文章类型：论文|刊号（ISSN）：2630-4740 /（中图刊号）：715GL012

鄂西南某山区县降雨型地质灾害气象预警模型研究

郭琪[1]　郑琦玉[2]
1 湖北天宏检测科技集团有限公司
2 湖北省防雷中心
DOI:10.32629/eep.v9i8.3293

[摘　要] 降雨是诱发山区地质灾害的主要外部动力,构建降雨型地质灾害气象预警模型对区域防灾减灾具有重要意义。本文以鄂西南某山区县为研究区,将确定性系数(CF)法与逻辑回归(LR)、支持向量机(SVM)、决策树(DT)和BP神经网络四种机器学习模型耦合构建潜势度评价模型,并结合有效降雨量模型建立气象预警体系,对2020年汛期典型降雨过程进行回溯验证。结果表明：曲率、土地利用类型和坡度是地质灾害发育的主导因子；CF-DT模型综合性能最优,准确率为0.6442、AUC为0.6838；结合有效降雨量分级后,CF-DT模型对2020年6—7月40余起降雨诱发地质灾害的平均命中率达95%,显著优于其他模型。CF耦合机器学习与有效降雨量的方法可有效支撑区域降雨型地质灾害气象预警。
[关键词] 降雨型地质灾害；气象预警；确定性系数；机器学习
中图分类号：P694　文献标识码：A

Study on a Meteorological Early Warning Model for Rainfall-Induced Geological Hazards in a Mountainous County of Southwestern Hubei

Qi Guo[1]　Qiyu Zheng[2]
1 Hubei Tianhong Testing Technology Group Co., Ltd
2 Hubei Province Lightning Protection Center

[Abstract] Rainfall is the main external trigger of geological hazards in mountainous areas, and constructing a meteorological early warning model for rainfall-induced geological hazards is of great significance for regional disaster prevention and mitigation. Taking a mountainous county in southwestern Hubei Province as the study area, this study couples the Certainty Factor (CF) method with four machine learning models, namely logistic regression (LR), support vector machine (SVM), decision tree (DT), and BP neural network, to construct a potentiality assessment model. Combined with an effective rainfall model, a meteorological early warning system is established, and retrospective validation is conducted for typical rainfall events during the 2020 flood season. The results show that curvature, land use type, and slope are the dominant factors controlling the development of geological hazards. The CF-DT model has the best overall performance, with an accuracy of 0.6442 and an AUC of 0.6838. After incorporating effective rainfall grading, the CF-DT model achieves an average hit rate of 95% for more than 40 rainfall-induced geological hazards during June-July 2020, significantly outperforming the other models. The method coupling CF with machine learning and effective rainfall can effectively support regional meteorological early warning for rainfall-induced geological hazards.
[Key words] Rainfall-induced geological hazards; Meteorological early warning; Certainty factor; Machine learning

1 引言

我国山区面积广阔,地质环境脆弱,降雨尤其是持续性强降雨和短时强降雨是触发滑坡、崩塌、泥石流等地质灾害的最主要外部动力。随着极端降雨事件多发频发,降雨型地质灾害的风险防控形势日益严峻[1-2]。

传统预警多采用基于区域降雨阈值的经验方法[3-5],原理简单但存在两方面不足:一是阈值多基于单一降雨指标,忽视了研究区内部地质环境的空间差异,难以反映同一降雨条件下不同区域致灾能力的差别;二是阈值高度依赖历史灾情样本,在样本不足或地质条件复杂地区适用性有限。

Copyright © This work is licensed under a Commons Attribution-Non Commercial 4.0 International License.

将地质灾害易发性(潜势度)评价与气象诱发因素相结合,是当前区域预警的重要方向。其核心思想是：灾害发生是内在地质条件与外部降雨共同作用的结果。机器学习因非线性拟合能力强、预测精度高,被广泛用于潜势度评价[6-7]；有效降雨量模型则能较好地刻画前期降雨对岩土体含水状态的持续影响。二者融合有望实现更精准的区域气象风险预警。

基于此,本文以鄂西南某山区县为研究区,构建“确定性系数—机器学习—有效降雨量”耦合的气象预警方法：先以CF法量化因子敏感性并结合相关性分析与信息增益率筛选因子；再将CF与多种机器学习模型耦合建立潜势度评价模型并择优；最后引入有效降雨量模型进行气象诱发修正,形成分级预警,并以典型降雨过程回溯验证其有效性。

2 研究区与数据

2.1研究区概况

研究区位于鄂西南武陵山区腹地,属中山—低山地形,海拔高差大、沟谷深切,地势西北高、东南低。区内出露地层以沉积碎屑岩、碳酸盐岩为主,工程地质岩组复杂；受构造影响断裂发育、岩体破碎,为地质灾害孕育提供了有利条件。

研究区属亚热带季风性湿润气候,多年平均降水量1100—1500 mm,降水集中在5—9月汛期且多以暴雨形式出现,持续与短时强降雨频繁,属地质灾害高易发区。“地质环境脆弱”与“降雨强度大、历时长”叠加,使研究区成为开展降雨型地质灾害预警研究的典型区域。

2.2数据来源

本文数据主要包括三部分：地质灾害历史数据来自历史资料与野外调查,为点状矢量；基础地理与地质环境数据中,DEM为30 m分辨率公开地理空间数据,土地利用类型为10 m分辨率公开遥感数据,地貌、水系、道路、土壤等为30 m或1:100万公开基础地理数据,工程地质岩组与断层来自地质调查资料(矢量)；气象降雨采用0.1° ×0.1° 公开气象数据集。

数据预处理时,将栅格数据统一重采样至同一分辨率与投影坐标系,保证各因子空间对齐。地质灾害历史数据经去重、剔除不全记录后,共获得400余处灾害点作为正样本；并随机选取等量非灾害点作为负样本,构成训练与测试数据集。

3 研究方法

3.1确定性系数法

确定性系数(Certainty Factor,CF)法是一种基于条件概率的敏感性分析方法,通过对比某因子分级内灾害发生的条件概率与研究区灾害发生的先验概率,判断该分级对灾害的敏感程度。CF取值范围为[-1,1],正值表示有利于灾害发生,负值表示不利于发生,接近0则无显著影响。

将连续型因子按自然间断法或等间距法分级,统计各分级内灾害点数与单元总数并计算CF值。CF值既可直接表征因子致灾倾向,又可作为量化指标代入机器学习模型,实现经验敏感性与数据驱动的结合。

3.2机器学习潜势度评价模型

选取逻辑回归(LR)、支持向量机(SVM)、决策树(DT)和BP神经网络四种代表性模型,分别与CF法耦合构建CF-LR、CF-SVM、CF-DT和CF-BP潜势度评价模型。模型输入为经CF值量化的因子分级值,输出为各评价单元的灾害潜势度概率。

其中LR可解释性强,SVM适合非线性分类,DT对特征交互刻画良好,BP神经网络非线性逼近能力强。四种模型原理互补,通过性能对比可筛选出最适合研究区的潜势度评价方法。

3.3有效降雨量模型与预警分级

地质灾害发生不仅与当日降雨有关,还与前期持续降雨密切相关。前期降雨渗入岩土体后滞留于孔隙裂隙中,改变其含水状态与力学强度。为刻画这一过程,采用有效降雨量模型综合度量降雨,计算公式为：

$$R_e=R_0+\Sigma\alpha\hat{}i\times R_i(i=1,2,\ldots,n)$$

式中：R_e为有效降雨量；R_0为当日降雨量；R_i为前第i日降雨量；α 为时间衰减系数(取0.84)；n为有效降雨日数(取7天)。

在潜势度与有效降雨量基础上计算综合危险度：将有效降雨量按0～20、20～40、40～80、80～120、120～160、160～180、≥180 mm七档分别归一化为0.4、0.5、0.6、0.7、0.8、0.9、1.0,与各单元潜势度相乘得到危险度,再依阈值划分为无预警、黄色、橙色和红色预警四个等级。

4 结果与分析

4.1因子筛选与敏感性分析

初步选取高程、坡度、坡向、曲率、地貌、工程地质岩组、距断层距离、雨量等值面、距河流距离、NDVI、土地利用类型、距道路距离共12个候选因子。相关性分析显示NDVI与土地利用类型相关系数达0.55,且信息增益率分析表明NDVI重要性较低,故予以剔除,最终保留11个因子用于评价。

利用CF法对保留因子进行敏感性分析,主要因子分级CF值如表1所示。高程100—300 m区间CF值较高(0.88～0.92),表明低海拔丘陵河谷地带灾害高发,与人类活动集中、地形切割强烈有关；坡度0～10° 区间CF值为0.50,缓坡地带灾害相对集中；土地利用类型中建筑物CF值达0.75,显著高于其他类型,表明工程建设等人类活动诱发作用明显；距道路距离≤100 m区间CF值为0.71,进一步印证人类工程活动与灾害的空间耦合关系。综合来看,曲率、土地利用类型和坡度是研究区地质灾害发育的主导因子。

表1 主要影响因子的确定性系数(CF)分级值

因子	分级	CF 值
高程/m	100～200 / 200～300 / 400～500	0.92 / 0.89 / 0.73
坡度/°	0～10 / 10～20 / 20～30	0.50 / 0.14 / -0.28
土地利用类型	建筑物 / 农作物 / 树木	0.75 / -0.29 / -0.40
距道路距离	≤100 m / >100 m	0.71 / -0.18
距断层距离	≤100 m / >100 m	0.39 / -0.01
距河流距离	≤800 m / >800 m	0.35 / -0.06
雨量等值面/mm	1000～1100 / >1500	0.66 / -0.16

4.2潜势度评价模型对比

Copyright © This work is licensed under a Commons Attribution-Non Commercial 4.0 International License.

将经CF值量化的11个因子作为输入,分别训练四种模型,采用准确率、精确率、召回率、F1分数和AUC五项指标综合评价,结果如表2所示。

表2　四种潜势度评价模型的训练性能指标

模型	Accuracy	Precision	Recall	F1 Score	AUC
CF-LR	0.6923	0.6556	0.6413	0.6484	0.7626
CF-SVM	0.6827	0.6512	0.6087	0.6292	0.7436
CF-DT	0.6442	0.5849	0.6739	0.6263	0.6838
CF-BP	0.6971	0.7231	0.5109	0.5987	0.7419

由表2可知,四种模型在训练集上均具有一定判别能力。AUC指标CF-LR最高(0.7626),CF-DT相对较低(0.6838)；但召回率CF-DT达0.6739,对灾害样本识别能力较强。综合来看,CF-DT在准确率、召回率与稳定性间取得较好平衡,且规则化结构更便于业务预警中的解释与应用,故选作后续气象预警的潜势度基础模型。

4.3气象预警回溯验证

选取2020年6月28日至7月19日一次典型强降雨过程进行回溯验证,该时段共记录40余起降雨诱发地质灾害,代表性较好。将各灾害发生日的有效降雨量代入模型,计算灾害点所在网格的预警等级,以灾害点落入黄色及以上预警区的比例为命中率,结果如表5所示。

回溯验证中,CF-DT平均命中率达95%,明显高于CF-LR(88%)、CF-SVM(86%)和CF-BP(81%),表现最优。在有效降雨量较大、灾害集中发生的强降雨日,CF-DT仍保持较高命中率,体现了对极端降雨下群发灾害的稳定捕捉能力。从预警等级分布看,灾害点主要落入黄色和橙色预警区,表明分级预警体系能将大部分灾害有效纳入预警范围。

综上,CF耦合决策树与有效降雨量的方法在回溯验证中取得95%命中率,验证了该方法在区域降雨型地质灾害气象预警中的有效性与实用性。

5 讨论

本文方法的核心优势在于将地质易发性与气象诱发因素有机耦合。相较于单纯依赖降雨阈值的传统方法,CF耦合机器学习得到的潜势度能反映研究区内部地质条件的空间差异,使同一降雨下不同区域的致灾能力得到差异化表达,提升了预警的空间针对性;有效降雨量模型则量化了前期降雨的持续影响,弥补了仅考虑当日雨量的不足。

本研究仍存在局限:一是灾害样本规模有限,可能影响模型泛化能力；二是有效降雨量模型的衰减系数与有效降雨日数等参数多依赖经验取值,在不同地质环境与气候条件下的适用性有待校验；三是预警分级阈值尚需结合更多历史灾情优化。

未来可引入更长时序、更精细尺度的降雨与灾害数据以优化有效降雨量参数,探索深度学习等方法提升潜势度评价精度,并将方法拓展至更多地质环境差异显著的区域进行适用性检验,推动预警从“经验阈值”向“地质—气象耦合”范式转变。

6 结论

本文以鄂西南某山区县为研究区,构建了“确定性系数—机器学习—有效降雨量”相耦合的地质灾害气象预警方法,主要结论如下：

(1)利用相关性分析与信息增益率对12个候选因子进行筛选,剔除NDVI后保留11个因子；确定性系数敏感性分析表明,曲率、土地利用类型和坡度是影响研究区地质灾害发育的主导因子,人类工程活动与灾害空间分布具有显著耦合关系。

(2)CF与四种机器学习模型耦合的对比结果表明,CF-DT模型在准确率、召回率与稳定性之间取得较好平衡,且规则化结构便于业务解释,被选作潜势度评价基础模型。

(3)结合有效降雨量分级的气象预警回溯验证显示,CF-DT模型对2020年6—7月典型强降雨过程诱发地质灾害的平均命中率达95%,显著优于其他三种模型,验证了该方法在区域降雨型地质灾害气象预警中的有效性与实用性。

[参考文献]

[1]黄雨,何正迎.气候变化背景下降雨滑坡风险评估研究进展[J].中国地质灾害与防治学报,2025,36(2):1-13.

[2]许冲,邵霄怡.气候变化与地质灾害风险防控:降雨触发型地质灾害研究进展[J].中国地质灾害与防治学报,2025.

[3]Guzzetti F,Peruccacci S,Rossi M,et al.Rainfall thresholds for the initiation of landslides in central and southern Europe[J].Meteorology and Atmospheric Physics,2007,98(3-4):239-267.

[4]谢洋义,殷坤龙,杜娟,等.基于不同地质环境分区的重庆市滑坡降雨阈值模型研究[J].地质科技通报,2025,44(1):126-137.

[5]孙强,刘明军,张泰丽,等.诱发台风暴雨型滑坡的降雨阈值研究——以泰顺县为例[J].水文地质工程地质,2024,51(4):197-205.

[6]许冲,代克滨,薛智文,等.中国月尺度降雨型滑坡发生概率研究[J].地质力学学报,2025,31(5):960-971.

[7]席雪萍,闫勇,袁婷婷,等.基于风险区划和降雨阈值的地质灾害预警模型研究[J].矿产勘查,2025,16(3):513-520.

作者简介：

郭琪(1982--),女,湖北武汉人,工程师,主要从事雷电防护装置检测、雷电灾害风险评估及气象地质灾害预警技术研究。

郑琦玉(1987--),女,湖北孝感人,工程师,主要从事雷电防护与监测,气象灾害预警工作。

Copyright © This work is licensed under a Commons Attribution-Non Commercial 4.0 International License.