Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-879-RC3, 2018

Atmospheric Chemistry and Physics Discussions

Open Access

ACPD

Interactive comment

Interactive comment on "On the value of reanalyses prior to 1979 for dynamical studies" *by* Peter Hitchcock

E. P. Gerber (Referee)

gerber@cims.nyu.edu

Received and published: 9 November 2018

This is an interesting and well thoughout study about the potential value of earlier, pre-satellite era reanalysis records. It is important to quantify the potential value of this earlier period, as it is a major undertaking for a reanalysis center to provide pre-satellite reanalyses. With the exception of JRA-55 (and the ERA5 analysis, currently in production), most of the state-of-the-art full input reanalyses do not begin until 1979 (ERA-I, MERRA, CFSR) or 1980 (MERRA2).

I recommend publication of the manuscript pending consideration of the comments below. They are mostly minor, in that I leave them to the author's discretion, but I hope that responding to them would improve the impact of the paper. (An exception is that

Printer-friendly version

Discussion paper

the author does need to better define a few things, to ensure the results are reproduceable. But this will be easy to do.) My more philosophical question about the proposed metric for assessing the value of earlier reanalyses (see below) is perhaps trickier to fully answer, and might be something for future work. I think that the contributions of this paper are already worthy of publication. Given that it could be a subject for future research, I'll sign this review, as I would welcome discussion with author. Edwin Gerber

General comments

1) A few key elements of the procedure were not sufficiently documented. In particular, how were the SSW dates set, and how were the events classified in the spilits or displacements. I suspect this was done within the S-RIP Chapter 6 framework, assembled by Amy Butler. If so, I am not sure how to properly cite this information at this time, though they will be published. In any case, to reproduce these results, the reader does need to know the dates, and some insight on how they were obtained.

2) It would help the reader to adopt a consistent use of the nomenclature "full-input", "conventional-input", and "surface-input" throughout the paper. I appreciate that terms evolved in parallel to this research, but as a result of this time mismatch, they appear inconsistently through the text.

3) I very much appreciate the central result of the manuscript: equation (3) and surrounding discussion, which seeks to quantify the value of earlier records. I was admittedly surprised, however, that the metric indicates that there is considerable value to much of the data in the austral hemisphere, where we know that the large scale circulation is not consistently captured by the reanalyses. (In Gerber and Martineau, 2018, for example, we found that the southern annular mode indices in JRA-55, ERA-40, and NCEP-R1 share only a small fraction of the variance during the pre-satellite period, indicating that there is very little consensus on the large scale state of the austral hemisphere on synoptic time scales.)

I think the key is the assumption that reanalyses properly capture the dynamical uncertainty, \sigma_d, in both the satellite and pre-satellite periods. I think this effectively implies that we trust their climatological values and variance, even if the become untethered to observations.

To make my concern clear, consider the extreme case where the reanalyses are perfect in the satellite era (\alpha_s -> 0) and know absolutely nothing about the state of the atmosphere in the radiosonde era (\alpha_r -> \srqt{2}). In this case, f-> 2 and \delta -> [1-2*\beta]/[1+(1-\beta)*2]

When \beta becomes small (<0.5), you would still conclude that there is value in the reanalysis, even though it knows nothing about the state of the atmosphere. (The "real" \beta is about 0.6, so in this limiting case delta would be negative, and you would concluded there is no value in earlier records). But given that \alpha_s is not zero, and there is some limited skill in the radiosonde era, it's not hard to see why delta is positive. And by this logic, there would be considerable value in using the entire record from ERA-20C, where beta drops below 0.5!

My intuition if we want an *observationaly constrained* estimate of the uncertainty, then we should only include the information from the earlier period when \alpha_r < 1. That is, when uncertainty in the renanalyses reaches the level of dynamical uncertainty, then we can argue the reanalyses are sufficiently untethered from the real atmosphere to provide any additional information than you could obtain from simply running a forecast model untethered to observations.

I haven't thought this through enough to provide a good way to quantify the value of events when \alpha_r <1. It helps me to think of this interms of events (as with the SSW composites shown by the author.) Suppose you have N events from the satellite record. Looking at past events, the idea would be to quantify the additional information content of each radiosonde period event on an event-by-event basis. When the spread between reanalyses for an earlier event is equivalent to the spread between events in the satellite period (\alpha_r = \alpha s), the event is clearly of complete value

Printer-friendly version

Discussion paper

(\delta=1); it should be added fully. Now your composite is based on N+1 events, and the uncertainty drops accordingly.

If the spread between the reanalyses for the event, however, becomes equivalent to the climatological/dynamical spread (\alpha_r = 1) then I feel that there's no additional information to be gained than if you simply ran a free running model: this event should be given zero value. I am just not sure how to develop a meaningful way to interpolate inbetween these cases.

4) I appreciate that the comment above is weak on specific suggestions. To be more concrete, I would have appreciated more discussion of the different limits around equation (3). The limit where \alpha_s is small and \alpha_r approaches sqrt(2) was interesting to me, as it drove home this issue of whether we ought to trust a good model that is untethered to reality.

Another problematic limit is \alpha_r = \alpha_s. Here, you always use more data, even if it's all untethered to reality. (Based on my arguements above, the value of the reanalyses should be zero when \alpha_r or \alpha_s approaches 1.)

And not to be overly critical, Figure 4 was not easy to interpret. Consider using color or more simply marking the contours. (I know that I should have realized that diagonal is 1 by definition, but it took me time at first reading.) I also think that it's inappopriate to show such a range. Once \alpha_s or \alpha_r reach sqrt(2), nothing is tethered to observations, and I don't see how \delta is meaningful for values beyond this point.

5) There is a paper that can be cited for the Martineau data set:

Martineau, P., Wright, J. S., Zhu, N., and Fujiwara, M.: Zonal-mean data set of global atmospheric reanalyses on pressure levels, Earth Syst. Sci. Data, 10, 1925–1941, doi:10.5194/essd-10-1925-2018, 2018b.

Also, it's my understanding that MERRA2 has a DOI that should be cited, as it's very important for them to justify resources. I find the situation problematic, in that you got the data from a different source (which did cite this doi), but perhaps you could add the doi to the data section.

Global Modeling and Assimilation Office: MERRA-2 inst3_3d_asm_Np: 3d, 3-Hourly, Instantaneous, Pressure-Level, Assimilation, As- similated Meteorological Fields V5.12.4, https://doi.org/10.5067/QBZ6MG944HW0, http://disc.sci.gsfc.nasa.gov/mdisc/, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 2017-07-5, 2015.

Small comments by page:line

2:3-8 This would be a good time to differentiate and define full-, conventional- and surface-input reanalyses.

2:21 \citep[e.g.,][] (Also, I think that perhaps one should include the comma on e.g., since if you spelled out the phrase, it would be: for example, Matsuno 1971. But perhaps this is a case where American English is different from British.)

2:34 I might break this off as a full sentence, instead of using the semicolon.

3:11 I appreciate why the author states that they are constrained *primarily* by surface observations (as the reanalyses are given changes in radiative gases, etc.), but this sentence seemed a bit to vague.

3:25 consider rephrasing this sentence: I understood it completely, but had to re-read it a few times

4:16 This would be a place to explain how the dates were set, and how splits and displacements were classified, or at least point the reader to the necessary information.

4:24-26 This is a fascinating/perplexing result. I think it makes sense, though: ERA-20C does a good job of getting SSWs, but since it only gets the dates right for half of them, you are better off treating it as a free running model (i.e. not fixing dates to reality) than trying to make it conform with what actually happened in our atmosphere. ERA-

Printer-friendly version

Discussion paper

20C provides the challenge to your metric in equation 3: I do think you would argue that it's worth while using the entire record, even if just assume it knows nothing about the actual state of the atmosphere. That's what motivated my thoughs on comment 3 above.

4:29 splits and displacements need to be defined

5:1 second half of the line is awkwardly phrased

5:17 I'm concerned that the zonal mean wind at 60 N and 10 hPa is decidely not Gaussian, and rather skewed towards negative values.

5:18. There is a sentence between when you introduce \sigma_d and _o and define them. Consider moving the first sentence of the next paragraph up, to define the variables, before discussing the central limit theorem.

6:10 In the limit where the reanalysis error is small relative to the dynamical uncertainty, isn't f small, and delta about equal to 1?

8:20-24 This sentence is long. Consider breaking at the ;, and then being more clear what agreement you mean to refer to.

Fig. 8: I assume the Fourier analysis was done on the deaseasonalized winds, as there's no discernable annual cycle peak here!

10:21 and Figure 7 e,f I am confused how you can estimate \delta without knowing \beta. It only seems to decouple from \beta when f is small. And in this limit, \delta would be close to 1 (and it's sort of a trivial result: you trust everything.)

In the figure, the value of delta varies considerably (changing sign!) so you must have some finite value of \beta. What is 0.6?

11:1-2 This sentence could be split up, giving you two sentences, enough to justify a paragraph!

11:13-15 It was hard for me to see this important result. At some latitudes (c. 65 in panel a of Fig 10, or near 75 in panel b), the uncertainty bounds on the full record were larger than than for the satelite record. So clearly there wasn't always a 20% reduction.

I'm was also rather struck by the fact that the dashed curved in panel b of Figure 10 approaches the edge of the confidence interval on the "all" composite. Does this mean that they were almost statistically different, or would this only apply when the confidence intervals themselves separate.

From a practical standpoint, if I wanted to ask whether my model was signficantly different from our best estimate of observations, which error bound should I use?

11:30 Might be good to emphasize "has been quantified in equation (3)."

12:2 Related to some comments above, when the dynamical uncertainty dominates, doesn't this imply that you trust everything?

12:15 An opportunity to use surface-input nomenclature.

15:18 I am not sure how to see this in Figure 9. Doesn't the fact that \delta is consistently greater than 0 for ERA-20C imply that there's always value to be found from this reanalysis?

15:26-30 Could be opportunity to highlight that your message has been heard, and ERA5 hopes to go back to 1950.

Fig 4: See my general comment (4) above. I think this figure could be improved.

Fig 5: Your notation differs a bit here, \sigma_d vs. \sigma_dyn. It's clear enough for the reader, but consistency is best.

Fig 7: caption has the wrong symbol. I would have appreciated more detail here in how the bottom panels were computed.

Fig. 8: I find that the log scale makes comparison very difficult. Would it be possible

to show the ratio of the differences in the power spectra? This is a number that would presumable vary from 0 to about 2 for all timescales. (It would be 2 if the limit that the models become untethered from observations. I guess it could become larger if there are systematic biases.)

If nothing else, the reference time series of JRA-55 gets buried by the other lines: consider bringing it up to the top. (If you produced this plot with matlab, but want to keep it first in the legend, a solution is to just print it again.)

Figure 10: Are these 95% confidence intervals?

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879, 2018.