

# 基于多元回归模型的斜坡影响因子分析

章 强<sup>1,2</sup>, 董建辉<sup>3</sup>

(1.中国地质科学院探矿工艺研究所,四川 成都 611734; 2.中国地质调查局地质灾害防治技术中心,四川 成都 611734; 3.成都大学建筑与土木工程学院,四川 成都 610106)

**摘要:**目前,对斜坡影响因子进行量化的评价方法均有各自的局限性。在基于监测数据的基础上提出多元回归模型进行量化斜坡影响因子的方法。通过将确定的斜坡影响因子与回归模型的融合,在一定程度解决了斜坡稳定性评价过程中影响因子的选择和量化的问题,有利于以后对斜坡的认识,建立准确的斜坡稳定性分析模型。最后,将模型应用于某水电站水库堆积体斜坡,对该斜坡的影响因子进行量化,最后分析汶川地震前后的影响因子变化。

**关键词:**多元回归;堆积体斜坡;影响因子;时间效益;降雨;库水位

**中图分类号:**P642 **文献标识码:**A **文章编号:**1672-7428(2018)08-0102-05

**Study on Slope Influence Factors Based on Multiple Regression Model/QIN Qiang<sup>1,2</sup>, DONG Jian-hui<sup>3</sup>** (1.Institute of Exploration Technology, CAGS, Chengdu Sichuan 611734, China; 2.Technical Center for Geological Hazard Prevention and Control, CGS, Chengdu Sichuan 611734, China; 3.School of Architecture and Civil Engineering, Chengdu University, Chengdu Sichuan 610106, China)

**Abstract:** At present, the quantitative evaluation methods based on slope influence factors have their own limitations. Based on the monitoring data, multiple regression models are proposed to quantify slope influence factors. Through the fusion of the determined slope influence factors and the regression model, the problem of selecting and quantifying the influence factors in the process of slope stability evaluation is solved to a certain extent, which is beneficial to the establishment of an accurate slope stability analysis model. Finally, the model is applied in the slope of a reservoir accumulation body of a hydropower station, and the influence factors of the slope are quantified. The changes of the influence factors before and after Wenchuan earthquake are analyzed.

**Key words:** multiple regression; accumulation slope; influence factor; time efficiency; rainfall; reservoir water level

## 0 引言

自 20 世纪 80 年代以来,我国的地质灾害进入了一个新的活跃期,特别是“5.12”汶川地震以后,地质灾害频频发生,规模也越来越大,特别是滑坡灾害的存在,严重威胁着当地群众生命财产安全和阻碍当地经济的发展。目前国内研究人员积极运用各种试验和方法开展探索滑坡时空的预测、滑坡灾害的致灾因素和变形破坏机理等方面研究,此类型的研究对于滑坡灾害的预防及治理研究有着极其重要的社会意义和经济价值。

但由于滑坡灾害本身是一个复杂的系统,有着较多的不确定性因素,滑坡时空预测的研究仍是领域内学者们的重要研究方向之一。随着我国基础设施建设的高规模、高规格、高速发展,如建筑场地和道路建设、矿产开采、水利水电工程开发,这些工程建设经常遇到边坡问题逐渐演化成滑坡灾害问题,因此

对于其致灾因素也越来越关注,也越来越重视。目前运用数学理论模型进行边坡稳定性分析的文献越来越多<sup>[1-5]</sup>,对于运用多元回归分析进行滑坡预测分析也有了一定的深度和广度<sup>[6-8]</sup>,也有学者运用多元非线性回归分析对单体滑坡进行了滑动时间预测<sup>[10-12]</sup>。但是运用多元非线性回归分析对斜坡的影响因子研究相对较少,通过本文的研究做到摸清滑坡灾害的影响因素的变化,从而进一步了解滑坡的各种特性。

## 1 多元回归模型

多元回归模型是以时间序列模型为基础展开的,而统计分析理论作为时间序列模型的基础,应首先明确某个事件的自变量和因变量之间存在关联,通过寻找一个能表达自变量与因变量间的函数关系来阐述。变形监测模型中运用回归模型的案例较多,

收稿日期:2018-07-01

作者简介:章强,男,壮族,1982年生,勘查技术与工程专业,长期从事地质灾害调查及水文地质调查、地质灾害危险性评估、地质灾害治理勘查设计工作,四川省成都市郫都区现代工业港(北区)港华路139号,89760486@qq.com。

其原理是利用监测数据来明确回归方程中的自变量系数,目前常用的方法有最小二乘法,利用该方法建立的模型属于统计学模型中的“后验模型”<sup>[13]</sup>。

根据实际的监测数据建立回归模型,变形组(自变量) $(X_{1t}, X_{2t}, X_{3t}, \dots, X_{kt}; Y)$  ( $t=1, 2, 3, \dots, N; N \geq k$ )的  $N$  组观测数,  $Y$  为变形效应量。

建立线性回归方程:

$$Y = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_k X_k \\ = B_0 + \sum_{i=1}^k B_i X_i \quad (1)$$

将观测数据进行最佳拟合,求出自变量系数 $(B_0, B_1, B_2, \dots, B_k)$ ,建立自变量 $(X_1, X_2, X_3, \dots, X_k)$ 与  $Y$  变形效应量之间的数学模型,这是典型多元回归方程。在实际案例运用过程中证明,多元回归方程中大部分情况是非线性的。多元线性回归方程可以是指数、多项式等等形式,以下以指数和多项式为列。

(1) 对于指数形式的回归模型

$$\ln W = B_0 + B_1 \ln Z_1 + B_2 \ln Z_2 + \dots + B_i \ln Z_i \quad (2)$$

只要令  $Y = \ln W, X_1 = \ln Z_1, X_2 = \ln Z_2, \dots, X_i = \ln Z_i$ , 则该指数形式回归模型就转化成多元线性回归问题。

(2) 对于多项式的回归模型

$$Y = B_0 + B_1 Z_1 + B_2 Z_2 + B_3 Z_1^2 + B_4 Z_2^2 + B_5 Z_1 Z_2 \dots \quad (3)$$

只要令  $X_1 = W_1, X_2 = W_2, X_3 = W_1^2, X_4 = W_2^2, X_5 = W_1 W_2, \dots$ , 则该多项式回归模型就转化成多元线性回归问题。

经典多元回归方程的计算值和实测值的拟合程度以及预报值的精度检验指标主要有剩余平方和  $Q$ 、回归平方和  $U$ 、复相关系统  $R$  以及剩余标准差  $S$  等。

## 2 工程案例

### 2.1 影响因子选取

对于回归模型运用到斜坡中,其关键就是选择合理的影响因子和正确的数学表达统计模型。首先要对监测数据进行一定的预处理,剔除误差,再利用回归理论建立影响因子(自变量)和变形量(因变量)之间的数学统计模型,通过数学统计模型的试算,进一步得到斜坡的变形规律以及影响因子的比重。

本文以某水电站水库堆积体斜坡为例,该斜坡从 2006 年开始进行变形观测,一共埋设 7 个深部位

移监测孔,编号分别为 IN-1~IN-7,其中 IN-1 为自动化监测孔,此次暂未统计深部位移监测孔 IN-1,因此总共有 6 个变形量(因变量),详见图 1。根据对该堆积体斜坡的现场调研和相关分析,自 2006 年以后该斜坡的外部条件没有任何改变,造成堆积体斜坡产生变形的主要影响因素为库水位升降、降雨、时间效益等 3 个影响因素,统计的时间段为 2006-2012 年,其时间段经过了 2008 年“5.12”汶川地震,详见图 2、图 3。

从变形成因分析,可以将变形量(自变量)分为 3 部分:库水位分量  $\delta_H$ 、降雨分量  $\delta_J$ 、时间效益分量  $\delta_T$ ,即:

$$\delta(\delta_x, \delta_y \text{ 或 } \delta_z) = \delta_H + \delta_J + \delta_T \quad (4)$$

### 2.2 多元回归建模步骤

根据以上对堆积体斜坡的影响因子选取以及回归模型的相关叙述,整理堆积体斜坡的深部位移监测数据和监测数据的相关影响因素资料,建立一个预测精度高及最佳拟合的多元回归模型,而且相关影响因子系数对于变形监测数据则具有很强的解释性,这也是本文所探索的核心内容。多元回归模型的技术路线如图 4 所示。

(1) 整理变形监测资料。整理堆积体斜坡的变形监测数据、观测的影响因子的监测数据。

(2) 数据的可靠性检验和预处理。在对堆积体斜坡的监测体系进行日常观测中,难免会出现差错和误差等一些额外的情况,因此有必要对监测数据进行可靠性检验,主要是分析监测数据的连续性和关联性,结合实际情况选择不同的可靠度检验方法,如果发现异常点应采取相关处理措施进行预处理。

(3) 影响因子选择。根据现场调查以及资料收集,确定主要影响因子集合,对其做多元回归分析计算。

(4) 确定回归方程形式。在明确影响因子之后,选择多元线性回归方程的形式进行试算,可以是指数、多项式等等形式。

(5) 检验回归模型。回归方程形式确定之后,应对每个回归方程进行相关可靠性检验,如有不能满足检验的回归方程或组合,则立刻调整回归方程并重新进行检验。

(6) 经过回归方程的试算以及相关检验,最终确定合适的回归模型。

(7) 根据回归模型,进行相关解释。

### 2.3 多元回归建模

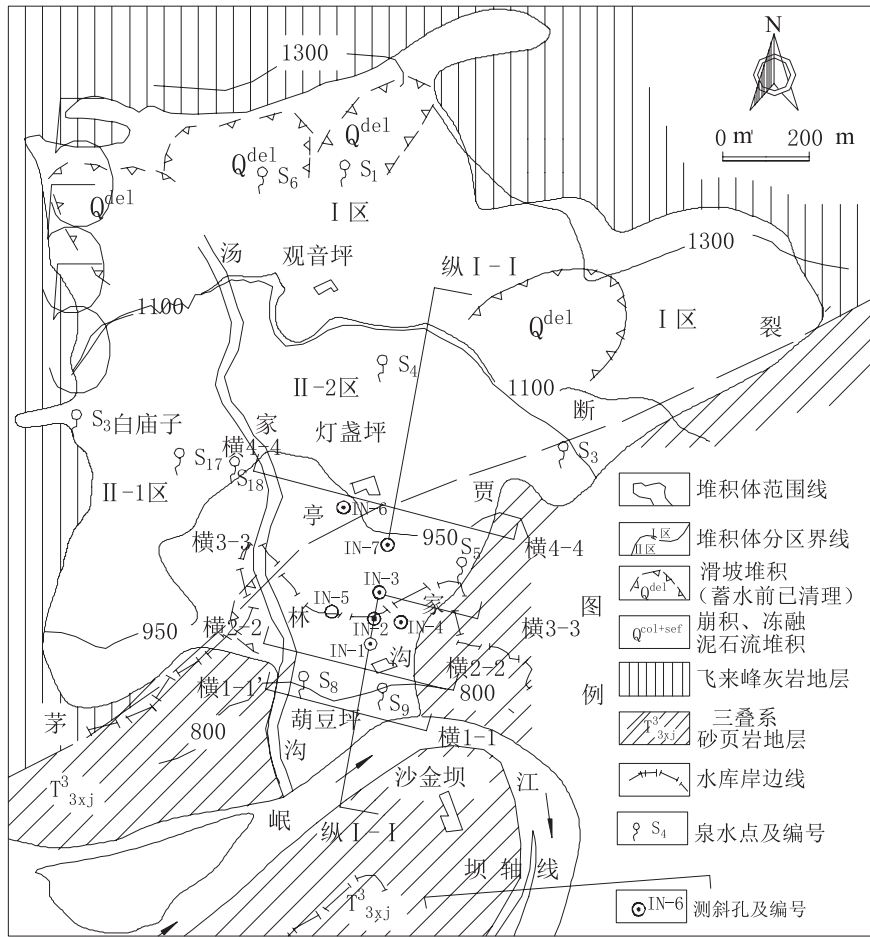


图1 某水电站水库堆积体斜坡的深部位移(测斜孔)监测平面布置图

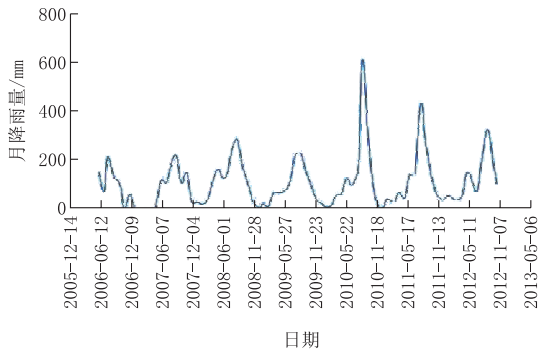


图2 月降雨量-时间变化图

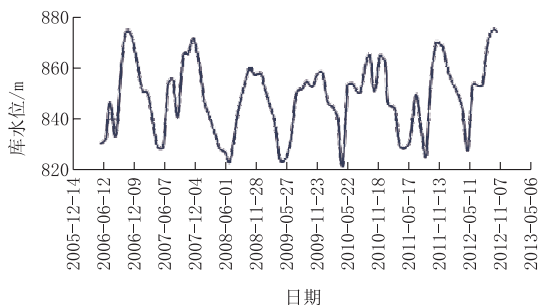


图3 库水位-时间变化图

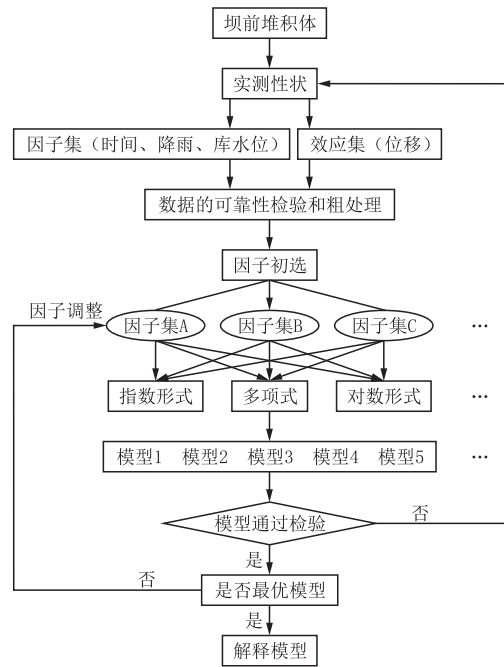


图4 多元回归模型流程图

针对某一水电站水库堆积体斜坡的 6 个深部位移监测孔数据,首先进行数据的可靠性检验和预处理分析。由于本文所用的 6 个深部位移监测孔均为人工观测孔,故地表的孔口累计位移误差较大,所以首先应对其进行平滑处理。将每年的监测数据按照自然月份分为 12 个周期,然后对其监测数据进行分段趋势预处理,结果如图 5 和图 6 所示。从观测时间段来看,可以将监测数据分为汶川地震前和汶川地震后两个时间段,即 2006 年 5 月—2008 年 4 月和 2008 年 5 月—2012 年 10 月。

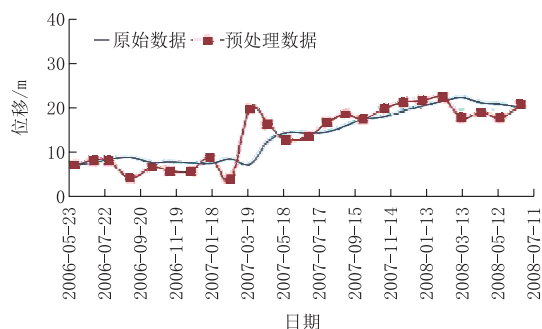


图 5 IN-2 地表位移预处理—时间关系曲线(地震前)

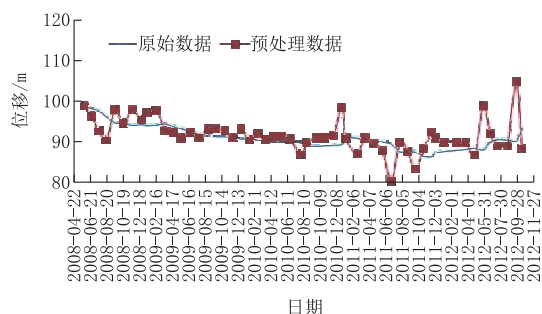


图 6 IN-2 地表位移预处理—时间关系曲线(地震后)

根据多元回归建模步骤所示,本文试算不同的回归方程形式,结果发现预测精度高及最佳拟合的多元回归模型为指数函数。回归方程表达式如下所示:

$$\ln Y = aX_{11} + bX_{22} + cX_{33} + C \quad (5)$$

根据影响因子的分析,我们可以假设:Y 为深部位移孔口的累计位移量,以每个月的监测数据作为统计数据;X<sub>11</sub> 为时间效益因子, X<sub>11</sub> = ln(1 + 0.01X<sub>1</sub>), X<sub>1</sub> 为观测次数,单位为月数;X<sub>22</sub> 为降雨量因子, X<sub>22</sub> = lnX<sub>2</sub>, X<sub>2</sub> 为每个月的降雨量;X<sub>33</sub> 为库水位因子, X<sub>33</sub> = lnX<sub>3</sub>, X<sub>3</sub> 为每个月月底的库水位。

我们采用 Eviews 专业统计软件进行相关回归方程统计以及可靠性检验工作,以 IN-2 为例,计算结果见表 1。

根据 IN-2 的回归方程的 Eviews 计算结果,回

表 1 回归方程输出结果

变 量	系 数	标准误差	t 统计值	概率(t 统计值)
C	9.651395	11.148020	0.865750	0.3969
X <sub>11</sub>	6.365613	0.463427	13.733790	0.0000
X <sub>22</sub>	0.038174	0.015496	2.463435	0.0230
X <sub>33</sub>	-1.182007	1.651612	-0.715669	0.4825
R <sup>2</sup>	0.912378	平均依赖值		2.549019
修正 R <sup>2</sup>	0.899235	风险依赖值		0.437147
回归值	0.138766	阿凯克信息标准		-0.961045
回归方程	0.385120	施瓦兹准则		-0.764702
对数似值	15.532530	汉南-奎因准则		-0.908955
F 统计值	69.417860	德宾-沃森统计值		1.182216
概率(F 统计值)	0.000000			

归模型的拟合结果如下:

$$\ln \hat{Y} = 9.651 + 6.364 \ln X_{11} + 0.0282 \ln X_{22} - 1.182 \ln X_{33} \quad (6)$$

(0.86) (13.73) (2.46) (-0.72)

结合表 1 和公式 6,其中括号内的数值表示 t 值,将根据假定的 0.05 显著性来判断 F 值,从计算结果来看回归系数的估计值,结果基本无显著性,其中模型结果的拟合度(R<sup>2</sup>)相当高,其数值为 0.91,这说明选择的模型较合理;F 值为 69.42,说明影响因子可能存在着共线性。我们利用 klein 判别法来分析影响因子之间的共线性。从表 2 所提供的影响因子之间相关系数来判断,各影响因子之间的关联系数都很小,我们认为各影响因子之间不相关。

表 2 “5.12”汶川地震前 IN-2 相关系数矩阵

	X <sub>11</sub>	X <sub>22</sub>	X <sub>33</sub>
X <sub>11</sub>	1.000000	0.125387	0.025224
X <sub>22</sub>	0.125387	1.000000	-0.207581
X <sub>33</sub>	0.025224	-0.207581	1.000000

## 2.4 回归数据分析

根据多元回归模型流程图,我们同样可以推导出其余的 5 个深部位移监测孔的孔口累计位移与各影响因子的多元回归方程,其计算结果详见表 3 和图 7~图 9。

如图 7~图 9 所示,从该水电站水库堆积体斜坡的影响因子与深部位移监测点孔口累计位移的回归方程系数关系分析出:

- (1)汶川地震后相对于汶川地震前的时间效益影响有所变小。
- (2)汶川地震后相对于汶川地震前的时间降雨量影响有所变小。



表3 汶川地震前后多元回归方程系数统计

监测点	回 归 系 数			
	C	$X_{11}$	$X_{22}$	$YX_{33}$
IN-2	9.651/5.022	6.364/-0.205	0.0282/0.00232	-1.182/-0.0701
IN-3	-44.164/4.166	9.452/-0.064	-0.579/-0.00645	6.692/0.0852
IN-4	3.968/6.099	1.701/0.333	-0.541/-0.00512	-0.218/-0.0253
IN-5	-15.586/4.030	0.817/0.205	-0.0596/-0.010	2.723/0.0737
IN-6	-17.399/5.969	6.415/0.131	-0.0841/0.00232	3.206/-0.178
IN-7	0.902/7.209	1.377/0.170	-0.00355/0.00195	0.247/-0.281

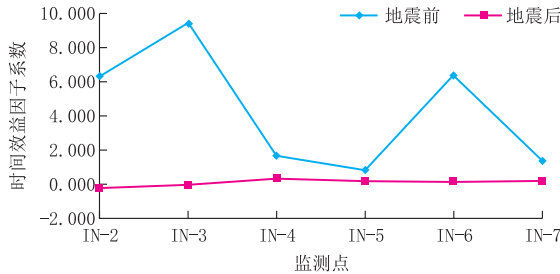


图7 各监测点时间效益因子系数曲线图

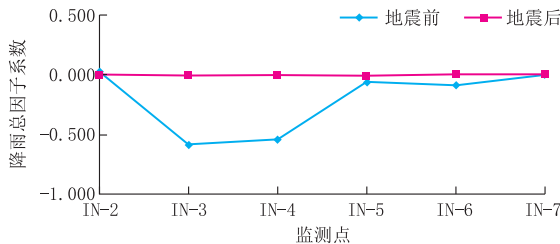


图8 各监测点降雨量因子系数曲线图

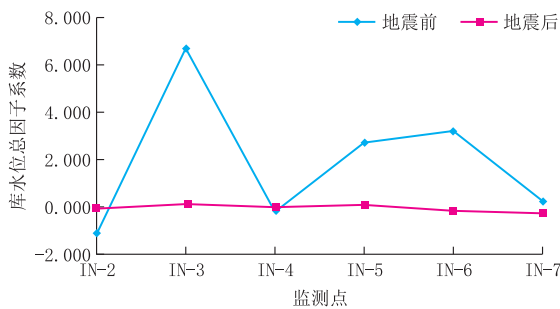


图9 各监测点库水位因子系数曲线图

(3)汶川地震后相对于汶川地震前的时间库水位影响有所变小。

(4)绝大部分监测点回归方程中的C值变小。

(5)汶川地震后降雨量的影响在增加,而时间效益和库水位的影响在下降,应当进一步加强降雨的观测。

### 3 结论

结合工程案例,对该水电站水库堆积体斜坡的影响因子进行了分析和选取,建立了相关多元回归

模型把影响因子与监测位移量进行了量化,经过检验和可靠度分析,其拟合结果较为合理,分析了影响因子在汶川地震的变化,应当进一步加强相关影响因子的监测。该方法在一定程度上解决斜坡影响因子量化的问题,但要使用准确的数学模型来模拟研究斜坡稳定性问题,并达到采用精确的预报模型来预测影响因子的发展规律还相当困难,特别是斜坡的稳定性受外界条件影响较大,如何在该模型中增加影响因子、影响因子确切取值及通过模型分析计算结果结合斜坡实际情况相分析、比较等问题,有待于今后进一步的深入研究。

### 参考文献:

- [1] 祝传兵.边坡稳定性分析与评价[D].云南昆明:昆明理工大学,2005.
- [2] 祝传兵.云南怒江流域河谷岸坡结构特征及稳定性评价[C]//中国地质学会工程地质专业委员会.第九届全国工程地质大会论文集,2012.
- [3] 马还援,宋传中,向钊,等.多元回归分析在安徽省广德县区域边坡稳定性分析中的应用[J].城市地质,2006,(2):42-45.
- [4] 任高峰,张世雄.最大似然估计与一般多元回归在边坡稳定性分析中的应用[J].矿业研究与开发,2004,(6):28-29,94.
- [5] 周洪燕,李满意,司洪涛.基于多元回归数学分析模型某区域内斜坡稳定性研究[J].长春工程学院学报(自然科学版),2015,16(1):73-76.
- [6] 何萍,丁原章.基于CF多元回归模型的地震滑坡危险性研究——以香港屯门为例[J].地震,2013,33(2):52-62.
- [7] 詹威威,黄润秋,裴向军,等.沟道型滑坡—碎屑流运动距离经验预测模型研究[J].工程地质学报,2017,25(1):154-163.
- [8] 白建光,许强.三峡水库塌岸预测的多元回归法[J].内蒙古农业大学学报(自然科学版),2009,30(2):194-198.
- [9] 崔绍峰,王玉良.基于多元回归分析的堤防滑坡预报研究[J].中国水运(下半月),2012,12(10):66-67.
- [10] 韦浩.多元回归分析法在滑坡空间预测中的应用[D].陕西西安:长安大学,2011.
- [11] 王震宇,孟陆波.滑坡预报的多元回归分析方法[J].中国地质灾害与防治学报,2003,(3):24-26.
- [12] 兰恒星,伍法权,王思敬.基于GIS的滑坡CF多元回归模型及其应用[J].山地学报,2002,(6):732-737.
- [13] 董建辉.地震作用下库岸堆积体斜坡动力响应及变形机理研究[D].四川成都:成都理工大学,2016.