Reply to Editor comments

General comments:

Editor:

"*Thank you for your revisions. The language changes look better and Figure 18 seems much improved. However, I still think it is important for you to make a closer comparison with the results of Yang et al. (2022), who use the same model at the same high resolution, but grow a much smaller LGM ice sheet. Since you want to make the case that very high resolution simulations are valuable, this point should not be overlooked.*

Best regards,
Alex"

Authors:

Yang et al. (2022): *"Previous GrIS simulations constrained by observations suggest that during the last deglaciation, the GrIS may hold 4.6 m to 5.1 m SLE more ice relative to present-day [16, 17]. Our simulated magnitudes show discrepancies with the previous studies, likely because the simulations* ***have not been constrained by any observation****. In addition, a simplified glacial index method could also contribute to such a discrepancy. We realize that the simulated ice volume of the GrIS varies depending on the choice of model parameters and also the prescribed climate forcing."*

The authors of that study are aware and acknowledge that their LGM volume estimates are most likely too low and unrealistic, and that higher estimates by Simpson et al. (2009) and Lecavalier et al. (2014) are most likely more realistic. As they mention themselves, this relates to the fact that the simulations by Yang et al (2022) are not constrained by observations, but only ran in the paleo using a glacial-index scheme for paleoclimate forcing.

Regardless of the model resolution used, if the parameterizations are not widely tested to maximize fit with available empirical constraints, the results could easily be heavily biased, especially when only conducting one-at-a-time parameter perturbations on 6 sensitivity simulations, as is done in that study.

High model resolutions will help produce more realistic GrIS dynamics (Aschwanden et al., 2016), and in turn will increase chances of reaching good model-data fit. However, biased forcings and parameterizations of internal dynamics and mass-exchange processes necessarily dominate, and can produce all kinds of results. For instance, our own paleo ensemble (5 km, PISM, 100 simulations, latin-hypercube sampling) features plenty of simulations with GrIS lLGM volume anomalies between 0 and 4 m SLE, with some matching numbers from Yang et al. (2022) (see Fig. 10). Well, these agree less well with empirical data on known GrIS LGM extent than our lLGM best-fit simulations: which suggest 6 – 7.5 m SLE. That is the result of model-data comparison.

Instead, the study by Yang et al. (2022) has a different focus: running transient paleo-to-future simulations compared against future runs starting from equilibrium spinups. Their goal was thus not to calibrate paleo simulations against paleo data.

A more appropriate study to compare our results against is the model reconstruction by Lecavalier et al. (2014), one of the few paleo GrIS modelling studies who ran an ensemble of GrIS wide simulations and calibrated their model and parameterizations through quantitative comparisons against widespread empirical constraints. Comparison between their and our results is made in section *"GrIS volume and thickness during the lLGM",* where we acknowledge that our higher numbers, relative to theirs, could be correct or not, and require more future testing and investigation.

Consequently, we do not think the suggested comparison with Yang et al. (2022) is appropriate, as would most likely agree the authors of that study.