# Response to Referee Comment (RC2) on

Understanding the model representation of clouds based on visible and infrared satellite observations (https://doi.org/10.5194/acp-2021-5)

We appreciate the comments of the reviewer and have revised the manuscript to address all remarks. Our response is provided in black and the review in blue.

*The manuscript presents satellite and model comparisons from 2 days during a 30-day ICON-D2 hindcast to motivate the use of visible and infrared analysis in tandem when assessing model clouds. Then statistics from the full 30 days are shown to illustrate systematic model deficiencies. An attempt is made to understand the source of these deficiencies by focusing on cloud parameters and parameterizations within ICON. Tweaks to these schemes are used to motivate possible ways to improve ICON.*

*There is a lot of back and forth in the study design between weather models, radiative transfer calculations, and satellite observations. This does not seem to be atypical for the satellite community, but for those of us on the cloud process/modeling side who would seem to benefit most from this study, this back and forth presents an opportunity for confusion. My biggest concern with the manuscript is not the methods, per se, but the logic of their presentation. I think the overall experimental design needs to be made much clearer. Why are the steps taken the right ones and taken in the right order? I had to sketch out sequencing of the study for myself after reading the manuscript a second time to make sense of things. Even then, some aspects of the manuscript felt out of place.*

We revised the abstract, introduction and various other parts to make the story/approach of the paper clearer.

Major Concerns

1. *The abstract contains lots of ambiguous sentences that simply can't stand on their own. For example, the second to last sentence means something very specific to the authors (and to the reader after reading the manuscript) but seems very unclear to the uninitiated. The same could be said of the final sentence and many others.*

The last two sentences were removed (together with most other sentences referring to surface radiation, see response to point 3 below) and the rest of the abstract was revised.

2. The paragraph starting on L65 seems very important, but has similar issues to the abstract. This is a somewhat roundabout study which focuses on a number of different things, so I think this paragraph which is intended to describe the logic of the methodology deserves to be better. I would start by reiterating the goal of the study (which I infer to be): "the meteorologically forecasting relevant quantities for PV generation will rely on assimilating clouds well and on accurate cloud

simulation. This study is therefore aimed at improving our general representation of clouds in models by assessing current model performance relative to satellite observations. Etc"

Thank you for your suggestion, we understand your concern. We have rephrased the introduction, including this paragraph.

3. *Section 3.3 seems unnecessary. Maybe I'm missing something important, but the result of this section seems logical and the figure unsurprising.*

We understand that this part should not be in Section 3. However, we still think that for some readers, e.g. those with a data assimilation background that were so far only concerned with remote sensing observations in the infrared and microwave part of the spectrum, the difference in the correlation could be valuable information. Therefore we have removed the model equivalents from the plot and moved it to Sect. 2.2. There it fits well as it shows, based on observations, the consequences of the idealised plots for the cloud signals in infrared and visible channels.

4. *The exact logic of section 4.1 needs to be explained. It's not clear precisely how I should interpret this figure in general. For example, if one of your test cases exactly recreated the OBS but REF didn't, it's not exactly clear to me what the conclusion would be. What if REF and REF-Grid were exactly the same? Should this analysis be used to draw conclusions about the success or failure of ICON or of the forward model? I don't need answers to these questions, exactly, but rather am trying to illustrate my lack of understanding of the logic of this section.*

This section is intended to show the relative contributions of different model clouds (different phases, subgrid or grid-scale) to the reflectance distribution and to identify the main suspects responsible for deviations between the observed and modelled distribution. Fig. 9 allows for comparing the contribution of different cloud types (e.g. REF - REF-grid for the subgrid clouds) to the deviation between OBS and REF. If the contribution of a certain cloud type is much smaller than the deviation it is unlikely that tuning (i.e. slightly changing) these clouds in the model or making changes related to them in the forward operator will be an effective way to reduce the deviations. If these changes should be made in the operator or the model is the topic of sections 4.2 and 4.3.

The most relevant information that can be gained from Fig. 9 is probably that subgrid clouds are important, as REF and REF-grid are really different.

We added several sentences to 4.1. to explain the intention of Fig. 9.

5. *I am left wondering how sensitive the conclusions are to the cloud morphology of summer over northern Europe. Presumably column precipitate mass is mostly liquid during these months which leads you to the conclusions that liquid is ultimately important for (if nothing else) solar reflectance. Do you feel your results are generally applicable in the context of a weather model that may need to simulate lots of different cloud states over the course of a year?*

The general conclusion of this paper is that visible channels can provide important additional information for model evaluation as they saturate later than infrared channels. For the specific (and relevant) situation investigated here it is the water content of liquid clouds that can be constrained by the visible reflectance. For situations in which ice clouds are more predominant the late saturation should still be helpful for constraining the full (liquid plus frozen) water content, but it may be more problematic to attribute deficiencies in the full content to problems with water or ice clouds. In such cases it may be helpful to additionally consider the $1.6\mu$m channel available on many satellite instruments, which allows for distinguishing between water and ice clouds.

These remarks have been added to the results section.

## Minor Concerns

*L96: How did you determine what is physically plausible?*

As explained in the text we target parameters whose values are unknown, i.e. not constrained by some physical law/observation. Previous to this study we talked to the developers of the different parameterizations of DWD about the plausibility of the values used for these experiments. We have individually discussed in section 2.1 which parameters have been changed and the effect on the parameterizations.

*L100: These seem arbitrarily chosen. How were these chosen before the study or were they chosen as a result of initial data analysis?*

The experiments target the parameterizations that directly influence clouds: microphysics, convective parameterization, subgrid parameterization and cloud number concentration. Some of them were chosen based on initial results.

*L126: Similarly, why these seven (especially for VI and VII)?*

See comment above; we chose simulation VI, because we found way too many ice clouds and we were not sure if the signal was caused by subgrid or grid scale clouds. Simulation VII was motivated because the two-moment scheme reflected too much radiation, and therefore we reduced the amount of subgrid clouds.

*L176-L191: You mean the effective radii calculated by the ICON radiation scheme, correct? Not the geometric radii?*
Yes, you are right, we mean the effective radii. However, we use the effective radii computed in the microphysics scheme. We have added a clarification in the text.

*L196: You mean you followed the procedure of Meirink by replacing MODIS with your ICON radiances? Why is this a necessary step? Without it, might you have usefully inferred a model bias?*

You are right. Basically, we remove a systematic bias, where the model produces too many very thick clouds. We can do that because the correct calibration of SEVIRI VIS006 solar reflectances is unknown and calibration of visible radiances is challenging.

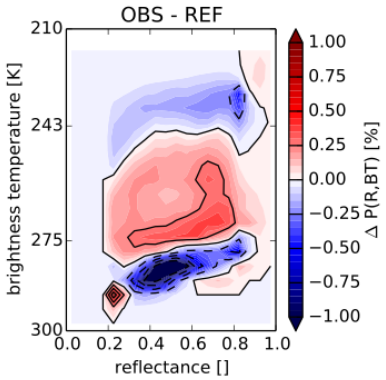
Meirink et al. (2018) showed that SEVIRI reflectances are underestimated by 8% compared to MODIS observations. MODIS observations should be more accurate.

*L242 and L273: Use of the word "exemplarily" feels a little out of place.*

We removed "exemplarily" in L242 and L273 accordingly from the text

*Fig. 6: Do "difference plots" help to highlight anything that isn't obvious simply by showing observation and simulation results side by side?*

We also looked at difference-plots, but we think they do not reveal any important additional information in this case. As an example here is Fig. 8c) - Fig. 8d):




*General: It feels as though there are a lot of acronyms that have been defined but are not used very much. You may not need to define as many as you do.*

We have reduced the number of acronyms and hopefully, wherever it was possible.

*Section 4.2: I don't feel as though I have sufficient background knowledge to judge this section.*

L461: Why shouldn't they be included?

We meant for the evaluation of model clouds (changed in the manuscript) they should not be included, because the high AOD in these cases can affect the visible reflectances strongly.