Atmos. Chem. Phys. Discuss., 11, C12647–C12655, 2011
www.atmos-chem-phys-discuss.net/11/C12647/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Five blind men and an elephant: can NASA Aura measurements quantify the stratosphere-troposphere exchange of ozone flux?" *by* Q. Tang and M. J. Prather

Q. Tang and M. J. Prather

qtang@cornell.edu

Received and published: 2 December 2011

We thank reviewer 2 for their careful reading and constructive suggestions. Most all of these will be implemented in our revised paper, and we anticipate that the reviewer will find it improved.

Overall Comments

We basically agree with the reviewer's principal tenet that the Aura instruments were never intended, nor had the prime capability to resolve the UT/LS region with enough skill to infer STE of ozone. (This statement and our paper would likely be very different if HIRDLS had not been damaged in launch.) Thus we agree to change the tone of

the paper and propose renaming it as "Five blind men and the elephant: what can the NASA Aura ozone measurements tell us about stratosphere-troposphere exchange?" Thus we take the view of what we can learn from the Aura combined-instrument ozone data regarding STE ozone fluxes.

Yes, Figure 1 shows that CTM stratospheric errors propagate and destroy the tropospheric comparison. What was unclear to us when we started was that the excellent comparisons with tropospheric ozone between TES and some CTMs (e.g., GEOS-Chem) were due to the fact that these models did not predict stratospheric ozone and hence had no such trouble. So indeed this was a surprise to us (and I suspect many trying to simulate full ozone profiles), and it taught us the importance of the a priori in the stratosphere for tropospheric ozone. This problem is separate from the issue of TES observations identifying folds.

Yes, we will reverse the order of the TES discussion as suggested, specifically noting that DOF = $\sim$1 to start with. We will emphasize the use of the CTM as a transfer standard and integrator to connect the four Aura ozone measurements. And, simply point out the difficulty (as expected) in resolving tropopause folds (TFs) with nadir ozone sounders. The measurements themselves do not directly measure STE ozone fluxes. We try to use the measurements to evaluate the model and narrow the uncertainties in the simulated STE fluxes. Since previous studies suggest that Aura data can provide useful information regarding STE processes (at least on a case basis), we choose these datasets in this work. In the revised version, we will state in the introduction that STE is never an easy task for space remote sensing and Aura (except for fully functional HIRDLS) is not expected to have this capability. The fact that we find Aura ozone data do have some skill in catching STE structures is beyond its designed scope and hence a bonus.

We concur with most all of the detailed fixes suggested by the reviewer and will make those changes where possible in accord with reviewer's suggestions. We discuss below only those we cannot readily fix as suggested.

Specific comments ABSTRACT: and the inconsistency . . .

Yes, this is a good point, and we propose the following change to the abstract:

"The statistics of exact-matching CTM-Aura comparisons identify the model's high biases in the lower stratosphere and the inconsistency amongst different instruments in the tropical upper troposphere and mid-latitude lower stratosphere, where STE processes are most evident in the model but where the instruments cannot provide very good information on individual ozone profiles."

26899 l 5 I thought an important driver . . .

Yes, different definitions of STE O_3 flux are used in these studies, and this is certainly an important source of the different fluxes. Other factors are also important, such as different diagnostic methods and model differences. However, when we diagnose the STE O_3 flux by different methods in our model consistently (e.g., flux across 100 ppb, 200 ppb, across 100 hPa, taken up at the surface) the differences are not that large and thus we believe there are other problems here. To avoid confusion, we add "definitions" to this sentence ". . .partly due to the different definitions and diagnostic methods. . .".

L 17 The Aura instruments were never designed to resolve trop folds – at best, designed to see their effects (in an integrated sense).

Agreed, it will be revised.

26901 7 'modeled profiles generally match sondes' – broad statement in Tang and Prather 2010 and also broad statement here. Really would love to see scatter plots, histograms, statistics.

We tried very hard in Tang and Prather (ACP 2010) to develop an objective criterion for the CTM matching of the ozone sondes in terms of strat-trop folds. We examined visually over a thousand CTM-sonde comparisons, and eventually restricted ourselves to the major fold regions (35° S–40° N, still over 600 comparisons) and then plotted 4 samples in that paper's Figure 1. The text notes that about 50% were good matches

(Fig 1ab), and 30%, "C" grades (Fig 1c), and 20%, "F"s (Fig 1d). These grades and percentages of the 600+ matches should have been put in the figures as well. We looked at many different objective algorithms, but the grading criteria (or equivalently the scatter plot of some simple metric) never worked, i.e., the subjective (visual) comparison of the profiles was not reflected in the objective one. Also note that the CTM vs. sonde and vs. O_3 DIAL in the TRACE-P mission using the same CTM (Wild et al., 2003 JGR D21, 8826) produced some excellent comparisons of mapping out folds, but no easy scatterplot describing it. This is in contrast to the statistical measures of near-tropopause O_3 profiles from the Logan analysis of sondes that is straightforward to compare with. We welcome any suggestions on how to do objectively grade the simulation of folds.

L 14 – they aren't exactly 'ozone instruments' – they are four instruments that measure ozone (among other thing)

Yes. Changed to "with four instruments (. . .) observing ozone on board".

L19 – 'many studies' – there is no need to ...

The reviewer makes a good point about the lack of precision on single profiles with L2 data, and this we clearly deal with in our comparisons (Fig 3,4,A1-7). We have re-read the short paragraph and believe it to be a simple factual statement that L3 data is averaged as stated (not all readers may be so familiar with the data) and that it certainly cannot be used to test the simulation of relatively short-lived meteorological features like folds. We believe that there are very few studies like this one that have used two full years of L2 swath data without averaging in a direct model comparison, while there have been many comparisons using L3 (which we agree is appropriate for much of the science).

26902 HIRDLS – you probably should refer to the appropriate data quality document.

In the beginning of this section, we cited the data quality document for HIRDLS V5 data (HIRDLS Team 2010). If there are other documents that we missed, we will gladly include them.

26902 MLS – mention single profile precision in the altitude range in question, and refer to data quality document.

Yes. Single profile precision is a piece of useful information. Before "(Livesey et al., 2007)" we add "The precision of single profile is 40 ppb at 215–100 hPa."

26903 OMI – other TCO products need to be mentioned.

We do not feel this is needed here since we discussed the other OMI-MLS TCO products in Tang Prather 2010 (p.9683). Also, we could find no available L3 TCO product to compare directly with although we know some are being prepared.

More specifically, we should have in 2010 cited the Schoeberl paper (2007, JGR D24S49) in which MLS stratospheric columns are mapped to L3 OMI maps using forward trajectories to get the TCO. Unfortunately, this does not seem to be a standard product. We have since had discussions with Ziemke about regenerating the equivalent L3 TCO product for a few months to do a comparison between our e90-based tropopause TCO using OMI profiles and the Ziemke product using OMI total columns and MLS stratospheric profiles. If the reviewer thinks this information should appear in the text we can so revise it.

26903 TES – someplace you are going to . . .

Agreed. L19 changed to ". . .surface to 0.1 hPa, but in the troposphere there is only 1–2 degrees of freedom for the signal (DOFS) (Nassar et al., 2008; Zhang et al., 2010)."

26906 – at the end of the discussion about TES . . .

Agreed, this is a bit complicated, but will be revised according to suggestions.

26908 – even when introducing the case studies. . .

We intentionally chose a much stricter criteria than did the satellite validation work because our focus is on rapidly changing processes, as we demonstrate in the model simulation (Fig. 2). The 1-hour coincidence requirement is a somewhat arbitrary choice, and we should not interpret the number of cases found as having significance, but at least we found enough cases to make some interesting comparisons. For these case studies, the sondes are used to examine the model's ability of resolving folds, not just reproduce the mean upper tropospheric ozone. So, a shorter time lag between sonde and simulation is desired.

26909 – you also said that TES had only 1 DOF in the troposphere. Isn't the Pressure/Loc plot misleading?

We will change the text to explain that the TES Pressure/Loc plot is shown because TES footprints are too sparse to allow it infer STE from TCO anomalies alone (as for OMI) and thus the combination of changes in upper-trop ozone (even with DOF = $\sim$1–2) with the expected (modeled) geographical pattern of the fold, may provide a convincing match and validation of the modeled fold. Admittedly, the TES vertical information requires careful interpretation.

26910 – does using the OMI operator make a difference?

The OMI operator changes CTM TCO, but the CTM TCO patterns are still similar to those of OMI. Since here we only qualitatively infer folds from the geographic anomalies, the conclusions are the same w/ or w/o applying the operator.

26910 – these numbers for MLS either make sense with respect to single profile precision and data quality or they don't. Do they? For an appropriate average, does MLS give the correct mean? Is this a 2 or 3 sigma number?

These comparisons between MLS and CTM generally agree with the single profile precision, that is 40 ppb at 215–100 hPa. There are outliers that may be caused by the model biases in this region. Table 2 and 3 give the monthly mean biases and RMS errors for different latitude bins which are in good agreement with previous validation with sondes (e.g., Jiang et al., 2007). This study attaches more importance to the patterns associated with STE events rather than the absolute magnitudes. It is hard to assess the statistical significance (2 or 3 sigma) for these case studies as the numbers are too few.

It is important to separate out likely model biases/errors (Tables 2 3) from the single profile noise. Thus, we will add the qualifier before the last sentence: "As expected from the single-profile precision of MLS in this region, some unphysical..."

26911 HIRDLS doesn't see below clouds ...

Agreed. We are impressed that there is some skill for the HIRDLS data in this region. We will add "... below the tropopause and are obscured by the presence of high clouds in the troposphere."

26911 Pretty shocking to me to read that TES shows 'skill' in detecting boundary-layer O3, since the sensitivity is very low there. In fact there is a huge effort to combine TES/OMI information because by the UV-vis + ir has theoretical sensitivity to the troposphere.

It was a surprise to us too. But the comparison does show agreement between TES and CTM in that region. Probably it is not sensed by TES. A further analysis on these profiles show that TES retrieval greatly reduces the a priori in the upper and middle troposphere due to likely folds (shown in simulation) in that region and consequently puts high ozone in the lowermost troposphere to give TCO matching the observations. This sentence ("TES sense the high anomaly...") will be rephrased as "TES shows the high anomaly at 37° N lowermost troposphere because the possible folding structure aloft, indicated by the simulation, is reflected as a significant, smooth inversion in TES middle and upper tropospheric retrievals and thus in compensation high values in the lowermost troposphere to give reasonable TCO."

We are aware of the study combining TES and OMI and look forward to this new tropospheric ozone product.

Figures 3 and 4 You need ...

Yes, the figures will be revised to make them clearer.

26912 – not just noise, lack of sensitivity and vertical resolution...

Yes, this will be revised.

26913 – 'cannot provide observations of all the ...

Yes, this was clearly a "no-duh" statement and will be revised as: "The Aura datasets alone are therefore capable for only a few specific cases (refs). Combined with the 4-D hindcasts here or with a data assimilation system, they may lead to a general, comprehensive integration of the global STE flux, but more work is needed."

26914 – 'suggesting low sensitivities and great noise' ...

Yes, changed as suggested.

26914 - 'The CTM-TES comparisons are almost always improved...' ...

Agreed.

Re: the poor syntax and grammar. Oops, these will be revised according to the comments.

Reference:

Jiang, Y. B., et al. (2007), Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, J. Geophys. Res., 112, D24S34, doi:10.1029/2007JD008776.

Schoeberl, M. R., et al. (2007), A trajectory-based estimate of the tropospheric ozone column using the residual method, J. Geophys. Res., 112, D24S49, doi:10.1029/2007JD008773.

Tang, Q., and M. J. Prather (2010), Correlating tropospheric column ozone with tropopause folds: the Aura-OMI satellite data, Atmos. Chem. Phys., 10(19), 9681–9688, doi:10.5194/acp-10-9681-2010.

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 26897, 2011.