中国地质灾害与防治学报
THE CHINESE JOURNAL OF GEOLOGICAL HAZARD AND CONTROL
中国科技核心期刊
《中国科学引用文数据库》来源期刊
Caj-cd规范获奖期刊

基于随机森林赋权信息量模型的地质灾害易发性评价

黄　成，晏祥省，梅红波，周翠琼，黄　格

Susceptibility analysis of geological hazards based on the random forest weighted information value model: A case study of Shidian County,Yunnan Province

HUANG Cheng, YAN Xiangsheng, MEI Hongbo, ZHOU Cuiqiong, and HUANG Ge

在线阅读 View online: https://doi.org/10.16031/j.cnki.issn.1003-8035.202401013

DOI: 10.16031/j.cnki.issn.1003-8035.202401013
黄成, 晏祥省, 梅红波, 等. 基于随机森林赋权信息量模型的地质灾害易发性评价——以云南省施甸县为例[J]. 中国地质灾害与防治学报, 2025, 36(3): 151-159.
HUANG Cheng, YAN Xiangsheng, MEI Hongbo, et al. Susceptibility analysis of geological hazards based on the random forest weighted information value model: A case study of Shidian County,Yunnan Province[J]. The Chinese Journal of Geological Hazard and Control, 2025, 36(3): 151-159.

基于随机森林赋权信息量模型的地质灾害易发性评价
——以云南省施甸县为例

黄 成[1,2,3,4,5], 晏祥省[2,3,4,5], 梅红波[6], 周翠琼[2,3,4,5], 黄 格[6]
(1. 昆明理工大学国土资源工程学院, 云南 昆明 650093; 2. 自然资源部滇中高原生物多样性与地球关键带野外科学观测研究站, 云南 昆明 650216; 3. 自然资源部高原山地地质灾害预报预警与生态保护修复重点实验室, 云南 昆明 650216; 4. 云南省高原山地地质灾害预报预警与生态保护修复重点实验室(筹), 云南 昆明 650216; 5. 云南省地质环境监测院, 云南 昆明 650216; 6. 中国地质大学(武汉)资源学院, 湖北 武汉 430074)

摘要: 传统的信息量模型在进行地质灾害易发性评价时, 通常只是简单地将各个评价因子的信息量值累加, 而忽略了因子之间权重的差异, 这在一定程度上影响了易发性分区的科学性和合理性。为了克服这个问题, 文章以云南省施甸县为例, 引入了随机森林模型来计算各评价因子的权重。在构建了合适的评价指标体系后, 逐一计算每个因子的信息量及其权重, 然后进行加权求和。按照等间隔分级法, 将研究区域划分为极高、高、中、低 4 个易发性等级。为了验证模型的准确性, 选取了近 3 年内该区最新调查-重点区域地质灾害精细化调查与风险评价成果得到的地质灾害隐患点与易发性分区进行叠加, 并通过隐患点密度和 ROC 曲线进行精度检验对比分析。对比研究结果发现, 引入随机森林赋权后, 极高易发内隐患点密度由 1.754 升至 1.926, *AUC* 值从 0.809 升至 0.847。研究结果表明: 在单一信息量模型中引入随机森林进行赋权能有效表达因子间的权重差异, 提升地质灾害易发性分区的精度, 在实际应用中具有更高的准确性。

关键词: 随机森林; 信息量模型; 地质灾害; 易发性评价
中图分类号: P694 **文献标志码:** A **文章编号:** 1003-8035(2025)03-0151-09

Susceptibility analysis of geological hazards based on the random forest weighted information value model: A case study of Shidian County,Yunnan Province

HUANG Cheng[1,2,3,4,5], YAN Xiangsheng[2,3,4,5], MEI Hongbo[6], ZHOU Cuiqiong[2,3,4,5], HUANG Ge[6]
(1. *School of Land and Resources Engineering, Kunming University of Technology, Kunming, Yunnan* 650093, *China*; 2. *The Observatory on Biodiversity & Critical Zone in the Central Yunnan Plateau, the Ministry of Natural Resources* (*OBCYP-MNR*), *Kunming, Yunnan* 650216, *China*; 3. *Key Laboratory of Geohazard Forecast and Geoecological*

收稿日期: 2024-01-05; **修订日期:** 2024-07-24 **投稿网址:** https://www.zgdzzhyfzxb.com/
基金项目: 云南省地质灾害隐患识别中心建设(云财资环〔2021〕23 号); 云南省地质灾害精细化调查与风险评价综合遥感(云财资环〔2020〕68 号); 云南省地质灾害综合防治体系建设专项计划(2013-2020)(云政发〔2013〕108 号)
第一作者: 黄 成(1981—), 男, 河南信阳人, 地质工程专业, 硕士, 高级工程师, 主要从事水工环地质、地质灾害时空大数据分析与应用、地质灾害隐患识别与综合遥感研究等工作。E-mail: hch2377@163.com
通讯作者: 周翠琼(1972—), 女, 广西崇左人, 水工环地质专业, 本科, 正高级工程师, 主要从事水工环地质调查研究工作。E-mail: 407528806@qq.com

Restoration in Plateau Mountainous Area, *MNR*, *Kunming*, *Yunnan* 650216, *China*; 4. *Yunnan Key Laboratory of Geohazard Forecast and Geoecological Restoration in Plateau Mountainous Area*, *Kunming*, *Yunnan* 650216, *China*; 5. *Yunnan Institute of Geological Environment Monitoring*, *Kunming*, *Yunnan* 650216, *China*; 6. *School of Earth Resources*, *China University of Geosciences*, *Wuhan*, *Hubei* 430074, *China*)

Abstract: Traditional information value models for evaluating geological hazard susceptibility typically involve simply summing the information values of various evaluation factors, without considering the differences in weight among these factors. This can affect the scientific rigor and rationality of susceptibility zoning to some extent. To address this issue, this paper takes Shidian County of Yunnan Province as an example and introduces the random forest model to calculate the weights of each evaluation factor. After constructing an appropriate evaluation index system, the information value and weight of each factor are calculated individually, followed by a weighted summation. According to the equal interval classification method, the study area is then divided into four susceptibility levels--extremely high, high, medium, and low. To verify the accuracy of the model, the latest geological hazard hidden points identified through detailed investigations and risk assessments over the past 3 a were overlaid with the susceptibility zones. The accuracy was analyzed through hazard point density analysis and ROC curve comparison. Based on the comparison of research results, after introducing the random forest weighting, the density of extremely high-risk hidden hazard points increased from 1.754 to 1.926, and the *AUC* value improved from 0.809 to 0.847. The research results indicate that introducing random forest for weighting in a single information quantity model can effectively reflects the weight differences among factors, enhancing the precision of geological disaster susceptibility zoning. This method shows higher accuracy in practical applications.

Keywords: random forest; information model; geological hazard; susceptibility evaluation

0 引言

地质灾害是指地球动力活动和环境变化异常引起的自然灾害,通常地质灾害包括:滑坡、崩塌、泥石流、地裂缝、地面塌陷、地面沉降等[1]。根据国家市场监督管理总局、中国国家标准化管理委员会 2021 年发布的《地质灾害危险性评估规范》,我国的地质灾害可以按照造成的伤亡和经济损失大小划分为:大型、中型和小型。多种类型的地形地貌与气候为地质灾害的发生孕育了条件。大部分地质灾害具有隐蔽性、突发性、多成因性、巨大破坏性和预测难度大等特点[2-3]。

云南省复杂的地质构造与地质环境条件致使滑坡、泥石流、崩塌等地质灾害多发易发[4]。整体上地质灾害的分布密度呈现出西部大于东部,北部大于南部的特点[5]。位于云南省西部边陲的施甸县,由于其具备广泛分布的西南山区独特的中高山峡谷地形地貌条件,其垂直高差悬殊,各类地质灾害频发,对该地区人民的生命和财产造成严重威胁,制约着当地社会经济的快速发展。如何不断提升地质灾害隐患识别能力是当前地质灾害防治的一项重要工作。

易发性评价是预测和防治地质灾害发生的有效方法,易发区是指满足地质灾害发生的地质构造和地形地貌等条件,易于发生地质灾害的区域,主要依据地质环境条件,参考地质灾害现状,进行易发风险等级的划定[6]。目前,地质灾害易发性分析的主要手段有:经验模型、统计分析模型和模式识别模型。经验模型主要运用的方法为模糊分析和层次分析法(AHP)[7-8]。这类方法基于专家经验,能够较好地反映地质灾害发生的实际机理,操作简单直观。但是由于依赖于主观判断,结果存在一定的不确定性和局限性,难以量化各因素的相对重要性。通常用于数据缺乏或难以获取的地区,可利用有经验的专家知识开展定性分析。统计分析模型主要运用的方法为频率比法[9]、确定性系数法[10]、信息量法[11-12]和证据权法[13-14]。统计分析模型基于大量实测数据,能够客观反映各因素对地质灾害发生的影响程度,结果可量化。但是该类模型需要大量可靠的历史数据作为样本,对数据质量和分布特征要求较高,难以体现复杂的发生机理。因此这类方法通常用在有充足历史灾害数据的地区,进行地质灾害的定量评估。模式识别模型主要运用到神经网络和随机森林模型等各种集成方法[15-20]。这类方法能够挖掘复杂的灾害发生规律,多种集成算法可以提高预测准确性。同样,这类方法的前提就是需要大量高质量训练数据,模型“黑箱”

特性使结果解释性较差，对参数调优要求高。在有丰富地质灾害数据支撑的情况下，可应用此方法对地质灾害进行精细化分析。

在地质灾害易发性评价中，模型本身并无优劣之别，关键在于针对不同地区特征选择不同的模型和方法。由于施甸县地形地貌类型复杂，涉及的致灾因子较多，单一信息量模型的应用不能够表现出不同评价因子之间的权重大小，因此本文采用了随机森林模型客观赋权的方法，能够很大程度上更加客观地反映出不同评价因子对地质灾害发生的促进作用，提高滑坡易发性分区结果的准确性，便于决策者对不同易发性分区实施针对性防治措施。

1　研究区概况

施甸县隶属云南省保山市，位于怒江东岸，南北走向的横断山脉中，东与昌宁县隔枯柯河而望，南与临沧市以勐波罗为界，西与龙陵县隔怒江相对，北与隆阳区相连，区内河流众多，植被类型丰富。该县三面江河环绕，以山地为主，总的地势为南北走向，北高南低，地质构造复杂，受到多个构造体系的影响，有多条主要断裂。研究区属中亚热带气候，以低纬度高原季风气候为主，且具有山势高峻、河谷深切等特点，垂直气候分异显著。施甸县的气候和水文特征，为该县的农业、林业、水利、旅游等产业提供了有利的条件，也带来了一些灾害风险，如干旱、洪涝、泥石流等。由于人类长期生产生活活动，大部分原生植被已遭破坏，除高远山区保存部分较好森林外，其余大部分属次生云南松林、灌丛等自然植被类型，还存在部分人工植被类型。施甸县地质构造极为复杂，其处于青藏滇缅印尼巨型“歹”字型构造体系和经向构造体系的交叉区，其中以断裂构造占居主要地位，由之构成了区内的基本构造格架，总体构造形迹特征是由一系列近南北向的高角度压性、压扭性断裂及其派生的次一级近东西向、北西向、北东向的横张断裂和紧密褶皱所组成。区内主要断裂有蚌冬断裂、姚关断裂、摆榔断裂、大海子—施甸断裂、胡家山—乌木龙断裂及金鸣—大庄断裂等，施甸县工程地质条件简图如图 1 所示。

施甸县主要地质灾害涉及崩塌、泥石流、滑坡、地裂缝及地面塌陷。其中以滑坡数量居多，其次为泥石流，其他类型的灾害点分布较少。2014 以来，施甸县发生地质灾害灾险情共 82 起，造成直接经济损失 405 万元，间接经济损失 5 117 万元。

施甸县共有滑坡灾害 238 处，滑坡灾害在施甸县全

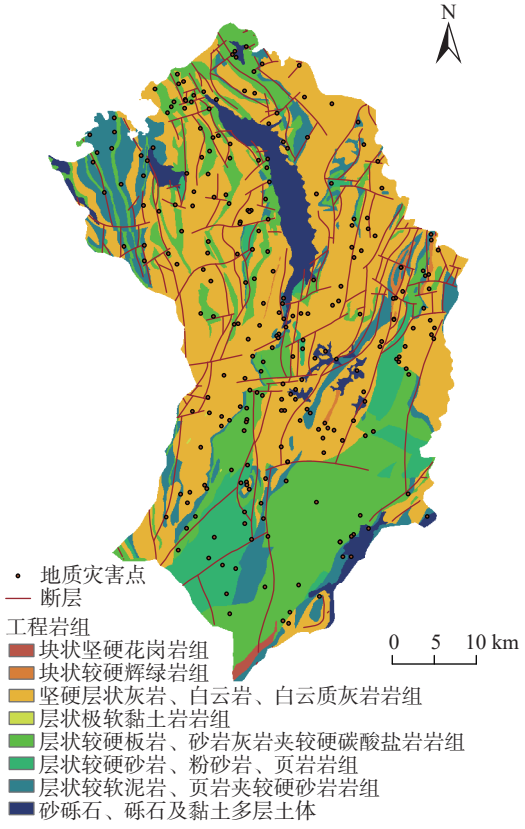

图 1　工程地质条件简图

Fig. 1　Schematic diagram of engineering geological conditions

县均有发生，怒江及支流沿岸和勐波罗河流域为滑坡最为发育和密集的地区。从数量上看，区内滑坡以小型和中型为主，小型滑坡 214 个、中型滑坡 22 个，大型滑坡 2 个，无巨型滑坡。区内滑坡的主要控滑面以松散层与基岩接触面为主，其次为层理面，节理裂隙面，片理或劈理、断层及老滑面。

施甸县崩塌不甚发育，且规模较小，均为中型和小型，中型 1 个、小型 3 个。但具一定的分布规律，均分布于县乡公路沿线及陡峭山坡地段，崩塌的发育程度与岩土体类型密切，主要发育于志留系、泥盆系、石炭系的灰岩、泥质灰岩、页岩岩组。

研究区内泥石流分布较少，小型泥石流 10 处、中型泥石流 1 处。主要分布于勐波罗河流域及姚关河流域，泥石流灾害主要发生在旧城乡、姚关镇和木老元乡，其中以旧城乡乡政府驻地芒梗沙河泥石流危害最为严重。由于区内地质构造复杂，新构造运动频繁，岩体破碎，滑坡、崩塌发育，加之区内分布大量易风化的软弱岩层及软硬相间的岩层，残坡积层较厚，为泥石流的形成提供了丰富的松散碎屑固体来源。区域内泥石流的发育程度与岩土体类型相关，寒武系板岩、变质砂岩岩组，奥陶系粉砂岩、页岩、粉砂岩夹页岩、泥灰岩岩

组，志留系页岩、粉砂岩、泥岩岩组，泥盆系泥质灰岩、泥灰岩夹页岩岩组，石炭系泥质灰岩、玄武岩、页岩岩组及三迭系泥岩、粉砂岩、页岩夹粉砂岩岩组，是区内泥石流发育的岩组，泥石流松散固体物质主要来源于由此产生的崩积、坡积物和沟床堆积物，滑坡和崩塌是提供泥石流物源的主要方式。

2 模型方法

2.1 信息量模型

信息量概念最早可以追溯到 20 世纪 40 年代，最初仅用于通信领域，在 1985 年左右晏同珍、殷坤龙逐渐将该理论应用于地质灾害的危险性评价中[21－23]。

信息量模型是利用统计的方法，对每个影响因子内的各个子类所对应的信息量值进行计算，从而反映出其对地质灾害发育的影响程度。这是一种客观高效的评价方法，在地灾评价中得到广泛应用。同时利用信息量值的总和可以表示某一评价单元发生地质灾害的可能性大小，从而进行易发性分区。其公式为：

$$I=\sum_{i=1}^{m}I(x_i,H)=\sum_{i=1}^{m}\frac{\frac{N_i}{N}}{\frac{S_i}{S}} \tag{1}$$

式中：I——某一评价单元内总信息量；

m——评价因子数量；

x_i——各个评价因子的等级；

$I(x_i,H)$——各评价因子表征的信息量；

N_i——存在地质灾害单元中分布有评价因子x_i的数量；

N——灾害点分布的单元总数；

S_i——评价因子x_i的评价单元总数量；

S——研究区的评价单元总数量。

2.2 随机森林模型

随机森林是一种集成学习模型，它由多棵决策树$\{h(X,\Theta k),k=1,2,\cdots\}$组成。参数集$\{\Theta k\}$是独立同分布的随机向量，在给定自变量$X$的情况下，最优预测结果由多棵决策树模型的结果整合投票得出。国内外有大量学者从理论和应用领域等方面，证实了随机森林模型在各个领域内的可靠性，该模型能兼容并清洗原数据中的噪声与离群点，也使得该模型成为目前受到普遍认可的机器学习模型之一[24－26]。

随机森林具有一项特殊的优点：可以在依据基尼指数或信息增益等方法选择最优划分属性的同时计算各评价因子的相对权重。在决策树中最佳分割通过不纯度衡量，不纯度则由基尼指数法计算。计算评价因子(k)在分割节点时基尼指数损失(D_{Gk})，将整个森林模型中所有节点的损失(D_{Gk})求和后对每棵决策树求平均，所得即为因子(k)的相对重要性。可按下式计算：

$$P_k=\frac{\sum_{h=1}^{n}\sum_{j=1}^{t}D_{Gkhj}}{\sum_{k=1}^{m}\sum_{h=1}^{n}\sum_{j=1}^{t}D_{Gkhj}} \tag{2}$$

式中：P_k——第k个评价因子相对权重；

m——评价因子总数；

n——分类树棵数；

t——单棵树节点数；

D_{Gkhj}——第k个因子在第h棵树的第j个节点的基尼指数损失值。

2.3 随机森林赋权信息量模型

传统的信息量模型存在一定缺陷：未区分不同评价因子的影响程度，影响易发性评价结果。因此将 2 种模型结合，引入随机森林模型计算出各因子的权重，将权重赋给各因子的信息量值。公式如下：

$$y=\sum_{j=1}^{m}P_jx_j \tag{3}$$

式中：y——综合信息量值；

m——评价因子的总个数；

P_j——评价因子x_j对应的权重。

3 评价指标选取

根据不同的影响因子建立地质灾害易发性评价体系，是进行科学合理的滑坡预测的先行条件。然而，地质灾害的发生与各种复杂的自然因素有关，并不可能把所有因素都考虑进去，高维特征会造成数据冗余，不利于评价的精度，也会增加模型计算成本。因此，选择少量适合的、具有代表性的因子参与模型评价非常关键。

首先，地形因素是最基本的影响因素。坡度决定了地表的稳定性，而地形起伏度反映了地表应力状态，二者直接决定了滑坡、崩塌等重力型灾害的易发性。

其次，不同的岩土体组合特征和构造活动是诱发地质灾害的内在动力机制。不同的岩性和构造环境会产生不同类型和强度的灾害，因此必须将其纳入评价指标体系。

再者，水文因素如河网密度也是重要考虑因素。丰

富的河流网络会造成地表水文条件的复杂变化，从而加剧地质灾害的发生。

此外，人类活动也是不可忽视的影响因素。道路建设改变了地表形态，而植被覆盖的变化直接影响了地表的稳定性。这些人为因素都会对地质灾害的发生产生显著影响。

因此，本文选取坡度、地形起伏度、工程岩组、距断层距离、河网密度、路网密度、植被覆盖度 7 个因子进行易发性评价，这 7 个因子的选择立足于施甸县地质灾害的发生机理，兼顾了自然因素和人为因素，能较全面地反映研究区的地质灾害风险状况。各评价因子分布见图 2。

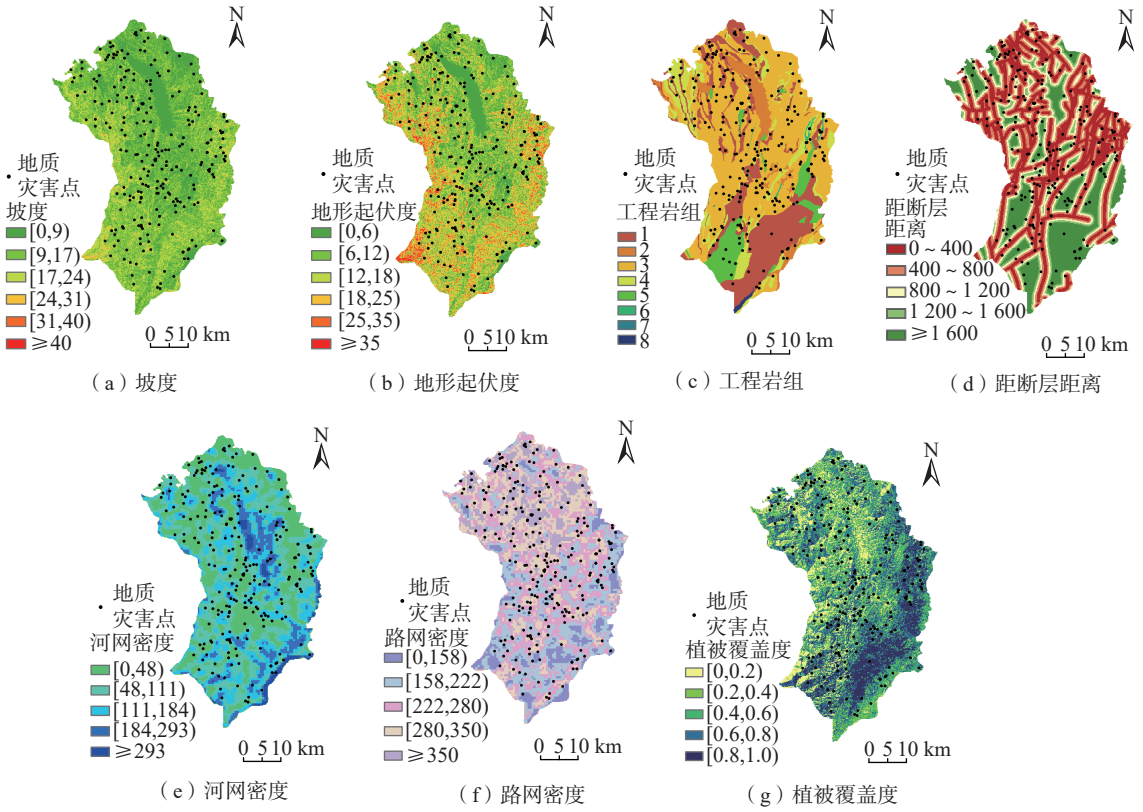

图 2　评价因子

Fig. 2　Evaluation factors

注：工程岩组中 1 为层状较硬板岩、砂岩灰岩夹较硬碳酸盐岩岩组；2 为砂砾岩、砾石及黏土多层土体；3 为坚硬层状灰岩、白云岩、白云质灰岩岩组；4 为层状较软泥岩、页岩夹较硬砂岩岩组；5 为层状较硬砂岩、粉砂岩、页岩岩组；6 为块状较硬辉绿岩组；7 为层状极软黏土岩岩组；8 为块状坚硬花岗岩组砂砾岩、砾石及黏土多层土体。下表 2 中同。

坡度因子根据高程数据在 ArcGIS 中提取得来，见图 2（a）。地形起伏度用于表征地貌形态，利用相应区域内高程的最大值和最小值，计算差值即可得该区域内的地表起伏度值，见图 2（b）。工程岩组决定了岩石的力学强度，同时也影响着岩石的抗风化和侵蚀能力，见图 2（c）。地质构造和地质灾害的发育之间存在着密切的关系。断层发育使得当地应力集中，导致区域岩土体破碎和底层错动起伏，这将直接导致地质灾害的发生，见图 2（d）。水系对地质灾害的影响用河网密度反映，河网密度来源于高程数据，见图 2（e）。路网密度通过高程数据提取路网后计算得出，见图 2(f)。植被覆盖度常用于衡量地表植被覆盖程度，采用 Landsat8 OLI 数据得出研究区的归一化植被指数，再根据归一化植被指数值进行波段运算得到研究区的植被覆盖度，见图 2（g）。具体数据来源见表 1。

4　地质灾害易发性评价及精度检验

4.1　随机森林权重和加权信息量值计算

根据发生施甸县的地质灾害发育情况，选择施甸县 253 个地质灾害点作为正例样本。同时使用同等数量的 1km 外的随机点作为负例样本。在随机森林中使用 80% 的数据为训练集，剩下 20% 为测试集，并由此计算出各个评价因子的权重，权重分布图如图 3 所示。

在将经过随机森林模型计算后的权重与信息量值

表 1 地质数据来源

Table 1 Sources of geological data

	数据	数据来源及比例尺	数据格式
1	坡度	ASTER GDEM v3 30 m分辨率数字高程数据	GeoTIFF
2	地形起伏度	ASTER GDEM v3 30 m分辨率数字高程数据	GeoTIFF
3	工程岩组	云南省重点区域地质灾害精细化调查与风险评价(2022)	MapGIS
4	断层	云南省重点区域地质灾害精细化调查与风险评价(2022)	MapGIS
5	河网密度	ASTER GDEM v3 30 m分辨率数字高程数据	GeoTIFF
6	路网密度	ASTER GDEM v3 30 m分辨率数字高程数据	GeoTIFF
7	植被覆盖度	Landsat 8 OLI_TIRS遥感影像	Tif
8	地质灾害	云南省重点区域地质灾害精细化调查与风险评价(2022)	MapGIS

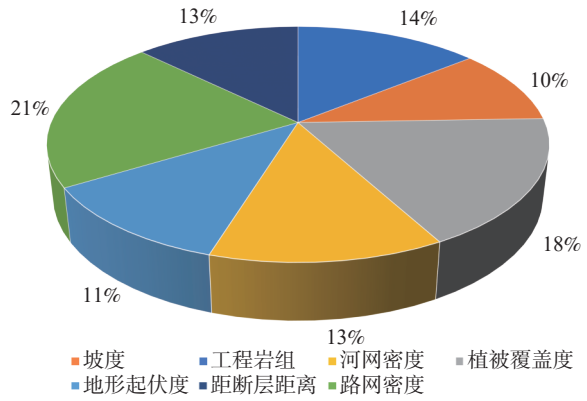

图 3 评价因子权重图

Fig. 3 Evaluation factors weight graph

进行线性叠加后,分别对各评价因子的各个分级进行信息量值和加权信息量值统计,得到的信息量值分别如表 2 所示。

4.2 地质灾害易发性评价及分区

在 ArcGIS 中进行因子图层加权叠加后,得到研究区地质灾害易发风险概率图,再利用等间隔法将整个区域划分为 4 类易发区:极高易发区(0.75 ~ 1)、高易发区(0.5 ~ 0.75)、中易发区(0.25 ~ 0.5)和低易发区(0 ~ 0.25),最终获得施甸县地质灾害易发性评价分区(图 4)。

利用 2022 年施甸县重点区域地质灾害精细化调查与风险评价成果的 176 个隐患点进行经度检验,分别统计信息量模型和随机森林赋权信息量模型分区结果中各类易发区的灾害点数量并计算灾害密度(表 3—4)。

在单一信息量模型得出的易发性分区中,极高易发

表 2 评价因子分级区间信息量值和加权信息量值

Table 2 Information value and weighted information value for classification intervals of evaluation factors

评价因子	分级区间	信息量值	权重	加权信息量
坡度/(°)	[0,9)	0.166 7	0.102 8	0.017 1
	[9,17)	0.267 4		0.027 5
	[17,24)	−1.516 0		−0.155 8
	[24,31)	−1.933 7		−0.198 8
	[31,40)	–		–
	[40,78]	–		–
地形起伏度/m	[0,6)	−0.131 1	0.115 2	−0.015 1
	[6,12)	0.612 8		0.070 6
	[12,18)	−0.175 2		−0.020 2
	[18,25)	−0.685 5		−0.079 0
	[25,35)	−1.290 3		−0.148 6
	[35,147]	–		–

(续表)

评价因子	分级区间	信息量值	权重	加权信息量
工程岩组	1	−0.090 5	0.141 6	−0.012 8
	2	−0.916 4		−0.129 8
	3	0.103 7		0.014 7
	4	0.151 2		0.021 4
	5	−0.460 0		−0.065 1
	6	2.123 6		0.300 7
	7	–		–
	8	–		–
距断层距离/m	0 ~ 400	0.233 3	0.126 9	0.029 6
	400 ~ 800	0.016 4		0.002 1
	800 ~ 1 200	0.032 0		0.004 1
	1 200 ~ 1 600	−0.643 5		−0.081 7
	>1 600	−0.221 3		−0.028 1
河网密度	[0,48)	−0.127 0	0.127 5	−0.016 2
	[48,111)	0.316 7		0.040 4
	[111,184)	−0.215 0		−0.027 4
	[184,293)	−0.322 7		−0.041 1
	[293,618]	−0.902 4		−0.115 1
路网密度	[0,158)	−1.447 4	0.210 1	−0.304 1
	[158,222)	−0.013 7		−0.002 9
	[222,280)	0.006 3		0.001 3
	[280,350)	0.194 7		0.040 9
	[350,590]	0.386 0		0.081 1
植被覆盖度	[0,0.2)	0.573 3	0.175 9	0.100 9
	[0.2,0.4)	0.521 2		0.091 7
	[0.4,0.6)	0.034 4		0.006 1
	[0.6,0.8)	−0.118 2		−0.020 8
	[0.8,1.0)	−1.440 6		−0.253 4

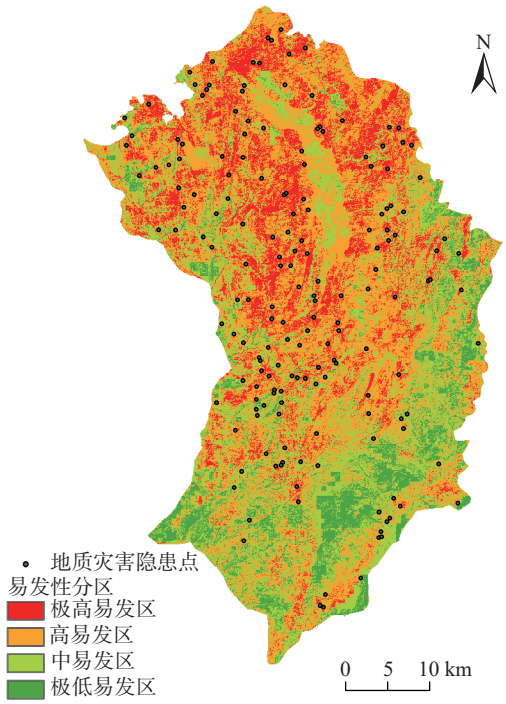

图 4　施甸县地质灾害易发性评价分区

Fig. 4　Zoning map of geological hazard susceptibility evaluation in Shidian County

表 3　信息量模型各评价分区面积和地质灾害隐患点占比

Table 3　Aera and proportion of geological hazard points in each evaluation zone using the information value model

易发性分区	隐患点个数/个	分区面积/km²	隐患点比例/%	分区比例/%	隐患点密度
极高易发区	108	703.01	61.36	34.99	1.754
高易发区	36	634.70	20.45	31.59	0.647
中易发区	25	465.71	14.20	23.18	0.613
低易发区	7	205.59	3.98	10.23	0.389

表 4　随机森林赋权信息量模型各评价分区面积和地质灾害隐患点占比

Table 4　Aera and proportion of geological hazard points in each evaluation zone using the random forest weighted information value model

易发性分区	隐患点个数/个	分区面积/km²	隐患点比例/%	分区比例/%	隐患点密度
极高易发区	112	663.63	63.64	33.03	1.926
高易发区	40	668.09	22.73	33.25	0.683
中易发区	23	428.31	13.07	21.32	0.613
低易发区	1	248.98	0.57	12.39	0.046

区的灾害点密度(隐患点数量占比/易发性分区面积占比)为 1.754，随机森林赋权信息量模型得出的极高易发区灾害点密度为 1.926，后者更有利于地质灾害风险的精准防控。

4.3　精度检验

为了进一步对比两种评价方式的准确性和可靠性，可通过受试者特征曲线(ROC)来检验模型评价结果的精度。

ROC 曲线，也被称作“敏感性曲线”或“受试者特征曲线”。由于其不受临界条件限制，能够反映出分析方法的特异性与敏感性的关系，且实验精度较高，故常被应用于地质灾害易发性评价中。ROC 曲线的 *AUC* 值(曲线下的面积)越大，证明预测结果越准确。

单一信息量模型和随机森林赋权信息量模型易发性评价的 ROC 曲线图如图 5 所示。

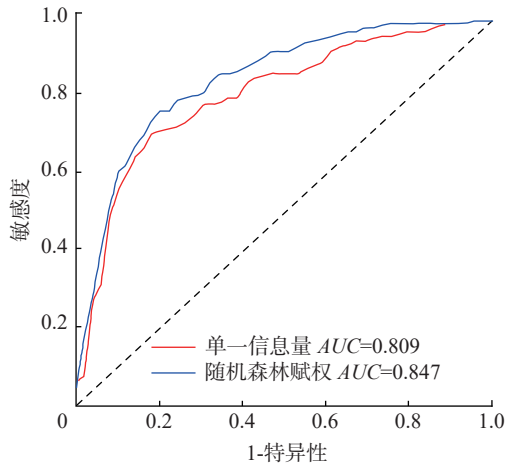

图 5　施甸县易发性评价 ROC 曲线图

Fig. 5　ROC curve of susceptibility evaluation in Shidian County

由图 5 可知，研究区随机森林赋权信息量模型得到的易发性评价结果的 ROC 曲线下面积 *AUC* 值为 0.847，高于单一信息量模型的 0.809。由此可见，引入随机森林方法计算评价因子权重有利于提高施甸县地质灾害易发性评价的精度。

5　结论

(1)本文采用随机森林赋权信息量模型对云南省施甸县进行了地质灾害易发性评价，总体评价精度较高(*AUC* 值为 0.847)，并且高于单一信息量模型(*AUC* 值为 0.809)。极高易发区分布较为分散，但相对集中在县域的西北部及中西部，高易发性区域地质灾害分布特征相似，多沿道路发育，受岩土体类型控制明显。

(2)通过将施甸县近 3 a 精细化调查的 176 个隐患点叠加到各分区进行对比分析，随机森林赋权后的极高易发区隐患点数量和隐患点密度更高，更有利于地质灾害风险的精准防控。

参考文献(References):

[1] 齐信,唐川,陈州丰,等.地质灾害风险评价研究[J].自然灾害学报,2012,21(5):33 - 40.[QI Xin, TANG Chuan, CHEN Zhoufeng, et al. Research of geohazards risk assessment[J]. Journal of Natural Disasters, 2012, 21(5): 33 - 40.(in Chinese with English abstract)]

[2] 马彦霞.地质灾害易发性风险评估及预测分析[J].世界有色金属,2021(15):144 - 145.[MA Yanxia. Risk assessment and prediction analysis of geological disaster[J]. World Nonferrous Metals, 2021(15): 144 - 145.(in Chinese with English abstract)]

[3] 刘业森,张晓蕾,郭良.自然灾害调查数据的多尺度异常检测方法研究及应用[J].地球信息科学学报,2017,19(12):1653 - 1660.[LIU Yesen, ZHANG Xiaolei, GUO Liang. Study and application of the method of multi-scale outliers detection of natural disaster investigation data[J]. Journal of Geo-Information Science, 2017, 19(12): 1653 - 1660.(in Chinese with English abstract)]

[4] 王宇.云南省地质灾害防治与研究历史评述[J].灾害学,2019,34(3):134 - 139.[WANG Yu. Historical review of geological disaster prevention and research in Yunnan Province, China[J]. Journal of Catastrophology, 2019, 34(3): 134 - 139.(in Chinese with English abstract)]

[5] 王宇.云南省崩塌滑坡泥石流灾害及防治[J].地质灾害与环境保护,1998,9(4):38 - 41.[WANG Yu. Hazards collapse landslide and debris flow in Yunnan and their control[J]. Journal of Geological Hazards and Environment Preservation, 1998, 9(4): 38 - 41.(in Chinese)]

[6] 唐亚明,张茂省,李林,等.滑坡易发性危险性风险评价例析[J].水文地质工程地质,2011,38(2):125 - 129.[TANG Yaming, ZHANG Maosheng, LI Lin, et al. Discrimination to the landslide susceptibility, hazard and risk assessment[J]. Hydrogeology & Engineering Geology, 2011, 38(2): 125 - 129.(in Chinese with English abstract)]

[7] 于开宁,吴涛,魏爱华,等.基于AHP-突变理论组合模型的地质灾害危险性评价——以河北平山县为例[J].中国地质灾害与防治学报,2023,34(2):146 - 155.[YU Kaining, WU Tao, WEI Aihua, et al. Geological hazard assessment based on the models of AHP, catastrophe theory and their combination: A case study in Pingshan County of Hebei Province[J]. The Chinese Journal of Geological Hazard and Control, 2023, 34(2): 146 - 155.(in Chinese with English abstract)]

[8] 冯卫,唐亚明,马红娜,等.基于层次分析法的咸阳市多灾种自然灾害综合风险评价[J].西北地质,2021,54(2):282 - 288.[FENG Wei, TANG Yaming, MA Hongna, et al. Comprehensive risk assessment of multi-hazard natural disasters in Xianyang City based on AHP[J]. Northwestern Geology, 2021, 54(2): 282 - 288.(in Chinese with English abstract)]

[9] SHANO L, RAGHUVANSHI T K, METEN M. Landslide susceptibility mapping using frequency ratio model: The case of Gamo highland, south Ethiopia[J]. Arabian Journal of Geosciences, 2021, 14(7): 623.

[10] 赵伯驹,李宁,幸夫诚,等.基于I-CF模型的四川德格县滑坡危险性评价与区划[J].中国地质灾害与防治学报,2023,34(5):32 - 42.[ZHAO Boju, LI Ning, XING Fucheng, et al. Landslide geological hazard assessment based on the I-CF model of Dege County in Sichuan Province[J]. The Chinese Journal of Geological Hazard and Control, 2023, 34(5): 32 - 42.(in Chinese with English abstract)]

[11] 阮沈勇,黄润秋.基于GIS的信息量法模型在地质灾害危险性区划中的应用[J].成都理工学院学报,2001,28(1):89 - 92.[RUAN Shenyong, HUANG Runqiu. Application of GIS-based information model on assessment of geological hazards risk[J]. Journal of Chengdu University of Technology, 2001, 28(1): 89 - 92.(in Chinese with English abstract)]

[12] 李文彦,王喜乐.频率比与信息量模型在黄土沟壑区滑坡易发性评价中的应用与比较[J].自然灾害学报,2020,29(4):213 - 220.[LI Wenyan, WANG Xile. Application and comparison of frequency ratio and information value model for evaluating landslide susceptibility of loess gully region[J]. Journal of Natural Disasters, 2020, 29(4): 213 - 220.(in Chinese with English abstract)]

[13] 许冲,戴福初,徐锡伟.基于GIS平台与证据权的地震滑坡易发性评价[J].地球科学,2011,36(6):1155 - 1164.[XU Chong, DAI Fuchu, XU Xiwei. Earthquake triggered landslide susceptibility evaluation based on GIS platform and weight-of-evidence modeling[J]. Earth Science, 2011, 36(6): 1155 - 1164.(in Chinese with English abstract)]

[14] 白光顺,杨雪梅,朱杰勇,等.基于证据权法的昆明五华区地质灾害易发性评价[J].中国地质灾害与防治学报,2022,33(5):128 - 138.[BAI Guangshun, YANG Xuemei, ZHU Jieyong, et al. Susceptibility assessment of geological hazards in Wuhua District of Kuming, China using the weight evidence method[J]. The Chinese Journal of Geological Hazard and Control, 2022, 33(5): 128 - 138.(in Chinese with English abstract)]

[15] 郭子正,殷坤龙,付圣,等.基于GIS与WOE-BP模型的滑坡易发性评价[J].地球科学,2019,44(12):4299 - 4312.[GUO Zizheng, YIN Kunlong, FU Sheng, et al. Evaluation of landslide susceptibility based on GIS and WOE-BP model[J]. Earth Science, 2019, 44(12): 4299 - 4312.(in Chinese with English abstract)]

[16] 刘帅,朱杰勇,杨得虎,等.基于斜坡单元与随机森林模

型的元阳县崩滑地质灾害易发性评价［J］. 中国地质灾害与防治学报，2023, 34(4)：144－150.［LIU Shuai, ZHU Jieyong, YANG Dehu, et al. Evaluation of geological hazard susceptibility of collapse and landslide in Yuanyang County using slope units and random forest modeling［J］. The Chinese Journal of Geological Hazard and Control, 2023, 34(4)：144－150.（in Chinese with English abstract）］

［17］ 曾斌，吕权儒，寇磊，等. 基于Logistic回归和随机森林的清江流域长阳库岸段堆积层滑坡易发性评价［J］. 中国地质灾害与防治学报，2023, 34(4)：105－113.［ZENG Bin, LYU Quanru, KOU Lei, et al. Susceptibility assessment of colluvium landslides along the Changyang section of Qingjiang River using Logistic regression and random forest methods［J］. The Chinese Journal of Geological Hazard and Control, 2023, 34(4)：105－113.（in Chinese with English abstract）］

［18］ 李国营，刘平，张凯，等. 量纲统一在滑坡易发性评价中的影响分析［J］. 水文地质工程地质，2024, 51(3)：118－129.［LI Guoying, LIU Ping, ZHANG Kai, et al. Analysis of the influence of dimensional unity in landslide susceptibility assessment［J］. Hydrogeology & Engineering Geology, 2024, 51(3)：118－129.（in Chinese with English abstract）］

［19］ HUANG Cheng, LI Fang, WEI Lei, et al. Landslide susceptibility modeling using a deep random neural network［J］. Applied Sciences, 2022, 12(24)：12887.

［20］ PARK S, CHOI C, KIM B, et al. Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea［J］. Environmental Earth Sciences, 2013, 68(5)：1443－1464.

［21］ 殷坤龙，晏同珍. 滑坡预测及相关模型［J］. 岩石力学与工程学报，1996, 15(1)：1－8.［YIN Kunlong, YAN Tongzhen. Landslide prediction and relevant models［J］. Chinese Journal of Rock Mechanics and Engineering, 1996, 15(1)：1－8.（in Chinese）］

［22］ 殷坤龙. 滑坡灾害风险分析［M］. 北京：科学出版社，2010.［YIN Kunlong. Risk analysis of landslide disaster［M］. Beijing: Science Press, 2010.（in Chinese）］

［23］ 殷坤龙，朱良峰. 滑坡灾害空间区划及GIS应用研究［J］. 地学前缘，2001, 8(2)：279－284.［YIN Kunlong, ZHU Liangfeng. Landslide hazard zonation and application of GIS［J］. Earth Science Frontiers, 2001, 8(2)：279－284.（in Chinese with English abstract）］

［24］ 郝国栋. 基于随机森林模型的商南县滑坡易发性评价［D］. 西安：西安科技大学，2019.［HAO Guodong. Landslide susceptibility assessment based on random forest model in Shangnan County［D］. Xi'an: Xi'an University of Science and Technology, 2019.（in Chinese with English abstract）］

［25］ 张书豪，吴光. 随机森林与GIS的泥石流易发性及可靠性［J］. 地球科学，2019(9)：3115－3134.［ZHANG Shuhao, WU Guang. Debris flow susceptibility and its reliability based on random forest and GIS［J］. Earth Science, 2019(9)：3115－3134.（in Chinese with English abstract）］

［26］ 刘永垚，第宝锋，詹宇，等. 基于随机森林模型的泥石流易发性评价——以汶川地震重灾区为例［J］. 山地学报，2018, 36(5)：765－773.［LIU Yongyao, DI Baofeng, ZHAN Yu, et al. Debris flows susceptibility assessment in Wenchuan earthquake areas based on random forest algorithm model［J］. Mountain Research, 2018, 36(5)：765－773.（in Chinese）］