) Line 181, another mentioning of the “good stray-light control”. Please be more specific. In Zhao et al. (2016) “good” is defined as low AMF dependence up to 81.6 degrees SZA, or within 1% up to AMF=5.5 9) Lines 197-198, “bi-weekly” means two weeks? Are you referring to the fact that the SBUV total ozone data are selected within the box centered on the station location, +/2 degrees in latitude and +/-20 degrees in longitude, and then distance weighted to create the station overpass? What is the uncertainty of SBUV total ozone overpass over Toronto? When comparing to satellite overpass data, do you use the satellite data uncertainty in the estimate of the agreement with Brewers? 10) Line 200, the reference to “+/- 1%” is one or 2 standard deviation? This number is based on the monthly averaged comparisons. How does it compare to the results in Table 5 where one standard deviation is provided based on 3-month averaged data? 11) Lines 251-232. Please explain why the instrument with more points would not dominate the forming of the baseline. Is it in reference to the previous method where three Brewers are used to establish a baseline? In the 3d party method, the baseline is derived for each instrument separately, therefore the 3d party instrument represents the “baseline”? 12) Line 268. In this method, B and C are shared between the instruments. In case one of the instruments have a stray light contribution that is larger than in the other two instruments, would it create the offset in the B and C coefficients? Is there a weighting method used to determine these coefficients? 13) Line 288. Would the effective absorption cross-section value change with the solar zenith angle due to the presence of the stray light? Do you restrict data comparisons to SZAs that have

limited impact of the stray light? 14) Line 319 – “only good quality satellite data are used in the analyses”. What are these criteria? Please discuss the QA criteria (flags) in section 2.4. 15) Line 341. What was the reason to select 3-months averages for the presentation of results? 16) Line 357-358. What is the 3-month mean TO and mean air mass in Toronto in each season? Is it comparable to 330 DU and mu=2? 17) Lines 364-369. Figure 2 suggests a drift in Brewer #14 between 1999-2004 and in Brewer #8 between 2007-2013. Were the drifts corrected in the data archived in WOUDC? According to Table 2, Brewer # 14 was calibrated in 2000, 2005, 2008, and 2013. If the drift is detected between independent calibrations, is there a method to post correct the data prior to the latest calibration reference? Brewer #145 shows a large spike in both errors with the opposite sign. What caused it? 18) Line 383, “although empirical correction methods have been applied, the residual effect still exists”. Figure 3 shows sudden changes in biases in 2016 and 2017, winter season. Does it have anything to do with this Brewer calibration in 2015? Can you explain instrument issues in this section while discussing Figure 3? 19) Line 412, Was eq(3) use to derive errors in the ETC and ozone absorption cross-sections? Was total ozone from Pandora used for this assessment? 20) Line 426, When issues with ozone absorption cross-section for Brewer 145 are discussed, what period if referred to? It is not clear from Figure 5(d) 21) Line 453, Table 5 results need to be discussed in greater detail. For example, if all comparison periods are included in the assessment of the BrT’s errors relative to satellite overpass data, the mean bias increases to 0.625 %, which is larger than the BrT-D bias. Another interesting fact is that OMITO3 shows the largest bias from both BrT and BrT-D, whereas OMDOAS bias is much lower. TROPOMI bias is negative wrt BrT-D, and it is almost of the same magnitude as of the OMDOAO3, but of the opposite sign. Are the OMDOAO3 and TROPOMI biases related to TRPOMI higher spatial resolution or their respective ozone absorption cross-sections? There seems to be a difference in relative biases for BrT and BrT-D, where BrT-D is often higher (although the difference is not statistically significant). Is there any reason for this? It would be of interest to know of each Brewer calibration results and how much the calibration was changed.

22) Line 463, " the same long-term stability of the Brewer reference instruments when compared with Pandora or satellite instruments". Figure 7 indicates that the mean bias (by eye) of Brewers in 1999-2004 is near 0% relative to MERRA-2, then it changes to $\sim$2% in 2005 (MERRA-2 change?). The bias in 2005-2015 shows a slow $\sim$ 1% drift. There is a step-change in 2015 and then it rises to $\sim$1% in 2017. Brewer #15 is the lowest in 2006. Brewer #08 is the lowest in 2017-2018 Are all these differences related to the MERRA-2 changes of assimilated data? 23) Line 476, after October 2004 instead of 2014? 24) Line 505-508. The issue with strong temperature dependence in Brewer #15 is discussed. The optical frame was fixed in 2017. Was the data prior to 2017 corrected? 25) Lines 508-510. Wavelength drift in Brewer #145 is discussed. It would make sense to mention instrumental issues while discussing results in Figures 3 and 6. 26) Line 513 – Hardware replacement issues, ie. ND filter and mercury bulb. What is the recommendation to the BrT and BrT-D data reprocessing? Please make sure to refer here to Appendix B.

Data availability section: There is no link to TROPOMI data Brewer data for triad is not accessible through WOUDC.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2020-324, 2020.