Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-198-AC1, 2019

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.

Sihan Li et al.

sihan.li@ouce.ox.ac.uk

Received and published: 18 February 2019

Response to Referee #1

General response:

Thank you very much for these comments. We feel very encouraged from this review and have made the best attempt to respond to these constructive comments.

Response to item 1:

We thank the reviewer for these constructive comments and agree with the assessment that these issues should be clarified in the main text. We have added a few sentences in

the main text (in Section 2.3 Perturbed parameters) to clarify the parameters perturbed " the parameters reside in the global model as well as the regional model, and are set to the same values in HadAM3P and HadRM3P in the experiments used in this study, thus any improvement in regional biases is considered to have been improved through the improvement of the RCM itself".

The RCM is embedded within the GCM, and nesting is one-way where the GCM provides boundary conditions for the RCM. The purpose of the global TOA flux constraints in Phase 1 was to make sure the large scale boundary conditions for the RCM are realistic. Within weather@home the purpose of the regional model is to provide higher resolution output over the area of interest, rather than to feedback scale dependent features to the GCM. The comparisons with observations are carried out at both HadAM3P resolution and HadRM3P resolution, and the regional model biases are calculated with respect to PRISM datasets. We have changed the wording in the main text (l.333 in the original manuscript), which now reads "biases of the regional model outputs are calculated with respect to PRISM". We agree with the reviewer that, in a topographically complex region such as the NWUS, model resolution does play an important role in model performance, and perhaps the model parameters should be resolution dependent, i.e., they should be set to different values in the GCM and RCM. This is indeed an important issue that needs to be further explored. In follow-on work, we have performed additional PPEs, where the parameter values are set to be different in HadAM3P and HadRM3P, and we will attempt to address the resolution-dependency of parameter values in a following paper using those PPEs.

Response to item 2:

Thank you for these comments. We agree with the reviewer that by perturbing parameters, the model response to future scenarios could be different from the standard version of the model. As part of our ongoing work, we are looking at the different climate sensitivities of different model variants in future projections. However, it was never the intention of these experiments to come up with model variants to be used in

coupled model simulations. Parametric uncertainty exists in the ocean component as well. If an ocean component were to be coupled to the tuned AMIP mode, a systematic parameter refocusing would be required to ensure consistency across the full climate system. It is likely that different atmospheric parameter settings would be selected due to potential interactions with oceanic parameters.

Response to item 3:

Thank you for pointing this out. We agree the phrase “perturbed parameter” should be used here instead of “perturbed physics”, since the latter could be interpreted as switching different physics schemes rather than change parameter values. We have changed the main text throughout to only use the phrase “perturbed parameter”.

Response to item 4:

We appreciate this comment and have added a couple of sentences at the end of paragraph 10 to explain the differences and similarities between our approach and history matching. Now it reads “The method we adopt in this study borrows the idea of ‘iterative refocusing’ from the third category, where the parameter values are refined through phases of experiments, but our method differs from history matching in that we do not employ a formal statistical framework based on the definition of implausibility”. We believe now it’s much clearer in the main text that the approach adopted in our study does not strictly follow history matching.

Response to item 5:

We respectfully point out that the temperature and precipitation biases in HadAM3P are shown in Figure 6 and Figure 7 respectively.

Response to item 6:

Thank you for these comments. We agree with the reviewer that regional temperature and precipitation are strongly anti-correlated in JJA, and this does suggest a physical link that can be exploited to find parameter combinations acceptable for both vari-

ables. Multivariate parameter sensitivity is indeed one of the research questions we will explore with our additional PPEs, which are follow up experiments of our current manuscript.

Response to item 7:

The interaction terms are not the property of emulators. In the extended Fourier Amplitude sensitivity analysis (FAST; Saltelli et al., 1999), the fraction of the total variance due to interactions is computed from the parameter contribution to the residual variance (variance not accounted for by the main effects). The relative importance of the interaction term is dependant on the metrics of interest. For example, the interaction term is non-trivial for the root mean squared error of JJA 850-hPa U (V) wind component in PP simulations over the regional model domain, with respect to SP simulations, as seen in Figure 1 and Figure 2 below.

Figure 1. Sensitivity analysis of the root mean squared error of JJA 850-hPa U wind component in PP simulations over the regional model domain, with respect to SP simulations, via the FAST algorithm of Saltelli et al. (1999).

Figure 2. Sensitivity analysis of the root mean squared error of JJA 850-hPa V wind component in PP simulations over the regional model domain, with respect to SP simulations, via the FAST algorithm of Saltelli et al. (1999).

Response to item 8:

We appreciate this comment and have added a few sentences to clarify about Fig. 5 “ The SP simulations have warm and dry biases over NWUS and mid-latitude land in general (as shown in Fig. 4, Fig. 6 and Fig. 7). In JJA all the selected PP model variants show considerably different results compared with the SP-cooler and wetter, i.e. reduced biases and improved model performance”. We have also updated Fig. 5 to include the results from different initial conditions to demonstrate that varying model parameters has more influence on the result than varying initial conditions, which helps

clarify item 12 as well.

Response to item 9:

Respectfully, Fig. 6i shows that the PP simulations are warmer than the SP, which indicates the cold biases in DJF gets slightly better in the PP simulations. We agree with the reviewer that reducing region/season-specific biases is difficult because of its implications for other seasons and regions. Parameter refinement exercises like ours are faced with the common dilemma where model revisions yield improvement in one field or regional pattern or season, but degradation in another, or improvement in the model climatology but degradation in the interannual variability representation etc. That is why we feel strongly that any parameter refinement process is and should be tailored to the scientific objectives of the experiments, because there is no one fit for all solution; more importantly, that's why we are advocating that parameter refinement process should be more explicit and transparent, because choices and compromises made during the refinement process may significantly affect model results and influence evaluations against observed climate, hence should be taken into account in any interpretation of model results, especially in multi-model intercomparison studies to help understanding of model differences.

Response to item 10:

We appreciate this comment and have added a few sentence as suggested to explicitly state this point "Furthermore, looking at biases in seasonal mean temperature and precipitation is insufficient to fully assess model performance. As a follow-up step to this study, we recommend a process-based model evaluation and physical explanation of model improvements to further refine the parameter space that provides improvements (e.g., reduce summer biases) for the right physical reasons. For example, more accurate representation of clouds in the model could lead to better simulated downward solar radiation at the surface, as well as better simulated surface energy and water balance."

Response to item 11:

In the main text, we showed model outputs in comparison with CRU datasets, simply so that temperature and precipitation are compared with the same data source. In the supplementary information, we showed the biases of precipitation compared with other datasets, which includes GPCC (Fig. S18). Now we have added biases of precipitation compared with CMAP and TRMM in the supplementary information of the revised manuscript.

Response to item 12:

Thank you for these comments regarding initial conditions. Indeed our original intent of having multiple ICs is so that the results would be representative of the parameter perturbations instead of reflecting the influence of any particular IC. Previous work (Bellprat et al., 2011, Figs. 1-4; Covey et al., 2011, Fig. 4) suggests that varying model parameters has more influence on climate than varying initial conditions. For the metrics we used (e.g. seasonal mean temperature and precipitation), our results show the same in the updated Fig. 5. To answer if internal variability affects different parameter sets differently, we will need to generate identical large initial condition ensembles for each parameter set, which is an interesting research question but beyond the scope of this study.

In previous work (Li et al., 2015) we have compared HadRM3P-HadAM3P results with NARCCAP, and found that that the coupled HadRM3–HadCM3 (HRM3-HadCM3 in NARCCAP convention) demonstrated similar skills in simulating temperature and precipitation as HadRM3P– HadAM3P, even though HadAM3P is an atmosphere only model and SSTs are specified whereas HadCM3 is a coupled ocean–atmosphere model. This similarity between HadRM3P–HadAM3P and HadRM3–HadCM3 suggests that the dynamical coupling between ocean and atmosphere in NARCCAP did not explain most of the difference between HadRM3P–HadAM3P and the various NARCCAP RCM–GCM pairings but that the differences were due mainly to the atmospheric

dynamics.

Minor comments:

Response to comment on l.381:

Thank you for pointing this out. Indeed it would be helpful to add some explanation of 'ranges of acceptability' in the main text. We have added a brief explanation of what the 'ranges of acceptability' at the beginning of Section 3.1 in the revised manuscript.

Response to comment on l.392:

Apologies for the sloppiness here, this has been fixed.

Response to comment on l.423:

This has been changed.

Response to comment on l.577:

We respectfully point out that in SP, the model has dry biases over northern South America, equatorial Africa, and south Asia in JJA. In PP simulations, the dry biases are stronger compared with SP (Fig. 6k), so the biases are increased in PP simulations.

Response to comment on l.725:

Thank you for pointing this out. We have now stated earlier in the paper (l.241 in the original manuscript) that we used LHS space-filling design.

Response to minor comment on single document that contains all the supplementary figures and text:

We will upload a single document containing all the supplementary figures and text as revised manuscript.

Reference:

Bellprat, O., Kotlarski, S., Lüthi, D. and Schär, C., 2012. Exploring perturbed physics ensembles in a regional climate model. Journal of Climate, 25(13), pp.4582-4599.

Covey, C., Brandon, S., Bremer, P.T., Domyancis, D., Garaizar, X., Johannesson, G., Klein, R., Klein, S.A., Lucas, D.D., Tannahill, J. and Zhang, Y., 2011. A new ensemble of perturbed-input-parameter simulations by the Community Atmosphere Model. Technical report, Lawrence Livermore National Laboratory, Livermore, CA.

Li, S., Mote, P.W., Rupp, D.E., Vickers, D., Mera, R. and Allen, M., 2015. Evaluation of a regional climate modeling effort for the western United States using a superensemble from weather@ home. Journal of Climate, 28(19), pp.7470-7488.

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

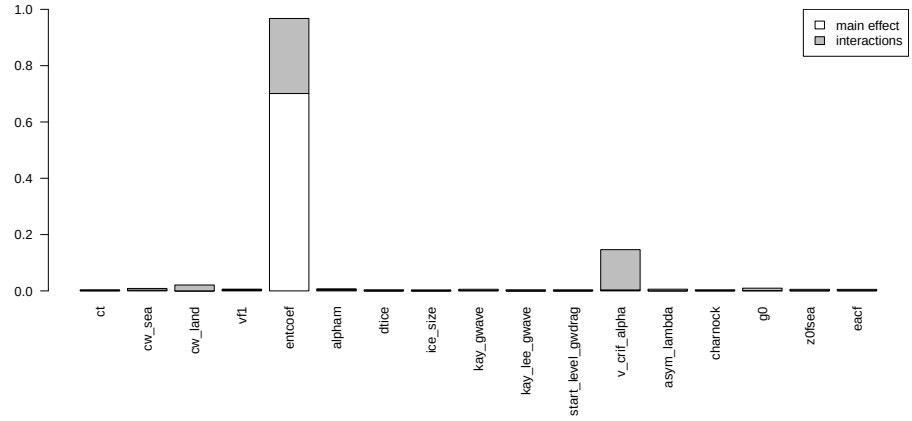
1.0
0.8
0.6
0.4
0.2
0.0
main effect
interactions
ct
cw_sea
cw_land
vf1
entcoef
alpham
dtice
ice_size
kay_gwave
kay_lee_gwave
start_level_gwdrag
v_crif_alpha
asym_lambda
charnock
g0
z0fsea
eacf

Fig. 1.

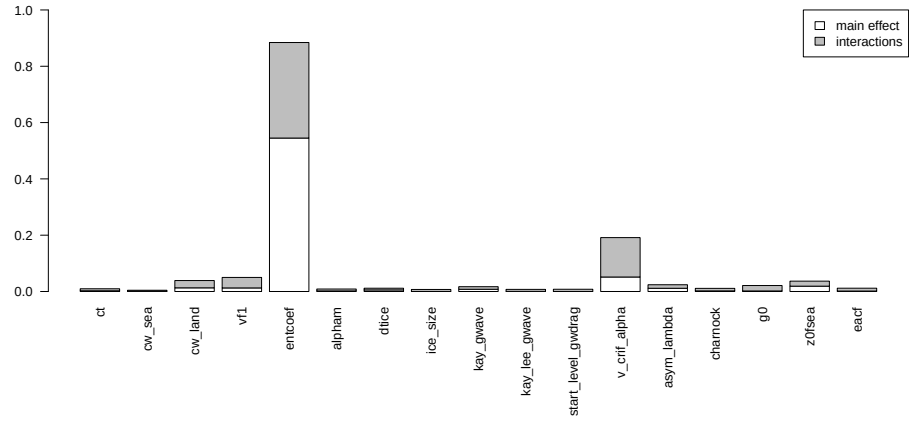
1.0
0.8
0.6
0.4
0.2
0.0
main effect
interactions
ct
cw_sea
cw_land
vf1
entcoef
alpham
dtice
ice_size
kay_gwave
kay_lee_gwave
start_level_gwdrag
v_crif_alpha
asym_lambda
charnock
g0
z0fsea
eacf

Fig. 2.