Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2019-104-RC2, 2019

Open Access Earth System Science Data Discussions

Interactive comment on "Cloud_cci AVHRR-PM dataset version 3: 35 year climatology of global cloud and radiation properties" *by* Martin Stengel et al.

Anonymous Referee #2

Received and published: 12 August 2019

Summary of paper: The authors describe a data set of global cloud properties based on AVHRR observations, available since 1982. The data set is an update of version 2, with the main changes the use of artificial neural networks for cloud mask and cloud phase detection, and additional cloud radiative properties. Both versions 2 and 3 are evaluated against the best available retrievals from other satellite and ground-based products. Standard verification metrics indicate overall improvement in most cloud properties, with some deterioration in ice cloud top height. The cloud radiative properties compare well against CERES observations.

Review: This paper is generally very well written. It is mostly complete and useful

information for anyone wishing to use this data set. Comments mostly concern some further clarification sought and perhaps slightly different presentation of the differences between version 2 and 3. The DOI links to a suitably presented web page describing the data. Overall, the recommendation is to accept this paper with minor corrections.

Minor comment:

Statistical significance. It is not immediately clear whether any of the differences in skill between v2 and v3 reported in the tables are statistically significant, although the large sample suggests these are. However, it should be possible and it would help the reader if the maps in figures 1, 2, 5, 6, 7, and 8 could include (i) difference/bias maps and (ii) stippling/hashing/shading for statistically significant differences. Most of these maps are visually similar and might hide key differences due to the colour scale used. A different way of presenting the various data sets, including additional maps of bias and statistically significant differences, would help inform the reader how the new data set compares against existing data sets.

Line-by-line comments:

p2. l28-30. This sentence is difficult to read, especially the first part.

p2. l33. "limitations". Limitations to do what?

p3. l3-7. Please provide references for the WCRP GEWEX data and the ISCCP DX data.

p3. l10-14. "based on the rationale above". It is clear why these data are required, compared to the MODIS/CERES and GEWEX data sets. However, what are the other data sets based on AVHRR lacking (PATMOS-x, CLARA-A2, Cloud-cci) that this paper will address with Cloud cci 3? A sentence on p2, line 19-21 would help clarify the shortcomings of those existing data sets.

p3. l20-31. This paragraph pre-empts the findings ("superior") and methodology. The relevant information is better placed in section 2.1.

p4. l2. Please add that table 1 contains all abbreviations used throughout the text. CER had not been introduced in the main text prior to p6 and it took a moment to figure out its meaning.

p4. l16. "much larger set". How do the two sets of training data compare? Did both v2 and v3 use CALIOP, but v2 just used fewer overpasses?

p6. l16. A "lower" CTP mean is not explained by more very low-level clouds, which have higher CTP. It appears that over the West Pacific and Maritime Continent, mean CTP has generally increased, which could be due to detection of more low-level clouds. Please re-consider this statement.

p7. l1. Regarding the validation, did the authors consider performing the validation separately for daytime and nighttime observations? The algorithms use different channels and the authors consider nighttime COT and CER "experimental". It would be useful to understand the algorithm performance for different times of the day.

p7. l3. Please, briefly explain how the collocation is carried out. In particular, what is the impact of the temporal mismatch between CALIOP and AVHRR? And what is the impact of the mismatch in footprint?

p7. l22-23. Why would improved identification of liquid clouds lead to reduced POD for ice clouds? This suggests that some ice clouds are now erroneously identified as liquid. Does that mean there are more "false alarms" in terms of liquid cloud detection?

p8. l5-15. It would be helpful to consider the results from Tables 4, 5, and 6 through a visual comparison, as done in Figure 4. A scatter plot (or 2D histogram) of CTH, LWP, and IWP comparing the data set with the "truth" could help identify where biases are most likely to occur. For instance, the CTH bias of ice cloud could be mostly due to the highest clouds, even at high COT, as these might have a region of low extinction coefficient near cloud top, that would lead to higher CTH in CALIOP. A scatter plot could show this clearly. Similarly, LWP and IWP are highly skewed variables and the metrics

presented could be affected by a few outliers. A scatter plot or 2D histogram (perhaps shown on a logarithmic scale) could indicate whether LWP and IWP estimates are typically good, or whether there is a consistent bias across cloud types of all LWP and IWP values.

p9. p10. p11. Please rename standard deviation to "root mean squared error", which is presumably what is reported.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-104, 2019.