Response to Interactive comment on by A. Voigt on "The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6" by Mark J. Webb et al.

Reviewer comments below are shown in bold and our responses are in italics.

Dear Aiko,

The authors provide a concise and well-written presentation of the CFMIP experiments proposed for CMIP6, which will continue the successful CFMIP activities over the last 15 years. I enjoyed reading the paper, in particular the historical context given in the introduction, and find that it nicely presents the scientific motivation and chosen simulation strategy at a level amenable to both CFMIP experts and climate scientists with other backgrounds. I recommend publication in GMD after my following minor comments have been addressed.

Thank you for your careful consideration of our manuscript and for these helpful comments.

Line 217, amip-future4K simulations: Why is the CMIP3 SST pattern used and not an updated pattern from CMIP5 AOGCM runs?

We will add the following to Appendix C:

"We have retained the SST forcing based on the CMIP3 coupled models because we consider it more important to be able to compare CMIP5 and CMIP6 models forced with the same SST pattern than to use a pattern which is consistent with, say, the CMIP5 coupled response."

Line 227: I am very glad to hear that the CMIP-3 aquaplanet simulations will be extended to 10 years. This will be beneficial for studies of extratropical dynamics, for which internal variability is larger than in the tropics.

Thank you.

Line 238, amip-m4k simulations: I am wondering to what extent some models might have problems with SSTs below freezing? Maybe this might require code changes in some models in case they employ a fixed lower threshold for the SST used in the calculation of surface fluxes? Such a problem would, of course, not occur for the p4K simulations?

We will add the following to section 2.2:

"In models which employ a fixed lower threshold near freezing for the SST used in the calculation of the surface fluxes, this should ideally also be reduced by 4K."

Lines 279: The authors might consider to also refer to Voigt and Shaw (2015, Nature Geoscience) here for the extratropical circulation. The study showed that cloud-

radiative feedbacks contribute substantially to the poleward jet shifts under 4K warming in aquaplanet simulations.

We will add that reference.

Line 266, lwoff experiments: Just an idea, but I though it's worthwhile bringing it up here: While the surface cloud effect is stronger in the shortwave than the longwave domain, the longwave can still substantial. I am wondering whether an experiment with clear-sky heating in the atmosphere and all-sky heating at the surface would be even better to isolate the effect of atmospheric cloud-radiative heating. I suspect it's too late to change the experimental protocol, and maybe there is a reason why lwoff is still better. If so, it might be worthwhile to briefly discuss this.

This is an interesting idea thank you. However, to be recommended for CMIP by CFMIP, we generally require new experiments to have been piloted and ideally written up with at least one GCM previously. The lwoff are experiments currently proposed are very similar to those piloted by Fermepin and Bony, 2014, and technically easier to implement than what is proposed. If such an experiment can be demonstrated to provide new insights which are relevant to the objectives of CFMIP then we will certainly consider it in the future.

We will add the following to the manuscript:

"An alternative method (proposed by A. Voigt) was also considered, in which clear-sky heating rates would be applied in the atmosphere while retaining the all-sky fluxes at the surface. Although this approach would potentially isolate the effects of cloud heating in the atmosphere more cleanly than the lwoff experiments proposed here, it is yet to be demonstrated in a pilot study, and is considered more technically difficult to implement than the lwoff experiments, which are very similar to those piloted by Fermepin and Bony, 2014."

Line 342: Non-linearity was also shown in the CMIP5 ensemble by Meraner et al. (2013, GRL, doi:10.1002/2013GL058118). Meraner et al. showed non-linear climate sensitivity across the multi-model CMIP5 ensemble, whereas the other cited work used single models if I am not mistaken. So maybe worthwhile including here?

We will add that reference.

Line 368: Maybe specify the reason why the CFMIP2/CMIP5 runs did not allow such an estimate. I.e., I assume that one would use SST-driven simulations for this and that the usual amip period is too short to reliably calculate feedbacks?

We will modify this line to read:

"The previous CFMIP-2/CMIP5 design was unable to diagnose the time-variation of feedbacks of explicit relevance to the historical period, because this requires the removal of the time varying forcing."

Sect. 2.7: The time slice experiments ask an interesting question but given that 8 experiments are demanded, I was wondering how they ought to be combined to answer the questions in mind. Maybe the authors can give an example?

We will add two examples in the text of how these experiments are combined:

"The timeslice experiments can be combined in various ways to isolate the climate response to each individual aspect of forcing and warming. For example the response to SST pattern change is given by taking the difference between a4SST and piSST-pxK, and the plant physiological response is found by taking the difference between piSST-4xCO2 and piSST-4xCO2-rad."

Line 487: I would be curious to know about the reasons to no longer ask for cfSites output in the aquaplanet ensemble and amip-future4K. Is it the lack of observational data to compare to, or a choice to avoid asking for too much data?

We will amend the manuscript to read:

"We have dispensed with the cfSites outputs in the aquaplanet and amip-future4K experiments because these have been less widely used compared to those from the other experiments."

Line 685: and –> an

We will amend that.

Figure 1: Why does the vertical cloud bar on the right side not include the lwoff simulations?

This point was also raised by F. Brient. We will extend the arrow to include the lwoff experiment.

For some of the proposed simulations the link to clouds, which are the prime motivation for CFMIP, is not very evident and maybe could be made clearer? I am thinking of the simulations in Sect. 2.7 (time slice experiments) and Sect. 2.5 (nonLinMIP).

The primary objective of CFMIP is to inform future assessments of cloud feedbacks through improved understanding of cloud-climate feedback mechanisms and better evaluation of cloud processes and cloud feedbacks in climate models. However, the CFMIP approach is also increasingly being used to understand other aspects of climate change, and so a second objective has now been introduced, to improve understanding of circulation, regional-scale precipitation, and non-linear changes. For this reason, not all experiments need to be relevant to clouds. We will modify the text in the abstract and introduction to state this explicitly.