Response to Reviewers' comments to manuscript essd-2020-57

"Annual 30-meter Dataset for Glacial Lakes in High Mountain Asia from 2008 to 2017"

Dear Editors and Reviewers:

Thank you a lot for your kind and careful reviewing. Your suggestions give us important and constructive perspective on this manuscript, and help to improve the manuscript greatly. We have fully considered all the comments of you, and have substantially revised our manuscript according to your comments. A point-by-point response to the outstanding comments raised is attached to this manuscript. The major changes are summarized as follows:

1. We have put considerable effort to update data of glacial lakes for the ten year records, and manually append their attribute information. The related statistics, figures and analysis in the article have also been modified based on the new lake inventory.

2. We have comprehensively investigated the existing works about glacial lake inventory in the Section 1. Introduction, and quantitatively analysed and compared these inventories with ours in the Section 6. Discussions, to clearly show benefits and challenges remaining in this study.

3. Detailed explanations about the mapping of some problematic ice-covered lakes have been given in the Section 3.2. The number of missed or misclassified lakes that need to be manually corrected were also described clearly.

4. A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript, the error bars for the lake area, and confidence intervals for the estimated trends were also added throughout the paper.

5. We have carefully modified the language deficiencies, imprecise expressions, and provided more detailed interpretations and conclusions.

The changes have been highlighted in colored text in the revised manuscript. In the following, we provide point by point responses to the outstanding comments and suggestions provided by the Anonymous Reviewers. We are indebted to you for your outstanding and constructive comments, which greatly helped us to improve the technical quality and presentation of our manuscript. Once again, thank you very much for your comments and suggestions.

Response to Comments by Reviewer#1:

1. *Chen et al. mapped glacier lakes in High Mountain Asia in annual time steps from 2008 to 2017 from Landsat imagery. They used an automated image segmentation algorithm to pre-classify >40,000 Landsat images between late spring and early fall. Ten experts manually refined the annual lake inventories. The authors further calculated trends of lake abundance and area in the entire study region and several subregions thereof, and show that the growth in lake area and number could be largely due to many new lakes that occupy the lower size spectrum. Hotspots of total lake growth are the Central Himalaya, Eastern Himalaya, and Western Himalaya. The authors also correlate the size-distribution of glacier lakes with a set of environmental predictors such as temperature change, debris cover, or glacier size, to explain where glacier lakes have formed in the landscape. Debris cover and trends in warming and precipitation could help to explain the size-distribution and growth of glacier lakes, though regional variances make a more rigorous explanation elusive.*

Response:

We greatly appreciate the suggestions of the Reviewer for his/her accurate summary of the main contributions of our work and for the outstanding recommendations provided. Following the Reviewer's very pertinent recommendations, in the revised manuscript and the following text we have made modifications according to these questions. We thank again the Reviewer for his/her careful handling of our manuscript and for the constructive suggestions provided, which greatly helped us improve the technical quality and presentation of our manuscript.

2. *Chen et al. have compiled an important, and possibly the most extensive, data set of glacier lakes in the HMA to date. This inventory could be useful for scientists and practitioners in many disciplines interested in changes of the cryosphere of the HMA. Each year has >14,000 lake polygons, so that it is difficult for me to judge how accurately these lakes have been mapped. (1)Yet, visually comparing the inventories of 2016 and 2017, I noticed that the 2017 data set misses >2 km² of lakes in a subset 0.3° x 0.2° large. The figure in the attachment shows 11 yellow polygons from 2016 that are not overlapped by blue polygons in 2017. Yet all these lakes are clearly visible in the Landsat scene from 2017-09-26, which is the basemap here. It is surprising that the 'ten trained experts' (L150-151) have not detected this error. Of course, my cross-check was not systematic, and it could be that I just accidentally came across one of very few undetected lakes in this study. Nevertheless, (2)I strongly recommend that the authors again check the inventories for such issues, because a larger quantity of undetected lakes could in fact contribute to the calculated statistics, and the conclusions drawn in this study.*

Response:

We gratefully thank the Reviewer for his/her comments and suggestions. In this study, we produced the glacier lake dataset in High Mountain Asia in annual time steps from 2008 to 2017. The number of glacial lakes in each year is more than 12,000, totally, the ten year records have nearly 140,000 lake polygons. To improve work efficiency, obtain the accurate boundary of the glacial lakes, and meanwhile, minimize the subjective judgment errors of the operatives, we applied a systematic glacial lake detection method that combined two steps from initial glacial lake extraction and subsequently manual refinement of lake mapping results.

The initial glacial lake extraction using GEE can make sure that approximately half of all the lakes are automatically extracted. To make annual data more complete and accurate, manual inspection and refinement of individual glacial lake is necessary to supplement some missing lakes and correct the mapping errors such as mountain shadows and river segments. The glacial lake vector over the entire HMA for the years from 2008 to 2017 has been rechecked and reedited individually through dynamic cross-validation by ten trained experts. Besides, the related attribution (i.e., lake type, elevation, distance to the nearest glacier terminus, area and perimeter) were manually added for each lake. It should be noted that the type of each individual glacial lake was carefully judged according to the different formation mechanisms or growth stages. Thus the whole processing is time consuming and require a considerable amount of work. We are terrible sorry for missing lakes that larger than 2 km^2 in the Landsat scene from 2017-09-26. Although the lakes >2 km^2 are not that much, they may still contribute to the calculated statistics, and the conclusions drawn in some sub-regions.

Therefore, following the Reviewer's very pertinent recommendations, we have carefully examined the lake data for the ten year by supplementing glacial lakes that have an area of greater than 0.04 km^2. The area of 0.04 km^2 is chosen as the threshold for re-revision of the glacial lake dataset for three reasons: **i)** glacial lakes with small sizes are more likely to be confused with surroundings due to its less effective spectral, textural and spatial information in comparison with those of relatively large glacial lakes, which results in the **large uncertainty of area** (Please see the Section 4, small glacial lakes with the area of less than 0.04km^2 have the mean area uncertainty of about 25.7 %). **ii)** small glacial lakes are highly variable in their locations, shapes and size, and also the optimal images with valid observations over the potential glacial lake area is very limited **in each year**. The image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from image quality, ice and shadow that obscured the lakes, this creates a great deal of uncertainty about the number and extent of small glacial lakes **for the such an annually resolved inventory**. **iii)** given the amount of work required to update the inventory and also revise our manuscript before the deadline.

Our team have been working hardly to update data of glacial lakes larger than 0.04 km^2 in the ten year, and add their attribute information. Correspondingly, the related statistics, figures and analysis in the article have also been modified based on the new lake inventory. We tried our best to improve the data and manuscript, and hope that the revised manuscript will meet with approval. As an annual time steps data over the HMA region for the ten-year time period, the produced data may not be complete and perfect enough, yet our team will continuously update and share more and better glacial lake data in the future.

3. *In any case, the tremendous effort that the authors have spent to map several tens of thousands of glacier lakes is, unfortunately, not reflected in the quality of the manuscript. I first wish to emphasize some major points in each section of this manuscript, and outline further details below.*

Response:

We gratefully thank the Reviewer for his/her constructive comments. This paper focus on obtaining the comprehensive knowledge of the distribution, area of glacial lakes, and also quantification of variability in their size and type at high resolution in HMA. We develop a HMA Glacial Lake Inventory (Hi-MAG) database to characterize the annual coverage of glacial lakes from 2008 to 2017 at 30 m resolution. This is the first glacial lake inventory across the HMA with annual temporal resolution, it can provide details for different types of glacial

lakes and evolution patterns. Although the method of lake mapping was automatic, for quality control every lake polygon was inspected and was manually edited where needed. The related attribution were also added for each lake. This is a huge amount of work for the mapping of nearly 140,000 glacial lakes over ten time periods from 2008 to 2017.

In the last version of manuscript, the mapping work was not fully reflected in the quality of the manuscript indeed. Therefore, in the revised manuscript, we have made considerable efforts to fill that defect, which is mainly manifested in the four aspects including **supplement of glacial lake data, quantitative uncertainty analysis of lake area and adding confidence intervals for trends, details about mapping procedure, comparison with other lake inventories**. We have made the following improvements in our study:

1) We have been working hardly to update data of glacial lakes larger than 0.04 km^2 in the ten year, and add their attribute information. Based on the new lake inventory, we have also carefully modified the tables, figures, the related analysis and made reliable conclusions.

2) A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript. Given the mapping uncertainty of glacial lakes and thus for the further improvement of reliability of calculated statistics, the error bars for the lake area, and confidence intervals for the estimated trends were also added throughout the paper.

3) Detailed descriptions about the key procedures for glacial lake mapping were provided, such as the amount of work required for the manual editing, and mapping of some problematic ice-covered lakes.

4) To clearly show benefits and challenges remaining in this study, and also for more rigorous cross-validation, we firstly comprehensively investigated the existing works about glacial lake inventory. Then in the **Section 6.1** of revised manuscript, we have conducted a deep analysis and comparison between our Hi-MAG dataset and glacial lake inventory from Wang et al., 2020, who mapped glacial lake larger than 0.0054 km^2 in a much larger High Mountain Asia region in two time periods (1990 and 2018), and made their data public recently. The comparative discussions with other studies (e.g. Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a) were used as a general reference.

Furthermore, we have carefully considered the major points in each section, and more points of criticism line-by-line provided by the Reviewer, and have substantially revised our manuscript according to your comments in the following response and revised manuscript.

4. *The introduction is very brief and the many dozens of glacier lake inventories that have been compiled in this region before remain largely unmentioned. These inventories are also largely disregarded in the discussion though they could help to discuss the benefits and challenges remaining in this study.*

5. *L37-39: The authors could use more suitable references to support their arguments: Salerno et al. used 10 m satellite images and have mapped lakes <0.001 km^2; Brun et al. have studied glaciers, not lakes; for the inventories with 'narrow geographic scope', it would be good have a reference for some of the major region such as Karakoram, Sikkim, Bhutan, Nepal, Tien Shan, Nyainqentanglha, etc., to stress this 'patchwork' of glacier lake inventories more clearly. The ~2003 inventory of ICIMOD (Maharjan et al., 2018) and the multi-temporal inventories by Zhang et al. (2015) and by Wang et al. (2020) are not*

mentioned at all here, though these studies addresses many of the issues raised here. What about the study from Pekel et al. (2016), who mapped surface water at high resolution globally? In summary, the authors could acknowledge previous work more than in the present form, given that many teams of researchers had aimed to map glacier lakes in the entire HMA (or parts thereof) before. . .

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out these issues. Many previously published researches have devoted to the glacial lakes mapping with remotely sensed data over the different regions of HMA, their inventories have proved to be important data resources for the spatiotemporal analysis of glacial lakes and better understanding the response of glacial lakes to climate change and glacier retreat. In the last version of manuscript, the introduction section is very brief indeed, which lacked the detailed descriptions and related references about many glacial lake inventories complied in the HMA region before. Following the Reviewer's very pertinent recommendations, in the revised manuscript, we have comprehensively investigated the existing works in the **Section 1. Introduction**. The detailed modifications are as follows:

Many previously published researches have devoted to the glacial lakes mapping with remotely sensed data over the different regions of HMA. Some works mainly focus on the investigation of the development of relatively large glacial lakes. Rounce et al. identified 131 glacial lakes in Nepal in 2015 that are greater than 0.1 km^2 (Rounce et al., 2017). Li et al. compiled an inventory of glacial lakes ($\geq$ 0.01 km^2) with a spatial resolution of 30 m in the Karakoram mountains (Li et al., 2020). Aggarwal et al. shared a new dataset of glacial and high-altitude lakes that have an area > 0.01 km^2 for Sikkim, Eastern Himalaya from 1972–2015 (Aggarwal et al., 2017). Ukita et al. constructed a glacial lake inventory of Bhutan in the Himalaya from the period 2006-2010 based on high-resolution PRISM and AVNIR-2 data from ALOS. Considering small lakes present less of a GLOF risk. They set 0.01 km^2 as the minimum lake size (Ukita et al., 2011). Ashraf et al. used Landsat-7 ETM+ images for the 2000-2001 period to delineate glacial lakes greater than 0.02 km^2 in the Hindukush-Karakoram-Himalaya (HKH) Region of Pakistan (Ashraf et al., 2012). Because small glacial lakes experience highly variable in their shape, location, and occurrence, and were clearly sensitive to the warming climate and glacier wastage, a growing number of scholars have paied attention to the abundance of small glacial lakes. Salerno et al. provided a complete mapping of glacial lakes (including lake size less than 0.001 km^2) and debris-covered glaciers with 10-m spatial resolution in the Mount Everest region in 2008 (Salerno et al., 2012). Wang et al. utilized Landsat TM/ETM+ images for the years 1990, 2000 and 2010 to map glacial lakes with area more than 0.002 km^2 in Tien Shan Mountains (Wang et al., 2013a). Luo et al. examined glacial lake changes (lake area >0.0036 km^2) for the entire western Nyainqentanglha range for the five periods between 1976 and 2018 using multi-temporal Landsat images (Luo et al., 2020). International Centre for Integrated Mountain Development (ICIMOD) provided comprehensive information about the glacial lakes (greater than or equal to 0.003 km^2) of five major river basins of the Hindu Kush Himalaya (HKH) using Landsat images for the year 2005 (Sudan et al., 2018). Nie et al.

mapped the distribution of glacial lakes across the entire Himalaya in the year of 2015 using a total of 348 Landsat images at 30 m resolution. They set the minimum mapping unit to 0.0081 km^2 (Nie et al., 2017). Zhang et al. presented a database of glacial lakes larger than 0.003 km^2 in the Third Pole for the years ~1990, 2000, and 2010 (Zhang et al., 2015). All these researches greatly help to fill the data gap of glacial lakes information in the HMA region. At the global scale, Pekel et al. used millions of Landsat satellite images to record global surface water over the past 32 years at 30-metre resolution (Pekel et al., 2016b), many large and visible glacial lakes were also included. More recently, Shugar et al. mapped glacial lakes with areas >0.05 km^2 around the world using 254,795 satellite images from 1990 to 2018 (Shugar et al., 2020). Wang et al. developed a glacial lake inventory (with size of larger than 0.0054 km^2) across the High Mountain Asia at two time periods (1990 and 2018) using manual mapping on 30 m Landsat images (Wang et al., 2020b). They firstly introduce glacial lake inventory at such a large-scale, the data shared will be served as a baseline for the further studies related to water resource assessment or glacier hazard risk.

Besides, in the **Section 6.1 Comparison with other lake dataset** of the revised manuscript, we have deeply analysed and discussed about the major differences between the existing inventories and ours, to clearly show benefits and challenges remaining in this study. The detailed modifications are as follows:

6.1 Comparison with other lake datasets

We compared our dataset with that of Wang et al., 2020 for the closest period (2017 from Hi-MAG database and 2018 from Wang et al. 2020) over the spatial extent of our HMA region. The differences in the total number and area between these two dataset are 6206 and 223.97 km^2, respectively. We also found that 2077 glacial lakes with a total area of 178.77km^2 in our Hi-MAG dataset were not detected by Wang et al. The main reasons for the missed glacial lakes in Hi-MAG are because the interference of some bad observations (cloud or snow), glacial lake dried up or outburst, or located in the middle of the river.

To test the spatial correlation of glacial lakes distribution in two datasets, we compared the statistics in glacial lake number and area aggregated on a 0.1°×0.1° grid for the HMA regions. The results for the total glacial lake area, areas for glacial lakes larger than 0.04km^2, and number for glacial lakes larger than 0.04km^2 are depicted in Figure IV. A clear and strong correlation can be observed for all the statistics between Hi-MAG dataset and glacial lake data by Wang et al. Most of the points being distributed around the 1:1 line, which shows that there is great consistency in the results.

In order to quantitatively and systematically evaluate the accuracy of our product, we implemented a stratified random sampling (Song et al., 2017b;Stehman and SV, 2012), where the glacial lakes were divided into four strata. The sample sizes are the spatial resolution (30 m) of the data, and the strata are designed as: C0W0. both the results are non glacial lakes; C0W1. non glacial lake for Chen's and glacial

lake for Wang's; C1W0. glacial lake for Chen's and non glacial lake for Wang's; C1W1. Both the results are glacial lakes.

A total of 4,000 points were randomly selected, as shown in Figure V. The sample number for C1W1 and C1W0 are 1300 and 700, respectively, which is almost the same ratio between the total areas for the two strata (1450.50 km^2 vs 732.77 km^2). Because of the approximate total area with C1W0, we also randomly selected 700 samples from stratum C0W1. The rest 1300 samples are from C0W0. Every validation sample was visually examined using Landsat imagery and higher-resolution imagery in Google Earth. Sample pixels were interpreted by a regional glacial lake mapping expert, and ambiguous samples were cross-validated by a second observer. If a sample is difficult to interpret, it was marked as ambiguous sample and excluded for the accuracy assessment. The sample number estimates were produced for each of the four strata (Table I), and these strata totals were then summed to obtain the total accuracy.

For the 1300 pixel samples that were considered to be non glacial lakes by both datasets, after the pixel by pixel verification, 1215 were indeed non glacial lakes, while 37 were the missed glacial lakes. In contrary, 1260 out of 1300 pixels belongs to the class of glacial lake, and 25 pixels were misclassified as glacial lakes by the two inventories. 307 error pixels were found in the results from Wang et al., constitute about half of the total validation number. For the glacial lakes identified only by our inventory, 678 out of 700 were corrected classified. Our results yielded high overall classification accuracy (88%), user's accuracy (97%), and producer's accuracy (82%) for glacial lake classification using Landsat data.

Hi-MAG was also compared with other Landsat-based lake inventories (Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a). Hi-MAG lake number was 7268 higher and area was 644.26 km^2 higher than the estimation for the Tibetan Plateau (Zhang et al., 2015). The largest discrepancy is in the Gangdise, Himalaya and Nyainqentanglha Mountains in 2010. Across the Himalaya, we found 476.09 km^2 of glacial lakes, 4.57% more than previous estimates in 2015 (Nie et al., 2017). In addition, we qualitatively compared the lake extent between publicly available high-resolution Global Surface Water (GSW) dataset and our Hi-MAG database summed by mountain range in 2015. GSW data can be accessed at https://global-surface-water.appspot.com/download. For the sake of a reliable comparative analysis, lake polygons in the Hi-MAG dataset were converted into the grid format, and glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km). Hi-MAG detected more glacial lakes in the Himalaya, Eastern Hindu Kush, and Tien Shan, and fewer in Eastern Pamir and Western Kunlun Shan. Fig. A5 illustrates the differences between our Hi-MAG glacial lake results and GSW-derived lake area for the whole HMA region.

The glacial lake area observed in our lake dataset in the Eastern Pamir and Western Kunlun Mountains does not conform to the mapped surface water in the GSW for these sub-regions. While there are numerous glacial lakes from an open water perspective, actually part of them are river

Additionally, the Himalaya, Eastern Hindu Kush, and some other Tien Shan host thousands of glacial lakes that are not readily observable in the GSW product. Large discrepancies in mountainous glacial lake estimates preclude a significant consistency between GSW and our Hi-MAG lake data over the HMA region. The region with the highest consistency between GSW and Hi-MAG product is interior Tibet. There is little agreement for Tien Shan, where the weather is rainy and snowy in the region above 3000 m, and large amounts of ancient glacial deposits have been accumulated. Here glacial lakes are featured by small size, due to the influence of source glaciers and lake beds, as well as the water depth and sediment inflow, glacial lakes appear heterogeneous reflectance in the image. Errors could exist in datasets produced by automated classification, but we also did a detailed manual editing, so we were not relying exclusively on automatic classification. Karakoram regions seem to have fewer glacial lakes in our estimate, owing to the overestimation of surface water on debris covered glaciers in the GSW dataset.

The little agreement between our Hi-MAG glacial lakes data and GSW data is mainly due to its lack of systematic glacial lake inventories and mapping capabilities. The lake dynamics and differing climate contexts within HMA may also lead to inconsistencies between the sub-regions. Hi-MAG might have made better use of the optimum satellite imaging season to map glacial lakes, potentially resulting in more complete mapping by avoiding conditions — such as periods of lake ice — that may confound mapping.

6. *The methods do not introduce the problem of ice-covered lakes that occurs with increasing elevation. Though the authors manually corrected for missing or misclassified lakes, it remains unclear, how many lakes needed to be corrected manually after automatic classification, and how many lakes are possibly missing in this data base.*

    L151-153: How many lakes (in total, or percent) were manually added to the automatically derived lake inventories? How many needed manual refinement?

Response:

We gratefully thank the Reviewer for his/her outstanding recommendations provided. In this study, to reduce the influence of seasonal lake fluctuations for the mapping, one effective solution is to map glacial lakes and measure their long-term changes during stable seasons when lake extents are minimally affected by meteorological conditions and glacier runoff. Here the selected time series of Landsat data were generally from July to November. During this period of each year, the Landsat imagery featured less perennial snow coverage. The lakes also reached their maximum extent, specifically around the end of the glacier ablation season (June to August). These criteria were used for the selection of imagery for the initial glacial lake extraction using GEE, manual inspection and refinement of individual glacial lake were then conducted to visually inspect and correct the mapping errors. Hi-MAG might have made better use of the optimum satellite imaging season to map glacial lakes, potentially resulting in more complete mapping by avoiding conditions— such as periods of lake ice— that may confound mapping. For the images without valid observations over the potential glacial lake area, then optimal mapping time needs to be broaden during the whole year.

However, some lakes may be permanently coved by ice with the increasing elevation. Besides, it is also difficult to map glacial lakes in environments with similar reflectance from the adjacent land surfaces, Tien Shan is such a typical area, where mapping errors could exist in the dataset. On these conditions, a detailed manual editing is essential to map these glacial lakes. So we were not relying exclusively on automatic classification. Here the ice-covered glacial lakes that were missed in the automated processing were manually edited using additional high-quality scenes downloaded from USGS Earth Explorer website during the whole year with assistance of images from adjacent years.

As for the automatic classification using GEE, there are four main procedures including (i) the clipped Landsat images by the extent of the glacier buffers and assembled into a time-series dataset; (ii) identified and masked some poor quality observations such as cloud, cloud shadow, topographic shadow; (iii) calculated MNDWI; and (iv) extracted the potential glacial lake areas by applying adaptive MNDWI threshold. Based on the automated processing, nearly 60% glacial lakes in each year can be correctly classified, of the other lakes that were not properly classified, 30% were missed and 10% were misclassified (mainly mountain shadows and river segments). For such a large-scale area that characterized by various and complex climatic, geological and terrain conditions, this classification method is simple but effective, the results are also reasonable since it provides very low commission errors. As shown in Figure I, glacial lake outlines extracted using the automated classification method in our study fit the real boundary of the glacial lake very well, while manually delineated glacial lake outlines are largely influenced by people's subjective experience and manual operation, resulting in overestimation for most part of a glacial lake region, and underestimation for a few areas. Moreover, results from manual digitization show poor performance for the delineation of the some glacial lakes with complex curved shapes.

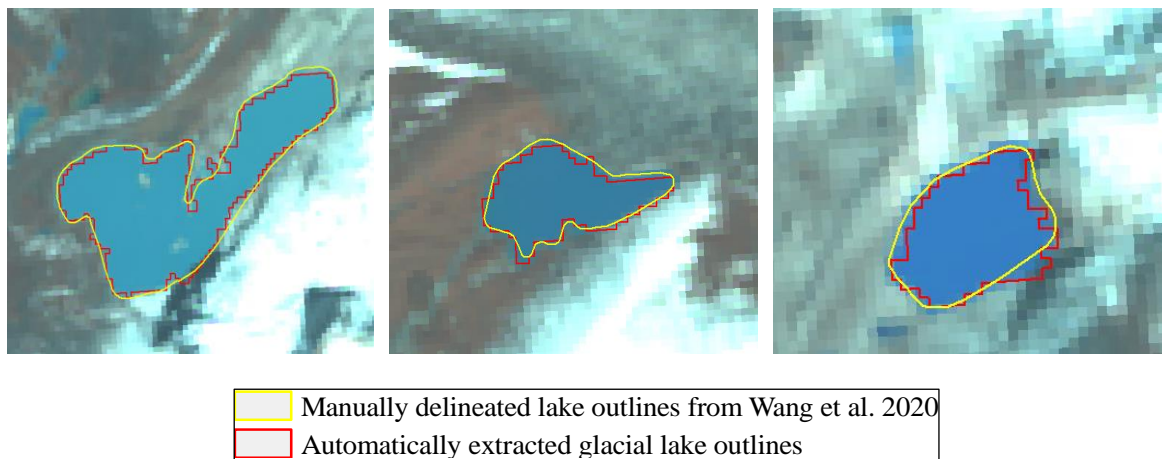|	
	Manually delineated lake outlines from Wang et al. 2020
	Automatically extracted glacial lake outlines

Figure I. Examples of extraction of glacial lake outlines using the automated classification method in our study and manual digitization from Wang et al. 2020.

However, an estimated 30% errors in the initially classified map were from the missing small glacial lakes. Glacial lakes with small sizes are more likely to be confused with surroundings due to its less effective spectral, textural and spatial information in comparison with those of relatively large glacial lakes. Because the automated method is mainly based on the spectral features for the glacial lake mapping over the large mountainous area, the spectral information provided by the small glacial lakes will be incomprehensive and insufficient for the accurate detection of this kind of glacial lakes under various and complex environmental conditions.

Therefore, visual interpretation and manual editing is still an effective way to ensure the high accuracy of lake inventories.

Finally, we have added some text in the **Section 3.2 Adaptive glacial lake mapping method** to give a detailed explanations about the mapping of some problematic ice-covered lakes that occurs with increasing elevation. The number of missed or misclassified lakes that need to be manually corrected were also described clearly.

The detailed modifications are as follows:

==Based on the automated processing, nearly 60% glacial lakes in each year can be correctly classified, of the other lakes that were not properly classified, 30% were missed and 10% were misclassified. For such a large-scale area that characterized by various and complex climatic, geological and terrain conditions, this classification method is simple but effective, the results are also reasonable since it provides very low commission errors.== To ensure the quality of inventory, strict quality control was conducted to visually inspect and correct the mapping errors after the automated processing using GEE. ==False lake features, mainly identified as mountain shadows and river segments, were manually removed by overlapping mapped lake shorelines on the source Landsat imagery and higher-resolution imagery in Google Earth.== For missing glacial lakes, such as some lakes may be covered by ice and clouds for years, grow at steep glacier tongues, and lakes show heterogeneous reflectance with the surrounding backgrounds, the lake boundaries were edited further using ArcGIS.

7. *The text and the figures in the results (Chapter 5.1 and 5.2) show dozens of numbers and trends, regarding lake areas, abundance, and changes thereof, but remain without error bars or confidence intervals throughout. Yet the authors acknowledge that mapping errors for small lakes in particular can be large, so that it must be cautioned against interpreting these trends. The panels in Fig. 5, for example, show considerable variance, and it remains unclear whether the regression models used to calculate the trends, account for this variance.*

Response:

We greatly appreciate the Reviewer's valuable comments very much. In the Section 5.1 and 5.2 of the last version of manuscript, lake area and calculated trends lacked the related error bar and confidence intervals, respectively, which are not rigorous indeed. In the revised manuscript, we have calculated and added the absolute area error and relative error for the lake area, and significance level for the estimated trends throughout the paper.

Besides, as for the estimation of area error, most of the large glacial lakes (area $\geqslant$ 0.04km^2) have the relative error of about 7%. We also measured glacial lake down to 0.0081 km^2 (nine pixels in Landsat imagery), where relative errors calculated were ~50%. It can be found that this systematic error was more significant for the small-sized glacial lakes. We gratefully thank the Reviewer for pointing out important issue about interpreting the trends for small glacial lakes cautiously. We have carefully modified the whole manuscript to make sure the statistics and related analyses all focus on large-sized and all sizes of glacial lakes over the whole HMA region or different mountain ranges to discuss spatial and temporal variability for the 10-year record, and meantime, to avoid analysing and explaining the trends and number for small-sized

glacial lakes. Fig. 5 shows the inter-annual variations in the number of small glacial lakes for different HMA mountain regions. Given the large uncertainty of area and number of small glacial lakes, and in order to make the analysis and logic of the whole article more rigorous, we have removed the Section 5.2 in the revised manuscript. Please see the revised manuscript for more detailed information.

8. *In Chapter 5.3 and the discussion, the authors search for environmental predictors that could contribute to the (changing) size-distribution of glacier lakes. Yet many of these correlations, for example with glacier slope or elevation, are presented in tables, but remain untouched in the manuscript. Other correlations such as local glacier mass balances (e.g. Brun et al., 2017) are disregarded, though these could be useful in the light of shrinking glaciers and growing lakes. In general, the discussion does not go far beyond showing these correlations; discussion with previous studies that aimed to explain the size-distribution of glacier lakes remains very limited.*

Response:

To the extent possible we have tried to improve the explanation and context provided for our results comparing lake distribution with primary glacial, geomorphic and climatic factors. Our main objective in this section is not to provide fundamental new understanding of lake formation processes (which would be outside the scope of this journal), but rather, utilise our unprecedented large lake dataset to confirm relationships and processes that have previously been suggested based on regional observations and studies. We have improved the reference to these previous studies and reworded sentences that were vague or unclear to the reviewer.

9. *Throughout the manuscript the authors use subjective qualifiers such as 'small', 'large', or 'relatively' that could be systematically filled with the data that they have produced. The orthography, grammar, and style could still deserve much improvement, which is surprising to see given the number of co-authors involved in the writing part of this study.*

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. In the revised manuscript, we have put considerable effort to revise the whole text, to avoid using these subjective qualifiers such as 'small', 'large', or 'relatively', and improve the presentations of the relevant data. Moreover, we have also carefully modified the language deficiencies, imprecise expressions, some simplistic interpretations and style, and make the manuscript in a better design.

10. *These major points should be fully addressed in a revised manuscript. Below I detail line-by-line some more points of criticism. Some technical corrections are also included.*

    L17: 'Atmospheric warming' instead of 'climate change'?

Response:

Thanks for this good suggestion. We have replaced the words "climate change" by "Atmospheric warming" in the revised manuscript.

11. *L19: 'incomplete': Can this study claim completeness?*

We apologize for using this inappropriate expression. Our research is also incomplete in some respects, for instance, the incomplete delineation of some quite small glacial lakes (<0.04km^2) over ten year period; besides, although our Hi-MAG dataset was produced in annual time steps from 2008 to 2017 (which has high temporal resolution), ten years is still a short span, shorter than typical glacier response times to the atmospheric warming, lake expansion can decouple from short-term climate fluctuations. Therefore, inspired by this original report, our research group will continue to update and optimize the released glacial lake data in the near future.

To make this expression much more accurate and rigorous, we have rewritten the related sentences in the revised manuscript. The detailed modifications are as follows:

There is a pressing need for obtaining the comprehensive knowledge of the distribution, area of glacial lakes, and also quantification of variability in their size and type at high resolution in HMA.

12. *L21-22: 'rapid', 'moderate', 'large', 'faster': please be more specific. From annual lake inventories, it should be straightforward to calculate, for example, the total lake area, number, and absolute changes during the study period.*

We gratefully thank the Reviewer for his/her constructive comments. By calculating the corresponding statistics of the number and area of glacial lakes, we have carefully revised the contents about 'rapid', 'moderate', 'large' and 'faster' to make them more specific. The detailed modifications are as follows:

Annual increase in lake number and area are 306 glacial lakes and 12 km^2, respectively, and maximum increased lake number occurred in 5400 m elevation, which increased by 249.

13. *L23-24: 'Proglacial lake dominated areas'; 'unconnected lake dominated areas': unclear what and where these areas are.*

We are so sorry for these unclear expressions. Proglacial lake dominated areas mainly refer to the areas (over 1°×1° grid) that around more than half of the glacial lake area consisted of proglacial lakes. In this study, the Nyainqentanglha and Central Himalaya are the typical representatives of the proglacial lake dominated areas. Similarly, in the regions of Eastern Tibetan Mountains and Hengduan Shan, the unconnected glacial lakes occupied over half of the total lake area in 1°×1° grid, were dominantly occupied.

Besides, to make it much more clearly, in the revised manuscript, we have modified this sentence by providing more details about what and where the proglacial lake dominated and unconnected lake dominated areas are. The detailed modifications are as follows:

Proglacial lake dominated areas such as the Nyainqentanglha and Central Himalaya, where around more than half of the glacial lake area (summed over 1°×1° grid) consisted of proglacial lakes showed obvious lake area expansion, while in the regions of Eastern Tibetan Mountains and Hengduan Shan,

14. L23: 'significant': how do the authors measure significance here?

Response:

We greatly appreciate the Reviewer's valuable comments very much. For the glacial lake area time series in each $1°\times1°$grid, we applied the Theil-Sen estimator to derive the slope (annual change) of the trend. The change estimates satisfy the 90% confidence interval for the slope. While in the last version of manuscript, areas showed 'significant' lake area expansion originally meant that the areas have fast or obvious expansion. For example, in the Nyainqentanglha and Central Himalaya, where around more than half of the glacial lake area (summed over $1°\times1°$ grid) consisted of proglacial lakes, showed obvious lake area expansion. From Fig. 5(b), the annual area change in these areas are larger than $+0.30\ \mathrm{km^2\,a^{-1}}$, present deep red in color bar.

   Moreover, we are sorry for the improper use of the word 'significant'. We have used "obvious" instead of "significant'" to make it easy to understand. The detailed modifications are as follows:

   "where around more than half of the glacial lake area (summed over $1°\times1°$ grid) consisted of proglacial lakes showed obvious lake area expansion,"

15. L26: 'a main contributor': to? Increase in lake area?

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. Proglacial lakes contributed approximately 62.87% of total area increase (56.67 $\mathrm{km^2}$) over HMA (Table A1 and A2), serving as a main contributor to the increase in lake area. We have supplemented this content in the revised manuscript.

16. L27: 'an overlooked element'. Not sure whether this can be called a novel finding, given that the prevalence of small lakes in glacier lake inventories has been emphasized before, e.g. in (Khadka et al., 2018; Nagai et al., 2017; Shukla et al., 2018)

Response:

We gratefully thank the Reviewer for providing us with important references from Khadka et al., 2018, Nagai et al., 2017, and Shukla et al., 2018, they definitely have a certain guiding significance for our study. In these studies, the delineated small glacial lakes were widespread in the different sub-regions of HMA. Results shown that small glacial lakes are highly variable both in the number and spatial extent, and are quite sensitive to the atmospheric warming. The importance of these small glacial lakes has become increasingly prominent, and people pay more and more attention to them. Therefore, in the last version of manuscript, the related statement about small glacial lakes are 'an overlooked element' to recent lake evolution in HMA is not a novel finding **indeed**. In our study, to further emphasize that small glacial lakes play an important role for the total number changes and area expansion of glacial lakes in **the**

entire HMA, we have modified the words 'an overlooked element' to 'an important element' in the revised manuscript.

17. L29: 'is' instead of 'are'?

Response:

Thanks for this suggestion. We have replaced 'are' by 'is' in the revised manuscript.

18. L30: New sentence instead of comma?

Response:

Thanks for this good suggestion. We have corrected the related sentences to make it clearly in the revised manuscript. The detailed modifications are as follows:

<mark>"it can be used for studies on the complex interactions between glacier, climate and glacial lake, glacial lake outburst floods, potential downstream risks and water resources."</mark>

19. L34: 'significant': how do the authors measure significance here?

Response:

We greatly appreciate the comment of the Reviewer, and we are terrible sorry again for the inappropriate use of 'significant', which may confuse the readers. Similar to your comment 14, the word 'significant' here is different from the concept of statistical significance, it means that obvious or widespread, namely, atmospheric warming has resulted in widespread glacier retreat and downwasting in many mountain ranges of the HMA. In the revised manuscript, we have completely modified this sentence to make it easy to understand. The detailed modifications are as follows:

<mark>"Atmospheric warming has resulted in widespread glacier retreat and downwasting in many mountain ranges of the HMA (Brun et al., 2017b;Bolch et al., 2012b),"</mark>

20. L35: 'incompletely': same argument as in the abstract. Do the authors mean that glacier lakes had been mapped at large intervals?

Response:

We gratefully thank the Reviewer for his/her valuable comments and suggestions. Yes, the word 'incompletely' does mean in the previously published researches, glacial lakes had been mapped at relatively large time intervals, not at annual time scale. That is, glacial lakes have been incompletely documented at small time intervals in the previous studies. To explain the 'incompletely' much more clearly, we have corrected this sentence in the revised manuscript. The detailed modifications are as follows:

<mark>"yet glacial lakes have been incompletely documented at small time intervals."</mark>

21. L36-37: Please add a reference.

Response:

Thanks for this suggestion. We have added a reference after this sentence. The detailed modifications are as follows:

"Glacial lake development varies according to climatic, cryospheric, and lake-specific conditions, such as basin geometry that is either connected to glaciers or unconnected, and the length of the lake/glacier contact (Zhao et al., 2018)."

22. L52: 'atmospheric warming' instead of 'climate warming'?

Response:

Thanks for this good suggestion. We have replaced the words "climate warming" by "atmospheric warming" in the revised manuscript.

23. L45-56: What is the difference between i) and iv)? For example, given the situation of a lake that is dammed by a moraine, and its parent glacier calving into it: is this a lake 'usually connected to the glacial tongue and dammed by unconsolidated or ice-cemented moraines' (L46) or a lake 'bounded by a lateral moraine on one side and damming glacier ice on the other side'? Furthermore, I miss the category of purely glacier-dammed lakes, such as the one at Kyagar glacier (Round et al., 2017). Maharjan et al. (2018) list more than 20 ice-dammed lakes in the Himalayas.

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. In this study, all the glacial lakes were manually classified into four categories according to their position relative to the parent glacier or their formation mechanisms. Among these four types of glacial lakes, proglacial lakes are located next to the glacier terminus and receive melt water directly from their mother glaciers. These lakes usually connected to the glacier tongue and dammed by **glacier ice, unconsolidated or ice-cemented moraines** (mixture of ice, snow, rock, debris and clay, etc.), as shown in Figure I(a). Ice-marginal lakes are not very common in HMA, these lakes are generally distributed on one side of the glacier tongue, which means that the lake is dammed by the glacier ice on this side. While on the other side, it is bounded by a **lateral moraine** (Figure I(b)). With the increase of atmospheric warming and accelerated melting of glacier, some glacier tributary gradually detaches from a main trunk glacier. The detached location, where glacier melting has been particularly intense, in some case is also likely to form ice-marginal lakes. Based on the above analyses, if a lake is dammed by a moraine, and its parent glacier is calving into it, which means that the lake is contacted with its parent glacier terminus, this lake can be called a proglacial lake. Besides, in our study, purely glacier-dammed lakes are typically one kind of glacial lakes formed by the advance of glaciers and dammed by glacier ice, thus usually directly connected to the glacier tongue (as shown in Figure I(c), Kyagar glacial-dammed lake). Although the dam composition and structure is slightly different between the proglacial lakes and glacier-dammed lakes, dam of proglacial lakes is composed by the ice, rock, clay, etc., while the dam of glacier-dammed lakes is made up of almost pure glacier ice, because they are all located in the front of the glacier tongue and driven by the mother glacier, in the process of appending attributed information to each glacial lake, glacier-dammed lakes were merged into class of proglacial lakes.

Maharjan et al., 2018 adopted a two-level classification that is based on lake dam type and lake forms for the glacial lakes in the Hindu Kush Himalaya. The first level, based on dam type,

broadly classifies the lakes into four types: 1) moraine-dammed lake, 2) ice-dammed lake, 3) bedrock-dammed lake, and 4) other dammed lake. The second level, based on lake form, classifies the lakes into seven subtypes. **Ice-dammed lakes** are defined as lakes dammed by glacier ice, including lakes on the surface of a glacier or lake dammed by glaciers in the tributary/trunk valley, or between the glacier margin and valley wall, or at the junction of two glaciers. The criteria for classifying a glacial lake by Maharjan are completely different from our study. Given that our study area covers the entire HMA region, which includes ensemble of mountain ranges and favoured different properties of glacial lakes, all the lakes were classified mainly with respect to their position relative to the glacier and their formation mechanisms.
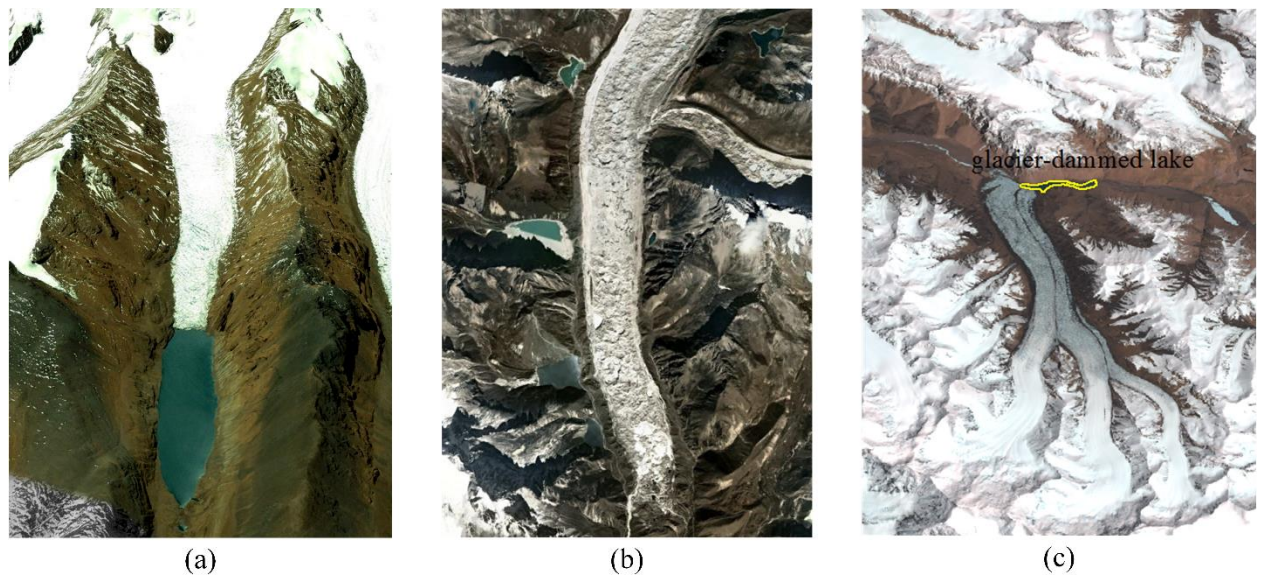

(a)   (b)   (c)

Figure I. Examples of (a) pro-glacial lakes, (b) ice-marginal lakes, and (c) glacier-dammed lakes (Kyagar glacial-dammed lake).

It should be noted that, in the last version of manuscript, the composition of dam also includes the glacier ice, which has been omitted in the definition of proglacial lakes. In addition, ice-marginal lakes and glacier-dammed lakes were also not introduced completely and clearly. Therefore, in the revised manuscript, we have thoroughly modified the concepts of different categories of glacial lakes in the Section I. Introduction. The detailed modifications are as follows:

Lakes were manually classified into four categories according to their position relative to the parent glacier or their formation mechanisms (Fig. A1): i) proglacial lakes, usually connected to the glacier tongue and dammed by glacier ice, unconsolidated or ice-cemented moraines (mixture of ice, snow, rock, debris and clay, etc.), proglacial lakes are located next to the glacier terminus and receive melt water directly from their mother glaciers; ii) supraglacial lakes - this is where ponds form in depressions on low-sloping parts of the surface of a melting glacier and are dammed by ice or the end-moraine or stagnating glacier snout; iii) unconnected glacial lakes, which are glacial lakes not directly connected to their parent glaciers at the present time but which, to some extent, may be fed by at least one of the glaciers located in the basin and may have been (but not necessarily are) recently detached from ice contact due to glacial recession. Although not directly connected with the parent glaciers, these glacial

lakes are also the outcome of glacier melting in response to ==atmospheric warming==, they can supply fresh water to major river systems of the HMA region, and their changes have significant scientific and socio-economic implications (Nie et al., 2017;Song et al., 2016); and ==(iv) ice-marginal lakes, these lakes are generally distributed on one side of the glacier tongue, which means that the lake is dammed by the glacier ice on this side. While on the other side, it is bounded by a lateral moraine. With the increase of atmospheric warming and accelerated melting of glacier, some glacier tributary gradually detaches from a main trunk glacier. The detached location, where glacier melting has been particularly intense, in some case is also likely to form ice-marginal lakes. We note that such ice-marginal lakes are very common in some parts of the world (e.g., Alaska (Armstrong and Anderson, 2020;Capps et al., 2011)) but are not common in HMA. Besides, purely glacier-dammed lakes are typically one kind of glacial lakes formed by the advance of glaciers and dammed by almost pure glacier ice. Although the dam composition and structure is slightly different between the proglacial lakes and glacier-dammed lakes, because they are all located in the front of the glacier tongue and driven by the mother glacier, in the process of appending attributed information to each glacial lake, glacier-dammed lakes were merged into class of proglacial lakes.==

24. L55-56: 'Alaska': reference missing

Response:

Thanks for this suggestion. We have added the corresponding references in the revised manuscript. The detailed modifications are as follows:

"We note that such ice-marginal lakes are very common in some parts of the world ==(e.g., Alaska (Armstrong and Anderson, 2020;Capps et al., 2011))== but are not common in HMA."

25. L57: 'for potential automatic mapping errors': The authors did not mention before that they mapped lakes automatically.

Response:

We gratefully thank the Reviewer for his/her comments and suggestion. As for this sentence, we did not really mention before that the lakes were mapped automatically. We have modified this sentence to make it much more accurate in the revised manuscript. The detailed modifications are as follows:

"Every lake was cross-checked manually ==for its boundary and attribution.=="

26. L58: What are 'systematic errors'?

Response:

Thanks for this good suggestion. In this study, systematic errors refer to the error estimate of glacial lake area. Assuming an uncertainty of 1 pixel for the detected glacial lake boundaries, we calculated the lake area error for the whole HMA region according to the equation (1). To

avoid conceptual confusion and for better understanding, we have corrected 'systematic errors' to 'lake area error' in the revised manuscript.

27. L60: The chapter on the study area contains very limited information on the 'study area'. Could be expanded, including the climate and topography of the HMA. This could prepare readers for the observed variability of glacier lake abundance.

Response:

We greatly appreciate the Reviewer's valuable comments very much. For preparing the readers for the observed variability of glacial lakes in the HMA region, following the Reviewer's very pertinent recommendation, we have rewritten and expanded the chapter on the study area by adding the map of the HMA and comprehensive descriptions about the climate and topography. The detailed modifications are as follows:

2.1 Study area

The term HMA refers to a broad high-altitude region in South and Central Asia that covers the whole Tibetan Plateau and adjacent mountain ranges, including the Eastern Hindu Kush, Western Himalaya, Eastern Himalaya, Central Himalaya, Karakoram, Western Pamir, Pamir Alay, Northern/Western Tien Shan, Dzhungarsky Alatau, Western Kunlun Shan, Nyainqentanglha, Gangdise Mountains, Hengduan Shan, Tibetan Interior Mountains, Tanggula Shan, Eastern Tibetan Mountains, Qilian Shan, Eastern Kunlun Shan, Altun Shan, Eastern Tien Shan, Central Tien Shan, Eastern Pamir (Fig. 1 and Fig. 4a). It extends from 26°N to 45°N and from 67°E to 105°E, the altitude of the plateau is about 4500 m on average (Baumann et al., 2009). It's made up of alternating mountains, valleys and rivers, the terrain is fragmented, showing a decreasing trend from Northwest to Southeast. HMA has a series of East-West mountains that occupy most of the area. Among these, Tanggula Shan lies in the central part of the HMA, with an altitude of over 6000 m. The height of fifteen highest mountains in the Himalayas are more than 8000 m, while the peaks of the mountains in the northern plateau are more than 6500 m. The North-South mountains are mainly distributed in the southeast of the plateau and near the Hengduan Mountain area. These two groups of mountains constitute the geomorphic framework and control the basic pattern of the plateau landform. Continuous and discontinuous permafrost have developed on the higher land and north facing slopes.

HMA is the source of several of Asia's major rivers, including the Yellow, Yangtze, Indus, Ganges, Brahmaputra, Irrawaddy, Salween, and Mekong. They play a crucial role in downstream hydrology and water availability in Asia (Immerzeel and Bierkens, 2010). Most glaciers in the Tibetan Plateau are retreating, except for the Western Kunlun (Neckel et al., 2014;Kääb et al., 2015) and the Karakoram, where a slight mass gain is occurring (Bolch et al., 2012a;Gardner et al., 2013). Moreover, glaciers in different mountain ranges show contrasting patterns. Local factors (e.g., exposure, topography, and debris coverage) may partly account for these differences but the spatial and temporal heterogeneity of both the climate and degree of climate change may be the main reason. Glacial lakes are formed and developed temporally with the retreat or thinning of glaciers and are directly or indirectly fed by glacier meltwater, they are located within 10 km from the nearest glacier terminus (Zhang et al., 2015;Wang et al., 2013b).

The HMA climate is under the combined and competitive influences of the East Asian and South Asian monsoons and of the westerlies (Schiemann et al., 2009). This unique geographical position produces an azonal plateau climate characterized by strong solar radiation, low air temperatures, large daily temperature variations and small differences between annual mean temperatures (Yao et al., 2012).

The annual mean temperature is 1.6 °C, with the lowest temperature of $-1 - -7$ °C occurring in January and the highest temperature of $7 - 15$°C occurring in July. The cumulative precipitation is about 413.6 mm a year.
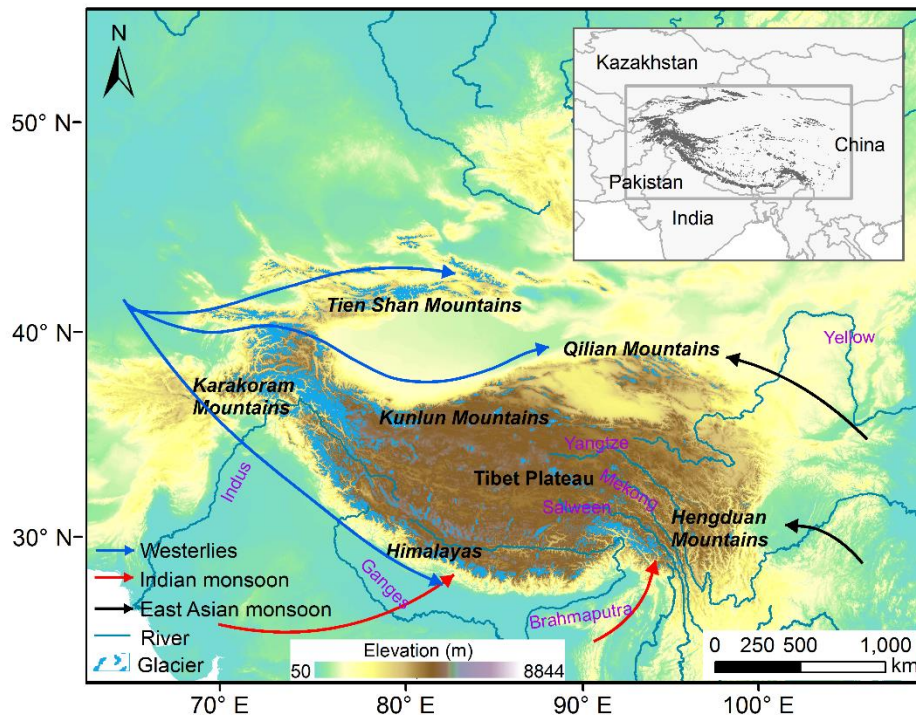

Fig. 1. The location of the High Mountain region of Asia (HMA). Glacier outlines from the Randolph Glacier Inventory (RGI v5.0), the Second Chinese Glacier Inventory (CGI2) and the GAMDAM inventory are drawn in sky blue.

28. L62: 'etc': Please write out all regions; Fig 4a is mentioned before Fig 1.

Response:

Thanks for your suggestion. In the revised manuscript, we have written out the names of all the mountain ranges in the HMA and adjusted the position of the Fig. 1 and Fig. 4a. The detailed modifications are as follows:

"The term HMA refers to a broad high-altitude region in South and Central Asia that covers the whole Tibetan Plateau and adjacent mountain ranges, including the Eastern Hindu Kush, Western Himalaya, Eastern Himalaya, Central Himalaya, Karakoram, Western Pamir, Pamir Alay, Northern/Western Tien Shan, Dzhungarsky Alatau, Western Kunlun Shan, Nyainqentanglha, Gangdise Mountains, Hengduan Shan, Tibetan Interior Mountains, Tanggula Shan, Eastern Tibetan Mountains, Qilian Shan, Eastern Kunlun Shan, Altun Shan, Eastern Tien Shan, Central Tien Shan, Eastern Pamir (Fig. 1 and Fig. 4a)."

29. L62: 'For this study': are there other studies where glacier lakes formed differently?

Response:

We are terrible sorry for making this mistake. In both our study and other studies, glacial lakes are all formed and developed temporally with the retreat or thinning of glaciers and are directly

or indirectly fed by glacier meltwater. We have deleted the words 'For this study' in the revised manuscript.

30. L63: 'within 10 km' instead of 'within a 10 km'?

Response:

Thanks for your suggestion. We have used 'within 10 km' instead of 'within a 10 km' in the revised manuscript.

31. L64: 'Approximately 40,481 Landsat series satellites': did the authors mean 'Landsat images from the three Landsat missions'? The term 'approximately' seems a bit odd here, given this exact number.

Response:

We greatly appreciate the suggestion of the Reviewer. Landsat series satellites scenes mean images from the three Landsat missions including Landsat 5 TM, Landsat 7 ETM+ and Landsat 8 OLI. In the revised manuscript, we have modified this sentence by explaining 'Landsat series satellites' clearly and giving the exact number of images used in this study. The detailed modifications are as follows:

==40,481 satellite images including Landsat 5 TM imagery during 2008 to 2011, Landsat 7 ETM+ imagery in 2012, and Landsat 8 OLI during 2013 to 2017,== were available in GEE and were used to produce the annual glacial lake maps over the entire HMA (Fig. 2).

32. L69: What do the authors mean by 'year-round'? Why did they not fill the years 2004, 2005, 2006, and 2007 with images from ETM+, similar to approach for 2012?

Response:

We gratefully thank the Reviewer for his/her constructive comments. Here the 'year-round' Landsat 5 TM data means all the available Landsat 5 TM images in each year. For the years before 2008, all the available Landsat 5 TM images in a single year do not fully cover the entire HMA region, please see Figure II for the examples of the annual Landsat 5 TM dataset available from 2004 to 2007 in GEE.

Similar for the images used for 2012 glacial lake mapping, Landsat 7 ETM+ images can be used as auxiliary data into the extraction of glacial lake outlines over the whole HMA region before 2008. However, the SLC-off condition of Landsat ETM+ introduces artefacts because the slatted appearance of the original images is occasionally carried into the glacial lake map in 2012. Techniques to fill the SLC-off gaps exist, but these create artificial values that will result in false detections of water. It is noted water mapping using multi-temporal time series images at large scales usually avoided the use of such techniques. Therefore, Landsat 7 ETM+ data with intensive slatted appearance is not really a good and suitable data for the classification of numerous of glacial lakes.

In this study, because the only useable data source for the year of 2012 is from Landsat 7 ETM+, to ensure continuity of annual data from 2008 to 2017, we have tried our best to extract the glacial lakes from the 2012 ETM+ images as accurately as possible. Errors caused by

striped gaps of Landsat ETM+ were manually corrected using additional high-quality scenes during the whole year with assistance of images from adjacent years.
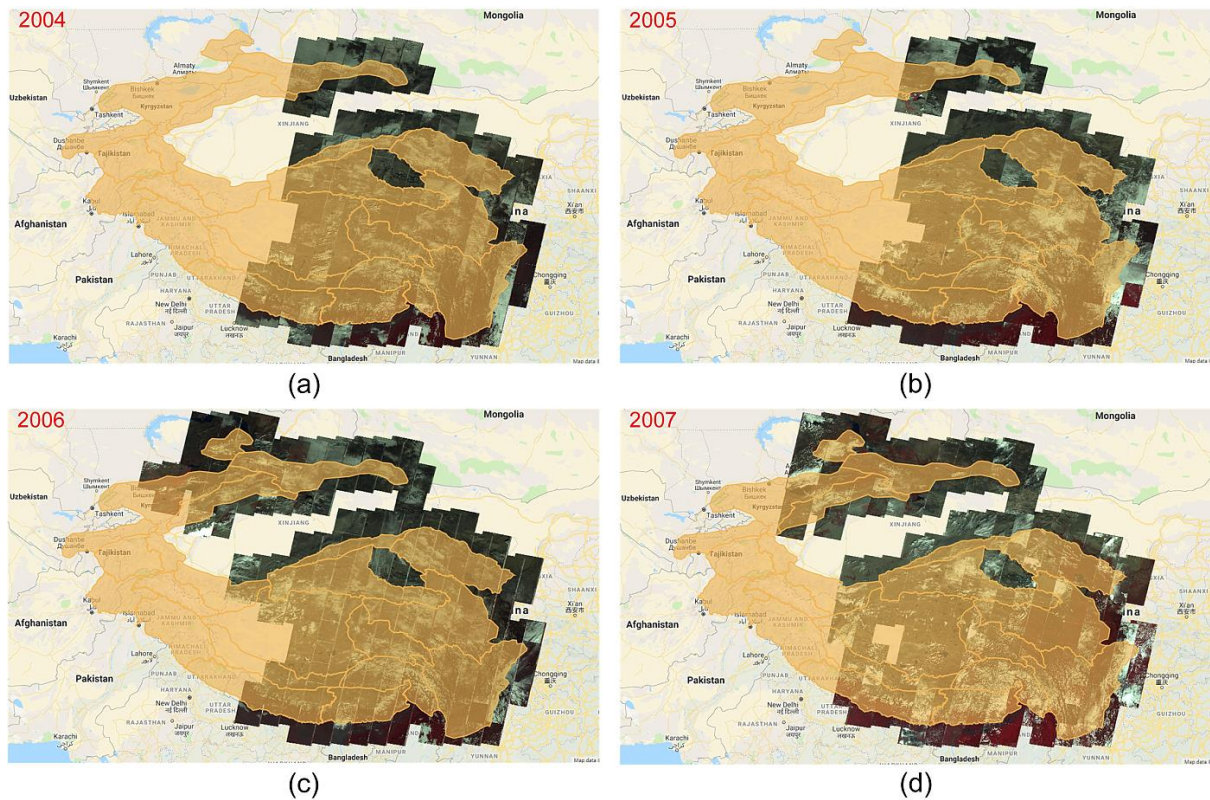

Figure II. All the available Landsat 5 TM dataset covering the whole HMA region for the years of (a) 2004; (b) 2005; (c) 2006 and (d) 2007.

Finally, in the revised manuscript, we have carefully modified related content by giving more detailed explanations about 'year-round' and why did not fill the years 2004, 2005, 2006, and 2007 with images from ETM+. The detailed modifications are as follows:

For the years before 2008, all the available Landsat 5 TM data in each year (e.g., 2004, 2005, 2006, and 2007) do not fully cover the HMA region.

Therefore, Landsat 7 ETM+ data with intensive slatted appearance is not really a good and suitable data for the classification of numerous of glacial lakes. In this study, because the only useable data source for the year of 2012 is from Landsat 7 ETM+, to ensure continuity of annual data from 2008 to 2017, we have tried our best to manually extract the glacial lakes from the 2012 ETM+ images as accurately as possible.

Errors caused by striped gaps of Landsat ETM+ were manually corrected using additional high-quality scenes during the whole year with assistance of images from adjacent years.

33. L73-74: 'in false detections of water' instead of 'in the false water detections'?

Response:

Thanks for this good suggestion. We have replaced 'in the false water detections' by 'in false detections of water' in the revised manuscript.

34. L72-73: 'were accurately classified but usually misclassified otherwise': not fully clear, please elaborate more clearly. Again, this sentence mentions classification, but any methodological background has not been given yet. The authors could consider to shift such sentences at appropriate locations down in the text.

Response:

We gratefully thank the Reviewer for his/her outstanding recommendations provided. In this study, lakes out of the gaps 'were accurately classified but usually misclassified otherwise' can be interpreted as lakes out of the gaps were accurately classified, but if glacial lakes are covered by gaps, they will be misclassified. Besides, this sentence is related with the lake classification indeed, which should belong to the content of Section 3. Methods other than 2.2 Dataset. We have moved such sentences to the **3.2 Extraction of glacial lake outlines** in the Section 3 of the revised manuscript.

35. L85-86: 'during stable seasons when lake extents are minimally affected by meteorological conditions and glacier runoff", which is 'from July to November': not sure whether this statement is true. Studies on supraglacial ponds (Miles et al., 2017b, 2017a) show that changes occur preferentially in that part of the year.

Response:

We greatly appreciate the Reviewer's valuable comments very much. To reduce the influence of seasonal lake fluctuations for the mapping, one effective solution is to map glacial lakes and measure their long-term changes during stable seasons. Zhang et al., 2015 conducted an inventory of glacial lakes in **the Third Pole** for ~1990, 2000, and 2010 using Landsat TM/ETM+ data (Zhang et al., 2015). The Landsat data selected for lake delineation were generally from the months July through November of ~1990, 2000, and 2010. During this period in each year, Landsat images featured less perennial snow coverage, and lake has a larger area following glacier-induced runoff and monsoon-induced precipitation. Nie et al., 2017 generated glacial lake extents in **the Himalaya** for 1990, 2000, 2005 and 2010 using automatic object-oriented mapping and human inspection methods, and in 2015 from Landsat 8 images using an advanced automated adaptive lake mapping method (Nie et al., 2017). Similarly, to reduce the effect of seasonal variability, images in similar seasons were selected, mainly in autumn and early winters (September to December), when glacial lake change is minor after the ablation period of glaciers. As the most highly variable glacial lakes in the study area, supraglacial lakes change preferentially in the year. For example, as shown in Miles et al., 2017a and Miles et al., 2017b, supraglacial pond cover in the **Langtang valley** shows an increase during the pre-monsoon, **rises to a peak in the early monsoon (June to July)**, drops during the post-monsoon and decreases to negligible during the winter months (Miles et al., 2017a;Miles et al., 2017b). Based on the above analyses about the selected lake mapping time in the different regions, in this study, the selected time series of Landsat data were generally from July to November when lake extents are minimally affected by meteorological conditions and glacier runoff. Most of the lakes also reached their maximum extent, specifically around the end of the glacier ablation season (June to August). If valid observations can not be obtained during this time span, then optimal mapping time needs to be broaden.

Furthermore, we gratefully thank the Reviewer for providing us with important references from Miles et al., 2017a and Miles et al., 2017b, they definitely have a certain guiding significance for our study. We have added these articles as references and providing some

explanations about the selected optimal mapping time in the **Section 3.1 Satellite imagery selection strategy** of the revised manuscript. The detailed modifications are as follows:

Here the selected time series of Landsat data were generally from July to November based on the analyses of mapping time of glacial lakes in the different regions. During this period of each year, the Landsat imagery featured less perennial snow coverage. Change in area of glacial lake is minor and has a large area following the glacier runoff and precipitation (Zhang et al., 2015;Nie et al., 2017).

As the most highly variable glacial lakes in the study area, supraglacial lakes change preferentially in the year, showing an increase in area during the pre-monsoon, and rising to the peak area in the early monsoon (June to July) (Miles et al., 2017a;Miles et al., 2017b).

36. L89: 'most-monsoon': did the authors mean 'post-monsoon'?

Response:

We are terrible sorry for this inaccurate expression. 'most-monsoon' do indeed mean 'post-monsoon'. We have corrected 'most-monsoon' to 'post-monsoon' in the revised manuscript.

37. L95: 'cloud score functions in GEE': did the authors make sure that this cloud score function works well in mountain headwaters with high snow and ice cover? This is a major challenge for cloud detection algorithms (Zhu and Woodcock, 2012).

Response:

We gratefully thank the Reviewer for his/her outstanding recommendations provided. We realize that this is a very important point that we may not explain clearly in the last version of manuscript. In this study, the selected time series of Landsat data for the glacial lake mapping were generally from July to November. During this period, the Landsat imagery featured less perennial snow coverage. Change in area of glacial lake is minor and has a large area following the glacier runoff and precipitation. This is the initial condition for the selection of images for glacial lake mapping in the HMA region. Although the selected image seasons are slightly different due to the meteorological conditions in different regions, they all comply with the same criterion that lake area were in clear-sky images and has small snow coverage, which will ensure the initial reliability of the mapping glacial lakes through GEE cloud computing platform. For the annual delineation of glacial lakes, the number of selected images is quite limited based on this prerequisite. If no valid observations can be obtained, then optimal mapping time needs to be broaden in the whole year.

To further increase data availability, and also as the basis for data selection in the periods beyond the optimum mapping time, we set two criteria for the selection of imagery with valid observations over the potential glacial lake area by using the cloud score functions in GEE, including (i) cloud cover is less than 20% in the 10 km buffer around each glacier outlines of a Landsat scene, or (ii) less than 20% cloud cover for the entire scene. As the reviewer pointed out that cloud score functions in GEE may face a big challenge to detect clouds in mountain headwaters with high snow and ice cover, large amounts of snow and ice are also likely to be identified as clouds using these cloud score functions. However, in this study, it is a much stricter criteria to filter out more images with lots of cloud-lookalike (snow/ice) bad observations and finally choose images with valid observations over the

potential glacial lake area. In addition to the selection of images for the extraction of glacial lakes by setting optimal mapping time and criteria of cloud score mentioned above, manual inspection and refinement of individual glacial lake is still an essential way to ensure the quality of inventory using high-quality scenes that were downloaded manually from USGS Earth Explorer website.

Finally, to give more comprehensive analysis about how cloud score functions in GEE were used for the image selection, we have completely modified the related content in **Section 3.1 Satellite imagery selection strategy** of the revised manuscript. The detailed modifications are as follows:

As the most highly variable glacial lakes in the study area, supraglacial lakes change preferentially in the year, showing an increase in area during the pre-monsoon, and rising to the peak area in the early monsoon (June to July) (Miles et al., 2017a;Miles et al., 2017b). Although the selected image seasons are slightly different due to the meteorological conditions in different regions, they all comply with the same criterion that lake area were in clear-sky images and has small snow coverage. This will ensure the initial reliability of the mapping glacial lakes through GEE cloud computing platform. If no valid observations can be obtained, then optimal mapping time needs to be broaden during the whole year.

To further increase data availability, and also as the basis for data selection in the periods beyond the optimum mapping time, we set two criteria for the selection of imagery with valid observations over the potential glacial lake area by using the cloud score functions in GEE, including (i) cloud cover is less than 20% in the 10 km buffer around each glacier outlines of a Landsat scene, or (ii) less than 20% cloud cover for the entire scene. The cloud score functions in GEE may face a big challenge to detect clouds in mountain headwaters with high snow and ice cover, large amounts of snow and ice are likely to be identified as clouds. However, in this study, it is a much stricter criteria to filter out more images with lots of cloud-lookalike (snow/ice) and finally choose images with good observations.

38. L95: 'a partial Landsat scene': unclear – please reformulate.

Response:

Thanks for this suggestion. Here 'a partial Landsat scene' refers to a part of a Landsat image that located in the 10 km buffer around each glacier outlines. We have modified this related sentence to make it clearly in the revised manuscript. The detailed modifications are as follows:

(i) cloud cover is less than 20% in the 10 km buffer around each glacier outlines of a Landsat scene.

39. L97: 'criteria' instead of 'criterions'?

Response:

Thanks for this suggestion. We have replaced 'criterions' by 'criteria' in the revised manuscript.

40. L97: 'needs to be broaden': by how many months? Possibly into the next year? Previous studies (Maharjan et al., 2018; Veh et al., 2018) have shown that some lakes could be veiled by clouds and ice for years.

We greatly appreciate the Reviewer's valuable comments very much. During the image processing using GEE, the automated selection of images for the extraction of glacial lakes is based on the optimal mapping time and criteria of cloud score. For the images without valid observations over the potential glacial lake area, then optimal mapping time needs to be broaden **during the whole year**.

As for the lakes that could be covered by clouds and ice for years, which will results in mapping errors in the dataset. On these conditions, a detailed manual editing is essential to map these glacial lakes. So we were not relying exclusively on automatic classification. Here the cloud or ice-covered glacial lakes that were missed in the automated processing were manually edited using additional high-quality scenes downloaded from USGS Earth Explorer website during the whole year with assistance of images from adjacent years.

Besides, in the section 3.1 and 3.2 of the revised manuscript, we have modified related content to clearly show the broaden time and how to map lakes that coved by ice and cloud for years. The detailed modifications are as follows:

If no valid observations can be obtained, then optimal mapping time needs to be broaden during the whole year.

For missing glacial lakes, such as some lakes may be covered by ice and clouds for years, grow at steep glacier tongues, and lakes show heterogeneous reflectance with the surrounding backgrounds, the lake boundaries were edited further using ArcGIS.

41. L98-100: Please reformulate this sentence, possibly splitting into two.

Response:

Thanks for this suggestion. We have reformulated this sentence by splitting it into two sentences. The detailed modifications are as follows:

Although the selected image seasons are slightly different due to the meteorological conditions in different regions, they all comply with the same criterion that lake area were in clear-sky images and has small snow coverage. This will ensure the initial reliability of the mapping glacial lakes through GEE cloud computing platform.

42. L104-106: Repetition, consider deleting.

Response:

Thanks for this suggestion. We have deleted this sentence in the revised manuscript.

43. L109: 'and SLC-off gaps'

Response:

Thanks for this suggestion. We have added the word 'and' before 'SLC-off gaps' in the revised manuscript.

44. L110: 'a more computationally efficient way': than?

Response:

We are so sorry for this inappropriate expression. We have modified this sentence to make it much rigorous. The detailed modifications are as follows:

Fmask has the advantage of being able to process a large number of images ==in a computationally efficient way==.

45. L109-L111: I suggest to switch the order of these two sentences.

Response:

Thanks for this good suggestion. We have switched the order of these two sentences in the revised manuscript.

46. L112: 'slopes (larger than 10°)': How did the authors make sure that this threshold does not mask water bodies that have grown at steep glacier tongues since the SRTM DEM was shot in 2000?

Response:

We greatly appreciate the Reviewer's valuable comments very much. Topographic shadows were masked using the slopes (larger than 10°) and shaded relief maps (value less than 0.25) calculated from SRTM data (Li and Sheng, 2012). This will remove considerable mountain shadows that have the similar spectral reflectance with water bodies. However, SRTM DEM was generated in 2000, which is different from the acquisition time of Landsat images used for the glacial lake mapping in this study, the derived slopes and shaded relief cannot fully represent the conditions on the date a given Landsat scene is acquired. As a consequence, some lakes that have grown at steep glacier tongue may be masked, and some mountain shadows that interfere with the mapping results of glacial lakes from GEE still remain, leading to the fact that glacial lakes in some steep areas are omitted, and some residual shadows are misclassified as glacial lakes. On these conditions, in order to further improve the accuracy of delineated glacial lake outlines, manual refinement of initial lake mapping results is necessary to supplement some missing lakes and correct the mapping errors such as mountain shadows using time series of high-quality Landsat images downloaded manually from USGS Earth Explorer website. For missing glacial lakes that have grown at steep glacier tongues, the lake boundaries were edited further using ArcGIS. False lake features, mainly identified as mountain shadows, were manually removed by overlapping mapped lake shorelines on the source Landsat imagery and higher-resolution imagery in Google Earth.

Besides, in the revised manuscript, we have added some content in the **Section 3.2 Extraction of glacial lake outlines** to clearly illustrate mapping errors induced by slope threshold of SRTM DEM and how to correct the mapping errors. The detailed modifications are as follows:

==However, SRTM DEM was generated in 2000, which is different from the acquisition time of Landsat images used for the glacial lake mapping in this study, the derived slopes and shaded relief cannot fully represent the conditions on the date a given Landsat scene is acquired. As a consequence, some lakes that have grown at steep glacier tongue may be masked, and some mountain shadows that interfere with==

47. L112: 'value less than 0.25': Unclear what this value means in physical quantities.

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out these issues. According to the results from Li and Sheng, 2012, DEM-derived surface gradients of most glacial lakes are less than $10°$, and the shaded reliefs are larger than 0.25 (Li and Sheng, 2012;Quincey et al., 2007). Generally, a lake is a flat surface at a constant water level, and the surface gradients across the lake are very slight. In contrast, shadow features are usually located on the dark side of the high mountains, where the sunlight is blocked, so their surface gradients are great and the terrain reliefs are small. Therefore, slope and shaded relief maps derived from DEMs can be used to mask the areas of topographic shadows. With these topographic features, the mountain shadow areas are defined as the places whose surface slopes are larger than $10°$ and shaded relief values are less than 0.25. Furthermore, we have added some text in the revised manuscript to give a clear explanation about meaning of 0.25 in physical quantities. The detailed modifications are as follows:

Topographic shadows are located in the areas where the sunlight is blocked, generally on the dark side of the high mountains, their surface gradients are great and the terrain reliefs are small. Therefore, topographic shadows were masked using the slopes (larger than $10°$) and shaded relief values (less than 0.25) calculated from SRTM data (Li and Sheng, 2012;Quincey et al., 2007).

48. L113: The paper from Quincey is from 2007.

Response:

We are terrible sorry for this mistake, we have corrected the year of this paper in the revised manuscript.

49. L115: Maybe 'acquired. Some'?

Response:

Thanks for this suggestion. We have modified it as 'acquired. Some' in the revised manuscript.

50. L116: 'remain' instead of 'remains', and 'was' instead of 'were'?

Response:

Thanks for this suggestion. We have replaced 'remains' by 'remain' and 'were' by 'was' in the revised manuscript.

51. L118: 'the relative stability': unclear, please reformulate.

L118: '0.0081 km² ': the reason for choosing this minimum mapping unit is not fully clear.

Response:

We gratefully thank the Reviewer for his/her constructive comments. The minimum number of pixels that used to define a glacial lake in the image was inconsistent in different studies. For example, Zhang et al., 2015 set the smallest detectable glacial lakes in the Third Pole of larger than 0.0027 km² (three connected pixels) using the Landsat TM/ETM+ data. Nie et al., 2017 selected 0.0081 km² (nine connected pixels) as the minimum mapping unit to map glacial lakes in the Himalaya. Other studies set the minimum threshold areas as 0.001km² (Salerno et al., 2012), 0.002km² (Wang et al., 2013a), 0.0036km² (Luo et al., 2020), 0.0054km² (Wang et al., 2020b), 0.01km² (Li et al., 2020) for the identification of glacial lakes and analysis of their spatial and temporal variations.

A smaller minimum mapping unit will detect a greater number of glacial lakes, however, it will also bring proportionally larger uncertainty than large lakes at the same resolution. This is because the uncertainty of measurement depends on the linear error and the perimeter of glacial lakes (Salerno et al., 2012). Our results demonstrate that a lake area covering fewer than nine lake water pixels have an area error of larger than 50% (Please see the **Section 4 Cross-validation and uncertainty estimate** in the revised manuscript). Given the area uncertainty of glacial lakes and the spatial resolution of Landsat data, in this study, glacial lakes larger than nine pixels ($\geqslant$ 0.0081 km²) were considered as the minimum mapping unit.

Finally, we have rewritten the related content to provide a fully clear reasons about the selection of 0.0081km² as the minimum mapping unit. The detailed modifications are as follows:

==To define a glacial lake in the image, the minimum number of water pixels was inconsistent in different studies. For example, Zhang et al., 2015 set the smallest detectable glacial lakes in the Third Pole of larger than 0.0027 km² (three connected pixels) using the Landsat TM/ETM+ data. Nie et al., 2017 selected 0.0081 km² (nine connected pixels) as the minimum mapping unit to map glacial lakes in the Himalaya. Other studies set the minimum threshold areas as 0.001km² (Salerno et al., 2012), 0.002km² (Wang et al., 2013a), 0.0036km² (Luo et al., 2020), 0.0054km² (Wang et al., 2020b), 0.01km² (Li et al., 2020) for the identification of glacial lakes and analysis of their spatial and temporal variations. A smaller minimum mapping unit will detect more glacial lakes, however, the uncertainty it brings is also larger than large lakes at the same resolution (Salerno et al., 2012). Our results demonstrate that a lake area covering fewer than nine water pixels have an area error of larger than 50% (Please see the Section 4). Given the area uncertainty of glacial lakes and the spatial resolution of Landsat data, in this study, glacial lakes larger than nine pixels ($\geqslant$ 0.0081 km²) were considered as the minimum mapping unit;==

52. L127: 'polygon include' instead of 'vector includes'?

Response:

Thanks for this suggestion. We have replaced 'vector includes' by 'polygon include' in the revised manuscript.

53. L128: 'distance to the nearest glacier terminus': how is this distance measured? Along the flow path or the Euclidian distance? What happens if there is no glacier upstream, given that the buffer around the glacier can extend into adjacent, yet glacier-free catchments?

Response:

We greatly appreciate the Reviewer's valuable comments very much. The distance from a glacial lake to the nearest glacier is calculated using the 'near' tool embedded in ArcGIS 10.6, which specifies a radius to search for candidate near glacier polygons, and then estimates the nearest Euclidian distance among them. For the glacial lakes that have close relationship with their parent glaciers, such as proglacial lakes, supraglacial lakes and ice-marginal lakes, the calculated distance to the nearest glacier using this analysis tool is reasonable since there is glacier in the upstream and the parent glacier is the nearest glacier to the glacial lake in the current catchment. For the unconnected glacial lakes that are not directly contacted with glacier, their interaction with glacier is also not that close compared with other types of glacial lakes, the lake water source are partly provided by glacier and partly by rainfall. Sometimes, maybe there is no glacier in the upstream, the distance calculated by 'near' tool is likely to be the distance to the nearest glacier in the adjacent catchments.

Besides, in the revised manuscript, we have modified related text to clearly show how to measure the distance between glacial lake and the nearest glacier. The detailed modifications are as follows:

(v) manual inspection and refinement of individual glacial lake were conducted and the related attribution (i.e., lake type, elevation, Euclidian distance to the nearest glacier terminus, area and perimeter) were added for each lake.

54. L135-146: It remains unclear, how the authors calculated the perimeter of a given lake in a given year, if there is more than one suitable satellite image in this year. Please elaborate more clearly.

Response:

We gratefully thank the Reviewer for his/her outstanding recommendations provided. We realize that this is a very important point that we may not explain clearly. If there are more than one suitable satellite images in a year, image with the least cloud cover will be selected for the calculation of the perimeter of a given lake. We have added some content in the revised manuscript to make it clearly. The detailed modifications are as follows:

Outputs per lake polygon include the information about lake type, elevation, Euclidian distance to the nearest glacier terminus, area and perimeter. Noted that if there are more than one suitable satellite images in a year, image with the least cloud cover will be selected for the calculation of the area and perimeter of a given lake.

55. L136: 'versus image date': the image dates are not consistent across the study region, not least because of atmospheric disturbances. What do the authors mean here?

Response:

We are terrible sorry for this improper expression, which easily make the readers to be misunderstood. Based on the final generated lake inventory data, in this study, the slope of

linear regression of lake area (over the grid cell of 1°×1°) **versus mapping year** was used to qualify the **yearly lake area changes**. Here the 'image date' does not represent the specific acquisition time of the image chosen for the mapping of glacial lakes, but the mapping year from 2008 to 2017. We have modified the related text in the revised manuscript to make it clearly. The detailed modifications are as follows:

Based on the final generated lake inventory data, we used the slope of linear regression of lake area (over the grid cell of 1°×1°) versus <mark>mapping year</mark> to qualify the yearly lake area changes during the study period.

56. L137: delete 'on'? 'chooses the median'?

Response:

Thanks for this suggestion. We have deleted 'on' and replaced 'chooses median' by 'chooses the median' in the revised manuscript.

57. L140-141: Li et al. (2019) had reported similar variance in the total annual lake area in the Tian Shan, and I am still wondering whether these changes are real (which means that a substantial area of glacier lakes must have shrunk in some years) or whether this is the result of 'adverse weather conditions, varying lake characteristics and image quality'. Given that the authors in this study manually checked and mapped each lake, it is surprising that the current maps still demand smoothing. At least the authors should be able to explain how much of the annual variance can be accounted to any of these three effects, and other effects such as ice or shadow on lakes, which the authors have not mentioned yet.

Response:

We greatly appreciate the suggestion of the Reviewer. In this study, we developed an annual 30-m dataset for glacial lakes in HMA region during 2008-2017. The slope of linear regression of lake area (over the grid cell of 1°×1°) versus mapping year was used to qualify the yearly lake area changes during the study period. Different linear trends estimation methods have been applied for the use in linear fitting of time series of data. Among these, Theil-Sen estimator chooses the median slope among all the derived fitted lines, can effectively represent long-term area changes due to its robustness for the trend detection and insensitivity to outliers. Although all the lake were manually checked and edited, due to the limitation of available images and other factors, the conditions of glacial lake mapping are not perfectly consistent for each year. For example, the image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from varying lake characteristics, image quality, ice and shadow that obscured the lakes, which all contributed to the detection error in the lake extent and their annual variation. In order to make the glacier lake area consistent year by year to the degree possible and estimate reliable annual changes in lake area, a high-performance Theil-Sen estimator is necessary for fitting and smoothing of the annual time series of data (Kumar, 1968;Song et al., 2018).

As for the **quantitative evaluation** of the influences from these three effects, and other effects such as ice or shadow on the annual variance of glacial lakes, generally, these factors do cause the errors in the detected glacial lake outlines and calculated area, which is objective and acceptable due to the nature of limited remote sensing data. While for this study, because we used long time series (10 years) data for the estimation of annual lake area change, and also

the errors only accounts for a small proportion of the total glacial lake area for each year, errors in the observed lake area due to the different effects do not apparently affect the trend statistical results. Quantitative analysis of the area error for every glacial lake caused by any of the above factors requires the examination and verification of all the dataset over a ten-year period, which is a large amount of work. Given the workload and the limitation of revision time, quantitative statistics can not be obtained at present, but a detailed analysis for measuring the area error and annual variance of glacial lakes resulted from the above factors is needed urgently in our future work.

Finally, in the revised manuscript, we have completely modified the related content to give detailed explanations about why the current maps demand smoothing and the influences from three effects, and other effects such as ice or shadow on the annual variance of glacial lakes. The detailed modifications are as follows:

Although all the lake were manually checked and edited, due to the limitation of available images and other factors, the conditions of glacial lake mapping are not perfectly consistent for each year. For example, the image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from varying lake characteristics, image quality (Bhardwaj et al., 2015;Thompson et al., 2012), ice and shadow that obscured the lakes, which all contributed to the detection errors in the lake extent and their annual variation. Generally, these errors are objective and acceptable as a result of the nature of the limited remote sensing data. For this study, because we used long time series (10 years) data for the estimation of annual lake area change, and also the errors only accounts for a small proportion of the total glacial lake area for each year, errors in the observed lake area caused by the different effects do not apparently affect the trend statistical results.

58. L146: maybe 'elimination of the effect from differences', or rephrase appropriately. And what do the authors mean with 'differences in the sensor capabilities'?

Response:

Thanks for this good suggestion. We have modified the 'elimination the effect of differences' as 'elimination of the effect from differences' in the revised manuscript. Here 'differences in the sensor capabilities' refer to the differences in the performance of different Landsat sensors for the mapping of glacial lakes, which can be manifested in many aspects such as the radiometric resolution, spectral resolution, spatial resolution and revisit time. For example, Landsat 8 has advantages in lake mapping over its ancestors. First of all, the 12-bit radiometric resolution and the improved sensitivity of Landsat 8's OLI sensor result in higher SNRs, essential to water body identification. Water bodies are dark objects with a reflectance usually lower than 10% in optical bands, and their mapping benefits greatly from these improvements. NDWI image segmentation is an effective water body mapping method, and histogram segmentation algorithms are widely used. Although NDWI enhances the difference between water bodies and land background, it is more sensitive to noise than the original green and NIR bands. Compared to Landsat 7 ETM+ (an 8-bit instrument), Landsat 8 OLI greatly improves its SNRs in the green and the NIR bands, from 37 to 304 and 34 to 201, respectively (Morfitt et al., 2015).

In the revised manuscript, we have rewritten the related sentence to make it clearly. The detailed modifications are as follows:

59. *L159: So n also contains m? Why is the perimeter not shown as an own variable?*

   L160: remove 'areal'.

Response:

Thanks for this good suggestion. Here n represents the number of pixels on the boundary of a glacial lake, which approximately equals to the ratio of the perimeter of a lake to the spatial resolution. m is the area of a pixel in the Landsat image (m^2). Therefore, n does not contains m. For better understanding, the perimeter can be expressed as its own variable, then the equation (1) is changed into:

$$A_{er} = 100 \cdot ((p/a)^{1/2} \cdot a^2) / A_{gl}$$

Where p refers to the perimeter of a glacial lake, a is the spatial resolution of the Landsat sensor (m), A_{gl} is the lake area (m^2) and the factor 100 is there to convert to percentage.

   However, this expression looks much more complicated in form than before and may be not ease of use. Therefore, we used the number of pixels to define the perimeter of a glacial lake, rather than the perimeter itself.

   Finally, in the revised manuscript, we have modified the description and explanation of each variable in the formula to make them much clearly and easier to understand. The detailed modifications are as follows:

   Where n refers to the number of pixels on the boundary of a glacial lake, approximated by the ratio of the perimeter length and spatial resolution, m is the area of a pixel in the Landsat image (m^2), A_{gl} is the lake area (m^2) and the factor 100 is there to convert to percentage.

60. *L170: The authors could explain more clearly what they mean by 'precision' here. It seems that 'precision' is a conversion of uncertainty, so that the issue of high uncertainties for small lakes remains.*

   L171: 'Krumwiede et al. (Krumwiede et al., 2014)': remove double reference.

   L173-174: Unclear what the authors wish to say here. What do they mean with 'predictably much better'? And how can they compare precision and accuracy, if these measures are not on the same scale?

Response:

We greatly appreciate the Reviewer's valuable comments very much and we are terrible sorry for these inappropriate expressions about 'precision', 'accuracy' and 'uncertainty'. As for the first point about what is the meaning of 'precision' here, it is a conversion of area uncertainty indeed, rather than the accuracy of glacial lake mapping results. Secondly, the reference regarding Krumwiede et al., 2014 appears repeatedly in the text and citation, and should be kept one. Finally, as for this sentence 'the precision of measurement (such as ability to detect area changes) is predictably much better than the accuracy', here the 'precision' and 'accuracy' are not on the same scale. What we originally meant was that because the type of data used in

this study are all Landsat images, the accuracy of detecting annual area change of glacial lake is better than uncertainty of lake area in a single year, this is just an inference, has not been confirmed by a scientifically systematic argument yet. The related words about '**predictably** much better' are really not rigorous, also the use of 'precision' and 'accuracy' are quite confusing, which easily make the readers to be misunderstood. Based on the above analysis, we have decided to delete these speculative sentences and inaccurate expressions in the revised manuscript.

61. L175: The authors could consider plotting the distributions of uncertainties (instead of a table) to highlight differences in the tails of uncertainties in each year? Do the differences in accuracy (e.g. lowest in 2010, highest in 2016) scale with the total area or number of glacier lakes?

Response:

We greatly appreciate the suggestion of the Reviewer. Assuming an uncertainty of 1 pixel for the detected glacial lake boundaries, we calculated the systematic errors of each glacial lake over the whole HMA region (Table 1). For the years between 2008 and 2017, generally, statistical results of area uncertainty (%) is similar, the area uncertainty of each glacial lake ranged from 0.30% to 50%, with the mean value of all the glacial lakes falling around the 17%, and standard deviation around 11%. The maximum and mean value of area uncertainty for the glacial lakes in 2010 are the lowest, while for the year of 2016, the corresponding statistics are highest, this can be attributed to the different factors. The maximum of area uncertainty of glacial lakes is related with the shape and size of a certain lake (as can be seen from equation (1)), but its mean value is equal to the sum of the area uncertainties of each glacial lake divided by the total number, which depend on the total number of glacial lakes in a year, and also the shape and area of each lake

In order to highlight differences in the tails of area uncertainties in each year, following the Reviewer's very pertinent recommendation, in the revised manuscript we have added a new figure (Fig. 4a) that plotted the distributions of uncertainties from 2008 to 2017. Fig. 4b was also added for the better visualization of the relationship of area uncertainties against areas of all the glacial lakes in HMA (taking the results in 2017 as an example).

Besides, to give more detailed and comprehensive explanations about differences in the tails of area uncertainties of glacial lakes in each year, we have added some content in the **Section 4 Cross-validation and uncertainty estimate** of the revised manuscript. The detailed modifications are as follows:

Assuming an uncertainty of 1 pixel for the detected glacial lake boundaries, we calculated the systematic errors for the whole HMA region (Fig. 4). For the year between 2008 and 2017, the area uncertainty of each glacial lake generally ranged from 0.30% to 50%, with the mean value falling around the 17%, and standard deviation around 11% (Fig. 4a). The maximum and mean value of area uncertainty for the glacial lakes in 2010 are the lowest, while for the year of 2016, the corresponding statistics are highest, this can be attributed to the different factors. The maximum of area uncertainty of glacial lakes is related with the shape and size of a certain lake (as can be seen from equation (1)), but its mean value is equal to the sum of the area uncertainties of each glacial lake divided by the total number, which depend on the total number of glacial lakes in a year, and also the shape and area of

each lake. Besides, a close relationship can be found between the area uncertainties and the sizes of the glacial lakes (Fig. 4b). Most of the large glacial lakes (area ≥ 0.04km^2) have the mean area uncertainty of about 7%. This systematic errors were more significant for the small-sized glacial lakes. We measured glacial lake down to 0.0081 km^2 (nine pixels in Landsat imagery), where systematic errors calculated by equation (1) were ~50%.
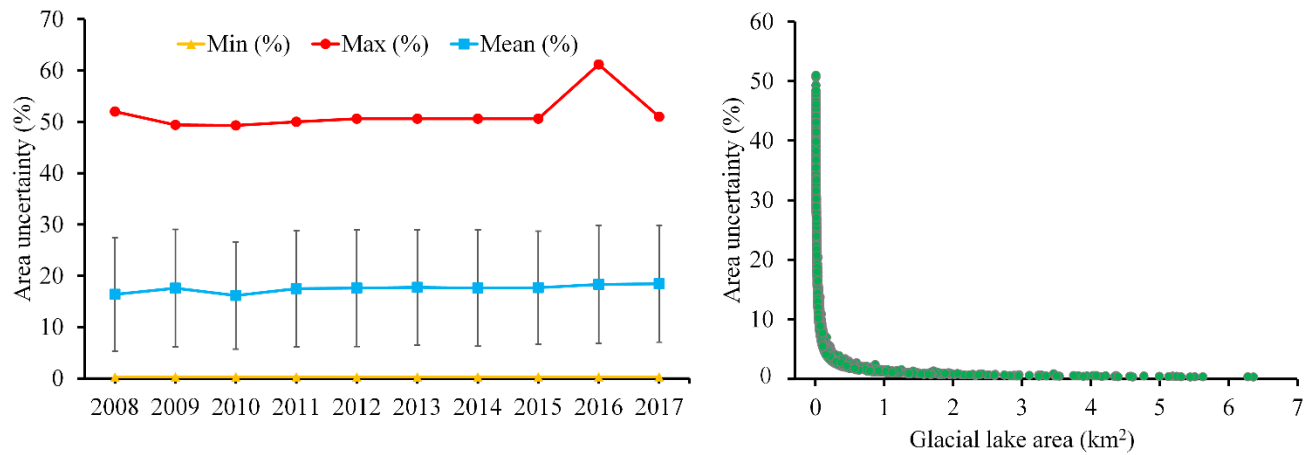

Fig. 4. (a) Statistics of area uncertainty (%) of glacial lakes for the years from 2008 to 2017. (b) Relationship between area uncertainties and areas of all the glacial lakes in HMA in 2017.

62. L178: The information from this sentence should go into the abstract.

Response:

Thanks for this good suggestion. We have moved the information from this sentence into the abstract of the revised manuscript. The detailed modifications are as follows:

It can be observed that glacial lakes exhibited total area increases of 90.14 km^2 between 2008-2017, a +6.90% change relative to 2008 (1305.59 ±213.99 km^2).

63. L179: 'A linear least-squares': Why did the authors refrain from using the Theil-Sen regression here?

Response:

We gratefully thank the Reviewer for his/her constructive comments. Based on the final generated lake inventory data, we used the slope of linear regression of lake area (**over the grid cell of 1°×1°**) versus mapping year to qualify the yearly lake area changes during the study period. The approach to change analysis was predicted using a Theil-Sen estimator, which chooses median slope among all the derived fitted lines to smooth the annual time series of data. **For the whole HMA region**, we applied a linear least-squares fit to all the data to derive the annual increment of glacial lake area and number for the 10-year record. Due to the robustness for the trend detection and insensitivity to outliers, Theil-Sen estimator can effectively represent long-term area changes, it is also useful for the elimination of the effect

from differences in the sensor performance for the mapping of glacial lakes. Therefore, compared with conventional linear trends (such as linear least-squares), Theil-Sen estimator show better performance and is much more reliable to estimate year-to-year changes for the annual layers.

In the last version of manuscript, we are terrible sorry for our careless use of inconsistent regression fitting methods. In the **Section 5.1 Distribution of various types and sizes of glacial lakes** of revised manuscript, we have used the Theil-Sen regression method to estimate the yearly changes in lake area and number over the whole HMA region, Fig. 3a and related statistics have also been modified. The detailed modifications are as follows:

A Theil-Sen regression fit to all the data showed a mean expansion rate of 12 km^2 a^{-1} for the 10-year record (Fig. 4a). Meanwhile, the estimated changes in glacial lake number from 2008 (12,593 lakes) to 2017 (15,348 lakes) showed an average increase of 306 lakes a^{-1}.
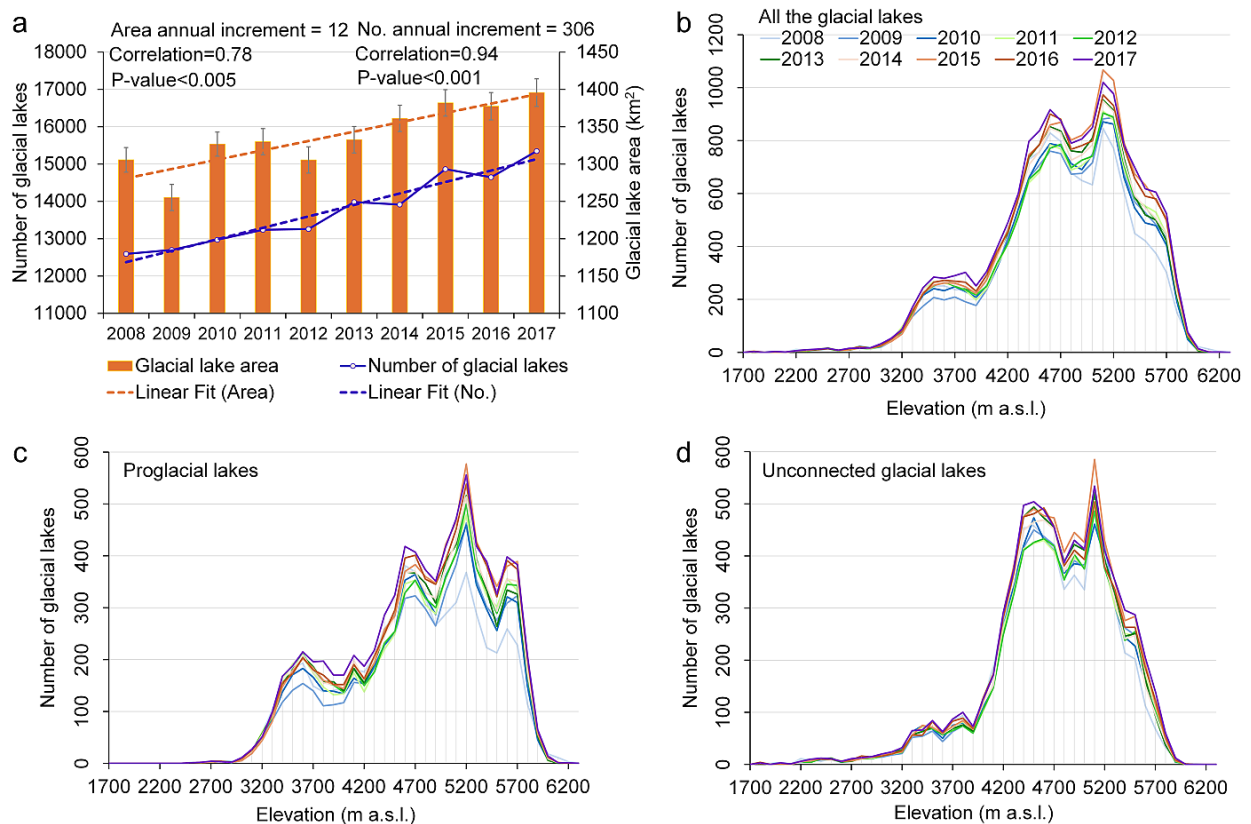

Fig. 4. Annual glacial lake number and area. (a) Total number and area of glacial lakes for HMA between 2008-2017. The annual increment is the slope of the trend of annual lake area and number. Altitudinal distribution (100-m bin sizes) of lake numbers for **(b)** all glacial lakes, **(c)** proglacial lakes, and **(d)** unconnected glacial lakes.

64. L179 & L181: '18 km^2 a^{-1}' & '380 lakes a^{-1}': Please add confidence bands. In principle, all numbers and changes in Chapter 5.1 need error bounds because of the mapping uncertainty.

Response:

We greatly appreciate the suggestion of the Reviewer. As for your first comment about adding the confidence bands for annual increase in lake area (18 km^2 a^{-1}) and number (380 lakes a^{-1}), we have added the computed P-value of the derived trends for the time-series glacial lake area and number to Fig. 3a. Besides, following the Reviewer's very pertinent recommendation, in the **Section 5.1 Distribution of various types and sizes of glacial lakes** of revised manuscript, we have completely modified the related statistics and trends by adding the error bars to area (both in the text and Fig. 3a), and confidence intervals to the trends and changes throughout the text. Please see the revised manuscript for more detailed information.

65. *L182: 'Fig. A3': unclear why this figure does not show all lake sizes. This could support the argument of the authors that smaller lakes have increased more in number than large lakes. On the other hand, do these two plot suggest that most of the increase in lake area is tied to the growth of large lakes? Data points in A3 could also have uncertainty bars given the mapping errors?*

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. According to the response to your Comment 2, our team have been working hardly to update data of glacial lakes larger than 0.04 km^2 in the ten year records, and add their attribute information. Based on the new lake inventory, the related statistics, figures and analysis in the article have also been modified accordingly, which mainly focus on the relatively large lakes (area larger than 0.04 km^2). Meanwhile, we have weaken the analyses and conclusions about lakes that have an area less than 0.04 km^2 to the highest degree possible. Therefore, Fig. A3 has been deleted in the revised manuscript.

66. *L183-184: 'The increase of proglacial lakes was concentrated above 4900 m (Fig. 3c)': unclear whether this clearly follows from this plot, could be also at 4600 or 5600 m asl. Plotting the absolute differences of lake number for each elevation band could help*

Response:

We greatly appreciate the Reviewer's valuable suggestions very much. In the last version of manuscript, this conclusion may be not easily drawn from this plot indeed. Following the Reviewer's very pertinent recommendation, we have plotted the absolute differences of lake number (2008-2017) for the proglacial lakes along each elevation band, as shown in Figure III. From Figure III, we can clearly see that the number of proglacial lakes increased slightly between 3700m and 4200m during 2008-2017, the increase of was mainly concentrated above **4900 m**, indicated as the red line.
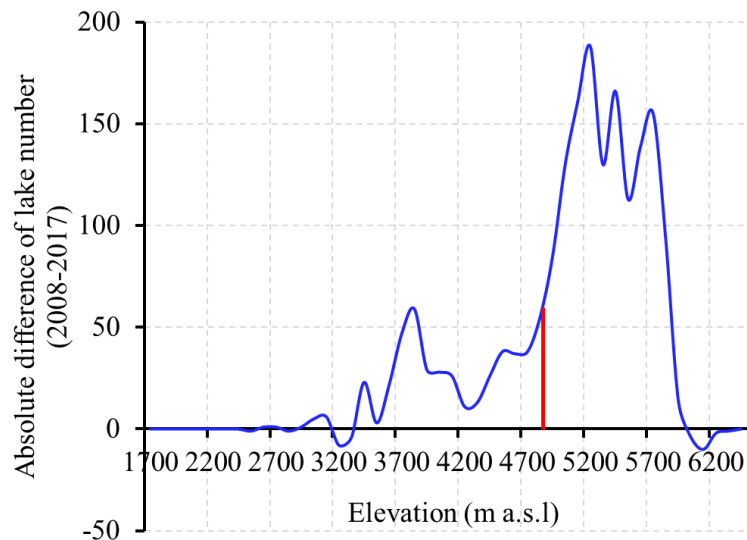
Figure III. Absolute differences of lake number (2008-2017) for the proglacial lakes along each elevation band. The red line indicates that the increase of lake number was mainly concentrated above 4900 m.

67. L186: 'supraglacial lakes': why are these not shown here? Given the downwasting of debris-covered glacier in the Himalayas, for example, this type of lakes could account for a substantial amount of the increase in total lake number, no?

Response:

We gratefully thank the Reviewer for his/her outstanding recommendations provided. With the accelerate downwasting of debris-covered glacier in the HMA region, a large amount of glacial lakes have emerged and expanded. Supraglacial lakes is one type of glacial lakes that formed in depressions on low-sloping parts of the surface of a melting glacier and are dammed by ice or the end-moraine or stagnating glacier snout. Typically, these lakes are highly dynamic, even ephemeral, and are quickly filled by glacier meltwater and quickly evaporated or drained off. They are **seasonal lakes**, which continue to change in their distribution and inundation area during a year. Mapping of supraglacial lakes creates great challenge for this study because the results largely depend on the annually available and valid observational data. If **images of different seasons or periods in a year** were used, the extraction results of supraglacial lakes are quite different.

In this study, we produced the glacier lake dataset in HMA region **in annual time steps** from 2008 to 2017. The optimal images with valid observations over the potential glacial lake area is very limited in each year. Moreover, the image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from image quality, ice and shadow that obscured the lakes, all greatly influence the results in the lake extent and number of supraglacial lakes. For such an annually resolved inventory, the obtained information about supraglacial lakes may be unstable. Therefore, the related statistics, figures and analysis in this section are mainly focused on the proglacial lakes and unconnected glacial lakes. As a consequence, supraglacial lakes were not shown here for the analysis of their altitudinal distribution and changes of lake numbers.

68. L190-Fig 3a: why have some 1,000 lakes vanished between 2008 and 2009?

We greatly appreciate the Reviewer's valuable comments very much. Based on the new lake inventory, the number of glacial lakes in the whole HMA region are 12593 for 2008, and 12698 for 2009, respectively, increased by a total of 105 glacial lakes.

69. L194-195: Which 'non-parametric trend analysis' did the authors use?

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. Here 'non-parametric trend analysis' refers to Theil-Sen regression analysis introduced in the Section 3.3. We are so sorry for this unclear expression, and we have modified this sentence to make it much clearly. The detailed modifications are as follows:

Annual changes in glacial lakes were further analyzed spatially using a 1°×1° grid over 22 mountain regions (Fig. 5a) using <mark>Theil-Sen regression analysis.</mark>

70. L200: 'large patches': unclear formulation, please elaborate.

Response:

Thanks for this good suggestion. We have modified the related sentence to make it clearly. The detailed modifications are as follows:

Glacial lakes in Nyainqentanglha, Gangdise Mountains <mark>exhibited area loss and gain in some regions</mark>.

71. L201-203: What is the reason for the two bubbles with the large decrease in lake area (Fig. 4b), while all surrounding bubbles show a positive trend? Is this because one (or more) large lake(s) shrunk or burst out?

Response:

We gratefully thank the Reviewer for his/her constructive comments. The two bubbles with the large decrease in lake area are indeed caused by the shrunk of few large glacial lakes. Subtle changes in the large glacial lakes can have a big impact on the annual area changes of glacial lakes over the grid cell of 1°×1°. In the last version of manuscript, some huge glacial lakes with an area of larger than $6.5km^2$ have also been mapped in the Hi-MAG database, this is because these lakes all located in the range of 10 km from the nearest glacier terminus according to the definition of glacial lake in HMA. However, in the dataset produced by Wang et al., 2020, these huge lakes were removed. For the sake of a reliable comparative analysis between different studies under the equal conditions, glacial lakes with an area of larger than $6.5km^2$ were not considered for the calculation and analyses in the revised manuscript. Based on the new lake inventory, Fig. 4 and related statistics and analysis have also been completely modified, the abnormal bubbles in Fig. 4 no longer exist. The detailed modifications of Fig.4 are as follows:
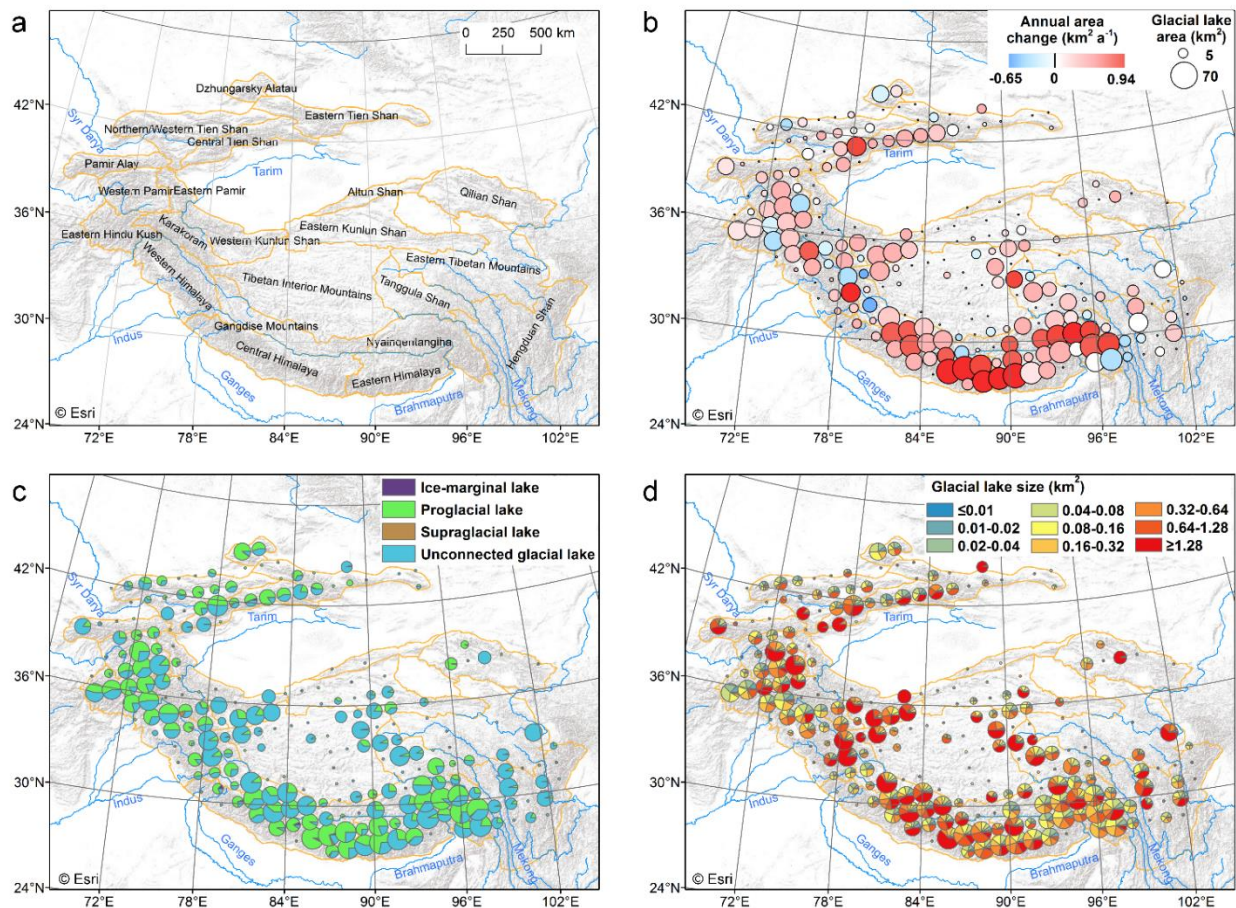
72. L210-211: What is happening in regions with shrinking lakes? Do some of these fall dry or burst out?

Response:

We greatly appreciate the Reviewer's valuable comments very much. As for the regions with negative lake growth, during the past ten years, many glacial lakes experienced shrinkage, dry up or outburst events. For example, some seasonal and highly dynamic lakes such as supraglacial lakes, their shapes and extent continue to change during a year. Because of the high sensitivity to regional climate and glacier runoff, their life is short, even ephemeral, and are quickly filled by glacier meltwater and quickly evaporated or drained off. Consequently, some supraglacial lakes drop during the post-monsoon and decreases to negligible during the winter months. Unconnected glacial lakes have been detached from ice contact due to glacial recession. As the interaction with the glacier gradually weakened, part of the water source supplied by glaciers is reduced. Also, combined with the effects from atmospheric warming and decrease of precipitation, regions mainly consist of unconnected glacial lakes exhibit the negative area change.

The above is just some general and global analysis. Based on the results obtained in this study, for a better understanding of how many glacial lake burst out, how many glacial lakes

decreased in area, as well as the reason and degree of shrinkage of lakes in the particular regions, it is necessary for us to conduct some more in-depth discussions and analyses in our future studies.

Finally, in the **Section 5.1 Distribution of various types and sizes of glacial lakes** of revised manuscript, we have added some content to provide more detailed information about what is happening in regions with shrinking lakes. The detailed modifications are as follows:

==In the negative lake growth (shrinkage) regions of Eastern Tibetan Mountains and Hengduan Shan, the unconnected glacial lakes were dominantly occupied. As the interaction with the glacier gradually weakened, part of the water source supplied by glaciers is reduced. Also, combined with the effects from atmospheric warming and decrease of precipitation, regions mainly consist of unconnected glacial lakes show the decreasing trend in area.==

73. L211-214: Very nice findings. These numbers could find their way into the abstract?

Response:

Thanks for this good suggestion. We have moved the information from this sentence into the abstract of the revised manuscript. The detailed modifications are as follows:

==Our results demonstrate proglacial lakes are a main contributor to recent lake evolution in HMA, accounting for 62.87% of the total area increase (56.67 km^2). Proglacial lakes in the Himalaya ranges alone accounted for 36.27% of the total area increase (32.70 km^2).==

74. L215: 'the largest area growth of lakes occurred in areas with relatively large proportion of small glacial lakes': is there a way to measure this notion?

Response:

We greatly appreciate the Reviewer's valuable comments very much. In the last version of manuscript, the related statement is not rigorous indeed, which needs further support from detailed statistics. Therefore, following the Reviewer's very permanent recommendations, we have calculated the area of different sizes of glacial lakes in 2017 for the regions with large area growth of rate (Table I). From Table I, it can be found that in the areas that show rapid lake area increases (higher than 0.23 km^{2}a^{-1}), glacial lakes with a size of less than 0.16 km^2 occupied more than 30% of total area. Especially for some of these regions, the area of small glacial lakes (≤0.16 km^2) even accounts for 69.47% of the total area.

In addition, we have modified this sentence by providing more statistical support in the revised manuscript. The detailed modifications are as follows:

==We also noted the large area growth of lakes occurred in areas with relatively large proportion of small glacial lakes, mainly due to rapid growth of existing lakes and new lake formation (Fig. 5d). For example, in some areas of Central and Eastern Himalaya, Nyainqentanglha that have large annual increase in lake area (higher than 0.23 km^{2}a^{-1}), glacial lakes with a size of less than 0.16 km^2 occupied more than 30% of total area. Especially for the region in Nyainqentanglha, the area of small glacial lakes (≤0.16 km^2) even accounts for 69.47% of the total area (Table A3).==

Table A3. Area of different sizes of glacial lakes in 2017 for some regions with large area growth of rate. The unit of area is km^2.

Lake grid ID (Mountain ranges)	69 (N)	116 (CH)	274 (WH)	71 (N)	48 (H)	74 (N)	72 (N)	14 (EH)	13 (EH)	39 (CH)	15 (EH)
≤0.01km^2	0.18	0.28	0.20	0.16	0.22	0.16	0.17	0.23	0.16	0.32	0.33
0.01km^2-0.02km^2	0.85	1.51	1.29	0.71	1.43	1.08	1.37	1.45	1.45	1.49	2.46
0.02km^2-0.04km^2	1.69	2.16	2.22	1.79	3.24	2.09	2.29	2.24	2.06	2.72	4.14
0.04km^2-0.08km^2	1.78	3.19	2.98	2.20	5.30	3.38	4.45	2.77	2.69	3.66	7.16
0.08km^2-0.16km^2	1.91	5.38	3.87	2.86	4.81	4.03	5.06	3.75	4.33	5.00	13.16
0.16km^2-0.32km^2	1.81	4.53	2.23	2.76	4.62	5.55	5.81	2.91	3.90	5.66	11.62
0.32km^2-0.64km^2	1.01	5.37	1.77	1.79	3.88	1.75	3.81	5.72	3.99	7.13	12.37
0.64km^2-1.28km^2	0.00	2.94	0.00	1.38	2.82	2.96	4.43	0.96	7.10	8.97	7.74
≥1.28km^2	0.00	7.22	4.19	0.00	11.46	3.17	2.59	6.07	1.40	6.06	12.00
Total area (km^2)	9.22	32.58	18.76	13.66	37.76	24.17	29.99	26.10	27.09	41.00	70.96
Total area (≤0.16km^2)	6.41	12.52	10.57	7.72	14.99	10.74	13.35	10.45	10.69	13.18	27.24
Total area % (≤0.16km^2)	69.47	38.43	56.32	56.56	39.70	44.45	44.52	40.03	39.47	32.15	38.39
Annual area increase (km^{2}a^{-1})	0.23	0.28	0.28	0.29	0.32	0.41	0.42	0.49	0.70	0.74	0.94

75. L216-217: 'many large-lake dominated areas exhibited decreased or nearly unchanged lake extent': not sure whether Fig. 4 fully supports this conclusion. In the Eastern Himalayas, for example, small glacier lakes account for less than 10 % for the entire lake area in that region (Table A4), so that most of the growth must be tied to the largest lakes, right?

Response:

We gratefully thank the Reviewer for his/her comments and suggestion. According to the response to your Comment 71, in the last version of manuscript, some huge glacial lakes with an area of larger than 6.5km^2 have also been mapped in the Hi-MAG database. Subtle decrease or changes in the large glacial lakes can have a big impact on the annual area changes of glacial lakes over the grid cell of 1°×1°. Therefore, for the areas that large-lake dominantly occupied, decreased or nearly unchanged lake extent can be observed. In the updated version of Hi-MAG database, glacial lakes with an area of larger than 6.5km^2 were all deleted, and this conclusion is correspondingly weakened.

76. L218-222: Unclear content in these sentences, please reformulate

Response:

Thanks for this suggestion. Similar to the response to your Comment 75, the conclusions and content shown in these sentences are also weakened based on the updated version of lake inventory. Consequently, we have deleted these sentences in the revised manuscript.

77. L230-Fig4b: not sure whether annual trends are useful here because, as far as I understood, the authors assume that lake area scales linearly with time in each 1∘x1∘ bin. Yet the authors provide neither a <u>measure of uncertainty for these annual changes</u>, nor is it clear <u>whether these changes occur linearly in time</u> or that the rate of change could be affected by <u>outliers</u>. <u>Total changes in glacier lake area between 2008 and 2017 could probably make more sense here</u>.

Response:

We greatly appreciate the Reviewer's valuable comments very much. We realize this is a very important point that we may not explain clearly in the last version of manuscript.

According to your comment about whether the glacial lake area occur linearly with time, also for the long time series of analysis of lake changes over a large scale area, Zhang et al., 2019 studied the regional differences of lake evolution across China during 1960s–2015, and confirm that multi-decadal lake measurements shown **significant linear regression trends, which was then used for the estimation of lake area change per year** (Zhang et al., 2019). In this study, annual changes in glacial lakes were analyzed spatially using a 1°×1° grid over 22 mountain regions using Theil-Sen regression analysis. Theil-Sen estimator chooses the median slope among all the derived fitted lines, can effectively represent long-term area changes due to its robustness for the trend detection and **insensitivity to outliers**. As for the measure of uncertainty for these annual changes, a **nonparametric Mann-Kendall trend test was used to detect and further confirm the statistical confidence or significance level by the linear regression results**. All the estimated trends fall within the **90% confidence intervals**. The upper and lower change estimates that satisfy the 90% confidence interval for the slope were also derived over the whole HMA region, ==please see Fig. A2== for the detailed information.

Just as the reviewer pointed out, the calculated rate of changes in the glacial lake area is the approximate estimations of the linear trend. Total changes in glacier lake area between 2008

and 2017 is measured in absolute values, and probably more accurate in quantitative degree. While in this study, we produced an annually resolved inventory, aiming to investigate the detailed evolution patterns and annual time series of glacial lake changes in the whole HMA for the past several years. The total changes was only calculated based on the two time periods (2008 and 2017), which not fully considered the mapping results of glacial lake in annual time steps from 2008 to 2017. More importantly, for the roughly weigh and comparison the **annual loss and gain** of glacial lake area in different regions, we used robust linear regression to calculate the yearly lake area changes over each grid cell of 1°×1°, which is not possible with simple two-stage lake mapping.

Finally, in the revised manuscript, we have added some content to give more detailed information about the measure of uncertainty for the annual changes in glacial lake area and significance test of the linear regression results. The detailed modifications are as follows:

==For the glacial lake area time series in each 1°×1°grid, we applied the Theil-Sen estimator to smooth the annual time series of data and derive the slope (annual change) of the trend. A Mann-Kendall trend test was used to detect and further confirm the statistical confidence by the linear regression results. All the estimated trends fall within the 90% confidence intervals. The upper and lower change estimates that satisfy the 90% confidence interval for the slope were also derived over the whole HMA region (Fig. A2).==
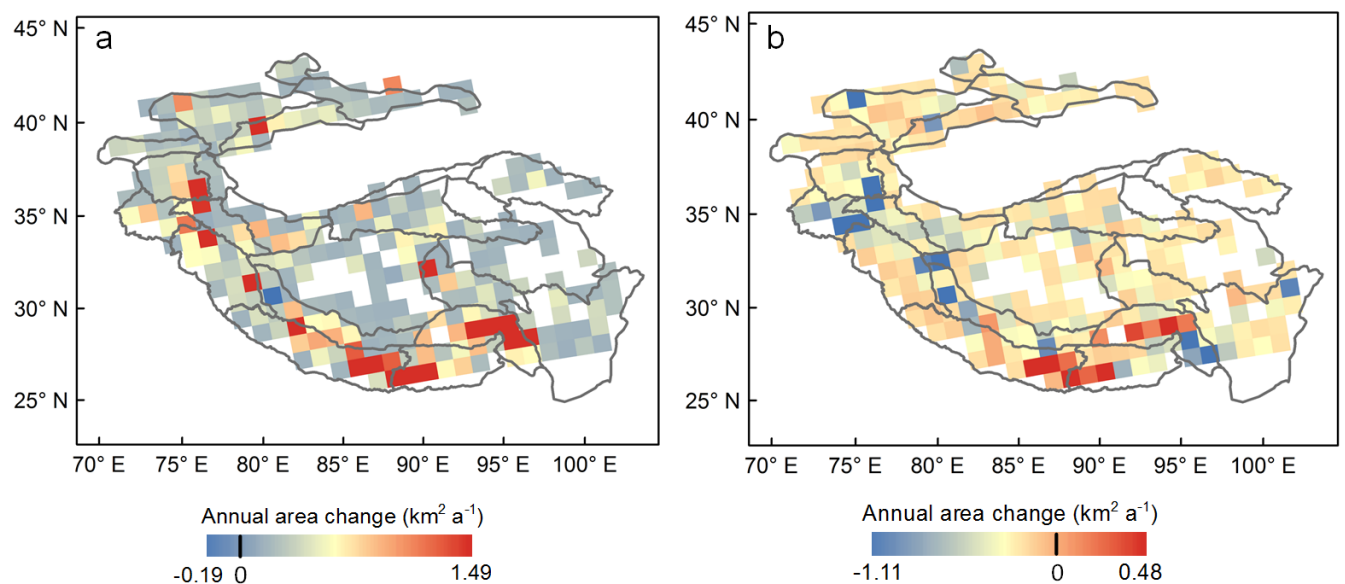

Fig. A2. Annual changes in lake area between 2008 and 2017 on a 1°×1° grid. The (a) upper and (b) lower slopes represent the 90% confidence interval.

78. L203: where does Fig. A4 show that the increase in lake area is 'due to retreat and thinning of debris-covered glaciers'?

Response:

We are terrible sorry for the mistake about the wrong location of Fig. A4. Fig. A4 shows the density (number per 100 km²) distribution of glacial lakes in 2017, and should be put after the

sentence 'exhibiting both a high density of 46 glacial lakes per 100 km^2 in 2017'. The detailed modifications are as follows:

Between 2008 and 2017 Central Himalaya's glacial lake area increased by 27.09 km^2 (Table A1), exhibiting both a high density of 47 glacial lakes per 100 km^2 in 2017 ==(Fig. A4)== and rapid growth, +0.94 km^2 a^{-1}, in lake area due to retreat and thinning of debris-covered glaciers (Song et al., 2016).

79. L207-208: Fig. 4c does not show any trends

Response:

We are sorry for missing another figure (Fig. 4b) here. In the last version of manuscript, it can be seen from Fig. 4b and Fig. 4c that glacial lakes exhibited different expansion trends for different lake types and supraglacial and ice-marginal lakes have relative few coverage areas comparing with proglacial and unconnected lakes. In the revised manuscript, we have added Fig. 4b and Fig. 4c after this sentence to make it easy to understand. The detailed modifications are as follows:

==We found that glacial lakes exhibited different expansion trends for different lake types and supraglacial and ice-marginal lakes have relative few coverage areas comparing with proglacial and unconnected lakes (Fig. 5b and Fig. 5c).==

80. L220-227: OK – and what about the Pamir Alay, Western & Eastern Kunlun, and Eastern Himalaya, where the contribution to the total growth in lake area is much smaller from small lakes compared to large lakes, according to Table A2? What I learned from this table is, that there is some variability in the lake-size distribution, which itself could be more emphasized. I have more the feeling that there are few (or no?) general rules that could help explaining the regional distribution of glacier lakes and associated growth patterns. If that feeling is wrong, I would appreciate simple correlation plots that check for such notions.

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. According to the response to your Comment 2, our team have been working hardly to update data of glacial lakes larger than 0.04 km^2 in the ten year records, and add their attribute information. Based on the new lake inventory, the related statistics, figures and analysis in the article have also been modified accordingly, which mainly focus on the relatively large lakes (area larger than 0.04 km^2). Meanwhile, we have weaken the analyses and conclusions about small glacial lakes that have an area of less than 0.04 km^2 to the highest degree possible. Therefore, Table A2, Table A3 and this paragraph have been deleted in the revised manuscript.

Just as the reviewer pointed out that the variability in the lake-size distribution itself could be more emphasized, rather than digging into few general rules that explaining the regional distribution of glacier lakes and associated growth patterns. Following the Reviewer's very pertinent recommendations, we have made efforts to thoroughly modify the whole text and paid more attention to analyze the spatiotemporal variability in the area and number of the glacial lake itself during the past ten years.

81. L230: Fig 4b: Colors around zero could be changed to white and negative value to a darker shade of blue for more contrast in the bubbles.

Response:

Thanks for this good suggestion. For better contrast between positive and negative values in the bubbles, we have modified Fig. 4b by changing the colors around zero to white and negative value to a darker shade of blue. The detailed modifications are as follows:
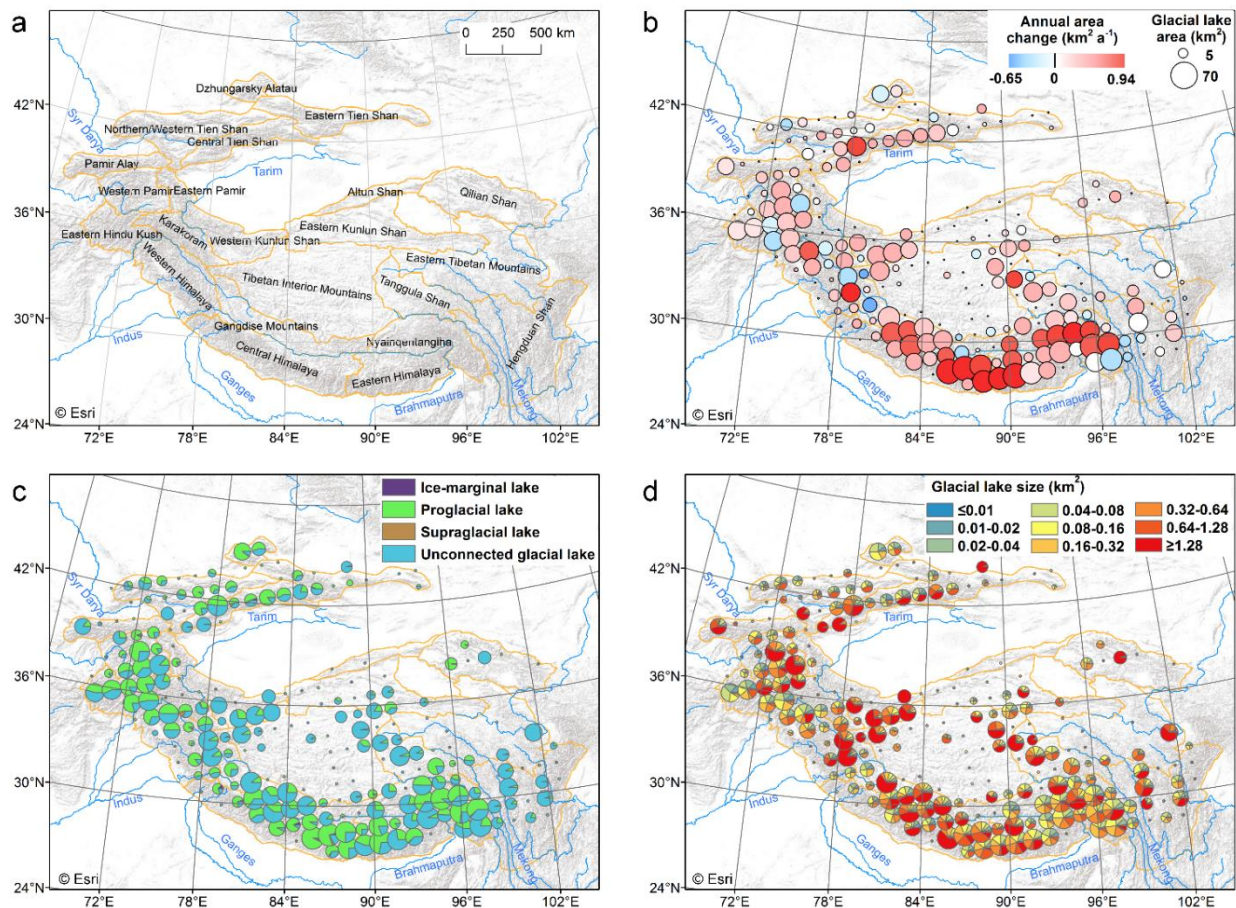

Fig. 5. Glacial lake area changes and area distribution. (a) Geographic coverage of mountain ranges in HMA. **(b)** Annual rate of change in lake area (2008-2017) on a 1°×1° grid. The size of the circle for the area in 2017. **(c)** Proportional areas of four types of glacial lakes in 2017. **(d)** Area of different sizes of glacial lakes in 2017. The terrain basemap is sourced from Esri (© Esri).

82. L242-244: Could it be that the number of lakes in the smallest size class (< 0.01 km^2) remains constant in most regions because these lakes climb into the next higher size class from year to year? If this holds true, could this mean that the lowest size class tells us more about the annual production rate of lakes in that region?

L247-248: Binning should be equal, so that size classes are comparable. Now, the smallest class spans 0,0919 km^2, the medium 0.01 km^2, and the largest 0.02 km^2. Linear regression should have confidence bounds, given that the positive trends in some panels (e.g. K, CT, WH, EHK) could be due to outliers. All panels could have the same Y-axes to show differences in lake abundance between regions

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. Similar to the response to your Comment 2, Comment 65 and Comment 80, based on the new lake inventory, the related statistics, figures and analysis in the article have been modified accordingly, which mainly focus on the relatively large lakes (area larger than 0.04 km^2). Meanwhile, we have weaken the analyses and conclusions about small glacial lakes that have an area of less than 0.04 km^2 to the highest degree possible. Therefore, information about **Section 5.2 Inter-annual variability of small glacial lakes for different mountain regions** have been deleted in the revised manuscript. But in our future work, our team will make great effort to supplement and improve the analysis about tiny glacial lakes, and continuously optimize and share glacial lake data with long time scales.

83. L256: R^2=0.53: I couldn't find this number in Table A5. Also, which type of correlation coefficient is calculated here? The caption of table A5 also talks about R, not R^2.

Response:

Apologies for this mistake. The Tables A5 and A6 contain the Pearson coefficients of correlation (R), and not coefficients of determination (R^2). This has now been made consistent between the text and tables, with R reported only.

84. L257: 'larger ice-contact proglacial lakes imply larger calving-front interactions': source?

Response:

Sentence has been clarified and supporting references provided.

85. L259: What and how long is 'a typical glacier response time'?

L260-261: Content of this sentence not fully clear, please reformulate.

Response:

Typical glacier response times cover a large range starting from 10 years, to over 150 years. We have added text on this and provided a reference.

86. L264: 'geomorphic, topographic and climate parameters': more specifically?

Response:

This specificity is given in the tables and figures immediately referred to in the next sentences.

87. L266: what does this correlation tell us?

Response:

The correlation reconfirms, and demonstrates over a much larger scale that pro and supraglacial lakes predominately develop on debris-covered glaciers. Text on this, with supporting reference, has been added.

88. L267: Not the best statistical approach to exclude data points that we don't like.

Response:

The point being made here is that we would expect the Karakorum to behave differently as it is a known anomaly. A sentence explaining this came later, but has now been moved to make this point earlier. We deliberately report statistical results both with/without including the Karakorum, so we don't believe the reviewers criticism is justified.

89. *L267-269: What has the correlation between debris cover and glacier length to do with glacier lakes? Where is a proof for the notion that 'low-gradient glaciers favour supraglacial and proglacial lake formation'?*

Response:

In the previous paragraph we highlight the link between glacier lakes and debris-covered glaciers. Here we take that one step further, and note that debris cover (and therefore lake growth) is most pronounced on long, flat glaciers. This is not a new finding, but is worth noting here as we can use our large-scale data to quantitatively demonstrate these relationships. We have added references to papers that have also noted this association. A statistically significant correlation is also seen between glacier length and lake area (Table A6).

90. *L272: 'Some adjacent regions': which ones exactly? 'comparable', 'large', 'substantial': please be more specific.*

Response:

This statement was referring to data given in Table A6. The sentence has been revised, and specific examples from Table A6 are now given.

91. *L273: 'longer-term climate trends': on which scales? Years, decades, centuries?*

Response:

The statement refers to the trends shown in Figure 6, i.e, multi-decadal. The text has been clarified.

92. *L274: 'rapid warming': warming seems to be present across the entire study region except the Karakoram. 'decreased precipitation': precipitiation is unchanged in the Eastern and Western Himalaya, where lakes have grown most in the study period. The line of argumentation is not clear here.*

Response:

The warming has not been uniform, and the point made in the previous sentences is that in some areas where warming has been most rapid (e.g. Eastern Himalaya), lakes are large despite other factors not being so ideal (e.g., a lower abundance of debris covered glacier tongues). Conversely, in Western Himalaya we see the opposite – less warming and less lakes, despite lots of long debris covered glaciers.

93. *L275-277: Unclear, please be more specific ('further west and north', 'similar'?) and reformulate.*

Response:

A specific example has been added to the text. The reader can refer to the cited figures for further examples.

94. L279-282: Vague conclusion – these arguments could explain anything, in the light of the high regional variability in all the predictors that the authors have shown. There is a large body of literature that has worked on the size distribution and regional growth pattern of glacier lakes (e.g. Song et al., 2016, 2017). However, this literature is largely left untouched here. By the way, until here, no references have been given for the datasets used in the correlations.

Response:

Concluding statement for this section has been rewritten. The point that we try to highlight is the close association between surpa- and proglacial lake formation with large debris covered glaciers. This is not a new finding, but our large dataset enable us to confirm observations previously reported at regional scales. We now refer to the earlier work of Song et al., and in particular the processes of lake enlargement on debris covered tongues that were reported three.

References/sources for all datasets used in the correlations are now provided in Table A5. In addition, the primary glacial datasets are cited directly in the text.

95. L282: Dan and CLAGUE, 2011: wrong citation.

Response:

Citation has been removed.

96. L291: 'datasets'.

Response:

Thanks for this suggestion. We have replaced 'dataset' by 'datasets' in the revised manuscript.

97. L293: compared to which dataset? Reminds me also a bit of comparing apples with oranges, given these studies used different minimum mapping units, sizes of the study regions, buffers around lakes, types of lakes analysed and so on.

Response:

We greatly appreciate the Reviewer's valuable comments very much.

Firstly, we are so sorry for the wrong location of the citation again. Compared with the results from Zhang et al., 2015, Hi-MAG lake number was 7268 higher and area was 644.26 km^2 higher than the estimation for the Tibetan Plateau. We have modified this sentence in the revised manuscript to make it clearly. The detailed modifications are as follows:

Hi-MAG lake number was 7268 higher and area was 644.26 km^2 higher than the estimation for the Tibetan Plateau (Zhang et al., 2015).

Secondly, when we used to produce the glacial lake database and wrote the last version of manuscript, no relevant data were released publicly. Because the dynamic cross-validation procedures are essential to maximize the quality of glacial lake change detections, we have

only made a rough comparison of the statistical results between our Hi-MAG database and other studies. However, at present, there are no uniform definition and mapping standard for the glacial lakes yet, which result in the minimum mapping units, buffer distance around lakes and types of lakes used are different for these studies. Also the extent of the study regions are not completely consistent, just as the Reviewer pointed out, the comparison with other lake dataset is not strict indeed.

To clearly show benefits and challenges remaining in this study, and also for more rigorous cross-validation, in the **Section 6.1** of revised manuscript, we have conducted a deep analysis and comparison between our Hi-MAG dataset and glacial lake inventory from Wang et al., 2020, who mapped glacial lake larger than 0.0054 km^2 in a much larger High Mountain Asia region in two time periods (1990 and 2018), and made their data public recently. The comparative discussions with other studies (e.g. Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a) can be used as a general reference. Please see the revised manuscript for more details.

98. L296: 'a much larger lake area': how much larger?

Response:

We are so sorry for this inappropriate expression. We have deleted this sentence in the revised manuscript.

99. L300: How did the authors convert the gridded data from GSW to polygons that match the Hi-MAG database?

Response:

We gratefully thank the Reviewer for his/her careful reading of the manuscript and for pointing out this issue. GSW and Hi-MAG database have different data formats, which are grid and vector, respectively. To match the formats of the two datasets, we have converted the lake polygons in the Hi-MAG dataset to the gridded data. In the **Section 6.1 Comparison with other lake datasets** of revised manuscript, we have added some text to explain the conversion between two dataset clearly. The detailed modifications are as follows:

==For the sake of a reliable comparative analysis, lake polygons in the Hi-MAG dataset were converted into the grid format, and glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km).==

100.   L300-301: 'more glacial lakes in the Himalaya, Eastern Hindu Kush, and Tien Shan, and fewer in Eastern Pamir and Western Kunlun Shan': please be more specific.

Response:

Thanks for this good suggestion. According to the response to your Comment 97, because the minimum mapping units, buffer distance around lakes and types of lakes used are not completely consistent for the different studies, we have made a **qualitative comparison** of the statistical results between our Hi-MAG database and study from Pekel et al., 2016. In the **Section 6.1** of revised manuscript, more detailed **quantitative comparisons** were implemented between Hi-MAG database and glacial lake inventory produced by Wang et al., 2020. Besides, we have also modified this sentence to make it formulated clearly. The detailed modifications are as follows:

In addition, we qualitatively compared the lake extent between publicly available high-resolution Global Surface Water (GSW) dataset and our Hi-MAG database summed by mountain range in 2015.

Hi-MAG detected more glacial lakes in the Himalaya, Eastern Hindu Kush, and Tien Shan, and fewer in Eastern Pamir and Western Kunlun Shan.

101.   L304: 'significant number': how do the authors measure significance here?

Response:

We are terrible sorry again for the inappropriate use of 'significant', which may confuse the readers. The word 'significant' here is different from the concept of statistical significance, it means that 'numerous' or 'a large amount of' glacial lakes. In the revised manuscript, we have modified this expression to make it easy to understand. The detailed modifications are as follows:

While there are numerous glacial lakes from an open water perspective,

102.   L305: 'only part of lakes are formed by glacier meltwater': how do the authors measure the contribution of glacier meltwater to the total volume of a glacier lake?

Response:

We greatly appreciate the Reviewer's valuable comments very much. We realize this is a very important point we may not explain clearly. In the last version of manuscript, what we originally meant was that although GSW detected more glacial lakes in some sub-regions of HMA, actually part of them are river segments, which is not formed by glacier meltwater. We have completely modified this sentence in the revised manuscript to make it clearly. The detailed modifications are as follows:

While there are numerous glacial lakes from an open water perspective, actually part of them are river segments.

103.   L304: the GSW does not 'define' glacier lakes per se, but maps surface water. Please clarify.

Response:

Thanks for this good suggestion. We have carefully modified this sentence in the revised manuscript to clarify that GSW does map surface water rather than glacial lakes. The detailed modifications are as follows:

The glacial lake area observed in our lake dataset in the Eastern Pamir and Western Kunlun Mountains does not conform to the mapped surface water in the GSW for these sub-regions.

104.   L310: 'errors could exist in either dataset': a surprising notion given that 10 independent experts have mapping glacier lakes for the Hi-MAG database. Also, what causes these 'similar reflectance from the adjacent land surfaces' to be so special in the Tien Shan compared to all other regions?

Response:

We gratefully thank the Reviewer for his/her comments and suggestion.

Tien Shan Mountains, which are located inland, are the only large mountains in the world that surrounded by huge deserts. It has a temperate continental arid climate, in the cold season, mountains, basins and valleys are all covered by deep snow below the altitude of 3000m above sea level, while in summer, it favors the formation of rainy and snowy climate in the region above 3000 m (Wang et al., 2013a;Shen et al., 2009). Being in this frost zone, large amounts of ancient glacial deposits have also been accumulated (Glazirin, 2010). The unique climatic and terrain conditions have developed glacial lakes that are abundant at small sizes. Also because of the characteristics of the source glacier and lake bed, and due to the water depth and sediment influx, these lakes appear turbid over a wide spectral range, this produces heterogeneous reflectance from the lakes and adjacent land surfaces. The regional heterogeneity in image bring great challenges for automated detection of glacial lakes. Errors could exist in datasets produced by automatic classification. For the further refinement of individual glacial lake, in this study, a detailed manual editing was conducted by ten independent experts, serving as an effective supplement to initial glacial lake extraction. So we were not relying exclusively on the results of automatic classification.

However, for the generation of such an annually resolved glacial lake inventory, the optimal images with valid observations over the potential glacial lake area is very limited in each year. Also the influences from image quality, ice and shadow that obscured the lakes, all greatly influence the extraction results of the lake extent and thus inevitably produced the detection errors, especially for the problematic regions like Tien Shan Mountains, where glacial lakes have outlines that remain unclear and are easily mixed with the glacier snow/ice cover. By combining with other types of optical or SAR data during the same periods, further improvement of the extraction results of glacial lakes in Tien Shan will be an urgent issue for us to consider in the next step.

Finally, in the revised manuscript, we have rewritten the related content to give more information about why the errors could exist in either dataset and why Tien Shan is a special region that glacial lakes show similar reflectance with the adjacent land surfaces. The detailed modifications are as follows:

There is little agreement for Tien Shan, where the weather is rainy and snowy in the region above 3000 m, and large amounts of ancient glacial deposits have been accumulated. Here glacial lakes are featured by small size, due to the influence of source glaciers and lake beds, as well as the water depth and sediment inflow, glacial lakes appear heterogeneous reflectance in the image. Errors could exist in datasets produced by automated classification, but we also did a detailed manual editing, so we were not relying exclusively on automatic classification.

105.   L311-312: 'Karakoram regions have fewer glacial lakes in our estimate': because? The GSW has overestimates surface water on debris covered glaciers, may this helps explaining the difference?

Response:

We gratefully thank the Reviewer for his/her valuable comments and suggestions. In the GSW dataset, the transient melt ponds and streams on the debris-covered glaciers are also identified here as permanent water surfaces, result in the overestimation of the surface water in the Karakoram regions. Therefore, it looks like that our Hi-MAG dataset represented fewer glacial lakes in the Karakoram regions. In the revised manuscript, we have added some text to give

more explanations about why Karakoram regions have fewer glacial lakes in our estimate. The detailed modifications are as follows:

==Karakoram regions seem to have fewer glacial lakes in our estimate, owing to the overestimation of surface water on debris covered glaciers in the GSW dataset.==

106. L343-346: 'relatively good performance for areas having simple lake characteristics and environmental backgrounds'; 'diverse climatic conditions, physical properties and surrounding environments': please be more specific.

Response:

Thanks for this good suggestion. To make the related expressions much more specific and clearly, we have rewritten these sentences in the revised manuscript. The detailed modifications are as follows:

==Such methods result in relatively good performance for lake areas that remain clear and having homogeneous reflectance in the image, but do not allow for continental-scale glacial lake mapping that have spectral interference from the other objects such as glaciers, snow, clouds, turbidity and sedimentation characteristics of glacial lake itself, or the atmospheric interference and terrain effects.==

107. L360: Please add the year when these images were shot. In each panel, the authors could also show the lake polygons mapped for this year for each lake type.

Response:

Thanks for this good suggestion. We have added the year of these images in the caption of Fig. A1, and modified Fig. A1 by overlaying lake polygons for each lake type. The detailed modifications are as follows:
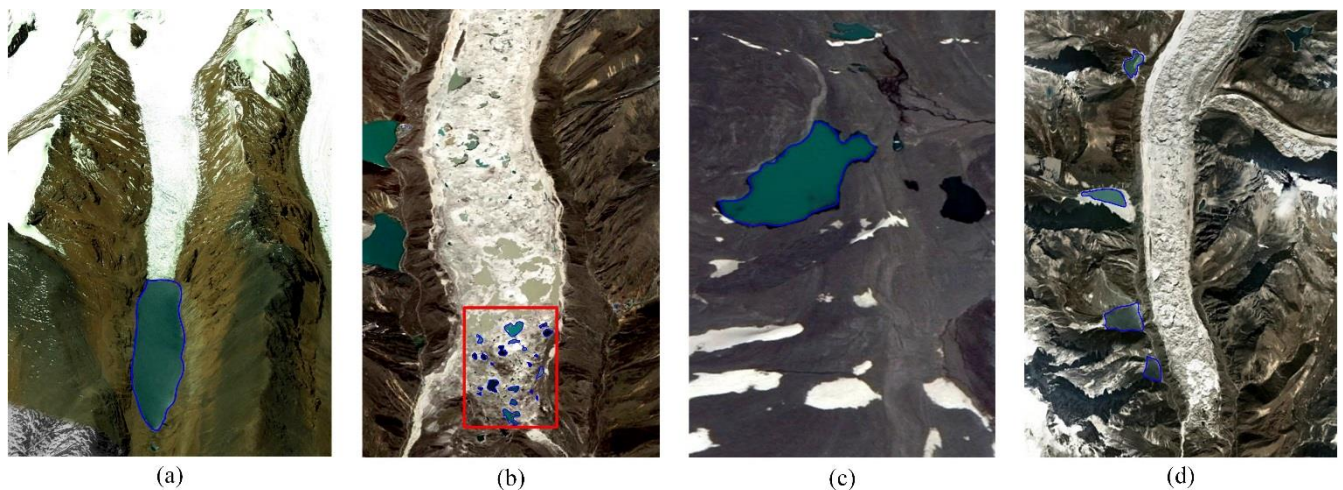

(a) (b) (c) (d)

Fig. A1. Examples of the various types of glacial lake found in the HMA: (a) pro-glacial lakes, which are connected to the parent glacier and usually impounded by a debris dam (usually a moraine or ice-cored moraine); (b) supraglacial lakes (denoted by the red rectangle) which develop on the glacier surface; (c) unconnected glacial lakes; and (d) ice-marginal lakes that distributed on the edge of a glacier. ==Background images were acquired from © Google Earth, and were shot in 2009, 2011, 2012 and 2014, respectively. Glacial lake outlines for each type are shown in blue color.==

108.   L366: Please avoid using rainbow color scales

Response:

Thanks for this good suggestion. We have modified Fig. A2 by using another color scale. The detailed modifications are as follows:
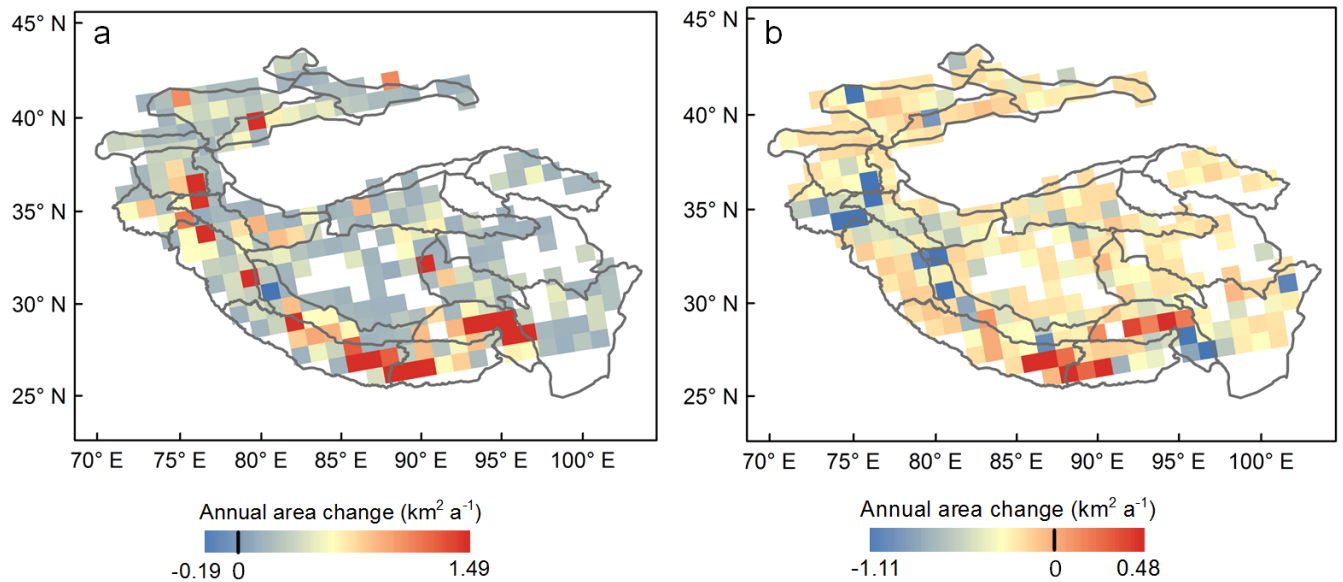

Fig. A2. Annual changes in lake area between 2008 and 2017 on a 1°×1° grid. The (a) upper and (b) lower slopes represent the 90% confidence interval.

109.   L406: Why does this need to be acknowledged given that these researchers are coauthors in this study?

Response:

We are terrible sorry for this mistake. These researchers are co-authors indeed, and do not need to be acknowledged in this section. We have deleted this sentence in the revised manuscript.

Response to Comments by Reviewer 2:

Chen et al. used Google Earth Engine to map glacial lakes in High Mountain Asia (HMA) from 2008 to 2017 with Landsat imagery. Their data is given in annual time steps, which so far is the highest temporal resolution of glacial lake inventory for the HMA. Thus, this kind of dataset if with fine quality could be particularly useful for scientific researches in changes of the cryosphere of the HMA as well as for related assessments and evaluations on the hydrological responds and glacial lake outburst flood risks under a changing climate in the HMA.

We greatly appreciate the suggestions of the Reviewer for his/her accurate summary of the main contributions of our work and for the outstanding recommendations provided. Following the Reviewer's very pertinent recommendations, in the revised manuscript and the following text we have made modifications according to these questions. We thank again the Reviewer for his/her careful handling of our manuscript and for the constructive suggestions provided, which greatly helped us improve the technical quality and presentation of our manuscript.

1. *Noticed that recently there published a similar dataset produced by Wang et al. (2020), which includes two periods (1990 and 2018) of glacial inventory for the HMA and currently is also under review for the ESSD. The later one used a more traditional method and probably involved more extensive manually inspection during their investigations. When comparing these two datasets for the closest period (2017 of Cheng et al. and 2018 of Wang et al. 2020), I found there is a very large discrepancy (Table 1) in their results, although they have claimed that they used similar (not the same) area threshold (0.0081 km^2 of Chen et al. 2020 and 0.0054 km^2 of Wang et al. 2020) and distance threshold (within a 10 km from the nearest glacier terminus) for the lake mapping.*

Table 1: Differences between two datasets of HMA glacial lake inventory

Source	Year	Total Numbers	Area_Sum (km²)	Area_Max (km²)	Area_Mean (km²)	Area_Min (m²)	Notes
Wang et al. 2020	2018	28953	1955.939	6.465	0.067	5400	Altai mountains included
Wang et al. 2020	2018	22219	1672.479	6.465	0.075	8100	Altai mountains excluded and area larger than 8100 m²
Chen et al. 2020	2017	14477	1635.939	26.598	0.113	8100	Altai mountains excluded

In general, dataset of Chen et al. has missed a quite number of glacial lakes when comparing their results with the Wang et al. 2020's. In addition, they excluded the Altay and Sayan mountains, which actually should be included for an inventory study for the HMA. Even if excludes Altay form Wang et al.'s result, there still exist remarkable discrepancy in total numbers (nearly 7800 lakes) and total area (~37 km^2) between each other (Table 1). I agree with the Reviewer #1 that the missing inventory by Chen et al. is far from the range of uncertainty, but it was indeed errors due to the lack of systematic experts' check through the results, which were greatly depended on the methods they used when applying their procedures in Google Earth Engine.

The authors should pay more attention to how to control the quality of their dataset. Although the Google Earth Engine offer the opportunity to calculate lake inventory with a higher temporal resolution, I still strongly recommend they check through their result for

Response:

We greatly appreciate the Reviewer's valuable comments and suggestions. This paper focus on obtaining the comprehensive knowledge of the distribution, area of glacial lakes, and also quantification of variability in their size and type at high resolution in HMA. We develop a HMA Glacial Lake Inventory (Hi-MAG) database to characterize the annual coverage of glacial lakes from 2008 to 2017 at 30 m resolution. This is the first glacial lake inventory across the HMA with annual temporal resolution, it can provide details for different types of glacial lakes and evolution patterns. Although the method of lake mapping was automatic, for quality control every lake polygon was inspected and was manually edited where needed. The related attribution were also added for each lake. This is a huge amount of work for the mapping of nearly 140,000 glacial lakes over ten time periods from 2008 to 2017.

In the last version of manuscript, the mapping work still have some deficiencies and was not fully reflected in the quality of the manuscript indeed. Therefore, we have made considerable efforts to fill these defects, which is mainly manifested in the four aspects including **supplement of glacial lake data, detailed comparison with other lake inventories, quantitative uncertainty analysis of lake area and adding confidence intervals for trends, details about the whole mapping procedure.** We have made the following improvements in our study:

1) We have been working hardly to update data of glacial lakes in the ten year, and add their attribute information. Based on the new lake inventory, we have also carefully modified the tables, figures, the related analysis and made reliable conclusions.

In this study, we produced the glacial lake dataset in High Mountain Asia in annual time steps from 2008 to 2017. The number of glacial lakes in each year is more than 12,000 totally, the ten year records have nearly 140,000 lake polygons. To improve work efficiency, obtain the accurate boundary of the glacial lakes, and meanwhile, minimize the subjective judgment errors of the operatives, we applied a systematic glacial lake detection method that combined two steps from initial glacial lake extraction and subsequently manual refinement of lake mapping results.

The initial glacial lake extraction using GEE can make sure that approximately half of all the lakes are automatically extracted. To make annual data more complete and accurate, manual inspection and refinement of individual glacial lake is necessary to supplement some missing lakes and correct the mapping errors such as mountain shadows and river segments. The glacial lake vector over the entire HMA for the years from 2008 to 2017 has been rechecked and reedited individually through dynamic cross-validation by ten trained experts. Besides, the related attribution (i.e., lake type, elevation, distance to the nearest glacier terminus, area and perimeter) were manually added for each lake. It should be noted that the type of each individual glacial lake was carefully judged according to the different formation mechanisms or growth stages. Thus the whole processing is time consuming and require a considerable amount of work. We are terrible sorry for the large discrepancy compared with dataset from

Wang et al., 2020, which may contribute to the difference in the calculated statistics, and the conclusions drawn in some sub-regions.

Therefore, following the Reviewer's very pertinent recommendations, we have carefully examined the lake data for the ten year by supplementing glacial lakes that have an area of greater than 0.04 km^2. The area of 0.04 km^2 is chosen as the threshold for re-revision of the glacial lake dataset for three reasons: **i)** glacial lakes with small sizes are more likely to be confused with surroundings due to its less effective spectral, textural and spatial information in comparison with those of relatively large glacial lakes, which results in the **large uncertainty of area** (Please see the Section 4, small glacial lakes with the area of less than 0.04km^2 have the mean area uncertainty of about 25.7 %). **ii)** small glacial lakes are highly variable in their locations, shapes and size, and also the optimal images with valid observations over the potential glacial lake area is very limited **in each year**. The image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from image quality, ice and shadow that obscured the lakes, this creates a great deal of uncertainty about the number and extent of small glacial lakes **for the such an annually resolved inventory**. **iii)** given the amount of work required to update the inventory and also revise our manuscript before the deadline.

Moreover, as for the definition of the geographical location of the study area, generally, the term HMA refers to a **broad high-altitude region in South and Central Asia** that covers the whole Tibetan Plateau and adjacent mountain ranges, including the Eastern Hindu Kush, Western Himalaya, Eastern Himalaya, Central Himalaya, Karakoram, Western Pamir, Pamir Alay, Northern/Western Tien Shan, Dzhungarsky Alatau, Western Kunlun Shan, Nyainqentanglha, Gangdise Mountains, Hengduan Shan, Tibetan Interior Mountains, Tanggula Shan, Eastern Tibetan Mountains, Qilian Shan, Eastern Kunlun Shan, Altun Shan, Eastern Tien Shan, Central Tien Shan, Eastern Pamir **(Yoon et al., 2019;Brun et al., 2017a;Zhao et al., 2014)**. It extends from 26°N to 45°N and from 67°E to 105°E, the altitude of the plateau is about 4500 m on average.

Until now, there are still not a uniform standard about the spatial location of HMA region. When we first started working on the glacial lake mapping in the HMA region and required accurate mountain ranges division data, Tobias Bolch kindly provided us the mountain boundary shapefile in High Mountain Asia, which excluded Altay and Sayan mountains, can be downloaded from the website geo.uzh.ch/~tbolch/data/regions_hma_v03.zip. Just as the Reviewer pointed out that, the generalized concept of HMA could contain the Altay and Sayan mountains, and a comprehensive delineation of glacial lakes in these regions over the ten year periods is needed urgently in our future work.

Our team have been working hardly to update data of glacial lakes larger than 0.04 km^2 in the ten year, and add their attribute information. Correspondingly, the related statistics, figures and analysis in the article have also been modified based on the new lake inventory. We tried our best to improve the data and manuscript, and hope that the revised manuscript will meet with approval. As an annual time steps data over the HMA region from 2008 to 2017, the produced data may not be complete and perfect enough, yet our team will continuously update and share more and better glacial lake data in the future.

Finally, to provide more information about our study area, in the **Section 2.1 Study area** of the revised manuscript, we have rewritten and expanded this chapter by adding the location map of the HMA and comprehensive descriptions about the climate and topography. Please see the revised manuscript for more details.

2) We have comprehensively investigated the existing works about glacial lake inventory, and conducted deep comparisons between our Hi-MAG dataset and glacial lake inventory from Wang et al., 2020, and other studies from Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a.

To clearly show benefits and challenges remaining in this study, we firstly comprehensively investigated the existing works about glacial lake inventory.

Many previously published researches have devoted to the glacial lakes mapping with remotely sensed data over the different regions of HMA. Some works mainly focus on the investigation of the development of relatively large glacial lakes. Rounce et al. identified 131 glacial lakes in Nepal in 2015 that are greater than 0.1 km^2 (Rounce et al., 2017). Li et al. compiled an inventory of glacial lakes ($\geq$0.01 km^2) with a spatial resolution of 30 m in the Karakoram mountains (Li et al., 2020). Aggarwal et al. shared a new dataset of glacial and high-altitude lakes that have an area > 0.01 km^2 for Sikkim, Eastern Himalaya from 1972–2015 (Aggarwal et al., 2017). Ukita et al. constructed a glacial lake inventory of Bhutan in the Himalaya from the period 2006-2010 based on high-resolution PRISM and AVNIR-2 data from ALOS. Considering small lakes present less of a GLOF risk. They set 0.01 km^2 as the minimum lake size (Ukita et al., 2011). Ashraf et al. used Landsat-7 ETM+ images for the 2000-2001 period to delineate glacial lakes greater than 0.02 km^2 in the Hindukush-Karakoram-Himalaya (HKH) Region of Pakistan (Ashraf et al., 2012). Because small glacial lakes experience highly variable in their shape, location, and occurrence, and were clearly sensitive to the warming climate and glacier wastage, a growing number of scholars have paied attention to the abundance of small glacial lakes. Salerno et al. provided a complete mapping of glacial lakes (including lake size less than 0.001 km^2) and debris-covered glaciers with 10-m spatial resolution in the Mount Everest region in 2008 (Salerno et al., 2012). Wang et al. utilized Landsat TM/ETM+ images for the years 1990, 2000 and 2010 to map glacial lakes with area more than 0.002 km^2 in Tien Shan Mountains (Wang et al., 2013a). Luo et al. examined glacial lake changes (lake area >0.0036 km^2) for the entire western Nyainqentanglha range for the five periods between 1976 and 2018 using multi-temporal Landsat images (Luo et al., 2020). International Centre for Integrated Mountain Development (ICIMOD) provided comprehensive information about the glacial lakes (greater than or equal to 0.003 km^2) of five major river basins of the Hindu Kush Himalaya (HKH) using Landsat images for the year 2005 (Sudan et al., 2018). Nie et al. mapped the distribution of glacial lakes across the entire Himalaya in the year of 2015 using a total of 348 Landsat images at 30 m resolution. They set the minimum mapping unit to 0.0081 km^2 (Nie et al., 2017). Zhang et al. presented a database of glacial lakes larger than 0.003 km^2 in the Third Pole for the years ~1990, 2000, and 2010 (Zhang et al., 2015). All these researches greatly help to fill the data gap of glacial lakes information in the HMA region. At the global scale, Pekel et al. used millions of Landsat satellite images to record global surface water over the past 32 years at 30-metre resolution (Pekel et al., 2016b), many large and visible glacial lakes were also included. More recently, Shugar et al. mapped glacial lakes with areas >0.05 km^2 around the world using 254,795 satellite images from 1990 to 2018 (Shugar et al., 2020). Wang et al. developed a glacial lake inventory (with size of larger than 0.0054 km^2) across the High Mountain Asia at two time periods (1990 and 2018) using manual mapping on 30 m Landsat images (Wang et al., 2020b). They firstly introduce glacial lake inventory at such a large-scale, the data shared will be served as a baseline for the further studies related to water resource assessment or glacier hazard risk.

For more rigorous cross-validation and assessment of the data quality, in the **Section 6.1** of revised manuscript, we have conducted a deep analysis and comparison between our Hi-MAG

dataset and glacial lake inventory from Wang et al., 2020, who mapped glacial lake larger than 0.0054 km^2 in a much larger High Mountain Asia region in two time periods (1990 and 2018), and made their data public recently. The comparative discussions with other studies (e.g. Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a) were also used as a general reference.

Firstly, we compared our dataset with that of Wang et al., 2020 for the closest period (2017 from our Hi-MAG database and 2018 from Wang et al. 2020) over the spatial extent of our HMA region. It should be noted that in the last version of database, some huge glacial lakes with an area of larger than 6.5km^2 have also been mapped in the Hi-MAG dataset, this is because these lakes all located in the range of 10 km from the nearest glacier terminus according to the definition of glacial lake in HMA. However, in the dataset produced by Wang et al., 2020, these huge lakes were removed. For the sake of a reliable comparative analysis between different studies under the equal conditions, glacial lakes with an area of larger than 6.5km^2 were deleted in our new lake inventory. Besides, **serval hundreds of glacial lakes in the dataset from Wang et al. are located outside of 10km buffer to the glacier terminus, which were also excluded for the comparison**.

The differences in the total number and area between these two dataset are 6206 and 223.97 km^2, respectively. **We also found that 2077 glacial lakes with a total area of 178.77km^2 in our Hi-MAG dataset were not detected by Wang et al.** As for the lakes we have not mapped, 96.1% of the glacial lakes are smaller than 0.04km^2, which means that 3.9 % (323 glacial lakes) of the difference in the total number is composed by the glacial lakes larger than 0.04 km^2. The main reasons for these missed 323 glacial lakes in our dataset are because the interference of some bad observations (cloud or snow), glacial lake dried up or outburst, or located in the middle of the river, which are summarized and explained in the following:

i) The limited or lack of high-quality images in a whole year. Many glacial lakes were always covered by cloud or snow, and no effective observation data are available, as shown in Figure I.
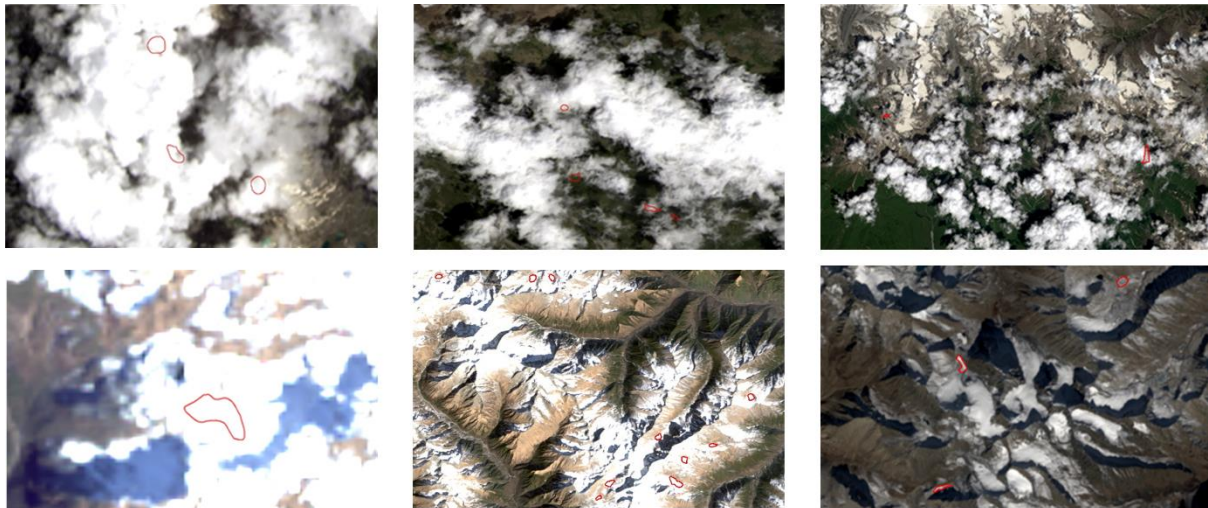

Figure I. Some glacial lakes that are always covered by cloud (first row) or snow (second row) in the true color composites (Bands:7, 4, 3) of the Landsat 8 OLI images for the year of 2017. The red contours refer to the 2018 glacial lake outlines digitized by Wang et al.

ii) Many glacial lakes dried up or outburst, and thus vanished in the image of 2017 (Figure II).
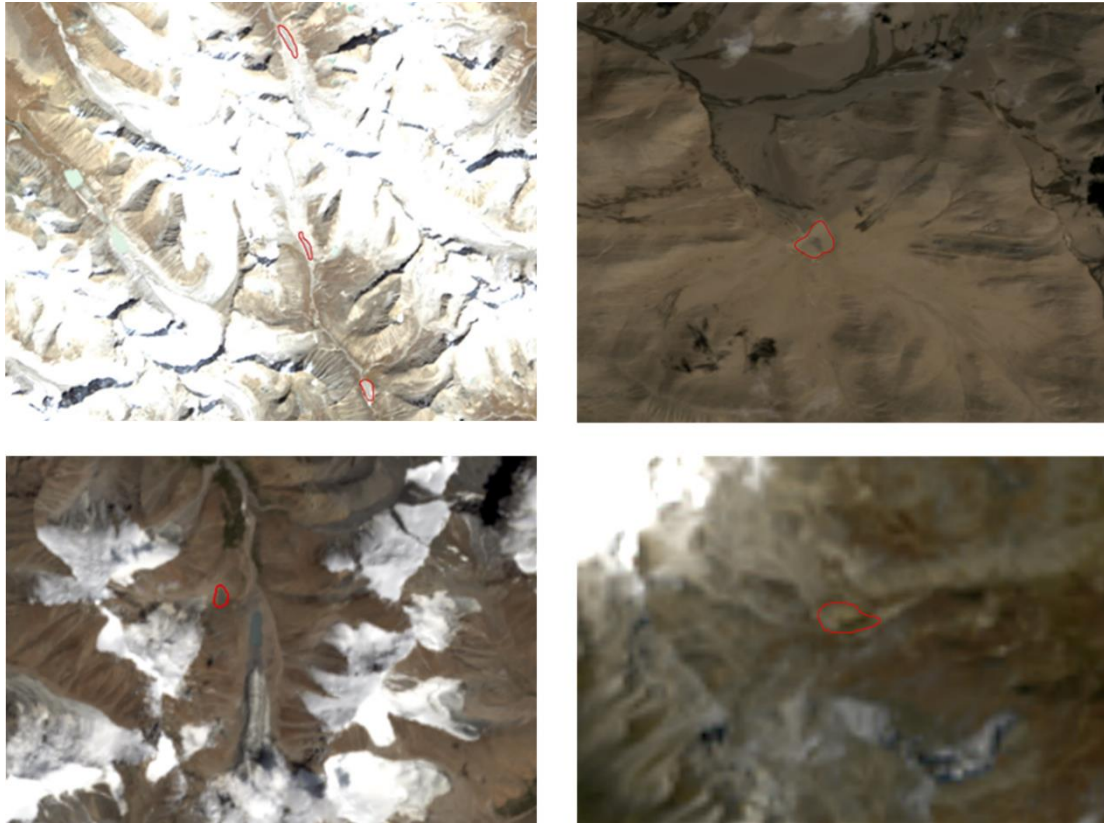
Figure II. Some glacial lakes dried up or outburst in 2017. Background images are the true color composites (Bands:7, 4, 3) of the Landsat 8 OLI images for the year of 2017. The red contours refer to the 2018 glacial lake outlines digitized by Wang et al.

iii) Lakes or ponds located in the middle of the river, which were not judged to be glacial lakes in our Hi-MAG database (Figure III).
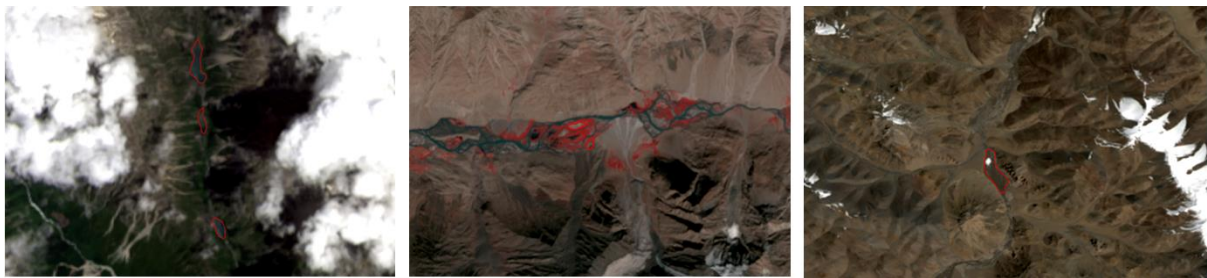

Figure III. Lakes or ponds located in the middle of the river, and were not judged to be glacial lakes in our inventory. Background images are the true color composites (Bands:7, 4, 3) of the Landsat 8 OLI images for the year of 2017. The red contours refer to the 2018 glacial lake outlines digitized by Wang et al.

To test the spatial correlation of glacial lakes distribution in two datasets, we compared the statistics in glacial lake number and area aggregated on a $0.1°×0.1°$ grid for the HMA regions. The results for the total glacial lake area, areas for glacial lakes larger than $0.04km^2$, and number for glacial lakes larger than $0.04km^2$ are depicted in Figure IV. A clear and strong correlation can be observed for all the statistics between our Hi-MAG dataset and glacial lake data by Wang et al. Most of the points being distributed around the 1:1 line, which shows that there is great consistency in the results.
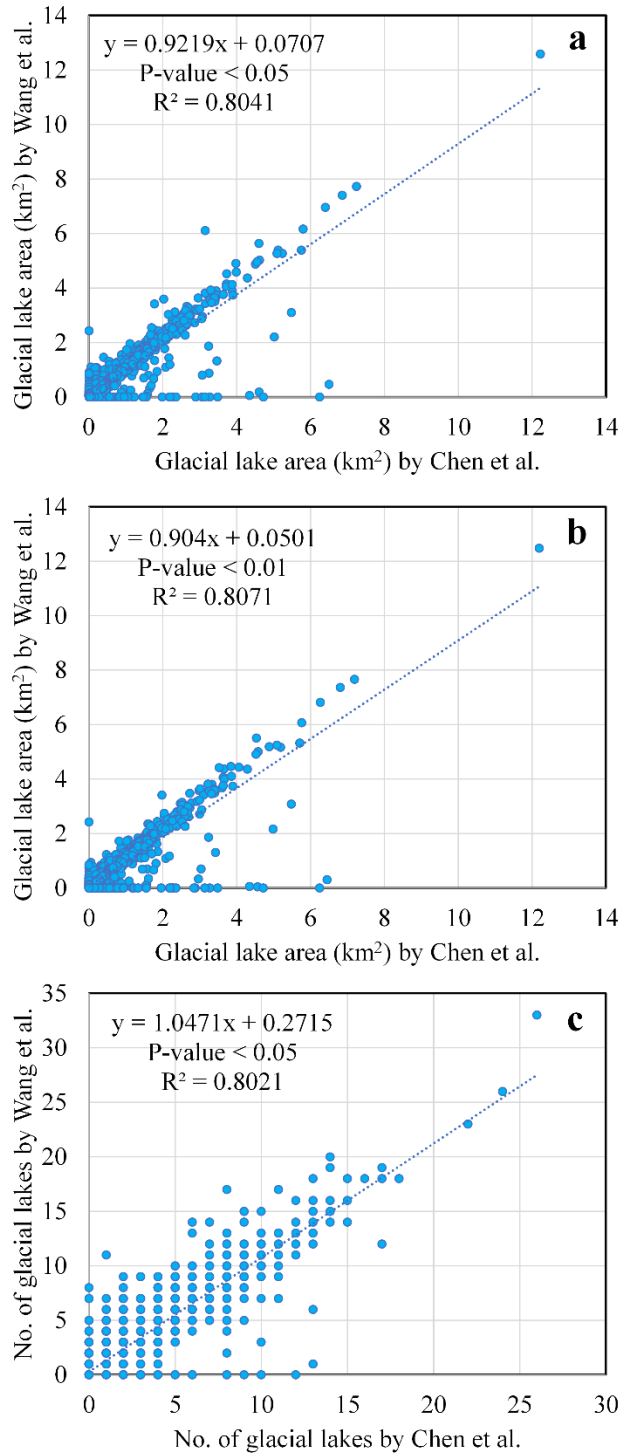
Figure IV. Comparison the results of (a) total glacial lake area, (b) areas for glacial lakes larger than 0.04km^2, and (c) number for glacial lakes larger than 0.04km^2 summed over a 0.1°×0.1° grid between Hi-MAG database and inventory by Wang et al., 2020.

In order to quantitatively and systematically evaluate the accuracy of our product, we implemented a stratified random sampling (Song et al., 2017b;Stehman and SV, 2012), where the glacial lakes were divided into four strata. The sample sizes are the spatial resolution (30 m) of the data, and the strata are designed as: C0W0. both the results are non glacial lakes;

C0W1. non glacial lake for Chen's and glacial lake for Wang's; C1W0. glacial lake for Chen's and non glacial lake for Wang's; C1W1. Both the results are glacial lakes.

A total of 4,000 points were randomly selected, as shown in Figure V. The sample number for C1W1 and C1W0 are 1300 and 700, respectively, which is almost the same ratio between the total areas for the two strata (1450.50 km^2 vs 732.77 km^2). Because of the approximate total area with C1W0, we also randomly selected 700 samples from stratum C0W1. The rest 1300 samples are from C0W0.
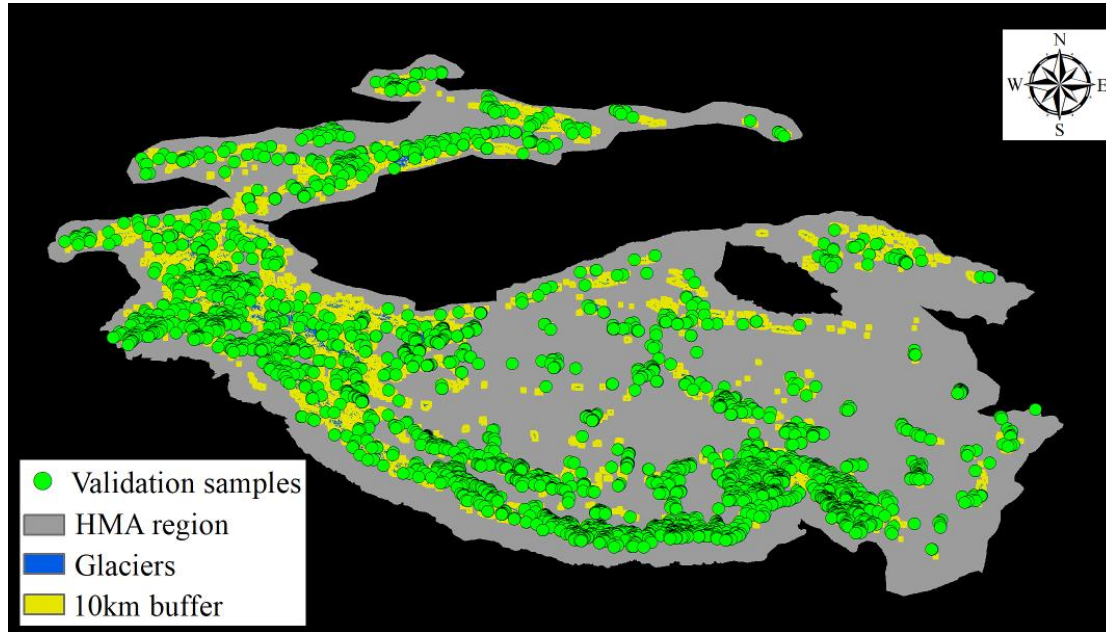

Figure V. Distribution of validation samples selected using stratified random sampling. Blue polygons are glacier outlines taken from the Randolph Glacier Inventory (RGI v5.0). Yellow polygons refer to buffer area within 10 km of glacier terminals.

Every validation sample was visually examined using Landsat imagery and higher-resolution imagery in Google Earth. Sample pixels were interpreted by a regional glacial lake mapping expert, and ambiguous samples were cross-validated by a second observer. If a sample is difficult to interpret, it was marked as ambiguous sample and excluded for the accuracy assessment. The sample number estimates were produced for each of the four strata (Table I), and these strata totals were then summed to obtain the total accuracy.

For the 1300 pixel samples that were considered to be non glacial lake by both datasets, after the pixel by pixel verification, 1215 were indeed non-glacial lakes, while 37 were the missed glacial lakes. In contrary, 1260 out of 1300 pixels belongs to the class of glacial lake, and 25 pixels were misclassified as glacial lake by the two inventories. 307 error pixels were found in the results from Wang et al., constitute about half of the total validation number. For the glacial lakes identified only by our inventory, 678 out of 700 were corrected classified. Our results yielded high overall classification accuracy (88%), user's accuracy (97%), and producer's accuracy (82%) for glacial lake classification using Landsat data, which further confirm the improved quality of the Hi-MAG database.

Table I. Statistical results of stratified random sampling.

Strata	Total pixel number	Total area (km²)	Sample number	Sample No. of non glacial lake	Sample No. of glacial lake	No. of ambiguous sample

C0W0	2,022,448,650	1,820,203.78	1300	1215	37	48
C0W1	925,449	832.90	700	307	362	31
C1W0	814,196	732.77	700	21	678	1
C1W1	1,611,668	1,450.50	1300	25	1260	15

Based on the new version of glacial lake inventory, we also made comparisons of the statistical results between our Hi-MAG database and other studies from Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a. It should be noted that these inventory data have not been released publicly yet, except for GSW dataset produced by Pekel et al. Pekel et al. used millions of Landsat satellite images to map global surface water, many transient melt ponds and streams on the debris-covered glaciers were also identified here as permanent water surfaces, meanwhile, large quantity of small glacial lakes were missing. For the sake of a reliable comparative analysis, glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km). We therefore conducted a rough comparisons with statistical results from Nie et al., 2017;Zhang et al., 2015, and qualitative comparison of the lake extent between GSW dataset and our Hi-MAG database summed by mountain range in 2015.

Hi-MAG lake number was 7268 higher and area was 644.26 km^2 higher than the estimation for the Tibetan Plateau (Zhang et al., 2015). The largest discrepancy is in the Gangdise, Himalaya and Nyainqentanglha Mountains in 2010. Across the Himalaya, we found 476.09 km^2 of glacial lakes, 4.57% more than previous estimates in 2015 (Nie et al., 2017). In addition, we qualitatively compared the lake extent between publicly available high-resolution Global Surface Water (GSW) dataset and our Hi-MAG database summed by mountain range in 2015. GSW data can be accessed at https://global-surface-water.appspot.com/download. For the sake of a reliable comparative analysis, lake polygons in the Hi-MAG dataset were converted into the grid format, and glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km). Hi-MAG detected more glacial lakes in the Himalaya, Eastern Hindu Kush, and Tien Shan, and fewer in Eastern Pamir and Western Kunlun Shan. Fig. A5 illustrates the differences between our Hi-MAG glacial lake results and GSW-derived lake area for the whole HMA region.

The glacial lake area observed in our lake dataset in the Eastern Pamir and Western Kunlun Mountains does not conform to the mapped surface water in the GSW for these sub-regions. While there are numerous glacial lakes from an open water perspective, actually part of them are river segments. Additionally, the Himalaya, Eastern Hindu Kush, and some other Tien Shan host thousands of glacial lakes that are not readily observable in the GSW product. Large discrepancies in mountainous glacial lake estimates preclude a significant consistency between GSW and our Hi-MAG lake data over the HMA region. The region with the highest consistency between GSW and Hi-MAG product is interior Tibet. There is little agreement for Tien Shan, where the weather is rainy and snowy in the region above 3000 m, and large amounts of ancient glacial deposits have been accumulated. Here glacial lakes are featured by small size, due to the influence of source glaciers and lake beds, as well as the water depth and sediment inflow, glacial lakes appear heterogeneous reflectance in the image. Errors could exist in datasets produced by automated classification, but we also did a detailed manual editing, so we were not relying exclusively on automatic classification. Karakoram regions seem to have fewer glacial lakes in our estimate, owing to the overestimation of surface water on debris covered glaciers in the GSW dataset.

Finally, we have added the related content in the **Section 1. Introduction**, to give a comprehensive introduction about the existing glacial lake inventories. In the **Section 6.1** of revised manuscript, we have conducted a deep analysis and comparison between our Hi-MAG

dataset and glacial lake inventory from Wang et al., 2020 and other studies from Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a. Please see the revised manuscript for more details.

3) A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript, the error bars for the lake area, and confidence intervals for the estimated trends were also added throughout the paper.

A thorough and quantitative uncertainty analysis of lake area was added in the Section 4 of revised manuscript. In order to highlight differences in the tails of area uncertainties in each year, we have added a new figure (Fig. 4a) that plotted the distributions of uncertainties from 2008 to 2017. Fig. 4b was also added for the better visualization of the relationship of area uncertainties against areas of all the glacial lakes in HMA (taking the results in 2017 as an example).

Assuming an uncertainty of 1 pixel for the detected glacial lake boundaries, we calculated the systematic errors for the whole HMA region (Fig. 4). For the year between 2008 and 2017, the area uncertainty of each glacial lake generally ranged from 0.30% to 50%, with the mean value falling around the 17%, and standard deviation around 11% (Fig. 4a). The maximum and mean value of area uncertainty for the glacial lakes in 2010 are the lowest, while for the year of 2016, the corresponding statistics are highest, this can be attributed to the different factors. The maximum of area uncertainty of glacial lakes is related with the shape and size of a certain lake (as can be seen from equation (1)), but its mean value is equal to the sum of the area uncertainties of each glacial lake divided by the total number, which depend on the total number of glacial lakes in a year, and also the shape and area of each lake. Besides, a close relationship can be found between the area uncertainties and the sizes of the glacial lakes (Fig. 4b). Most of the large glacial lakes (area $\geq 0.04 \text{km}^2$) have the mean area uncertainty of about 7%. This systematic errors were more significant for the small-sized glacial lakes. We measured glacial lake down to 0.0081 km^2 (nine pixels in Landsat imagery), where systematic errors calculated by equation (1) were ~50%.

Fig. 4. (a) Statistics of area uncertainty (%) of glacial lakes for the years from 2008 to 2017. (b) Relationship between area uncertainties and areas of all the glacial lakes in HMA in 2017.

Besides, given the mapping uncertainty of glacial lakes and thus for the further improvement of reliability of calculated statistics, the error bars for the lake area, and confidence intervals for the estimated trends, and measure of uncertainty for the annual changes were also added throughout the paper.

4) Detailed descriptions about the key procedures for glacial lake mapping were provided in the Section 3.2, such as advantages of automated mapping, and the amount of work required for the manual editing.

For the development of HMA Glacial Lake Inventory (Hi-MAG) database, we applied a systematic glacial lake detection method that combined two steps from initial glacial lake extraction and subsequently manual refinement of lake mapping results. As for the automatic classification using GEE, there are four main procedures including (i) the clipped Landsat images by the extent of the glacier buffers and assembled into a time-series dataset; (ii) identified and masked some poor quality observations such as cloud, cloud shadow, topographic shadow; (iii) calculated MNDWI; and (iv) extracted the potential glacial lake areas by applying adaptive MNDWI threshold. Based on the automated processing, nearly 60% glacial lakes in each year can be correctly classified, of the other lakes that were not properly classified, 30% were missed and 10% were misclassified (mainly mountain shadows and river segments). For such a large-scale area that characterized by various and complex climatic, geological and terrain conditions, this classification method is simple but effective, the results are also reasonable since it provides very low commission errors. As shown in Figure VI, glacial lake outlines extracted using the automated classification method in our study fit the real boundary of the glacial lake very well, while manually delineated glacial lake outlines are largely influenced by people's subjective experience and manual operation, resulting in overestimation for most part of a glacial lake region, and underestimation for a few areas. Moreover, results from manual digitization show poor performance for the delineation of the some glacial lakes with complex curved shapes.

    ☐ Manually delineated lake outlines from Wang et al. 2020
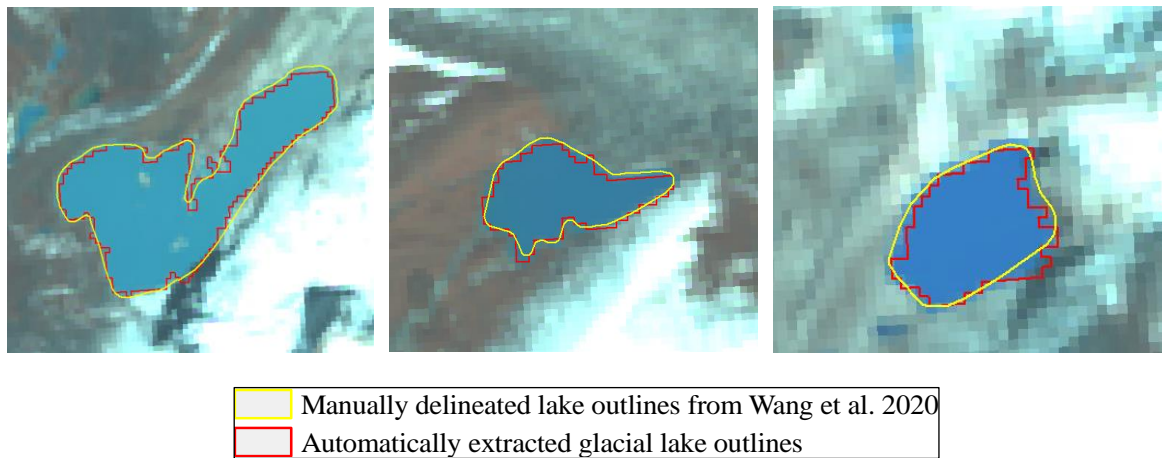   ☐ Automatically extracted glacial lake outlines

Figure VI. Examples of extraction of glacial lake outlines using the automated classification method in our study and manual digitization from Wang et al. 2020.

The estimated 30% errors in the initially classified map were from the missing small glacial lakes. Glacial lakes with small sizes are more likely to be confused with surroundings due to its less effective spectral, textural and spatial information in comparison with those of relatively large glacial lakes. Because the automated method is mainly based on the spectral features for the glacial lake mapping over the large mountainous area, the spectral information provided by the small glacial lakes will be incomprehensive and insufficient for the accurate detection of this kind of glacial lakes under various and complex environmental conditions. Therefore, visual interpretation and manual editing is still an effective way to ensure the high accuracy of lake inventories. In the updated version of dataset, we have carefully examined the lake data for the ten year by supplementing glacial lakes that have an area of greater than 0.04 km^2, and

add their attribute information. We tried our best to improve the quality of glacial lake dataset, and make sure it is greatly optimized.

<u>Finally, in the **Section 3.2 Adaptive glacial lake mapping method** of revised manuscript, we have added some text to give detailed explanations about the whole mapping procedures. The amount of work required for the manual editing were also described clearly. Please see the revised manuscript for more details.</u>

2. *The current attribute table of this inventory is too sample, that it even did not give an ID for each lake. Glacial lakes should be indexed with unique ID that could be used to connect with RGI or GLIMS glacier inventory dataset. In addition, the abbreviations used in the dataset (PGL, UCL and SGL) were totally not mentioned in the main text of the paper, we don't know what they meant.*

Response:

We are terrible sorry for the incomplete attribute information of our inventory. We have carefully modified the attribute table of the new version of inventory by adding the unique ID for each glacial lake (attribute item GL_ID), for example, GL075720E40943N that formed by 'GL'+ 'longitude of centroid' + 'latitude of centroid', and retain three decimal places. Besides, the abbreviations of glacial lake type including PGL, UCL, SGL and IGL were all replaced by their full name, i.e. proglacial lake, unconnected glacial lake, supraglacial lake and ice-marginal lake.

3. *For a dataset paper, it should avoid including any further analysis on the data (for example, the inter‐annual variability of lake area presented in the section 5.2), especially when their results exist large uncertainty or errors.*

Response:

We greatly appreciate the suggestion of the Reviewer. According to the response to your Comment 1, the dataset was updated by supplementing glacial lakes that have an area of greater than 0.04 km^2 for the ten year. As for the estimation of area error, most of the large glacial lakes (area $\geqslant 0.04 km^2$) have the mean relative error of about 7%. We also measured glacial lake down to 0.0081 km^2 (nine pixels in Landsat imagery), where relative errors calculated were ~50%. It can be found that this systematic error was more significant for the small-sized glacial lakes. Given the large uncertainty in the area and number of small glacial lakes, we should avoid including any further analysis on the data, especially for small glacial lakes that have large uncertainty. Therefore, we have carefully modified the whole manuscript to make sure the statistics and related analyses all focus on large-sized and all sizes of glacial lakes over the whole HMA region or different mountain ranges to discuss spatial and temporal variability for the 10-year record, and meantime, to avoid analysing and explaining the trends and number for small-sized glacial lakes. Fig. 5 shows the inter-annual variations in the number of small glacial lakes for different HMA mountain regions, we have removed the Section 5.2 in the revised manuscript.

Aggarwal, S., Rai, S. C., Thakur, P. K., and Emmer, A.: Inventory and recently increasing GLOF susceptibility of glacial lakes in Sikkim, Eastern Himalaya, Geomorphology, 295, 39-54, https://doi.org/10.1016/j.geomorph.2017.06.014, 2017.

Armstrong, W. H., and Anderson, R. S.: Ice-marginal lake hydrology and the seasonal dynamical evolution of Kennicott Glacier, Alaska, Journal of Glaciology, 1-15, 2020.

Ashraf, A., Naz, R., and Roohi, R.: Glacial lake outburst flood hazards in Hindukush, Karakoram and Himalayan Ranges of Pakistan: implications and risk analysis, Geomatics, Natural Hazards and Risk, 3, 113-132, 10.1080/19475705.2011.615344, 2012.

Baumann, F., Jinsheng, H. E., Karsten, S., Peter, K., and Thomas, S.: Pedogenesis, permafrost, and soil moisture as controlling factors for soil nitrogen and carbon contents across the Tibetan Plateau, Global Change Biology, 15, 3001–3017, 2009.

Bolch, T., Kulkarni, A., Kääb, A., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., and Scheel, M.: The state and fate of Himalayan glaciers, Science, 336, 310, 2012.

Brun, F., Berthier, E., Wagnon, P., Kääb, A., and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nature Geoscience, 10, 668-673, 2017.

Capps, D. M., Wiles, G. C., Clague, J. J., and Luckman, B. H.: Tree-ring dating of the nineteenth-century advance of Brady Glacier and the evolution of two ice-marginal lakes, Alaska, The Holocene, 21, 641-649, 2011.

Gardner, A. S., Moholdt, G., Cogley, J. G., Wouters, B., Arendt, A. A., Wahr, J., Berthier, E., Hock, R., Pfeffer, W. T., and Kaser, G.: A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009, Science, 340, 852, 2013.

Glazirin, G. E.: A century of investigations on outbursts of the ice-dammed lake merzbacher (Central Tien Shan), Austrian Journal of Earth Sciences, 103, 171-179, 2010.

Immerzeel, W. W., and Bierkens, M. F. P.: Seasonal prediction of monsoon rainfall in three Asian river basins: the importance of snow cover on the Tibetan Plateau, International Journal of Climatology, 30, 1835–1842, 2010.

Kääb, A., Treichler, D., Nuth, C., and Berthier, E.: Brief Communication: Contending estimates of 2003-2008 glacier mass balance over the Pamir-Karakoram-Himalaya, Cryosphere, 9, 557-564, 2015.

Kumar, S. P.: Estimates of the Regression Coefficient Based on Kendall"s Tau, Publications of the American Statistical Association, 63, 1379-1389, 1968.

Li, D., Shangguan, D., and Naveed, A. M.: Glacial Lake Inventory Derived from Landsat 8 OLI in 2016–2018 in China–Pakistan Economic Corridor, International Journal of Geo-Information, 9, 294, 2020.

Li, J., and Sheng, Y.: An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas, International Journal of Remote Sensing, 33, 5194-5213, 2012.

Luo, W., Zhang, G., Chen, W., and Xu, F.: Response of glacial lakes to glacier and climate changes in the western Nyainqentanglha range, Science of The Total Environment, 735, 139607, https://doi.org/10.1016/j.scitotenv.2020.139607, 2020.

Miles, E. S., Steiner, J., Willis, I., Buri, P., Immerzeel, W. W., Chesnokova, A., and Pellicciotti, F.: Pond Dynamics and Supraglacial-Englacial Connectivity on Debris-Covered Lirung Glacier, Nepal, Frontiers in Earth Science, 5, 10.3389/feart.2017.00069, 2017a.

Miles, E. S., Willis, I. C., Arnold, N. S., Steiner, J., and Pellicciotti, F.: Spatial, seasonal and interannual variability of supraglacial ponds in the Langtang Valley of Nepal, 1999–2013, Journal of Glaciology, 63, 88-105, 10.1017/jog.2016.120, 2017b.

Morfitt, R., Barsi, J., Levy, R., Markham, B., Micijevic, E., Ong, L., Scaramuzza, P., and Vanderwerff, K.: Landsat-8 Operational Land Imager (OLI) Radiometric Performance On-Orbit, Remote Sensing, 7, 2208-2237, 2015.

Neckel, N., Kropáček, J., Bolch, T., and Hochschild, V.: Glacier mass changes on the Tibetan Plateau 2003-2009 derived from ICESat laser altimetry measurements, Environmental Research Letters, 9, 468-475, 2014.

Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y., and Song, C.: A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015, Remote Sensing of Environment, 189, 1-13, 2017.

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418-422, 10.1038/nature20584, 2016a.

Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418-422, 2016b.

Quincey, D. J., Richardson, S. D., Luckman, A., Lucas, R. M., Reynolds, J. M., Hambrey, M. J., and Glasser, N. F.: Early recognition of glacial lake hazards in the Himalaya using remote sensing datasets, Global & Planetary Change, 56, 137-152, 2007.

Rounce, D., Watson, C. S., and Mckinney, D.: Identification of Hazard and Risk for Glacial Lakes in the Nepal Himalaya Using Satellite Imagery from 2000–2015, Remote Sensing, 9, 654, 2017.

Salerno, F., Thakuri, S., D'Agata, C., Smiraglia, C., Manfredi, E. C., Viviano, G., and Tartari, G.: Glacial lake distribution in the Mount Everest region: Uncertainty of measurement and conditions of formation, Global & Planetary Change, 92-93, 30-39, 2012.

Schiemann, R., Lüthi, D., and Schär, C.: Seasonality and interannual variability of the westerly jet in the Tibetan Plateau region, Journal of Climate, 22, 2940-2957, 2009.

Shen, Y., Wang, G., and Ding, Y.: Changes in Merzbacher Lake of Inylchek Glacier and glacial flash floods in Aksu River Basin, Tianshan during the period of 1903-2009, Journal of Glaciology & Geocryology, 31, 993-1002, 2009.

Shugar, D. H., Burr, A., Haritashya, U. K., Kargel, J. S., Watson, C. S., Kennedy, M. C., Bevington, A. R., Betts, R. A., Harrison, S., and Strattman, K.: Rapid worldwide growth of glacial lakes since 1990, Nature Climate Change, 10.1038/s41558-020-0855-4, 2020.

Song, C., Sheng, Y., Ke, L., Nie, Y., and Wang, J.: Glacial lake evolution in the southeastern Tibetan Plateau and the cause of rapid expansion of proglacial lakes linked to glacial-hydrogeomorphic processes, Journal of Hydrology, 540, 504-514, 2016.

Song, X. P., Potapov, P. V., Krylov, A., King, L. A., Bella, C. M. D., Hudson, A., Khan, A., Adusei, B., Stehman, S. V., and Hansen, M. C.: National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey, Remote Sensing of Environment, 190, 383-395, 2017.

Song, X. P., C, H. M., V, S. S., V, P. P., Alexandra, T., F, V. E., and R, T. J.: Global land change from 1982 to 2016, Nature, 560, 639-643, 2018.

Stehman, and SV: Impact of sample size allocation when using stratified random sampling to estimate accuracy and area of land-cover change, Remote Sens Lett, 2012.

Sudan, B. M., Pradeep, K. M., Wu, L., Gao, X., Finu, S., Rajendra, B. S., Narendra, R. K., Samjwal, R. B., Sharad, J., Sonika, S., and Prasant, B.: The status of glacial lakes in the Hindu Kush Himalaya-ICIMOD Research Report 2018/1, ICIMOD Research Report 2018/1, 2018.

Ukita, J., Narama, C., Tadono, T., Yamanokuchi, T., Tomiyama, N., Kawamoto, S., Abe, C., Uda, T., Yabuki, H., and Fujita, K.: Glacial lake inventory of Bhutan using ALOS data: methods and preliminary results, Annals of Glaciology, 52, 65-71, 2011.

Wang, X., Ding, Y., Liu, S., Jiang, L., Wu, K., Jiang, Z., and Guo, W.: Changes of glacial lakes and implications in Tian Shan, central Asia, based on remote sensing data from 1990 to 2010, Environmental Research Letters, 8, 575-591, 2013a.

Wang, X., Ding, Y. J., Liu, S. Y., Jiang, L. H., Wu, K. P., Jiang, Z. L., and Guo, W. Q.: Changes of glacial lakes and implications in Tian Shan, central Asia, based on remote sensing data from 1990 to 2010, Environmental Research Letters, 8, 575-591, 2013b.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Liu, S., Zhang, Y., Jiang, Z., and Tang, Z.: Glacial lake inventory of High Mountain Asia (1990-2018) derived from Landsat images, Earth System Science Data, 1-23, doi: https://doi.org/10.5194/essd-2019-212, 2020.

Yao, T., Thompson, L., Yang, W., Yu, W., Yang, G., Guo, X., Yang, X., Duan, K., Zhao, H., and Xu, B.: Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings, Nature Climate Change, 2, 663-667, 2012.

Yoon, Y., Kumar, S. V., Forman, B. A., Zaitchik, B., and F., K., Yonghwan, Qian, Yun, Rupper, Summer, Maggioni, Viviana, Houser, Paul, Kirschbaum, Dalia, Richey, Alexandra, Arendt, Anthony, Mocko, David, Jacob, Jossy, Bhanja, Soumendra, and Mukherjee, Abhijit.: Evaluating the Uncertainty of Terrestrial Water Budget Components Over High Mountain Asia, Frontiers in Earth Science, 7, 1-21, 2019.

Zhang, G., Yao, T., Xie, H., Wang, W., and Yang, W.: An inventory of glacial lakes in the Third Pole region and their changes in response to global warming, Global & Planetary Change, 131, 148-157, 2015.

Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C., Yang, K., Piao, S., Sheng, Y., Yi, S., and Li, J.: Regional differences of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes, Remote Sensing of Environment, 221, 386-404, 2019.

Zhao, L., Ding, R., and Moore, J. C.: Glacier volume and area change by 2050 in high mountain Asia, Global and Planetary Change, 122, 197-207, 2014.

Annual 30-meter Dataset for Glacial Lakes in High Mountain Asia from 2008 to 2017

Fang Chen[1,2,3], Meimei Zhang[1], Huadong Guo[1,2,3], Simon Allen[4,5], Jeffrey S. Kargel[6], Umesh K. Haritashya[7], C. Scott Watson[8]

[1]Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Beijing 100094, China.

[2]State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Beijing 100094, China.

[3]Hainan Key Laboratory of Earth Observation, Aerospace Information Research Institute, Chinese Academy of Sciences, Sanya 572029, China.

[4]Department of Geography, University of Zurich, Zurich, 8057, Switzerland.

[5]Institute for Environmental Sciences, University of Geneva, Geneva 1205, Switzerland.

[6]The Planetary Science Institute, Tucson, Arizona, 85719, USA.

[7]Department of Geology, University of Dayton, Dayton, Ohio, 45469, USA.

[8]Department of Hydrology & Atmospheric Sciences, University of Arizona, Tucson, Arizona, 85721, USA.

Correspondence to: Meimei Zhang (zhangmm@radi.ac.cn)

Abstract. Atmospheric warming is intensifying glacier melting and lake development in High Mountain Asia (HMA), which could increase glacial lake outburst flood hazards and impact water resource and hydroelectric power management. There is a pressing need for obtaining the comprehensive knowledge of the distribution, area of glacial lakes, and also quantification of variability in their size and type at high resolution in HMA. Here, we developed a HMA Glacial Lake Inventory (Hi-MAG) database to characterize the annual coverage of glacial lakes from 2008 to 2017 at 30 m resolution using Landsat satellite imagery. It can be observed that glacial lakes exhibited total area increases of 90.14 km^2 between 2008-2017, a +6.90% change relative to 2008 (1305.59 $\pm$ 213.99 km^2). Annual increase in lake number and area are 306 glacial lakes and 12 km^2, respectively, and maximum increased lake number occurred in 5400 m elevation, which increased by 249. Proglacial lake dominated areas such as the Nyainqentanglha and Central Himalaya, where around more than half of the glacial lake area (summed over $1° \times 1°$ grid) consisted of proglacial lakes showed obvious lake area expansion, while in the regions of Eastern Tibetan Mountains and Hengduan Shan, the unconnected glacial lakes occupied about over half of the total lake area in each grid, exhibited stability or slight reduction. Our results demonstrate proglacial lakes are a main contributor to recent lake evolution in HMA, accounting for 62.87% (56.67 km^2) of the total area increase. Proglacial lakes in the Himalaya ranges alone accounted for 36.27% (32.70 km^2) of the total area increase. Regional geographic variability of debris cover, together with trends in warming and precipitation over the past few decades, largely explain the current distribution of supra- and proglacial lake area across HMA. The Hi-MAG database is available at: https://doi.org/10.5281/zenodo.4059181 (Chen et al.,

2020), it can be used for studies on the complex interactions between glacier, climate and glacial lake, glacial lake outburst floods, potential downstream risks and water resources.

1 Introduction

High Mountain Asia (HMA) consisting of the whole Tibetan Plateau and adjacent mountain ranges such as Himalaya, Karakoram, and Pamirs, covers the largest area of mountainous glaciers globally. Atmospheric warming has resulted in widespread glacier retreat and downwasting in many mountain ranges of the HMA (Brun et al., 2017b;Bolch et al., 2012b), which favor the formation and development of a large amount of glacial lakes, yet glacial lakes have been incompletely documented at small time intervals. Glacial lake development varies according to climatic, cryospheric, and lake-specific conditions, such as basin geometry that is either connected to glaciers or unconnected, and the length of the lake/glacier contact (Zhao et al., 2018).

Many previously published researches have devoted to the glacial lakes mapping with remotely sensed data over the different regions of HMA. Some works mainly focus on the investigation of the development of relatively large glacial lakes. Rounce et al. identified 131 glacial lakes in Nepal in 2015 that are greater than 0.1 km^2 (Rounce et al., 2017). Li et al. compiled an inventory of glacial lakes ($\geq 0.01 \text{ km}^2$) with a spatial resolution of 30 m in the Karakoram mountains (Li et al., 2020). Aggarwal et al. shared a new dataset of glacial and high-altitude lakes that have an area $> 0.01 \text{ km}^2$ for Sikkim, Eastern Himalaya from 1972–2015 (Aggarwal et al., 2017). Ukita et al. constructed a glacial lake inventory of Bhutan in the Himalaya from the period 2006-2010 based on high-resolution PRISM and AVNIR-2 data from ALOS. Considering small lakes present less of a GLOF risk. They set 0.01 km^2 as the minimum lake size (Ukita et al., 2011). Ashraf et al. used Landsat-7 ETM+ images for the 2000-2001 period to delineate glacial lakes greater than 0.02 km^2 in the Hindukush-Karakoram-Himalaya (HKH) Region of Pakistan (Ashraf et al., 2012). Because small glacial lakes experience highly variable in their shape, location, and occurrence, and were clearly sensitive to the warming climate and glacier wastage, a growing number of scholars have paied attention to the abundance of small glacial lakes. Salerno et al. provided a complete mapping of glacial lakes (including lake size less than 0.001 km^2) and debris-covered glaciers with 10-m spatial resolution in the Mount Everest region in 2008 (Salerno et al., 2012). Wang et al. utilized Landsat TM/ETM+ images for the years 1990, 2000 and 2010 to map glacial lakes with area more than 0.002 km^2 in Tien Shan Mountains (Wang et al., 2013a). Luo et al. examined glacial lake changes (lake area $>0.0036 \text{ km}^2$) for the entire western Nyainqentanglha range for the five periods between 1976 and 2018 using multi-temporal Landsat images (Luo et al., 2020). International Centre for Integrated Mountain Development (ICIMOD) provided comprehensive information about the glacial lakes (greater than or equal to 0.003 km^2) of five major river basins of the Hindu Kush Himalaya (HKH) using Landsat images for the year 2005 (Sudan et al., 2018). Nie et al. mapped the distribution of glacial lakes across the entire Himalaya in the year of 2015 using a total of 348 Landsat images at 30 m resolution. They set the minimum mapping unit to 0.0081 km^2 (Nie et al., 2017). Zhang et al. presented a database of glacial lakes larger than 0.003 km^2 in the Third Pole for the years ~1990, 2000, and 2010 (Zhang et al., 2015). All these researches greatly help to fill the data gap of glacial lakes information in the HMA region. At the global scale, Pekel et al. used millions of Landsat satellite images to record global surface water over the past 32 years at 30-metre resolution (Pekel et al., 2016b), many large and visible glacial lakes were also included. More recently, Shugar et al. mapped glacial lakes with areas $>0.05 \text{ km}^2$ around the world using 254,795 satellite images from 1990 to 2018 (Shugar et al., 2020). Wang et al. developed a glacial lake inventory

(with size of larger than 0.0054 km^2) across the High Mountain Asia at two time periods (1990 and 2018) using manual mapping on 30 m Landsat images (Wang et al., 2020b). They firstly introduce glacial lake inventory at such a large-scale, the data shared will be served as a baseline for the further studies related to water resource assessment or glacier hazard risk.

In summary, a homogeneous, annually resolved inventory and analysis of the spatial and temporal extent of different types of glacial lakes over the entire HMA has still been lacking. We developed a HMA Glacial Lake Inventory (Hi-MAG) database to characterize the annual coverage of glacial lakes from 2008 to 2017 at 30 m resolution. 40,481 Landsat scenes were processed using Google Earth Engine (GEE) cloud computing to delineate glacial lakes (located within 10 km from the nearest glacier terminus) larger than nine (e.g., 3 x 3) pixels (0.0081 km^2) (Nie et al., 2017).

Lakes were manually classified into four categories according to their position relative to the parent glacier or their formation mechanisms (Fig. A1): i) proglacial lakes, usually connected to the glacier tongue and dammed by glacier ice, unconsolidated or ice-cemented moraines (mixture of ice, snow, rock, debris and clay, etc.), proglacial lakes are located next to the glacier terminus and receive melt water directly from their mother glaciers; ii) supraglacial lakes - this is where ponds form in depressions on low-sloping parts of the surface of a melting glacier and are dammed by ice or the end-moraine or stagnating glacier snout; iii) unconnected glacial lakes, which are glacial lakes not directly connected to their parent glaciers at the present time but which, to some extent, may be fed by at least one of the glaciers located in the basin and may have been (but not necessarily are) recently detached from ice contact due to glacial recession. Although not directly connected with the parent glaciers, these glacial lakes are also the outcome of glacier melting in response to atmospheric warming, they can supply fresh water to major river systems of the HMA region, and their changes have significant scientific and socio-economic implications (Nie et al., 2017;Song et al., 2016); and (iv) ice-marginal lakes, these lakes are generally distributed on one side of the glacier tongue, which means that the lake is dammed by the glacier ice on this side. While on the other side, it is bounded by a lateral moraine. With the increase of atmospheric warming and accelerated melting of glacier, some glacier tributary gradually detaches from a main trunk glacier. The detached location, where glacier melting has been particularly intense, in some case is also likely to form ice-marginal lakes. We note that such ice-marginal lakes are very common in some parts of the world (e.g., Alaska) (Armstrong and Anderson, 2020;Capps et al., 2011) but are not common in HMA. Besides, purely glacier-dammed lakes are formed by the advance of glaciers and dammed by almost pure glacier ice. Although the dam composition and structure is slightly different between the proglacial lakes and glacier-dammed lakes, because they are all located in the front of the glacier tongue and driven by the mother glacier, in the process of appending attributed information to each glacial lake, glacier-dammed lakes were merged into class of proglacial lakes.

Every lake was cross-checked manually for its boundary and attribution. We defined an uncertainty of 1 pixel for the detected glacial lake boundaries, and calculated the lake area error for the whole HMA region. We also assessed the inventory for climatic and geomorphological influences on lake distribution across HMA.
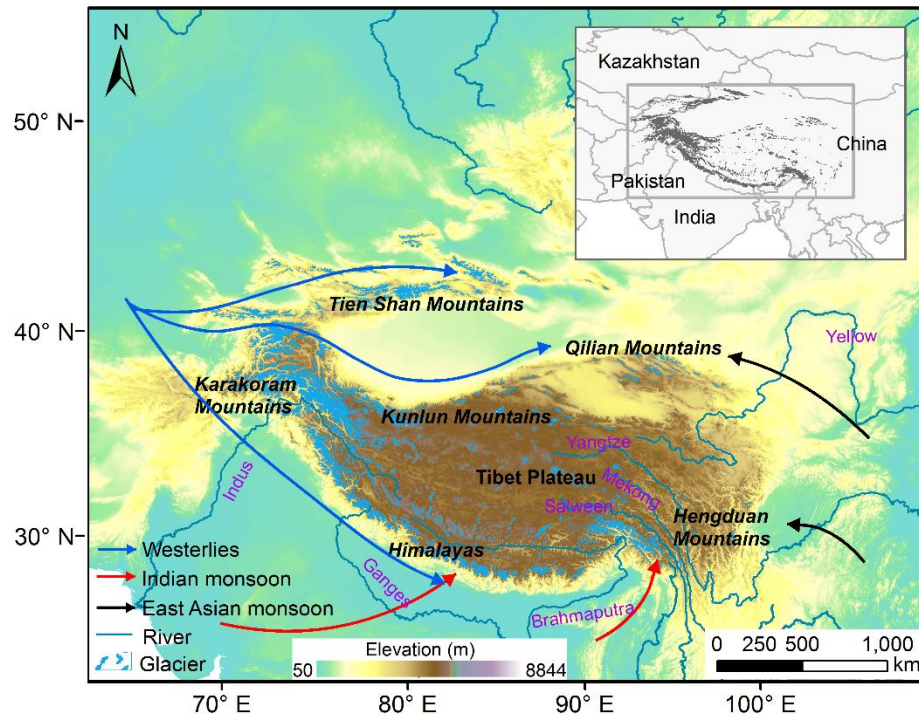
2 Study area and data

2.1 Study area

The term HMA refers to a broad high-altitude region in South and Central Asia that covers the whole Tibetan Plateau and adjacent mountain ranges, including the Eastern Hindu Kush, Western Himalaya, Eastern Himalaya, Central Himalaya, Karakoram, Western Pamir, Pamir Alay, Northern/Western Tien Shan, Dzhungarsky Alatau, Western Kunlun Shan,

105 Nyainqentanglha, Gangdise Mountains, Hengduan Shan, Tibetan Interior Mountains, Tanggula Shan, Eastern Tibetan Mountains, Qilian Shan, Eastern Kunlun Shan, Altun Shan, Eastern Tien Shan, Central Tien Shan, Eastern Pamir (Fig. 1 and Fig. 6a). It extends from 26°N to 45°N and from 67°E to 105°E, the altitude of the plateau is about 4500 m on average (Baumann et al., 2009). It's made up of alternating mountains, valleys and rivers, the terrain is fragmented, showing a decreasing terrain from Northwest to Southeast. HMA has a series of East-West mountains that occupy most of the area.

110 Among these, Tanggula Shan lies in the central part of the HMA, with an altitude of over 6000 m. The height of fifteen highest mountains in the Himalayas are more than 8000 m, while the peaks of the mountains in the northern plateau are more than 6500 m. The North-South mountains are mainly distributed in the southeast of the plateau and near the Hengduan Mountain area. These two groups of mountains constitute the geomorphic framework and control the basic pattern of the plateau landform. Continuous and discontinuous permafrost have developed on the higher land and north facing slopes.

115 HMA is the source of several of Asia's major rivers, including the Yellow, Yangtze, Indus, Ganges, Brahmaputra, Irrawaddy, Salween, and Mekong. They play a crucial role in downstream hydrology and water availability in Asia (Immerzeel and Bierkens, 2010). Most glaciers in the Tibetan Plateau are retreating, except for the Western Kunlun (Neckel et al., 2014;Kääb et al., 2015) and the Karakoram, where a slight mass gain is occurring (Bolch et al., 2012a;Gardner et al., 2013). Moreover, glaciers in different mountain ranges show contrasting patterns. Local factors (e.g., exposure, topography,

120 and debris coverage) may partly account for these differences but the spatial and temporal heterogeneity of both the climate and degree of climate change may be the main reason. Glacial lakes are formed and developed temporally with the retreat or thinning of glaciers and are directly or indirectly fed by glacier meltwater, they are located within 10 km from the nearest glacier terminus (Zhang et al., 2015;Wang et al., 2013b).

The HMA climate is under the combined and competitive influences of the East Asian and South Asian monsoons and of

125 the westerlies (Schiemann et al., 2009). This unique geographical position produces an azonal plateau climate characterized by strong solar radiation, low air temperatures, large daily temperature variations and small differences between annual mean temperatures (Yao et al., 2012). The annual mean temperature is 1.6 °C, with the lowest temperature of $-1 - -7$ °C occurring in January and the highest temperature of $7 - 15$°C occurring in July. The cumulative precipitation is about 413.6 mm a year.

Fig. 1. The location of the High Mountain region of Asia (HMA). Glacier outlines from the Randolph Glacier Inventory (RGI v5.0), the Second Chinese Glacier Inventory (CGI2) and the GAMDAM inventory are drawn in sky blue.

2.2 Dataset

40,481 satellite images including Landsat 5 TM imagery during 2008 to 2011, Landsat 7 ETM+ imagery in 2012, and Landsat 8 OLI during 2013 to 2017, were available in GEE and were used to produce the annual glacial lake maps over the entire HMA (Fig. 2). Here, when Landsat 5 or 8 data were available, Landsat 7 ETM+ imagery with SLC-off gaps were generally excluded due to their artefacts induced by the slatted appearance of the original images, but were exclusively used for the glacial lake mapping in 2012 since no other Landsat data were acquired that year. For the years before 2008, all the available Landsat 5 TM data in each year (e.g., 2004, 2005, 2006, and 2007) do not fully cover the HMA region.

The SLC-off condition of Landsat ETM+ introduces artefacts because the slatted appearance of the original images is occasionally carried into the glacial lake map in 2012. Techniques to fill the SLC-off gaps exist, but these create artificial values that will result in false detections of water (Chen et al., 2011). Considering the strong spatial and temporal variability of glacial lakes like supraglacial lakes, techniques which merge data from one or more SLC-off fill scenes for generation of a gap-free image require careful use, even using the thousands of Landsat ETM+ images. It is noted water mapping using multi-temporal time series images at large scales usually avoided the use of such techniques (Mueller et al., 2016). Therefore, Landsat 7 ETM+ data with intensive slatted appearance is not really a good and suitable data for the classification of numerous of glacial lakes. In this study, because the only useable data source for the year of 2012 is from Landsat 7 ETM+, to ensure continuity of annual data from 2008 to 2017, we have tried our best to manually extract the glacial lakes from the 2012 ETM+ images as accurately as possible.

73
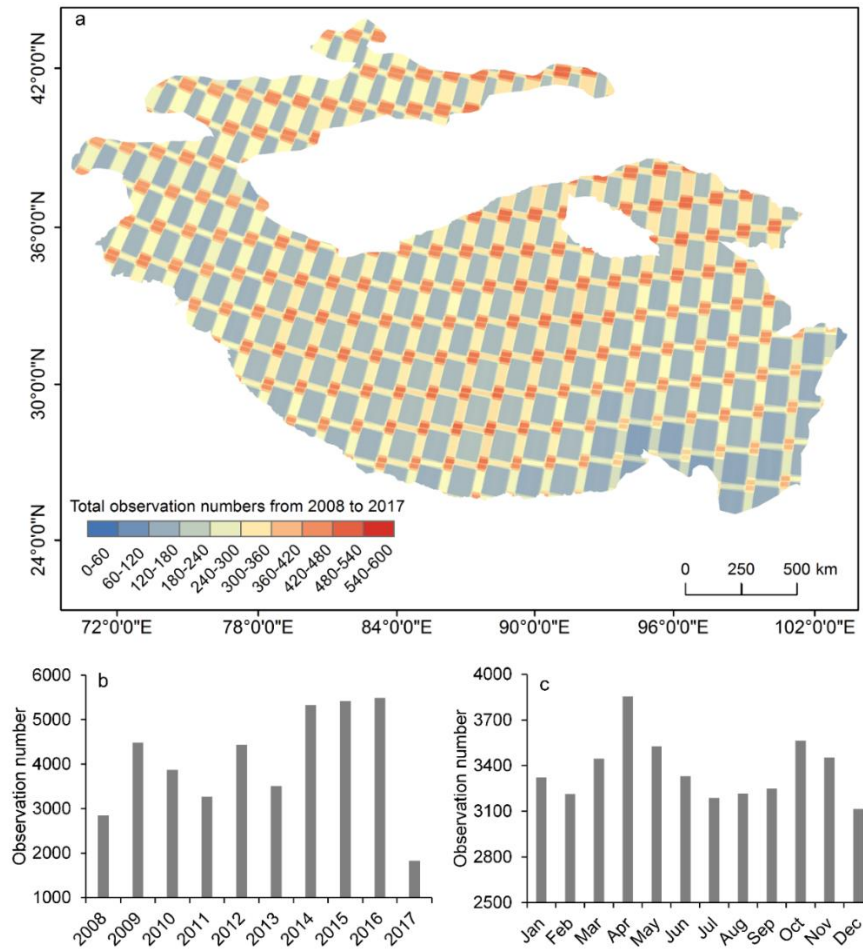
Fig. 2. (a) The distribution of total observation numbers from all GEE Landsat scenes by (b) year and (c) month.

3 Methods

3.1 Satellite imagery selection strategy

To reduce the influence of seasonal lake fluctuations for the mapping, one effective solution is to map glacial lakes and measure their long-term changes during stable seasons when lake extents are minimally affected by meteorological conditions and glacier runoff. Here the selected time series of Landsat data were generally from July to November based on the analyses of mapping time of glacial lakes in the different regions. During this period of each year, the Landsat imagery featured less perennial snow coverage. Change in area of glacial lake is minor and has a large area following the glacier runoff and precipitation (Zhang et al., 2015;Nie et al., 2017;Chen et al., 2017). The lakes may also reached their maximum extent around the end of the glacier ablation season (June to August) (Gardelle et al., 2013;Liu et al., 2014)- except in the central and eastern Himalaya, where peak ablation extends into post-monsoon September and October. In monsoon-affected areas such as Nepal and Bhutan, monsoon cloud cover in July to mid-September means that most of those areas are covered by clear-sky images only from late September to November. Southeast Tibet regions are problematic not only because the observation season is short but abundant cloud cover, which is formed by the warm humid airflow raised by topography (Haritashya et al., 2018;Qiao et al., 2016;Zhang et al., 2020). As the most highly variable glacial lakes in the study area, supraglacial lakes

74

change preferentially in the year, showing an increase in area during the pre-monsoon, and rising to the peak area in the early monsoon (June to July) (Miles et al., 2017a;Miles et al., 2017b). Although the selected image seasons are slightly different due to the meteorological conditions in different regions, they all comply with the same criterion that lake area were in clear-sky images and has small snow coverage. This will ensure the initial reliability of the mapping glacial lakes through GEE cloud computing platform. If no valid observations can be obtained, then optimal mapping time needs to be broaden during the whole year.

To further increase data availability, and also as the basis for data selection in the periods beyond the optimum mapping time, we set two criteria for the selection of imagery with valid observations over the potential glacial lake area by using the cloud score functions in GEE, including (i) cloud cover is less than 20% in the 10 km buffer around each glacier outlines of a Landsat scene, or (ii) less than 20% cloud cover for the entire scene. The cloud score functions in GEE may face a big challenge to detect clouds in mountain headwaters with high snow and ice cover, large amounts of snow and ice are likely to be identified as clouds. However, in this study, it is a much stricter criteria to filter out more images with lots of cloud or cloud-lookalike objects (snow/ice) and finally choose images with good observations.

3.2 Extraction of glacial lake outlines

For the development of HMA Glacial Lake Inventory (Hi-MAG) database, we applied a systematic glacial lake detection method that combined two steps from initial glacial lake extraction and subsequently manual refinement of lake mapping results. The main procedures for glacial lake mapping using Landsat data are (Fig. 3): (i) the Landsat top of atmosphere data were clipped according to the extent of the glacier buffers and assembled into a time-series dataset; (ii) poor quality observations were identified - these included areas affected by cloud, cloud shadow, topographic shadow and SLC-off gaps. Here we used the Fmask routine (Zhu and Woodcock, 2012) to detect the clouds and cloud shadows in an imagery. Fmask has the advantage of being able to process a large number of images in a computationally efficient way. Topographic shadows are located in the areas where the sunlight is blocked, generally on the dark side of the high mountains, their surface gradients are great and the terrain reliefs are small. Therefore, topographic shadows were masked using the slopes (larger than 10°) and shaded relief values (less than 0.25) calculated from SRTM data (Li and Sheng, 2012;Quincey et al., 2007). This will remove considerable mountain shadows that have the similar spectral reflectance with water bodies. However, SRTM DEM was generated in 2000, which is different from the acquisition time of Landsat images used for the glacial lake mapping in this study, the derived slopes and shaded relief cannot fully represent the conditions on the date a given Landsat scene is acquired. As a consequence, some lakes that have grown at steep glacier tongue may be masked, and some mountain shadows that interfere with the mapping results of glacial lakes from GEE still remain, leading to the fact that glacial lakes in the steep areas are omitted, and residual shadows are misclassified as glacial lakes. As for the SLC-off gaps in the ETM+ images, lakes out of the gaps were accurately classified, but if glacial lakes are covered by gaps, they will be misclassified. Errors caused by striped gaps of Landsat ETM+ were manually corrected using additional high-quality scenes during the whole year with assistance of images from adjacent years. (iii) the modified Normalized Difference Water Index (MNDWI) was calculated (Hanqiu, 2006); (iv) the potential glacial lake areas were extracted by applying adaptive MNDWI threshold (Li and Sheng, 2012). To define a glacial lake in the image, the minimum number of water pixels was inconsistent in different studies. For example, Zhang et al., 2015 set the smallest detectable glacial lakes in the Third Pole of larger than 0.0027 km^2 (three connected pixels) using the Landsat TM/ETM+ data. Nie et al., 2017 selected 0.0081 km^2 (nine connected pixels) as the

minimum mapping unit to map glacial lakes in the Himalaya. Other studies set the minimum threshold areas as $0.001km^2$ (Salerno et al., 2012), $0.002km^2$ (Wang et al., 2013a), $0.0036km^2$ (Luo et al., 2020), $0.0054km^2$ (Wang et al., 2020b), $0.01km^2$ (Li et al., 2020) for the identification of glacial lakes and analysis of their spatial and temporal variations. A smaller minimum

205   mapping unit will detect more glacial lakes, however, the uncertainty it brings is also larger than large lakes at the same resolution (Salerno et al., 2012). Our results demonstrate that a lake area covering fewer than nine water pixels have an area error of larger than 50% (Please see the Section 4). Given the area uncertainty of glacial lakes and the spatial resolution of Landsat data, in this study, glacial lakes larger than nine pixels ($\geq$ 0.0081 km^2) were considered as the minimum mapping unit; and (v) manual inspection and refinement of individual glacial lake were conducted and the related attribution were

210   added for each lake.

    Based on the automated processing, nearly 60% glacial lakes in each year can be correctly classified, of the other lakes that were not properly classified, 30% were missed and 10% were misclassified. For such a large-scale area that characterized by various and complex climatic, geological and terrain conditions, this classification method is simple but effective, the results are also reasonable since it provides very low commission errors. To ensure the quality of inventory, strict quality

215   control was conducted to visually inspect and correct the mapping errors after the automated processing using GEE. False lake features, mainly identified as mountain shadows and river segments, were manually removed by overlapping mapped lake shorelines on the source Landsat imagery and higher-resolution imagery in Google Earth. For missing glacial lakes, such as some lakes may be covered by ice and clouds for years, grow at steep glacier tongues, and lakes show heterogeneous reflectance with the surrounding backgrounds, the lake boundaries were edited further using ArcGIS. Furthermore, a cross-

220   check and modification was conducted for each glacial lake based on the lake mapping results in conjunction with multi-temporal Landsat imagery. Here all the Landsat imagery that used for the inspection were downloaded manually from USGS Earth Explorer website (https://earthexplorer.usgs.gov/). Outputs per lake polygon include the information about lake type, elevation, Euclidian distance to the nearest glacier terminus, area and perimeter. Noted that if there are more than one suitable satellite images in a year, image with the least cloud cover will be selected for the calculation of the area and perimeter of a

225   given lake. Meanwhile, each mountain range was characterized individually by utilizing the mountain boundary shapefile in High Mountain Asia (geo.uzh.ch/~tbolch/data/regions_hma_v03.zip).
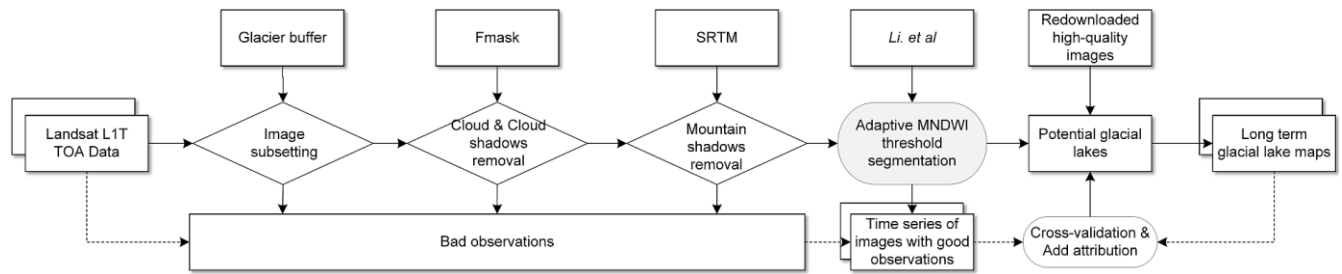

Fig. 3. Diagram of the glacial lake mapping workflow.

230   **3.3 Yearly lake area changes calculations**

    Based on the final generated lake inventory data, we used the slope of linear regression of lake area (over the grid cell of $1°\times1°$) versus mapping year to qualify the yearly lake area changes during the study period. The approach to change analysis

was predicted using a Theil-Sen estimator, which chooses the median slope among all the derived fitted lines, can effectively represent long-term area changes due to its robustness for the trend detection and insensitivity to outliers, it is also useful for the elimination of the effect from differences in the sensor performance for the mapping of glacial lakes (Kumar, 1968;Song et al., 2018). Although all the lake were manually checked and edited, due to the limitation of available images and other factors, the conditions of glacial lake mapping are not perfectly consistent for each year. For example, the image dates are not consistent across the whole HMA region because of atmospheric disturbances, also the influences from varying lake characteristics, image quality (Bhardwaj et al., 2015;Thompson et al., 2012), ice and shadow that obscured the lakes, which all contributed to the detection errors in the lake extent and their annual variation. Generally, these errors are objective and acceptable as a result of the nature of the limited remote sensing data. For this study, because we used long time series (10 years) data for the estimation of annual lake area change, and also the errors only accounts for a small proportion of the total glacial lake area for each year, errors in the observed lake area caused by the different effects do not apparently affect the trend statistical results. For the glacial lake area time series in each 1°×1°grid, we applied the Theil-Sen estimator to smooth the annual time series of data and derive the slope (annual change) of the trend. A Mann-Kendall trend test was used to detect and further confirm the statistical confidence by the linear regression results. All the estimated trends fall within the 90% confidence intervals. The upper and lower change estimates that satisfy the 90% confidence interval for the slope were also derived over the whole HMA region (Fig. A2).

4 Cross-validation and uncertainty estimate

Accuracy assessment of the mapping results is difficult due to the lack of field measurements of glacial lakes in the continental-scale area like HMA. To obtain the quality controlled data, the glacial lake vector over the entire HMA for the years from 2008 to 2017 has been rechecked and reedited individually through dynamic cross-validation by ten trained experts, which is a time consuming process but are essential to maximize the quality of glacial lake change detection.

For the estimation of the uncertainty of glacial lake area, a key influence factor is the spatial resolution of satellite data. In this study, the uncertainty of the glacial lake area was estimated as an error of ±1 pixels on either side of the delineated lake boundary. The percentage error of area determinations, A_{er}, then is proportional to sensor resolution and is given by (Krumwiede et al., 2014):
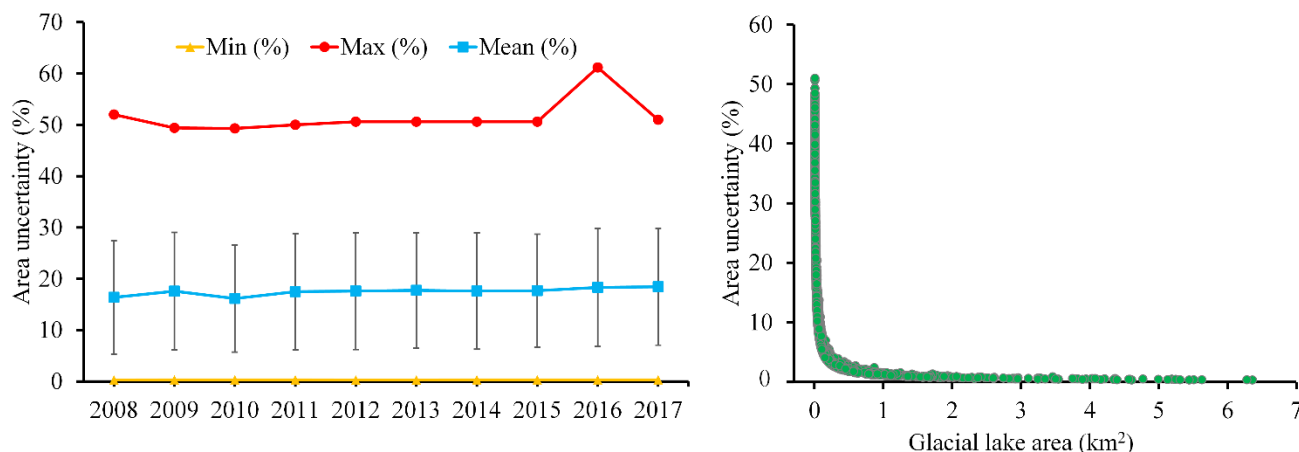
$$A_{er}=100 \cdot (n^{1/2} \cdot m) / A_{gl} \tag{1}$$

Where n refers to the number of pixels on the boundary of a glacial lake, approximated by the ratio of the perimeter length and spatial resolution, m is the area of a pixel in the Landsat image (m^2), A_{gl} is the lake area (m^2) and the factor 100 is there to convert to percentage.

Assuming an uncertainty of 1 pixel for the detected glacial lake boundaries, we calculated the systematic errors for the whole HMA region (Fig. 4). For the year between 2008 and 2017, the area uncertainty of each glacial lake generally ranged from 0.30% to 50%, with the mean value falling around the 17%, and standard deviation around 11% (Fig. 4a). The maximum and mean value of area uncertainty for the glacial lakes in 2010 are the lowest, while for the year of 2016, the corresponding statistics are highest, this can be attributed to the different factors. The maximum of area uncertainty of glacial lakes is related with the shape and size of a certain lake (as can be seen from equation (1)), but its mean value is equal to the sum of the area uncertainties of each glacial lake divided by the total number, which depend on the total number of glacial lakes in a year,

270

275

5 Results

5.1 Distribution of various types and sizes of glacial lakes

The area coverage of glacial lakes increased by 90.14 km^2 between 2008-2017, a 6.90% increase relative to 2008 (1305.59 $\pm$ 213.99 km^2) (Fig. 5a). A Theil-Sen regression fit to all the data showed a mean expansion rate of 12 km^2 a^{-1} for the 10-year record (Fig. 5a). Meanwhile, the estimated changes in glacial lake number from 2008 (12,593 lakes) to 2017 (15,348 lakes) showed an average increase of 306 lakes a^{-1}. The steeper percentage increase in lake number (22.33%) compared to a slower area expansion (8.79%) based on their linear fit trends showed that many small glacial lakes formed over this decade. The number of lakes increased most rapidly above 4400 m a.s.l., especially above 5300 m (Fig. 5b). The increase of proglacial lakes was concentrated above 4900 m (Fig. 5c). Unconnected glacial lakes grew very slightly in total area below 4400 m (Fig. 5d), but increased at higher elevations. Glaciers are retreating and thinning at ever-higher elevations (Nie et al., 2017), causing the formation of new supraglacial lakes at high-elevation, expansion of existing ice-contact lakes, and detachment of glaciers from some lakes.
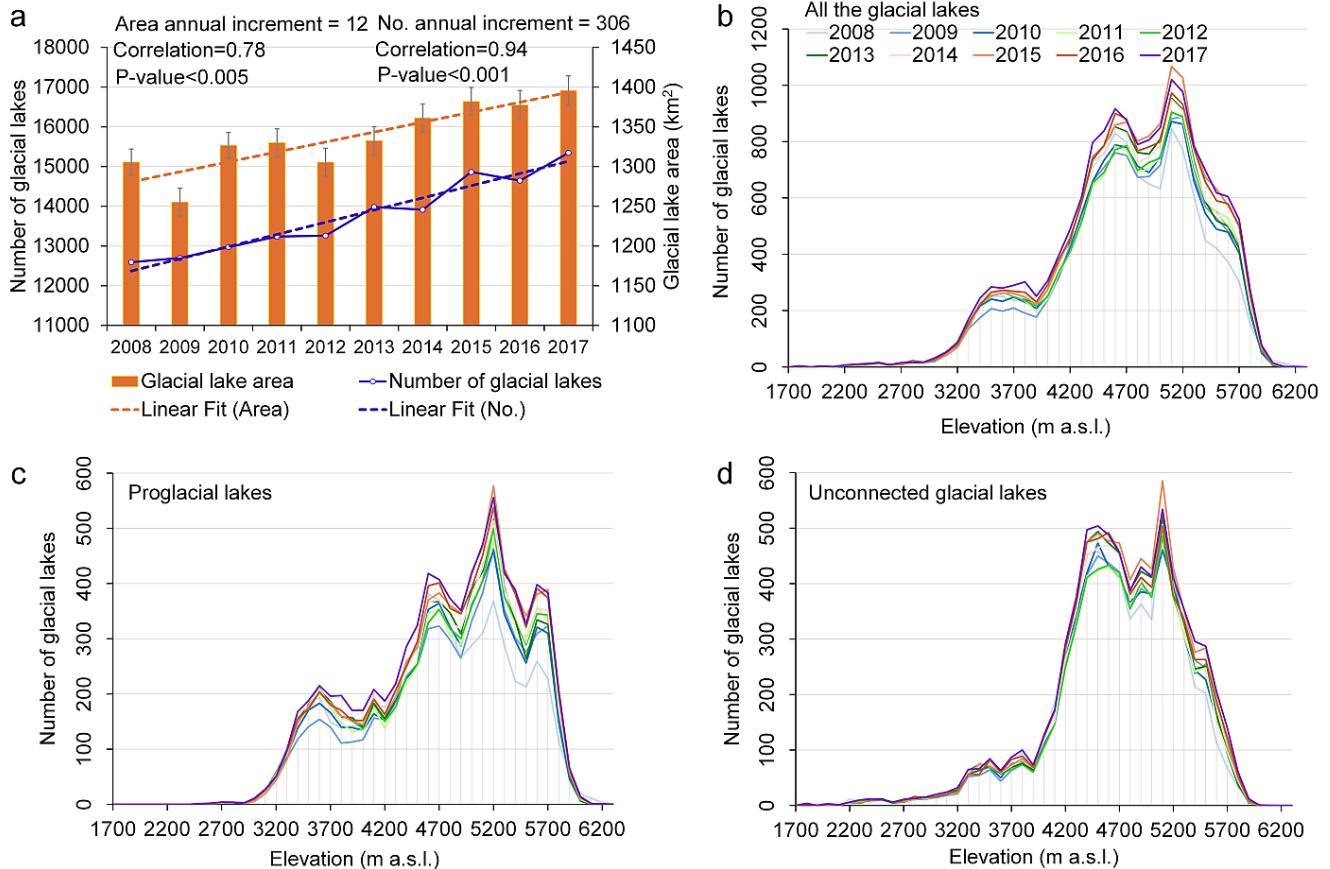
280

285

Fig. 5. Annual glacial lake number and area. (a) Total number and area of glacial lakes for HMA between 2008-2017. The annual increment is the slope of the trend of annual lake area and number. Altitudinal distribution (100-m bin sizes) of lake numbers for **(b)** all glacial lakes, **(c)** proglacial lakes, and **(d)** unconnected glacial lakes.

Annual changes in glacial lakes were further analyzed spatially using a $1° \times 1°$ grid over 22 mountain regions (Fig. 6a) using Theil-Sen regression analysis. An analysis on mountain-wide lake area loss/gain from 2008 to 2017 was conducted (Table A1). Negative or undiscernible changes in glacial lake area were observed in the Eastern Tien Shan, Eastern Hindu Kush, Hengduan Shan and Eastern Tibetan Mountains (Fig. 6b), thus reducing overall increasing glacial lake area in HMA. The Eastern Hindu Kush lost 2.8 km^2 of lake area (Table A1), with the negative area change (-0.43 km^2 a^{-1}) near 35°N, 73°E. Glacial lakes in Nyainqentanglha, Gangdise Mountains exhibited area loss and gain in some regions. In contrast, Central and Eastern Himalaya, Central Tien Shan showed rapid lake area increases. Between 2008 and 2017 Central Himalaya's glacial lake area increased by 27.09 km^2 (Table A1), exhibiting both a high density of 47 glacial lakes per 100 km^2 in 2017 (Fig. A3) and rapid growth, +0.94 km^2 a^{-1}, in lake area due to retreat and thinning of debris-covered glaciers (Song et al., 2016). Moderate area gains occurred along most of the Western Kunlun and Tanggula Shan, e.g., +0.38 km^2 a^{-1} in Tanggula Shan. Glacial lake area in Pamir Alay, Eastern Pamir and Eastern Kunlun Shan was spatially and temporally invariant across the whole observation record.

We found that glacial lakes exhibited different expansion trends for different lake types and supraglacial and ice-marginal lakes have relative few coverage areas comparing with proglacial and unconnected lakes (Fig. 6b and Fig. 6c). In the Nyainqentanglha and Central Himalaya, around half of the glacial lake area consisted of proglacial lakes, where most growth

79

occurs. In the negative lake growth (shrinkage) regions of Eastern Tibetan Mountains and Hengduan Shan, the unconnected glacial lakes were dominantly occupied. As the interaction with the glacier gradually weakened, part of the water source supplied by glaciers is reduced. Also, combined with the effects from atmospheric warming and decrease of precipitation, regions mainly consist of unconnected glacial lakes show the decreasing trend in area. Proglacial lakes contributed approximately 62.87% (56.67 km^2) of total area increase over HMA (Table A1 and A2). Proglacial lakes in the Central Himalaya, Eastern Himalaya, and Western Himalaya, accounted for 36.27% (32.70 km^2) of the total area increase. In general, proglacial lakes are a main contributor to recent lake evolution in HMA.

We also noted the large area growth of lakes occurred in areas with relatively large proportion of small glacial lakes, mainly due to rapid growth of existing lakes and new lake formation (Fig. 6d). For example, in some areas of Central and Eastern Himalaya, Nyainqentanglha that have large annual increase in lake area (higher than 0.23 km^{2}a^{-1}), glacial lakes with a size of less than 0.16 km^2 occupied more than 30% of total area (Table A3). Especially for the region in Nyainqentanglha, the area of small glacial lakes (≤0.16 km^2) even accounts for 69.47% of the total area.
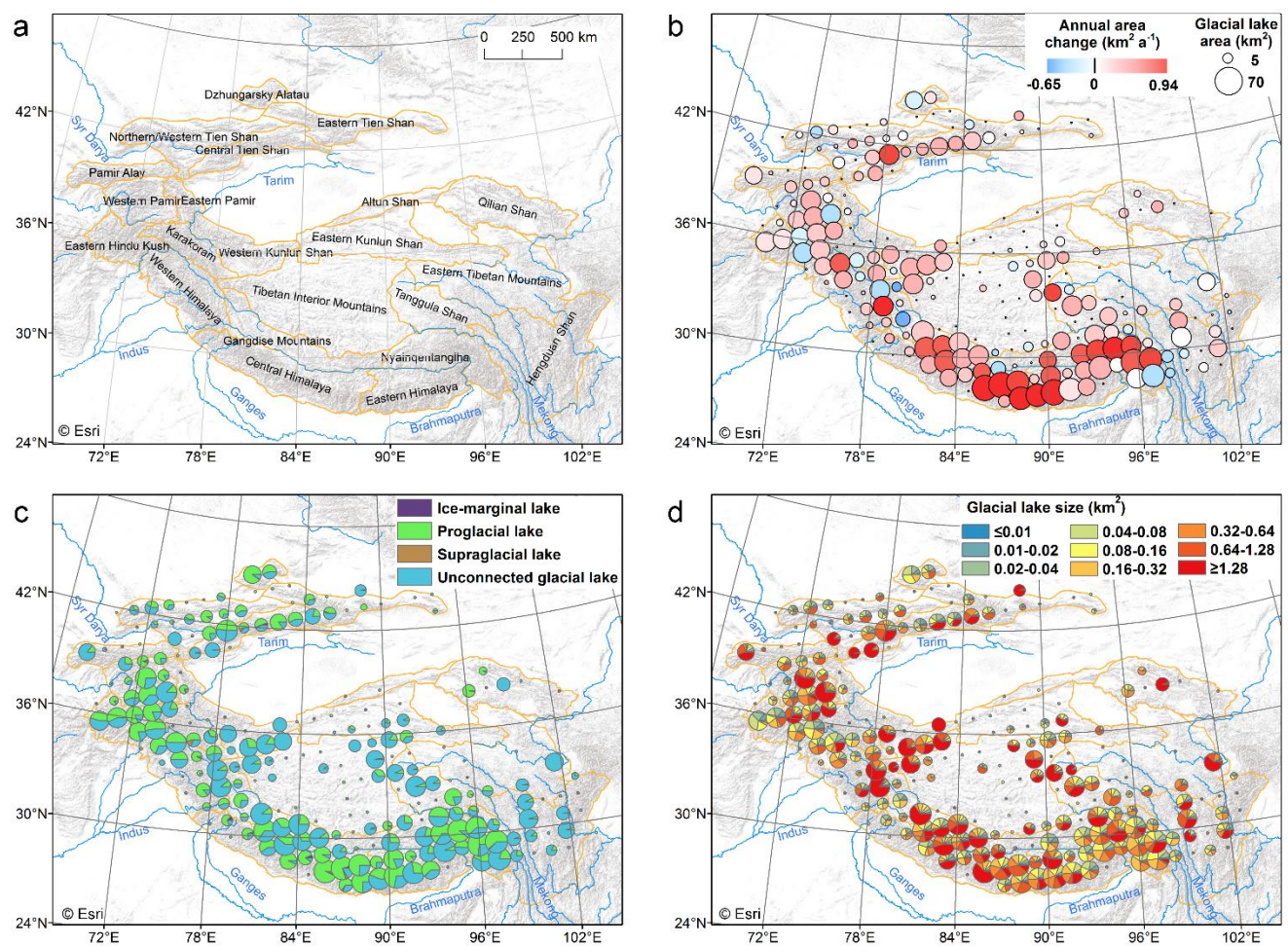

Fig. 6. Glacial lake area changes and area distribution. (a) Geographic coverage of mountain ranges in HMA. (b) Annual rate of change in lake area (2008-2017) on a 1°×1° grid. The size of the circle for the area in 2017. (c) Proportional areas of four types of glacial lakes in 2017. (d) Area of different sizes of glacial lakes in 2017. The terrain basemap is sourced from Esri (© Esri).

5.2 Influencing factors of current distribution of glacial lakes

To explore potential factors that have influenced glacial lake distribution across HMA, we focus on proglacial and supraglacial
lakes, for which the changes are closely related with glaciers and expansion is most rapid. Proglacial lakes frequently develop
from the enlargement and coalescence of one or more supraglacial lakes (Haritashya et al., 2018;Thakuri et al., 2016).
Proglacial and supraglacial lake development from 2008 to 2017 is significantly correlated to initial lake area in 2008 (R=0.82,
Table A4); larger ice-contact proglacial lakes imply a larger water body in contact with the calving-front of the glacier, and
more rapid retreat (Truffer and Motyka, 2016;King et al., 2019).

For the years before 2008, the year-round Landsat 5 TM data in many years do not fully cover the HMA region. In this
study we constructed the inventory over a ten-year time period, which is shorter than typical glacier response times which
start from a minimum of 10 years for short, steep glaciers, to over 150 years for long debris-covered glaciers (Scherler et al.,
2011). Hence, lake expansion is not expected to couple with short-term climate trends, particularly for debris covered glaciers
(Haritashya et al., 2018). In the inclusion of mass balance forcing of glacial lake changes, the same questions about the
response times also occur. Hence, rather than focus on the short term evolution of lake expansion, we investigated if climate
and other factors have influenced the overall distribution of lake area, as observed in 2017. To investigate factors influencing
the predominance of proglacial and supraglacial lakes, geomorphic, topographic and climate parameters were correlated with
lake area over a 1°×1° grid, and aggregated (mean or summed) for HMA regions. A statistically significant positive correlation
exists between lake area and debris-covered glacial area (after Scherler et al. 2018) across HMA (R=0.36, Table A4),
confirming the predominance of pro- and supraglacial lakes forming on debris covered glacier tongues (Nie et al. 2017).
Correlations and significance levels strengthen if the Karakoram is excluded (Table A5). The Karakoram is known as an
anomaly of positive glacier mass balances and glacier advances (Gardelle et al., 2012) and also has an anomalously small area
of proglacial lakes. Glacier length (RGI-Consortium, 2017) and debris cover are strongly correlated (R=0.85, Table A4),
reflecting abundant debris on most large, low-gradient valley glacier tongues in HMA; in turn, there is a statistically significant
direct correlation between glacier length and lake area (R=0.32, Table A4), as these tongues provide the ideal conditions for
the coalescence of supraglacial ponds and formation of large proglacial lakes (Shaun D and John M, 2000;Nie et al., 2017).
Glaciers are generally longest and most heavily debris covered in the Hindu Kush-Himalaya (Fig. 7a and Fig. 7b).

Some regions have comparable amounts of large debris-covered glaciers but substantial differences in total lake area and
area-growth rates (for example, Central Himalaya compared to Central Tian Shan or Western Pamir, Table A5) . Regional
differences in multi-decadal climate trends could play a role, with Nyainqentanglha, Central and Eastern Himalayan regions
all characterized by rapid warming and decreased precipitation since 1979 (Fig. 7c and Fig. 7d), favoring negative glacial
mass balances (Brun et al., 2017a). This plausibly explains why lake area is typically larger in these regions relative to adjacent
regions further to the west and north (e.g. Western Himalaya) despite often similar glacier characteristics (in terms of debris
cover and glacier length) (Fig. 7e and Fig. 7f). Further, there is very little debris covered area but rapid warming in Eastern
Himalaya, where proglacial lakes are abundant (Fig. 7f). These results emphasize that the distribution of supra and proglacial
lakes across HMA are primarily associated with the presence of large debris covered glaciers, but regional variability in
warming and precipitation trends over the past few decades have likely also had some influence (Haritashya et al.,
2018;Scherler et al., 2018;Dan and CLAGUE, 2011;Bo et al., 2019). These results are consistent with previous findings at

regional scales, that have demonstrated a rapid expansion of proglacial lakes on debris-covered glaciers, with expansion in the upstream direction demonstrated to occur primarily through a process of subsidence at of the lake-contact debris covered glacier tongue (Harrison et al., 2018;Song et al., 2016;Song et al., 2017a).
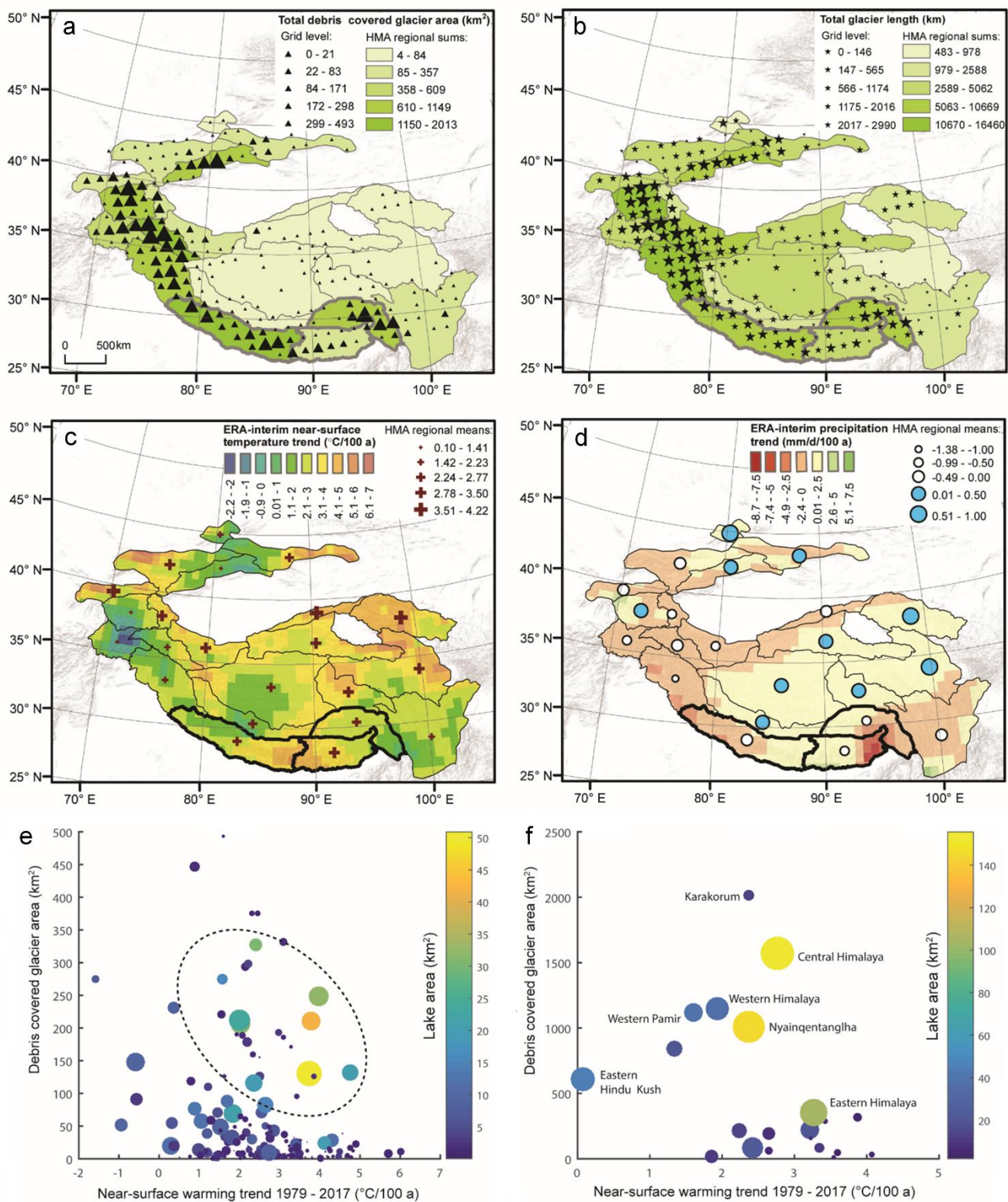
Fig. 7. Geomorphic and climatic influences on lake distribution. (a) Debris-covered area and **(b)** glacier length aggregated on a 1°×1° grid. Linear trends in **(c)** temperature and **(d)** precipitation calculated for 1979-2017 from ERA-Interim, including aggregated means over HMA regions. Relationship between total debris-covered area, near-surface temperature warming, and proglacial and supraglacial lake area of 2017 in **(e)** 1°×1° grid tiles and **(f)** HMA regions. Some regions discussed in the text are labelled. The lake coverage is high in areas of both rapid warming and high debris cover (E, dashed ellipse). Dot sizes are proportional to lake number. See Table A4 for details on data sources.

6 Discussions

6.1 Comparison with other lake datasets

We compared our dataset with that of Wang et al., 2020 for the closest period (2017 from Hi-MAG database and 2018 from Wang et al. 2020) over the spatial extent of our HMA region. The differences in the total number and area between these two dataset are 6206 and 223.97 km^2, respectively. We also found that 2077 glacial lakes with a total area of 178.77km^2 in our Hi-MAG dataset were not detected by Wang et al. The main reasons for the missed glacial lakes in Hi-MAG are because the interference of some bad observations (cloud or snow), glacial lake dried up or outburst, or located in the middle of the river.

To test the spatial correlation of glacial lakes distribution in two datasets, we compared the statistics in glacial lake number and area aggregated on a 0.1°×0.1° grid for the HMA regions. The results for the total glacial lake area, areas for glacial lakes larger than 0.04km^2, and number for glacial lakes larger than 0.04km^2 are depicted in Fig. A4. A clear and strong correlation can be observed for all the statistics between Hi-MAG dataset and glacial lake data by Wang et al. Most of the points being distributed around the 1:1 line, which shows that there is great consistency in the results.

In order to quantitatively and systematically evaluate the accuracy of our product, we implemented a stratified random sampling (Song et al., 2017b;Stehman and SV, 2012), where the glacial lakes were divided into four strata. The sample sizes are the spatial resolution (30 m) of the data, and the strata are designed as: C0W0. both the results are non glacial lakes; C0W1. non glacial lake for Chen's and glacial lake for Wang's; C1W0. glacial lake for Chen's and non glacial lake for Wang's; C1W1. Both the results are glacial lakes.

A total of 4,000 points were randomly selected, as shown in Fig. A5. The sample number for C1W1 and C1W0 are 1300 and 700, respectively, which is almost the same ratio between the total areas for the two strata (1450.50 km^2 vs 732.77 km^2). Because of the approximate total area with C1W0, we also randomly selected 700 samples from stratum C0W1. The rest 1300 samples are from C0W0. Every validation sample was visually examined using Landsat imagery and higher-resolution imagery in Google Earth. Sample pixels were interpreted by a regional glacial lake mapping expert, and ambiguous samples were cross-validated by a second observer. If a sample is difficult to interpret, it was marked as ambiguous sample and excluded for the accuracy assessment. The sample number estimates were produced for each of the four strata (Table A6), and these strata totals were then summed to obtain the total accuracy.

For the 1300 pixel samples that were considered to be non glacial lakes by both datasets, after the pixel by pixel verification, 1215 were indeed non glacial lakes, while 37 were the missed glacial lakes. In contrary, 1260 out of 1300 pixels belongs to the class of glacial lake, and 25 pixels were misclassified as glacial lakes by the two inventories. 307 error pixels were found in the results from Wang et al., constitute about half of the total validation number. For the glacial lakes identified only by our inventory, 678 out of 700 were corrected classified. Our results yielded high overall classification accuracy (88%), user's accuracy (97%), and producer's accuracy (82%) for glacial lake classification using Landsat data.

405   Hi-MAG was also compared with other Landsat-based lake inventories (Nie et al., 2017;Zhang et al., 2015;Pekel et al., 2016a). Hi-MAG lake number was 7268 higher and area was 644.26 km^2 higher than the estimation for the Tibetan Plateau (Zhang et al., 2015). The largest discrepancy is in the Gangdise, Himalaya and Nyainqentanglha Mountains in 2010. Across the Himalaya, we found 476.09 km^2 of glacial lakes, 4.57% more than previous estimates in 2015 (Nie et al., 2017). In addition, we qualitatively compared the lake extent between publicly available high-resolution Global Surface Water (GSW)

410   dataset and our Hi-MAG database summed by mountain range in 2015. GSW data can be accessed at https://global-surface-water.appspot.com/download. For the sake of a reliable comparative analysis, lake polygons in the Hi-MAG dataset were converted into the grid format, and glacial lakes in the GSW were further extracted using the range of glacier buffer (10 km). Hi-MAG detected more glacial lakes in the Himalaya, Eastern Hindu Kush, and Tien Shan, and fewer in Eastern Pamir and Western Kunlun Shan. Fig. A6 illustrates the differences between our Hi-MAG glacial lake results and GSW-derived lake

415   area for the whole HMA region.

        The glacial lake area observed in our lake dataset in the Eastern Pamir and Western Kunlun Mountains does not conform to the mapped surface water in the GSW for these sub-regions. While there are numerous glacial lakes from an open water perspective, actually part of them are river segments. Additionally, the Himalaya, Eastern Hindu Kush, and some other Tien Shan host thousands of glacial lakes that are not readily observable in the GSW product. Large discrepancies in mountainous

420   glacial lake estimates preclude a significant consistency between GSW and our Hi-MAG lake data over the HMA region. The region with the highest consistency between GSW and Hi-MAG product is interior Tibet. There is little agreement for Tien Shan, where the weather is rainy and snowy in the region above 3000 m, and large amounts of ancient glacial deposits have been accumulated. Here glacial lakes are featured by small size, due to the influence of source glaciers and lake beds, as well as the water depth and sediment inflow, glacial lakes appear heterogeneous reflectance in the image. Errors could exist in

425   datasets produced by automated classification, but we also did a detailed manual editing, so we were not relying exclusively on automatic classification. Karakoram regions seem to have fewer glacial lakes in our estimate, owing to the overestimation of surface water on debris covered glaciers in the GSW dataset.

        The little agreement between our Hi-MAG glacial lakes data and GSW data is mainly due to its lack of systematic glacial lake inventories and mapping capabilities. The lake dynamics and differing climate contexts within HMA may also lead to

430   inconsistencies between the sub-regions. Hi-MAG might have made better use of the optimum satellite imaging season to map glacial lakes, potentially resulting in more complete mapping by avoiding conditions — such as periods of lake ice — that may confound mapping.

6.2 Known issues and planned improvements

There are several important issues and limitations to the datasets produced and methods used within this study that are

435   important to highlight to potential users. (i) Bodies of water smaller than nine connected pixels (e.g., 1 x 9 pixels or 3 x 3 pixels, corresponding to 30 x 270 m or 90 x 90 m, respectively), those obscured by frozen water surface and loose moraines or hidden by terrain shadows were not included. Broken floating ice or isolated moraine that stand in open water for some times were mapped. Supraglacial lakes such as melt ponds developed on the surface of glaciers present particular challenges because of their small size and highly dynamic properties. Most supraglacial lakes are transient or seasonal, or at least fluctuate

440   seasonally, as they commonly drain and may refill, but in fact this short-duration seasonal water more generally is likely to be underestimated because of temporal discontinuities in the archive and gaps caused by persistent cloud cover. (ii) The spatial

and temporal information reported in the Landsat dataset used in this study complements that acquired in the past. Nevertheless, the biggest limitation to glacial lake mapping from these data are undoubtedly the geographic and temporal discontinuities of the Landsat archive itself. Historical data over the entire HMA before 2008 can be recovered partly from the Landsat 4 TM/MSS, Landsat 5 TM, Landsat 7 ETM+, and partly from SPOT, and other satellite systems, etc., although data access is not always at the full, free and open level of Landsat. In this regard, ASTER is freely accessible and has higher resolution than Landsat, but the temporal coverage is very limited in most of HMA. Other Landsat-like moderate resolution multi-spectral sources could be also used to improve and extend the temporal sampling. For example the European Space Agency's Sentinel 2a satellite launched in 2015 and provides optical imagery at 10 m resolution (Wang et al., 2020a), which will benefit future research combing all available satellite observations with GEE cloud computing power would make long-term monitoring of changes to HMA's glacial lakes and inland waters possible.
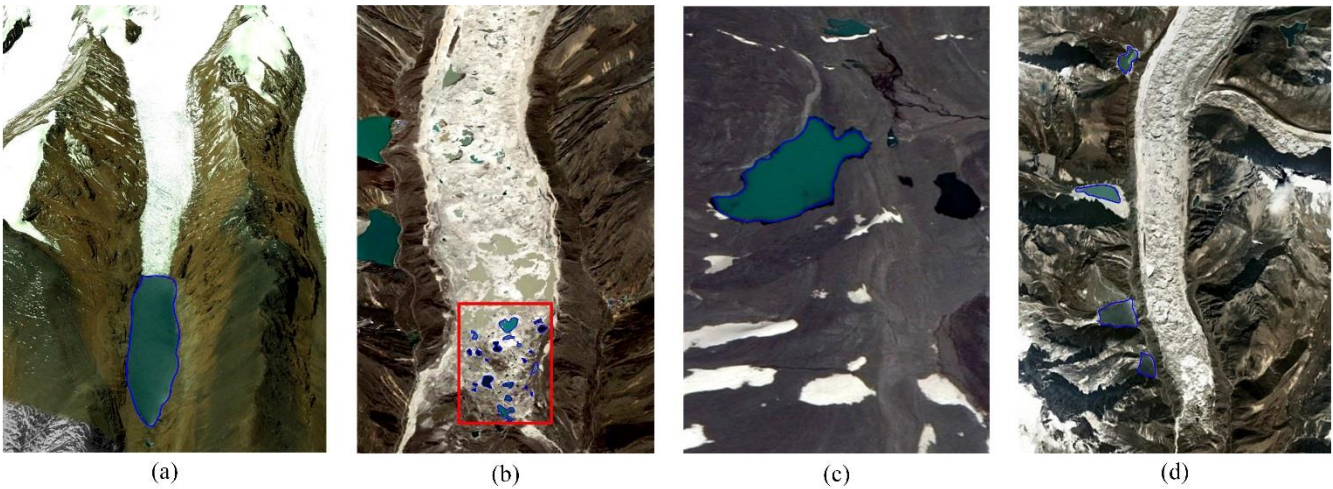
7 Data availability

The Hi-MAG database is distributed under a Creative Commons Attribution 4.0 License. The data can be downloaded from the data repository Zenodo at https://doi.org/10.5281/zenodo.4059181 (Chen et al., 2020).
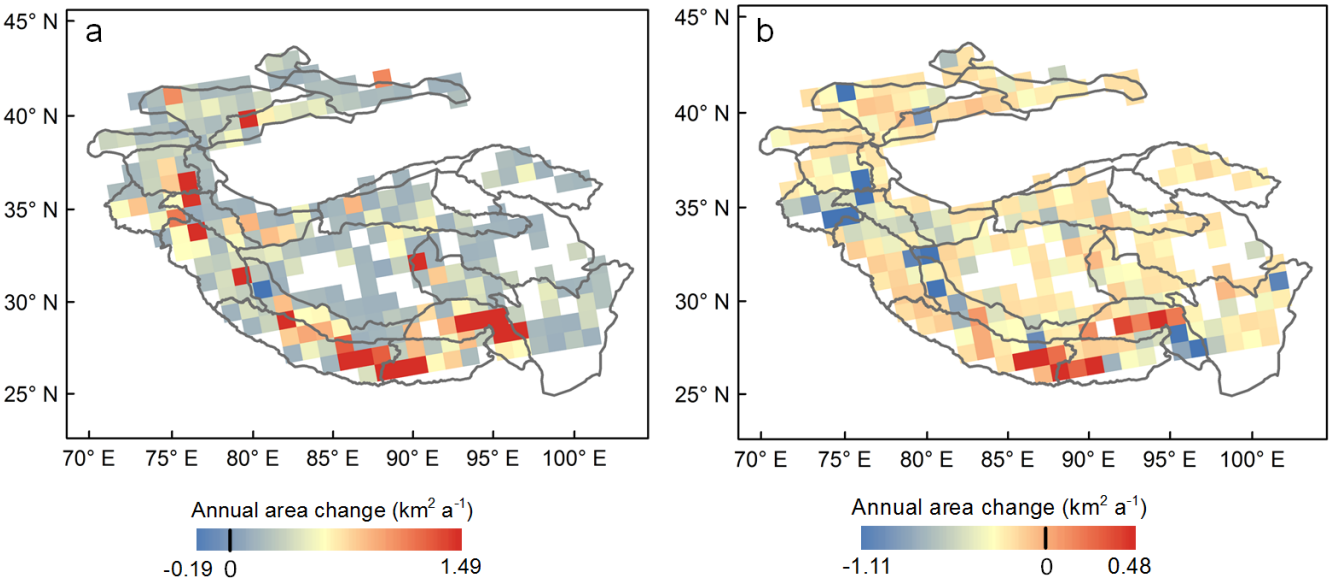
8 Conclusions

In conclusion, the Hi-MAG dataset and others have turned to Earth observation satellite data, especially Landsat imagery, to provide a more consistent delineation of large-scale glacial lake changes. Some remote-sensed glacial lake mapping methods have enabled local-scale area estimation or spatial representation of lake extent and change. Such methods result in relatively good performance for lake areas that remain clear and show homogeneous reflectance in the image, but do not allow for continental-scale glacial lake mapping that have spectral interference from the other objects such as glaciers, snow, clouds, turbidity and sedimentation characteristics of glacial lake itself, or the atmospheric interference and terrain effects. Automated methods for the extraction of glacial lakes over the large-scale areas are further developed in our work. However, visual interpretation and manual editing is still an effective way to ensure the high accuracy of lake inventories and appended attributed information for further analysis. Based on an error of ±1 pixels on the lake boundary, the area uncertainty of each glacial lake ranged from 0.30% to 50% for the year between 2008 and 2017, and the mean area uncertainty of 17% in the entire HMA region.

Mapping of glacial lakes across the Tibetan Plateau and adjoining ranges reveals a complex pattern of lake occurrence and growth/shrinkage. During the past ten years, 2755 glacial lakes with a total area of 90.14 km^2 were increased in the HMA region. Proglacial lakes contributed 62.87% of that increase. We found that most of areas in HMA have experienced rapid expansions, Central and Eastern Himalaya, Central Tien Shan showed the most lake area increases (up to +0.94 km^2 a^{-1}). Negative area changes were observed in the Eastern Tien Shan, Eastern Hindu Kush, Hengduan Shan and Eastern Tibetan Mountains. The number of lakes grown very rapidly above 4400 m a.s.l. Proglacial lake growth is proceeding at high elevations of above 4900 m, but glacier retreat and lake disconnections are also starting to occur at higher elevations, causing the number and area of both classes to increase. At low elevations, few glaciers remain where proglacial lakes can form, and already detached lakes lack growth mechanisms. Overall, continued growth of glacial lakes can be expected, particularly where large debris covered tongues remain.

This freely-downloadable, detailed Hi-MAG dataset can also be used in future studies to provide a sound and consistent basis on which to quantify critical relationships and processes in HMA, including glacier-climate-lake interactions, glacio-hydrologic models, glacial lake outburst floods and potential downstream risks and water resources.

Appendix A

<div align="center">(a)   (b)   (c)   (d)</div>

Fig. A1. Examples of the various types of glacial lake found in the HMA: (a) pro-glacial lakes, which are connected to the parent glacier and usually impounded by a debris dam (usually a moraine or ice-cored moraine); (b) supraglacial lakes (denoted by the red rectangle) which develop on the glacier surface; (c) unconnected glacial lakes; and (d) ice-marginal lakes that distributed on the edge of a glacier. Background images were acquired from © Google Earth, and were shot in 2009, 2011, 2012 and 2014, respectively. Glacial lake outlines for each type are shown in blue color.

Fig. A2. Annual changes in lake area between 2008 and 2017 on a 1°×1° grid. The (a) upper and (b) lower slopes represent the 90% confidence interval.

Fig. A3. Density (number per 100 km^2) distribution of glacial lakes in 2017.
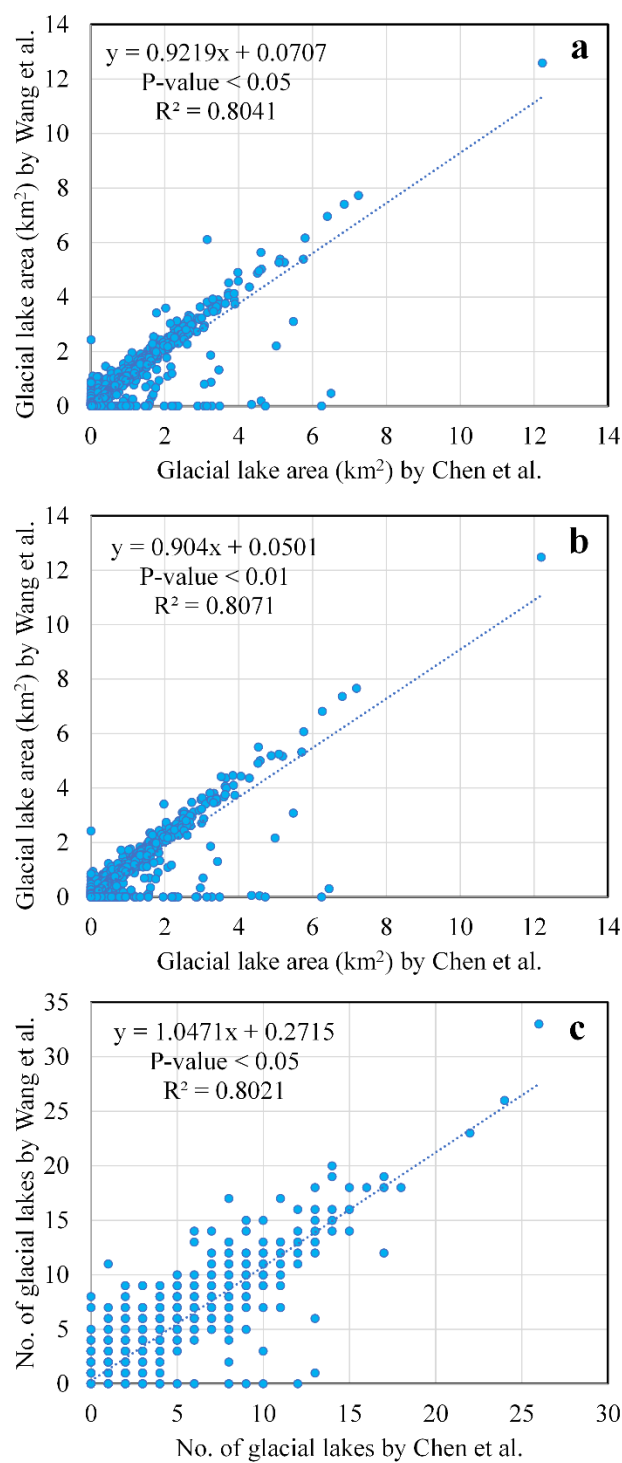
Fig. A4. Comparison the results of (a) total glacial lake area, (b) areas for glacial lakes larger than 0.04km², and (c) number for glacial lakes larger than 0.04km² summed over a 0.1°×0.1° grid between Hi-MAG database and inventory by Wang et al., 2020.

495
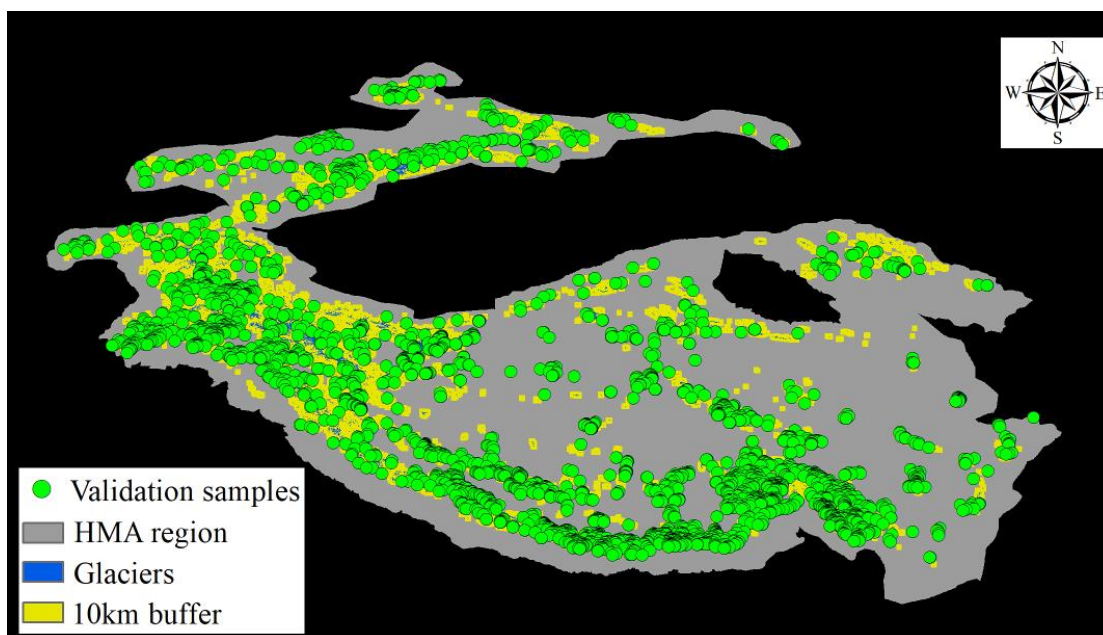
500 **Fig. A5. Distribution of validation samples selected using stratified random sampling. Blue polygons are glacier outlines taken from the Randolph Glacier Inventory (RGI v5.0), the Second Chinese Glacier Inventory (CGI2) and the GAMDAM inventory. Yellow polygons refer to buffer area within 10 km of glacier terminals.**
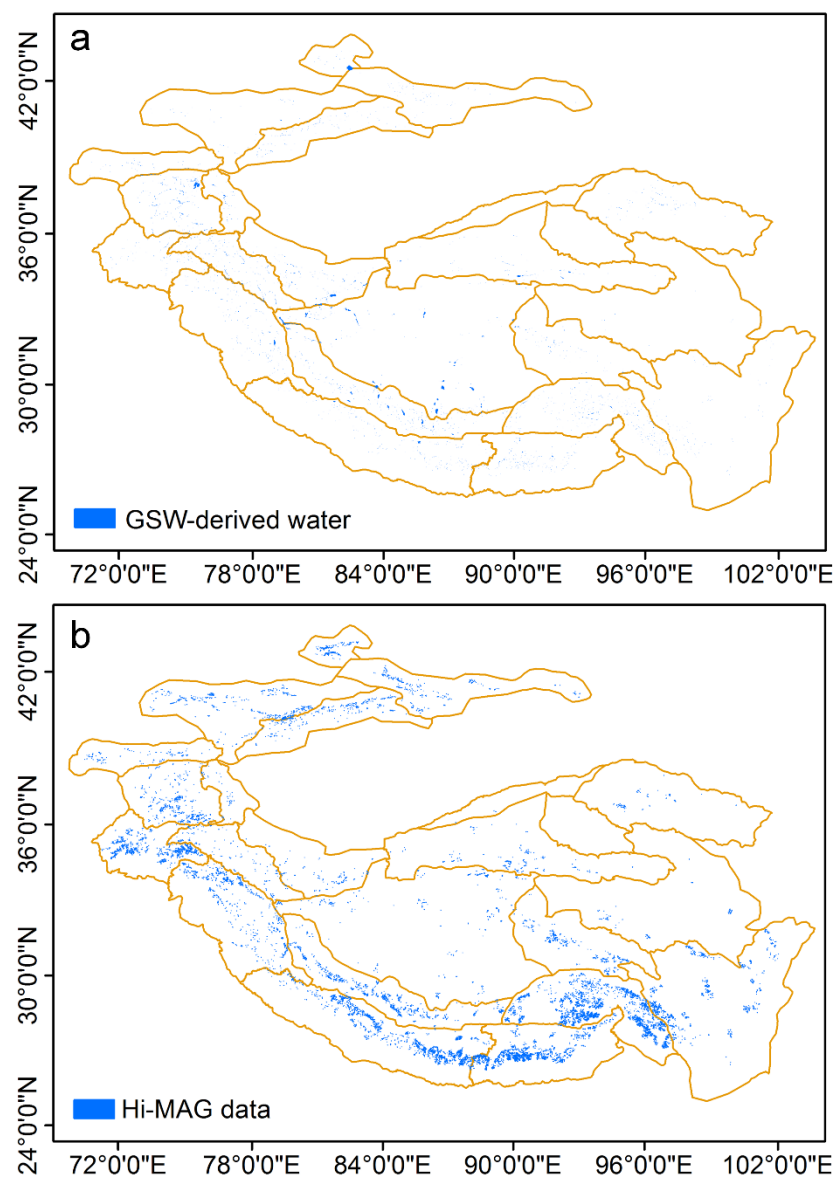
505 **Fig. A6. Comparison of the glacial lake measured in the global maps of (a) Pekel et al. and (b) our Hi-MAG data.**

Table A1. Mountain-wide glacial lake number and area per year and total loss/gain from 2008 to 2017. The unit of area is km².

Mountain range	2008		2009		2010		2011		2012		2013		2014		2015		2016		2017		Total gain/loss (2008-2017)	
	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area	No.	Area
Eastern Hindu Kush	1227	73.63	856	61.40	1113	68.91	1036	67.94	1015	59.80	1206	67.90	1179	71.30	1172	66.36	1260	65.77	1399	70.83	172	-2.80
Western Himalaya	894	77.65	856	70.23	756	71.25	747	71.38	746	68.95	779	76.00	874	82.76	823	78.33	999	79.51	1005	87.96	111	10.31
Eastern Himalaya	1634	164.12	1667	163.10	1675	166.49	1836	172.87	1859	171.67	1910	176.54	1943	174.93	2159	190.96	1954	179.09	1943	179.80	309	15.68
Central Himalaya	1312	182.54	1697	179.04	1577	188.35	1847	195.77	1782	195.56	1728	196.72	1850	198.52	2049	206.80	2096	206.41	2182	209.63	870	27.09
Karakoram	167	18.89	214	16.78	206	17.91	159	14.90	152	13.55	182	14.76	163	15.41	182	17.02	250	18.26	219	18.60	52	-0.29
Western Pamir	481	75.53	486	77.64	526	83.27	548	80.86	495	79.17	537	79.32	550	81.20	557	83.96	570	81.76	624	75.82	143	0.29
Eastern Pamir	38	5.06	45	5.69	48	5.81	43	4.99	56	5.75	50	5.43	50	5.26	49	5.65	58	5.66	56	5.81	18	0.75
Pamir Alay	124	10.82	79	9.76	100	10.93	132	11.79	129	11.60	127	11.74	128	11.66	137	11.79	130	11.54	131	12.10	7	1.28
Northern/Western Tien Shan	474	36.16	358	37.20	499	36.87	541	41.71	522	36.04	518	36.98	551	46.06	530	38.28	512	37.38	626	40.91	152	4.75
Central Tien Shan	307	29.19	241	29.43	305	32.65	334	35.44	340	35.41	335	32.36	333	36.07	337	35.51	441	35.97	471	36.69	164	7.50
Eastern Tien Shan	247	15.83	230	17.24	241	15.99	245	15.64	251	15.75	259	16.15	250	16.07	297	17.31	241	17.99	245	18.31	-2	2.48
Western Kunlun Shan	112	40.27	134	38.30	121	36.92	119	41.41	111	39.84	136	37.09	108	37.62	120	41.62	126	40.54	132	44.71	20	4.44
Eastern Kunlun Shan	180	11.92	193	12.83	265	20.62	237	16.46	232	15.74	246	15.90	244	16.72	290	18.49	255	17.95	248	16.90	68	4.98
Gangdise Mountains	852	99.15	908	96.46	848	104.24	886	96.88	961	102.48	895	95.75	954	96.28	991	94.64	1053	96.37	1116	99.83	264	0.68
Hengduan Shan	909	62.25	1077	64.70	967	65.38	867	61.29	927	62.75	1064	66.53	948	63.82	1023	65.45	942	62.93	961	63.35	52	1.10
Tibetan Interior Mountains	335	52.96	308	44.76	349	47.59	334	46.42	318	46.94	334	47.73	313	46.17	335	48.12	318	44.74	315	47.97	-20	-4.99
Eastern Tibetan Mountains	57	12.74	86	13.74	86	14.39	107	15.38	84	13.65	100	14.35	93	14.93	76	13.58	85	14.53	66	13.60	9	0.86
Tanggula Shan	468	39.45	318	40.99	344	44.28	327	43.24	347	42.17	363	43.40	376	44.97	461	45.53	372	45.94	363	44.97	-105	5.52
Qilian Shan	82	7.96	78	8.67	91	9.68	77	9.23	76	9.19	85	9.36	90	10.38	86	9.90	80	9.99	67	8.97	-15	1.01

Dzhungarsky Alatau	290	13.82	218	11.57	233	11.75	272	12.60	259	12.48	269	12.68	267	12.90	205	11.80	240	12.44	264	12.80	-26	-1.02
Nyainqentanglha	2401	275.33	2647	255.23	2614	272.95	2539	273.28	2594	266.60	2856	275.46	2646	277.70	2977	280.35	2660	292.24	2911	285.68	510	10.35

Table A2. Mountain-wide annual glacial lake area from 2008 to 2017 for proglacial lakes and unconnected lakes. Supraglacial and ice-marginal lakes have relative few coverage areas and are not listed details in the table. The unit of area is km^2. Acronyms are used to represent the name of mountain ranges to save space.

		EHK	WH	EH	CH	K	WP	EP	PA	N/WT	CT	ET	WK	EK	G	H	TIM	ETM	T	Q	DA	N	Total
Proglacial lakes	2008	46.71	34.58	91.45	133.06	10.99	34.07	4.56	4.13	17.77	14.52	9.12	5.15	1.58	24.01	11.40	3.59	0.76	9.73	3.74	11.19	133.29	605.40
	2009	37.98	30.03	92.97	130.09	10.60	34.27	5.30	2.79	16.39	14.07	8.11	5.69	1.72	19.97	12.58	2.97	1.54	10.12	4.71	9.04	122.95	573.90
	2010	41.68	29.73	95.11	137.49	10.55	35.72	5.02	3.57	18.09	16.70	9.18	5.73	5.66	22.27	13.55	4.72	1.68	11.52	5.25	9.24	136.04	618.48
	2011	42.77	32.88	100.58	143.53	9.70	34.82	4.62	4.04	20.27	15.64	8.66	4.73	3.69	20.76	12.14	3.80	1.74	10.63	4.71	10.18	136.69	626.59
	2012	37.32	31.30	99.54	143.02	9.00	34.87	5.12	3.91	18.58	16.76	8.83	4.69	3.66	21.78	12.83	3.73	1.74	10.60	4.66	10.02	133.29	615.25
	2013	41.67	30.58	100.97	141.25	9.22	34.57	4.71	4.08	18.28	16.72	9.20	4.95	3.51	20.56	13.55	3.99	1.67	10.60	4.74	10.07	136.58	621.47
	2014	43.88	36.12	98.12	145.22	9.96	35.06	4.71	4.15	20.33	17.37	8.88	5.04	4.40	21.88	12.91	4.38	1.75	10.68	5.01	10.21	136.95	637.00
	2015	40.49	33.88	108.45	151.17	10.68	37.63	5.03	4.02	18.98	16.26	9.53	5.08	5.01	22.66	13.53	4.44	1.62	11.30	4.67	9.43	138.12	652.00
	2016	40.52	34.20	103.31	150.59	10.92	36.13	5.02	3.83	18.12	17.63	8.31	4.67	4.03	23.29	13.03	4.08	1.74	10.75	4.73	9.99	147.83	652.73
	2017	43.53	35.52	103.75	152.52	11.33	37.60	5.16	4.03	20.07	18.56	8.52	5.36	4.21	23.33	13.59	4.41	1.69	10.49	4.71	10.15	143.58	662.07
Unconnected lakes	2008	26.82	41.86	70.98	43.82	6.33	41.24	0.47	6.70	18.38	14.56	6.69	35.10	10.35	74.83	50.60	49.08	11.98	29.5	4.23	2.57	140.42	686.53
	2009	23.35	39.37	69.41	45.41	4.81	43.00	0.39	6.98	20.79	15.36	9.07	32.56	11.07	76.41	52.10	41.80	12.21	30.87	3.97	2.49	131.63	673.04
	2010	27.08	40.69	70.78	46.96	5.64	47.24	0.80	7.37	18.78	15.87	6.79	31.20	14.81	81.97	51.83	42.87	12.71	32.76	4.44	2.49	136.12	699.22
	2011	25.00	37.62	71.65	47.79	4.44	45.79	0.31	7.75	21.42	19.76	6.96	36.65	12.78	76.13	49.16	42.63	13.64	32.59	4.52	2.34	136.05	694.98
	2012	22.41	36.80	71.23	47.75	3.74	44.13	0.50	7.69	17.44	18.48	6.82	35.15	12.04	80.71	49.92	43.22	11.92	31.53	4.54	2.37	132.53	680.92
	2013	26.10	44.59	74.81	50.85	4.09	44.52	0.59	7.67	18.68	15.55	6.83	32.09	12.34	75.20	52.98	43.75	12.69	32.77	4.62	2.51	138.07	701.31
	2014	27.27	45.69	75.92	48.65	4.31	45.85	0.43	7.52	25.69	18.54	7.08	32.55	12.29	74.40	50.92	41.74	13.19	34.26	5.38	2.60	140.03	714.33
	2015	25.76	43.60	81.11	50.30	5.76	46.08	0.47	7.76	19.29	19.15	7.66	36.54	13.44	71.98	51.92	43.66	11.96	34.20	5.23	2.30	141.40	719.59
	2016	25.14	44.34	74.76	50.30	6.06	45.39	0.49	7.70	19.24	17.77	9.57	35.81	13.91	73.09	49.91	40.66	12.79	35.15	5.26	2.35	143.38	713.11
	2017	27.19	51.42	74.96	51.59	6.54	37.95	0.46	8.06	20.81	17.52	9.66	39.34	12.68	76.53	49.76	43.57	11.91	34.45	4.27	2.56	141.22	722.44

Table A3. Area of different sizes of glacial lakes in 2017 for some regions with large area growth of rate. The unit of area is km².

Lake grid ID (Mountain ranges)	69 (N)	116 (CH)	274 (WH)	71 (N)	48 (H)	74 (N)	72 (N)	14 (EH)	13 (EH)	39 (CH)	15 (EH)
≤0.01km²	0.18	0.28	0.20	0.16	0.22	0.16	0.17	0.23	0.16	0.32	0.33
0.01km²-0.02km²	0.85	1.51	1.29	0.71	1.43	1.08	1.37	1.45	1.45	1.49	2.46
0.02km²-0.04km²	1.69	2.16	2.22	1.79	3.24	2.09	2.29	2.24	2.06	2.72	4.14
0.04km²-0.08km²	1.78	3.19	2.98	2.20	5.30	3.38	4.45	2.77	2.69	3.66	7.16
0.08km²-0.16km²	1.91	5.38	3.87	2.86	4.81	4.03	5.06	3.75	4.33	5.00	13.16
0.16km²-0.32km²	1.81	4.53	2.23	2.76	4.62	5.55	5.81	2.91	3.90	5.66	11.62
0.32km²-0.64km²	1.01	5.37	1.77	1.79	3.88	1.75	3.81	5.72	3.99	7.13	12.37
0.64km²-1.28km²	0.00	2.94	0.00	1.38	2.82	2.96	4.43	0.96	7.10	8.97	7.74
≥1.28km²	0.00	7.22	4.19	0.00	11.46	3.17	2.59	6.07	1.40	6.06	12.00
Total area (km²)	9.22	32.58	18.76	13.66	37.76	24.17	29.99	26.10	27.09	41.00	70.96
Total area (≤0.16km²)	6.41	12.52	10.57	7.72	14.99	10.74	13.35	10.45	10.69	13.18	27.24
% of Total area (≤0.16km²)	69.47	38.43	56.32	56.56	39.70	44.45	44.52	40.03	39.47	32.15	38.39
Annual area increase (km²a⁻¹)	0.23	0.28	0.28	0.29	0.32	0.41	0.42	0.49	0.70	0.74	0.94

515

Table A4. Summary of correlation coefficients (*R*) for key lake, topographic, geomorphic and climatological parameters, calculated within 1°×1° grid cells across HMA. Correlation coefficients are bold where *p* < 0.05; (*) indicates *p* < 0.01.

	Lake area (2008)	Lake area (2017)	Lake change (2008 – 2017)	Glacier (gl.) area ^	Debris-covered gl. area	Total gl. length	Mean gl. slope	Mean gl. elevation	Temperature change^^ 1979 – 2017	Precipitation change 1979 – 2017
Lake area (2008)	1.00									
Lake area (2017)	**0.99***	1.00								
Lake change (2008 – 2017)	**0.82***	**0.87***	1.00							
Glacier (gl.) area	**0.23***	**0.24***	**0.22***	1.00						
Debris-covered gl. area	**0.35***	**0.36***	**0.34***	**0.85***	1.00					
Total gl. length	**0.32***	**0.32***	**0.28***	**0.90***	**0.85***	1.00				
Mean gl. Slope	0.07	0.07	0.05	0.02	**0.18**	0.06	1.00			
Mean gl. Elevation	0.12	0.14	**0.17**	0.11	0.00	0.05	**-0.28***	1.00		
Temperature change 1979 – 2017	-0.09	-0.07	0.10	**-0.17**	**-0.25***	**-0.27***	0.00	0.07	1.00	
Precipitation change	-0.03	-0.01	0.09	-0.13	**-0.16**	**-0.15**	**-0.18***	**0.15**	**0.16**	1.00

520 ^ Glacier data is derived from the Randolph Glacier Inventory (RGI Consortium, 2017), except for debris cover (after Scherler et al. 2018). Climate data is for ERA Interim

^^ ERA-Interim near surface temperature and precipitation fields for the period 1979 – 2017 were obtained from the KNMI climate explorer (https://climexp.knmi.nl).

Table A5. Regional summary of key topographic, geomorphic and climatological parameters compared to pro- and supraglacial lake area in 2017. Correlation coefficients are bold where $p < 0.05$; (*) indicates $p < 0.01$.

Region	Total area (km^2)	Lake area (km^2)	Glacier (gl.) area (km^2)	Debris-covered gl. area (km^2)	Total gl. length (km)	Mean gl. slope (°)	Mean gl. elevation (m)	Temperature change 1979 – 2017 (°C/century)	Precipitation change 1979 – 2017
Central Himalaya	254886	155.7	8678	1567	10669	26	5542	2.77	-0.25
Central Tien Shan	105456	19.0	7270	842	7415	27	4181	1.35	0.05
Dzhungarsky Alatau	37542	10.3	521	18	978	24	3615	1.85	0.74
Eastern Himalaya	164785	104.7	2838	357	3614	24	5484	3.26	-0.84
Eastern Hindu Kush	95404	43.6	2938	609	5062	25	4856	0.08	-0.86
Eastern Kunlun Shan	256729	4.2	2995	45	3384	24	5389	3.60	0.06
Eastern Pamir	39605	5.2	2118	291	2364	27	5064	3.42	-0.50
Eastern Tibetan Mountains	333123	1.8	312	12	483	24	5345	3.55	0.73
Eastern Tien Shan	140900	8.7	2332	193	3977	28	3974	2.65	0.17
Gangdise Mountains	154884	23.2	1271	80	2570	24	5892	2.42	0.33
Hengduan Shan	372649	13.6	1281	212	2048	23	5278	2.24	-0.13
Karakoram	83644	11.7	21474	2013	16460	31	5399	2.37	-0.35
Northern/Western Tien Shan	187275	20.1	2262	223	4138	23	3943	3.22	-0.36

Nyainqentanglha	172746	144.6	7047	1011	8710	25	5282	2.37	-1.00
Pamir Alay	71845	4.0	1847	319	3441	25	4109	3.88	-0.27
Qilian Shan	201699	4.7	1598	30	2588	26	4847	4.07	0.51
Tanggula Shan	145064	10.6	1841	84	1893	21	5521	3.34	0.46
Tibetan Interior Mountains	526111	4.4	3815	59	4179	23	5927	2.64	0.31
Western Himalaya	189494	36.3	7986	1149	11974	24	5180	1.93	-1.24
Western Kunlun Shan	123388	5.3	8457	159	8108	26	5642	3.22	-0.55
Western Pamir	109239	37.9	8417	1118	11640	27	4844	1.61	0.08
Lake area: Correlation Coefficient (*R*)			0.21	**0.50**	0.36	0.01	0.23	-0.17	**-0.49**
Exl. Karakoram			**0.49**	**0.72***	**0.52**	0.10	0.25	-0.18	**-0.50**

Table A6. Statistical results of stratified random sampling.

Strata	Total pixel number	Total area (km²)	Sample number	Sample No. of non glacial lake	Sample No. of glacial lake	No. of ambiguous sample
C0W0	2,022,448,650	1,820,203.78	1300	1215	37	48
C0W1	925,449	832.90	700	307	362	31
C1W0	814,196	732.77	700	21	678	1
C1W1	1,611,668	1,450.50	1300	25	1260	15

References

Aggarwal, S., Rai, S. C., Thakur, P. K., and Emmer, A.: Inventory and recently increasing GLOF susceptibility of glacial lakes in Sikkim, Eastern Himalaya, Geomorphology, 295, 39-54, https://doi.org/10.1016/j.geomorph.2017.06.014,
545   2017.
Armstrong, W. H., and Anderson, R. S.: Ice-marginal lake hydrology and the seasonal dynamical evolution of Kennicott Glacier, Alaska, Journal of Glaciology, 1-15, 2020.
Ashraf, A., Naz, R., and Roohi, R.: Glacial lake outburst flood hazards in Hindukush, Karakoram and Himalayan Ranges of Pakistan: implications and risk analysis, Geomatics, Natural Hazards and Risk, 3, 113-132,
550   10.1080/19475705.2011.615344, 2012.
Baumann, F., Jinsheng, H. E., Karsten, S., Peter, K., and Thomas, S.: Pedogenesis, permafrost, and soil moisture as controlling factors for soil nitrogen and carbon contents across the Tibetan Plateau, Global Change Biology, 15, 3001–3017, 2009.
Bhardwaj, A., Singh, M. K., Joshi, P. K., Snehmani, Singh, S., Sam, L., R.D.Gupta, and RajeshKumar: A lake detection
555   algorithm (LDA) using Landsat 8 data: A comparative approach in glacial environment, International Journal of Applied Earth Observations & Geoinformation, 38, 150-163, 2015.
Bo, Zhao, Yunsheng, Wang, Yonghong, Luo, Ruifeng, Liang, Jia, and Li: Large landslides at the northeastern margin of the Bayan Har Block, Tibetan Plateau, China, Royal Society Open Science, 6, 180844, 2019.
Bolch, T., Kulkarni, A., Kääb, A., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., and Scheel, M.:
560   The state and fate of Himalayan glaciers, Science, 336, 310, 2012a.
Bolch, T., Kulkarni, A., Kaab, A., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., and Scheel, M.: The State and Fate of Himalayan Glaciers, Science, 336, 310-314, 2012b.
Brun, F., Berthier, E., Wagnon, P., Kääb, A., and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nature Geoscience, 10, 668-673, 2017a.

565 Brun, F., Berthier, E., Wagnon, P., Kaab, A., and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nature Geoscience, 10, 668-673, 2017b.

Capps, D. M., Wiles, G. C., Clague, J. J., and Luckman, B. H.: Tree-ring dating of the nineteenth-century advance of Brady Glacier and the evolution of two ice-marginal lakes, Alaska, The Holocene, 21, 641-649, 2011.

Chen, F., Zhang, M., Tian, B., and Li, Z.: Extraction of Glacial Lake Outlines in Tibet Plateau Using Landsat 8 Imagery
570 and Google Earth Engine, IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 10, 4002-4009, 2017.

Chen, J., Zhu, X., E.Vogelmann, J., Gao, F., and Jin, S.: A simple and effective method for filling gaps in Landsat ETM+ SLC-off images, Remote Sensing of Environment, 115, 1053-1064, 2011.

Dan, H. S., and CLAGUE, J. J.: The sedimentology and geomorphology of rock avalanche deposits on glaciers,
575 Sedimentology, 58, 1762-1783, 2011.

Gardelle, J., Berthier, E., and Arnaud, Y.: Slight mass gain of Karakoram glaciers in the early twenty-first century, Nature Geoscience, 5, 322-325, 10.1038/ngeo1450, 2012.

Gardelle, J., Berthier, E., Arnaud, Y., and Kääb, A.: Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999–2011, The Cryosphere, 7, 1263-1286, 10.5194/tc-7-1263-2013, 2013.

580 Gardner, A. S., Moholdt, G., Cogley, J. G., Wouters, B., Arendt, A. A., Wahr, J., Berthier, E., Hock, R., Pfeffer, W. T., and Kaser, G.: A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009, Science, 340, 852, 2013.

Glazirin, G. E.: A century of investigations on outbursts of the ice-dammed lake merzbacher (Central Tien Shan), Austrian Journal of Earth Sciences, 103, 171-179, 2010.

Hanqiu, X.: Modification of normalised difference water index (NDWI) to enhance open water features in remotely
585 sensed imagery, International Journal of Remote Sensing, 27, 3025-3033, 2006.

Haritashya, U. K., Kargel, J. S., Shugar, D. H., Leonard, G. J., Strattman, K., Watson, C. S., Shean, D., Harrison, S., Mandli, K. T., and Regmi, D.: Evolution and Controls of Large Glacial Lakes in the Nepal Himalaya, Remote Sensing, 10, 798, 2018.

Harrison, S., Kargel, J. S., Huggel, C., Reynolds, J., Dan, H. S., Betts, R. A., Emmer, A., Glasser, N., Haritashya, U. K.,
590 and Klimeš, J.: Climate change and the global pattern of moraine-dammed glacial lake outburst floods, Cryosphere, 12, 1-28, 2018.

Immerzeel, W. W., and Bierkens, M. F. P.: Seasonal prediction of monsoon rainfall in three Asian river basins: the importance of snow cover on the Tibetan Plateau, International Journal of Climatology, 30, 1835–1842, 2010.

Kääb, A., Treichler, D., Nuth, C., and Berthier, E.: Brief Communication: Contending estimates of 2003-2008 glacier
595 mass balance over the Pamir-Karakoram-Himalaya, Cryosphere, 9, 557-564, 2015.

King, O., Bhattacharya, A., Bhambri, R., and Bolch, T.: Glacial lakes exacerbate Himalayan glacier mass loss, Scientific Reports, 9, 2019.

Krumwiede, B. S., Kamp, U., Leonard, G. J., Kargel, J. S., and Walther, M.: Recent Glacier Changes in the Mongolian Altai Mountains: Case Studies from Munkh Khairkhan and Tavan Bogd, Global Land Ice Measurements from Space
600 2014, Springer, Berlin Heidelberg, 2014.

Kumar, S. P.: Estimates of the Regression Coefficient Based on Kendall"s Tau, Publications of the American Statistical Association, 63, 1379-1389, 1968.

Li, D., Shangguan, D., and Naveed, A. M.: Glacial Lake Inventory Derived from Landsat 8 OLI in 2016–2018 in China–Pakistan Economic Corridor, International Journal of Geo-Information, 9, 294, 2020.

605 Li, J., and Sheng, Y.: An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: A case study in the Himalayas, International Journal of Remote Sensing, 33, 5194-5213, 2012.

Liu, J. J., Cheng, Z.-L., and Su, P.-C.: The relationship between air temperature fluctuation and Glacial Lake Outburst Floods in Tibet, China, Quaternary International, 321, 78-87, 2014.

Luo, W., Zhang, G., Chen, W., and Xu, F.: Response of glacial lakes to glacier and climate changes in the western
610 Nyainqentanglha range, Science of The Total Environment, 735, 139607, https://doi.org/10.1016/j.scitotenv.2020.139607, 2020.

Miles, E. S., Steiner, J., Willis, I., Buri, P., Immerzeel, W. W., Chesnokova, A., and Pellicciotti, F.: Pond Dynamics and Supraglacial-Englacial Connectivity on Debris-Covered Lirung Glacier, Nepal, Frontiers in Earth Science, 5, 10.3389/feart.2017.00069, 2017a.

615 Miles, E. S., Willis, I. C., Arnold, N. S., Steiner, J., and Pellicciotti, F.: Spatial, seasonal and interannual variability of supraglacial ponds in the Langtang Valley of Nepal, 1999–2013, Journal of Glaciology, 63, 88-105, 10.1017/jog.2016.120, 2017b.

Morfitt, R., Barsi, J., Levy, R., Markham, B., Micijevic, E., Ong, L., Scaramuzza, P., and Vanderwerff, K.: Landsat-8 Operational Land Imager (OLI) Radiometric Performance On-Orbit, Remote Sensing, 7, 2208-2237, 2015.

620 Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnowa, S., and Ip, A.: Water observations from space: Mapping surface water from 25years of Landsat imagery across Australia, Remote Sensing of Environment, 174, 341-352, 2016.

Neckel, N., Kropáček, J., Bolch, T., and Hochschild, V.: Glacier mass changes on the Tibetan Plateau 2003-2009 derived from ICESat laser altimetry measurements, Environmental Research Letters, 9, 468-475, 2014.

625 Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y., and Song, C.: A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015, Remote Sensing of Environment, 189, 1-13, 2017.

Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418-422, 10.1038/nature20584, 2016a.

Pekel, J. F., Cottam, A., Gorelick, N., and Belward, A. S.: High-resolution mapping of global surface water and its long-term changes, Nature, 540, 418-422, 2016b.

630 Qiao, L., Wanqin, G., Yong, N., Shiyin, L., and Junli, X.: Recent glacier and glacial lake changes and their interactions in the Bugyai Kangri, Southeast Tibet, Annals of Glaciology, 57, 61-69, 2016.

Quincey, D. J., Richardson, S. D., Luckman, A., Lucas, R. M., Reynolds, J. M., Hambrey, M. J., and Glasser, N. F.: Early recognition of glacial lake hazards in the Himalaya using remote sensing datasets, Global & Planetary Change, 56, 137-152, 2007.

635 RGI-Consortium: Randolph Glacier Inventory - A Dataset of Global Glacier Outlines: Version 6.0: Technical Report, Global Land Ice Measurements from Space, Digital Media, Colorado, USA, Digital Media, 2017.

Rounce, D., Watson, C. S., and Mckinney, D.: Identification of Hazard and Risk for Glacial Lakes in the Nepal Himalaya Using Satellite Imagery from 2000–2015, Remote Sensing, 9, 654, 2017.

640 Salerno, F., Thakuri, S., D'Agata, C., Smiraglia, C., Manfredi, E. C., Viviano, G., and Tartari, G.: Glacial lake distribution in the Mount Everest region: Uncertainty of measurement and conditions of formation, Global & Planetary Change, 92-93, 30-39, 2012.

Scherler, D., Bookhagen, B., and Strecker, M. R.: Spatially variable response of Himalayan glaciers to climate change affected by debris cover, Nature Geoence, 4, 156-159, 2011.

645 Scherler, D., Wulf, H., and Gorelick, N.: Global Assessment of Supraglacial Debris Cover Extents, Geophysical Research Letters, 45, 11798-11805, 2018.

Schiemann, R., Lüthi, D., and Schär, C.: Seasonality and interannual variability of the westerly jet in the Tibetan Plateau region, Journal of Climate, 22, 2940-2957, 2009.

Shaun D, R., and John M, R.: An overview of glacial hazards in the Himalayas, Quaternary International, 65, 31-47, 2000.

650 Shen, Y., Wang, G., and Ding, Y.: Changes in Merzbacher Lake of Inylchek Glacier and glacial flash floods in Aksu River Basin, Tianshan during the period of 1903-2009, Journal of Glaciology & Geocryology, 31, 993-1002, 2009.

Shugar, D. H., Burr, A., Haritashya, U. K., Kargel, J. S., Watson, C. S., Kennedy, M. C., Bevington, A. R., Betts, R. A., Harrison, S., and Strattman, K.: Rapid worldwide growth of glacial lakes since 1990, Nature Climate Change, 10.1038/s41558-020-0855-4, 2020.

655 Song, C., Sheng, Y., Ke, L., Nie, Y., and Wang, J.: Glacial lake evolution in the southeastern Tibetan Plateau and the cause of rapid expansion of proglacial lakes linked to glacial-hydrogeomorphic processes, Journal of Hydrology, 540, 504-514, 2016.

Song, C., Sheng, Y., Wang, J., Ke, L., Madson, A., and Nie, Y.: Heterogeneous glacial lake changes and links of lake
expansions to the rapid thinning of adjacent glacier termini in the Himalayas, Geomorphology, 280, 30-38, 2017a.

660   Song, X. P., Potapov, P. V., Krylov, A., King, L. A., Bella, C. M. D., Hudson, A., Khan, A., Adusei, B., Stehman, S. V.,
and Hansen, M. C.: National-scale soybean mapping and area estimation in the United States using medium
resolution satellite imagery and field survey, Remote Sensing of Environment, 190, 383-395, 2017b.

Song, X. P., C, H. M., V, S. S., V, P. P., Alexandra, T., F, V. E., and R, T. J.: Global land change from 1982 to 2016,
Nature, 560, 639-643, 2018.

665   Stehman, and SV: Impact of sample size allocation when using stratified random sampling to estimate accuracy and area
of land-cover change, Remote Sens Lett, 2012.

Sudan, B. M., Pradeep, K. M., Wu, L., Gao, X., Finu, S., Rajendra, B. S., Narendra, R. K., Samjwal, R. B., Sharad, J.,
Sonika, S., and Prasant, B.: The status of glacial lakes in the Hindu Kush Himalaya-ICIMOD Research Report 2018/1,
ICIMOD Research Report 2018/1, 2018.

670   Thakuri, S., Salerno, F., Bolch, T., Guyennon, N., and Tartari, G.: Factors controlling the accelerated expansion of Imja
Lake, Mount Everest region, Nepal, Annals of Glaciology, 57, 245-257, 2016.

Thompson, S. S., Benn, D. I., Dennis, K., and Luckman, A.: A rapidly growing moraine-dammed glacial lake on
Ngozumpa Glacier, Nepal, Geomorphology 145-146, 1-11, 2012.

Truffer, M., and Motyka, R. J.: Where glaciers meet water: Subaqueous melt and its relevance to glaciers in various
675   settings, Reviews of Geophysics, 54, 220-239, 2016.

Ukita, J., Narama, C., Tadono, T., Yamanokuchi, T., Tomiyama, N., Kawamoto, S., Abe, C., Uda, T., Yabuki, H., and
Fujita, K.: Glacial lake inventory of Bhutan using ALOS data: methods and preliminary results, Annals of Glaciology,
52, 65-71, 2011.

Wang, N., Chen, F., Yu, B., and Qin, Y.: Segmentation of large-scale remotely sensed images on a Spark platform: A
680   strategy for handling massive image tiles with the MapReduce model, ISPRS Journal of Photogrammetry and
Remote Sensing, 162, 137-147, 2020a.

Wang, X., Ding, Y., Liu, S., Jiang, L., Wu, K., Jiang, Z., and Guo, W.: Changes of glacial lakes and implications in Tian
Shan, central Asia, based on remote sensing data from 1990 to 2010, Environmental Research Letters, 8, 575-591,
2013a.

685   Wang, X., Ding, Y. J., Liu, S. Y., Jiang, L. H., Wu, K. P., Jiang, Z. L., and Guo, W. Q.: Changes of glacial lakes and
implications in Tian Shan, central Asia, based on remote sensing data from 1990 to 2010, Environmental Research
Letters, 8, 575-591, 2013b.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Liu, S., Zhang, Y., Jiang, Z., and Tang, Z.: Glacial lake inventory
of High Mountain Asia (1990-2018) derived from Landsat images, Earth System Science Data, 1-23, doi:
690   https://doi.org/10.5194/essd-2019-212, 2020b.

Yao, T., Thompson, L., Yang, W., Yu, W., Yang, G., Guo, X., Yang, X., Duan, K., Zhao, H., and Xu, B.: Different glacier
status with atmospheric circulations in Tibetan Plateau and surroundings, Nature Climate Change, 2, 663-667, 2012.

Yoon, Y., Kumar, S. V., Forman, B. A., Zaitchik, B., and, F., K., Yonghwan, Qian, Yun, Rupper, Summer, Maggioni,
Viviana, Houser, Paul, Kirschbaum, Dalia, Richey, Alexandra, Arendt, Anthony, Mocko, David, Jacob, Jossy,
695   Bhanja, Soumendra, and Mukherjee, Abhijit.: Evaluating the Uncertainty of Terrestrial Water Budget Components
Over High Mountain Asia, Frontiers in Earth Science, 7, 1-21, 2019.

Zhang, G., Yao, T., Xie, H., Wang, W., and Yang, W.: An inventory of glacial lakes in the Third Pole region and their
changes in response to global warming, Global & Planetary Change, 131, 148-157, 2015.

Zhang, G., Yao, T., Chen, W., Zheng, G., Shum, C., Yang, K., Piao, S., Sheng, Y., Yi, S., and Li, J.: Regional differences
700   of lake evolution across China during 1960s–2015 and its natural and anthropogenic causes, Remote Sensing of
Environment, 221, 386-404, 2019.

Zhang, M., Chen, F., Tian, B., Liang, D., and Yang, A.: High-frequency glacial lake mapping using time series of sentinel-
1A/1B SAR imagery: An assessment for the southeastern tibetan plateau, International Journal of Environmental
Research and Public Health, 17, 1072, 2020.

705   Zhao, H., Chen, F., and Zhang, M.: A Systematic Extraction Approach for Mapping Glacial Lakes in High Mountain
        Regions of Asia, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11, 2788-
        2799, 10.1109/JSTARS.2018.2846551, 2018.
      Zhao, L., Ding, R., and Moore, J. C.: Glacier volume and area change by 2050 in high mountain Asia, Global and
        Planetary Change, 122, 197-207, 2014.
710   Zhu, Z., and Woodcock, C. E.: Object-based cloud and cloud shadow detection in Landsat imagery, Remote Sensing of
        Environment, 118, 83-94, 2012.
      Chen, F., Zhang, M., Guo, H., Allen, S., Kargel, J., Haritashya, U., Watson, S.: Annual 30-meter Dataset for Glacial
        Lakes   in   High   Mountain   Asia   from   2008   to   2017   (Hi-MAG)   [Dataset],   Zenodo,
        https://doi.org/10.5281/zenodo.4059181.
715