Comments to "Historically consistent mass loss projections of the Greenland ice sheet" by Rahlves et al.

1 General comments

In this paper, the authors used the ice-sheet model CISM for model initialisation, historical run and sea level projections. Simulations are designed to explore and quantify the impact of processes such as SMB-elevation feedback, initialisation strategies, and forcing strategies. Specifically, the authors conducted initialisation simulations with not only reanalysis SMB (ERA5), but also Earth System Model outputs to produce a suite of basal sliding coefficient products. Once reaching (quasi) steady state, a 1000 year relaxation simulation is done for every model setup. In the projetions, ERA5 initialised model are applied SMB anomalies while ESMs initialised model are applied direct ESMs SMB outputs. Furthermore, the authors explored the impact of outlet glacier frontal retreat. I think the simulations are very interesting and compensated some questions that were not possible to explore in the model intercomparison simulations ISMIP6.

However, I have some concerns about the manuscript: (1) I find some conclusions not evidently supported by the simulations. For example, 'While discharge from outlet-glaciers remains a substantial factor, the future evolution of the ice sheet is governed by mass loss due to changes in SMB'. First, this conclusion might be forcing dependent, i.e. if you use another ESM output as SMB forcing, does this conclusion stand? Second, the authors prescribed ice front retreat in the simulations to compare, although, ice discharge is not the same with mass loss due to frontal retreat (calving), and frontal retreat is only one of many factors that can cause changes in ice discharge (others are like basal sliding, dynamical thinning). (2) Methods description is quite general. For example, what does medium sensitivity mean and how is it defined? For the mechanisms that the authors are try to explore, such as smb-elevation feedback, it would be helpful to present the equations. (3) The writing is easy to follow, but the logics are not always clear. For example, after reading the introduction, I don't see how these pieces of work connected with each other to form one piece of story. There are other mechanics not explored in ISMIP6 (e.g. sliding laws, initialisation methods) if the target is to compensate ISMIP6. Here, I by no means suggesting more experiments, but organising and streamline the structure. (4) Some conclusions are presented without evidence. For example, the mesh convergence study was conducted and suggested 16km resolution is enough, but no figure to present this. I'm happy that the authors have done mesh convergence study to ensure a proper numerical setup, but I'd like to see these results presented (maybe in supplementary).

I gave more detailed comments in the following section. I suggest major revision of the manuscript.

2 Specific comments

Topic: remove the period dot

Abstract: Quite lengthy with too much details?

Line27: Please check the citations. I don't think Shepherd et al., 2012 suggested the contributions from SMB and ice dynamics separately. This recent article is more appropriate to be cited here: The IMBIE Team. Mass balance of the Greenland Ice Sheet from 1992 to 2018. Nature 579, 233–239 (2020). https://doi.org/10.1038/s41586-019-1855-2

Line49-66: The description of different initialisation methods is a bit confusing. The authors used 'consistency' and 'self-consistent' to describe the advantage of paleo spin-up approach. In my opinion, as long as the model solve the equations correctly, the output variables should be consistent. Does 'consistency' means 'an equilibrium state'? Or 'consistent' with the paleo climate? It would be helpful to explain the issues (and reasons causing them) explicitly.

Line74-79: 'sample uncertainties'→'quantify uncertainties'?

Line74-79: This paragraph summarized the study, although not streamlined. Initialisation methods, SMB, ice dynamics of outlet glaciers are mentioned. However, the uncertainty study is then on climate forcing and 'modeling choices'. The connections between the pieces marked by '.... Additionally, . We.... Moreover...' are not obvious. This might be improved by a better narrative.

Line91: Is 4km resolution enough for the simulations? Did the authors conduct a mesh convergence study to ensure the model behaviour is not restricted by the choice of grid size? I refer the authors to this article: Cornford SL, Martin DF, Lee V, Payne AJ, Ng EG. Adaptive mesh refinement versus subgrid friction interpolation in simulations of Antarctic ice dynamics. Annals of Glaciology. 2016;57(73):1-9. doi:10.1017/aog.2016.13

Line96-97: How is ocean forcing set at the ice-ocean boundary?

Line97-98: 'All floating ice is assumed to calve immediately'. Is removing all floating ice a reasonable setup for Greenland projections? How much of a difference would it cause? Could you provide some evidence (citations) to justify this approach?

Line103-105: Can you explain in detail the SMB-elevation feedback with equations and/or add the relevant references?

Line107 (section 2.2): Is the runoff depth averaged or depth integrated? The sentence suggest location of ice front position is prescribed as a function of maximum ice front position, far field ocean temperature and runoff. How is maximum ice front defined? How do you adjust the ice front positions (e.g. do you remove or calve ice out of the prescribed boundary)? Again, adding some equations will help. It will also help when you later on explain uncertainty caused by the choice of parameters, i.e. what are the values of parameters? What are the ranges of these parameter values in the sensitivity simulations? What are the calibrating variables (Line 112 mentioned 'calibrated values', but 'calibrated values' of what variable? Sea level contribution, mass loss, or other index?)? Equations and tables may be helpful.

Line161: delete the first 'approach'.

Line166: Could the authors add a figure with total SMB applied to the ice sheet for the two ensembles? The discription of this paragraph is a bit confusing. Did you use absolute values of SMB and ST from the ESMs? Why is it necessary to divide the reference part and anomaly part?

Line176: unit is missing

Line177: What are the four forcing modes? I believe this is not mentioned before. Also, I cannot quite put the numbers together 10(ESMs)*3(SSPs)*3(sensitivity params)*3(grids)*2(init methods)=540 simulations. How are 304 and 192 calculated? What does 'full forcing' mean? I think here a table would be helpful.

Line184: 'ACCESS1.3' was not explained before this location.

Line204: To my understanding, the two experiments here are using two different initial geometries: one with ESM spinup, the other with ERA spinup, same basal friction product (ERA spinup), and same SMB forcing (ESM) for the 1000 year relaxation. In this case, the first sentence of this paragraph is a bit misleading (second half). Furthermore, the authors concluded that 'the initial geometry is not decisive for the resulting ice sheet geometry, and it's the mismatch between SMB and friction that leas to a deviation from the ERA5-initialization', which I think is questionable. Here is my argument: When we look at Figure 3a and b, they shows a similar pattern of surface elevation difference comparing to ERA5 initialisation. The 1000 year relaxation of ERA5 used geometry from ERA5 spinup, basal friction from ERA5 spinup, **SMB from ERA5**; the two experiments here used geometries from ERA5 and ESM spinups (slightly different as the authors said), basal friction from ERA5 spinup, **SMB from ESM**. The differences can't be dominated by geometries, because the one using ERA5 and shows the same pattern with the other. The only explanation would be different SMB products caused the geometry differences in this 1000 year relaxation. Maybe I have misunderstood the setup of experiments, in which case, the authors are welcome to correct me, and maybe explain a bit more in the text.

Line223: Fig. 4b→ Fig. 4

Line281-285: I find describing parameters as low, medium and high hard to understand in depth. Could you give the numbers, maybe as well as the calving front locations when applying different parameters?

Line300: Usually we cite figures in order, so the citation of Fig. 10 would normally be before the citation of Fig. 11.

Line307: Period dot missing. How does SMB-induced upstream counteracts the effective removal of ice via discharge?

Paragraph from Line310: I guess the proportion numbers will change according to the applied SMB products of ESMs?

Line319: How did you calculate discharge from outlet-glaciers? How are the proportions in Figure 10 calculated in the simulations with different SMB forcings?

Line326: If you have done simulations with other ESMs and have the relative numbers, could you present them in the main text or supplementary? For example, the MPI-ESM1-2-HR run results mentioned here are not presented (in Figures/tables).

Line360: I think the number should be 40mm rather than 80mm according to the reference. However, Goelzer et al., substract control simulations from projections, while this study used projection results directly, so I'm not sure how comparable are these two numbers.

Paragraph of Line385: The arguments of this paragraph is hard to get.

(1) This sentence 'control experiments only offer....rather than providing conclusive information about its future response to historical forcing.' seems to indicate historical period is not reconstructed confidently, while the last paragraph conclude the historical run matches well with observations.

(2) The argument on committed sea-level contribution in historical time is also confusing. To estimate committed sea-level contribution by 2015, it

was suggested a constant value per year is necessary; then after 'However' suggest the annual mean constant is hard to get due to interannual variability; in the end proposed a solution is to use moving mean over time. Why can't we apply SMB with interannual variability to model simulation? If the model need yearly SMB, can't we simply remove the interannual variability by annual mean?

Paragraph of Line415: Differences caused by mesh resolution has been mentioned, and 16km is suggested to be enough. I suggest to present the mesh convergence results in the supplementary as evidence.

Line 416-420: Are this argument relevant or supported by this study? The argument about parameterization seems to be contradict with the argument that 16km mesh grid is enough?

The last paragraph is also not very relevant to this research.