** Please find our response highlighted in blue to the reviewer's comments.

This study aims at modelling potential future GLOFs from 21 Nepalese glacial lakes identified as potentially dangerous in previous study by Bajracharya et al. (2020). This study uses regression model to estimate plausible ranges of mean lake depths and so lake volumes and peak discharges. Further, exposed elements within the limits of modelled inundation areas are mapped and potential damage is assessed. Undoubtedly, such results are of value for disaster risk management authorities and I appreciate holistic approach going beyond GLOF modelling itself. While I'm very much in favor of GLOF hazard / risk assessment studies that consider a range of scenarios and I really appreciate the amount of work done, I have some thoughts for further improvements in the hazard assessment part.

Response: Thank you very much for the reviewer's overall positive feedback.

Similar approach has been developed and employed by Veh et al. (2020) across the whole Himalaya. While thousands of lakes were considered in their study and the approach was suitable, I would expect bit more site-specific input data in case study of 21 lakes. For instance, moraine dam geometry (height and width) could be used to estimate max. breach depth and so max. flood volume that could differ substantially compared to the assumption of 100% lake volume release which is: (i) in general very unlikely for large lakes; and (ii) physically not even possible in many cases (when the height of a damming moraine is less than max. depth of the lake or the geometry of the dam is very flat).

Response: I believe the reviewer is referring to Veh & Walz (2020), Proceedings of the National Academy of Sciences, 117(2), 907-912. We have indeed adopted the same approach developed and employed by Veh & Walz (2020). However, in addition to hazard evaluation that Veh & Walz (2020) focused on, our focus extends to exposure and impact assessments of GLOFs. To achieve this, our study relies on several key components. These include remote sensing techniques for accurate glacial lake area delineation, Bayesian regression models for deriving relationships between lake water depth and peak discharge (Veh & Walz, 2020), state-of-the-art flood modelling technology supported by parallelised high-performance computing, and object-based GLOF exposure and impact evaluation using open-source data. Not covered in Veh & Walz (2020), our study conducts high-resolution inundation simulations for various GLOF scenarios using flood modelling technology. This facilitates subsequent object-based assessments of GLOF exposures and impacts. The results provide a comprehensive and detailed evaluation of potential exposures and impacts stemming from these PDGLs. While much of the prior work has focused on the initial step of GLOF risk evaluation, specifically hazard assessment for glacial lakes, like Veh & Walz (2020), our study advances the field by addressing the second stage, which involves exposure and impact evaluation. The insights gained from this study can empower authorities not only with knowledge of where threats exist but also with an understanding of the expected magnitude and location of impacts.

I agree that incorporating additional data, such as moraine dam geometry (height and width), would enhance the estimation of maximum breach depth and flow volume. However, it is challenging to obtain these data for all 21 lakes. Therefore, to encompass all potential glacial lake outburst scenarios, we have also considered less severe conditions. Specifically, scenarios where 25%, 50%, and 75% of the lake water volume is released have been examined. The outcomes of these less severe scenarios have been compared to the worst-case conditions, where 100% of the lake water is released, as discussed in Section 4.3.1.

Further, the procedure of random selection of 1000 scenarios and subsequent calculation of inundation frequencies and median of max. inundation depths for each lake is not appropriate

because these scenarios are not equally probable. Reflecting on frequency-magnitude relationships of common GLOF triggers (various mass movements), low to moderate magnitude GLOFs are more frequent and more likely while extreme GLOFs are rare and less likely. Instead of selecting the scenarios randomly, my suggestion is to select them on purpose to cover the full range, with assigned weights (or ideally probabilities).

Response: Thank you so much for pointing out this question. When taking the median of the maximum values, we default to assuming each scenario is equally probable. However, as highlighted in the reviewer's comments, this assumption is not correct. Therefore, we assigned weights to each scenario based on probabilities. The weight of each scenario is determined by its occurrence probability, i.e., the proportion of times its peak discharge does not exceed that of other scenarios relative to the total number of scenarios. A smaller proportion indicates a lower likelihood of occurrence, while a larger proportion indicates a higher likelihood. The weight of each scenario is calculated by dividing the proportion by the total proportion of all possible scenarios. Subsequently, the final flood inundation probability and maximum water depth are derived by multiplying each scenario's results by its respective weight. Based on these derived flood inundation probability and maximum water depth values, exposure and impact evaluations have been conducted for these 21 lakes. Section 4.3, which covers flood inundation simulation, exposure, and damage assessment, has been reanalysed and rewritten.

Since the modelling part lacks any validation, this is where frequency-magnitude relationship can come into play. I wonder whether employing your approach over past GLOFs can yield “typical extremity” of GLOFs (if you standardize the extremity of your scenarios for each lake on the dimensionless scale from 0-1)? While it is mentioned in Discussion section that the incompleteness of data about past GLOFs prevents the authors from attempting validation, I wonder whether any single GLOF characteristic (e.g., breach depth, flood volume, peak discharge, inundation area, etc.) could be used to validate the flood modeling results and estimate “typical extremity” of GLOFs in Nepal? Such an analysis could guide the weighting of your scenarios.

Response: On one aspect, we did not consider the GLOF outburst frequency because the underlying database for frequency–magnitude relations typically is very poor. Here, we considered different scenarios, i.e., the release of lake water at different predetermined proportions. Under the predetermined proportion, we examined a plausible range of values for lake volumes and peak discharges for each glacial lake, ensuring comprehensive coverage of all potential glacial lake outburst scenarios.

For validation, we appreciate the reviewer's understanding regarding the lack of historical event records, which is a common issue with GLOF inundation simulations. Additionally, it is noteworthy that our proposed framework utilises the fully physically based hydrodynamic model HiPIMS, intricately designed to capture the highly transient and complex hydrodynamic processes induced by events such as dam breaks and flash floods. HiPIMS has been successfully validated for these extreme flow conditions. The adoption of this model enhances our confidence in simulating the spatial-temporal processes of GLOF inundation, ultimately contributing to improved hazard evaluation results. Regarding the "typical extremity" of GLOFs, the weight assigned to each scenario is determined by its occurrence probability (see above). We are uncertain whether this addresses the reviewer's comments, and we would greatly appreciate further guidance and clarification on the term "typical extremity".

Overall, I’m in favor of recommending this study for further processing and subsequent publication after some modifications are considered. My suggestions to the authors are: (i) to

consider dam geometry when estimating max. flood volume; (ii) to consider the validation of this approach with some of the past GLOFs in the country (and obtaining “typical extremity”); (iii) to consider avoiding the use of random selection of scenarios which may be misleading.

Response: Thank you very much for the valuable comments from the reviewer. If there are any additional considerations beyond our responses provided above, we hope the reviewer will let us know.