

Interactive comment on “Baseline Evaluation of the Impact of Updates to the MIT Earth System Model on its Model Parameter Estimates” by Alex G. Libardoni et al.

Anonymous Referee #2

Received and published: 14 May 2018

The authors compare the most recent version of the MIT Earth System Model (MESM) with an earlier version and document the impact of updating changes in external forcing over the historical period and of the land surface scheme. They present results from an 1800-member ensemble simulation over the industrial period and compare the model results to two observational metrics of change (ocean heat uptake, surface air temperature pattern). They also present a 372-member ensemble in idealized forcing scenarios to establish the links between TCR, thermosteric sea level rise, ECS, and ocean heat uptake. The illustration of these links in Figure 5 is nice and interesting. On the less positive side, it is surprising that the authors do not invoke a larger suite of observational constraints to estimate probability density distributions of TCR and other

C1

model outcomes. The method section requires more work and there is also a lack of discussion in section 5.

1) I miss a description of the basic components and parameterizations of the model in the method section 3.

I miss a section that describes model spin-up and the setup for the different model simulations, including external forcing factors.

Further, it is not evident from the description why the model is called “Earth System Model”. For example, are biogeochemical cycles included? Does dynamic vegetation affect albedo? Is it an ESM or rather an Earth System Model of Intermediate Complexity?

I also miss a brief description of the metric used to compare model and data and how they are used to derive probability distribution. It is not sufficient to refer the reader to the literature (Libardoni and Forest 2011).

2) Section 3: The authors vary three parameters – ocean diffusivity, an aerosol forcing scaling, and the strength of the cloud feedback determining ECS and constrain the models with two parameters.

2a) There is little information in the method section what these parameters specifically influence. The aerosol forcing scaling is unclear. Does this mean that all aerosol forcings are lumped together and scaled with a constant time invariant factor? How are different uncertainties applying to different aerosol classes (e.g. sulfate versus soot) considered or not and what is the justification for this approach. Please discuss caveats related to your assumption of a scaling factor.

2b) Effective ocean diffusivity is a very loose term. Is this diapycnal, vertical or horizontal diffusivity or does the parameter refer to the diffusivity associated with Gent-McWilliams parameterization? The subscript v of K_v points to vertical diffusivity. I would hope that this parameter reflects diapycnal diffusivity as diapycnal diffusivity co-

C2

governs ocean overturning strength and thus surface-to-deep heat transport. In any case, I am puzzled about the range sampled. Diapycnal diffusivity in coarse resolution, dynamic ocean models is typically of order $0.1 \cdot 10^{-4} \text{ m}^2 \text{ s}^{-1}$. Here diffusivity is varied in steps of $1 \cdot 10^{-4} \text{ m}^2 \text{ s}^{-1}$ and a very wide range up to $64 \cdot 10^{-4} \text{ m}^2 \text{ s}^{-1}$ is used. The upper value is even much larger than applied in classical box-diffusion models ($1\text{-}2 \cdot 10^{-4} \text{ m}^2 \text{ s}^{-1}$); in box-diffusion models the entire vertical transport (mixing, advection, convection) is parameterized by diffusion only. What is the justification for this large sampling range? As a minor point, please use SI units for diffusivity. Further, I though Gent-McWilliams parameterization is included in the MIT model. If yes, why is the Gent-McWilliams diffusivity not varied or is this parameter linked with the “effective diffusivity”?

2c) ECS is typically used to abbreviate Equilibrium Climate Sensitivity. Here, an effective climate sensitivity is introduced and termed ECS. What represents this effective climate sensitivity?

3) Section 3: I question somewhat the application of only two observational metrics to constrain ECS, TCR, and sea level rise. Namely, pattern of surface air temperature change and “linear” ocean heat uptake are used as constraints by the authors. In my opinion, there is a lack of observational constraints to probe the timescales of deep ocean overturning (e.g. 14C). Thus it appears not surprising that the diffusivity parameter remains not well constrained.

There is also a lack of metrics to probe the spatial pattern of heat uptake. This is particularly important as the thermal expansion coefficient varies by almost an order of magnitude in the ocean. Thus it matters, where the heat is taken up to estimate sea level rise.

As another focus of the study is on TCR, it would also be nice to invoke additional metrics on thermocline ventilation as for example available by observation-derived fields of CFCs and bomb-produced 14C.

C3

4) Page 5 to page 7, results, The description of the difference in input forcing is useful, but in my opinion misplaced. Solar and ozone forcings are model drivers (or forcings) and distinct from a particular model version. These forcings should be described in the method section where the simulations and the applied external forcings are to be described.

5) P6, line 3ff; Q-flux adjustment: Does this mean that the authors apply temperature flux correction to their model? This should be explained in the method section.

6) Section 4: I miss a figure comparing the modelled pattern of the median (or mean or best-guess version) with the observed pattern of surface air temperature change and similar for the global ocean heat uptake and its spatial pattern (and may be for upper air temperature) to illustrate how well the model is able to capture the observations.

7) Page 12, line 7: How well does the polynomial fit represent the model results?

8) Page 12, line 14: Why is the PDF for the TCR not directly estimated from the 372-member ensemble? Does the fitting add additional uncertainties to the procedure of estimating TCR?

9) Discussion and conclusion: While the authors suggest that their approach should serve as a template for other groups, they fail to mention that similar, and sometime much more comprehensive approaches of parameter calibration, have been undertaken by other groups. They also fail to compare their estimate of TCR and ECS with published estimate and to put their findings in the context of the wider literature. See for example, Collins et al., IPCC, 2013 for the most recent assessment of TCR and ECS values by IPCC. Of course there are recent updates of these estimates and there are also many other studies that determine model parameters such as vertical ocean diffusivity. Examples that come immediately in my mind are Holden et al., *Clim. Dyn.*, 2010, Richardson; *Nat. Clim.Change*, 2016, Schmittner et al., *GBC*, 2009, Steinacher et al., *Science*, 2013 or Steinacher and Joos, *Biogeosciences* 2016. It is the task of the authors to identify the recent literature to provide a relevant discussion.

C4

P1, Line 22: typo: sensitivity

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-54>, 2018.