Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-515-AC2, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "How well do stratospheric reanalyses reproduce high-resolution satellite temperature measurements?" *by* Corwin J. Wright and Neil P. Hindley

Corwin J. Wright and Neil P. Hindley

c.wright@bath.ac.uk

Received and published: 25 August 2018

1. General comments

the figures are excellent as long as they are seen on a large screen: many require improvements to be readable on a print-out.

The reviewer provides specific comment about figures 1d, 6, and 12-15 below. We have addressed these individually, and discuss the changes made in the specific comments section.

I found that Section 4 is too long and technical. It could be summarized, moving some content to the annexes (e.g. Figure 5 and most of section 4.2). The really significant information in section 4.2 is about the specification of the sensing volumes of the instruments but this is not sufficiently detailed (see specific comment below).

We agree with this comment, which both reviewers made in some form. The current ordering arose due to the evolution of the paper: as originally planned it did not include the material after section 7 and as such was more focused on the method. We have now moved the technical details of the OIF, MIF and Core to a new Appendix, and replaced this section with a brief overview of the three components.

In my opinion this study has only one weakness: the vertical and horizontal dimensions should be separated in the comparison between full sampling and single-point sampling of AIRS observations (section 5 and appendix B). It is found that the added cost of full sampling is not justified when comparing reanalysis temperatures with HIRDLS data and seldom justified when comparing them with COSMIC and SABER data. This is an interesting (and comforting) outcome. But considering the viewing geometry of AIRS, the finding that it requires full sampling is quite trivial in the vertical dimension. For example in the case of constituent measurements, any comparison with nadir-looking instruments requires preliminary convolution of the model output by the vertical averaging kernels of the observations. Hence we are left wondering if full-sampling of AIRS in the horizontal dimension has any impact on the comparison. Of course there is no point in re-processing the whole dataset, but this question deserves either some discussion or (better) a sensitivity test. Here I would recommend to pick one year from one of the higher-resolution reanalyses and compare its fully (3D) AIRS-sampled dataset with another one where a more "usual" sampling is applied, i.e. 1D sampling in the vertical and bilinear interpolation in the horizontal.

This is an important question, and one that should have been considered in more detail in the original study. To properly assess this, we have re-run our AIRS sampling routine, using settings intended to simulate a 1D-only AIRS sampling while still allowing us to re-use the same software and thus to not introduce new inconsistencies. Specifically, we have re-run the AIRS analysis as described in the original text, except with no rotation in the horizontal or vertical and with horizontal weighting functions of width $\sim$100 m in both directions. Since the Core routine always centres the sampled points at the centre of the measurement volume and the grid spacing required for multiple points is of order kilometres, this results in a single column of points for each sample calculation, which are then summed using the same vertical weighting functions used for the original 3D analysis to produce equivalent synthetic measurements. We have then run this over all reanalyses for the year 2011.

We find that while a 1D approach does achieve most of the gains of the 3D approach, more than 50% of samples retain an above-instrument-precision difference, of up to 5 K. This has been added to the text both in the main body and as an additional section of the relevant appendix.

2 Specific comments and minor corrections

1. The abstract should explain in a few words the "full-3D sampling approach" to contrast it with the single point approximation.

The phrase "(i.e. one which takes full account of the instrument measuring volume)" has been added at the first mention of full-3D sampling.

P.2, lines 1-4: You state that the biases between the reanalysis and observed states are due not only to the multiplicity of assimilated datasets but also to the need to favour the model state for reasons of numerical stability and dynamical balance. This is a very interesting insight but it requires supporting references.

A relevant citation has been added.

P.2, line 29: please give a few words about the different diagnostics used in sections 6–9.

Done!

P.3, line 20: "they preserve"

'They' has been replaced with 'this', which fixes this problem.

P.3, line 21: re-phrase the sentence, e.g. "... while they are suppressed by the methods used to optimise the standard AIRS Level 2 product..."

The above change should also resolve this issue.

P.3, lines 25-26: it makes no sense to describe values derived from perturbations to synthetic data as "measurements" - please use better wording.

The sentence has been removed, as it doesn't add to the paper anyway.

P.4, line 7: Please provide specific references for the across-line-of-sight and along-line-of-sight resolutions of AIRS.

We believe this refers to COSMIC, for which we have added a suitable reference (Hindley et al, ACP 2015)

P.4, lines 29–31 and Figures 1–2: these are very helpful and informative figures but they require some details about the methods used to approximate the sensitivity of the *instruments. Note also similar question below (p.10, lines 28–32).*

See response below.

Figure 1d: it is not possible to distinguish between solid and dotted lines (except looking on a screen with very high zooming)

We have replaced this with a solid line, and modified the caption accordingly.

P.5, lines 13–14: "Each of them is widely used in the scientific community for a variety of purposes" - not yet for ERA-5 which was released very recently.

Clarified.

P.5 lines 15 and 16: please define "upper-atmospheric data" in this context. Consider using the word "upper" between quotes.

We have added the phase "(in this context, stratospheric and mesospheric)" to clarify our meaning.

P.5 lines 20-21: it is easy to be more specific. Consider: "COSMIC is assimilated by all reanalyses except for JRA-55 and JRA-55C, AIRS by most..."

Rephrased as suggested.

P.5 lines 21-22: The words "Beyond these details" and "extremely" are not necessary, and Fujiwara et al. (2017) is an introductory paper - not a special issue. Consider: "The S-RIP introductory paper (Fujiwara et al., 2017) provides a detailed summary of the key features of each reanalysis".

Rephrased as suggested.

P.5 line 26: I think that there is a description of ERA-5 either in the ECMWF newsletter or (better) in a dedicated ECMWF technical report. Please check.

We have added a reference to ECMWF Newletter 147, which summarises key information about ERA-5.

P.10, line 24: I am not an export in satellite viewing geometries, but this really puzzles me: when vertical viewing angles are defined from instrument nadir, limb scanning instruments should be defined as 90deg - not zero !?

This has been changed to 90 degrees. Our code specified a value of zero degrees as described in the original text, but also had the horizontal and vertical volume-width parameters for each limb sounder specified the wrong way around, which cancelled out the conceptual error in angle to produce the same final averaging volume (hence why we didn't spot the error in tests!).

P.10, lines 28–32: sensing volume parameters are an important input for this study, yet no sufficient details are given about this. Are the standard deviations in each dimension a constant for each instrument? If so, this should be written in a table. If not, what do these standard deviations depend upon? Latitude, longitude, date? Or do they differ for each observation depending on its context (e.g. surface albedo)? Please provide appropriate references for each instrument. Note also similar question above (p.4, lines 29–31).

A paragraph has been added to the definition of the OIF explaining the assumptions implicit in this specification.

P. 12, line 23: please take this opportunity to define the SPA acronym (and capitalize the first letters).

Done!

P.13, lines 10–11 and also p.32 lines 4–5: see general comment above – is this difference between SPA and full sampling due only to the vertical distribution of sensitivities or also to their horizontal distribution?

See discussion in 'General comments' section, above.

P.13, line 12: footnote is not necessary

Removed.

P.13, line 17: delete extraneous words "and Appendix A)"

Removed.

P.13, line 18: "for COSMIC and SABER data, in particular..."

Fixed.

Figure 6: it is not possible to distinguish between solid and dotted lines (except looking on a screen with very high zooming).

The caption has been modified to make clear that it essentially overlies the solid line.

Figure 6 and 7: Despite the explanation in the caption of Figure 6, the last line of text annotation (e.g. "CO=0.99C+3" or "SA=1.01E+-2") is unclear (especially for ERA-I

and ERA-5 where the "E" looks like scientific notation). It would be simpler to directly write the values of gradient and the intercept separated by a comma, e.g. "(0.99,3)" or "(1.01,-2)".

The annotations have been replaced with the form y=mx+c - this provides a compromise between the two forms and is clearer than the original.

P. 14, line 6: while discussing figure 6, please remind the reader that ERA-5 is not compared with HIRDLS because you study only the post-2010 subset of ERA-5 while HIRDLS ended in 2008.

A footnote has been added clarifying this.

P.15, lines 3–7: this is easy and interesting to check: are most outlying COSMIC profiles located close to the poles?

In fact, the data we present later in the paper already tests this, although we had not made the mental link! A cross-reference has now been added to the later section on geographically-localised comparisons, where we see that at 30 km altitude this does not appear to be a significant factor.

P.18, lines 5–9: this attempt to qualitatively discuss SSW interannual variability is inadequate. Since this topic is largely out of scope, it should be sufficient to simply list the largest SSWs while dropping lines 6–8: "The SSWs of January 2006, January 2009, March 2010 and January 2012 (Butler et al., 2017) are clearly visible at both altitudes, and all datasets show a near-identical response at the 30km level. However... "

Changed as suggested.

P.19, end of section 7: Figure 10 is not discussed at all. This should be done (e.g. there *is a large spread in annual cycles at 50 km) or else this figure should be dropped.*

A couple of sentences have been added describing this figure in more detail.

*P.19, line 14: "...acts as a 'true' estimate which the *reanalyses* are attempting to approximate."*

Fixed.

P.20, line 19: It may be worth mentioning that this "ability of the reanalyses to reproduce the observational record" is relatively low at 70km.

We have added the clause "however, it is also low in absolute terms due to the difficulties of modelling this atmospheric region and the limited data constraints available." to clarify this.

Figures 12–15: quite difficult to visualize (especially on paper) due to the monochrome colormaps.

This was an attempt to use a colourscale based on the SRIP multi-model-mean colour, which is a fairly murky brown. However, we agree these figures are hard to read, and have replaced them with a diverging red-yellow-blue colour table to emphasise the features better.

Captions of Figures 12–13: re-phrase "...indicate boundary full and partial SABER coverage".

Added the word 'of' to fix the sentence.

P.23, line 14: "... resolving..."

Fixed.

P.23, end of 23: on Figure 15 one also notes significantly larger RMSD in the winter polar latitudes. This should be highlighted and may be shortly discussed.

Mentioned and some discussion added.

Title of sub-section 10.2: this should not be identical to the title of section 10. For 10.2 I suggest "Co-located Cluster Analysis".

Changed as suggested.

P.29 line 19 This is still part of section 10. Replace "Section 10 suggests..." by "The previous section suggests..."

Changed as suggested - this was a hangover from an earlier version with separate sections for these topics.

P.26 line 31: remove extraneous ")"

Fixed.

P.28 caption of Figure 18: this is an unusual graph in our field and it plays an important role in the paper, so it is important to provide a clear and complete caption (i.e. "see text for details" is insufficient). Please repeat that the co-located measurement pairs are all at 30km, horizontal distance does not imply any information, and the ordering is chosen purely to produce a simple tree.

A more detailed description has been added to the caption, including all the requested points.

P.29 line 13: "... always required when comparing to AIRS,...": this may be true only w.r.t. the vertical dimension whcih would be trivial (see above)

See discussion in 'General comments' section, above.

P.29 line 14: "... and required in equatorial regions and regions of high gravity wave activity ..." . This is only a conjecture i.e. it has not been demonstrated in this study. I suggest to tone down the conclusion: "... and may be required in ..."

Changed as suggested.

P.29 lines 27–28: as I understand them your results are not about the variability between pairs of datasets but rather about the agreement between these pairs. If this is correct, consider replacing "...variability...significantly less..." with "...agreement... significantly better..."

Agreed and changed.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-515, 2018.