Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-119-RC2, 2017

Atmospheric Measurement Techniques Discussions

Open Access

Interactive comment on "Intercalibration between HIRS/2 and HIRS/3 channel 12 based on physical considerations" *by* Klaus Gierens et al.

Anonymous Referee #2

Received and published: 26 July 2017

The paper presents a reasonable attempt at a modest task, but fails to employ the necessary rigour and draws conclusions not justified by the presented evidence. In essence, the main conclusion of the paper is not justified. The material may be publishable with a weaker conclusion. Generally, the lack of rigour is somewhat worrisome; at one point the authors explicitly state that they do "not know" something that can be quite readily found out.

I find it problematic to call this a physics-based approach considering that the physical reasons for the differences in the measurement are mostly ignored. I know that modelling the instrument properly is a major undertaking and I do not expect the authors to do so for this paper, but I think the term "physics-based" promises too much. There are many other differences between HIRS/2 and HIRS/3, and even between HIRSes in the

Printer-friendly version

Discussion paper

same series, such as calibration, footprint size, and many technical details within the instrument, that should lead one to expect differences.

The authors show the differences between HIRS/2 and HIRS/3, but the SRFs are not consistent within the other HIRSes either. How do their differences compare to "usual" inter-HIRS differences?

Please see a list of recommendations for substantial revision below.

Page 3, line 6-7: Almost certainly the author mean the complete spectral response function for the entire optical path, not only for the filter. The overall spectral response function takes into account (in fact, is a multiplication of) all components in the optical path, including mirror, filter, detector, etc. In fact, some of those may degrade over the period of an orbit such that the overall SRF may change. In any case, I think the authors should use the term channel spectral response function (SRF) rather than filter response function, to avoid confusion about what is meant.

Page 3, line 16-17: “channel 12 weighting functions do not reach the ground”, this statement may be too general. It is true for the situation described by the authors, but is it also true in polar or high mountain conditions?

Page 3, line 16-17: “brightness temperatures from infrared sounders are only reported for cloud-free situations”. This statement is confusing. The authors must be talking about the simulations, not the measurements, which are all-sky top-of-atmosphere radiances. I suggest replacing "reported" by "calculated".

Page 3, line 30: I’m confused. Are you including the Sodankylä and Manus data for training, or merely for validation/comparing?

Page 4, line 12-15: If data are clearly wrong they shouldn’t be shown in the first place. I don’t think showing this tail in this figure adds any value to the paper. Just state in the text that a sanity check has discarded x/y, alternately retained (y-x)/y of the radiosonde profiles.

Page 4, line 22-23: The authors find Δ=6.5K, Shi and Bates find 8 K. What causes the 1.5 K discrepency?

Page 5, line 11: The precision of this number is very high. How is it calculated?

Page 5, line 23-28: The authors conclude that "One additional independent variable is clearly insufficient to capture all this variability.". As the authors show in Figure 5, the additional variable is not independent, but has a very strong correlation. They then proceed with the bilinear regression in equation (1) anyway, despite having concluded it is insufficient. This is somewhat surprising. Have the authors considered using a multilinear regression with additional HIRS channels, perhaps including those that add information on temperature (oxygen channels)?

Page 5, line 27 and footnote: I don't think the specific software is relevant (except in the case of open-source software, where the license may require a mention+citation in the acknowledgements, but that is not applicable here)

Page 6, line 9-12: I do not agree with the author's interpretation of the result. It would be easier if they would plot the difference as a function of reference, but even as shown it is evident there are large differences exceeding 4 K.

Page 6, line 15: Why are you showing such a generic weighting function? Please show actual weighting functions corresponding to typical radiosonde profiles for each station, both summer and winter, for all channels. That probably belongs in the methods section of the paper.

Page 6, line 23: Although it will happen less frequently, channel 12 in both cases may also see the ground sometimes. What does a typical Sodankylä winter water vapour Jacobian look like for those channels?

Page 6, lines 24–27: Certainly, in HIRS L1B data as obtained from NOAA CLASS, ground observations are /not/ flagged. Flagging in HIRS data is purely based on events on the spacecraft, such as bad calibration, mirror movements, etc. See the NOAA KLM

User's Guide and NOAA POD User's Guide. It should be easy to check whether the larger scatter is due to ground influence, by calculating corresponding Jacobians. You should also investigate what the winter vs. summer points look like for the Sodankylä method, and/or colour-code the points by IWV, which can be estimated from radiosonde profiles. And such a figure is more usefully plotted by Δ vs. reference, like was done for Figures 1 and 2 (this comment also applies to other scatter plots)

Page 7, lines 1-9: I don't understand what point the authors are trying to make here. What is the relevance or implication of this observation?

Page 7, line 19: For both the present paper and GE17, what are the uncertainties associated with the best estimate of those parameters?

Page 7, line 27: Why don't you show those data in a density plot as well, like for the other figures? Scatter plots have marked disadvantages (see Carr et al., 1987).

Page 8, lines 6-7: I disagree with the conclusion. You have shown that two independent methods give /consistent/ results, but that does not imply that the data series can be considered as homogeneous. You have shown the weighting functions for both NOAA-14 channel 12, NOAA-15 channel 12, and your estimated pseudo-channel 12, in Figure 7. To me, that figure shows it is /impossible/ to produce a homegeneous data series. The authors have reached this result — but apparently not this conclusion — in two independent ways. That is a useful (but modest and unsurprising) result. If a user wants to produce a climate data record (such as for UTH) spanning the entire HIRS series, s/he will in any case need to take note of all the inter-satellite differences (not only between HIRS/2 and HIRS/3, but between each and every copy of HIRS, none of which are identical) in their forward modelling.

Page 8, lines 26–28: Again — this conclusion is not justified. Just because there exists a correlation between A and B, does not imply that B can be fully predicted from A. It is trivial to construct a counterexample: just take B=2*A, then add random noise to both A and B. They remain strongly correlated but it is impossible to predict B from A.

Printer-friendly version

Discussion paper

Page 9, lines 4–6: For the reasons explained above, I consider this conclusion is not justified.

Page 9, line 13-17: I think where you obtained code and data is important information in the methods/data section and should not be tucked away in a couple of single-line paragraph near the end. The methods section also needs an illustration of the SRFs for both NOAA/14-12 and NOAA/15-12.

Page 12, Figure 2: consider showing a hexplot instead, as Carr et al. (1987) have proposed as a superior alternative to the scatterplot and hexagons have less distortion than squares or rectangles. See figure 10 in Carr et al. (1987).

Page 12, Figure 2: please use a perceptually uniform colourmap (see Borland and Taylor, 2007). The chosen colourmap goes from white to dark blue to light blue to dark red, which leads to misleading inferences. It was probably designed for showing diverging data, such as data that are either negative or positive. In recent years there has been considerable progress in the development of perceptually (nearly) uniform colourmaps that are now included with all popular data visualisation tooklits.

Page 13, Figure 3: Those figures are very hard to read. The colours are too light: green, cyan, and yellow are hard to read against a white background. It appears the authors may have used a plotting tool that defaults a black background, and have changed the background colour without changing the line colours. The bottom panel is completely impossible to read, because (1) there are too many lines, (2) the line colours, and (3) the reuse of colours. As the individual profiles are not important, I would suggest the authors plot the data as a large quantity of very thin black lines, which will give a visual effect of looking "dense" when there are many lines together, thus somewhat having the effect of a density plot. I have previously used lines of 0.01mm thickness to this end, with a visually pleasing result when plotting hundreds of lines. Although individual profiles will not be distinguishable in this case, that is not relevant for the message conveyed by this figure.

There appear to be some data problems in these figures as well. I don't know if fast variations such as 2000 07 01 12 in the middle panel between 90 and 80 kPa are correct, but the blocky structure at low RH, low pressure at the bottom panel clearly isn'.

Page 13, legend: The word "average" can be confusing because many people use it to mean "arithmetic mean". The authors may wish to choose a phrasing like "values near the median" or so.

Page 16, Figure 6: The authors should show the independent data on the x-axis and the dependent on the y-axis; need to swap the axis on this figure.

Page 16, Figure 7: Do you mean averaging kernels or weighting functions? In any case, they need be shown earlier as this is a fundamental instrument property, and thus belongs in the instrument description.

Page 17, Figure 8: The bottom of the label "number of events" is off the figure.

Textual:

- Page 1, line 20, replace "providing" by "provide" - Page 2, line 1: replace "an HIRS" with "a HIRS" - Page 2, line 14: replace "sufficent" by "sufficient" - Page 5, line 21: replace "sqare" with "square" - Page 7, line 15: replace "amazing" with "remarkable" - Page 8, line 22: Remove "Now the idea is, to", replace "take" by "Take"

References:

Carr, Daniel B., Richard J. Littlefield, W. L. Nicholson, and J. S. Littlefield. "Scatterplot matrix techniques for large N." Journal of the American Statistical Association 82, no. 398 (1987): 424-436.

Borland, David, and Russell M. Taylor Ii. "Rainbow color map (still) considered harmful." IEEE computer graphics and applications 27, no. 2 (2007).

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-119, 2017.