Global vs local glacier modelling: a comparison in the Tien Shan

Lander VAN TRICHT, Harry ZEKOLLARI, Matthias HUSS, Daniel FARINOTTI, Philippe HUYBRECHTS

General comments

This paper produces with a lot of heavily calibrated model outputs some comparisons between local glacier-specific models and global-scale models. The main goal of the research questions is very unclear in relation to a real scientific outcome of the paper. This results from shortcomings in acknowledging and dealing with known processes and problems in this scientific field, ranging from highly parameterized models, over big uncertainties in climatic forcing reanalysis and future scenarios, to regional and global glacier mass balance estimation uncertainties. As a consequence, it is confusing to see some results presented in a publication of a journal addressing experts in glaciology, like 'The Cryosphere', with statements mentioned in the abstract section like '*Our findings thus suggest that when modelling small to medium-sized glaciers the emphasis should be on having a reliable reconstruction of the glacier geometry rather than focusing on a detailed representation of ice flow and mass balance processes'*. Here, the overgeneralized interpretation that in my opinion is insufficiently addressing these issues reads as if the authors' main goal is to explain finally the glaciological community that it is important to observe glacier geometry and volume. I think it is well known that the existing ice geometry and volume is fundamental for doing any long-term observations on glaciers and estimating their future state and we do not need a paper telling us this simple fact.

The "Models and Data" section is extremely vague and lacking important information, leaving it up to the reader gather these crucial pieces of information in a large number of other papers; sometimes this also generates confusion in relation to **which version of a data/model** or **which type of calibration** is used. Some examples of missing information is mandatory to be presented within the manuscript and/or in a supplement:

- **Which are the time periods** (for each glacier) over which the models are calibrated against *in situ* measurements? **How many *in situ* measurements** (ll. 135-137) are actually used for each glacier? This could also be a collection of figures from previous papers, but in a modeling study like this it must be shown. It would for example show that accumulation measurements are sparse or even lacking for some glaciers. Accumulation rates are not only uncertain but also potentially evolving over long-term periods, is this somehow considered?
- Calibration of GloGEMFlow against the geodetic data of Hugonnet (ll. 150-154): the authors mention a calibration following Huss and Hock (2015), but they calibrate their parameters (precipitation gradient, degree-day factors and air temperatures) only against the bias. **Calibration of three or more parameters against a single number (the bias) is not very meaningful,** as it is always possible to adjust the model output until it matches the reference (geodetic) value. What about the spatial residuals (RMS)? The model could potentially be over-estimating mass balance by +999 m w.e. at certain locations, and under-estimating by -999 elsewhere! Thus, the **full results of calibration must be shown**, including residuals w.r.t. the single measurements. The modeled mass balance must be made available for download – that is the only way to see that the modeled mass balance is realistic.

The paper claims a main role of initial thickness in controlling the modeled glacier evolution, as opposed to the mass balance forcing and the choice of ice flow model complexity.

- ll. 373-375, *'The findings reveal that the initial ice thickness (volume) reconstruction strongly dictates the glacier's future evolution, with larger initial volumes resulting in more ice remaining by 2050 and the end of the 21st century (Figure 7).'*.

  However, the spreads in Fig. 3 look at least as large (and more temporally variable) as the spreads in Fig. 7 (excluding the red Millan line, which is reflecting a very large – probably wrong – ice thickness estimate). This would contradict the authors' statement: **the role of mass balance and ice dynamics appears to have the same order of magnitude as the initial thickness.** It would be useful to have a quantitative metric (e.g. R^2 or similar) of the relative importance of mass balance, ice dynamics and ice thickness (e.g., some summary number computed from Figure 10).

- ll. 461-462, *'One of the key outcomes of this study is the pronounced sensitivity of the projected future glacier volume to the selection of the ice thickness reconstruction employed for model initialisation (Figure 10).'*

  Fig. 10 in the sense as it is interpreted in the paper. Only for Grigoriev (purple diamonds in Fig. 10d) the difference is very high, and most of the difference likely comes from the (very wild) thickness estimate of Millan. Fig. 10 is showing a fundamental role of ice thickness (versus choice of mass balance and ice flow model). It would be useful to show an actual number summarizing the relative importance - the magnitudes in panels (a) and (b) appear to be similar to panel (d). Also, obviously the importance of initial ice thickness will be greatest for the near-future (as the glaciers need time to adapt) and decrease over time. In Fig. 10, for 2075 and 2100 the dots are not really higher in panel (d) than in panels (a-c), negating the main conclusion of the manuscript. Thus, initial ice thickness is maybe equally important as mass balance and ice flow, but not "mostly [important]" (l. 502).

- What about the importance of the future climate? Do the mentioned climate models (ll. 155-158) have any clue about local weather and its trends? The authors mention a "bias correction" procedure (l. 160) performed *'in order to align the output of the GCMs with observed climate data'.* How big of a change is that? How does GCM precipitation compare to measurements, is it any good? One could argue that in fact the input weather can be the largest uncertainty for the future, far exceeding that of thickness: especially in a regional model, the (virtually unknown) spatial variability of precipitation (on the ~1 km to ~100 km scale) could add vast amounts of variability in the glacier evolution, not captured by any of the used models. Thus **all the models would be similar (in part) because they all share a common lack of a key variability in the input.** Furthermore, the authors mention that the GCMs are aligned with observations (l.160) where it is assumed that the ERA5 dataset should represent the observations. However, calling ERA5 an observational dataset is incorrect and unsubstantiated as ERA5 has a lot of uncertainty and does not necessarily represent real situations, especially not in Central Asia (e.g. Zandler et al. 2019, Guo et al. 2021, Barandun and Pohl, 2023). Or are meteorological station data meant? This should be made clearer and which datasets the GCM runs are calibrated against. Furthermore, the reference as of why the specific CMIP6 climate model runs were selected seems inaccurate. The referenced paper seems to talk about hot biases but is not presenting a comparison of model runs that would then lead to the selection of the specific model runs. This would also need some justification/clarification or at least/best a comparison of these model runs so that the reader can get an impression of how different the model runs are. Optimally, this also includes the data after bias calibration.

Impact of mass balance:

The authors are comparing the results of mass balance forcing computed using the simple SEB of Oerlemans (2001) and using the degree-day approach of GloGEMFlow. **The Oerlemans SEB model** is calibrated against stakes and snow pits, but it **is still just a highly simplified melt model!** For example, the **effect of sublimation** – significant both as an energy sink and a mass balance component across dry Central Asia – **is ignored by both models**. Using a full energy balance model (like the EBFM or COSIPY) as "local" model would be more meaningful, as it would really show the impact of including vs excluding the actual mass balance processes. Also, for a better understanding of the model comparison, **it would be important to at least provide a list of the processes which are included in each model** – what about the future evolution of albedo and of debris cover of the glaciers? I suspect it might be included in the simple, regional GloGEMFlow but not in the local Oerlemans SEB model?

Impact of ice flow:

- Surely the *'3-dimensional higher-order thermomechanical ice-flow model (3D HO-model)'* does not have only two parameters *('the enhancement factor and the basal sliding parameter'*, l. 183)? The manuscript shows that it is possible to run the 3D HO-model and obtain a result which is similar to that of the simpler GloGEMFlow; but I expect **it is also possible to obtain a very different result while still using reasonable parameter values:** what about the impact on future glacier evolution of the other parameters in the model? **A sensitivity study is expected here**, or, if already performed in Van Tricht and Huybrechts (2023, still a preprint!), its main results must be summarized here. Unless the authors can claim that all other parameters have a negligible impact on future glacier evolution, a reasonable uncertainty estimate contributed by the other parameters must also be considered and is potentially significant (e.g. within Fig. 10).

- Can any of the used models simulate dynamic glacier instabilities? **A large fraction of glaciers in Central Asia are known to be of surge type**, while the 6 selected glaciers (to the best of my knowledge) are currently not. The accuracy of a model on regional scale will certainly be affected by the presence of surging glaciers, which can radically alter the hypsography, topography and surface morphology of a glacier, leading to large variations in mass balance not directly linked to the climate. This would limit the validity of the study's results to small, stable glaciers. It is also important to consider the temporal evolution of such instabilities – for instance:
    - Increasingly prevalent glacier instabilities, possibly linked to a large-scale cold-polythermal-temperate transition.
    - Formerly unstable glaciers, made stable (no longer surge-type) by the changed morphology following retreat.

The main author uses mainly self-citations of his studies, which he probably did during his PhD. However, all the older studies from well-known Central Asian colleagues, published in the past like the studies of Glazirin, Aizen, Dyurgerov, Dikih, etc. are nearly not mentioned. Even if the scientist is not able to read old Russian literature, it is necessary to find a way to include the findings of this literature for many aspects of glacier evolution in the region were already assessed and reported.

Overall, the outlined points render this paper very vague, and the reader is left guessing what exactly is being compared to what.

Specific comments

- II. 144, This citation is an Egusphere abstract, where the 'comprehensive information' is not findable as the information in the abstract is very limited.
- II. 149-150, ERA-5 reanalysis in Central Asia should not be taken as a gold standard as recent studies have shown (e.g. Zandler et al. 2019, Guo et al. 2021, Barandun and Pohl, 2023,).
- II. Table 1: Description of ice thickness datasets. Why are the measurements of the ice volume not shown in Table 1? Please give some more details about GPR measurements and how they have been inter-and extrapolated.
- II. Table 1: The paper uses the results from Millan et al., 2022. Why do the authors take this paper as a comparison, if they obviously do **not agree** with the ice thickness results presented in Millan et al. 2022. The authors already have the results from Farinotti et al. 2019a and all the results from Model 1 to Model 4.
- ll. 228-229, The observation on Kara-Batkak suggest a shorter response time and faster approach to equilibrium. Please show the calculated response time for each glacier.
- II. 232, According to the interpretation of your figure (Fig. 3), this is not true. See Ashu-Tor and other glaciers in Figure 3. You mention that regional mass balance and local mass balance are the same. This is in my view not correct and not supported in the figure, e.g. certain glaciers like Ashu-Tor in Figure 3, where the difference of the volume at 0.2 is around 20 years.
- II. 233-235, If you compare a local to a regional mass balance model, then the focus should still be on the differences of the individual glaciers and not on an aggregated ice volume. This is because the differences at individual glaciers can actually reveal similar or different behaviour, whereas the aggregated ice volume is affected anyways from the calibration.
- II. 298-301, Give more details exactly for this calibration process?
- II.303-305, This sentence is not conclusive. What means *'…peculiar local characteristics that are difficult to capture'.* Please be more specific.
- II. 329, Surface elevation from SRTM is well known to have large uncertainties because of penetration of the radar waves into the firn area.

- ll. 371-375, This result (*'the initial ice thickness (volume) reconstruction strongly dictates the glacier's future evolution')* is expected when comparing glaciers with significantly different ice volumes regarding their disappearance time. But glacier evolution is more than disappearance time and this statement seems unsubstantiated with the lack of information and considerations about process descriptions and uncertainties in climate forcing as outlined in the previous comments.
- ll. 427-428, What are the extrapolation techniques, which are mentioned in the caption of Figure 9?
- ll. 438-440, This statement clearly shows that not only geodetic mass balance will deliver the best results, but instead, also local data is fundamental to understand the local differences which are normally **not presented in models like GloGEM**. Inclusion of special processes such as long-term change of temperature conditions within glaciers, changes in runoff based on changes in temperature or mass balance, changes in flow regimes, sublimation effects, long-term changes in pore conditions of accumulation areas are often leading to changes, which are not covered at all in these models but are significantly changing the behaviour of the reaction of a glacier. The authors should be more careful with their statements. Simplification is sometimes ok but here the authors oversimplify.
- ll. 443-444, This is still only an assumption if you calibrate your parameters in a way that both models fit. However, this is not showing all the uncertainties if you would change your parameters particularly in the complex model.
- ll. 461-467, This is commonly accepted and taught at universities. There is no need to have a paper telling specialists in this research field that existing glacier volume is fundamental for studying glacier change.
- ll. 480-481, The author writes: '*highlights that a global-scale flowline model is capable of accurately simulating glacier dynamics and evolution'*. This quite generic claim is just wrong. "[...] accurately simulating glacier dynamics [...]" would include the simulation of observed processes such as seasonal velocity changes and glacier surges.

References

Barandun, M. and Pohl, E., 2023. Central Asia's spatiotemporal glacier response ambiguity due to data inconsistencies and regional simplifications. The Cryosphere, 17(3): 1343-1371.

Guo, H., Bao, A., Chen, T., Zheng, G., Wang, Y., Jiang, L. and De Maeyer, P., 2021. Assessment of CMIP6 in simulating precipitation over arid Central Asia. Atmospheric Research, 252: 105451.

Zandler, H., Haag, I. and Samimi, C., 2019. Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions. Scientific Reports, 9(1): 15118.