Second Referees comments

Review: Predicting the Risk of Glacial Lake Outburst Floods in Karakorum

Based on remote sensing as well as field data the authors aim to support their hypothesis that GLOFs can be anticipated from remotely sensed surge velocities as well as knowledge of lake geometries and lake expansion. They find a link between glacier velocity and lake volumes, and subsequently a relation between lake volume (or depth) and the risk of disastrous drainage. They do this based on 3 well known case studies in the Karakoram and in the end expand this to the global scale for the Discussion. The manuscript is in general well written and language is clear. Literature has to be revised at times and is not always completely appropriate with respect to basic conclusions in the Introduction, that are at the basis of the arguments then later followed. More problematic is that the hypothesis isn't clearly laid out at the start and hence the presentation of Results becomes a bit confusing, which would require some restructuring. Additionally, there is also a general lack of scrutiny on error margins especially for field measurements, which has implications on interpretation of results as I outline below, as these uncertainties propagate into your results. As I describe further down as well (and as the authors themselves admit throughout the text), 'prediction' itself isn't possible with this approach (if it will ever be) and hence the title is in my view misleading, and I would strongly consider to revise that.

Thank you for noting that the manuscript is generally well-written. However, as you suggest, some restructuring improves the flow of the arguments, and we have taken up your suggestions. As for the title, we have changed 'Predicting' to 'Anticipating,' which is more appropriate.

Most importantly however, I think that there remains lack of clarity on why surge velocities need to be part of the predictive mechanism and even the initial hypothesis that links them to the thickness of the ice dam. This oversimplifies a complex problem and the physical reasoning escapes me. This then also leads to strong interpretation of results, that are not as clear cut as the text makes them seem. All this in mind I think the manuscript needs a careful major rehaul before it is suitable for The Cryosphere (while the importance of the topic itself and the general approach I do think merit consideration for this Journal). I explain this in more detail with a number of major concerns below, followed by quite a short list of minor issues encountered throughout the text.

Thank you for raising the issue of glacier surging as a mechanism for predicting the formation of glacial lakes and subsequent GLOFs. We have ensured that the revised text does not suggest that glacier surges can be used to predict the formation of lakes. We had already noted that the physics of the problem is complex, and a complete understanding of the processes that enable the modeling of the controls on GLOFs is not imminent. Rather, some simplifying assumptions related to empiricisms would be valuable in providing guidance with reference to anticipating GLOFs. We have revised the text to ensure this is stated clearly.

Major comments including conceptual/methodological issues:

Title: I am concerned that 'predicting the risk' is a very strong term as this will unlikely ever work to a degree that people understand what 'predicting' means and we shouldn't suggest we can predict it (you say that yourself in the Discussion). You could change it to what you use later to ‘Enhance the predictive capabilities for Glacial lake outburst risk in the Karakoram’

We have changed ‘Predicting’ to ‘Anticipating,’ as the latter word better relates to the requirement to provide timely warnings.

L14: Not all glacier snouts produce lakes and there is no evidence that proglacial lakes lead to advancements or surges! Please amend.

Thank you for noting this error, which was due to translation into English. The text has been revised.

L15: Recent research is conflicted on this topic and maybe actually showing the opposite - that ice dammed lakes drain less frequently (Veh et al., 2023). I would therefore remain very cautious with such kind of assertions to set the scene for your research.

The argument at this point relates to the Karakorum. However, we have modified the text to provide a caveat that GLOFs may decrease as global ice decreases.

L32: While mass gain has indeed been observed this trend is already over since approximately a decade – see (Jackson et al., 2023) for a number of studies referring to this. It would be prudent to note here that by now we consider this anomalous process to be over.

Thank you for the suggestion to provide a fuller picture of glacier behaviour, including outside of the period of time we are considering. We have amended the text, adding the Jackson reference, amongst others.

L34: The statement that there is a link between the anomalous mass balance and surges is spurious – what do you base this on (or which study specifically)? Also there has to my knowledge been no evidence for an increase in surge frequency (but only our ability to detect them).

It is indeed challenging to establish a definitive correlation between glacier mass balance and glacier surging due to the complexity of glacial dynamics and the multifaceted factors influencing both phenomena. However, it is overly strong to label the relationship as ‘spurious.’ Several studies provide evidence suggesting a link between mass balance and surge events. Bazai et al. (2021; 2022) have demonstrated that surges often follow periods of increased mass balance. Specifically, their research highlights that a higher mass balance can result in increased glacier velocity and surging activity. For example, Bazai et al. (2021) showed that the Shishper and Chilinji glaciers exhibited significant mass transfer from the accumulation zone to the ablation zone, leading to surges between 1999 and 2002.

Heïdi Sevestre (2015) and Bhambri (2017) also support this notion, indicating that surges are often associated with changes in mass balance. Their studies suggest that an increase in mass balance contributes to glacier thickening and enhanced basal sliding, which can trigger surging behaviour. Sevestre (2015) particularly emphasizes the role of mass redistribution within the glacier in surge dynamics. Additionally, Bhambri et al. (2019) observed that recent surges of the Shishper and

Chilinji glaciers involved the transfer of a higher mass compared to previous surges, indicating a possible correlation between increased mass balance and surge activity. While there is no conclusive evidence for an increase in surge frequency, improvements in detection capabilities might account for the perceived increase in surge occurrences. Enhanced satellite imagery and remote sensing techniques have made it easier to monitor and identify surge events that might have gone unnoticed in the past. This latter point is now noted in the Discussion.

L40: All three studies cited here for 'glacier avalanche increase' are based on few specific events, not trends (since so far we do not have long records). Hence you cannot speak of an increase of these events based on this evidence or need to adapt language to caution our lack of records so far.

This relevant text has been moved and now refers to an apparent increase. The issue of the lack of long-term records is better referred to within the Discussion, wherein we have acknowledged the lack of long-term records.

L41: I am not sure what you refer to when you speak of 'positive variation in regional climate feedback' with respect to lakes. You mean 'more lakes leading to more mass loss leading to yet more lakes'? Then spell this out, and you'd need to cite studies that show an increase in lakes leading to more mass loss (e.g. (Zhang et al., 2023)

Thank you for noting this error in English expression. The text has been completely rewritten.

L54ff: While (Carrivick & Tweed, 2016) is of course a global compilation, there is since a more recent overview globally (Veh et al., 2022) as well as for HMA specifically (Shrestha et al., 2023), both showing a much larger number of events and providing more accurate statistics for the numbers you are quoting here. Since we need to be mindful of the advancement of research, it would be prudent to refer to these updated studies here.

Thank you for the additional references that have been added to the revised text.

L96: 'Data' is generally not a subsection of Methodology but at least on equal footing. I would hence rename this section 2 to 'Data and Methods'. Also, you lack an introduction of your specific field locations and why you chose those but rather jump directly on the three lakes in L98ff. It doesn't require much, but an introduction of your field sites here briefly is required, possibly with an overview map as Figure 1 or an inset elsewhere. L146f actually is a kind of introduction in this regard that should be placed earlier.

We have renamed section 2 (now section 3) as 'Data and Methods' and sub-section 3.1 as 'Remote Sensing data'. A new section 2 has been added to the revised text to introduce the study area.

L167: Please explain how you 'estimated' this length.

We have added text to the revised method to indicate that lengths were measured using GIS 3D interpolation to UAV data and ground GPS points. These GPS points were then overlaid with the GIS 3D interpolation lines for enhanced accuracy.

L199: What are ‘other’ dimensions you get to with this method, please explain.

A reference to Fig 3b has been added to the text wherein the other dimensions are indicated.

L193ff and Figure 2: I am a bit confused by this section leading into results. You introduce the tetrahedral and pentrahedral shape of the dam, but never refer to the second but just the first. Why? You then also present results already in Figure 2b, even though we are still in Methods (especially your very last part on the issue of overestimated h). It is important to keep this apart and not mix the two sections.

Thank you for highlighting the confusion in the text at this point. We have restructured the text to ensure that the pentahedral approximation is detailed in the Method rather than just being introduced within the Results section.

Finally, I think it would be important to rather use the space in methods how you were able to determine the D, E and Z (and Y, which is wrongly denoted twice as Y1 in the figure but should be Y1 and Y2!) in the field. How accurate are your GPS measurements? How accurately are you able to determine the position of the dam crest and base? The errors that are quite normal from such field measurements will then propagate into your volume estimates, which in a complex equation like equation 1 you show here, can become complex themselves. It is important to note them however to judge your comparisons you then make against volume estimates from the DEMs.

The two lengths Y are variable, but in Fig. 3b, they were drawn as both equal in lengths such that both lengths are denoted as Y1. However, as Y can vary in both cases, we have changed the notation in Fig. 3b and made text changes to make the point clear. We have added text to indicate the maximum errors that might be associated with measurements obtained from remote sensing images or the estimation of lake depth from DEMs. For lakes typically 1 km in length, the errors are <3% and <14%, respectively. NB: The uncertainty in estimating the lake volumes using remote sensing, DEM, and UAV is already given in Table 1.

In turn, you also need to specify the uncertainty of your UAV derived DEM, which in turn translates into an error range for your volume estimate.

The error in the DEMs used was already reported in the original submission. The errors in lake volume determination were addressed in the prior reply to the reviewer.

L256f: You here come up with the physical hypothesis you try to prove after you present the results. That is confusing and should be the other way around, the hypothesis should be introduced in the Introduction already, when you set out to argue why you look at the datasets you choose.

We have broken the reviewer's comments down into sections to make it easier for the reader to follow. It was not our intention to introduce a hypothesis at this point rather than the text the reviewer refers to is an interpretation of the data within Fig. 4. Nonetheless, it would be better if hypotheses underpinning this interpretation were stated upfront, so we have introduced two hypotheses in the Introduction section, as follows:

> "Consequently, herein, we explore the main hypothesis that lake volume is related to glacier velocity. As lake volume can dictate the characteristics of a GLOF, a secondary hypothesis was addressed that ice-dammed lakes can exhibit geometries similar to regular geometric shapes, such that in the absence of a detailed lake volume date, lake volumes might be estimated from geometric consideration. "

I then however do not quite follow your hypothesis. You argue that ice thickness at the lake dam decreases as velocities of the surge peak. But lake dams are in the receiving zone of the surging glacier (mostly, definitely in all the three cases here), where velocity peaks coincide with thickening of the tongue. Ice thins in the quiescent phase after the surge (where again we see many GLOFs of course in all three cases, for many years to come). This leads me then to the main concern I have with the prediction plan based on just velocity here –

As the review infers, sites for ice-dammed lakes are formed during glacier surges when ice thickness increases. However, at that time, lake volumes were generally low as there was insufficient time for meltwater to build up, and crevasses allowed for ready drainage. As the surge reduces, ice velocities reduce, and lake volume increases due to meltwater increments and the closure of crevasses. We are not suggesting that lake volume or GLOFs can be predicted using only one variable – glacier surge speed – only that there is a potential correlation that provides additional insight into lake formation and draining.

I would argue that lakes fill following closure of the subglacial drainage, from crevassing (as you argue) but also with changing lake water temperatures (depending on inflow temperatures and atmospheric temperatures and snow melt and glacier melt forcing. These variables are of course very hard to constrain, but with your approach you are packing these complex processes into glacier velocity alone – which is sometimes changing without any perceptible effect on lake properties as well. This way you gloss over other potential drivers, which could become a problem when wanting to be truly predictive. I think this is finally all reflected in your results.

We agree that changing water temperature may be a factor in lake growth, and the controls for these factors are complex. We make no attempt to suggest that these complexities are 'packed' into just one variable; instead, the speed in the velocity versus lake volume data actually reflects the role of these unquantified variables. We have not 'glossed' over these other controls; many are noted in the Introduction. Rather, we are trying to present the evidence that one readily measurable control – ice velocity – may have an inverse relationship with lake volume. The importance of other controls was already acknowledged in the first paragraph of the Discussion.

Naturally there are GLOFs always after surges (Figure 3), but that just follows from the damming.

GLOFs do not always occur after surges, and surely, it is not satisfactory to state that they occur 'just' due to ice damming. Such a conclusion negates any attempt to understand the controlling processes better.

I can't see any clear relation between speed up and occurrence just from these results. In Figure 4 you finally of course have a very large spread. For Shisper alone there seems to be no change of volume with velocity (I suspect here it is simply constrained by topography also since around later drainage events the surge had stopped completely), at Kyagar the volumes varies widely with very similar velocities and I see no geometric relationship at all. It finally is a deficiency here that you do not show the uncertainty bars around velocities (especially since these mean values spread quite a lot during the surge itself on top of the error introduced from the SfM approach) or the volumes, which would further put your relations into context (and put the geometric relationship in question).

In presenting these results, we have taken a cautious approach, stating in the original text that the results are 'clearly not definitive'. In the Discussion, we acknowledged that the data scatter means that the trend line is not statistically significant. Nonetheless, the trend line is always negative, and we have suggested that further research might seek to better populate diagrams similar to Fig. 5 (formerly Fig. 4) to examine this relationship further.

L265ff: Maybe this is just me, but I would find it more easy to follow if you first present results on your ability to represent volumes of lakes with trigonometric considerations and then follow this by the much more complex relation with surge velocities. As it stands now you jump between one and the other, which leaves it quite hard for the reader to follow the final reasoning. When it comes to volume estimates you furthermore now have results spread between Figure 2 and Figure 5, with a completely different story in between. It would be advisable to bring this together to then come to a clear conclusion simply on your ability to estimate volumes accurately. Once that is done, it will also be more reliable to interprete your results from Figure 4. Relating to your volume estimates I am finally surprised that you do not refer to earlier considerations in this direction, e.g. taken in (Cook & Quincey, 2015), who summarize multiple attempts in literature in this direction.

Thank you for these observations. We have restructured the Results section to include sub-headings separating the surge velocity considerations from the trigonometric considerations. We have added the reference to Cook & Quincey, 2015.

L316ff and Figure 6: You have used multiple data to assess volumes/lake dimensions but do not specify what your lake geometries are based here now. They look very detailed, is this all taken from UAV generated DEMs and field measurements (for depth)?

The methodology was defined in the 'Data and Method' section, to which we have now added the uncertainty in both length measurements and the derivation of lake volumes.

It needs to be absolutely clear what data finally leads you to what conclusion, especially when you draw up ranges that may be used later for breach assessments but whose accuracy change widely depending on what data you base them on. I am also not entirely sure, how to interpret the red lines in Figure 6 – are these the actual depths/volumes at the respective GLOF events as presented in Table 1? Make that clear in the caption to Figure 6.

The red lines connect the elevation of the lake level at the closest time to the GLOF with the lake's volume at this time. To make this clear, text has been added to Fig. 7 (previously Fig. 6).

Also you then show relations for times when ice thickness wasn't well known right (e.g. older Kyagar events)? Doesn't you assumption of a constantly same geometry become weaker away from your UAV-DEM date? You have deposition, sediment drainage as well as different dam heights every time. This will affect the final calculation of n/pressure and hence I suspect give you much more spread in 6d, which to me now seems surprisingly well lined up for Khurdopin and Kyagar.

Lake volume data are in Table 1 and, as shown in Fig. 7 (previously Fig. 6), are obtained from survey data and not from geometric approximation.

L328: Following from above to your conclusion that 500kPa is the threshold – this is simply 50 m of water table equivalent. Isn't that simply confirming that 'a lot of water is needed to cause a splash'? Below 50 m in most cases we will have small GLOFs (which have occurred often unrecorded in all lakes and one could argue aren't really GLOFs but simply high flow drainage events) while when you get to 50 m you slowly get to volumes that cause considerable flow. But why would that be specific to surging glaciers or the region? Again this will then depend on ice dam properties (temperature, thickness) which bring in a number of other unknowns and can't just be summarized with a single value. Even your results show that for Shsiper it may be >1000 kPA, while for Khurdopin 700 kPA, quite big differences.

The pressure on the ice dam face derives simply from the water depth at the dam face. We have not sought to determine that 'a lot of water is needed to cause a splash' as the reviewer suggests. Instead, in the initial submission, as in the revision, we have sought to provide some indication of the relative depth (new Fig. 7d) at which GLOFs occur, such that warnings might be provided downstream in a timely fashion. There is nothing in the literature to indicate what might be considered a minimum sudden discharge to be labelled a GLOF. We have noted that a few of the recorded small discharges occurred well below a potential threshold of non-dimensional depth above which substantial GLOFs have occurred in the Karakorum (new Fig 7d). In the text, we have made no case that such behaviour is specific to the surging glaciers in the region, so we are unsure why the referee raises this point. Rather, as the referee acknowledges below, we have extended the analysis to include other glaciers worldwide.

L340f: It is nice that at this stage you manage to pull in examples from elsewhere to support your volume/threshold theory. This of course gives your approach some strength. This leads however to two observations:

- Maybe it would be prudent to leave away the surge aspect from your study altogether, to me the velocity story doesn't quite hold up and also does not have a clear connection to the critical threshold. It even further weakens your prediction capabilities as making the step from 'surge velocity u' to 'critical threshold n' becomes spurious and doesn't hold as a clear indication (i.e. you can't say 'if Shisper exceeds a velocity of x m/d a GLOF may be more likely').

The reviewer makes some useful points here. However, we have not tried to relate the surge speed directly to the critical depth, so we do not understand why the reviewer would suggest that we have. We argue cautiously that there might be a relationship between surge speed and lake volume and indicate that further work is required to test this idea.

- As you note yourself in L391f, you can't 'predict' above a certain threshold (hence also my concern with the title), you can just indicate higher likelihood of catastrophic drainage. This itself is useful but the recommendation to always report that as imminent threat has its drawbacks, as people will become less sensitive to warnings if they are reported often without any resulting events. Hence I rather suggest to use this as a general support in risk assessment procedures (paired with other means).

We accept that 'predict' was too strong a word and have changed it to 'anticipate' in the title and in the main text. We acknowledge that human nature can mean that warnings are not considered by the general public, but the reviewer seems to be implying that there is no point in issuing warnings. Whether the imminent risk is considered within risk assessment procedures or used to directly warn people would be based on specific countries' own procedures.

- In L380 you get back to the volume estimation – I agree that for a first order measure this is useful, but to make this a considerable improvement over simply seeing lake area from space (much easier), the reader requires much more scrutiny on your volume estimates. How accurate are the depth/dam dimension variables you have to come to good volume estimates here and how does that propagate to your n computation to make it robust? This is currently missing from the text and with the data you show from lakes elsewhere could be expanded upon for a thorough Discussion.

The resolution of the remote sensing and the accuracy of depth measurements were quantified in the original submission. In the revision, we have added text to clarify the uncertainty in lake volume estimates, which is already stated in Table 1. It is well known that linear measurement errors are additive or can cancel out. However, we have added text to the Discussion to emphasize the fact that mensuration errors can now be quite small given modern techniques.

L410f: The Conclusion as it stands is too short. It also hinges much weight on the linear regression against one means to estimate volume, while I do not think that was the initial aim of the paper (prediction and GLOFs rather than lake volumes). This suggests that the general aim needs sharpening and the conclusion should also converge on that. From your Conclusions I would also like to see to what degree your field data was essential in this process as this clearly impacts the potential to upscale. For this a more thorough descriptions of said field data including their uncertainties is however required (even if parts of this has been presented in other publications).

Thank you for these helpful comments. We have restructured the Conclusions, putting the geometric estimation conclusions after the comments on the lake draw-down. We have added a comment on the value of field data. The uncertainties of the field data are detailed in the Method.

Minor:

L1: While 'Karakorum' spelling is of course possible, I would consider going with the standard in English in literature by now and go for 'Karakoram'

We have standardized spelling to 'Karakoram'.

L40: 'has increased'

Text altered.

L92: '…research foci described in this study aims to leverage …'

Text altered.

L142: 'measured'

Corrected.

L147: 'have occurred'

Text altered.

L150: 'Sentinel 2', capitalized

Corrected.

L165: 'the empty lake'

Text corrected.

L197: 'outlines'

Corrected.

L198: remove first 'to'

Corrected.

L205: 'a value which can be …'

Text altered.

L240/Figure 3: typo in the figure – 'Kayager'. Also what are the numbers in 3d? I assume days after 1/1/1990? You would need to say that but it is also a very hard way of reading your data. Dates would be much more useful.

Kyager has been corrected, and the figure panel has been redrawn. However, adding dates would cramp the lettering on the x-axis.

References:

Thank you for the additional references which we have cited in the revised manuscript.