Dear editor and authors

In this manuscript, Shrestha et al., presented a comprehensive database of Glacier Lake Outburst Floods (GLOFs) in High Mountain Asia (HMA). The authors combined three databases of glacier lakes in this work. The compilation of the GLOF database is largely based on literature review of articles from different sources including peer-reviewed papers, book chapters, technical reports as well as online news articles. The authors also take local knowledge into consideration, which is believed to be necessary considering the issue with under or over reporting of hazards in rural areas in HMA. However, this also brings challenges to the reliability of the sources. The novelty of this database lies in its inclusion of impacts of these GLOFs downstream, which could be difficult to quantify, fact-check and describe in a single .csv file. The authors have also created an interactive map and dashboard for visualization and quick check for non-academic users. I praise the efforts that have been put on compiling such a database, which has great value in creating vulnerability assessments and hazard adaptation plans for mountain communities.

However, I also have a few major comments on how the article is written regarding ESSD guideline and how the data is archived, and quality controlled.

1) First, the authors spent a lot of efforts on analyzing and interpreting the data in Sect. 3 and 4. Of course, an overall statistic of the data could be included. However, I believe, since the article is about presenting a dataset, the emphasis should be put on elaborating the methods used to produce the data, the choices of the variables, analyzing the quality, uncertainties, and limitations in the data and how it could be useful in other studies.
2) The methodology about how the lake dataset and GLOFs dataset were derived are not detailed enough. I understand that the lake data was compiled from three earlier datasets from different years. But it is not clear how they are different in terms of coverage and quality, if the three datasets are merged or used for GLOFs happened in different years separately, and if they are merged what the rule for merging is. For the GLOF dataset, the authors stated briefly that the data was derived by reviewing articles and interviews from different sources and verified by satellite imageries, and that false reported events are removed. But I think more details are need in describing and discussing this process as from this description the readers have no idea how reliable these derived data are. And the removed cases need more vigorous justification. Since this article is about the dataset not the interpretation of the data more discussion could be put in these aspects.
3) In terms of the datasets, the authors indicate that the dataset is publicly available on ICIMOD data portal (https://doi.org/10.26066/RDS.1973283). I assume this will be the main platform for downloading the data. However, the HMAGLOFDB_v1.0.csv file downloaded from there is not accompanied by either a metadata file, a description file or the HMAGLOFDB_removed.csv file. Thus, the user who downloads the data from there has no idea what each column in the data file means. I later realize that those files are included on the Github repository. But in the ICIMOD data port there is no mentioning of the Github repository. It would be nice to centralize these different bits of data or at least link them together to be more user friendly, especially for non-academic users since they are considered as important stakeholders in the article.

4) My biggest concern is the criteria for choosing the mechanism involved in lake breach or drainage. It is not clear to me how the authors decided to adopt a definite mechanism or mark it as unknown. I did not check all the GLOF events in the data files. But for the lakes I checked (next to Kyagar glacier and Khurdopin glacier) there seem to be some issues with this.
   The 34 GLOFs of the lake next to Kyagar glacier are well documented in different articles, which are cited in the HMAGLOFDB_v1.0.csv file. The GLOFs before 2018 were believed to be triggered by ice-dam failure linked to subglacial drainage. And the one in 2018 was more likely to be linked to overtopping. However, the entry for the mechanism is unknown. I don't know what the authors' reasoning behind that.
   The GLOF next to Khurdopin glacier are marked as caused by 'englacial tunnels. However, Bazai et al. (2022) used a subglacial hydrology model to simulate the sudden drainage. Clearly, Bazai et al. (2022) thought the drainage was likely to be linked to subglacial drainage system. The authors of this manuscript have cited Bazai et al. (2022) but decided to adopt englacial tunnel as the mechanism.
5) There are many events that only recorded in 'this study'. We have no idea how they are identified and quality controlled.

Besides, the major issues I also have a few minor comments:

L90 : (Chen, et al., 2021) -> Chen et al. (2021). There are also some other places that the citations are written not according to the convention.

L115: This kind of statement or practice does not seem to be very rigorous. The authors have excluded many cases that might have been caused by debris flows but include cases that have happened far away from any glacier just because they appear in a landscape that was most likely glaciated at one point?

L120-124: The information in this part is not included in the dataset but only describes how the data is analyzed for the discussion in Sect.4. Following my major comment no.1 I don't know if this should be put here or be included at all.

L193-194: In which place did the other 0.8% happen?

L 339-342: Need some references here.

Fig. 2: It is not clear why this figure should be included and why the pictures of these glaciers are selected. These pictures are not a part of the database; or should they be included as a part of the database? Something could be a reference is the Norwegian Water Directorate GLOF map (http://glacier.nve.no/Glacier/viewer/GLOF/en/).

Fig. 3 Maybe it is better to use another colormap to be color-blind friendly.

Fig. 4 A Should the cause of lake appearing be 'glacier melt' or 'glacier retreat'? Or it means something else?

Fig. 6 Not quite sure what the x-axes represent.