Klose et al. investigates the millennial-scale commitment of the Antarctic Ice Sheet to global sea level from 21st century emissions scenarios using an ensemble of two continental-scale ice sheet models. Their results demonstrate that a multi-meter sea-level commitment of a low emissions scenario (SSP1-2.6) cannot be ruled out over millennial timescales, highlighting the difference between what they define as the transient “realised” sea level contribution from the long-term “committed” sea level contribution. Under high emissions (SSP5-8.5), the sea level contribution is as high as 40 m over 7000 years, with significant loss of the EAIS.

The study is a comprehensive and well-written and has potential to be a useful contribution to *The Cryosphere*. One of the main points of novelty of this study is that the millennial-scale projections are performed with two different ice sheet models in a consistent manner. However, the approaches used to initialise these ice sheet models are very different, resulting in quite substantial bias in one of the models. It does not seem clear to me that the authors have considered to what extent the inter-model uncertainties they describe are due to these large differences in the ice sheet initial state. I suggest that the authors address this aspect in particular to improve the work. This review is divided into general comments and specific comments.

General comments

Two different approaches are used to initialise the ice sheet models. For PISM, an ensemble of spin-up simulations were run and the model output were scored based on the fit of modelled floating and grounded ice area, ice thickness, and surface velocity to Bedmap2 and Rignot et al. (2011) velocities. Fig S1 indicates that the initial state of PISM for the projection is hundreds of meters too thick for most of WAIS, and hundreds of meters too thin for most of EAIS. For Kori-ULB, a nudging procedure is implemented to minimise model drift from Bedmachine-Antarctica. The authors should elaborate on why these two different methods are employed as well as the potential impacts on their results.

Differences in the input datasets, such as bed topography, could alone account for some of the ice sheet model differences (e.g. Wernecke et al., 2022), but this is generally not discussed other than with regard to present-day climatologies.

The results section jumps straight into the projection experiments, but I think it is worth commenting on the historical simulations. Notably, the two models show differences in the direction of SLE change over the course of the historical run, with one of the models showing a basal melt rate of nearly double the other (Fig S2). This is important context for the transient ice-sheet response.

As a specific example of the impact of the initial state, for SSP1-2.6, the two different ice sheet models display different short-term and long-term behaviour. In general, Kori-ULB simulations show a higher centennial-scale sea level contribution than the PISM simulations, but this reverses for many of the simulations by the year 3000 (judging from Fig 2c). To what extent is this slower but eventually larger response of PISM due to its initial bias in ice thickness? One of the key findings highlighted in the abstract is this large sea level contribution under SSP1-2.6 of up to 6 m, but would a PISM model with a thinner initial WAIS produce the same result? This is worth exploring with a few sensitivity experiments.

Specific comments

Could you provide a table of experiments run, both in terms of forcings and ice sheet model parameter values?

Line 180: capitalise "Initialisation"

Line 184-185: NorESM1-M is CMIP5? "ocean and atmosphere anomalies" refers to anomalies of this particular GCM?

Fig 1a: Should the top part of the curve have a steeper slope? (i.e. accelerated mass loss)

Line 192: By atmospheric climatologies, do you mean from the RCMs or the CMIP forcing?

Line 204: So scoring for Bedmap2 for PISM, but Bedmachine for Kori-ULB?

Line 214: "balanced"

Line 230: "GCMs"

Line 240-243: Do you average at or over a particular depth?

Line 270: Is the reason for the difference in parameter values that they produce better fit to observations for those particular ice sheet models?

Line 278: Is the negative SLE change from a model that has positive SLE change over the historical period? It is worth specifying if the models that show modern ice mass loss have a negative or positive SL contribution.

Line 280: I would think the difference in initial state is a larger contributor to the differences between the two models than the dynamic response to the forcings.

Line 284-286: PISM is too thick in the ASE to start with.

Line 297-300: The models don't include hydrofracture parameterization, correct? Does surface temperature of ice shelves reach threshold for melt pond formation? e.g. van Wessem et al. (2023)

Line 303: "(ISMIP6)"

Line 315: But you have some simulations that show the opposite

Line 339-342: This sentence is confusing.

Figure 2: I suggest changing the colors because SSP1-2.6 is generally dark blue, and SSP5-8.5 is generally dark red, but here the colors refer to different GCMs.

Fig 3: Clarify the time of commitment. Are these means of both models?

Line 391: "changes"

Line 399: By regional warming, do you mean of air temperature?

Line 406: Is this realistic that the Ross Ice Shelf doesn't collapse under sustained SSP5 forcing? It could be an artifact of the basal melt parameterization, or the fact that there is no hydrofracture parameterization implemented.

Line 409: For how much warming? And is this due to the initial condition (i.e. PISM has more of WAIS to lose…)?

Line 413: What do you mean by "uncertainty"?

Line 477: The projection experiments are consistent, but the initialisations are not.

Line 567-593: Initial state discussion focuses on the choice of present atmospheric forcing, but what about the initialisation procedure, which seems to result in quite different ice thicknesses?

I suggest that you reduce the number of supplemental figures by consolidating Fig S4 to Fig S19.

References

Wernecke, A., Edwards, T. L., Holden, P. B., Edwards, N. R., & Cornford, S. L. (2022). Quantifying the impact of bedrock topography uncertainty in Pine Island Glacier projections for this century. *Geophysical Research Letters*, *49*(6), e2021GL096589.

van Wessem, J. M., van den Broeke, M. R., Wouters, B., & Lhermitte, S. (2023). Variable temperature thresholds of melt pond formation on Antarctic ice shelves. *Nature Climate Change*, 13(2), 161-166.