Atmos. Chem. Phys. Discuss., 14, C12701–C12711, 2015
www.atmos-chem-phys-discuss.net/14/C12701/2015/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Comparisons of polar processing diagnostics from 34 years of the ERA-Interim and MERRA reanalyses" *by* Z. D. Lawrence et al.

Z. D. Lawrence et al.

zlawrenc@nmt.edu

Received and published: 19 March 2015

We thank Simon for his detailed comments and very helpful suggestions. Our responses to his comments and questions are:

Major Comments
P. 31366, lines 6-10
The absence of comparison with independent measurements is the main limitation of this study. It can not be justified by the scarcity of such measurements, because many observational datasets include temperature and are not used in data assimilation. Two

important examples are

- *the ground-based observations collected by the NDACC network: these include Lidar and ozonesondes (which also measure temperature);*
- *the satellite limb sounders such as UARS-MLS, Envisat-MIPAS or Aura-MLS.*

Several comparisons between such instruments and meteorological analyses can be found in the literature. I think that such comparisons are simply beyond the scope of this paper, but this limitation should nonetheless be mentioned in the Introduction, with proper references to available datasets and published comparisons. I suggest to remind it in the Conclusion as well, because this is an important venue for further research in our field.

We certainly agree that reanalysis comparisons with independent measurements are extremely important when they are possible. In the context of polar processing and the diagnostics that are commonly used, however, there is very little potential for such comparisons. We were unclear in our introduction by making it sound as if there are few independent observations available. What we are really trying to say is that there is a scarcity of independent measurements that are applicable for polar processing diagnostics, which require large-scale temperature and potential vorticity data that currently cannot be provided by satellite or ground-based measurements.

This was a topic of considerable interest at the most recent S-RIP meeting back in September 2014 when several attendees asked if the primary goal of the S-RIP project was to determine the accuracy of reanalyses. From the perspective of those of us who attended the meeting (Zachary, Gloria, and Michelle), it seemed like the consensus was to aim for comparisons with observations when possible, but just as importantly, we should also consider the agreement between reanalyses as valuable information since agreement (or lack thereof) gives users of the data an idea of uncertainty.

We have added a paragraph near the end of our introduction that we hope better explains our reasoning, and the scope of this study.

P. 31367, lines 2-5
If I understand well, the Potential Vorticity from MERRA is interpolated from 42 pressure levels to the model levels, followed by an interpolation to the isentropic levels where all vortex diagnostics are calculated. It looks like significant information could be lost in this process, while vorticity could be derived from the MERRA wind fields which are distributed directly on the model levels. Please check that the 42 pressure levels have a vertical resolution similar to the model levels, and/or that the diagnostics derived from p-levels PV are sufficiently close to diagnostics derived from model-levels PV. It also looks like the vertical integration of APSC and Avort (to VPSC and Vvort) is done on the vertical grid of isentropic levels (p.31370 line 2). What is its vertical resolution? Using a grid coarser than the model grid could introduce unnecessary errors in the integration. If this is the case, are such errors negligible?

We would have used MERRA PV on the model levels and grid if it were available. Several years ago, we tried to obtain model level MERRA PV from GMAO, but unfortunately we were told that it was not archived.

You are correct that two interpolations are performed on the MERRA PV data: one to make the reduced-resolution PV match the model levels and grid, and another to interpolate to isentropic surfaces. Even though this procedure undoubtedly introduces some error, we argue that it is still best to use the PV provided in the MERRA dataset since it does come directly from the model, and the errors introduced in calculating PV from the MERRA winds are likely to be larger than those introduced by the interpolations. We also think that this procedure is what most data users are likely to do, since it is arguably easier to interpolate one field to match the model grid and levels than it is to interpolate all of the model level data to the reduced resolution grid and levels, or calculate PV from MERRA winds. This process of matching the pressure-level PV with the model levels has been used before (e.g., in Manney et al., 2011). We have added

this citation in the MERRA portion of the Data Analysis section. Furthermore, in our experience, the interpolation of MERRA PV preserves the polar vortex edge gradients very well in comparison to GEOS 5.2.0 and 5.9.1 PV data that are, unlike MERRA, provided on the original model grid. Since we use the interpolated PV for vortex diagnostics that primarily depend on well-defined vortex edges, our diagnostics are unlikely to be affected significantly by the interpolated PV. Sentences about this preservation of the vortex edge have been added to the text (also at the end of the MERRA portion of the Data Analysis section).

The resolution of the isentropic levels we use for the vertical integrations that calculate V_{PSC} and V_{vort} is comparable to, but perhaps slightly coarser than, that of MERRA and ERA-I. According to the Knox equation for altitude, the isentropic levels we use (390, 410, 430, 460, 490, 520, 550 and 580 K) are all roughly 1.1 km apart (this information has been added in the text). V_{PSC} and V_{vort} are commonly calculated with altitude approximations - some more simple than others. For instance, Rieder and Polvani (2013) calculate V_{PSC} from A_{PSC} only on two pressure levels (50 and 30 hPa). These diagnostics are not very sensitive to the method used to calculate the volumes, but they are still commonly used because of the strong correlation of V_{PSC} (and V_{PSC}/ V_{vort}) with ozone loss. A sentence about this insensitivity has also been added to the revised paper.

P. 31371, lines 8-16
This description of the calculation of trajectories is not sufficient to allow reproducibility of the results. Livesey (2013) is a simple web page which does not provide the source code, only a brief description and output datasets using (if I understand well) MERRA fields. It looks like Livesey (2013) included diabatic motion, but from section 3.4 we understand that it is not the case here. I think that the explanations on p.31379 lines 15-21 should be transferred here. Even so, some key questions must be addressed: Was LTD code fed with daily wind fields or more frequent analyses, e.g. 6-hourly? In the first case, explain why the errors due to daily update are negligible; in the second

case, update the description of downloaded datasets in sections 2.1 and 2.2.

We use the 6-hourly wind and temperature data from MERRA and ERA-I to calculate the trajectories. This information and other clarifications about the trajectory code have been added to the text. We did not update the description of the datasets in sections 2.1 and 2.2, since we mention that 12UT data are used by default "except where specified otherwise" and "unless stated otherwise."

Section 3.1
*It is explained that in this paper, the word "bias" designates the mean difference between a diagnostic extracted from MERRA and the same diagnostic extracted from ERA-Interim. This is *very* confusing because in the context of analysis evaluation, the evaluation of the "bias" uses an observational dataset as reference and is a proxy for the systematic error present in the analyses. It is explained that the words "relative bias" are an attempt to clarify the concept. This makes the text even more confusing in my view, because in the context of analysis evaluation the "relative bias" is a dimensionless ratio between the absolute bias and some value representing the investigated quantity. The differences discussed in this paper, on the other hand, have the same units as the diagnostics themselves, do not use any independent observations and are not meant to evaluate the validity of either dataset. The typical reader first looks quickly at a paper, reading the titles of the figures and the sections to understand the scope of the paper. This choice of words will unavoidably lead her to believe that some comparison with observations is performed, while this is precisely not the case (see major comment 1 above). Furthermore discussing the units of a "relative bias" is totally counter-intuitive. I strongly recommend to replace throughout the whole manuscript, "bias" and "relative bias" by "mean difference".*

We regret the confusion and clash of terminologies. Our intention was for "relative bias" (in the context of this paper) to mean the "direction" of the diagnostics - in other words, to point out when one dataset is higher/lower in a given diagnostic with respect to another dataset. In hindsight, this was a poor choice of vocabulary due to

the standard statistical definitions of "bias" and "relative bias." We feel, however, that simply replacing these words with "mean difference" does not accurately describe the quantities we use. Therefore, we now refer to these quantities as monthly Comparison Period Average Differences, or monthly CPADs, to indicate that we're taking monthly means of differences averaged over some period of years (i.e., they are not just mean differences). Hopefully this is more explicit, and much less confusing. We have also removed most mentions of "bias" in the text; only the occurrences of "bias" that were relevant are left in.

P31373, line9; p31374 line 15; p31375 lines 15-16
Why are most figures shown at 580K (fig. 3,4,6,7) while fig. 5 is shown at 490K ? The text mentions repeatedly that several mean diagnostic differences depend on altitude, but this is not shown on any figure. I suggest to show this dependence explicitly for some well-chosen diagnostic, ideally through a vertical profile of this mean difference. For example p31375 lines 15-16: if below 520K there is no convergence towards better agreement, why not show it? It would be interesting to show a case where the disagreement persists even after 2002.

The "Number of Days T < T_{PSC}" diagnostic is shown at 490 K because it is the level with the greatest number of days below T_{PSC} for both datasets, and because it gives an idea of the T_{min} differences at 490 K in addition to those shown at 580 K.

We have added a new figure (Figure 9) to show the average differences of the A_{NAT} diagnostic at all levels up to 580 K over the different comparison periods. We chose to show the altitude dependence for this diagnostic because it fits naturally with the discussion of V_{PSC} later in the paper.

Minor Comments
P. 31363, lines 1-12: consider adding some newer references.

The references given in the first few lines are used because they are definitive papers on these topics, but we have added newer references where they are relevant.

P. 31363, line 27: is the word "myriad" really necessary?

It has been removed.

P. 31367, line 13: replace words "In this case", e.g. by "Here"

Changed as suggested.

P. 31368, lines 4-5: for clarity, mention already here the year of introduction of COSMIC GPSRO data in the reanalyses.

Done.

P. 31369, line 10 (also line 24): how is it possible to examine "daily" minimum temperature with only one instantaneous field per day (i.e. at 12:00 UT per sections 2.1 and 2.2)? Please clarify.

We have clarified that these diagnostics are “Daily 12 UT.”

PP.31369-31371: section 2.4 is too long (especially taking into account major comment 3 and next comment suggesting an additional figure). Consider splitting it into "basic" diagnostics (up to P.31370 line 12) and "advanced" diagnostics (VTC, TT195, CT195).

We have split section 2.4 into a “Temperature and Vortex Diagnostics” section, and an “Advanced Dynamical Diagnostics” section at the recommended location.

P. 31370, lines 14-23: The definition of VTC provided here is new. It should be illustrated with a dedicated figure. I recommend "snapshot maps" showing situations with VTC close to 1 and <= 0, for date(s) of special interest (e.g. the initialization dates of the trajectories used for figures 15 and 16). Since it will probably not be possible to distinguish the PV (and temperature) isocontours by both reanalyses, the figure could mention the two numerical values of VTC for each example.

We have added an example figure (now Figure 2) to demonstrate our VTC diagnostic, and included a couple sentences in the text describing this figure.

P. 31370, line 27: vortex split and SSW are two related but different events. Do you mean here simply "vortex split events" ? Same for p.31371 line 26: do you mean e.g. "major SSW with a vortex split" ?

We have clarified p. 3137 line 27 by removing “SSW” and p. 31371 line 26 as suggested.

P. 31373, line 5: provide a reference about this likely impact of AIRS data

We have added references to McNally et al., 2006, and Rienecker et al., 2011. Both of these discuss the large volume of usable measurements from AIRS, and their impact on the assimilation systems.

P. 31373, line 17: "clearly demonstrate" - consider replacing by "clearly show"

Changed as suggested.

P. 31374, lines 9-10: I do not understand "...below either threshold...". Please clarify.

We have clarified this sentence to better explain that we use the ice temperature threshold for the Antarctic because it is a lower (i.e., more sensitive) threshold, which makes the differences clearer for the Antarctic where temperatures often stay below T_{ice} for long periods of time.

P. 31375 line 19: "mixing of air ...and..." -> "mixing of air ...with..."

Changed as suggested.

P. 31376 line 11: Is paper by Livesey et al. already submitted? If no, consider removing this reference as there is already one for this topic; if yes, please update the reference.

The paper by Livesey et al. was submitted recently and should be in press for ACPD very soon. This citation has been updated accordingly.

P. 31376 line 17-20: the blue and red lines on Fig.10 are so close that the differences can not be discussed there in a credible manner. Consider deleting these lines and

discussing the differences directly with fig.11.

We feel that these lines are important to demonstrate just how similar the sunlit vortex area diagnostic is from both MERRA and ERA-I. We have modified Figure 10 (now Figure 12), and Figures 6 and 8 for consistency (now Figures 7 and 10), to make the small differences we reference easier to see and understand. In the top panel, ERA-I's line is almost always above MERRA's, which indicates that ERA-I's polar vortex tends to be filled with slightly more sunlight. In the bottom panel, the months May through June show ERA-I's line above MERRA's, while mid-September through October shows MERRA's line above ERA-I's, indicating that the differences change over the season.

P. 31377, lines 8-10: static stability was not discussed in section 2.4. This sentence is not clear and seems not useful to me.

This sentence has been removed.

P. 31378 lines 4-8: These are indeed important caveats on the impact of vertical integration, time averages and smoothing errors. They should be mentioned in the conclusions as well.

We have reiterated this point in bullet 3 of our conclusions.

P. 31379 line 9: "...ERA-I could bias model runs..." -> "...ERA-I could lead model runs..."

Changed as suggested.

P. 31379 lines 15-21: move to end of section 2.4 (see major comment 3 above).

We have left these lines in this section because we think they fit better here, but we have added clarifications to section 2.4.

P. 31381 lines 1-2: It seems to me that very similar results between ERA-I and MERRA can not "lend confidence in transport calculations using winds from these two reanalyses". They simply show that both reanalyses were well constrained by the same datasets. The only way to have confidence in transport calculations is using indepen-

dent observations (e.g. of chemical tracers).

We have clarified this sentence to say that the similar results lend confidence that using either dataset in transport calculations would give comparable outcomes.

Table 1: I do not understand the difference between "System" (ATOVS, TOVS), "Instrument" which flew on several satellites (e.g. AMSU) and "Satellites" (GOES)

We have removed the "Types" table column. We originally intended for this column to help distinguish what each of the acronyms are (e.g., the name of an instrument versus the name of a satellite), but it is not crucial for the paper.

Figure 2: I guess that the x-axis tickmarks and monthly labels are for the 1st day of each month. If this is the case, please edit the labels to "1Nov", "1Dec" etc. Same for figs. 6,8,10

We have made these changes in addition to those described above, and have changed the figure titles for Figures 6, 8 and 10 (now Figures 7, 10, and 12) for consistency with the next comment.

Figures 3,7,9,11: these figures look very similar and one easily confuses them while reading the text. I suggest to add as figure title (bold font) the name of the diagnostic difference shown on the plot (as for most other figures).

All of these figures (now Figures 4, 8, 11, and 13, respectively) now have titles.

Figure 4: too small, not readable. Please re-arrange the layout (the web page layout of ACP requires wide figures). Legend: please write the three values of sPV used to draw the vortex edge.

We agree that this figure (now Figure 5) is difficult to read in the landscape layout of ACPD, but we think (and hope) it will work much better for the portrait layout of the final ACP publication. The vortex edge sPV values have been added to the figure caption.

Figure 10: please indicate on the legends the units (fraction of hemisphere area)

Figure 11: please indicate on the legends the units (% of hemisphere area?)

We have changed Figure 10 (now Figure 12) to have units of '% of a hemisphere' to be consistent with the units in Figure 11 (now Figure 13), and have included these units in the figure caption. This has also been done for Figure 6 (now Figure 7) to be consistent with Figure 7 (now Figure 8). All text that refers to fraction of a hemisphere has been updated accordingly.

Interactive comment on Atmos. Chem. Phys. Discuss., 14, 31361, 2014.