Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2019-1175-RC1, 2020

Atmospheric Chemistry and Physics Discussions
Open Access
EGU

Interactive comment on "On the Climate Sensitivity and Historical Warming Evolution in Recent Coupled Model Ensembles" *by* Clare Marie Flynn and Thorsten Mauritsen

Anonymous Referee #1

Received and published: 17 February 2020

This paper is an analysis of the differences between CMIP5 and CMIP6. It has relatively low information content, but I don't see anything technically wrong with it. My comments are therefore relatively minor.

Overall, having read the paper, it was hard to determine what the top-line scientific message of the paper was. This problem is epitomized in the abstract. Per that, the most important conclusion is that the difference in ensemble averages between CMIP5 and CMIP6 could not have arisen by chance. Is that really the most important point in the paper? I do think it's a reasonable point to make, but it's a simple means test and it's hard to rationalize writing an entire paper to make that point.

Printer-friendly version

Discussion paper

The authors also claim in the abstract that "These results suggest that changes in model treatment of mixed-phase cloud processes and changes to Antarctic sea ice representation are likely causes of the shift towards larger ECS." But this is not proven anywhere in the paper — in fact, it's only mentioned twice and labeled as "speculation".

Line 73: The authors subtract the time series of the PI control run from the abrupt 4xCO2 run to account for drift of the underlying model run. This assumes that the modes of unforced variability are the same for 150 years in both runs, but I'm wondering how reasonable this is. Couldn't the huge RF from 4xCO2 change the internal modes of variability after 100 years of warming? I think it would be useful for the authors to comment on whether the results would be any different if the authors just subtracted the time-mean of the piControl run.

Line 93: I do not think it is correct to say that natural climate variability is a source of observational uncertainty. That needs to be rephrased.

Line 105: how is forcing calculated from the Gregory method? Is it just the Y intercept? Or are they making any other adjustments for fast responses?

I'm somewhat confused by section 3.2. Estimating the significance of differences between the means of two populations is a standard statistical problem. i.e, you can use some variant of a t-test to do that. I would think that this calculation should be described in one or two sentences. I do not understand why the authors have addresses this question with the approach in this section. There may be a reason I'm missing — if so, they should detail it.

Related to the previous comment, I don't understand why the median being larger than the mean matters. In their calculation, they reject ECS values > 10,000 K. What if they rejected values of ECS > 10 K? 10 K is only slightly less improbable, and I think choosing this lower threshold would reduce the difference between mean and median.

I don't doubt the result that the difference is significant, but I do think this section needs to be rethought and perhaps reduced to a simple t-test.

Section 4.2: In this section, they compare zonal average feedbacks between the CMIP5 and CMIP6 ensembles. They need to add to the section a determination of where (at which latitudes) the differences between the ensemble averages are statistically significant. Then they can modify the discussion accordingly. For example, I'm not sure any of the differences in Fig. 6 are significant.

They should probably add a reference to Zelinka et al.'s new paper on the difference between the CMIP5 and 6 ensembles (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GL085782). It would be good to put the results of this paper into context with those results.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2019-1175, 2020.