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

Overall Recommendation

This paper presents a novel method to simulate AVHRR radiances of the NOAA satellites from the Suomi-NPP (VIIRS) imager. The validity of the simulation method is confirmed by comparing the cloud products based on simulated data against CALIPSO cloud products as well as against CLARA cloud products based on original AVHRR radiances.

The topic of this paper is of great interest. The proposed methods are key to developing data records for climate applications. The simulation method confirms that simulated radiance can be used to increase the frequency of available observations per day, as well as to extend the data record in time by using observations from future (or may be even past) satellite imagers.

The manuscript needs to explain better how the presented work is related to other initiatives with respect to developing homogenised records of satellite radiances from observations of multiple satellites. Moreover, the manuscript appears to have been written in haste and therefore needs to be checked and revised carefully. Among others, the authors need to review the wording of their manuscript and assure to use similar terminology throughout the manuscript. Similarly, the style of figures and tables may be harmonized better for the sake of improving the presentation of the work.

The manuscript needs some major, but mainly minor revisions before it can be published. Below some general remarks followed by a chronological list of minor points of criticisms are given.

General Criticisms

Abstract

The story line may be presented clearer. For non-insiders the objective of the paper is not clear enough. Could you have another look at it and try to introduce better the purpose of the paper.

Introduction

The author focuses on using the method to simulate AVHRR radiance on extending their data records with observations of future satellite, thus extending the number of years covered. However, the method can also be applied to increase the number of satellites observations per day, for example by adding simulated AVHRR radiances from MODIS data, can the authors elaborate on the possibilities of that?

Handling changes of the SBAFs over time

The authors derive their SBAFs using collocations between VIIRS and AVHRR during the years 2012 and 2013, offers just a brief or limited view of a situation, providing a limited rather than a comprehensive or long-term perspective. Instrument spectral response functions, like those from VIIRS, are subject to changes over time. Can the authors add a discussion explaining how to handle changes in the SBAFs over time? Note, within the FIDUCEO project much emphasis was given to quantifying channel degradation and reconstructing changes in channel spectral response over time (Rüthrich et al., 2019, https://doi.org/10.3390/rs11101165).

Using collocations of two instruments to derive SBAF is not only mitigating differences in SRF but also considering any radiometric biases these measurements may have. And these biases may not be static in time, therefore the relationship derived between the instruments may not hold for another period outside of the training. Using collocations, you are basically inter-calibrating the measurements, not just deriving spectral band adjustment factors.

Deleted: SRF, but

If one needs to account only for the SRF differences, it is advisable to simulate both AVHRR and VIIRS measurements from hyperspectral measurements such as IASI (for IR channels) and SCIAMACHY (VIS channels) and use either of the methods suggested in the paper to derive SBAF, as you already mentioned in the manuscript.

Link to GSICS and FIDUCEO

The work presented is related to work done by the Global Space-based Inter-Calibration System (GSICS) international partnership, an initiative launched in 2005 by the World Meteorological Organisation (WMO) and the Coordination Group for Meteorological Satellites (CGMS). Further the work has many elements touched upon in the framework of the FIDUCEO (Fidelity and Uncertainty in Climate Data Records from Earth Observations) Horizon 2020 project funded by the European Union. Within GSICS and FIDUCEO good practices and common terminologies for doing harmonization and homogenization - all satellites are forced to look like a "reference sensor", AVHRR on NOAA-19 in this paper - of level-1 observations. Hereto:

- Can the authors explain in their introduction how the presented work relates to what is being done within GSICS and FIDUCEO?
- Can the authors use the terminology suggested in GSICS and FIDUCEO, and avoid confusing terminology in different fora?

For illustrative purposes here a schematic representation of the homogenization principles

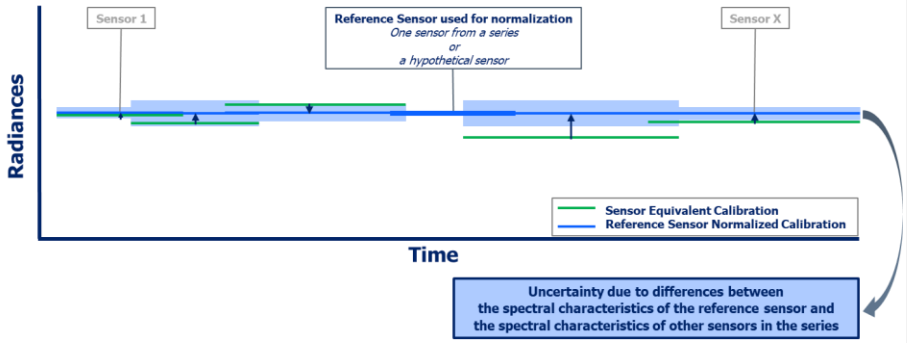

Table styles
The paper comprises many tables. The tables all have different styles (width, font size, font type, border styles, shading styles): Although I understand it is difficult to completely align table styles, I still suggest trying to homogenize the styles of the tables in the manuscript as much as possible, following the style suggested by egusphere.

Figure style and quality
The paper comprises many figures. The figures different in font type, font size, resolution, eg font size of Figure 5 is much larger than of Figure 9. Could the authors try to align look and feel of the figures, to make the paper optically more attractive. As above, please try to follow style suggested by egusphere.

Minor Criticisms

Introduction

Line 45 "Reanalysis datasets are undoubtedly capable of providing the best possible description of the Earth's atmospheric and surface state evolution, at least over the last 3-5 decades

Also good to mention here that reanalysis is designed to be physically consistent. This is an advantage and a disadvantage at the same time, as this consistence may mast out actual information that is not well covered by the current physical description.

Please see (Roebeling et al., 2025, BAMS, accepted) who write:
Note that reanalysis products from ECMWF, like ERA5, are indeed designed to be internally consistent but are not independent, and thus one needs to be cautious when using reanalysis data for studies asking for multiple data records that are independent of each other

Methodology

Line 75 "..estimated that SBAFs can explain up to 80 % of the variance …"
The number of 20% of variance not explained surprises me. It is a rather large number, that I was not expecting for narrow-band visible and near-infrared imagers. Could you give examples of the channels of instrument pairs were 20% of the variance is not

explained. In addition, I would expect that the bulk of the channels are clean channels (in which both instruments little absorption lines) and the explained variance is much higher. Please elaborate.

Line 81 "These can be derived from direct inter-comparisons of spatio-temporally collocated measurements from the two sensors"

A good reference for this sentence would be Meirink et al. (2013) https://doi.org/10.5194/amt-6-2495-2013

Line 90 "Notice that our target is the third version of this sensor (AVHRR/3) as carried by the NOAA-19 satellite"
Consider referring to the AVHRR/3 sensor on NOAA-19 as "the reference sensor"

Line 95 "In this study, we are not interested in simulating AVHRR channel 3A, as shown in Table 1. The reason 95 is that satellites carrying the VIIRS sensor follow an afternoon orbit, a sun-synchronous path with a daytime equator crossing shortly after noon."

Could this still become of interest in future work? This would for example be the case a future version of CLARA would try to increase the diurnal observation frequency and add simulated observation of the 'reference sensor' based on MODIS.

Line 129 "The CLARA-A3 CDR is based on AVHRR data with a coarser resolution than the nominal horizontal resolution. The archived global AVHRR dataset is stored in a format called Global Area Coverage (GAC) with a horizontal resolution of approximately 4 km (Kidwell, 1991)."

I suggest shortening this sentence to
"The CLARA-A3 CDR is based the archived global AVHRR dataset is stored in a format called Global Area Coverage (GAC) with a horizontal resolution of approximately 4 km (Kidwell, 1991)."

Line 129 "..an equivalent format."

I suggest rephrasing this part to
"..an equivalent horizontal resolution."

Line 139 "..various SBAF relations, …"

Do I understand well that NASA developed a set of different SBAFs? What is the reason for having different SBAFs, is it related to the surface underneath, or atmospheric absorption features? Please elaborate.

Line 160: Figure 2

For clarity, could you also show the local crossing time of Suomi-NPP in this figure?

Line 180: "The linear SBAF regression methods utilized"

From this line I understand that the paper evaluates two SBAF methods. Could you use, throughout the paper, the same terminology for these two methods, as you do nicely in the next section, but did not earlier in the paper, eg.

1. SBAFs derived from linear regression (referred to as SBAF-Linear1a and SBAF Linear1b)
2. SBAFs derived from a MultiLayer Perceptron (MLP) neural network (referred to as SBAF-NN)

Line 209: " eg for harmonizing"

Do you mean here harmonization of homogenizing. The difference between the two terms are explained in the FIDUCEO project, see https://www.fiduceo.eu/vocabulary

Harmonisation
A harmonised satellite series is one where all the calibrations of the sensors have been made consistent with (a) reference dataset(s) which can be traced back to known reference sources, in an ideal case back to SI. Each sensor is calibrated to the reference in a way that maintains the characteristics of that individual sensor such that the calibration radiances represent the unique nature of each sensor. This means that two sensors which have been harmonised may see different signals when looking at the same location at the same time where the difference is related to known differences in the responses of each sensor such as differences in the sensors spectral response functions. Harmonisation can be achieved to within an uncertainty that should be estimated, and the uncertainty contributes to the component of uncertainty that is common across the whole record of a given sensor.

Homogenisation
Unlike harmonisation, homogenisation is where all satellites are forced to look the same such that when looking at the same location at the same time they would (in theory) give the same signal. In reality the signals from different sensors would be different and homogenisation is adding in corrective terms to each satellite to make them look the same. It is likely that these corrective terms will not be 100% effective and that the process of homogenisation will add in scene dependent errors to the uncertainty budget which may be difficult to assess.

Line 210: Difference between Linear-1a and Linear-1b
Could you provide more information about the difference between Linear-1a and Linear-1b? I ask because in Tables B1, B2, B3, and B4 both methods appear. The authors indicate that Linear-1b separates day, night, and twilight, does this mean you separated your data pairs in three groups to do the Linear-1b regression, if so, what criteria were used for separating the data pairs?

Line 226: "Definition and training of the MLP network "

To be consistent in terminology consider using

"Definition and training of the SBAF-NN"

Line 300: "Table 2 … "

Should this not be another table? Table 2 shows the spectral bands of VIIRS much earlier in the paper, but no results. Please check cross references to tables and figures throughout the manuscript.

Tables 3, 4, 5, 6
These are very technical tables; I propose to move these to an Appendix.

Figure 3, 4, 5
The plots in these figures do not show regression statistics (slope, offset, and correlation), making that I can only compare the results qualitatively. The authors provide some of these statistics in the Tables B1, B2, B3, and B4. What are the considerations to leave slope, offset, and correlation out of these tables.

Note that much of the differences in the scatterplot result from collocation and synchronisation differences and do not talk so much about the performance of the SBAF correction. In that sense Figure 6 is a much better figure to illustrate the difference between the methods. Basically, it suffices to only present Figure 6, and leave out Figure 3, 4, 5. Can the authors comment on this?

Tables B1, B2, B3, and B4 present the most important quantitative results of the paper. I propose to either make these tables part of the main text, or to refer to Appendix B in the captions of Figure 3,4,5, or make the Tables B1, B2, B3, B4 in Appendix B part of the main text?

Figure 6
This is a very informative figure. Similar as above, also here good to refer to table that presents the statistics to this figure, i.e. B1, B2, B3, or B4. Do the results shown in Figure 6 represent the results of for the radiance validation scores for ALL cases (Table B4)?

Table 7, 8, 9
The result of CFC bias and CTH bias do not match very well. The authors do not describe this in the text. I understand that the Calipso observation do only become representative when a small COT threshold is set, to ascertain that low concentration of small particles in thin air is excluded. If this is done, the results become more robust. Can to authors explain this in the text and only keep in the statistics for CFC bias (COT > 0.2) CTH bias (COT > 0.4).

Table 7, 8, 9
Why are the COT criteria for CFC bias (COT > 0.2) different from the one for CTH bias (COT > 0.4)?

Line 414 "The SBAF-NN method has the best overall scores for the VIIRS/VGAC simulations validated against all CALIPSO-detected clouds."
Judging from the statistics in Table 7 none of the methods seems to stick out with better results. It appears even that VGAC no SBAF has very good results. Please explain.

Figure 7, 8, 9
These figures present scatterplots of data pairs of cloud property retrieval from two different satellites. Although the authors did everything to collocate and synchronise the data pairs, much of the observed scatter is simply the result so collocation and synchronisation differences and does not talk so much about the performance of the SBAF correction. To mitigate this, it would be better to present the results are frequency distributions of differences as is done in Figure 6. Can the authors comment why they have chosen to use scatterplot instead of frequency distributions to present the results?

Table A1 - A4
Table A1 - A4 presents the slope and offset of the linear regression on the training sets for Linear-1a and Linear-1b. The tables mention channel numbers and not channel wavelength. As the manuscript mostly refers to the wavelength and not the number of the channel, can the authors add the wavelength in brackets for each channel?