Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-129-RC2, 2017

Geoscientific Model Development Discussions

Interactive comment on "The Path to CAM6: Coupled Simulations with CAM5.4 and CAM5.5" *by* Peter A. Bogenschutz et al.

Anonymous Referee #2

Received and published: 16 September 2017

The manuscript documents two versions of CAM, each a step towards the development of CAM6. The first one has new microphysics, aerosol, and ice nucleation changes, the second one implements pdf-based unified cloud parameterization to replace the turbulence, shallow convection, and warm cloud macrophysics. The newest model has lower cloud forcing bias, better nino3.4, better precip diurnal cycle, subtropical wind stress and less of a double ITCZ. It also has worse Amazon precip, colder SSTs in general, and sea ice that is too thin. The 20th century simulation ends up colder in new model (too cold), due to stronger aerosol indirect, and the authors propose a fix. The authors conclude that model is ready to integrate into NCAR's next generation model.

The documentation of changes in model simulation during the development stages of the CMIP6 model, along with the attribution of these changes to specific elements

of the coupled model, is clearly an important contribution to understanding climate change processes and our models' ability to simulate them. This manuscript requires the major revisions outlined below in the general and detailed comments in order to be ready for publication in GMD.

General Comments ————

1) The issue of whether the 5.4 and 5.5 simulations were tuned: The authors state throughout the manuscript that the 5.4 and 5.5 models were tuned for radiative balance but not for SST, and so improvements are due to the former and degradations are due to the latter. The authors need to explain why an untuned version of the model is ready for publication.

2) Attribution claims are unwarranted. The entire moist physics and cloud forcing was replaced from 5.3 to 5.4, say, and the text speculates about which change in the model was responsible for which change in the behavior. examples are: i) p8 line 26: "can be attributed to interactions of CLUBB with the ZM scheme and/or feedbacks to/from the coupled system with CAM5.5" ii) p9 line 1 : "the difference in precipitation simulation is likely linked to differences in parameterized physics as opposed to biases in the large-scale circulation"

3) Statements about differences among different model version results are qualitative in nature in many instances ("subtantially", "much too....") and are not put in the context of internal variability or significance. This makes it difficult to understand the changes.

4) Section 4.5 is an outstanding example of how to report on and discuss attribution. The other sections of the manuscript could benefit from this type of discussion or reporting of results of sensitivity experiments.

Detailed comments —–

p3 - discussion of implementations of CLUBB into global models. AM3 wrote the code to do it and tested it, but did not adapt and will not use in CMIP runs. please make this clear.

page 5 line 24 and then line 31 says 5.5 tuning.... abstract says not tuned. is it or not? [if not, this paper not ready yet for evaluation]

page 5 line 27 - autoconv of ice to snow increased to increase cloud? seems counter-intuitive on the face of it. or does increasing Dcs mean decreasing autoconversion? if so, please make this clear.

page 6 line 1 - also needs a few more words - why suppressing turb mixing tends to increase low cloud? won't mixing more moisture for instance tend to increase low cloud (ie., Lock's coupled BL idea)?

page 6 parag with line 3 - why insist that the tuning was only partial? please remove this discussion here and elsewhere. again, if the tuning was only partial the model is not ready for publication.

p7 l6 - is 5.5 equillibrated? looks like still rising.

p7 l15 - the periods chosen for 5.4 and 5.5 seem to be periods when global sfc temp is rising, almost as fast as it is falling in the earlier period. also - is it the author's assertion that the "offset" from pre-industrial to present day (period of simulation vs period of validation data) does not have geographical structure? please elaborate a bit more about the offset.

p7 line 18 - please discuss significance of the differences in RMSE and pattern correlation. is 0.94 really an improvement over 0.93?

p7 line 21 - i see a southern ocean difference in 5.3 of approximately +15 or so w/m***2 and in 5.4 of approximately -15 w/m**2. certainly a difference in behavior but not a "significant reduction" of bias. please correct this statement, or please quantify the zonal mean difference, say and report on a significance test.

p7 end of discussion of SWCF - the authors comment on the transition from 5.4 to 5.5

without commenting on what looks like a degradation in the global mean difference.

p8 l2-3 - please remove or restate this sweeping assessment. as discussed below, it is not clear that each model is an improvement over the previous.

p8 l9 - is this speculation about the autoconversion? if so, please state that it is probably or possibly. if not, please report on experiments that were done to isolate the impact of this parameter, and their applicability to the simulations reported here.

p8 l13 - which metric is being referred to as skill score? please rephrase this to refer to all or one of the metrics provided in the figure.

p8 l31 - which panel is this sentence referring to? 5.5? please rephrase this sentence.

p9 l2 - 5.5 also seems to have degraded the precip over the maritime continent.

p9 l14 - the authors state that the reason for the improved diurnal cycle "appears to be" due to CLUBB - please either show or report on results of atmosphere only (or coupled, if you have them) sensitivity studies demonstrating this. especially since this point is mentioned again (p9 l 23).

p9 l18 - summer precip in JJA in Amazon seems to have almost gone away in JJA - any speculation? is this an important result?

p9 l29 - please say mean error and rmse instead of skill score. also - are these "successive degradations" significant statistically?

p9-10, discussion of figure 7 - There are several issues with this discussion that warrant addressing by the authors. a) why would the "but it wasn't tuned" impact the SST and not the precip, for instance? b) line 29 p9 says "worsening", but line 33 says "not substantially worse". is it worse (statistically) or is it not? please clarify. c) the statement in line 33 of p9 that the cloud forcing can be adjusted to fix SST bias is in stark contrast to the statement on p10 line 3 that states that 5.4 cloud forcing is better but SST is worse, perhaps due to compensating biases in the ocean.

the entire discussion of the attribution of (perhaps) SST errors that increase with model advance is vague and speculative. please re-purpose this discussion. either show that a re-consideration of the tuning will improve the SST and that the eventual CMIP 6 model will address this, or wait until those results are in to publish the manuscript.

p10 l12 - please explain why the stress errors cause SST errors and not the other way around. or state that they are related. also please quantify the difference (the "degradation" to 5.4 and the "improvement" to 5.5 in the subtropics).

p10 l19-27 - same issue as with the SST. i see one region of "degradation", and the statements about fixing it with tuning need to be reconsidered.

p10 l27 - this is not an attribution of the error, but speculation in a broad sense. please either show some sensitivity results or remove the statement.

p10 l30 - please provide an observational estimate for comparison. either include an extra set of panels in the figure or provide the number at 35N, 1 km depth. please also provide an estimate of variability - is 23 Sv different from 26?

p11 l1 - please remove the statement about the likely cause for the change in the AMOC. there is no basis provided for this. the authors could be correct, or it could be a myriad of reasons, including some offered by the authors when discussing SST or wind stress differences.

p11 l29 - fig 12 adds little to this discussion. please consider removing it.

p12 l3 - the lifecycle phases are not mentioned or indicated in the figure. please either amend the figure or remove the mention of the lifecycle phases from the text.

p12 l10 - studies have shown that the addition of a (better) shallow conv parameterization increased the amplitude of MJO variability in atmos only simulations. please refer to one of these. shouldn't CLUBB be expected to help in this regard? what do atmos only simulations show the impact of CLUBB to be?

p12 l23 - syntax "a couple exceptions"

p13 l4,11 repeat the same information

p14 l20 this statement (each one improved over the other) was also stated in the results section but was not clearly borne out by the figure or the metrics reported. if the results section includes some discussion of variability or significance this statement would be appropriate.

p14 l21 please rephrase "big"

p14 l24 - please modify discussion of tuning. it is confusing to state that 5.5 and 5.4 were tuned for radiation but not for SST, so its not surprising that they are worse, but they aren't that much worse. in the text it is suggested that the degradation of SST is due to cloud forcing. please modify this entire discussion. and please discuss why an untuned model is ready for reporting in the literature.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-129, 2017.