Anonymous referee #3

RC: referee comment

AR: author response

AC: author's changes in manuscript

Received and published: 8 January 2018

Review of "The Community Cloud retrieval for Climate (CC4CL). Part I: A framework applied to multiple satellite imaging sensors", by Sus et al.

RC: The manuscript introduces a valuable approach to establish a common passive cloud retrieval applicable to a series of standard polar orbiters in order to create data sets usable for climatological studies. This would be an important step for the community and the usability of satellite products outside the satellite community. I also understand and acknowledge the need to base such an approach on well established methods instead of more experimental approaches as suggested by one of the other referees. The general presentation is of good/excellent quality. My two co-referees have elaborated on a number of specific technical and scientific details already. I want to focus on a more general weakness.

What exactly is the focus of this manuscript? If I missed important, clear, early statements in the existing text, I apologize. If not, the reader needs this guideline. In many places important details can not be given and are not explained owed to the shear extent of this project. In most cases the reader is then correctly referred to other publications where the methods of CC4CL are introduced. This way the purpose of the manuscript at hand becomes more and more unclear while reading through it. First impression is that the general method will be explained. But then the core retrieval techniques are explained elsewhere (McGarragh). Then a technical explanation of the ANN cloud mask is started, but it stays too short to be fully comprehensible. After the introduction of example cases Fig 3-5 and cross sections Fig 9-16, I expected an in-depth discussion of reason for differences and a quantitative validation (section titles containing "validation") or cross-comparison of all products, but the discussion stays very general and mostly describes differences. Proper validation is again shown elsewhere (Stengel).

The limited original content of this manuscript (correct me, if I'm wrong) is not reflected by the title and manuscript length (e.g. 8 figures 8-15 with very comparable content and not too surprising differences between active and passive sensor, but no quantitative validation). The authors should clarify the purpose of this manuscript and shorten parts published elsewhere even stronger. I suggest to consider these general points and a revision of the manuscript.

AR: We appreciate the comments of referee 3 and agree that some clarification is required to explain the purpose of the paper. Please note that this has also been pointed out by referee #2, so our answer here has been copied from our comments to reviewer #2.

This paper's main purpose is to present a new cloud retrieval framework (CC4CL). It is a two part publication that contains a detailed description of the retrieval algorithm in part II. Part I should not be seen as a validation paper, but rather contains a section that provides the reader with an overview of the functionality of CC4CL, including generic strengths and weaknesses. The goal is to inform the reader of potential applications of this data in future research. The four case studies aim to illustrate the strengths and weaknesses of CC4CL through detailed, direct (i.e. with very little averaging), and collocated comparisons with independent CALIOP data. The Stengel paper, as the reviewer correctly mentions, contains a true validation of CC4CL, but to include such an in-depth analysis here would have substantially increased the paper's length. We think that keeping part I concise and focused better serves its purpose as an introduction to the functionality and generic applicability of CC4CL. For readers who might be interested in a validation of CC4CL after reading part I, we refer to the Stengel paper in the text.

However, we will replace "validation" with "examination" or "analysis" throughout the text. The reviewer is correct that no true validation study has been carried out here, and we rephrase in order to avoid misunderstandings.

Specific major issues:

RC: p3, line 27: "Moreover, the resulting time series are carefully validated ... (ISCCP, PATMOS-x, CM SAF, and MODIS Collection 6), reanalysis and model data (ERAInterim and EC-Earth), ground-truth synoptic observations, and CALIOP lidar data."

My understanding was that I would see that in this manuscript: You will only show CALIOP comparisons, will you? Could you please clarify.
AR: Yes, we only compared with CALIOP. We will add a reference here to the Stengel paper, and also a reference to our internal product validation report.
AC: "Moreover, the resulting time series were carefully validated against well-established climatologies (ISCCP, PATMOS-x, CM SAF, and MODIS Collection 6), reanalysis and model data (ERAInterim and EC-Earth), ground-truth synoptic observations, and CALIOP lidar data (Stengel et. Al, 2017, PVIR)."

RC: p8, section 4.3: I think you cannot call this chapter "validation". There is no systematic validation, only a few selected case studies, which mainly show the problems and no systematic quantitative validation. Four case studies of time height cross sections are shown only to present that lidar cth does not have much to do with passive cth? I also expected CER and COT validation somewhere.
AR: We agree with the reviewer and will, as mentioned above, replace "validation" with "comparison". As the reviewer mentions, this comparison shows the generic strengths and weaknesses of CC4CL, which certainly relates to the processing of passive imager data. However, the reader should appreciate the basic functionality of CC4CL and we find that these local comparisons are well suited for that purpose. Please also note that the CALIOP COT information is less reliable than CTH, which is why we did not compare with COT.
AC: "Comparison with CALIOP"

RC: p13, line 15: You mean, proper quantitative validation is shown in another paper ...Stengel et al 2017 ESSD? The retrieval method was shown in two other papers as well... McGarragh 2017 at JAS and AMT. Remind me about the reason for this manuscript?
AR: Please see our comments above.

Minor issues:

RC: p 2, line 33: AVHRR was not introduced before, was it?
AR: The reviewer is correct. Will rephrase.
AC: “Compared to the Advanced Very High Resolution Radiometer (AVHRR), MODIS has several…”

RC: p 2, line 53: What is r?
AR: The Pearson correlation coefficient.
AC: “(up to a Pearson correlation coefficient r = 0.94)”

RC: p 2, line 65: How can CTP and CTH be underestimated at the same time? Can you please comment?
AR: We agree that the use of the word “underestimates” twice suggests that CLARA-A2 is wrong in both cases. However, whereas CALIOP data are considered to be “truth” data, we will now simply state that CLARA-A2 has a *lower* CTP than the other retrievals, which is not an underestimation, just a different retrieval outcome.
AC: “Comparing CLARA-A2 to PATMOS-X, MODIS C6 and ISCCP, global CTP is lower by 4–90 hPa…”

RC: p 2, line 66: What is a "cloud phase bias ... of 9%"? Cloud phase? Liquid and ice? Or cloud cover?
AR: This refers to the fraction of liquid clouds.

RC: p 2, line 68: Low or high bias?
AR: We will specify.
AC: “+ 197 m”

RC: p 2, line 102: It would be nice to say at this early stage what the purpose of this particular manuscript is in ESA Cloud_cci? And what other parallel publications contribute? Later on, the reader gets the impression that everything relevant is introduced elsewhere.
AR: We agree that this needs clarification. We copied our answer to reviewer #1, who made a similar comment.
AC: “The European Space Agency has established the ESA Climate Change Initiative program (ESA CCI, 2015; Hollmann et al., 2013) in order to advance knowledge of the climate system through the generation of satellite based data records utilizing European and non-European assets. The CCI project’s primary focus is the production of thirteen Essential Climate Variables (ECVs) covering ocean, atmospheric, and land geophysical variables. With these data records CCI is aiming to fulfil highest climate requirements from the Global Climate Observing System (GCOS). The study presented here is part of the ESA CCI for clouds (ESA Cloud_cci), which has the objective to develop a state-of-the-art open-source community cloud retrieval algorithm being capable of processing passive satellite imager data for several decades. Both in part I and part II of this paper, we present the processing framework as developed within ESA Cloud_cci (CC4CL, part I), the detailed mechanisms of the optimal estimation retrieval (part II), and provide an initial assessment of the strengths and weaknesses of derived cloud parameters (part I). With CC4CL several decades of passive imaging satellite data have been processed and are made available to the user. The resulting climate data records (CDR) are presented in Stengel et al., 2017.”

RC: p 6, line 27: If this is the only description of ANNCOD available, you might at least want to cite Kox et al . 2014 (AMT, 7, doi:10.5194/amt-7-3233-2014) who introduced the idea and described in much more detail.
AR: Please also our answers to reviewer #1, who asked for a more detailed introduction of ANNCOD, which we will provide. Kox et al. developed an approach similar to ours for retrieving Cirrus COT and CTH, but we do not think that they introduced our idea for cloud masking.

RC: p6, line 51ff: This is all a slightly vague description, if it isn't detailed somewhere else. Why do you need ... after viewing angle dependency correction ... a whole set of thresholds? ANNCOD already gives an answer on the question cloud or no-cloud, doesn't it?
AR: That would mean that the ANNCOD perfectly reproduces CALIPSO data, which is not the case. The thresholds were necessary to avoid overestimation of cloud cover due to the sensitivity of the passive sensors. With the passive sensor we measure reflectance and temperatures, in contrast to CALIPSO which is independent of both. Strongly reflecting surfaces and/or difficult illumination conditions will create ambiguities. Especially under difficult illumination conditions such as twilight, and over ice/snow surfaces we needed to increase the thresholds to avoid overestimation and decrease the false alarm rate (knowing that we might miss some clouds). We made a skill analysis with CALIOP to find the most suitable thresholds. The viewing angle correction has nothing to do with this, but more or less you can see this as a sun-zenith and surface correction of the retrieval.

RC: p7, Figure 2: y-axis. It is PEC not 1-PEC shown, isn't it? Does the graph show that, at your threshold you are only correct by about 50%?? Please discuss.
AR: The y-axis shows 100 – PEC [%]. The graph shows that the uncertainty increases to about 50 % at the threshold. This makes sense, as an ANNCOD value close to its threshold indicates that no clear distinction between cloud/no cloud can be made, thus the highest uncertainty. The larger the difference between ANNCOD and its threshold, the lower the associated cloud mask uncertainty.

RC: p10, line 10: "consistent". You could also say its all over the place, with different physical reasons in any single column. This is not a validation. You even tried to correct cth for cc4cl and still have big problems.
AR: We do find that Figure 9 shows very similar retrieval results of CTH for all three sensors, except in sector 2. We are referring here to the agreement *amongst* sensors, not between sensors and CALIOP data.

RC: p11, Figure 7: Please make the labels consistent with the rest of the manuscript: n18->avhrr, myd->modis ...
AR: We will modify labels accordingly.