"Sensitivity analysis of peak water to ice thickness and temperature: A case study in the Western Kunlun Mountains of the Tibetan plateau"

Summary

This study investigates how uncertainties in glacier ice thickness estimates 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.

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*"

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.
    - 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.
    - 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.
    - 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.

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.

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.

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.

6. 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.

Specific Comments

Title and abstract:

- "Western Kunlun Mountains of the Tibetan plateau" - "plateau" should be capitalized as "Plateau"
- Define SSP in the abstract (first mention)
- “*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.
- Avoid using “*significantly”* if there is not a statistical test applied (is repeated four times in the current version).

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.
- "Determining precisely the timing and magnitude" -> "Determining the precise timing"
- 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.
- Suggest clarifying why the Western Kunlun was chosen (beyond availability of datasets).

Section 2 (Methods):

- Fig1. Indicate the source of the basemap. Please also indicate that the green area corresponds to the Tarim Interior River Basin
- "These are located at very high elevations" -> The glaciers...
- "*mean elevation change rate of -9.6 m yr^{-1}*" -> Double check the units
- "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
- "T*his 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)
- I don't see the contradiction of having a "however" in the following sentence
- "with the help of five different models, selected from" - should be "using five different models selected from"
- "One of the common approach in between these models" - should be "One common approach among these models"
- "*The first is the consensus for 2019*". This may be misleading as this is year of the study
- "*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.
- "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?
- Good comparison of FARI19 vs MIL22, but the correction for temporal mismatch is only briefly explained. How sensitive are the results to this correction?
- OGGM description: This section reads like a model manual. It could be shortened, focusing on features relevant for this study.
- "The model starts by using the outlines from the RGI" - awkward phrasing
- 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.
- "Due to the fact that Open Global Glacier Model" - wordy, could be "Because OGGM"
- "we first calculate, from satellite or model-based observation" - should be "observations"
- "If the model cannot converge toward a consistent value" - should be "converge on"
- 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.
- 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 interannual variability and highlight long-term trends. Please clarify the procedure and consider adding either a supplementary figure or additional explanation in the text.
- L165-182. Paragraph too long. Considering shortening by the main idea.
- 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.
- The bias correction process of the GCM used in missing in section 2.6.
- 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.

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.
- 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.
- 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.
- Sections 3.1–3.2: Both begin with descriptive sentences that duplicate figure captions. These could be removed to improve conciseness.
- Wording corrections:
- “It appears clearly…” → “It is clear that” / “It clearly appears that.”
- “Indeed, increasing the total ice volume will not significantly advance” → remove “Indeed.”
- “We choosed to adjust” → “We chose to adjust.”
- Literature placement: The statement “It is worth mentioning that previous work (Gao et al., 2010)” would fit better in the Discussion.
- Units: Use mm/yr or mm/day for precipitation, and retain m^3/s for runoff, to align with hydrological conventions (Fig. 3 and text).
- Supplementary material: Fig. S1 is only mentioned at the end of the Results without discussion; either integrate it into the narrative or remove it.

Section 4 (Discussion):

- Having a few subheadings would help to follow the discussion more easily
- "Future work could explore more quantitatively the sensitivity" - should be "explore the sensitivity more quantitatively"
- The contrast between SSP1-2.6 and SSP5-8.5 is well described but somewhat repetitive of the Results section.
- The discussion of satellite missions and data sharing is interesting but feels speculative and less connected to the case study.
- 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.

Technical and Formatting Issues:

- Minor errors (“choosed” → “chose”; “did not used” → “did not use”).
- Several instances of comma splices
- Inconsistent use of past vs. present tense in methodology sections
- Elimination of wordy constructions and passive voice where possible