**Review Report on TC-2020-178: "Statistical emulation of a perturbed basal melt ensemble of an ice sheet model to better quantify Antarctic sea level rise uncertainties"**

The manuscript proposes a new way to build a statistical surrogate for an expensive Antarctic ice sheet model which can quickly generate future SLR values given a basal melt rate anomaly trajectory, which represents the trajectory of sub-shelf ocean forcing. The proposed approach parameterizes the anomaly trajectories using sigmoid curves and builds a Gaussian process emulator for the relationship between the parameters of the sigmoid curves and the SLR values in the target year. This paper addresses the problem of accounting for ocean forcing uncertainties in future Antarctic ice sheet projections, one of the long-standing issues in the ice model community, and hence has scientific merit that warrants publication in *the Cryosphere*. However, I think the following major and minor points that I list below need to be addressed or at least seriously discussed in the manuscript before being considered for publication.

<u>Major Points</u>

1. I am mainly concerned about how realistic the smoothed the basal melt rate anomaly trajectories are and, if not, how the unrealistic (perhaps oversmoothed) forcing trajectories affect the realism of the final SLR projections. For example, while the authors argue that the basal melt rate anomaly trajectories from Timmermann and Hellmer (2013) and Cornford et al. (2015) in Figure 5 are accurately captured by the sigmoid functions, I can see that a lot of mid-range temporal patterns are smoothed out. For example, in the second (Ross Island) panel there are some clear discrepancies between the fitted sigmoid curves and the original trajectories and the fitted sigmoid curves clearly underestimate the basal melt rate anomalies in the end. Will the overly smoothed trajectories lead to vastly different SLR distributions compared to the unsmoothed trajectories? My worry is that using smoothed forcing trajectories may result in notably smaller SLR variations (as the resulting simulated SLR trajectories might be also overly smoothed) than the variations that would have been obtained without smoothing the basal melting rate anomalies. One easy way to check if this is the case is to obtain a few ice sheet model runs using the original basal melting rate anomalies from Timmermann and Hellmer (2013) and Cornford et al. (2015) and see how the final results differ from the runs based on the smoothed trajectories.

2. If the smoothing indeed leads to underestimation of the SLR uncertainties, one way to solve the issue might be to add some additional noise generated from temporally dependent processes such as the ARMA model to the simulated SLR trajectories. The parameters for the ARMA model might be estimated by comparing the SLR projections generated based on the original basal melting rate anomalies and those generated based on the corresponding sigmoid curves.

3. In Lines 273-275, the authors mention that 'least-squares optimization' is done to find the best fit. However I cannot find what variables are actually used in the 'least-squares optimization' here. Are they the simulated SLR trajectories and some observational data? Or are they the fitted sigmoid curves and the original basal melting rate anomalies? Judging based on the caption in Figure 6, I think it is the latter. Then I think the issue can be easily solved by expanding the plausible ranges and also running more ice sheet model runs and obtaining more emulated runs accordingly so that the envelop of the colored curves shown in Figure 6 well-contain the black curves. I am not sure why the authors are relying on some ad-hoc procedure to fix the issue instead of expanding the plausible ranges.

## Minor Points

1. Related to the major point #1 above, there is an existing method to emulate the future projections for different forcing scenarios (Catruccio et al. 2014). I think it will be ideal to compare the proposed method with this approach, but it might require too much effort to repurpose this method for ice sheet projection. I will leave the decision to the authors, but I think it is at least worth mentioning this approach as a possible future direction.

2. The authors use Kennedy and O'Hagan (2001) as the main reference for Gaussian process-based emulation, but that idea should be attributed to Sacks et al. (1989). In fact the main contribution of Kennedy and O'Hagan (2001) is more on the calibration side rather than the emulation side.

3. Related to the major point #3 above, having estimated parameter values that are at or outside of the plausible parameter ranges for a model ensemble is a well-known issue in computer model calibration literature (see. e.g., Brynjarsdóttir and O'Hagan,2014, Chang et al., 2016, Salter et al., 2019). In fact, this is a typical example of a 'terminal case' mentioned in Salter et al. (2019).

## References


Chang, W., Haran, M., Applegate, P.J., and Pollard, D. (2016) Improving ice sheet model calibration using paleoclimate and modern data, *the Annals of Applied Statistics*, 10 (4), 2274-2302

Jenný Brynjarsdóttir and Anthony O'Hagan 2014 *Inverse Problems* 30 114007

Sacks, Jerome; Welch, William J.; Mitchell, Toby J.; Wynn, Henry P. Design and Analysis of Computer Experiments*. Statistical Science* 4 (1989), no. 4, 409-423. doi:10.1214/ss/1177012413. https://projecteuclid.org/euclid.ss/1177012413

Castruccio, Stefano, David J. McInerney, Michael L. Stein, Feifei Liu Crouch, Robert L. Jacob, and Elisabeth J. Moyer. "Statistical emulation of climate model projections based on precomputed GCM runs." *Journal of Climate* 27, no. 5 (2014): 1829-1844. (website: http://www.rdcep.org/research-projects/climate-emulator, source code: https://github.com/RDCEP/climate_emulator)

Salter, M. J., Daniel B. Williamson, John Scinocca & Viatcheslav Kharin (2019) Uncertainty Quantification for Computer Models With Spatial Output Using Calibration-Optimal Bases, *Journal of the American Statistical Association*, 114:528, 1800-1814