Comment on acp-2021-5
Matthew Igel (Referee)

Referee comment on "Understanding the model representation of clouds based on visible and infrared satellite observations" by Stefan Geiss et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-5-RC2, 2021

Review of “Understanding the model representation of clouds based on visible and infrared satellite observations” by Geiss, Scheck, de Lozar, and Weismann.

Review by Matthew Igel.

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.

Major Concerns

- 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 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"
- Section 3.3 seems unnecessary. Maybe I’m missing something important, but the result of this section seems logical and the figure unsurprising.
- 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.
- 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?

**Minor Concerns**

L96: How did you determine what is physically plausible?

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

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

L176-L191: You mean the effective radii calculated by the ICON radiation scheme, correct? Not the geometric radii?

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?

L242 and L273: Use of the word “exemplarily” feels a little out of place.

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

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.

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?