Reviewer 2 (Jason Amundson)

General comment

[RC2.01] This study investigates how mass balance data used for calibrating global glacier model projections affects model outcomes. Until recently, global glacier projections relied on either limited mass balance data from a small subset of glaciers or regional mass balance observations. Recent work has produced a data set of geodetic mass balances for all glaciers worldwide, thus enabling glacier-specific model calibration. In this study the authors demonstrate that the glacier-specific and regional mass balance tuning procedures produce similar glacier projections at the regional scale, but that they (can) deviate significantly at the scale of individual glaciers. The agreement at the regional scale provides some confidence in sea level rise projections; the disagreement at small scales motivates further work aimed at refining global glacier models in order to better understand evolving water resources and natural hazards, which are more local in nature. In addition, the paper is also one of the first to use the CMIP6 scenarios, and therefore complements a recent paper by Rounce et al. This work is somewhat tangential to my main research interests. With that in mind, I found it a little difficult to wrap my head around the various global glacier models that are under development and considered in this paper (i.e., GloGEM, OGGM, and PyGEM). I think the paper could benefit from a clearer description of how the models differ from each other. The models are described in the text, but some of the description just points to previous publications. I'm not sure that lengthier descriptions would be necessary. Perhaps a table that lays out how the models handle mass balance and ice dynamics, while also including some of the strengths and weaknesses of each? Other than that I felt that the paper was fairly easy to read and will make an important contribution to global glacier modeling.
We thank the reviewer for his positive review and are pleased to read that the manuscript was also easy to follow for a scientists that is not directly involved in the field of (large-scale) glacier evolution modelling. To allow for a better grasp of what the different large-scale glacier models do, what they have in common, and what is different, we will follow the suggestion by the reviewer and add a supplementary table (Table S2) that provides an overview.

Specific comments

[RC2.02] L 33–36: I assume that these results come from using the glacier-specific mass balance observations? It is a little ambiguous because the previous sentences discuss regional vs. glacier-specific observations.
This is a good point and was indeed not very clear in the original abstract. We will now update the abstract by explicitly including the information that all these results are based on the mass balance calibration with glacier-specific observations: *We project the 2015-2100 global glacier loss to vary between 25±15% (GloGEM) and 29±14% (OGGM) under SSP1-2.6 to 46±26% and 54±29% under SSP5-8.5 (ensemble median, with 95% confidence interval; **calibration with glacier-specific observations**)*

[RC2.03] L 280: Should this be "21st century"?
Yes, this should indeed have been '21st century'. We will update this to: *…but here the calibration approach has an important effect on the 21st **century** transient evolution towards this deglaciation:…*

[RC2.04] L 288: What is meant by "similar state independent"? It seems that something is missing from this sentence.
The sentence was unclear due to a missing comma and will be updated to: *…since a lot of the regions lose a large part of their mass, evolving to a similar (almost ice-*

free) state, independent of the calibration methodology (Table S 3).
(i.e., it is not “state independent”, but rather a “state”, which is “independent of the calibtation methodology”)

[RC2.05] L 291: I see the reference to Huss and Hock, but I also think that a brief description of the method/definition of discharge is warranted here. If I understand correctly, the discharge includes precipitation plus melt over the initial glacier area (i.e., the watershed area is fixed and the discharge includes the sum of glacier runoff and nonglacier runoff). The calculation does not take into account changes in evapotranspiration over the catchment, which could be significant, especially for glaciers that experience substantial retreat through the course of the simulations, which I think should be stated.

Yes, this is an accurate description, which we will incorporate when updating the manuscript as follows:
…calculated over the initial glacier area, i.e. with fixed watershed area, following the method presented in Huss and Hock (2018), accounting for glacier and non-glacier runoff, but not including possible changes in evapotranspiration…

[RC2.06] L 331–334: Another way to say this is that the large glaciers have long response times and therefore they are farther out of equilibrium with climate. For that reason, in a warming climate I would expect to see large glaciers tending to be in the lower left corners of Figures 5a,b.

Yes, this is true indeed. We will add this information in the updated sentence:
…the little remaining regional ice volume is typically located in the largest glaciers (which have longer response times and thus take longer to disappear), which here tend to have a mass balance that is more negative than the regional one…

[RC2.07] L 385: Suggest deleting “logically”.

We will omit “logically” here.

[RC2.08] L 462: Didn’t you already state previously that Rounce et al., 2023 would be referred to as PyGEM?

Yes, we did indeed already include this information earlier in the manuscript. The updated sentence will now be:
Our modelled regional and global glacier volume changes agree well with those simulated with PyGEM by Rounce et al. (2023).

[RC2.09] L 468: There is a typo here or maybe a missing sentence? "...PyGEM. Table 1Figure 7A noteworthy distinction..."

Thank you for spotting this. This will be corrected in the new manuscript.