Anonymous Referee #1

In the manuscript "Assessing the glacier projection uncertainties in the Patagonian Andes (4-56 S) from a catchment perspective", a large set of glacier model simulations obtained with the OGGM model is used to assess the uncertainty in glacier runoff and glacier melt projections. The study looks at the effect of different glacier outlines, different glacier volume estimates, various historical climate datasets, different GCM, different emission scenarios and different bias correction methods. Each of these different datasets are discussed and it is shown how they vary, mainly focusing on their spatial patterns. Using a random forest regression method, the relative importance of each of these "model choice" uncertainties is examined. The study concludes that the reference climate is the most important source of uncertainty for a range of glacio-hydrological signatures, even for signatures that represent a signal beyond the reference period.

Overall, I think that the study presents an impressive amount of model run comparisons at a large regional scale and shows clear insights into spatial differences of glacier volume and runoff changes in the Patagonian Andes. Moreover, the effect of each source of uncertainty on the glacio-hydrological signatures nicely illustrates their variable importance for different aspects of the change in glacier runoff. However, I feel that some parts of the manuscript could be improved, such as the use of "catchments", and the discussion section, which should describe more the implications and possible hypotheses of the findings, rather than only a summary of the findings and comparison with other studies. Please find below a list of more detailed comments, also explaining these two examples.

R: Thank you for your thoughtful and constructive feedback on our manuscript. We appreciate your positive assessment of the extensive model run comparisons and the insights into spatial differences in glacier volume and runoff changes in the Patagonian Andes. We will consider your suggestions for improvement, particularly regarding the use of "% of catchment area" and the discussion section. We recognise the importance of providing a more in-depth discussion of the implications and possible hypotheses of our results, rather than simply summarising and comparing them with other studies.

We will make the necessary revisions to improve the clarity and depth of the discussion section to ensure a more thorough exploration of the implications of our findings. In addition, we will prioritise the use of 'glacier area (at inventory date)' rather than '% of catchment area' in the manuscript to ensure appropriateness and clarity for potential future readers. Please see our detailed responses to these general comments below.

Abstract

The introduction of "catchments" starts here in the abstract. Without stating what "catchment" represents in this study, it is quite hard to follow. The reader doesn't know if the study is about a few catchments, or many, and how they are delineated/defined. Accordingly, all the statements with "xxx% of the catchment area" are difficult to interpret. On a more general note, it did not became clear why catchments are used in the study? Most of the results are rather described on the basis of the hydrological zones. In the study area description, it is noted "847 catchments were selected" without any further information. On the basis of what were these catchments defined? Since the study is not using any downstream information or non-glacierized catchment info ("only using glacierized grid

cells"), I think the "catchment" part should not be part of this study. Results could then be just presented as xx% of glaciers show this and that, while also keeping the aggregation level of the hydrological zones.

R: We acknowledge the concerns raised about the introduction of "catchments" and the resulting confusion for the reader. We recognise the importance of clarity in our presentation, and therefore we agree that the use of "% of catchment area" in the comparison between the different sources of uncertainty may add unnecessary complexity to the study. Consequently, we have decided to remove these values from the manuscript.

In the revised version, we plan to provide a simpler presentation by i) focusing only on "glacier areas (at inventory date)" in our analysis and ii) adding glacier area-weighted averages to summarise the values. These will be incorporated into Figures 5-7 and 10 and will contribute to a clearer and more concise presentation of our results. Finally, we plan to keep the aggregation level of the hydrological zones to explore the spatial differences over a wide area that includes different climates and geographical features.

Introduction

L44 undergoing "a" shift – shouldn't shift not be an increase, as otherwise there is no increase in risk? Or does it relate to larger volumes that are stored before bursting out?

R: The specific sentence on GLOFs will be removed to improve the continuity with the following paragraph.

L65 "have had to be used" – have been used? Also check the "but" which is more an additional problem than a contrast?

R: The specific sentence will be removed

I would suggest to move table 1 to the SI. The abbreviations in the first column are not yet clear and in general the aim/message of the table did not become clear in the introduction

R: Thanks for the comment. Table 1 will be moved to the Supplementary Material. The abbreviations in the first column will be defined in the table caption.

L93 "to project the evolution of the glacier area" – the model is not only used for projecting the area?

R: The specific sentence says "to project the evolution of each glacier (area > 1 km^2)". The glacier model outputs used are not only the glacier area, but also the volume and hydrological outputs (glacier runoff and melt from formerly glacierised areas, see Section 3.4).

Study area

L104-105 "crucial", "essential" in the same sentence, maybe one is enough?

R: The sentence will be summarised: "the seasonal melting of glaciers is essential for the long-term sustainability of the local ecosystems and coastal human populations".

Methods

Step 1 of the calibration procedure – I wondered if it should be discussed what the effect is of only calibrating the melt factor/temperature sensitivity using geodetic mass balances. If the precipitation is off (e.g. too little) then the temperature sensitivity might be smaller, to compensate?

R: It is true that the calibrated temperature sensitivity parameter depends strongly on the assumed precipitation, and thus on the chosen precipitation factor (Pf; see e.g. Schuster et al. (2023). However, for most glaciers we have only one robust observation, i.e., the 20-year geodetic mass-balance average. Therefore, we decided to set the Pf to 1, as some regional climate datasets used in this study (e.g. PMET and CR2MET) already include a bias correction process to correct for potential precipitation underestimation, and therefore we expect the range of "true" precipitation to be covered by the different products. This will be added to the discussion.

Step 2 – the description is a little bit unclear, is the gradient adjusted? Or what is meant with "a mass balance residual to add"?

R: Sorry for the confusion, the gradient of the mass balance is not adjusted. In fact, the whole mass balance profile is shifted so that the resulting apparent mass balance is zero, to satisfy the equilibrium assumption. We will make this clearer in the manuscript.

Step 3 – maybe the equation could be given here? I think it is similarly known as equation 1, but helps the reader understanding what is needed for the inversion and at what scales it is applied

R: Thanks for the suggestion. We think that the equations for the inversion are not that relevant for the paper and are already very well documented in Maussion et al. 2019. However, we will add the equation for the ice velocity to illustrate the control of the ice creep parameter A (which is calibrated in this step).

Step 4 – what is a "constant mass balance run"?

R: A constant mass balance run means that we define a mass balance over a period of time (e.g. the average mass balance of all the years in the period), this is our constant mass balance. Then we use this constant mass balance and let the glacier evolve for a few years. We will clarify this in the manuscript.

Figure 2 – I think it would be helpful to add the number of the steps in the figure. And what is the arrow from reference climate to bed inversion? It may refer to one step before that?

R: Thanks for the suggestion, we will add roman numerals in Fig. 2 to refer to the calibration steps in Section 3.1

L259 "according to the location of the glacier terminus" – I am not sure to understand this addition? Maybe some more explanation of glaciers that are crossing catchment/hydrological zone borders could be added here

R: Thanks for the comment. This will be clarified according to your suggestion: "For area and volume, we calculated the relative and absolute differences for each catchment and hydrological zone defined in Fig. 1.

To calculate these differences, we aggregated glacier area and volume for a given catchment by selecting all glaciers with their terminus location within that catchment. It is assumed that, if the inventory outlines are correct, all the water flowing out of the glacier will flow via its terminus.".

L291 here the text could benefit from some explanation/ careful discussion how streamflow metrics can be used to apply on aggregated glacier runoff data (i.e. glacier runoff is not the same as downstream streamflow and so their effect on the aquatic ecosystem is not 1:1 comparable)

R: Thanks for the comment. We will add a sentence to highlight the limitations of our analysis: "However, our analysis of glacier runoff should not be considered as downstream streamflow because our simulations only considered the initially glacierised area and did not include the interaction with other hydrological fluxes (e.g., evaporation and infiltration)"

Table 2 misses information about which period was used to calculate the signatures, apart from the ones that explicitly state "ref and future period"

R: We will add a "Period" column for each signature.

Header of 3.5 – maybe choose the titles of section 3.3 and 3.5 in such a way that it is more clear what their different content is

R: The original header (Uncertainty analysis) will be renamed to "Hydrological importance of sources of uncertainty"

L299 – where comes 329 from?

R: Out of 847 catchments, only 329 catchments have one or more glaciers in both inventories (RGI6 and RGI7). This will be clarified in the text.

In the calculation of RMSE, what is the baseline? i.e. how is RMSE calculated?

R: Thank you for your question. The permutation feature importance measures the change in model performance (in this case, the Root Mean Square Error; RMSE) after the values of a single model feature have been permuted (also known as shuffled), with more important features resulting in greater decreases in performance when permuted. The baseline corresponds to the model performance before the permutation of the model features (in this case, the six sources of uncertainty). This will be clarified in the manuscript.

Results

Figure 2: possibly remove the blue/gray background so the results are better visible

R: We think that this comment refers to Fig 3, as there are no results in Fig 2. We will add black outlines and a zero line to the bar plot in Fig 3b to prevent the bars being confused with the background.

L335 – why was the glacier thickness divided by the catchment area, while all other variables are focused on the glaciers/glacier grid cells?

R: In the case of ice volume, which is derived from ice thickness and glacier outline, we decided to normalise the value by catchment area to facilitate the comparison between catchments - hydrological zones.

L374 – maybe I missed it, but what are "the main catchments"?

R: The main catchments correspond to the catchments that have an area greater than 5,000 km^2 (Fig. 1), which account for the 68% of the total catchment area. All maps have the following sentence to indicate the main catchments: "The names in grey correspond to the names of the main catchments"

L408 "historical sources" – climate?

R: Yes. This will be clarified in the text ("Considering the prolongation of historical climate conditions, 26% ± 9% of the total glacier ice is committed to melt in the long term")

L435-L440 Does isolating the melt from glacier runoff result in more or less uncertainty? There is a hint that melt only has an effect of temperature, but it also states that precipitation can compensate for the change in melt? The 60% suggests that 40% of the glacier runoff is generated from off-glacier melt, or liquid precipitation on+off glacier?

R: Compared to total runoff (Fig. S2), the uncertainty (standard deviation of the annual glacier melt) was reduced because the reference climate, which is the main source of climate uncertainty, had less influence, as the strongly past climate-dependent liquid precipitation is only included in the glacier runoff components (on- and off-glacier melt and on- and off-glacier liquid precipitation). This is mentioned in the Discussion section (L553-L556). Therefore, the 60% suggests that 40% of the glacier runoff is generated by off-glacier melt and on- and off-glacier liquid precipitation.

Discussion

For the discussion in general, I found it sometimes hard to follow what was described in the different parts. The first part discusses the uncertainties, but the other two parts as well. Maybe by less repetition of the results, and more focusing on the implications of the results would help in restructuring the discussion. I think the comparison with other studies is mostly well described, but would also benefit from an additional thought on what are the implications of that it agrees well or not.

R: Both reviewers have raised concerns about the discussion section. To address this, we have proposed a new structure for this section. This revised structure aims to minimise repetition of results and to consolidate previous sections on uncertainty. In addition, we plan to incorporate the suggestions of both reviewers by focusing more on the implications of the results, rather than simply comparing them with other studies. This approach will improve the clarity and coherence of the discussion and allow readers to better understand the significance of our findings.

The proposed new structure is as follows: Section 5.1 ("Changing glacier hydrology") will present regional projections of changing glacier hydrology and compare them with results from previous studies. Section 5.2 ("Hydrological importance of data uncertainty") will discuss the hydrological importance of data uncertainty, summarising comparisons and highlighting the importance of sources of uncertainty, emphasising spatial differences and the importance of domain characteristics, as suggested by Reviewer 2. In addition, the importance of model calibration is addressed, as suggested by Reviewer 1. Finally, Section 5.3 ("Limitations and global implications") discusses limitations such as unconsidered sources and potential global implications suggested by both reviewers.

I was wondering if there should be maybe a discussion on the way such glacier models are calibrated? If all model runs are calibrated equally well (RGI areas and glacier volumes are used for calibrating), then how come the results are so different, especially regarding reference climates? These parameters propogate in all other simulations, right? This must then come from processes that are not captured when looking at annual and long-term metrics only when calibrating model parameters? Or maybe there are other processes that are relevant to discuss for improving glacier modelling?

R: We recognise the need to discuss how glacier models are calibrated and the potential factors that contribute to differences in results despite the use of the same calibration approach. This will be addressed as part of the new section 5.2 entitled "Hydrological importance of data uncertainty" (see previous response). In this section, we will explore the implications of model calibration. In section 5.3 ("Limitations and global implications"), we will discuss the need to use additional data to calibrate other relevant processes that may be critical for improving the accuracy of glacier modelling. The proposed new discussion aims to provide insights into the complexities of glacier modelling and strategies for improving its reliability.

L484 "acquisition dates" – this sentence could benefit from more explanation about what the difference is in both RGI outlines with respect to acquisition dates.

R. Thanks for the comment. The sentence will be modified with further explanation: "The decrease in glacier area in the major ice masses from both inventories may be due to improved outlines from the local inventories and differences in the acquisition year of glacier outlines (Fig. 3d). Furthermore, these differences show the effect of climate and water warming on the complex dynamics of these glaciers (Minowa et al., 2021; cited in the main text)".

L487 "this threshold" – do you mean a 10% threshold for volume?

R: Yes, the sentence compares the relative differences between different sources of glacier inventory and ice thickness estimates.

L494 – Isn't this paragraph suggestion a contradiction? It would help starting the paragraph like that.

R: The main topic of the specific paragraph is the uncertainty associated with the reference climate (in specific precipitation). The content of this paragraph will be summarised and included in the subsection: "Hydrological importance of data uncertainty".

L510 – a space needs to be inserted before "The selected…."

R: Thanks for the catch. A space will be inserted there.

L512 “the ten selected….of the assessed warming – not sure how this sentence fits in?

R: The specific sentence will be removed from the discussion to avoid repetition of the methodology/results.

L526 Why are the “older” estimates named after the newer estimate and referred to as “recently”?

R: Thanks for the question. The “recently” was misplaced there and will be removed from the sentence.

L539 “of its current volume” – although it is clear that the study only discusses glaciers, the confusing use of “catchments” means that here there “glacier” needs to be added

R: Thanks for the comment. The reference to "catchments" will be removed from the sentence to clarify the source of the volume loss.

L594 – possibly add that Mackay et al. deals with glacio-hydrological modelling in Iceland

R: Thanks for the comment. We will add this to clarify the scope of the antecedent.

Conclusions

L610 – “differences”, “different”, “varied” in one sentence – consider rephrasing.

R: The specific sentence will be rephrased: “The six sources of data uncertainty showed differences of varying magnitude”.

L613 isn’t it a separate point, the one on reference climate being most important for uncertainty?

R: We decided to aggregate the results of Section 4.1 (“Analysis of sources of uncertainty”) as one point of the conclusions. L613 compares the relative differences between glacier inventories and volume data sources (Section 4.1.1), while L614 refers to future sources of climate uncertainty (climate, SSP, GCM and bias correction method; Section 4.1.2 and Section 4.1.3).

L622 “tended to converge towards an overall decrease” – decrease of what?

R: Decrease of glacier runoff and melt. This will be clarified in the conclusions: “glacier runoff and melt projections aggregated by SSP scenario tended to converge towards an overall decrease throughout the 21st century”

L625-628 double?

R: Yes, the specific sentence indicated that regardless of the spatial scale of interest (catchment or glacier area), the main source of uncertainty was the reference climate. This sentence will be clarified as the “catchment area” will be removed from the Results and Discussion sections.

L629 – what is meant with “local”?

R: The specific word will be removed from the sentence.

L634 “relative contribution of non-glacial water sources” – shouldn’t future studies not focus on the dynamics of these sources, rather than the relative contribution? And what should future studies do with these non-glacial water sources? The follow up sentence does not directly fit here, i.e. there is some gap between knowing a relative contribution and understanding other catchment stores. Maybe a general sentence on extending the scope from glaciers to downstream hydrology fits better?

R: Thanks for the suggestion. We agree that a general sentence about broadening the scope is more appropriate than focusing only on non-glacial water sources. In the revised version, we will modify this as suggested: “future studies should address sources of uncertainty not considered in this study (…) and extend the scope from glaciers to downstream hydrology. Downstream hydrology can play a critical role in the seasonal and interannual water release during dry seasons (Drenkhan et al., 2022), attenuating the consequences of glacier shrinkage (e.g., Somers et al., 2019)”