Glacial Lake Observatory (GLO): Annual dataset of glacial lakes in Nepal and transboundary catchments (2017–2024)

Lauren D. Rawlins[1], C. Scott Watson[1*], Rakesh Bhambri[2], Nitesh Khadka[3,4], Mohan B. Chand[5]

Response to Reviewer 2

This study presents a new annual glacial lake dataset for Nepal and adjacent transboundary catchments covering the period 2017–2024. The authors use multi-sensor satellite imagery from Sentinel-1 and Sentinel-2 to map lake extent and quantify changes in lake area over time. Image processing was conducted in Google Earth Engine, and lake detection was performed using a DeepLabV3 deep learning model trained on previously published manual inventories. The dataset is validated against manually digitized lakes and existing inventories, showing generally strong performance, particularly for larger lakes.

I really enjoyed reading this paper. I think it is a strong contribution and will be very useful for the community. The manuscript is well written, clearly structured, and the methods are explained clearly.

That said, I believe there are a few aspects that should be clarified so that users of the dataset can better understand the methodological limits and avoid potential misinterpretations.

We thank the reviewer for their overall positive comments and that they enjoyed reading paper.

We appreciate there are some aspects to be clarified, including methodological and dataset limitations, which align with reviewer 1, and we have now included a new limitations section (**Sect. 5.3.2 Recommended use and detection limits**), which is provided on page 2.

We have further expanded on specific points below.

Specific comments:

1 – it would be helpful to be more explicit about detection limits and methodological biases. While the nominal threshold is 0.001 km^2, the validation shows that small lakes have much higher relative uncertainty. It would strengthen the paper to **clearly state what lake sizes can be reliably monitored for change analysis**, and where **uncertainty becomes too large for robust interpretation**. Similarly, lake surface ice and partial freezing are mentioned as sources of error, but it would help to specify when and where this is most problematic (e.g., elevation bands, specific basins, certain months). Without this, users might overinterpret trends in small or high-elevation lakes. A short paragraph summarizing recommended use and limitations would be very useful.

Thank you for your comment. We agree that uncertainty does increase for lakes approaching this threshold of 0.001 km^2. To further expand on this uncertainty and reliability of lake sizes for change analysis (including for both Sentinel-1 and -2), we have added some further robustness analysis (lake area and elevation) and included associated methods in Sect. 3.4.2, its results in Sect.4.1.2 and further discussion in Sect. 5.3.2 (see added text below). This is an expansion of some limitations that were originally mentioned in Sect. 5.1, which have now been collated into this new section.

In short, the results of this robustness analysis show a size dependence of 0.028 km^2 and 0.084 km^2 for Sentinel-2 and Sentinel-1 respectively is required to achieve robust classification performance (F1 $\geq$ 0.85). Low F1 scoring lakes were associated with smaller lakes at slightly higher elevations, as one would expect. In terms of controls on F1 score, size was shown to have a strong positive and significant relationship, whilst elevation had a weak non-significant negative effect.

Whilst we do acknowledge some limitations in Sect. 5.1, we have now expanded this into a new section **Sect. 5.3.2. Recommended use and detection limits**. See full text below.

In terms of whether specific basins posed problems for lake delineation, no significant effect of basin on F1 score was detected, and the sensor-specific difference in F1 did not vary significantly among basins, so this was not explored further.

Sect 3.4.2:

'Further validation analysis was also undertaken to assess classification robustness as a function of lake size and elevation for both Sentinel-1 and Sentinel-2 data. For size-dependent performance, F1 scores were examined across lake area and used to identify the approximate minimum lake size at which robust classification ($F1 \geq 0.85$) was achieved. For the influence of elevation on classification performance, median elevation was extracted for each lake polygon (-10 m buffer) using the AW3D30 DSM (Tadono et al., 2014) and a multiple linear regression performed to assess the effects of lake area and median elevation on F1 score for the sensor-comparable years of 2020.'

Sect 4.1.2:

'This size dependence of lakes is also used to approximate the threshold at which robust classification performance of the model is reached. For Sentinel-2, a robust performance score ($F1 \geq 0.85$) was generally achieved for lakes of above 0.028 km^2 across the study period, whilst for complimentary data provided by Sentinel-1, a larger size threshold of 0.084 km^2 was required to reach the same classification performance ($F1 \geq 0.85$), indicating lower robustness for smaller lakes in the SAR-derived dataset. Although the overall elevation distributions of both datasets were similar (median elevations ~5,020–5,060 m), lower-F1 lakes did tend to occur at slightly higher elevations and were generally smaller in area.

However, the multiple linear regression showed that lake area was the dominant control on F1 score for both Sentinel-1 and Sentinel-2. For Sentinel-2, lake area alone explained 42% of F1 variability ($R^2 = 0.42$; $\beta = 0.19$; $p < 0.001$), with a strong positive significant relationship whereby a 10-fold increase in lake area was associated with an increase in F1 score by ~0.19. The addition of elevation provided little improvement to the model although it did have a weak negative but significant influence ($\beta = -2.84 \times 10^{-5}$; $p = 0.02$). Sentinel-1 showed a similar positive relationship between F1 and lake area ($R^2 = 0.24$; $\beta = 0.19$; $p < 0.001$), however lake area explained less of the variance in F1 when compared to optical, indicating lower overall predictability of SAR classifications. For elevation, this had no significant independent effect on Sentinel-1 performance ($\beta = 2.84 \times 10^{-5}$; $p = 0.33$). Overall, Sentinel-1 was associated with lower and more variable F1 scores than Sentinel-2, consistent with the lower robustness of the SAR-derived dataset, particularly for smaller and more geometrically complex lakes.'

Sect 5.3.2:

5.3.2 Recommended use and detection limits

The dataset produced here provides a consistent and reproducible inventory of glacial lakes across the Nepal-transboundary region across the eight-year study period, making it well suited to regional-scale assessments of lake distribution, presence/absence, and broader interannual change. As with any regional-scale inventory, data confidence varies according to lake size and sensor type. Although a minimum mapping threshold of 0.001 km^2 was applied in this study, classification performance remains largely size dependent. For Sentinel-2, robust performance ($F1 \geq 0.85$) was generally achieved for lakes of above 0.028 km^2, whilst for complimentary data provided by Sentinel-1, a larger size threshold of 0.084 km^2 was required to reach the same classification performance ($F1 \geq 0.85$). A multiple linear regression model further showed that lake area was a dominant control on F1 score for both sensors, whilst elevation exerted a very small secondary influence. For Sentinel-2, the effect was weakly negative and statistically significant ($p = 0.02$), but of negligible magnitude, whereas for Sentinel-1 this was not significant. Together, these results indicate that reduced performance in small, often high-elevation lakes is driven primarily by lake size and the difficulty of complete delineation, rather than elevation itself acting as a strong independent control.

That said, small, irregular and high-elevation lakes do remain particularly challenging due to increased mixed pixel boundaries of water, snow and ice, combined with steep terrain and seasonal partial freezing (Kumar and Vijay, 2026; Qayyum et al., 2020; Watson et al., 2018). This likely explains the slightly lower agreement observed when lake comparisons included the Kumar et al. (2025) inventory, which contained a greater proportion of smaller lakes near the minimum mapping threshold (0.001 km^2). Additionally, some smaller lakes are highly ephemeral (e.g., supraglacial lakes) and may partially drain or disappear during the compositing period. In this case, while it would appear as a false negative when compared to Kumar's dataset, it is in fact a true negative, as the lakes were absent at the time of observation. In this context, higher-resolution imagery such as PlanetScope may support improved delineation for small or irregular lakes by resolving these mixed-pixel edge effects (Xu et al., 2024). However, this capability relies on quota-limited free research-programme data access, or paid access to the commercial dataset (https://www.planet.com/industries/education-and-research/) and may be most effectively applied to targeted training data or validation.

In terms of sensor specifics, the two sensors provide strengths for glacial lake monitoring. Sentinel-2 is shown to be better suited to the accurate delineation of individual lake boundaries, including smaller lakes, and should therefore be considered the primary dataset for more detailed and complete lake mapping. Sentinel-1 provides a valuable complimentary perspective by offering more consistent, all-weather observations, assisting with supplementing optical data gaps (i.e., the capture of lakes during monsoon or cloudy periods) and improve visibility of snow-or-ice covered lakes (Fig. S4; Fig. S5). It is noted, however, that these Sentinel-1 delineations are more prone to fragmented or irregular lake outlines, particularly for smaller lakes, likely reflecting known SAR challenges including speckle noise, look direction, incidence angle and backscatter variability over water and rugged terrain (Khan et al., 2025; Miles et al., 2017). At the same time, the Sentinel-1 dataset provided more temporally consistent lake counts, which is a likely reflection of a combination of factors: (i) the narrower acquisition compositing window meaning lakes were more likely to be captured at a similar time each year; (ii) a greater omission of smaller lakes, reducing year-to-year variability in the seasonal changes of smaller lakes, that are then intermittently captured by Sentinel-2; (iii) the use of median SAR composites which are more likely to exclude small, highly dynamic lakes (Kumar and Vijay, 2026). As such, Sentinel-1 should be interpreted as reflecting a more conservative subset of lake detections, rather than completeness. Finally, the annual Sentinel-1 and Sentinel-2 mosaics were generated from different acquisition windows to balance the observational capabilities and data availability of each sensor, so some minor differences between the two datasets may reflect differences in temporal sampling. Additionally, the Sentinel-1 model was trained using lake outlines derived from optical imagery (Kumar et al., 2025; Zhang et al., 2024) which may introduce some minor boundary mismatch between training labels and the SAR signal. Given the current lack of SAR-based training data, this approach is a reasonable compromise for Sentinel-1 model development and highlights an important avenue for future model refinement using SAR-based training labels. Beyond the overall strong performance shown of both sensors in this study, and the known limitations, ongoing work will set to further refine the classification models, with additional training and inference planned for the wider Himalayan region.

2- the different compositing windows for Sentinel-1 (July–August) and Sentinel-2 (May–November) introduce a potential sampling bias. Even if justified by cloud cover and image availability, the two sensors are not sampling the same seasonal window, which could influence median lake extent and comparability between sensors. A short discussion of how much this might affect inter-sensor comparisons would increase confidence in the results. Thank you for your comment. We acknowledge that the annual mosaics for Sentinel-1 and -2 were generated for different time windows and therefore some lake extent variability may occur. These windows were selected to optimise data availability under sensor-specific constraints: Sentinel-1 allows consistent post-monsoon acquisition irrespective of cloud cover, whereas Sentinel-2 requires a broader window to obtain sufficient cloud-free observations in this monsoon-affected region.

Sentinel-1 and -2 datasets were processed independently and are intended to be complementary (rather than comparable). Here, Sentinel-1 data contributes temporally (i.e., during cloudy conditions) when optical observations are limited, whilst Sentinel-2 provides greater spatial lake boundary precision (e.g., smaller lake delineation) during cloud-free conditions.

Some additional text has been added to Sect. 3.1.1, Sect 3.1.2 and Sect. 5.3.2 (see full Sect 5.3.2. above).

We have now added in some additional justifications to the following:
Sect. 3.1.1:
'A smaller date range was used compared to Sentinel-2 processing (Sect. 3.1.2), as the Sentinel-1 backscatter is not affected by cloud cover, allowing for more images to be available for compositing during the post-Monsoon period when lakes are typically ice-free and near their seasonal maximum extent. Whilst some lake extent variability may occur due to these time window differences per sensor, Sentinel-1 data is intended to be complimentary by contributing temporally when optical observations are limited.'

And Sect. 3.1.2:
'All available Sentinel-2 images (L1C) were used within the date range of May to November for years 2017 to 2024. This broader date range was used to coincide with the Monsoon season and late-ablation period, when the surface of glacial lakes is typically unfrozen, while also maximising the availability of cloud-free observations and reducing spatial data gaps in this persistently cloudy region.'

3- Sentinel-1 in mountainous terrain is affected by layover, foreshortening, and radar shadow, which can strongly influence backscatter and geometric representation in steep valleys. **It would be helpful to clarify whether SAR visibility constraints were explicitly accounted for, for example by masking radar shadow/layover areas or using terrain-flattened products.** In landslide studies, we have seen that training SAR-based models using inventories digitised from optical imagery can degrade performance if SAR visibility is not considered. There are approaches to compute SAR visibility masks in GEE (e.g., https://doi.org/10.3390/rs12111867). This might not be essential here, but **acknowledging this limitation** and potentially considering it in future refinements could strengthen transparency. Related to this, since the training data are derived from optical inventories, it would be useful to briefly comment on **potential cross-sensor discrepancies in geometry or boundary definition** and whether these could influence Sentinel-1 training quality.

Masking of radar shadow and layover areas was performed as part of the processing and we have now added this text. Example areas of masking are visible in (**e.g. Fig. S4**).

Under Sect. 3.1.1 we have modified the text to read:

'Sentinel-1 Ground Range Detected (GRD) images (2017–2024) were processed to analysis ready data format following the framework of Mullissa et al. (2021) in GEE, which included speckle filtering, radiometric terrain normalisation, border noise correction, and masking of radar shadow and layover areas (e.g. Fig. S4).'

We agree that there will be some discrepancies between the optical-derived training data and the SAR data, which could reduce the quality of the Sentinel-1 model training. However, it is not possible to quantify this without SAR-specific training data, which to our knowledge is not available. Since the model learns generalised patterns rather than exact pixel-level boundaries, we expect the overall impact on Sentinel-1 training quality to be limited.

We have modified the following text under 3.2 where we introduce the training data:

'These data were merged, incorporating all lakes from Zhang et al. (2024a), and then supplemented with any additional lakes from Kumar et al. (2025), which were typically smaller supraglacial lakes that were not present in the Zhang et al. (2024a) data. The lake outlines were used to generate training data (512-pixel image chips) with the corresponding 2020 Sentinel-1 and -2 image composites. There was no systematic coregistration offset between the lake outlines and the imagery we used for training. However, local shifts may occur, particularly in Sentinel-1 data, since the lake outlines were derived from optical imagery. Nonetheless, since the models learn generalised patterns rather than exact pixel-level boundaries, we expect the overall impact to be limited.'

4- The reliance on ArcGIS Pro for model training and deployment may pose practical limitations for some users, particularly given the strong open-data framing of the study. Since a Python file and trained model weights are provided, it would be helpful to clarify how portable these resources are outside the ArcGIS ecosystem. For example, can the trained DeepLabV3 model be readily deployed in standard open-source environments (e.g., PyTorch/TensorFlow workflows), or does it require ArcGIS-specific dependencies?

Our models can be called through ArcGIS Pro or ArcGIS API for Python, but cannot be directly loaded into a fully open source PyTorch workflow. Although all data derived from the models is open, we acknowledge model access is a potential limitation for users wishing to modifying the underlying weights or retrain. We were unable to decouple our models from ArcGIS in this case. We have modified the text under 'Code and data availability'.

'Deep learning models and the GLO glacial lake dataset presented in this manuscript are hosted on the Zenodo data repository at https://doi.org/10.5281/zenodo.17802334 (Rawlins et al., 2025). Our trained deep learning models rely on ArcGIS for deployment.'

5 – personal curiosity. Can you comment if you had any false positives and in which occasions? Deploying over such large areas is quite challenging.

Thank you for your interest. Yes, false positives (and negatives) were present in both Sentinel-1 and Sentinel-2 inventories, although they generally represent a small proportion of lake-related pixels overall. For Sentinel-2, false positives (FP; i.e., lakes mapped by our deep learning model and not in reference) were generally associated with more smaller lakes that were not in the reference datasets, accounting for 9.8% of lake pixel counts. For false negatives (FN: i.e., lakes/pixels not mapped by our deep learning but present in reference), our deep learning better tracked glacial lake outlines, particularly for larger lakes, where reference data may have over-captured or over-generalised such boundaries (hence FN in our data). Also some lakes which had non-distinct boundaries due to snow, ice cover or shadow were also fully or partially labelled FN, but these were typically smaller lakes Generally, smaller lakes are more ephemeral and so may have partially or fully drained in our classification, therefore appearing as a false negative (not present) – this is acknowledged in Sect. 5.1. FN equated to 7.2% of lake pixel counts.

For Sentinel-1, again there was good overall agreement, albeit slightly less than Sentinel-2. For FP, our deep learning captured lakes which were partially ice or snow covered were captured more frequently (see Fig. S4), which were not fully captured (if at all) in optical-based reference datasets – also see Fig.(c-d) and (g-h) below. FP accounted for 8.7% of total captured lake pixels. For FN, as discussed in Sect.5.3.2, lakes which had poorer F1 scores and were maybe not captured were smaller in size, surrounded by more complex terrain or be ephemeral in nature (e.g., in particular, supraglacial lakes). FN accounted for 21.7% of total lake pixels, a much larger percentage than Sentinel-2. This therefore reflects our suggestions of using preferential use of Sentinel-2 as the primary dataset, but complimented with data from Sentinel-1, which may have captured some of the more difficult-to-optically-observe lakes. Some examples of TP, FP and FN are now also given in Figure S5 (see below).

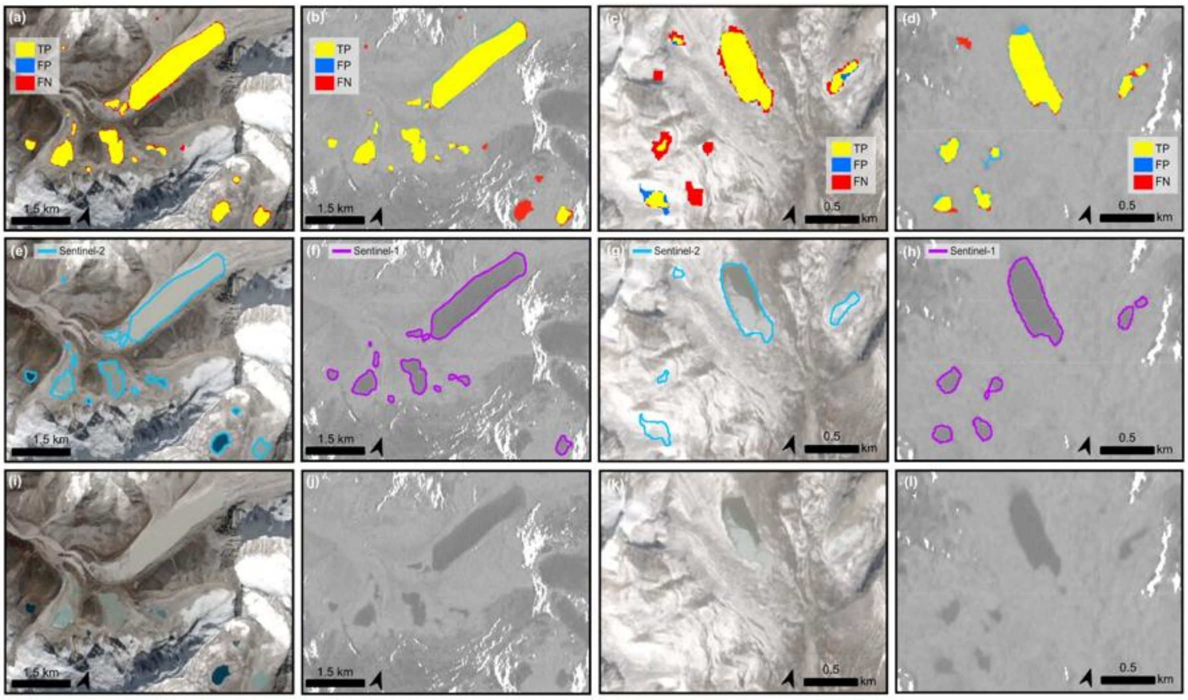

Figure S5: Examples of deep learning classification performance relative to reference inventories for Sentinel-2 and Sentinel-1 (2020) and subsequent lake delineations overlain on Sentinel-2 RGB and Sentinel-1 SAR training imagery. (a-d) True positives (TP), false positives (FP), and false negatives (FN), with left-hand pair (a,b) showing better performance in Sentinel-2, whereas right-hand pair (c,d) showing better performance in Sentinel-1 where snow and ice are more prevalent; (e-h) Corresponding model lake outlines overlain on 2020 training imagery; (i-l) corresponding imagery only.

PDF comments:

Table 2 – not sure the format adds useful info here

Whilst we appreciate this comment, we believe that including the data format for each variable ensures clarity in the dataset structure and supports the open, reproducible and interpretable use of the dataset across different platforms.

Line 198 – Maybe polynomial could reveal nonlinear trends here.

We agree that non-linear trends will be apparent for some glacial lakes; however, these are generally apparent on multi-decadal timescales (e.g. Haritashya et al., 2018). Our primary goal was to estimate annual rates of change, whilst accounting for outliers. Applying polynomials to ~8 data points could potentially introduce false trends or spurious reversals of lake area change at the end of the timeseries, whilst also making it difficult to estimate rates of change. We therefore omit polynomials in this analysis but acknowledge the benefit of accounting for non-linear trends as our dataset extends.

Line 213 – How do you deal with this if for example your median includes such changes? What could be the effect of this?

Thank you for your comment. Median compositing can smooth those short-term changes occurring within the compositing period, therefore expansion, drainage or small shifts in lake position may be underrepresented or omitted (particularly for lakes such as those supraglacial in origin, which are typically more dynamic). This may lead to some reduced detection of lakes or the potential for slight boundary mismatches when compared with other inventories. However, it is important to note that often regional inventories (including our Nepal-transboundary here and the validated inventories we use) are derived from composite imagery and are likely to represent the seasonal state of a lake rather than its short-lived minimum/maximum state over single date/s.

Additionally, as we are only using a single composited year for comparison against the two external inventories, any temporal mismatches are constrained to that year. We believe this median compositing method remains appropriate for regional-scale lake assessments (and for further wider-scale Himalayan inventories).

Line 221 – Maybe adding IoU would be nice

Whilst IoU values are provided in Table S2 and Table S3 in the Supplementary Information, we agree that its inclusion in the main body of text is a valuable addition. These values have now been added into Sect. 4.1.1 and Sect. 4.1.2 for both Sentinel-1 and -2. See added text below.

Sect. 4.1.1

'First, comparisons with Zhang et al. (2024a) found good agreement in detections between the datasets, resulting in F1 scores of 0.82 for Sentinel-1 and 0.87 for Sentinel-2 (Table S2), with corresponding IoU values of 0.70 and 0.78, respectively.'

Sect. 4.1.2

'When compared against the combined Zhang et al. (2024a) and Kumar et al. (2025) inventories the F1 scores were slightly lower, with 0.79 and 0.85 for Sentinel-1 and -2 respectively, with corresponding IoU values of 0.66 and 0.74. Again, these results demonstrate good spatial agreement across datasets of differing sensors.'

Line 310 – blank

We could not see this sticky note unfortunately.

Line 325 – Can you be more specific (differ slightly in temporal trends)

We have now amended the text to the following:

'Collectively the two lake datasets derived from SAR and optical Sentinel sensors (respectively) show broadly comparable spatial patterns across the Nepal-transboundary but differ slightly in their interannual detection behaviour temporal trends likely due to sensor-specific characteristics. Sentinel-2 captured 23% more lakes over the time period with a greater number of unique lakes and year-to-year variability, whilst Sentinel-1 provided more consistent annual detection lake counts with lower year-to-year variability.'

Also, we have added further interpretation of why there is differ in temporal trends in the new Sect. 5.3.2. Text as follows:

'Sentinel-1 derived lakes provide valuable imaging through all-weather conditions, assisting with supplementing data gaps that may be left by optical images (i.e., the capture of lakes during monsoon or cloudy periods) and even the capture of snow-covered lakes which are omitted in Sentinel-2 (Fig. S4). It is noted, however, that these Sentinel-1 delineations are more prone to fragmented or irregular lake outlines, particularly for smaller lakes, likely reflecting known SAR challenges including speckle noise, look direction, incidence angle and backscatter variability over water and rugged terrain (Khan et al., 2025; Miles et al., 2017). At the same time, the Sentinel-1 dataset does provide more temporally consistent counts, which is a likely reflection of a combination of factors: (i) the narrower acquisition compositing window meaning lakes were more likely to be captured at a similar time each year; (ii) a greater omission of smaller lakes, reducing year-to-year variability in the seasonal changes of these smaller lakes, that are then intermittently captured by Sentinel-2; (iii) the use of median SAR composites which are more likely to exclude small, highly dynamic lakes (Kumar and Vijay, 2026). As such, this temporal variability reflected for Sentinel-1 should be interpreted as reflecting a more conservative subset of lake detections, rather than completeness.'

Line 326 – This is interesting. S1 should be great in detecting water. Do you have any thoughts on why? See which one you miss is terms of size might help seeing whether size matters here. Otherwise, it'd be interested to see whether this is due to geometric distortions (e.g., shadowing)

Thank you for your comment – we agree, in principle, that Sentinel-1 SAR should be effective for water detection. However, in our study the main limitation is not the detection of open water itself, but the reliable detection and delineation of small, irregular or topographically complex glacial lakes because of steep

Himalayan terrain. Our validation indicates that size is a dominating factor, which is now backed up by further robust performance analysis which shows that Sentinel-1 requires a larger lake-size threshold than Sentinel-2 to achieve robust classification performance (F1 $\geq$ 0.85).

Line 453 - I'd just say performance

Changed to performance.

Line 474 – what would this be in this case (minimum mapping threshold)?

Here we are referring to our minimum mapping threshold (0.001 km^2), with the Kumar (2025) dataset containing a greater proportion of small lakes that are close to this threshold. The sentence has been restructured for clarity.

'Notably, when including the Kumar et al. (2025) inventory for our classification comparison, overall F1 and IoU scores slightly decreased for both sensors, including Sentinel-2, relative to comparisons using the Zhang et al. (2024a) inventory alone. This reduction is likely due to the Kumar inventory containing a greater proportion of more smaller lakes near our the minimum mapping threshold (0.001 km^2), which are inherently more difficult to detect and that were less accurately classified due to spatial resolution constraints and mixed pixel effects (e.g. Fig. 5).'

References

Haritashya, U., Kargel, J., Shugar, D., Leonard, G., Strattman, K., Watson, C., Shean, D., Harrison, S., Mandli, K., and Regmi, D.: Evolution and Controls of Large Glacial Lakes in the Nepal Himalaya, Remote Sens., 10, 798, https://doi.org/10.3390/rs10050798, 2018.

Khan, I., Vardaman, M., Johnston, J., and Jacobs, J. M.: Glacial lake mapping: Progress, challenges, and future directions, Arct. Antarct. Alp. Res., 57, 2592365, https://doi.org/10.1080/15230430.2025.2592365, 2025.

Kumar, A., Mal, S., Schickhoff, U., and Dimri, A. P.: Basin-scale spatio-temporal development of glacial lakes in the Hindukush-Karakoram-Himalayas, Glob. Planet. Change, 245, 104656, https://doi.org/10.1016/j.gloplacha.2024.104656, 2025.

Kumar, R. and Vijay, S.: Automated satellite-based glacial lake inventory and change detection in High Mountain Asia, Sci. Rep., 16, 5760, https://doi.org/10.1038/s41598-026-35446-0, 2026.

Miles, K. E., Willis, I. C., Benedek, C. L., Williamson, A. G., and Tedesco, M.: Toward Monitoring Surface and Subsurface Lakes on the Greenland Ice Sheet Using Sentinel-1 SAR and Landsat-8 OLI Imagery, Front. Earth Sci., 5, https://doi.org/10.3389/feart.2017.00058, 2017.

Xu, X., Liu, L., Huang, L., and Hu, Y.: Combined use of multi-source satellite imagery and deep learning for automated mapping of glacial lakes in the Bhutan Himalaya, Sci. Remote Sens., 10, 100157, https://doi.org/10.1016/j.srs.2024.100157, 2024.

Zhang, T., Wang, W., and An, B.: Heterogeneous changes in global glacial lakes under coupled climate warming and glacier thinning, Commun. Earth Environ., 5, 374, https://doi.org/10.1038/s43247-024-01544-y, 2024.