Citation (APA)
Wang, Z., Shi, L., Qiu, Y., Jia, G., Li, X., Leppäranta, M., Li, L., & Menenti, M. (in press). Vector dataset of rivers, lakes/reservoirs and their spatial connectivity in High Mountain Asia. *Big Earth Data*. https://doi.org/10.1080/20964471.2026.2688654

Big Earth Data

ISSN: 2096-4471 (Print) 2574-5417 (Online) Journal homepage: www.tandfonline.com/journals/tbed20

Vector dataset of rivers, lakes/reservoirs and their spatial connectivity in High Mountain Asia

Zhaocheng Wang , Lijuan Shi , Yubao Qiu, Guoqiang Jia , Xiaoting Li , Matti Leppäranta , Lanhai Li & Massimo Menenti

To cite this article: Zhaocheng Wang , Lijuan Shi , Yubao Qiu, Guoqiang Jia , Xiaoting Li , Matti Leppäranta , Lanhai Li & Massimo Menenti (26 Jun 2026): Vector dataset of rivers, lakes/reservoirs and their spatial connectivity in High Mountain Asia, Big Earth Data, DOI: 10.1080/20964471.2026.2688654

To link to this article: https://doi.org/10.1080/20964471.2026.2688654

Published online: 26 Jun 2026.

BIG EARTH DATA
https://doi.org/10.1080/20964471.2026.2688654

DATA NOTE

OPEN ACCESS Check for updates

Vector dataset of rivers, lakes/reservoirs and their spatial connectivity in High Mountain Asia

Zhaocheng Wang[a,b,c], Lijuan Shi[c,d], Yubao Qiu[b,c], Guoqiang Jia[b,c], Xiaoting Li[b,c], Matti Leppäranta[b,c,e], Lanhai Li[f] and Massimo Menenti[c,g,h]

[a]School of Computer and Information Security, Guilin University of Electronic Technology, Guilin, China; [b]International Research Center of Big Data for Sustainable Development Goals (CBAS), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China; [c]Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China; [d]The College of Computer and Information Engineering, The Henan Key Laboratory of Big Data Analysis and Processing, Henan University, Kaifeng, China; [e]Institute of Atmospheric and Earth Sciences, University of Helsinki, Helsinki, Finland; [f]The National Key Laboratory of Ecological Security and Sustainable Development in Arid Region, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China; [g]The Netherlands Department of Geoscience and Remote Sensing, Civil Engineering, Delft University of Technology, Delft, The Netherlands; [h]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

ABSTRACT

Significant changes in snow, glaciers, and the water system across High Mountain Asia (HMA) under global warming have intensified regional water system instabilities. Accurate mapping of the spatial distribution and connectivity of rivers and lakes is essential for understanding and managing the regional water balance and ecosystems. However, existing surface-water datasets lack clear distinctions and connectivity information between flowing rivers and lakes. In this study, three vector datasets over the HMA were developed—Rivers, Lakes, and River–Lake spatial Connectivity—based on the European Space Agency's (ESA) WorldCover data and the 2020 Global Surface Water dataset from the European Commission's Joint Research Centre (JRC). Rivers and lakes were extracted separately using Geographic Information System tools for all basins in the HMA and incorporated into a surface water time series. The analysis results showed that the total area of rivers and lakes in the HMA increased by 23.5% from 2000 to 2019, with lakes accounting for 78.3% of this expansion, while river areas expanded by 40.8%. Glacier-fed rivers and lakes contributed 75.5% of the total water area increase, although non-glacier-fed waters exhibited higher relative growth (44.8%) than glacier-fed ones (20.4%). In the endorheic basins, a consistent annual increase in lake area from 2000 to 2019 was observed: the area of hydrologically connected lakes and rivers expanded from 20,782 km^2 to 25,396 km^2. In contrast, the area of the

ARTICLE HISTORY
Received 10 November 2025
Accepted 23 May 2026

KEYWORDS
High Mountain Asia; rivers and lakes; water body's connectivity; vector dataset; surface water change

CONTACT Yubao Qiu qiuyb@aircas.ac.cn Aerospace Information Research Institute, Chinese Academy of Sciences, No. 9 Dengzhuang South Road, Haidian District, Beijing 100094, China

isolated lakes grew from 9,550 km^2 to 12,970 km^2. Statistical analysis revealed a significant correlation between the lake area expansion and precipitation. The datasets developed in this study provide the fundamental basis for analyzing changes in rivers and lakes in the HMA, along with their connectivity, supporting regional water resource management, ecological conservation, and sustainable agricultural and pastoral development.

1. Introduction

The High Mountain Asia (HMA), centered with the Qinghai-Tibetan Plateau, spans over mountainous regions ranging from 2,000 to 8,844 (Mt. Everest) meters above sea level. It is home to the largest reservoirs of snow, ice, and water outside the Earth's polar regions (Loomis et al., 2019) and serves as a critical high-altitude water resource for mid- to low-latitude regions of Asia. This area contains various types of water storages, including glaciers, lakes, and permafrost, featured with the name of "Asian Water Tower" (Yao et al., 2019; Zhu et al., 2019). Rivers and lakes (including reservoirs) form the integral of the region's water resources, connect surrounding areas, and support the daily needs of over 1.4 billion people in downstream regions. A precise understanding of the spatial distribution and dynamic changes in these waters is essential for effective water resource management, ecological conservation, and social and economic development in the HMA and its drainage basin.

However, this critical water system is undergoing rapid changes due to ongoing climate warming. Climate change is a rapidly reshaping hydrological process in the HMA that influences glacier melt, precipitation patterns, and river dynamics. These changes have increased the complexity and frequency of flood events across the region, posing increasing risks to ecosystems and infrastructure (Bai et al., 2025; Liu & Wang, 2026). Over the past few decades, HMA has experienced significant warming, resulting in reduced snow cover, shrinking glaciers, and permafrost (Wang et al., 2021). These changes have led to alterations in the surface water areas of rivers and lakes, disrupting the balance of water resources. Comprehensive discoveries have been made by a number of researchers from 2010 to 2025 for the fast expansion of the water surface area. Zhang et al. (2017) indicated that from the 1970s to 2010, 99 new lakes emerged on the plateau, and 81% of the existing lakes expanded, leading to a total area increase of 18.4% (7,240 km^2). Qiao et al. (2019) found that the total area of lakes in the endorheic basin expanded from 31,858.0 km^2 in 1976 to 38,292.6 km^2 in 2013. Song et al. (2019) reported an increase in the area of lakes in the inner Tibetan Plateau from 2000 to 2015, with 95.6% of the expansion occurring in lakes larger than 1 km^2. A further study indicated that the number of lakes larger than 1 km^2 remained relatively stable from ~2000 to 2018, with only a slight increase from 1,174 to 1,424, while the total area of these lakes grew from 41,000 ± 443.7 km^2 to 50,000 × 10^4 ± 791.4 km^2. The increase in the number of lakes (>1 km^2 in area) is primarily due to lake expansion, where lakes grow in size from less than 1 km^2 to more than 1 km^2 (Zhang et al., 2019). The rapid and exceptional expansion of lakes in the HMA poses significant risks to the water resources and ecosystems.

These changes in water bodies over HMA need to be properly understood for water and ecosystem governance and decision-making. Recent studies indicate that lake expansion on the Qinghai-Tibetan Plateau is closely linked to increased precipitation and rising temperatures, with the resulting melting of glaciers and permafrost. For example, Phan et al. (2013) found that 244 of 891 analyzed lakes received direct glacial runoff, highlighting the influence of glacial meltwater on lake hydrological processes. Liu et al. (2025) highlighted that one of the main drivers behind the expansion of lakes is the melting of ground ice near the permafrost surface. Accelerated melting of glaciers and permafrost has also significantly increased surface runoff in some regions. For instance, Wang et al. (2021) discovered that since 1997, the accelerated melting of glaciers and permafrost has led to an approximately 80.9% increase in surface runoff in the upper reaches of the Yarlung Zangbo River Basin in the southeastern Qinghai-Tibetan Plateau. In addition, the increase in regional precipitation and sustained warming over the past decade have jointly contributed to lake expansion (Zhang et al., 2020). Specifically, the water level of Qinghai Lake is positively correlated with surface runoff and precipitation, and negatively correlated with evaporation (Li et al., 2023).

Overall, climate warming on the Qinghai-Tibetan Plateau intensified the hydrological cycle, with the increase in water volume occurring much faster than the rise in net atmospheric water flux (Ayantobo & Wei, 2025; Wang et al., 2022). River responses to climate change in the HMA vary significantly across basins and seasons (Khanal et al., 2021). River discharge has accelerated across parts of the HMA, with significant increases observed in nearly 10% of monitored river reaches (Flores et al., 2025). Recent research further indicates that climate change is increasing flood complexity and exposure risks across the HMA, with more populations and infrastructure becoming vulnerable to flood hazards (Bai et al., 2025).

These climate-driven hydrological changes not only alter water storage but also strongly affect river systems and their internal structures. In river network systems, the interconnection between river branches affects water flow distribution, sediment transport, and deposition, and consequently influences flood events, sediment distribution, and the maintenance of river biodiversity (Lei et al., 2023; Nyberg et al., 2023). Glaciers play a dominant role in controlling sediment flux in rivers (Li et al., 2024; Zhang et al., 2023). As sediment flux increases, riverbed migration becomes more frequent, and the risk of flood disasters intensifies (Li et al., 2021). This poses a challenge to the stability of river ecosystems. The loss of connectivity between rivers and lakes can lead to a significant reduction in lake wetland areas, although the emergence of new land provides opportunities for agricultural development (Li et al., 2024; D. Wang et al., 2021).

The research thus far tells the truth of the temperature increase, water cycle acceleration (e.g., increased precipitation), and the induced systematic changes in water storage threatening water and ecosystem security over HMA. However, the detailed aspects of these changes are not clearly understood at the scale of rivers, lakes, and their interactions. Rapid variations in surface water bodies require high-frequency observations (weekly or daily) for accurate monitoring. In addition, the connectivity of rivers and lakes in the HMA is vital for understanding regional shifts in water resources and their impact (Chauhan et al., 2023; Xu et al., 2024).

Numerous global and regional datasets have been developed to map surface water bodies. Notable examples are the integrated SRTM data from the Space

Shuttle Radar Topography Mission and Moderate Resolution Imaging Spectroradiometer (MODIS) 250-meter resolution data to develop a global land surface water mask with a 250-meter spatial resolution (Carroll et al., 2009), and the permanent and seasonal water areas of global rivers and lakes using Landsat satellite image data from 1972 to 2022 with manual visual inspection and an automated computer algorithm (Nyberg et al., 2024; Pi et al., 2022; Zhang et al., 2017). However, these datasets have limitations in distinguishing between rivers and lakes, and suffer from a lack of connectivity characteristics of HMA surface waters. Some of the datasets include the location of reservoirs but not the area of water bodies (Fan et al., 2024). Recently, Sui et al. (2025) developed an HMA water body dataset based on satellite remote sensing data and provided valuable information on the regional water distribution.

Overall, most existing surface water datasets primarily provide raster-based water masks or lake inventories, which neither explicitly separate river and lake systems, nor represent their connectivity. This limitation hinders a comprehensive understanding of river–lake interactions across the HMA. To address this gap, this study uses two widely used water body datasets as inputs to develop and process a set of vector-based datasets of rivers and lakes and their spatial connectivity across the HMA. In this study, river–lake spatial connectivity refers to the geometric or topological relationships between rivers and lakes rather than hydrological connectivity. The proposed dataset explicitly separates river and lake systems and establishes river–lake spatial connectivity relationships, thereby providing a more complete representation of the regional surface water network. This dataset supports automated calculations and spatial analyses of water bodies, enables river and lake separation, and provides essential spatial connectivity information between the rivers and lakes. The resulting data will be instrumental in extracting more precise information on water body changes related to weather and hydrological processes and will provide critical data for surface water resource management and analysis of their changes in HMA.

2. Data

The study area of the HMA spans 22°N to 46°N and 62°E to 110°E, with elevations ranging from 2,000 to 8,844 m. The vector data for rivers and lakes and their spatial connectivity within this region were compiled using existing datasets, such as the vector dataset for river systems originating in Eurasia (Liang et al., 2020) and the 2020 ESA Land Cover Data Product WorldCover (Zanaga et al., 2022). Additional information on the datasets used is provided in the subsections below.

2.1. *Vector dataset for river network system*

The vector dataset for the river network system originating in Eurasia was developed using 30-meter resolution surface water data extracted from Landsat imagery in 2010, the Global River Width Dataset (GRWL), and the 30-meter resolution digital elevation data from NASA's Shuttle Radar Topography Mission (SRTM). This dataset offers a comprehensive representation of rivers and lakes (including reservoirs) across the Eurasian continent, with an accuracy of 93.96% (Liang et al., 2020).

The dataset was available for downloading from the National Tibetan Plateau Data Center (https://data.tpdc.ac.cn/home). Hereafter, referred to as the River Dataset 2010 (RD2010), provides basic historical vector data of rivers and lakes separately in the HMA.

2.2. ESA WorldCover

The WorldCover 2020 product provided by the European Space Agency (ESA) represents the 2020 status of the land surface at a spatial resolution of 10 m. It was produced from Sentinel-1 and Sentinel-2 satellite data and verified by ground measurements and visual reviews (Zanaga et al., 2022). The dataset includes ten land feature types, including water bodies, which can be downloaded from the ESA WorldCover website (https://esa-worldcover.org/en) or Zenodo (https://zenodo.org/records/5571936).

Water body data extracted from WorldCover 2020 were used to create a water surface dataset, which is hereafter referred to as WorldCover Surface Water 2020 (WSW2020). This was employed to supplement the gaps in the water surface in RD2010. This dataset has an overall accuracy of approximately 74.4%, and its classification errors may introduce some uncertainty when distinguishing water bodies from other land-cover types. In addition, the dataset can be affected by factors such as the terrain and shadows. Therefore, integrating this dataset with RD2010 can improve the accuracy and spatial detail of surface water mapping, while reducing misclassification in the HMA region.

2.3. JRC Global Surface Water

The JRC Yearly Water Classification History (v1.4) is a raster dataset that classifies global water-surface changes from 1984 to 2022. It was developed by the Joint Research Centre of the European Union (JRC) based on high-resolution images of the Landsat satellite series (Pekel et al., 2016). The dataset categorizes water bodies into four types by analyzing the water occurrence in each pixel over the course of the year: no data, no water, seasonal water, or permanent water. The dataset is freely available for downloading and accessing via Google Earth Engine, which is primarily used for verification and time-series analysis.

2.4. HydroSHEDS

HydroSHEDS, derived from NASA's Shuttle Radar Topography Mission (SRTM), provides global river networks, watershed boundaries, and hydrological feature information with a spatial resolution of 30 m (Lehner & Grill, 2013). Developed primarily by the International Rivers Network, this dataset was downloaded from the HydroSHEDS website (https://www.hydrosheds.org/). In this study, the HydroRIVERS river distribution dataset and HydroBASINS watershed vector dataset from HydroSHEDS were used. Specifically, HydroRIVERS served as the reference for river data preparation, whereas HydroBASINS facilitated further subdivision of the HMA watersheds.

2.5. Randolph Glacier Inventory

The Randolph Glacier Inventory (RGI) v6.0 is a global glacier dataset that provides information on the distribution of glaciers as well as their area, volume, length, orientation, and elevation, supports climate change research and water resource management (RGI Consortium, 2017). Developed collaboratively by the international glacier research community, this dataset covers approximately 200,000 glaciers worldwide. In the present study, version 6.0, which was used to identify glaciers within river basins, enabled the classification of basins into glacier-fed or runoff-fed cases. Because existing surface water body datasets often include seasonal snow cover on glaciers, RGI helps to generate glacier masks for water body extraction.

2.6. Classified drainage areas based on water-supply patterns

The area of the present water resource study was based on the HMA river basin. The basin dataset was extracted using water resource recharge patterns (Lu et al., 2020). Four types of basins were identified: Glacier-fed and Runoff-fed Drainage Area (GRDA), Glacier-fed and Runoff-free Drainage Area (GDA), Glacier-free and Runoff-fed Drainage Area (RDA), and the Glacier-free and Runoff-free Drainage Area (NGRDA).

In this study, the influence of glaciers on rivers and lakes was determined based on the basin classification dataset developed by Lu et al. (2020). Basins classified as GRDA and GDA were considered glacier-fed basins, whereas those classified as RDA and NGRDA were considered non-glacier-fed basins. Therefore, rivers and lakes located within glacier-fed basins were classified as glacier-fed water bodies. It should be noted that the term "runoff-free" adopted in this study from the work of Lu et al. (2020) refers to the absence of clearly identifiable surface runoff inputs under the spatial resolution and basin classification framework of the original input datasets rather than the complete absence of any water supply. Water bodies within the NGRDA regions may still receive water from direct precipitation, groundwater recharge, seasonal meltwater, or other local hydrological processes.

Figure 1 outlines the basin boundaries in HMA and illustrates the spatial distribution of glaciers, rivers, and lakes within these basins, and the corresponding basins are listed in Table 1. The dataset is available from Science DB (https://www.scidb.cn/en). It is primarily used to determine basin boundaries and, additionally, in combination with river and lake water body datasets, to help distinguish the water supply recharge patterns across the HMA basins.

3. Methods

The data production process includes three primary steps: creating river datasets, lake datasets, and datasets for connected river-lake systems. The detailed workflow is illustrated in Figure 2.

(1) Creation of the River network System dataset.
    Using the Intersection, Erase, and Merge tools in common GIS software, the extracted water body data were refined to retain only river features, resulting in a dedicated river vector dataset, named River network System dataset. This process ensured accurate identification and documentation of all rivers within the HMA.

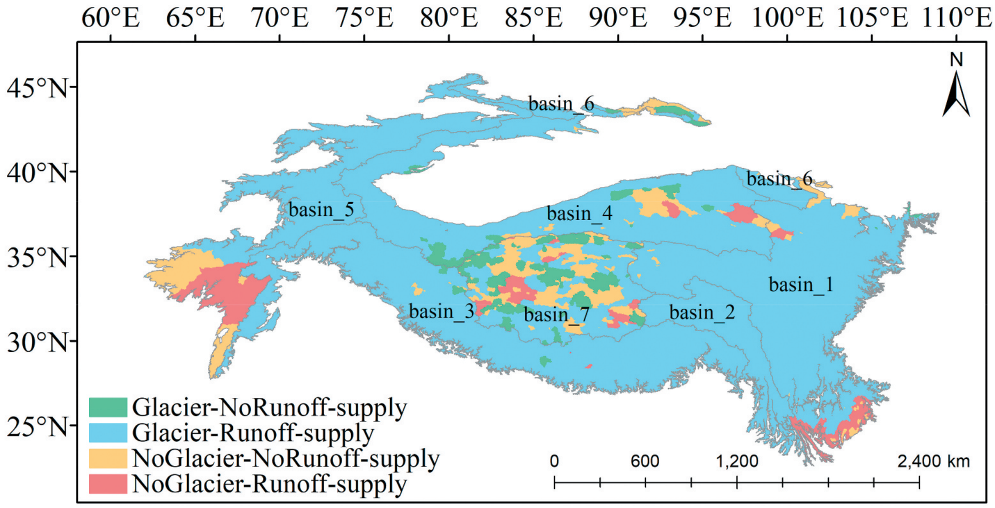


Figure 1. Classification map of surface water supply types across different basins in the High Mountain Asia region (Lu et al., 2020).

Table 1. Range of each basin in Figure 1.

Basin	Range
Yellow–Yangtze (basin_1)	Parts of the Yellow River and the Yangtze River
Mekong (basin_2)	Parts of the Mekong River
Brahmaputra–Ganges–Indus (basin_3)	Parts of the Brahmaputra, Ganges, and Indus Rivers
Tarim (basin_4)	Parts of the Tianshan region, Tarim Basin, and Qaidam Basin
Amu Darya–Syr Darya (basin_5)	Parts of the Amu Darya River and the surrounding areas in the western part of HMA
Qilian Mountains–Eastern Tianshan (basin_6)	Parts of the Tianshan region and the northern edge of HMA
Inner (basin_7)	Parts of the endorheic basin

(2) Creation of Lake/Reservoir dataset.
Building upon the RD2010 dataset and incorporating it with the WSW2020 dataset for HMA, we developed a lake/reservoir dataset. This step involved the meticulous removal of misclassified features such as rivers, glaciers, and mountain shadows from the WSW2020 dataset to enhance the accuracy of the LR dataset.

(3) Creation of River-Lake Spatial Connectivity dataset.
Water body data were identified for areas where rivers were connected to lakes and reservoirs. By creating buffer zones for the lake boundaries and intersecting them with the vector domain of the nearest river, it was observed that lakes were hydrologically connected to rivers. After careful manual verification and supplementation, we finalized the third dataset that represents river–lake spatial connectivity, named the river–lake spatial connectivity dataset.

Following Steps (1)–(3), we constructed datasets of river networks, lakes/reservoirs, and river-lake spatial connectivity. Subsequently, a new temporal source field named *Source* was appended to the attribute table to record the temporal origin of each geographic feature. This field contains three categories: 2010 (derived exclusively from the RD2010

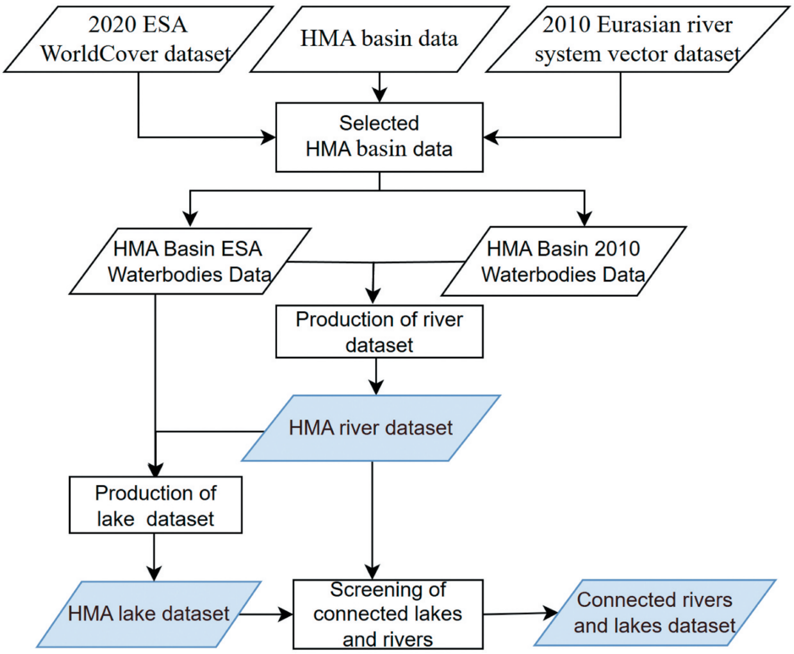


Figure 2. Workflow for creation of the integrated river network dataset, lake/reservoir dataset, and river-lake spatial connectivity dataset in the High Mountain Asia, representing the years 2010 and 2020.

dataset), 2020 (derived exclusively from the WSW2020 dataset), and "merged" (generated by integrating the two datasets). These metadata enable users to distinguish the original data sources of each polygon clearly.

3.1. Construction of river system datasets

The river system dataset was constructed as follows: To minimize the uncertainties introduced during the integration process, multiple quality control steps were applied, including the removal of misclassified features (e.g., glaciers and shadows), cross-referencing with RGI, and visual inspection using high-resolution Google Earth imagery.

(1) ESA land cover data were obtained from the European Space Agency and used to characterize the surface features of the basins classified by HMA water supply patterns. Using the basin boundaries defined by the HMA surface water-supply patterns (Lu et al., 2020), the ESA land cover data were mosaicked and clipped using GIS tools to generate raster images for the corresponding basin areas. Water body classes were then extracted from these raster datasets and converted into a vector format, yielding detailed boundary and shape information of the rivers and lakes/reservoirs suitable for spatial overlay analysis with WSW2020.
(2) Spatial overlay analyses were performed Using the WSW2020 and the RD2010. Intersecting and non-intersecting segments were identified based on spatial relationships (e.g., intersections and disjoints). For the intersecting areas, GIS tools were

used to merge and integrate the overlapping vectors, producing the final river vector dataset for these regions (hereafter referred to as W1).

(3) For non-intersecting river vectors, the considerations included additional rivers, lakes/reservoirs, glaciers, and mountain shadows present in the ESA water data were considered. The lake and shadow vectors were manually removed using Google Earth. Glacier vectors were eliminated using the Randolph Glacier Inventory (RGI) as a mask, yielding the final non-intersecting ESA river vectors (hereafter referred to as W2). Subsequently, W1, W2, and the non-intersecting segments of the 2010 river vectors were merged to form a comprehensive HMA river-vector dataset.

(4) To enhance the accuracy and reliability of this dataset, the merged results were further validated and refined using the RGI glacier data and the HydroRIVERS global river network dataset. The final product was an updated and integrated HMA river vector dataset that encompassed river vectors for both 2010 and 2020.

3.2. Construction of lake datasets

The lake dataset was constructed as follows: Similar quality control procedures were used to ensure reliability.

(1) Water body data (WSW2020) extracted from the ESA WorldCover, including lakes, reservoirs, and rivers, were overlaid with the HMA river vector dataset developed in Section 3.1. River features were removed to avoid duplication, resulting in a dataset that primarily contained lakes/reservoirs, wetlands, and minor glacier and terrain shadow features.

(2) RGI glacier data were used to eliminate glacier features from the lake dataset. By comparing the RGI and ESA data, we observed that most glacier features were accurately removed. The remaining potential glacier remnants and terrain shadows were further assessed to identify non-water body areas.

(3) The lake dataset was verified using Google Earth imagery to remove the residual glaciers and terrain shadows. This process refined the HMA lake-vector datasets for 2010 and 2020.

3.3. Construction of river-lake spatial connectivity datasets

Based on the river and lake vector datasets developed in Sections 3.1 and 3.2, we constructed a river–lake spatial connectivity dataset for the HMA. As direct connections between rivers and lakes are not always explicitly represented in vector datasets, a combination of automated spatial analysis and manual validation was applied to detect potential hydrological linkages.

3.3.1. Automated spatial connectivity detection

Rivers and lakes may be connected through minor tributaries, wetlands, or seasonal channels, which are often under-represented in vector datasets. To capture these potential linkages, a buffer analysis was applied around each lake/reservoir to approximate its area of influence and identify the intersecting river segments. Buffer zones ranging from

10 m to 400 m at 10 m intervals were generated for each lake/reservoir. Overlay analysis using river vector data indicated that when the buffer distance was less than 100 m, the number of intersections between lakes and rivers increased rapidly, whereas beyond this threshold, the growth rate gradually stabilized. This suggests that a 100 m buffer can capture the most hydrologically meaningful connections while reducing the overestimation caused by spatial proximity. Therefore, a 100 m buffer distance was selected as a balanced and robust threshold for defining river–lake spatial connectivity, used to construct the initial spatial connectivity dataset, and serve as the basis for subsequent analyses.

3.3.2. Manual refinement processing

To improve the completeness of the dataset, manual inspections were conducted using Google Earth imagery captured at nearly the same time. This step enabled the identification of tributaries in the vicinity of lakes that were not captured by automated analysis, thereby enhancing the representation of subtle but significant hydrological linkages.

Through this combined approach, we finalized the river–lake spatial connectivity vector dataset for the HMA, providing a comprehensive basis for subsequent hydrological analyses.

4. Validation and change analysis of the dataset

4.1. Validation and comparison assessment

4.1.1. Comparison with images in Google Earth

The dataset included 418,206 lakes and a region-wide vector river network. Among them, 405,797 lakes were less than 0.1 km^2.

Representative regions were selected from three developed waterbody datasets for comparative analysis using Google Earth images. The results are shown in Figure 3.

For the river vector dataset, Figure 3(a) shows that the data for a section of the Yangtze River Basin exhibit clear contours and a complete overall structure, indicating high accuracy in river delineation.

For the lake vector dataset, Figure 3(b) shows a comparison of Qinghai Lake and its surrounding areas. The enlarged view on the right demonstrates that the outlines of the Qinghai Lake and the adjacent small water bodies are distinct and complete. Moreover, the boundaries between lakes and rivers are clearly defined, facilitating effective differentiation between various water body types.

For the river–lake spatial connectivity dataset, Figure 3(c) shows that the developed vector data accurately reflect the hydrological connections between rivers, lakes, and reservoirs. The dataset successfully encompasses river segments that are connected to lakes, ensuring a comprehensive representation of the spatial connectivity of water bodies.

4.1.2. Comparison with public water body datasets

The dataset was further compared spatially with the JRC Global Surface Water dataset and National Fundamental Geographic Information Database at a 1:250 000 Scale (Liu et al., 2016). As shown in Figure 4, our dataset demonstrates improved spatial detail compared

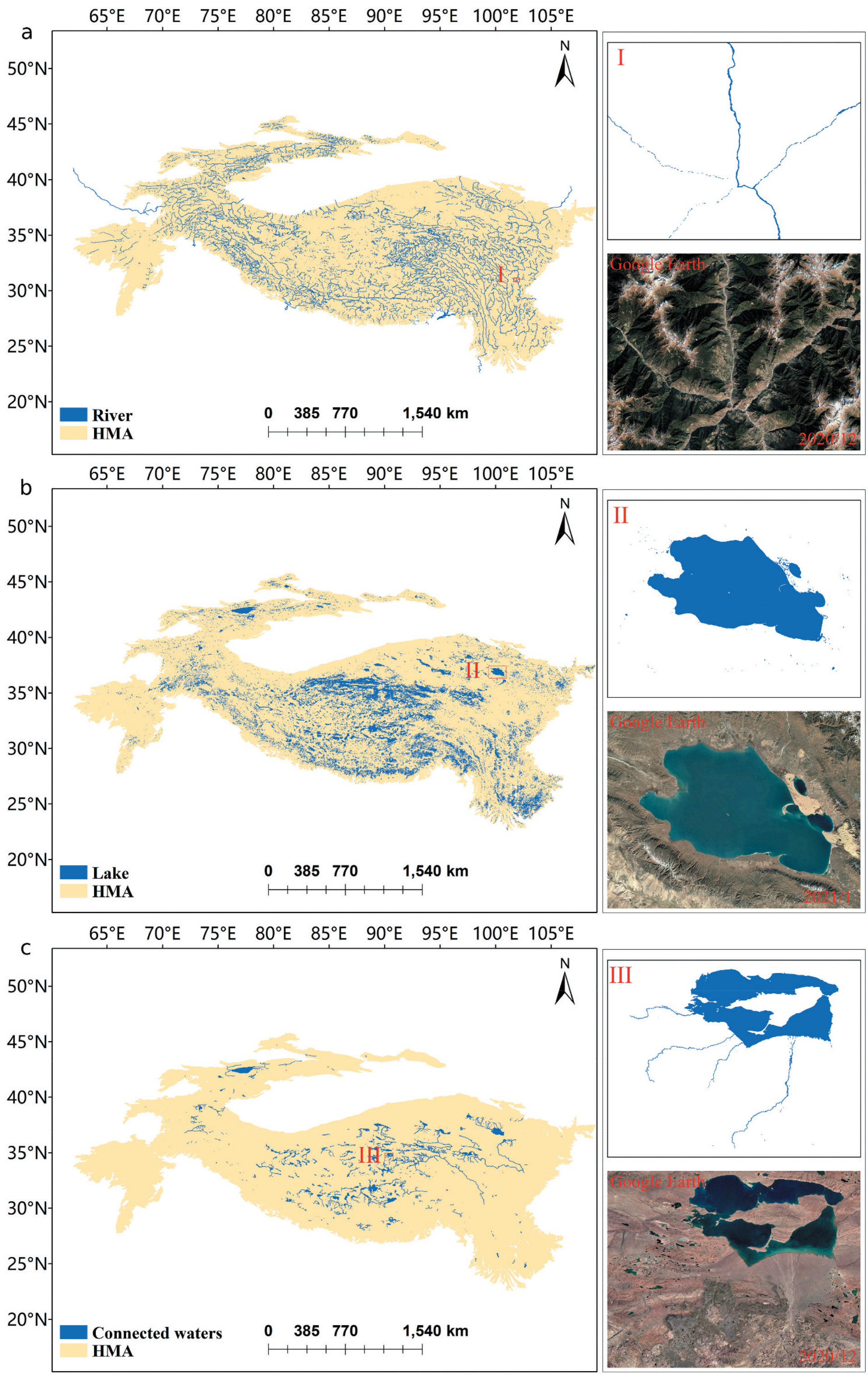


Figure 3. Comparison of the High Mountain Asia: (a) River dataset, (b) Lake datasets, and (c) River–lake spatial connectivity vector datasets with Google Earth imageries. Regions I, II, and III correspond to selected areas from the Yangtze River Basin, Yellow River Basin, and endorheic basin, respectively.

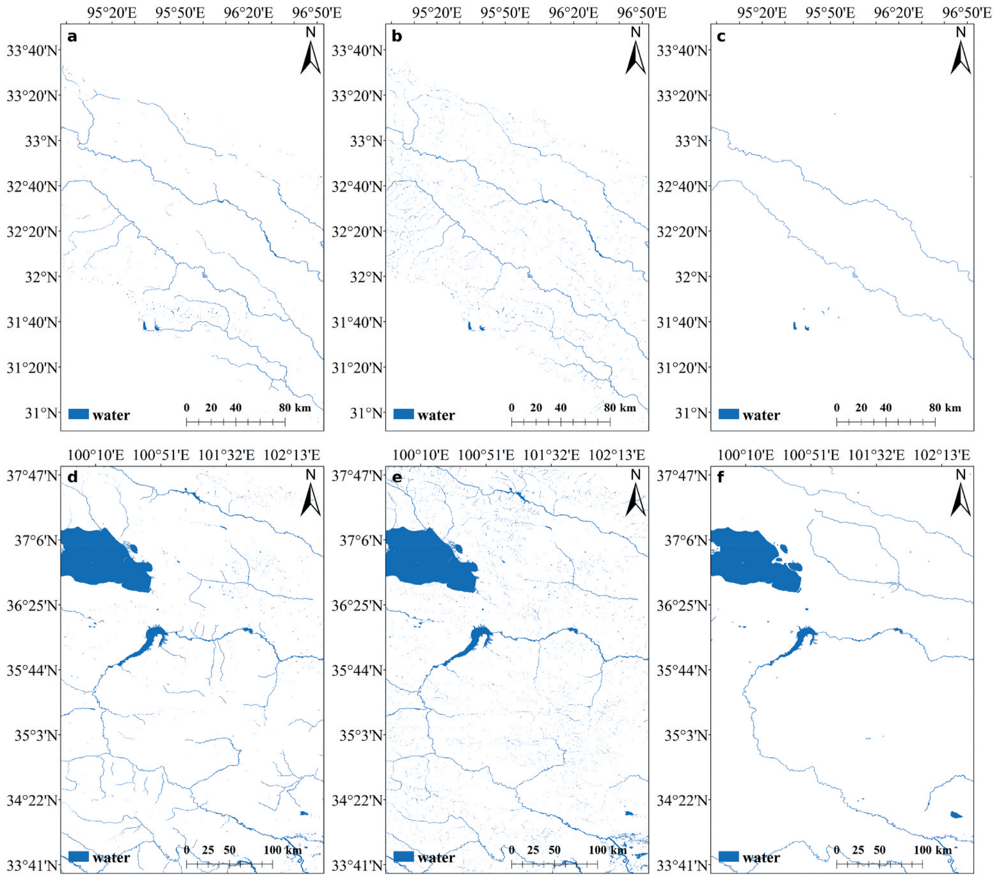


Figure 4. Comparative visualizations of the water body datasets: (a, d) Dataset generated in this study, (b, e) JRC Global Surface Water dataset, and (c, f) National fundamental geographic information database. Panels (a–c) cover the Mekong River Basin, while panels (d–f) cover the Yellow River Basin.

with both the JRC and national databases. Notably, the national database lacks representation of river tributaries and small lakes/reservoirs. Compared with the JRC dataset, the developed dataset effectively removed most glacier-related features, fewer river discontinuities, and more complete river connectivity with clearer contours.

Overall, the spatial distribution of rivers and lakes in our dataset closely aligns with that of the JRC data. Using the JRC data as a benchmark, we assessed the consistency between our dataset and the JRC dataset in the various HMA basins (the basins are shown in Figure 1). The statistical results (Table 2) indicated an overall consistency of 89.39% between the two datasets. The high spatial consistency with independent datasets (JRC) and visual agreement with Google Earth imagery further supported the reliability of the data integration and processing workflow.

The observed differences arise mainly from distinct data sources and production methodologies. The ESA WorldCover dataset, used as one of the primary inputs in this study, has a higher spatial resolution (10 m), enabling a more precise delineation of water body boundaries and details, and primarily represents permanent water bodies. In contrast, the

Table 2. Comparison of water body areas between this study and the JRC dataset across HMA basins.

Basin	This study (km^2)	Overlapping area (km^2)	JRC water area (km^2)	JRC permanent water area (km^2)	Permanent water proportion (%)	Consistency (%)	Basin level
Basin_1	14,626.9	14,487.3	15,328.7	11,208.0	73.1%	94.5%	Lev02
Basin_2	1,436.5	1,412.0	1,518.5	908.5	59.8%	93.0%	Lev02
Basin_3	8,040.4	7,933.8	9,706.1	5,758.4	59.3%	81.7%	Lev02
Basin_4	10,223.7	10,132.9	13,348.5	8,324.7	62.4%	75.9%	Lev02
Basin_5	8,503.7	8,457.9	9,043.4	7,965.2	88.1%	93.5%	Lev02
Basin_6	688.2	679.3	822.1	603.8	73.4%	82.6%	Lev02
Basin_7	33,712.2	33,688.9	36,137.4	31,170.0	86.3%	93.2%	Lev02
Total	77,231.5	76,792.0	85,904.6	65,937.6	76.8%	89.4%	

JRC Global Surface Water dataset incorporates extensive temporal information derived from long-term observations, capturing the full range of water occurrence within each year, and explicitly distinguishing between permanent and seasonal water. This allows for the analysis of long-term dynamics, although the spatial resolution is relatively coarse.

Further analysis showed that basins with a lower proportion of permanent water bodies in the JRC dataset also exhibited lower consistency with our dataset. For example, Basin_1 (the Yellow and Yangtze Rivers) and Basin_2 (Mekong River) have minor glacier coverage, as confirmed by RGI glacier data, and seasonal variations have a limited impact on the identification of permanent water bodies, leading to minor differences between our dataset and the JRC data. In contrast, Basin_3 (the Brahmaputra, Ganges, and Indus Rivers), Basin_4 (the Tianshan region, Tarim Basin, and Qaidam Basin), and Basin_6 (the Tianshan region) exhibited lower consistency. These regions are characterized by extensive glacier coverage, which likely contributes to seasonal fluctuations in the water extent in the JRC dataset due to glacier meltwater. In this study, the basin delineation was based on the HydroBASINS dataset. Level 2 (lev02) classification represents large drainage basins at the subcontinental scale (e.g., the Indus, Ganges, and Mekong River basins).

4.2. Analysis of changes of rivers and lakes

4.2.1. Changes of river and lake area with various recharge types

To further analyze the dynamic changes of rivers and lakes in the study area, the differences in area changes under various recharge types were investigated. Using our water body data as the base mask, the JRC river and lake data for the years 2000–2019 in the HMA were extracted. The classification of glacier-fed and non-glacier-fed water bodies followed the basin-based recharge type dataset proposed by Lu et al. (2020), in which basins containing glacier contributions were defined as glacier-fed. These data were classified and statistically analyzed based on the surface water recharge methods, as shown in Figure 5, which illustrates the changes in the JRC water body areas (permanent and seasonal) and permanent water body areas.

Table 3 shows that in 2000, the total HMA lake area was 56,790.1 km^2. Glacier-fed lakes (G-Lake) accounted for 48,561.5 km^2, and non-glacier-fed lakes (NG-Lake) accounted for 8,228.6 km^2. By 2019, the total lake area had increased to 68,763.1 km^2, G-Lake area to 57,118.6 km^2, and NG-Lake area to 11,644.5 km^2. The total area of the HMA River also increased. In 2000, it was 8,140.6 km^2, with a G-River area covering 8,013.0 km^2, and an

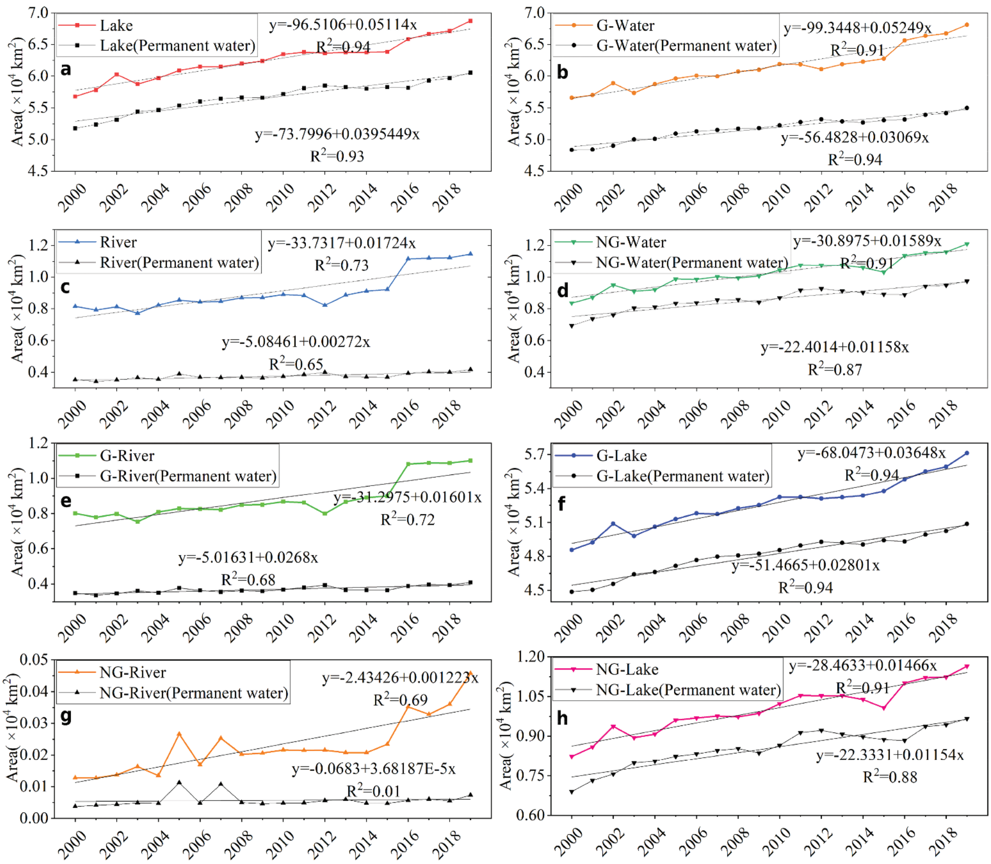


Figure 5. The temporal variations in the areas of rivers, lakes, reservoirs, and permanent water bodies in High Mountain Asia from 2000 to 2019, as derived from the JRC water dataset under different water supply modes. "G" denotes glacier-fed water bodies, while "NG" represents non-glacier-fed water bodies. (a, c) Total lake and river areas; (b, d) Glacier-fed and non-glacier-fed water bodies, (e, g) Glacier-fed rivers and non-glacier-fed rivers, and (f, h) Glacier-fed lakes and non-glacier-fed lakes.

Table 3. Comparison of water body area increases in different recharge types in the High Mountain Asia from 2000 to 2019.

Type	Area in 2000 (km^2)	Proportion in 2000 (%)	Area in 2019 (km^2)	Proportion in 2019 (%)	Change (2000–2019) (%)
Lake	56,790.1	–	68,763.1	–	+21.1%
G-Lake	48,561.5	85.5%	57,118.6	83.1%	+17.6%
NG-Lake	8,228.6	14.5%	11,644.5	16.9%	+41.5%
River	8,140.6	–	11,462.6	–	+40.8%
G-River	8,013.0	98.4%	11,005.4	96.0%	+37.3%
NG-River	127.7	1.6%	457.3	4.0%	+258.2%
Total	64,930.7	–	80,225.7	–	+23.5%

NG-River area covering 127.7 km^2. By 2019, the total river area had grown to 11,462.6 km^2, with a G-River area of 11,005.4 km^2 and NG-River area of 457.3 km^2.

Overall, both the river and lake areas showed a large increasing trend, with a total growth of 23.5% over 19 years. The lake and reservoir areas accounted for 78.3% of the total growth, while the river area increased by 40.8%, or approximately twice as much as the lake area. The increase in glacier-fed rivers and lakes accounted for 75.5% of the total growth, whereas the increase was greater in non-glacier-fed rivers and lakes (44.8%) than in glacier-fed rivers and lakes (20.4%). To understand why the area increase was much larger in rivers than in lakes, we statistically analyzed permanent water bodies in the HMA from 2000 to 2019 using JRC data. The results revealed that the HMA Rivers experienced large seasonal variations. In non-glacier-fed areas, seasonal variations in water bodies remained pronounced, indicating that other sources of water supply (such as precipitation, other surface water, and groundwater recharge) may have played a significant role in surface runoff changes in this region. The amplitude of annual variation was similar in the permanent lakes and lakes (Figure 5(a)).

Considering the long-term trend of river and lake areas from 2000 to 2019, both glacier-fed and non-glacier-fed water bodies exhibited an increasing trend, with glacier-fed water bodies showing a growth rate that was more than three times larger than that of non-glacier-fed water bodies (Figure 5(a)). This trend may be closely related to climate warming, such as increased glacier melt, leading to the expansion of glacier-fed water bodies. Additionally, the results indicated that 2016 was a specific year with higher water body changes than in other years. This sudden increase may be related to the accelerated glacier melting and increased precipitation.

4.2.2. Changes of river and lake area considering the connectivity

To consider the impact of river and lake connectivity, this study analyzed the changes in connected water bodies in the HMA endorheic basin (Figure 6). The extent of the basin was determined based on data from Zhang et al. (2013). Overall, the lake and river areas in the endorheic basin show an increasing trend. As in Table 4, The lake area increased from 30,332.0 km^2 in 2000 to 38,365.6 km^2 in 2019, and the river area was doubled, from 902.1 km^2 to 1,846.3 km^2. The area of connected lakes (Lake_con) increased from 20,781.7 km^2 to 25,395.6 km^2, indicating that lake expansion could have been influenced by river input. Unconnected lakes (Lake_uncon) increased from 9,550.3 km^2 to 12,970.0 km^2 mainly due to the expansion of the existing lakes, but could also be related to other factors such as increased precipitation. In HMA endorheic lake areas, glacier melt is an important factor for lake expansion, and G-Lakes account for 67.7% of the total lake area increase.

To analyze the reasons for the growth of rivers and lakes in the HMA endorheic basin, we used MODIS daily cloud-free snow cover products over the Tibetan Plateau from 2003 to 2019 (Qiu et al., 2016) and daily gridded precipitation data from 2001 to 2019 (Han et al., 2023) to analyze the changes in annual average snow cover area and precipitation. The results (Figure 6(e,f)) show that the snow cover area had no trend, whereas precipitation increased significantly. This suggests that the main factor for the growth of rivers and lakes in the endorheic basin is precipitation. Indeed, changes in annual precipitation were closely synchronized with variations in water body areas. For example, the average annual precipitation in 2012 has decreased by 14% since 2011. Consequently, the number of connected rivers decreased by 15.6% and the number of glacier-fed rivers declined by 14.6%. In contrast, compared with 2015, the average annual precipitation in 2016 increased by 27%. Consequently, the number of connected rivers increased by 52.2%, and the number of glacier-fed rivers increased by 50.5%. The slight decrease in lake area in 2015 may be related to the 2015/2016 El Niño phenomenon (Lei et al., 2019; Zhang et al., 2020).

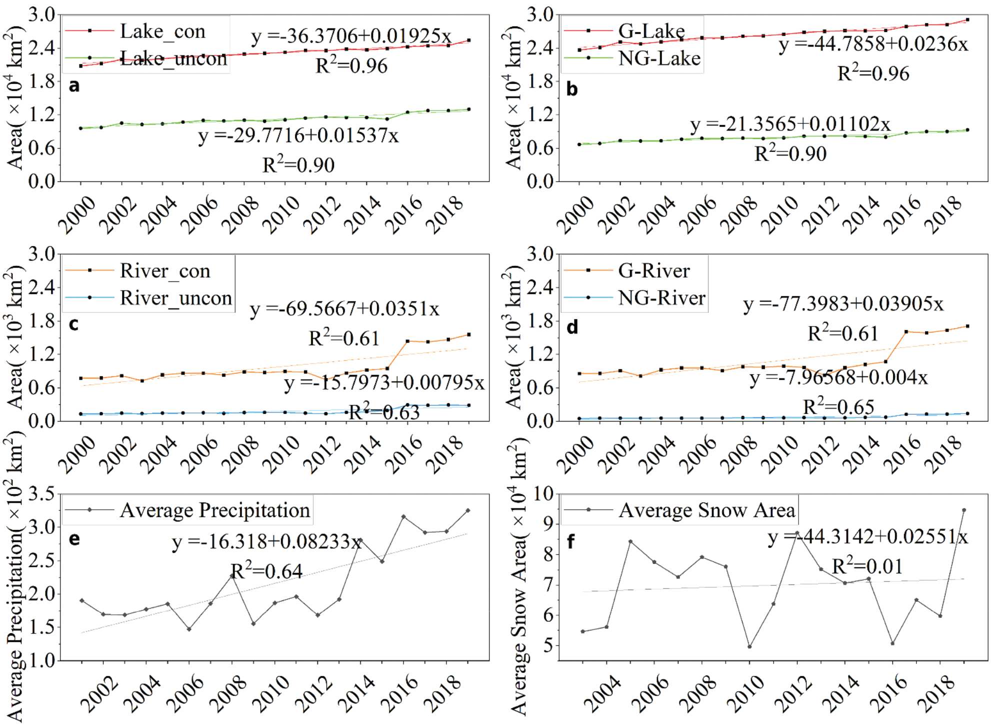


Figure 6. The temporal variations in rivers and lakes within the endorheic basins of High Mountain Asia. (a, c) Temporal variations in the areas of connected lakes and connected rivers during 2000–2019. (b, d) Area changes in lakes and rivers grouped by water-resource recharge mode during 2000–2019. (e) Temporal changes in precipitation during 2001–2019. (f) Temporal changes in snow cover area during 2003–2019.

Table 4. Changes in water body area in endorheic basins in High Mountain Asia from 2000 to 2019.

Type	Area in 2000 (km^2)	Area in 2019 (km^2)	Change (2000–2019)
Total Lake	30,332.0	38,365.6	+26.5%
Lake_con	20,781.7	25,395.6	+22.2%
Lake_uncon	9,550.3	12,970.0	+35.8%
River	902.1	1,846.3	+104.7%
River_con	771.1	1,560.2	+102.3%
River_uncon	131.0	286.1	+118.4%
Total	31,234.0	40,211.9	+28.8%

5. Discussion

5.1. Limitations of the dataset

In the HMA, the presence of mountain shadows, snow, ice, and other cryosphere elements poses challenges for accurate water body extraction from remote sensing data (Khalid et al., 2021). The data utilized in this study were developed for global-scale applications and may not fully capture the unique topographical and environmental characteristics of the HMA, potentially affecting their regional accuracy. Furthermore, owing to the spatial and temporal resolution limitations of the input datasets (30 m for RD2010 and 10 m for WSW2020), the resulting river network inevitably exhibits polygon fragmentation and discontinuities, particularly for narrow streams and complex terrains. Consequently, river segments connected to lakes may be omitted, particularly for small lakes or those in regions with limited high-quality optical observations. Additionally, buffer analysis was employed in this study to provide a preliminary delineation of river–lake spatial connectivity rather than a fully resolved hydrological topology. However, in reality, the connectivity between rivers and lakes is often influenced by seasonal variability and anthropogenic activities, such as reservoir construction, which have not been fully accounted for in the current dataset. These limitations suggest that further refinement of both the input datasets and connectivity characterization is necessary in future work.

In Section 4.2, the proposed dataset was used as a mask to extract annual JRC water body data for the period 2000–2019, which were subsequently used to analyze temporal variations in surface water. This approach may have resulted in a small number of omitted water bodies in our dataset, which were not included in the analysis. Nevertheless, given that this study focuses on regional-scale patterns of water body dynamics, the influence of such omissions on overall trends is expected to be limited.

5.2. Applications and extensions of the dataset

The dataset developed in this study provides robust support for spatial delineation, automated identification, and area estimation of rivers and lakes, thereby enhancing the efficiency of spatial information extraction. In particular, the river-lake spatial connectivity information serves as a fundamental dataset for studying changes in connected water bodies. Understanding the connectivity between water bodies aids in understanding fundamental river-lake geomorphology (Nyberg et al., 2023), assessing the impacts of connectivity alterations on human activities (Lei et al., 2023), and evaluating associated ecological and environmental responses (Li et al., 2024). The proposed fundamental dataset offers information on the extent of rivers and lakes, facilitates rapid distinction and extraction of water body types, and identifies basic spatial connectivity among water bodies, thereby enabling swift assessment of regional hydrological changes. For instance, by integrating our dataset with the JRC data, we conducted a preliminary analysis to classify water bodies in the HMA. Further integration with environmental variables, such as precipitation, evapotranspiration, glacier dynamics, snow cover, and snow water equivalent, can support comprehensive evaluations of water resource changes across different basins in the HMA.

Moreover, this dataset can function as a base map for water body studies with high temporal resolution, suitable for periodic (e.g., every 12 days) automated water body extraction using Sentinel satellite data (Liang et al., 2024), thereby reducing interference from land features and enhancing the accuracy and efficiency of the target water body extraction.

Constructed in vector data format, this dataset exhibits excellent flexibility and scalability. Researchers can freely combine these three types of data to generate thematic datasets, such as non-connected rivers and isolated lakes. Additionally, integration with other vector-based dataset data from other studies allows for further refinement of the classifications; for example, combining basin-scale water supply type data to distinguish rivers and lakes with different sources of replenishment. To represent the extent of rivers and lakes, this dataset offers precise spatial positioning and comprehensive classification attributes, serving as a base map for the integration and updating of high-resolution, multi-source water body vector data in future research.

Beyond the example analysis presented in this study, the dataset also has potential applications in cryosphere-related research on HMA. Detailed river network and river–lake spatial connectivity information can support studies of glacial meltwater routing and downstream hydrological connectivity. In combination with glacier-fed and non-glacier-fed classifications, the dataset enabled differentiation of hydrological responses associated with glacier meltwater contributions. Furthermore, it can facilitate investigations of cryosphere–hydrology–ecosystem interactions, such as assessing how changes in glacier meltwater supply influence lake expansion, river connectivity, and downstream aquatic habitats. The dataset may also be integrated with glacier, snow cover, and permafrost datasets to explore the impact of cryospheric change on regional water resources and ecosystem dynamics.

Recently developed datasets, such as those presented by Sui et al. (2025), provide high-resolution surface-water mapping products with improved capability for detecting small water bodies. These datasets offer opportunities for future integration to further enhance the quality and completeness of the results.

6. Conclusion

This study developed a comprehensive vector dataset encompassing rivers, lakes, reservoirs, and their spatial connectivity across HMA. The primary objective was to achieve effective spatial separation between rivers and lakes and to provide systematic information on their spatial connectivity. Based on the 2010 vector dataset of river systems across the Eurasian continent and the 2020 ESA WorldCover land cover classification data, this study employed vector operations, river-lake classification, and identification of spatial connectivity between Rivers, Lakes to produce three vector datasets for the HMA region, Rivers, Lakes (including reservoirs), and River–Lake Spatial Connectivity, which integrates the water surface information from both 2010 and 2020 data. In constructing the river dataset, spatial intersections and area filtering between the ESA WorldCover data and 2010 river data were conducted to delineate precise river vector boundaries. The supplementary data were used for refinement and correction to enhance the completeness and continuity of the river network. For the lake and reservoir datasets, a stepwise exclusion process was implemented by subtracting the river data and incorporating RGI glacier vectors and Google Earth imagery for visual verification, effectively eliminating misclassified glaciers, mountain shadows, and other non-lake/reservoir features. The river-lake spatial connectivity dataset was generated through multilevel buffer analysis of lakes/reservoirs combined with manual verification to accurately identify the connected waterbodies. Comparisons with Google Earth imagery revealed high spatial accuracy in the distribution of water bodies across all three datasets. Compared to the JRC Global Surface Water dataset, an overall consistency of 89.39% was achieved. Notably, this new dataset distinguishes between rivers and lakes, presents more complete river networks in certain regions, explicitly removes glaciers and shadows, and provides fundamental information on river–lake spatial connectivity.

The constructed dataset was then applied to analyze water body changes in the HMA from 2000 to 2019. The total water body area exhibited a 23.5% increase with lakes and reservoirs, accounting for 78.3% of that. The river areas increased by 40.8%, and the lakes/reservoirs increased by 21.1%. Waterbodies fed by glacial melt contributed to 75.5% of the total area increase, with non-glacial-fed waters experiencing a greater area increase (44.8%) than glacial-fed cases (20.4%). The areas of glacial-fed rivers and lakes were significantly larger than those of their non-glacial-fed counterparts, and rivers exhibited more pronounced seasonal variations. Overall, glacial-fed water bodies demonstrated higher areas and growth rates than non-glacial-fed water bodies. In the endorheic basin, analysis of connected water bodies from 2000 to 2019 showed a consistent annual increase in the total lake area. The area of connected lakes expanded from 20,782 km^2 to 25,396 km^2, while that of disconnected lakes grew from 9,550 km^2 to 12,970 km^2.

The dataset developed in this study provides a fundamental resource for spatial extraction and automated calculation of river and lake/reservoir water bodies, facilitating more precise detection of their changes. It serves as a crucial data source for surface water resource management and change analysis in the HMA. The river, lake/reservoir, and river-lake spatial connectivity vector datasets can support future research involving high-resolution, high-frequency remote sensing for waterbody monitoring in HMA, enabling efficient batch extraction of waterbody extents and serving as baseline vector data for waterbody studies in the region.

Acknowledgements

We sincerely thank the providers of the datasets that made this study possible. The vector dataset of the river systems originating from Eurasia to the Arctic Ocean was obtained from the National Tibetan Plateau Data Center. The ESA WorldCover 10 m 2020 product, provided by the European Space Agency, is essential for identifying permanent water bodies. We also acknowledge the JRC Global Surface Water dataset developed by the European Commission Joint Research Centre and HydroSHEDS data from the World Wildlife Fund (WWF) for hydrological referencing. The Randolph Glacier Inventory V6.0 (RGI) is critical for identifying the glacier-fed basins. The basin classification data based on water resource supply modes were provided by the Science Data Bank. We are also grateful to the teams and institutions that provided analytical data and computational tools that were instrumental in performing spatial statistics, precipitation analysis, and snow cover assessments throughout this study. We gratefully acknowledge Professor Dongfeng Li from Peking University for his insightful suggestions on the formulation of this paper.

Author contributions

CRediT: **Zhaocheng Wang:** Data curation, Formal analysis, Methodology, Validation, Visualization, Writing – original draft; **Lijuan Shi:** Methodology, Supervision, Writing – review & editing; **Yubao Qiu:** Conceptualization, Methodology, Project administration, Supervision, Writing – review & editing; **Guoqiang Jia:** Writing – review & editing; **Xiaoting Li:** Writing – review & editing; **Matti Leppäranta:** Writing – review & editing; **Lanhai Li:** Writing – review & editing; **Massimo Menenti:** Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Open Research Project of the Xinjiang Key Laboratory of Water Cycle and Utilization in Arid Zone, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences [No. XJYS0907-2023-03], the the National Natural Science Foundation of China [No. 42101389], the National Key Research and Development Program of China [No. 2022YFF0711700, No. 2021YFE0116800], and the Chinese Academy of Sciences (CAS) President's International Fellowship Initiative [No. 2021VTA0007].

Notes on contributors

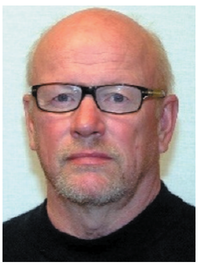
Zhaocheng Wang is a master degree candidate at the School of Computer and Information Security, Guilin University of Electronic Technology. His research interests include surface water extraction by remote sensing technology, and spatiotemporal analysis, with a focus on the dynamics and spatial connectivity of rivers and lakes.

Lijuan Shi is a Professor at Henan University, supported by the High-Level Talent Program. She received her Ph.D. degree in Cartography and Geographic Information Systems from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (CAS), Beijing, China, in 2018. She was a Visiting Research Scientist at the Finnish Meteorological Institute (FMI) from 2016 to 2017 and a Postdoctoral Researcher at the Aerospace Information Research Institute (CAS) from 2018 to 2024. Her research focuses on remote sensing retrieval and physical modeling of snow, ice, water, and atmospheric moisture, as well as intelligent geospatial information processing and remote sensing algorithm development.

Yubao Qiu is a Research Professor at the Aerospace Information Research Institute of the Chinese Academy of Sciences (AIR-CAS). He currently leads the Digital Environment Unit under the International Research Center of Big Data for Sustainable Development Goals (CBAS). Since obtaining his Ph.D. in 2008, Dr. Qiu has specialized in microwave remote sensing, digital environment and sustainable development, and data science. From 2012 to 2015, he worked as a Scientific and Technical Officer at the Secretariat to the Group on Earth Observations (GEO) in Geneva, where he supported water and cold regions information services, ecosystems, and data sharing. In 2007, Dr. Qiu was a Visiting Research Scientist at the Helsinki University of Technology (TKK). In 2017, he began co-chairing the High Mountain and Polar Cold Regions Working Group under the Digital Belt and Road Program (DBAR-HiMAC). Now, he also co-leads the GEO Cold Regions Initiative (GEOCRI). He is an editorial board member for several international journals, including *Innovation Informatics* and *International Journal of Digital Earth*.

Guoqiang Jia is a Assistant Researcher with the Aerospace Information Research Institute, Chinese Academy of Sciences (AIR-CAS), and International Research Center of Big Data for Sustainable Development Goals (CBAS), Beijing. He received the D.Sc. degree in physical geography from the University of Chinese Academy of Sciences, Beijing, China in 2021. His research interests include hydrological remote sensing, rainfall-induced landslide modeling and assessment, and relevant impact and response to climate change.

Xiaoting Li is an Assistant Researcher at the Aerospace Information Research Institute, Chinese Academy of Sciences (CAS), and International Research Center of Big Data for Sustainable Development Goals (CBAS), Beijing. She received her Ph.D. degree in Earth and Human Settlement Environment Science and Engineering from Xi'an Jiaotong University, Xi'an, China in 2024. She was a Visiting Student at the Department of Ecology and Evolutionary Biology in Cornell University from 2022 to 2023. Her research focuses on cold region environmental and climate change, and remote sensing monitoring of vegetation dynamics, soil freeze-thaw processes, and snow cover.



Matti Leppäranta is Professor Emeritus of Geophysics at the University of Helsinki, Helsinki, Finland. He received the M.Sc. degree in mathematics and the Ph.D. degree in physical oceanography with the University of Helsinki, Helsinki, Finland, in 1976 and 1981, respectively. He has worked as a Research Physicist with the Finnish Institute of Marine Research, Helsinki, from 1974 to 1991. Since 1992, he has been a Professor with the University of Helsinki. He visited Cold Regions Research and Engineering Laboratory, Hanover, NH, USA, for his Post-Doctoral Researcher in 1983–1984, he has been a Visiting Teacher with UNIS, Norway, in 1996–2014, and spent sabbatical in Hokkaido University, Sapporo, Japan, in spring semester 2003 and fall semester 2012. He presently works with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. He has published books *Drift Of Sea Ice* (2004, Second Edition 2011) and *Freezing Lakes And Evolution Of Their Ice Cover* (2015, Second Edition 2023). His primary research field has been geophysics of sea ice and lake ice, based on field studies, remote sensing, and mathematical modelling.

Lanhai Li is a Professor in the College of Geoinformatics, Zhejiang University of Technology. With more than 35 years of research, consulting and teaching experience, he has worked in a variety of positions for institutes, universities, international organizations and private companies in China, England, Thailand, Japan and Canada. Dr. Li's primary research interest focuses on soil and water resources management, natural hazards and climate change impact assessment. He has extensive working experience in hydrological modeling, risk and reliability assessment of hydrogeological disasters and climate change impact assessment. He has undertaken applied research projects that integrate the mathematical modelling, data-base management and geographic information systems into decision support tools. He has published more than 200 papers in peer reviewed journals and three monographs during academic research. The publications have been widely referred by academic research and engineering design, and his research and consulting work have been recognized by a number of awards for excellence in research and outreach, and by invitations from Northern American, Asian and African organizations for seminars and workshops as well as hosting international conferences.

Massimo Menenti received the Laurea degree in physics from the Università di Roma, Rome, Italy, in 1972, and the Dr. Sc. degree in environmental sciences from the Wageningen University, The Netherlands. He has held senior research and faculty positions at Wageningen University, at the École Nationale Supérieure des Physique de Strasbourg (ENSPS), France. the National Research Council of Italy and the Delft University of Technology, The Netherlands. He is currently a professor at the Institute of Tibetan Plateau Research, Chinese Academy of Sciences, China, at the National Key Laboratory of Remote Sensing and Digital Earth, Aerospace Information Research Institute, Chinese Academy of Sciences, China and at the Delft University of Technology, The Netherlands. His research interests include land surface processes, remote sensing of land-atmosphere interactions, surface energy balance modeling, and hydrological cycle analysis. He has made ground-breaking and early contributions to the development of remote sensing methodologies for water resource management, climate change studies, and land surface temperature retrieval. His research integrates multi-source satellite observations with numerical models to improve environmental monitoring and prediction.

ORCID

Yubao Qiu http://orcid.org/0000-0003-1313-6313

AI disclosure statement

We used GPT-5 (OpenAI) only for minor language polishing during manuscript preparation. The use of the tool did not influence the scientific content, data analysis, results, or conclusions of this work. All intellectual contributions and interpretations are solely those of the authors.

Data availability statement

The dataset generated in this study is deposited in the repository of the National Cryosphere Desert Data Center (NCDDC) (https://doi.org/10.12072/ncdc.nieer.db6895.2025) and is also publicly accessible via Zenodo (https://zenodo.org/records/20236480).

References

Ayantobo, O. O., & Wei, J. (2025). Quantifying climate-induced changes in the Tibetan Plateau's hydrologic cycle through atmospheric moisture depletion and restoration times. *Water Resources Research*, *61*(2), e2024WR037644. https://doi.org/10.1029/2024WR037644

Bai, Y., Li, D., Wangchuk, S., Kettner, A., Zhao, Y., Deng, R., Liu, Y., Xiao, C., Ni, J., & Cui, P. (2025). Flood complexity and rising exposure risk in high mountain Asia under climate change. *Science Bulletin*, *70*(10), 1601–1604. https://doi.org/10.1016/j.scib.2025.01.055

Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, *2*(4), 291–308. https://doi.org/10.1080/17538940902951401

Chauhan, T., Devanand, A., Roxy, M. K., Ashok, K., & Ghosh, S. (2023). River interlinking alters land-atmosphere feedback and changes the Indian summer monsoon. *Nature Communications*, *14*(1), 5928. https://doi.org/10.1038/s41467-023-41668-x

Consortium, R. G. I. (2017). *Randolph glacier inventory-a dataset of global glacier outlines: Version 6.0: Technical report, global land ice measurements from space*. Boulder, Colorado, USA. Digital Media. doi: 10.7265/N5-RGI-60.

Fan, C., Song, C., Wang, J., Sheng, Y., Lin, Y., Yu, C., Safat Sikder, T., Crétaux, J.-F., Liu, K., Chen, T., Zeng, F., & Ke, L. (2024). Emerging global reservoirs in the new millennium: Abundance, hotspots, and total water storage. *Science Bulletin*, *69*(14), 2179–2182. https://doi.org/10.1016/j.scib.2024.04.043

Flores, J. A., Gleason, C. J., Brown, C., Vergopolan, N., Lummus, M. M., Stearns, L. A., Li, D., Andrews, L. C., Basnyat, D., Brinkerhoff, C. B., Ducusin, R., Feng, D., Friedmann, E., He, X., Girotto, M., Kumar, S. V., Lammers, R. B., Lamsal, G., Maina, F. Z., . . . Subedi, D. (2025). Accelerating river discharge in High Mountain Asia. *AGU Advances*, *6*(4), e2024AV001586. https://doi.org/10.1029/2024AV001586

Han, J., Miao, C., Gou, J., Zheng, H., Zhang, Q., & Guo, X. (2023). A new daily gridded precipitation dataset for the Chinese mainland based on gauge observations. *Earth System Science Data*, *15*(7), 3147–3161. https://doi.org/10.5194/essd-15-3147-2023

Khalid, H. W., Khalil, R. M. Z., & Qureshi, M. A. (2021). Evaluating spectral indices for water bodies extraction in Western Tibetan Plateau. *Egyptian Journal of Remote Sensing and Space Science*, *24*(3), 619–634. https://doi.org/10.1016/j.ejrs.2021.09.003

Khanal, S., Lutz, A. F., Kraaijenbrink, P. D. A., van den Hurk, B., Yao, T., & Immerzeel, W. W. (2021). Variable 21st century climate change response for rivers in high mountain Asia at seasonal to decadal time scales. *Water Resources Research*, *57*(5), e2020WR029266. https://doi.org/10.1029/2020WR029266

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Lei, Y., Yao, T., Sheng, Y., Yang, K., Yang, W., Li, S., Yu, Y., Jiang, Y., & Yu, Y. (2023). Unprecedented lake expansion in 2017-2018 on the Tibetan Plateau: Processes and environmental impacts. *Journal of Hydrology*, *619*, 129333. https://doi.org/10.1016/j.jhydrol.2023.129333

Lei, Y., Zhu, Y., Wang, B., Yao, T., Yang, K., Zhang, X., Zhai, J., & Ma, N. (2019). Extreme lake level changes on the Tibetan Plateau associated with the 2015/2016 El Niño. *Geophysical Research Letters*, *46*(11), 5889–5898. https://doi.org/10.1029/2019GL081946

Li, B., Wan, R., Yang, G., Yang, S., Dong, L., Cui, J., & Zhang, T. (2024). Centennial loss of lake wetlands in the Yangtze Plain, China: Impacts of land use changes accompanied by hydrological connectivity loss. *Water Research*, *256*, 121578. https://doi.org/10.1016/j.watres.2024.121578

Li, D., Lu, X., Overeem, I., Walling, D. E., Syvitski, J., Kettner, A. J., Bookhagen, B., Zhou, Y., & Zhang, T. (2021). Exceptional increases in fluvial sediment fluxes in a warmer and wetter High Mountain Asia. *Science*, *374*(6567), 599–603. https://doi.org/10.1126/science.abi9649

Li, D., Zhang, T., Walling, D. E., Lane, S., Bookhagen, B., Tian, S., Overeem, I., Syvitski, J., Kettner, A. J., Park, E., Koppes, M., Schmitt, R. J. P., Sun, W., Ni, J., & Ehlers, T. A. (2024). The competing controls of glaciers, precipitation, and vegetation on high-mountain fluvial sediment yields. *Science Advances*, *10*(48), eads6196. https://doi.org/10.1126/sciadv.ads6196

Li, X., Zhang, T., Yang, D., Wang, G., He, Z., & Li, L. (2023). Research on lake water level and its response to watershed climate change in Qinghai Lake from 1961 to 2019. *Frontiers in Environmental Science*, *11*, 1130443. https://doi.org/10.3389/fenvs.2023.1130443

Liang, W., Qiu, Y., Xie, P., Shi, L., & Wang, X. (2020). Vector dataset for river systems originating in Eurasia to the Arctic Ocean. *IOP Conference Series Earth and Environmental Science*, *502*(1), 012035. https://doi.org/10.1088/1755-1315/502/1/012035

Liang, Y., Qiu, Y., Jia, G., Zhao, N., Zhong, W., & Cui, H. (2024). A dataset of egg transaction prices in China's market from 2014 to 2021. *China Scientific Data*, *9*(3), 1–7. https://doi.org/10.11922/11-6035.nasdc.2023.0006.zh

Liu, H., & Wang, L. (2026). TPRD3km: An observation-constrained high-resolution runoff data for third Pole transboundary rivers. *Big Earth Data*, *10*(1), 656–684. https://doi.org/10.1080/20964471.2025.2585701

Liu, J., X, L., Zhang, Y., & Yang, M. (2016). Technical innovation and application for linkage updating of national fundamental geographic information database at 1: 250 000 scale. *Bulletin of Surveying and Mapping*, (4), 1–4. https://doi.org/10.13474/j.cnki.11-2246.2016.0109

Liu, S., Zhou, H., Zhao, L., Wang, L., Zou, D., Hu, G., Qiao, Y., Sun, Z., Li, Z., & Qiao, Y. (2025). The ground ice melting has accelerated the lake expansion in the hinterland of the permafrost region on the Qinghai-Tibetan Plateau. *Journal of Hydrology: Regional Studies*, *58*, 102268. https://doi.org/10.1016/j.ejrh.2025.102268

Loomis, B. D., Richey, A. S., Arendt, A. A., Appana, R., Deweese, Y.-J. C., Forman, B. A., Shean, D. E., Sabaka, T. J., & Shean, D. E. (2019). Water storage trends in High Mountain Asia. *Frontiers in Earth Science*, *7*. https://doi.org/10.3389/feart.2019.00235

Lu, J., Qiu, Y., Wang, X., Liang, W., Xie, P., Shi, L., Menenti, M., & Zhang, D. (2020). Constructing dataset of classified drainage areas based on surface water-supply patterns in High Mountain Asia. *Big Earth Data*, *4*(3), 225–241. https://doi.org/10.1080/20964471.2020.1766180

Nyberg, B., Henstra, G., Gawthorpe, R. L., Ravnås, R., & Ahokas, J. (2023). Global scale analysis on the extent of river channel belts. *Nature Communications*, *14*(1), 2163. https://doi.org/10.1038/s41467-023-37852-8

Nyberg, B., Sayre, R., & Luijendijk, E. (2024). Increasing seasonal variation in the extent of rivers and lakes from 1984 to 2022. *Hydrology and Earth System Sciences*, *28*(7), 1653–1663. https://doi.org/10.5194/hess-28-1653-2024

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*(7633), 418–422. https://doi.org/10.1038/nature20584

Phan, V. H., Lindenbergh, R. C., & Menenti, M. (2013). Geometric dependency of Tibetan lakes on glacial runoff. *Hydrology and Earth System Sciences*, *17*(10), 4061–4077. https://doi.org/10.5194/hess-17-4061-2013

Pi, X., Luo, Q., Feng, L., Xu, Y., Tang, J., Liang, X., Bryan, B. A., Cheng, R., Fensholt, R., Brandt, M., Cai, X., Gibson, L., Liu, J., Zheng, C., Li, W., & Bryan, B. A. (2022). Mapping global lake dynamics reveals the emerging roles of small lakes. *Nature Communications*, *13*(1), 5777. https://doi.org/10.1038/s41467-022-33239-3

Qiao, B., Zhu, L., & Yang, R. (2019). Temporal-spatial differences in lake water storage changes and their links to climate change throughout the Tibetan Plateau. *Remote Sensing of Environment*, *222*, 232–243. https://doi.org/10.1016/j.rse.2018.12.037

Qiu, Y., Guo, H., Chu, D., Zhang, H., Shi, J., Shi, L., Laba, Z., & Laba, Z. (2016). MODIS daily cloud-free snow cover products over Tibetan Plateau (2002–2015). *China Scientific Data*, *1*(1). https://doi.org/10.11922/csdata.170.2016.0003

Song, C., Sheng, Y., Zhan, S., Wang, J., Ke, L., & Liu, K. (2019). Impact of amplified evaporation due to lake expansion on the water budget across the inner Tibetan Plateau. *International Journal of Climatology*, *40*(4), 2091–2105. https://doi.org/10.1002/joc.6320

Sui, Y., Hu, Z., Zhang, K., Yan, D., Su, Y., Xu, J., Wang, R., & Feng, M. (2025). A comprehensive water bodies dataset of High-Mountain Asia. *Scientific Data*, *12*(1), 1350. https://doi.org/10.1038/s41597-025-05659-5

Wang, D., Wang, X., Huang, Y., Zhang, X., Zhang, W., Xin, Y., Qu, M., & Cao, Z. (2021). Impact analysis of small hydropower construction on river connectivity on the upper reaches of the great rivers in the Tibetan Plateau. *Global Ecology and Conservation*, *26*, e01496. https://doi.org/10.1016/j.gecco.2021.e01496

Wang, J., Chen, X., Liu, J., & Hu, Q. (2021). Changes of precipitation-runoff relationship induced by climate variation in a large glaciated basin of the Tibetan Plateau. *Journal of Geophysical Research Atmospheres*, *126*(21), e2020JD034367. https://doi.org/10.1029/2020JD034367

Wang, J., Wang, L., Li, M., Zhu, L., & Li, X. (2022). Lake volume variation in the endorheic basin of the Tibetan Plateau from 1989 to 2019. *Scientific Data*, *9*(1), 611. https://doi.org/10.1038/s41597-022-01711-w

Xu, F., Zhang, G., Woolway, R. I., Yang, K., Wada, Y., Wang, J., & Crétaux, J.-F. (2024). Widespread societal and ecological impacts from projected Tibetan Plateau lake expansion. *Nature Geoscience*, *17*(6), 516–523. https://doi.org/10.1038/s41561-024-01446-w

Yao, T., Wu, G., Xu, B., Wang, W., Gao, J., & An, B. (2019). Asian water tower change and its impacts. *Bulletin of Chinese Academy of Sciences (Chinese Version)*, *34*(11), 1203–1209. https://doi.org/10.16418/j.issn.1000-3045.2019.11.003

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., & Fritz, S. (2022). *ESA worldcover 10 m 2021 v200*.

Zhang, G., Luo, W., Chen, W., & Zheng, G. (2019). A robust but variable lake expansion on the Tibetan Plateau. *Science Bulletin*, *64*(18), 1306–1309. https://doi.org/10.1016/j.scib.2019.07.018

Zhang, G., Yao, T., Piao, S., Bolch, T., Xie, H., Chen, D., Zhang, H., O'Reilly, C. M., Shum, C. K., Yang, K., Yi, S., Lei, Y., Wang, W., He, Y., Shang, K., Yang, X., & Zhang, H. (2017). Extensive and drastically different alpine lake changes on Asia's high plateaus during the past four decades. *Geophysical Research Letters*, *44*(1), 252–260. https://doi.org/10.1002/2016GL072033

Zhang, G., Yao, T., Xie, H., Kang, S., & Lei, Y. (2013). Increased mass over the Tibetan Plateau: From lakes or glaciers? *Geophysical Research Letters*, *40*(10), 2125–2130. https://doi.org/10.1002/grl.50462

Zhang, G., Yao, T., Xie, H., Yang, K., Zhu, L., Shum, C. K., Bolch, T., Yi, S., Allen, S., Jiang, L., Chen, W., & Ke, C. (2020). Response of Tibetan Plateau lakes to climate change: Trends, patterns, and mechanisms. *Earth-Science Reviews*, *208*, 103269. https://doi.org/10.1016/j.earscirev.2020.103269

Zhang, T., Li, D., East, A. E., Kettner, A. J., Best, J., Ni, J., & Lu, X. (2023). Shifted sediment-transport regimes by climate change and amplified hydrological variability in cryosphere-fed rivers. *Science Advances*, *9*(45), eadi5019. https://doi.org/10.1126/sciadv.adi5019

Zhu, L., Ju, J., Qiao, B., Yang, R., Liu, C., & Han, B. (2019). Recent lake changes of the Asia water tower and their climate response: Progress, problems and prospects. *Chinese Science Bulletin*, *64*(27), 2796–2806. https://doi.org/10.1360/TB-2019-0185