We would like to thank the reviewer for positive and helpful comments and suggestions. Our responses and actions are indicated in blue below:

Review

An extensible perturbed parameter ensemble (PPE) for Community Atmosphere Model version 6. Trude Eidhammer and Co-Authors.

This paper presents a clear description of the generation of a perturbed parameter ensemble (PPE) for a widely used global atmospheric model. Model sensitivity to key parameters is described. The use of the PPE to develop emulators characterizing the model's parametric dependencies is presented, and the emulators are applied to tuning. The paper is important in establishing new methods for model analysis and development, as well for its characterization of often poorly documented parametric dependencies. Only minor revisions are suggested.

Minor Revisions

1. The tuning results summarized in Figs. 9 and 10 point towards the possibility of using objective methods for model tuning going forward. Could the authors comment briefly on the possibilities of tuning towards a weighted function of a range of outputs, as would typically arise subjectively during a model development process? Have any experiments along these lines been undertaken?

It is in essence what is being done in this paper with the rejection sampling. Each output has its own defined permissible range, and so each is "weighted" differently. This is also what is done in Watson-Parris et al (2021, GMD, https://doi.org/10.5194/gmd-14-7659-2021).

2. In Eq. (1), as the Euclidean distance d characterizes each ensemble member m, would it be useful to denote as dm rather than simply d?

Thank you for this suggestion. We have renamed d as d_m. We also realized that there was a mistake in the description of the equation:

For each individual ensemble j, we calculate the average distance of each parameter i in ensemble j to parameter i in the other ensembles. Then d is the sum of all Euclidean distances in ensemble j divided by number of parameters (pa) and ensembles (en):

The sentence should instead be as follow:

For each individual ensemble m**,** we calculate the average distance of each parameter i in ensemble m to parameter i in the other ensembles, j. Then d_m is the sum of all Euclidean distances in ensemble m divided by number of parameters (pa) and ensembles (en):

3. ll. 139-142: Instead of as stated in the text, does clubb_C6thl refer to liquid water potential temperature flux, while clubb_C6rt is total water flux?

You are correct. We have corrected the description, and the text is now as follows:

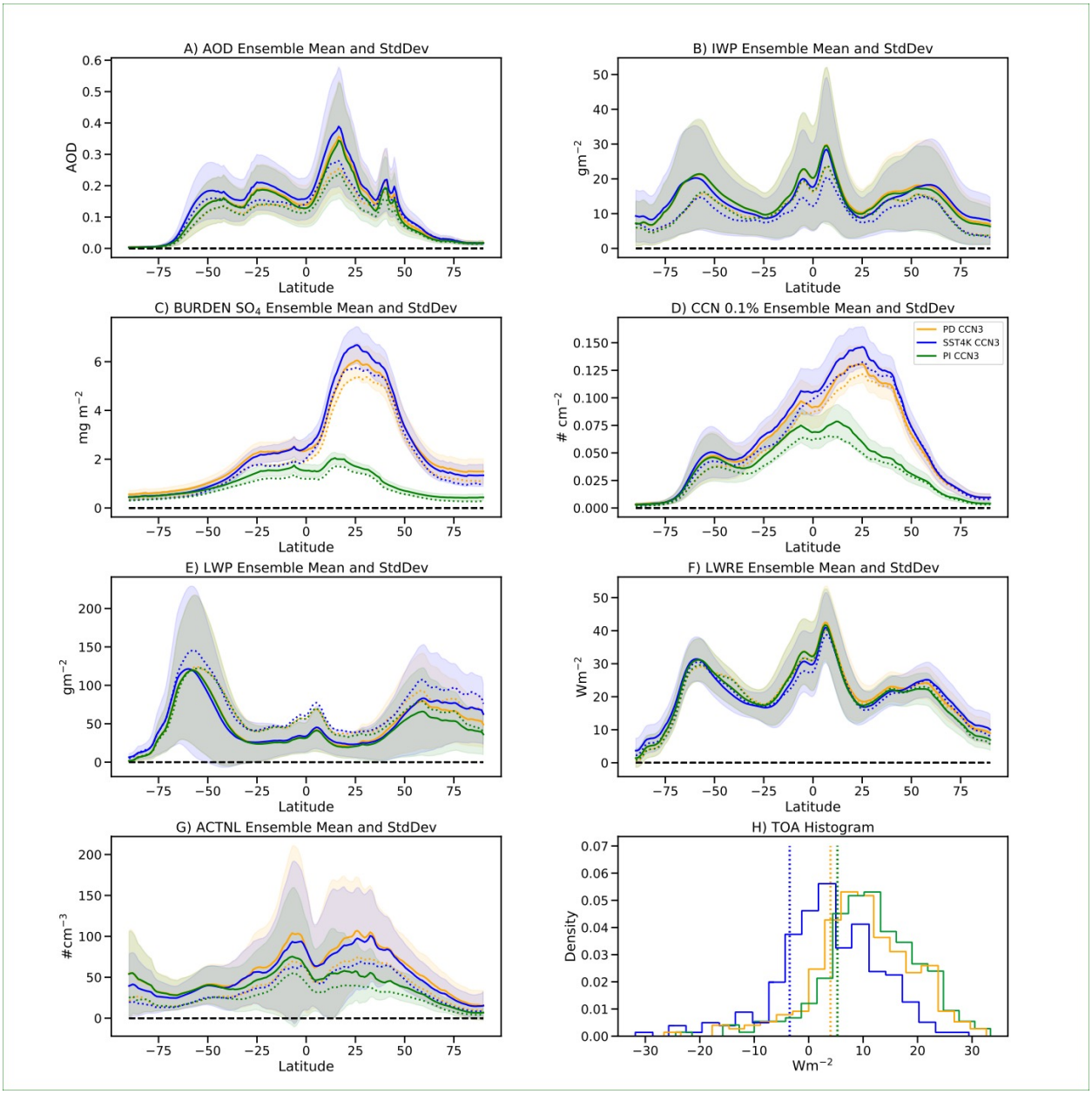
The parameters clubb_C6rt (clubb_C6thl) and clubb_C6rtb (clubb_C6thlb) are the low and high skewness of Newtonian damping of the total water flux (potential temperature flux). Decreasing these parameters tends to boost fluxes, producing a more well mixed layer, with minor effects on cloud brightness. The low skewness especially impact stratocumulus while the high skewness especially impacts cumulus. Similar parameters were perturbed simultaneously so clubb_C6rt=clubb_C6thl and clubb_C6rtb=clubb_C6htlb.

4. Longwave and shortwave cloud forcing are now more typically referred to as longwave and shortwave cloud radiative effective effect, to avoid confusion with forcing as a concept related to change in atmospheric composition.

We have changed all LWCF/SWCF to LWRE/SWRE

5. Fig. 3H: I was looking for 3 histograms, one each for PD, SST4K, and PI. But there seem to be more. Perhaps this is just a perception issue related to overlapping colors. Perhaps better to outline the histograms rather than shade for clarity?

We have changed the figure to illustrate the 3 histograms better by outlining them instead of using the shading as suggested. See new figure below



6. l. 359: Guo et al. (2015) used a GFDL model, not CAM, which could also explain the differing responses to CLUBB tuning. See also the discussion of CLUBB tuning for CAM in Appendix A of Bogenschutz et al. (2013).

The Guo et al (2015) paper we refer to is this one, which uses CAM:

Guo, Z., Wang, M., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., A. Bogenschutz, P., Gettelman, A., and Zhou, T.: Parametric behaviors of CLUBB in simulations of low clouds in the Community Atmosphere Model (CAM), Journal of Advances in Modeling Earth Systems,
7, 1005–1025, https://doi.org/https://doi.org/10.1002/2014MS000405, 2015

Perhaps the reviewer is thinking about the Guo et al. (2014) paper, which use the GFDL model:

Guo, H., J. C. Golaz, L. J. Donner, P. Ginoux, and R. S. Hemler (2014), Multivariate probability density functions with dynamics in the GFDL atmospheric general circulation model: Global tests, J. Clim., 27(5), 2087–2108.

Since the Guo et al. (2015) paper uses the CAM5, we keep the text as is.

l. 121: “aerosol” -> “aerosol effective” *We have rephrased the statement to:*

“By performing these three sets of simulations with the same parameter sets, not only can we evaluate the output spread by perturbing parameters, but we can also evaluate the cloud feedback (difference between PD and SST4) and aerosol forcing (difference between PD and PI). Here the aerosol forcing is the aerosol effective forcing after adjustments of atmospheric temperature and humidity.”

l. 123: “3 and 5” -> “3 and 5 yrs” Corrected.
l. 413: “relative” -> “relatively” Corrected.