

Earth Syst. Dynam. Discuss., author comment AC4
<https://doi.org/10.5194/esd-2020-85-AC4>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on CC1

Benjamin M. Sanderson et al.

Author comment on "The potential for structural errors in emergent constraints" by
Benjamin M. Sanderson et al., Earth Syst. Dynam. Discuss.,
<https://doi.org/10.5194/esd-2020-85-AC4>, 2021

While the three comments from reviewers so far seem to complain that the article in question is more like a review article than an article describing new research, I think this is a good reason to publish it, not a negative. While I am not a researcher who develops earth system models, in most areas of applied science there are far too many "original" research articles, and far too few articles discussing the strengths and weaknesses of how very complex models like ESMs are being developed and applied in making forecasts. After CWIP6 it is certainly time for climate system modelers to step back and assess the value of trying to analyze multi-model ensembles, especially when the results of so many models are so different. Needless to say, policy makers probably do not have the slightest idea of how to evaluate these differences.

Many thanks to the commenter for their response to our study. We agree that often in literature which assesses high level ensemble relationships and distributions, there is a lack of consideration of the fundamental assumptions which might be common between models in the ensemble, and the degree to which these assumptions might influence our confidence in projections. It was certainly our aim in this study to highlight - for the particular application of emergent constraint literature, how such consideration could result in overconfidence in some studies published to date.

The gist of the current article under review is that the key focus in developing climate science models, like all scientifically-based models, must be on the scientific basis for each major process modeled. Thus, rather than try to use emergent constraints to help make forecasts more accurate, one should focus on improving the scientific basis for how each process is modeled.

Agreed

To me, this implies trying to get more scientific consensus as to how each process should be modeled, and not to think that diverse methodologies are helpful.

We would disagree on this point. Diverse methodologies to represent a process are invaluable to represent uncertainty in a process and how it might evolve under future climate conditions. In some ways, the greatest danger is the perception of diversity in methodology when in fact there are a number of processes in ESMs where there are

common or similar parameterisation schemes across the ensemble, effectively differing only in parameter values (e.g. soil respiration-temperature relationships). The ensemble can thus produce very strong internal relationships (because both future and past changes are a function of the same small set of parameters), but the constrained result excludes uncertainty associated with aspects of the process which are not represented in the current set of models.

This is especially true when the number of degrees of freedom is too limited in order to get reasonably accurate backcasts for each process in all past time periods. This issue exists for other types of scientific models, and even applies to more policy oriented models like the integrated assessment models used to create climate change mitigation scenarios.

Agreed (especially for IAMs). For Earth System Models, hindcasts are standard, but we agree that there is the strong capacity for compensating errors in reproducing historical climate change (in forcing, feedback and heat uptake - for example, in the case of transient global mean temperature changes), such that the data is insufficient to constrain some of the key degrees of freedom which control the dynamics of the system under mitigation. There is also - as you suggest, a tendency to focus on aspects of historical performance which the models can reproduce - while ignoring aspects which the models cannot reproduce. A recent talk by Cristian Proistosescu on a paper in preparation discussed an excellent example of this in the failure of current generation models to represent recent tropical warming patterns, which have the capacity to bias estimates of global climate sensitivity constrained by recent warming (<https://www.youtube.com/watch?v=5cO1TQtGyHc&t=17s>)

My takeaway from the article is that having fewer earth systems models that include more consensual methodologies for each process modeled would be better than the current situation. More cooperation between climate scientists, then, would be helpful for make each process modeled in a sufficiently detailed and more scientifically accurate way, rather than the more approximate approaches that different modeling groups currently take.

There is an important discussion to be had on how to approach the modeling problem going forward to best represent uncertainty in projections. We would argue that a diversity of approaches is useful, perhaps essential, but the current ad-hoc sampling of structural assumptions, lack of consistent parameter sampling and treatment of the multi-model archive as a distribution of equally plausible models represents a genuine problem. Weighting strategies might address some of these issues (see Brunner 2020, Sanderson 2017), as could operational parameter perturbation experiments in climate modeling (see Rostron 2020). But, as the commenter suggests, there is potential to think beyond what is possible in the current multi-center approach, and Fisher and Koven (2020) laid out a perspective for land surface modeling which considered the representation of structural process uncertainty within a common framework.

To summarize, I think the implication of this article, whether or not the authors would want to say this, is that there are too many different modeling groups in the world, and thus too many different models, each of which is too simplistic, and which don't agree with each other sufficiently to have much comfort in the overall progress of the field.

We don't see a problem with diversity of approaches *per se*, but we do agree that not enough focus is given to the structural assumptions within models - and this problem is particularly acute in the application of emergent constraints to an ensemble of models which in many cases share assumptions. From a climate risk perspective, our position is that greater focus could be placed on what current models might be missing, rather than

on further constraining a distribution of projections of future change which is already a collection of best-estimates and ill-suited to the assessment of tail risks for future impacts.

Unfortunately, there is a tendency in most fields for too many research groups to exist, often supported by their national governments, which tends to prevent greater scientific cooperation and agreement between model experts because they focus their work on their own models. Competition is not always a plus whether in dealing with science or other social issues. The right balance needs to be achieved between cooperation and competition to maximize scientific progress, and the field of earth system modeling is no exception.

We broadly agree - that there are huge benefits to open, collaborative science which are often impeded by national interests and funding strategies. We would note also that there are some fine existing examples of collaborative science in climate modeling in the field, which successfully span borders - aided substantially by the progress of github and community development tools over the last decade (see the Community Earth System Model <https://github.com/ESCOMP/CESM> , the FATES land surface model <https://github.com/NGEET/fates-release> , and the FAIR simple climate model <https://github.com/OMS-NetZero/FAIR>).

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, 11(4), 995-1012.

Sanderson, B. M., Wehner, M., & Knutti, R. (2017). Skill and independence weighting for multi-model assessments. *Geoscientific Model Development*, 10(6), 2379-2395.

Rostron, J. W., Sexton, D. M., McSweeney, C. F., Yamazaki, K., Andrews, T., Furtado, K., ... & Tsushima, Y. (2020). The impact of performance filtering on climate feedbacks in a perturbed parameter ensemble. *Climate Dynamics*, 55, 521-551.

Fisher, R. A., & Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *Journal of Advances in Modeling Earth Systems*, 12(4), e2018MS001453.