The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2019-309-RC2, 2020

The Cryosphere
Discussions

Interactive comment on "Evaluating permafrost physics in the CMIP6 models and their sensitivity to climate change" *by* Eleanor J. Burke et al.

David Lawrence (Referee)

dlawren@ucar.edu

Received and published: 26 March 2020

This paper evaluates the simulation of permafrost in the CMIP6 (and CMIP5) models. The overall conclusion is that, perhaps not surprisingly, there isn't a huge difference between CMIP6 and CMIP5 models in terms of their representation of permafrost, with perhaps a small amount of improvement. The paper utilizes existing metrics and introduces several new metrics, including a metric for thawed volume, to assess the models. Though this paper does not generate much by way of new insight into the sources of the modeled biases, it is still a useful exercise to establish the current state of permafrost simulations in coupled climate models. Overall, the paper is well-written and will make a good contribution to the literature.

I mostly have a collection of specific points focused on clarity. However, I list a few more significant points here:

1. The authors made a decision to include results for any model that had uploaded the requisite data for the analysis. I wonder if this is really the best strategy, however. For many models, there are two versions that, as far as I am aware, are very similar. Examples include the BCC, CESM, CNRM, Hadley Centre, GISS, and MIROC models. I would suggest that the authors consider showing results for only one of each of these pairs of models. I didn't study every figure, but in general, what I saw was that slightly different versions of models from the same institution didn't look very different to each other, especially for many of the large-scale metrics assessed here. Doing this would significantly reduce the number of shown models, which would then make the figures and analysis more rapidly accessible to readers. In the limited number of cases where there is a significant difference in model behavior across models from the same center, it could be noted in the text or as a supplemental figure. I realize this would be a decent amount of work to redraw every figure and table, but . . .

2. The LS3MIP community specifically requested land-only simulations so that they could be used in direct comparison to the coupled model experiments. This paper would be stronger and more meaningful if the land-only simulations were included in the analysis. This would allow a separation between the role of climate biases and biases that arise due to representation of land processes. Again, I realize that this would be a significant amount of work, but it is work that someone in the community really needs to do, and this paper would be a perfect place for it.

Minor points:

1. Table 1: NorESM uses CLM5. Does HTESSEL have a version number? All the other land models have version numbers.

2. Line 184: Doesn't the Chadburn et al. (2017) method also supply a probability of permafrost for each grid cell?

3. Line 208: You note that an advantage of the D metric is that it enables taliks to be identified. I agree that this would be an advantage of D, but I think most taliks form at depths deeper than the 2m level, so I believe due to the 2m restriction, most taliks would be missed, if the models are simulating them.

4. Line 220: Maybe make it clear that this is an estimate of 'present-day' Dtot and Ftot.

5. Figure 2: Any idea if the 20-year average biases are 'robust'. That is, to what extent would internal model decadal variability affect these biases? Maybe it is beyond the scope of the paper, but there are lots of large-ensemble papers in the literature that could likely help make at least a qualitative assessment, or a few models have already submitted several ensemble members for the historical period.

6. Figure 2 caption: "and is not available for every ensemble member". I think you mean for each model. Even though you don't really utilize ensembles in this analysis, probably best to keep the terminology correct so as to avoid confusion.

7. Figure 2: Why are you only showing bias against the PFbenchmark and not the CCI-PF data as well?

8. Figure 3: As above, perhaps replace 'ensemble' with 'multi-model'.

9. Would be helpful to show the observational estimates of Sdepth,eff on Figure S1.2.

10. Given the challenges in determining snow depth in observations, I do wonder how accurate the Sdepth,eff dataset really is across the pan-Arctic. Nothing that you can do about this, but maybe should qualify statements here and there across the paper, noting that snow depth can be highly spatially-variable is difficult to measure and/or assimilate due to this strong local heterogeneity due to aspect, snow redistribution, snow-vegetation interactions, etc.

11. Line 306: Sentence starting with "This means" was confusing to me. Restate?

12. In fact, that whole paragraph seems confusing and would benefit from a rewrite.

13. Line 324 and elsewhere: Do you ever define the "MOHC" models? I couldn't find this acronym defined anywhere, but also could have missed it.

14. The degradation of snow insulation from CCSM4/CESM1 to CESM2 is interesting because the fresh snow density parameterization was deliberately changed to more accurately reflect observations and detailed snow models, and which leads to higher snow density and better permafrost and ALT simulation, at least when forced with GSWP3. See discussion in Lawrence et al., 2019 and van Kampenhout et al., 2019, Figure 1. The new parameterization definitely improves snow densities over ice sheets. Would be interesting to see if the CLM5 land-only simulations show similar relationship as in Figure 6 of this paper. Anyway, not really a comment that needs to be addressed, but just makes me wonder what is going on?

15. Lines 325-328: It was unclear to me in the assessment of the MIROC and CESM models what biases you think are canceling out. Reword?

16. Figure 8: Can you clarify what you are only including the obs grid cells where the model simulates permafrost. Seems to me that it makes this figure harder to understand. Perhaps would be better to keep obs same in all plots, but then report the number of sites where the model doesn't simulate permafrost for any given temperature range . . . or something like that.

17. Figure 10: I think NorESM should be hatched.

18. Line 388: I don't understand what you mean by considerable variation here. Variation across models? Also, not clear why the variability in D at warmer temperatures leads to excessive Dtot and underestimated Ftot? Couldn't variability in D go either way in terms of biasing towards Dtot or Ftot?

19. Line 420: Could also consider reporting the projected permafrost loss per degree of warming of permafrost zone temperature, as in Slater and Lawrence (2013). Gets around the Arctic amplification diversity across models problem, but then, debatably,

may not be as policy-relevant.

20. Line 451: '. . . slightly different models' → 'slightly different set of models'

21. Line 465. It's true that permafrost might thaw more quickly if abrupt that processes were included, but I think the main point of Turetsky et al. is that the carbon consequences of these abrupt thaw processes could be large . . . even if the actual area affected is actually quite small. Maybe should clarify.

22. I'm not convinced by the back of the envelope calculation at the end of the paper. Implicit in the D diagnostic is the fact that the seasonal thaw length is extended as well as a deepening of the active layer (and possible talik formation, though as noted above, maybe not much talik within 2m of surface). So, not sure that it is appropriate to simply multiply the change in D by the carbon stock to get the C vulnerable to decomposition. It's quite a bit more complicated than that. Koven et al. (2015) attempted to make this calculation, I think. Further, just because carbon is made vulnerable to decomposition, it doesn't mean that that is a committed carbon loss. Some or even a lot of that carbon could stay in the soil for a long time due to the slow decomposition rates in the still cold and moist soils. Perhaps it would be better to just remove this brief analysis.

Refs:

Koven, C.D, E. A. G. Schuur, C. Schädel, T. Bohn, E. J. Burke, G. Chen, X. Chen, P. Ciais, G. Grosse, J. W. Harden, D. J. Hayes, G. Hugelius, E. E. Jafarov, G. Krinner, P. Kuhry, D. M. Lawrence, A. H. MacDougall, S. S. Marchenko, A. D. McGuire, S. M. Natali, D. J. Nicolsky, D. Olefeldt, S. Peng, V. E. Romanovsky, K. M. Schaefer, J. Strauss, C. C. Treat, M. Turetsky, 2015. A simplified, data-constrained approach to estimate the permafrost carbon-climate feedback. Phil. Trans. R. Soc. A, doi.org/10.1098/rsta.2014.0423.

Lawrence, D.M. R.A. Fisher, C.D. Koven, K.W. Oleson, S.C. Swenson, G. Bonan, N. Collier, B. Ghimire, L. van Kampenhout, D. Kennedy, E. Kluzek, P.J. Lawrence, F. Li, H.

Li, D. Lombardozzi, W.J. Riley, W.J. Sacks, M. Shi, M. Vertenstein, W.R. Wieder„ C. Xu, A.A. Ali, A.M. Badger, G. Bisht, M. van den Broeke, M.A. Brunke, S.P. Burns„ J. Buzan, M. Clark, A. Craig, K. Dahlin, B. Drewniak, J.B. Fisher, M. Flanner, A.M. Fox, P. Gentine, F.Hoffman, G. Keppel-Aleks, R., Knox, S. Kumar, J. Lenaerts, L.R. Leung, W.H. Lipscomb, Y. Lu, A., Pandey, J.D. Pelletier, J. Perket„ J.T. Randerson, D.M. Ricciuto, B.M. Sanderson, A. Slater, Z.M. Subin, J. Tang, R.Q. Thomas, M. Val Martin, and X. Zeng, 2019. The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. JAMES, doi.org/10.1029/2018MS001583.

Van Kampenhout, L., J.T.M. Lenaerts, W.H. Lipscomb, W.J. Sacks, D.M. Lawrence, A.G. Slater, and M.R. van den Broeke, 2017. Improving the representation of polar snow and firn in the Community Earth System Model. JAMES, 9, 2583-2600, doi.org/10.1002/2017MS000988.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-309, 2020.