基于机器学习方法改进 IVM-RF 耦合模型的崩滑灾害危险性评价：以延安市志丹县为例

屈鹏鑫[a,b]，谢婉丽[a,b*]，刘琦琦[a,b]，王昱琛[a,b]

西北大学 a. 地质学系；b. 大陆演化与早期生命全国重点实验室，西安 710069

第一作者，E-mail：1279806472@qq.com

* 通信作者，E-mail：xiewanli@nwu.edu.cn

2024-10-08 收稿；2024-12-21 修回；2024-12-23 接受

基金项目：国家自然科学基金项目（42372320；41972292）；陕西省创新能力支撑计划项目（2021TD-54）；陕西省重点研发计划项目（2022ZDLSF06-03）；西安市科技计划项目（24LLRHZDZX0019）

中图分类号：P642.2　　文章编号：2096-8523（2025）03-0280-16　　doi：10.19509/j.cnki.dzkq.tb20240583

摘　要：为了给延安市志丹县防灾减灾和风险管控提供数据支持，给相似地区危险性评价提供参考依据，补充在崩塌和滑坡灾害危险性评价中未考虑累积降雨方面的影响，在 10，20，50 和 100 a 一遇 4 种不同降雨工况下对研究区进行了危险性评价。以栅格单元作为评价单元，结合灾害特征及区域孕灾背景，通过皮尔逊相关系数法，选取高程、坡度、坡向、曲率、岩土体类型、距河流距离、距道路距离、归一化植被指数 8 个评价因子，采用信息量模型进行了易发性评价并分析了致灾因子与灾害分布关联性。利用计算机语言，自动处理前期因子数据的分析、转化、管理和出图等流程，改进了信息量（IVM）-随机森林（RF）耦合模型，实现了模型的自动循环迭代对比选择，通过 ROC（受试者工作特征曲线，receiver operating characteristic curve）曲线对比了 2 种易发性模型精度。在耦合模型评价结果的基础上进行危险性评价，用皮尔逊Ⅲ型曲线估算研究区 10，20，50 和 100 a 一遇 4 种不同工况下降雨量，并进行危险性分区。对于易发性分区结果，信息量-随机森林耦合模型评价结果的 *AUC* 值（ROC 曲线下面积）为 0.87，优于 IVM 模型的评价结果；对于危险性分区结果，从 10 a 一遇到 100 a 一遇降雨工况的高和极高危险区面积都逐级增加。研究表明，改进的耦合模型评价方法不仅简化了操作还提高了精度，耦合模型确实拥有更好的评价精度和预测能力。

关键词：崩塌；滑坡；信息量（IVM）模型；随机森林（RF）模型；危险性评价；机器学习；延安市志丹县

黄土高原地形地貌多样、地质环境复杂、生态环境脆弱，加之黄土解理裂隙发育、水敏性强，由于前期不合理人类工程活动影响及近年来频发的极端降雨的诱发，崩塌和滑坡等地质灾害事件大量增加。虽然前人[1-3]关于黄土崩塌和滑坡的结构特征、形成机理、破坏模式、灾害防治和其危险性评价方面做了大量的研究工作，但在极端降雨频发的背景下如何开展黄土高原地区高精度地质灾害危险性评价，准确预测地质灾害影响范围，规划指导地质灾害防控工作，从而降低因地质灾害造成的生命财产损失[4-7]，还待进一步研究。

危险性评价方法分为定性与定量 2 类，定性评价法面世较早，其代表性方法为层次分析法[8-10]，随着技术的发展，定量评价法也开始推广并应用，其代表性方法有信息量法（information value method，简称 IVM）[11-12]、逻辑回归[13]、确定性系数法[14]、频率

著录信息：屈鹏鑫，谢婉丽，刘琦琦，等. 基于机器学习方法改进 IVM-RF 耦合模型的崩滑灾害危险性评价：以延安市志丹县为例［J］. 地质科技通报，2025，44（3）：280-295.

QU Pengxin，XIE Wanli，LIU Qiqi，et al. Collapse and landslide risk assessment based on machine learning improved IVM-RF coupling method：A case study of Zhidan County，Yan'an City［J］. Bulletin of Geological Science and Technology，2025，44（3）：280-295.

© Editorial Office of Bulletin of Geological Science and Technology. This is an open access article under the CC BY-NC-ND license.　https://dzkjtb.cug.edu.cn

比法[15-16]、证据权法[17]等。研究发现[18-19]，相较于单一模型，耦合模型精度更高，且克服了单一模型存在的诸多问题。樊芷吟等[20]通过信息量模型与逻辑回归模型耦合的方法开展了地质灾害易发性评价，克服了信息量模型难以获得诱发因素贡献度的问题；罗路广等[21]采用确定性系数与逻辑回归耦合模型进行了九寨沟景区滑坡易发性评价，解决了确定性系数模型未能考虑各评价因子对滑坡易发性影响的差异性问题。近年来，随着 3S 技术与计算机科学的发展，神经网络法[22-23]、支持向量机[24]和随机森林（random forest，简称 RF）[25]模型等机器学习法[26]也被广泛应用于危险性评价工作中。神经网络法作为一种基于所给图像和案例的学习方法，近年来多用于灾害遥感识别和灾害点预测方面[27-30]。支持向量机则多用于小面积、小比例尺（镇级）且灾害点相对较少的区域，当面对大面积及灾害点过多的情况时，支持向量机模型很难给出精确的超平面来划分灾害点，从而降低了危险性评价的精度[31-35]。随机森林是通过随机抽样，形成决策树后生成因子权重的，比较契合地质灾害这种多因素耦合作用下发生的案例[36-38]。但机器学习方法为主的耦合模型在不同地区的适用性与实用性还有待提升。

降雨作为崩塌和滑坡灾害的重要诱发因素，在常规的危险性评价工作中，一般直接导入研究区域的极端日降雨量分布数据，而忽略了降雨入渗导致地质灾害发生的滞后性，因此，日降雨量作为评价因子并不适应于连续阴雨天的地区，而应采用区域不同工况累计降雨量为数据源进行分析，以提高评价准确性。刘磊[39]基于确定性模型，通过模拟不同降雨工况开展了重庆市万州区的的滑坡灾害的危险评价研究；严明等[40]针对陕南秦巴山区，在易发性基础上基于极值降雨假设法分别计算大雨、暴雨、大暴雨和特大暴雨 4 种工况下危险性区划；但针对黄土高原地区不同降雨工况下的评价研究还亟待加强。

本研究以陕西省延安市志丹县为研究区，结合前人研究及地质环境背景，选取了高程、坡度、坡向、曲率、岩土体类型、距河流距离、距道路距离、归一化植被指数（*NDVI*）等 8 个评价因子，结合计算机编程语言 Python，使用 IVM 模型和 IVM-RF 耦合模型分别对志丹县崩塌和滑坡易发性进行了评价。绘制了各评价因子分级信息量值与灾害点占比图，对因子信息量结果与致灾影响因素进行了详细分析，依次得到该地区易发性分布。通过对比分析 2 种方法所得的易发性分区图，用 *AUC* 值（ROC 曲线下面积）对比其精度。在高准确性易发性评价基础上，基于研究区月均降雨量数据，通过皮尔逊Ⅲ型曲线估算志丹县 10，20，50 和 100 a 一遇 4 种工况下的降雨量，得到志丹县 4 种降雨工况条件下的崩塌和滑坡灾害危险性评价结果。将 RF 模型的高能性与鲁棒性与便于数据分析的 IVM 模型相结合，评价精度得到提升的同时，也降低了 RF 模型的前处理难度，并解决了 IVM 法难以获取因子权重的问题。编程插件开源，非技术人员也可直接使用。通过对危险性评价结果进行分析，可为同类型地区地质灾害危险评价工作提供参考依据。

1　研究区概况

陕西省延安市志丹县地处陕北黄土高原腹地，地理坐标为东经 108°11′56″-109°03′48″，北纬 36°21′33″-37°11′49″，总面积 3 781 km^2。地表支离破碎，梁峁起伏，沟壑纵横，可分为土石山区、黄土丘陵沟壑区和河谷区；主要出露中−新生代地层，包括白垩系、新近系和第四系；境内整体呈向 NW 缓倾的单斜构造，岩层倾角 1°～3°，无大的褶皱和断裂，但节理裂隙构造较为发育；县内中、新生代地壳垂直形变不明显，褶皱、断裂不发育，地震活动水平低。志丹县共有崩塌和滑坡灾害点 44 处，点数分布较分散，灾害点多分布在洛河、周河及杏子河及其支流中、下游两侧。县境内保安街道、顺宁镇和金丁镇灾害点密度较高。其中保安街道灾害点最多，达 15 处；最少为旦八镇，灾害点只有 1 处（图 1）。

2　数据来源与易发性评价方法

2.1　数据来源及评价单元划分

结合野外实际调查，研究区共有滑坡 27 处，崩塌 17 处；地形地貌下所有因子均基于 Alos-DEM 利用 SAGA7.0 软件地形分析模块计算和提取所得；基础地质下的岩土体类型和断层褶皱来自中国国家地质资料数据中心；气象水文下的河流来自中国 1∶400 万主要基础数据集，降雨量来自中国气象数据网；人类活动下的道路来自中国 1∶400 万主要基础数据集；植被覆盖数据（*NDVI*）来自地理空间数据云 Landsat-8 遥感影像（表 1）。

根据志丹县当地的崩塌和滑坡灾害分布范围与栅格评价单元的特点，参考汤国安等[41]的经验

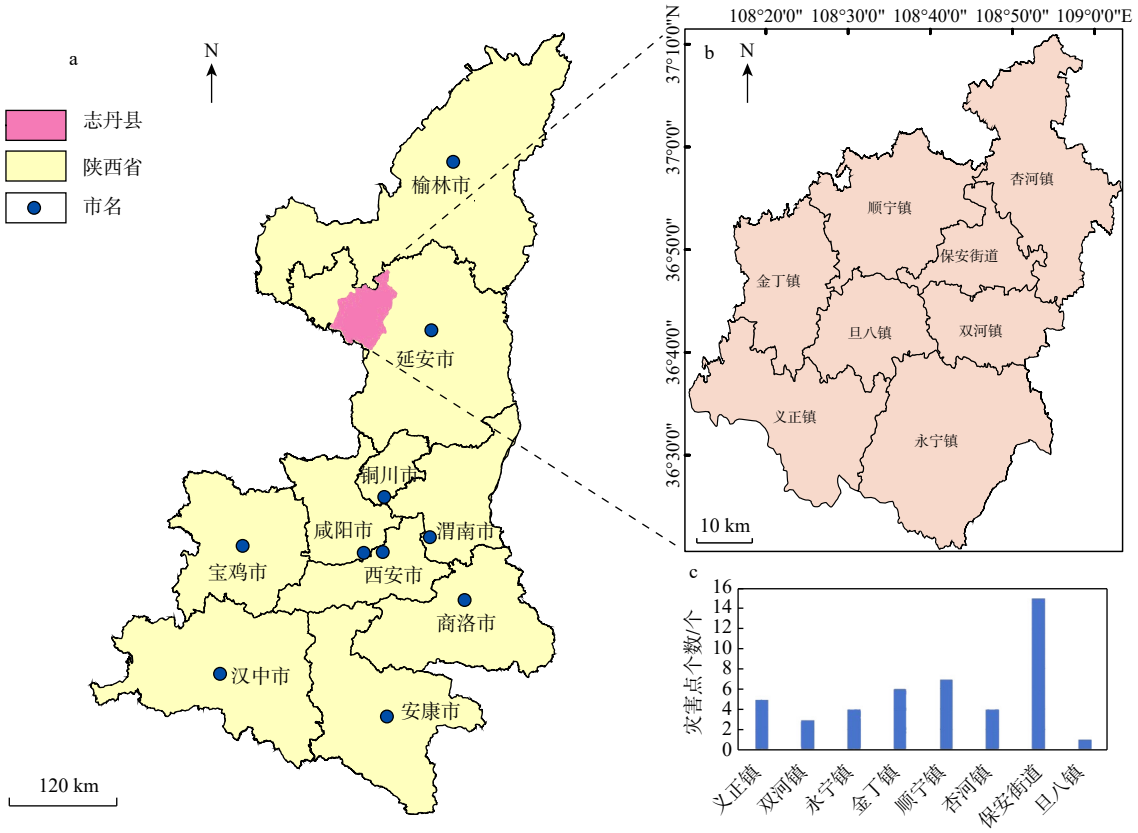


图 1　延安市志丹县地理位置示意图 (a，b) 及乡镇灾害点数量图 (c)

Fig. 1　Geographical location (a, b) of Zhidan County,Yan'an City and distribution of disaster sites in townships (c)

表 1　数据来源

Table 1　Data sources

一级因子	二级因子	数据来源	数据精度
地形地貌	高程 坡度 坡向 曲率 地形起伏度	基于SAGA7.0软件地形分析模块和Alos-DEM计算和提取	12.5 m
基础地质	岩土体类型 断层褶皱	中国国家地质资料数据中心	500 m
气象水文	河流	中国1：400万主要基础数据集	400 m
	降雨量	中国气象数据网	30 m
人类活动	道路	中国1：400万主要基础数据集	400 m
植被覆盖	*NDVI*	地理空间数据云Landsat-8遥感影像	30 m

NDVI. 归一化植被指数, 下同

公式:

$$G_S = 7.49 + 6 \times 10 - 4 \times S - 2.0 \times 10 - 6 + 2.9 \times 10 - 15 \times S^2 , \quad (1)$$

式中: G_S 为建议评价栅格单元大小; S 为评价工作比例尺分母。

经计算, 选定 25 m×25 m 大小的栅格单元 DEM 数据为评价单元。通过 ArcGIS 软件, 将整个研究区范围划分为 6 044 031 个基本栅格单元。

2.2　评价因子选取

根据志丹县特殊的地形地貌、岩土体条件、人类工程活动等, 选取高程、坡度、坡向、曲率、地形起伏度、岩土体类型、距河流距离、距道路距离和 *NDVI* 9 个指标作为易发性评价指标。以 Python 编程语言为基础, 调用 ArcPy 程序包编写了一套基于 ArcGIS 平台的因子前处理插件。通过插件的编写, 可以快速系统地完成数据分析、数据转换、数据管理和地图导出等工作, 从而高效地完成各评价因子的前处理工作, 大大提高了数据处理的速度(图 2)。

2.2.1　地形地貌因子

研究区高程 [1 058, 1 726]m, 高差较大。志丹县处于黄土高原地带, 整体高程较高, 其中 [1 377, 1 457]m 占比最大, 此高程范围植被较少, 岩土体风化侵蚀程度较强(图 3a); 坡度分布范围为 [0°, 68°], 其中 [15°, 21°)占比最大, 此坡度较平缓, 容易受降雨影响形成沟壑(图 3b); 研究区的坡向为 [−1°, 360°], 其中 [67.5°, 112.5°)占比最大, 此坡向朝东, 日照充足, 植被覆盖度较好(图 3c); 曲率分布范围为 [−29.44, 28.32], 其中 [−2.72, 0)占比最大, 此曲率表明地表向下凹陷, 地形形态较复杂(图 3d); 地形起伏度越大, 地貌差异程度越高(图 3e)。

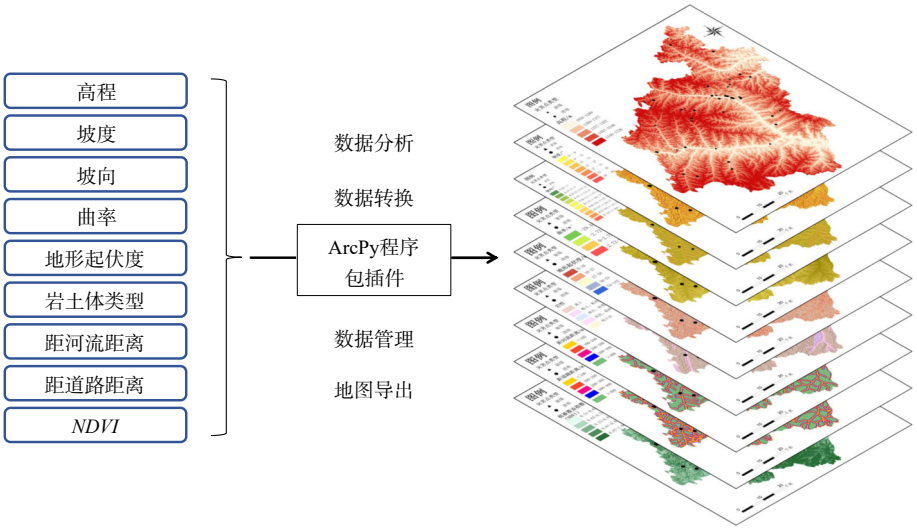


图 2　ArcPy 程序包插件处理因子图

Fig. 2　ArcPy package add-in processing factors diagram

2.2.2　基础地质因子

根据区内岩土颗粒间有无牢固联结为依据，将区内岩土体划分为岩体和土体 2 类。按照建造类型、结构类型结合岩石强度，岩体又可划分为次软–半坚硬碎屑岩组、软硬相间碎屑岩组。土体按工程地质性质又可分为砂砾类土、黄土和红黏土(图 3f)。

2.2.3　气象水文因子

距河流距离的大小对滑坡发生有很大的影响，境内主要河流为洛河、周河、杏子河，以 100 m 为间隔将其划分为 5 级。河流对地质灾害的影响主要是坡脚的侵蚀，形成坡体下部的临空面。此外，水体加速岩土体饱和，降低了坡体的抗滑力，最终造成灾害(图 3g)。

2.2.4　人类活动因子

距道路距离的远近对滑坡影响有所不同，根据研究区特殊的地形地貌，以 100 m 为间隔将其划分为 5 级，切坡建房和修建道路是常见的现象，这会导致斜坡坡脚荷载减少，从而更容易引发灾害(图 3h)。

2.2.5　植被覆盖因子

志丹县植被指数范围为 [0.21, 1.00]。其中占比最多为 [0.81, 0.87)，此区间表示植被覆盖率较大，而生长在坡面上的植被具有防止水土流失和保持稳定的作用(图 3i)。

2.3　评价因子的独立性检验

如果评价指标间线性相关，则评价结果的精确性会严重降低，因此需要保证各个指标相互独立。本研究通过皮尔逊系数(pearson correlation coefficient，简称 PCC)来判断各评价指标间的相关程度。假设各因子数据与现实结果之间的关系为 (x_i,y_i)，则相关系数计算公式为：

$$PCC=\frac{\sum_{i=1}^{n}(x_i-\bar{x})\sum_{j=1}^{n}(y_j-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{j=1}^{n}(y_j-\bar{y})^2}},\qquad(2)$$

式中：x_i 为第 i 个观测值在变量 x 上的值；y_j 为第 j 个观测值在变量 y 上的值；$\bar{x}$ 为变量 x 的平均值，即所有 x_i 值的平均；$\bar{y}$ 为变量 y 的平均值，即所有 y_j 值的平均；n 为观测值的总数；$\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}$ 为 x 的标准差；$\sqrt{\sum_{j=1}^{n}(y_i-\bar{y})^2}$ 为 y 的标准差。标准差越大，说明数据分散程度越大；反之说明数据分散程度越小；$\sum_{i=1}^{n}(x_i-\bar{x})\sum_{j=1}^{n}(y_j-\bar{y})$ 为 x 与 y 的协方差，其反映的是 2 个变量总体的误差。如果 2 个变量之间不存在线性关系，协方差为 0。

本研究中，为保证因子前处理过渡到后续的易发性评价代码的完整性，通过 Python 语言直接读取灾害点与评价因子的交集并对其进行皮尔逊系数相关性分析，生成的因子相关性热力如图 4 所示：当 $0<PCC\leqslant0.3$ 时为弱相关；$0.3<PCC\leqslant0.5$ 为中等相关；$0.5<PCC\leqslant1.0$ 为强相关[40]。由图 4 可知，坡度与地形起伏度因子之间相关性值为 0.94>0.5，因此，在后续的评价过程中对地形起伏度进行剔除，使用剩余的 8 个因子完成易发性评价流程。

2.4　崩滑易发性评价方法

2.4.1　IVM 模型

在易发性评价中，信息量模型能够将评价因子

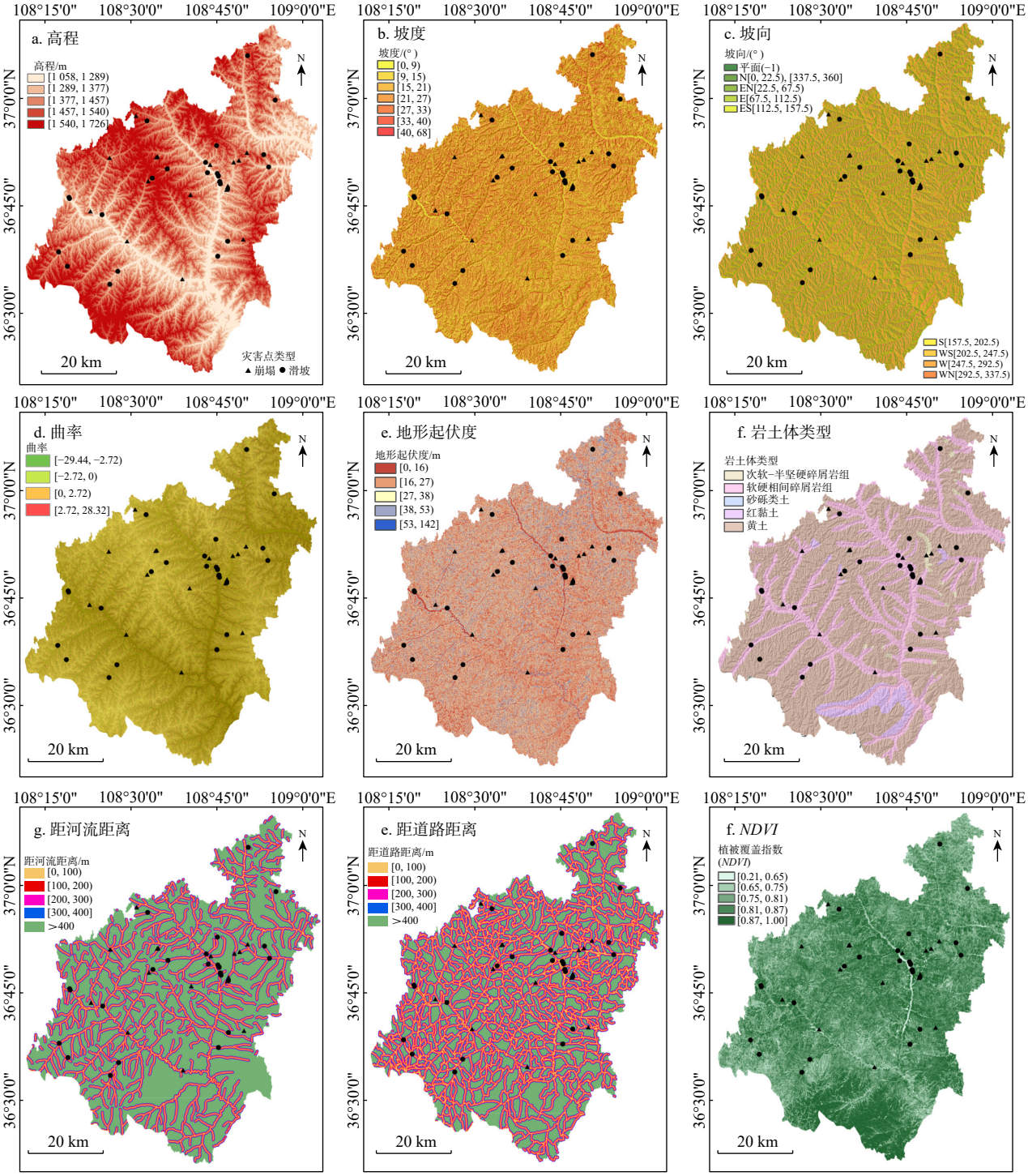


图 3　评价因子分级图

Fig. 3　Evaluation factors classification diagram

对灾害发生贡献率量化，即某个因子中的某分级的信息量值越高，在此分级的区域中灾害也就越容易发生。信息量法的计算公式如下：

$$I(X_i,Y)=\ln\frac{N_i/N}{S_i/S}, \tag{3}$$

式中：I 为信息量值；X_i 表示评价因子；Y 为单因子信息量值；Ni 为二级因子内的崩塌和滑坡灾害点个数；N 为研究区内崩塌和滑坡灾害点总数；S_i 为二级因子分级栅格所占面积；S 为评价区域栅格总面积。

2.4.2　RF 模型

随机森林模型是利用“多棵决策树”进行训练并预测的一种分类、回归预测模型。多棵分类回归决策树共同构成了“森林”。决策树是一种弱监督集成学习算法，其学习流程与分析二分类直观思维类似。

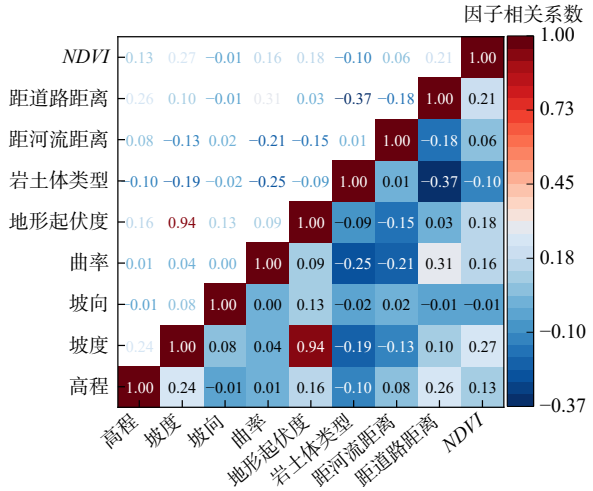


图 4　各因子相关性热力图

Fig. 4　Heatmap of factor correlations

在易发性评价工作中，随机森林模型可以通过在评价工作数据集中以抽取、处理、放回的方式得道一颗决策树，最终形成一片决策树“森林”的方式，得出各评价因子在崩、滑灾害中的贡献程度。

上述多棵树的选取方式则依据 Boostrap 自助抽样法，即多次有放回的重复抽样随机原始数据，形成 k 个不同数据集构成的决策树数据子集，且所有树都具有相同的分布；每棵树同时独立对各数据子集的特征因子，即评价因子向量中随机抽取一部分特征形成该特征随机向量；即对于第 k 棵树，生成一个随机向量 K，独立于过去生成的随机向量，并且使用训练集和 k 生长树形成分类器。在生成大量树后，筛选出最适合的向量的过程为随机森林学习。在崩塌和滑坡易发性分析中，其整合输出的最终特征向量代表因子的权重(图 5)。

2.4.3　IVM-RF 耦合模型的构建

首先构建样本数据，灾害点即为正样本，在 IVM 模型易发性分区的低和极低易发区处随机生成的非灾害点即为负样本，提取样本的 8 个因子属性，组成正负样本合集。将所有样本数据的多数作为训练集，剩余少数作为测试集，进行随机森林模型运算。通过 Python 语言编码设置，自动进行循环迭代对比选择，调整相应参数，可自动选取最小误差下的最大特征值以及最大特征值下的决策树数量，从而达到较高的精度，以及较好的预测效果(图 6)。最后将 IVM-RF 模型得出的预测结果进行易发性分区。

3　易发性评价结果

3.1　基于 IVM 模型易发性评价

3.1.1　IVM 模型处理结果

将评价因子的分级，通过 GIS 软件的“多值提取至点”工具，提取进灾害点数据表中，并结合每类因子下不同分级的栅格数量，最终逐一计算出研究区内各评价因子的信息量值(表 2)。

3.1.2　因子信息量结果与致灾影响因素分析

高程的信息量值和灾害点占比最高为 [1 058, 1 289)m，表明志丹县 [1 058, 1 289)m 高程范围相比于其他高程更可能发生崩塌和滑坡灾害，这是此范围内大量的切坡修路导致的(图 7a)。坡度范围在 9°以下及 [15, 40)°之间的崩滑灾害数量明显多于其他坡度，而对于志丹县来说坡度为 [33, 40)°时相较于其他坡度更容易发生崩滑灾害，这是由于人们大量切坡建房(图 7b)。崩滑灾害坡向分布只有 EN 分布最少，从信息量来看，ES 和 S 2 个坡向引发崩滑灾

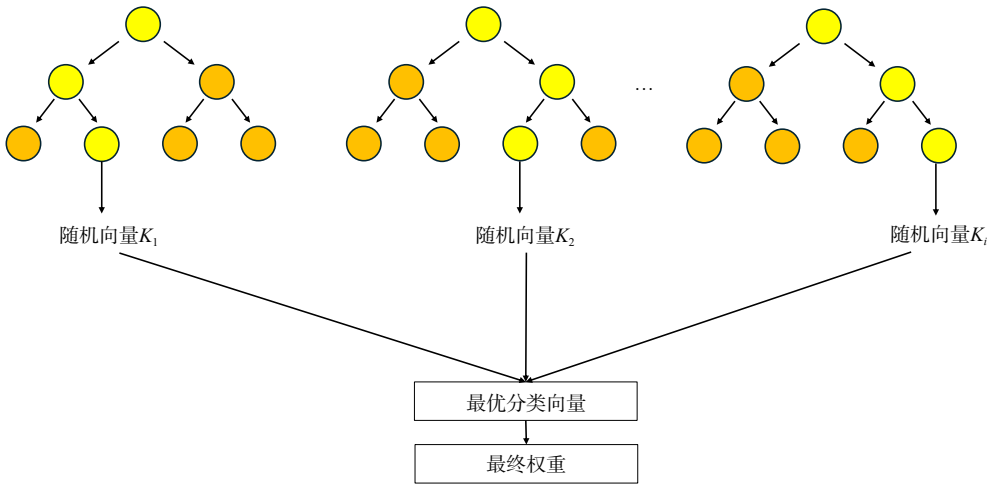


图 5　随机森林计算权重流程图

Fig. 5　Flowchart of random forest weight calculation

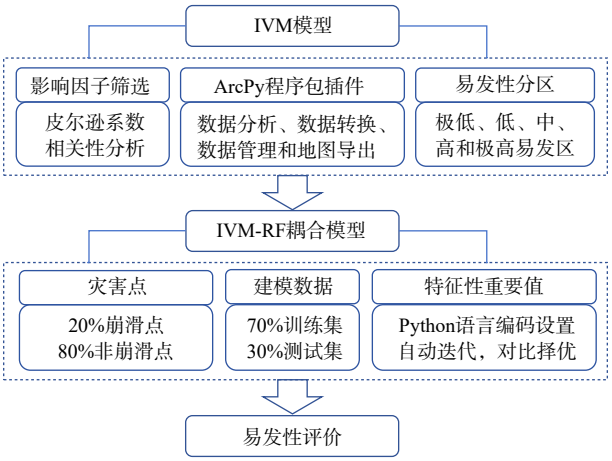


图 6　IVM-RF 耦合模型流程图

Fig. 6　Flowchart of IVM-RF coupled model

害的概率更大，这是因为光照的不均匀导致土体差异化引起的(图 7c)。曲率位于 [−29.44, −2.72)的信息量值最高，表明研究区受人类活动的影响，凹形坡更易发生崩滑灾害(图 7d)。岩土体类型位于软硬相间碎屑岩组的崩滑灾害点数量和信息量值均最高，表明研究区内软硬相间碎屑岩组更易发生灾害，地表水冲刷导致岩土体松散不堪(图 7e)。研究区内距河流距离<100 m 的区域，信息量值和灾害点占比都最大，表明距离河流越近的地方越容易发生灾害(图 7f)。研究区有 36% 的灾害分布于距道路距离<100 m 的区域，且随着距道路距离增大，发生灾害的可能性逐渐减少(图 7g)。灾害的 *NDVI* 在 [0.81, 0.87)区间的数量占研究区灾害总数的 39%，*NDVI* 为 [0.21, 0.56)区间的信息量值最高，发生灾害的可能性更大，说明区内植被的影响不可忽略(图 7h)。

3.2　基于 IVM-RF 耦合模型易发性评价

首先建立正样本和负样本合集，正样本是志丹县的崩塌滑坡灾害点(44 个)，负样本为没发生崩塌滑坡的样本点。为满足样本容量与模型计算的需求，负样本按照正负样本比例 1∶4 在 IVM 模型易发性评价结果的极低和低易发区随机生成 176 个点。数据样本包含训练集和测试集，训练集和测试集占比分别为 70% 和 30%；标记灾害点为 1，非灾害点为 0。

影响随机森林模型的参数有 2 个：最大特征数和决策树数量。通过 Python 语言编码设置，可自动选取最大特征值和决策树数量，最后选择的最优组合为：最大特征数为 3，决策树数量为 100。将 IVM-RF 模型得出的预测结果进行特征重要性排序(图 8)。

由图 8 可知：坡度对灾害发生的影响最大，特征重要性为 0.161 7；曲率的影响最小，特征重要性为 0.060 8；8 个评价因子对崩滑灾害都有一定的影响。对于黄土高原，小于 45°的斜坡在连续降雨情况下，坡面被水流侵蚀，水分渗入加快，内部逐渐不稳，极易发生崩滑灾害。

3.3　不同评价模型评价精度对比

如图 9，表 3 所示，研究区按照自然间断法划分为极低、低、中、高和极高易发区。极高易发区和高易发区主要位于志丹县中部的沿河区域。IVM 模型结果显示：中易发区的灾害点占比和面积占比均最大，分别为 56.82% 和 41.07%；从灾害点密度来看，从低易发区到极高易发区呈逐渐增大趋势，但极低易发区的 0.006 0 个/km^2 高于低易发区的 0.001 9 个/km^2(表 3)。信息量模型的灾害点密度数据变化不规律，而 IVM-RF 耦合模型得到的易发性分区的结果分布就较为合理。极低易发区和低易发区主要分布在高海拔远离河道的地区，该区域植被覆盖率较高且受地表流水影响较小；高和极高易发区主要分布在河谷阶地地区，该区域受流水的侵蚀导致崩塌和滑坡更易发生。从灾害点占比来看，极低易发区到极高易发区的占比从 2.27% 增加到 36.36%，变化趋势明显。面积占比也从极低、低、中、高、极高易发区逐渐降低；灾害点密度从极低易发区的 0.000 5 个/km^2 逐级增加到极高易发区的 0.331 3 个/km^2。

ROC 曲线和 *AUC* 值可作为模型准确性评估的重要量化指标。如图 10 所示，2 种模型的 *AUC* 值都大于 0.5，说明 2 种模型评价结果都可行；而 IVM-RF 模型的 *AUC* 值 0.87 大于 IVM 模型的 0.66，表明 IVM-RF 模型的精度更高、预测能力更好，能够为志丹县崩滑灾害防治提供数据支持。后续危险性评价是以 IVM-RF 耦合模型的易发性分区结果为基础进行的。

4　地质灾害危险性评价

4.1　危险性灾害诱发因素与计算流程

地质灾害危险性评价是在易发性评价基础上叠加时间概率问题以及如降雨和地震等诱发因子。根据国家标准《中国地震动参数区划图》(GB18306-2015)[42]，志丹县地震动峰值加速度 *PGA*=0.05 *g*，反应谱特征周期为 0.35 s。因此，危险性评价时不考虑地震影响，只考虑降雨量。通过近 22 a 的月最大降雨量均值计算出皮尔逊Ⅲ型曲线的变差系数 C_v 和偏态系数 C_s，然后查阅皮尔逊Ⅲ型频率曲线离均系

表 2　评价因子分级及信息量值

Table 2　Classification and IV values of evaluation factors

因子	分级	灾害点个数/个	分级栅格数量/个	分级栅格面积/m^2	信息量值
高程/m	[1058，1289)	16	608485	380302991	1.284253630
	[1289，1377)	7	1317352	823345276	−0.314832308
	[1377，1457)	8	1630754	1019220989	−0.394719147
	[1457，1540)	12	1552184	970115147	0.060125067
	[1540，1726]	1	935256	584534972	−1.918183422
坡度/(°)	[0，9)	7	953159	595724375	0.008765236
	[9，15)	4	1242953	776845625	−0.816314099
	[15，21)	10	1303633	814770625	0.052311650
	[21，27)	8	1086974	679358750	0.010925392
	[27，33)	7	767382	479613750	0.225562246
	[33，40)	7	474786	296741250	0.705682791
	[40，68]	1	215144	134465000	−0.448670753
坡向/(°)	N[0，22.5)，[337.5，360]	6	665049	415655625	0.214535566
	EN[22.5，67.5)	1	748633	467895625	−1.695612058
	E[67.5，112.5)	3	944041	590025625	−0.828920490
	ES[112.5，157.5)	8	667542	417213750	0.498476051
	S[157.5，202.5)	9	664204	415127500	0.621272065
	WS[202.5，247.5)	6	673845	421153125	0.201396174
	W[247.5，292.5)	7	954627	596641875	0.007226279
	WN[292.5，337.5)	4	723506	452191250	−0.275177659
	平面(−1)	0	2584	1615000	0
曲率	[−29.44，−2.72)	7	755668	472292335	0.240945191
	[−2.72，0)	19	2350223	1468889970	0.104809896
	[0，2.72)	14	2187230	1367018525	−0.128696869
	[2.72，28.32]	4	750910	469318545	−0.312354189
岩土体类型	黄土	14	4131521	2582200634	−0.764706755
	红黏土	0	205738	128586332	0
	砂砾类土	0	1074	670650	0
	软硬相间碎屑岩组	29	1644572	1027857804	0.884696905
	次软−半坚硬碎屑岩组	1	61126	38203954	0.809694166
距河流距离/m	[0，100)	25	1192609	745380334	1.057614410
	[100，200)	5	1010958	631848502	−0.386578551
	[200，300)	4	875906	547441357	−0.466327796
	[300，400]	0	751641	469775319	0
	>400	10	2212917	1383073863	−0.476845456
距道路距离/m	[0，100)	16	663742	414839058	1.197331278
	[100，200)	8	641810	401131058	0.537786529
	[200，300）	5	626302	391438738	0.092242079
	[300，400]	3	609756	381097518	−0.391809817
	>400	12	3502421	2189013003	−0.753666199
NDVI	[0.21，0.65)	5	126663	79164373	1.690544739
	[0.65，0.75)	10	790323	493951624	0.552780697
	[0.75，0.81)	8	1503227	939516840	−0.313291013
	[0.81，0.87)	17	2009315	1255822463	0.150300547
	[0.87，1.00]	4	1614503	1009064074	−1.077850971

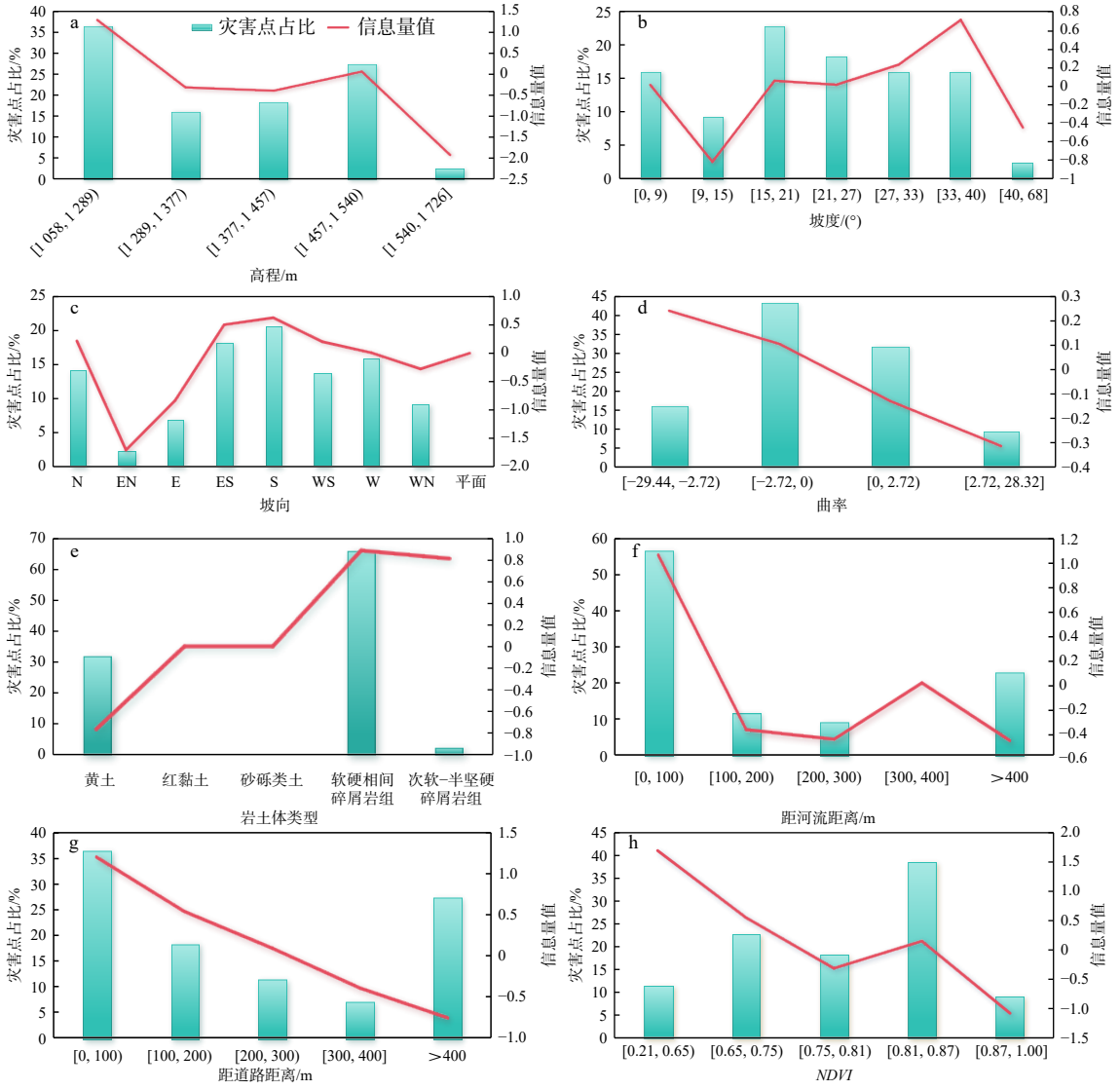


图 7　各因子分级信息量值与灾害点占比图

Fig. 7　Information values and disaster point ratio of factor classifications

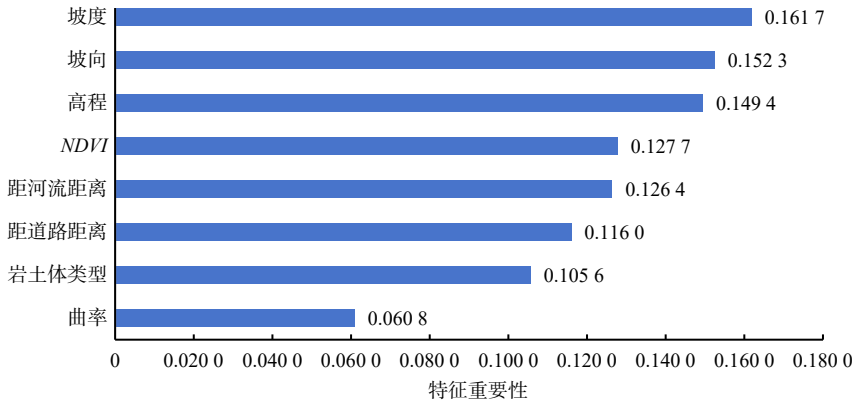


图 8　特征重要性图

Fig. 8　Feature importance plot

数(ϕ)表，得出不同概率下月最大降雨量。

由图 11 可知：志丹县 2001−2022 年月均降雨量较大主要发生在汛期 7，8 和 9 月。

由图 12 可知，志丹县 2021−2022 年的月最大降雨量忽增忽减，变化不规律，但峰值出现在 2021 年 9 月。

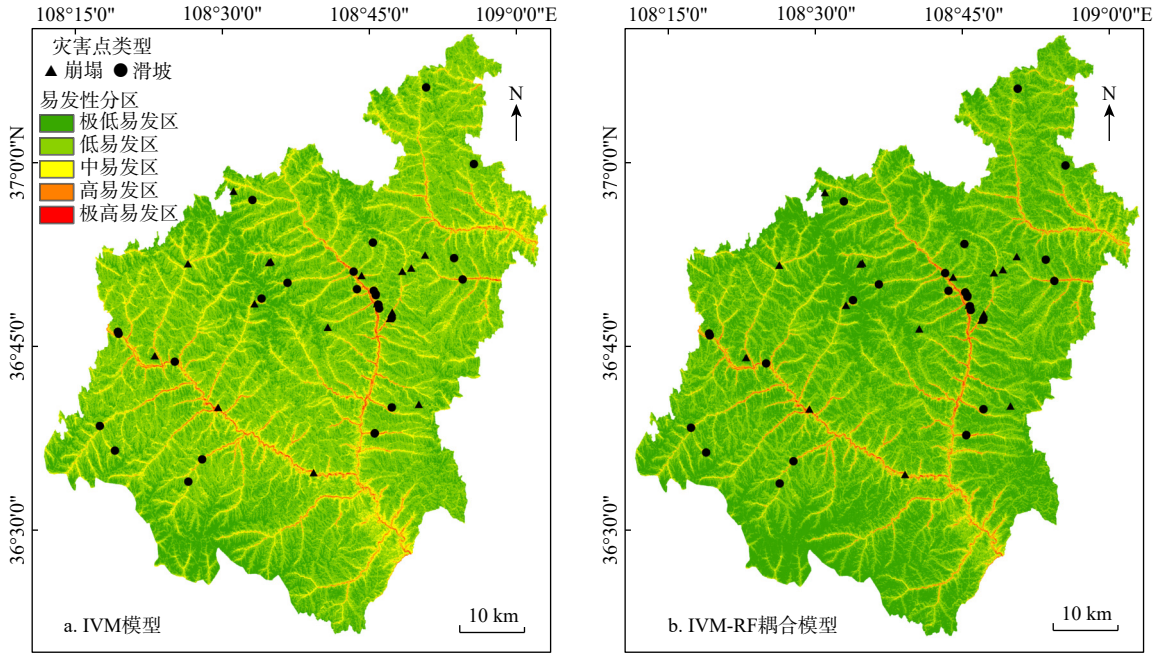


图 9　IVM 模型 (a) 和 IVM-RF 耦合模型 (b) 的崩滑易发性分区图

Fig. 9　Landslide susceptibility zoning of IVM model (a) and IVM-RF coupled model (b)

表 3　IVM 模型和 IVM-RF 耦合模型的崩滑易发性分区统计

Table 3　Statistics of landslide susceptibility zoning for IVM model and IVM-RF coupled model

评价模型	易发性分区	灾害点个数/个	灾害点占比/%	面积/km^2	面积占比/%	灾害点密度/(个·km^{-2})
IVM模型	极低易发区	6	13.64	1004.2651	26.56	0.0060
	低易发区	2	4.55	1075.2974	28.44	0.0019
	中易发区	25	56.82	1552.7663	41.07	0.0161
	高易发区	5	11.36	130.1442	3.44	0.0384
	极高易发区	6	13.64	18.5271	0.49	0.3239
IVM-RF耦合模型	极低易发区	1	2.27	1830.3821	48.41	0.0005
	低易发区	5	11.36	1237.8994	32.74	0.0040
	中易发区	10	22.73	403.1599	10.66	0.0248
	高易发区	12	27.27	261.2671	6.91	0.0459
	极高易发区	16	36.36	48.2915	1.28	0.3313

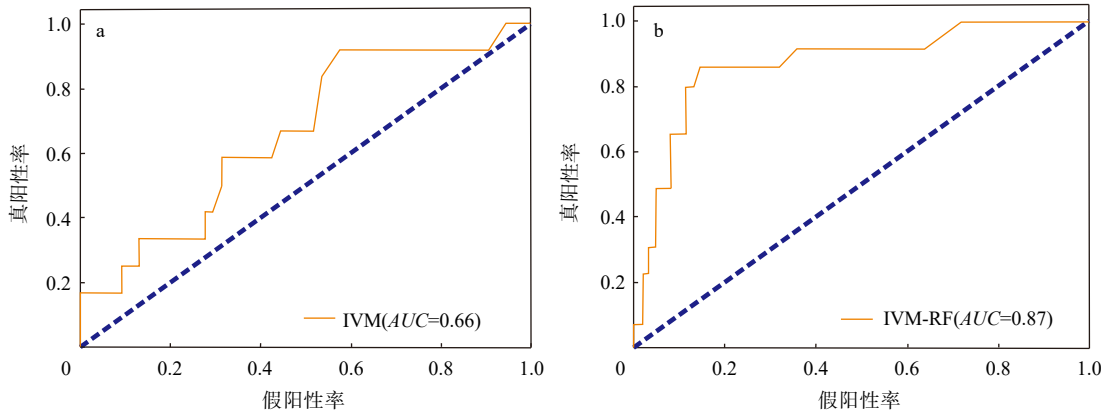


图 10　IVM 模型 (a) 和 IVM-RF 耦合模型 (b) 的 ROC 曲线图（*AUC* 为 ROC 曲线下面积）

Fig. 10　ROC curves of IVM model (a) and IVM-RF coupled model (b)

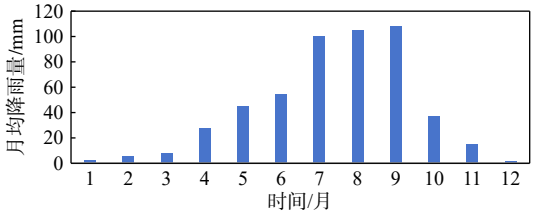


图 11　2001−2022 年研究区月均降雨量

Fig. 11　Monthly average rainfall (2001−2022) in the study area

通过图 12 计算出月最大降雨量均值 $\bar{r}$ =146.04, 代入公式(4)和(5), 计算 C_s, C_v:

$$C_v=\frac{1}{\bar{r}}\sqrt{\frac{\sum_{i=1}^{22}(r_i-\bar{r})^2}{22-1}}=0.261\,451\,433,\quad(4)$$

$$C_s=\frac{\sum_{i=1}^{22}\left(\frac{r_i}{\bar{r}}-1\right)^3}{(22-3)C_v^3}=0.655\,216\,656,\quad(5)$$

式中: r_i 为第 i 个降雨量, mm; $\bar{r}$ 为月最大降雨量均值, mm。

查阅皮尔逊Ⅲ型频率曲线离均系数(ϕ)表, 不同降雨工况(概率)下的月最大降雨量如表 4 所示。

4.2　危险性评价

基于皮尔逊Ⅲ型曲线估算出 4 种不同工况下降雨所导致的危险性指数:

$$H_i=\frac{Y_i}{Y_{max}}P_i,\quad(6)$$

$$P_i=\frac{L_i}{L_{max}},\quad(7)$$

式中: H_i 为第 i 个评价单元危险性指数; Y_i 为第 i 个评价单元易发性指数; Y_{max} 为最大易发性指数; P_i 为第 i 个评价单元灾害发生概率; L_i 为第 i 个评价单元月降雨量, mm; L_{max} 为 2001−2022 年月最大降雨量, mm。

4.3　评价结果

研究区 2001−2022 年月最大降雨量为 233.66 mm (2021 年 9 月)。利用 ArcGIS 在易发结果上叠加 4 种降雨工况的灾害发生概率, 计算出危险性指数, 通过自然断点法将其分成 5 个等级(表 5), 根据表 5 绘制出 4 种不同降雨工况的地质灾害危险性分区图(图 13)。

根据图 13, 14 和表 6: 从 10, 20, 50 和 100 a 一遇降雨工况来看, 极低危险区面积占比整体逐渐递减, 这是因为随着降雨量的增大, 雨水入渗和地表侵蚀加剧, 土体受侵蚀影响严重使得危险性提升, 从而使极低危险区面积下降, 符合实际规律; 低危险区面积呈波动状态, 但整体呈递增趋势, 这是因为随着不同降雨工况的降雨量增大, 部分极低危险区提升为

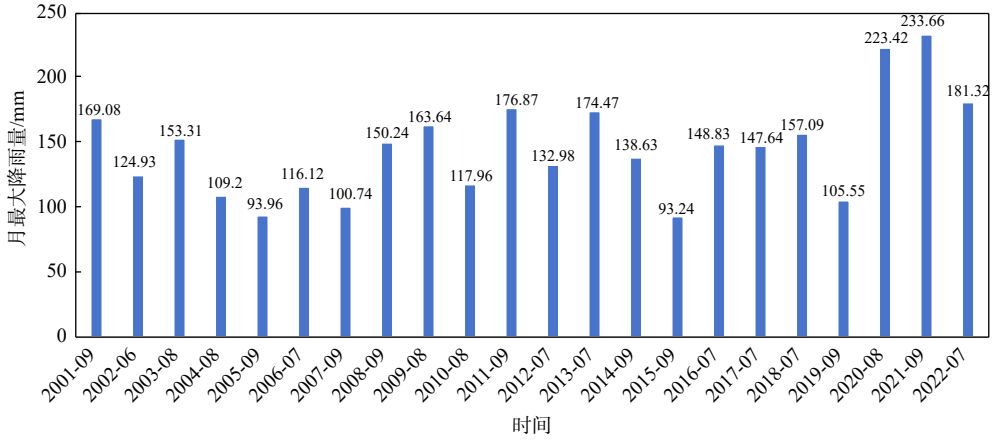


图 12　2001−2022 年研究区各年月最大降雨量图

Fig. 12　Annual maximum monthly rainfall (2001−2022) in the study area

表 4　延安市志丹县不同降雨工况（概率）的月最大降雨量

Table 4　Monthly maximum rainfall at different probabilities in Zhidan County, Yan'an City

降雨工况概率	离均系数ϕ值	月最大降雨量$R=\bar{r}(1+\phi C_v)$/mm
10 a一遇 10%	1.33	196.822 548 5
20 a一遇 5%	1.82	215.531 908 4
50 a一遇 2%	2.40	237.677 681 5
100 a一遇 1%	2.81	253.332 452 0

表 5　危险性等级分级

Table 5　Hazard level classification

危险性等级	极高危险区	高危险区	中危险区	低危险区	极低危险区
危险性指数 H_i	>0.7	(0.55，0.7]	(0.45，0.55]	[0.33，0.45]	<0.33

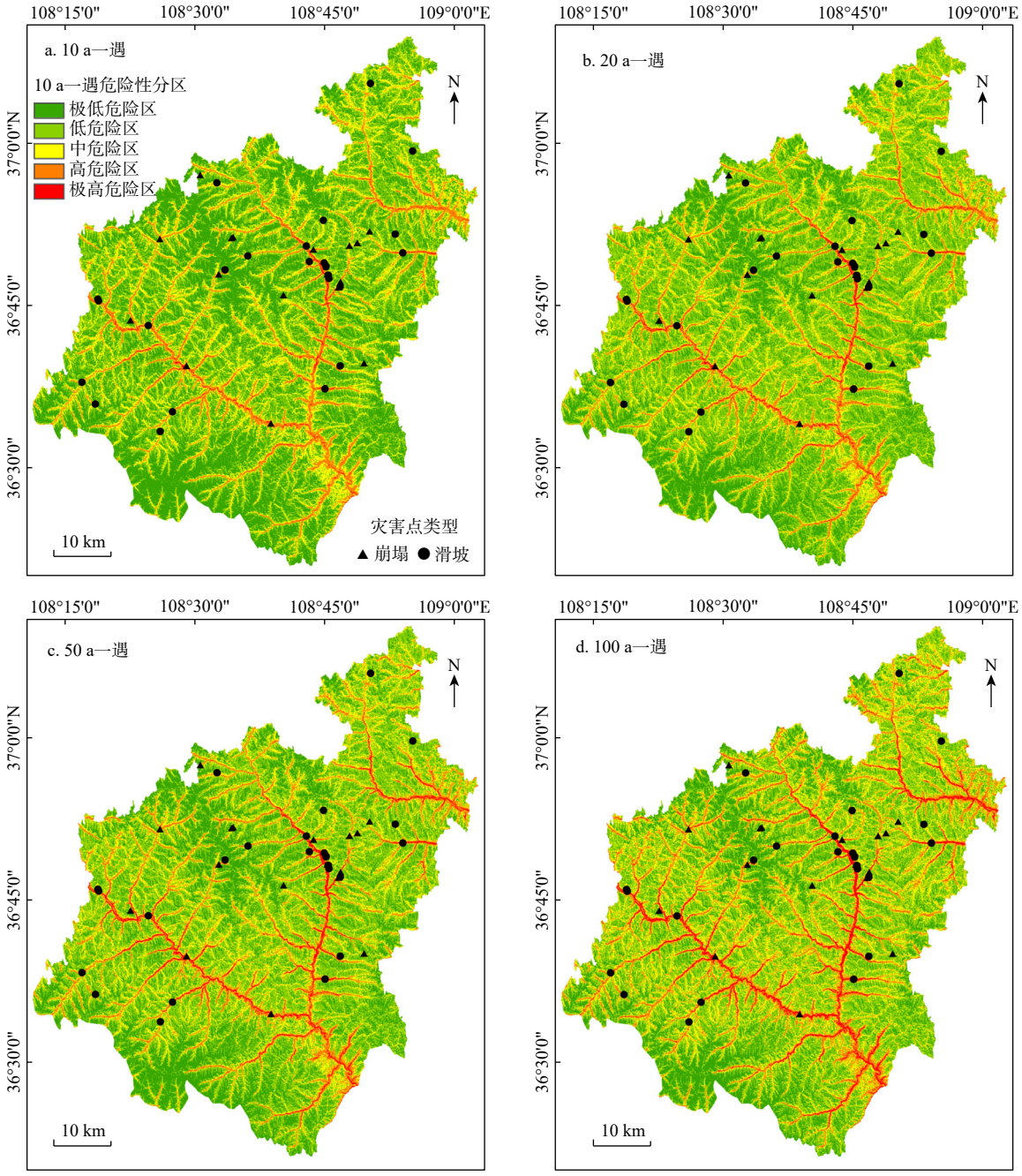


图 13　4 种降雨工况下危险性分区图

Fig. 13　Hazard zoning map under four rainfall conditions

低危险区，同时原有的部分低危险区提升为中、高或极高危险区，从而造成低危险区面积忽增忽降；中、高和极高危险区面积都逐渐递增，尤其是极高危险区面积递增幅度较大，从 10 a 一遇降雨工况的 0.79% 增加到 100 a 一遇降雨工况的 3.63%。

综上所述，随着不同降雨工况的降雨量增加，极低和低危险区面积总和减少，中、高和极高危险区面积增加较大，极高危险区面积增幅最为明显。该结果符合实际情况，表明本研究在统计月降雨量情况下通过不同降雨工况得到的危险性分区比较合理。

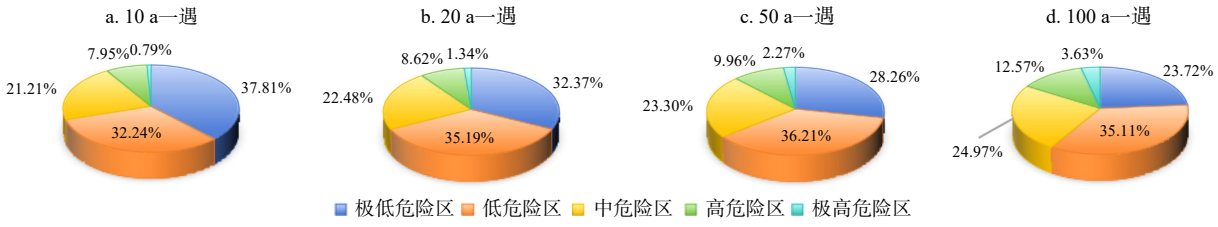


图 14　4 种降雨工况危险性分区面积占比图

Fig. 14　Area proportion of hazard zoning under four rainfall conditions

表 6　不同降雨工况下危险性分区面积统计

Table 6　Hazard zoning statistics under different rainfall conditions　面积/km²

降雨工况	极低危险区	低危险区	中危险区	高危险区	极高危险区
10 a一遇	1 429.47	1 219.17	801.90	300.67	29.80
20 a一遇	1 223.97	1 330.22	850.05	326.10	50.65
50 a一遇	1 068.58	1 369.08	881.11	376.57	85.66
100 a一遇	896.72	1 327.48	944.21	475.43	137.50

5　结　论

(1)以延安市志丹县域内崩塌和滑坡地质灾害点为研究对象, 以栅格单元作为评价单元, 选取了高程、坡度、坡向、曲率、岩土体类型、距河流距离、距道路距离、*NDVI* 8 个易发性评价因子。通过信息量模型, 绘制了各因子分级信息量值与灾害点占比图, 并对因子信息量结果与致灾影响因素进行了分析。对比分析 2 种模型的易发性分区图、灾点密度和 ROC 曲线图, 得到了 IVM 模型的 *AUC* 值为 0.66, 小于 IVM-RF 模型的 0.87。相比于 IVM 模型, IVM-RF 模型的分区结果更合理, 其中极高、高和中易发区覆盖了研究区的洛河、周河、杏子河等河谷边缘区域, 预测效果较好。

(2)通过调用 ArcPy 程序包编写了一套基于 ArcGIS 平台的因子前处理插件。通过插件可以快速系统地完成数据分析、数据转换、数据管理和地图导出等工作, 从而高效地完成各评价因子的前处理工作, 大大提高了数据处理的速度。通过 Python 语言编码设置, 改进了 IVM-RF 耦合模型评价方法, 实现了能够自动进行循环迭代对比选择, 自动选取最小误差下的最大特征值, 从而解决了耦合模型评价过程繁琐的问题。

(3)通过 IVM-RF 耦合模型的易发性区划结果, 根据延安市志丹县多年的降雨量数据, 用皮尔逊Ⅲ型曲线估算研究区 10, 20, 50 和 100 a 一遇的降雨量, 并进行危险性分区。从 10, 20, 50 和 100 a 一遇降雨工况来看, 极低危险区面积整体逐渐递减; 低危险区面积忽增忽降呈波动状态; 中、高和极高危险区面积都逐渐递增, 且递增幅度较大, 尤其是极高危险区面积, 从 10 a 一遇到 100 a 一遇降雨工况分别为: 0.79%, 1.34%, 2.27% 和 3.63%。本研究可为志丹县防灾减灾和风险管控提供数据支持, 为当地以及相似地区危险性评价提供了参考依据。

所有作者声明不存在利益冲突。

The authors declare that no competing interests exist.

参考文献:

[1] 谢婉丽, 王延寿, 马中豪, 等. 黄土湿陷机理研究现状及发展趋势 [J]. 现代地质, 2015, 29(2): 397-407.
XIE W L, WANG Y S, MA Z H, et al. Research status and prospect of loess collapsibility mechanism[J]. Geoscience, 2015, 29(2): 397-407. (in Chinese with English abstract)

[2] 滕宏泉, 谢婉丽, 盖海龙, 等. 分形分维理论在地质灾害发育及空间分布规律中的应用: 以长安区滑坡、崩塌地质灾害为例 [J]. 地质灾害与环境保护, 2016, 27(1): 44-50.
TENG H Q, XIE W L, GAI H L, et al. Application of fractal dimension theory in geological hazard development and spatial distribution law: Taking Chang'an district landslide and collapse as an example[J]. Journal of Geological Hazards and Environment Preservation, 2016, 27(1): 44-50. (in Chinese with English abstract)

[3] 谢婉丽, 滕宏泉, 杜蕾, 等. 大西安地区滑坡分类及其特征分析 [J]. 地球环境学报, 2018, 9(1): 79-88.
XIE W L, TENG H Q, DU L, et al. An analysis of landslide classifi cations and characteristics in Xi'an[J]. Journal of Earth Environment, 2018, 9(1): 79-88. (in Chinese with English abstract)

[4] 赵欢, 田伟平, 齐洪亮, 等. 陕西公路边坡灾害危险性分区研究 [J]. 灾害学, 2016, 31(4): 75-81.
ZHAO H, TIAN W P, QI H L, et al. Research on the risk regionalization of highway slope disasters in Shaanxi Province[J]. Journal of Catastrophology, 2016, 31(4): 75-81. (in Chinese with English abstract)

[5] ZHENG H, DING M T. Spatiotemporal changes of landslide susceptibility in response to rainfall and its future prediction: A case study of Sichuan Province, China[J]. Ecological Informatics, 2024, 84: 102862.

[6] YUAN X Y, LIU C, NIE R H, et al. A comparative analysis of certainty factor-based machine learning methods for collapse and landslide susceptibility mapping in Wenchuan County, China[J].

Remote Sensing, 2022, 14(14): 3259.

[7] BARREDO J, BENAVIDES A, HERVÁS J, et al. Comparing heuristic landslide hazard assessment techniques using GIS in the Tirajana Basin, Gran Canaria Island, Spain[J]. International Journal of Applied Earth Observation and Geoinformation, 2000, 2(1): 9-23.

[8] 刘帅, 朱杰勇, 杨得虎, 等. 不同降雨工况条件下的崩滑地质灾害危险性评价 [J]. 地质科技通报, 2024, 43(2): 253-267.
LIU S, ZHU J Y, YANG D H, et al. Geological hazard risk assessment of collapse and landslide under different rainfall conditions[J]. Bulletin of Geological Science and Technology, 2024, 43(2): 253-267. (in Chinese with English abstract)

[9] ALAMREW B T, KASSAWMAR T, MENGSTIE L, et al. Combined GIS, FR and AHP approaches to landslide susceptibility and risk zonation in the Baso Liben district, northwestern Ethiopia[J]. Quaternary Science Advances, 2024, 16: 100250.

[10] BISWAKARMA P, JOSHI V, ABDO H G, et al. An integrated quantitative and qualitative approach for landslide susceptibility mapping in West Sikkim district, Indian Himalaya[J]. Geomatics, Natural Hazards and Risk, 2023, 14(1).

[11] SONG R X, WANG S Z, SONG L F, et al. Landslide susceptibility assessment based on slope unit and information value method in Changbai Mountain district[J]. Journal of Civil and Transportation Engineering, 2024, 1(2): 39-44.

[12] BHANDARI B P, DHAKAL S, TSOU C Y. Assessing the prediction accuracy of frequency ratio, weight of evidence, Shannon entropy, and information value methods for landslide susceptibility in the siwalik hills of Nepal[J]. Sustainability, 2024, 16(5): 2092.

[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, 20(3): 49-51.
XING Q J, ZHAO C Y, GAO K C, et al. Logical regression analysis on the hazard of landslide based on GIS[J]. Geography and Geo-Information Science, 2004, 20(3): 49-51. (in Chinese with English abstract)

[14] 许冲, 戴福初, 姚鑫, 等. 基于 GIS 与确定性系数分析方法的汶川地震滑坡易发性评价 [J]. 工程地质学报, 2010, 18(1): 15-26.
XU C, DAI F C, YAO X, et al. GIS platform and certainty factor analysis method based Wenchuan earthquake-induced landslide susceptibility evaluation[J]. Journal of Engineering Geology, 2010, 18(1): 15-26. (in Chinese with English abstract)

[15] 李郎平, 兰恒星, 郭长宝, 等. 基于改进频率比法的川藏铁路沿线及邻区地质灾害易发性分区评价 [J]. 现代地质, 2017, 31(5): 911-929.
LI L P, LAN H X, GUO C B, et al. Geohazard susceptibility assessment along the Sichuan-Tibet railway and its adjacent area using an improved frequency ratio method[J]. Geoscience, 2017, 31(5): 911-929. (in Chinese with English abstract)

[16] KLAI A, KATLANE R, HADDAD R, et al. Landslide susceptibility mapping by frequency ratio and fuzzy logic approach: A case study of Mogods and Hedil (Northern Tunisia)[J]. Applied Geomatics, 2024, 16(1): 91-109.

[17] BANDA K, PIETERSEN K, HAMUTOKO J, et al. Using a comparative of DRASTIC and Bayesian weights of evidence approach to assess transboundary aquifer vulnerability in a data scarcity region: Tuli-Karoo aquifer[J]. Journal of Hydrology: Regional Studies, 2024, 55: 101930.

[18] 李嘉昊, 谢婉丽, 严明, 等. 基于 PCA 与改进 AHP-CRITIC 法的地质灾害风险评价研究: 以神木市为例 [J]. 地球环境学报, 2023, 14(4): 472-487.
LI J H, XIE W L, YAN M, et al. Research on geological hazard risk assessment based on PCA and improved AHP-CRITIC method: A case study of Shenmu, Shaanxi Province[J]. Journal of Earth Environment, 2023, 14(4): 472-487. (in Chinese with English abstract)

[19] CHANG L L, ZHANG R, WANG C S. Evaluation and prediction of landslide susceptibility in Yichang section of Yangtze River Basin based on integrated deep learning algorithm[J]. Remote Sensing, 2022, 14(11): 2717.

[20] 樊芷吟, 苟晓峰, 秦明月, 等. 基于信息量模型与 Logistic 回归模型耦合的地质灾害易发性评价 [J]. 工程地质学报, 2018, 26(2): 340-347.
FAN Z Y, GOU X F, QIN M Y, et al. Information and logistic regression models based coupling analysis for susceptibility of geological hazards[J]. Journal of Engineering Geology, 2018, 26(2): 340-347. (in Chinese with English abstract)

[21] 罗路广, 裴向军, 黄润秋, 等. GIS 支持下 CF 与 Logistic 回归模型耦合的九寨沟景区滑坡易发性评价 [J]. 工程地质学报, 2021, 29(2): 526-535.
LUO L G, PEI X J, HUANG R Q, et al. Landslide susceptibility assessment in Jiuzhaigou scenic area with GIS based on certainty factor and logistic regression model[J]. Journal of Engineering Geology, 2021, 29(2): 526-535. (in Chinese with English abstract)

[22] ZENG B, CHEN X X. Assessment of shallow landslide susceptibility using an artificial neural network[J]. Arabian Journal of Geosciences, 2021, 14(6): 499.

[23] LUCCHESE L V, DE OLIVEIRA G G, PEDROLLO O C. Investigation of the influence of nonoccurrence sampling on landslide susceptibility assessment using artificial neural networks[J]. Catena, 2021, 198: 105067.

[24] HAN X M, ZHAO X, WU Y C, et al. A least squares-support vector machine for learning solution to multi-physical transient-state field coupled problems[J]. Engineering Applications of Artificial Intelligence, 2024, 138: 109321.

[25] ZUO L L, LIU G H, FANG Z, et al. Assessing the destabilization risk of ecosystems dominated by carbon sequestration based on interpretable machine learning method[J]. Ecological Indicators, 2024, 167: 112593.

[26] MELATI D N, UMBARA R P, ASTISIASARI A, et al. A comparative evaluation of landslide susceptibility mapping using machine learning-based methods in Bogor area of Indonesia[J]. Environmental Earth Sciences, 2024, 83(3): 86.

[27] 刘艺梁, 殷坤龙, 刘斌. 逻辑回归和人工神经网络模型在滑坡灾害空间预测中的应用 [J]. 水文地质工程地质, 2010, 37(5): 92-96.
LIU Y L, YIN K L, LIU B. Application of logistic regression and artificial neural networks in spatial assessment of landslide haz-

ards[J]. Hydrogeology & Engineering Geology, 2010, 37(5): 92-96. (in Chinese with English abstract)

[28] WANG W, WU G F, SONG X Y. The application of neural networks to comprehensive prediction by seismology prediction method[J]. Acta Seismologica Sinica, 2000, 13(2): 210-215.

[29] 田乃满, 兰恒星, 伍宇明, 等. 人工神经网络和决策树模型在滑坡易发性分析中的性能对比 [J]. 地球信息科学学报, 2020, 22(12): 2304-2316.
TIAN N M, LAN H X, WU Y M, et al. Performance comparison of BP artificial neural network and CART decision tree model in landslide susceptibility prediction[J]. Journal of Geo-Information Science, 2020, 22(12): 2304-2316. (in Chinese with English abstract)

[30] SHAHIN A M , JAKSA B M , MAIER R H . Artificial neural network applications in geotechnical engineering[J]. Australian Geomechanics, 2001, 36(1): 49-62.

[31] 黄发明, 殷坤龙, 蒋水华, 等. 基于聚类分析和支持向量机的滑坡易发性评价 [J]. 岩石力学与工程学报, 2018, 37(1): 156-167.
HUANG F M, YIN K L, JIANG S H, et al. Landslide susceptibility assessment based on clustering analysis and support vector machine[J]. Chinese Journal of Rock Mechanics and Engineering, 2018, 37(1): 156-167. (in Chinese with English abstract)

[32] RIAZ M T, RIAZ M T, REHMAN A, et al. An integrated approach of support vector machine (SVM) and weight of evidence (WOE) techniques to map groundwater potential and assess water quality[J]. Scientific Reports, 2024, 14: 26186.

[33] GAO X C, BAI W T, DANG Q L, et al. Learnable self-supervised support vector machine based individual selection strategy for multimodal multi-objective optimization[J]. Information Sciences, 2025, 690: 121553.

[34] 马志江, 陈汉林, 杨树锋. 基于支持向量机理论的滑坡灾害预测: 以浙江庆元地区为例 [J]. 浙江大学学报(理学版), 2003, 30(5): 592-596.
MA Z J, CHEN H L, YANG S F. Prediction of landslide hazard based on support vector machine theory: A case study of Qingyuan area in Zhejiang Province[J]. Journal of Zhejiang University (Sciences Edition), 2003, 30(5): 592-596. (in Chinese with English abstract)

[35] ZHU Y T, GU C S, DIACONEASA M A. A missing data processing method for dam deformation monitoring data using spatiotemporal clustering and support vector machine model[J]. Water Science and Engineering, 2024, 17(4): 417-424.

[36] 郭衍昊, 窦杰, 向子林, 等. 基于优化负样本采样策略的梯度提升决策树与随机森林的汶川同震滑坡易发性评价 [J]. 地质科技通报, 2024, 43(3): 251-265.
GUO Y H, DOU J, XIANG Z L, et al. Susceptibility evaluation of Wenchuan coseismic landslides by gradient boosting decision tree and random forest based on optimal negative sample sampling strategies[J]. Bulletin of Geological Science and Technology, 2024, 43(3): 251-265. (in Chinese with English abstract)

[37] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[38] 李苍柏, 肖克炎, 李楠, 等. 支持向量机、随机森林和人工神经网络机器学习算法在地球化学异常信息提取中的对比研究 [J]. 地球学报, 2020, 41(2): 309-319.
LI C B, XIAO K Y, LI N, et al. A comparative study of support vector machine, random forest and artificial neural network machine learning algorithms in geochemical anomaly information extraction[J]. Acta Geoscientica Sinica, 2020, 41(2): 309-319. (in Chinese with English abstract)

[39] 刘磊. 三峡水库万州区库岸滑坡灾害风险评价研究 [D]. 武汉: 中国地质大学(武汉), 2016.
LIU L. Assessment of landslide risk along Wanzhou area in Three Gorges reservoir[D]. Wuhan: China University of Geosciences (Wuhan), 2016. (in Chinese with English abstract)

[40] 严明, 谢婉丽, 何亚南, 等. 基于极值降雨假设法的城镇地质灾害风险性评价耦合模型研究: 以安康市岚皋县官元镇为例 [J]. 灾害学, 2023, 38(4): 219-227.
YAN M, XIE W L, HE Y N, et al. Research on the coupling model of urban geological hazard risk assessment based on the assumption of extreme rainfall: A case study of Guanyuan Town, Langao County, Ankang City[J]. Journal of Catastrophology, 2023, 38(4): 219-227. (in Chinese with English abstract)

[41] 汤国安. 地理信息系统教程 [M]. 2 版. 北京: 高等教育出版社, 2019.
TANG G A. Gis Tutorial [M]. Second Edition. Beijing: Higher Education Press, 2019. (in Chinese)

[42] 国家质量监督检验检疫总局, 中国国家标准化管理委员会. 中国地震动参数区划图: GB 18306-2015[S]. 北京: 中国标准出版社, 2016.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Seismic ground motion parameters zonation map of China: GB 18306-2015[S]. Beijing: Standards Press of China, 2016. (in Chinese)

Summary for “基于机器学习方法改进 IVM-RF 耦合模型的崩滑灾害危险性评价：以延安市志丹县为例”

Collapse and landslide risk assessment based on machine learning improved IVM-RF coupling method: A case study of Zhidan County,Yan'an City

QU Pengxin[a,b]，XIE Wanli[a,b*]，LIU Qiqi[a,b]，WANG Yuchen[a,b]

a. Department of Geology；b. State Key Laboratory of Continental Evolution and Early Life, Northwest University, Xi'an 710069, China

First author，E-mail：1279806472@qq.com

* Corresponding author， E-mail：xiewanli@nwu.edu.cn

Abstract：[**Objective**] To provide data support for disaster prevention, mitigation and risk management in Zhidan County, while providing reference for risk assessment in similar regions and addressing the research gap in collapse and landslide risk assessment that neglects cumulative rainfall impacts, this study conducted risk assessment under four different rainfall conditions: 10, 20, 50 and 100 years recurrence intervals. [**Methods**] Using Zhidan County as the study area with grid unit as evaluation element, we selected eight evaluation factors through Pearson correlation coefficient method: elevation, slope, aspect, curvature, rock/soil type, distance to river, distance to roads, and normalized vegetation index. The information model (IVM) was employed to evaluate the susceptibility and analyze the correlations between causative factors and disaster distribution. Through programming implementation, the analysis, transformation, management and drawing of the previous factor data are automatically processed, the IVM-RF (random forest) coupled model is improved, and the automatic cycle iteration comparison selection of the model is realized. The accuracy of the two susceptibility models is compared by ROC curve. Based on the evaluation results of the coupled model, the risk assessment was carried out. The Pearson type Ⅲ curve was used to estimate the rainfall in the study area under four different conditions of 10, 20, 50 and 100 years, and the risk zoning was carried out. [**Results**] For the results of susceptibility zoning, the *AUC* value of the evaluation results of the information-random forest coupling model is 0.87, which is better than the evaluation results of IVM. For the results of risk zoning, the area of high and extremely high-risk areas have gradually increased from 10 to 100 years. The improved coupled model evaluation method not only simplifies the operation but also improves the accuracy. [**Conclusion**] According to the survey results, the coupled model does have better evaluation accuracy and prediction ability.

Key words: collapse；landslide；information value method；random forest model；risk assessment；machine learning；Zhidan County, Yan'an City

Received：2024-10-08；Revised：2024-12-21；Accepted：2024-12-23

doi: 10.19509/j.cnki.dzkq.tb20240583