https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

Open Access Earth System Science Data Discussions

A worldwide event-based debris-flow barrier dam dataset from 1800 to 2023

Haiguang Cheng[a,b,c], Kaiheng Hu[a,c*], Shuang Liu[a,c], Xiaopeng Zhang[a,b,c], Hao Li[a,b,c], Qiyuan Zhang[a,c], Lan Ning[a,b,c], Manish Raj Gouli[a,b,c], Pu Li[a,c], Anna Yang[a,b,c], Peng Zhao[a,b,c].

[a]Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, China;

[b]University of Chinese Academy of Sciences, Beijing 100049, China;

[c]State Key Laboratory of Mountain Hazards and Engineering Resilience, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu, Sichuan, China.

*Correspondence to:

Kaiheng Hu (khhu@imde.ac.cn)

Abstract: Debris flows, as a special kind of landslides, often block rivers to form barrier dams and trigger a series of disasters such as upstream aggradation and outburst floods. The understanding of debris-flow barrier dams (DFBDs) is poor, mostly due to existing researches focusing on individual events and a lack of summarization of multiple DFBD events. The existing global or regional datasets of landslide barrier dams (LDs) contain only a few cases of DFBDs, and ignore the differences between DFBDs and other landslide barrier dams (LDs), such as the dams of rock slide, debris avalanche, or earth slide. To fill this gap, we reviewed 2519 literatures and media reports with high quality. Focusing on identified debris-flow damming events, a rigorous data review and validation process was conducted using Google Earth. A systematic approach was employed to prioritize conflicting information from various data sources. Consequently, a global dataset was compiled, encompassing 555 historical DFBDs from 1800 to 2023.

This pioneering global dataset includes five categories and 36 attributes, detailing DFBDs. It captures basic information (location, the date of formation, etc.), dam characteristics (height, length, volume, etc.), lake characteristics (area, capacity, length), debris flow characteristics (velocity, discharge, volume, etc.), and failure characteristics (peak discharge, loss of life, etc.). Our dataset elucidates that DFBDs exhibit key features of instability, complete blockage, and overtopping failure. The number of such dams has notably increased, especially in China. 15 % of channels showed recurrent debris flows, resulting in DFBDs that make up 35 % of all DFBDs. Further analysis recommends the *Ls* (*AHV*) model is recommended for priority use, followed by the *DBI* model, for the stability assessment of DFBDs. Compared to other barrier dam datasets, our dataset is more targeted, lays a greater emphasis on the review of raw data, and stresses the unification of terminology and concepts (such as blockage modes and stability), ensuring the consistency and accuracy of the data. The dataset and results in this work may help to deepen the understanding of DFBD formation, distribution, and evolution. The DFBD dataset can be accessed through this link: https://doi.org/10.5281/zenodo.13382846 (Cheng et al., 2024).

1 Introduction

Debris flows, composed of fine and coarse-grained components, boulders, woody, and water, are a rapid two-phase flow with non-zero yield stress (Hungr et al., 2014). When debris flows carry large amounts of sediment flowing rapidly in a valley, they may accumulate in a narrow river channel and form a barrier dam, that is, debris-flow barrier dam (DFBD) (Fan et al., 2020; Yin et al., 2016; Yu et al., 2022; Zhang et al., 2022). The formation of such barrier dams not only changes the original hydrogeological conditions, but may also results in secondary disasters, such as floods, landslides, and even larger debris flows, posing a serious threat to human society and the natural environment (Cui et al., 2016; Hu et al., 2011; Liu et al., 2019). For example, on August 7, 2010,

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

triggered by heavy rainfall, a large-scale debris flow broke out in Luojiayu and Sanyanyu Gully in Zhouqu County, China. After the debris flow passed through Zhouqu City, it blocked the Bailong River. The water level in the upper reaches rose sharply, which submerged half of Zhouqu City, resulting in 1364 casualties and 401 missing (Chong et al., 2021; Hu et al., 2010).

Currently, researches on DFBDs mainly focus on a single event (Hu et al., 2010; 2011), or physical and numerical experiments conducted with a single event as the prototype, focusing on the research of river obstruction by debris flows (Chen et al., 2022; Dang et al., 2009; Ruan et al., 2021). In terms of properties and scale, debris flows that form barrier dams are typically large-scale and cohesive, with high density and uniformity, exhibiting considerable resistance to erosion (Chen et al., 2019; Ruan et al., 2021). In terms of topography, the rivers and valleys blocked by DFBDs are generally narrow, with steep terrain slopes (Song et al., 2023; Wang et al., 2017; Yu et al., 2022).

Isolated studies of individual DFBD events cannot reflect the overall distribution characteristics. However, statistical analysis of a great number of historical data on barrier dam disasters can help to clarify this issue. Some scholars have conducted extensive researches on parameters such as geometric characteristics, breaching, longevity, and stability of barrier dams by establishing datasets (Casagli et al., 2003; Dong et al., 2014; Fan et al., 2012; 2017; Peng and Zhang, 2012a; b; Stefanelli et al., 2015; 2016). However, there are relatively few cases of DFBDs in these datasets. The conclusions drawn from these barrier dam datasets may not be applicable to DFBDs. Therefore, there is an urgent need to establish a global comprehensive dataset specifically for DFBDs, laying a data foundation for in-depth research on such dams in the future, which is one of the goals of this study.

After the formation of a barrier dam, timely predictions of the stability of the dam and the outburst peak discharge are the keys to formulating disaster reduction measures, and it are also hot topics in barrier dams-related

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

researches (Azimi et al., 2015; Casagli and Ermini, 1999; Korup, 2004). Based on the statistical analysis methods, some scholars analyzed the influence of dam structure characteristics, dam material characteristics, hydrological characteristics, and other factors on the stabilities of dams, and established some models for evaluating barrier dam stability (Dong et al., 2011; Ermini and Casagli, 2003). Other studies based on historical statistical cases, summarized parameter models for the peak discharge, in order to achieve rapid prediction of peak discharge of barrier dam breach (Azimi et al., 2015; Hakimzadeh et al., 2014; Hooshyaripor et al., 2014). However, these studies did not strictly differentiate the barrier dams, focusing more on LDs. Considering that DFBDs have unique characteristics compared to LDs (Cheng et al., 2007a; b; Dang et al., 2009; Ruan et al., 2021), the applicability of stability and peak discharge models, originally designed for LDs, to DFBDs remains unclear. This constitutes the second key issue to be explored in this study.

This study establishes a dataset containing 555 DFBDs worldwide by exploring 2519 literatures and media reports. This dataset contains information of DFBDs on the formation time, location, geometric characteristics, longevity, peak discharge, failure characteristics, blockage modes, failure mechanisms, stability, loss of life, etc. A detailed analysis was conducted on the spatiotemporal distribution, blockage modes, failure mechanisms, longevity, and stability of DFBDs. The applicability of stability and peak discharge models, of LDs, for DFBDs was discussed. Compared with other datasets, our dataset stands out for its emphasis on the unity of terminology and concepts, as well as the review and validation of raw data, to ensure consistency and accuracy of the data.

2 Data and method

2.1 Data sources

In the process of building the dataset, we adopted a comprehensive and

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

systematic approach to collect and analyze data. The data sources, totally 2519, mainly included peer-reviewed scientific literature, data released by government agencies, proceedings from professional conferences, and reports from authoritative news media. We placed special emphasis on selecting publications that have a high reputation and professionalism in the field to ensure the accuracy and authority of the data. To ensure the breadth and diversity of the data, we have made every effort to consult academic journal literature from different countries and regions (China, Japan, Taiwan, the United States, Italy, etc.) to obtain records of DFBDs from various perspectives. Additionally, recognizing that media reports offer real-time and first-hand information on the formation and impact of DFBDs, we meticulously collected and reviewed coverage from mainstream media.

Many barrier dam datasets have been established by compiling significant historical events. Although they have included a limited number of DFBDs and the related information is not comprehensive, they have provided us with a wealth of clues that facilitate the collection of information on DFBDs. These datasets are one of the main data sources in this study (Table 1). We conducted a rigorous screening of barrier dams in the existing dataset and further supplemented and refined the information related to the screened DFBDs. In addition, most of the cases in our dataset were sourced from individual studies of debris flow river-blocking events from various regions.

Table 1. The most relevant inventories and datasets of DFBDs

NO.	Sources	Region, Country	Information on the inventory	Total number of dams	Number of DFBDs
1	Costa and Schuster (1991)	World	Including location, triggering mechanism, landslide type, size, failure time and mechanism, etc.	463	98
2	Ermini and Casagli (2003)	World	Most of the collected cases are from the western USA, Japan, and Italy. A common feature is a high	350	3

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

			frequency of damming events along or near the active margin areas of tectonic plates.		
3	Tong (2008)	World	Including country, date of formation, dam volume, lake area,	84	16
4	Liu et al., (2019)	China	Including, location, date of formation, dam volume, dam height, lake volume, triggering factors	287	4
5	Yan (2016)	China	Including, location, date of formation, and impacts.	226	38
6	Chai et al., (1995)	China	Including location, date, landslide volume, Longevity, triggering factors and impacts.	147	7
7	Schuster and Costa (1986)	World	Containing the failure mechanism, and longevity.	187	3
8	Tacconi Stefanelli et al., (2016)	Italy	Containing the date of formation, location, failure mechanism, dam height, dam length, dam volume, lake area, lake volume, longevity, and impacts.	300	33
9	Fan et al. (2020)	world	Containing the date of formation, location, failure mechanism, dam height, dam length, dam volume, lake area, lake volume, longevity, impacts, and loss of life.	1886	34

2.2 Dataset content

The statistical principles used for data collection are as follows: We collected all DFBDs if they had clear records of the country and location in the literatures and media reports. Therefore, it can be considered that data collection process is free of subjective bias, and the dataset has

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

statistical significance. The descriptions were classified and organized in five categories: basic information, debris flow characteristics, dam characteristics, lake characteristics, and failure characteristics. The five categories included 36 attributes to characterize DFBDs, as detailed in Table 2. The basic information includes the name, country, longitude and latitude, time of dam formation, trigger, reference, and reliability. The debris flow characteristics include the debris flow channel slope gradient, debris flow channel length, debris flow gully basin area, debris flow density, debris flow velocity, debris flow peak discharge, and debris flow volume. The dam characteristics include the blockage modes, dam volume, dam height, dam length, dam width, dam material, longevity, stability, and the controls. The lake characteristics include lake length, lake area, and lake volume. The failure characteristics include failure mechanism, breach depth, breach top width, breach bottom width, breaching time, peak discharge, average discharge, and loss of life. Detailed explanations for each attribute are provided in Table 2.

Incorporating relevant information on debris flows into this dataset can provide a more comprehensive record of detailed information on DFBD events. In addition, this category provides convenience for potential users, enabling them to cross-validate and compare this dataset with other datasets, such as "Two multi-temporal datasets to track debris flow after the 2008 Wenchuan earthquake" (Wang et al., 2022), and "The ITAlian rainfall-induced LandslIdes Catalogue" (Peruccacci et al., 2023). This cross-use and mutual corroboration enhances the reliability of data and the universality of application.

Table. 2 Data present in the DFBD dataset with units.

Category	Attribute	Symbol	Description	Unit
Basic information	EFBD_ID	*ID*	Unique identifier for each individual DEBD, starting at 1.	[-]
	Name	*Na*	Names of DFBDs.	[-]
	Country	*Cou*	Name of the country in which the DFBD formed, as listed by the U.S. Board of Geographic Names or included in The Times	[-]

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

			Atlas of the World, 7th edition, 1988.	
	Longitude	*Lon*	Longitude of the reported events.	[°, WGS 1984]
	Latitude	*Lat*	Latitude of the reported events.	[°, WGS 1984]
	Date of formation	D_f	The date the DFBD was formed, if known.	[yyyy/mm/dd]
	Trigger	*Tri*	Main factor that initiated the debris flows.	[-]
	Reference	*Re*	Sources of information about individual DFBD.	[-]
	Reliability	*R*	The reliability proposed in this dataset is used to describe the credibility of the data, which is divided into low reliability, medium reliability, and high reliability.	[-]
Debris flow characteristic	Debris flow channel slope gradient	S_{df}	The change rate of height difference in unit horizontal distance along the flow direction of debris flow channel	[%]
	Debris flow channel length	L_{df}	The distance of debris flow movement path in the channel.	[km]
	Debris flow gully basin area	A_{df}	The total area of the ground surface that directly or indirectly collects water flow into the debris flow channel	[km^2]
	Debris flow density	C_{df}	The weight of a debris flow per unit volume before the debris flow rushes into the main river channel.	[$g\ cm^{-3}$]
	Debris flow velocity	V_{df}	The velocity of debris flow movement along the channel before the debris flow rushes into the main river channel.	[$m\ s^{-1}$]
	Debris flow peak discharge	Q_{df}	The maximum discharge of a debris flow just before it blocks a river.	[$m^3\ s^{-1}$]
	Debris flow volume	V_{df}	The volume of debris flow rushed into the river channel.	[$10^6\ m^3$]
Dam characteristic	Blockage mode	*BM*	Blockage of river course by debris flow. Here, there are	[-]

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

			three modes of complete blockage, partial blockage and submerged dam blockage (Fig. 1(b)).	
	Dam volume	V_d	The part of the debris flow volume that blocks the river (Fig. 1(a)).	[10^6 m^3]
	Dam height	H_d	The vertical altitude difference from the river bed to the overflow point on the barrier dam (Fig. 1(a)).	[m]
	Dam length	L_d	The crest length of the barrier dam measured perpendicular to the major valley axis (Fig. 1(a)).	[m]
	Dam width	W_d	The base width of the landslide dam measured parallel to the main valley axis (Fig. 1(a)).	[m]
	Dam material	*DM*	General type of material that constitutes the DFBD.	[-]
	Controls	*Con*	Any physical modifications made to the DFBD to help minimize volume of impounded water, artificially lower height, change the geometry of dam, or prevent erosion upon overtopping.	[-]
	Longevity	*T*	The time from formation to failure.	[day]
	Stability	*Sta*	Stability refers to the real-time state of the dam.	[-]
Lake characteristic	Lake length	L_l	Length of backwater ponded behind dam, measured upstream from dam (Fig. 1(a)).	[m]
	Lake area	A_l	The surface area of water ponded behind the DFBD (Fig. 1(a)).	[km^2]
	Lake volume	V_l	The volume of water ponded behind the DFBD (Fig. 1(a)).	[10^6 m^3]
Failure characteristic	Failure mechanism	*FM*	The mechanism that led to dam failure or breach. Where more than one failure mechanism was involved, the most severe was selected. Here, there are three types mechanisms of	[-]

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

		overtopping (OT), piping (PP), and slope failure (SF) (Fig. 1(c)).	
Breach depth	H_b	The vertical distance from the lowest bottom of the breach to the original lowest point on the landslide dam crest (Fig. 1(a)).	[m]
Breach top width	W_t	The width of the breach at the height of the dam crest (Fig. 1(a)).	[m]
Breach bottom width	W_b	The width of the bottom of the breach (Fig. 1(a)).	[m]
Breaching time	T_b	The period from the inception to the completion of the breaching process (Singh and Snorrason 1984).	[hour]
Peak discharge	Q_p	The peak discharge of outburst flood after dam failure.	[$m^3 s^{-1}$]
Average discharge	Q_a	The average discharge of outburst flood after dam failure.	[$m^3 s^{-1}$]
Loss of life	LF	The number of people who lost their lives in the DFBD incident	[-]

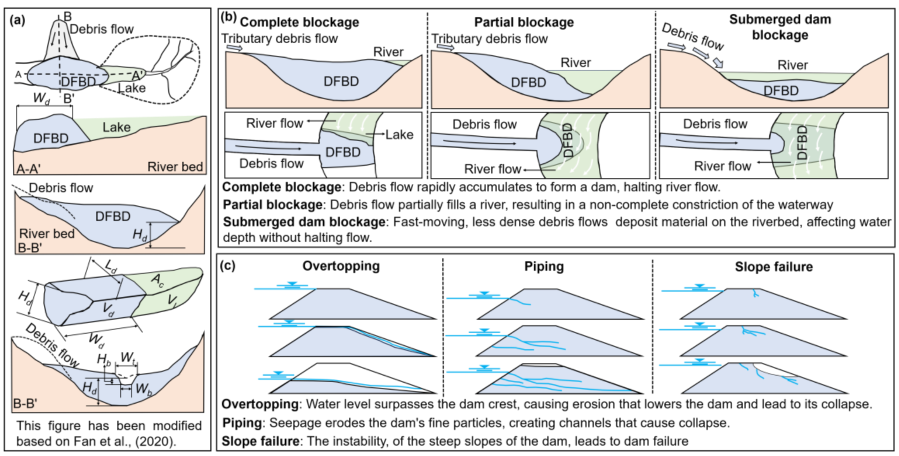

Figure.1 (a) The geometric characteristics; (b) the blockage modes; (c) failure mechanism.

2.3 Data processing procedure

We followed three meticulously designed steps: First, we conducted a comprehensive search for DFBD events; second, we reviewed and validated

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

key data; and lastly, we carried out data complementation (Fig. 2). In each step, we adhered to strict standards to ensure that all data included in the dataset underwent a thorough review to eliminate potential biases and errors.

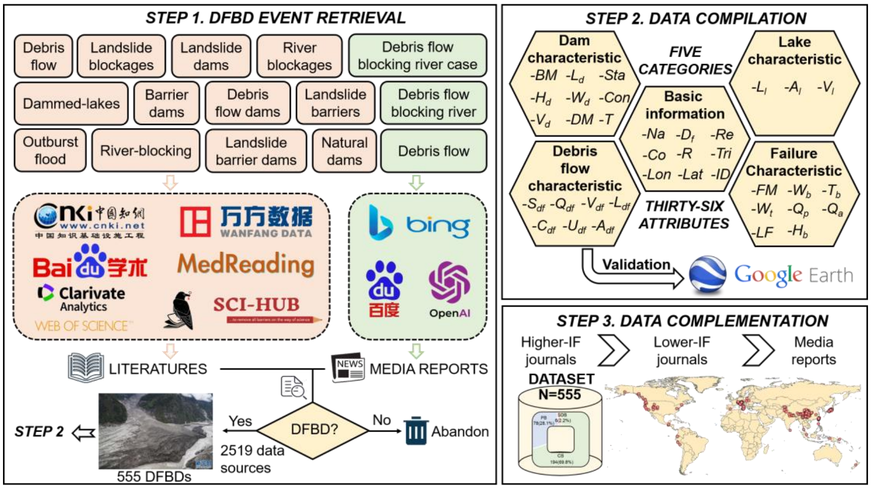

Figure. 2 Procedure for the compilation of the dataset.

In the first step, we were committed to extensively retrieving DFBD events, carefully consulting a wide range of academic literature and media reports in both Chinese and other languages. The collection of Chinese literature mainly relied on online service platforms such as CNKI, Wanfang, Baidu Scholar, and Medreading. For non-Chinese literature, our search encompassed international academic databases such as Google Scholar and Web of Science, complemented by other reputable academic resources. It is important to note that the term 'Barrier dams' can vary depending on the authors, regions, and research focuses, leading to the use of different terms in parallel. Common synonyms found in the literature include 'landslide blockages', 'landslide dams', 'stream/river blockages', 'natural dams', 'landslide barriers', and 'dammed-lakes'. To collect DFBDs as comprehensively as possible, we used a range of keywords including 'debris flow', 'landslide blockages', 'landslide dams', 'river blockages', 'natural dams', 'landslide barriers', 'dammed-lakes', 'barrier dams', 'debris flow dams', 'landslide barrier dams', 'outburst flood', 'river blocking' for literature search. We also combined these with regional and national terms,

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

such as 'Chinese debris flow barrier dams', to enhance the search's relevance. To ensure search consistency and reduce errors, we used unified search keywords for both Chinese and non-Chinese literature. In addition, we inputted keywords such as 'debris flow blocking river', 'debris flow dam incident', and 'debris flow' into search engines such as Baidu and Google, as well as ChatGPT, and combined regional information to retrieve relevant dam events. Among the identified barrier dam cases, those of DFBDs were selected, included in this dataset, and the term 'debris-flow barrier dam (DFBD)' was adopted. We repeated this step multiple times during 2022 to 2024 to ensure that any updated DFBD events are included in our dataset.

In the second step, we conducted a thorough review and validation of the information gleaned from literature and media reports. It is important to note that we applied different validation methods to treat various attributes of DFBDs, as detailed below:

-***DFBD-ID***. We sorted the DFBDs by the first letter of their names.

-Name. When compiling this attribute, we adhered strictly to the original names of DFBDs as documented in the data sources. These names typically derive from the channels of the debris flows or the rivers that were obstructed. It is worth noting that a DFBD may have different names in different data sources. To avoid information redundancy and confusion, we have carefully checked using Google Earth based on geographical coordinates and the date of formation. We have identified and eliminated the DFBD events that were reported repeatedly in different data sources due to naming differences, ensuring that each event is uniquely named and recorded only once in our dataset.

-***Country***. When country information was reported, it was verified and confirmed through Google Earth. Once errors are found, we used the inquiry results from Google Earth. When data sources lack explicit national information, we determine the event's country using geographic coordinates or prominent

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

landmarks with Google Earth. This process ensures the accuracy and reliability of the country information recorded in our dataset, improving the overall quality of the data.

-***Longitude*** and ***Latitude.*** When latitude and longitude information were provided in the data sources, they were validated through Google Earth. If there were any deviations between the latitude and longitude information in the data sources and the Google Earth validation results, or if the data sources did not provide latitude and longitude information, we determined their latitude and longitude information based on Google Earth.

- ***Date of formation.*** For events with a long history, we relied on the formation dates provided in the data sources. For more recent events, we verified their formation dates using Google Earth. If there was a discrepancy between the reported formation dates and the Google Earth review results, we excluded this attribute from our dataset.

- Trigger, ***Debris flow channel slope gradient***, ***Debris flow channel length***, ***Debris flow gully basin area***, ***Debris flow density***, ***Debris flow velocity***, ***Debris flow peak discharge***, ***Debris flow volume***, ***Dam volume***, ***Dam height***, ***Dam length***, ***Dam width***, ***Dam material***, ***Controls***, ***Lake length***, ***Lake area***, ***Lake volume***, ***Failure mechanism***, ***Breach depth***, ***Breach bottom width***, ***Breach top width***, ***Breaching time***, ***Peak discharge***, ***Average discharge***, ***Loss of life***, and ***Reference***. These attribute data followed reports from different data sources. If there was conflicting information among different data sources, we performed the third step, namely data complementation.

-***Blockage mode.*** Costa and Schuster (1988) proposed six blockage modes: types I, II, III, IV, V, and VI. In this dataset, we reclassified the DFBD blockage modes based on the event description in the data sources. We regarded type I as partial blockage, types II, III, IV, and V as complete blockage, and type VI as submerged dam blockage. This classification method is helpful to record and understand the different blocking

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

characteristics of DFBDs more precisely.

-Longevity. We divided the DFBDs from different data sources into two types: dams that have been clearly reported to have failed, and the other is the dams that have not been clearly reported to have failed. For the former, we included the longevity attribute according to the report. For the second type of dams, according to the latitude and longitude information, Google Earth was used to view the latest remote sensing images to confirm the current status of the dams. If it is found that the dams no longer exist, we would compare and analyze the remote sensing images at different time to determine the duration of their existence, that is, longevity.

-Stability. We evaluated the stability based on its real-time status. When the literature and media reports indicated that dams have failed, we judged them as unstable dams. On the contrary, if the data sources claimed that the dams were still exist, we used Google Earth to further confirm their actual status. Once the images of the dams on Google Earth were found to show that they still exist, we classified them as stable dams; if they were not found or confirmed that they no longer exist, they were classified as unstable dams.

-***Reference***. We kept detailed records of data sources to ensure traceability and transparency of information. For academic literature, we detailed key information such as the publisher, date of publication, title, author name, and unique identifier DOI. As for media reports, we also meticulously recorded the URL links of the reports so that users could directly access the original reports.

-***Reliability***. We determine the credibility of debris-flow barrier dam events based on the number of data sources. When there is only one literature or one news report on a DFBD, we define the reliability of this event as low. When a DFBD event is reported by two data sources, the reliability of this event is medium. When an event is reported by three or more data sources, we consider the reliability of the event to be high.

The third step is data complementation. In the situations when there is

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

conflicting information among different data sources, we have adopted a hierarchy of information sources based on perceived reliability to resolve the issue: Priority was given to literature published in journals with higher impact factors; next were publications in journals with lower impact factors; lastly, media reports were considered. According to this priority rule, we have incorporated the conflicting information into our dataset to ensure the accuracy and reliability of the data.

3 Results

3.1 Reliability

Based on an in-depth review of 2519 literatures and news reports, we have recorded 555 DFBD events. To evaluate the reliability of these events, we have introduced the key attribute of "reliability". According to our analysis, the 555 DEBD events have a high reliability, with a total of 494 events, accounting for 89 % of the total. In addition, there are 48 events with medium reliability, which make up 8.7 % of the total, and there are only 13 events with low reliability, accounting for only 2.3 % (Fig. 3).

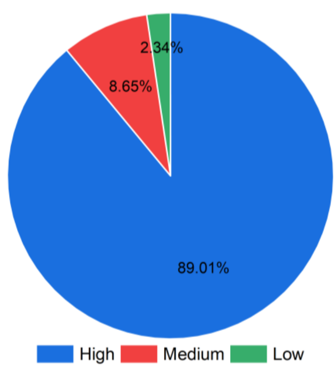

Figure 3. Event reliability.

3.2 Spatiotemporal distribution of the DFBDs

The 555 DFBDs lie in different countries, including 39 dams in Italy, 43 dams in Japan, 376 dams in China, 33 dams in the United States, and a total of 64 dams in other counties (Figs. 4(a) and 4(b)). Most of the DFBDs are distributed along the edge of tectonic plates. These areas are characterized by frequent earthquakes, obvious valley topography, broken strata and poor geological conditions due to plate stacking (Coviello et al., 2019; Zhang et al., 2023; Zhao et al., 2020; Zhao et al., 2023). For example, Italy is located at the junction of the Eurasian Plate and the African Plate, with frequent plate activity and geological disasters such as earthquakes. The Mediterranean climate throughout the country has a high rainfall, making it easy to form DFBDs caused by earthquakes and rainfall (Loche et al., 2022; Stefanelli et al., 2015; Tiranti et al., 2019). South Asia is also a high-risk area of DFBD disasters. The collision between the Indian Ocean Plate and the Eurasian Plate makes it easy for countries such as India and Nepal near the Himalayas to form DFBD disasters induced by earthquakes (Ruiz-Villanueva et al., 2017; Walsh et al., 2012). In addition, Japan is located at the junction of the Eurasian plate and the Pacific plate, and the reason for the frequent occurrence of DFBDs is similar to that of India and Nepal (Fan et al., 2020).

Since the 1900s, the number of global DFBDs has shown an overall upward trend. Between the 1900s and 1960s, the number experienced fluctuation and increase, but the growth range was relatively small (Fig. 4(c)). During this period, society's awareness and attention to such disasters were insufficient, resulting in limited records and reports. Between the 1960s and 1990s, the number of DFBDs showed a more significant increase. However, between the 1990s and 2000s, the number of reported DFBDs worldwide significantly decreased compared to the previous decade by approximately 1.5 times. Since the year 2000 to the present, the number of DFBDs has increased significantly, particularly reaching a peak in the last ten years. Global climate change may be one of the key factors leading to an increase in debris flows (Ma et al., 2024; Sharma et al., 2024; Yu et al., 2021). With the rise in global temperatures, extreme weather events such as heavy rainfall, droughts, and

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

floods have become more frequent. These extreme weather conditions are highly likely to induce the formation of debris flows and the blocking of rivers to form dams.

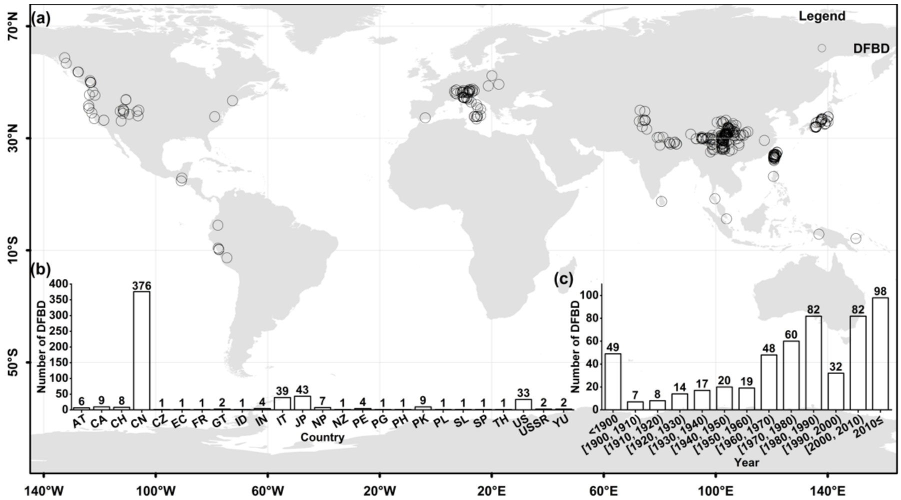

Figure.4 Worldwide DFBDs spatiotemporal distribution. (a) Spatial distribution; (b) temporal distribution.

The number of DFBDs in China is far greater than that in other countries. This is due to the fact that two-thirds of China's territory is mountainous, with significant terrain undulations, high mountains, and deep valleys, which are conducive to the occurrence of steep gullies and provide favorable topographical conditions for formation the of debris flows. In addition, some mountainous areas experience frequent heavy rains in summer, coupled with the melting of ice and snow, providing an abundant water source for debris flows, which is beneficial for blocking rivers with debris flows. As shown in Fig. 5(a), China's DFBDs are mainly concentrated in Sichuan Province, where there are frequent earthquakes along the Longmenshan Fault Zone, Xianshuihe Fault Zone, and Anninghe Fault Zone. (Cui et al., 2009; 2010; 2011; Fan et al., 2012; 2017; Gorum et al., 2010; Xu et al., 2009; Zhang et al., 2023).

Since the early 1960s, China vigorously promoted railway and highway construction, but mountain disasters such as debris flows have become increasingly prominent. In response to these challenges, the Chinese Academy

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

of Sciences established the Institute of Glaciology and Cryopedology in Lanzhou and the Institute of Mountain Hazards and Environment in Chengdu, respectively, to conduct systematic research on debris flow disasters. The establishment of these institutes marked that Chinese scholars began to focus on the phenomenon of debris flows blocking the rivers. In particular, the 2008 Wenchuan earthquake triggered many debris flows, further exacerbating the formation of DFBDs. These events not only had a huge impact on the local area, but also made the research and prevention of debris flow blocking rivers reach an unprecedented height. After the Wenchuan earthquake, Chinese scholars have paid more attention to the research and prevention of DFBDs, and the number of reported DFBDs increased significantly (Fig.5(b)).

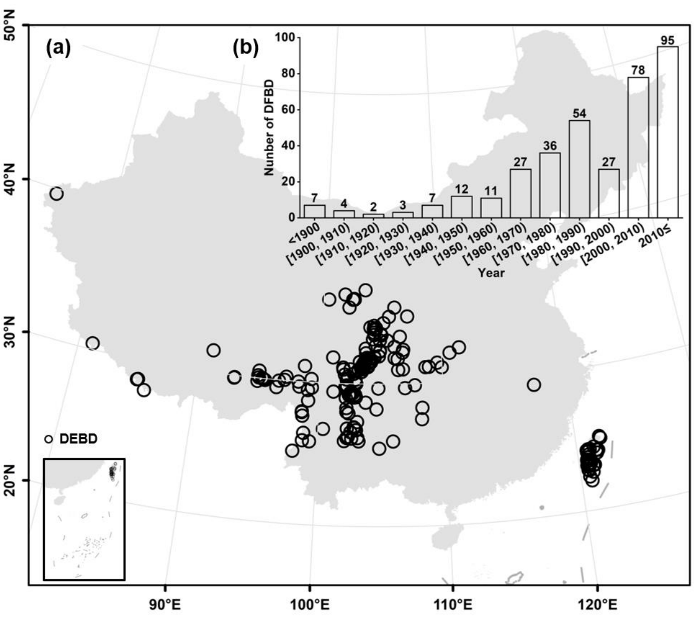

Figure. 5 Chinese DFBDs spatiotemporal distribution. (a) Spatial distribution; (b) temporal distribution.

3.3 DFBD blockage modes and failure mechanisms

The identification of DFBD blockage modes is helpful in predicting the river blocking situation, which helps us to better understand the formation

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

mechanisms of DFBDs, and further provides scientific basis for the prevention and management (Chen, 2019; Dang et al., 2009; Yu et al., 2022).

According to the statistics of this dataset, the blockage modes of 278 DFBDs are recorded. In these cases, 194 dams (69.8 %) exhibit complete blockage mode, 78 dams (28.1 %) exhibit partial blockage mode, and only 6 dams (2.2 %) exhibit submerged dam blockage mode (Fig. 6(b)). Figure. 6(b) indicates that complete blockage is the most common mode of river blockage caused by debris flows. It should be noted that DFBDs with submerged dam blockage modes have a high concealment and are not easily detected by direct observation. Therefore, there may be cases that have not been reported, suggesting that the actual number of submerged dams may be underestimated.

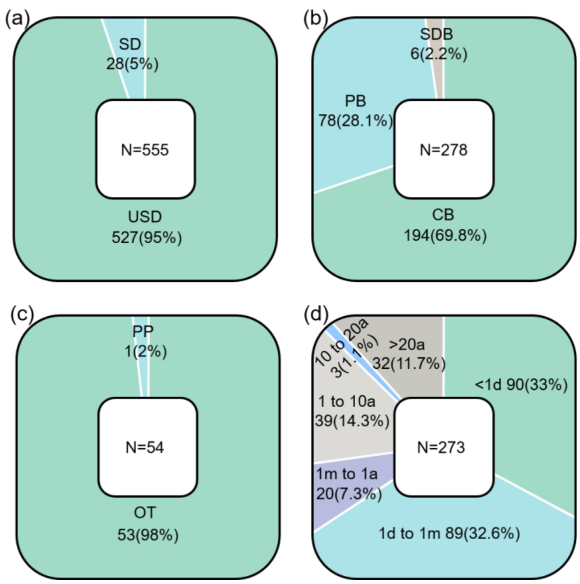

Figure. 6 The stability, blockage modes, failure mechanisms, and longevity of DFBDs. (a) Stability; (b) blockage modes; (c) failure mechanisms; (d) longevity. Notation: In Fig.6(a), SD means stable dam, and USD means unstable dam; in Fig. 6(b), CB means complete blockage, PB means partial blockage, and SDB means submerged dam blockage; in Fig. 6(c), OT means overtopping, and PP means piping.

We have compiled the failure mechanisms of 54 DFBDs, among which the overtopping (OT) accounts for an overwhelming 98 % (Fig. 6(c)). There is only one case for the piping (PP). The fluidity of debris flows inherently limits the

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

height of DFBDs but increases their widths, leading to DFBDs with smaller upstream and downstream slopes (Ruan et al., 2021). Additionally, the DFBDs are almost completely saturated, and the soil of the dams has strong water retention, especially more pronounced in the cases of barrier dams formed by viscous debris flows. Therefore, there are no cases of the SF in our dataset. On the other hand, before the debris flow merges into the main river, the solid-liquid phase materials inside it have been fully mixed after long-distance transportation, and the content of fine particles is high, making no obvious connected pores and seepage channels inside the DFBD. Even if there is seepage, due to the long seepage channel (the dam width is big), it is difficult to form a complete seepage channel in a short time. So, the probability of PP in DFBDs is relatively low.

3.4 Stability and longevity

Out of 555 DFBDs, only 28 dams are still existed, indicating that the stability of DFBDs is relatively poor (Fig. 6(a)). Moreover, all the existing dams exhibit complete blockage modes, suggesting that the dams that exhibited partial blockage modes and submerged dam blockage modes are unstable.

As illustrated in Fig.6(d), the DFBDs tend to fail within a brief period. 33 % of the dams failed within one day, and 32.6 % of the dams failed within one day to a month. About 73 % of DFBDs have a longevity of less than a year. Due to the generally lower height of DFBDs (Ruan et al., 2021), the reservoir behind the dam is more likely to reach full capacity quickly. In addition, the rheology of the DFBDs is high, and its overall structural integrity is poor (Iverson, 1997). The surface of the dam has a weaker ability to resist overflow erosion, and once the water reaches the top of the dam, it can quickly erode the materials of the DFBDs. Consequently, the longevity of DFBDs is relatively short.

DFBDs are characterized by rapid outbursts and require close monitoring and concern. The formation of DFBDs and subsequent rapid

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

failure can trigger a series of secondary disasters, which often causes significant economic losses and casualties. According to our statistics, at least 5065 deaths have been attributed to the 41 DFBD events recorded in this dataset. However, events with known casualties represent only 7 % of the total DFBD events. This suggests that the actual global death toll could be significantly higher than currently estimated, possibly exceeding our imagination.

3.5 The phenomenon of repeatedly river blockage

It can be found that it is a common phenomenon for debris flows to occur multiple times in the same channel and to form multiple DFBDs. For example, from 2017 to 2018, at least four debris flows occurred in the Sedongpu basin, upstream of the Yarlung Tsangpo River in eastern Tibet, China, which repeatedly blocked the Yarlung Tsangpo River (Huai et al., 2022; Tong et al, 2018; Zhang et al., 2022) (Fig. 7). On December 21, 2017, a glacial debris flow erupted in the Sedongpu Valley and blocked the Yarlung Tsangpo River (Fig. 7(d)), and the DFBD breached three days later. On July 26, 2018, the Yarlung Tsangpo River also experienced a temporary blockage, and the resulting barrier lake was not large, posing no serious threat (Fig. 7(c)). However, on October 17, 2018, a disaster chain composed of an ice-rock avalanche and glacial debris flow formed a DFBD (Figs. 7(a) and 7(b)), which impounded a massive barrier lake with a volume of approximately 0.605×10^9 m^3, and the maximum water depth in front of the dam was about 79.43 m (Jin, 2019). Three days later, the DFBD breached, but there was still a narrow spillway, and the possibility of re-blockage remained. Therefore, a small-scale glacier debris flow on October 29, 2018, caused further blockage of the river, forming a DFBD with a volume of 0.326×10^9 m^3 and a maximum depth of approximately 0.77 m (Figs. 7(a) and 7(b)). The two DFBDs in 2018 posed a serious threat to the upstream

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

village of Gala.

According to the records in this dataset, 555 DFBDs have predominantly occurred in 426 different gullies or river channels. It is particularly noteworthy that 63 gullies have shown a high activity, experienced multiple debris flows, and formed 192 DFBDs. This phenomenon reveals an important issue: For the channels that are in an active debris flow phase, timely engineering measures should be taken to prevent repeated blockages of rivers that lead to the formation of DFBDs.

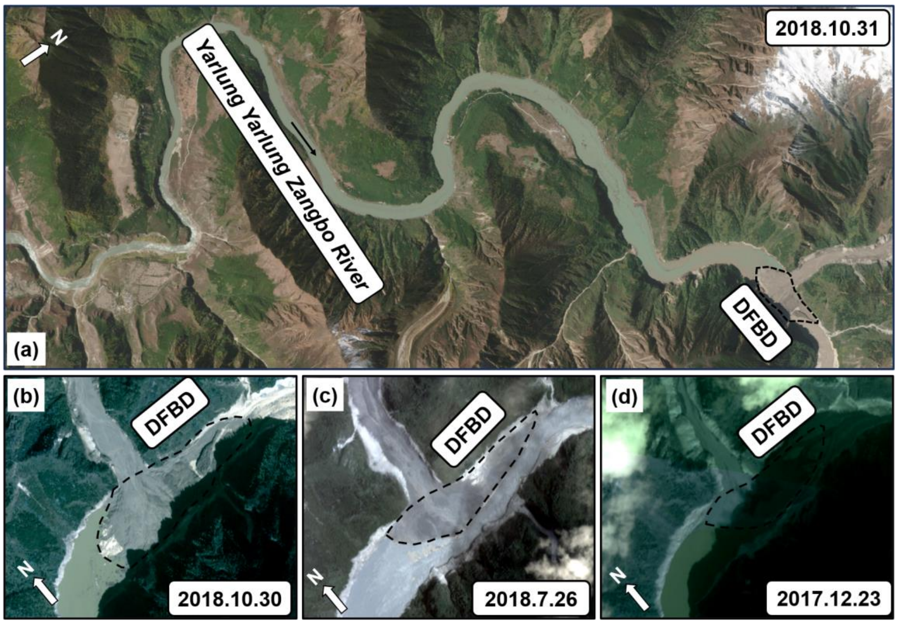

Figure. 7 Repeated DFBDs due to glacial debris flows generated in the Sedongpu catchment, upper Yarlung Tsangpo, eastern Tibet. (a) Remote sensing image after the events on October 17th and 29th, 2018 (October 31st, 2018); (b) remote sensing image after the event of October 17, 2018 (October 30, 2018); (c) remote sensing image of July 26, 2018; (d) remote sensing image after the event of December 22, 2017 (December 23, 2017). The remote sensing image (a) is sentinel-2 (https://dataspace.copernicus.eu/) and (b-d) are sourced from PlanetScope (https://www.planet.com/).

4 Discussion

4.1 Applicability of LD stability models to DFBDs

Due to the limitation of investigation conditions and the threat of dam break

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

risk, it is difficult to obtain the geological, physical, and mechanical parameters of the barrier dams. Nonetheless, dam geometric characteristics can be obtained in time through satellite optical images or UAV aerial photography. Therefore, it is of great practical significance to use geometric characteristic parameters to quickly evaluate and predict the stability of barrier dams (Table 3), which can provide timely decision support for emergency disposal.

It should be pointed out that the stability models based on morphological parameters in Table 3 ignore the category of barrier dam and are mainly focused on LD cases. However, Ruan et al. (2021) pointed out that the DFBDs are quite different from the LDs. Until now, there are no specific models or criterions for judging the stability of DFBDs. When the LD stability models in Table 2 are used to analyze the stability of DFBDs, can they better distinguish the stability? Which model is more suitable for the stability analysis of DFBDs? This work has not been studied, but it is necessary.

Because some parameters of some cases in the dataset are missing, the number of calculation cases used for each model is not exactly the same. Additionally, different research works and regions adopt varying definitions and criteria for dam stability, which has led to the development of various stability identification models. LD stability models, with their parameters and concepts, are highly prone to confusion, causing significant inconvenience in the stability assessment of barrier dams. This paper identifies dam stability based on the dam’s real-time status, and to avoid inconsistencies in the criteria and models for stability judgment, all verification models selected in Table 3 are based on the stability identification indicators of the dam’s real-time status.

Table 3 Stability prediction mode for LDs.

Model	Stability			Reference
	Stable	Uncertain	Unstable	
$BI=\lg\left(\frac{V_d}{A_c}\right)$	>5	(4,5)	(3,4)	Canuti et al. (1900)

$DBI = \lg\left(\frac{A_c H_d}{V_d}\right)$	<2.75	(2.75,3.08)	>3.08	Ermini and Casagli (2003)
$II=\lg(\frac{V_d}{V_l})$	>0	-	<0	Casagli and Ermini (1999)
$I_e = -1.554 + 2.317 lgV_l - 2.828 lgL_d - 2.336 lgW_d$	<0	-	>0	Wu et al. (2021)
$L_s(AHWL) = -2.22\lg A_c - 3.76\lg H_d + 3.17\lg L_d + 2.85\lg W_d + 5.93$	>0	-	<0	Dong et al. (2011)
$L_s(AHV) = -4.48\lg A_c - 9.31\lg H_d + 6.61\lg V_d + 6.39 - 2.336\lg W_d$	>0	-	<0	

Notation: The definition of parameters is shown in Fig.1.

Table 4 Calculation results of LD stability models to DFBDs.

Models	Number of cases	Number of misjudged cases	Number of accurate cases	F/%	R_c/%	R/%
BI	49	6	25	12.24	53.06	51.02
DBI	50	8	36	16	74	72
II	44	11	31	25	75	70.45
I_e	47	38	9	80.85	19.15	19.15
$L_s(AHWL)$	49	12	35	24.49	75.51	71.43
$L_s(AHV)$	50	0	43	0	100	86

Referring to Zhong and Shan (2019), the calculation results of each evaluation method were compared and analyzed by using the misjudged rate F, the conservative accuracy rate R_c, and the absolute accuracy rate R, respectively. Among them, the misjudgment rate F refers to the probability that the DFBD is actually unstable, but the calculation result is stable. In practical application, the model with a low misjudgment rate should be selected as far as possible; the absolute accuracy rate R refers to the probability that the actual status of the dam is completely consistent with the calculated result; the conservative accuracy R_c refers to the probability that the actual status of the dam is stable and the calculated result is unstable, and the absolute accuracy R is added to the result. The calculation results are shown in Fig. 8 and Table 4.

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

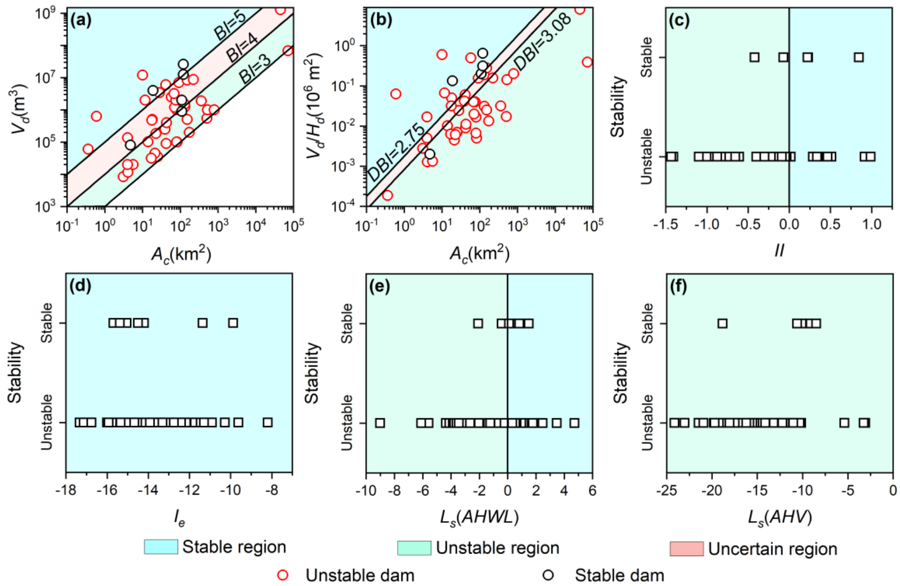

Figure.8 Stability calculation results. (a) Results of *BI* model; (b) results of *DBI* model; (c) results of *II* model; (d) results of I_e model; (e) results of L_s (*AHWL*) model; (f) results of L_s (*AHV*) model.

According to Fig. 8 and Table 4, it can be seen that the absolute accuracy rate *R* of the calculation results of the *BI* model and the I_e model are relatively low (*BI*: 51.02 %; I_e: 19.15 %), indicating that the *BI* model and the I_e model are not suitable for determining the stability of DFBDs. The absolute accuracy rate *R* and conservative accuracy rate R_c of the calculation results of the *DBI*, *II*, and L_s (*AHWL*) models are good (>70 %). Among the three models, the *DBI* model has a lower misjudgment rate *F* (only 16 %), indicating that compared to these three models, the *DBI* model has a higher applicability to DFBDs. L_s (*AHV*) model has the highest absolute accuracy rate *R*, conservative accuracy rate R_c, and the lowest misjudgment rate *F*. Considering all the above, it is recommended to prioritize the use of the L_s (*AHV*) model, followed by the *DBI* model, in evaluating the stability of DFBDs. It is not recommended to use the *BI*, *II*, and I_e models.

4.2 Applicability of LD peak discharge models to DFBDs

As an important parameter in the dam breach, the peak discharge is a prerequisite for the risk assessment of barrier dams and the simulation of downstream flood routing, which directly determines downstream disaster (Stuart-Smith et al., 2021; Zhong et al., 2021). Therefore, it is particularly important to accurately and quickly predict the outburst flow peak discharge after the formation of the barrier dams (Bazai et al., 2021; Dubey and Goyal, 2020; Vilca et al., 2021).

The peak discharge depends on the failure mechanism and the characteristics of the dam itself (Costa and Schuster, 1988; Latrubesse et al., 2020), such as the geotechnical properties of dams (Pisaniello et al., 2015; Schuster, 2000). However, it is difficult to obtain the complete parameters required for the calculation of the dam break dynamic models in a short time. The empirical models based on historical statistical cases have been widely used. Table 5 lists the empirical models used to predict the peak discharge of LD breaches around the world. It should be noted that the empirical models in Table 5 are based on some barrier dam cases, most of which are LDs. These empirical models do not distinguish the type of barrier dam and have good applicability to LDs. There are no peak discharge models for DFBDs.

Table 5. Empirical models used for LD peak discharge prediction

NO. Model	Model	Source
M1	$Q_p=0.763\times(H_w\times V_w)^{0.42}$	Costa (1985)
M2	$Q_p=1.122\times V_l^{0.57}$	Costa (1985)
M3	$Q_p=672\times V_l^{0.56}$	Costa (1985)
M4	$Q_p=2.634\times(V_l\times H_d)^{0.44}$	Costa (1988)
M5	$Q_p=0.0158\times P_e^{0.41}$	Costa and Schuster (1988)
M6	$Q_p=1.6\times V_l^{0.46}$	Walder and O'Connor (1997)
M7	$Q_p=6.7\times H_w^{1.73}$	Walder and O'Connor (1997)
M8	$Q_p=0.6971\times H_d^{1.5}\times V_l^{0.25}$	Hakimzadeh et al. (2014)
M9	$Q_p=0.54\times(V_l\text{-}H_d)^{0.5}$	Hagen (1982)
M10	$Q_p=13.4\times H_d^{1.89}$	Singh and Snorrason (1984)
M11	$Q_p=1.776\times V_l^{0.47}$	Singh and Snorrason (1984)
M12	$Q_p=0.607\times V_w^{0.295}\times H_w^{2.24}$	Froehlich (1995)
M13	$Q_p=0.4\times g^{0.5}\times(H_w+0.3)^{2.5}$	Kirkpatrick (1977)
M14	$Q_p=16.6\times H_w^{1.85}$	SCS (1981)
M15	$Q_p=19.1\times H_w^{1.85}$	USBR (1988)
M16	$Q_p=48\times H_w^{1.85}$	USBR (1988)
M17	$Q_p=3.85\times(H_w\times V_w)^{0.41}$	MacDonald and Langridge-Monopolis

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

		(1984)
M18	Q_p=0.72×$V_w^{0.53}$	Evans (1986)
M19	Q_p=0.0443×$g^{0.5}$×$V_w^{0.365}$×$H_w^{1.405}$	Webby (1996)
M20	Q_p=0.0068×$g^{0.5}$×$V_w^{0.543}$×$H_w^{0.871}$	Hooshyaripor et al. (2014)
M21	Q_p=0.0166×$(g×V_w)^{0.5}$×H_w	Azimi et al. (2015)

Note: H_w is depth of the breach (m); V_w is the released water volume (m^3); V_l is the volume of the barrier lake (m^3); H_d is dam height (m); P_e is the potential energy of water body; g is acceleration of gravity (m s^{-2})

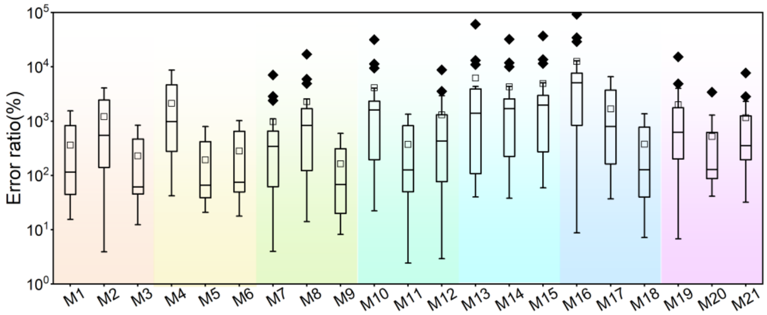

Figure.9 Error ratio (ER) of peak discharge calculated from the different models in Table 5, where ER = |Pv−Av|/P, Pv is the predicted value, Av is the actual value.

Figure 9 shows that 21 LD peak discharge models exhibit poor applicability to DFBDs, and the calculated values are consistently higher than the actual DFBD peak discharges. This may be due to significant differences in the geometry and material composition between the LDs and DFBDs (Ruan et al., 2021). Establishing a peak discharge model suitable for DFBDs is a key issue to be solved in the future. This dataset can provide rich cases and basic data to help solve this problem.

4.3 Comparison with barrier dam datasets

Some studies have established datasets of barrier dams through the collation and reconstruction of historical events. These datasets contain a large number of barrier dams, for example, Tacconi Stefanelli et al., (2016), which summarized 300 Italian barrier dams during field investigations, through air photo interpretation, and by estimating using historical and bibliographic information. Schuster and Costa (1986) established the first dataset containing 187 barrier dams worldwide by

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

reviewing literature from various regions. Fan et al. (2020) compiled a comprehensive dataset encompassing 1,887 dams, achieved by integrating various datasets.

Compared with other dam datasets, our dataset only includes 555 DFBDs, and the number of dams is not dominant. However, other datasets mainly focused on LDs, with less attention and collection on DFBDs. Our dataset is highly targeted, only focusing on DFBDs. Additionally, some datasets are obtained by summarizing other datasets, while our dataset places greater emphasis on the review and validation of raw data, rather than being a simple summary of other datasets. The DFBDs in this dataset are mostly derived from case studies scattered in various regions (Dang et al., 2009; Wei et al., 2018; Yin et al., 2016), which requires us to review extensively from original literature rather than merely superficial dataset compilation, thus avoiding the errors that might arise from a simple dataset aggregation. For example, during the data collection process, we identified a common issue where some individual case study documents confused the concepts of dam length and width (for instance, Tian et al., 2023). After correcting these errors, we included the correct data for dam length and width in the dataset. Furthermore, this paper integrates data from different sources to provide a comprehensive perspective, precisely describing the characteristics of DFBDs with 36 attributes. This dataset is the first of its kind dedicated to DFBDs.

The relationship between the shape and size of a barrier dam and the size of the valley it blocks is one of the most widely accepted classifications of barrier dam blockage modes, proposed by Costa and Schuster (1988). But some scholars have conducted a geomorphological classification of DFBD blockage modes based on hydrodynamics, dam size, and the width of the main valley, identifying three modes: submerged dam blockage, partial blockage, and complete blockage (Fig.1). This classification criterion is more in line with the

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

characteristics of DFBDs (Chen et al., 2019; Dang et al., 2009; Yu et al., 2022; Zou et al., 2020). The classification criteria of Costa and Schuster (1988) may be more applicable to LDs. However, other datasets still use Costa and Schuster (1988) classification criteria to categorize DFBD blockage modes, leading to confusion in terminology and inconsistency in criteria in subsequent researches. Therefore, this dataset has re-identified the blocking patterns of DFBDs.

The stability of a barrier dam is a dynamically changing process, and some scholars have defined the stability of a barrier dam from different perspectives: Korup (2004) defined it from the perspective of the dam's longevity, considering a barrier dam to be stable if the barrier lake exists for more than 10 years. Liao et al. (2022) and Xu (2020) believed that if no breach occurs within one year, it can be regarded as stable. However, some scholars have defined the stability of a barrier dam from the perspective of the dam's real-time condition, considering it is an instantaneous definition. When specifically analyzing a barrier dam, if the barrier lake still exists or has been filled due to the accumulation of gravel and sediment, it can be considered stable (Casagli and Ermini, 1999; Ermini and Casagli, 2003; Tacconi et al., 2016). It is evident that there is considerable divergence in the understanding of the stability of barrier dams. This divergence not only leads to confusion that different stability criteria being applied to different dams within the same datasets, but it also poses significant challenges to the research on the stability of barrier dams. Based on real-time status to assess the stability of the dams, it is possible to differentiate between failed and not failed barrier dams. Therefore, our dataset judges the stability based on the real-time status of the DFBDs, re-evaluates, and compiles the stability of all DFBDs.

4.4 Limitations in this work

While this dataset offers valuable data, it acknowledges certain limitations

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

in specific aspects. Firstly, the dataset contains some ancient events, and the authenticity of the historical records may be difficult to review fully, especially when it comes to details such as the geometric characteristics of DFBDs. Due to various limitations, some attribute information of the DFBDs is still lacking in completeness, such as the data on the failure characteristics and debris flow characteristics. In addition, we must honestly admit that this dataset does not cover all DFBD events. In the process of data collection, it is inevitable that some literature or reports might be missing, and some unreported events are not included. At present, the dataset is only an initial attempt, and although there are still shortcomings, it is already a relatively comprehensive and well-documented dataset of DFBDs.

In future work, we plan to refine this dataset from two perspectives. First, we plan to interpret the geometric characteristics of the DFBDs and lakes using remote sensing imagery. Second, we aim to uncover unreported DFBDs through field investigations. Additionally, we look forward to and welcome active participation from experts and contributors in various fields to jointly promote the continuous improvement and expansion of the dataset through interdisciplinary collaboration and the integration of multi-source data.

5 Data availability

The data can be freely downloaded via Zenodo at https://doi.org/10.5281/zenodo.13382846 (Cheng et al., 2024).

6 Conclusion

In this study, we meticulously reviewed 2519 high-caliber literature and media reports, successfully identifying 555 global DFBD events spanning from 1800 to 2023. This effort culminated in the creation of the inaugural DFBD dataset, marking a significant advancement in the field. Our dataset described the characteristics of DFBDs using five categories and 36 attributes, including

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

basic information (latitude and longitude, etc.), debris flow characteristics (debris flow velocity, debris flow peak discharge, etc.), dam characteristics (dam height, dam volume, etc.), lake characteristics (lake area, lake volume, etc.), and failure characteristics (peak discharge, loss of life, etc.). We not only conducted strict review and verification of these information using Google Earth, but also developed a method to resolve conflicts between information from different data sources. Considering the current lack of unified standards for distinguishing river blocking modes and the confusion surrounding stability concepts, this dataset reassessed and reintegrated river blocking modes and the stability of DFBDs. The results show that since the 1960s, the number of DFBDs has increased rapidly, which may be related to global climate degradation. The most common blockage mode for DFBDs is complete blockage (69.8 %), and the most common failure mechanism is overtopping (98 %). Moreover, DFBDs tend to have relatively poor stability, with about 73 % of DFBDs failing within one year after formation. The phenomenon of repeated river blocking is very common, with about 15 % of rivers experiencing multiple debris flows, leading to river blockage and the formation of 192 DFBDs, accounting for 35 % of the total number of DFBDs. Based on the data included in this dataset, the applicability of LD stability models and peak discharge models to DFBDs is discussed. Discussion indicated that the *Ls* (*AHV*) model and *DBI* model perform well in the stability assessment of DFBDs. However, the peak discharge models of LDs are not suitable for DFBDs.

Although this dataset does not have an obvious advantage in the number of cases, our dataset is the first of its kind dedicated to DFBDs. We place special emphasis on the unification of terminology and concepts, as well as the review of raw data, to ensure data consistency and accuracy. We believe that this dataset can provide a rich set of foundational data for researches related to debris flow river blocking, and enhance understanding of DFBDs. Of course, there are still some limitations that need to be improved. We will continue to improve and update this dataset in future work.

Author contributions

Conceptualization, Kaiheng Hu; Methodology, Kaiheng Hu and Haiguang Cheng; Validation, Xiaopeng Zhang, Hao Li, and Qiyuan Zhang; Formal Analysis, Kaiheng Hu, Shuang Liu, and Haiguang Cheng; Data Curation, Kaiheng Hu and Haiguang Cheng; Writing-Original Draft Preparation, Haiguang Cheng; Writing-Review & Editing, Kaiheng Hu, Shuang Liu, Haiguang, Cheng, Manish Raj Gouli, Pu Li, Lan Ning, Anna Yang; Visualization, Haiguang Cheng and Peng Zhao; Supervision, Kaiheng Hu; Funding Acquisition, Kaiheng Hu.

Competing interests

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

Disclaimer

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

Acknowledgements

This work was jointly supported by and Key R&D Program of Tibet Autonomous Region (XZ202301ZY0039G), the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0902), the Science and Technology Research Program of Institute of Mountain Hazards and

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

Environment, Chinese Academy of Sciences (IMHE-ZDRW-01), Sichuan Science and Technology Program (2024NSFSC0072).

Reference

Amante, C., Eakins, B. W.: ETOPO1 global relief model converted to PanMap layer format, NOAA-National geophysical data center, boulder, CO, USA, pp. 10, 2009.

Azimi, R., Vatankhah, A. R., and Kouchakzadeh, S.: Predicting peak discharge from breached embankment dams, E-Proc., 36th IAHR world congress, Hague, Netherlands, 2015.

Bazai, N. A., Cui, P., Carling, P. A., Wang, H., Hassan, J., Liu, D., Zhang, G., and Jin, W.: Increasing glacial lake outburst flood hazard in response to surge glaciers in the Karakoram, Earth-Science Reviews, 212, 103432, 2021.

Canuti, P., Casagli, N., and Ermini, L.: Inventory of landslide dams in the Northern Apennine as a model for induced flood hazard forecasting//Managing Hydro-geological Disasters in a Vulnerable Environment for Sustainable Development, Perugia: CNR–GNDCI Publ, 189-202, 1900.

Casagli, N. and Ermini, L.: Geomorphic analysis of landslide dams in the Northern Apennine, Transactions of the Japanese Geomorphological Union, 20(3), 219-249, 1999.

Casagli, N., Ermini, L., and Rosati, G.: Determining grain size distribution of the material composing landslide dams in the Northern Apennines: sampling and processing methods, Engineering geology, 69(1-2), 83-97, 10.1016/S0013-7952(02)00249-1, 2003.

Chai, H. J., Liu, H. C., and Zhang, Z. Y.: The catalog of Chinese landslide dam events, J. Geol. Hazards Environ. Preserv., 6 (4), 1-9, 1995.

Chen, H., Ruan, H., Chen, J., Li, X., and Yu, Y.: review of investigations on hazard chains triggered by river blocking debris flows and dam break floods, Front. Earth Sci, 10:830044, 10.3389/feart.2022.830044, 2022.

Chen, K. T., Chen, X. Q., Niu, Z. P., and Guo, X. J.: Early identification of river blocking induced by tributary debris flow based on dimensionless volume index, Landslides, 16, 2335-2352, 2019.

Cheng, H. G., Hu, K. H., Liu, S., Zhang, X. P., Li, H., Zhang, Q. Y., Ning, L., Manish, R. G., Li, P., Yang, A. N.: A worldwide event-based debris-flow barrier dam dataset from 1800 to the 2023, [Data set]. Zenodo, https://doi.org/10.5281/zenodo.13382846, 2024.

Cheng, Z. L., Dang, C., Liu, J. J., and Gong, Y. W.: Experiments of debris flow damming in Southeast Tibet. Earth Sci. Front, 14 (6), 181-185. https://doi.org/10.1016/S1872-5791(08)60010-X, 2007a.

Cheng, Z. L., Geng, X. Y., Dang, C., and Liu, J. J.: Modeling experiment of break of debris-flow dam, Wuhan Univ. J. Nat. Sci, 12 (4), 588-594, 2007b.

Chong, Y., Chen, G., Meng, X., Yang, Y., Shi, W., Bian, S., Zhang, Y., and Yue, D.: Quantitative analysis of artificial dam failure effects on debris flows - A case study of the Zhouqu '8.8' debris flow in northwestern China, The Science of the total environment, 792, 148439, 2021.

Costa, J. E. and Schuster, R. L.: The formation and failure of natural dams, Geological Society of America Bulletin, 100(7), 1054-1068, 1988.

Costa, J. E., and Schuster, R. L.: Documented historical landslide dams from around the world. U.S., Geological Survey Open-File Report, 91-239, 486, 1991.

Costa, J. E.: Flood from dam failures, U. S. geological survey open-file report, 85-560, 1985.

Coviello, V., Arattano, M., Comiti, F., Macconi, P., and Marchi, L.: Seismic characterization of debris flows: Insights into energy radiation and implications for warning, Journal of Geophysical Research: Earth Surface,124, 1440-1463, 2019.

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

Cui, P., Dang, C., Zhuang, J. Q., You, Y., Chen, X. Q., and Scott, K. M.: Landslide-dammed lake at Tangjiashan, Sichuan province, China (triggered by the Wenchuan Earthquake, May 12, 2008): risk assessment, mitigation strategy, and lessons learned, Environmental Earth Sciences, 2010.

Cui, P., Han, Y., Chao, D., and Chen, X.: Formation and treatment of landslide dams emplaced during the 2008 Wenchuan earthquake, Sichuan, China, in: Evans, S., Hermanns, R., Strom, A., and Scarascia-Mugnozza, G.: Natural and artificial rockslide dams. lecture notes in Earth Sciences, 133, 2011.

Cui, P., Lei, Y., Hu, K., Zhou, G. G. D., Zhu, X., and Chen, H.: Amplification mechanism and hazard analysis for Zhouqu giant debris flow, International Journal of Erosion Control Engineering, 9(3), 71-79, DOI:10.13101/ijece.9.71, 2016.

Cui, P., Zhu, Y. Y., Han, Y. S., Chen, Q., and Zhuang, J.: The 12 May Wenchuan earthquake-induced landslide lakes： distribution and preliminary risk evaluation, Landslides, 6(3), 209-223, 2009.

Dang, C., Cui, P., and Cheng, Z. L.: The formation and failure of debris flow-dams, background, key factors and model tests: case studies from China, Environmental Geology, 57(8), 1901-1910, DOI:10.1007/s00254-008-1479-6, 2009.

Dong, J. J, Lai P. J., Chang, C. P., Yang, S. H., Yeh, K. C., Liao, J. J., and Pan, Y. W.: Deriving land-slide dam geometry from remote sensing images for the rapid assessment of critical parameters related to dam-breach hazards, Landslides, 11(1), 93-105, 2014.

Dong, J. J., Tung, Y. H., Chen, C. C., Liao, J. J., and Pan, Y. W.: Logistic regression model for predicting the failure probability of a landslide dam, Engineering Geology, 117(1-2), 52-61, 2011.

Dubey, S., and Goyal, M. K.: Glacial lake outburst flood hazard, downstream impact, and risk over the Indian Himalayas, Water Resources Research, 56, 2020.

Ermini, L., and Casagli, N.: Prediction of the behaviour of landslide dams using a geomorphological dimensionless index, Earth Surface Processes & Landforms, 28(1), 31-47, 2003.

Evans, S. G.: The maximum discharge of outburst floods caused by the breaching of man-made and natural dams, Canadian Geotechnical Journal, 23, 1986.

Fan, X. M., Dufresne, A., Subramanian, S. S., et al.: The formation and impact of landslide dams-State of the art, Earth-Science Reviews, 203, 103116, 2020.

Fan, X. M., van Westen, C. J., Xu, Q., Gorum, T., and Dai, F.: Analysis of landslide dams induced by the 2008 Wenchuan earthquake, Asian Earth Sci, 57, 25-37, 2012.

Fan, X., Xu, Q., van Westen, C. J., Huang, R., and Tang, R.: Characteristics and classification of landslide dams associated with the 2008 Wenchuan earthquake, Geoenvironmental Disasters, 4(1), 12, 2017.

Froehlich, D. C.: Peak outflow from breached embankment dam, Journal of Water Resources Planning and management, ASCE, 121(1), 1995.

Gao, J. L.: A preliminary study of blocking river of the Gaojiagou debris flow in Wenchuan, Sichuan, Acta Geologica Sichuan, 36(01), 111-113+117, 2016 (in Chinese).

Gorum, T., Fan, X., Westen, C.J., Huang, R., Xu, Q., Tang, C., and Wang, G.: Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake, Geomorphology, 133, 152-167, 2010.

Guo, X. J., Cui, P., Xiang, L. Z., Zhou X. J., and Yang W.: Research on the debris flow hazards in Gaojia Gully and Shenxi Gully in 2011, JOURNAL OF CATASTROPHOLOGY, 27(03), 81-85, 2012(in Chinese).

Hagen, V. K.: Re-evaluation of design and dam safety, 14th International

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

Commission on Large Dams Congress. Rio de Janerio, 1982.

Hakimzadeh, H., Nourani, V., and Amini, A. B.: Genetic programming simulation of dam breach hydrograph and peak outflow discharge, Journal of Hydrologic Engineering, 19(4), 757-768, 2014.

Hooshyaripor, F., Tahershamsi, A., and Golian, S.: Application of copula method and neural networks for predicting peak outflow from breached embankments, J. Hydro-Environ. Res., 8(3), 292-303, 2014.

Hu, K. H., Ge, Y., Cui, P., Guo, X. J., and Yang, W.: Preliminary analysis of extra-large-scale debris flow disaster in Zhouqu County of Gansu Province, Mountain Res. 28, 628–634, 2010(In Chinese).

Hu, K. H., Wei, F. Q., and Li Y.: Real-time measurement and preliminary analysis of debris-flow impact force at Jiangjia Ravine, China, Earth Surface Processes and Landforms, 36(9), 1268-1278, 2011.

Hungr, O., Leroueil, S., and Picarelli, L.: The Varnes classification of landslide types, an update. Landslides, 11 (2), 167-194, 2014.

Iverson, R. M.: The physics of debris flows, Reviews of Geophysics, 35(3), 245-296, 1997.

Jin, X.: Review and reflections on emergency response countermeasures for barrier lakes in Jinsha river and Yarlung Zangbo river, Yangtze River, 50(3), 5-9, 2019 (in Chinese).

Kirkpatrick, G. W.: Guidelines for evaluating spillway capacity, Water Power Dam Constr., 29(8), 29–33, 1977.

Korup O.: Geomorphometric characteristics of New Zealand landslide dams, Engineering Geology, 73(1), 13-35, 2004.

Latrubesse, E. M., Park, E., Sieh, K., Dang, T. D., Lin, Y. N., and Yun, S.: Dam failure and a catastrophic flood in the Mekong basin (Bolaven Plateau), southern Laos, 2018. Geomorphology, 362, 107221, 2020.

Liao, H. M., Yang, X. G., Lu, G. D., Tao, J., and Zhou, J. W.: A geotechnical index for landslide dam stability assessment, Geomatics, Natural Hazards and Risk, 13:1, 854-876, 2022.

Liu, W., Carling, P. A., Hu, K., Wang, H., Zhou, Z., Zhou, L., Liu, D., Lai, Z., and Zhang, X.: Outburst floods in China: A review, Earth Sci. Rev. 197, 2019.

Loche, M., Alvioli, M., Marchesini, I., Bakka, H., and Lombardo, L.: Landslide susceptibility maps of Italy: Lesson learnt from dealing with multiple landslide types and the uneven spatial distribution of the national inventory, Earth Sci. Rev., 104125, 2022.

Ma, C., Chen, Y., Hu, K., Du, C., Dong, J., Lyu, L. Q.: Climate warming triggered a glacial lake outburst flood and debris flow events in an Alpine Watershed, Western Himalayas, Tibet Plateau, Bull Eng Geol Environ, 83, 201, 2024.

MacDonald, T. C., and Langridge-Monopolis, J.: Breaching characteristics of dam failures, Journal of hydraulic engineering, 110, 567-586, 1984.

Morgenstern, R., Wolter, A., Cox, S. C., et al.: The New Zealand landslide dam database, v1.0, Landslides, 21, 121-134, 2024.

Peng, M., and Zhang, L. M.: Analysis of human risks due to dam break floods-part 2:application to Tangjiashan landslide dam failure, Nat Hazards, 64(2), 1899-1923, 2012b.

Peng, M., and Zhang, L. M.: Breaching parameters of landslide dams, Landslides, 9,13-31, 2012a.

Peruccacci, S., Gariano, S. L., Melillo, M., Solimano, M., Guzzetti, F., and Brunetti, M. T.: The ITAlian rainfall-induced LandslIdes CAtalogue, an extensive and accurate spatio-temporal catalogue of rainfall-induced landslides in Italy, Earth Syst. Sci. Data, 15, 2863–2877, https://doi.org/10.5194/essd-15-2863-2023, 2023.

Pisaniello, J. D., Dam, T. T., and Tingey-Holyoak, J. L.: International small dam safety assurance policy benchmarks to avoid dam failure flood disasters in developing countries, Journal of Hydrology, 531, 1141-1153, 2015.

Ruan, H., Chen, H., Li, Y., Chen, J. G., and Li, H., B.: Study on the downcutting rate of a debris flow dam based on grain-size distribution, Geomorphology,

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

391, 107891, 2021.

Ruiz-Villanueva, V., Allen, S., Arora, M., Goel, N. K., and Stoffel, M.: Recent catastrophic landslide lake outburst floods in the Himalayan Mountain range, Progress in Physical Geography: Earth and Environment, 41(1), 3-28, 2017.

Schuster, R. L., and Costa, J. E.: Landslide Dams: Processes, Risk, And Mitigation, Proc., (Seattle, U.S.A.: Apr. 7, 1986), R.L. Schuster (Ed.), New York, U (3), pp. ISBN 0-87262-524-9, 1986.

Schuster, R. L.: Dams built on pre-existing landslides, In: GeoEng 2000-Geotechnical and Geological Engineering: International Society for Rock Mechanics and Rock Engineering, Melbourne, Australia. pp. 1537-1589, 2000.

SCS (Soil Conservation Service): Simplified dam-breach routing procedure, U.S. Dept. of Agriculture, Washington, DC, 1981.

Sharma, A., Sajjad, H., Roshani, and Rahaman, M. H.: A systematic review for assessing the impact of climate change on landslides: research gaps and directions for future research, Spat. Inf. Res., 32, 165-185, 2024.

Shi, Z., Zhang, G., Peng, M., Zhang, Q. Z., Zhou, Y. Y., and Zhou, M. J.: Experimental investigation on the breaching process of landslide dams with differing materials under different inflow conditions, Materials (Basel, Switzerland), 15(6), 2022.

Shroder Jr, J. F., and Weihs, B. J.: Geomorphology of the lake shewa landslide dam, badakhshan, afghanistan, using remote sensing data, Geografiska Annaler, Series A, Physical Geography 92, 469-483, 2010.

Singh, K. P. and Snorrason, A.: Sensitivity of outflow peaks and flood stages to the selection of dam breach parameters and simulation models, Journal of hydrology, 68, 295-310, 1984.

Song, Z., Fan, G., Chen, Y., and Liu, D.: Identification method of river blocking by debris flow in the middle reaches of the Dadu River, Southwest of China, Water, 15, 4301, 2023.

Stefanelli, C. T., Catani, F., and Casagli, N.: Geomorphological investigations on landslide dams, Geoenvironmental Disasters, 2(1), 21, 2015.

Stefanelli, C. T., Segoni, S., Casagli, N., and Catani, F.: Geomorphic indexing of landslide dams evolution, Engineering Geology, 208,1-10, 2016.

Stuart-Smith, R. F., Roe, G. H., Li, S. J., and Allen, M. R.: Increased outburst flood hazard from Lake Palcacocha due to human-induced glacier retreat, Nature Geoscience, 14, 85-90, 2021.

Tian, Y., Jiang, L., and Guo, J.: Numerical simulation of glacier avalanche-river blocking-outburst in Sedongpu gully of the Yarlung Zangbo River, Journal of Geology, 47(02):196-202, 2023(in Chinese).

Tiranti, D., Nicolò, G., and Gaeta, A. R.: Shallow landslides predisposing and triggering factors in developing a regional early warning system, Landslides, 16, 235-251, 2019.

Tong, L., Tu, J., Pei, L., Guo, Z., Zheng, X., Fan, J., Zhong, C., Liu, C., Wang, S., He, P. and Chen, H.: Preliminary discussion of the frequently debris flow events in Sedongpu basin at Gyalaperi peak, Yarlung Zangbo river, Journal of Engineering Geology, 26(6),1552-1561, 2018 (in Chinese).

Tong, Y. X.: Quantitative analysis for stability of landslide dams. Master thesis, National Central University, Taiwan, 2008 (in Chinese).

Vilca, O., Mergili, M., Emmer, A., Frey, H., and Huggel, C.: The 2020 glacial lake outburst flood process chain at Lake Salkantaycocha (Cordillera Vilcabamba, Peru), Landslides, 18, 2211-2223, 2021.

Walder, J., and O'Connor, J.: Methods for predicting peak discharge of floods caused by failure of natural and constructed dams, Water Resources Research, 33(10), 2337-2348, 1997.

Walsh, L.S., Martin, A.J., Ojha, T.P., and Fedenczuk, T.: Correlations of fluvial Knick zones with landslide dams, lithologic contacts, and faults in the southwestern Annapurna Range, central Nepalese Himalaya, Journal of

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

Geophysical Research, 117, 2012.

Wang, L., Chang, M., Dou, X., Ma, G., and Yang, C.Y.: Analysis of River Blocking Induced by a Debris Flow, Geofluids, 1-8, 2017.

Wang, L., Chang, M., Le, J., and Zhang, N.: Two multi-temporal datasets to track debris flow after the 2008 Wenchuan earthquake, Sci Data, 9, 525, 2022.

Webby, M. G.: Discussion of 'Peak outflow from breached embankment dam.' by D. C. Froehlich, J. Water Resource. Plan. Manage., 10.1061/(ASCE)0733-9496, 122, 4(316), 316-317, 1996.

Wei, R., Zeng, Q., Davies, T., and Yin, Q.: Geohazard cascade and mechanism of large debris flows in Tianmo gully, SE Tibetan Plateau and implications to hazard monitoring, Engineering Geology, 233, 172-182, 2018.

Wu, H., Nian T. k., and Shan Z. G.: Investigation of landslide dam life span using prediction models based on multiple machine learning algorithms, Geomatics, Natural Hazards and Risk, 14:1, 2023.

Wu, H., Shan, Z. G., Ni, W. D., et al.: Study on the application of rapid evaluation model of landslide dam stability//proceedings of 2021 National Engineering Geology Annual Conference. Beijing: Science Press, 2021.

Wu, H., Trigg, M. A., Murphy, W., Fuentes, R.: A new global landslide dam database (RAGLAD) and analysis utilizing auxiliary global fluvial datasets, Landslides, 19, 555-572, 2022.

Xu, F.: A rapid evaluation model of the stability of landslide dam, Journal of Natural Disasters, 29(02), 54-63, 2020 (in Chinese).

Xu, Q., Fan, X. M., Huang, R. Q., Westen, C. V.: Landslide dams triggered by the Wenchuan Earthquake, Sichuan Province, south west China, Bulletin of Engineering Geology and the Environment, 68(3), 373-386, 2009.

Yan, R.: Secondary disaster and environmental effect of landslides and collapsed dams in the upper reaches of Minjiang River, Master thesis, Sichuan University, 2006.

Yin, Y., Cheng, Y., Liang, J., and Wang, W.: Heavy-rainfall-induced catastrophic rockslide-debris flow at Sanxicun, Dujiangyan, after the Wenchuan Ms 8.0 earthquake, Landslides, 13 (1), 9-23, 2016.

Yu, B., Yang, C., and Yu, M.: Experimental study on the critical condition of river blockage by a viscous debris flow, Catena, 213, 106198, 2022.

Yu, G. A., Yao, W., Huang, H. Q., and Liu, Z.: Debris flows originating in the mountain cryosphere under a changing climate: A review. Progress in Physical Geography: Earth and Environment, 45(3), 339-374, 2021.

Zhang, Q. Y., Hu, K. H., Liu, S., Liu, W. M.: Controls on the regional distribution of landslide dams and implications for fluvial landform evolution in Bhutan and its surrounding area, Journal of Mountain Science, 20(8), 2107-2132, 2023.

Zhang, X. P., Hu, K. H., Liu, S., Nie Y., and Han Y.: Comprehensive interpretation of the Sedongpu glacier-related mass flows in the eastern Himalayan syntaxis, Journal of Mountain Science, 19(9), 1672-6316, 2022.

Zhao, B., Su, L. J., Xu Q., Li W., Xu C., and Wang, Y. S.: A review of recent earthquake-induced landslides on the Tibetan Plateau, Earth-Science Reviews, 244, 104534, 2023.

Zhao, Y., Meng, X., Qi, T., Qing, F., Xiong, M., Li, Y., Guo, P., and Chen, G.: AI-based identification of low-frequency debris flow catchments in the Bailong River basin, China, Geomorphology, 359, 107125, 2020.

Zhong, Q. M., and Shan, Y. B.: Comparison of rapid evaluation methods for barrier dam's stability, Yangtze River, 50(04), 20-24+64, 2019 (in Chinese).

Zhong, Q., Wang, L., Chen, S., Chen, Z. Y., Shan, Y. B., Zhang, Q., Ren, Q., Mei, S. Y., Jiang, J. D., Hu, L., and Liu, J. X.: Breaches of embankment and landslide dams - State of the art review, Earth-Science Reviews, 12, 103597, 2021.

Zhou, G. G. D., Zhou, M. J., Shrestha, M. S., Song, D. R., Choi, C. E., Cui, K. F. E., Peng, M., Shi, Z. M., Zhu, X. H., and Chen, H. Y.: Experimental

https://doi.org/10.5194/essd-2024-382
Preprint. Discussion started: 30 September 2024

Open Access Earth System Science Data Discussions

investigation on the longitudinal evolution of landslide dam breaching and outburst floods, Geomorphology, 334, 29-43, https://doi.org/10.1016/j.geomorph.2019.02.035, 2019.

Zou, Q., Cui, P., Jiang, H., Wang, J., Li, C., and Zhou, B.: Analysis of regional river blocking by debris flows in response to climate change, Science of The Total Environment, 741, 140262, https://doi.org/10.1016/j.scitotenv.2020.140262, 2020.