

基于贝叶斯网络的堰塞坝稳定性快速评价模型

杜镇瀚¹, 钟启明^{2,3}, 董海洲⁴, 单熠博²

(1. 南京水利科学研究院大坝安全与管理研究所, 江苏 南京 210024;
2. 南京水利科学研究院岩土工程研究所, 江苏 南京 210024; 3. 水利部水库大坝安全重点实验室, 江苏 南京 210029;
4. 河海大学土木与交通学院, 江苏 南京 210098)

摘要:为了科学、合理地评价堰塞坝稳定性,收集了国内外 313 个已溃和未溃堰塞坝的基础信息,建立了案例数据库;考虑堰塞坝形态特征、物质组成以及堰塞湖水动力条件包含的 8 个影响参数,基于贝叶斯网络提出了两种新的堰塞坝稳定性快速评价模型,分别为节点缺失条件下的简化模型和节点完整条件下的精细化模型。这两种新模型与国内外典型的堰塞坝稳定性快速评价模型的对比结果表明:基于贝叶斯网络的简化模型的绝对准确率和保守准确率分别为 83.33% 和 94.44%,精细化模型的绝对准确率和保守准确率分别为 80.77% 和 96.15%,验证了贝叶斯网络模型的可靠性和科学性,且贝叶斯网络模型可以实时输入和更新样本集,从而不断提高模型预测精度。

关键词:堰塞坝;稳定性评价;贝叶斯网络

中图分类号:P642 **文献标志码:**A **文章编号:**1006-7647(2023)04-0037-09

Rapid evaluation method of landslide dam stability based on Bayesian network//DU Zhenhan¹, ZHONG Qiming^{2,3}, DONG Haizhou⁴, SHAN Yibo² (1. Dam Safety Management Department, Nanjing Hydraulic Research Institute, Nanjing 210024, China; 2. Department of Geotechnical Engineering, Nanjing Hydraulic Research Institute, Nanjing 210024, China; 3. Key Laboratory of Reservoir Dam Safety of the Ministry of Water Resources, Nanjing 210029, China; 4. College of Civil and Transportation Engineering, Hohai University, Nanjing 210098, China)

Abstract: To perform scientific and rational evaluation on landslide dam stability, essential information was undertaken from 313 collapsed and intact landslide dams worldwide to establish a case database. Eight influential factors related to the morphological characteristics and material composition of landslide dams, hydrodynamic conditions of barrier lakes were considered. Based on the analysis of Bayesian networks, two novel and expedited methods for assessing landslide dam stability were introduced, including a simplified model with missing nodes and a detailed model with intact nodes. The comparison between typical models of rapid evaluation method of landslide dam stability from both home and aboard shows that the simplified Bayesian network model can achieve an absolute accuracy rate of 83.33% and a conservative accuracy rate of 94.44%. Similarly, the detailed model demonstrates absolute and conservative accuracy rates of 80.77% and 96.15%, respectively. The reliability and scientificity of the Bayesian network model are verified. The Bayesian network models can incorporate and update real-time sample sets, and thereby the predictive accuracy of the models can be increased perpetually.

Key words: landslide dam; stability evaluation; Bayesian network

在一定地质地貌条件下,由降水、地震等原因引起山崩、滑坡、泥石流等堵截山谷、河道,贮水而形成的湖泊统称为堰塞湖,山谷或河道内的堆积体称为堰塞坝^[1-3]。由于我国辽阔的疆土面积以及独特的地理、地质条件,位于强地震活动带的青藏高原、中

南以及西南山区成为堰塞湖的高发区^[4-5]。2008 年 5 月 12 日,四川汶川发生 8.0 级特大地震,产生了 257 处堰塞坝,其中规模最大、潜在威胁最严重的是唐家山堰塞坝^[6-7];据测算,上游堰塞湖最大蓄水量达 3.16 亿 m³,而且恰逢雨季,增大了堰塞坝的溃决

基金项目:国家自然科学基金区域创新发展联合基金重点项目(U2040221);中央级公益性科研院所基本科研业务费专项资金资助项目(Y320005)

作者简介:杜镇瀚(1998—),男,博士研究生,主要从事堰塞湖风险评估方法研究。E-mail:zhdu@nhri.cn
通信作者:钟启明(1981—),男,教授级高级工程师,博士,主要从事土石(堰塞)坝灾害模拟与防控研究。E-mail:qmqzhong@nhri.cn

风险,对下游 170 万人民生命财产安全造成了巨大的威胁^[8]。堰塞坝的寿命可以持续几分钟到几千年不等,根据 352 座堰塞坝的寿命统计发现^[9],29.8%的堰塞坝寿命小于 1 d,68.2%的寿命小于 1 月,84.4%的寿命小于 1 a。堰塞湖一旦溃决,往往造成严重的洪水灾害,给下游人民生命财产安全带来严重威胁^[10]。但有的堰塞坝形成后结构稳定,至今仍未溃决,如位于塔吉克斯坦因 1911 年地震形成的 Usoi 堰塞坝,体积约为 2 200 万 m³,高度约为 600 m,堰塞湖库容约为 170 亿 m³,是世界上现存库容最大的堰塞湖。

研究表明,堰塞坝的稳定性与其物质组成^[11]、结构特征^[12]、堰塞湖水动力条件^[13]及当地的降水条件和次生灾害^[14-15]相关。合理的稳定性评价有利于科学的应急响应,减少社会财产损失、保障人民生命安全和降低环境生态破坏,因而具有重要的理论意义和实际价值。我国现行的水利行业规范 SL 450—2021《堰塞湖风险等级划分标准》根据堰塞湖规模、堰塞坝物质组成和堰塞坝高度,将堰塞坝危险级别划分为极高危险、高危险、中危险和低危险^[16];Casagli 等^[11]基于亚平宁北部山区的 70 个堰塞坝案例,最早提出了堰塞坝稳定性快速评价方法——堆积指标法(BI 法),评价指标为堰塞坝体积和堰塞湖流域面积;随后,Ermini 等^[17]基于 84 个堰塞坝案例,在堆积指标法基础上引入堰塞坝高度,提出了无量纲堆积体指标法(DBI 法)。目前,BI 法和 DBI 法仍是世界上使用最为广泛的堰塞坝稳定性快速评价方法。另外,Korup^[18]、Dong 等^[19]、Stefanelli 等^[20]、Shan 等^[21]、石振明等^[22]也基于各自收集的堰塞坝案例提出了一系列的稳定性快速评价方法。但上述评价方法都是基于多元回归或逻辑回归,关系式一经确定,无法根据新增的案例实现动态更新,且计算结果依赖于表达式所需参数的完整性,在实际应急抢险中操作性不强,因此亟须提出可量化的堰塞坝稳定性快速评价方法。

1 贝叶斯网络

贝叶斯网络是一种表示变量间概率关系的有向无环图,可以看作是概率论和图论的结合。网络中的节点表示随机变量,节点之间的有向连接表示随机变量之间的条件因果关系,指向节点 A 的节点称为节点 A 的父节点,而节点 A 称为其父节点的子节点^[23]。每个节点都对应一个条件概率,表示该变量与父节点之间的关系强度,没有父节点的用先验概率进行信息表达^[24]。贝叶斯网络通过概率关系来描述变量之间的不确定性,根据贝叶斯理论,联合概

率可表示为^[25]

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | P_a(X_i)) \quad (1)$$

式中: $P_a(X_i)$ 为 X_i 父节点的先验概率(如果 X_i 不存在父节点,那么 $P_a(X_i)$ 即为 $P(x_i)$); n 为节点数量。

贝叶斯网络被广泛应用于工程问题中^[26-27],堰塞坝稳定性问题虽然属于确定性事件,但短期内是否发生溃决属于不确定性概率问题,且堰塞湖多发生于深山中,基础信息短期收集困难,而贝叶斯网络有效地结合了概率论和图论,可以从不完全或不确定的信息中做出推理,适用于表达和分析不确定事件和概率性事件。

对于贝叶斯网络模型的建立,可以分为以下步骤:①网络节点的确定。可以从目标节点入手,通过因果关系和专家经验得到观测节点和中间节点。②网络结构的确定。贝叶斯网络结构的构建是一个复杂的过程,需要对某个领域反复研究和不断完善,最终确定最合理的网络结构。③节点值的划分。每个节点都有多个阈值区间,即不同的节点状态,可以是连续的或是离散的,可以对输入数据进行定性判定。④节点概率的确定。确定节点的条件概率的方法有两种:一是基于样本的数据,通过训练集的样本学习得到节点的概率分布,这种方法简单且准确,但需要有一定比例的测试集进行模型验证;二是基于专家经验,专家通过以往经验进行概率确定,这种方法具有一定的主观性,且准确性难以保障。

本文将贝叶斯网络应用于堰塞坝稳定性分析,首先确定稳定性影响参数,根据因果关系采用专家经验法建立堰塞坝稳定性评价网络的结构,并确定贝叶斯网络各节点及其状态分布;然后借助贝叶斯建模软件 Netica 进行数据学习得到各节点的条件概率,从而得到完整的贝叶斯网络模型;最后以数据库测试集对模型进行准确性验证,并将模型与国内外评价方法进行对比。

2 堰塞坝稳定性贝叶斯网络

2.1 堰塞坝稳定性影响参数选取

堰塞坝形成后一般会在短期内发生溃决,且多发于偏远山区,不易发现,因此堰塞坝稳定性影响参数在短期内可获取尤为关键。近年来,国内外学者基于已溃和未溃堰塞坝资料,采用统计学方法提出数学表达式和判别准则。表 1 给出了国内外常用堰塞坝稳定性评价模型,通过分析发现目前国内外学者大多基于堰塞坝地貌学指标、物质组成和堰塞湖的水动力条件对堰塞坝的稳定性进行快速评价。但此类评价模型太依赖于参数的完整性,若缺失某参

表 1 典型的堰塞坝稳定性快速评价模型

模型名称	来源文献	案例数/个	数学表达式	判别准则
I_{BI} 模型	Casagli 等 ^[11]	70	$I_{BI} = \lg(V_d/10^6 A_b)$	$I_{BI} > 5$ 为稳定区; $4 < I_{BI} \leq 5$ 为不确定区; $3 \leq I_{BI} < 4$ 为不稳定区
I_{DBI} 模型	Ermini 等 ^[17]	84	$I_{DBI} = \lg(A_b H_d/V_d)$	$I_{DBI} \leq 2.75$ 为稳定区; $2.75 < I_{DBI} \leq 3.08$ 为不确定区; $I_{DBI} > 3.08$ 为不稳定区
I_s 模型	Korup ^[18]	83	$I_s = \lg(H_d^3/V_1)$	$I_s > 0$ 为稳定区; $-3 < I_s \leq 0$ 为不确定区; $I_s \leq -3$ 为不稳定区
I_a 模型		110	$I_a = \lg(H_d^2/A_b)$	$I_a > 3$ 为稳定区; $I_a \leq 3$ 为不稳定区
I_r 模型		108	$I_r = \lg(H_d/H_r)$	$I_r > -1$ 为稳定区; $I_r \leq -1$ 为不稳定区
I_{L1} 模型	Dong 等 ^[19]	43	$I_{L1} = -2.551\lg Q - 3.641\lg H_d + 2.991\lg W_d + 2.731\lg L_d - 3.87$	$I_{L1} > 0$ 为稳定区; $I_{L1} \leq 0$ 为不稳定区
I_{L2} 模型		43	$I_{L2} = -2.221\lg A_b - 3.761\lg H_d + 3.171\lg W_d + 2.851\lg L_d + 5.93$	$I_{L2} > 0$ 为稳定区; $I_{L2} \leq 0$ 为不稳定区
I_{L3} 模型		84	$I_{L3} = -4.481\lg A_b - 9.311\lg H_d + 6.611\lg(10^6 V_d) + 6.39$	$I_{L3} > 0$ 为稳定区; $I_{L3} \leq 0$ 为不稳定区
I_{L4} 模型	Shan 等 ^[21]	27	$I_{L4} = -0.2641\lg I + 1.1661\lg V_d - 1.551\lg A_b - 0.168\lg S_d - 4.847$	$I_{L4} > 0$ 为稳定区; $I_{L4} \leq 0$ 为不稳定区
I_{L5} 模型		115	$I_{L5} = -0.1981\lg I + 1.3871\lg V_d - 1.4321\lg A_b + 4.169M_1 - 8.674$	$I_{L5} > 0$ 为稳定区; $I_{L5} \leq 0$ 为不稳定区

注: I_{BI} 为堆积体指标; V_d 为堰塞坝体积; A_b 为堰塞湖流域面积; I_{DBI} 为无量纲堆积体指标; H_d 为堰塞坝高度; I_s 为蓄水量指标; I_a 为流域面积指标; I_r 为水力梯度指标; V_1 为堰塞湖库容; H_r 为堰塞湖流域最高点与堰塞坝底部高程的差值; I_{L1} 、 I_{L2} 、 I_{L3} 、 I_{L4} 、 I_{L5} 均为逻辑回归指标; Q 为堰塞湖上游来流峰值流量; W_d 为堰塞坝宽度; L_d 为堰塞坝长度; S_d 为颗粒组成指标; I 为堰塞体最大水力梯度, 即 H_d/W_d ; M_1 为颗粒特征参数, M_1 的取值方法见文献[21]。

数便无法进行稳定性评价,且大多数模型仅考虑较少因素,导致判别准确性难以保证。贝叶斯模型可以考虑较多因素不同程度的影响,且可以在参数缺失的情况下进行快速评价,综合考虑了堰塞坝高度、长度、宽度、体积、物质组成,以及堰塞湖流域面积、库容和最大水力梯度 8 个影响参数。本次稳定性评价研究选取的状态是堰塞坝形成后未溃决并保持相对稳定来流时的情况,并未考虑暴雨、余震以及山体二次崩塌等极端情况下的影响。由于堰塞湖不存在绝对稳定的状态,因此评价结果适用于堰塞湖形成后的短期评估。

2.2 网络结构确定

网络结构的确定一般依据专家经验、机器学习,以及专家经验和机器学习相结合 3 种方法,但在节点参数较多且没有合适算法的前提下,机器学习的效率较低且建立的网络结构精确性不高^[28]。本文在前人经验的基础上,采用专家经验法建立网络模型,从堰塞坝形态特征、物质组成和堰塞湖水动力条件 3 个方面分析因果关系,建立堰塞坝稳定性贝叶斯网络结构(图 1)。

a. 堰塞坝形态特征。堰塞坝高度和堰塞湖库容是堰塞坝稳定性的重要影响因素^[20]。堰塞坝宽度(指顺河向长度)影响最大水力梯度,宽度越大则漫顶过流水体的动力势能消减的越多^[16];最大水力梯度是指堰塞坝最大水头(坝高)与坝宽的比值。当堰塞坝上、下游水位恒定时,堰塞坝长度(指堰塞坝横河向最大长度)越小,其受两岸山体约束越大,

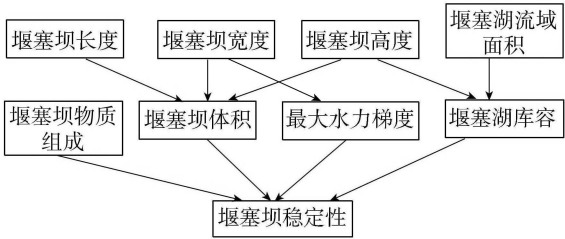


图 1 堰塞坝稳定性贝叶斯网络结构

稳定性相对越好^[16];而堰塞坝体积 V_d 影响其自重,由自身的长、宽、高决定^[21]。

b. 堰塞坝物质组成。堰塞坝物质组成决定了自身抵抗水流冲蚀的能力,一般而言,堰塞坝组成物质的颗粒粒径越大,抗冲蚀能力越强。堰塞坝物质组成用土与块石的比例表征,将土体质量分数超过 70% 的堰塞坝称为土质型堰塞坝,块石质量分数超过 70% 的称为块石型堰塞坝,介于两者之间的称为混合型堰塞坝^[29]。

c. 堰塞湖水动力条件。堰塞湖库容决定其蓄水能力,蓄水过程即水势能的积累过程,主要受限于坝高,对于堰塞坝稳定与否起着决定性作用^[18,30];流域面积决定了上游来流量和水能^[17,21],来流量直接影响堰塞湖蓄水速度和水流冲蚀力。堰塞湖来流量和库容共同影响漫坝时水流的冲刷力大小、持续冲刷的历时以及冲刷洪量,是致使堰塞坝发生溃决的外在使动因素^[16]。

2.3 网络节点状态确定

通过上述对堰塞坝稳定性影响因素的分析,选

用堰塞坝高度、长度、宽度和体积、最大水力梯度表征堰塞坝的形态特征,选用堰塞湖库容和流域面积表征堰塞湖水动力条件,并综合考虑堰塞坝的物质组成。Liao 等^[30]对堰塞坝高度、长度、宽度以及堰塞湖库容做出了 4 个等级的定量界定;SL 450—2009《堰塞湖风险等级划分标准》将堰塞坝根据最大水力梯度大小分为低危险、中危险、高危险、极高危险 4 个等级;程根伟等^[31]根据堰塞湖上游流域面积将堰塞坝分为无风险、近期低风险和高风险 3 个等级,参考前人研究成果确定各节点状态和阈值区间,见表 2。

表 2 堰塞坝稳定性网络节点状态及阈值

节点名称	符号	状态	阈值
堰塞坝高度	H_d	低	(0,15]
		中	(15,30]
		高	(30,70]
		极高	(70,∞)
堰塞坝长度	L_d	短	(0,160]
		中	(160,300]
		长	(300,500]
		极长	(500,∞)
堰塞坝宽度	W_d	窄	(0,300]
		中	(300,500]
		宽	(500,900]
		极宽	(900,∞)
堰塞坝体积	V_d	小	(0,2]
		中	(2,5]
		大	(5,10]
		巨大	(10,∞)
最大水力梯度	I	小	(0,0.05]
		中	(0.05,0.1]
		大	(0.1,0.2]
		巨大	(0.2,∞)
堰塞湖库容	V_l	小	(0,1]
		中	(1,10]
		大	(10,100]
		巨大	(100,∞)
流域面积	A_b	小	(0,200]
		中	(200,400]
		大	(400,1000]
		巨大	(1000,∞)
堰塞坝物质组成	M_c	土质型	土质型
		混合型	混合型
		块石型	块石型
堰塞坝稳定性	Y	稳定	稳定
		不稳定	不稳定

表 3 313 个国内外堰塞坝案例信息统计

参数	案例数/个	缺失率/%	最大值	最小值	平均值	中位值	方差	偏态系数	峰态系数
H_d	313	0	800 m	1 m	42.2 m	25 m	60.8	7.1	78.1
W_d	313	0	2700 m	5 m	320.3 m	200 m	327.5	3.0	14.8
L_d	313	0	3800 m	5 m	604.2 m	400 m	640.7	2.4	6.9
M_c	313	0	0.8	0	0.1	0.1	0.1	3.1	10.7
V_d	252	19.49	13 亿 m ³	0	1 560 万 m ³	170 万 m ³	89.0	12.4	173.5
V_l	171	45.37	160 亿 m ³	0	15 280 万 m ³	270 万 m ³	1 241.6	12.3	157.0
A_b	189	39.62	17 3484 km ²	0.5 km ²	2 221.8 km ²	53 km ²	17 768.8	9.6	90.7

3 贝叶斯网络节点量化

贝叶斯网络的量化就是找出基本节点的概率和各节点之间相互关系的条件概率,主要量化方法有:统计数据、物理模型或经验模型、经验判断和逻辑推理。稳定性评价模型中的条件概率表通过样本数据的学习得到,选用中国、意大利和日本等国家共 313 个堰塞坝案例^[1,6,11,17,19,21,32-33]作为案例数据库(篇幅有限,数据省略)。但由于各种因素的限制,案例数据并不完整,其中 186 个案例不稳定,127 个案例稳定,详见表 3。根据参数数据的完整程度分别建立了贝叶斯简化模型和精细化模型,其中完整数据案例是指每个参数节点的数据都完整,存在 1 个或者多个空缺数据的案例称为不完整数据案例。

3.1 贝叶斯简化模型

从 313 个堰塞坝案例中选取 248 个案例作为训练集进行数据学习,剩下的 65 个案例作为测试集以评价模型准确性。训练集中堰塞坝体积、堰塞湖库容和流域面积的缺失率分别为 17.34%、41.94% 和 38.71%。由于学习案例中存在不完整数据案例,因此采用期望极大算法^[25]来处理缺失数据情况下的概率计算。这里以堰塞坝稳定性 Y 为观测变量,观测的结果是稳定或者不稳定,概率分布为 $P(Y| \theta)$,其中 $\theta = \{H_d, W_d, L_d, I, V_d, V_l, A_b, M_c\}$,表示模型参数;缺失的节点数据 Z 为隐变量,表示未观测到的某节点的结果,假设 Y 和 Z 的联合概率为 $P(Y, Z| \theta)$,条件概率分布也为 $P(Y, Z| \theta)$;首先以案例中计数概率作为参数的初值 $\theta_{(0)} = \{(H_{d(0)}, W_{d(0)}, L_{d(0)}, I_{(0)}, V_{d(0)}, V_{l(0)}, A_{b(0)}, M_{c(0)})\}$ 开始迭代,记 θ_i 为第 i 次迭代参数的 θ 估计值,计算 $Q(\theta, \theta_i)$ 函数^[25]:

$$Q(\theta, \theta_i) = \sum_Z \lg P(Y, Z | \theta) P(Z | Y, \theta_i) \quad (2)$$

式中: $P(Z|Y, \theta_i)$ 为在给定观测数据 Y 和当前参数估计 θ_i 下隐变量数据 Z 的条件概率分布。接着求使 $Q(\theta, \theta_i)$ 极大化的 θ ,确定第 $i+1$ 次迭代的参数估计值

$$\theta_{i+1} = \arg \max_{\theta} Q(\theta, \theta_i) \quad (3)$$

当 $\| \theta_{i+1} - \theta_i \| < \varepsilon$ 或 $\| Q(\theta_{i+1}, \theta_i) - Q(\theta_i, \theta_i) \| < \varepsilon$ (ε 为迭代收敛条件) 时停止迭代, 并输出此时的 θ^* 作为输出的网络条件概率分布, 否则置 $i=i+1$ 继续迭代。将数据库案例输入 Netica 中进行数据学习并设 $\varepsilon=0.01\%$, 直到找到最大似然贝叶斯网络, 最终共迭代 213 次得到最终的先验(条件)概率分布情况, 即贝叶斯简化模型(图 2), 模型可以通过输入参数指标, 得到实时的堰塞坝的稳定情况。模型中每个节点都有一个概率表, 其中各节点概率表最上端的是各节点的名称(影响参数); 表中间部分是节点的状态分类及其概率分布情况; 图中箭头表示的是各节点之间的父子因果关系; 其中初始节点表中概率分布为先验概率分布; 中间节点表中概率分布为迭代得到的条件概率分布; 目标节点用来评价堰塞坝的稳定性。

3.2 贝叶斯精细化模型

简化模型的训练中存在大量不完整数据, 为了多角度考虑贝叶斯方法的实用性和可行性, 选用数据库中 82 个信息完整堰塞坝案例作为案例集建立精细化模型进行对比验证, 将其中 56 个案例作为训练集进行数据学习, 剩下的 26 个案例作为测试集以评价模型准确性。需要说明的是, 由于完整数据有限, 为了充分量化各节点状态, 对中间节点(堰塞坝体积、最大水力梯度、堰塞湖库容)的状态进行调整, 将表 2、图 2 中的“中”和“大”两个状态合并为“大”, 见表 4。

由于完整数据案例学习中没有隐变量的存在, 在此采用共轭梯度下降法^[34]迭代计算各节点的概率值, 梯度下降学习以负对数似然函数为目标函数, 搜索贝叶斯网络参数空间。给定一个初始的贝叶斯网络, 它可以通过自主学习来计算最陡梯度的方向, 进而搜索更具优势的贝叶斯网络, 直至得到最佳的贝叶斯模型。以目标节点的条件概率作为目标函数 $f(\theta)$, 其对应梯度函数 $g(\theta)=\nabla f(\theta)$, 首先以案例中

表 4 更新后的堰塞坝稳定性网络节点状态及阈值

节点名称	符号	状态	阈值
堰塞坝体积	V_d	小	(0,2]
		大	(2,10]
		巨大	(10,∞)
水力梯度	I	小	(0,0.05]
		大	(0.05,0.2]
		巨大	(0.2,∞)
堰塞湖库容	V_l	小	(0,1]
		大	(1,100]
		巨大	(100,∞)

计数概率作为参数的初值 $\theta_{(0)} = \{H_{d(0)}, W_{d(0)}, L_{d(0)}, I_{(0)}, V_{d(0)}, V_{l(0)}, A_{b(0)}, M_{c(0)}\}$ 开始迭代, 取函数搜索方向 $p_i = -\nabla f(\theta_i)$, 求步长 λ_i , 使

$$f(\theta_i + \lambda_i p_i) = \min_{\lambda \geq 0} f(\theta_i + \lambda p_i) \tag{4}$$

令 $\theta_{i+1} = \theta_i + \lambda_i p_i$, 并计算此时的目标函数 $f(\theta_{i+1})$, 当 $\| \theta_{i+1} - \theta_i \| < \varepsilon$ 或 $\| f(\theta_{i+1}) - f(\theta_i) \| \leq \varepsilon$ 时停止迭代, 并输出此时的 θ^* 作为输出的网络条件概率分布, 否则令 $i=i+1$ 继续迭代。将 56 个完整案例输入 Netica 中并设置 $\varepsilon=0.01\%$, 经过 10 次迭代后, 得到贝叶斯精细化模型(图 3)。

至此, 堰塞坝稳定性分析模型已经建立完成, 针对不同类型的数据案例, 分别采用期望极大算法和梯度下降法进行模型构建。相比于其他稳定性评价方法, 贝叶斯网络模型可以通过实时输入和更新样本集, 实现节点的动态学习, 随着案例数据的增加和相关知识的学习, 可以获得实时的节点概率变化情况, 从而实现模型的实时更新; 除此之外, 贝叶斯模型还可以处理缺失数据(隐变量), 即在给定新案例时, 可以为未知变量寻找合适的节点状态, 使得模型更为精确。

4 模型验证与对比

4.1 模型测试集验证

以数据库中剩余 65 个案例作为测试集对贝

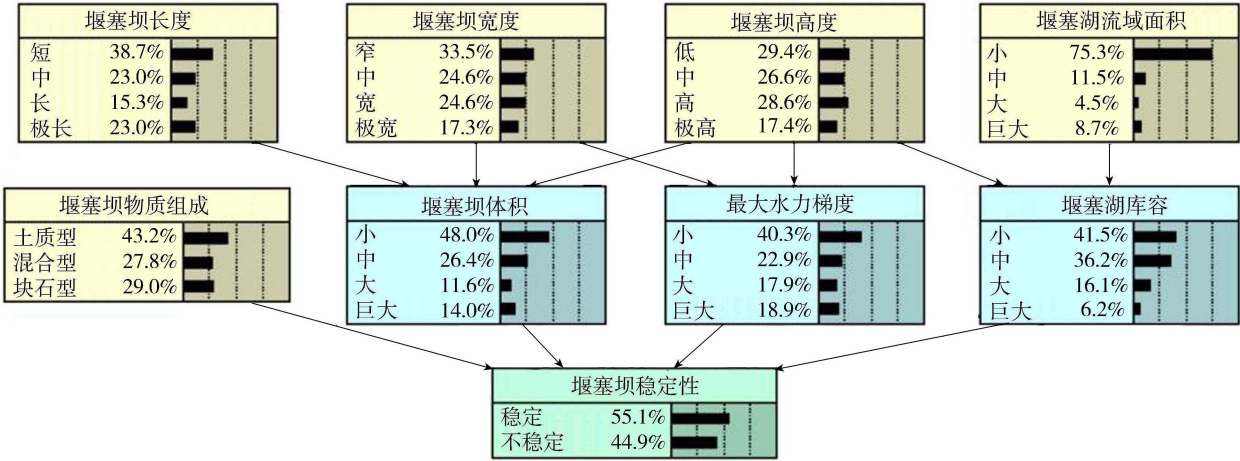


图 2 堰塞坝稳定性评价贝叶斯简化模型

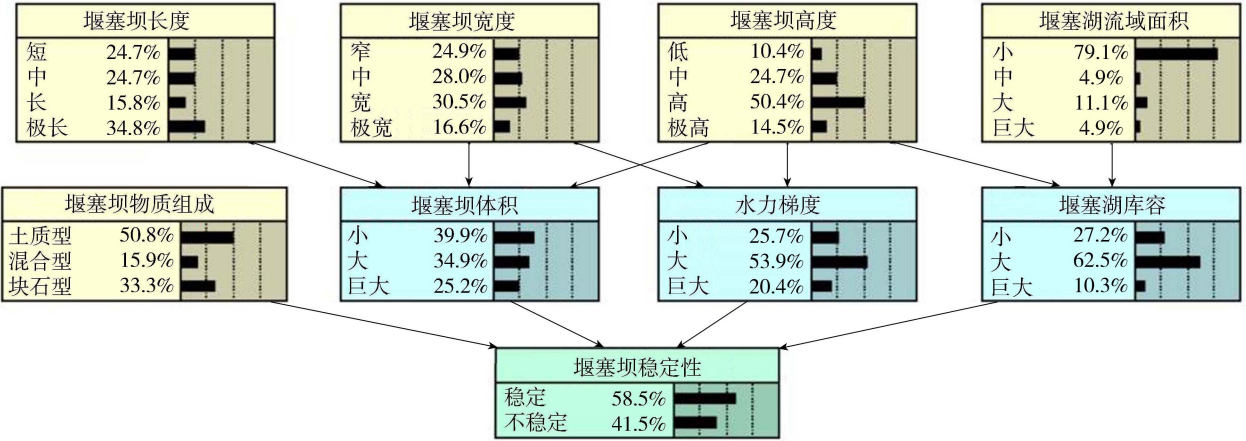


图3 堰塞坝稳定性评价贝叶斯精细化模型

叶斯简化模型进行准确性验证。为了验证模型评价的合理性,采用绝对准确率、保守准确率对模型进行分析。其中,绝对准确率是指模型判断结果和案例实际结果一致的概率;保守准确率指的是模型判断结果为不稳定但案例实际结果为稳定的概率与绝对准确率的和。由于贝叶斯模型得到是概率性结果且目标节点(堰塞坝稳定性)仅有两个状态,因此以概率更大的状态作为结果输出(如不稳定的概率大于50%即以该堰塞坝不稳定作为结果输出)。将测试集案例输入模型并通过 Netica 计算分析发现,输出结果与真实结果相同的有 53 例,因此精细化模型的绝对准确率为 81.54%;输出结果为不稳定但真实结果为稳定的有 9 例,因此模型保守正确率为 95.38%。

为了验证贝叶斯精细化模型的准确性,以完整数据集案例集中剩余 26 个堰塞坝案例作为模型测试集进行验证分析,输出结果与真实结果相同的有 21 例,因此精细化模型的绝对准确率为 80.77%;输出结果为不稳定但真实结果为稳定的有 4 例,因此模型保守正确率为 96.15%,较简化模型有略微的变化,但由于训练案例的有限性(精细化模型训练集不到简化模型训练集的1/4),精细化模型的稳定性不如简化模型,仍需后续训练案例的补充。随着训练集的不断增加,模型的准确性将会越来越高。

4.2 白格堰塞坝案例验证

白格堰塞坝位于金沙江上游右岸西藏自治区江达县波罗乡白格村境内,2018 年 11 月 3 日,在第一次滑坡的滑源区后缘岩土体再次发生失稳破坏,并再次堵塞金沙江形成堰塞湖,溃决后将下游生命财产安全造成严重威胁。白格“11·3”滑坡体堆积物平面上呈舌状,主要堆积在上一次堰塞体上,靠近金沙江右岸。堰塞坝高约 86 m、长约 600 m、宽约 1200 m,体积可达 283 万 m³,其主要由土、碎石和块石共同组成,内部结构密实;堰塞湖蓄满库容可达

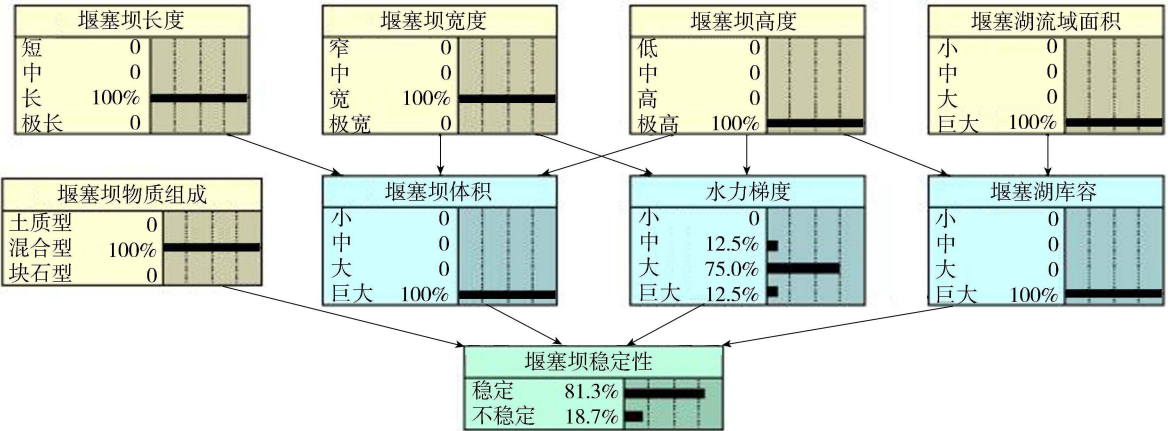
7.9 亿 m³,上游汇水面积约为 17 万 km²[35]。将相关参数分别输入稳定性评价简化、精细化贝叶斯网络中,可得到白格堰塞坝溃决概率为 81.3% 和 99.6% (图 4),坝体极易溃坏,因此亟须进行工程处置,并迅速进行下游风险人群紧急撤离,模型的高溃坝率与实际情况相符,也体现了贝叶斯网络方法的可行性。

4.3 国内外模型对比

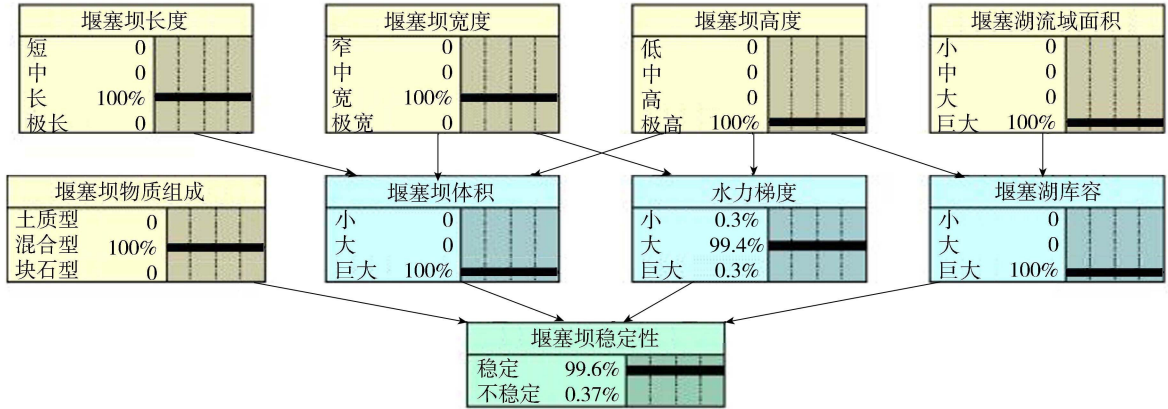
为了进一步分析贝叶斯模型的先进性和适用性,选择国内外典型的堰塞坝稳定性评价模型进行对比分析,由于选用参数不同,因此选择简化模型测试集中共 18 个堰塞坝案例(表 5 中 1~18 号案例),将贝叶斯简化模型与其他典型模型的计算结果进行对比,见表 6。选择以完整数据集案例集中共 26 个堰塞坝案例(表 5 中所有案例),将贝叶斯精细化模型与其他典型模型的计算结果进行对比,见表 7。

从表 6 可以看出, I_{B1} 模型、 I_{DB1} 模型和 I_s 模型的计算结果都存在不确定组数,影响其准确率,尤其是 I_s 模型对绝大部分案例都无法给出确定性判断。贝叶斯简化模型绝对准确率和保守准确率分别为 83.33% 和 94.44%,与国内外典型稳定性评价模型相比绝对准确率最高,保守准确率仅次于 I_a 模型。从表 7 可以看出,贝叶斯精细化模型绝对准确率和保守准确率为 80.77% 和 96.15%,从绝对准确率上看,最好的是 I_{15} 模型, I_{12} 模型和精细化模型次之,较差的是 I_s 模型;从保守准确率来看,精细化模型仅次于 I_a 模型。

从准确率可以看出两种贝叶斯模型的计算结果都是偏保守的,这样的结果对应急响应更稳妥。综合来看,贝叶斯网络评价模型判别准确率较高,由于选用的训练集案例较少,而贝叶斯网络具有独特的优势就是模型的开放性,在原有的模型基础上添加新的数据后,训练集可以实时更新,模型的判别也将更加准确。



(a) 贝叶斯简化模型



(b) 贝叶斯精细化模型

图 4 白格堰塞坝溃决概率分析

表 5 26 个国内外典型堰塞坝案例

序号	国家	名称	高度/m	长度/m	宽度/m	最大水力 梯度	堰塞坝体 积/10 ⁶ m ³	堰塞湖库 容/10 ⁶ m ³	流域面 积/km ²	物质组成	稳定性
1	中国	石板沟	52.5	450	800	0.07	15	11	3 546	块石型	不稳定
2	中国	孙家院子	50	180	400	0.13	1.6	5.6	1 785	混合型	不稳定
3	中国	易贡	110	2 500	2 500	0.04	300	3 000	13 533	混合型	不稳定
4	中国	燕子岩	10	40	20	0.50	0.01	0.03	18	块石型	不稳定
5	中国	肖家桥	62	200	200	0.31	2.42	20	231	混合型	不稳定
6	中国	肖家沟	15	500	100	0.15	0.75	0.23	7.23	土质型	不稳定
7	中国	小岗剑下游	30	120	400	0.08	0.45	7	378	混合型	不稳定
8	中国	小岗剑上游	95	300	300	0.32	2	12	376	块石型	不稳定
9	中国	红松	50	100	60	0.83	0.26	1	10	土质型	不稳定
10	印度	Birehi Ganga	274	760	2 750	0.10	286	460	253	混合型	不稳定
11	意大利	Valderchia	9	110	160	0.06	0.1	0.005 95	4.5	土质型	不稳定
12	意大利	Val Vanoi	40	500	1 000	0.04	10	18.2	167	土质型	不稳定
13	意大利	Tovel	45	1 300	1 700	0.03	40	7.37	40.4	块石型	稳定
14	意大利	Schiazzano	15	40	65	0.23	0.02	0.008 83	5.6	土质型	不稳定
15	意大利	Scanno	33.1	500	2 000	0.02	17	26	95	块石型	稳定
16	意大利	Sernio	43	300	930	0.05	2	22	891	土质型	不稳定
17	意大利	Rovina	15	400	900	0.02	2	1.2	17.2	混合型	稳定
18	意大利	Ronchi	20	160	190	0.11	0.3	0.471	24.5	土质型	稳定
19	中国	新街村	20	350	200	0.10	0.7	2	3 546	土质型	不稳定
20	中国	唐家湾	30	300	600	0.05	4	15.2	1 395	土质型	不稳定
21	中国	唐家山	120	612	802	0.15	20.37	316	3 550	土质型	不稳定
22	意大利	Tramarecchia	20	200	450	0.04	1.5	0.58	40.8	混合型	稳定
23	意大利	Tenno	50	900	650	0.08	10	5	19.3	土质型	稳定
24	意大利	Signatico	30	450	620	0.05	8.37	8	151.5	土质型	不稳定
25	意大利	Prato Casarile	40	200	450	0.09	1.75	0.3	1.5	土质型	稳定
26	日本	Abe River	30	500	650	0.05	4.88	4.7	19	块石型	稳定

表 6 贝叶斯简化模型与其他典型模型计算结果对比							
模型	案例数/个	判错组数/个	准确组数/个	保守组数/个	不确定组数/个	绝对准确率/%	保守准确率/%
I_{BI} 模型	18	2	9	1	6	50.00	55.56
I_{DBI} 模型	18	1	11	4	2	61.11	83.33
I_s 模型	18	0	1	0	17	5.56	5.56
I_a 模型	18	0	14	4	0	77.78	100.00
I_{12} 模型	18	2	15	1	0	83.33	88.89
I_{13} 模型	18	4	13	1	0	72.22	77.78
I_{15} 模型	18	2	15	1	0	83.33	88.89
贝叶斯简化模型	18	1	15	2	0	83.33	94.44

表 7 贝叶斯精细化模型与其他典型模型计算结果对比							
模型	案例数/个	判错组数/个	准确组数/个	保守组数/个	不确定组数/个	绝对准确率/%	保守准确率/%
I_{BI} 模型	26	2	14	2	8	53.84	61.54
I_{DBI} 模型	26	2	14	8	2	53.85	84.62
I_s 模型	26	0	1	0	25	3.85	3.85
I_a 模型	26	0	18	8	0	69.23	100.00
I_{12} 模型	26	2	22	2	0	84.62	92.31
I_{13} 模型	26	5	20	1	0	76.92	80.77
I_{15} 模型	26	2	23	1	0	88.46	92.31
贝叶斯精细化模型	26	1	19	4	0	80.77	96.15

5 结 语

收集国内外 313 个已溃和未溃堰塞坝案例,并考虑堰塞坝的形态特征、物质组成和堰塞湖水动力条件,结合贝叶斯理论提出了节点缺失条件下的简化模型和节点完整条件下的精细化模型。采用国内外已溃和未溃堰塞坝案例对两种模型进行验证,并以白格堰塞坝作为实例进行溃决概率分析。然后与国内外典型的堰塞坝稳定性快速评价模型进行对比,从各模型的计算结果看出,基于贝叶斯网络的简化模型和精细化模型表现优异,表明本文提出的基于贝叶斯网络的堰塞坝稳定性快速评价模型的可靠性和科学性。此外,贝叶斯网络模型可以实时输入和更新样本集,即实现节点的动态学习;还可以很好地处理缺失数据情况的案例,从而实现广泛数据学习。随着案例的不断丰富,模型的预测精度将不断提升,因此具有良好的发展前景。

参考文献:

[1] COSTAJ E,SCHUSTER R L. The formation and failure of natural dams [J]. GSA Bulletin, 1988, 100 (7) : 1054-1068.

[2] 杜镇瀚,钟启明,董海洲,等. 堰塞湖风险评估研究综述 [J]. 河海大学学报(自然科学版),2022,50(4) :13-25. (DU Zhenhan, ZHONG Qiming, DONG Haizhou, et al. A review of risk assessment studies on dammed lake [J]. Journal of Hohai University (Natural Sciences),2022,50

(4) :13-25. (in Chinese))

[3] 钟启明,陈生水,邓翌. 堰塞坝漫顶溃决机理与溃坝过程模拟 [J]. 中国科学: 技术科学,2018,48 (9) : 959-968. (ZHONG Qiming, CHEN Shengshui, DENG Zhao. Breach mechanism and numerical modeling of barrier dam due to overtopping failure [J]. Sclctia Sinica: Technologica,2018,48(9) :959-968. (in Chinese))

[4] 刘宁,杨启贵,陈祖煜. 堰塞湖风险处置 [M]. 武汉:长江出版社,2016.

[5] 罗辉,邓创,赵高文. 2008—2017 年中国典型滑坡堰塞坝(湖)灾害事件统计与初步分析 [J]. 水利水电科技进展,2020,40 (1) : 17-24. (LUO Hui, DENG Chuang, ZHAO Gaowen. Statistics and preliminary analysis for typical landslide dams in China in past decade [J]. Advances in Science and Technology of Water Resources, 2020,40(1) :17-24. (in Chinese))

[6] CUI Peng, ZHU Yingyan, HAN Yongshun, et al. The 12 May Wenchuan earthquake-induced landslide lakes: distribution and preliminary risk evaluation [J]. Landslides,2009,6(7) :209-223.

[7] 李相南,陈祖煜. 两种溃坝模型在唐家山堰塞湖溃决模拟中的对比 [J]. 水利水电科技进展,2017,37(2) :20-26. (LI Xiangnan, CHEN Zuyu. Comparison of two models for back analysis of dam breach of Tangjiashan Barrier Lake [J]. Advances in Science and Technology of Water Resources,2017,37(2) :20-26. (in Chinese))

[8] LIU Ning, ZHANG Jianxin, LIN Wei, et al. Draining Tangjiashan barrier lake after Wenchuan earthquake and the flood propagation after the dam break [J]. Science in China Series E: Technological Sciences, 2009, 52 (4) : 801-809.

[9] SHEN Danyi, SHI Zhenming, PENG Ming, et al. Longevity analysis of landslide dams [J]. Landslides, 2020, 17 (8) : 1797-1821.

[10] 任强,陈生水,钟启明,等. 堰塞坝的形成机理与溃决风险 [J]. 水利水电科技进展,2011,31(5) :30-34. (REN Qiang, CHEN Shengshui, ZHONG Qiming, et al. Formation mechanism and breaching failure risk of barrier dams [J]. Advances in Science and Technology of Water Resources, 2011,31(5) :30-34. (in Chinese))

[11] CASAGLI N, ERMINI L. Geomorphc analysis of landslide dams in the Northern Apennine [J]. Transactions of the Japanese Geomorphological Union,1999,20(3) :219-249.

[12] 刘怀湘,王兆印,刘乐. 河床结构对堰塞坝稳定性的影响研究 [J]. 水力发电学报,2011,30(3) :98-103. (LIU Huaixiang, WANG Zhaoyin, LIU Le. Study on the influence of riverbed structures on the landslide dam stability [J]. Journal of Hydroelectric Engineering,2011,30(3) :98-103. (in Chinese))

[13] 石振明,刘思言,彭铭. 堰塞坝坝体材料渗流特性研究现状及展望 [C] //《工程地质学报》编辑部. 2014 年全

- 国工程地质学术大会论文集. 北京: 科学出版社, 2014: 97-102.
- [14] 彭铭, 蒋明子. 滑坡涌浪作用下堰塞坝稳定性研究 [C]//中国地球物理学会. 2017 中国地球科学联合学术年会论文集(十五). 北京: 中国和平音像电子出版社, 2017: 27-29.
- [15] 周亦良, 姚令侃, 艾洪舟, 等. 地震力与共振动水压力综合作用下冰碛堰塞坝失稳机制研究[J]. 岩石力学与工程学报, 2017, 36(7): 1726-1735. (ZHOU Yiliang, YAO Lingkan, AI Hongzhou, et al. Instability of moraine dams under the combined action of seismic forces and resonant hydrodynamic pressures [J]. Chinese Journal of Rock Mechanics and Engineering, 2017, 36(7): 1726-1735. (in Chinese))
- [16] 中华人民共和国水利部. 堰塞湖风险等级划分标准: SL 450—2009[S]. 北京: 中国水利水电出版社, 2009.
- [17] ERMINI L, CASAGLI N. Prediction of the behaviour of landslide dams using a geomorphological dimensionless index[J]. Earth Surface Processes and Landforms, 2003, 28(1): 31-47.
- [18] KORUP O. Geomorphometric characteristics of New Zealand landslide dams [J]. Engineering Geology, 2004, 73(1/2): 13-35.
- [19] DONG J J, TUNG Y H, CHEN C C, et al. Logistic regression model for predicting the failure probability of a landslide dam[J]. Engineering Geology, 2011, 117(1/2): 52-61.
- [20] STEFANELLI C T, SEGONI S, CASAGLI N, et al. Geomorphic indexing of landslide dams evolution [J]. Engineering Geology, 2016, 208: 1-10.
- [21] SHAN Yibo, CHEN Shengshui, ZHONG Qiming. Rapid prediction of landslide dam stability using the logistic regression method [J]. Landslides, 2020, 17(12): 2931-2956.
- [22] 石振明, 程世誉, 张清照, 等. 堰塞坝稳定性快速评估模型: 以小岗剑(上)堰塞坝为例[J]. 水利与建筑工程学报, 2020, 18(2): 95-100. (SHI Zhenming, CHENG Shiyu, ZHANG Qingzhao, et al. A fast model for landslide dams stability assessment: a case study of Xiaogangjian (upper) landslide dam [J]. Journal of Water Resources and Architectural Engineering, 2020, 18(2): 95-100. (in Chinese))
- [23] 张连文, 郭海鹏. 贝叶斯网引论[M]. 北京: 科学出版社, 2006.
- [24] CHEN Jing, FU Jingqi. Fire alarm system based on multi-sensor Bayes network[J]. Procedia Engineering, 2012, 29: 2551-2555.
- [25] 李航. 统计学习方法[M]. 2 版. 北京: 清华大学出版社, 2019.
- [26] KHAKZAD N, KHAN F, AMYOTTE P. Quantitative risk analysis of offshore drilling operations: a Bayesian approach[J]. Safety Science, 2013, 57: 108-117.
- [27] ZHU Yan, PENG Ming, ZHANG Peng, et al. Warning decision-making for landslide dam breaching flood using influence diagrams[J]. Frontiers in Earth Science, 2021, 9: 679862.
- [28] 胡笑旋, 杨善林, 马溪骏. 面向复杂问题的贝叶斯网建模方法[J]. 系统仿真学报, 2006, 18(11): 3242-3246. (HU Xiaoxuan, YANG Shanlin, MA Xijun. Bayesian networks modeling methods for complex problems [J]. Journal of System Simulation, 2006, 18(11): 3242-3246. (in Chinese))
- [29] SHAN Yibo, CHEN Shengshui, ZHONG Qiming, et al. Development of an empirical model for predicting peak breach flow of landslide dams considering material composition[J]. Landslides, 2022, 19(6): 1491-1518.
- [30] LIAO Haimei, YANG Xingguo, XU Fugang, et al. A fuzzy comprehensive method for the risk assessment of a landslide-dammed lake [J]. Environmental Earth Sciences, 2018, 77(22): 750.
- [31] 程根伟, 范继辉, 程尊兰. 四川 5·12 地震灾区滑坡堰塞坝溃决灾害风险评估[J]. 西南民族大学学报(自然科学版), 2008, 34(6): 1086-1090. (CHENG Genwei, FAN Jihui, CHENG Zunlan. Breaking risk evaluation of earthquake-induced lakes in Sichuan, China [J]. Journal of Southwest University for Nationalities (Natural Science Edition), 2008, 34(6): 1086-1090. (in Chinese))
- [32] STEFANELLI C T, CATANI F, CASAGLI N. Geomorphological investigations on landslide dams [J]. Geoenvironmental Disasters, 2015, 2(1): 21.
- [33] ZHANG Limin, PENG Ming, CHANG Dongsheng, et al. Dam failure mechanisms and risk assessment [M]. Singapore: John Wiley & Sons Singapore Pte. Ltd., 2016: 419-449.
- [34] NEAPOLITAN R E. Learning Bayesian networks [M]. Chicago: Pearson Prentice Hall, 2004.
- [35] 陈祖煜, 陈生水, 王琳, 等. 金沙江上游“11·03”白格堰塞湖溃决洪水反演分析[J]. 中国科学: 技术科学, 2020, 50(6): 763-774. (CHEN Zuyu, CHEN Shengshui, WANG Lin, et al. Back analysis of the breach flood of the “11·03” Baige barrier lake at the Upper Jinsha River [J]. Scientia Sinica: Technologica, 2020, 50(6): 763-774. (in Chinese))

(收稿日期: 2021-11-21 编辑: 骆超)