Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-198-RC1, 2018

Geoscientific Model Development Discussions

Interactive comment on "Improving climate model accuracy by exploring parameter space with an O(10^5) member ensemble and emulator" *by* Sihan Li et al.

Anonymous Referee #1

Received and published: 23 December 2018

Review

Improving climate model accuracy by exploring parameter space with an O(105) member ensemble and emulator by Li et al.

The paper describes an approach to identify parameters in the atmosphere and land-surface models that control warm and dry biases in summer in the northwestern US with an aim of finding parameter combinations that reduce these regional biases. The authors construct PPEs by perturbing 17 parameters simultaneously using the weather@home modeling framework. They use an 'iterative refocusing' approach to find parameter combinations that have satisfactory performance over the NWUS while also maintaining acceptable performance globally.

The PPE approach, in general, is extremely powerful in providing insights into model development and improvements. Parametric uncertainty is also one key area which remains relatively unexplored. So in my opinion, this paper makes an important contribution. The paper is well-structured and clear. The assumptions are clearly stated and the results sufficiently support the interpretations and conclusions. The abstract does provide a concise and complete summary. I also highly appreciated the fact that the authors clearly explain what roles each of the important parameters play in the model. Overall, I strongly feel that the paper is worthy of publication in GMD. I, however, have a few concerns that I would like the authors to address.

General comments:

Item 1: About the parameters perturbed: it is not made clear whether the parameters reside in the global model or the regional model. If the parameters belong to HadAM3P, which I think is the case, then the statement (l.206) "PPEs for parameter refinement with the aim of improving regional climate models (RCMs)" needs to be rephrased. The improvement is in terms of regional biases but the RCM in itself is not improved. If the parameters belong to the RCM, then I don't fully understand the global TOA flux constraints in Phase 1.

The purpose of using a regional model setup is unclear. Assuming that the parameters belong to HadAM3P and since the comparisons with observations are carried out at HadAM3P resolution, what role does the regional model (HadRM3) play here? In a topographically complex region such as the NWUS, model resolution likely plays an important role in model performance, but it's not discussed in the paper despite using a regional model. It also appears that the nesting is one-way wherein the RCM is nested within the global model, which suggests that any improvements in the RCM are not felt by the global model. These issues need clarification in the main text.

Item 2: The parameter refocusing is entirely carried out using AMIP-style simulations.

Printer-friendly version

Discussion paper

What are the implications of this for using these model variants in coupled model simulations? By perturbing parameters, the model response to future scenarios could be amplified or dampened relative to the standard version of the model, which may lead to differences in behavior between atmosphere-only and coupled model simulations. Is tuning a model in an AMIP mode for better regional performance useful for coupled model simulations?

Item 3: I wonder if the phrase ‘perturbed parameter’ should be used here instead of ‘perturbed physics.’ The authors perturb parameters instead of switching between schemes with different physics. For instance, Shiogama et al. (2014) generate multi-parameter, multi-physics PPEs by perturbing parameters in different versions of MIROC. So, while the words ‘physics’ and ‘parameter’ are used interchangeably in the context of PPEs, perhaps it is useful to make that distinction. [Shiogama, H., Watanabe, M., Ogura, T., Yokohata, T. and Kimoto, M., 2014. MultiâĂŘparameter multiâĂŘphysics ensemble (MPMPE): a new approach exploring the uncertainties of climate sensitivity.ÂăAtmospheric Science Letters,Âă15(2), pp.97-102.]

Item 4: The authors argue that they use ‘history matching’ citing McNeall et al. and Williamson et al. Doesn’t history matching require a formal statistical framework that is based on the definition of Implausibility? My reading of your approach is that it borrows ideas of from history matching (e.g., iterative refocusing) but doesn’t strictly follow it. It would be useful to have a couple of sentences explaining the differences and similarities between your approach and history matching.

Item 5: (l.99) Biases in the regional model are shown in Figure S1, but it would be interesting to show what these biases look like in HadAM3P.

Item 6: Fig. S3 suggests that regional temperature and precipitation metrics are strongly anti-correlated across the PPE. Doesn’t that suggest a physical link between these variables that can be exploited to find parameter combinations acceptable for both variables?

Item 7: I am a bit surprised that the sensitivity analysis shows very little interaction between the parameters especially at 10-year timescales. My reading of the paragraph starting at l.468 is that the small interaction terms are the property of your emulators. Is that the case? While entrainment coefficient and the ice fall speed are often the dominant parameters, I wonder if inadequate accounting of parameter interactions might prevent you from finding parts of parameters space that might be plausible.

Item 8: Fig. 5 – I am not sure what the purpose of Fig. 5 is since observations are not included in the figure. Yes, the selected 10 models show a substantially smaller spread compared to phases 1 and 2, but in JJA all the selected models show considerably different results compared to the SP and whether that's an improvement or not is not discussed in the text.

Item 9: Fig 6i indicates that the cold bias in DJF gets slightly worse in the PP simulations compared to SP even though there are improvements in the other seasons and regions. This needs to be mention in the paper since NWUS is the target region. On l.492, you do mention that the DJF-Pr bias worsens slightly in Phase 3. It suggests that reducing region/season-specific biases is difficult because of its implications for other seasons.

Item 10: Following on from the point above and the description on lines l.513-529 and l.662, it is clear that a more process-level evaluation is required to find parts of parameter space that truly provide improvements for right physical reasons. The authors should state explicitly (maybe at l.662) that looking at seasonal mean biases in temperature and precipitation is insufficient to fully assess model performance and a more process-based analysis could strengthen the validity of the chosen parameter sets.

Item 11: (Fig. 8) Why use only reanalysis datasets for precipitation? Aren't GPCC, CMAP, TRMM better observational precipitation products for model validation?

Item 12: (l.717) The number of initial conditions members, 10 or smaller, seems small especially because the performance is evaluated at regional scales and at climate

(~10-year) timescales. Recent work has demonstrated that large initial conditions ensembles show very different trends on longer timescales. Can the authors please comment on how their parameter refinement might be affected by model's internal variability characteristics? Is variability in the model strongly coupled to the observed SSTs for the NWUS?

Minor comments:

l.381: It would be useful to have a brief explanation of what the 'ranges of acceptability' are for TOA fluxes here. I understand that this has been described in the Appendix and in Fig. 1 caption, but this is critical information and should be concisely described here.

l.392: missing 'a' in Yamazaki

Fig. S3 caption should read 'Same as Fig. S2'

l.423: should be 'lead to decreased outgoing ...'

l.577: should 'increased' be 'improved'?

l.725: Please mention that you use LHS space-filling design earlier in the paper, perhaps at l.241.

It would be helpful to have a single document that contains all the supplementary figures and text.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-198, 2018.