

文章编号: 2096-1618(2020)05-0573-06

基于 Logistic 模型的四川山洪流域危险性评价

叶帮苹¹, 冯汉中^{1,2}, 刘志红³, 张小丽⁴, 罗辉¹

(1. 四川省气象台, 四川 成都 610072; 2. 高原与盆地暴雨旱涝灾害四川省重点实验室, 四川 成都 610072; 3. 成都信息工程大学资源环境学院, 四川 成都 610225; 4. 成都市郫都生态环境局, 四川 成都 611730)

摘要: 为对四川山洪灾害的危险性分布进行研究, 利用山洪危险点、历史灾害点和相关的地形、植被和土壤等因子, 采用二元 Logistic 回归方法, 建立四川小流域山洪危险性评价模型, 并根据模型计算得出四川各流域的山洪危险性评价结果。结果表明: 山洪灾害的因子按重要程度由高到低依次为土壤的汇流系数、地形起伏度、流域圆度、平均植被指数和水流累计量。而各因子中, 除土壤的汇流系数与山洪灾害的暴发成正相关, 其余均为负相关; 四川山洪灾害风险划分为风险低、风险较高、风险高和风险很高 4 类, 各风险类型所占全省面积比例分别为 47.05%、11.34%、16.19% 和 25.39%。其中, 四川山洪灾害高风险地区主要分布在盆周山区、凉山州和攀枝花地区, 分布范围较广。低风险区为四川盆地、甘孜北部和阿坝北部。

关键词: 气象学; 气象防灾减灾; 山洪灾害; Logistic 模型; 危险性评价

中图分类号: P694

文献标志码: A

doi: 10.16836/j.cnki.jcui.2020.05.016

0 引言

降雨是诱发山洪灾害的直接因素和激发条件^[1]。据研究分析, 山洪^[2-4]及降雨诱发的泥石流^[5]、滑坡灾害^[6]的发生与降雨量、降雨强度和降雨历时关系密切。降雨量大, 多数情况下意味着雨强高、激发力强, 在一定的下垫面条件下, 易产生溪河洪水灾害、泥石流和滑坡灾害。降雨强度大, 降雨迅速汇聚成地表径流引发溪河洪水, 而泥石流的发生与前 10 min 和 1 h 的降雨强度关系十分密切^[7]。

同样, 陡峭的地形和狭窄山谷和峡谷^[8]等不利的地形地物条件, 也是山洪灾害发生的重要因素。四川省地跨青藏高原、横断山脉、云贵高原、秦巴山地、四川盆地等几大地貌单元, 地势西高东低, 由西北向东南倾斜^[9]。最高点是西部的大雪山主峰贡嘎山, 海拔高达 7556 m。地形复杂多样。以龙门山—大凉山一线为界, 东部为四川盆地及盆缘山地, 西部为川西高山高原及川西南山地。山地、高原和丘陵约占全省土地面积的 97.46%^[10], 且除四川盆地的平原和丘陵外, 大部分地区岭谷高差均在 500 m 以上, 极易发生山洪灾害。因此, 开展小流域山洪灾害的危险性评价, 可为防灾减灾提供技术支持, 进而减小山洪灾害带来的危害。

1 资料

山洪资料: 四川省 2116 个山洪沟普查资料和

2011—2018 年四川全省历史山洪灾害资料。地形资料: 分辨率为 90 m 的四川省 DEM 数据。植被资料: MODIS 卫星 2010—2014 年的平均植被指数。土壤资料: 中国科学院南京土壤研究所发布的 1:100 万中国土壤属性数据库。

以汇流面积 < 200 km² 的小流域为研究单元, 基于四川区域 DEM 数据, 提取汇流面积在 50 ~ 150 km² 的小流域 5600 余个。因山洪主要发生在地势起伏、坡度较大的山地和丘陵区域, 坡度小于 12.5° 的区域发生山洪的危险性较低, 所以, 研究以缓坡^[11]的中间值 10° 作为阈值, 筛选出流域平均坡度大于 10° 的流域, 以此流域分布区域作为评价的基本研究区域。筛选出的小流域平均面积为 87 km², 共计 4600 余个。研究区域见图 1。

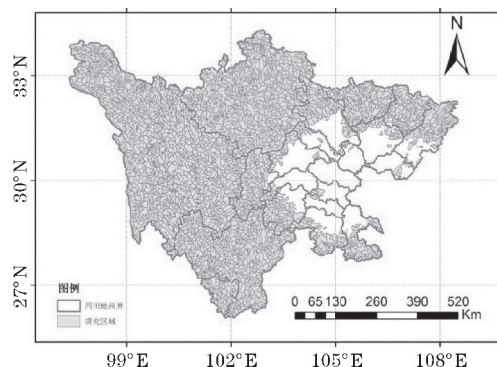


图1 研究区域图

2 研究方法

山洪灾害(是否发生)属于二元变量问题, 只有发

收稿日期: 2019-10-29

基金项目: 四川省重点实验室基金资助项目(2017-青年-06); 中国气象局预报员专项基金资助项目(CMAYBY2019-097)

生与不发生两种情况,且影响因素复杂。二元 Logistic 回归模型^[13-17]是一个概率型非线性的二元统计模型。该模型回归方法相较于主观性方法具有假设简单^[18]、客观性强^[19]、不受人为因素干扰等特点,因而使用二元 Logistic 回归模型作为山洪危险性评价的基本模型。

研究以四川小流域(汇流面积<200 km²的子流域)为最小研究单元,通过已知的山洪危险点和山洪灾害相关的地形、植被和土壤因子,使用 Logistic 回归方法,建立山洪危险性评价模型。建模方法如下:

$$Z=\ln(\frac{P}{1-P})=A+B_1X_1+B_2X_2+\cdots+B_nX_n$$
$$P=\frac{EXP(Z)}{1+EXP(Z)}$$

式中, P 为山洪灾害发生概率,取值为 $[0,1]$,输入参数为有山洪危险的样本; $(1-P)$ 为山洪灾害不发生的概率,输入参数为无山洪危险的样本; X_1, X_2, \cdots, X_n 是影响山洪灾害发生的因子, B_1, B_2, \cdots, B_n 是各影响因子对应的回归系数。

建模过程中的无山洪危险样本是在无山洪危险点流域中随机选取的。通过多次回归和验证确定回归模型。最后通过 ARCGIS 软件将各山洪影响因子代入模型进行叠加分析,得出四川省山洪危险性区划。

3 山洪流域危险性评价

3.1 山洪流域危险性评价因子确定与提取

以小流域(面积<200 km²)为最小研究单元,分别按流域统计各山洪的影响因子。基于判识的小流域,

分析可能影响山洪发生的因子,包括地形、植被和土壤等因子。山洪灾害主要影响因子及统计要素见表 1:

表 1 山洪灾害主要影响因子及统计要素

因子类型	因子名称	统计要素
地形因子	流域特征	流域圆度
	坡度	流域平均
	地形起伏度	流域平均
	水流累计量	流域平均
植被因子	植被指数	流域平均
	植被覆盖度	流域平均
土壤因子	汇流系数	流域平均

3.1.1 地形因子提取

地形因子中,研究除了使用常用的坡度和地形起伏度等影响因子,还引入水流累计量和代表流域形状特征的几何学变量——圆度。

圆度的几何学定义:设平面上一个封闭图形(内部无空洞)的面积为 S ,周长为 C ,则定义该图形的圆度 Afa 为

$$Afa=\frac{4\pi S}{C\cdot C}$$

圆度 Afa 越接近 1,则封闭图形越接近圆,即流域形状越接近圆。研究假设不同圆度的流域对流域内降水的汇流有着不同的影响,流域圆度可能是山洪形成的重要因子之一。

根据表 1 的统计要素,提取每个地形因子对应的流域数据,为模型的建立提供基础数据。地形因子分布见图 2。

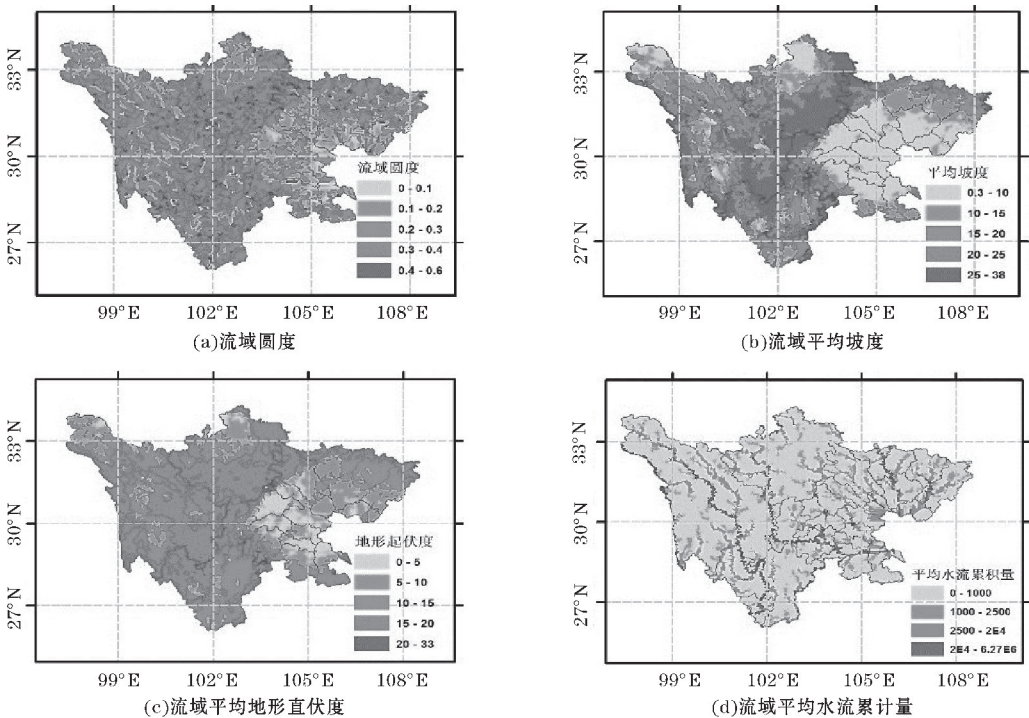


图 2 地形因子分布图

3.1.2 植被和土壤因子提取

山洪形成过程中除了地形因子可以作为主要的动力因子外,不同植被对流水的拦截作用和不同土壤^[12]的下

渗作用也会影响河流的汇流。因此,选取植被指数、植被覆盖度和土壤属性中的汇流系数作为引发山洪灾害发生的重要影响因子。植被和土壤因子分布图见图 3。

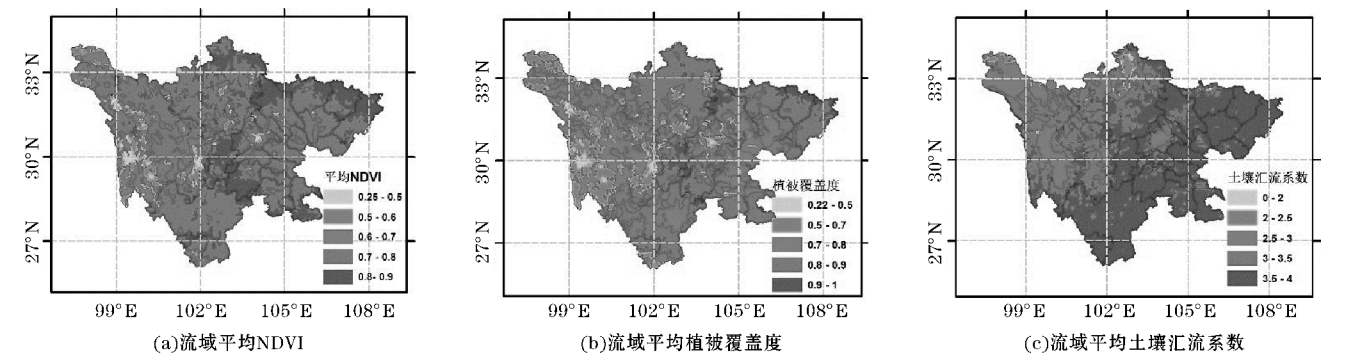


图 3 植被和土壤因子分布图

3.2 山洪流域危险性评价模型建立

使用有山洪沟的小流域作为二元 Logistic 回归模型中的有山洪危险样本,其余小流域为无山洪危险流域样本。回归模型中,以通过数字高程模型 DEM 提取出来的 4600 余个 小流域为分析样本,其中,山洪沟所在的流域为 2000 余个(作为危险流域参与建模)。为了更加合理地选择模型输入中的无危险样本,研究从剩下的无山洪危险点的流域中,随机选取 2000 个流域作为无危险流域样本参与建模。

汇流系数、平均植被指数、流域平均植被覆盖度、流域最大水流累积量、流域平均水流累计量、流域坡度平均、流域坡度余弦平均、平均地形起伏度等 9 项因子。通过对回归因子相关性分析,筛选出相关性≤0.61 的因子进行回归,并通过显著水平为 0.05 的 Wals 检验进行因子重要性分析,最后确定出 Afa(流域圆度)、AccMean(流域平均水流累计量)、DrainMean(土壤平均汇流系数)、Qifudu(流域平均地形起伏度)和 NDVIMean(流域平均植被指数)等 5 个因子。各因子间的相关系数见表 2。

参与建模的自变量因子包括流域圆度、流域平均

表 2 模型中各因子间相关系数

	Afa	AccMean	DrainMean	Qifudu	NDVIMean
Afa	1.000	0.052	-0.305	-0.132	0.109
AccMean	0.052	1.000	-0.250	0.130	0.253
DrainMean	-0.305	-0.250	1.000	-0.442	-0.610
Qifudu	-0.132	0.130	-0.442	1.000	0.085
NDVIMean	0.109	0.253	-0.610	0.085	1.000

通过多次回归和检验,最后选取模型拟合度最优的一组作为模型结果。本研究最终确定山洪危险性评价 Logistic 模型如下:

$$P = EXP(-5.874x_1 - 78.46x_2 + 22.581x_3 - 10.34x_4 - 7.498x_5 - 8.032) / (1 + EXP(-5.874x_1 - 78.46x_2 + 22.581x_3 - 10.34x_4 - 7.498x_5 - 8.032))$$

式中,P 为山洪灾害流域危险性值,x₁ 代表 Afa,x₂ 代表 AccMean,x₃ 代表 DrainMean,x₄ 代表 Qifudu,x₅ 代表 NDVIMean

显著水平为0.05的 Wals 检验是表征各因子的重要性,当 Wald 值越大时表明自变量的回归系数越不可能为 0,即对危险性评价的重要程度越高。通过 Lo-

gistic 分析提供的卡方检验结果如表 3 所示,除降水外,影响四川省山洪灾害的重要程度依次为 DrainMean、Qifudu、Afa、NDVIMean 和 AccMean。

表 3 卡方检验结果

自变量	回归系数	Wals	Sig.
Afa	-5.874	172.004	0.000
AccMean	-78.46	38.540	0.000
DrainMean	22.581	726.960	0.000
Qifudu	-10.34	485.833	0.000
NDVIMean	-7.498	144.920	0.000
常量	-8.032	258.474	0.000

根据回归模型的参数分析,得出以下结论:

- (1)流域圆度越大,流域发生山洪的危险性相对较小,进一步反映出较长的汇流时间,不利于山洪发生;相反,椭圆形或长条形的流域更有利于山洪发生。
- (2)不同的土壤类型,决定了地表对雨水的不同作用,土壤的汇流系数越大,汇流作用就越大,反之则越小。因此,土壤的汇流系数也是影响山洪灾害最主要的因子之一。
- (3)除土壤的汇流系数与山洪灾害的暴发成正相关,流域圆度、水流累积量、地形起伏度和植被指数均为负相关。

3.3 山洪流域危险性评价模型检验分析

山洪危险性评价模型中,其卡方检验结果(表 3)显示,各因子均通过显著水平为 0.05 的 Wald 检验,模型结果具有统计学意义;模型整体拟合度(表 4)显示, Cox & Snell R 方差达到 0.468, Nagelkerke R 方差达到 0.624,通常情况下认为 Cox & Snell R 方差达到 0.4 以上、Nagelkerke R 方差达到 0.5 以上即判断回归模型能较好地模拟所有样例数据;回归模型自检准确率(表 5)显示,模型对山洪灾害发生的预测准确率达 84.5%,对山洪灾害不发生的预测准确率达到 81.1%,总预测准确率为 82.8%。

表 4 模型拟合度

-2 对数似然值	Cox & Snell R 方差	Nagelkerke R 方差
2744.345a	0.468	0.624

表 5 模型自检准确率

已观测		已预测		
		0	1	百分比校正/%
是否危险	0	1454	338	81.1
	1	286	1558	84.5
总计百分比				82.8

而通过对 321 个山洪历史灾害点分析,有 203 次山洪灾害发生在高风险区,占历史灾害的 63%。

文献[20]研究认为,四川盆地及周边地区是山洪灾害的高风险值地区。研究除了盆地区域与其差别较大外,盆周山区和凉山等地区的高风险区分布基本一致。主要原因是,研究预先通过小流域平均坡度,对山洪的研究区域进行了筛选,两者研究区域有所差别。

3.4 四川山洪流域危险性评价结果

二元 Logistic 回归模型的数学结果只是从统计学的角度验证了本研究确定的危险性评价模型的可行性,山洪灾害的长期监测预警更重要的是获得空间上

的山洪危险性区划。

通过 ARCGIS 空间分析功能按上述回归模型,将流域圆度、流域平均水流累计量、土壤平均汇流系数、流域平均地形起伏度和流域平均植被指数代入模型计算,获得各流域的山洪灾害概率。Logistic 回归模型的结果是事件发生的概率值,范围是 0 到 1,通过山洪风险分类标准将结果划分为风险低、风险较高、风险高和风险很高 4 类,最后得到四川山洪危险性评价结果(图 4)。山洪风险等级分类标准和所占面积见表 6。

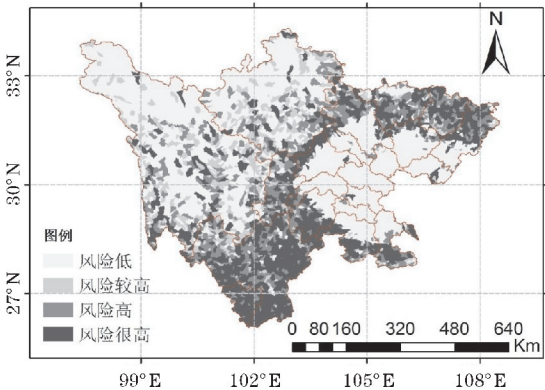


图 4 山洪危险性评价结果

表 6 山洪风险等级分类标准和所占面积

	风险低	风险较高	风险高	风险很高
概率区间	[0,0.4)	[0.4,0.6)	[0.6,0.8)	[0.8,1]
面积/万 km ²	22.77	5.51	7.87	12.34
全省面积占比/%	47.05	11.34	16.19	25.39

通过山洪危险性区划分析,攀枝花、凉山州、龙门山脉一带和盆地东北部的山洪危险性较高,高风险区(发生概率大于 0.6)共约 20.21 万 km²,约占全省总面积的 41.58%。四川盆地、甘孜北部和阿坝北部为低风险区。

4 结论

研究以四川小流域为最小研究单元,通过已知的山洪危险点和山洪灾害相关的地形、植被和土壤因子,使用 Logistic 回归方法,建立山洪危险性评价模型,并根据模型计算得出四川省各流域的山洪危险性评价结果,得出以下结论:

- (1)在山洪因子选取的过程中,除了选择常用的坡度、地形起伏度和植被指数等影响因子外,还选择了几个特殊的因子:水流累积量、流域圆度和土壤汇流系数。其中,流域圆度是地形因子中的一个,该因子借鉴于一个几何学变量——圆度。该变量的引用,是先假设山洪的形成与流域的形状有关。而模型最后的结果也证实了这个假设的成立。土壤汇流系数是土壤众多

属性中的一个,与水流累积量等地形因子一样,该属性能对山洪的形成产生重要影响。

(2)在影响四川山洪灾害的因子中,按重要程度由高到低依次为土壤的汇流系数、地形起伏度、流域圆度、平均植被指数和水流累积量。而除土壤的汇流系数与山洪灾害的暴发成正相关,流域圆度、水流累积量、地形起伏度和植被指数均为负相关。

(3)四川山洪灾害风险划分为风险低、风险较高、风险高和风险很高4类,各风险类型所占全省面积比例分别为47.05%、11.34%、16.19%和25.39%。其中,四川山洪灾害高风险地区主要分布在盆周山区、凉山州和攀枝花地区,分布范围较广;低风险区为四川盆地、甘孜北部和阿坝北部。

(4)研究使用山洪普查的灾害点进行建模。而普查点的主观影响较大,如果能都收集足够的历史山洪灾害样本进行建模,再结合降水资料进行分析,结果会更加符合实际情况。

研究使用二元 Logistic 回归模型对四川流域的山洪危险性进行评价,模型中使用到的地形因子、植被和土壤因子,都比较容易获取。虽然受到山洪灾害的样本个数和质量的限制,但由于 Logistic 回归模型自身的特点,结果具有客观性较强、研究方法的适用性和可推广性较高等特点。

参考文献:

- [1] 章德武,谌宏伟. 山洪灾害致灾因子分析与防治措施[J]. 中国水运(下半月刊),2011,11(3): 146-147.
- [2] 熊俊楠,龚颖,程维明,等. 西藏自治区近30年山洪灾害时空分布特征[J]. 山地学报,2018,36(4):557-570.
- [3] John H. Puffer,ChrisLaskowich. Volcanic diapiers in the Orange Mountain flood basalt:New Jersey,USA [J]. Journal of Volcanology and Geothermal Research,2012(2):237-238.
- [4] M Rico, G Benito, A Barnolas. Combined palaeo-flood and rainfall-runoff assessment of mountain floods(Spanish Pyrenees)[J]. Journal of Hydrology,2001,245(1).
- [5] 杨红娟,韦方强,马振峰,等. 四川省泥石流灾害的时空分布规律和降水特征[J]. 灾害学,2017,32(4):102-107.
- [6] 李云君,刘志红,吕远洋,等. 四川省滑坡灾害气象预警模型建立与验证[J]. 地球信息科学学报,2017,19(7):941-949.
- [7] 张志彤. 我国山洪灾害特点及其防治思路[J]. 中国水利,2007(14):14-15.
- [8] Phillips A D. Flash flooding across the southern Appalachian Mountains:an abbreviated climatology [J]. 2012.
- [9] 何云玲,张一平. 纵向岭谷区生态环境对气候变化的适应性评价[J]. 山地学报,2009,27(3): 300-305.
- [10] 肖术连,鲁长江,杨璐遥. 山地救援与城市救援的对比研究[J]. 城市与减灾,2017(1):42-45.
- [11] 张志永,程郁春,程丽,等. 三峡库区万州段消落带植被及土壤理化特征分析[J]. 水生生态学杂志,2016,37(2):24-33.
- [12] Marchi L, Borgia M, Preciso E, et al. Characterisation of selected extreme flash floods in Europe and implications for flood risk management [J]. Journal of Hydrology(Amsterdam), 2010, 394(1-2):118-133.
- [13] 邢秋菊,赵纯勇,高克昌,等. 基于GIS的滑坡危险性逻辑回归评价研究[J]. 地理与地理信息科学,2004(3):49-51.
- [14] 田春山,刘希林,汪佳. 基于CF和Logistic回归模型的广东省地质灾害易发性评价[J]. 水文地质工程地质,2016,43(6):154-161+170.
- [15] 唐川,马国超. 基于地貌单元的小区域地质灾害易发性分区方法研究[J]. 地理科学,2015,35(1):91-98.
- [16] 聂婷,肖荣波,王国恩,刘云亚. 基于Logistic回归的CA模型改进方法——以广州市为例[J]. 地理研究,2010,29(10):1909-1919.
- [17] Bursac Z, Gauss C H, Williams D K, et al. Purposeful selection of variables in logistic regression [J]. Source Code for Biology and Medicine, 2008,3(1):17.
- [18] 费晓燕,柳锦宝,屈伯强,等. 四川省降雨诱发滑坡灾害的气象预警模型[J]. 水土保持通报,2017,37(5):315-321.
- [19] 杨可明,郭海会,钱小丽,等. 结合信息熵改进的证据权法与滑坡危险度区划[J]. 地理与地理信息科学,2013,29(3):104-108.
- [20] 丁文峰,杜俊,陈小平,等. 四川省山洪灾害风险评估与区划[J]. 长江科学院院报,2015,32(12):41-45.

Risk Assessment of Sichuan Mountain Flooding based on Logistic Model

YE Bangping¹, FENG Hanzhong^{1,2}, LIU Zhihong³, ZHANG Xiaoli⁴, LUO Hui¹

(1. Sichuan Meteorological Observatory, Chengdu 610072, China; 2. Rainstorm, droughts and floods disaster of plateau and basin of Sichuan emphasis laboratory, Chengdu 610072, China; 3. College of Resources and Environment, CUIT, Chengdu 610225, China; 4. Chengdu Pidu Ecological Environment Bureau, Chengdu 611730, China)

Abstract: In order to study the risk distribution of flash floods in Sichuan, we first collected factors such as danger points, historical disaster points of flash flood flash flood danger points, flash flood historical disaster points, and terrain, vegetation, and soil related to flash flood disasters. Then using binary logistic regression method, a model of flash flood risk assessment of small watershed in Sichuan was established and the results of flash flood risk assessment in Sichuan Basin were calculated based on the model. The results show that: Among the factors of the flash flood disasters in Sichuan, they are ranked in order of importance: soil confluence coefficient, topographic relief, watershed roundness, vegetation index, and water flow accumulation. And except that the soil confluence coefficient is positively correlated with the mountain flood disaster, others are negatively correlated; Sichuan mountain flood disaster risks are divided into four categories: low, medium, high, and higher. The proportions of each type of risk in the province are 47.05%, 11.34%, 16.19%, and 25.39%. Among them, the high-risk areas of mountain flood disasters are mainly distributed in the basin-mountainous mountainous areas, Liangshan Prefecture and Panzhihua area, with a wide range of distribution. The low-risk areas are the Sichuan Basin, northern Ganzi, and northern Aba.

Keywords: meteorology; meteorological disaster prevention and mitigation; flash flood; Logistic model; risk assessment