

## ***Interactive comment on “Introducing the Probabilistic Earth-System Model: Examining The Impact of Stochasticity in EC-Earth v3.2” by Kristian Strommen et al.***

### **Anonymous Referee #2**

Received and published: 3 May 2019

General comments =====

The authors present a well-written manuscript about the introduction of stochastic schemes into the atmosphere and land-surface scheme of EC-Earth, testing the impact on the various schemes on the mean biases as well as on variability. Stochasticity in a climate or earth-system model context is still rather novel and the work presented here is a valuable addition in the field. I therefore recommend the manuscript for publication in GMD after some necessary revision.

Specific comments =====

Title: the title mentions the "Probabilistic Earth-System Model" while indeed all the  
C1

work has been done entirely with the atmosphere component of EC-Earth only. In my eyes the title this suggests more than what the manuscript delivers and to avoid any too far reaching expectation I'd therefore suggest to modify the title to something more adequate.

Section 2.1, 1st paragraph: there is no need to go into details about EC-Earth's ocean component or coupler because they are irrelevant in this work.

Section 2.2, eq 1: what is "the i'th physics parameterization scheme"? You mention the schemes explicitly towards the end of Sec 3.1, why don't you list them here already?

Section 2.2, eq 1: are the tendencies for all variables perturbed by the same  $r$ ? Or are there different perturbations for the different variables?

Section 2.2, eq 1: are the perturbations constant in time, or do they vary from one timestep to the next?

Section 2.2, end of 1st paragraph: you say "perturbation is limited between 0 and 2" but shouldn't that be between -1 and 1? Even with  $r=-1$  we still get that  $\hat{P}$  has the same sign as  $P$ , or?

Section 3.1, p.6 l.13: why do you use ERA-40 or ERA interim for the evaluation? These re-analysis datasets belong to the same model "family" as EC-Earth and may therefore share common biases, in particular where the re-analysis are largely a model product not constrained by observations. I would prefer the evaluation to be done against a different re-analysis dataset, but that may be too much to ask for at this stage. In any case, it would be good if the similarity between EC-Earth and the model used for the re-analysis would be clearly stated.

Section 3.4: how good is it to compare only the last 10 (5) years from the last ensemble member against observations? Is that short time period really sufficient to confirm or reject the fitness of EC-Earth? I don't understand why you don't average 10-20 year from all ensemble members (start dates) and compare against the corresponding time

period of the re-analysis (see also comment above about choice of re-analysis data).

Section 4,5,6: when you talk about cloud water do you mean the gridpoint value of the in-cloud value? I would suspect it's the gridpoint value, and if that's the case then the reduction in cloud cover would imply an even larger increase in the in-cloud water content with all the consequences for optical thickness and cloud microphysics. Good if you could clarify on this point.

Section 4.2,5.2,6.2 and Fig 8: I don't see a point in making a timeseries of the energy budget, the interesting aspect for a climate model is how well it simulates the average flux and its variability compared to observations (re-analysis). For that reason it would make more sense to average the biases in the fluxes and present them as a table or barplot similar to how you did in Figs 1 and 2.

Section 7: what is the motivation for selecting the Hadley circulation and the QBO as test cases for the stochasticity? Why not NAO or PNA as other prominent modes of atmospheric variability?

Section 7.1 Fig 9: It is not easy to distinguish the different colors of the dots (I am slightly color blind) and you should consider presenting the data in a different way, e.g. by using different symbols for the different members/time periods.

Section 7.2: what is the reason for evaluating QBO only for the last ensemble member? Why not using the results from all start dates?

Section 8, 1st sentence: you cannot blame the absence of a more process-oriented analysis on the lack of available data because you have designed the experiment and its output.

Section 8.1, p.14 l.7: it's not clear to me how an increased in-cloud liquid water content could steepen the near-surface humidity gradient. Could you explain better what you mean?

Section 8.2, p.15 l.13: what do you mean with runoff being a key driver of land-

C3

atmosphere interaction? Isn't runoff simply the difference between P-E and the amount of water absorbed by soil? It's a residual, not a driver, or?

Section 8.2, p.15 l.22: why do you call runoff a tuning parameter in the LAND case? Runoff is an important diagnostics of the model that can be used to tune the model, but runoff itself isn't a tuning parameter.

Code availability: EC-Earth is licensed and not openly available, it's not sufficient for any presumptive user to request access. I would suggest you check the guidelines of GMD that regulate code availability and re-phrase this section. ([https://www.geoscientific-model-development.net/about/code\\_and\\_data\\_policy.html](https://www.geoscientific-model-development.net/about/code_and_data_policy.html))

Technical corrections =====

Sorry, I am not a native English speaker and cannot really comment on the language, but two things struck my eye:

p.2 l.12: "However" seems to be inappropriate in this sentence.

p.15 l.15: shouldn't it rather be "...none of the schemes is able...?"

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-337>, 2019.

C4