Atmospheric
Measurement
Techniques
Open Access

Discussions

EGU

Interactive comment on "Accuracy, precision, and temperature dependence of Pandora total ozone measurements estimated from a comparison with the Brewer triad in Toronto" *by* Xiaoyi Zhao et al.

Xiaoyi Zhao et al.

xizhao@atmosp.physics.utoronto.ca

Received and published: 31 October 2016

Thank you to referee #1 for your helpful comments. We have revised the manuscript based on your suggestions. Please note the page and line numbers in our response are referring to the numbers in the revised manuscript (but the referee's comments refer to numbers in the AMTD version).

p. 6, line 7. "model" should be replaced by "define" because word "model" usually means simulation of the ozone with a software and thus the beginning sentence in this section is confusing for the reader. "We define the small scale variability in the Pandora and Brewer observed TOC time series ()..." as a linear combination of the

true ozone variability and the . . ..”

Done. See p. 6 line 7 in the revised manuscript.

p. 6 lines 14-15, variance around annual mean or monthly mean or daily mean?

The meaning of variance X depends on the type of data selected. For example, if we use monthly TCO, then the variance X is around the monthly mean TCO. For our case, we used a two-year dataset, which means our variance X is around the annual mean TCO. However, we also defined Eqn. 7 and 8 to remove the daily variance in the TCO. So, after the removal of the daily mean, the variance X (residual ozone) is simply around zero (see Fig. 4).

We included an explanation for this point on p.6 line 15.

P 6 , line 26 is it daily mean difference?

No, for our dataset, it was difference between Pandora and Brewer high-frequency TCO measurements (averaged in 3 min bin). This was addressed before in Table 2.

p7, line 7. to remove variability - add daily in front of variability

Done.

p 7. Line 12-14. Are TCO measurements selected for each day and then compared to the daily mean for that day (mean(TCO lof-f(t)), where t is the time of individual TCO

measurement for each day)?

Yes. We modified the sentence to clarify this point (see p. 7 line 11-12).

p.7 line 22-24. Does equation (8) fitted to Brewer and Pandora data separately or all data are aggregated and sorted by time to fit this equation?

The daily M_{low-f} fitting (see Eq. 8) was done by fitting Brewer and Pandora data for that day together. Thus for the M_{low-f} for Brewer and Pandora share the same B and C terms, but have different offset term (A_B and A_P).

We modified the sentence to clarify this point (see p.7 lines 22-23).

p. 7 line 26 Add (Eq. 7 and Eq. 8) after " residual ozone "

Done (see p.7 line 29). We only add Eq. 7 here, because Eq. 7 is the general equation for both type 1 and 2 calculation, but Eq. 8 is only for type 2. Including both of them in the sentence is a bit misleading.

p.7 line 27 – please give more details about "statistical variable estimation method" you use and reference. If you need more space you can use Appendix. It is hard to follow your approach without knowing how it was applied. The point here is that the paper should be written in such way that anyone can repeat your calculations and get the same answer.

Thanks for the suggestion. It was referring to Eq. 6. So we modified the sentence to clarify it (see p.7 lines 30 to p.8 line 2).

C3

p. 7 line 28 Add Table before 2

Done (see p.8 line 2).

p. 7 lines 30-32. Please add more description on how you get your estimates of random error. Do you use Eq. 7 or Eq. 8 to calculate variance for B and P to get σM_b and σM_p? Do you use difference between dB and dP where dB and dP are residuals to get $\sigma M_b - \sigma M_p$? How do you match B and P data in time? Not sure if the difference in panels a) and c) in Figure 2 is caused by the fit of Pandora time matching to Brewer measurements or interpolation of infrequent Brewer measurements to Pandora high frequency measurements.

We use Eq.s 7 and 8 to calculate the residual ozone dM_b and dM_p, and then use the calculated residual ozone to calculate the difference $dM_b - dM_p$. Next, we use the calculated dM_b, dM_p, and $dM_b - dM_p$ to calculate their variances $\sigma^2(dM_b)$, $\sigma^2(dM_p)$, and $\sigma^2(dM_b - dM_p)$. Those variance terms are used in Eq. 6 to estimate the random uncertainties. We have modified the symbols in Eq.s 7 and 8, and Table 2 to make the description more clear. We also included a sentence giving an example of the calculation (see p.7 lines 30 to p.8 line 2 and Table 2).

To make this work fairly for both Brewer and Pandora, we decided not to fit Pandora times to match the Brewer or vice versa. As shown in Table 2, the high frequency M_b and M_p data are the paired Brewer and Pandora data in the fixed 3 min bin. Each bin should have at least 1 M_b and 1 M_p measurement, but for example, if there are two M_p measurements in one bin, then they will be averaged. This was mentioned in Table 2 and Section 2.

Printer-friendly version

Discussion paper

C4

p.8 , line 17-19. Is the lower random uncertainty for Pandora a result of the more frequent measurements as compare to Brewer?

Pandora does have a higher sampling frequency than the Brewer. In each 3 minute bin, we should have 1 Brewer measurement, but about 2-3 Pandora measurements. Statistically, if those 2-3 measurements are considered as repeated measurements for the same quantity, the uncertainty of the mean will be 30-40 % lower than the Pandora random uncertainty from a single measurement.

We tested using a single Pandora measurement in each 3 min bin instead of the averaged value (for example, if we have two Pandora data points in one bin, we only use one of them). The result shows that the Pandora random uncertainties change by less than 0.1 DU and thus Pandora still has lower random uncertainties than Brewer. This is because the statistical uncertainty estimation method is not sensitive to the averaging effect. The random uncertainties were calculated from the variances of residual ozone, and so they reflect the deviation of high-frequency ozone measurements from the low-frequency signal (daily mean or the 2nd order fitting). By subtracting the low-frequency signal, the averaging effect is largely removed in the residual ozone. We also tested several time bins, such as 5 and 10 min, and the results were consistent with those for the 3 min bin (as long as the number of data points is large).

P.8, line 21 – this is the first time that it is explicitly explained that residuals are used for analyses. It is better to do it at the beginning of the section 3.2.

Done. (Answered in previous, see p.7 lines 30 to p.8 line 2 in the revised manuscript).

p.8 line 23. For the type 1 fit the residual variability is a natural variability in ozone

during that day, yes? Not clear in the text.

The sentence has been modified to clarify this point. The estimated variability is the variability for the period of the dataset. Thus in our case (a two year dataset), the estimated variability is comparable to yearly variability. See p.8 line 29.

p. 11 lines 17-18. Herman et al. used temperature and ozone climatology, but interpolated climatological ozone profile to the observed TO in order to capture day-to-day variability.

Thanks for the information, and we think the reviewer intended to mention p. 12 lines 17-18 in the AMTD version. We have revised the sentence to include this information, please see p. 12 lines 27-28.

p. 13 lines 30-32. How does global ozone variability compares to ozone variability over Toronto?

The ozone nature yearly variability over Toronto is moderate comparing to high latitude sites. In fact, this also benefits our comparison. This paragraph is discussing OMI-TOMS vs. Pandora rather than ozone variability. Balis et al. (2007) reported that the corresponding percent difference between OMI-TOMS and Brewer for the years 2005–2006 is 0.61% as noted in the paper. They also reported (see the Fig. 1 in Balis et al. 2007) a difference close to zero in the 40-50° N band (Toronto is at 44° N). This is consistent with our corrected Pandora data which has a difference of -0.19 % $\pm$ 1.00 %. We have added this in the paper, see p. 14 lines 11-14.

Table 3 . It might be useful to have a temperature sensitivity for each Brewer listed in the Table. It could help to understand variability in results of comparisons (i.e. Figures 7, 10 and 11) when comparing Pandora to different Brewers.

The numbers are included in the Table 3.

Figure 1. The colors overlap, and it is hard to discern data for individual Brewer or Pandora. Would it help to plot comparisons as difference (from the mean)?

We have revised the Fig. 1. An extra panel is included to show the difference (from the OMI-TOMS).

Figure 2. panel c) The difference in random uncertainties shown in red (Pandora #104) and black symbols (Pandora #103) is the largest when comparisons are done with Brewers #8,14 and 15, but not with # 145, 187, 191. Is it something to do with the first set of Brewers being single and the last three in the plot are double instruments? Is it discussed in the paper?

In general, the results from double Brewer instruments are more consistent with each other. In fact, there were no coincident measurements between Pandora #104 and Brewer #145 (see Table 1). The similarity for Brewer #187 vs. Pandoras is due to small sample size. So on Fig. 2c, we can only confirm the difference (red minus black symbols) for Brewer #191 is smaller when compared to that for Brewers #8, 14, and 15. More coincident measurements with BrT-D are needed to determine whether this is due to the Brewer instrument design.

You discussed differences in the uncertainty estimates due to lower sample available for comparisons of Pandora #104. However, when comparing to Brewer 191 both Pandora retrieve the same uncertainty when using method type 1. Please provide an explanation. Is it something to do with Brewer temperature sensitivity of stray light interference...

Only the data with ozone AMF less than 3 were used for the statistical random uncertainty estimation work. Thus the stray light effect should be negligible in this estimation. We have included explanations on p.7 lines 25-27.

Both Brewer single and double instruments use the same wavelength range, thus theoretically, the temperature dependence for the Brewers should be identical. We can see some differences in the estimated Pandora RTDFs (compared to different Brewers), but we do not think it was dominated by the stray light effect. Figure 4c shows when comparing to Brewer 191, both Pandoras retrieve similar uncertainties (although the sample sizes are different). This is due to two facts. First, the number of coincident data points is good enough for the comparison for the estimation work. Second, the estimation of Pandora uncertainties is related to the quality of Brewer data. As we reported in the manuscript, Brewer 191 was one of the most reliable instruments during the comparison period.

Figure 5. panel b) although the seasonal variability is reduced after applying temperature correction, there seems to be a large spread in the remaining data. Since Brewer #14 is a single Brewer – there may be a stray light interferences that contribute to the range of the daily Delta Ozone differences (vertically grouped dots). Is it possible to repeat these comparisons but use double Brewer data to test is the spread will be reduced?

C8

Only the data with ozone AMF ≤ 3 were used in the temperature sensitivity study. So the spread on Fig. 5b was not due to stray light interference. It could be due to large changes in the effective ozone temperature over one day. Please note that the Teff we used is a daily value (which was calculated from ECMWF data on 18:00 UTC for each day). So, the retrieval of Teff from TCO measurements (a goal of the Pandora research group) is important for a further improvement of Pandora TCO. Similar plots to Fig. 5 for each pair of Pandora and Brewer were made, but we didn't find any obvious decrease in the spread.

Figure 7 and corresponding discussion on page 11. The tests are described, but no further conclusion is made on how these test change the values or uncertainty bars. Can you please add further discussion of how the choice for different combined and individual data sets can impact the derived Pandora RTDFs and biases. It can be discussed in regards to the instrumental parameters (single vs double) and sampling limitations.

In general, the size of the error bars from the tests with a small number of data points is larger. For this 2-year data period, the derived RTDFs from BrT-D instruments are lower (0.241-0.246 % K^{-1}) than the ones derived from BrT instruments (0.262-0.290 % K^{-1}). However, with the large uncertainties on the estimated RTDFs and the bias, we could not conclude whether this is due to the different instrument designs or a sampling issue. In the future, a longer comparison period may help to draw a further conclusion as suggested by the referee. We have included some of the above discussion on p.12 lines 8-10.