Review of egusphere-2025-379
Extension of AVHRR-based climate data records: Exploring ways to simulate AVHRR radiances from Suomi-NPP VIIRS data

By: Karl-Göran Karlsson, Nina Håkansson, Salomon Eliasson, Erwin Wolters, Ronald Scheirer

Reviewer 2

General comments

This study aims to simulate AVHRR radiances, with subsequently produced cloud products, from VIIRS radiances with similar satellite orbital configuration. The stated goal is extension of the CLARA CDR, and I believe the potential benefits and uses would extend beyond that. The authors' arguments are well constructed with conscious use of independent validation data, and the writing is strong. The clear presentation of results with added exploration of cloud parameter validation is appreciated.

I think the one major area of improvement falls in the area of the linear SBAF methodology explanation. It is stated that the linear SBAF is determined by looking at simultaneous observations, but it isn't clear to me how the authors would be able to separate spectral bias from radiometric bias with this approach. That is, the inter-calibration sequence has three components to account for, 1) radiometric bias, 2) spectral bias, 3) retrieval biases. The authors mitigate possible retrieval biases with careful angle-matched colocation, and then seem to claim that any remaining bias is explained by spectral differences. If that's true, then the authors would be suggesting that AVHRR and VIIRS have absolute radiometric consistency that is stable across the record. If this assumption is the case, then the authors should make that clear and share their justification.

Regarding discussion of the NASA SBAF tool: It should be noted that the NASA-derived SBAFs are limited by the fact that IASI covers the continuous spectral range of 3.60 - 15.50 µm, and thus computations that consider spectral response below 3.60 µm must rely on assumptions, which are likely imperfect. This should not be framed as a "problem" (Line 146) with the NASA-derived SBAFs, but rather a knowledge limitation due to an observation gap. The AVHRRs and VIIRS have significant response below 3.6 µm, thus computing SBAFs for these would be questionable to begin with. That is, if the use of such would not be recommended, then their impact should not be framed as "negative" (Line 146).

I recommend acceptance after the above minor issues are addressed, as well as the minor specific comments below.

Author response: *We thank the reviewer for these appreciating words. We notice that the main issues to discuss and improve upon aligns well with comments from Reviewer 1, so for parts of our reply we will refer to (or copy) the response to Reviewer 1. We provide our answers and explanations below.*

Specific Comments:

Line 87: "satCORPS" should be identified as "satellite cloud and radiation property retrieval system (SatCORPS)"

Author response: *Yes, we have changed this now.*

Line 139-141: Given the limited spectral range of IASI data combined with the effects of solar contribution in the 3B channel range, "serious deviations" should not necessarily be unexpected (see discussion above).

Author response: *Yes, we admit it was a too strong statement. We propose the following instead on line 139 and onwards:*

"This study initially tested various SBAF relations, primarily sourced from NASA (2016). Results were acceptable for most AVHRR channels, but for some channels (especially channel 3B at 3.7 µm), we encountered problems in using the VIIRS-based simulations. For example, night-time cloud detection significantly overestimated the low-level cloud amount. The cloud detection method used (CMAPROB, described by Karlsson et al., 2020) is a probabilistic method using all AVHRR channels. AVHRR channel 3B is considered the most crucial channel for this method's performance, especially at night. Only minor deviations from the original AVHRR channel 3B radiances significantly affect the results at night. The encountered problems were likely caused by the limitation of IASI not observing radiances for wavelengths shorter than 3.62 µm. Since the AVHRR channel 3B and VIIRS band M12 spectral responses both allow for significant contributions at wavelengths shorter than 3.62 µm (see Figure 2), this limitation can be substantial, especially since this affects in particular the contribution from reflected solar radiation, which rapidly increases with decreasing wavelengths. An effort to describe these contributions using Radiative Transfer Model (RTM) calculations was applied in the satCORPS tool, but this was made using assumptions, making results more uncertain."

Line 146: See discussion above regarding "problems" and "negatively."

Author response: *We changed the sentence on line 146 to the following:*

"Due to the uncertainties encountered for the NASA-derived SBAFs for this channel, we decided to proceed by calculating SBAFs from collocated AVHRR- and VIIRS-observed radiances."

Table 1: What's the reason for using Training 1, Training 2, and Validation as a naming convention rather than the more typical Training, Validation, and Testing?

Author response: *We cannot see that there is an obvious naming convention to be used here. For example, the "Testing" option is unclear to us (e.g., what's the difference from "Validation"?). However, we believe that what confuses the Reviewer is the indication that we use two different training datasets. But we think that everything should be clear from the statements made on lines 181-183 (repeated here):*

"The linear SBAF regression methods utilized the entire training dataset (dataset 1 and 2). For the Neural Network (NN) approach, training dataset 1 was used for the actual training and dataset 2 was used as a "during training validation dataset" to decide when to stop the training. The radiance validation dataset was used to evaluate the performance of all SBAF approaches."

Thus, the confusion probably comes from the fact that Neural Network training was done in two steps (as explained), where the second step involved an evaluation procedure to decide when to stop the training. We believe that the explanations on lines 181-183 should suffice for the understanding of the training, testing, and validation procedures.

Line 233: Please define "channel quota."

Author response: *It is simply one channel divided by another (e.g. M10/M6 or Ch3a/Ch1). An example will be given in brackets on line 233. However, we will now use the notation "channel ratio" in the manuscript to clarify it even further.*

Table 5: Were other activation functions tested, e.g., ReLU?

Author response: *No, we just kept the activation function we normally use. Different activation functions could be tested in the future, but we do not expect that to make a large difference.*

Lines 278-281: The authors should explain why > 0.20 and > 0.4 were specifically used. Was a sensitivity analysis done?

Author response: *Reviewer 1 had the same question. This is our reply:*

The COT threshold of 0.2 for validation of CFC is well-documented, e.g., in the CLARA-A2 validation paper by Karlsson and Håkansson from 2018. Notice the improvement from an earlier value of 0.3 as reported by Karlsson and Johansson from 2013 (https://amt.copernicus.org/articles/6/1271/2013/) that was seen when going from CLARA-A1 to CLARA-A2. The value has improved further for CLARA-A3 but only marginally. Down to this threshold value, clouds are generally detected well by the current CMAPROB method, i.e., with 50 % or higher probability, and the use of this threshold consequently gives the best and most reliable validation scores for CFC compared to not doing any COT thresholding at all. But, if we are able to detect thinner and thinner clouds, it doesn't automatically also lead to improvements CTH improvements, simply because the complexity for the semi-transparency correction increases for thinner clouds.

The threshold to be used for CTH validation is currently more arbitrarily chosen. It should be higher than 0.2 but not drastically higher, since cloud detection efficiency increases rapidly for COTs larger than 0.2 (as illustrated by Karlsson and Håkansson, 2018, Figure 6). A value of 0.4 was found reasonable to at least avoid some of the uncertainties with very thin clouds. At the same time, the threshold should not be too high, so that it does not give justice to all efforts to correct for semi-transparency (i.e., it would not make sense to only look at opaque clouds). We could have produced results for a wide range of COT thresholds, but we found it appropriate to stay with the established method that we have recently used regularly for evaluation of results from CLARA-A2 and CLARA-A3 in standard EUMETSAT validation and review studies. We also don't think that results should be restricted to only include COT-thresholded results, since it always has a value to also see unfiltered results, as it is generally easier to see improvements when introducing new or upgraded methods (at least for CFC).

We also introduced the following comment on line 400:

"The use of different COT thresholds here is motivated by the wish to avoid the thinnest clouds detected by the cloud masking procedure, since these are always the most difficult clouds to deal with for the CTH retrieval."

As a last remark on this issue, we remind the reviewer that we have made a small change to the content of tables 7-9 after discovering that the category "No SBAF" was not adequately described in the first version of the manuscript. As a consequence, two alternative interpretations of "No SBAF" have now been introduced in the revised tables. For further clarification, the reviewer is asked to read the discussion of tables 7-9 in the reply to Reviewer 1.

Line 301: Table 2 shows the spectral relationship of VIIRS and AVHRR channels. It does not explicitly show how radiances compare unless such is broadly inferred from the spectral information (which would be scene-dependent). The authors should clarify how the table shows the radiance comparison, or is this a mistake?

Author response: *This sentence is definitely wrong. Thanks for discovering, it is now removed.*

Lines 543-544: This concluding statement regarding "implicit object type identification" is excellent.

Author response: *Thank you very much!*

Lines 545-547: I agree that using more VIIRS channels in the NNet would lead to better-simulated AVHRR channels with improved consideration of specific scene types. Is there a reason the authors did not use additional training channels to being with? There is seemingly no need to limit selection to comparable channels when it comes to training purposes.

Author response: *The reviewer is absolutely right in that theoretically one could have used all available VIIRS channels for the training. But for practical reasons, adding more input will also need more processing time and more processing resources. However, we did test various setups and the chosen one seemed to be the most appropriate setup. The various tests were actually only with fewer channels than what was finally used. We started with the channels we definitely needed, then added the 8.5 µm channel to have one more channel during night where cloud products initially looked worse. Finally, we added the 1.6 µm channel to get one more visual channel for day/twilight. The processing time is maybe not the biggest problem, but with every new channel added there is a need to do validation to make sure we do not get unexpected results and new problems.*

Lines 572-573: It should be mentioned that the basic radiances also would need to maintain stability.

Author response: *This was also one of the major points brought up by Reviewer 1. We are now adding the following sentences in the Conclusions section:*

"

In order to check the validity of the derived spectral adjustments on a longer timescale, it is suggested to regularly repeat the validation efforts on derived cloud products based on new observations from active sensors on EarthCARE (Illingworth et al., 2015) and similar missions in the future. In addition, regular checks of Simultaneous Nadir Observations (SNOs) at high latitudes between VIIRS and IASI (and its successors) can also help in deducing the infrared bands spectral responses' stability. The possibility to make similar checks for VIIRS and METimage visible bands seems unfortunately not possible from Troposphere Ozone Monitoring Instrument (TROPOMI, Veefkind et al., 2012) due to very limited spectral coverage of the AVHRR-heritage channels. However, it is possible that measurements from the Ocean Color Instrument (OCI) sensor on the Plankton, Aerosol, Cloud, and ocean Ecosystem (PACE, Gorman et al., 2019) mission and the Earth surface Mineral dust source Investigation (EMIT, Connely et al., 2021) sensor onboard the International Space Station (ISS) can be used. Some reference measurements from the Sentinel 3 Ocean and Land Colour Instrument (OLCI, Donlon et al., 2012) measurements could also be used. Although not purely hyperspectral, OLCI contains 21 bands in the 0.4-1.0 µm range and has an equator overpass time at 10:00 AM."