Reviewer 3

In this study, Gimenes et al. use OGGM and two ice thickness products to assess the impact of geometric data and temperature bias on the magnitude and timing of peak runoff in the Western Kunlun Mountains. For future projections, the authors employ one GCM simulation extending to 2300 under SSP1-2.6 and SSP5-8.5 scenarios.

The topic is certainly interesting and relevant, but the study provides limited innovation in terms of methodology or analysis. The level of depth is rather modest, and several aspects of the methods and datasets require clarification. Even for a Brief Communication, I believe these points should be in depth addressed. While the results show that initial ice thickness strongly influences runoff, this outcome is somewhat expected. The “so-what” aspect of the study is currently missing and should be emphasised more clearly through deeper analysis and contextualization as well as richer figures. Given the potential of the topic, I recommend major revisions, with the hope that the comments below will help strengthen the manuscript.

Authors: We would like to thank Reviewer 3 their detailed comments that helped improving the paper clarity and quality.

Major comments

• Introduction:

The second part of the introduction does not really serve as an introduction to the study itself, but rather as a general overview of glacier modelling approaches. It would be more relevant to provide additional context on which variables most strongly influence glacier runoff and why ice thickness and temperature bias are key factors controlling peak water so that these are chosen. This could then naturally lead into the motivation and significance of the present study. Why is the Western Kunlun Mountains are chosen should also be addressed.

Authors: We provided more information on parameters controlling glacier runoff in the second paragraph of the introduction at L18-20. The choice of the region was addressed in the introduction of the revised version (see earlier comments to Reviewer #1 & #2).

• Methods

- The authors mention that the ice thickness data from Millan et al. (2022) were corrected to match the consensus estimate (reported for 2003), but the procedure is not entirely clear. Was the ice thickness change from 2000–2019 (Hugonnet et al., 2021) applied across the full domain, or only partially (e.g. 2003–2017)? It would also be helpful to clarify what the reported average value of 4.0 km³ (line 90) represents.

Authors: The correction procedure was further explained at L91-97; ice thickness change was applied on the full time domain 2000-2019 and we added percentage of change in ice mass compared to original estimates (see response to comments of reviewer #2):

“we corrected MIL22 estimates to get corresponding thicknesses in the year 2000. As correction, we simply subtracted average glacier thickness changes of Hugonnet et al. (2021) between 2000 and 2019, which is obtained from DEM differencing, to MIL22 thicknesses. While solving temporal ambiguities of the thickness models is complicated, since DEM sources and in-situ data are not properly dated, this correction may be a step toward roughly matching the timing of the consensus estimate. Overall, differences after correction for glacier mass change average 6 km3 compared to the original MIL22 estimate, representing 1 % of the latter.”

- The calibration section needs revision. It currently reads as a mix of “limitations” and “comparison with previous studies”, which weakens the message and the perceived robustness of the study. Why not rerunning OGGM using the new configuration to ensure consistency with the latest developments?

Authors: We are now using the new configuration of OGGM for the revised version, so the calibration section was accordingly shortened.

- The discussion of the temperature bias (lines 180-181) is too brief, as it is a central element of the title. Is this bias the same as that derived during the calibration? Was the same procedure repeated but with a uniform bias applied over the full simulation period? Please clarify this.

Authors: We clarified the experiment at L172-175 as following:

“To this aim, we conceive an ensemble of projection runs (also using MIL22 dataset) adding a uniform temperature bias (ranging from -5 to 5°C) over the full simulation period, meaning that a this bias is added to the temperatures time series used by the model to calculate surface mass balance (additionally to any bias calculated for the mass balance calibration, see section 2.3).”

• Results and discussion:

- Figures 2 and 3 could be further optimised. Axis limits are inconsistent, which makes comparison difficult. Colour choices are sometimes awkward; using the IPCC standard colour palette would improve clarity. Labels are not always self-explanatory, making the caption crucial for interpretation, and overall label thickness should be increased for readability.

Authors: Figures 2 & 3 were further optimized accordingly to the comment, improving overall clarity. However, we cannot apply the exact IPCC colors for SSPs since one of the co-authors has color-blindness.

- The discussion lacks comparison with previous studies or observational runoff datasets that could provide at least partial validation. As a result, the broader context and implications of the results remain unclear.

Authors : As detailed above, the purpose of this study is not to obtain an exact value of runoff, but rather to understand the sensitivity of peak water to geometry and to potential temperature biases that may be present in climate models. This objective is now clearly stated in lines 32–38. Note, however, that we have nevertheless included a comparison with runoff estimates in the Tarim region (Gao et al., 2010) for the period 1961–2006 in lines 273–279.

Minor comments and suggestions:

Abstract

• Line 1: Sensitivity to what? Please specify (-> ice thickness and temperature?).

Authors: Ice thickness and temperature are already mentioned in the following sentence and adding it again would be too repetitive.

• Line 4: Replace Predicted with Projected.

Authors: Done.

Introduction

• Line 11: Add a reference for the Sea Level Rise component.

Authors: The reference is at the end of the sentence of L11, so there is no change in the revised version.

• Line 12: Remove also.

Authors: Done.

Data and Methods

• Line 37: Replace with “more precisely, the Western Kunlun Mountains.” (or West?)

Authors: Done.

• Line 38: Clarify why RGI v6.0 is used to define the study area (maybe better add this reference in the next sentence).

Authors: RGI v6.0 was used in Farinotti et al., 2019 and Millan et al., 2022 to invert ice thickness. It is therefore used again in this study. A clarification and reference was added at L43: *“This study specifically targets a group of 160 glaciers with a total surface area of approximately 2900 km2, calculated with the RGI v6 (RGI Consortium, 2017) (Fig. 1).”*

• Line 45: The phrase “these glaciers” currently implies 10% of 70%. Please reword for clarity.

Authors: Agreed, we clarified that by replacing *“these glaciers*” by “*the remaining 20%*” at L50.

• Line 50: A thinning rate of −9.6 m yr^{-1} seems unrealistically high? Please verify the source or units.

Authors: We double checked this value and realized we had read the wrong rate. Mean elevation change rate is now of -0.23 ± 0.04 m yr−1.

• Region 13 is generally thinning, but the Karakoram is an exception within RGI13, what does this imply for interpreting the Hugonnet (2021) regional value? Why not taking the Karakoram region from Hugonnet (2021) data separately?

Authors: We think that adding the Hugonnet et al. (2021) regional value provides context on the general thinning that is occurring at the scale of region 13, as a point of comparison with the Karakoram anomaly illustrated by values from Brun et al. (2017) previously mentioned in the text (L53-54). There is no mentioned value for only the Karakoram in the published paper of Hugonnet et al. (2021); the acceleration of thinning in this area is mainly illustrated in extended data Figure 7 of the same study.

• Figure 1: The inset zoom on the study area in HMA is too small. Please revise and consider adding the ice thickness dataset already in the panel (not only in the caption). The figure should be interpretable without relying on the caption. If ice thickness measurements exist in the study region, they could be included.

Authors : We have added the source of each dataset inside the Figure 1. The inset panel is already hiding some glaciers of the region, and is only meant to provide the rough location of the mountain range.

• Line 63: Add a reference for this statement. This could also be moved to the introduction to justify the regional focus.

Authors: We repeated the reference for this statement at L66. Concerning moving to the introduction, we have already accounted, in response to earlier comments which leading to a reshaping of the Introductory part.

• Line 67: Replace Predicted with Projected or Simulated.

Authors: Done, replaced predicted with simulated.

• Line 68: The consensus (or community) estimate was published in 2019 but represents 2003 (I thought actually more 2000), not 2019.

Authors: Agreed, see response to similar comment by Reviewer #2.

• Line 73: Clarify the phrase “the use of glacier by.”

Authors: We clarified that with the phrasing *‘‘the use of flowline inversions "glacier by glacier"’’* at L77-78.

• Line 84: Confirm whether Farinotti et al. (2019) used measurements.

Authors: We added information about the use of measurements in Farinotti et al. (2019) from the GlaThiDa v2 dataset at L87-88 : *“Farinotti et al. (2019) used thickness measurements from the Glacier Thickness Database (GlaThiDa) v2 (WGMS, 2016), with less than 50 glaciers being covered within this RGI region”.*

• Lines 84-89: The correction method for the Millan dataset is unclear — was the Hugonnet thickness change simply added?

Authors: We clarified the correction, see response to major comment on this issue.

• The OGGM initialization description is too detailed for a Brief Communication. Focus on “We use OGGM, a model…” and emphasize what was new or modified for this study.

Authors: We shortened the section with the new model framework, see response to Reviewer #2’s comment on this issue.

• Lines 108–112: Clarify which dataset was ultimately used and for what time period.

Authors: We clarified in section 2.4 ‘Climate forcing’ that W5E5 data is used for dynamical spin-up or historical runs and that CMIP6 GCMs are used to simulate glaciers evolution from 2020 until 2300.

• Lines 113–127: This section reads more like discussion than methods. It could be shortened substantially. Based on line 121, it even seems that part of the research question has already been answered.

Authors: This section was removed as we are using a new model framework in the revised version, calibration section is now at L108-110.

• Lines 154–159: This belongs more logically in the description of the reanalysis dataset. Ensure consistent verb tense (past or present) throughout each section.

Authors: This section was also removed since we now use the dynamical spin-up procedure in the revised version.

• Consider first presenting the individual glacier runoff results before aggregating to the regional scale.

Authors: We added a description of an individual glacier runoff composition at the end of section 2.3’ OGGM’, L116-118:

“Finally, the model gives as simulation outputs glacier volume, length and area, as well as glacier runoff. Considering a fixed glacier area including glacierized and increasingly non-glacierized terrain, the annual total runoff computed in OGGM is derived as the sum of snow melt on now ice-free area, the ice and

seasonal snow melt on glacier, the liquid precipitation on glacier, and the liquid precipitation on now ice-free area (e.g. in Fig. S1).”

• There is some mixing of climate data, geometry, and model description; please reorder for clarity.

Authors: We did reorder the section order for clarity in the revised version : 2.3 OGGM, 2.4 Climate forcing, 2.5 Integration of ice thickness dataset and 2.6 Peak water calculation.

• Why was only one GCM used, given that six are available? The paper itself acknowledges that using a single GCM limits accuracy — please justify this choice.

Authors: We are now using a multi-GCMs ensemble, see response to comment of Reviewer #1 for more details.

Results

• Figure 2: It is unclear what the “multiplication factor” represents. Also, clarify the meaning of “temperature bias.” Ensure consistent y-axis ranges across panels and apply IPCC colours for SSP1-2.6 and SSP5-8.5.

Authors: We changed “multiplication factor” to “initial ice volume multiplication factor”. We think that the signification of temperature bias was further explained in section 2.4 (see response to previous comment). y-axis are now consistent across panels of figure 2 and we also changed the color palette.

• Figure 3: Move the figure closer to the relevant text. The colour scheme used to distinguish Farinotti et al. (2019) and Millan et al. (2022) is identical to that used for SSP scenarios in Figure 2 — please revise for clarity. Label font size should be increased.

Authors: Agreed. We moved the figure closer, label size was increased and color scheme changed. We did not use the exact IPCC colors for SSPs because it made it difficult to read for one of our color-blind co-authors.

• The content of panel C is not intuitive, why not show annual mean temperature and total precipitation instead?

Authors: The corresponding panel was changed in the revised version and is now temperature and precipitation from the different used GCMs.

• Lines 241–246: Move to the discussion section.s

Authors: Done.

Discussion

• Consider subdividing this section into subsections for clarity.

Authors: Done, discussion was divided into four subheadings.

• Is Millan et al. (2022) systematically thicker across the entire glacier, or mainly in the accumulation area? If the latter, would that difference significantly affect runoff?

Authors: MIL22 is thicker than FARI19 across the entire glacier, both at low and high elevations (see Figure 1c). We added a sentence to describe this at L97-99.

• Line 280: Please clarify what “sufficient” refers to.

Authors: “sufficient” was replaced by ‘‘warm enough’’ in this paragraph now at L257.

• Lines 297-304: The outlook is vague. The discussion should place the results in a broader context and compare them with previous studies. Many results are rather intuitive (e.g. thicker ice leads to delayed and higher peak water). It would be more appropriate to revisit the research questions rather than expand on new remote-sensing products, which remain uncertain if and when they will come.

Authors : In response to an earlier comment, the discussion on data sharing and satellite mission is now removed. We further disagree with this comment since the discussion already placed the study in a broader context. We start by (1) providing interpretation to the sensitivity analysis results, (2) compare the calculated runoff with other studies on longer time period, (3) discussing limitation of the methodology used to assimilate ice thickness in global models, (4) discuss the broader context of ice thickness uncertainties and global model geometry limitations. No changes.

• Why is there no conclusion section? Please include one to summarise the key findings and implications, even in a couple of sentences.

Authors: We added a conclusion in the revised version:

“This study highlights the strong sensitivity of peak water timing and magnitude to uncertainties in initial glacier thickness and temperature biases in climate models. In regions where ice thicknesses are highly uncertain, such as the Western Kunlun mountains, peak water can be delayed by a decade, while its magnitude can change by up to 27% depending on the data source used under SSP-5.8.5. With the same scenario, peak water date can be brought forward by roughly a decade for each degree of temperature bias in the climate forcing data used. Finally, our results emphasize that accurate estimates of glacier geometry are crucial for robust projections of future water availability.”