Interactive comment on "Radiometric consistency assessment of hyperspectral infrared sounders" *by* L. Wang et al.

Anonymous Referee #1

Received and published: 15 July 2015

1 General Comments

This manuscript describes the comparison of observations from 4 hyperspectral infrared sounders. Very important work for the satellite calibration community, and to support the generation of Fundamental Climate Data Records, not only from these instruments, but from other satellite instruments, which can be inter-calibrated against them.

However, there seem to be some shortcomings in the methodology adopted. In particular:

1. The analysis lacks a full uncertainty analysis of the comparison methods, and their results. This makes it difficult to judge their significance.
A: In response to reviewer's comments, extensive efforts have been made in the revised manuscript. First, sensitivity studies have been performed to examine the collocation method and criteria. Second, a section has been added to discuss the results with previous studies. We don't find the large differences between them. Third, a lot of detailed information has been added as the reviewer suggested. It will help the readers to judge our method.

2. The methodology is not always well justified - for example the collocation criteria.

A: Please see our response to another reviewer's comments on collocation criteria. Using CrIS-IASI/B dataset as an example, we performed sensitivity tests to check how the BT difference (CrIS-IASI) at 900 cm^{-1} changes as a function of relaxing the criteria as this reviewer suggested. First, the all the collocated data are filtered by limiting the standard deviation to mean ratio of the VIIRS radiances (M16) less than 0.05% but relaxing other collocation criteria (FOV distance to 13.0 km, time difference to 10 minutes, no angle limitation at all). Figure R1 shows the BT differences varying with FOV distance in the range from 0 to 13.0 km. Figure R2 shows the BT difference varying with time distance in the range from 0 to 150 seconds. Figure R3 shows the BT differences varying with angle distance (abs(cos(zen1))-cos(zen2))) in the range from 0 to 0.002.

Based on the above figures and Figure 5 in the paper, the following conclusions can be drawn. First, the FOV distance and collocation environment uniformity are the major factors that control the scattering pattern of BT differences and sample number. Second, the time differences of all the samples are less than 150 seconds. The 120 second threshold value only filtered out a few points. Third, since we only use the FOVS in two fields of regard (FOR) that have smallest scan angles (±1.67 degree), the CrIS FOVs' satellite local zenith angles range from 0.6-3.36 degree and IASI 2x2 FOVs angles range from 1.3-2.7 degree. As a result, all the samples meet the angle difference criteria. We made these points clear in the revised manuscript.

3. More care should be taken to minimise the potential for systematic biases in the instrument comparisons. For example, if the same time period is not used for each dataset, their double differences could alias components of the seasonal cycles in instrument calibration or scene radiances if their biases are radiance- dependent.
A: We agree. In the revised manuscript, we made this point clear.

4. The authors should be more specific when reporting their results - for example, specifying exact dataset versions, date and time periods. They often use more expressions such as "less than 0.02K in all bands", rather than specifying the exact magnitude of the differences - e.g. "with an rms difference of 0.01K". It is, however, appreciated that the relative difference between the channels can be radiance-dependent, which means that such summary statistics as "CrIS is 0.06K warmer than IASI on average" could be misleading. The manuscript would also benefit from more discussion on the following questions:
A: Except to the Abstract, we added a lot detailed information to discuss the statistics in the revised manuscript. Please see our responses in the following.

5. Can you make recommendations to how to address the differences in the instruments' datasets for different applications?

A: We added a section to discuss the results, which is related to how to address the differences for different applications. Basically, it can be done use the double difference method relative to RTM simulations.

6. Can we distinguish between spectral and radiometric differences?
A: Basically, the spectral differences should show the ringing patterns (BT differences jump up and down) around the spectral spectra absorption regions (e.g, CO_2, water vapor, and other trace gases). Here we can't find these features from our results. We think that these BT differences (espiecally between CrIS and IASI) are mainly dominant by radiometric differences.

7. Are differences simple functions of radiance in each spectral band?
A: BT differences versus BTs have been demonstrated in Figures 11 and 12. BT differences versus radiances should be similar to them.

8. Are the differences stable in time?
A: We continued this study for whole satellite overlapped period. So far, we find the BT differences are relatively stable when comparing one year's results with two years'. The two years' inter-comparison results can be found at,

Wang L., Yong Han; Yong Chen; Xin Jin; Denis A. Tremblay, 2015: Soumi NPP CrIS Radiometric Calibration Stability Assessment: A Perspective from Two Years' Inter Comparison with AIRS and IASI. *2015 EUMETSAT Metrological Satellite Conference*. Toulouse, France, September 21-25 2015.

9. Is there any relative scan angle dependence between the instruments?
A: Scan angle dependent bias can be caused by instrument polarization. So far, based on our best knowledge, only AIRS performed polarization corrections. It is a challenging task to characterize each instrument's own scan-dependent bias. We have tried to identify CrIS scan angle dependent bias using CrIS-VIIRS inter-comparison and pitch-maneuver datasets. However, the results are not encouraging. For scan angle dependent bias between the instruments, it is much harder to identify it based on inter-calibration methods.

10. The analysis of CrIS' performance is limited to the normal mode. It is now routinely operated in full resolution mode. What is the expected impact of this change?
A: This change doesn't impact on data user because the ground processing software - Interface Data Processing Segment (IDPS) was updated to process the full resolution raw data records (RDR) into the normal resolution sensor data records (SDR) in order to keep the consistency of operational datasets. In other words, the official CrIS SDR is still in normal resolution. Before and after these changes, we did some comparisona, the normal resolution data are consistent during this transition.

2 Specific Comments

2.1 Introduction

Line 101: Applying strict collocation criteria will certainly reduce the uncertainty on each collocation, but will not necessarily reduce the uncertainty on their ensemble mean. It should also be pointed out that it is important to ensure the full dynamic range of radiances are covered. Fig. 9 suggests this is OK - but nonetheless, it is worth highlighting in the text.
A: We agree, it reads as,
"With enough collocated measurements that is able to cover instruments' measure dynamic range , we believe that the strict collocation can reduce collocation uncertainties according to our previous studies and thus can effectively identify the measurement differences at an instrument calibration level if the size of comparison samples is statistically large enough (Wang et al. 2009a; Wang et al. 2009b)."

2.2 Methods

Line 209: The collocation criteria in Table 2 appear to be plucked from the either. Their choice (and the associated trade offs) should be justified in the text.
A: We agree. A paragraph has been added to discuss the collocation criteria.

Line 221: What is the impact of relaxing the time difference threshold to 15 minutes? For example, biases could be introduced if there is a systematic time difference between the observations. This could be quantified, considering typical rates of change of scene radiances.
A: We did some sensitivity studies, given the uniform collocation environments, the scattering pattern of BT differences are not sensitive to time differences.

Line 234: Using only uniform FOVs does not avoid these uncertainties - but it may reduce them. It could, however, be better to quantify these uncertainties (e.g. using a higher resolution collocated imager) and combine all FOVs in a statistically optimal way, in a way similar to the approach adopted within GSICS for the inter-calibration of geostationary imagers using collocations with hyperspectral sounders.

A: If we understood correctly, the reviewer suggested that we use scene uniformity as weights to perform the weighted-average to compute the statistics. Tobin et al. from UW used this method to compare CrIS with IASI. However, both methods give the similar results. The patterns and magnitude of BT differences varying with wavenumbers are close to each other, indicating that the method used to compute statistics are not sensitive. So we still use our old way.

Line 237: Using the high resolution imager will only be appropriate for the channels in the IR window. Channels in strongly absorbing H2O/CO2 bands will not be sensitive to the same variability - especially when caused by low clouds or surface conditions. The authors should consider using high resolution imagery from different VIIRS channels for the same purpose.

A: First, there are only three IR M bands are fully overlapped with CrIS, i.e. M13 (3.7μm) M15 (10.7μm), M16 (12.0μm). They are all in the window regions. NPP VIIRS does not have water vapor and CO_2 absorption bands. We agree that channels in strongly absorbing H2O/CO_2 channels will not be sensitive to the same variability - especially when caused by low clouds or surface conditions. However, the homogeneity from IR window channel is the strictest way to check the scene uniformity, though some samples are filtered out. In other words, if the IR channels are uniform, the strongly absorbing H2O/CO2 should also be uniform.

Line 249: Applying a threshold to this ratio of radiances will reject proportionally more cold collocations. The authors should consider applying a constant threshold to the standard deviation instead to allow more collocations at the low radiance end of the scale.

A: We disagree. The purpose of using the standard deviation to mean ratio is to treat the cold scenes and warm scenes in the same way. More often than not, the standard deviation of the scenes is modulated by the mean. In other words, warm scenes have larger standard deviation than cold scenes. If we use the constant standard as a threshold, there will be more cold scenes than warm scenes. We think that the standard deviation to mean ratio is a fair way to select the scenes.

Line 281: It would be helpful to also know the rms difference, as this can more easily be used in an uncertainty analysis than a limit such as "less than 0.02K", which requires an assumption of the statistical distribution of the differences.

A: We are a little confused. Figure 7b gives their spectral differences of $IASI_{CrIS}$ - CrIS in term of brightness temperature. CrIS and IASI spectra are simulated by LBLRTM using an identical atmospheric profile as inputs. $IASI_{CrIS}$ is the IASI spectrum that is converted onto CrIS wavenumber grids. Here we just want to show the uncertainty of this conversion. Figure 7 shows the ringing pattern with the mean of zero. This is not a statistical plot but purely the spectral differences (two spectra) varying with wavenumber. The 0.02 K value is the worst scenarios caused by the method. Is the RMS that this reviewer wanted is the value averaged along wavenumber? However, we do describe this part more specifically.

Line 294: This sentence implies the calculations are actually done in radiances - at least for CrIS vs AIRS. Is this the case for all instrument pairs? If so, it should be made more clear earlier in the article, as other sections seem to imply the comparison is done in brightness temperature space.

A: For CrIS and IASI, first, the IASI radiance spectra converted onto the CrIS spectra grids. Second, they are converted into BT spectra using Planck function. Third, their spectral differences are computed in term of BT.

For CrIS and AIRS, first, their spectra radiance averaged at 25 spectral regions. Second, they are converted into BT at each region's central wavenumber through Planck function. Third, their differences are computed in term of BT.

We made it clear in the revised manuscript.

2.3 Results and Discussion

Line 333: Again, please be more specific. Instead of "less than 0.2K", please specify their mean and/or rms difference.
A: There are a total of 1305 mean and standard deviation values at 1305 spectral channels. It is very difficult to specify these values in a table. If the reviewer strongly insists, we can provide table as a supplement.

Line 338: It should be pointed out that, given sufficient number of collocations, the radiometric noise on each sample should cancel out - and therefore not contribute to the mean difference.
A: Good Point. We added this point.

Line 349: More care should be taken to minimise the potential for systematic biases in the analysis of these double differences. In particular, exactly the same time period should be for each dataset to ensure components of the results do not include aliased signals of seasonal cycles in instrument calibration or scene radiances if their biases are radiance-dependent.
A: Good Point. It has been added.

Line 364: The fact there is no strong scene dependence in the difference could imply that their spectral calibration is consistent. However, the channels in the H_2O and CO_2 bands appear to have statistically significant slopes (at k=3 level). This may indicate a spectral calibration issue - or may be due or deficient spectral matching.

A: We agree that there are statistically-significant slopes for two H_2O and CO_2 channels. However, we disagree that it is caused by spectral calibration issue. Usually, the bias caused by the spectral calibration should have ringing patterns with large magnitude. The non-perfection non-linearity correction can also cause the scene-dependent bias. We reworded this part.

Line 368: The authors could suggest how the CrIS-IASI radiometric consistency be analysed at low latitudes, without direct collocations.
A: Agree. Several efforts have made through double differences relative to RTM simulations. We discuss it in the discussion section.

Line 373: How are the 3 datasets combined together? A simple mean? Or weighted according to the number of collocations and their geographic distribution?
A: This just a simple mean. It has been clearly described.

Line 380: The number of collocations included in the sample should be specified to justify the statistical significance of the results.
A: We added the number into the section 3.3 "Data Processing".

Line 393: How are these 25 spectral regions defined? Are they the same as used by Tobin?
A: This is different from UW and CNES. The 25 spectral regions are defined to minimize the effects of AIRS band channels.

Line 400: What is the rationale for choosing different spectral bands for analysis in Figures 14 and 15, compared to Figures 11 and 12? The authors should follow the same approach here - and fit a trendline to the differences, giving its slope and uncertainty.

A: We added fitting line, fitting coefficients, and their uncertainties in Figure 14.

Line 412: Logically the time series analysis should precede the other analyses, as it is only because the instruments' calibrations are relatively stable that the results can be combined in this way. More statistics of the time series of the instruments' differences should be included.

A: This is because we want to emphasize that the seasonal variation in Figure 15 can be explained by the scene dependent bias features in Figure 14.

Line 414: Again, please be specific about the magnitude of the seasonal variations.
A: Agree. We added the magnitude of the seasonal variations in the revised manuscript.

2.4 Conclusions

Line 433: It would be interesting to speculate on whether these differences could be related to difficulties in calibrating this band when CrIS is processed in "normal" spectral resolution mode.
A: During the ground software and instrument operation transition, extensive validation efforts have been made to make sure the CrIS normal resolution data consistent before and after transition. Based on validation results, we do not see any issues after satellite was switched into FSR mode. These differences can be caused by two instruments' radiometric calibration, e.g. non-linear correction, blackbody emissivity, PRT temperature drift.

2.5 Tables

Table 3

I found this type of table very useful, as it contains quantitative information and would also like to see counterpart tables for CrIS-IASI-A and -IASI-B.

A: There are a total of 1305 channels for inter-comparison. It is very hard to produce this table in the paper. We can provide the table as supplement document if this reviewer insists.

The specific period of data analysed should also be mentioned.
A: We added it in.

This table seems to have too many decimal places. The STDEV suggest that two dec- imal places should be sufficient. However, the number of samples should also be specified to allow the statistical significance of these values to be established.
A: We think that four decimal places clearly show the results so we still keep it.

It may be helpful to summarise the mean rms difference of all spectral ranges.
A: RMS can be roughly estimated from sqrt(mean^2.0+stdev^2.0) There are already too many rows in table 3 and I don't want make it too crowded.

2.6 Figures

Figure 1 is hardly referred to in the discussion, yet it provides a useful overview of the spectral coverage of the sensors. It would be helpful to add the wavelength on a second x-axis for readers without extensive experience of spectroscopy.
A: We added wavelength as a second x-axis on the top.

Figure 7: Again, it is difficult to judge from this plot the typical magnitude of the errors introduced by the spectral conversion process, as the line density saturates. It would be helpful to specify the rms difference in each band.
A: Please see our response on the explanation of Figure 7.

Figure 8: It would be helpful to mention that the blue dots show the mean difference in each of the 25 spectral regions. And even more helpful if error bars were also plotted to show the uncertainties of these differences.

A: Just like to Figure 7, Figure 8 gives the $CrIS_{AVG}$-$AIRS_{AVG}$ (single_value – single_value) at 25 spectral regions from the model-simulated CrIS and AIRS spectra. AIRS bad channels are set as missing values and then interpolated in order to match real observations. Here we just want to estimate the uncertainties caused by the average method. There are no mean and standard deviation values but only 25 values at 25 spectra regions.

Figure 9 and 10: I would find it helpful to indicate the standard errors on the biases. By my calculations these would be 15mK (SD 0.5K, n 1000). But I'm sure you can do a better calculation. The scale on the middle panels could then be zoomed to show the significance of these differences.
A: We don't understand why the standard errors on the biases are significant. It is just another way to show statistics. The bias plus standard deviation is more straightforward here.

Figures 11 and 12: Although the values of the slopes look very small, it still represents -0.12K and -0.34K over the range shown for the window channel. For completeness, it would be good to include the fitted intercept (and its uncertainty).
A: We added the intercept as well as its uncertainty as suggested.

Figure 14: The summary statistics should be shown, including rms difference, slope, intercept and their uncertainties.
A: We added the regression coefficients as well as its uncertainty as suggested.

3 Technical Corrections

3.1 Abstract

The authors should specify the period of data used in the comparison in the abstract.

A: Agree. It was added in.

3.2 Introduction

Line 66: I suggest removing "newly-launched", as that will soon be out-of-date.
A: Agree. It was removed.

Line 85: The text should be reworded to make it clear that there is not only one 300 km x 300 km area which contains all orbit crossing points.
A: We double-checked the reference. For IASI/A and IASI-B inter-comparison, the authors used the overlapped scenes with different view zenith angles. For CrIS and IASI inter-comparison, they still used the SNO observation but with relatively large areas. Here we referred to the later one.

Line 86-87: This sentence is not sufficiently unambiguous. How are they averaged?
A: We changed the word "averaged" to "collected". In their method, through collecting a large ensemble of collocated samples, it is believed that the spatial collocation errors are random and Gaussian distributed. A weighted mean difference between the two sensors was computed independently for each spectral channel by using the spatial variability of each SNO as the weights.

Line 99: Typo: hypespectral

A: corrected. This was caused by the conversion from word file to PDF file.

3.3 Instruments and Datasets

Lines 114-119: This paragraph repeats much of the material from the introduction and should be rationalised.
A: We agree. We removed some sentence with specific numbers in the introduction Part.

Line 123: It does not seem to be relevant to discuss this requirement here. It could be removed.
A: We agree. This sentence was removed.

Lines 129-133: This is probably more detail than is necessary and could be reduced.
A: We tried our best to make this part clear and organized.

Line 140: Typo: "Fields of Regard" not "Fields of Regards".
A: Corrected.

Line 154: If only the apodized spectra are used in this study, then the spectral resolution before appodization is not relevant.

A: Agree. We added one sentence to point out the spectral resolution after apodization.

Line 157: This sentence implies there was a change in the CrIS hardware or control software to switch to FSR mode. If it was just a change in the processing software, it should be made clear here.
A: Both have changed. The ground processing software was updated almost at the same time as the operational mode to process FSR RDR into the normal resolution SDR. We made this point clear.

3.4 Methods

Line 181: The fact that the satellites are at different altitudes is not really relevant to this discussion.
A: We just want to show two satellite on the same obit but on different altitude. So we still kept this sentence.

Line 184: Figures 9 and 10 show the SNOs do cover a wide BT dynamic range. So it would be more accurate to say here that they cannot be compared over their full dynamic range.
A: Agree. We changed it.

Line 190: Similarly, the word "relatively" could be omitted here.
A: Agree. It was removed.

Line 208: It would be helpful to add "(not shown)" in the discussion of the CrIS-IASI swath overlaps.
A: We added "nadir" when we discussed the Figure 3.

Line 250: It should be noted that the use of brightness temperature difference amplifies the apparent differences for cold scenes.

A: Agree. We added this point when we discussed the results in shortwave bands.

Line 260: It should be noted that these plots nicely how CrIS' lower spectral resolution smooths the line features and limits the observed range of radiances.
A: Agree. We added this point.

Line 312: This sentence is not clear and should be re-written. Also typo: "happed"
A: This sentence was reworded.

3.5 Results and Discussion

Line 334: I do not see the relevance of this statement. NWP centres typically include their own bias correction schemes.
A: This sentence was removed as suggested.

Line 343: It should be made clear that this discussion on the large spectral resolution in the SWIR band refers to CrIS.
A: Yes. We added it.

Line 365: typo: “implies” = “applies”.
A: Corrected.

Line 407: The cause could be described simply as “differences on the instruments’ non-linearity corrections”.
A: Agree. We modified this sentence.

3.6 References

Line 493: The article cited by Jouglet et at. does not appear in the proceedings of this conference.
A: Sorry. There is no article available for this study. It is the best reference that we can find.

3.7 Figures

Figure 2: Typos: ARIS=AIRS, 1 March = 1-3 March.
A: Corrected.

Figure 3 and 4: A color scale should be included, if only in the caption (e.g. blue=200K, red=300K).
A: We added it in Figure captions.

Figures 6, 9 and 13: Spectral gaps should not be interpolated by continuous lines!
A: These figures was re-made to remove continue lines.

Figure 9 and 10: These figures appear to be fuzzy - both when the PDF is viewed online and printed.
A: The red and blue shadow areas are composed by ~1000 single spectra here. We just want to show that the sampling spectra are enough to cover instrument dynamic range. This is the reason that spectra can’t been clearly seen.

Figures 11 and 12: Again, the y-scales could be zoomed on the top and bottom panels.
A: Basically, we are trying to show BT differences in the same scale for all channels, which clearly shows the different characteristics for different channels. Thus, we still keep the old plots.

Figure 13: The y-scale on the lower panel could be zoomed.
A: It was zoomed from [-1.0K, 1.0K] to [-0.8K, 0.8K]

Figure 15: ~~The y-scale could be zoomed.~~
A: **It was zoomed from [-1.0K, 1.0K] to [-0.6K, 0.6K]**