https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

High-resolution mapping of glacial lake expansion in Kyrgyzstan (2016–2024) using Sentinel-2 imagery

Valentine Piroton[1], Adam Emmer[2], Hans-Balder Havenith[1]

[1]Department of Geology, University of Liège, Liège, 4000, Belgium
[2]Department of Physical Geography and Geoecology, Faculty of Science, Charles University, Prague, Czechia

Correspondence to: Valentine Piroton (v.piroton@uliege.be)

Abstract. Glacial lakes are sensitive indicators of glacier retreat and are important for regional hydrological and hazard assessments. This study presents a high-resolution (10 m) inventory of glacial lakes in Kyrgyzstan for 2022–2024, derived from Sentinel-2 imagery using a semi-automated workflow complemented by manual refinement. A Random Forest classification model trained on the 2022–2024 inventory was applied to 2016–2017 imagery to reconstruct historical lake distributions and analyze changes over an eight-year period. The 2024 inventory includes 2 592 lakes covering 77.6 km^2, mostly small (<0.05 km^2) lakes located between 3 500 and 3 800 m elevation. Between 2016 and 2024, the number of lakes increased by 10.5 % and total lake area by 8.7 %, driven primarily by the formation of small, high-altitude proglacial lakes, whereas larger, lower-elevation lakes remained largely stable. Supraglacial lakes exhibited slight area increases and, on average, an upward shift in elevation, whereas proglacial, and glacier-detached lakes showed minimal changes. Regional trends reveal pronounced heterogeneity, with Issyk-Kul, Batken, and Talas emerging as regions of new lake formation. Comparison with global datasets confirms completeness and reliability of our inventories. These results highlight the ongoing influence of glacier retreat on lake formation and expansion in Kyrgyzstan, providing a robust baseline for hazard assessment, water resource management, and future cryospheric monitoring.

1 Introduction

Glacier retreat is an ongoing worldwide phenomenon, leading to the formation and expansion of numerous glacial lakes (Frey et al., 2010; Carrivick and Tweed, 2016; Harrison et al., 2018; Shugar et al., 2020). These lakes respond rapidly to glacier mass loss and indirectly reflect climate-driven changes in glacier-covered landscapes. They also represent important water resources and potential hazards, particularly concerning the risk of catastrophic downstream flooding (Clague and Evans, 2000; Huggel et al., 2002; Mergili et al., 2013; Emmer et al., 2014; Haeberli et al., 2016; Otto, 2019). Glacial lakes in contact with glacier ice also efficiently transmit heat to the ice front, accelerating melting and calving, and contributing to the rapid disintegration of stagnant glacier tongues over decades (Bolch et al., 2012).

Sudden releases from moraine- or ice-dammed lakes cause glacial lake outburst floods (GLOFs), producing extreme peak discharges and flood volumes (Evans and Clague, 1994; Carrivick and Tweed, 2016; Bat'ka et al., 2020). The likelihood and

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

magnitude of such events depend on dam integrity, total drainable volume, and slope stability (Sattar et al., 2021). GLOFs are among the most hazardous glacier-related events, alongside landslides, ice avalanches, and slope instabilities. A total of 1 997 GLOFs occurred between 1901 and 2017 in the Andes, the Pacific Northwest, Iceland, the European Alps, Scandinavia, and High Mountain Asia, with the latter accounting for 24 % GLOFs worldwide (Veh et al., 2022). Indeed, Central Asia, and particularly the Tien Shan, is highly susceptible to GLOFs due to the high concentration of mid-latitude glaciers there (Aizen et al., 2006; Takeuchi et al., 2014). More than 80 % of Kyrgyzstan is mountainous, hosting over 2 000 glacial lakes, at least 300 of which are considered at risk of outburst (Janský et al., 2010). Steep terrain, active seismicity, and residents' high dependence on meltwater amplify the socio-environmental vulnerability of downstream communities.

The growth of glacial lakes in Kyrgyzstan mirrors the widespread glacier retreat documented across High Mountain Asia, including the Tien Shan, where glaciers have been shrinking rapidly in recent decades (Shugar et al., 2020). During that time, scientific interest in glacial lakes and GLOFs has increased globally. Inventories have been compiled for high mountain regions including the Andes, the European Alps, and Central Asia (Strozzi et al., 2012; Mergili et al., 2013; Emmer et al., 2016; Buckel et al., 2018; Chen et al., 2020; Zhang et al., 2022). In Kyrgyzstan, regional-scale inventories exist for selected mountain ranges, such as the Kyrgyz Range (Kattel et al., 2020; Daiyrov and Narama, 2024) and the Terskey Range (Kattel et al., 2020; Daiyrov et al., 2022). Although these studies provide valuable information, they are geographically limited, often focusing only on the northern Tien Shan. Recent initiatives such as the High Mountain Asia glacier lake inventory (Hi-MAG) and the global inventory of glacial lakes (GIGLak) now offer broader spatial coverage across Central Asia and worldwide, respectively. Although they provide near-complete coverage of glacial lakes, they still exhibit coarse spatial resolution, limited seasonal targeting, do not consistently include vectorized lake polygons, and/or are not manually inspected. These limitations underscore the need for updated, high-resolution glacial lake inventories to enable national-scale monitoring of lake dynamics and GLOF hazards.

Given the remote and often inaccessible nature of glacial lakes, remote sensing plays a central role in their detection and monitoring (Kääb et al., 2005; Bolch and Kamp, 2006; Capps et al., 2010; M. Zhang et al., 2020; B. Zhang et al., 2021). High-resolution optical satellite imagery, including Sentinel-2 and Landsat-8, has proven effective for lake mapping and many studies have applied (semi-)automated delineation methods (Frey et al., 2010; Bolch et al., 2011; Liu et al., 2016; Veh et al., 2018; Shugar et al., 2020). In particular, the Normalized Difference Water Index (NDWI) has emerged as a robust indicator for identifying water bodies, allowing accurate delineation of both proglacial and supraglacial lakes (Huggel et al., 2002; Bolch et al., 2011; Zhang et al., 2022). Machine learning approaches, such as Random Forest classifiers, have further enhanced mapping accuracy by efficiently discriminating water from surrounding terrain and detecting smaller lakes often overlooked in coarser datasets (Chen et al., 2020; Wangchuk and Bolch, 2020).

Building on these advances, we here integrate multispectral satellite data, NDWI-based semi-automated detection, Random Forest classification, and expert-guided manual refinement to produce the first comprehensive polygon-based glacial lake inventory for Kyrgyzstan. Our specific objectives are to:

1. conduct a multi-temporal analysis of glacial lake distribution and surface area between 2016 and 2024;

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

2. quantify changes in lake number and extent over this period; and
3. identify hotspot regions where lake expansion is most pronounced and where GLOF risk is likely to increase.

In addition, we evaluate the extent to which large-scale automated glacial lake inventories can accurately capture lakes at risk of GLOFs. By comparing our results with global datasets, we demonstrate the added value of expert-guided refinement within automated classification workflows, which remains essential for achieving the level of detail and accuracy required for hazard-oriented inventories. The resulting dataset contributes both to scientific understanding of cryospheric changes and to practical applications in hazard monitoring and adaptation strategies in High Mountain Asia.

2 Study area

Kyrgyzstan is a landlocked country in Central Asia covering 199 951 km^2 and sharing borders with Kazakhstan to the north, Uzbekistan to the west, Tajikistan to the south, and China to the east. The topography of Kyrgyzstan is dominated by the Tien Shan Mountains, bordered to the south by the Pamir Mountains. Elevations range from approximately 400 m in the Fergana Basin to 7 439 m at Jengish Chokusu (Pik Pobeda), the highest peak in the country, located on the China–Kyrgyzstan border (Fig. 1). More than 80 % of the territory lies above 1 500 m elevation, and a large proportion exceeds 3 000 m, creating steep elevation gradients that strongly influence climate, hydrology, and ecosystems.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

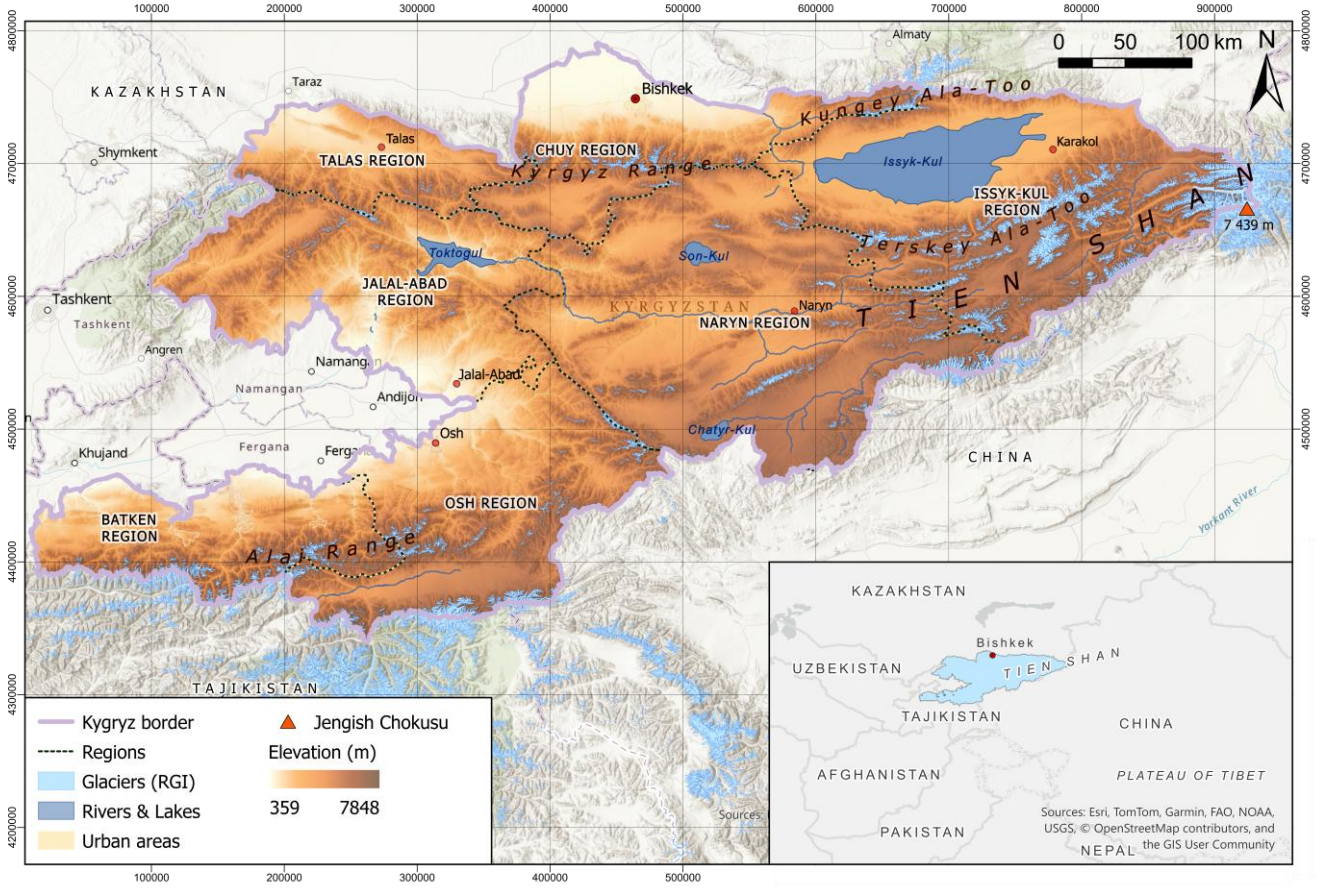


Figure 1. Hillshade and elevation map of Kyrgyzstan showing major mountain ranges, glacier outlines (RGI v7.0), main rivers and lakes, and the highest peak of Kyrgyzstan (Jengish Chokusu). Administrative regions are indicated for spatial reference (black dashed lines). Elevation data are from the AW3D30 DEM. Base map: Esri World Hillshade (Esri, World Hillshade; see References). Sources: Esri, Vantor, Airbus DS, USGS, NGA, NASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA, Intermap, TomTom, Garmin, FAO, NOAA, USGS and the GIS user community | Powered by Esri.

The climate of Kyrgyzstan is strongly continental, influenced by its landlocked location and the surrounding mountain belts, which isolate the country from maritime influences (Akimaliev et al., 2013). The country experiences hot, dry summers in the lowlands and cold, snowy winters nationwide. Alpine conditions prevail in high mountain regions. Mean July temperatures vary from 17–40 °C in valleys to around 4 °C at high altitudes, and severe frosts occur throughout the country in winter (Kulikov and Schickhoff, 2017; Tomaszewska and Henebry, 2018). Annual precipitation is highly variable, from less than 150 mm in some interior basins (e.g., Issyk-Kul) to over 1 000 mm in the Fergana Valley (Adyshev et al., 1987).

Snow accumulation and glacier melt dominate the hydrological cycle in Kyrgyzstan's high mountains, producing strong seasonal contrasts between summer runoff and low winter flows. This variability, combined with intense solar radiation and the influence of dry continental air masses, creates pronounced environmental gradients over short distances (Akimaliev et al., 2013). Seasonal snow and glacier melt is the primary water source for rivers and lakes, making the region highly sensitive to

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

climatic changes. Rising temperatures and declining snow cover in recent decades have accelerated glacial melt, increasing both the volume of glacier-fed lakes and the frequency of GLOFs. The steep and rugged terrain further shapes local microclimates and contributes to hazards such as landslides, avalanches, and rockfalls, particularly in areas experiencing active glacial retreat.

Land cover in Kyrgyzstan reflects the strong vertical climatic zonation: agricultural lands are concentrated in valleys and lowlands (<10 % of total land) and forests occupy about 5 % of the territory, whereas the majority of the land (>50 %) is used as rangeland (CACILM/ADB, 2010). Rural livelihoods and national agriculture depend heavily on meltwater availability, making the stability of glaciers and glacial lakes a central concern for both environmental and socio-economic resilience. Kyrgyzstan hosts several thousand glaciers, particularly in the northern and central Tien Shan; these glaciers continuously feed glacial lakes, which play a critical role in regional hydrology by supplying meltwater to major rivers such as the Naryn, a key tributary of the Syr Darya. In addition, these high-altitude lakes sustain extensive irrigation schemes and hydroelectric power generation. However, the combination of steep topography, seismicity, and accelerating glacial melt makes many glacier-fed lakes prone to GLOFs, which threaten downstream communities, infrastructure, and agricultural areas.

Historical examples of GLOFs illustrate their socio-economic impacts. The 2008 Zyndan West Lake GLOF in the Issyk-Kul region resulted in fatalities, infrastructural damage, and agricultural losses (Narama et al., 2010). The 1998 Shakhimardan event at the border between Uzbekistan and Kyrgyzstan was one of the deadliest GLOFs of the last century in Central Asia (Petrakov et al., 2020); this GLOF was presumably triggered by an outburst from small lakes and initiated a suite of chain reactions. The failure of the upper lake likely caused the overtopping of lower lakes, followed by intense channel erosion, sediment entrainment, and the transformation of the flood wave into a debris flow (Petrakov et al., 2020). A further example is the 2012 Teztor event in the Ala-Archa basin (Erokhin et al., 2018), which was caused by drainage through intra-moraine channels, showing that, when enhanced by heavy precipitation and rapid melting, even small outbursts can trigger destructive debris flows. Amongst many others, these cases highlight the importance of continuous monitoring and high-resolution mapping of glacial lakes to assess and mitigate GLOF risks.

The geographic extent of this work corresponds to areas across the territory of Kyrgyzstan at elevations ≥3 000 m to focus on high-mountain glacial lake environments. This study area was selected because of the high density of glaciers, the abundance of proglacial lakes, and the growing potential hazard associated with GLOFs.

3 Data sources

3.1 Satellite imagery

This study uses multispectral imagery from the Sentinel-2 mission of the Copernicus Program. Sentinel-2A, launched in 2015, initially operated alone until Sentinel-2B joined in 2017, enabling a five-day revisit frequency when both satellites are active. Level-2A surface reflectance products were used herein to provide atmospherically corrected imagery suitable for large-scale glacial lake mapping. Two distinct temporal windows were considered:

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

(1) July to October of 2016 and 2017, the earliest full-coverage period after the launch of Sentinel-2; and

(2) July to October of 2022–2024, representing recent conditions for our glacial lake inventory.

Both intervals correspond to the regional ablation season, when seasonal snow cover is minimal and glacial lakes are most clearly observable. For clarity, these periods are hereafter referred to as the 2016 and 2024 inventories, respectively, with each temporal window encompassing potential seasonal variability and providing a conservative basis for comparison. For each period, only scenes with less than 5% cloud cover were retained, as selected using the Copernicus Data Space Browser. In areas where a single tile did not meet these conditions, multiple adjacent Sentinel-2 tiles were merged to obtain optimal coverage, minimizing both cloud contamination and residual snow. The complete set of tiles used for the analysis is shown in Fig. 2, which illustrates the Sentinel-2 tiling system over Kyrgyzstan.

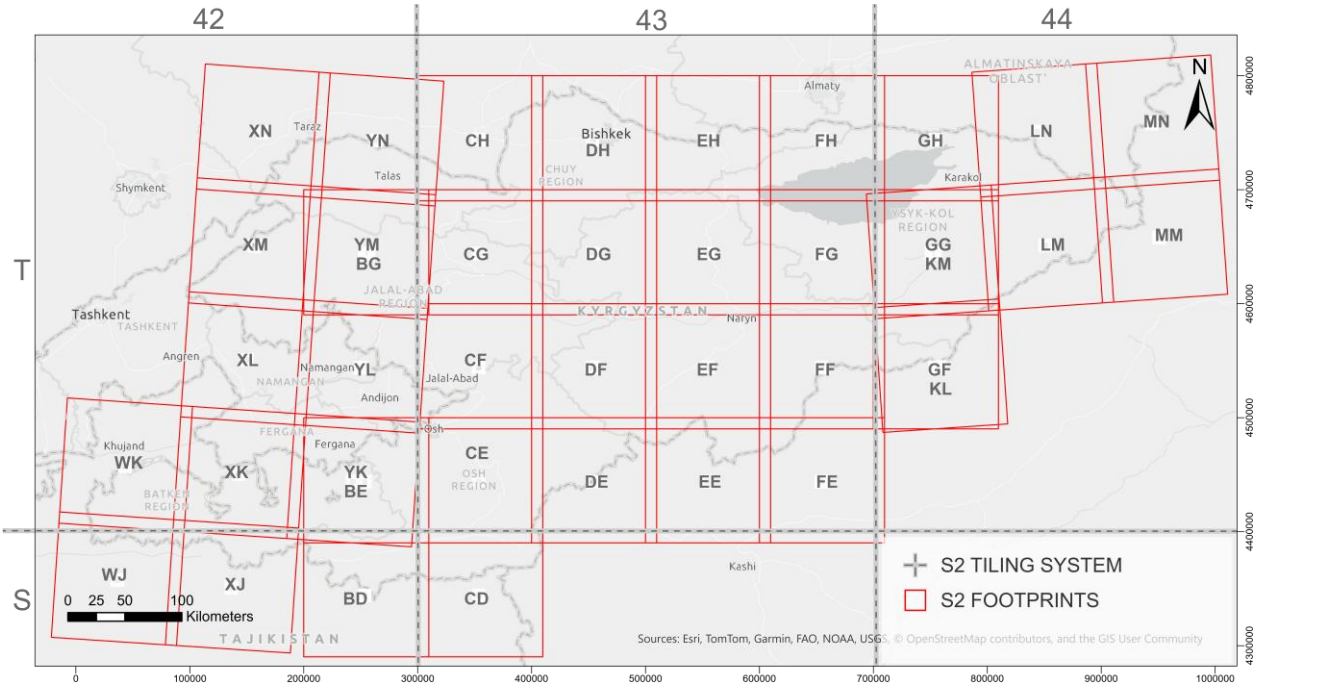


Figure 2. Ground footprints of the Sentinel-2 Level-2A scenes for 2016–2017 and 2022–2024, with the highlighted tiling system shown along the top and left borders. Bold two-letter pairs indicate tile IDs. The background shows the Kyrgyzstan national boundary for reference. Sources: Esri, TomTom, Garmin, FAO, NOAA, USGS | Powered by Esri.

The 2022–2024 dataset comprised 41 images, mostly from Sentinel-2B, with limited contributions from Sentinel-2A (e.g., 21 August and 25 October 2023). The 2016–2017 dataset also contained 41 images, primarily from Sentinel-2A (Fig. 3). For both periods, four bands were used: B2 (Blue), B3 (Green), B4 (Red), and B8 (NIR). This dual-period dataset provides consistent, high-resolution coverage of Kyrgyzstan's alpine regions, enabling robust delineation of glacial lake boundaries and detection of subsequent changes.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

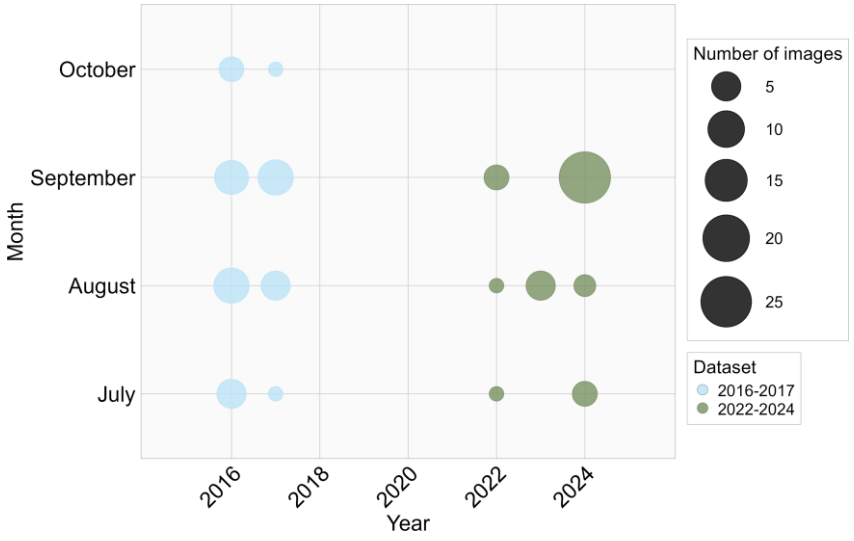


Figure 3. Number of Sentinel-2 Level-2A scenes used per year and month (July–October) over Kyrgyzstan for the periods 2016–2017 and 2022–2024.

3.2 Reference data

Several ancillary datasets on elevation, glacier outlines, river networks, and regional boundaries were used to provide contextual information for glacial lake mapping, interpretation, and analysis. Glacier outlines were obtained from the Randolph Glacier Inventory (RGI v7.0; Maussion et al., 2022; RGI Consortium, 2023), providing a globally consistent reference for glacier extent *circa* 2000. This dataset enables the differentiation of proglacial lakes from other water bodies and supports analyses of lake–glacier interactions.

Topographic context was provided using a high-resolution digital elevation model (DEM) from the Advanced Land Observing Satellite World 3D 30 m dataset (AW3D30; Tadono et al., 2016; JAXA, 2025). AW3D30 has 1″ (~30 m) spatial resolution and is derived from the 5-m mesh version of ALOS World 3D. The DEM was generated from optical stereo images acquired by the Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) between 2006 and 2011. PRISM captured three-directional stereo imagery with a 2.5 m nadir resolution, enabling precise 3D topographic data (González-Moradas and Viveen, 2020; Li et al., 2022). The AW3D30 dataset, released free of charge in 2016 and updated in 2024 to fill voids, provides a vertical RMSE of approximately 4.4 m (Tadono et al., 2016). It was delivered in 1° × 1° tiles, offering a reliable reference for hydrological modeling, glacier-lake delineation, and post-classification filtering.

Administrative and political regions were obtained from the Region Line Boundaries, Kyrgyzstan, 2016 dataset (Stanford, 2012). These line features were primarily derived from paper topographic maps provided by the State Service of Cartography and Geodesy of the Kyrghyz Republic, with supplementary Vector Map Level 0 data. These datasets provide a spatial framework for aggregating glacial lake inventories and interpreting regional patterns.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Rivers and streams were represented using the World Linear Water dataset (Esri, Garmin International, Inc., 2025). This dataset offers global coverage of linear hydrographic features that are too narrow to be represented as polygons. It ensures detailed representation of drainage networks and downstream connectivity of glacier-fed lakes.

Major lakes were incorporated from the Natural Earth World Lakes, 1:10 million dataset (Patterson and Kelso, 2012). This polygon layer includes natural and artificial lakes worldwide, ranked by relative importance, and was derived from the CIA World Data Bank 2 with supplementary satellite-based updates. It provides additional reference points for evaluating glacial lake inventories and hydrological connections.

All reference datasets were projected in ArcGIS Pro to a common coordinate system (UTM Zone 43N) and resampled where necessary to ensure spatial consistency with the Sentinel-2 imagery. Together, these datasets—including glacier outlines, the DEM, administrative boundaries, rivers, and major lakes—establish a comprehensive spatial framework that supports high-resolution mapping, temporal analysis, and visualization of glacial lakes in Kyrgyzstan.

4 Methods

4.1 Methodological overview

The methodological workflow of this study, implemented within Python-based environments, was designed to accurately map glacial lakes in Kyrgyzstan and analyze their spatio-temporal evolution. It follows a two-step approach combining expert-based manual refinement with automated machine learning techniques (Fig. 4). By manually correcting the inventory from 2022–2024, we ensured a high-quality reference dataset because spectral indices alone (e.g., NDWI, normalized difference snow index) are often insufficient to reliably delineate lakes in complex mountainous terrain (Shugar et al., 2020). This approach provides robust training samples while maintaining reproducibility and scalability, essential for consistent temporal comparisons and quantitative assessments of lake changes.

We first established a baseline inventory of glacial lakes for 2022–2024 using Sentinel-2 Level 2A imagery. Spectral index and topographic constraints were used to generate preliminary lake outlines, which were then carefully refined through extensive manual editing to ensure accurate delineation. This manually corrected inventory provided a consistent set of reference polygons, serving as training samples for supervised classification.

We then used these training samples to train and optimize a Random Forest classifier, subsequently applied to 2016 Sentinel-2 imagery to produce a second lake inventory. By training the model on recent, high-quality 2022–2024 imagery with superior atmospheric correction, we ensured robust and reliable samples, justifying this backward application. Nonetheless, our comparison between the two inventories enabled quantitative assessment of the spatial and temporal dynamics of glacial lakes across the Kyrgyz mountains.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

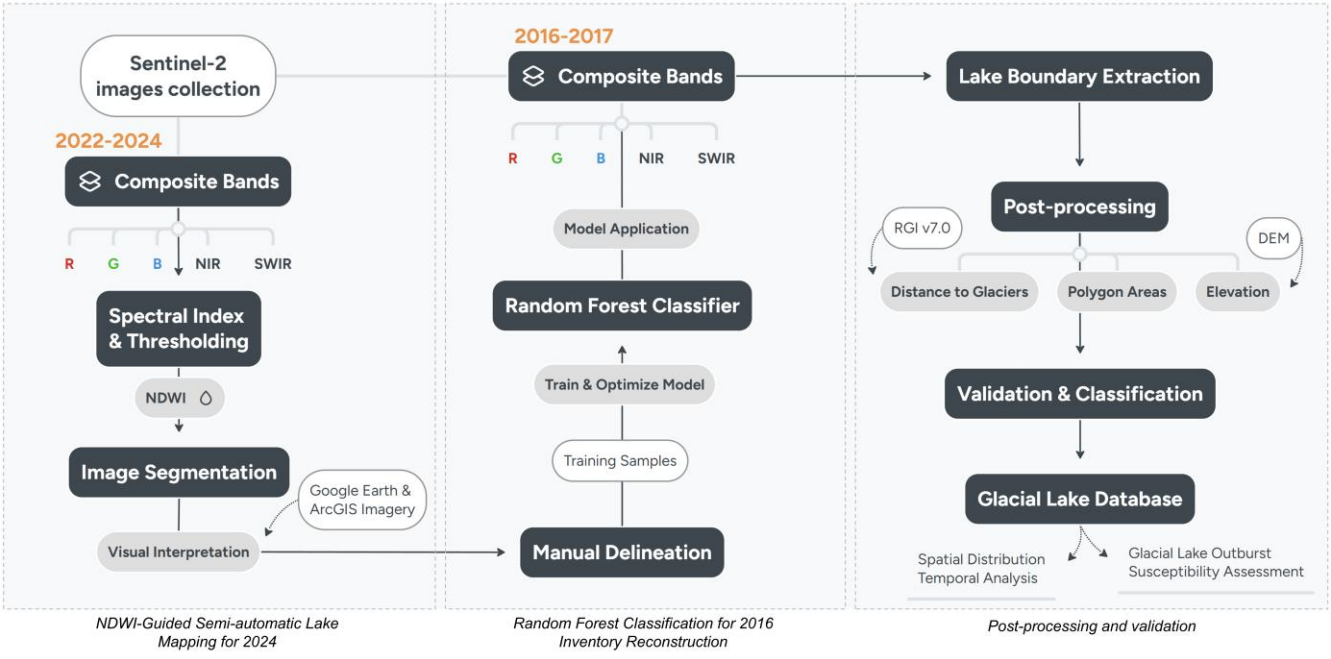


Figure 4. Workflow of the methodological approach combining the 2022–2024 reference inventory and the 2016-2017 supervised classification for glacial lake evolution analysis.

4.2 NDWI-guided semi-automatic mapping for the 2024 glacial lake inventory

Our inventory of glacial lakes was constructed using remote sensing data within a Python-based environment in a geographic information system (ArcGIS Pro 3.3.2). Forty-one atmospherically corrected Sentinel-2 Level 2A images, acquired during the summer seasons from 2022 to 2024, were selected and downloaded via the Copernicus Browser (https://browser.dataspace.copernicus.eu). We used the red, green, blue, and NIR bands, which have a spatial resolution of 10 m, enabling detection of lakes larger than 0.003 km^2 (Kattel et al., 2020). After merging these bands in the 41 images, the four rasters were clipped to the border of Kyrgyzstan. The processed rasters were used to calculate the NDWI as:

$$\text{NDWI} = \frac{(\text{GREEN} - \text{NIR})}{(\text{GREEN} + \text{NIR})} \quad (1)$$

Here, GREEN and NIR represent surface reflectance in the bands B3 and B8, respectively (McFeeters, 1996; Zhang et al., 2018). NDWI values range from −1 to +1. Values close to +1 indicate the presence of water, intermediate values provide a relative measure of soil moisture, and negative values correspond to dry areas or non-water surfaces.

Before applying this workflow to the entire country, the robustness of the semi-automated NDWI-based approach was first assessed on a test area of interest (AOI) of 8 935 m^2 corresponding to the Kyrgyz Range near Bishkek (Fig. 5) as a blind evaluation of the workflow. Within this AOI, preliminary polygons were generated from NDWI thresholds tested through an

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

iterative trial-and-error procedure, including both small and large lakes without applying any minimum area, elevation, or glacier-proximity constraints. Morphological filtering and grouping by region were applied to reduce noise and merge contiguous pixels. The resulting polygons were then compared to manually predelineated lakes within the AOI to quantify the proportion of correctly identified lakes (>87.64 %) and the accuracy of their boundaries. This step allowed the identification of optimal NDWI thresholding and filtering parameters, ensuring that the semi-automated workflow produced reliable preliminary delineations while minimizing false positives and artefacts. The selected NDWI threshold (>0.07) differs from the initial zero threshold established for Landsat images (McFeeters, 1996) but is broadly consistent with the value of 0.1 used in a previous Sentinel-2-based study (Kaplan and Avdan, 2017). Once these parameters were established, the semi-automated workflow was applied across the entire Kyrgyzstan dataset. The resulting water mask was then processed following the post-processing and validation procedures described in Sect. 4.4.

4.3 Supervised Random Forest classification for the 2016 inventory reconstruction

The second stage of the workflow aimed to produce a glacial lake inventory for 2016 using a supervised classification approach. The manually refined 2022–2024 inventory served as a high-quality reference, providing labeled polygons of lake and non-lake areas. Training samples were extracted from these polygons, creating a balanced dataset with 20 000 pixels for each class to capture spectral and topographic variability while avoiding spatial bias or overrepresentation of large lakes. A Random Forest classifier was implemented in Python (Visual Studio Code) and trained using these samples. Hyperparameters were optimized to maximize recall for the lake class while maintaining overall accuracy, reflecting our priority of detecting all possible glacial lakes in the study area. The model produced per-pixel class probabilities, and an optimal probability threshold was determined from validation data to balance the detection of lakes against the minimization of false positives. The input for classification was the preprocessed 2016 Sentinel-2 composite raster (Blue, Green, Red, NIR), prepared to match the 2022–2024 imagery in terms of atmospheric correction, band merging, and clipping to the Kyrgyzstan boundary. Due to the large spatial extent, classification was performed in blocks using windowed reading to ensure memory efficiency and computational stability. Each block was processed independently, and the probability threshold was applied to generate a binary lake mask. The resulting raster was then submitted to the post-processing and validation workflow detailed in Sect. 4.4 to ensure spatial, spectral, and topographic consistency, producing the final 2016 lake inventory.

4.4 Post-processing and accuracy assessment

Post-processing and validation were performed to ensure the spatial, spectral, and topographic consistency of the glacial lake inventories for both periods. The filtered 2024 inventory raster was converted into polygons, with a minimum area threshold of 0.003 km^2. Polygons were further filtered by proximity to glaciers (within 30 km, based on the RGI v7.0) and by elevation (>3 000 m) using the AW3D30 DEM. Finally, all polygons were visually inspected to correct residual misclassifications, remove false positives (particularly river segments and shaded areas), and ensure the geometric integrity of the mapped lakes.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

This semi-automated workflow ensured a high-quality reference inventory suitable for subsequent training of the 2016 classification model.

For the 2016 Random Forest-based inventory, post-processing focused on threshold optimization and accuracy assessment. Model performance was evaluated by comparing predicted lake locations to the manually mapped reference polygons. Class probabilities generated by the model were adjusted using an optimal probability threshold derived from validation data, prioritizing recall for the lake class to reduce omission errors. A confusion matrix was produced, and standard metrics—including accuracy, precision, recall, and F1-score—were calculated to evaluate model performance. The post-processing workflow also included morphological filtering and grouping by region to clean artefacts and refine lake boundaries, followed by polygon conversion, minimum area filtering (0.003 km^2), and spatial/topographic filtering based on glacier proximity, elevation, and slope (to exclude steep or shaded areas). Highly elongated objects, likely rivers, were removed to reduce false positives.

We quantitatively evaluated the inventories by systematically comparing each automated dataset to its corresponding manual reference. Specifically, the NDWI-based inventory was assessed against the manual 2024 inventory, whereas the Random Forest-based inventory was compared to the manual 2016 inventory. Metrics such as the number of reference lakes detected, the proportion of correctly predicted lakes, and standard classification measures (precision, recall, and F1-score) were calculated to quantify both omission and commission errors. This approach ensured that only physically plausible lakes were retained while maintaining computational efficiency. Together, these post-processing and validation steps support the refinement and expert-checked comparison of automated and manual inventories, providing a robust methodological framework for analyzing glacial lake dynamics across the Kyrgyz mountains.

4.5 Estimation of area uncertainties

The spatial uncertainties on lake boundaries were quantified following the approach of Hanshaw and Bookhagen (2014), which estimates the one-sigma boundary error based on image resolution and polygon geometry. For each lake polygon, the boundary uncertainty (1σ, in m^2) was calculated as:

$$\text{Error } (1\sigma) = \frac{P}{R} \times 0.6872 \times \frac{R^2}{2} \quad (2)$$

where P is the polygon perimeter (in meters), R is the spatial resolution of the input imagery (10 m for Sentinel-2) and 0.6872 is a correction coefficient reflecting the assumption that approximately 69 % of pixels may contain positional errors. This method estimates how small positional inaccuracies along the lake boundary translate into total area uncertainties, providing a consistent geometric error value for each polygon.

To express these uncertainties in terms of lake area, the estimated boundary error was converted to square kilometers and analyzed as a function of lake size class. Smaller lakes, characterized by higher perimeter-to-area ratios, are expected to exhibit

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

proportionally higher area uncertainties than larger lakes. This procedure provides a consistent, quantitative estimate of the potential delineation error inherent to the spatial resolution and polygon complexity.

4.6 Quantitative and qualitative characterization of glacial lakes

4.6.1 Lake size distribution

The inventory considered only natural lakes formed or influenced by glacial processes. Large non-glacial lakes in Kyrgyzstan were excluded due to their disproportionate size and different formation processes: Issyk-Kul, Son-Kul, and Chatyr-Kul are tectonic lakes occupying structural depressions, while Toktogul is an artificial reservoir (Fig. 1). Kel-Suu, although partially glacially fed, was also excluded because its tectono-glacial origin and size make it atypical relative to most glacial lakes.

The statistical distribution of glacial lake sizes was first examined to determine whether the dataset could be treated as a single continuous population. Outlier detection and Hartigan's Dip Test (Hartigan and Hartigan, 1985) indicated a unimodal distribution (D = 0.0041, p = 0.9951), indicating that our inventory of glacial lakes forms a continuous size spectrum. The interquartile range was also used to detect extreme values, confirming the treatment of the dataset as a single continuous population. Complementary cumulative distribution functions (CCDFs) of lake areas were computed in RStudio using the poweRlaw package, representing the probability that a lake has an area greater than or equal to a given value x. This approach is suitable for highly right-skewed distributions dominated by small lakes, as is the case here (see Sect. 5.1). Scaling behavior was quantified by fitting a power-law model to the CCDF for lakes exceeding an empirically determined minimum area threshold ($x_{\min}$):

$$P(x) \propto x^{-\alpha}, x \geq x_{\min} \quad (3)$$

where $P(x)$ is the probability of observing a lake of area x and α is the scaling exponent describing the decrease in lake frequency with increasing size. The lower threshold $x_{\min}$ was estimated iteratively to maximize agreement between the observed CCDF and the fitted power law using the Kolmogorov–Smirnov (KS) statistic (Clauset et al., 2009). Only lakes with areas $\geq x_{\min}$ were included, ensuring that the exponent reflects the scaling of larger lakes rather than being biased by mapping or detection limits. This method enables the quantitative comparison of glacial lake size distributions across regions, lake types, and neighboring mountain ranges.

4.6.2 Temporal changes

Temporal changes between 2016 and 2024 were assessed by matching lakes from the two inventories using centroid proximity and percent spatial overlap in ArcGIS Pro. Each 2024 lake polygon was paired with the 2016 counterpart showing the highest intersection-over-union ratio, allowing identification of lakes that had disappeared, newly formed, merged, fragmented, or remained stable.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Relative area change was calculated as:

$$\text{Area change} = \frac{\text{Area}_{2024} - \text{Area}_{2016}}{\text{Area}_{2016}} \times 100 \quad (4)$$

Changes were classified into five categories: strong loss (≤−50 %), moderate loss (−50 to −10 %), stable (−10 to +10 %), moderate gain (+10 to +50 %), and strong gain (≥+50 %). This classification balances the representation of surface change while minimizing the influence of small delineation uncertainties. Spatial patterns of lake change were analyzed by region and elevation.

4.6.3 Lake classification

For both inventories, basic lake metrics were extracted from vector polygons, including surface area, perimeter, centroid coordinates, and elevation. To support a geomorphological classification, each lake was assigned to one of three categories based on its spatial and functional connection to glaciers (RGI v7.0):

(i) supraglacial lakes, located directly on glacier surfaces (lake centroid within the glacier polygon);
(ii) proglacial lakes, situated at the glacier front and in direct contact with the ice margin;
(iii) glacier-detached lakes, which include lakes formerly or indirectly influenced by glaciers but no longer in direct contact.

This simplified classification relies on the presence or absence of a hydrological or spatial connection to a glacier. We did not perform a detailed identification of damming mechanisms (e.g., moraine-, bedrock-, or landslide-dammed lakes) due to the insufficient 10 m resolution of Sentinel-2 imagery and the substantial manual effort required.

Summary statistics were compiled, including the number of lakes per region, total and mean lake areas, overall change in area between the study periods, and mean elevation (see Table 1). Detailed results, including spatial patterns and temporal changes, are presented in the following section.

5 Results

5.1 General characteristics of the 2024 glacial lake inventory

The 2024 glacial lake inventory represents the first high-resolution (10 m) polygon-based mapping of glacial lakes across Kyrgyzstan derived from Sentinel-2 imagery (Fig. 5). After filtering by minimum area and elevation thresholds, a total of 2 592 lakes were identified in 2024, covering 77.58 ± 5.09 km^2, with a national mean lake area of 29 930 ± 19 634 m^2.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

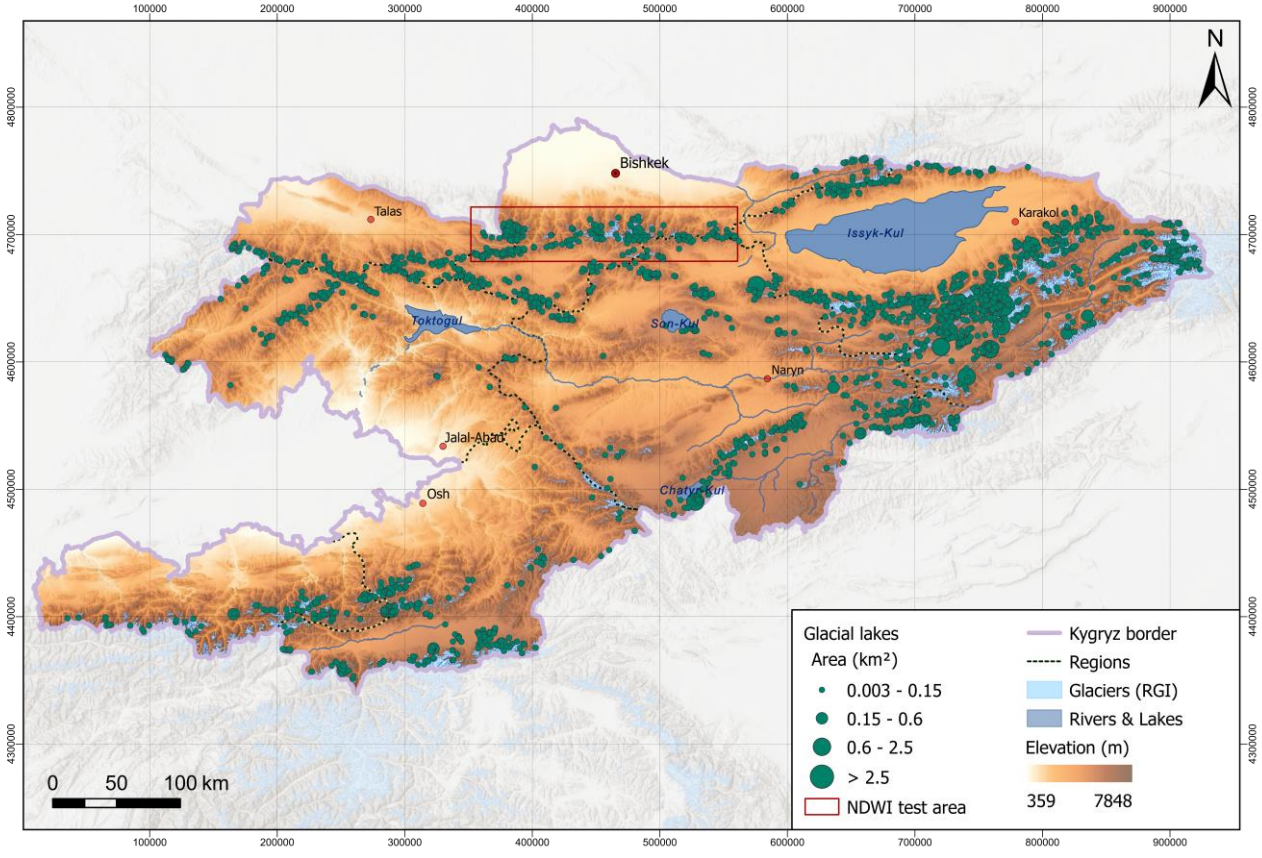


Figure 5. Distribution of glacial lakes in Kyrgyzstan in 2024. Hillshade and elevation map of Kyrgyzstan showing major mountain ranges, glaciers (RGI v7.0), main rivers, and lakes. Administrative regions are indicated for spatial reference. Elevation data are from the AW3D30 DEM. Base map: Esri World Hillshade (Esri, World Hillshade; see References). Sources: Esri, Vantor, Airbus DS, USGS, NGA, NASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA, Intermap, and the GIS user community | Powered by Esri.

The distribution of glacial lake areas is highly right-skewed, being dominated by small lakes. The median lake area in 2024 was 0.008 5 ± 0.000 08 km^2, with first and third quartiles of 0.004 97 ± 0.000 05 km^2 and 0.020 5 ± 0.000 19 km^2, respectively. Most lakes (91 %) were smaller than 0.05 km^2. Only 0.9 % exceeded 0.5 km^2, with the largest reaching 4.68 ± 0.04 km^2. The absolute number of lakes per size class highlights both common, small lakes and rare, large lakes (Fig. 6a), consistent with the CCDF of lake areas (Fig. 6b), which is 1 at the minimum mapped size (0.003 km^2) and decreases with lake area. For areas of the CCDF above the empirical threshold (x_{min} = 0.026 km^2; green dotted line, Fig. 6b), the distribution follows a power law with exponent α = 2.197 (red dashed line), consistent with neighboring ranges (Himalayas, $\alpha \approx 2.20$; Upper Indus Basin, $\alpha \approx 2.26$; Khan et al., 2025). The steeper slope compared to the Karakoram ($\alpha \approx 1.9$), a nearby major Central Asian range known for its extensive glaciation, indicates a faster decline in large lake frequency (Khan et al., 2025). Lake elevations range from 3 000 m (i.e. the minimum elevation threshold) to 4 810 m a.s.l. The elevation histogram (Fig. 6c) reveals that most lakes occur between ~3 500 and 3 800 m a.s.l., concentrated between the first (Q1 = 3 457 m) and third

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

quartiles (Q3 = 3 822 m) with a median elevation of 3 636 m. The distribution is approximately unimodal, with a slight asymmetry toward higher elevations, reflecting the dominant occurrence of lakes within the mid- to upper-mountain zone of Kyrgyzstan.

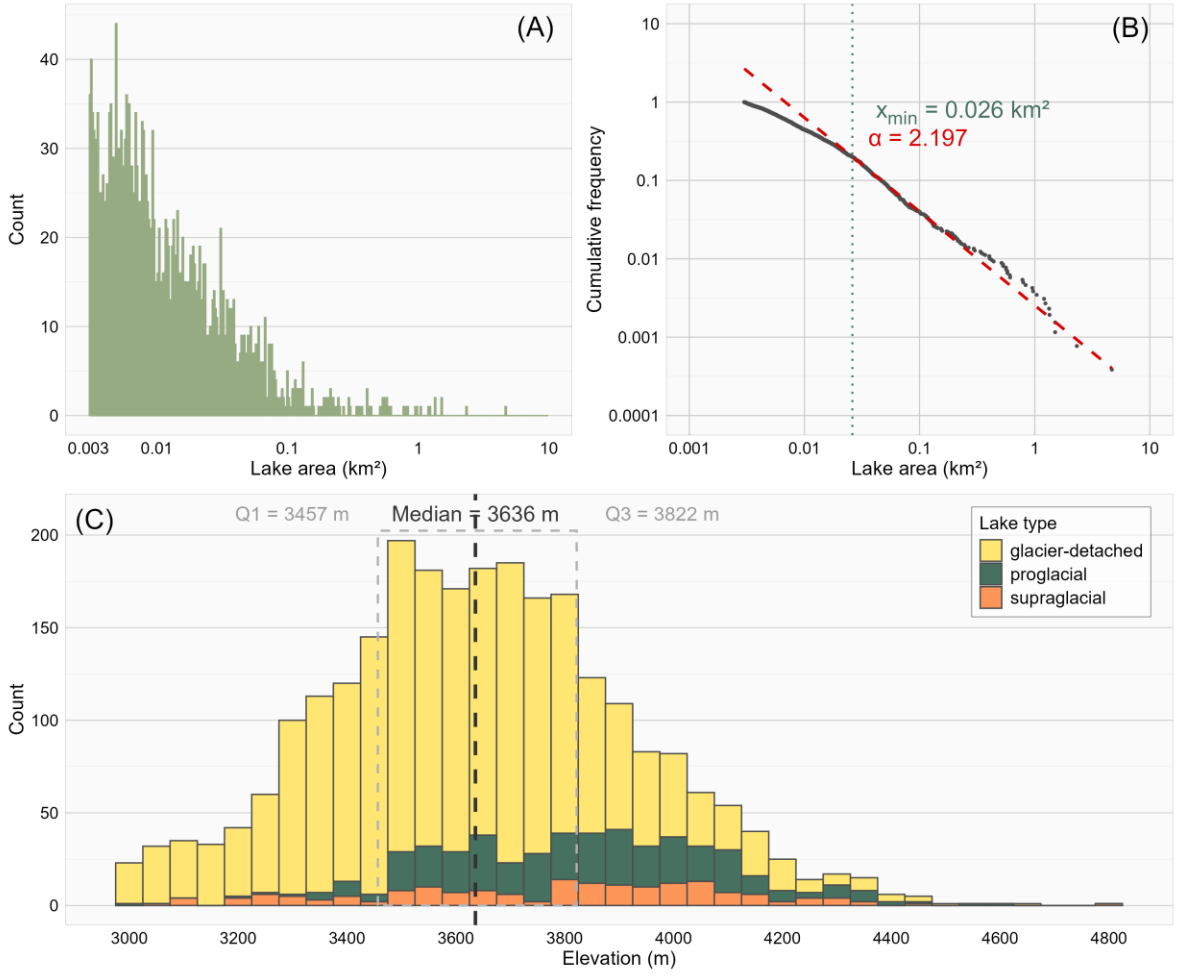


Figure 6. Distribution of glacial lakes in Kyrgyzstan in 2024. (a) Histogram of lake occurrence as a function of area (log scale on *x*-axis). (b) Log-log plot of the complementary cumulative distribution function of lake areas, showing the empirical data (black), the fitted power-law model (red dashed line), and the lower cutoff x_{min} (vertical green dotted line). (c) Elevation histogram with the interquartile range (Q1–Q3) indicated by the dashed gray rectangle, and the median by the black dashed line.

Regionally, Issyk-Kul hosts the largest number of glacial lakes (1 123), accounting for ~43 % of the national total, followed by Naryn (448 lakes, ~17 %) (Fig. 7a). In contrast, Batken and Jalal-Abad contain fewer lakes (96 and 152, respectively) but exhibit high mean lake areas (26 665 ± 1 948 m^2 and 20 085 ± 1 863 m^2), reflecting the presence of larger proglacial or moraine-dammed lakes (Fig. 7b). Chuy and Osh display the highest lake densities (63.43 and 57.19 lakes per 1 000 km^2, respectively), whereas Issyk-Kul shows the lowest density (52.58 lakes per 1 000 km^2) due to the dominant coverage of the main lake, Lake Issyk-Kul.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

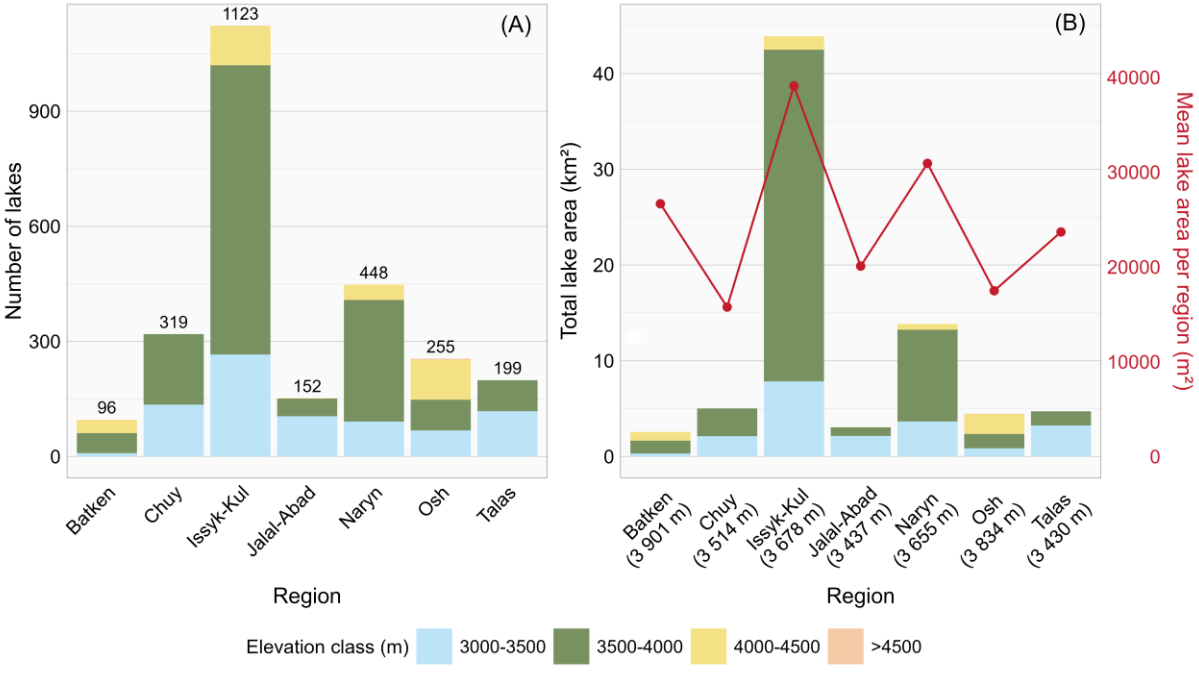


Figure 7. (a) Number of glacial lakes and (b) total lake area in km² per region in Kyrgyzstan for 2024, classified by elevation. The secondary red axis shows the mean lake area per region (m²). Total lake counts are annotated above the bars in panel (a).

Inventoried lakes were classified as supraglacial, proglacial, or glacier-detached (Fig. 8a). Glacier-connected lakes (supraglacial and proglacial) are likely still influenced by glaciers, whereas glacier-detached lakes are no longer directly affected by glacier ice or meltwater. Nationwide, supraglacial lakes account for 170 lakes (6.6 %), proglacial lakes for 366 lakes (14.1 %), and glacier-detached lakes for 2 056 lakes (79.3 %) (Fig. 8b). The distribution of lakes varies across regions, reflecting glacier coverage, regional elevation, and past glacier dynamics. Issyk-Kul contains the largest number of glacier-connected lakes (271), consistent with its high glacier fraction (9.34 %). Batken and Osh have similar glacier coverage (~5 %) and comparable densities of glacier-connected lakes (0.026–0.027 lakes/km² glacier). Despite their different regional sizes, this result suggests that glacier size rather than regional area primarily limits lake formation. The lower-elevation regions Talas (3 430 m) and Jalal-Abad (3 437 m) host very few glaciers (0.9 % and 0.4 %) and mostly glacier-detached lakes (~84 %). Nonetheless, Talas maintains a relatively high density of lakes per area (0.017 lakes/km²), whereas Jalal-Abad shows a much lower density (0.005 lakes/km²). Higher regions such as Naryn (3 655 m) and Chuy (3 514 m) also have high proportions of glacier-detached lakes (~80 %) but retain a moderate density of total glacier lakes. Figure 8c presents the total lake area by type and region. Glacier-connected lakes generally occur at higher elevations (Fig. 6c): supraglacial and proglacial lakes occur at mean elevations of ~3 780 m and ~3 825 m, respectively, whereas glacier-detached lakes occur at a lower mean elevation of ~3 600 m. Proglacial lakes are largest (~0.046 ± 0.002 3 km²), followed by glacier-detached (~0.029 ± 0.001 9 km²) and supraglacial lakes (~0.010 ± 0.001 3 km²).

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

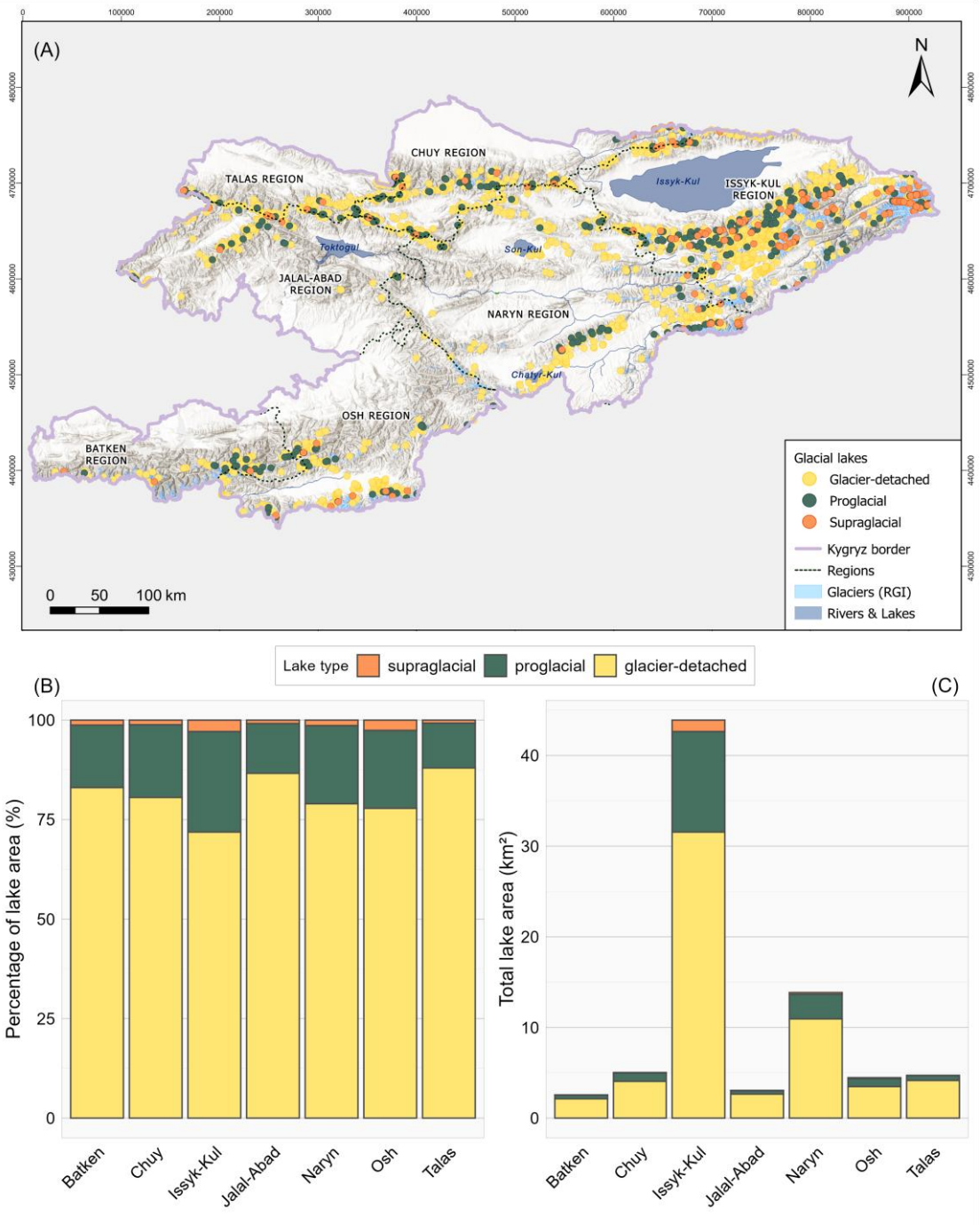


Figure 8. Spatial and quantitative distribution of glacial lakes in Kyrgyzstan. (a) Hillshade map showing the locations of supraglacial, proglacial, and glacier-detached lakes. (b) Stacked bar chart showing the proportion (%) of lake area by type within each administrative region. (c) Bar chart of total lake area (km^2) by lake type per region. Base map: Esri World Hillshade (Esri, World Hillshade; see References). Sources: Esri, Vantor, Airbus DS, USGS, NGA, NASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA, Intermap, and the GIS user community | Powered by Esri.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

5.2 Temporal evolution of glacial lakes between 2016 and 2024

Comparison between the 2016 and 2024 inventories highlights temporal changes in lake number, area, and regional distribution (Fig. 9). A total of 2 345 lakes were mapped in 2016, compared to 2 592 in 2024, representing a net increase of 10.5 % in lake count. Total lake area increased from 71.34 ± 5.49 km^2 to 77.58 ± 5.09 km^2 (+8.74 %), whereas mean lake area slightly decreased, reflecting the formation of numerous small lakes. Elevation patterns remained largely stable, with most lakes situated between ~3 500 and 3 800 m a.s.l.; the national mean elevation slightly increased from 3 628 m to 3 644 m, likely associated with high-altitude supraglacial and proglacial lake formation. The largest relative increases in lake number and area occurred above ~4 000 m a.s.l., consistent with documented upward shifts of the summer snowline/equilibrium line altitude (ELA) in the Tien Shan (~3 800–4 100 m a.s.l.) and widespread glacier retreat in recent decades (Sorg et al., 2012; Kapitsa et al., 2020).

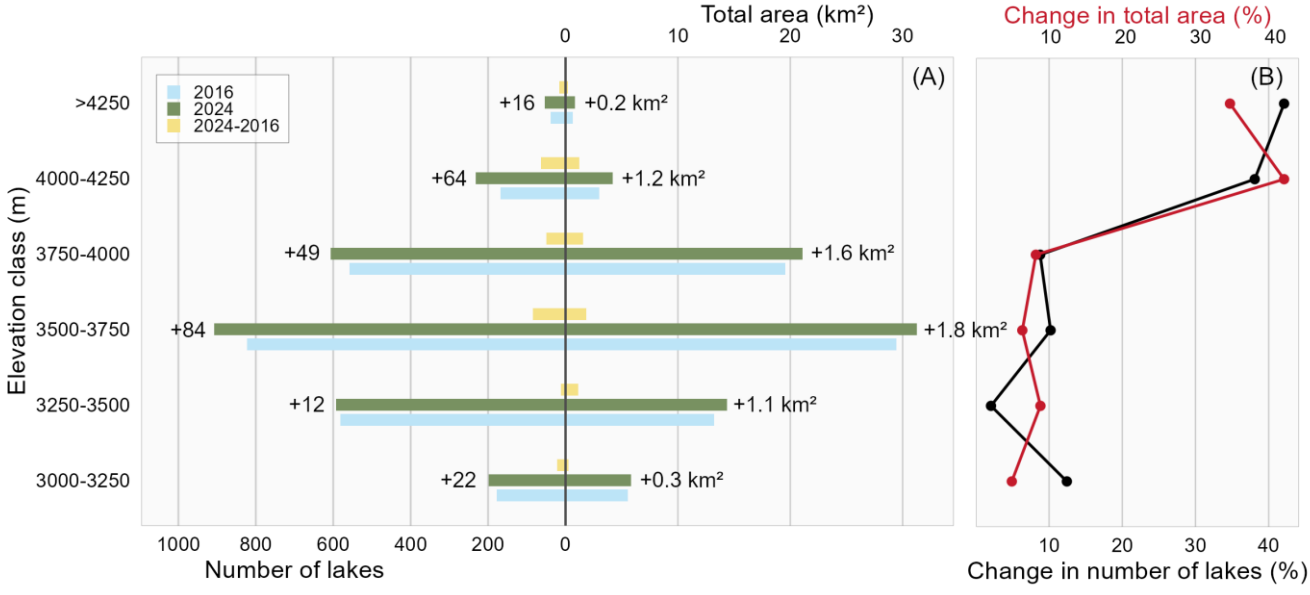


Figure 9. Distribution and change of glacial lakes by elevation class between 2016 and 2024. (a) Number of lakes (primary *x*-axis) and total lake area in km^2 (secondary *x*-axis) for each elevation class in 2016 (blue), 2024 (green), and their difference (yellow). Absolute changes in lake number and total area (km^2) are written in black. (b) Relative changes expressed as percentages for both lake number (black, primary *x*-axis) and total lake area (red, secondary *x*-axis) across the same elevation classes. Elevation classes are left-inclusive and right-exclusive.

Table 1 summarizes regional statistics. At the regional scale, Jalal-Abad, Talas, and Chuy exhibited the strongest relative increases in lake number (+19–25 %), whereas Osh was the only region to experience a decline in lake number (−5.6 %) despite a moderate gain in lake area, suggesting that existing water bodies expanded or coalesced. Issyk-Kul, Naryn, and Batken showed smaller increases in lake number (+7–11 %). Mean lake area slightly declined in all regions except Osh, where it increased markedly. Lake density patterns mirror these trends, with Chuy and Osh exhibiting the highest densities (~60 lakes per 1 000 km^2) and Issyk-Kul the lowest (~25 lakes per 1 000 km^2), mainly due to the coverage of Lake Issyk-Kul. Changes

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

in lake size distributions indicate that the smallest lakes (<0.05 km^2) experienced the largest relative gains, reflecting ongoing glacier retreat and the formation of new proglacial lakes. Regions with many lakes (e.g., Naryn, Issyk-Kul) continue to host larger lakes on average, whereas regions with fewer lakes (e.g., Chuy, Talas) remain dominated by smaller lakes.

Table 1. Regional summary of glacial lake statistics in Kyrgyzstan between 2016 and 2024. The Δ Area [km^2] (Δ Number [%]) column reports absolute changes in total lake area and relative changes in lake number between 2016 and 2024. The absolute values represent the difference in mapped area between 2016 and 2024, and the percentages indicate changes in lake count relative to 2016.

Region	Number of lakes		Area [km^2]		Δ Area [km^2] (Δ Number [%])	Mean lake area [m^2]		Mean elevation [m]		Density [lakes per 10^3 km^2]	
-	*2016*	*2024*	*2016*	*2024*	-	*2016*	*2024*	*2016*	*2024*	*2016*	*2024*
Batken	86	96	2.40 ± 0.21	2.56 ± 0.19	0.16 (+11.63 %)	27 929 ± 2 396	26 665 ± 1 948	3 880	3 901	35.80	37.50
Chuy	268	319	4.42 ± 0.52	5.03 ± 0.53	0.61 (+19.03 %)	16 487 ± 1 928	15 765 ± 1 673	3 512	3 513	60.65	63.43
Issyk-Kul	1 017	1 123	40.54 ± 2.66	43.91 ± 2.41	3.36 (+10.42 %)	39 865 ± 2 616	39 099 ± 2 142	3 670	3 678	25.08	25.58
Jalal-Abad	122	152	2.65 ± 0.27	3.05 ± 0.28	0.41 (+24.59 %)	21 690 ± 2 235	20 085 ± 1 863	3 437	3 437	46.10	49.79
Naryn	420	448	13.15 ± 0.98	13.85 ± 0.87	0.70 (+6.67 %)	31 314 ± 2 338	30 925 ± 1 936	3 655	3 654	31.93	32.34
Osh	270	255	4.09 ± 0.48	4.46 ± 0.42	0.37 (−5.56 %)	15 142 ± 1 795	17 485 ± 1 657	3 669	3 833	66.04	57.19
Talas	162	199	4.09 ± 0.37	4.71 ± 0.39	0.62 (+22.84 %)	25 262 ± 2 268	23 687 ± 1 962	3 427	3 429	39.59	42.22
Total	2 345	2 592	71.34 ± 5.49	77.58 ± 5.09	6.23 (+10.53 %)	30 423 ± 2 341	29 929 ± 1 963	3 628	3 644	32.87	33.41

Quantitative analysis of the 3 097 glacial lakes tracked between 2016 and 2024 reveals clear spatial and statistical trends across Kyrgyzstan. Over half of all lakes (54 %) remained stable over the study period, whereas 23.9 % newly appeared and 15 % disappeared. Fragmentation and fusion represent minor processes (5.3 % and 1.9 %, respectively). The total of 3 097 tracked lakes exceeds the total number of lakes counted in the 2016 or 2024 inventories because it includes all lakes that were stable, emerged, vanished, or underwent fragmentation/fusion, providing a comprehensive overview of lake dynamics over the eight-year period. Newly emerged lakes occurred at slightly higher elevations (mean 3 676 m a.s.l.) than those that disappeared (3 614 m) or remained stable (3 622 m), confirming high-altitude environments as the main formation areas, particularly in the Kyrgyz Range (Chuy) and Terskey Mountains (Issyk-Kul) (Fig. 10a). Elevation shows a weak Spearman correlation with area change (ρ = 0.096), suggesting that elevation has minimal influence on the magnitude of lake expansion. Among classified lakes, 27.4 % experienced moderate or strong gains, compared to 13.1 % with moderate or strong losses. Stable lakes accounted for about 20.7 %, whereas 38.8 % could not be classified, i.e. either newly emerged or vanished. These results highlight that, overall, expansion processes dominate over shrinkage or disappearance.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Marked regional contrasts are observed. Newly appeared lakes are most frequent in Batken (25.9 %) and Issyk-Kul (25.3 %), the latter hosting numerous small proglacial lakes in the Terskey Range. In contrast, the largest proportion of lakes vanished in Osh (27.6 %), likely due to drainage, merging, or hydrological instabilities in steep glacial valleys. A moderate proportion of lakes also vanished in Naryn (16.4 %). Other areas in which lakes vanished included the eastern border with China and south-southwest of Lake Issyk-Kul (Fig. 10a), whereas lakes that emerged near Lake Issyk-Kul were mainly concentrated southeast of the lake, toward Karakol. Furthermore, gains in individual lake area – including both moderate and strong gain categories – dominated most regions, peaking in Talas (54.3 %) and Batken (55.7 %), followed by Chuy, Jalal-Abad, Osh, Issyk-Kul, and Naryn (49.6–40.4 %) (Fig. 10b). Stable lakes remained important, especially in Jalal-Abad (47.0 %). Losses in individual lake area were limited and mostly affected small or transient water bodies due to coalescence, partial drainage, or local hydrological changes, with the largest proportions observed in Osh (30.5 %), Naryn (27 %), and Issyk-Kul (25.3 %) (Fig. 10b).

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

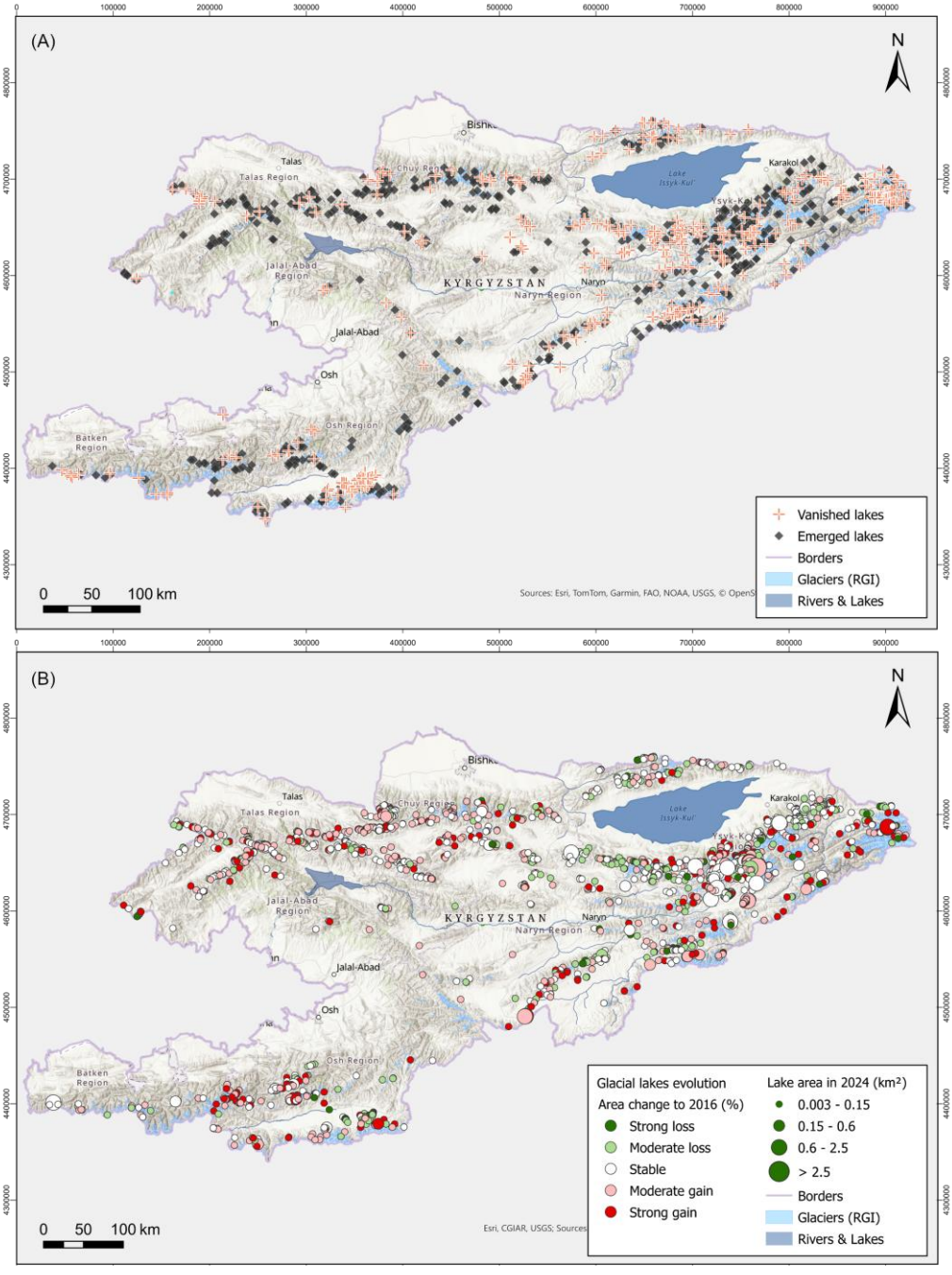


Figure 10. Temporal changes of glacial lakes in Kyrgyzstan between 2016 and 2024. (a) Spatial distribution of lakes that emerged (black diamonds) or vanished (orange crosses). (b) Percentage change in lake area relative to 2016, classified into five categories from strong loss (green) to strong gain (red). Circle sizes correspond to lake surface area in 2024 (km²), illustrating both the magnitude of area changes and absolute lake size. Base map: Esri World Hillshade (Esri, World Hillshade; see References). Sources: Esri, Vantor, Airbus DS, USGS, NGA, NASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA, Intermap, and the GIS user community | Powered by Esri.

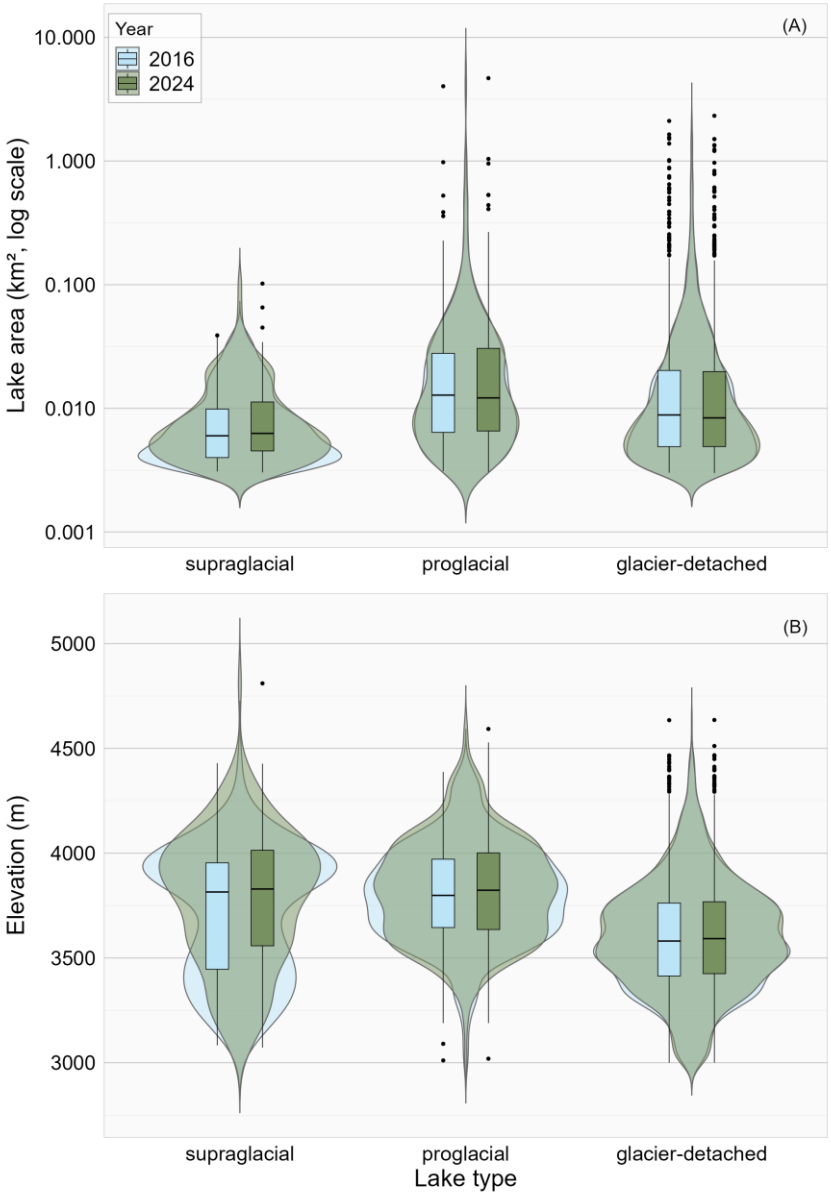
https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

5.3 Lake type and elevation distributions

Building on the 2024 inventory described above, Fig. 9 illustrates temporal changes in lake size and elevation between 2016 and 2024. The violin plots in Fig. 11 highlight consistent contrasts among lake categories and reveal subtle shifts over time. The clearest case of lake evolution is that of supraglacial lakes: the lower range of their area distribution (0.005–0.007 km^2) was truncated, indicating a reduction in the number of small supraglacial lakes within the study period. Supraglacial lakes remain the smallest, with median areas increasing slightly from 0.006 0 ± 0.001 2 km^2 in 2016 to 0.006 3 ± 0.001 1 km^2 in 2024, and are located at the highest elevations (median 3 815–3 830 m). Their elevation distribution also evolved, shifting from a bimodal profile in 2016 to a more uniform shape in 2024. This reflects a relative depletion of lakes between 3 000 and 3 500 m, accompanied by increased lake density around 3 550–3 800 m and decreased lake density above 4 050 m. Overall, this indicates a gradual upward shift and homogenization of supraglacial lake elevations. This pattern is consistent with the regional rise of the ELA and the upward migration of glacier ablation zones documented across the Tien Shan (Sorg et al., 2012; Kapitsa et al., 2020). The relative depletion of lower-elevation supraglacial lakes and the emergence of new ones at higher altitudes directly mirror the ongoing lowering of glacier surfaces and the transition of meltwater storage toward higher-elevation ice surfaces.

In contrast, proglacial and glacier-detached lakes maintained broadly stable distributions. A slight narrowing of proglacial lake elevations is observed between 3 600 and 4 000 m, whereas their area distribution remained largely unchanged. Proglacial lakes showed slightly larger median areas than supraglacial lakes (~0.013 ± 0.001 6 km^2) and occur at elevations around ~3 820 m. Glacier-detached lakes are generally found at lower altitudes (median ~3 590 m) and display minimal changes in their median values between 2016 and 2024. Overall, this comparison indicates that the lake size and elevation distributions remained largely stable for proglacial and glacier-detached categories, whereas the most noticeable changes affected supraglacial lakes, reflecting reductions in small lake abundance and an upward shift in elevation, consistent with the results in Fig. 9b.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026



475

Figure 11. Violin plots showing the distribution of (a) lake area (log scale) and (b) elevation by lake type for 2016–2017 (blue) and 2022–2024 (green). The boxplots embedded within each violin represent the interquartile range and median values. Violin shapes indicate kernel density estimates.

5.4 Validation and error assessment

480 Our internal validation provided contrasting insights into the performance of the semi-automated NDWI-based (NDWI_2024) and Random Forest Classifier (RFC_2016) workflows. Performance metrics (precision, recall, and F1 score) were derived from confusion matrices computed from the intersection between predicted and reference lake polygons. True positives (TP)

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

correspond to correctly detected lakes, false positives (FP) to erroneously segmented water bodies, and false negatives (FN) to missed lakes that were present in the reference dataset (Table 2). Precision indicates the proportion of detected lakes or pixels that correctly match reference lakes (low FP), recall measures the proportion of reference lakes or pixels that are captured (low FN), and F1 provides a balanced measure of both.

Table 2. Detection performance of NDWI_2024 and RFC_2016 glacial lake inventories compared to manual references. TP, true positives; FP, false positives; FN, false negatives (all in pixels). N_ref, number of lakes in the reference inventory; N_cmp, number of lakes in the compared inventory; N_intersect, number of correctly detected lakes (polygon level). Precision, Recall, and F1 score are reported at the polygon level.

Inventory		*Pixel-level metrics*			*Polygon-level metrics*					
Reference	**Compared**	**TP**	**FP**	**FN**	**N_ref**	**N_cmp**	**N_intersect**	**Precision**	**Recall**	**F1**
Manual_2024	NDWI_2024	59 617	532 344	26 453	2 592	5 002	1 517	0.10	0.69	0.18
Manual_2016	RFC_2016	74 452	3 297	4 740	2 345	2 103	2 071	0.96	0.94	0.95

The semi-automatic NDWI-based extraction for 2024, when compared to the manually corrected reference inventory of 2 592 lakes, shows that 1 517 lakes were successfully detected among 5 002 polygons, corresponding to a recall of 69 %, precision of 10 %, and an overall F1 score of 0.18. Performance was strongly size-dependent: for small lakes (<0.05 km^2), we obtained moderate recall (31 %) and low precision (19 %), whereas for medium (0.05–0.5 km^2) and large lakes (>0.5 km^2), we achieved higher recall (61–75 %) but suffered from severe over-segmentation (precision of only 7 % for large lakes). These results highlight the intrinsic limitations of single-index thresholding under complex spectral and topographic conditions. As noted by Shugar et al. (2020), glacial lakes may display highly variable spectral signatures depending on suspended sediment content, brash ice, and turbid water, and shadows or terrain effects can further reduce detection accuracy. Adjusting thresholds to capture faint water bodies inevitably increases false positives, whereas stricter thresholds lead to the omission of true lakes. Consequently, extensive manual correction was indispensable to achieve a reliable 2024 inventory.

In contrast, the Random Forest classification applied to the 2016 imagery performed remarkably well when evaluated against the manually delineated inventory. Of 2 345 reference lakes, 2 071 were correctly identified among 2 103 predicted polygons, yielding a precision of 95.8 %, a recall of 94.0 %, and an F1 score of 0.95. The model maintained robust performance across all lake-size classes (F1 = 0.92–0.97), indicating that training on a high-quality manual dataset from 2024 successfully captured spectral and morphological signatures of glacial lakes that were transferable to the earlier imagery. This result validates the use of a manually refined reference inventory as an effective training base for temporal extrapolation of glacial lake distributions.

Additionally, the spatial uncertainty of lake boundaries was also quantified using Eq. 2. Across the 2024 manually corrected inventory, the mean boundary uncertainty was 0.001 96 km^2, with a median of 0.001 33 km^2. For the RFC_2016 inventory, the mean and median uncertainties were slightly higher, at 0.002 34 km^2 and 0.001 58 km^2, respectively. Small

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

lakes (<0.05 km^2) exhibited larger relative uncertainties due to their higher perimeter-to-area ratios, whereas medium and large lakes generally showed lower per-lake errors. These results indicate that even though most lakes were accurately delineated, minor boundary deviations persist, particularly for small and/or complexly shaped lakes. Graphs illustrating the spatial uncertainties on lake boundaries by size class are provided in the Supplementary Material.

Overall, these tests demonstrate that although single-index thresholding can support preliminary detection, it lacks the reliability required for inventory production without manual editing in cases of spectral variability, shadowing, and complex water boundaries. Conversely, machine-learning classification, when trained on a precise and representative dataset, provides accurate and internally consistent results suitable for scientific analysis of multi-temporal glacial lake evolution, with both high detection performance and quantified spatial precision of lake outlines.

6 Discussion

6.1 Evolution of glacial lakes in Kyrgyzstan between 2016 and 2024

The comparison of the 2016 and 2024 inventories reveals clear trends in the evolution of glacial lakes across Kyrgyzstan. Overall, lake numbers and total area increased, driven primarily by the emergence of small, high-altitude proglacial and supraglacial lakes, whereas larger, lower-elevation lakes remained largely stable. This reflects the progressive development of new depressions exposed by glacier retreat, particularly in the Terskey and Kyrgyz ranges. The spatial distribution of these changes is highly heterogeneous. Regions with extensive glacier coverage, such as Issyk-Kul, continue to host numerous small glacier-connected lakes, further highlighting the role of glacier retreat in creating new depressions that accumulate meltwater. Interestingly, Jalal-Abad exhibited the largest relative increase in lake number (+24.6 %) despite having the lowest glacier density and the smallest mean glacier size (Table X), consistent with the expectation that smaller glaciers tend to react earlier and faster to climate signal. New lakes consistently formed in high-elevation environments (Fig. 10a), whereas losses were predominantly observed in small, transient water bodies located in steep valleys or hydrologically unstable terrain, as in Osh and Naryn (Fig. 10b).

Trends specific to lake type are also apparent. Supraglacial lakes showed slight increases in median area and an upward shift in elevation, consistent with the retreat of ice surfaces. Proglacial lakes remained stable in both size and elevation, likely because gains from supraglacial-to-proglacial transitions were broadly balanced by losses as some proglacial lakes became glacier-detached (Fig. 11). Glacier-detached lakes, disconnected from contemporary glacier influence, remained mostly stable. This highlights that glacier retreat primarily drives the formation of new high-altitude lakes rather than causing widespread changes in existing detached lakes. In general, expansion dominated over shrinkage, suggesting that ongoing glacier retreat continues to foster the formation of new lakes in supraglacial and proglacial settings, and that losses are spatially localized. These patterns indicate that glacial lake evolution in Kyrgyzstan is controlled by the interplay of glacier retreat, topography, and local hydrology (Daiyrov et al., 2022). High-elevation proglacial areas remain the primary sites of new lake formation until topography becomes the key limitation, as observed for example in the Tropical Andes where glaciers have lost more

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

than 50 % of their LIA area and the topographic potential for new lake formation is now fairly limited (Emmer et al., 2020). In contrast, larger glacier-detached lakes persist in lower-elevation or more stable terrain. This narrative aligns with broader observations in glaciated mountain regions, where glacier retreat preferentially triggers the formation of small, high-elevation lakes while leaving larger, low-elevation reservoirs relatively unaffected (e.g., Wang et al., 2020; Izagirre et al., 2025).

6.2 Comparison with global datasets

To evaluate the completeness and spatial reliability of our final inventories, the manually refined lake datasets for 2016 and 2024 were compared with three widely used regional and global references (Fig. 12):

(1) GIGLak (Song et al., 2025), a globally consistent 30 m Landsat-based inventory (1984–2020);

(2) Hi-MAG (Chen et al., 2020), an annual 30 m glacial lake product for High Mountain Asia (2008–2017), from which we considered the 2017 layer for comparison; and

(3) Wang et al. (2020), a manually derived 30 m HMA-wide inventory constructed from 668 Landsat scenes (1990 and 2018) by visual interpretation within 10 km of glaciers.

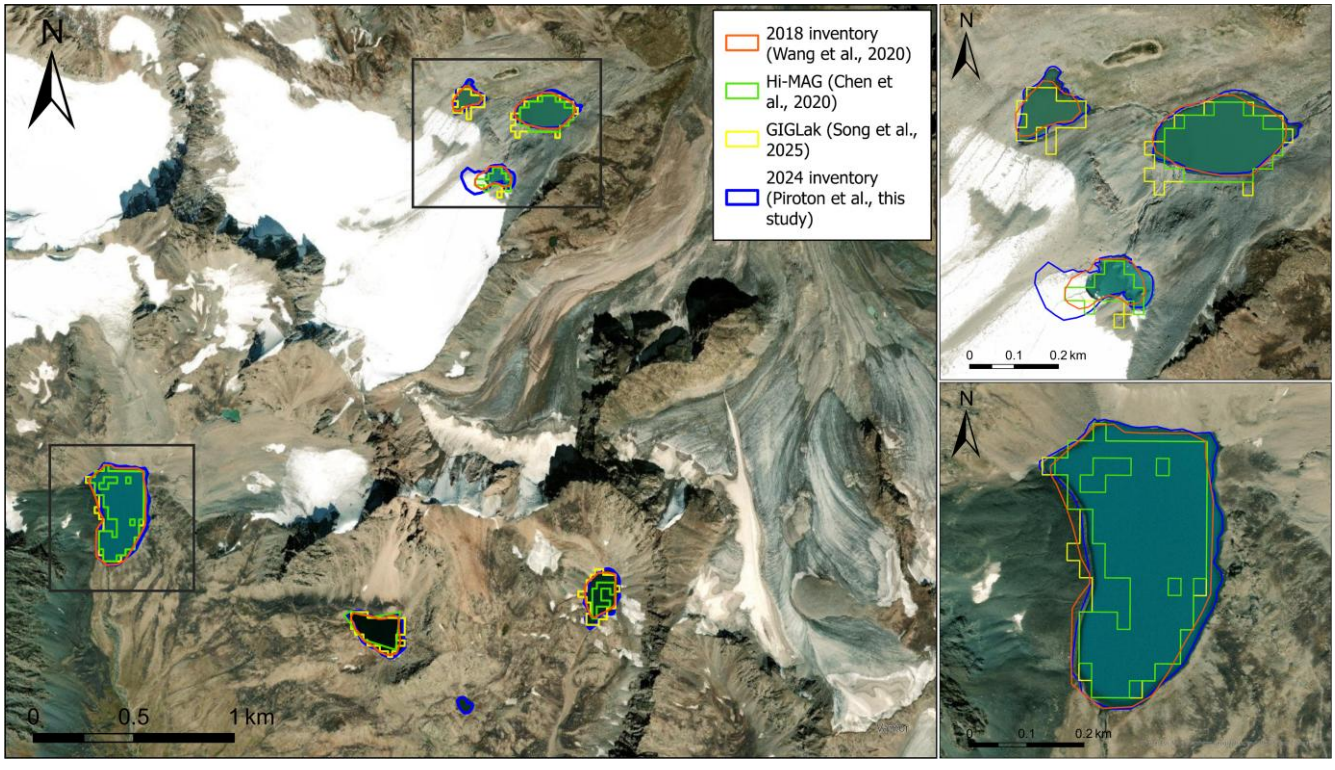


Figure 12. Comparison between the 2024 glacial lake inventory produced in this study (blue) and three global reference inventories: Wang et al. (2020, orange), the Hi-MAG inventory of Chen et al. (2020, green), and the GIGLak inventory of Song et al. (2025, yellow). The lake outlines generated in this study show closer agreement with current lake extents, whereas several small lakes were not captured in the global datasets. The background image corresponds to the Esri Imagery basemap as of 30 August 2014 and may not fully reflect the most recent conditions. Basemap: Esri World Imagery. Source: Esri, Maxar, Earthstar Geographics, and the GIS User Community | Powered by Esri.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Table 3 summarizes both pixel-level metrics (true positives, false positives, false negatives) and polygon-level metrics (number of intersecting lakes, proportion of reference lakes captured, proportion of compared lakes matched, precision, recall, F1 score).

Table 3. Comparison of our manually refined 2016 and 2024 glacial lake inventories against GIGLak, Hi-MAG, and Wang et al. (2020), showing pixel-level (TP, FP, FN) and polygon-level metrics (lake count, intersecting lakes, % reference covered, % compared matched, precision, recall, F1 score).

Inventory		*Pixel-level metrics*			*Polygon-level metrics*							
Reference	**Compared**	**TP pixels**	**FP pixels**	**FN pixels**	**N_ref**	**N_cmp**	**N_intersect**	**% of reference covered**	**% of compared matched**	**Precision**	**Recall**	**F1**
Manual_2024	GIGLak	46 238	26 682	40 316	2 592	1 342	1 032	52.9 %	62.5 %	0.63	0.53	0.58
	Hi-MAG	50 198	8 826	36 356	2 592	716	677	57.8 %	84.4 %	0.85	0.58	0.69
	Wang, 2020	58 116	14 004	28 438	2 592	1 041	916	67.4 %	80.2 %	0.81	0.67	0.73
Manual_2016	GIGLak	44 849	28 146	34 453	2 345	1 342	1 060	56.7 %	61.6 %	0.61	0.57	0.59
	Hi-MAG	50 028	8 910	29 164	2 345	716	702	63.2 %	84.8 %	0.85	0.63	0.72
	Wang, 2020	58 109	14 493	21 083	2 345	1 041	937	73.4 %	80.2 %	0.80	0.73	0.77

This comparison highlights four key points. First, both of our manually refined inventories detected substantially more lakes than the global or HMA-wide products, particularly small, high-altitude, and/or complexly shaped lakes. This is reflected by the high precision and recall values (up to 0.85 and 0.73, respectively), high percentages of reference lakes covered, and high percentages of compared lakes matched. High TP counts and moderate FP/FN counts further confirm that the manual approach effectively identifies true lakes while minimizing omissions and false detections.

Second, some discrepancies are attributable to temporal mismatches: GIGLak represents multi-year Landsat composites, Hi-MAG corresponds to 2017, and Wang et al. (2020) to 2018, whereas our Kyrgyz inventories refer specifically to 2016 and 2024. Rapid lake evolution—emergence, expansion, or drainage—makes perfect agreement impossible. Comparing both years provides insight into temporal variability and highlights the importance of frequent, high-resolution regional mapping.

Third, the use of Sentinel-2 imagery at 10 m resolution, compared to 30 m Landsat imagery, allows more precise delineation of small and/or narrow lakes. This improves polygon-level metrics, increases TP pixels, reduces FN pixels, and enhances overall precision, recall, and F1 scores.

Finally, manual refinement of lake boundaries, conducted through systematic inspection of the entire area rather than solely verifying pre-identified lakes (as in some previous HMA inventories), enables the detection of previously missing lakes and the removal of false positives. This approach ensures that both the 2016 and 2024 inventories capture small, complexly

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

shaped, and/or high-altitude lakes that are often omitted in global datasets, achieving higher precision and recall. The results demonstrate that careful manual editing combined with high-resolution imagery provides a more complete and spatially consistent inventory, albeit limited to Kyrgyzstan here, whereas the global products offer wider geographic coverage.

Overall, the manually refined 2016 and 2024 inventories constitute a robust reference for national- and sub-regional-scale glacial lake monitoring, with polygon metrics confirming the high reliability and spatial completeness of the dataset.

6.3 Advantages and limitations of the dataset

The inventories produced in this study provide a high-resolution, spatially consistent reference for glacial lakes across Kyrgyzstan, combining semi-automated detection, machine-learning classification, and systematic manual refinement. The 10 m spatial resolution enables the detection of small (≥0.003 km^2), complexly shaped, and/or high-altitude lakes often omitted from global datasets (Sattar et al., 2025). Consistency between the 2016 and 2024 inventories allows robust analysis of lake evolution, and the quantified spatial uncertainties on lake boundaries provide confidence intervals for area estimates, supporting hydrological modeling, hazard assessment, and early-warning applications. Compared to global inventories, these datasets reduce omission errors, improve the detection of previously unmapped lakes, and provide valuable training or validation data for future machine-learning applications. Importantly, they will be shared with the Ministry of Emergency Situations of the Kyrgyz Republic to support monitoring and management of glacial lakes.

Nonetheless, several limitations remain. Reliance on optical Sentinel-2 imagery makes detection sensitive to clouds, shadows, ice, or debris cover. Temporal extrapolation assumes stable spectral–morphological lake signatures, which may vary with turbidity, sediment load, or seasonal ice, potentially reducing classification accuracy (Shugar et al., 2020; Zhang et al., 2024). Very small lakes, narrow riverine features, or rapidly changing water bodies may still be missed or misclassified. Although manual correction improves accuracy, it is time-consuming and may limit scalability. Finally, the dataset does not yet integrate radar or multi-sensor imagery, which could help mitigate persistent cloud or snow cover.

6.4 Perspectives for future monitoring

The 2016 and 2024 inventories highlight several directions for future work. A next key step involves integrating lake evolution with glacier dynamics, climatic forcing, and geomorphological setting to better understand the processes controlling lake formation and disappearance. Linking these inventories with temperature and precipitation trends, glacier mass balance datasets, and permafrost indicators would allow quantitative assessments of how cryospheric changes influence lake development and regional hydrology. Beyond detection, the characterization of lake-damming types (moraine, ice, bedrock, or landslide) remains a crucial objective for assessing stability and potential GLOF susceptibility. Such analyses require higher-resolution optical and radar data (e.g., PlanetScope, WorldView, Sentinel-1), as well as field verification, which are challenging to implement over large national extents but essential for targeted hazard assessments.

Although hazard and risk evaluation were not the primary objectives of this study, the inventories established here provide the foundation for future GLOF risk modeling and mitigation planning. Integrating lake dynamics with topographic and

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

downstream exposure data could support regional early-warning systems and guide infrastructure development in vulnerable valleys. Continued monitoring using higher-resolution, multi-sensor observations and (semi-)automated workflows will be critical for understanding the evolving cryosphere–hydrosphere interactions in Central Asia and for translating scientific insights into effective risk management strategies (Emmer, 2024; Wang et al., 2025).

7 Conclusions

This study provides a comprehensive, high-resolution (10 m) mapping and temporal analysis of glacial lakes across Kyrgyzstan, combining semi-automated NDWI-based delineation with machine learning-based reconstruction. Beyond simple mapping, it provides new insights into the spatial, typological, and temporal evolution of glacial lakes in a rapidly changing cryospheric environment. Key findings and perspectives are summarized below.

Inventory completeness and accuracy: The 2024 inventory represents the most detailed and consistent dataset to date, with 2 592 glacial lakes covering 77.6 km^2. Its strength lies in the manual correction of thousands of small lakes that automated methods often fail to detect or accurately delineate due to topographic shadows, debris cover, and seasonal variability. The Random Forest model trained on this dataset successfully reconstructed the 2016 lake distribution, confirming that a carefully curated reference inventory enables robust temporal transfer of classification models. In contrast, single-index NDWI detection alone required extensive manual refinement due to spectral variability and topographic effects.

Temporal trends: Between 2016 and 2024, the total number and area of lakes increased by 10.5 % and 8.7 %, respectively. Growth was dominated by the formation of new, small lakes at high elevations, particularly in the Terskey and Kyrgyz ranges. These small lakes, though individually minor, represent the most dynamic and numerous components of Kyrgyzstan's glacial lake system and thus play a key role in early-stage glacier–hydrology interactions. The largest relative increases occurred above ~4 000 m a.s.l., probably reflecting the regional rise of the ELA. Overall, lake expansion exceeded shrinkage during the study period, reflecting ongoing glacier retreat.

Spatial and typological patterns: The lake distribution remains strongly elevation-dependent. Glacier-connected lakes (supraglacial, proglacial) prevail above 3 500 m, whereas larger glacier-detached lakes dominate at lower altitudes. The subtle increase in supraglacial lake size and upward shift in elevation mirror the lowering of glacier surfaces and glacial retreat. Such changes are consistent with regional deglaciation trends and highlight the sensitivity of small lakes to micro-topographic and hydrological conditions.

Regional variability: The spatial distribution of change was highly heterogeneous. Issyk-Kul and Batken regions showed dense clusters of glacier-connected lakes, whereas Chuy and Talas were characterized by newly emerging small lakes. In Osh, both some lakes expanded and others disappeared, reflecting complex geomorphological and hydrological controls. This variability underlines the need for region-specific monitoring strategies rather than uniform national approaches.

Implications for hazard and water resource management: The continued formation of small, high-altitude lakes emphasizes the importance of long-term glacial lake observation to anticipate potential GLOFs and assess water resource dynamics. The

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

manually validated 2024 inventory provides a valuable baseline for such efforts and a unique training resource for future remote sensing and machine learning studies across Central Asia.

In summary, the combined use of semi-automated NDWI delineation and machine-learning classification, anchored by a manually corrected reference dataset, enables robust multi-temporal glacial lake mapping in complex mountainous terrain. Our analysis revealed that the smallest and most numerous glacial lakes are the most affected by glacial retreat, yet are the most challenging to detect automatically. Their inclusion in the 2024 inventory marks a significant step toward comprehensive monitoring of glacier–lake systems in Kyrgyzstan. This refined dataset will be instrumental for improving automated mapping approaches and for advancing our understanding of cryospheric responses to climate change in Central Asia.

8 Data availability

The data set comprises two shapefiles (.shp) containing the glacial lake inventories of the Kyrgyz mountains for 2016–2017 and 2022–2024. The data described in this study, including post-processed and validated lakes, can be accessed at the Zenodo repository under https://zenodo.org/records/17869915 (Piroton et al., 2025).

9 Code availability

The scripts and workflows used for Sentinel-2 image processing, NDWI computation, and Random Forest classification are available from the corresponding author upon reasonable request.

Author contributions.

Conceptualization: VP. Data curation: VP. Formal analysis: VP. Investigation: VP. Visualization: VP Writing– original draft preparation: VP. Writing– review and editing: AE and HBH. Funding acquisition: AE. All the authors have read and agreed to the published version of the manuscript.

Competing interests.

The authors declare that they have no conflict of interest.

Disclaimer.

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgments.

The original manuscript was refined using ChatGPT and later checked by a native speaker. The authors would like to thank Robert Dennen for proofreading the article. AE acknowledges the support from the HINTERLANDS project (High mountains

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

in the Anthropocene: from landscape dynamics to hazards and risks; PRIMUS/25/SCI/005) and the Johannes Amos Comenius Programme (P JAC), project No. CZ.02.01.01/00/22_008/0004605, Natural and anthropogenic georisks, realized at the Charles University, Faculty of Science.

References

Adyshev, M. M., Kashirin, F. T., Umurzakov, S. U., Almaev, T. M., Voronina, A. F., Grigorenko, P. G., Dzhamgerchinov, B. D., Zabirov, R. D., Zinkova, Z. Y., Izmailov, A. E., Isabaeva, V. A., Kravchenko, A. V., Mamytov, A. M., Makhrina, L. I., Moldokulov, A. M., Murzaev, E. M., Otorbaev, K. O., Popova, L. I., Yar-Mukhamedov, G. K., Yashina, V. V., and Chernova, L. I.: Атлас Киргизской ССР [Atlas of the Kyrgyz SSR (vol. I)] (in Russian), Fabrika #4, Moscow, 1987.

Aizen, V. B., Kuzmichenok, V. A., Surazakov, A. B., and Aizen, E. M.: Glacier changes in the Central and Northern Tien Shan during the last 140 years based on surface and remote-sensing data, Ann. Glaciol., 43, 202–213, doi:10.3189/172756406781812465, 2006.

Akimaliev, D. A., Zaurov, D. E., and Eisenman, S. W.: The geography, climate and vegetation of Kyrgyzstan, in: Medicinal Plants of Central Asia: Uzbekistan and Kyrgyzstan, edited by Eisenman, S. W., Zaurov, D. E., and Struwe, L., Springer, New York, pp. 1–3, doi:10.1007/978-1-4614-3912-7_1, 2013.

Baťka, J., Vilímek, V., Štefanová, E., Cook, S. J., and Emmer, A.: Glacial Lake Outburst Floods (GLOFs) in the Cordillera Huayhuash, Peru: Historic Events and Current Susceptibility. Water, 12(10), 2664, 1-17, doi:10.3390/w12102664, 2020.

Bolch, T., Kulkarni, A., Kääb, A., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., Scheel, M., Bajracharya, S., and Stoffel, M.: The state and fate of Himalayan glaciers, Science, 336, 310–314, doi:10.1126/science.1215828, 2012.

Bolch, T., Peters, J., Yegorov, A., Pradhan, B., Buchroithner, M., and Blagoveshchensky, V.: Identification of potentially dangerous glacial lakes in the northern Tien Shan, Nat. Hazards, 59, 1691–1714, doi:10.1007/s11069-011-9860-2, 2011.

Bolch, T., and Kamp, U.: Glacier mapping in high mountains using DEMs, Landsat and ASTER data, Grazer Schriften der Geographie und Raumforschung, 1, 37–48, 2006.

Buckel, J., Otto, J. C., Prasicek, G., and Keuschnig, M.: Glacial lakes in Austria – Distribution and formation since the Little Ice Age, Glob. Planet. Change, 164, 39–51, doi:10.1016/j.gloplacha.2018.03.003, 2018.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Capps, D. M., Rabus, B., Clague, J. J., and Shugar, D. H.: Identification and characterization of alpine subglacial lakes using interferometric synthetic aperture radar (InSAR): Brady Glacier, Alaska, USA, J. Glaciol., 56(199), 861–870, doi:10.3189/002214310794457254, 2010.

Carrivick, J. L., and Tweed, F. S.: A global assessment of the societal impacts of glacier outburst floods, Glob. Planet. Change, 144, 1–16, doi:10.1016/j.gloplacha.2016.07.001, 2016.

CACILM/ADB: Central Asia atlas of natural resources, 221 pp., https://think-asia.org/handle/11540/155, 2010.

Chen, F., Zhang, M., Guo, H., Allen, S., Kargel, J., Haritashya, U., and Watson, C. S.: Annual 30-meter dataset for glacial lakes in High Mountain Asia from 2008 to 2017, Earth Syst. Sci. Data, 4275164, 1–29, doi:10.5194/essd-2020-57, 2020.

Clague, J. J., and Evans, S. G.: A review of catastrophic drainage of moraine-dammed lakes in British Columbia, Quaternary Sci. Rev., 19, 1763–1783, doi:10.1016/S0277-3791(00)00090-1, 2000.

Clauset, A., Shalizi, C., and Newman, M.: Power-law distributions in empirical data, SIAM Rev., 51, 661–703, doi:10.1137/070710111, 2009.

Daiyrov, M., and Narama, C.: Constantly renewing glacial lakes in the Kyrgyz Range, northern Tien Shan, Nat. Hazards Earth Syst. Sci., doi:10.5194/nhess-2024-160, 2024.

Daiyrov, M., Kattel, D. B., Narama, C., and Wang, W.: Evaluating the variability of glacial lakes in the Kyrgyz and Teskey ranges, Tien Shan, Front. Earth Sci., 10, doi:10.3389/feart.2022.850146, 2022.

Emmer, A.: Understanding the risk of glacial lake outburst floods in the twenty-first century, Nature Water, 2, 608-610, doi: 10.1038/s44221-024-00254-1, 2024.

Emmer, A., Harrison, S., Mergili, M., Allen, S., Frey, H., and Huggel, C.: 70 years of lake evolution and glacial lake outburst floods in the Cordillera Blanca (Peru) and implications for the future, Geomorphology, 365, 107178, doi:10.1016/j.geomorph.2020.107178 , 2020.

Emmer, A., Vilímek, V., Huggel, C., Klimeš, J., and Schaub, Y.: Limits and challenges to compiling and developing a database of glacial lake outburst floods, Landslides, 13, 1579–1584, doi:10.1007/s10346-016-0686-6, 2016.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Emmer, A., Vilímek, V., Klimeš, J., and Cochachin, A.: Glacier retreat, lakes development and associated natural hazards in Cordilera Blanca, Peru, Environ. Sci. Eng., 231–252, doi:10.1007/978-3-319-00867-7_17, 2014.

Erokhin, S. A., Zaginaev, V. V., Meleshko, A. A., Ruiz-Villanueva, V., Petrakov, D. A., Chernomorets, S. S., Viskhadzhieva, K. S., Tutubalina, O. V., and Stoffel, M.: Debris flows triggered from non-stationary glacier lake outbursts: the case of the Teztor Lake complex (Northern Tian Shan, Kyrgyzstan), Landslides, 15, 83–98, doi:10.1007/s10346-017-0862-3, 2018.

Esri, DeLorme Publishing Company, Inc.: World Linear Water, Esri [data set], https://www.arcgis.com/home/item.html?id=273980c20bc74f94ac96c7892ec15aff, accessed on 12 December 2025.

Esri: World Hillshade [basemap], ArcGIS Online Living Atlas of the World, https://www.arcgis.com/home/item.html?id=1b243539f4514b6ba35e7d995890db1d, accessed on 12 December 2025.

Evans, S. G., and Clague, J. J.: Recent climatic change and catastrophic geomorphic processes in mountain environments, Geomorphology, 10, 107–128, doi:10.1016/0169-555X(94)90011-6, 1994.

Frey, H., Huggel, C., Paul, F., and Haeberli, W.: Automated detection of glacier lakes based on remote sensing in view of assessing associated hazard potentials, Grazer Schriften der Geographie und Raumforschung, 45, 261–272, doi:10.5167/UZH-128917, 2010.

González-Moradas, M. D. R., and Viveen, W.: Evaluation of ASTER GDEM2, SRTMv3.0, ALOS AW3D30 and TanDEM-X DEMs for the Peruvian Andes against highly accurate GNSS ground control points and geomorphological-hydrological metrics, Remote Sens. Environ., 237, 111509, doi:10.1016/j.rse.2019.111509, 2020.

Haeberli, W., Buetler, M., Huggel, C., Friedli, T. L., Schaub, Y., and Schleiss, A. J.: New lakes in deglaciating high-mountain regions – opportunities and risks, Clim. Change, 139, 201–214, doi:10.1007/s10584-016-1771-5, 2016.

Hanshaw, M. N., and Bookhagen, B.: Glacial areas, lake areas, and snow lines from 1975 to 2012: status of the Cordillera Vilcanota, including the Quelccaya ice cap, northern Central Andes, The Cryosphere, 8, 359–376, doi:10.5194/tc-8-359-2014, 2014.

Harrison, S., Kargel, J. S., Huggel, C., Reynolds, J., Shugar, D. H., Betts, R. A., Emmer, A., Glasser, N., Haritashya, U. K., Klimeš, J., Reinhardt, L., Schaub, Y., Wiltshire, A., Regmi, D., and Vilímek, V.: Climate change and the global pattern of moraine-dammed glacial lake outburst floods, The Cryosphere, 12, 1195–1209, doi:10.5194/tc-12-1195-2018, 2018.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Hartigan, J. A., and Hartigan, P. M.: The Dip Test of Unimodality. The Annals of Statistics, *13*(1), 70–84; doi:10.1214/aos/1176346577, 1985

Huggel, C., Kääb, A., Haeberli, W., Teysseire, P., and Paul, F.: Remote sensing based assessment of hazards from glacier lake outbursts: a case study in the Swiss Alps, Can. Geotech. J., 39, 316–330, doi:10.1139/t01-099, 2002.

Izagirre, E., Casassa, G., Dussaillant, I., Miles, E.S., Wilson, R., Rada, C., Faria, S.H. and Antiguedad, I.: Evolution of glacial lakes and southernmost GLOFs in the Cordillera Darwin and Cloue Icefields (Tierra del Fuego) between 1945–2024, Front. Earth Sci., 13:1641167, doi: 10.3389/feart.2025.1641167, 2025.

JAXA: ALOS global digital surface model 'ALOS World 3D–30m' (AW3D30), Japan Aerospace Exploration Agency, 2025.

Janský, B., Šobr, M., and Engel, Z.: Outburst flood hazard: case studies from the Tien-Shan Mountains, Kyrgyzstan, Limnologica, 40(4), 358–364, doi:10.1016/j.limno.2009.11.013, 2010.

Kääb, A., Huggel, C., Fischer, L., Guex, S., Paul, F., Roer, I., Salzmann, N., Schlaefli, S., Schmutz, K., Schneider, D., Strozzi, T., and Weidmann, Y.: Remote sensing of glacier- and permafrost-related hazards in high mountains: an overview, Nat. Hazards Earth Syst. Sci., 5, 527–554, doi:10.5194/nhess-5-527-2005, 2005

Kapitsa, V., Shahgedanova, M., Severskiy, I., Kasatkin, N., White, K., and Usmanova, Z.: Assessment of changes in mass balance of the Tuyuksu Group of Glaciers, Northern Tien Shan, between 1958 and 2016 using ground-based observations and Pléiades satellite imagery, Front. Earth Sci., 8(259), doi:10.3389/feart.2020.00259, 2020.

Kaplan, G., and Avdan, U.: Object-based water body extraction model using Sentinel-2 satellite imagery, Eur. J. Remote Sens., 50, 137–143, doi:10.1080/22797254.2017.1297540, 2017.

Kattel, D. B., Mohanty, A., Daiyrov, M., Wang, W., Mishra, M., Kulenbekov, Z., and Dawadi, B.: Evaluation of glacial lakes and catastrophic floods on the northern slopes of the Kyrgyz Range, Mountain Res. Dev., 40(3), 37-43, doi:10.1659/mrd-journal-d-19-00068.1, 2020.

Khan, I., Jacobs, J. M., Johnston, J. M., and Vardaman, M.: Mapping Glacial Lakes in the Upper Indus Basin (UIB) Using Synthetic Aperture Radar (SAR) Data, Preprints, doi:10.20944/preprints202509.0422.v1, 2025.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Kulikov, M., and Schickhoff, U.: Vegetation and climate interaction patterns in Kyrgyzstan: spatial discretization based on time series analysis, Erdkunde, 71, 143–165, doi:10.3112/erdkunde.2017.02.04, 2017.

Li, H., Zhao, J., Yan, B., Yue, L., and Wang, L.: Global DEMs vary from one to another: an evaluation of newly released Copernicus, NASA and AW3D30 DEM on selected terrains of China using ICESat-2 altimetry data, Int. J. Digit. Earth, 15, 1149–1168, doi:10.1080/17538947.2022.2094002, 2022.

Liu, Z., Yao, Z., and Wang, R.: Assessing methods of identifying open water bodies using Landsat 8 OLI imagery, Environ. Earth Sci., 75, 873, doi:10.1007/s12665-016-5686-2, 2016.

Maussion, F., Hock, R., Paul, F., Rastner, P., Raup, B., Zemp, M., and the RGI Consortium: The Randolph Glacier Inventory (RGI) version 7, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-4484, doi:10.5194/egusphere-egu22-4484, 2022.

Mergili, M., Müller, J. P., and Schneider, J. F.: Spatio-temporal development of high-mountain lakes in the headwaters of the Amu Darya River (Central Asia), Glob. Planet. Change, 107, 13-24, doi:10.1016/j.gloplacha.2013.04.001, 2013.

McFeeters, S. K.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, Int. J. Remote Sens., 17(7), 1425–1432, doi:10.1080/01431169608948714, 1996.

Narama, C., Duishonakunov, M., Kääb, A., Daiyrov, M., and Abdrakhmatov, K.: The 24 July 2008 outburst flood at the western Zyndan glacier lake and recent regional changes in glacier lakes of the Teskey Ala-Too range, Tien Shan, Kyrgyzstan, Nat. Hazards Earth Syst. Sci., 10(4), 647–659, doi:10.5194/nhess-10-647-2010, 2010.

Otto, J.-C.: Proglacial lakes in high mountain environments, in: Geomorphology of Proglacial Systems. Landform and sediment dynamics in recently deglaciated alpine landscapes, Geography of the Physical Environment, eds. Heckmann, T., and Morche, D., Springer International Publishing, Cham, 231–247, doi:10.1007/978-3-319-94184-4, 2019.

Patterson, T., and Kelso, N. V.: Natural Earth World Lakes, 1:10 million, Stanford Public, 2012.

Petrakov, D. A., Chernomorets, S. S., Viskhadzhieva, K. S., Dokukin, M. D., Savernyuk, E. A., Petrov, M. A., Erokhin, S. A., Tutubalina, O. V., Glazyrin, G. E., Shpuntova, A. M., and Stoffel, M.: Putting the poorly documented 1998 GLOF disaster in Shakhimardan River valley (Alay Range, Kyrgyzstan/Uzbekistan) into perspective, Sci. Total Environ., 724, doi:10.1016/j.scitotenv.2020.138287, 2020.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Piroton, V., Emmer, A., and Havenith, H.B.: Glacial lakes of Kyrgyzstan (2016 & 2024), Zenodo [data set], doi:10.5281/zenodo.17869915, 2025.

RGI 7.0 Consortium. Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 7.0. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. doi:10.5067/f6jmovy5navz, 2023.

Sattar, A., Goswami, A., Kulkarni, A. V., Emmer, A., Haritashya, U. K., Allen, S., Frey, H., and Huggel, C.: Future glacial lake outburst flood (GLOF) hazard of the South Lhonak Lake, Sikkim Himalaya, Geomorphology, 388, 107783, doi:10.1016/j.geomorph.2021.107783, 2021.

Sattar, A., Emmer, A., Lhazom, T., Rai, S.K., and Azam, M.F.: Flood risk from small mountain lakes. Comm. Earth Environ., 6, 785, doi:10.1038/s43247-025-02758-4 , 2025.

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, Nat. Clim. Change, 10, 939–945, doi:10.1038/s41558-020-0855-4, 2020.

Song, C., Fan, C., Ma, J., Zhan, P., and Deng, X.: A spatially constrained remote sensing-based inventory of glacial lakes worldwide, Sci. Data, 12, 464, doi:0.1038/s41597-025-04809-z, 2025.

Sorg, A., Bolch, T., Stoffel, M., Solomina, O., and Beniston, M.: Climate change impacts on glaciers and runoff in Tien Shan (Central Asia), Nat. Clim. Change, 2, 725–731, doi:10.1038/nclimate1592, 2012.

Stanford University / International Steering Committee for Global Mapping: Region line boundaries, Kyrgyzstan, Global Map Kyrgyz v2 dataset, Stanford Digital Repository, 2012, available at https://purl.stanford.edu/jy474rx9360, accessed 2025-10-21.

Strozzi, T., Wiesmann, A., Kääb, A., Joshi, S., and Mool, P.: Glacial lake mapping with very high resolution satellite SAR data, Nat. Hazards Earth Syst. Sci., 12(8), 2487–2498, doi:10.5194/nhess-12-2487-2012, 2012.

Tadono, T., Nagai, H., Ishida, H., Oda, F., Naito, S., Minakawa, K., and Iwamoto, H.: Generation of the 30 m-Mesh Global Digital Surface Model by ALOS Prism, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLI-B4, 157–162, doi:10.5194/isprsarchives-XLI-B4-157-2016, 2016.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Takeuchi, N., Fujita, K., Aizen, V. B., Narama, C., Yokoyama, Y., Okamoto, S., Naoki, K., and Kubota, J.: The disappearance of glaciers in the Tien Shan Mountains in Central Asia at the end of Pleistocene, Quaternary Sci. Rev., 103, 26–33, doi:10.1016/j.quascirev.2014.09.006, 2014.

Tomaszewska, M. A., and Henebry, G. M.: Changing snow seasonality in the highlands of Kyrgyzstan, Environ. Res. Lett., 13, doi:10.1088/1748-9326/aabd6f, 2018.

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.

Veh, G., Lützow, N., Kharlamova, V., Petrakov, D., Hugonnet, R., and Korup, O.: Trends, breaks, and biases in the frequency of reported glacier lake outburst floods, Earth's Future, 10(3), doi:10.1029/2021EF002426, 2022.

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 in 1990 and 2018 derived from Landsat images, Earth Syst. Sci. Data, 12(3), 2169–2182, doi:10.5194/essd-12-2169-2020, 2020.

Wang, X., Chen, W., Zhang, G., Emmer, A., Frey, H., Taylor, C., Huggel, C., Sattar, A., Zheng, G., Rashid, I., Carrivick, J.L., Veh, G., Allen, S., Steiner, J., Quincey, D., and Mergili, M.: Mitigating future glacial lake outburst floods in the Himalaya, Science Bulletin, doi:10.1016/j.scib.2025.11.024, 2025.

Wangchuk, S., and Bolch, T.: Mapping of glacial lakes using Sentinel-1 and Sentinel-2 data and a random forest classifier: strengths and challenges, Sci. Remote Sens., 2, 100008, doi:10.1016/j.srs.2020.100008, 2020.

Zhang, B., Liu, G., Zhang, R., Fu, Y., Liu, Q., Cai, J., Wang, X., and Li, Z.: Monitoring dynamic evolution of the glacial lakes by using time series of Sentinel-1A SAR images, Remote Sens., 13(7) 1313, doi:10.3390/rs13071313, 2021.

Zhang, G., Carrivick, J.L., Emmer, A., Shugar, D.H., Veh, G., Wang, X., Labedz, C., Mergili, M., Mölg, N., Huss, M., Allen, S., Sugiyama, S., and Lützow, N.: Characteristics and changes of glacial lakes and outburst floods. Nat. Rev. Earth and Environ., 5, 447-462, doi:10.1038/s43017-024-00554-w, 2024.

Zhang, M., Chen, F., and Tian, B.-S.: An automated method for glacial lake mapping in High Mountain Asia using Landsat 8 imagery, J. Mountain Sci., 15, 13–24, doi:10.1007/s11629-017-4518-5, 2018.

https://doi.org/10.5194/egusphere-2026-1442
Preprint. Discussion started: 27 March 2026

Zhang, M., Chen, F., Guo, H., Yi, L., Zeng, J., and Li, B.: Glacial lake area changes in High Mountain Asia during 1990–2020 using satellite remote sensing, Res., 2022, doi:10.34133/2022/9821275, 2022.

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, Int. J. Environ. Res. Public Health, 17, 1072, doi:10.3390/ijerph17031072, 2020.