Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-535-RC1, 2016

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Multi-model dynamic climate emulator for solar geoengineering" *by* Douglas G. MacMartin and Ben Kravitz

Anonymous Referee #1

Received and published: 1 July 2016

The authors propose a nonparametric emulator aimed at reproducing geoengineering scenarios. Using dynamical linear models, they propose a formulation of the emulator as a convolution of the forcing and the impulse responses, and test this approach for two geoMIP scenarios for some variables of interests.

The manuscript is overall well written and presents an interesting problem, but I believe that in its present form is not suitable for publication and, in order to be reconsidered, needs to be considerably improved in many parts. The proposed method would have considerable limitations if it is to be expanded beyond the narrow context of this work, e.g. annual averages for two model runs. Further, the validation setting is extremely limited, not based on any metric, and completely ignores the emulation uncertainty.

General comments

- The validation setting is extremely limited: the proposed approach is fit for the G1 scenario, and then used to extrapolate G2. Also, for the G1 scenario the emulator is likely to work well, since it consists of impulse functions. A considerable amount of work is needed to perform more tests under different forcing scenarios. While the geoMIP is limited in size, the CMIP5 or other large multi-model ensembles could be used to validate the forcing part of this emulator.
- The present version puts very little emphasis of the uncertainty of the scenario estimation. The validation essentially consists in eyeballing many plots of the emulator against the original computer model, with no attempt to quantify the fit or, most importantly, to asses how the internal variability of the model is reproduced by the emulator. The definition itself of 'climate variability' $n_i(t)$ is unclear. Are the authors assuming a white noise? Also, I would assume that this noise is independent for different variables, but it should be clearly stated.
- This approach will have significant limitations at finer temporal scales. The authors briefly discuss this when they mention how we can impose $h = h(\tau, m)$. This solution is not straightforward, as a nonparametric estimation of 12 different impulse responses will require more scenarios (surely more than two) to have reliable estimates. The authors somewhat acknowledge it when they state that additional simulations would be required, but in an off-the-shelf ensemble such as geoMIP, where no more scenarios are readily available, this is a strong limit of this approach. This will be become even more evident for finer temporal scales, e.g. weekly or daily data.
- The results and the discussion do not mention model differences, and most importantly what do they mean. Does the emulator estimate different impulse responses for different models? I would expect so, and I would expect these differ-

ences to convey information on how the models differ. For example, HadCM3 and HadGEM2-ES will likely display similar responses as both models are released from the Hadley Centre.

- The part on grid-scale emulation must be extended. Firstly, the methodology is unclear: a clear explanation of how were the EOFs selected must be presented, either in the main text or in the supplement. Secondly, as before, a more formal assessment of the pattern similarity is needed, as eyeballing figure 5 is not enough to convince that the emulator is performing well.

Specific comments

- Title: what the authors present is not a multi-model emulator, in the sense that it independently fits each model and does not assume interdependencies.
- pag. 1 l.16-17. The claim that the 'emulator prediction may be a more accurate estimate [...] of the models' response than an actual simulation' is very questionable. The emulator is not meant to replace a climate model, it's just a faster approximation that is used to explore the input space in a computationally efficient manner. While emulators are arguably a useful tool for calibration and, as in this case, scenario extrapolation, they cannot replace the physics of the climate model and they are useful only as long as the training set from the climate model is meaningful.
- pag 1. l.19-20. Actually, emulators are much more popular in model calibration and local sensitivity analysis of physical parameters then in projections of anthropogenic forcings. Only very recently this methodology have been extended to deal with forcings. This introductory part must be rewritten with a more extensive literature review on traditional emulators.

- pag. 4, eq (1) and onwards. It is somewhat inappropriate to represent the emulator as a convolution given that the authors are effectively using just annual averages. A reformulation in terms of discrete sums is necessary.
- pag. 4, line 101. $h(\tau)$ was never defined.
- pag. 6, line 161. Poor choice of pedix in $f_t(t)$, please reformulate.
- Figures. What is the unit measure of precipitation? Also, are the all figures expressed as anomaly with respect to a reference value? If so, what is it?

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-535, 2016.