Atmos. Chem. Phys. Discuss., 11, C6257–C6261, 2011
www.atmos-chem-phys-discuss.net/11/C6257/2011/
© Author(s) 2011. This work is distributed under
the Creative Commons Attribute 3.0 License.

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Evaluation of cloud fraction and its radiative effect simulated by IPCC AR4 global models against ARM surface observations" *by* Y. Qian et al.

Anonymous Referee #3

Received and published: 12 July 2011

General comments

In this study the authors explore long-term (i.e. multi-year) measurements of CF and TCF at various permanent ARM sites in the world, and discuss their potential use in the evaluation of the cloud-radiative model climate of the GCMs participating in IPCC AR4. Given the significant uncertainty in numerical predictions of future climate that is related to cloud representation in models, this topic is very relevant for the climate science community. In addition, any attempt to involve the wealth of available ARM data that could be used to this purpose should also be encouraged. In this study the authors make use of simultaneous measurements of cloud presence by multiple in-

dependent instruments. They go to great lengths in inter-comparing these different datasets, and provide a detailed analysis of their differences and the possible causes behind them. The introduction contains a good and quite thorough review of the topic of the evaluation of model clouds against observations, and the written text on the instrumentation and the measurements is quite clear and accessible also for non-experts in observational meteorology. All of this is commendable.

In my view two results of this study stand out, as they are the most relevant for the climate science community and in particular for model evaluation purposes. However, as presented in the current manuscript, these results either i) are still somewhat complicated or ii) their implications are not elaborated on as fully as possible. I think addressing both shortcomings is necessary, as they would significantly increase the value and relevance of this paper.

1) An important result is that the considerable differences that exist between the various independent measurements of CF and TC on a daily basis seem to reduce significantly in the monthly and annual means. The authors in principle do well in explaining all known causes for these differences. I am not convinced though that long time-averaging completely solves all comparability problems, as the authors seem to suggest when they then use the long-term means for model evaluation. Just establishing that the long-term means converge and get pretty close is not sufficient; while long time-averaging might perhaps solve the ergodic side of the problem (i.e. time-averaging gets equal to spatial averaging), compensating errors might still exist between say the effective detection threshold of the instruments and the cloud side effect in wide FOV instruments, just to name two examples. I don't see why the last two effects could not affect the monthly or annual means. For example, the cloud side effect in wide FOV instruments will be significant in deeper boundary-layer cumulus cloud fields; averaging over a month long of exclusively cumulus days will still show this impact. The same goes for threshold differences. So what is lacking in the current manuscript are some arguments for why the effects other than the ergodic one will not affect the longer-term

means; can the authors provide those?

My guess is that the application of instrument simulators are the only way to eliminate all comparability problems. Am I right in assuming that this was not an option because the IPCC AR4 runs did not require simulator output? And what do the authors think about the assumptions that still go into simulator models?

2) The authors manage to establish from the confrontation of long-term means of model cloud fraction and radiative fluxes with their observed equivalents that compensating errors are present in the interaction of clouds and radiative transfer as represented in most of the GCMs. Unfortunately, from these results we do not (and can not) learn exactly where in the system of interacting parameterizations this compensation takes place. This is also acknowledged by the authors in the final section (p.14963 lines 18-19). While I realize that this may be beyond the scope of this study, what is important though is to think about how this insight might actually be obtained, and to also discuss this. In my view, this exercise is key to solving this important problem in climate science. Signalling the problem of compensating errors is of course great, but in this case is actually not new; previous studies have shown this, and by now most modellers are already aware that such a compensation takes place in the cloud-radiation interaction. The actual reason for the existence of compensating errors is that each individual parameterization (e.g. cloud overlap, inhomogeneity, vertical structure, etc.) has not yet been sufficiently constrained by observations. Could the authors come up with a list of all these processes, and perhaps make proposals on how to cover each with observations? I think the addition of a discussion on this topic, probably in the summary section 6, would benefit this paper. It would certainly be constructive, as it would suggest a way forward and could also act as an outreach to the modelling community.

To summarize, given i) the relevance of the topic, ii) the good use of available observational data, and iii) the fair scientific quality of the study (both in writing and content) I think this manuscript should be acceptable for publication after some major revisions as mentioned above.

Specific comments

1) p14934, line 13-14: The cloud overlap assumption should also be mentioned here. Is it known what the overlap assumptions were in all participating GCMs?

2) p14934, line 15: "very minimal". Use of the word "very" is not scientific, please rephrase.

3) p14938, lines 8-10: This would be a good point to discuss the potential use of instrument simulators.

4) p14938, lines 19-21: "... examine ... at different time-scales". The evaluation at daily time-scale as announced here is in conflict with what is mentioned on the previous page (p14937, lines 19-21: "... we do not perform direct hour-by-hour or daily comparisons, ..."). This is confusing. Also, if your intention is to avoid daily comparisons from the start, why then still dedicate such a big part of the manuscript (Figs 1-2) to this?

5) p14939, section 2.1: Could you spend some sentences on what CF actually represents in the models? Is it the cloud fraction as used in the radiation scheme, or is it maybe the cloud fraction as part of a vertical transport model? These are not equivalent; the fraction as used in the radiation code should carry information on sub-grid scale cloud overlap, while the transport scheme cloud fraction does not (area-averaged versus volume-averaged; see e.g. Brooks et al., JAS, 2005). Accordingly, the vertical resolution also plays a role here. Would it not be better to evaluate the CF in models on the same vertical grid as the observations?

6) p14942, lines 22-23 "... averaged only for daytime hours between 8 and 17 LT": Would it make sense to preserve the diurnal time-dependence in the evaluation of the monthly and annual means that follows later? This might be informative, as it can indicate at which point of the day biases are biggest. In turn, this could help in attributing biases to specific parameterization schemes (boundary layer, deep convection, etc.).

7) p14943, line 23 "... affected by cloud sides ..": See also my first general comment

above. Although you mention and explain it here, after this point no attention is given to the potential impact of this effect on longer-term means.

8) p14957, line 15-16: This is pretty much the only point in the text where the overlap assumption is mentioned as a possible contributor to model-obs differences. In broken and irregular cloud fields this could actually be significant. Also, I don't think the ARSCL product can detect the true overlap, as i) the lidar can not penetrate through optically thick clouds, so that it misses all clouds behind the first, and ii) the cloud radar might be blind for small droplets in warm clouds.

9) p14964, first paragraph: The proposition to compare retrievals of cloud presence by ground-based and satellite-based instrumentation is interesting. Would the technique of cloud height - optical thickness histograms as originally used for the ISCCP dataset be worth mentioning here?

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 14933, 2011.