I have a short comment about this manuscript, but also the broader review process and tone of reviewer comments.

To start, the reviews for this manuscript reveal broader philosophical questions about the role of heuristic parameterizations vs more physically based approaches to simulate ice sheets. All of these questions are important topics of discussion, but I am increasingly concerned about the tone of reviews that veer away from critical discussions to attacks on authors. To be fair to reviewers, tone is hard to gauge in the printed word. This is especially true when considering the reality that reviewers may not be native English speakers and idioms and cultural norms do not always translate well to the page, the pandemic has been so challenging, funding is always an issue. All of this means we have less time and patience to engage meaningfully with each other in science. Keeping these caveats in mind, some of the reviewer comments read more as an attack on the authors or rehashing old debates than an uplifting engagement with the scientific content of the manuscript.

This is especially damaging for the Cryosphere because the preprint stage allows the scientific community to engage in public discussions about science during the review process. The discussion for Cryosphere papers is public, permanent and announces to early career scientists the type of community that we are. Nobody likes to be blasted by anonymous, unsigned reviews and these experiences can sour early career folks on the community when what we to provide is useful (if painful) feedback. I would much rather be part of a community that works together to share ideas that improve our science instead of pitting ourselves against each other in rhetorical combat. Whether intentional or not, the tone and professionalism of reviews is something that the editors and we as a community need to pay close attention to.
Coming back to the manuscript and the debate that has enraged at least one reviewer, one of the bigger concerns surrounding this manuscript is an ever present debate about the validity of heuristic parameterizations versus more “physics” based representations on processes. This debate goes back (at least) to the origin of weather and climate models with Edward Lorenz noting back in the 70s that weather models did an OK job just based on the fluid mechanics, but models did a better job when other processes that were not well resolved or even understood were included as “parameterizations”. Echo’s of this can be found throughout climate modeling history, including the great “flux correction” controversies of coupled atmosphere-ocean models. In glaciology the difference between an unphysical parameterization and physical approach is often time and history. Our field has a history of introducing working hypotheses based on little more than gut feeling and sheer force of will that eventually become enshrined in our literature because they either work or nobody managed to come up with anything better. The broader question swirling around the undercurrent is always whether it is best to simply ignore processes because they are hard to parameterize or use a crude parameterization that we know needs improvement just to see how important a process might be. And I think that the later is what this study is attempting. The authors are essentially whacking their model hard with a hammer to see what kind of response they get. I don’t see any problem with that approach as a motivator for whether more in depth studies are needed or not.

A theme that has been helpful for many of these past debates about physical processes, rests in asking what observations can be used to test the model or better constrain the parameters? I personally think that this link back to existing or needed observations is a useful question to expound upon in the manuscript. How do we know the model (or certain parameter combinations) are capturing some of the relevant physics. Are there specific predictions that the model can make that can be tested to (in)validate any of the model hypotheses?

Another important issue is one of numerical convergence. My view has long been that modelers (I include myself in this) need to more systematically demonstrate numerical convergence of models, but we often forget or assume that because it was done for a previous study that it doesn’t need to be repeated. I do think that a numerical convergence study, even if it is for a shorter period of time or limited portion of the domain would better allow the authors to demonstrate how robust the results are to numerical parameters.

Overall, I would like to plead with the reviewers, editors, authors and broader community to try to elevate our scientific criticisms to focus on concrete suggestions for how we best move forward to improve the science.