Reply to Reviewer 3

The manuscript presents comprehensive results from the model calibration effort to improve DOE EAM_v1. It explores the impacts of the model recalibration on the fidelity of model climate, and implications for cloud feedbacks and ERF. The documented technical details in parameter calibration are highly beneficial for the scientific community to understand the model at the process level. The calibration comprises four major parts that reflect the latest advancement in modeling cloud, convection, aerosol, and radiation. The manuscript is co-authored by leading scientists in the field. Overall, the model development processes and associated impacts on major science questions are thoroughly discussed. The work is appropriate for GMD and I recommend acceptance after the authors address the comments listed below.

Reply: We thank the reviewer for the positive comment.

1) As each experiment runs for 11 years, some more statistical analyses can be conducted. For example, for Fig. 3, are the differences significant compared with the natural variability in the control run? For those difference maps, the authors may consider wiping out those pixels with insignificant differences.

Reply: We have added the statistical significance test in Fig 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20, and 21.

2) L139-141, the logic is unclear here. Why the smaller cloud feedback and aerosol radiative forcing can lead to a better surface temperature simulation in a couple run?

Reply: We have revised the text to clarify this point:

"…, the recalibrated atmosphere model, denoted as EAMv1P, shows lower cloud and precipitation sensitivity to aerosol perturbation and to surface warming. Because the notable biases in E3SMv1's simulated surface temperature evolution are due to a combination of high ECS (from cloud feedback) and strong aerosol forcing (Golaz et al., 2019), EAMv1P may lead to improvements to the simulation of the 20th century temperature evolution and a lower estimate of ECS when running as part of the fully coupled E3SM."

3) L343, should be "reduced conversion".

Reply: We have changed it to "insufficient conversion".

4) Fig. 4c&5c, tuning in MP apparently impacts LWP, but not cloud fraction in the Sc regions. Is it simply due to the diagnostic cloud fraction scheme which is independent with cloud microphysics? If yes, should such a disconnection be targeted in the future model development?

Reply: Indeed generally speaking cloud fraction and liquid water path are computed by different parts of an ESM, which can introduce inconsistencies. In E3SMv1, CLUBB is the macrophysics scheme that diagnoses the liquid cloud fraction and computes the condensation of water vapor and evaporation of liquid condensate which contributes to the changes in LWP. The microphysics scheme computes all the cloud microphysical processes which also affect LWP. The two parameterizations are coupled using 5-minute sub-time-step. The results show that LWP and cloud fraction changes are quite consistent with each other in marine subtropical regions: Figure A below shows that EAMv1_MP produces larger lower tropospheric LWP and larger low cloud fraction in the Sc regions.

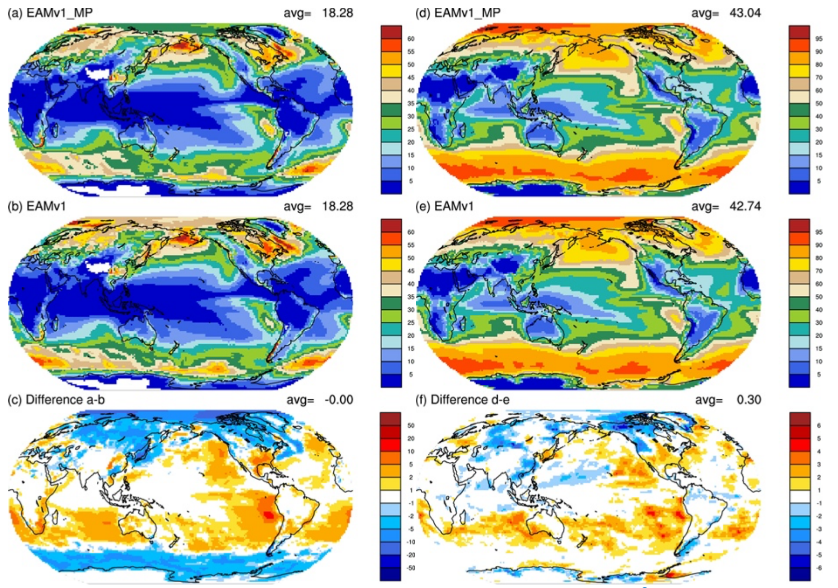

Figure A. Lower tropospheric (below 700 hPa) LWP (left; unit = g m^{-2}) and lower tropospheric cloud fraction (right; unit = %).

5) L527-534, please provide how EIS is calculated from the model.

Reply: EIS was introduced in (Wood and Bretherton, 2006). We computed the EIS following the CFMIP diagnostics code catalogue (Tsushima et al., 2017). We have added the reference in the manuscript.

6) Fig. 14c, it is a little surprising to see the altered cloud-rain autoconversion does not impact Eaci significantly over the subtropical warm cloud regions. Any explanation?

Reply: In Fig 14, EAMv1_MP does produce a weaker ERFaci. We have also attached the zonal mean plot below, which shows that ERFaci in EAMv1_MP is about 0.5 Wm^{-2} weaker than EAMv1 in the subtropics. However, we note that the EAMv1_MP includes 7 parameter changes as described in Table 3, and the autoconversion change is only one of them. Therefore, Fig 14c shows the effects of all 7 changes combined. We stated in Section 2.4 that "The effects and the mechanisms of each individual parameter adjustment require further investigation and will be documented in separate manuscripts."

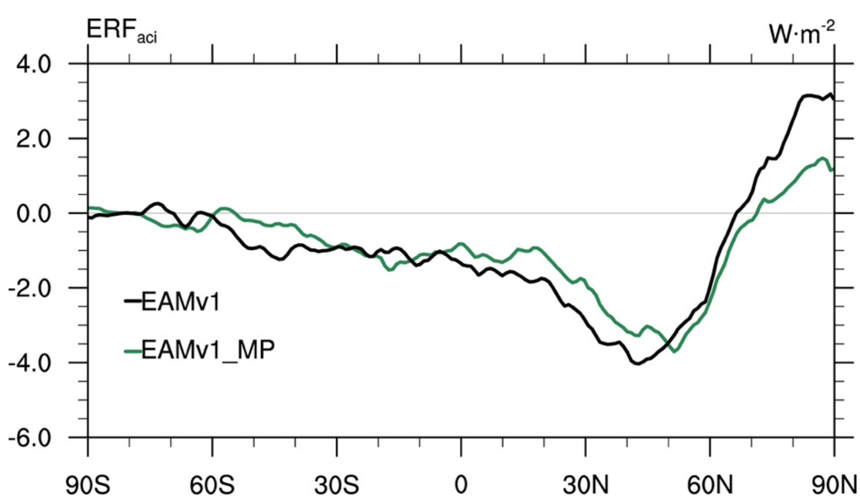

Figure B. Zonal mean of ERFaci (Wm^{-2}).

7) Fig. 15a, on about the same latitude, why aerosols induce opposite land temperature changes over the northeast Eurasia and the northwest North America?

Reply: This suggests that the surface temperature changes are not determined only by local energy balance. Other processes in the climate system such as large-scale circulation changes also play a role. We have added these sentences in the manuscript.

8) Fig. 14e and L963-965, it is unclear to me why EAMv1_ZM shows less sensitivity of Nc and Ni to aerosols, but produces stronger ERFaci?

Reply: By combining the present-day Nc and Ni in Table 6 and the relative change of Nc and Ni due to anthropogenic aerosols in Table 10, we can derive that EAMv1_ZM produces higher pre-industrial Nc and Ni than EAMv1. EAMv1_ZM also produces a larger Nc increase (4.79×10^{9} m^{-2}) due to anthropogenic aerosols than EAMv1 (4.58×10^{9} m^{-2}). These results are consistent with the stronger ERFaci produced by EAMv1_ZM. We have revised the text for clarification.

9) It may be beyond the scope of the study, but I am curious of the additivity of the impacts of different tuning parts. Would the total impacts from EAMv1P be a linear addition of those from each individual configuration? In other words, are there significant nonlinear interactions among those different configurations?

Reply: We stated in Section 4 that "the recalibrated model produces more improvements than the sum of the improvements from individual intermediate configuration, demonstrating the nonlinearity in the climate system and the necessity of combining all of the improvements that target different biases in different regimes."

10) Near the end of the paper, it is worth discussing what are the unresolved outstanding biases in EAMv1P and whether they are likely to be resolved in the next stage.

Reply: This paper attempts to carefully document the approach, strategy, and results of model calibration, including the strengths and weaknesses of the approach and results. Our tuning target is the climatology of bulk cloud properties such as CRE and cloud fraction and show improvements in many other aspects of the model. Further reducing the model biases by improving parameterizations, numerics, resolution, and calibration is an ongoing effort for the E3SM team. Incorporating process-oriented diagnostics in model development and calibration will be useful for ensuring that the model get the right answer for the right reason. We have added the discussion in Section 4.

Reference

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H. Y., Lin, W. Y., Lipscomb, W. H., Ma, P. L., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Eyre, J. E. J. R., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X. Y., Singh, B., Tang, J. Y., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H. L., Wang, S. L., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S. C., Yang, Y., Yoon, J. H., Zelinka, M. D., Zender, C. S., Zeng, X. B., Zhang, C. Z., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, J Adv Model Earth Sy, 11, 2089-2129, 10.1029/2018ms001603, 2019.

Tsushima, Y., Brient, F., Klein, S. A., Konsta, D., Nam, C. C., Qu, X., Williams, K. D., Sherwood, S. C., Suzuki, K., and Zelinka, M. D.: The Cloud Feedback Model Intercomparison Project (CFMIP) Diagnostic Codes Catalogue - metrics, diagnostics and methodologies to evaluate, understand and improve the representation of clouds and cloud feedbacks in climate models, Geosci Model Dev, 10, 4285-4305, 10.5194/gmd-10-4285-2017, 2017.

Wood, R., and Bretherton, C. S.: On the relationship between stratiform low cloud cover and lower-tropospheric stability, J Climate, 19, 6425-6432, Doi 10.1175/Jcli3988.1, 2006.