Review Felikson et al. (2022)

This study aims to constrain and improve projections of the Greenland ice sheet evolution by making use of changes in three different observed quantities (mass changes, velocity changes and dynamic thickness changes) that are a measure the Greenland ice sheet dynamics. So far, Bayesian calibration of Greenland ice sheet model projections have been based on a single observation type.

I think the study presents a very useful framework to improve ice sheet model projections based on observations. However, as also stated in the study, using three different observation types do not give a clear answer on what is the best observation type to constrain the ice sheet model behavior and there might be a better metric when different observation types are combined. Therefore I am wondering whether the study as presented in the present form is convincing enough that model projections have been improved by using Bayesian calibration using different observation types. In addition, I would like to propose some minor clarifications on the methodological choices that you can read below.

Main comments

You explain that you coarsen the resolution in order to remove the spatial correlation between the different observations. However, if you look at Figure 1, there is still a strong correlation between the observed velocity change and the dynamic thickness change, especially for the fast outlet glaciers. It looks like it is a workaround to justify the neglection of a covariance matrix. I am also wondering if you are not losing too much important, detailed information by coarsening the resolution only to justify that the different variables are not correlated. Especially velocity changes at 50 km resolution remove a lot of detail. Did you investigate the posterior probability distribution for higher spatial resolutions?

The model is initialized to the year 2007 and observations are used until 2015. What is the influence of the length of the observation time on the results? Could it be useful to use a longer observational time series, for instance by initializing the model to the year 2000 to increase the observational length?

You refer in the methodology to the Nias et al. paper in review, but there is no preprint available (or no reference in the reference list) so it is impossible to acquire the right information. For instance, the model ensemble details can be found in Nias et al. (in review). It would be interesting to have them in this manuscript as well. Also, you mention on L153 that the modeled mass change is aggregated within the same drainage basin as used for aggregating the observed mass change. Can you explain what that means, why and how that is done?

The goal is to narrow the committed Greenland ice sheet change projections. However, you are changing not only model parameters but also the SMB, which you define as a forcing. Can you still speak of committed sea-level change if you change the SMB forcing?

It would be nice to add some figures about the modelled Greenland ice sheet state. You show only 4 figures that basically focus only on the posterior probability distribution of committed GMSL rise due to melting of the Greenland ice sheet.

Minor comments

L61: It could be useful to give the values for the basal friction factor, the ice viscosity offset and the SMB offset for the 137 forward runs in the supplementary information. Could you add a justification for the offsets chosen?

L151: The modeled quantities are regridded to 50 km x50 km. What is the original resolution of the model results?

L173-L175: The model uncertainty is a linear function of the observational uncertainty and you assign different multiplication factors for the different calibration to match the peak in the posterior distribution. Could you elaborate a bit more on the choice of these values? Because it looks contradictory to what you say on L185, that the median and maximum a posteriori GMSL are far apart.

L207: That sounds like a logical consequence of the larger changes along the margin of the ice sheet.

L252: You discuss the firn thickness change as a potential bias. What is the modelled firn thickness change? Could you show a figure? Also you report a potential bias of 10 cm per year in the ice sheet interior. What is this number based on?

Typo's

L124: and and

L274: Open questions?

L284: correct 'to on the choice'

Figures

Figure 1: Something went wrong with the labels: please add/modify the labels (a), (b), (c), (d), (e) and (f). Please also adapt the colormap labels to make them uniform with the indication of the variable and the units between brackets. I would also add the resolution for the different gridded observations.

Figure 3: I do not see the blue curve with the Gaussian approximation.