Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-57-RC1, 2020

Interactive comment on "Annual 30-meter Dataset for Glacial Lakes in High Mountain Asia from 2008 to 2017" *by* Fang Chen et al.

Anonymous Referee #1

Received and published: 18 June 2020

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

Printer-friendly version

Discussion paper

could help to explain the size-distribution and growth of glacier lakes, though regional variances make a more rigorous explanation elusive.

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. Yet, visually comparing the inventories of 2016 and 2017, I noticed that the 2017 data set misses >2 km^2 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, 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.

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.

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.

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.

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.

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.

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.

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’?

L19: ‘incomplete’: Can this study claim completeness?

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.

L23-24: ‘Proglacial lake dominated areas’; ‘unconnected lake dominated areas’: unclear what and where these areas are.

L23: ‘significant’: how do the authors measure significance here?

L26: ‘a main contributor’: to? Increase in lake area?

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)

L29: ‘is’ instead of ‘are’?

L30: New sentence instead of comma?

L34: ‘significant’: how do the authors measure significance here?

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

L36-37: Please add a reference.

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

Printer-friendly version

Discussion paper

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...

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

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.

L55-56: 'Alaska': reference missing

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

L58: What are 'systematic errors'?

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.

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

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

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

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.

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?

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

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.

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.

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

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).

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

L97: 'criteria' instead of 'criterions'?

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.

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

L104-106: Repetition, consider deleting.

L109: 'and SLC-off gaps'

L110: ‘a more computationally efficient way’: than?

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

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?

L112: ‘value less than 0.25’: Unclear what this value means in physical quantities.

L113: The paper from Quincey is from 2007.

L115: Maybe ‘acquired. Some’ ?

L116: ‘remain’ instead of ‘remains’, and ‘was’ instead of ‘were’?

L118: ‘the relative stability’: unclear, please reformulate.

L118: ‘0.0081 km^2’: the reason for choosing this minimum mapping unit is not fully clear.

L127: ‘polygon include’ instead of ‘vector includes’?

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?

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.

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?

L137: delete ‘on’? ‘chooses the median’?

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.

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

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

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

L160: remove 'areal'.

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?

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?

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

L179: ‘A linear least-squares’: Why did the authors refrain from using the Theil-Sen-regression here?

L179 & L181: ‘18 km2 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.

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?

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.

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?

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

L194-195: Which ‘non-parametric trend analysis’ did the authors use?

L200: ‘large patches’: unclear formulation, please elaborate.

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?

L210-211: What is happening in regions with shrinking lakes? Do some of these fall dry or burst out?

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

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?

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?

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

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 measure of uncertainty for these annual changes, nor is it clear whether these changes occur linearly in time or that the rate of change could be affected by outliers. Total changes in glacier lake area between 2008 and 2017 could probably make more sense here .

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

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

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.

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.

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.

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.

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

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

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

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

L266: what does this correlation tell us?

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

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 favor supraglacial and proglacial lake formation'?

L272: 'Some adjacent regions': which ones exactly? 'comparable', 'large', 'substantial':

please be more specific.

L273: ‘longer-term climate trends’: on which scales? Years, decades, centuries?

L274: ‘rapid warming‘: warming seems to be present across the entire study region except the Karakoram. ‘decreased precipitation‘: precipitation 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.

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

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.

L282: Dan and CLAGUE, 2011: wrong citation.

L291: ‘datasets’.

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.

L296: ‘a much larger lake area’: how much larger?

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

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.

L304: ‘significant number‘: how do the authors measure significance here?

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?

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

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?

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?

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.

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.

L366: Please avoid using rainbow color scales.

L406: Why does this need to be acknowledged given that these researchers are co-authors in this study?

References

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, Nat. Geosci., 10(9), 668–673, doi:10.1038/ngeo2999, 2017.

Printer-friendly version

Discussion paper

Khadka, N., Zhang, G. and Thakuri, S.: Glacial Lakes in the Nepal Himalaya: Inventory and Decadal Dynamics (1977–2017), Remote Sens., 10(12), 1913, doi:10.3390/rs10121913, 2018. Li, J., Warner, T. A., Wang, Y., Bai, J. and Bao, A.: Mapping glacial lakes partially obscured by mountain shadows for time series and regional mapping applications, Int. J. Remote Sens., 40(2), 615–641, doi:10.1080/01431161.2018.1516314, 2019.

Maharjan, S. B., Mool, P., Lizong, W., Xiao, G., Shrestha, F., Shrestha, R., Khanal, N., Bajracharya, S., Joshi, S. and Shai, S.: The status of glacial lakes in the Hindu Kush Himalaya-ICIMOD Research Report 2018/1 (2018)., ICIMOD Res. Rep., (2018/1), 2018.

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, Front. Earth Sci., 5, 69, doi: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, J. Glaciol., 63(237), 88–105, doi:10.1017/jog.2016.120, 2017b.

Nagai, H., Ukita, J., Narama, C., Fujita, K., Sakai, A., Tadono, T., Yamanokuchi, T. and Tomiyama, N.: Evaluating the Scale and Potential of GLOF in the Bhutan Himalayas Using a Satellite-Based Integral Glacier–Glacial Lake Inventory, Geosciences, 7(3), 77, doi:10.3390/geosciences7030077, 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(7633), 418–422, doi:10.1038/nature20584, 2016.

Round, V., Leinss, S., Huss, M., Haemmig, C. and Hajnsek, I.: Surge dynamics and lake outbursts of Kyagar Glacier, Karakoram, The Cryosphere, 11(2), 723–739, doi:10.5194/tc-11-723-2017, 2017.

Shukla, A., Garg, P. K. and Srivastava, S.: Evolution of Glacial and High-Altitude Lakes in the Sikkim, Eastern Himalaya Over the Past Four Decades (1975–2017), Front. Environ. Sci., 6, 81, doi:10.3389/fenvs.2018.00081, 2018.

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, J. Hydrol., 540, 504–514, doi:10.1016/j.jhydrol.2016.06.054, 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, doi:10.1016/j.geomorph.2016.12.002, 2017.

Veh, G., Korup, O., Roessner, S. and Walz, A.: Detecting Himalayan glacial lake outburst floods from Landsat time series, Remote Sens. Environ., 207, 84–97, doi:10.1016/j.rse.2017.12.025, 2018.

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 Syst. Sci. Data, 1–23, doi:https://doi.org/10.5194/essd-2019-212, 2020.

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, Glob. Planet. Change, 131, 148–157, doi:10.1016/j.gloplacha.2015.05.013, 2015.

Zhu, Z. and Woodcock, C. E.: Object-based cloud and cloud shadow detection in Landsat imagery, Remote Sens. Environ., 118, 83–94, doi:10.1016/j.rse.2011.10.028, 2012.

Please also note the supplement to this comment:

https://essd.copernicus.org/preprints/essd-2020-57/essd-2020-57-RC1-supplement.pdf

---

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-57, 2020.