Geosci. Model Dev. Discuss., referee comment RC2
https://doi.org/10.5194/gmd-2021-298-RC2, 2021

Comment on gmd-2021-298

Anonymous Referee #2

Referee comment on "Better calibration of cloud parameterizations and subgrid effects increases the fidelity of the E3SM Atmosphere Model version 1" by Po-Lun Ma et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-298-RC2, 2021

This manuscript describes the process of retuning of version 1 of the atmospheric component EAM of E3SM climate model, which focused on parameters related to various cloud processes, and how the retuning impacted a range of quantities beyond the tuning targets, ranging from surface temperature in the present-day to aerosol forcing (via cloud adjustments) to cloud feedbacks.

General comments

The manuscript does a good job of explaining the reasoning that led to the strategy used in retuning, which relies on the bet that improving the representation of clouds will lead to improvements across the board. It’s nice to see that the bet pays off.

There’s rather a lot of detail in section 2, describing groups of parameters were tuned. There’s really a lot of detail in section 3, which explains how the returned model behaves with respect to a wide range of emergent phenomena. There is so much detail, in fact, that the manuscript works much better as a description of what was done than it does as an explanation of what was learned. If the authors’ goal is to document the strategy and its impacts they have succeeded, but if they aim to influence the ways in which readers undertake or understand model tuning they would be well advised to bring their ideas into sharper focus. Sharpening the manuscript will almost certainly involve relegating material to appendices or supplemental material.

A large proportion of the very many figures are of the form of Figure 4: six (well-constructed) maps showing the difference of v1 against observations, four maps showing the change induced by the four sets of parameter changes, then a map showing the aggregate change of the final retuning relative to the original. Readers are left to judge improvement by mentally subtracting the bias in the upper left plot from the change in the lower right plot. Could the lower plot be revised to show, for example, the improvement or degradation in the original bias as a result of the tuning?

Versions of the six-ma figure (e.g. Fig 7) with no observational constraint are harder for readers to assess.

Tuning relies on the ability to measure improvements in simulations, normally relative to observations. It's remarkable that the authors spend essentially no time discussing the sources of their observational constraints, or how uncertainty in these constraint is or isn't considered as part of the tuning strategy.

The tuning strategy used by the authors is somewhat traditional. Comparisons to other approaches (e.g. the automated calibration to process-scale constraints used by HiTune, doi:10.1029/2020MS002217 or the formal inference discussed in the Clima project, doi:10.1016/j.jcp.2020.109716) would no doubt be welcome.

More specific comments

Tuning of course involves the changing of specific parameters. The variable names in the specific computer code are perhaps too specific to be in the main text. This information, and indeed probably the original and changed values, could be summarized in one or more tables in an appendix.

Line 113: the current term of art is "perturbed parameter ensemble".

The subsections of section 2 are labeled as tropical clouds, low clouds, etc. In practice each section might also be categorized according the scheme whose parameters are being tuned. Indicating this (e.g. "Tropical clouds and the deep convection scheme") might guide readers' attention.

Section 2 describes the re-tuning in detail, including which parameters are re-tuned and why. General material (e.g. line 286-294) should be deferred or removed.

Line 220: The authors scale the temperature variance provided by one scheme by a factor of 2 before introducing it in another scheme. It's not clear whether this is a reasonable physical assumption. It if is the choice should be justified; if it's not the choice should be explained.

Section 2.2: how are the different cloud regimes identified in practice, during a simulation?

One reason for showing six panels in Figure 4 and its many analogs is to highlight the geographic distribution of the impacts of parameter changes. The authors might ask themselves if maps are the best way to show these differences in all cases.

Line 395-400 could be edited for clarity and to remove general material, as could lines 486-490.

Line 528: EIS is thought to control low cloud properties, not their feedback (sensitivity to surface temperature change).

The central point of lines 715-728 could no doubt be made more compactly and directly.

Line 739-740 are bewildering.

Figure 13 is hard for readers to interpret. Coding bias with shapes and variables with numbers is really quite unfriendly - bias would be better coded with size or shading, leaving shape to stand for quantities. But readers will also appreciate guidance in interpretation, since all the panels look very much the same to an unpracticed reader.

Line 831: Does the present analysis use the specific kernels of Zelinka 2012 and Pendergrass 2018? If so this should be made explicit. If the text refers to the ideas the original papers (e.g. doi:10.1175/2007JCLI2044.1) should be referenced.

Figure 18, especially panel a, is not particularly informative, since readers are asked to compare small changes in large numbers introduced by tuning.