Risk assessment of dammed lakes in China based on Bayesian network

Zhenhan Du
Hohai University
Qiming Zhong (qmzhong@nhri.cn)
Nanjing Hydraulic Research Institute https://orcid.org/0000-0002-4732-5174
Shengyao Mei
Nanjing Hydraulic Research Institute
Yibo Shan
Nanjing Hydraulic Research Institute

Research Article

Keywords: Dammed lake, Risk assessment, Failure probability, Flood loss, Risk level, Baige dammed lake

Posted Date: April 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1549465/v1

Abstract

Scientific risk assessment of dammed lakes is vitally important for emergency response planning. In this study, based on the evolution process of the disaster chain, the logic topology structure of dammed lake risk was developed. Then, a quantitative risk assessment model of dammed lake employing Bayesian network is developed, which includes the three modules of dammed lake hazard evaluation, outburst flood routing simulation, and loss assessment. In the proposed model, the network nodes of each module were quantified through statistical data, empirical model, logical inference, and Monte Carlo method. The failure probability of a dammed lake, and the losses of life and property because of its breaching were calculated. This can be multiplied to assess the risk a dammed lake imposes after the uniformization of each loss type. Based on the socio-economic development and longevity statistics of dammed lakes, a risk level classification method for dammed lakes is proposed. The Baige dammed lake, which emerged in China in 2018, was chosen for a case study and risk assessment was conducted. The calculated results showed that the comprehensive risk value of Baige dammed lake is 0.7339 under the condition without manual intervention, identifying it as the extra-high level according to the classification. These results are in accordance with the actual condition, which corroborates the reasonability of the proposed model. This model can quickly and quantitatively evaluate the overall risk of a dammed lake and provide a reference for decision-making in a rapid emergency response scenario.

1 Introduction

Triggered by precipitation events or earthquakes, dammed lakes often form in mountainous areas where water is stored after landslides have blocked valleys and rivers (Costa and Schuster, 1988; Zhong et al., 2018). If such a dammed lake breaks, it often causes serious flood disasters that pose a huge threat to the safety of people's lives and properties in downstream areas (Korup, 2002; Strom, 2010). A scientific risk assessment of these dammed lakes is conducive to an appropriate emergency response, which can help to reduce the losses of life and property.

According to the definition of risk, it can be quantified according to the expected value of consequences, which is the product of disaster frequency and expected losses (Helm, 1996). Since the 1980s, considerable research has focused on risk assessment of dammed lakes. After decades of development, great progress has been achieved in both theoretical and practical aspects (Costa and Schuster, 1988; Korup, 2002; Liu et al., 2019; Fan et al., 2019, 2020; Zhong et al., 2021). However, most of the studies focused on the formation, stability, and outburst floods of dammed lakes. The risk assessment of the dammed lake has not yet been developed into a complete system and a rigorous theoretical framework. Disaster evaluation and loss assessment are independent of each other and not highly quantified. A thorough study of risk assessment has not been conducted, particularly one that focuses on advancing disaster chain evolution because of the damming of rivers by landslides.

In this study, quantitatively risk assessment of dammed lakes is conducted based on Bayesian network, which mainly concludes four steps: (1) Hazard assessment of dammed lakes to quantify their failure probability. (2) Flood routing simulation to obtain the inundation situation of outburst flood. (3) Quantitative assessment and uniformization of losses in the inundated area. (4) Utilizing the product of failure probability and losses to assess the risk of dammed lakes based on the risk-grading standard.

2 Bayesian Network

The Bayesian network is a directed acyclic graph that represents the probability relationship between variables. It can be regarded as a combination of probability and graph theories (Pearl, 2003; Aguilera et al., 2011). The nodes in the network represent random variables and directed connections between nodes represent conditional causal relationships between random variables. Each node corresponds to a conditional probability that represents the strength of the relationship between that variable and its parent node. Nodes without a parent node are expressed informally with prior probabilities (Cai et al., 2020a). According to the theorem of the Bayesian network (Jensen, 2001), the joint probability can be calculated as:

$$P\left(x_1, \ldots, x_n\right) = \prod_{i=1}^{n} P\left(x_i | Pa\left(X_i\right)\right)$$

1

where $Pa(X_i)$ is the parent set of X_i. If X_i has no parents, the function is reduced to the unconditional probability of $P(x_i)$.

In recent years, Bayesian networks have been widely applied for engineering risk analysis (e.g., Smith, 2006; Khakzad et al., 2013; Zhu et al., 2021). To apply the Bayesian network to the risk assessment of dammed lakes, the causal relationship among the influencing factors should be primarily determined. In this study, based on the Bayesian software Netica Application 6.08 (2020), risk assessment of dammed lakes is conducted within the three modules of hazard evaluation of dammed lake, outburst flood routing simulation, and loss assessment. Figure 1 shows the network structure and its nodes.

3 Hazard Evaluation Of Dammed Lakes

As most dammed lakes failed within a short period after their formation (Shen et al., 2020), a reasonable hazard evaluation forms the prerequisite and foundation for a scientific risk assessment. Previous studies have shown that the hazard of a dammed lake is related to its morphological characteristics, material composition, and hydrodynamic conditions (Ermini and Casagli, 2003; Korup, 2004; Dong et al., 2011; Stefanelli et al., 2016; Shan et al., 2020).

3.1 Selecting network nodes

3.1.1 Morphological characteristics of landslide dams

The dam height (i.e., the vertical altitude difference from the valley floor to the lowest point on the landslide dam) determines the potential energy of the water body and the volume of the upstream dammed lake. The dam height is an important influencing factor of the stability of the landslide dam. Dam width (i.e., the base width of the landslide dam measured parallel to the main valley axis) affects the maximum hydraulic gradient. The maximum hydraulic gradient refers to the ratio of maximum water head (dam height) to dam width, which affects the permeability of the dam body itself. When upstream and downstream water levels are constant, the smaller the dam length (i.e., the crest length of the landslide dam measured perpendicular to the major valley axis), the greater its constraint by the mountains on both sides and the better its stability. The volume of the dam, which is determined by its own length, width, and height, affects its self-weight and global stability.

3.1.2 Material composition of landslide dams

The material composition of a landslide dam determines its anti-erosion ability. In general, the larger the size of particles that form the dam, the stronger its ability to resist erosion. In this study, the proportion of soil and stone is utilized to quantitatively represent the material composition. A landslide dam with soil (particle size < 60 mm) share exceeding 70% of the total weight is categorized as a soil dam, a dam with stone (particle size > 200 mm) share exceeding 70% total weight is categorized as a stone dam, and other landslide dams are categorized as mixed dams (Shan et al., 2022).

3.1.3 Hydrodynamic conditions of dammed lakes

The volume of the dammed lake determines its water storage capacity and plays a decisive role in landslide dam stability, which is generally limited by dam height and catchment area. Here, catchment area determines the upstream inflow and directly affects the impounding speed. In general, the larger the catchment area, the higher the risk.

3.2 Determination of network structure and node states

The determination of the network structure is generally based on three approaches: expert experience, machine learning, or a combination of both. However, based on the premise that there are many nodes and no appropriate algorithm, both the efficiency of machine learning and the accuracy of the established network structure are low. In this section, the expert experience method is employed to develop the network model. Dam height, dam length, dam width, dam volume, and maximum hydraulic gradient were selected to characterize the morphological characteristics of landslide dams. The lake volume and catchment area were selected to represent the hydrodynamic conditions of dammed lakes, and the material composition of the landslide dam was considered. According to the relationship among each node, the hazard topology network of dammed lakes is shown in Fig. 1, and the states and value ranges of each node are presented in Table 1 (Liao et al., 2018). Figure 1 shows that the network of hazard evaluation of dammed lake contains 9 nodes and 11 arcs.

Table 1
Node states and value ranges of the hazard network of dammed lakes

Node	State	Value range
Dam height (H_d) (m)	Short; Medium; Long; Very long	(0, 15]; (15, 30]; (30, 70]; (70, ∞)
Dam length (L_d) (m)	Low; Medium; High; Very high	(0, 160]; (160, 300]; (300, 500]; (500, ∞)
Dam width (W_d) (m)	Narrow; Medium; Wide; Very wide	(0, 300]; (300, 500]; (500, 900]; (900, ∞)
Dam volume (V_d) ($\times 10^6$ m^3)	Small; Medium; Large; Very large	(0, 2]; (2, 5]; (5, 10]; (10, ∞)
Hydraulic gradient (i)	Small; Medium; Large; Very large	(0, 0.05]; (0.05, 0.1]; (0.1, 0.2]; (0.2, ∞)
Lake volume (V_l) ($\times 10^6$ m^3)	Small; Medium; Large; Very large	(0, 1]; (1, 10]; (10, 100]; (100, ∞)
Catchment area (A_b) (km^2)	Small; Medium; Large; Very large	(0, 200]; (200, 400]; (400, 1000]; (1000, ∞)
Material composition (M_c)	Soil; Stone; Mixed	Soil > 70%; Stone > 70%; The others
Failure probability	Stable; Unstable	(0, 0.5]; (0.5, 1]

3.3 Quantifying the network

Quantifying the Bayesian network means finding the conditional probability relationship between nodes. In general, there are four sources of data for quantifying a Bayesian network: statistical data, experience-based judgment, existing physical or empirical models, and logical inference (Peng and Zhang, 2012a). The quantification of dammed lake hazard network nodes is mainly based on historical statistical data. To quantify network nodes, a database was compiled containing 313 documented dammed lake cases across the world (Schuster, 1985; Costa and Schuster, 1991; Chai et al., 1998; Tahata et al., 2002; Ermini and Casagli, 2003; O'Connor and Beebee, 2009; Nie et al., 2004; Yan, 2006; Tong, 2008; Cui et al., 2011; Dong et al., 2011; Peng and Zhang, 2012b; Shi et al., 2014; Stefanelli et al., 2015; Zhang et al., 2016; Jia et al., 2019; Zhang et al., 2019) (see Supplementary Table 1). These include 186 formed unstable cases and 127 formed stable cases. Of these documented cases, 151 occurred in Italy, 48 in Japan, 40 in China, 31 in the USA, and 43 in other countries (see Fig. 2). However, because of various objective reasons, information for certain cases is incomplete. Table 2 presents the statistics of all documented dammed lake cases in the database.

Table 2
Information statistics of all documented dammed lake cases in the database

Parameter	H_d (m)	W_d (m)	L_d (m)	M_c	V_d (×10^6 m^3)	V_l (×10^6 m^3)	A_b (km^2)
Cases	313	313	313	313	252	171	189
Loss rate	0	0	0	0	19.49%	45.37%	39.62%
Maximum	800	2700	3800	0.8	1300	16000	173484
Minimum	1	5	5	0	0	0	0.5
Mean value	42.2	320.3	604.2	0.1	15.6	152.8	2221.8
Median	25	200	400	0.1	1.7	2.7	53
Variance	60.8	327.5	640.7	0.1	89	1241.6	17768.8
Skewness	7.1	3	2.4	3.1	12.4	12.3	9.6
Kurtosis	78.1	14.8	6.9	10.7	173.5	157	90.7

In this study, 248 cases were selected from the database as training sets for data learning, where the missing rates of V_d, V_l, and A_b were 17.34%, 41.94%, and 38.71%, respectively. As the learning cases contain incomplete data cases, the expectation maximum algorithm is adopted to deal with the probability calculation in the case of missing data (Dempster et al., 1977). Here, the failure probability of the dammed lake (Y) is taken as observation variable, while the observation result can be stable or unstable, with a probability distribution of $P(Y|\theta)$. Herein, $\theta = (H_d, W_d, L_d, i, V_d, V_l, A_b, M_c)$, which represents the input parameters. The missing node data Z is an implicit variable that represents the unobserved outcome of a node, assuming that the joint probability of Y and Z is $P(Y, Z|\theta)$, and that the conditional probability distribution is $P(Z|Y, \theta)$. Firstly, the counting probability in the case is taken as the initial value of the parameter to start the iteration. $\theta^{(i)}$ is denoted as the estimation of parameter θ of the i-th iteration and the Q function can be calculated as:

$$Q\left(\theta, \theta^{(i)}\right) = E_Z\left[\log P(Y, Z|\theta)\, Y, \theta^{(i)}\right] = \sum_Z \log P(Y, Z|\theta) P\left(Z\middle|Y, \theta^{(i)}\right)$$

2

where $P(Z|Y, \theta^{(i)})$ is the conditional probability distribution of the implicit variable data Z under given observation data Y and current parameter estimation $\theta^{(i)}$. To find θ which can maximize $Q(\theta, \theta^{(i)})$, the estimated parameter $\theta^{(i+1)}$ of the $i+1$ iteration can be expressed as:

$$\theta^{(i+1)} = \underset{\theta}{\operatorname{argmax}} Q\left(\theta, \theta^{(i)}\right)$$

3

The iteration is repeated until its convergence, and the convergence condition is set as $||\theta^{(i+1)} - \theta^{(i)}|| < 0.01\%$. When the maximum likelihood Bayesian network is found, the final prior (conditional) probability distribution is obtained, and the Bayesian network model is obtained (see Fig. 3). The model can obtain the real-time failure probability of the dammed lake according to the input parameters.

To verify the accuracy of the model, the 65 remaining cases in the database were used as test set. Because the Bayesian model is a probabilistic result and the target node (the failure probability of the dammed lake) has only two states, the state with higher probability is taken as resulting output (i.e., when the probability of instability is greater than 50%, the instability is taken as resulting output). Testing the model with the test set cases showed that the output results are identical with the real results in 53 cases, with an absolute accuracy of 81.54%. Nine cases had unstable output results while the real results were all stable, hence, the conservative accuracy rate was 95.38%.

4 Outburst Flood Routing Simulation Of Dammed Lakes

Outburst flood routing simulation is an important module in the risk assessment of a dammed lake, and it is the pivot connecting hazard evaluation with loss assessment. This module can be divided into two parts: flood severity and duration. The node states and value ranges of outburst flood routing simulation of dammed lakes are shown in Table 3.

Table 3
Node states and value ranges of outburst flood routing simulations of dammed lakes

Node	State	Value range
Distance from dam site (km)	Near; Medium; Far; Very far	(0, 4.8]; (4.8, 12]; (12, 36]; (36, ∞)
Velocity (m/s)	Slow; Medium; Fast; Very fast	(0, 1]; (1,2]; (2, 4]; (4, ∞)
Water depth (m)	Shallow; Medium; Deep; Very deep	(0, 1.5]; (1.5, 3]; (3, 6]; (6, ∞)
Duration of routing (h)	Short; Medium; Long; Very long	(0, 1]; (1, 3]; (3, 9]; (9, ∞)
Flood rise time (h)	Short; Medium; Long; Very long	(0, 0.25]; (0.25, 1]; (1, 3]; (3, ∞)
Flood arrival time (h)	Short; Medium; Long; Very long	(0, 0.25]; (0.25, 1]; (1, 3]; (3, ∞)
Flood severity (m^2/s)	High; Medium; Low	(0,3]; (3, 7]; (7, ∞)

4.1 Flood severity

Flood severity is the product of flow velocity and water depth at a certain cross section, which forms an important parameter for measuring the damage to the inundation area. Clausen and Clark (1990) partitioned the flood severity, where the product of water depth and flow velocity less than or equal to 3 m^2/s is defined as the inundation zone, an area with a product greater than 3 m^2/s and less than or equal to 7 m^2/s is defined as the partial damage zone, and an area with a product greater than 7 m^2/s is defined as complete damage zone. Accordingly, in this study, flood severity is classified into three levels: high, medium, and low, respectively (see Table 3).

The nodes involved in flood severity mainly include water depth, flow velocity, distance to dam site, dam height, and lake volume. In Table 4, dam breach cases with flood characteristics are summarized and used to quantify each node (Zhou, 2007; Cui et al., 2009; Peng and Zhang, 2012c; Wang, 2018). Among these cases, the priori probabilities of distance to dam site and flow velocity can be obtained directly from the statistical probability, while the conditional probability of water depth can be obtained from the case data.

Table 4
Statistics of dam breach cases with flooding characteristics

Name	Dam height (m)	Lake volume ($\times 10^4$ m^3)	Distance to dam site (km)	Water depth (m)	Flow velocity (m/s)
Tangjiashan	120	31600	3.5	6.13	1.11
			25.1	0.99	0.71
			32.6	1.21	0.66
			45.9	2.45	0.89
			55.9	3.90	0.61
			64.0	1.10	0.43
			77.8	0	0
Hongshiyan	83	26000	2	12.7	9.81
			11.5	13.55	8.345
			26.5	12.6	9.385
			41.5	11.2	8.8
			54	9.65	8.32
			65.5	12.4	8.115
Yigong	110	300000	8.13	47.79	1.95
			28.33	47.09	1.51
Macaotan downstream	30	10	7.85	1.0	0.07
			13.04	1.05	0.05
Hongsong	50	100	5	1.28	0.65
			10.2	1.52	0.31
Xiaogangjian upstream	95	1200	6.53	5.7	0.16
			23.55	3.78	0.20
			25.85	3.44	0.19
Dongkoumiao	21.5	255	0.75	0.49	7.15
			1.75	0.5	7.39
Lijiaju	25	145	0.6	3.7	5.2
Shijiagou	30	85	0.8	5.26	5.9
Liujiatai	35.8	4054	4	3.02	19.53
			11	1.3	13.41
			22.5	0.39	4.75
Hengjiang	45	7879	3.5	1.09	5.94
			4.5	1.03	4.75
			5.5	0.96	3.86
			8.0	1.3	3.4
			12.5	1.99	4.17
Banqiao	24.5	49200	9	3.41	3.03
			28.5	1.52	1.52
			52.5	0.16	2.62
Sheyuegou	41	6.78	2	3.04	2.25
			6	2.16	1.75
Qingshan	14.1	21350	0.375	5.84	2.27
			1.27	3.05	2.88

Name	Dam height (m)	Lake volume (×10^4 m^3)	Distance to dam site (km)	Water depth (m)	Flow velocity (m/s)
			2.73	3.28	1.84
			4.6	3.6	1.6
			6.66	3.79	1.96
			9.21	3.17	1.87
			12.16	2.76	1.38
			14.66	2.7	1.06
			17.65	0.98	0.89
			21.82	0.94	0.55
			25.27	1.01	0.84

Flood severity is an important index for describing flood characteristics, and it is mainly affected by water depth and flow velocity (see Fig. 1). The conditional probability distribution of flood severity is calculated by using the zoning standard proposed by Clausen and Clark (1990), and the Monte Carlo simulation method (see Table 5).

Table 5
Conditional probability distribution of flood severity

Water depth	Shallow				Medium				Deep				Very deep			
Velocity	Slow	Medium	Fast	Very fast	Slow	Medium	Fast	Very fast	Slow	Medium	Fast	Very fast	Slow	Medium	Fast	Very fast
High (%)	0	0	0	0	0	0	0	34.95	0	0	62.21	100	100	100	100	100
Medium (%)	0	0	5.18	38.2	0	11.49	86.44	65.05	5.9	87.28	37.79	0	0	0	0	0
Low (%)	100	100	94.82	61.8	100	88.51	13.56	0	94.1	12.72	0	0	0	0	0	0

4.2 Flood duration

Flood duration mainly utilizes flood arrival time and flood rise time. Flood arrival time is the time from dam breaching to the time when the flood reaches the affected area, which is also one of the key influencing parameters for the evacuation of downstream risk population and movable property. Affected by dam height and lake volume, the flood arrival time can be calculated using the following empirical formula (Yellow River Water Conservancy Commission, 1977):

$$T_1 = K_1 \frac{x^{1.75}\left(10 - H_0\right)^{1.3}}{V_l^{0.2} H_d^{0.35}}$$

4

where T_1 is the flood arrival time; K_1 is the time coefficient, $K_1 = 0.7\times10^{-3}$; H_0 is the average base flow depth, which can be assumed as 2 m; x is the distance to dam site; V_l is the lake volume; H_d is the dam height.

The Monte Carlo simulation method was used to calculate the conditional probability of the flood arrival time. Taking the inundation area far from the dam site (12–36 km) as example, the probability distribution of flood arrival time is shown in Table 6.

Table 6
Conditional probability distribution of flood arrival time (distance to dam site: 12–36 km) (unit: %)

Node	Lake volume	Small				Medium				Large				Very large		
	Dam height	Low	Medium	High	Very high	Low	Medium	High	Very high	Low	Medium	High	Very high	Low	Medium	
Flood arrival time	Short	0	0	0	0	99.85	19.88	0	0	100	84.01	2.04	0	100	100	
	Medium	0	100	14.42	0	0.15	80.12	45.19	0	0	15.99	95.54	19.47	0	0	
	Long	22.13	0	85.58	20.64	0	0	54.81	67.29	0	0	2.42	80.53	0	0	
	Very long	77.87	0	0	79.36	0	0	0	32.71	0	0	0	0	0	0	

Flood rise time refers to the time interval between the time when the flood reaches the affected areas and the flood rising to a dangerous water depth, which is mainly affected by water depth, dam breach duration, and distance to dam site. It can be calculated by the following formula (Peng and Zhang, 2012a):

$$T_2 = B_t \cdot \alpha\beta\gamma$$

5

where T_2 is the flood rise time; B_t is the dam breach duration; α is the ratio of the time to peak and the dam breach duration, and a mean value of 0.2225 is suggested; $\beta = 1$ when the peak water depth h_p <1.5 m, and $\beta = 1.5/h_p$ when h_p >1.5 m; γ is the coefficient of energy consumption, the values of γ are assumed as 1, 1.05, 1.1, 1.2 when the distances to dam site are 0–4.8 km, 4.8–12 km, 12–36 km, and > 36 km. Taking the inundation area very far from the dam site (> 36 km) as example, the probability distribution of flood rise time is shown in Table 7.

Table 7
Conditional probability distribution of flood rise time (distance to dam site: >36 km) (unit: %)

Node	Water depth	Shallow				Medium				Deep				Very deep	
	Duration of flood routing	Short	Medium	Long	Very long	Short	Medium	Long	Very long	Short	Medium	Long	Very long	Short	Medium
Flood rise time	Short	94.57	0	0	0	99.85	19.88	0	0	100	84.01	2.04	0	100	100
	Medium	5.43	100	14.42	0	0.15	80.12	45.19	0	0	15.99	95.54	19.47	0	0
	Long	0	0	85.58	20.64	0	0	54.81	67.29	0	0	2.42	80.53	0	0
	Very long	0	0	0	79.36	0	0	0	32.71	0	0	0	0	0	0

The probability distribution of each node is the input and the Bayesian network of the evolution of dammed lake outburst flood is obtained (see Fig. 4).

5 Loss Assessment Because Of Dammed Lake Breaching

Outburst of a dammed lake causes irreversible harm to the downstream area. The loss caused by dammed lake breaching mainly includes losses of life and property.

5.1 Loss of life

When a dammed lake breach occurs, downstream population at risk should be evacuated to safe zones. Evacuation behavior includes a series of processes, such as decision-making, alarm, reaction, and evacuation. Whether an evacuation is successful depends on the required time and the available time for evacuation, which can be expressed by:

$$W_t + T_2 > T_n$$

6

where W_t is the warning time; T_2 is the flood rise time; T_n is the time required for the evacuation.

Loss of life is generally defined as the exposed population lost in the flood and is primarily determined by the level of achieved evacuation and flood characteristics. The loss of life Bayesian network is mainly composed of six nodes, such as dam breach time, warning time, evacuation distance, time required for evacuation, evacuation situation, and loss of life. The node states and value ranges of loss of life because of dammed lake breaching are shown in Table 8.

Table 8
Node states and value ranges of loss of life because of dammed lake breaching

Node	State	Value range
Dam breach time	Day; Evening; Night	(8:00–17:00]; (17:00–22:00]; (22:00–8:00]
Warning time (h)	Short; Medium; Long; Longer	(0, 0.25]; (0.25, 1]; (1, 3]; (3, ∞)
Evacuation distance (km)	Close; Medium; Far; Farther	(0, 0.1]; (0.1, 0.5]; (0.5, 1.5]; (1.5, ∞)
Required time for evacuation (h)	Short; Medium; Long; Longer	(0, 0.25]; (0.25, 1]; (1, 3]; (3, ∞)
Evacuation situation	Evacuated; Not-evacuated	Evacuated; Not-evacuated
Loss of life	Died; Alive	Died; Alive

Dam breach is generally an equal probability event, and it is assumed that the dam breach time is evenly distributed throughout a day. That is, the probability of dammed lake breach events occurring in the daytime (8:00–17:00), evening (17:00–22:00), and night (22:00–8:00) is 0.375, 0.208, and 0.417, respectively.

In addition, because of the limited location information of evacuation points, it is assumed that these events follow a uniform distribution.

Warning time refers to the time from when the alarm is issued to when the flood reaches the affected area, which is one of the key factors affecting the successful evacuation of the population at risk. The warning time consists of warning initiation time and flood arrival time. The warning initiation time is generally affected by the dam breach time. In general, it is easier to find the dam breach event in the daytime. The conditional probability of the warning time is calculated by the Monte Carlo method, and the probabilities of 0–0.25 h, 0.25–1 h, 1–3 h, and > 3 h are 0.218, 0.252, 0.259, and 0.271, respectively (see Table 9).

Table 9
Probability distribution of input parameters for the network of loss of life

Parameter	Dam breach time		
	8:00–17:00	17:00–22:00	22:00–8:00
Warning initiation time (h)	*U*(0, 1.5)	*U*(-0.5, 1)	*U*(-1, 0.5)
Warning propagation time (h)	*W*(3.5, 0.6)	*W*(2, 0.5)	*W*(1.3, 0.7)
Response time (h)	*W*(0.25, 1)		
Evacuation speed (km/h)	6	4.98	4.26
Note: *U*(*a*, *b*) obeys uniform distribution; *W*(*a*, *b*) obeys Weibull distribution.			

The time required for evacuation is the time for the downstream population at risk to be successfully evacuated to safe zones, which is mainly composed of warning initiation time, warning propagation time, response time, and evacuation time. The warning propagation time is the time elapsed between issuing and receipt of warnings by the population at risk, which is related to factors such as weather, dam breach time, and flood characteristics. Lindell (2008) suggested to use the Weibull distribution (see Eq. (7)) to depict the warning propagation time (see Table 9). Response time refers to the time required for the population at risk to receive the warning and initiate an emergency response, which is also depicted by the Weibull distribution (see Table 9). The evacuation time is the time required by the evacuation process, which is mainly related to the dam breach time, evacuation distance, and evacuation speed of the population at risk (see Table 9). The conditional probability of the time required for evacuation is calculated in combination with the Monte Carlo method, and the probabilities of 0–0.25 h, 0.25–1 h, 1–3 h, and > 3 h are 0.124, 0.523, 0.297, and 0.056, respectively.

$$P_t = 1 - e^{-at^b}$$

7

where P_t is the proportion of population corresponding to a certain state; a and b are the two coefficients in $W(a, b)$.

The evacuation situation is mainly affected by flood rise time, warning time, and evacuation time. Herein, the Monte Carlo simulation method is used to calculate the conditional probability distribution of the evacuation situation.

The population at risk in the affected areas may be in four states: successful evacuation, low, medium, and high flood severities. The loss rate of successfully evacuated population is assumed to be 0 (Jonkman, 2007), while the population in the high severity flood area has a high loss rate, with a hypothetic value of 0.9078 (Peng and Zhang, 2012a). Compared with the two extremes, low and medium severity floods are more uncertain, which is why they are quantified based on historical data. Jonkman (2007) proposed a lognormal distribution function with water depth as independent variable to simulate the rate of loss of life:

$$F_D(h) = \Phi\left[\frac{\ln(h) - \mu_D}{\sigma_D}\right]$$

8

where $F_D(h)$ is the loss rate for a certain flood severity when the water depth of the downstream affected area is h; Φ is the standard normal distribution function; μ_D and σ_D are mean and standard deviation, respectively.

Areas with low flood severity generally have slow flow velocity and shallow water depth, which pose little threat to the safety of the affected population. Low severity flood cases were collected from the datasets from Jonkman (2007) and Zhou (2007), and the lognormal distribution curve was utilized for data fitting (see Fig. 5). Based on the fitting curve, μ_D, σ_D, and R^2 (coefficient of determination) are 3.71, 1.382, and 0.497, respectively.

Areas with medium flood severity generally have deeper water depths and faster flow rates than areas with low flood severity, with an accompanying greater potential for damage to the affected population. Medium severity flood cases were collected based on datasets from Jonkman (2007), Zhou (2007), and McClelland and Bowles (2002), and the lognormal distribution curve was also utilized to fit the data (see Fig. 6). Based on the fitting curve, μ_D, σ_D, and R^2 are 1.624, 0.472, and 0.658, respectively.

Based on the Monte Carlo simulation method, the conditional probability distribution of the loss of life in the Bayesian network can be obtained (see Fig. 7).

5.2 Loss of property

The loss of property because of dammed lake breaching can be divided into direct and indirect economic losses. Of these, direct economic loss is the main component of the loss, which is generally represented by the direct economic loss rate. The economic loss rate refers to the ratio of the post-disaster value of disaster bearing bodies of various economic types after experiencing floods to their economic value before or in normal years (Das and Lee, 1988). This can be expressed by:

$$R_{\mathrm{Loe}}(i) = \frac{F_i - F_i^{'}}{F_i}$$

9

where $R_{\mathrm{Loe}}(i)$ is the direct economic loss rate of economy type i; F_i is the pre-disaster value of economy type i; $F_i^{'}$ is the post-disaster value of economy type i.

The economic loss rate is usually determined by investigating the flood losses in previous years to establish the relationship function between the loss rate, the water depth, and other factors; it can also be determined in reference to loss rates of other catchments. In this study, as an example, the actual loss rates of 14 different types of economic losses in China were collected as data sources (Kang et al., 2006). According to the trends of water depth and loss rates, loss of property can be divided into the four categories of loss of agriculture and fishery, loss of forestry and animal husbandry, loss of industry, commerce, and residential property, and loss of tertiary industry and infrastructure. In addition, various types of economic property values are taken as weights, however, there is a lack of available data. Therefore, the proportions of Gross Domestic Product (GDP) for each economic type are taken as weight to calculate the loss rate of each loss classification. This can be expressed as:

$$R_{\mathrm{Loe}}(i, h) = \sum_j \left[Gr_{ij} \cdot R_{\mathrm{Loe}}(i, j, h) \right]$$

10

where $R_{\mathrm{Loe}}(i, h)$ is the direct economic loss rate corresponding to the economy type i at water depth h; Gr_{ij} is the share of GDP of category j in economy type i, where i is the economic type classification, and j is the economic subcategory within category i; $R_{\mathrm{Loe}}(i, j, h)$ is the direct economic loss rate corresponding to the economy type j at water depth h.

Based on the economic statistics for China (Kang et al., 2006), the relationship between water depth and loss rate of each economy type can be calculated by Eq. (11), and the curve is obtained by fitting a power function to these data (see Fig. 8). The R^2 values of the four curves were 0.919, 0.959, 0.994, and 0.995, indicating a good fitting effect. It is worth mentioning that in this study, the basic information is collected from China; the same fitting method can be adopted for other countries. In addition, the economic loss rate changes little when the water depth exceeds 3.5 m. Therefore, the loss rate at a water depth of 3.5 m is taken as the upper limit of the loss rate.

$$R(h) = \begin{cases} x \cdot h^y & h \leqslant 3.5 \\ R(3.5) & h > 3.5 \end{cases}$$

11

where $R(h)$ is the direct economic loss rate corresponding to water depth; x and y are the fitting coefficients.

Today, most studies on the economic loss rate merely consider water depth. However, the economic loss rate is also related to warning time, flow velocity, and flood duration. In this study, based on the function of water depth and loss rate, the disaster-causing environment and other influencing factors are considered to revise the expression of the economic loss rate.

Agriculture and fishery are the economic activities people have relied on for survival since ancient times. Clearly, the deeper the water depth, the higher the loss, and most crops die when the water depth exceeds 0.5 m. Moreover, crops are sensitive to the duration of flood and have poor inundation tolerance. According to previous studies on the relationship between the duration of inundation and the loss rate in China (Wang, 2009; Yang et al., 2010), the suggested loss rate values for agriculture and fishery related to flood duration are shown in Table 10.

Table 10
Suggested loss rate values for agriculture and fishery activities considering flood duration

Water depth (m)	0-1.5				1.5-3				> 3			
Flood duration (h)	0–2	2–4	4–6	> 6	0–2	2–4	4–6	> 6	0–2	2–4	4–6	> 6
Loss rate (%)	65	80	95	100	85	95	100	100	95	100	100	100

Although forestry, animal husbandry, tertiary industry, and infrastructure have better water tolerance, they still have high loss rates because of the influence of flow velocity. The correction coefficient for flow velocity based on statistics for China can be expressed as (Liu et al., 2016):

$$f_v = \begin{cases} \frac{(m-1)(v-0.1)^2}{(v_2-0.1)^2} & v \leqslant v_2 \\ m & v > v_2 \end{cases}$$

12

where f_v is the correction coefficient for flow velocity; m is the reciprocal of economic loss rate $R_{\text{Loe}}(i,h)$ corresponding to water depth h; v is the flow velocity; v_2 is assumed to be 3 m/s based on the loss statistical information for China.

In addition to water depth and flow velocity, warning time is also a major factor for the loss assessment of industry, commercial, and residential property because of their shared characteristic of partial portability. In general, the longer the alarm time, the lower the economic loss rate (until a limiting value). The correction coefficient for waring time based on statistics in China can be expressed as (Liu et al., 2016):

{f_T}=\left\{ {\begin{array}{*{20}{c}} {\frac{{\left({n - 1} \right)T}}{{{T_1}}}+1}&{T \leqslant {T_1}} \\ n&{T>{T_1}} \end{array}} \right.
13

where f_T is the correction coefficient for warning time; n is the property transfer coefficient, which is assumed to be 0.65 when the rescue level is uncertain; T_1 is the time required for residents to move portable properties to a safe zone, which is assumed to be 1.5 h based on a questionnaire survey in China (Shi et al., 2009).

According to the function of water depth and loss rate, considering the correction coefficients of each influencing factor, the direct economic loss rate of dammed lake breaching can be expressed as:

{R_{{\text{Loe}}}}\left(i \right)={R_{{\text{Loe}}}}\left({i,h} \right) \cdot {f_d} \cdot {f_v} \cdot {f_T}
14

where $R_{\text{Loe}}(i)$ is the direct economic loss rate of the economy type i; f_d is the correction coefficient of flood duration (see Table 10).

In Eq. (14), only the correction coefficient corresponding to the node needs to be calculated, and all other correction coefficients are assumed to be 1. The Bayesian network for loss of economy is established in Fig. 9.

Based on the direct economic loss rate in each disaster-causing environment, regional loss of the economy can be expressed as:

{L_E}=\sum\limits_{i} {{F_i}} \cdot {R_{{\text{Loe}}}}\left(i \right) \cdot \left({{p_i}+1} \right) \cdot {\left({r+1} \right)^t}
15

where L_E is the regional loss of economy; F_i is the pre-disaster value of the economy type i; p_i is the indirect loss coefficient, which can be 15%, 15%, 35%, and 20% for loss of agriculture and fishery, loss of forestry and animal husbandry, loss of industry, commerce, and residential property, and loss of tertiary industry and infrastructure, respectively (Wahlstrom et al., 1999); r is the comprehensive annual economic growth rate, which can be determined by the local economic development; t is the year between the forecast year and the base year.

5.3 Comprehensive loss

The losses caused by dammed lake breaching are multifaceted and multi-dimensional, and the impact degree of each loss is different. Consequently, how the loss caused by the outburst flood can be comprehensively evaluated is also a very important issue. Losses of life and economy are normalized according to the influence degree of the two types of losses, to conduct quantitative analysis under a uniform standard for different losses. The normalized functions based on the statistics for China can be expressed as (Li et al., 2006):

\left\{ {\begin{array}{*{20}{c}} {F_l}=\frac{1}{{{5^{0.1}}}}{{\left({\log {\text{Lol}}} \right)}^{0.1}}} \\ {F_e}=\frac{1}{{{5^{0.2}}}}{{\left({\log \frac{{{\text{Loe}}}}{{10}}} \right)}^{0.2}}} \end{array}} \right.
16

where F_l and F_e are the indexes of losses of life and economy, respectively; Lol and Loe are the values of losses of life and economy, respectively.

Comprehensive loss assessment can be conducted by linear weighting of losses of life and economy:

F={F_l} \cdot {S_l}+{F_e} \cdot {S_e}
17

where F is the comprehensive loss; S_l and S_e are the weighting coefficients of losses of life and economy, which are assumed to be 0.875 and 0.125, respectively, according to the statistics for China (Li et al., 2006).

6 Risk Level Of A Dammed Lake

The risk level of a dammed lake is mainly determined by the failure probability of the dammed lake and the losses caused by the outburst flood. The Bayesian network for dammed lake risk assessment by integrating each module is shown in Fig. 10.

The classification of the risk level of a dammed lake requires quantitative definitions of failure probability and loss. There is no unified quantitative definition for the hazard of dammed lake failure. Certain documented disposal cases of dammed lakes in the past 20 years are presented in Table 11 (Cai et al., 2021), in which the average, median, and maximum disposal times were 13.37 d, 12 d, and 31 d, respectively. Further, based on 352 dammed lake cases around the world, Shen et al. (2020) reported that 60% of the cases failed within two weeks, 68.2% within a month, and 84.4% within a year; thus, the longevities of the dammed lakes often determine the danger the dammed lakes impose themselves. In this study, the longevities of dammed lakes of two weeks, one month, and one year are taken as boundaries around which extra-high, high, medium, and low risks for the hazard assessment of dammed lakes are defined.

Table 11
Statistics of disposal time of dammed lakes

Name	Formation time	Disposal start time	Disposal completion time	Duration of disposal (d)
Yigong	2000.04.09	2000.05.03	2000.06.03	31
Jingshan	2001.04.25	2001.05.02	2001.06.01	29
Qianjiapin	2003.07.13	2003.07.13	2003.07.21	8
Terano	2004.10.23	2004.11.05	2004.11.21	17
Higashitakezawa	2004.10.23	2004.11.06	2004.11.25	20
Qingyandong	2005.06.21	2005.06.23	2005.06.26	4
Tangjiashan	2008.05.12	2008.05.19	2008.06.01	12
Xiaojiaqiao	2008.05.12	2008.05.23	2008.06.06	14
Wenjiaba	2008.05.12	2008.05.23	2008.06.04	12
Xiaogangjian	2008.05.12	2008.05.21	2008.06.10	20
Jiweishan	2009.06.05	2009.06.06	2009.06.12	7
Houziyan	2009.08.06	2009.08.06	2009.08.11	6
Miaoba	2010.07.19	2010.07.19	2010.07.26	7
Zhouqu	2010.08.07	2010.08.08	2010.08.30	23
Yixialuohe	2012.09.03	2012.09.05	2012.09.09	5
Hongshiyan	2014.08.03	2014.08.06	2014.08.13	8
Sucun	2016.09.28	2016.10.01	2016.10.15	15
Ayongxiaoqiao	2018.08.01	2018.08.07	2016.08.12	6
Baige	2018.11.03	2018.11.04	2018.11.13	10

In general, the shorter the longevity of a dammed lake, the higher the hazard. Based on the dataset, one month is assumed to be a manageable time for the disposal of dammed lakes, so the failure probability of dammed lake within a month is used to determine the inverse proportional function.

P=\frac{{0.465}}{{{P_t}}}
18

where P is the failure probability of the dammed lake; P_t is the failure probability of the dammed lake within the time period of t. The boundaries of extra-high, high, medium, and low risks are calculated as 0.775, 0.682, and 0.551. That is, the failure probability of dammed lake is extra-high if it exceeds 77.5%, high if it exceeds 68.2%, medium if it exceeds 55.1%, and low if it is less than or equal to 55.1%.

The risk value is the product of the probability of dammed lake failure and the loss because of dammed lake breaching. Based on the results of hazard evaluation and outburst flood routing simulation of the dammed lake, while simultaneously considering the situation of socio-economic development in China, the risk level can be divided into four levels: risk level Ⅰ ($0 \leq I_r \leq 0.439$), risk level Ⅱ ($0.439 < I_r \leq 0.583$), risk level Ⅲ ($0.583 < I_r \leq 0.687$), and risk level Ⅳ ($0.687 < I_r \leq 1$) (see Table 12). The classification of the risk level of a dammed lake can provide a reference for the emergency response.

Table 12
Risk level classification for dammed lakes

Variable	Low	Medium	High	Extra-high
Indicators of each loss type	Loss of life ≤ 3 Loss of economic ≤ 10 million CNY	Loss of life ≤ 10 Loss of economic ≤ 50 million CNY	Loss of life ≤ 30 Loss of economic < 100 million CNY	Loss of life ≥ 30 Loss of economic ≥ 100 million CNY
Index of loss (I_l)	$0 \le I_l \le 0.796$	$0.796 < I_l \le 0.855$	$0.855 < I_l \le 0.887$	$0.887 < I_l \le 1$
Index of hazard (I_h)	$0 \le I_h \le 0.551$	$0.551 < I_h \le 0.682$	$0.682 < I_h \le 0.775$	$0.775 < I_h \le 1$
Index of risk (I_r)	$0 \le I_r \le 0.439$	$0.439 < I_r \le 0.583$	$0.583 < I_r \le 0.687$	$0.687 < I_r \le 1$
Risk level				

7 Case Study

To test the rationality and applicability of the risk assessment method, the case of Baige dammed lake in China was chosen, for which detailed measured information on dam breaching process and flood routing, as well as the statistics of losses are available. On November 3, 2018, a huge landslide occurred on the right bank of the Jinsha River near Baige Village, Tibet Autonomous Region, China, which blocked the Jinsha River and formed a dammed lake (Zhong et al., 2020) (see Fig. 11).

7.1 Failure probability of the Baige dammed lake

The “11.03” Baige landslide deposits were tongue-shaped in-plane, and the landslide dam was composed of a long-range debris flow, which had the height, length, width, and volume of 86 m, 600 m, 1200 m, 2.83×10^6 m^3, respectively (Cai et al., 2020b). The dam was mainly composed of soil gravel and block stone, with a ratio of soil to rock of about 7:3. The dammed lake had a maximum volume of 7.9×10^8 m^3, and the upstream catchment area is about 1.7×10^5 km^2 (Cai et al., 2020b; Zhong et al., 2020; Shan et al., 2022). After inputting the relevant parameters into the Bayesian network of hazard assessment, the failure probability of Baige dammed lake was calculated at 81.3% (see Fig. 12).

7.2 Loss assessment because of the breaching of Baige dammed lake

In this section, the evacuation situation and flood characteristics of each region are considered comprehensively, and the loss of life is quantitatively evaluated based on the possible fatality in each region. The inundation area because of Baige dammed lake outburst flood mainly involves both sides of the Jinsha River from the dam site to the Liyuan Hydropower station. Therefore, the loss assessment is conducted for this area, which involved the cities of Diqing and Lijiang in Yunnan Province.

After hours of erosion by overtopping flow, Baige dammed lake breached, and the time to peak occurred at about 18:00, November 13, 2018 (Zhong et al., 2020). Based on the measured data, the duration of the dam breach lasted for about 14 h. To facilitate the analysis, the water depth and flow velocity in the inundation area were calculated using the empirical formulas shown in Eqs. (19) and (20), respectively (Xie, 1993):

{h_x}={h_0}{\left[{\frac{1}{{1+\frac{{4{\lambda ^2}\left({2\bar {m}+1} \right)h_{0}^{{2\bar {m}+1}}}}{{\bar {m}{{\left({\bar {m}+1} \right)}^2}{i_0}V_{l}^{2}}}x}}} \right]^{\frac{1}{{2\bar {m}+1}}}}
19

where h_x is the water depth at the cross-section x km away from the dam site; h_0 is the water depth at the dam site; \bar {m} is the index of the cross section shape; i_0 is the longitudinal slope at the bottom of the river; λ is the coefficient of the cross section, which is approximately the width of the cross section.

{v_x}=\frac{{{Q_x}}}{{\lambda \cdot {h_x}}}
20

where v_x is the flow velocity rate at the cross-section x km away from the dam site; Q_x is the peak flow at the cross-section x km away from the dam site, which can be expressed as:

{Q_x}=\frac{{{V_l}}}{{\frac{{{V_l}}}{{{Q_p}}}+\frac{x}{{vK}}}}
21

where Q_p is the peak breach flow at the dam site; vK is the empirical coefficient of the flow velocity, vK = 7.15 for mountainous area, and vK = 3.13 for flat area.

The parameter values of each node are input into the network to obtain the rate of loss of life (see Fig. 13). Because of timely manual intervention, the population at risk was all evacuated before Baige dammed lake breaching; hence, the calculated loss of life is 0 (see Table 13). Table 13 also presents the calculated loss of life under the condition without manual intervention.

Table 13
Calculated loss of life in the inundation area because of the Baige dammed lake breaching

City name	Town name	Distance to dam site (km)	Water depth (m)	Flow velocity (m/s)	Risk population	Rate of loss of life[a] (%)	Loss of life[a]	Rate of loss of life[b] (%)	Loss of life[b]
Diqing	Benzilan	385.68	7.63	2.54	5419	0.14	8	0	0
	Tuoding	449.90	10.11	2.74	5285	0.14	7	0	0
	Wujing	476.73	8.95	4.17	2254	0.14	3	0	0
	Tacheng	489.10	10.82	1.75	4839	0.14	7	0	0
Lijiang	Judian	519.81	4.19	2.54	11755	0.043	5	0	0
	Liming	546.00	6.5	1.74	9902	0.021	2	0	0
	Shigu	582.28	3.88	1.85	12267	0.021	3	0	0
Total	Total	—	—	—	51721	—	35	—	0

Note: [a] Without manual intervention; [b] With manual intervention.

Through the emergency response, all population at risk were safely evacuated, but the flood has caused irreversible economic losses. Based on the statistics of the disaster situation of various industries in the inundation areas, the rates of loss of property for various industry types were calculated by the Bayesian network (see Fig. 13), and the total loss of property was calculated by Eq. (15). A comparison between calculated and actual loss of properties of the two affected cities in the inundation area is shown in Table 14. It is worth mentioning that, only the actual direct economic loss was available based on statistical data; therefore, the comparison was conducted on the direct economic loss. The comparison between the calculated and actual direct economic losses verified the rationality of the proposed model. In addition, the calculated total loss of property was also provided, and the risk index was calculated based on that value.

Table 14
Calculated loss of property in the inundation area because of the Baige dammed lake breaching

Affected area	Industry type	Pre-disaster value (CNY million)	The direct economic loss rate (CNY million)	The calculated direct economic loss (CNY million)	The actual direct economic loss (CNY million)	The calculated total loss of property (CNY million)
Lijiang	Agriculture and fishery	309.4	100%	309.4	2935.53	309.4
	Forestry and animal husbandry	5.8	91.2%	5.29		5.8
	Industry, commerce, and residential property	963	53.5%	515.21		814
	Tertiary industry and infrastructure	3476	77.3%	2686.95		3476
Diqing	Agriculture and Fishery	200	100%	200	3929.71	200
	Forestry and Animal husbandry	—	78.5%	—		—
	Industry, commercial, and residential property	70	37.1%	25.97		41
	Tertiary industry and infrastructure	5553	44.8%	2487.75		3493
Total	—	10577.2	—	6230.57	6865.24	8939.2

In summary, the failure probability of Baige dammed lake was 81.3%, resulting in the loss of 35 lives and an economic loss of CNY 8939.2 million under the condition without manual intervention. Hence, the risk index is 0.7339, which belongs to extra-high level classification.

8 Conclusions

Based on the Bayesian network, a quantitative risk assessment model for dammed lakes is developed, which includes modules of dammed lake hazard assessment, outburst flood routing simulation, and loss assessment. The risk classification method of the dammed lake is also proposed and applied to the Baige dammed lake. The following conclusions can be drawn:

(1) According to the disaster chain formed by dammed lake breaching, the risk topological structure is proposed, and parameters are selected according to the logical causality of each module; then, a Bayesian network for dammed lake risk assessment with 27 nodes and 41 arcs is developed.

(2) A database containing 313 documented dammed lake cases around the world was compiled. Considering morphological characteristics and material compositions of these landslide dams, as well as the hydrodynamic conditions of the dammed lakes, the network nodes are quantified. Then, the failure probability of dammed lakes and loss rates of life and economy because of dammed lake breaching are obtained.

(3) Based on the socio-economic development of China, as well as the longevity statistics of dammed lakes around the world, a new quantitative risk level classification method based on the losses of life and economy for dammed lakes is put forward. Taking 0.439, 0.583, and 0.687 as boundary values, the risk level of the dammed lake is divided into low, medium, high, and extra-high, which is more scientific and feasible than the previously used qualitative classification method.

(4) The results of the risk assessment show that the failure probability of the Baige dammed lake was 81.3%. Comparison between the calculated and actual losses of life under manual intervention, as well as the direct economic losses verified the rationality of proposed model. Without manual intervention, the loss of life would be 35, and the loss of property would be 8939.2 million CHY; hence, the calculated risk index is 0.7339, which belongs to the extra-high level.

Declarations

Data availability

Data used and obtained in the current work are available from the corresponding author upon request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (Grant No. U2040221). We would like to thank MogoEdit (https://www.mogoedit.com) for its English editing during the preparation of this manuscript.

References

1. Aguilera PA, Fernández A, Fernández R, Rumí R, Salmerón A (2011) Bayesian networks in environmental modelling. Environ Model Softw 12(26):1376–1388
2. Cai BP, Liu YH, Lin J, Liu ZK, Chang YJ, Jiang L (2020a) Application of Bayesian networks in reliability evaluation. Bayesian networks for reliability engineering. Springer, Singapore
3. Cai YJ, Luan YS, Peng WX, Li JQ, Xu Y (2021) Analysis on typical cases of barrier lake: cause, disaster and emergency response. Changjiang Press, Wuhan. (in Chinese)
4. Cai YJ, Cheng HY, Wu SF, Yang QG, Wang L, Luan YS, Chen ZY (2020b) Breaches of the Baige barrier lake: Emergency response and dam breach flood. Sci China: Technological Sci 63(7):1164–1176
5. Casagli N, Ermini L (1999) Geomorphic analysis of landslide dams in the Northern Apennine. Trans Japanese Geomorphological Union 20(3):219–249
6. Chai HJ, Liu HC, Zhang ZY (1998) Study on the categories of landslide-damming of rivers and their characteristics. J Chengdu Inst Technol 25(3):411–416 (in Chinese)
7. Clausen L, Clark PB (1990) The development of criteria for predicting dam-break food damages using modeling of historical dam failures. In: White W R (ed) International Conference on River Flood Hydraulics. Wiley, London, 369–380
8. Costa JE, Schuster RL (1988) The formation and failure of natural dams. Geol Soc Am Bull 100(7):1054–1068
9. Costa JE, Schuster RL (1991) Documented historical landslide dams from around the world. U.S. Geological Survey Open-File Report 91–239, Washington
10. Cui P, Zhu YY, Han YS, Chen XQ, Zhuang JQ (2009) The 12 May Wenchuan earthquake-induced landslide lakes: distribution and preliminary risk evaluation. Landslides 7(6):209–223
11. Cui P, Han YS, Chao D, Chen XQ (2011) Formation and treatment of landslide dams emplaced during the 2008 Wenchuan earthquake, Sichuan, China. Natural and Artificial Rockslide Dams, Edited by Evans SG, Hermanns RL, Storm A, Scarascia-Mugnozza G, Lecture Notes in Earth Sciences, 133: 295–321
12. Das S, Lee R (1988) A nontraditional methodology for flood stage-damage calculations. J Am Water Resour Assoc 24(6):1263–1272
13. Dempster A, Laird N, Rubin D (1977) Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc B 39(1):1–38
14. Dong JJ, Tung YH, Chen CC, Liao JJ, Pan YW (2011) Logistic regression model for predicting the failure probability of a landslide dam. Eng Geol 117:52–61
15. Ermini L, Casagli N (2003) Prediction of the behaviour of landslide dams using a geomorphological dimensionless index. Earth Surf Proc Land 28(1):31–47
16. Fan XM, Scaringi G, Korup O, West AJ, van Westen CJ, Tanyas H, Hovius N, Hales TC, Jibson RW, Allstadt KE, Zhang LM, Evans SG, Xu C, Li G, Pei XJ, Xu Q, Huang RQ (2019) Earthquake-induced chains of geologic hazards: Patterns, mechanisms, and impacts. Rev Geophys 57:421–503
17. Fan XM, Dufresne A, Subramanian SS, Strom A, Hermanns R, Stefanelli CT, Hewitt K, Yunus AP, Dunning S, Capra L, Geertsema M, Miller B, Casagli N, Jansen JD, Xu Q (2020) The formation and impact on landslide dams - State of the art. Earth Sci Rev 203:103116

18. Helm P (1996) Integrated risk management for natural and technological disasters. Tephra 15(1):4–13
19. Jensen FV (2001) Bayesian networks and decision graphs. Springer, New York
20. Jia H, Chen F, Pan D (2019) Disaster Chain Analysis of Avalanche and Landslide and the River Blocking Dam of the Yarlung Zangbo River in Milin County of Tibet on 17 and 29 October 2018. Int J Environ Res Public Health 16(23):4707
21. Jonkman SN (2007) Loss of life estimation in flood risk assessment: theory and applications. PhD Thesis of Delft University of Technology, Delft
22. Kang XW, Wu SH, Dai EH, Yang QY, Liu ZQ, Yang PG, Ma X, Zhao RF (2006) Loss and impact pre assessment of large scale flood disaster. Chin Sci Bull 51(S2):155–164 (in Chinese)
23. Khakzad N, Khan F, Amyotte P (2013) Quantitative risk analysis of offshore drilling operations: A Bayesian approach. Saf Sci 57:108–117
24. Korup O (2002) Recent research on landslide dams - a literature review with special attention to New Zealand. Prog Phys Geogr 26(2):206–235
25. Korup O (2004) Geomorphometric characteristics of New Zealand landslide dams. Eng Geol 73:13–35
26. Li L, Wang RZ, Sheng JB (2006) Study on evaluation models of severity degree of dam failure impact.Journal of Safety and Environment, (1):1–4. (in Chinese)
27. Liao HM, Yang XG, Xu FG, Xu H, Zhou JW (2018) A fuzzy comprehensive method for the risk assessment of a landslide-dammed lake. Environ Earth Sci 77:750
28. Lindell MK (2008) EMBLEM2: An empirically based large scale evacuation time estimate model. Transp Res Part A: Policy Pract 42(1):140–154
29. Liu WM, Carling PA, Hu KH, Wang H, Zhou Z, Zhou LQ, Liu DZ, Lai ZP, Zhang XB (2019) Outburst floods in China: A review. Earth Sci Rev 197:102895
30. Liu XX, Gu SP, Zhao YM, Lv WW, He L, He J 2016 Evaluation methods for economic loss due to dam-break flood using modified loss rate.Journal of Economics of Water Resources, 34(3):36–40 + 80–81. (in Chinese)
31. McClelland DM, Bowles DS (2002) Estimating life loss for dam safety risk assessment - a review and new approach. IWR report 02-R-3, Institute for Dam Safety Risk Management. Utah State University, Logan
32. Nie GZ, Gao JG, Deng Y (2004) Preliminary study on earthquake-induced dammed lake. Quaternary Sci 24(3):293–301 (in Chinese)
33. Norsys Software Corp. Netica Application 6.08. https://www.norsys.com/downloads/Netica_Win.exe/ Accessed 31 December 2020
34. O'Connor JE, Beebee RA (2009) Floods from natural rock-material dams. Cambridge University Press, New York
35. Pearl J (2003) Causality: models, reasoning and inference. Cambridge University Press, London
36. Peng M, Zhang LM (2012a) Analysis of human risks due to dam break floods - part 1: a new model based on Bayesian networks. Nat Hazards 64(1):903–933
37. Peng M, Zhang LM (2012b) Breaching parameters of landslide dams. Landslides 9(1):13–31
38. Peng M, Zhang LM (2012c) Analysis of human risks due to dam break floods - part 2: application to Tangjiashan landslide dam failure. Nat Hazards 64(2):1899–1923
39. Schuster RL (1985) Landslide dam in the Western United States. U.S. Geological Survey, Washington
40. Shan YB, Chen SS, Zhong QM (2020) Rapid prediction of landslide dam stability using the logistic regression method. Landslides 17(12):2931–2956
41. Shan YB, Chen SS, Zhong QM, Mei SY, Yang M (2022) Development of an empirical model for predicting peak breach flow of landslide dams considering material compositions. Landslides. doi: 10.1007/s10346-022-01863-1
42. Shen DY, Shi ZM, Peng M, Zhang LM, Jiang MZ (2020) Longevity analysis of landslide dams. Landslides 17(8):1797–1821
43. Shi Y, Xu SY, Shi C, Sun AL (2009) A review on vulnerability assessment of residential buildings in urban floods. J North China Inst Water Conservancy Hydroelectric Power 30(1):34–37 (in Chinese)
44. Shi ZM, Ma XL, Peng M, Zhang LM (2014) Statistical analysis and efficient dam burst modelling of landslide dams based on a large-scale database. Chin J Rock Mechan Eng 33(9):1780–1790 (in Chinese)
45. Smith M (2006) Dam risk analysis using Bayesian networks. In: Nadim F, Pöttler R, Einstein H, Klapperich H, Kramer S (eds), Proceedings of the 2006 ECI Conference on Geohazards. ECI Symposium Series No. 7, Engineering Conferences International (ECI), New York
46. Stefanelli CT, Segoni S, Casagli N, Catani F (2016) Geomorphic indexing of landslide dams evolution. Eng Geol 208:1–10
47. Stefanelli CT, Catani F, Casagli N (2015) Geomorphological investigations on landslide dams. Geoenvironmental Disasters 2:21
48. Strom A (2010) Landslide dams in Central Asia region. J Japan Landslide Soc 47(6):309–324
49. Tahata M, Mizuyama KN, Inoue N (2002) Natural reservoirs and disasters. Ancient and Modern Academy, Tokyo. (in Japanese)
50. Tong YX (2008) Quantitative analysis for stability of landslide dams. Master Thesis of National Central University, Taoyuan. (in Chinese)
51. Wahlstrom E, Loague K, Kyriakidis PC (1999) Hydrologic Response: Kaho'olawe, Hawaii. J Environ Qual 28:481–492
52. Wang SJ (2018) Study on dam break calculation and life loss assessment of landslide dam. Master Thesis of Xi'an University of Technology, Xi'an. (in Chinese)
53. Wang ZJ (2009) Study on methods for estimation of loss caused by dam break. PhD Thesis of Hohai University, Nanjing. (in Chinese)
54. Xie RZ (1993) Dam failure hydraulics. Shandong Science and Technology Press, Jinan. (in Chinese)
55. Yan R (2006) Secondary disaster and environmental effect of landslides and collapsed dams in the upper reaches of Minjiang River. Master Thesis of Sichuan University, Chengdu. (in Chinese)
56. Yang JM, Feng MQ, Lu YL, Han QQ (2010) Prediction of dam-break flood damage in Wenyuhe reservoir. J Wuhan Univ Technol (Information Manage Eng Edition) 32(4):598–601 (in Chinese)

57. Yellow River Water Conservancy Commission (1977) Preliminary discussion on calculation method of dam break flow. Institute of Water Conservancy Science, Yellow River Water Conservancy Commission, Zhengzhou. (in Chinese)
58. Zhang LM, Peng M, Chang DS, Xu Y (2016) Dam failure mechanisms and risk assessment. John Wiley & Sons, Singapore
59. Zhang LM, Xiao T, He J, Chen C (2019) Erosion-based analysis of breaching of Baige landslide dams on the Jinsha River, China, in 2018. Landslides 16(10):1965–1979
60. Zhong QM, Chen SS, Mei SA, Cao W (2018) Numerical simulation of landslide dam breaching due to overtopping. Landslides 16(6):1183–1192
61. Zhong QM, Wang L, Chen SS, Chen ZY, Shan YB, Zhang Q, Ren Q, Mei SY, Jiang JD, Hu L, Liu JX (2021) Breaches of embankment and landslide dams - State of the art review. Earth Sci Rev 216:103597
62. Zhong QM, Chen SS, Wang L, Shan YB (2020) Back analysis of breaching process of Baige landslide dam. Landslides 17(7):1681–1692
63. Zhou KF (2007) Study on analysis method for loss of life due to dam breach. Master Thesis of Nanjing Hydraulic Research Institute, Nanjing. (in Chinese)
64. Zhu Y, Peng M, Zhang P, Zhang LM (2021) Warning decision-making for landslide dam breaching flood using influence diagrams. Front Earth Sci 9:679862

Figures

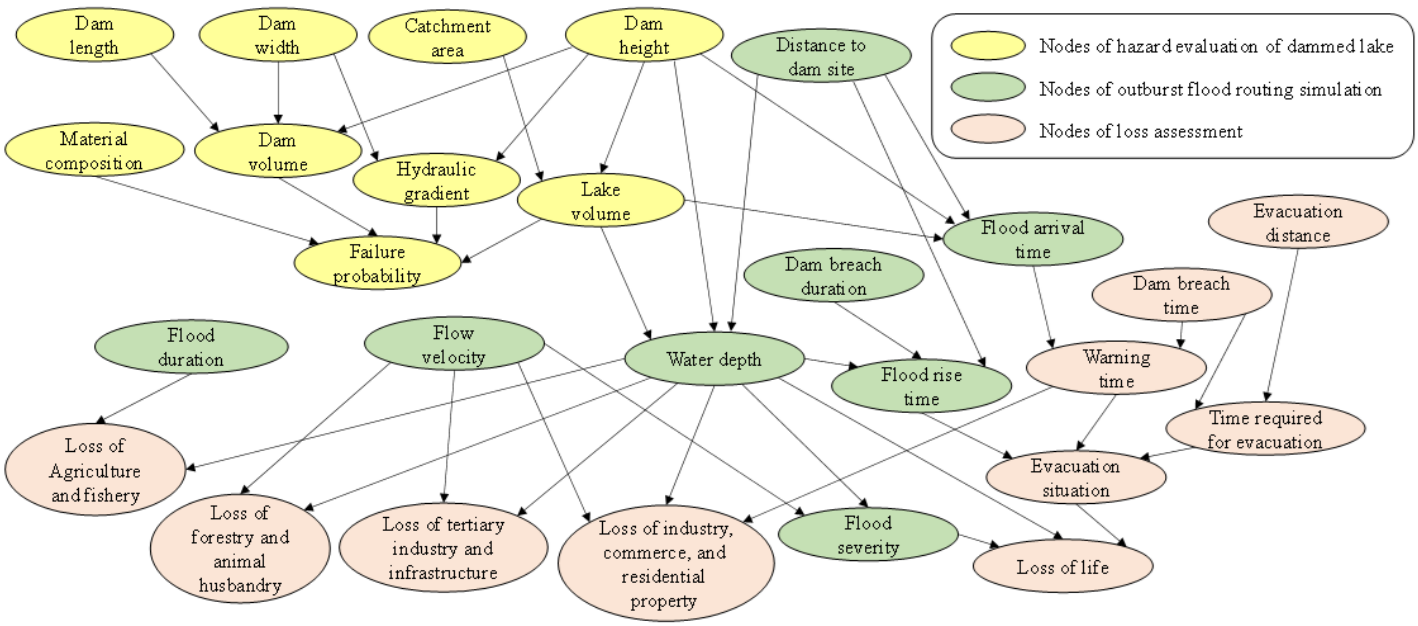

Figure 1

Topological map for dammed lake risk assessments

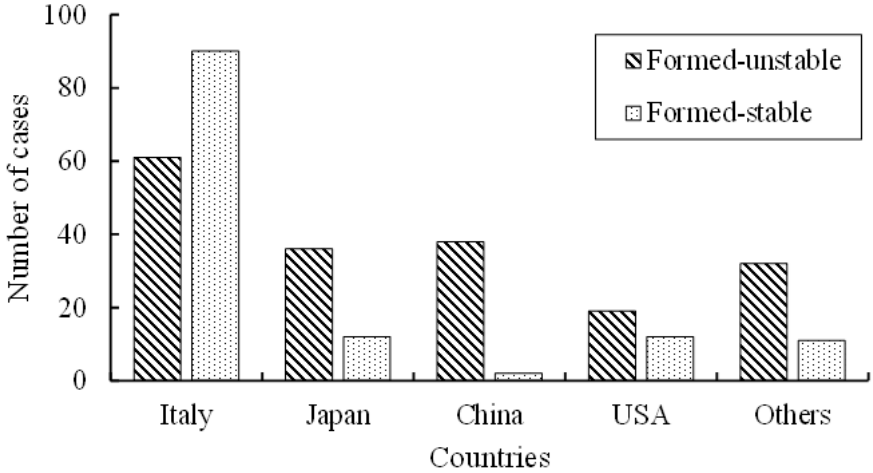

Figure 2

Documented formed unstable and formed stable cases around the world

Figure 3

Bayesian network model for hazard evaluations of dammed lakes

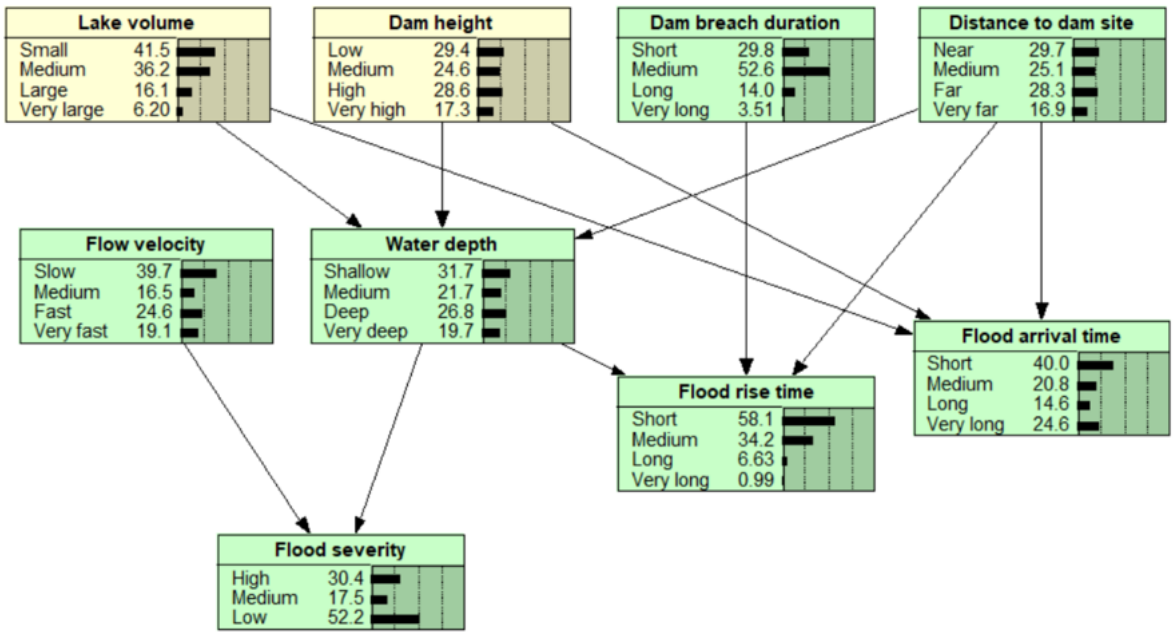

Figure 4

Bayesian network for outburst flood routing simulations of dammed lakes

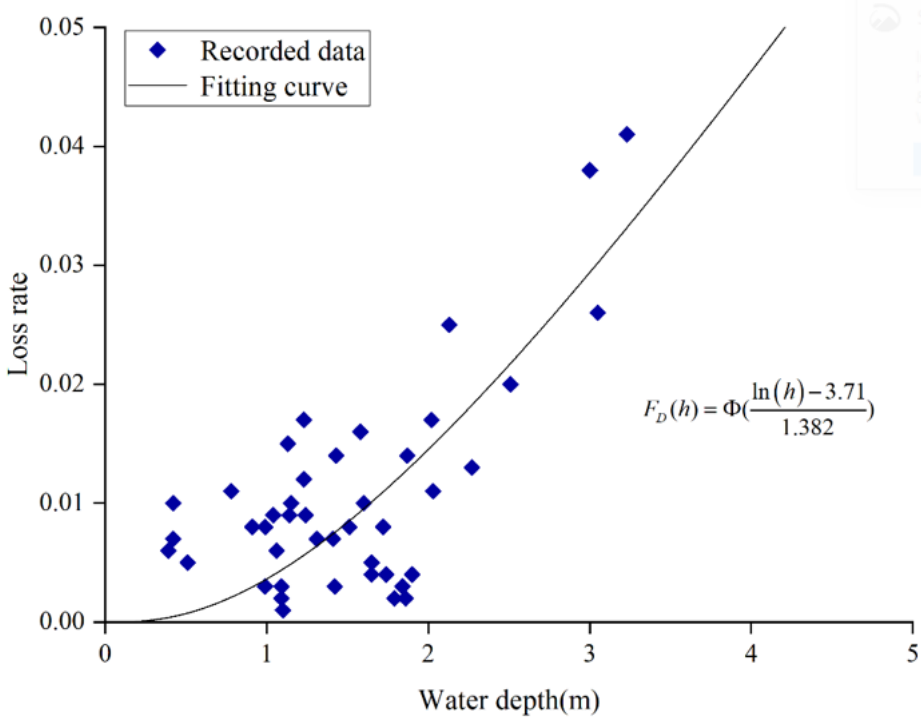

Figure 5

Fitting curve of the rate of loss of life for low severity floods

Figure 6

Fitting curve of the rate of loss of life for medium severity floods

Figure 7

Bayesian network for loss of life

Figure 8

Relationship between water depth and direct economic loss rate

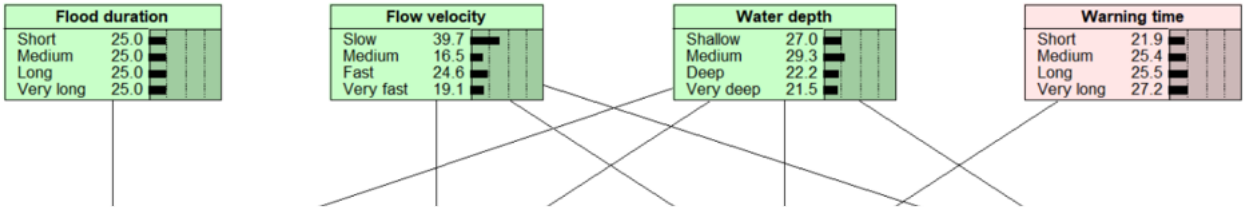

Figure 9

Bayesian network for loss of property

Figure 10

Bayesian network for dammed lake risk assessment

Figure 11

“11.03” Baige dammed lake (Photo credit: Xinhuanet.com)

Figure 12

Failure probability analysis of the Baige dammed lake

Figure 13

Loss assessment in the inundation area because of Baige dammed lake breaching

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1.docx