https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Open Access Earth System Science Data Discussions

Bridging the Data Gap: An Enhanced Global Inventory for Statistical Characterization and Hazard Assessment of Landslide Dams

Xiangang Jiang[1], Guoqiang Xiao[1], Tao Wen[1], and Guang Yang[1]

[1]School of Architecture and Construction, Sichuan Agricultural University, Chengdu 611800, China

Correspondence to: Xiangang Jiang(jxgjim@163.com)

Abstract. Landslide dams and their subsequent outburst floods represent cascading geohazards with profound socio-economic and morphological impacts. However, the widespread absence of dynamic breaching parameters in existing global inventories severely constrains quantitative hydrodynamic modeling and downstream risk assessment. To bridge this critical data void, this study presents a comprehensive global landslide dam dataset encompassing 902 rigorously vetted events spanning before 2020. Moving beyond traditional static cataloging, the assembled dataset integrates 11 fundamental morphological and triggering parameters with 6 highly transient breaching metrics. Notably, it significantly improves the data availability of historically scarce variables, including peak discharge, released water volume, and three-dimensional breach geometries. Spatially, the database achieves global coverage, with the highest data densities clustered within the Alpine-Himalayan and Circum-Pacific active belts. To objectively account for observational limitations and chronological biases across different technological eras, a point-by-point Data Quality Flag (DQF) system is incorporated into the dataset, transparently classifying the spatial, geometric, and hydrodynamic uncertainties for every cataloged event. This multi-dimensional and structurally transparent inventory provides a robust empirical foundation for future machine-learning-based hazard susceptibility mapping and physically-based dam-breach simulations. The dataset is publicly available at Zenodo https://doi.org/10.5281/zenodo.19198720 (Jiang et al. 2026).

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Open Access Earth System Science Data Discussions

1 Introduction

Landslide dams, alternatively termed barrier lakes or quake lakes, constitute a formidable geohazard arising when mass movements such as landslides, rock avalanches, or debris flows obstruct river channels, thereby impounding substantial upstream water volumes (Luo et al. 2025; Wu et al. 2025; Feng et al. 2025; Takayama et al. 2026; Z. Li et al. 2025; Silwal et al. 2024). These ephemeral structures are notoriously unstable, frequently culminating in abrupt breaching events that unleash devastating outburst floods, endangering downstream populations, infrastructure, and ecological systems (Luo et al. 2025; Takayama et al. 2026; Huang et al. 2026). Under the influence of global climate change, warming-induced increases in extreme rainfall, coupled with active seismicity and glacial degradation, have heightened the geomorphic instability of mountainous terrains, leading to a marked upsurge in landslide dam formations in recent decades (Lützow et al. 2023; Wu et al. 2025; Costa and Schuster 1988; Wu et al. 2022; Gao et al. 2025). Consequently, robust and comprehensive data are indispensable for driving probabilistic risk assessments, calibrating predictive models, and optimizing disaster mitigation strategies (Cheng et al. 2025; Fan et al. 2020; Li et al. 2025; Peng et al. 2026).

Despite progressive scholarly endeavors, prevailing global landslide dam databases manifest substantial shortcomings that impede rigorous statistical modeling. Seminal catalogs (Ermini and Casagli 2003; Costa and Schuster 1991)provided foundational typologies and geomorphic indices but are constrained by historical limitations and a general dearth of quantitative metrics on breach geometries or outflows. Subsequent regional and global augmentations (Wu et al. 2022; Shi et al. 2022; Tacconi Stefanelli et al. 2015) have substantially broadened the sample breadth. However, a pervasive data gap persists regarding the quantitative completeness of crucial hydrodynamic and breach morphology parameters (Lei et al. 2025; Li et al. 2025; Dufresne et al. 2023). Specifically, detailed records of breach geometries (e.g., breach top width, breach bottom width, breach depth) and extreme hydraulic outputs (e.g., peak discharge, released water volume) remain acutely scarce in mainstream inventories (Azmi and Thomson 2024; Lei et al. 2025). This lack of high-completeness parameterization inherently limits the training of machine-learning-driven classifications and the validation of predictive models (Wang et al. 2025; Li et al. 2025; Nasser et al. 2026; El Bilali and Taleb 2025).

The primary objective of this paper is to provide the scientific community with a detailed and transparent enhanced global landslide dam dataset encompassing 902 meticulously vetted cases spanning from 1800 to 2020. This dataset systematically integrates multi-source historical records and modern observational data. While achieving extensive spatiotemporal coverage and high-density parameter acquisition, it specifically fills critical data gaps in quantitative indicators such as breach morphology and hydrological extremes(Lei et al. 2025; Dufresne et al. 2023). To ensure data reliability and scientific applicability, this study implemented rigorous standardized quality control, conducted data validation by benchmarking against mainstream inventories, and objectively evaluated the inherent uncertainties and observational limitations of the dataset. Consequently, it establishes a solid data foundation for future geohazard prediction, probabilistic risk assessment, and the formulation of disaster mitigation strategies.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

2 Data sources and retrieval methods

2.1 Parameter completeness and variable composition

To transparently illustrate the data abundance of the constructed global landslide dam inventory, Figure 1 quantifies the data completeness (non-empty percentage) of the core parameters across the 902 vetted cases. The statistical distribution reveals a clear hierarchy in data availability, with fundamental geometric parameters constituting the most robust and complete component of the dataset. Specifically, the data completeness for dam height is exceptionally high, reaching approximately 97 %, followed closely by dam length (85 %), dam width (84 %), and dam volume (80 %). This high degree of completeness demonstrates that the fundamental geometric parameters constitute the attribute subset with the highest data density and the most comprehensive coverage within the inventory.

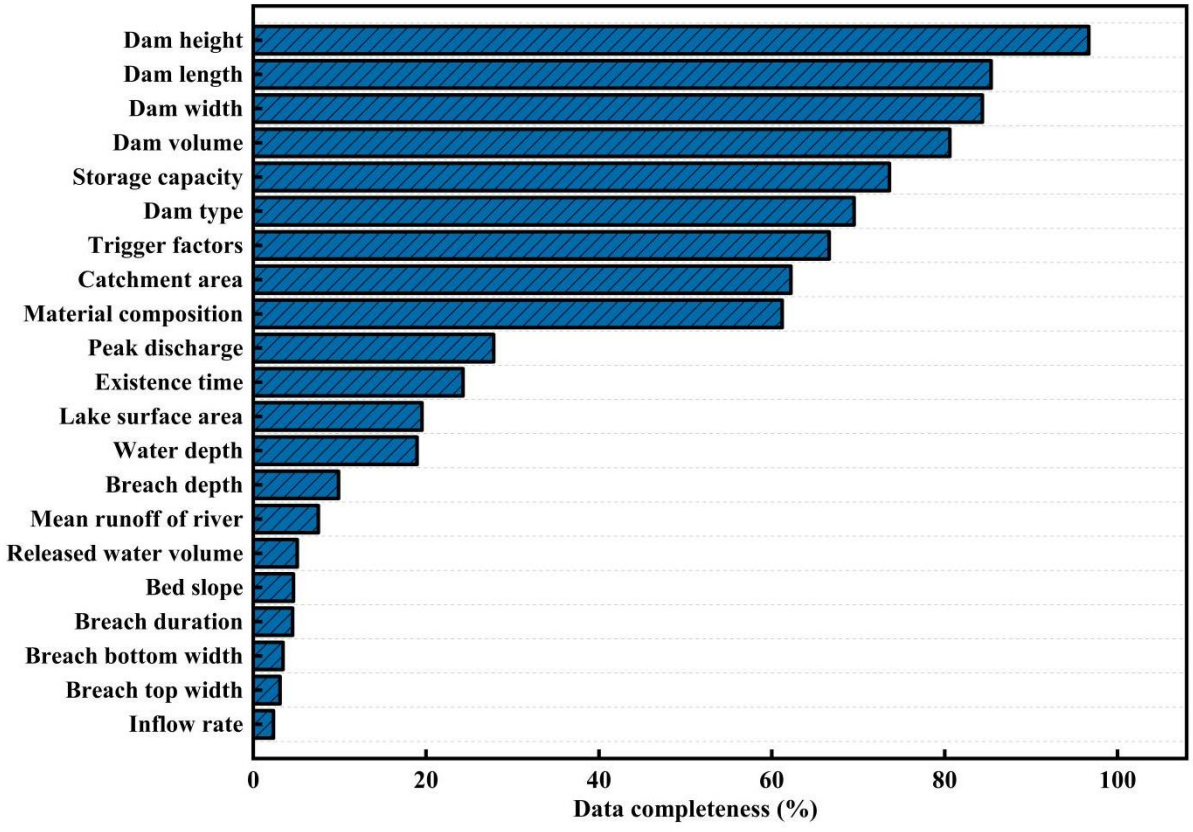


Figure 1. The completeness of core parameters

Beyond the fundamental geometric features, several critical environmental and categorical attributes maintain a moderately high level of completeness. Parameters such as storage capacity (74 %), dam type (70 %), trigger factors (66 %), catchment area (62 %), and material composition (61 %) provide essential contextual data. This stable data acquisition rate ensures a

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026


Open Access Earth System Science Data Discussions

sufficiently large and representative sample size to support robust statistical distributions and categorical grouping within the inventory.

Crucially, the dataset also incorporates highly transient hydrological and morphological parameters that are rarely systematically collected in prevailing mainstream databases (Lei et al. 2025; Dufresne et al. 2023). Variables such as peak discharge (27 %), existence time (24 %), lake surface area (20 %), and water depth (19 %) represent a secondary tier of availability. Furthermore, the dataset successfully archives extremely scarce metrics, including breach depth (10 %), mean runoff of river (8 %), released water volume (5 %), breach duration (5 %), and breach top/bottom widths (3 %). Although the data completeness for these specific breach parameters falls within the single-digit or low double-digit range, their inclusion still yields a substantial absolute number of valid records given the total base of 902 events.

2.2 Standardized screening and quality control

To ensure the reliability and analytical value of the dataset, we implemented a multi-stage standardization and quality control protocol (Figure 2). We first compiled a raw pool of approximately 2,000 preliminary records by systematically aggregating data from existing global and regional inventories, peer-reviewed literature, government technical reports, and remote sensing interpretations.

The initial filtering phase focused on deduplication and spatial verification. Because historical landslide dams are frequently documented under varying local toponyms or translated names across different sources, robust spatial and temporal cross-referencing was essential (Wu et al. 2022; Dufresne et al. 2023; Wu et al. 2020; Cheng et al. 2025). We utilized geospatial coordinates, formation dates, and key geometrical parameters (such as dam height and width) as primary cross-referencing keys. For records with conflicting coordinates for the same historical event, we pinpointed the precise blockage locations using high-resolution satellite imagery (Google Earth). Duplicate entries were merged to retain the most comprehensive parameter measurements, while ambiguous records lacking verifiable spatial anchors were excluded.

Subsequently, to guarantee the database's utility for geomorphological and hydrological modeling, we enforced a strict parameter completeness threshold. Cases were screened based on the completeness of core parameters (as outlined in Figure 1), and records containing fewer than three core parameters were removed. Furthermore, to ensure global interoperability, all retained physical parameters were standardized to consistent units. This completeness threshold and standardization workflow ensure that the final inventory possesses the data depth and consistency required for robust statistical benchmarking. Through this systematic process, the initial raw data pool was refined into a high-fidelity inventory of 902 validated landslide dam cases.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

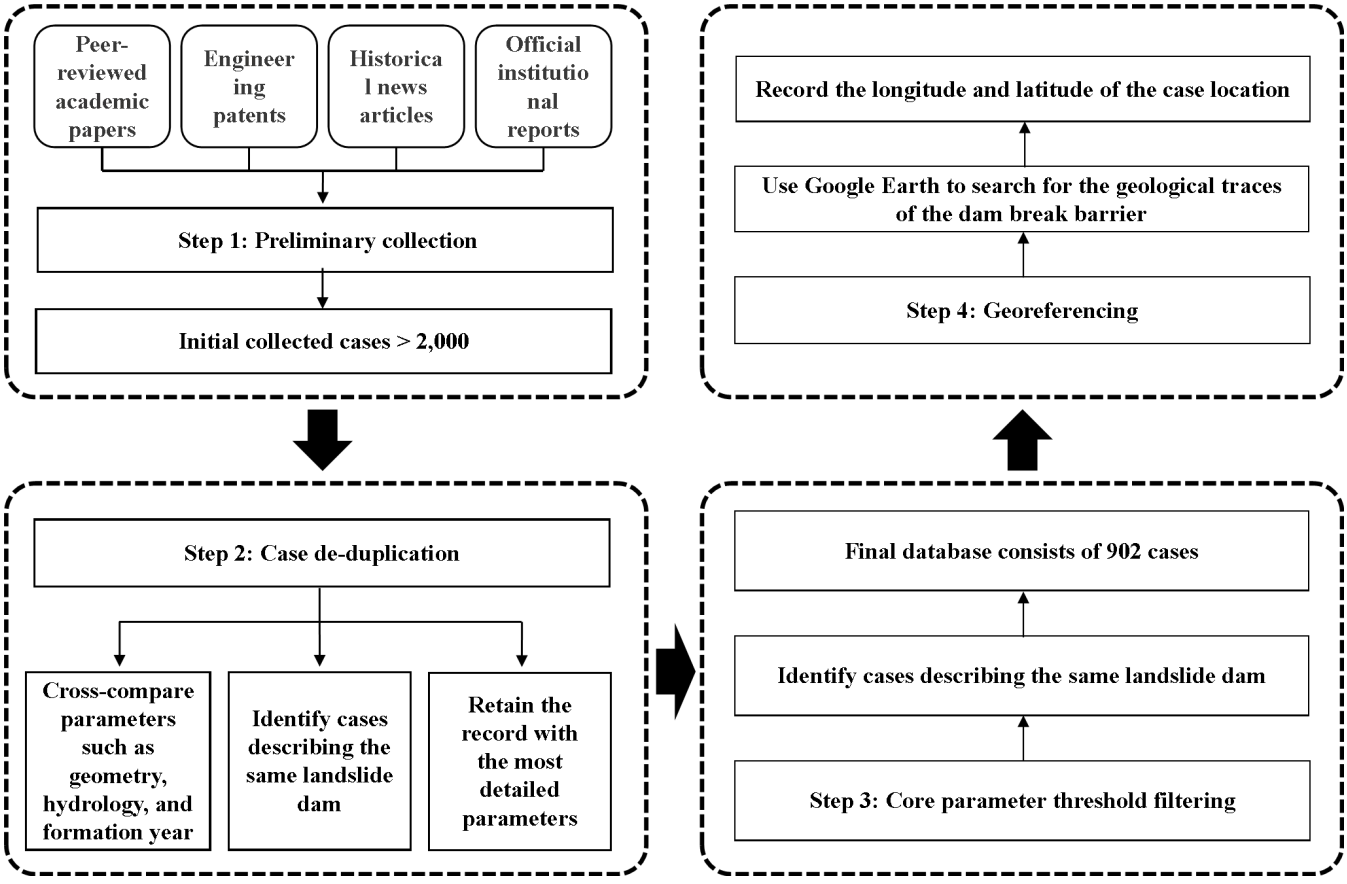


Figure 2. Methodological flowchart for the compilation of the global landslide dam database, detailing the sequential processes of multi-source data retrieval, standardized screening, and rigorous quality control.

2.3 Global Geospatial Positioning

Figure 3 illustrates the global spatial distribution of all landslide dam cases within the database. This demonstrates the extensive spatial coverage of our established inventory, with the geographical clustering of these events aligning rigorously with the world's most prominent mountainous terrains and seismically active tectonic boundaries (Chen et al. 2024; Wu et al. 2022; Korup et al. 2010; Gao et al. 2025; Stefanelli et al. 2015; Shafieiganjeh et al. 2022; Silwal et al. 2024).

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

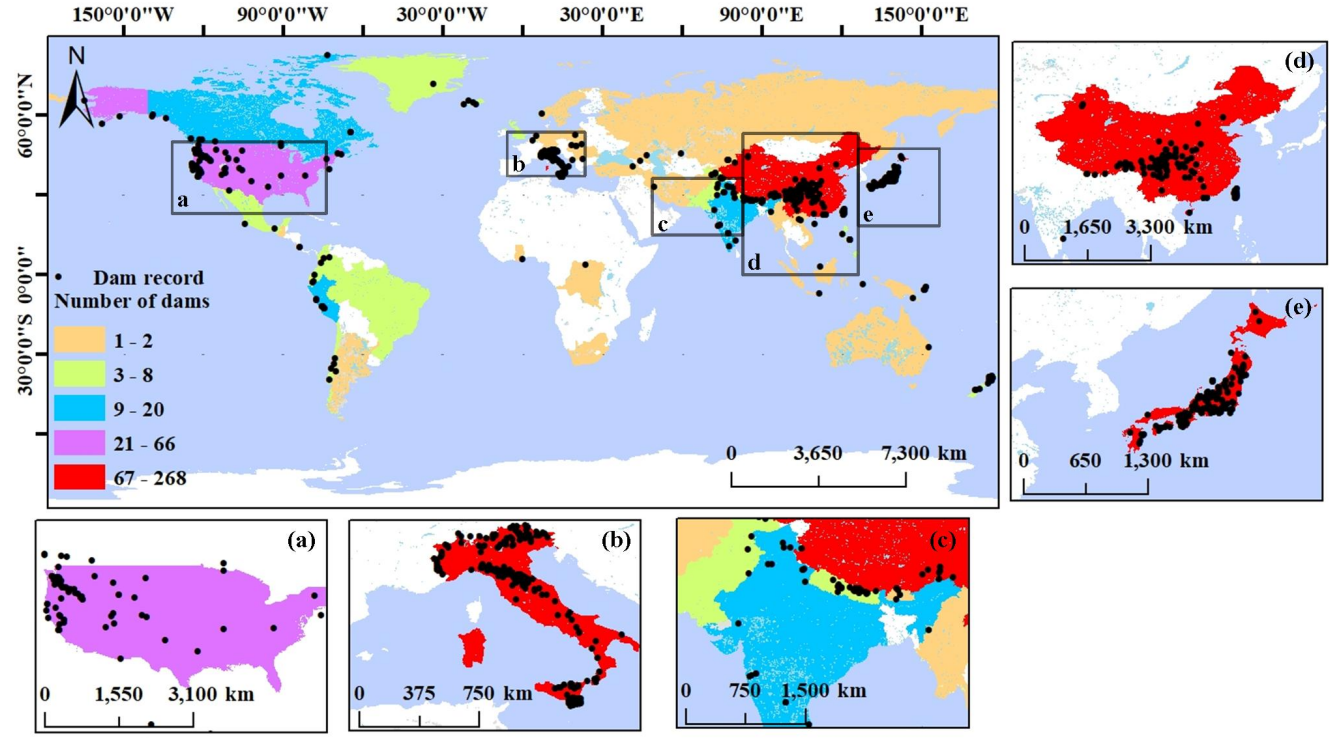


Figure 3. Global spatial distribution of the georeferenced landslide dam records. Insets (a)–(e) provide magnified views of regions with concentrated geospatial data points.

Specifically, these cases are overwhelmingly concentrated along the Alpine-Himalayan orogenic belt and the Circum-Pacific seismic belt. At a regional scale, the highest data densities are observed across East Asia—particularly along the steep margins of the Qinghai-Tibet Plateau in southwestern China, the mountainous terrains of Taiwan, and the Japanese archipelago. Secondary dense clusters are clearly evident in the European Alps, the western cordilleras of North and South America, and New Zealand (Korup 2005; Wu et al. 2020; Fan et al. 2020; Korup et al. 2010; Wu et al. 2024; Morgenstern et al. 2024).

From a spatial data perspective, the density mapping highlights the profound heterogeneity of the inventory's global coverage. The coordinate metadata reveal distinct high-density data clusters that contrast sharply with vast data gaps (zero-record areas) across continental interiors. By explicitly documenting the precise geographic extent of both data-rich and data-poor regions, this database provides a transparent spatial reference, enabling users to accurately assess spatial data availability prior to conducting large-scale geohazard modeling.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

3 Database structure and attribute composition

3.1 Categorization of database parameter

To facilitate efficient data retrieval, interdisciplinary integration, and standardized application, the comprehensive variables archived within the inventory are systematically structured into a clear logical hierarchy. As illustrated in Figure 4, the architecture of the dataset comprises five primary functional categories: Basic Info, Geometrical Parameters, Hydrological Parameters, Time Parameters, and Breach Parameters.

The "Basic Info" module catalogs the fundamental descriptive and contextual metadata for each case, encompassing geographical coordinates, trigger factors, material composition, dam type, and associated reference sources. The "Geometrical Parameters" and "Hydrological Parameters" modules quantify the physical dimensions of the dam body (e.g., dam height, length, and volume) and the boundary hydrodynamic conditions of the reservoir (e.g., storage capacity, inflow rate, and catchment area), respectively. The "Time Parameters" module records the chronological lifecycle of the events, tracking metrics from the formation year to the existence time.

Importantly, the "Breach Parameters" module does not function as an isolated category within the underlying database; rather, its constituent variables inherently fall under the fundamental geometrical, hydrological, and time modules. However, recognizing that the detailed documentation of highly transient breach processes is a distinct advantage of this inventory over prevailing mainstream datasets (Liu 2014; Dufresne et al. 2023), we have logically aggregated these critical indicators (i.e., breach depth, breach top width, breach bottom width, peak discharge, released water volume, and breach duration) into a specific thematic group within the framework of Figure 4. This targeted structural presentation not only visually highlights the dataset's unique value in supporting research on dam failure mechanics but also establishes a clear parametric foundation for the subsequent multidimensional benchmarking against other datasets.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

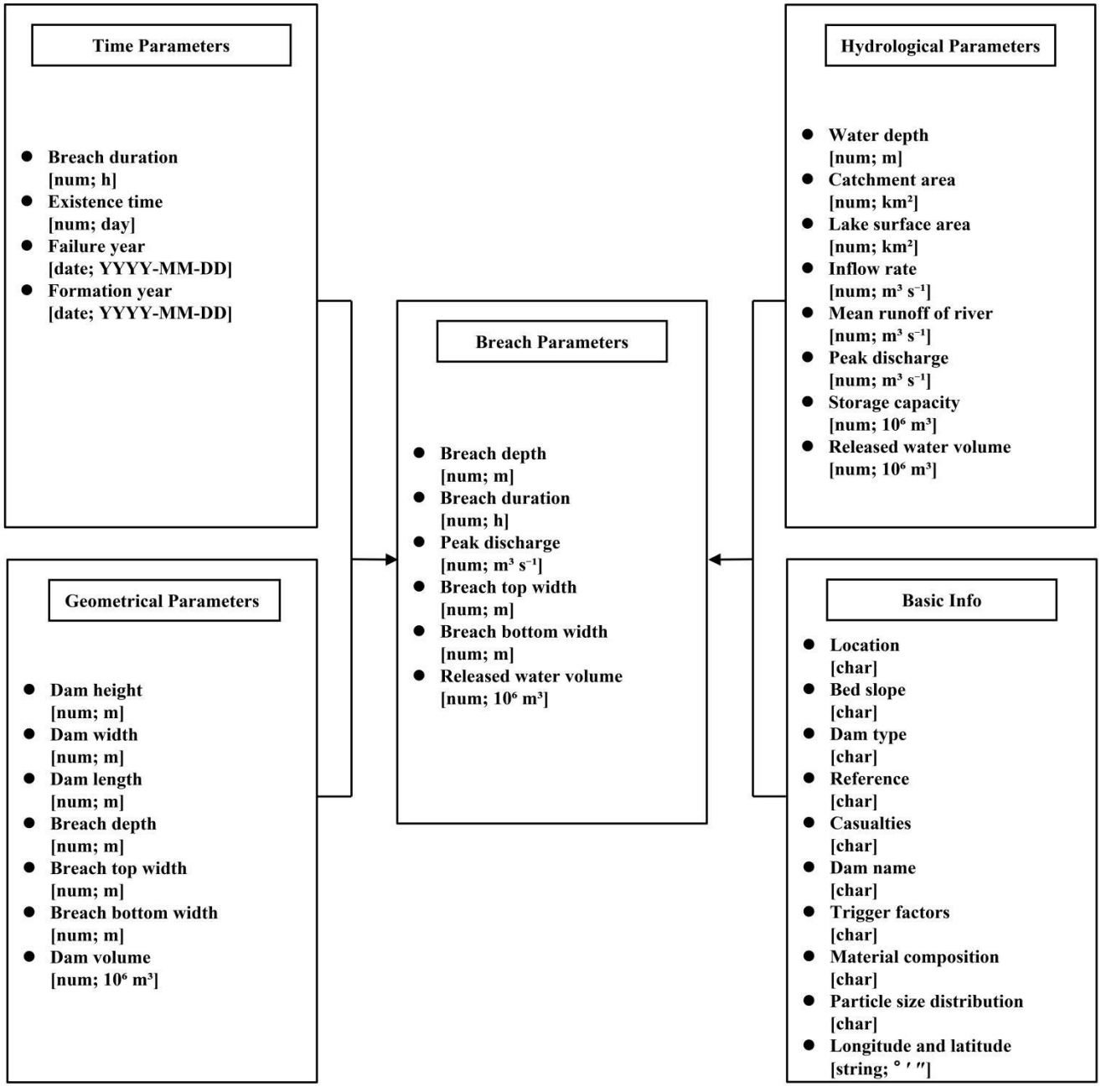


Figure 4. Logical classification and hierarchical structure of the database parameters.

3.2 Case-Level Data Richness

To quantitatively evaluate the informational depth and reliability of the compiled database, we analyzed the case-level data richness, defined as the total number of valid parameters recorded for each independent landslide dam event. As established

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

during the initial screening phase, all 902 validated records strictly meet the minimum threshold of containing at least three fundamental parameters.

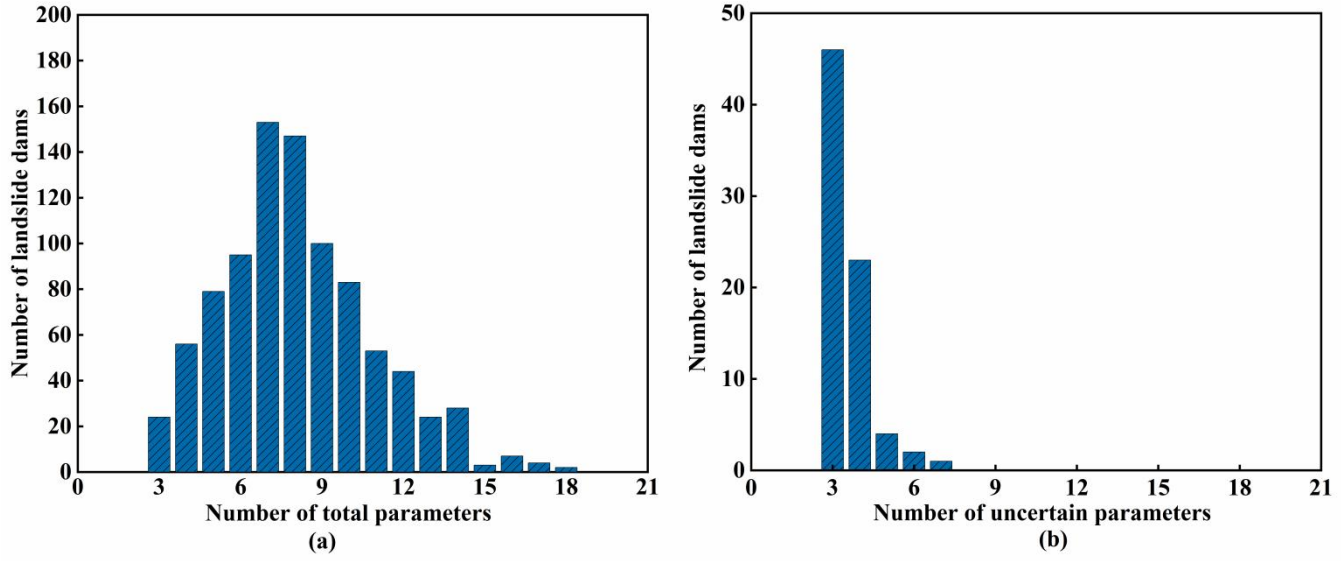


Figure 5. Frequency distributions characterizing case-level data richness and parameter certainty within the compiled database. Panel (a) displays the distribution of the total number of valid parameters recorded for each landslide dam event. Panel (b) illustrates the distribution of uncertain parameters per case.

Figure 5a illustrates the frequency distribution of total parameter availability across the inventory. The histogram exhibits a distinct, approximately normal unimodal pattern, spanning from the baseline of 3 to a maximum of 18 parameters per case. The distribution is decisively anchored by a peak at 7 and 8 parameters, which together form the informational core of the database. Furthermore, the vast majority of cases fall within the robust range of 6 to 11 parameters. From a data-quality perspective, this concentration indicates a high level of completeness. It implies that most records provide sufficient dimensional and contextual depth to support multivariate statistical analysis, rather than serving as mere spatial waypoints. The rightward tail of the distribution (cases comprising 12 to 18 parameters) represents a highly valuable subset of ultra-rich records. These exceptional cases also benefit from modern post-disaster field forensic surveys and advanced remote sensing (Singh et al. 2025; Liu et al. 2025; Rimal and Tiwary 2024; Fan et al. 2021, 2018; Xie et al. 2025; Dun et al. 2025; Song 2025).

Crucially, Figure 5b offers a complementary assessment of data certainty by tracking the distribution of uncertain parameters. While preserving historical uncertainty (e.g., historical estimations recorded as ranges rather than discrete values) is essential for data transparency, the graph reveals that this ambiguity is highly constrained (Cheng et al. 2025; Dufresne et al. 2023). Among the specific subset of cases containing uncertain metadata, the absolute majority are strictly limited to exactly 3 uncertain parameters. When juxtaposed with the high overall parameter counts (as seen in Panel a), this restricted presence of uncertainty strongly underscores the overall determinism and completeness of the database. It statistically demonstrates

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

that the informational fuzziness inherent to historical archives has been minimized, ensuring that the final inventory provides a robust, highly reliable empirical foundation for subsequent geohazard susceptibility modeling.

3.3 Distribution of qualitative attribute

In addition to quantitative metrics, qualitative categorical attributes provide crucial contextual information for geomorphological classification and structural assessment. Figure 6 summarizes the internal proportional distribution of fundamental qualitative indicators across the 902 cataloged events, encompassing dam types (Figure 6a), material compositions (Figure 6b), and triggering factors (Figure 6c).

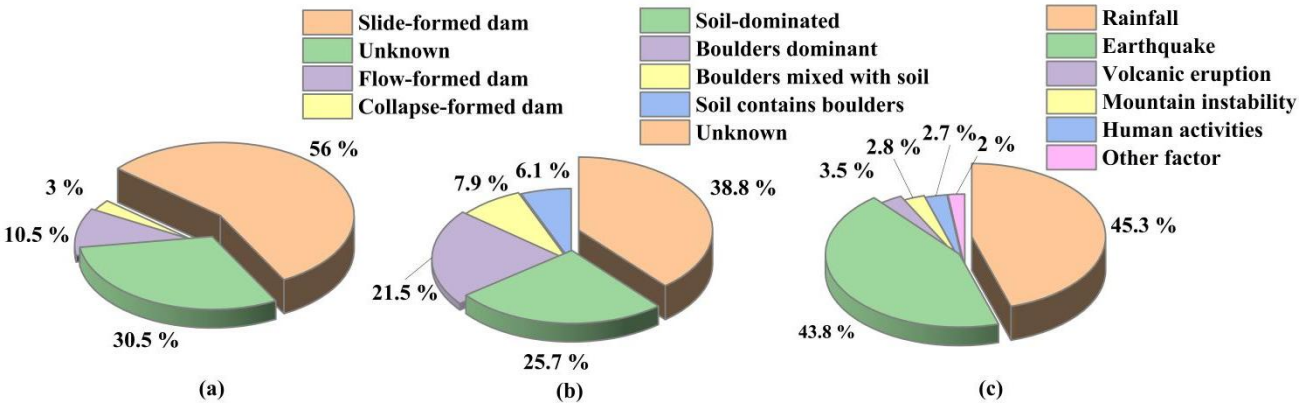


Figure 6. Statistical proportions of qualitative categorical attributes across the compiled database. Panel (a) illustrates the distribution of primary dam types, where "Unknown" denotes cases lacking explicit morphodynamic classification. Panel (b) displays the material compositions, where "Unknown" indicates historical records with unidentified internal geological structures. Panel (c) presents the primary triggering factors responsible for landslide dam formation.

Regarding the primary formation mechanisms (Figure 6a), slide-formed dams constitute the absolute majority, accounting for exactly 56 % of the total inventory. Flow-formed and collapse-formed dams represent smaller fractions at 10.5 % and 3 % (Fan et al. 2020; Wu et al. 2022; Cheng et al. 2025; Song et al. 2026; Wu et al. 2025), respectively. Notably, 30.5 % of the records lack explicit geomorphodynamic classification (categorized as "Unknown"), highlighting the inherent ambiguity prevalent in early historical literature (Dufresne et al. 2023; Costa and Schuster 1988).

Similarly, the material composition of the dams (Figure 6b) exhibits a diversified distribution. Among the cases with explicit records, soil-dominated structures form the largest known category (25.7 %), closely followed by boulder-dominant dams (21.5 %). Mixed-material structures, classified as "boulders mixed with soil" and "soil contains boulders," account for 7.9 % and 6.1 %, respectively. Consistent with the practical limitations of historical data retrieval, a substantial proportion (38.8 %) is categorized as "Unknown." This reflects the objective challenges encountered in determining the internal geological structures of early historical events or dams in inaccessible regions (Dai et al. 2023; Fan et al. 2021b; Liu 2014; F. Wang et al. 2018).

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

In terms of triggering factors (Figure 6c), the dataset demonstrates a clear dual-dominance. Rainfall and earthquakes act as the primary catalysts, responsible for 45.3 % and 43.8 % of the landslide dam events, respectively (Cheng et al. 2025; Wu et al. 2022). Minor triggers account for the remaining fraction, including volcanic eruptions (3.5 %), mountain instability (2.8 %), human activities (2.7 %), and other underlying factors (2 %). By explicitly quantifying these triggers alongside the structural characteristics, the database offers a transparent, macroscopic physical profile of global landslide dams, providing a viable foundation for researchers to conduct stratified statistical analyses and hazard susceptibility modeling.

4 Spatiotemporal coverage

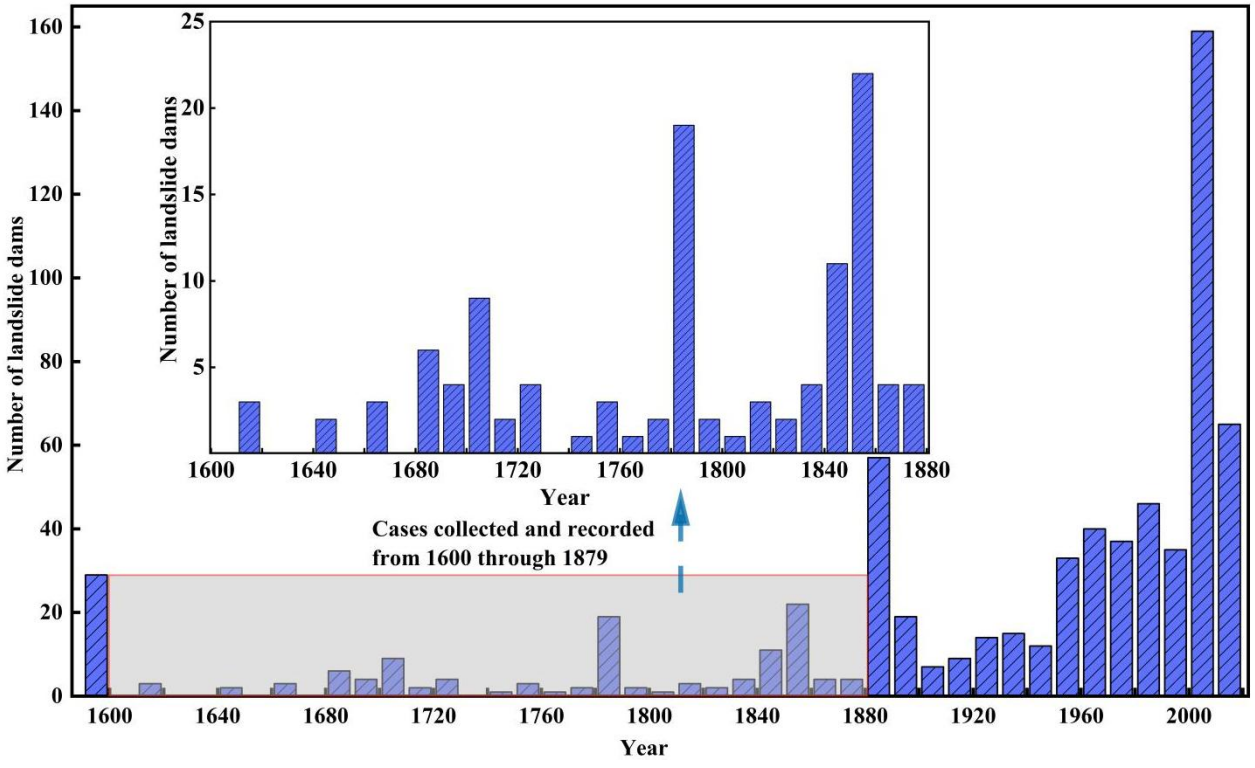
4.1 Temporal coverage and decadal records

The assembled database encompasses a vast temporal range of landslide dam formation years, spanning from antiquity to 2020. Figure 7 illustrates the temporal distribution of these events, where each bar (post-1600) represents the aggregate number of records within a ten-year interval (e.g., from 1621 through 1630). All documented events occurring prior to 1600 are consolidated into a single initial bin to establish a long-term chronological baseline.

The statistical distribution reveals a distinct, multi-stage increase in the volume of recorded events over time. From 1600 through the late 19th century, the records are relatively sparse and intermittent, primarily capturing catastrophic events that left significant geological traces or were preserved in historical archives. A notable transition occurs after 1880, marked by a steady rise in decadal records. The most substantial surge is observed from the mid-20th century onwards, with the decadal count peaking in the interval from 2001 through 2010 (exceeding 130 cases).

This progressive increase in record density reflects not only the enhanced preservation of scientific data and the maturation of global remote sensing technologies in the modern era (Tavus et al. 2024; Vassileva et al. 2023; Malheiro et al. 2023), but also aligns with the widely documented physical escalation of mountainous geohazards driven by global climate change and increasingly frequent extreme hydrometeorological events (Lützow et al. 2023). Consequently, this extensive temporal coverage provides a high-fidelity chronological dataset that captures both the evolution of measurement techniques and the long-term physical macro-trends of landslide dams.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026



205 **Figure 7. Decadal distribution of the formation years of cataloged landslide dams (up to 2020). Each bar represents a 10-year interval, with all cases prior to 1600 aggregated in the first column. The inset provides a magnified view of the records from 1600 through 1879.**

4.2 Regional data density

Figure 8 illustrates the quantitative distribution of the compiled landslide dam events at the national level. The integration of
210 the pie chart and the bar histogram provides a dual perspective on both the absolute frequencies and the relative categorical proportions of the inventory. A pronounced feature of this distribution is its extreme concentration. The dataset is overwhelmingly dominated by three specific nations: Italy (IT), Japan (JP), and China (CN). Quantitatively, Italy contributes the largest individual subset with 29.7 % of the total records, closely followed by Japan and China, which account for 24.7 % and 24.6 %, respectively. Cumulatively, this primary tier constitutes nearly 80 % (79.0 %) of the entire database, establishing
215 a heavily weighted spatial footprint.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

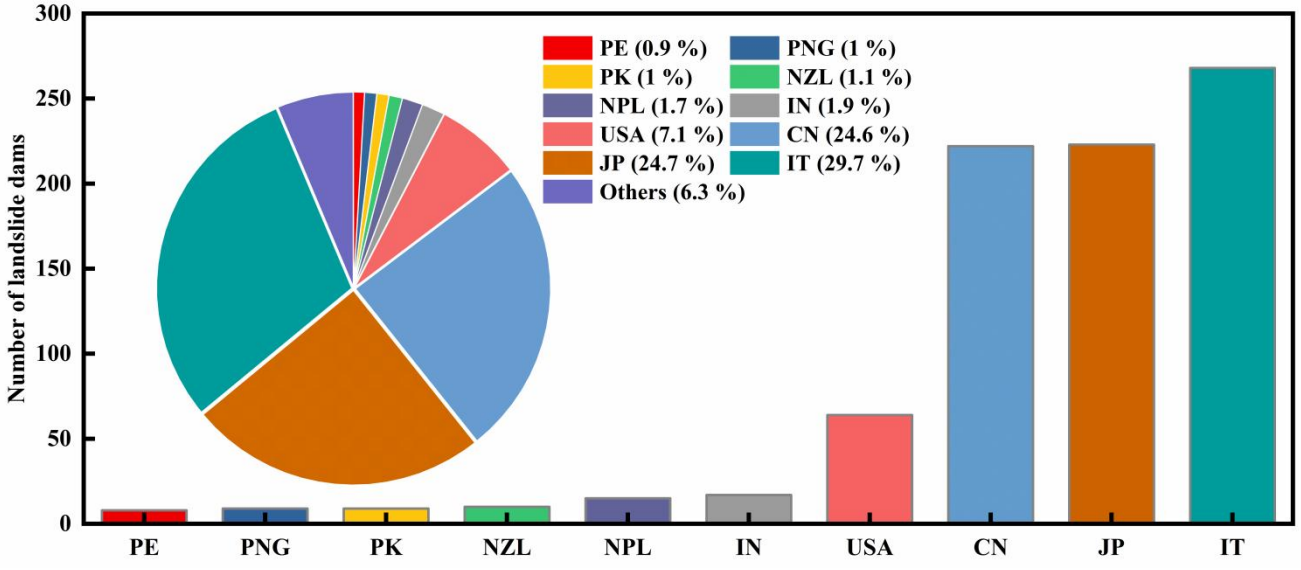
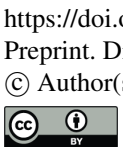


Figure 8. Global spatial distribution of landslide dam events across different nations. The bar and pie charts display the absolute case counts and their relative proportions, respectively. Data are predominantly concentrated in Italy (IT), Japan (JP), and China (CN), with additional records from the United States (USA), India (IN), Nepal (NPL), New Zealand (NZL), Pakistan (PK), Papua New Guinea (PNG), and Peru (PE).

Beyond this primary cluster, the United States (USA) forms a distinct secondary tier, representing 7.1 % of the documented cases. The remaining data points exhibit a typical long-tail distribution, spreading across several other countries with significantly lower frequencies. As visualized in the chart, nations such as India (IN), Nepal (NPL), New Zealand (NZL), Papua New Guinea (PNG), Pakistan (PK), and Peru (PE) each account for marginal proportions ranging from a minimum of 0.9 % to a maximum of 1.9 %. Furthermore, a minor fraction of 6.3 % is aggregated into the "Others" category, capturing all remaining countries with extremely sparse historical or modern records.

Explicitly quantifying this national-level disparity is essential for presenting an objective profile of the inventory's spatial architecture (Dufresne et al. 2023; Wu et al. 2022). Rather than implying underlying physical or geomorphological mechanisms, this highly skewed distribution serves as a direct indicator of regional data availability within the current compilation. For database users, this detailed national breakdown acts as critical metadata. It explicitly highlights the uneven geographic coverage, ensuring that researchers are fully informed of these inherent data imbalances. Consequently, users can make mathematically sound decisions when evaluating global completeness, extracting specific regional subsets, or designing spatially stratified sampling strategies to mitigate representation bias in large-scale geohazard modeling.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026



Open Access Earth System Science Data Discussions

5 Statistical Distribution of Quantitative Parameters

5.1 Statistical Distribution of Geometrical Parameters

Geometrical parameters define the spatial scale of a landslide dam and serve as the most fundamental quantitative inputs for hazard assessment and breach modeling. Figures 9 and 10 present the statistical distributions of linear geometric parameters (dam body and breach dimensions) and the total dam volume across the cataloged inventory, respectively.

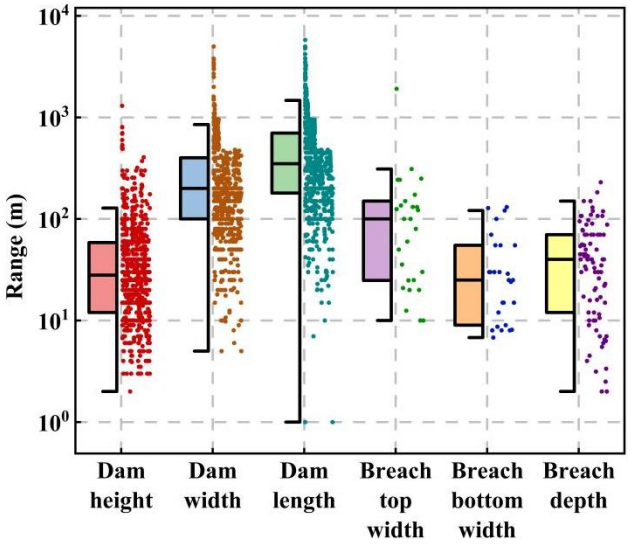


Figure 9. The variability and range of geometric parameters in the landslide dam dataset.

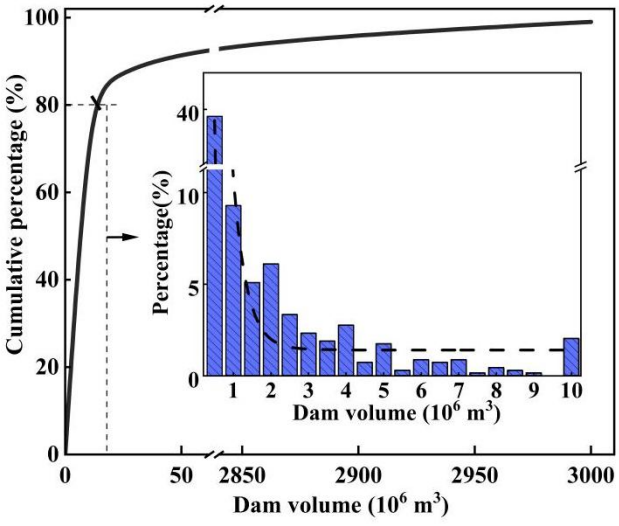


Figure 10. The volume of landslide dams and their orresponding proportion.

As illustrated in the logarithmic box plots (Figure 9), the linear dimensions exhibit a vast natural variance, spanning three to four orders of magnitude. For the overall dam body geometry, dam length and width exhibit the broadest ranges and highest median values, frequently extending from 10 to over 10^3 meters. Dam height generally scales lower than the planar dimensions but retains a substantial spread, with median values concentrated in the tens of meters and extreme outliers exceeding 10^3 meters. The extracted breach parameters (top width, bottom width, and depth) scale correspondingly smaller, primarily distributed between 1 and 10^2 meters. The preservation of these highly variable geometrical bounds accurately reflects the complex topographical constraints and diverse material volumes inherent in global landslide damming events (Dong et al. 2011; Jian et al. 2023; Li et al. 2020; Zhou et al. 2025).

Regarding the three-dimensional scale, the dam volume exhibits a highly right-skewed distribution (Figure 10). The cumulative percentage curve indicates that approximately 80 % of the cataloged events are concentrated in the lower-volume range. The inset frequency histogram further details the data structure within the 0 to 10 × 10^6 m^3 interval, showing that nearly 40 % of the cataloged dams have a volume of less than 1 × 10^6 m^3. In addition to these smaller-volume cases, the dataset also includes large-scale events with volumes reaching up to 3000 × 10^6 m^3. This continuous coverage from small- to large-scale events objectively demonstrates the broad volumetric bounds of the assembled database.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

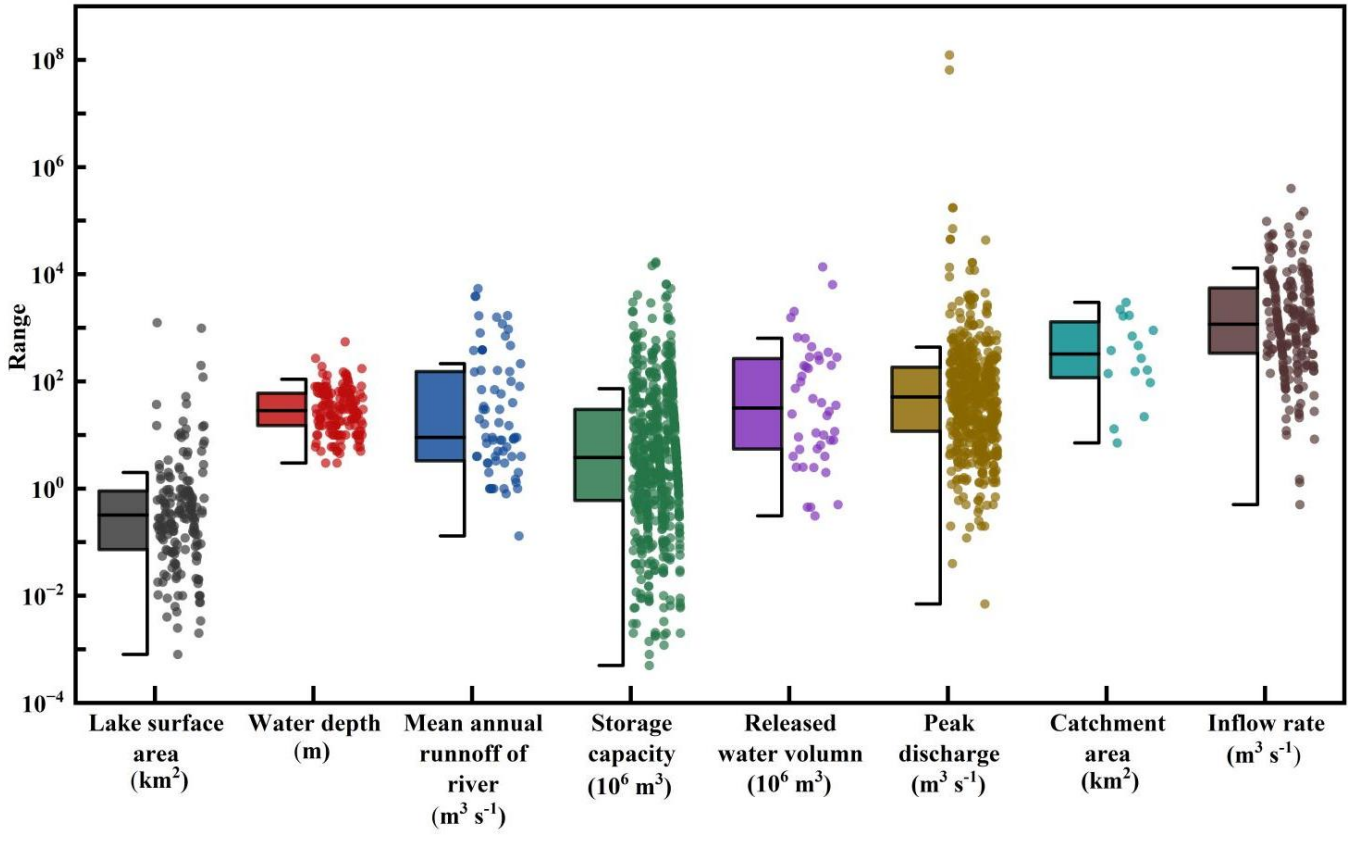


Figure 11. Statistical distribution of key hydrological parameters. The vertical axis represents the data ranges on a logarithmic scale. The plot displays lake dimensions, storage capacities, and flow-related metrics across the compiled database using superimposed boxplots and jittered data points.

5.2 Statistical Distribution of Hydrological Parameters

Figure 11 presents the statistical distribution of eight core hydrological parameters across the compiled landslide dam events, visualized through a combination of box-whisker plots and jittered data points. A defining characteristic of the hydrological profile is the exceptional range of the values, which span up to 12 orders of magnitude (from 10^{-4} to 10^{8}). This vast dispersion necessitates the use of a logarithmic scale to capture the intrinsic variability of landslide-dammed lake systems, ranging from small-scale mountain ponding to catastrophic inundation events.

Key morphological and capacity metrics, such as lake surface area and storage capacity (10^6 m^3), exhibit highly skewed distributions with median values typically concentrated within the 10^{-1} to 10^{1} range. Notably, the "Peak discharge" and "Released water volume" parameters display a high density of data points across several orders of magnitude, reflecting the diverse dynamic processes of dam breach and outburst floods recorded in the database (Rafiq et al. 2019; Risley et al. 2006; Cheng et al. 2025). Furthermore, parameters representing basin-scale characteristics, such as "Catchment area" (km^2) and "Mean annual runoff of river" (m^3 s^{-1}), provide essential boundary conditions for assessing the longevity and hazard

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

potential of the dams (Struble et al. 2021; Kumar et al. 2021; Li et al. 2011; Niazi et al. 2020). The overlapping yet distinct ranges of these metrics illustrate the multidimensional nature of the hydrological data, ensuring that the database can accommodate various modeling requirements, from steady-state seepage analysis to transient flood hydrograph simulation.

5.3 Statistical Distribution of Temporal Parameters

Figure 12 presents the statistical distribution of two critical temporal metrics: existence time and breach duration. The existence time, defining the total duration from initial river impoundment to the onset of dam failure, exhibits a highly right-skewed distribution(Figure 12a). As detailed in the inset focusing on the 0–10 day interval, the absolute highest frequency of failure events occurs within the first 24 hours (Shi et al. 2023). The main histogram expands this temporal profile up to 800 days, demonstrating a rapid decay in case frequency over time. Notably, the final consolidated bin on the horizontal axis represents a "long-tail" subset of landslide dams that remained stable for extended periods exceeding 800 days, explicitly capturing the extreme upper bounds of the inventory.

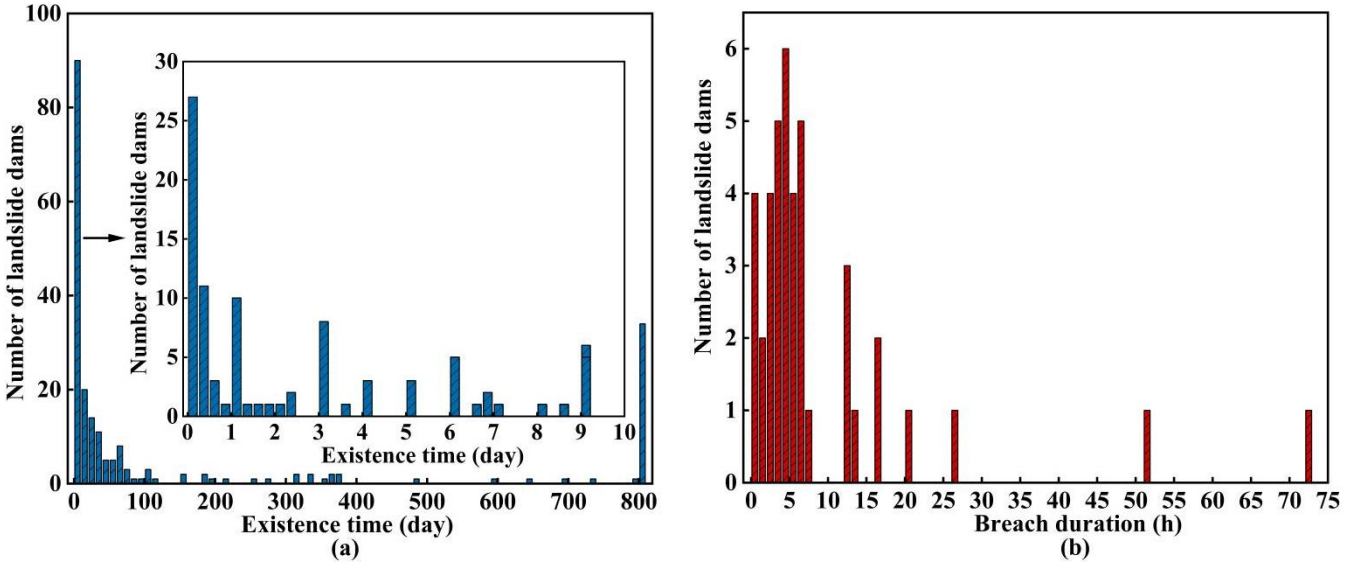


Figure 12. Statistical distribution of temporal parameters. (a) Frequency distribution of the existence time of cataloged landslide dams. The main plot spans an 800-day window, with the terminal bin representing all stable cases exceeding 800 days. The inset provides a detailed view of the high-frequency 0–10 day interval. (b) Frequency distribution of breach duration for the subset of cases with explicit temporal records.

The Figure 12b displays the frequency distribution of breach duration, which characterizes the time elapsed during the actual physical failure process. It is important to note that the sample size for this metric is substantially smaller (as indicated by the vertical axis scale), directly reflecting the observational difficulty of capturing continuous, high-resolution temporal data during sudden outburst events (Lei et al. 2025; Chen et al. 2026; Yi et al. 2025; Jónatas et al. 2025). For the documented cases, recorded breach durations range from 0.5 to approximately 35 hours, with the highest clustering of data points observed within the 1- to 7-hour window. By providing these two distinct temporal dimensions, the database transparently

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

presents both the pre-failure stability period and the subsequent breaching duration, complete with their respective empirical ranges and sample densities.

6 Uncertainty Analysis and Benchmarking

6.1 Data Uncertainty and Observational Limitations

Although rigorous standardized screening protocols were implemented during the database construction, quantifying the inherent uncertainties within historical landslide dam records remains critical for the objective application of this dataset. Given the extensive temporal span of the inventory (1800–2020), many historical and pre-instrumental records intrinsically lack exact numerical error margins (e.g., precise ± values). To avoid introducing artificial precision to these early events while strictly fulfilling the requirement for point-by-point error estimation, we incorporated Data Quality Flags (DQFs) into the database. These flags categorize the inherent observational and technological uncertainties of each cataloged event across three distinct dimensions

1. Uncertainty spatial: This dimension evaluates the precision of geographical coordinates based on data sources and geocoding resolution. "Low" uncertainty (high precision, estimated error < 1 km) is assigned to exact coordinates validated by modern GPS or high-resolution remote sensing. "Medium" uncertainty (1 to 10 km) is applied to approximate coordinates inferred from historical topographic maps or specific landmarks. "High" uncertainty (> 10 km) is reserved for vague spatial positioning reliant solely on qualitative historical texts or regional descriptions.

2. Uncertainty geometric temporal: This dimension assesses the confidence level of static parameters (e.g., dam height, lake surface area), utilizing the technological era of the event as a primary proxy for measurement accuracy. Post-2000 events receive a "Low" uncertainty flag due to the availability of high-resolution digital elevation models (DEMs), LiDAR, or uncrewed aerial vehicle (UAV) photogrammetry. Events between 1950 and 1999 are marked as "Medium", reflecting the use of early aerial photography and standard cartographic surveys. Pre-1950 and paleo-events are flagged as "High" uncertainty, as their dimensions rely heavily on historical archival descriptions or paleo-geomorphological reconstructions, carrying broader systematic estimation risks.

3. Uncertainty hydrodynamic: This dimension evaluates the reliability of highly transient and difficult-to-capture variables, such as peak discharge, released water volume, and breach duration. A "Low" flag is extremely rare and strictly assigned to cases with in-situ instrumental river gauge measurements. "Medium" uncertainty represents values rigorously back-calculated using established empirical hydrodynamic equations based on precise post-event breach geometries or flood mark leveling. "High" uncertainty is assigned to values broadly estimated from qualitative historical damage descriptions or paleo-flood deposits lacking physical calibration.

By providing this granular, multi-dimensional uncertainty metadata, the database strictly adheres to the principles of open science. It ensures that users are fully informed of the observational limitations, allowing them to rigorously filter the

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

inventory based on their specific modeling tolerances when conducting probabilistic risk assessments or training spatial algorithms.

6.2 Benchmarking Against Existing Inventories

To objectively evaluate the contribution and position of the newly compiled dataset within the broader scientific community, a systematic benchmarking analysis was conducted against four globally recognized landslide dam inventories: Costa and Schuster (1991), Tacconi Stefanelli et al. (2015), Wu et al. (2022), and Shi et al. (2022). The comparative analysis focuses on temporal coverage, parameter dimensionality, and data completeness.

Figure 13 compares the decadal distribution of the cataloged events between this study and the four reference databases. While macro-scale inventories like Shi et al. (2022) prioritize sheer volume (2783 cases) by encompassing a vast array of historical events spanning centuries, this study (902 cases) adopts a highly stringent vetting protocol strictly focused on parameter extractability. Consequently, while the absolute number of pre-20th-century cases in this study is lower, its temporal distribution perfectly captures the overarching exponential growth trend observed across all mainstream databases. Notably, the current dataset maintains a highly competitive and robust data density from the mid-20th century onwards, providing a solid statistical foundation for contemporary geohazard analysis without compromising data quality.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

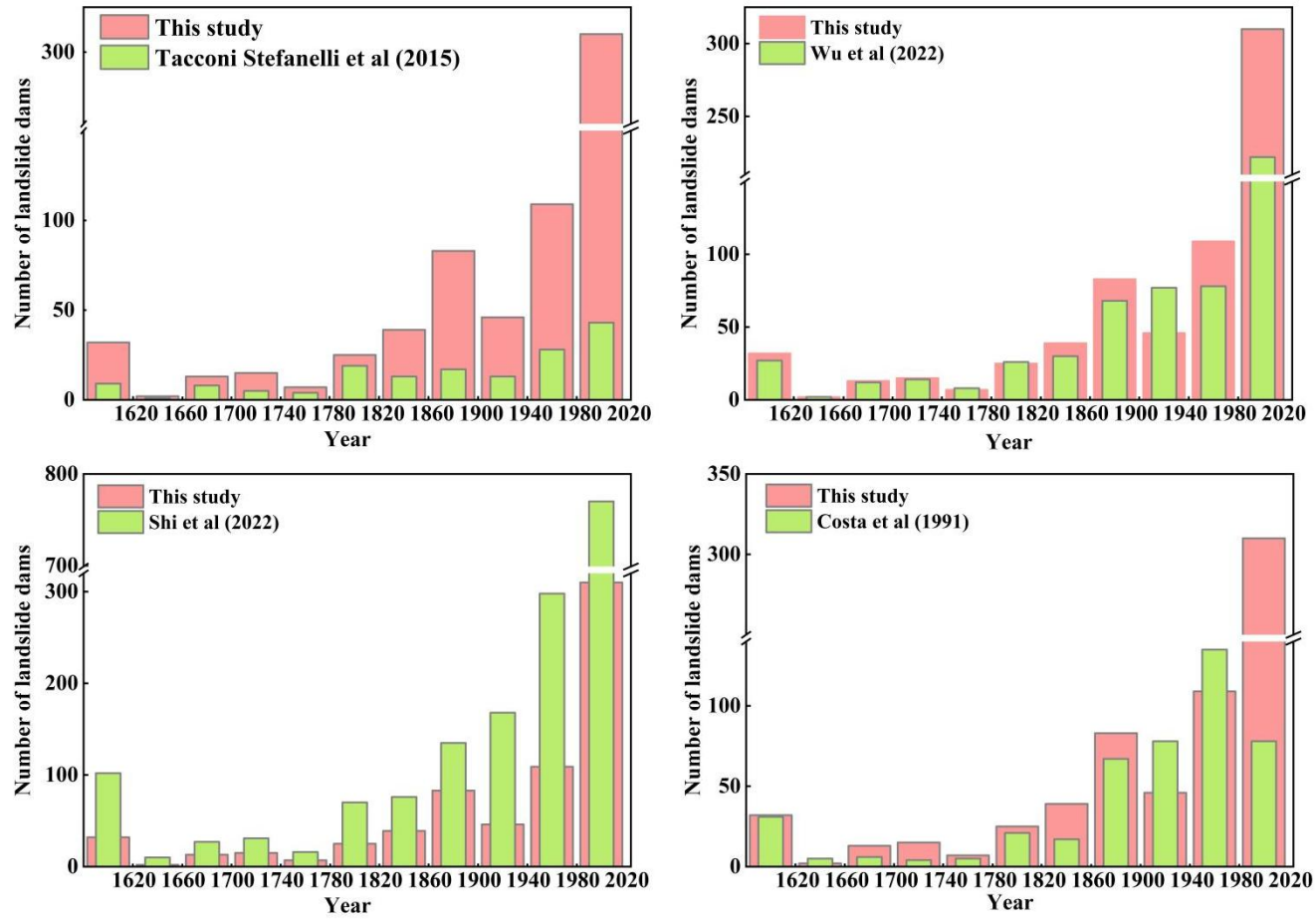


Figure 13. Comparison of the temporal distribution of landslide dam records between this study and four representative global inventories. The bar charts illustrate the decadal data density and the overarching growth trend of historical documentation across different datasets.

The most significant advancement of this database lies in its multidimensional parameter richness and structural completeness, as systematically detailed in Table 1 and Figure 14. Previous compilations have predominantly restricted their focus to static geometric features or qualitative classifications of the dam body. For instance, Tacconi Stefanelli et al. (2015) and Shi et al. (2022) catalog 10 and 8 types of dam parameters respectively, yet they entirely omit quantitative metrics related to the breaching process. Wu et al. (2022) introduces valuable dynamic flow metrics, but lacks the corresponding structural breach geometries. Although Costa and Schuster (1991) pioneered the inclusion of breach dimensions, their overall parameter count and sample size remain highly constrained.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Open Access Earth System Science Data Discussions

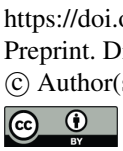
Table 1. Comparison between the dataset developed in this study and existing landslide dam datasets.

Dataset	Numb-er of cases	Types of landslide dam parameters	Landslide dam parameters	Types of breach parameters	Breach parameters
This study	902	11	Dam volume, Dam length, Dam width, Dam height, Material composition, Dam type, Trigger factors, Particle size distribution, Formation year, Existence time， Failure year	6	Peak dischrge, Released water volume, Breach duration, Breach top width, Breach bottom width, Breach depth
Tacconi Stefanelli et al (2015)	300	10	Dam height, Dam length, Dam width， Dam volume Level, Dam area, Dam evolution, Dam condition, Date of damming， Date of failure， Lake life time	0	-
Wu et al (2022)	779	9	LDam status, Duration from formation to flood, LDam-type, Dam materials, Interpreted cause, Reported cause, Dam height, Dam length, Dam width	3	Mean flow velocity, Peak flow velocity, Overflow/flood Time
Shi et al (2022)	2783	8	Dam height, Dam length, Dam width,Dam volume, Triggers, Dam,material, Stability, Date of damming	0	-
Costa and Schuster (1991)	463	7	Dam length, Dam width, Dam height, Dam type, Dam materials, Date， Time to failure	1	Breach dimensions

All parameter names are derived from the author's editing naming convention.

In stark contrast, this study establishes a comprehensive multi-dimensional framework encompassing 11 types of fundamental dam parameters and, crucially, 6 types of detailed breach parameters (including peak discharge, released water volume, breach duration, breach depth, and top/bottom widths). Figure 14 quantitatively visualizes this advancement: while the subplots for the reference databases (Figure 14b–e) exhibit massive data voids regarding dynamic and breaching characteristics, subplot (a) demonstrates unprecedentedly dense and continuous records for these highly transient variables. The systematic integration of both static morphologies and dynamic failure metrics effectively bridges a critical gap in the

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026



existing literature. By providing these essential boundary conditions, the presented dataset facilitates a crucial transition from static hazard cataloging to dynamic hydrodynamic routing and quantitative downstream risk modeling.

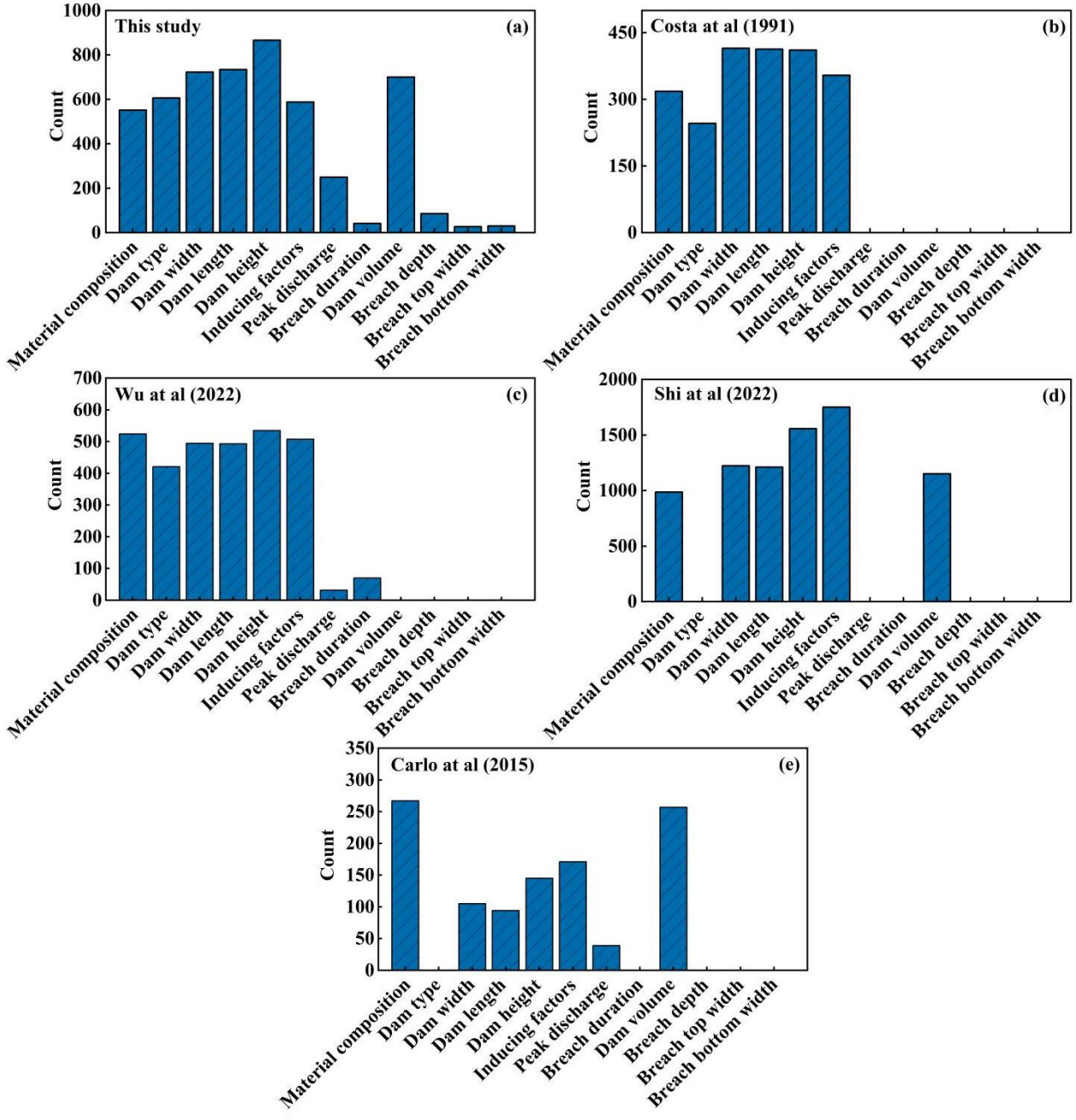


Figure 14. Quantitative comparison of parameter dimensionality and data completeness between this study (a) and four reference databases (b–e). The bar charts highlight the extensive coverage of dynamic breach metrics in the newly compiled dataset, in stark contrast to the data voids present in earlier inventories.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

7 Data availability

The global landslide dam database developed in this study has been archived and is freely accessible on Zenodo at https://doi.org/10.5281/zenodo.19198720 (Jiang et al. 2026). To ensure full cross-platform interoperability and comply with open data standards, the dataset is provided in non-proprietary comma-separated values formats. The repository consists of three primary components: (1) the core inventory, which stores all the multidimensional parameters and data quality flags for the 902 cataloged events; (2) a spatial metadata dictionary detailing the full names of countries and regions corresponding to the abbreviations used in the "Location" column; and (3) the detailed uncertainty classification standards, which explicitly define the evaluation criteria for the spatial, geometric-temporal, and hydrodynamic uncertainty flags (the DQF system). This dataset derives from an ongoing project and is subject to future refinements and continuous additions.

8 Conclusions

This study presents an enhanced and meticulously curated global landslide dam dataset, encompassing 902 verified cases spanning from 1800 to 2020. By systematically integrating multi-source historical archives and modern observational records, the inventory successfully addresses the pervasive data gaps regarding quantitative geomorphic and hydrodynamic parameters in existing catalogs. Specifically, the dataset compiles 11 fundamental morphological and triggering attributes alongside 6 highly transient breaching metrics. The inclusion of detailed three-dimensional breach geometries and extreme hydraulic outputs (e.g., peak discharge, released water volume) provides a highly complete parameterization that is historically scarce but critical for rigorous statistical modeling.

Statistical assessment of the database reveals a robust level of parameter completeness, with the vast majority of cases providing sufficient dimensional depth to support multivariate analysis rather than serving as mere spatial waypoints. The inventory comprehensively captures diverse geomorphodynamic contexts, documenting the dual-dominance of rainfall and earthquake triggers and a wide spectrum of material compositions. Furthermore, it quantifies the extreme temporal variability of landslide dams, presenting a bifurcated perspective that ranges from highly unstable structures failing within hours to long-term stable impoundments, thereby offering a multifaceted hydrological profile across multiple orders of magnitude.

To objectively manage the observational limitations and temporal survivorship biases inherent in historical records, this study implemented a point-by-point Data Quality Flag (DQF) system. This multi-dimensional framework systematically categorizes spatial location accuracy, geometric precision, and hydrodynamic reliability based on the technological era and measurement constraints of each event. By explicitly establishing these confidence intervals, the dataset provides a transparent and structured metric for data reliability, enabling researchers to apply physically meaningful boundary conditions.

Provided as an open-access resource, this structurally transparent inventory establishes a solid empirical foundation for the scientific community. It is specifically designed to facilitate advanced geohazard research, including the training of machine-

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

learning-driven susceptibility models, the calibration of physically-based dam-breach simulations, and the optimization of regional probabilistic risk assessments under changing environmental conditions.

Author contributions

XJ designed the research, wrote most of the papers, revised all the charts and participated in the data collection and the establishment of the database. GX analyzed all the data, created all the charts, wrote part of the paper, and participated in the data collection and the establishment of the database. TW has collected most of the data. GY participated in the chart drawing. All of the authors participated in the review and editing of the paper.

Competing interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgements

We sincerely thank those scholars who measured and recorded the data of the case, enabling us to collect and organize the data related to landslide dam from the literature and materials they have published. At the same time, we would like to express our gratitude to the topic editors and the anonymous referees for their valuable suggestions.

Financial support

This research was supported by the National Natural Science Foundation of China (grant no. 42177149).

Reference

Azmi, Monte, and Kyle Thomson. 2024. “Dam Breach Parameters: From Data-Driven-Based Estimates to 2-Dimensional M odeling.” *Natural Hazards* 120 (5): 4423–61. https://doi.org/10.1007/s11069-023-06382-3.

Chen, Guan, Jiacheng Jin, Xingmin Meng, et al. 2024. “Influence of Tectonic Effects on the Formation and Characteristics o f Landslide Dams on the NE Tibetan Plateau: A Case Study in the Bailong River Basin, China.” *Landslides* 21 (9): 2135–53. https://doi.org/10.1007/s10346-024-02273-1.

Chen, Su-Chin, Chi-Yao Hung, Pei-Yi Chen, et al. 2026. “Synchronized Multidisciplinary Observations in Large-Scale Dam Breach Experiments to Enhance the Understanding of Dam Failure Evolution.” *Water Resources Research* 62 (1): e2025W R040786. https://doi.org/10.1029/2025WR040786.

Cheng, Haiguang, Kaiheng Hu, Shuang Liu, et al. 2025. “A Worldwide Event-Based Debris Flow Barrier Dam Dataset from 1800 to 2023.” *Earth System Science Data* 17 (4): 1573–93. https://doi.org/10.5194/essd-17-1573-2025.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Costa, John E., and Robert L. Schuster. 1988. "The Formation and Failure of Natural Dams." *Geological Society of America Bulletin* 100 (7): 1054–68. https://doi.org/10.1130/0016-7606(1988)100%253C1054:TFAFON%253E2.3.CO;2.

Costa, John E., and Robert L. Schuster. 1991. "Documented Historical Landslide Dams from around the World." Proceedings of the International Symposium on Landslides.

Dai, Lanxin, Xuanmei Fan, Dan Wang, et al. 2023. "Electrical Resistivity Tomography Revealing Possible Breaching Mechanism of a Late Pleistocene Long-Lasted Gigantic Rockslide Dam in Diexi, China." *Landslides* 20 (7): 1449–63. https://doi.org/10.1007/s10346-023-02048-0.

Dong, Jia-Jyun, Yun-Shan Li, Chyh-Yu Kuo, et al. 2011. "The Formation and Breach of a Short-Lived Landslide Dam at Hsiaolin Village, Taiwan — Part I: Post-Event Reconstruction of Dam Geometry." *Engineering Geology* 123 (1–2): 40–59. https://doi.org/10.1016/j.enggeo.2011.04.001.

Dufresne, Anja, Xuanmei Fan, and Wolter Andrea. 2023. "Landslide Dams around the World - Case Studies to Global Datasets." May 1, EGU-5214. https://doi.org/10.5194/egusphere-egu23-5214.

Dun, Jiawei, Wenkai Feng, Xiaoyu Yi, et al. 2025. "Detection and Updation of Landslide Inventory before and during Impoundment in the Baihetan Reservoir Area Using Multi-Temporal InSAR Datasets." *Scientific Reports* 15 (1): 9889. https://doi.org/10.1038/s41598-025-94520-1.

El Bilali, Ali, and Abdeslam Taleb. 2025. "A Novel Approach for Predicting Peak Flow from Breached Dam: Coupling Monte Carlo Simulation, Hydrodynamic Model, and an Interpretable XGBoost Model." *Water Resources Management* 39 (3): 1177–94. https://doi.org/10.1007/s11269-024-04018-0.

Ermini, L., and N. Casagli. 2003. "Prediction of the Behaviour of Landslide Dams Using a Geomorphological Dimensionless Index." *Earth Surface Processes and Landforms* 28 (1): 31–47. https://doi.org/10.1002/esp.424.

Fan, Xuanmei, Anja Dufresne, Srikrishnan Siva Subramanian, et al. 2020. "The Formation and Impact of Landslide Dams – State of the Art." *Earth-Science Reviews* 203 (April): 103116. https://doi.org/10.1016/j.earscirev.2020.103116.

Fan, Xuanmei, Anja Dufresne, Jim Whiteley, et al. 2021a. "Recent Technological and Methodological Advances for the Investigation of Landslide Dams." *Earth-Science Reviews* 218 (July): 103646. https://doi.org/10.1016/j.earscirev.2021.103646.

Fan, Xuanmei, Anja Dufresne, Jim Whiteley, et al. 2021b. "Recent Technological and Methodological Advances for the Investigation of Landslide Dams." *Earth-Science Reviews* 218 (July): 103646. https://doi.org/10.1016/j.earscirev.2021.103646.

Fan, Xuanmei, Weiwei Zhan, Xiujun Dong, et al. 2018. "Analyzing Successive Landslide Dam Formation by Different Triggering Mechanisms: The Case of the Tangjiawan Landslide, Sichuan, China." *Engineering Geology* 243 (September): 128–44. https://doi.org/10.1016/j.enggeo.2018.06.016.

Feng, Zhen-yu, Jia-wen Zhou, Xing-guo Yang, Long-jin Tan, and Hai-mei Liao. 2025. "Prediction of Landslide Dam Stability and Influencing Factors Analysis." *Engineering Geology* 350 (May): 108021. https://doi.org/10.1016/j.enggeo.2025.108021.

Gao, Shaohua, Yang Gao, Yueping Yin, et al. 2025. "Characteristics of Massive Glacier-Related Watershed Geohazard Chains in the Eastern Himalayan Syntaxis, China." *Journal of Earth Science* 36 (3): 1181–97. https://doi.org/10.1007/s12583-024-0116-y.

Huang, Bolin, Zhen Qin, Bin Li, et al. 2026. "New Methodology for the Seamless Numerical Simulation of Flow-like Landsliding–Damming–Breaching Process: Application to the 2020 Shaziba Landslide, China." *Computers and Geotechnics* 190 (February): 107742. https://doi.org/10.1016/j.compgeo.2025.107742.

Jian, Fu-xian, Zheng-yin Cai, and Wan-li Guo. 2023. "Laboratory-Scale Investigation of the Material Distribution Characteristics of Landslide Dams in U-Shaped Valleys." *Journal of Mountain Science* 20 (3): 688–704. https://doi.org/10.1007/s11629-022-7664-3.

Jiang, X., T. Wen, and G. Xiao. 2026. "Bridging the Data Gap: An Enhanced Global Inventory for Statistical Characterization and Breach Prediction of Landslide Dams." Zenodo. https://doi.org/10.5281/zenodo.19198720.

Jónatas, Ricardo Jorge Lourenço, Rui Jorge Ferreira Aleixo, Sílvia Rute Caleiro Amaral, Solange Valente Mendes, Maria Teresa Fontelas Santos Viseu, and Rui Miguel Lage Ferreira. 2025. "Laboratory Application of Real-Time LiDAR Technology for Assessing Breach Morphology during the Failure of Earth Dams and Fluvial Dikes." *Journal of Hydraulic Engineering* 151 (6): 04025033. https://doi.org/10.1061/JHEND8.HYENG-14126.

Korup, Oliver. 2005. "Geomorphic Hazard Assessment of Landslide Dams in South Westland, New Zealand: Fundamental Problems and Approaches." *Geomorphology* 66 (1–4): 167–88. https://doi.org/10.1016/j.geomorph.2004.09.013.

Korup, Oliver, David R. Montgomery, and Kenneth Hewitt. 2010. "Glacier and Landslide Feedbacks to Topographic Relief in the Himalayan Syntaxes." *Proceedings of the National Academy of Sciences* 107 (12): 5317–22. https://doi.org/10.1073/pnas.0907531107.

Kumar, Vipin, Imlirenla Jamir, Vikram Gupta, and Rajinder K. Bhasin. 2021. "Inferring Potential Landslide Damming Using Slope Stability, Geomorphic Constraints, and Run-out Analysis: A Case Study from the NW Himalaya." *Earth Surface Dynamics* 9 (2): 351–77. https://doi.org/10.5194/esurf-9-351-2021.

Lei, Yunlong, Marwan A. Hassan, Giorgio Rosatti, et al. 2025. "Back-Analysis of the 2000 Yigong Dam Breach Flood Morphodynamics: Challenges and Promises." *Geomorphology* 472 (March): 109588. https://doi.org/10.1016/j.geomorph.2024.109588.

Li, Dongyang, Tingkai Nian, Hao Wu, Fawu Wang, and Lu Zheng. 2020. "A Predictive Model for the Geometry of Landslide Dams in V-Shaped Valleys." *Bulletin of Engineering Geology and the Environment* 79 (9): 4595–608. https://doi.org/10.1007/s10064-020-01828-5.

Li, Ming-Hsu, Rui-Tang Sung, Jia-Jyun Dong, Chyi-Tyi Lee, and Chien-Chih Chen. 2011. "The Formation and Breaching of a Short-Lived Landslide Dam at Hsiaolin Village, Taiwan — Part II: Simulation of Debris Flow with Landslide Dam Breach." *Engineering Geology* 123 (1–2): 60–71. https://doi.org/10.1016/j.enggeo.2011.05.002.

Li, Xiaojun, Xiaobo Zhang, Jun He, Yixiang Song, and Yanqi Li. 2025. "Enhancing Landslide Dam Stability Prediction: A Data-Driven Framework Integrating Missing Data Imputation and Optimal Threshold Discrimination." *Frontiers in Earth Science* 13 (July): 1642791. https://doi.org/10.3389/feart.2025.1642791.

Li, Zhihai, Xiangjun Pei, Zhigang Shan, et al. 2025. "The Kinematic Process and Geomorphological Impact of the Jiaobunong Paleolandslide Dam in the Eastern Himalayan Syntaxis." *Bulletin of Engineering Geology and the Environment* 84 (1): 34. https://doi.org/10.1007/s10064-024-04038-5.

Liu, Che-Hsin, Jui-Yi Ho, Horn-Ru Liao, et al. 2025. "Using Unmanned Aerial Vehicle Monitoring for the Landslide and Dammed Lake: A Case Study of the Xiuluan Area in Taiwan." *Bulletin of Engineering Geology and the Environment* 84 (6): 340–55. https://doi.org/10.1007/s10064-025-04361-5.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Liu, Ning. 2014. *Remediation and Integrated Management of Red Rock Rockfall Lake Reservoir*. 16 (10): 39–46.

Luo, H. Y., P. Shen, L. M. Zhang, and J. He. 2025. "Numerical Investigation on the Landslide Dam Formation in Landslide-River Interaction." *Computers and Geotechnics* 180 (April): 107118. https://doi.org/10.1016/j.compgeo.2025.107118.

Lützow, Natalie, Georg Veh, and Oliver Korup. 2023. "A Global Database of Historic Glacier Lake Outburst Floods." *Earth System Science Data* 15 (7): 2983–3000. https://doi.org/10.5194/essd-15-2983-2023.

Malheiro, Ana, Francisco Fernandes, and Helder I. Chaminé, eds. 2023. *Advances in Natural Hazards and Volcanic Risks: S haping a Sustainable Future: Proceedings of the 3rd International Workshop on Natural Hazards (NATHAZ'22), Terceira Is land—Azores 2022*. Advances in Science, Technology & Innovation. Springer Nature Switzerland. https://doi.org/10.1007/9 78-3-031-25042-2.

Morgenstern, Regine, Andrea Wolter, Simon C. Cox, et al. 2024. "The New Zealand Landslide Dam Database, v1.0." *Landsl ides* 21 (1): 121–34. https://doi.org/10.1007/s10346-023-02133-4.

Nasser, Mohammed, Eleyas Assefa, Siraj M. Assefa, Constantinos C. Sachpazis, and Lysandros Pantelidis. 2026. "Adaptive Multihazard Modeling Predicts Rainfall-Driven Dam Failure: A Case Study." *Scientific Reports*, ahead of print, February 23. https://doi.org/10.1038/s41598-026-36927-y.

Niazi, Fawad S., Aranzazu Pinan-Llamas, and Kamran Akhtar. 2020. "Lessons from the Case History of a Massive Landslid e Dam." *Geofluids* 2020 (November): 1–32. https://doi.org/10.1155/2020/8840629.

Peng, Ming, Sitong Ji, Rui Sun, Limin Zhang, Shuai Zhang, and Zewen Bai. 2026. "Assessing the Probabilistic Evolution of Cascading Hazards of Landslide-River Blockage-Dam Breaching-Flood by Integrating Physics-Based and Data-Driven Meth ods." *Reliability Engineering & System Safety* 269 (May): 112097. https://doi.org/10.1016/j.ress.2025.112097.

Rafiq, Mohammd, Shakil Ahmad Romshoo, Anoop Kumar Mishra, and Faizan Jalal. 2019. "Modelling Chorabari Lake Outb urst Flood, Kedarnath, India." *Journal of Mountain Science* 16 (1): 64–76. https://doi.org/10.1007/s11629-018-4972-8.

Rimal, Bhagawat, and Abhishek Tiwary. 2024. "Monitoring Hazards in Dam Environments Using Remote Sensing Techniqu es: Case of Kulekhani-I Reservoir in Nepal." *Earth* 5 (4): 873–95. https://doi.org/10.3390/earth5040044.

Risley, John C., Joseph S. Walder, and Roger P. Denlinger. 2006. "Usoi Dam Wave Overtopping and Flood Routing in the B artang and Panj Rivers, Tajikistan." *Natural Hazards* 38 (3): 375–90. https://doi.org/10.1007/s11069-005-1923-9.

Shafieiganjeh, Roshanak, Marc Ostermann, Barbara Schneider-Muntau, and Bernhard Gems. 2022. "Assessment of the Land slide Dams in Western Austria, Bavaria and Northern Italy (Part of the Eastern Alps): Data Inventory Development and Appl ication of Geomorphic Indices." *Geomorphology* 415 (October): 108403. https://doi.org/10.1016/j.geomorph.2022.108403.

Shi, Ning, Yanlong Li, Lifeng Wen, and Ye Zhang. 2022. "Rapid Prediction of Landslide Dam Stability Considering the Mis sing Data Using XGBoost Algorithm." *Landslides* 19 (12): 2951–63. https://doi.org/10.1007/s10346-022-01947-y.

Shi, Ning, Yanlong Li, Lifeng Wen, Ye Zhang, and Haiyang Zhang. 2023. "Longevity Prediction and Influencing Factor An alysis of Landslide Dams." *Engineering Geology* 327 (December): 107334. https://doi.org/10.1016/j.enggeo.2023.107334.

Silwal, Bishow Raj, Katsuichi Ota, and Kohki Yoshida. 2024. "Effects of Hydrothermal Activity and Weathering in the Acti ve Fault Area: Formation of Large Landslide and Landslide Dam Lake, Lake Nakatsuna, Nagano, Japan." *Natural Hazards* 1 20 (9): 9057–91. https://doi.org/10.1007/s11069-024-06567-4.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Singh, Amanpreet, Vipul Anand, K. H. V. Durga Rao, Shashivardhan Reddy P., and Prakash Chauhan. 2025. "Unveiling the Catastrophic Landslide-Induced Flash Flood in Teesta River, Sikkim: Insight from South Lhonak Glacial Lake." *Landslides* 22 (3): 837–55. https://doi.org/10.1007/s10346-024-02378-7.

Song, Chenguang. 2025. "A Novel Rapid Approach for the Stability Discrimination of Landslide Dams." *Bulletin of Engineering Geology and the Environment* 84 (4): 185. https://doi.org/10.1007/s10064-025-04209-y.

Song, Liang, Liqiang Ma, Zhiguo Chang, and Yunsheng Wang. 2026. "Distribution Patterns and Formation Mechanisms of Stable Landslide Dams in the Qinghai-Tibet Plateau: A Preliminary Study." *Scientific Reports* 16 (1): 5399. https://doi.org/10.1038/s41598-025-33618-y.

Struble, William T., Joshua J. Roering, William J. Burns, Nancy C. Calhoun, Logan R. Wetherell, and Bryan A. Black. 2021. "The Preservation of Climate-Driven Landslide Dams in Western Oregon." *Journal of Geophysical Research: Earth Surface* 126 (4): e2020JF005908. https://doi.org/10.1029/2020JF005908.

Tacconi Stefanelli, Carlo, Filippo Catani, and Nicola Casagli. 2015. "Geomorphological Investigations on Landslide Dams." *Geoenvironmental Disasters* 2 (1): 21–134. https://doi.org/10.1186/s40677-015-0030-9.

Takayama, Shoki, Kosei Ikeda, and Fumitoshi Imaizumi. 2026. "Longitudinal Evolution of Landslide Dam Geometries during Overtopping." *Landslides* 23 (1): 19–34. https://doi.org/10.1007/s10346-025-02645-1.

Tavus, Beste, Sultan Kocaman, Hakan A. Nefeslioglu, and Candan Gokceoglu. 2024. "Analysing Slope Dynamics of Kalecöy (Türkiye) Dam Reservoir with Sentinel-1 SAR Time Series and Sentinel-2 Spectral Indices." *Environmental Earth Sciences* 83 (17): 510–25. https://doi.org/10.1007/s12665-024-11807-8.

Vassileva, Magdalena, Mahdi Motagh, Sigrid Roessner, and Zhuge Xia. 2023. "Reactivation of an Old Landslide in North–Central Iran Following Reservoir Impoundment: Results from Multisensor Satellite Time-Series Analysis." *Engineering Geology* 327 (December): 107337. https://doi.org/10.1016/j.enggeo.2023.107337.

Wang, Danyan, Xingguo Yang, Jiawen Zhou, Zhenyu Feng, and Haimei Liao. 2025. "Longevity Prediction and Missing Data Treatment of Landslide Dams." *Journal of Mountain Science* 22 (7): 2640–53. https://doi.org/10.1007/s11629-024-9310-8.

Wang, Fawu, Austin Chukwueloka-Udechukwu Okeke, Tetsuya Kogure, Tetsuya Sakai, and Hisao Hayashi. 2018. "Assessing the Internal Structure of Landslide Dams Subject to Possible Piping Erosion by Means of Microtremor Chain Array and Self-Potential Surveys." *Engineering Geology* 234 (February): 11–26. https://doi.org/10.1016/j.enggeo.2017.12.023.

Wu, Hang, Mark A. Trigg, William Murphy, et al. 2024. "A Global-Scale Applicable Framework of Landslide Dam Formation Susceptibility." *Landslides* 21 (10): 2399–416. https://doi.org/10.1007/s10346-024-02306-9.

Wu, Hang, Mark A. Trigg, William Murphy, and Raul Fuentes. 2022. "A New Global Landslide Dam Database (RAGLAD) and Analysis Utilizing Auxiliary Global Fluvial Datasets." *Landslides* 19 (3): 555–72. https://doi.org/10.1007/s10346-021-01817-z.

Wu, Hang, Mark A. Trigg, William Murphy, and Raul Fuentes. 2025. "Analysis of the Fundamental Differences between Dam-Forming Landslides and All Landslides." *Geomorphology* 475 (April): 109665. https://doi.org/10.1016/j.geomorph.2025.109665.

Wu, Hang, Mark Trigg, and William Murphy. 2020. *A Global Scale Geospatially Located Landslide Dam Dataset*. March 23. https://doi.org/10.5194/egusphere-egu2020-5881.

https://doi.org/10.5194/essd-2026-107
Preprint. Discussion started: 2 April 2026

Xie, Lei, Wenbin Xu, and Yosuke Aoki. 2025. "Extracting a Decadal Deformation on Xiaolangdi Upstream Dam Slope Usin
g Seasonally Inundated Distributed Scatterers InSAR (SIDS - InSAR)." *International Journal of Applied Earth Observation
and Geoinformation* 138 (April): 104462. https://doi.org/10.1016/j.jag.2025.104462.

Yi, Shuang, Hao-si Li, Shin-Chan Han, et al. 2025. "Quantification of the Flood Discharge Following the 2023 Kakhovka D
am Breach Using Satellite Remote Sensing." *Water Resources Research* 61 (3): e2024WR038314. https://doi.org/10.1029/20
24WR038314.

Zhou, Yuanyuan, Zhenming Shi, Hongchao Zheng, and Chengzhi Xia. 2025. "Numerical Study on Morphology and Material
Spatial Distribution of Landslide Dams in Different Shaped Valleys." *Geomorphology* 478 (June): 109727. https://doi.org/1
0.1016/j.geomorph.2025.109727.