

## ***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.***

**Alex G. Libardoni et al.**

alex.libardoni11@gmail.com

Received and published: 2 July 2018

P1, L10: "absent an increase" is odd wording. P1, L12: This statement seems redundant to line 4-5 P1, L14-15: What causes these shifts? P1, L15-16: So if the land surface model has limited effect on temperature evolution, is it updates to the forcings that cause the differences in climate sensitivity estimates? It's not entirely clear what points the authors are trying to convey here. I suggest tightening up the abstract to highlight the significance.

Response: Per these comments, we have revised our abstract to make the summary clearer. We attribute the observed shifts in the parameter distributions to the changes

Printer-friendly version

Discussion paper



in model forcings. The land surface model impacts other components of MESM (i.e., carbon fluxes), but in the climate component used here, it has little impact.

P7, last paragraph: The authors raise interesting, but somewhat contradictory, points. They state that reducing the number of diagnostics from 3 to 2 has little impact on model parameter estimates, but then go on to state that CS estimates are lower when using 2 diagnostics. Why are the results insensitive to the upper-air diagnostic? Also, the constraint on  $K_v$  is not clear. Is there any update since what was shown their previous work (e.g. Libardoni and Forest 2011)? I suggest adding more details to these points to help the reader.

Response: We have cleared up these points by adding discussions into the manuscript. The main reason for omitting the upper-air diagnostic is the significant correlation between the upper-air temperature pattern and the surface temperature pattern as a result of the lapse rate and water vapor feedbacks. Each of these diagnostics reject similar regions of the parameter space for being inconsistent with the observed climate record, thus potentially double counting the same temperature response signal. Removing the upper-air diagnostic removes the risk of bias due to treating it as a statistically independent diagnostic.

There has not been any additional work on constraining  $K_v$  between our previous work and this manuscript. Currently, a second publication is in review (Libardoni et al., 2018, ASCMO) that investigates how including additional data in the model diagnostics improves the model parameter estimates. We show there that including additional data improves the model diagnostics and leads to better constraint on  $K_v$ . We chose not to incorporate any changes to the model diagnostics in this manuscript to provide a clean comparison of changes resulting from changing only the model version.

P8, L10: Can you show a plot of the ECS pdf for IGSM and MESM for comparison?

Response: For each of the distributions derived from the individual surface temperature datasets, we plotted the marginal PDFs for the full IGSM and MESM ensembles.

[Printer-friendly version](#)[Discussion paper](#)

In all five cases, the same changes are observed: higher climate sensitivity, nearly unchanged ocean diffusivity, and weaker negative aerosol forcing.

P9, L3-4: How do these new estimates of net aerosol forcing compare with other recent estimates?

Response: For EMICs like MESM, the net aerosol forcing is a model-specific parameter, making a clear comparison between studies and direct observations challenging. For example, the aerosol forcing pattern may account for different model forcings and be defined for different time periods. For example, while Andronova and Schlesinger (2001) scale the natural and anthropogenic aerosol direct and indirect forcings by adjusting the amplitude in 1990, the aerosol parameter in Knutti et al. (2002) is scaled in 2000 and represents the indirect aerosol effect and any other forcing not explicitly represented in the model. With these differences in mind, estimates of aerosol forcing from energy balance models and EMICs fall in the ranges  $-1.3$  to  $-0.54$   $\text{W/m}^2$  (Andronova and Schlesinger, 2001),  $-1.2$  to  $0$   $\text{W/m}^2$  (Knutti et al., 2002),  $-1.53$  to  $-0.33$   $\text{W/m}^2$  (Kriegler, 2005),  $-0.83$  to  $-0.19$   $\text{W/m}^2$  (Libardoni and Forest, 2011), and  $-1.7$  to  $-0.4$   $\text{W/m}^2$  (Skeie et al., 2014).

P10: L14: I'm a little unclear how ocean diffusivity fits in with the analysis. Why did the old ensemble cut off high values of  $K_v$ ? It is also relatively insensitive to the model updates compared to aerosol forcing and equilibrium climate sensitivity. Why is this? I recommend the authors streamline the results and discussion sections to include a summary of key points about each model parameter, the constraints and model sensitivities, and physical reasoning for the differences.

Response:  $K_v$  fits into the analysis because all three model parameters are estimated jointly, with the marginal PDFs calculated by integrating the joint PDF over the other two model parameters. Thus, changes due to the model and forcings can impact any of the three marginal distributions. As we point out in the edited manuscript, physical explanations for the changes in the ECS and aerosol distributions are more accessible

[Printer-friendly version](#)[Discussion paper](#)

than an explanation for  $K_v$ . However, because all three parameters are estimated together, changes in the other two parameters can impact our  $K_v$  estimates.

In both ensembles, values of  $K_v$  outside the of range of values sampled are assigned zero probability. This meant assigning zero probability for regions greater than  $5 \text{ cm/s}^{1/2}$  for the IGSM ensemble and  $8 \text{ cm/s}^{1/2}$  for the MESM ensemble. From the full MESM ensemble, we find non-zero, although small, probabilities of  $K_v$  between 5 and  $8 \text{ cm/s}^{1/2}$ . By accounting for the extra mass in the tail regions for the MESM ensembles, the  $K_v$  quantiles are pulled towards higher values.

We have added text to the manuscript explaining the points above. Further, we have re-ordered the results section to devote separate paragraphs for each parameter that explains the changes in the marginal distributions separately. This makes the key points stand out more clearly.

P12, L7: Why choose a third-order polynomial here? Is there sensitivity in the fits to the functional form? Would you expect similar results in terms of model differences using a 2nd order polynomial?

Response: The third-order polynomial was chosen for consistency with previous work to provide the most direct comparison possible between the surfaces derived for IGSM and MESM. In offline tests, we derived additional surfaces for first-, second-, and fourth-order polynomial fits and compared them to the TCR and SLR values calculated directly from the transient simulations. The first-order approximation leads to an unsatisfactory fit with gradients of TCR and SLR in the  $K_v$  direction that are too weak. The second-order fit produces curvature in TCR and SLR contours that are inconsistent with those calculated directly from the transient simulations. In particular, the 1.5 C contour for TCR using the second-order fit suggested that for a single  $K_v$  value, two different ECS values could be used. Further, the second-order fit shows that sea level rise greater than 14 cm is possible within the sampled domain, whereas none of the transient simulations had SRL that high. The third- and fourth-order fits both showed

[Printer-friendly version](#)[Discussion paper](#)

good agreement with the simulated results but were not without their flaws. The third-order fit showed some error in the 1.5 C TCR contour, where the fourth-order fit led to regions of SLR greater than 14 cm within the domain. We mention these tests in the revised manuscript and keep the third-order fit to maintain consistency with previous work. Improving this fit is a potential avenue of future research.

P12, L24-25: The authors state that the shift towards higher transient climate response is driven by higher climate sensitivity in MESM, but there is not enough explanation in my opinion as to why there is a larger CS in MESM compared to previous versions, how they compare (e.g. posterior distributions), and to what extent the updated forcings play a role.

Response: Through the points made above and the changes made to the manuscript, we believe that this has been addressed more clearly. We have added discussions for each parameter that specifies how the changes to the model forcings could lead to the observed shifts in the marginal distributions. Looking at the response surfaces, for any  $K_v$  value, an increase in ECS leads to larger TCR. Thus, given a constant  $K_v$  distribution, shifts towards higher ECS result in a shift towards higher TCR. With the relatively small changes in the  $K_v$  distribution from the subsampled MESM ensemble (see Table 2 of the manuscript), we find the assumption of constant  $K_v$  distribution needed for this argument justifiable.

P12-13: The conclusions provide a nice summary of the paper's key points. I suggest expanding the results section to include more in-depth discussion along these lines.

Response: As noted above, we have expanded the results section to provide more in-depth discussion of the reasons for the changes in the parameter and TCR estimates.

---

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

[Printer-friendly version](#)[Discussion paper](#)