The manuscript provides an in-depth evaluation of the uncertainty in glacio-hydrological modeling of the Patagonian Andes using the Open Global Glacier Model (OGGM). Through an extensive series of simulations (1920 in total), the authors have quantified how various model choices influence glacier melt and runoff projections. They examined six different model choices (referred to as sources of uncertainty), that relate to both historical datasets (different glacier outlines, glacier volumes, and reference climate data) and future climate forcings (various general circulation models, emission scenarios, and bias correction methods). Additionally, the relative importance of each source of uncertainty was evaluated using a random forest regression method. The analysis revealed that reference climate data is the most critical source of uncertainty, even for metrics related to future projections.

The authors have caried out a rather comprehensive assessment of model uncertainty, which is adequately documented in this manuscript. In response to the feedback during the first round of revision, the authors have reorganized the discussion section and improved the figures and general clarity throughout the text.

I deem the manuscript fit for publication after a minor, mainly textual, revision. Please consider the more detailed list of suggestions below.

Abstract

L24: area > 1km2 → with area > 1km2

L25: Here, I would suggest removing the mention of the number of catchments and hydrological zones. It does not add to the clarity of the abstract and it is not of key importance.

L25: Consider replacing "We used different glacier […]" by "As sources of uncertainty, we used different glacier […]" to prevent confusion about what is meant by *each source* in L30.

L30: "We used the permutation feature importance of random forest regression models to assess the relative importance of each source on the signatures of each catchment." → "We used the permutation feature importance of random forest regression models to assess the relative importance of each source <u>of uncertainty</u> on the signatures."

Introduction

L53-56: Having limited knowledge of South American geography, these lines appear confusing to me. How exactly are the 'Andean glaciers', 'Patagonian Andes' and, 'Southern Andes' related? You mention that the Glaciers in the Patagonian Andes are the dominant source of ice loss in the Southern Andes, and that all Andean glaciers combined are one of the largest contributors to sea level rise. Am I right to assume that the Patagonian glaciers thus make up most of the total number of glaciers in the Andes?

L54: mass loss → ice mass loss

L60: in rivers with important glacier cover → in rivers in glacierized catchments.
It is unclear what is meant by 'important glacier cover'. In addition, the introduction of the term 'catchments' in relation to rivers, at this stage, helps support your decision to study glacier evolution on catchment-scale rather than for each individual glacier.

L60-61: The subsentence ‘some of which […]’ downplays the rest of the sentence and seems contradicting to the first sentence of *Study Area* (L100). Consider rewriting to e.g., ‘However, recent studies have reported increased flows in rivers in glacierized catchments (Masiokas et al., 2019; Vries et al., 2023), with a growing number of rivers showing significant trends ($p < 0.01$) in the last decade (e.g., Santa Cruz; Pasquini et al., 2021).

L82: have examined the influence of → have simultaneously examined the influence of OR have compared the influence of OR have examined the respective influence.

L88: Unclear what is meant by ‘the climate model chain components’, this could be more explicit by stating e.g., ‘[…] the main source of uncertainty was associated to climate forcing data […]’

L92-97: The scenarios were tested […] → The resulting 1920 scenarios were simulated […]

Study Area

L106-107: I suggest to mention here why you aggregate the catchments into nine hydrological zones. I.e., ‘For this or that purpose/analysis, the catchments were aggregated into nine hydrological zones.’ This could help prevent the confusion that was brought up during the initial revision. Otherwise, the reasoning needs to be more explicitly mentioned in the method section.

L108: ‘[…], that showed a strong capacity to reproduce recent glacier changes.’ Technically, ‘that’ refers to spatial patterns in precipitation and temperature, which cannot explain glacier changes. Either rephrase to ‘[…], that showed a strong capacity to reproduce recent spatial variability in glacier changes.’ OR ‘‘[…]. Precipitation and temperature showed a strong capacity to reproduce recent glacier changes.’ or similar.

Figure 1: It is unclear to my why the names of the main catchments are included in all the figures. In my opinion it makes the figures more cluttered and less obvious. If you are not explicitly referring to these names, I suggest to remove them.

Methodology

L128: In fact, you are using the same version at the model used by Marzeion et al. (2012), since you set Pf to 1 (i.e., ignore it) and include a Tspinup (which is similar to their Tbias), correct?

L136: The term ‘positive degree-months’ has not been introduced with respect to Eq. 1.

L154: It is unclear what ‘this’ refers to. Consider rephrasing to ‘Not considering frontal ablation is an acknowledged shortcoming of our study […]’

L163: Although these lines are rephrased with respect to the initial submission, it is still unclear to me what is meant by the residual term. Based on Maussion et al. (2019), it appears that this is a time-based correction of the observed mass balance. Is that correct? I suggest to mention explicitly what you have done.

L173: The Tspinup appears very similar to Tbias which is more often used in mass balance equations like Eq. 1. In fact, in Appendix A the term temperature bias is also used. For consistency and compatibility with other studies, consider adding Tbias to Eq. 1 and mention in step i that this parameter is initially set to 0 for reference period simulations.

L195: I suggest to explicitly introduce the distinction between historical and future uncertainty sources from the start and ideally use the same or similar headers as in 4.1.

L224: PMET outperformed ERA5 […] → In an earlier study, PMET outperformed ERA5 […]

L247: This description of the different bias correction methods is rather abstract. Although I understand that (explaining) these methods is not the focus of the study, I suggest to clarify.

For instance (L246-248): '[…] method that combines quantile-based delta change and bias correction methods. Thus, it not only preserves the quantile changes predicted by climate projections, but also corrects the biases of modelled time series with respect to those of the reference time series.' Here, instead of explaining the method, the second sentence repeats the first sentence. In addition, 'but also corrects the biases of modelled time series with respect to those of the reference timeseries', is essentially the main aim of any bias correction methods, so it doesn't provide any real information on what distinguishes this method from the others.

L269: of → for

L282: I recommend to introduce the term 'commitment runs' here, as that term is used in the figures, but never explicitly coupled to these 16 runs.

L284: You also corrected annual glacier area, volume and specific mass balance for the additional 16 runs.

Table 1: first DJF → (December – February) OR (December, January, and February)

L306: As indicated by one of the earlier reviews, the term RMSE is confusing because this usually refers to an error between predicted and actual (observed) values (hence 'error'), while you are comparing a number of predictions to another prediction using one of the selected options for each source of uncertainty. Is that right?

In that case, does it matter which option (e.g., PMET as reference climate data set) is chosen as the initial selection? For example, imagine the projected glacier evolution to be very similar for ERA5, CR2MET and MSWEP, but very different for PMET: then the obtained total RMSE with respect to the PMET run would be larger than with respect to one of the others. I assume that the RMSE is always computed with respect to the previously selected option (instead of the initial one)? Maybe you can add a note on this in the text.

Results

L326: 'While […] both years).' This sentence reads confusing while you could just say that 84.7% of glacier area in RGI6 was acquired in 2000, the majority (~65%) of the data in RGI7 was acquired in 2001. There is no need to hide that on average, there is only one year difference between the acquisition dates in the two datasets. You come back to this in the discussion where you state that the inclusion of local data and different processing (correction) techniques are more important than the acquisition dates.

L336: The use of the term normalized thickness appears confusing to me because you haven't considered real ice thicknesses. Perhaps you can refer to it as normalized or scaled volumes, or 'ice volume per catchment area'.

L349: spatial climate diversity.

L352: When you refer to glacier area, is this based on RGI6, RIG7 or does it refer to the entire study domain?

L383: The term 'future climate uncertainty' is not properly introduced. At first, it was unclear to me that this refers to the combined product of four other sources of uncertainty. This section can easily be misunderstood as a preview of the discussion of the respective importance of the six sources of uncertainty on glacio-hydrological modeling. I suggest to introduce this term/analysis in section 3.3.

L405: The results suggest that ice loss will vary according to the different sources of uncertainty → The results indicate variable ice loss depending on model choices.

L407: Mention the term 'commitment run' in brackets.

Header 4.3: Use same header as in 3, Hydrological importance of sources of uncertainty.

L448: accumulated → explained

L455: 'Consistently, […]' Wasn't this to be expected?

Discussion

L475-477: This sentence is rather long and the last subsentence is unclear.

L490: may have peaked (2021 +- 15) → may have peaked in 2021 (+-15 year).

L496: more important for → more important than the volume data source for

L519: what was the other methodology that Watanabe et al. (2019) used? Could any aspect of their methodology explain the different differences between introduced uncertainties?

L533: when comparing your projections to those of Rounce et al. (2023), do you consider only the runs with ERA5 forcing or a combination of all historical climate data sets? I believe Rounce uses ERA5.

End of 5.4: In 5.2 you concluded that historical sources of uncertainty (ref climate followed by glacier attributes) are more important than future sources of uncertainty, while most other studies focus on only the future sources of uncertainty. This could be emphasized more.

Conclusion

L581: 'six sources of data uncertainty' → 'six sources of uncertainty associated with model choices'.

Appendix A

L614: Here you refer to Tspinup as the temperature bias. See earlier comment.

Supplementary Material

Table S2: Longitude – Latitude → Resolution (lat – lon)