Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-307-RC1, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Detailed characterisation of AVHRR global cloud detection performance of the CM SAF CLARA-A2 climate data record based on CALIPSO-CALIOP cloud information" *by* Karl-Göran Karlsson and Nina Håkansson

Anonymous Referee #1

Received and published: 13 November 2017

This manuscript evaluates the cloud mask of the CLARA-A2 climate data record (based on passive imagery from AVHRR polar orbiters) with collocated active cloud detections (CALIOP). Another, more general, paper has been published in ACP this year, and this AMT paper focuses exclusively on the cloud mask. This approach is sufficiently well justified, but the paper under review relies too much on the earlier publication (Karlsson et al., 2017; also to some extent on Karlsson et al., 2013) to explain the background. In order to qualify for publication in AMT, revisions need to be made to ensure that it can stand on its on, while not replicating too many of the science results.

In its current state, the paper is hard to review because some of the concepts are not explained sufficient well (specific examples are given below), and because details are left out. In addition, the manuscript is unnecessarily wordy in some places and has basic deficiencies with English/Grammar (for example, "were" is used instead of "where" throughout the manuscript; there are many run-on sentences; punctuation is used too sparingly; use of slang words such as "punish" for a statistical approach that are frequently used by the community, but should be used only where absolutely necessary). Before going into the copy/edit process at AMT, a native speaker should be consulted to ensure logical flow and readability of the manuscript overall.

Despite the criticism of the presentation quality, the content is interesting in that the cloud detection capability is studied as a function of optical thickness and region. Obviously, the POD (probability of cloud detection) depends on surface albedo and emissivity, mechanisms that are identified by the authors. Two comments here:

1) It should be stated more clearly where such findings have been made previously. The author make a point that the regional assessment is new, but there have been previous studies that focused on some of the problematic regions specifically in the Arctic with CALIOP that are not cited here (for example, studies by Gettelman, Kay, L'Ecuyer and a few others).

2) It remains unclear (partially because of the structural problems of the manuscript pointed out above) why there are some regions where cloud cover is overestimated by the passive imagers. One possible explanation is not sufficiently investigated: sub-grid resolution clouds that could be picked up by passive imagers but not by active imagers (if they are outside the FOV). There is some discussion of it, but it remains superficial. Also, active observations are portrayed as the ultimate "judge" for the performance of the cloud mask derived from passive observations, and they shouldn't be. As pointed out by the authors, active observations have their own limitations (sensitivity, FOV, day-vs-night contrasts). The truth is that active cloud observations afford a different perspective on clouds that happens to be less sensitive to the surface reflectivity and

emissivity than that of passive observations. This distinction (and the limitations of both approaches) should be made clear by the authors.

Sequential comments:

General language comments are provided below, but they are far from complete and should only serve as examples to go through the manuscript as a whole before submitting the revised version. Below, a few specific comments regarding the scientific content are given.

p2,L60: Why is CALIOP singled out as important for cloud observations, where in fact MODIS is flown in the A-Train as well. Wouldn't the MODIS observational record, in conjunction with CALIOP, lend itself to a similar study as the one presented here? Of course, its data record is much shorted, but on the other hand, MODIS and CALIOP are collocated all the time, by design.

p2,L60: The limitations of CALIOP (e.g., day time vs. night time detection, noise etc., strategies for thin cloud detection) should be discussed here.

p2,L70: The earlier study by Karlsson is cited here. It should be summarized in at least one paragraph since this paper needs to stand on its own. What was the scope of that manuscript? The extension by CALIOP, on the other hand, are well explained (with the caveats pointed out above).

p3,L79-87: This paragraph should be completely rewritten. The explanation of the field of view of the passive vs. the active instrument is vital for understanding this manuscript, yet it is incomplete. What is the GAC FOV vs. the FOV(passive) vs. the FOV(active, at native vs. aggregated resolution)? What data specifically are dropped? The best way to explain this would be through a simple illustration of the AVHRR pixels vs. the CALIOP FOV of single shots, as well as the aggregation of individual pixels/shots in the various products used in this study. Without this added figure, it will be hard to retrace the steps that were taken in this manuscript.

p3,L90: Which parameter retrievals? How is the radiance inter-calibration and data record homogenization done? Simply referencing Heidinger will not do because the specifics are missing. One of the clear requirements of AMT publications is that anybody reading the paper needs to be able to retrace the steps of a study from the original data to the findings. There is not sufficient detail provided here (or in other parts of the manuscript) to do that.

p4,l118-127: See comment above. These sections cannot be understood without better explanations of the FOVs, data aggregation and homogenization.

p4,l129-134: Provide description of specific NOAA orbits that were included (vs. those that were not). Also, why were MODIS observations NOT used? The minimum information for the NOAA observations are: (a) instrument/satellite names and short description; (b) orbit inclination and equator crossing time; (c) life time of satellite; (d) orbital shifts over time

p4,l148: The theoretical deliberations on cloud mask/cover are insufficiently backed by literature. The paper that comes to mind when talking about the meaning of a "small" or "thin" cloud is that by Koren ("How small is a small cloud"). A short literature study on the topic would be advisable, given that it is the main topic of this article.

p5,l184: "possibly punish AVHRR-based methods in an unfortunate and undeserved way. . .": three words (punish, undeserved, unfortunate) are inappropriate for a scientific publications. There are multiple occurrences of such "personalized" or "humanized" comments, which should all be translated into objective, rather than "punitive" language.

p5, l187: The optical thickness threshold of 5 for CALIPSO is higher than usually assumed. If it is necessary for this study to work with such a high threshold, it should be justified, and it should be explained how this is possible (referring to literature where this has been done, or with a dedicated sub-section in this manuscript where it is shown that the lidar does, in fact, allow to go to COD 5, and under which circumstances).

p6: There are multiple gaps on this page: The notion of “scores” (and different kinds) are used without sufficient (or any) explanation in this section, or in section 3.3. Too many questions remain, for example, which parameter of what satellite is validated with which other parameter, and how exactly as “score” (of any kind) is established. How is the aggregation done? Why are scores only plotted as a function of COD up to 1, where in fact CODs up to 5 are advertised? What is the “improvement”? If the figures are insufficiently explained, it is not possible to understand. What has been “transformed from cloudy to clear cases” (l212), and how is that done? What is the role of Kuiper vs. hit rate (should be spelled “hit rate”, not “hitrate”). Each of the bulleted items of the list on p6/p7 need to be explained and supported with formulae where appropriate. Here again, terms such as “punishing” should be avoided if at all possible. After this paragraph, the reviewer was unable to give this a thorough review because the basics for understanding the remainder of the manuscript were not established. The review is willing to review another version of the manuscript where this has been fixed.

p7,l265: This question is a great one, and at the center of this manuscript. However, the method description below is insufficient. Terms from machine learning (“overtrained”) are evoked without explanation how they relate to the manuscript content. Also, here again, CALIOP is represented as the “objective” instrument that AVHRR is validated by where possible – where in fact the two instrument just assess different aspects of a cloud (see comment above).

l278-l304: This seems to wordy and hard to follow since some of the concepts were not introduced.

l306: Now some of the orbits are introduced, but that is too late in the manuscript. In addition to NOAA-18 and NOAA-19, did other data go into the CDR under investigation?

p9,L350: insufficient introduction how systematic and random errors were establish

make it hard to understand Figure 7.

p9,l355-359: Add explanation why AVHRR gives higher cloud cover. It is easy to imagine a scenario where small cumulus clouds would be picked up by AVHRR (even if below its spatial resolution), but not by CALIOP products (for physical reasons). The statistical explanation given here does not seem to be complete and is hard to follow.

p10,369-371: What is Kuiper’s score, what’s the dominating mode in which case? At this point, some examples that help understanding one score vs. another are provided which is helpful, but that should be done (more systematically) earlier in the manuscript.

p11: “The cloud detection sensitivity is here as high as 1.5”; “all optically thick clouds”... Define what “high” and “thick” means (earlier in the manuscript).

p13, L495-500: Since specific orbits and satellites were not clarified, there’s confusion here as to what was actually compared/validated. If it was equally applied to the morning and afternoon orbits (the wording leaves this open), one has to wonder how this would work because CALIOP operated in the afternoon orbit. How can morning cloud cover be “compared” to afternoon cloud cover, considering the significant diurnal cycle of clouds in most regions?

Language comments: p1: “considerably” -> “considerable”

p1,l20: “were” -> “where”: multiple occurrences throughout the manuscript

p1: “geographically higher” -> use “surface elevation” instead?

p1,l23-l25: run-on sentence (multiple occurrences); at the very least use punctuation (in this case, a comma after “CDRs”) to break it up. Better still, re-write.

p1: “sensor families”: a bit unusual for science manuscript, consider revising “family”

p1: add comma before “which” (in most cases; multiple occurrences)

p1: “four decades...” -> “four decades, which qualifies them to be used in climate...”

p2,L51-53: “Linked to this. . .” Unclear: efforts by whom? stringent with regard to?

p2: The A-Train stands for “Afternoon constellation”, not “Aqua Train”

p2: “a project being a part of” -> fix language

p2,l70-73: run-on sentence

p2,L91: MODIS: Introduce upon first occurrence

p3,l117: “including” -> “detecting”

p4, l121: “notice” -> “note”

p4,l148-165 and following: Avoid “you”: Not only is this inconsistent with the style of this manuscript, but it is also not advisable for a science manuscript in general. This sounds more like a seminar or talk than a paper at this point in the manuscript. I recommend a complete re-write of this section, as well as a thorough discussion of the meaning of a “cloud mask” (see comment above).

p4,l148: “areal extension” -> “areal extent”

p5,l177-179: Revise English, hard to understand

p5,l184: “possibly punish AVHRR-based methods in an unfortunate and undeserved way. . .” - see comment above

p5,L199: “of which single shots that were removed. . .” fix English

p8,L314: “navigation” -> “geolocation”?

p10: “Validation results are probably underestimated” -> What does that mean?

p10: “compared to if only showing results based” -> fix English

---

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-307, 2017.