Response to Interactive comment on by Anonymous Referee #1 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 Referee,

In this paper, the authors state the goals and motivation of CFMIP, review the major accomplishments of previous CFMIPs, and describe the proposed experiments and diagnostics for CFMIP3. The coordinated experiments proposed for CFMIP3 will target a number of outstanding questions for which previous model intercomparisons were not equipped to address, in addition to sustaining a number of highly useful experiments from earlier MIPs that will help to characterize and understand the response of the CMIP6-generation of models to external forcing (in addition to help quantify the forcing itself). Advanced diagnostics (e.g., satellite simulators and high frequency tendency terms) will aid in dissecting model results, and the authors have proposed that they be used more broadly (e.g., COSP turned on for longer durations and in more experiments). The emphasis on (mostly) atmosphere-only simulations in CFMIP3 should hopefully make it appealing for modeling centers to take part in several of the experiments despite the high volume of requested diagnostics.

The scientific questions to be addressed by CFMIP3 are well articulated and the various proposed experiments seem well designed to address these questions, and will advance the community's knowledge. The presentation of the paper is not particularly concise, not are the figures particularly insightful, but the writing is clear and overall the presentation seems appropriate for a paper proposing a model intercomparison project. Thus, in my opinion the manuscript represents a substantial contribution to modelling science within the scope of Geoscientific Model Development, and I recommend publication following consideration of some minor comments detailed below.

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

Specific Comments:

*piSST and a4SST: It is not clear to me whether (a) monthly- and annually-varying SSTs from the relevant 30 years in the piControl run, or (b) a monthly-resolved climatology of SSTs over the relevant 30 years in the piControl run are prescribed in piSST. Same question for a4SST.

We will modify the text to read 'monthly- and –annually varying SSTs....' in the descriptions for piSST, a4SST, a4SSTice and a4SSTice-4xCO2. For consistency we well also refer the AMIP SSTs and sea ice in the amip-a4SST-4xCO2 description as monthly- and –annually varying.

*amip-piForcing: I'm curious whether there was any interest in performing a similar experiment, but with present-day (rather than preindustrial) forcing held fixed. An example application that occurs to me is that a model with large aerosol-cloud interactions would presumably have brighter clouds with smaller droplets downwind of aerosol sources if

the forcing were fixed at present-day, and its temperature-mediated changes in clouds might therefore be different than that occurring in an atmosphere with fewer aerosols. Having these two experiments would allow one to explore this effect (and others related to other forcing agents).

This is an interesting idea. 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. 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.

*Given its implications for understanding apparent state- or time-dependent changes in effective climate sensitivity, I was a little surprised to see no experiments designed to explore causes of nonlinearity in the Gregory plot, perhaps using warming experiments in which the SST pattern is fixed in time (with various patterns), similar to those conducted in Andrews et al, J. Climate (2015). Is there a reason for not proposing these, or are these effects already captured in other proposed experiments?

We do consider the causes of non-linearity in abrupt4xCO2 experiments to be an important area to be investigated. The experiments in Andrews at al (2015) were based on actual SSTs from individual models. Pilot studies are ongoing to devise future experiments for CFMIP relevant to this question based on SST pattern responses more representative CMIP5 ensemble mean. We plan to organise a pilot intercomparison based on this, although this might initially be arranged informally within CFMIP rather than as part of CFMIP/CMIP6.

```
*Line 126: should be "
...
meetings AND international
...
"
```

We will correct this.

*Line 426: "a4SST-4xCO2-all" should be "a4SSTice-4xCO2-all". There may be other instances of this; please verify that they are also changed.

We will correct this.

*Line 512: What is the reason for dispensing with the cloud tendency terms in CFMIP3?

We will add the following:

"We have dispensed with the cloud water tendency terms because these have been less widely used than the temperature and humidity tendencies."

*Lines 597-608: it is not clear to me why some of these have a CMIP5 prefix, a CFMIP

prefix, or no prefix at all (cfDay-3d). Why would a CMIP5 prefix be appropriate at all?

The different prefixes represent detail in the formal data request which is not required here. In the manuscript we will delete the prefixes to avoid confusion, and will add the following sentence:

"(Please note that in the full data request these variable groups are in many cases split into a number of sub-tables. As noted above, the formal data request provides the definitive specification of the model outputs.)"

*Line 611: should be "
...
for 140 years OF the piControl
...
"

We will correct that.

*Appendix A: I don't understand what is meant by "Lead coordinator". Is this the person who has "first dibs" on writing papers based on these experiments? Are interested investigators expected to contact this person to avoid duplicating work that others are doing with output from these experiments?

We will add the following to Appendix A

"We plan to scientifically analyze, evaluate and exploit the proposed experiments and diagnostic outputs, and have identified lead coordinators within CFMIP for different aspects of this activity. The lead coordinators are responsible for encouraging analysis of the relevant experiments as broadly as possible across the scientific community. While they may lead some analysis themselves, they do not have any first claim on analysing or publishing the results. All interested investigators are encouraged to exploit the data from these experiments. While investigators may wish to liaise with the lead coordinators to avoid duplicating work that others are doing, this is not a requirement."

*Figure 1: I think "CMIP6" should be deleted before "historical".

If it is supposed to be there, I don't understand why it is only there.

This is the correct naming. Please see Eyring et al. for the justification.

It is also unclear to me why the "Clouds" arrow only extends as far as abrupt-0p5xCO2. I think both the clouds arrow and the circulation and precipitation arrows should include all experiments, but in that case, what is the point of showing them?

In response to a comment from F. Brient, we will extend the cloud arrow to encompass the lwoff experiments. The timeslice experiments in the bottom group are designed to look at circulation and precipitation responses rather than cloud feedbacks.

*Table 1: should be "This IS a single ... "

We will correct that.

*Table 3: Several of the observational datasets end many years ago despite the fact that these satellites are still in orbit. Are there plans to extend these records, especially since the AMIP runs end in 2015?

We will add the following the end of Section 3.2:

"These datasets are periodically updated to include more recent data from the relevant satellites, many of which are still operational. Please refer to the CFMIP-OBS website for updates."