Reviewer 2

Summary

This study investigates how uncertainties in glacier ice thickness estima-tes and temperature projections affect the timing and magnitude of "peak water" - the maximum glacier runoff that occurs before glaciers shrink sufficiently to reduce water output. Using the Open Global Glacier Model, the authors compare two global ice thickness datasets (Farinotti et al. 2019 and Millan et al. 2022) for 160 glaciers in the Western Kunlun Mountains. They analyze sensitivity to initial ice volume and temperature bias under different climate scenarios. Key findings show that smaller initial ice volumes lead to earlier peak water occurrence, particularly under high-emission scenarios, and that the substantial differences between ice thickness datasets (40% volume difference) translate into important variations in predicted peak water timing and magnitude. Finally, this study highlights the large uncertainty stemming from poorly constrained ice thickness in High Mountain Asia and calls for improved datasets and integration of 2D/3D thickness estimates into models.

The topic is important and the approach is reasonable, but methodological limitations reduce the study's impact. The main finding is valuable, but the paper would benefit from methodological updates, clearer contextualization, and a sharper framing of its contribution. I recommend major revisions.

Authors: We would like to thank Reviewer 2 for the thorough and detailed comments that contributed in largely improving the paper quality and robustness.

Major Comments (not in order of importance)

1. Introduction. Good background on peak water and ice thickness uncertainty. However, the review of global models (flowline models, Elmer/Ice, emulators) is lengthy and somewhat tangential for a brief communication. I suggest focusing on (1) why the problem matters globally, and (2) regionally, (3) what is uncertain, (4) what the study does to address this. Considering that the study area is already state in the title, maybe it would better to have a few sentences of the "region of interest" here. There are several introduction-like sentences there: "*This basin stands out as one of the most vulnerable, susceptible to the impacts of climate, political, and socioeconomic changes*"

Authors: Agreed. We have now completely reshaped the introduction (L10-38), removing the review section on models, and focusing on (1) why the problem matters globally, and (2) regionally, (3) what is uncertain, (4) what the study does to address this.

2. Methodological choices. The study acknowledges using an "earlier calibration framework" and justifies this choice based on the research questions. I partly agree, but I think the value of the study lies not only in the comparison between F19 and M22, but also in providing new insights for local ecosystems and communities. If "*the new calibration methods now available in OGGM would allow a closer match to observed mass balance values and consequently more accurate estimates of real peak water timing and magnitude*" then why not use them? Specifically:
   - Using the default climate dataset (CRU) is not a sufficient justification. Given the sensitivity to climate input, I would expect some evidence that this product performs better than other datasets available in OGGM.

Authors: Agreed. We are now using the W5E5 dataset in the revised version since it is used as baseline climate in OGGM. Additionally W5E5 is employed as input for the bias-correction of ISIMIP3b CMIP6 GCMs, therefore it seems relevant for us to use it for bias-correction of the CMIP6 GCMs that we handled in this study.

- The argument that “the climate data used for the spin-up procedure are mostly the results of large interpolations” is true for CRU, but this is precisely why ERA5 or W5E5, now the OGGM defaults, would be more appropriate.

Authors: Agreed. We are now using W5E5 dataset for spin-up procedures.

- Stating that “no spin-up procedure was used because of the large uncertainties in climatic data in Asia” is not convincing, since spin-up is the default in OGGM despite large uncertainties being present in most mountain regions.

Authors: Agreed. We used the dynamical spin-up available in OGGM in all simulations, except for certain sensitivity tests. For the sensitivity experiments on initial ice volume, the prescribed volumes were much lower or higher than realistic values, preventing the iterative dynamical spin-up from converging to match the RGI glacier area given the resulting glacier geometry and the calibrated mass balance model. Similarly, for comparison purposes, the dynamical spin-up was not applied in the temperature bias sensitivity tests. Instead of the dynamical spin-up we have conducted historical simulations usg W5E5 data from 2000 until 2020 in order to initialize glacier state before future projections simulations. The text was modified accordingly at L112- 115 and L169-171 to describe these changes.

“We also make use of a feature from the latest version of OGGM which is a dynamic spin-up that can provide glacier initial state for the year 2020 by reconstructing its recent past while ensuring that modeled glacier area and observed one are matched within 1 % at the inventory date under historical climate forcing (Aguayo et al., 2023; Zekollari et al., 2024).“

“We do not use the dynamical spin-up in this set up since it can not converge to match the RGI area with the amount of simulated reduced or increased ice volume. Instead, after ice thicknesses assimilation we simulate glaciers evolution from 2000 to 2020 simply using historical W5E5 data to initialize glaciers before future projections (see section 2.4).”

- Glacier-specific observations (e.g., geodetic mass balance) could play an important role, as shown in Zekollari et al. (2024). Their use would be especially relevant here, since they are already applied to correct ice thickness.

Authors: Agreed. We now use the mass balance calibration on geodetic mass balance observations in the revised version. A description was added at L108-110 :

‘‘The surface mass balance model is calibrated on geodetic mass balance observations obtained from remote sensing (Hugonnet et al., 2021) for the 2000-2019 time period, using the W5E5v2.0 climate dataset as forcing (Lange et al., 2021). ’

3. Relying on a single GCM is a recognized weakness, especially given the availability of multiple ready-to-use GCMs in the OGGM Shop. Including an ensemble analysis (e.g., “Out of the 6 GCMs available, MRI-ESM2.0 ranks below the ensemble mean”) would substantially strengthen confidence in the results.

Authors: Agreed. This was modified accordingly. See earlier response to Reviewer #1.

4. Considering the almost identical results obtained under SSP126 and SSP534, I recommend carefully double-checking the simulations. Such similarity might indicate a potential issue in the climate forcing data. Since these scenarios differ substantially in their radiative forcing trajectories, one would normally expect more divergence in glacier and runoff responses, especially toward the end of the century.

Authors: Results were updated using the new modeling framework (see previous comments). We are therefore able to see more differences in changes in volume and area of glacier between SSP126 and 534. We also added in the paper that the latter is actually the SSP5-3.4 Overshoot (OS) scenario and that is is *''actually not an intermediate scenario since it explores the implications of a peak and decline in forcing during the 21st century (Lee et al., 2021)''* at L131-132. It is described the following way in the AR6 of IPCC: *"By 2300, warming under the SSP5-3.4-OS scenario decreases from a peak around year 2060 to a level very similar to SSP1-2.6."*. This is consistent with climate evolution represented in Fig. 3d of the revised manuscript. However, since the two scenario are still quite close, we only display SSP126 and SSP585 for the peak water in order to enhance readability, but Fig. 3c is available with SSP5-3.4 in the supplementary. Figure 3 was updated accordingly. *We chose those scenarios since they are only ones available in OGGM running until the year 2300.*

5. The research question is clear: how sensitive is peak water to uncertainties in initial glacier volume and temperature in a highly glacier-dependent region? While sensitivity analyses are not new, applying them systematically to the Western Kunlun using two global thickness products is a valuable case study. However, the novelty is somewhat limited, as similar experiments have been carried out elsewhere. The contribution could be strengthened by explicitly situating this study within the broader context of global peak water analyses and clarifying what unique insights this regional focus provides. For example, the 40% volume difference between datasets is substantial - discuss whether this represents a common local value or is particularly extreme. From Table 1 in Millan et al. (2022), it seems this would be a very extreme scenario, and therefore some context/discussion is needed.

Authors : Agreed. We have added in the introduction section a sentence situating this study within a broader context of other peak water studies and clarifying what new insights are brought here (L28-31) :

"*While several studies have investigated the estimation of peak water at regional scales \citep{pwhuss,caro}, there are, to date, few quantitative assessments of how uncertainties in initial glacier geometry and climate model biases influence its timing and magnitude*."

Furthermore, along with the reshaping of the introduction, we are now specifically stating that this mountain range was chosen because of the extreme difference in ice thickness models (L27-28).

Given the research questions, it is unclear why the authors chose to analyze total glacier runoff rather than glacier melt (snow + ice), since only the latter would be influenced by the choice of ice thickness sources, while precipitation, which is included in the total glacier runoff, remains identical in both scenarios.

Authors : We thank the reviewer for this comment. Indeed, since precipitations are the same across simulations, any difference in total glacier runoff necessarily reflects the impact of ice thickness and geometry on melt water production, storage, and release. Furthermore, our analysis focuses on total glacier runoff rather than melt alone because our objective is to quantify the hydrological implications of differences in initial ice thickness. Indeed, total runoff better represents what downstream hydrological systems experience. Furthermore, in the temperature bias sensitivity experiments, only air temperature was perturbed, while precipitation remained identical across runs. Nevertheless, changing temperature could potentially affect the partitioning between solid and liquid precipitation, thereby impacting seasonal snow accumulation and melt timing. Even though precipitation remains constant, this shift in phase and timing could modify the runoff regime and consequently the timing and magnitude of the peak water. Additionally, we find it more convenient to have results in terms of total glacier runoff for potential comparison in future studies.

Specific Comments Title and abstract:

- "Western Kunlun Mountains of the Tibetan plateau" - "plateau" should be capitalized as "Plateau"

Authors: Done.

- Define SSP in the abstract (first mention)

Authors: Done.

- "*Temperature bias also significantly influences runoff magnitude and the timing of peak water, especially under high-emission scenarios*" This feels not very informative for the abstract, what is the influence extent? Maybe it would be better to emphasize the regional importance of findings rather than general glacier science.

Authors: We have modified the abstract accordingly, to better provide the influence extent as follow : *" Temperature bias also notably influences the timing runoff by delaying peak water timing in the region by roughly 13 years for each bias degree".*

- Avoid using "*significantly"* if there is not a statistical test applied (is repeated four times in the current version).

Authors: The word was removed from the abstract and replaced by *''notably''*.

Section 1 (Introduction):

- "*glacier runoff is expected to rise and reach a maximum*" ... only in certain regions as in many places peak water was already reached.

Authors: Done, we added the following phrase (L13-14) to clarify that: "*(if it is not already reached as it is the case in many regions)*"

- "Determining precisely the timing and magnitude" -> "Determining the precise timing"

Authors: Done.

- The transition between ice thickness products and glacier models could be smoother. Suggest adding

1-2 sentences or split the paragraph according to the main ideas.

Authors: Agreed, introduction has been divided into four paragraphs and largely reshaped, see earlier response to major comment on the introduction.

- Suggest clarifying why the Western Kunlun was chosen (beyond availability of datasets).

Authors: Agreed, the choice of the region of interest is now explicitly in the introduction and addressed in a paragraph (L25-31) :

"The Western Kunlun Mountains are located at the norther edge of the Tibetan Plateau. This is a major glacierized region within the Tarim Interior River Basin (TIRB), where glacial meltwater provides contribution to downstream water resources (Immerzeel et al., 2020). Despite its importance, this region is subject to the largest discrepancies among existing ice thickness datasets, making it an ideal case to illustrate how geometry uncertainty influences modeled glacier runoff. While several studies have investigated the estimation of peak water at regional and global scales (Huss and Hock, 2018; Caro et al., 2025), there are to date, few quantitative assessments of how uncertainties in initial glacier geometry and climate model biases influence its timing and magnitude."

Section 2 (Methods):

- Fig1. Indicate the source of the basemap. Please also indicate that the green area corresponds to the Tarim Interior River Basin

Authors: We added this indications in the caption of Figure 1 : *“Basemap is a mosaic of images from Copernicus Sentinel-2 data generated via sentinel-hub (Sinergise Solutions d.o.o.). Green area on the insert map corresponds to the Tarim Interior River Basin.”*

- "These are located at very high elevations" -> The glaciers...

Authors: Done.

- “*mean elevation change rate of -9.6 m yr−1*” -> Double check the 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.

- “Tarim is recognized as one of the most significant water units in Asia, with a notably high contribution of glacier water yield compared to precipitation in the basin”. This needs some references, it would be ideally to have some numbers to support this

Authors: We added the reference of *(Immerzeel et al., 2020)* at the end of the sentence. Glacier and precipitation contributions are expressed in the referenced paper as components of a supply index, therefore we think it will not be very obvious to the reader to add it in the paper.

- *“This basin stands out as one of the most vulnerable, susceptible to the impacts of climate, political, and socioeconomic changes.*”. This should be part of the introduction to justify why this specific region was chosen beyond ice thickness dataset discrepancies (see also major comment on this)

Authors: Agreed, we addressed that in previous comments on introduction.

- I don’t see the contradiction of having a “however” in the following sentence

Authors: Agreed, we removed “however” from the following sentence at L67-69.

- "with the help of five different models, selected from" - should be "using five different models selected

from"

Authors: Done.

- "One of the common approach in between these models" - should be "One common approach among these models"

Authors: Done.

- "*The first is the consensus for 2019*”. This may be misleading as this is year of the study

Authors: Agreed, this was clarified with the phrasing *“The first is the consensus obtained in 2019”* at L72.

- “*with a total glacier ice volume around 35% higher than the consensus over the same surface area”, “The Himalayan region is indeed one of the most uncertain in terms of ice thickness inversion*”. Again, this could be used to strengthen the Introduction, and Section 2.2 could be more descriptive.

Authors: The first sentence mentioned in the commentary was indeed removed from section 2.2 and added to the introduction as *“Yet large discrepancies persist between these datasets, particularly in High Mountain Asia, where total ice volume estimates differ by up to 35% between products (Farinotti et al., 2019; Millan et al., 2022).”* at L22-24. We think that *“The Himalayan region is indeed one of the most uncertain in terms of ice thickness inversion*” should remain in the section 2.2 to further mention the lack of thickness measurements available to constrain the datasets.

- “Since the consensus model is dated from 2003”. Should not be 2000? As this is the inventory year? Or this region has a different inventory year?

Authors: we changed the phrasing to *“Since the consensus model is roughly dated from 2000”* at L89-90. However we also mentioned (L90-91) that the inventory year is 2010 in this region, but the DEM used for thickness inversion (SRTM) is from 2000-2001, so we deemed FARI19 thicknesses to be centered indeed on the year 2000.

- Good comparison of FARI19 vs MIL22, but the correction for temporal mismatch is only briefly explained. How sensitive are the results to this correction?

Authors: We clarified the procedure for the correction and its impact on MIL22 ice volume following sentences at L91-97 :

“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.”

- OGGM description: This section reads like a model manual. It could be shortened, focusing on features relevant for this study.

Authors: We significantly shortened the section with the new modeling framework (see responses to major comments).

- "The model starts by using the outlines from the RGI" - awkward phrasing

Authors : We changed the phrasing to *“OGGM is an open-source glacier evolution model that uses RGI outlines”* at L101.

- Using the default climate dataset (CRU) is not a justification and considering the impact of the selection of the climate I would expect some justification indicating that this product is performing better than the alternatives available in OGGM.

Authors: see previous response to major comment on climate dataset.

- "Due to the fact that Open Global Glacier Model" - wordy, could be "Because OGGM"

Authors: agreed, changed to “Since OGGM”.

- "we first calculate, from satellite or model-based observation" - should be "observations"

Authors: Done.

- "If the model cannot converge toward a consistent value" - should be "converge on"

Authors: Done.

- Section 2.5: Consider moving the runoff estimation to the OGGM description section as this is a direct output of OGGM, and not something you would derive from.

Authors: Agreed, we move the part *“Indeed, 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 (Fig. S1).”* to the end of the OGGM section, at L116-118.

- The description of how peak water is determined is unclear. It is not evident whether the calculation considers only the year of maximum runoff or also the duration of the “plateau” around peak water. Typically, the peak water year is identified after applying an 11-year running mean to smooth inter-annual variability and highlight long-term trends. Please clarify the procedure and consider adding either a supplementary figure or additional explanation in the text.

Authors: Agreed, we added clarifications on the calculation at L157-162, to explain that it also considers the duration of the plateau : *“While the principle of peak water is often presented as a single maximum value (Huss and Hock, 2015), our simulations often reach a maximum regional runoff "plateau", which remains constant for several years or decades. To measure the extent of this plateau, we empirically chose to define it as the top 10 % of simulated runoff values for SSP5-8.5 (and as the top 5 % for other SSPs). To pick one single date value for peak water timing, we selected the median date of the plateau’s temporal extent. Similarly for the associated quantity of water runoff, we use the average of the plateau’s values..”*

- L165-182. Paragraph too long. Considering shortening by the main idea.

Authors: We tried to shortened the paragraph in the revised version, but this is kind of contradictory with previous comment.

- 2.6 Climate data: Considering moving the historical data to this section to be more self-contained and also move it after the description of the model to know before hand you will be using future projections.

Authors: Agreed, climate data section was moved right after OGGM section and historical data was added in it.

- The bias correction process of the GCM used in missing in section 2.6.

Authors: Agreed, we added mention of the bias correction at L127-128:

“The reanalysis data from 2000-2019 is used as reference climatology for bias correction of the 5 GCMs, following the anomaly method implemented in OGGM”.

- I could not find any mention of the spatial resolution of the climate datasets used. Please provide this information. In addition, it only becomes clear in Fig. 3 that the simulations extend until the year 2300— this should be stated explicitly in the Methods section.

Authors: We added mention of spatial resolution (which is 0.5° for W5E5 and varies between 1.12° and 2.50° for GCMs) of climate data in section 2.4, as well as the simulations’ extent at L22.

Section 3 (Results):

- Figure 2 caption: Suggest rephrasing as: “Timing and runoff at peak water for varying ice volume fractions (a–b) and temperature biases (c–d).” Also specify directly in the axis labels that the multiplying factor applies to initial ice volume.

Authors: Agreed, we changed the caption to the suggested one and modified the axis labels of Fig. 2a and 2b to '*'Initial ice volume multiplying factor''.*

- Figure 3 layout: Consider stacking the panels vertically. The phrase "the glaciers of the set" is unnecessary—if all 160 glaciers in the region are modeled, simply state that.

Authors: Agreed, Figure 3 was changed with vertically stacked panels. We replaced *"the glaciers of the set"* by *''all 160 glaciers''* in the caption.

- Figure 3 variability: The origin of variability in the light lines is unclear. If this reflects glacier altitude or location differences, please clarify; if not relevant, consider removing these lines.

Authors: Figure 3.d was changed and now represents the evolution of temperature and precipitation in the region under different scenarios (multi-GCM mean is shown in bold, shading is the mean ±1 standard deviation of the GCM ensemble).

- Sections 3.1–3.2: Both begin with descriptive sentences that duplicate figure captions. These could be removed to improve conciseness.

Authors : We removed descriptive sentences from beginnings of sections 3.1 and 3.2

- Wording corrections:

- "It appears clearly..." → "It is clear that" / "It clearly appears that."

Authors: Done.

- "Indeed, increasing the total ice volume will not significantly advance" → remove "Indeed."

Authors: Done.

- "We choosed to adjust" → "We chose to adjust."

Authors: Done.

- Literature placement: The statement "It is worth mentioning that previous work (Gao et al., 2010)" would fit better in the Discussion.

Authors: Changed this literature placement and added to the discussion at L273.

- Units: Use mm/yr or mm/day for precipitation, and retain m3/s for runoff, to align with hydrological conventions (Fig. 3 and text).

Authors: Done, units were changed in Figures 2 & 3, and in text as well.

- Supplementary material: Fig. S1 is only mentioned at the end of the Results without discussion; either integrate it into the narrative or remove it.

Authors: Supplementary material was added to the narrative in the revised version at L239-242:

"(...) as indicates the evolution of annual runoff decomposed by its four different contributions (Fig. S1). The latter highlights that under SSP1-2.6 and after 2150, runoff is mostly sustained by snow melt of ice-free areas with FARI19, whereas ice melt on glacier remains the largest component with MIL22. Under SSP5-8.5 annual runoff is largely due to precipitation on ice-free areas with both datasets".

Section 4 (Discussion):

- Having a few subheadings would help to follow the discussion more easily

Authors: Agreed, we added four subheadings to the discussion : *''Peak water dynamics and sensitivity''*, *''Peak water dynamics and sensitivity''*, *''Uncertainties and limitations in ice thickness data''* and ''Model *limitations and implication for large-scale simulations''*.

- "Future work could explore more quantitatively the sensitivity" - should be "explore the sensitivity more quantitatively"

Authors: Done.

- The contrast between SSP1-2.6 and SSP5-8.5 is well described but somewhat repetitive of the Results section.

Authors : We reformulated the discussion section to avoid too much repetition with the results section at L245-272.

- The discussion of satellite missions and data sharing is interesting but feels speculative and less connected to the case study.

Authors : Agreed. We removed the sentences on future satellite missions and data sharing.

- While the study explores sensitivity to ice thickness and temperature, other sources of uncertainty (precipitation projections, model parameters, glacier dynamics, etc) receive limited attention and should be at least mentioned in the discussion.

Authors: Thank you for this comment. We added a sentence mentioning uncertainties related to glacier dynamics, precipitations and model parameters at L280-283.

Technical and Formatting Issues:

- Minor errors ("choosed" → "chose"; "did not used" → "did not use").

Authors: Done

- Several instances of comma splices

Authors: Done

- Inconsistent use of past vs. present tense in methodology sections

Authors: Done

- Elimination of wordy constructions and passive voice where possible

Authors: Done