DOI:10. 19657/j. geoscience. 1000-8527. 202509240079

环境因子联接方式和评价模型对区域滑坡防控效能的优化研究：以陕西安康市汉滨区为例

王 滔[1]，王小浩[1,*]，王梦谣[1]，吝哲峰[1]，麻 晟[2]

（1. 陕西核工业工程勘察院有限公司，陕西 西安 710054；2. 中国地质大学（北京）工程技术学院，北京 100083）

摘要：滑坡易发性评价（LSA）是区域地质灾害风险防控的关键环节。为优化资源配置并提升预警效率，本研究探究了环境因子联接方式与评价模型对LSA的影响规律，以期为提高风险防控效能提供理论依据。以陕西省安康市汉滨区主城区为例，基于260处滑坡及10种环境因子数据，采用原始输入、滑坡密度、频率比（FR）、信息量（IV）、确定性系数（CF）5种联接方式，分别与随机森林（RF）、人工神经网络（ANN）、极端梯度提升（XGBoost）3种机器学习模型耦合，构建15种组合。同时将FR、IV、CF的预测结果作为对比基准，采用分区统计、ROC曲线对上述18种结果进行检验。结果表明：（1）环境因子联接方式对“基于树”的模型（RF、XGBoost）无明显影响；（2）“函数拟合型”模型（ANN）对联接方式较为敏感，其精度取决于联接方式的优劣；（3）机器学习模型整体优于统计学模型，模型整体性能表现为：RF > XGBoost > ANN > FR = IV > CF。因此，从防控实践出发，在难以确定最佳联接方式时优先采用RF模型，可保证评价精度的同时，有效提升防控资源的靶向性。

关键词：滑坡；易发性评价；联接方法；评价模型；风险防控

中图分类号：P642. 22；P694 **文献标志码：**A **文章编号：**1000-8527(2026)02-0504-18

Optimizing Landslide Prevention Efficacy Based on Environmental Factor Connection Method and Model Selection: A Case of Hanbin District in Ankang, Shaanxi

WANG Tao[1], WANG Xiaohao[1,*], WANG Mengyao[1], LIN Zhefeng[1], MA Sheng[2]

(1. *Shaanxi Nuclear Industry Engineering Survey Institute Co., Ltd, Xi'an, Shaanxi* 710054, *China*;
2. *School of Engineering and Technology, China University of Geosciences (Beijing), Beijing* 100083, *China*)

Abstract: Landslide Susceptibility Assessment (LSA) plays a crucial role in regional geological hazard risk prevention and control. To optimize resource allocation and enhance early warning efficiency, this study investigates the influence of environmental factor integration methods and evaluation models on LSA performance, with the aim of providing a theoretical basis for improving risk management effectiveness. Focusing on the main urban area of Hanbin District, Ankang City, Shaanxi Province, we utilized data from 260 historical landslides

收稿日期：2025-06-23；**修回日期：**2025-10-11；**编辑：**戚开静。

基金项目：国有企业－政府委托项目（2021-［DZ］-268）。

作者简介：王滔，男，1981年出生，高级工程师，岩土工程专业，主要从事地质灾害、矿山地质环境调查与治理等相关工作。Email：250684741@qq. com。

＊通信作者：王小浩，男，1992年出生，工程师，地质工程专业，主要从事地质灾害、矿山地质环境调查与治理等相关工作。Email：1543919431@qq. com。

引用格式：王滔，王小浩，王梦谣，等. 环境因子联接方式和评价模型对区域滑坡防控效能的优化研究：以陕西安康市汉滨区为例［J］. 现代地质，2026，40（2）：504-521. DOI：10. 19657/j. geoscience. 1000-8527. 202509240079.

WANG Tao, WANG Xiaohao, WANG Mengyao, et al. Optimizing Landslide Prevention Efficacy Based on Environmental Factor Connection Method and Model Selection: A Case of Hanbin District in Ankang, Shaanxi［J］. GEOSCIENCE, 2026, 40（2）：504-521. DOI：10. 19657/j. geoscience. 1000-8527. 202509240079.

and ten environmental factors. Five integration methods, namely Original Input, Landslide Density, Frequency Ratio (FR), Information Value (IV), and Certainty Factor (CF), were coupled with three machine learning models: Random Forest (RF), Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost), resulting in 15 combined approaches. The predictive results of the FR, IV, and CF models alone were also included as benchmark comparisons. All 18 susceptibility results were evaluated using zonal statistics and ROC curve analysis. The findings indicate that: (1) The choice of connection method has no significant impact on tree-based models (RF and XGBoost); (2) The function-fitting model (ANN) is sensitive to the connection method, and its accuracy depends on the appropriateness of the method selected; (3) Machine learning models generally outperform statistical models, with overall performance ranked as: RF > XGBoost > ANN > FR = IV > CF. Therefore, from a practical prevention and control perspective, adopting the RF model when the optimal integration method is uncertain can ensure assessment accuracy while effectively improving the targeting of prevention resources.

Key words: landslide; susceptibility assessment; connection method; evaluation model; risk prevention

0　引　言

滑坡是斜坡岩土体在重力及其他外界因素作用下发生变形破坏现象，作为全球发育最广泛的地质灾害之一，严重威胁人类生命财产安全[1, 2]。开展准确、有效的滑坡易发性评价（Landslide Susceptibility Assessment, LSA）并进行滑坡易发性制图（Landslide Susceptibility Mapping, LSM）已成为研究热点[3-4]。

LSA 的建模过程通常涵盖滑坡编录、环境因子选择、因子空间联接、评价模型确定及性能评估等步骤[5-6]。其中，环境因子空间联接方式与评价模型选择是影响LSA结果可靠性的关键环节。系统揭示这两个环节中的不确定性及其影响规律，对于提升区域滑坡风险防控的精准性具有重要意义。

依据建模原理，常见LSA模型可分为知识驱动、数据驱动和确定性方法三类[7]。数据驱动模型通过分析历史滑坡与环境因子之间的统计或非线性关系进行预测，因其客观性强且适用于区域尺度评价，已成为当前主流方法[8-9]。该类模型主要包括两大类：一是统计学模型（如频率比FR、信息量IV）[10-11]；二是机器学习模型（如随机森林RF、人工神经网络ANN、极端梯度提升XGBoost）[12-14]。然而，目前关于最优模型选择仍缺乏共识，即使模型的预测精度相近，但在易发性分区上也可能存在显著差异[15]。

环境因子联接通过量化各因子分级内的滑坡发育程度，建立其与滑坡的空间联系。研究表明，适宜的联接方式对提高LSA具有重要的意义[6]。例如，在江西丘陵山区以及贵州省思南县山区，将频率比作为随机森林（RF）、支持向量机（SVM）模型输入，其效果优于原始输入[16-17]；在清江水电开发河谷区与青海沙塘川流域，使用确定性系数、信息量等联接方式，也显著提升人工神经网络（ANN）和逻辑回归（LR）的预测性能[18-19]。这些结果表明，基于统计学的联接方式对于等ANN、LR模型具有重要价值。然而，现有研究多局限于单一联接方式与模型类型（如LR、SVM），缺乏对不同联接方法与“基于树”等模型的系统比较。

因此，为系统揭示联接方式与评价模型的交互影响机制，本文以陕西省安康市汉滨区主城区为研究案例，综合5种环境因子联接方式（原始输入、滑坡密度LD、频率比FR、信息量IV、确定性系数CF）与3种机器学习模型（RF、ANN、XGBoost），构建15种评价方案。本研究旨在阐明联接方式对不同原理模型的敏感性差异，并且识别适用于区域滑坡风险防控的最优建模策略，以期为精细化防灾减灾提供科学依据。

1　研究区与数据源

1.1　研究区概况

汉滨区主城区位于陕西省安康市中部，汉江、月河穿其而过，坐标范围108°50′17″—109°06′22″E，32°38′35″—32°51′19″N，总面积约328.39 km²（图1），其中中部的城区面积约62.16 km²。区域海拔介于134～1516 m，整体地势南北高、中部低。

区域出露震旦系、寒武系—奥陶系、志留系、新生界及新生界地层，岩性以千枚岩、板岩、片

岩和灰岩为主。构造上处于南秦岭印支褶皱带与北大巴山加里东褶皱带的交接部位，断裂发育，其中月河断裂为中元古代以来的长期活动断裂，走向北西，倾向北东，倾角 50 ~ 70° [20]。

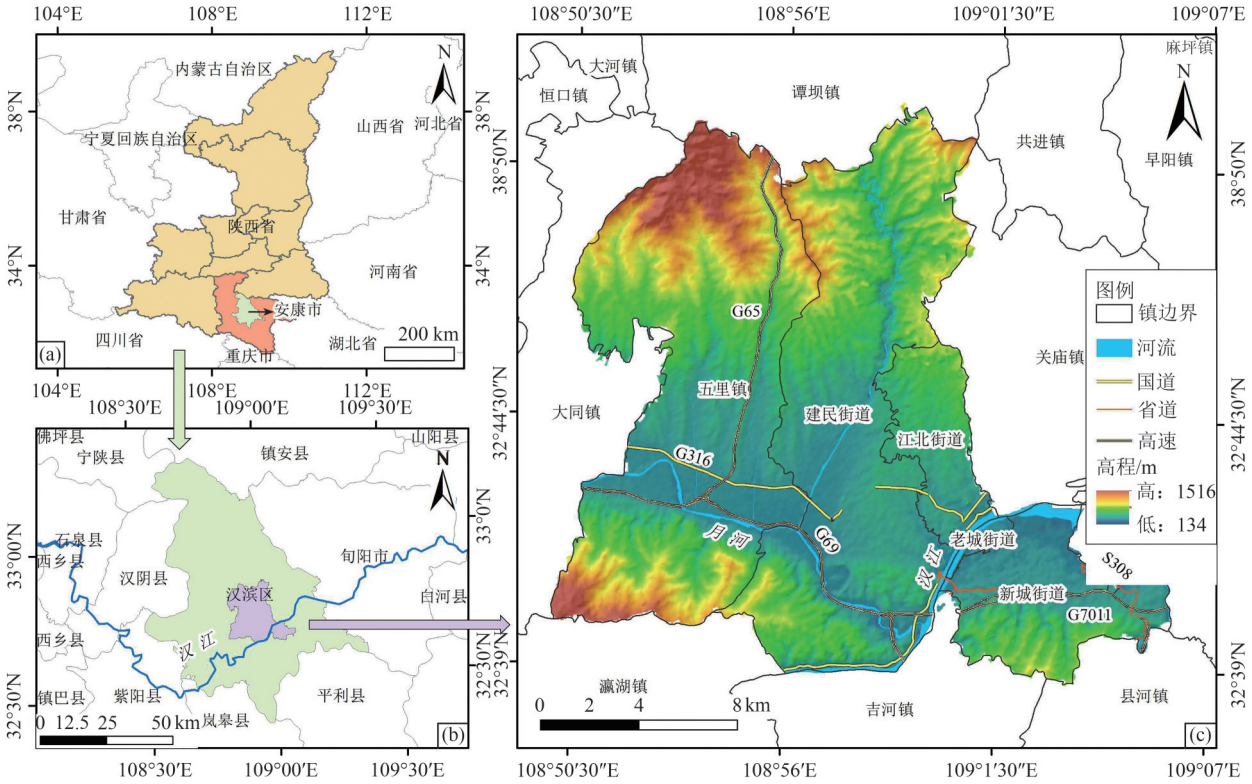


图 1 安康市（a）汉滨区（b）主城区（c）地理位置图

Fig. 1 Geographical location of the main urban area of Hanbin District in Ankang

1.2 数据源

本研究使用的数据主要包括环境因子与滑坡清单两类（表 1）。环境因子数据源自多源平台：30 m 分辨率 DEM 用于提取地形地貌因子；工程地质岩组与断裂信息从 1∶5 万地质图中提取；土地利用与归一化植被系数（NDVI）数据分别来自杨杰等 [21] 发布的中国 30 m 土地利用数据集与资源环境科学数据平台；基于 Open Street Map（2021 年）道路数据与第三次全国国土调查（2019 年）河流数据，通过欧氏距离算法生成距道路与河流距离因子。

表 1 安康市汉滨区主城区数据源及用途说明

Table 1 Data source and purpose description for the main urban area of Hanbin District in Ankang

数据类型	分辨率	来源	用途
DEM	30 m	地理空间数据云（https://www.gscloud.cn/sources/index?pid=302）	高程、坡度、坡向等提取
道路	—	开放街道地图（https://www.openstreetmap.org/）	距道路距离因子提取
河流	—	第三次全国国土调查（2019 年）	距河流距离因子提取
地质图	1∶5 万	全国地质资料馆（http://www.ngac.org.cn/）	工程地质岩组、断裂提取
土地利用	30 m	https://zenodo.org/records/8176941	土地利用因子
NDVI	30 m	资源环境科学数据平台（https://www.resdc.cn/DOI/DOI.aspx?DOIID=68）	NDVI 因子

滑坡数据主要来源于资料收集和遥感解译。考虑到中部城区地势平坦，高程变化微弱，灾害发育少，且人类工程活动强烈，微地貌复杂，30 m 的 DEM 难以有效反映人为微地貌特征，将其纳入易发性评价可能引入误差，故予以剔除。据 1∶10000 地质灾害风险调查与遥感解译，截至 2021 年共识别并野外验证滑坡 260 处（图 2），这些滑坡多发生于 6—9 月汛期，其主导诱发因素为强降雨，规模以小型为主，空间上沿河流附近区域呈线状分布。

此外，为独立验证模型泛化能力，本研究通过野外实地调查和资料收集，获取 2021—2023 年期间的 22 起滑坡灾情，该滑坡数据集与建模数据在时空上完全分离。时间上，这些新滑坡集中发生在 8—9 月；根据应急调查资料，其诱发因素同样为强降雨，这与建模数据的触发机制保持一致。图 3 为该 22 处滑坡点的空间分布（均采用后缘采

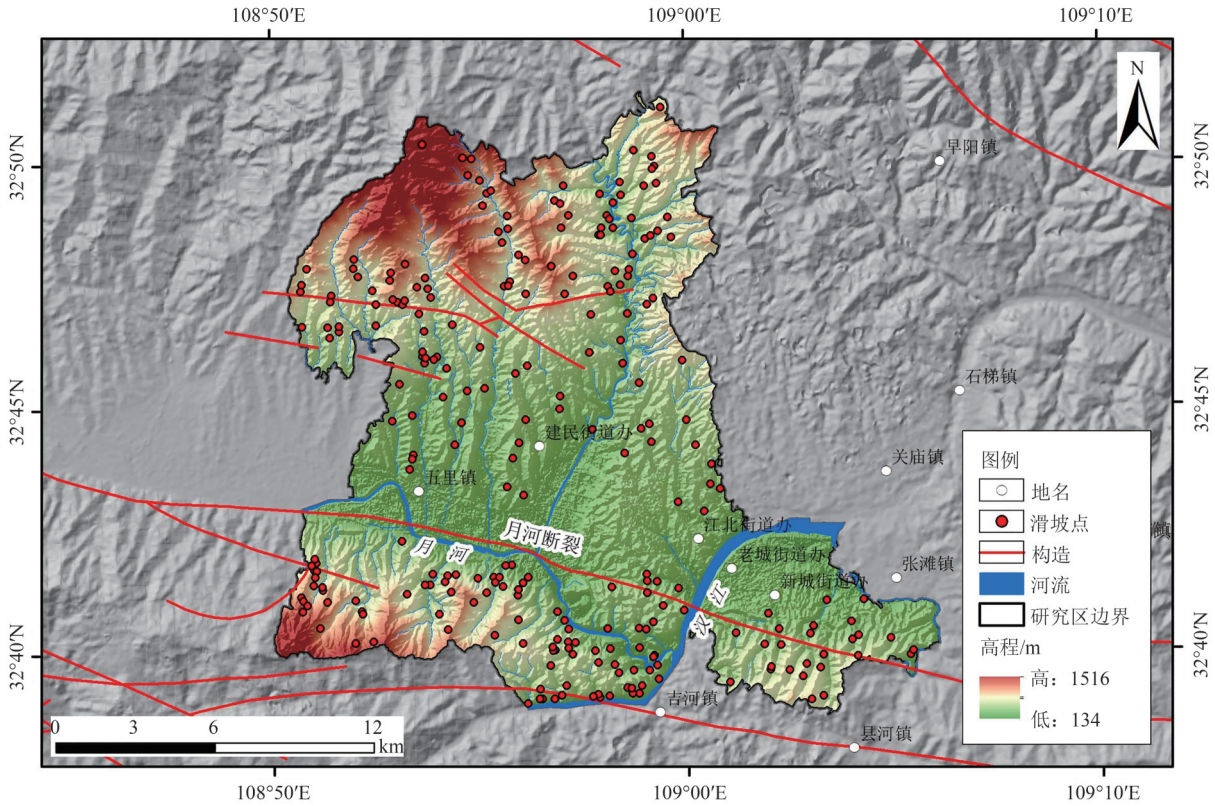


图 2　安康市汉滨区主城区滑坡空间分布图

Fig. 2　Spatial distribution of landslides in the main urban area of Hanbin District in Ankang

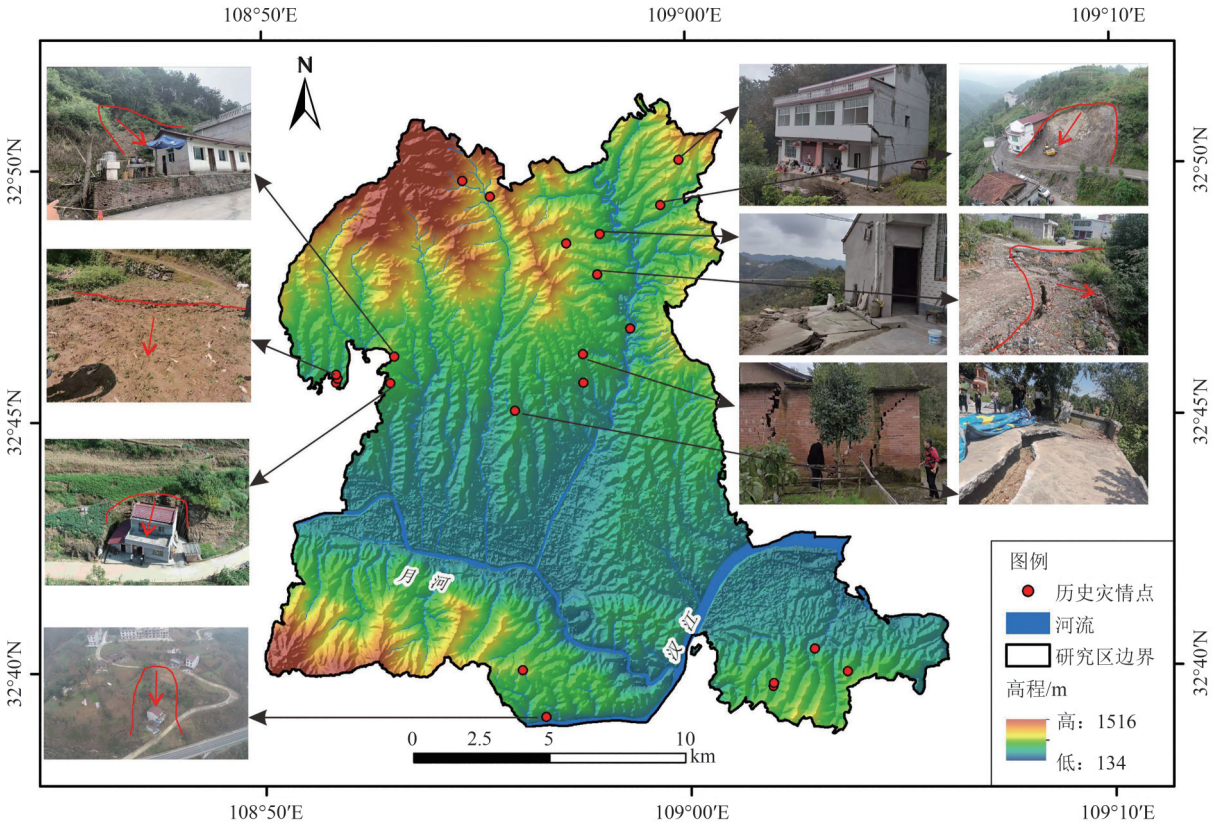


图 3　安康市汉滨区主城区历史灾情点空间分布与野外照片

Fig. 3　Spatial distribution of historical landslide incidents and field photographs from the main urban area of Hanbin District in Ankang

样）以及笔者实地考察的滑坡边界以及房屋、道路破坏情况。

2 研究方法

2.1 滑坡易发性建模流程

本研究的技术路线如图5所示，主要包括以下三个核心步骤：（1）数据准备与因子分析，基于滑坡编录与10项环境因子，构建评价数据库；（2）易发性建模，耦合5种环境因子联接方式与3种机器学习模型，生成15种易发性评价结果；（3）模型验证与对比，采用分区统计、ROC曲线及野外实例验证，系统评估不同建模策略的效能差异，以揭示联接方式与评价模型的交互影响规律，为滑坡风险防控提供理论支撑（图4）。

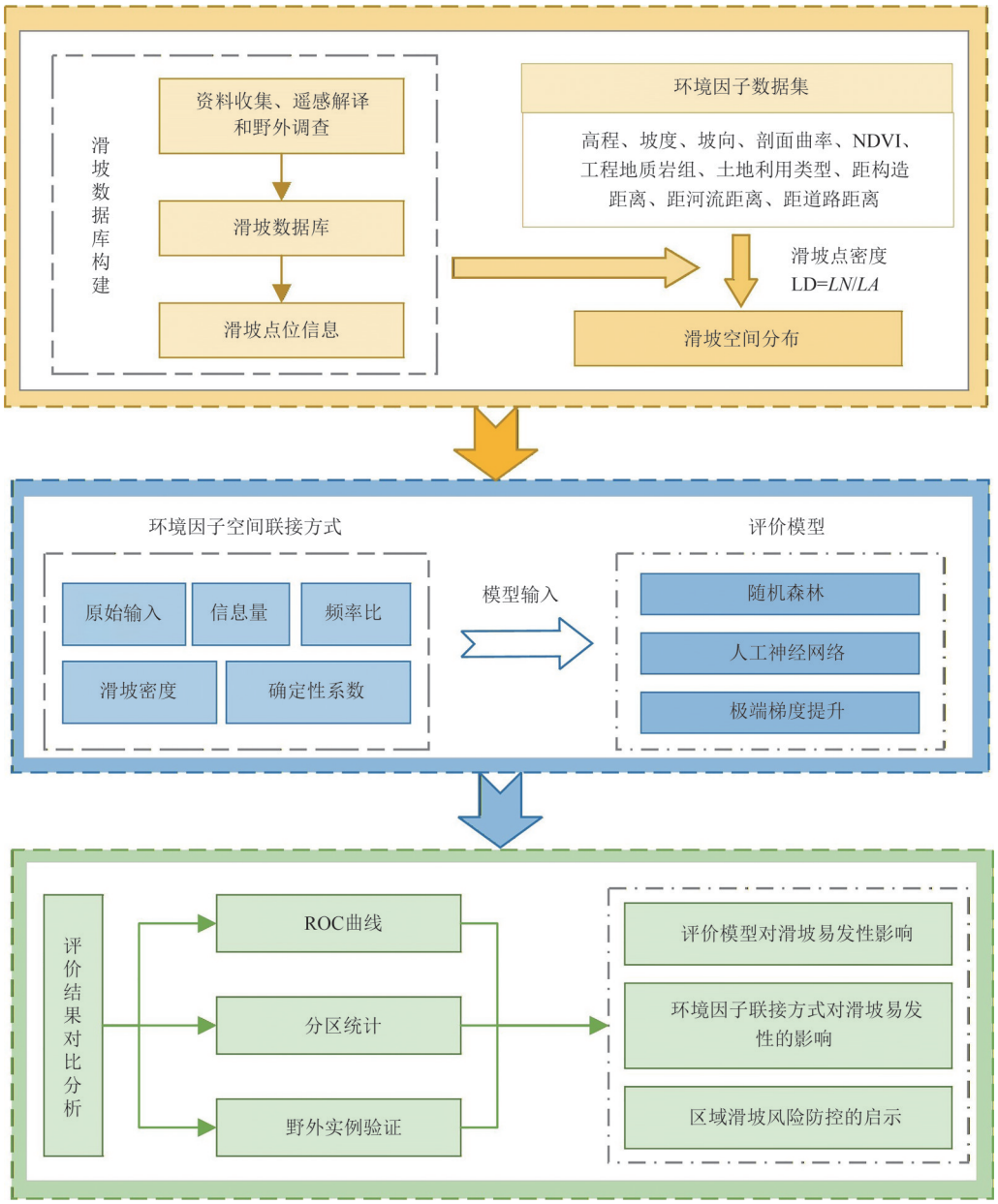


图4 滑坡易发性评价技术路线图

Fig. 4 Technical flowchart for landslide susceptibility assessment

2.2 环境因子联接方式

2.2.1 原始输入

原始输入将环境因子的原始数值作为模型输入。为满足模型要求，所有连续型因子均归一化至［0，1］区间（式（1）），离散型因子则进行哑变量处理[22]。

$$X_{norm}=\frac{X-X_{min}}{X_{max}-X_{min}} \tag{1}$$

式中：X代表原始数据，X_{norm}代表归一化结果，X_{min}代表最小值，X_{max}代表最大值。

2.2.2　滑坡密度（LD）

滑坡密度（Landslide Density，LD）通过计算各因子分区内滑坡数量与分区面积的比值，表征其与滑坡发生的相关性（式（2））[23-24]。

$$LD = LN/LA \tag{2}$$

式中：LN代表环境因子分级下滑坡发育数量，LA代表滑坡因子分级面积。

2.2.3　频率比（FR）

频率比（Frequency Ratio，FR）通过计算各因子分级内滑坡发生概率与区域平均概率的比值，量化其对滑坡发育的影响程度（式（3））[25]。FR > 1表示该分区有利于滑坡发生，同时FR亦可作为独立易发性评价模型使用。

$$FR_{ij} = \frac{N_{ij}/N}{M_{ij}/M} \tag{3}$$

式中：FR_{ij}代表第i个环境因子在j分区下的频率比值，N_{ij}为第i个环境因子j分区下滑坡发育数量，N为滑坡发育总数量；M_{ij}为第i个环境因子j分区的面积，M为研究区总面积。

2.2.4　信息量（IV）

信息量（Information Value，IV）用于评估各环境因子分区所提供的滑坡发生信息量（式（4））[26-27]，以反映环境因子对滑坡发生的非线性影响。IV > 0表示有利于滑坡发生。

$$IV_{ij} = \mathrm{In}\frac{N_{ij}/N}{M_{ij}/M} \tag{4}$$

式中各字母含义与2.2.3节相同。

2.2.5　确定性系数（CF）

确定性系数（Deterministic Coefficient，CF）基于概率统计函数，量化各因子对滑坡发生的确定性程度（式（5））[28]，$CF > 0$表示确定性增加。

$$CF = \begin{cases} \dfrac{PP_a - PP_s}{PP_a(1 - PP_s)} & PP_a \geqslant PP_s \\ \dfrac{PP_a - PP_s}{PP_s(1 - PP_a)} & PP_a < PP_s \end{cases} \tag{5}$$

式中：PP_a为某因子分级内滑坡发生的条件概率（滑坡数量与分级面积之比），PP_s为研究区内滑坡发生的先验概率（滑坡总数与研究区总面积之比）。

2.3　评价模型

2.3.1　随机森林（RF）

随机森林（Random Forest，RF）是一种集成机器学习模型，通过构建多棵决策树并采用投票机制输出最终结果[29-30]。该模型鲁棒性强、对异常值不敏感及抗过拟合等特点，在滑坡易发性预测中表现优异[31]。本研究基于python环境，实现对滑坡易发性的量化评估。

2.3.2　人工神经网络（ANN）

人工神经网络（Artificial Neural Network，ANN）是一种常用于分类与回归问题的非线性模型，由输入层、隐含层和输出层构成[32]。本研究采用反向传播神经网络，通过梯度下降算法最小化误差函数[33]。不断调整神经元之间的权重值和偏置，使输出逼近近似预期目标。

2.3.3　极端梯度提升（XGBoost）

极端梯度提升（eXtreme Gradient Boosting，XGBoost）是一种高效的梯度提升树算法，通过集成多个弱评估器构建强预测模型[14]。XGBoost支持并行计算，具有内存优化、处理稀疏数据能力强等优点，通常较传统线性模型具有更高精度。其模型形式为加法模型[34-35]。

2.4　精度检验方法

2.4.1　分区统计

滑坡易发性分区统计通过分析不同易发等级分区的面积以及滑坡发育情况，评估模型结果的合理性。合理的易发性分区应满足[36-37]：（1）极高易发区面积占比适中且需要包含多数滑坡；（2）极低和低易发区面积占比大而滑坡数量稀少；（3）滑坡发育密度随易发性等级降低而递减。

2.4.2　ROC曲线

ROC曲线是评价滑坡易发性模型精度的常用方法，能有效降低测试集选取带来的偏差，提升评估客观性[38-39]。ROC曲线纵轴为真阳性率，也称敏感性（*Sensitivity*）；横轴为假阳性率，也称“1—特异性”（1—Specificity），计算公式如下：

$$\mathrm{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{FR} + \mathrm{TN}} \tag{7}$$

式中：TP、FN分别为正确与错误识别的滑坡样本数；FP、TN分别为错误与正确识别的非滑坡样本数。通过计算ROC曲线下面积（Area Under Curve，AUC）评估模型性能，AUC值0.5～0.7表示预测能力较差，0.7～0.9表明较好，0.9～1.0为优秀[40]。

2.4.3 野外实例验证

为检验模型在实际风险防控中的适用性，本研究采用未参与建模的22处滑坡灾情点（2021—2023年）作为独立验证集。通过统计这些灾害点落入各易发等级的数量与密度，检验模型对未知滑坡的识别能力与实际应用价值。

3 滑坡易发性评价

3.1 环境因子分析与相关性检验

为提升评价的科学性，将滑坡发育稀少的中部城区（面积62.16 km^2）予以剔除，实际参与的评价面积为266.23 km^2。基于区域滑坡发育特征及相关研究[38, 41-42]，共选取10项环境因子（图5），涵盖地形地貌、工程地质、水文、地表覆盖与土地利用和人类工程活动等类别。所有因子统一在30 m栅格尺度下，采用自然间断点法（连续型因子）和实际类别（离散型因子）进行分级[38]。并计算了各因子分级内的滑坡密度、频率比、信息量与确定性系数等联接值（表2）。

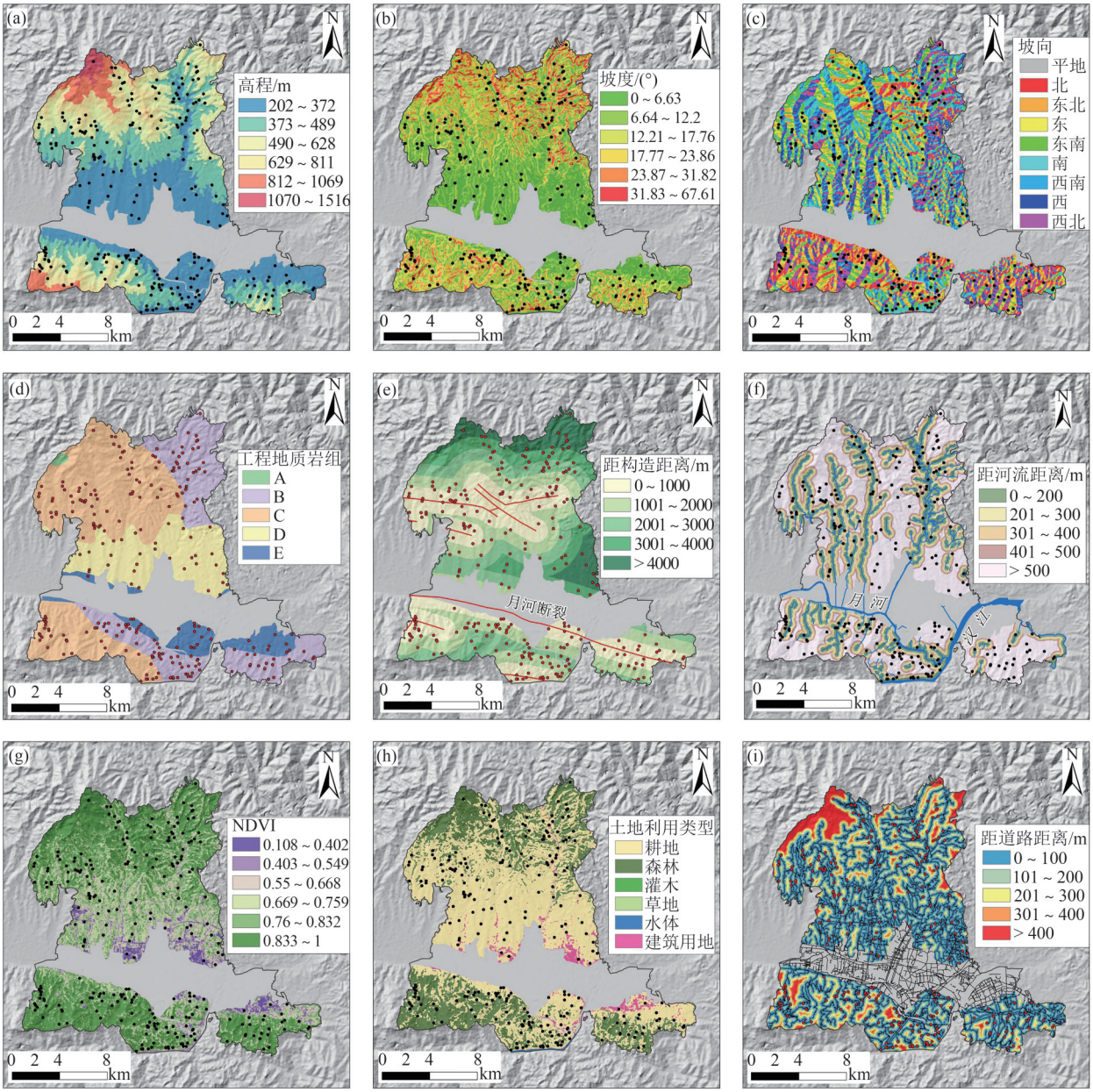


图5 安康市汉滨区主城区环境因子图（剖面曲率未列入）

Fig. 5 Environmental factors in the main urban area of Hanbin District in Ankang (excluding profile curvature)

A. 块状坚硬侵入岩；B. 较坚硬片状变质岩；C. 较软弱片状变质岩；D. 膨胀土；E. 一般黏性土

表 2　安康市汉滨区主城区环境因子分级统计与空间联接值

Table 2　Statistical classification and spatial connection of environmental factors for the main urban area of Hanbin District in Ankang

评价因子	分级	分级面积/km^2	灾害点数量/处	LD	FR	IV	CF
高程/m	202 ~ 372	91.58	100	1.092	1.118	0.112	0.122
	373 ~ 489	74.35	77	1.036	1.06	0.058	0.059
	490 ~ 628	54.38	63	1.159	1.186	0.171	0.155
	629 ~ 811	28.39	17	0.599	0.613	−0.489	−0.489
	812 ~ 1069	13.82	2	0.145	0.148	−1.911	−1.909
	1070 ~ 1516	3.69	1	0.271	0.277	−1.284	−1.282
坡度/(°)	0 ~ 6.63	60.88	25	0.411	0.42	−0.868	−1.195
	6.63 ~ 12.2	74.15	69	0.931	0.953	−0.048	0.008
	12.2 ~ 17.76	56.43	76	1.347	1.379	0.321	0.295
	17.76 ~ 23.86	42.57	49	1.151	1.179	0.165	0.185
	23.86 ~ 31.82	24.58	33	1.343	1.375	0.318	0.511
	31.82 ~ 67.62	7.60	8	1.052	1.077	0.074	−0.396
坡向	平地	0.17	0	0	0	—	0
	北	25.65	26	1.014	1.038	0.037	−0.085
	东北	30.00	37	1.233	1.263	0.233	0.260
	东	43.11	49	1.137	1.164	0.152	0.285
	东南	39.46	48	1.216	1.246	0.220	0.133
	南	31.08	31	0.997	1.021	0.021	0.084
	西南	33.08	23	0.695	0.712	−0.34	−0.479
	西	33.70	26	0.771	0.790	−0.236	−0.198
	西北	29.96	20	0.668	0.683	−0.381	−0.432
剖面曲率	−6.28 ~ −1.13	1.91	2	1.047	1.072	0.07	0
	−1.13 ~ −0.42	19.27	19	0.986	1.010	0.010	0.243
	−0.42 ~ −0.12	55.88	59	1.056	1.081	0.078	−0.068
	−0.12 ~ 0.19	126.51	110	0.869	0.89	−0.117	0.012
	0.19 ~ 0.69	55.22	59	1.068	1.094	0.090	0.001
	0.69 ~ 6.61	7.41	11	1.484	1.520	0.419	−0.188
工程地质岩组	块状坚硬侵入岩	0.98	0	0	0	—	0
	较坚硬片状变质岩	71.27	95	1.333	1.365	0.311	0.311
	较软弱片状变质岩	116.97	103	0.881	0.902	−0.103	−0.104
	膨胀土	53.20	34	0.639	0.654	−0.425	−0.424
	一般黏性土	23.79	28	1.177	1.205	0.186	0.186
距构造距离 (m)	0 ~ 1000	85.52	107	1.251	1.281	0.248	0.245
	1001 ~ 2000	72.34	76	1.051	1.076	0.073	0.035
	2001 ~ 3000	52.24	40	0.766	0.784	−0.243	−0.168
	3001 ~ 4000	28.41	21	0.739	0.757	−0.278	−0.279
	> 4000	27.68	16	0.578	0.592	−0.524	−0.528
距河流距离 (m)	0 ~ 200	81.34	73	0.897	0.919	−0.084	0.054
	201 ~ 300	35.47	42	1.184	1.212	0.192	0.233
	301 ~ 400	26.71	31	1.161	1.188	0.172	0.105
	401 ~ 500	24.00	27	1.125	1.152	0.141	0.008
	> 500	98.67	87	0.882	0.903	−0.102	−0.151
距道路距离 (m)	0 ~ 100	157.30	206	1.31	1.341	0.293	0.289
	101 ~ 200	56.92	38	0.668	0.684	−0.380	−0.369
	201 ~ 300	27.95	13	0.465	0.476	−0.742	−0.741
	301 ~ 400	10.60	2	0.189	0.193	−1.645	−1.644
	> 400	13.43	1	0.074	0.076	−2.577	−2.575
NDVI	0.108 ~ 0.402	7.58	2	0.264	0.270	−1.309	−0.615
	0.403 ~ 0.549	13.62	4	0.294	0.301	−1.201	−0.508
	0.550 ~ 0.668	23.84	27	1.133	1.16	0.148	0.318
	0.669 ~ 0.759	45.76	66	1.442	1.477	0.39	0.261

（续）表 2 安康市汉滨区主城区环境因子分级统计与空间联接值

（continued）Table 2 Statistical classification and spatial connection of environmental factors for the main urban area of Hanbin District in Ankang

评价因子	分级	分级面积/km²	灾害点数量/处	LD	FR	IV	CF
	0.76 ~ 0.832	67.08	73	1.088	1.114	0.108	0.212
	0.833 ~ 1	108.33	88	0.812	0.832	−0.184	−0.318
土地利用类型	耕地	177.29	215	1.213	1.242	0.217	0.235
	森林	80.15	44	0.549	0.562	−0.576	−0.671
	灌木	0.03	0	0	0	—	0
	草地	0.11	0	0	0	—	0
	水体	0.80	0	0	0	—	0
	建筑用地	7.83	1	0.128	0.131	−2.033	−2.034

（1）地形地貌因子。包括高程、坡度、坡向和剖面曲率，这里以高程为例进行分析。高程通过控制斜坡应力状态影响滑坡发育[43]。统计表明，滑坡数量随高程上升总体呈下降趋势，202 ~ 372 m 区间最为发育（100 处），1070 ~ 1516 m 区间最少（仅 1 处）。LD、FR、IV、CF 等联接值在 490 ~ 628 m 区间均达峰值，反映各联接方法在表征滑坡—高程非线性关系方面具有一致性。

（2）工程地质岩组。研究区较软弱片状变质岩区滑坡发育数量最多（103 处），而较坚硬片状变质岩区滑坡密度最高（1.33 处/km²）。

（3）距构造距离。构造对地质灾害发育控制作用明显，靠近构造带附近的岩体往往节理裂隙发育，较为破碎，斜坡整体稳定性降低[44-45]。表 2 结果显示，距构造 0 ~ 1000 m 范围内滑坡发育数量最多（107 处），LD 值亦最高（1.25 处/km²），并随距离增加逐渐降低。

（4）水文因子。距河流距离反映河流侵蚀对坡脚支撑的削弱作用[46]。统计表明，滑坡数量随河流距离增加总体减少，而滑坡密度在 201 ~ 300 m 处出现峰值（1.18 处/km²）。

（5）地表覆盖与土地利用。NDVI 作为地表植被覆盖的定量指标，其通过根系力学加固与调节地表入渗作用，影响斜坡稳定性[47]。从结果看 NDVI 值在 0.669 ~ 0.759 区间与滑坡具有较强的关系，LD 达 1.44 处/km²。土地利用类型综合反映了人类活动和自然背景下的交互作用[48]，其中，耕地区滑坡易发程度最高，LD 为 1.21 处/km²。

（6）人类工程活动。以距道路距离为表征，道路切坡改变原始地貌与应力状态，从而影响斜坡稳定[49]。结果显示，该因子对滑坡发育控制作用显著，距离越近，滑坡风险越高；0 ~ 100 m 范围内滑坡发育数量与密度均达最大值（206 处，1.31 处/km²）。

为排除因子间多重共线性对模型性能的影响，采用皮尔逊相关性系数（r）进行因子筛选（$|r|\leqslant 0.5$）[41]。因子相关性热力图（图 6）结果显示，10 个环境因子相关性较弱，无需对环境因子进行剔除。

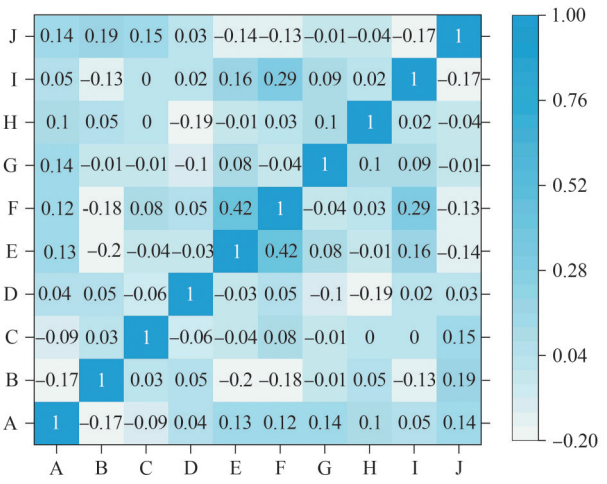


图 6 安康市汉滨区主城区环境因子皮尔逊相关性热力图

Fig. 6 Pearson correlation thermal map of environmental factors in the main urban area of Hanbin District in Ankang

A. 高程；B. 坡度；C. 坡向；D. 剖面曲率；E. NDVI；F. 土地利用类型；G. 距构造距离；H. 距河流距离；I. 距道路距离；J. 工程地质岩组

3.2 统计学模型评价结果

为系统比较不同评价模型对滑坡易发性结果的影响，本研究同时采用频率比（FR）、信息量（IV）与确定性系数（CF）三种统计学模型进行易发性分区。基于表 2 中各环境因子联接值，通过栅格叠加分别得到易发性指数，并采用自然间断点法[50-51]将结果划分为 5 个易发等级，分区结果见图 7（a1—a3）。

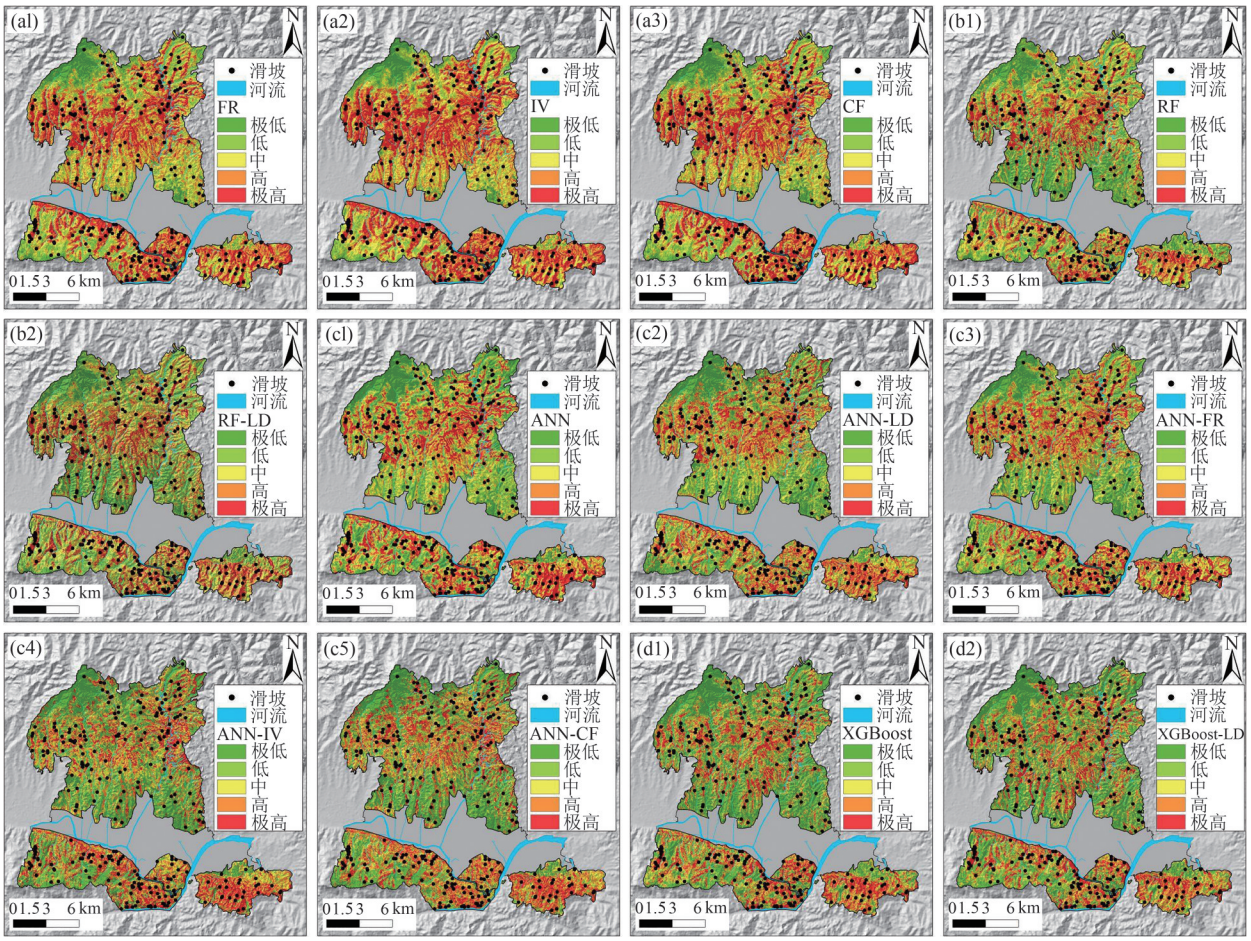


图 7　安康市汉滨区主城区滑坡易发性分区图

Fig. 7　Landslide susceptibility zoning map of the main urban area of Hanbin District in Ankang

a1. FR 模型；a2. IV 模型；a3. CF 模型；b1—b2. RF-based 模型；c1—c5. ANN-based 模型；d1—d2. XGBoost-based 模型

从空间分布来看，三种模型所划定的极高与高易发区分布较为一致，主要沿月河等河流两岸呈线状展布。在面积分配上，IV 模型所识别的极高与高易发区范围最大，FR 模型相对较小，CF 模型介于二者之间，反映了不同统计方法在风险边界识别上的差异。

3.3　机器学习模型评价结果

为保障模型的可比性，RF、ANN、XGBoost 模型均采用统一的缓冲控制采样法（Buffer controlled sampling，BCS）构建滑坡数据集（正负样本 1∶1，训练测试比例 7∶3）；正样本为 260 个滑坡点，负样本则依据空间约束条件，在历史滑坡 200 m 缓冲区外的区域随机选取[52-54]。为评估负样本随机抽样对模型性能的影响并验证采样方法的稳健性，本文独立重复了 10 次负样本采样与模型构架流程。结果如图 8（a）所示，三种机器学习模型的 AUC 值在重复试验中的波动范围均小于 0.02。这表明在既定的空间约束下，负样本的随机性并未对模型性能产生实质性的影响。基于此，本研究随机选择了其中一次负样本抽样结果，以构建后续的模型。该次抽样的负样本和正样本分布见图 8（b），可见负样本有效避开了滑坡影响区。

基于 python 的 sklearn 和 xgboost 库构建机器学习模型，依据数据特点，确定了各模型合适的参数：（1）RF，树数量 200；（2）ANN，隐藏层为（14，9），迭代次数 1000；（3）XGBoost，学习率 0.1，最大深度 4，树数量 200。

（1）RF 模型结果。

基于五种联接方式输入的 RF 易发性分区结果见图 7（b1—b2），由于除原始输入外，其余输入方式的易发性分区结果一致，故这里仅展示原始输入和 LD 联接的结果。与统计学模型相比，RF 所划定的极高与高易发区面积显著缩小，但空间分布仍集中于沟谷两侧，一致性较好。其分区结果在空间上更为精细，呈现破碎化分布。

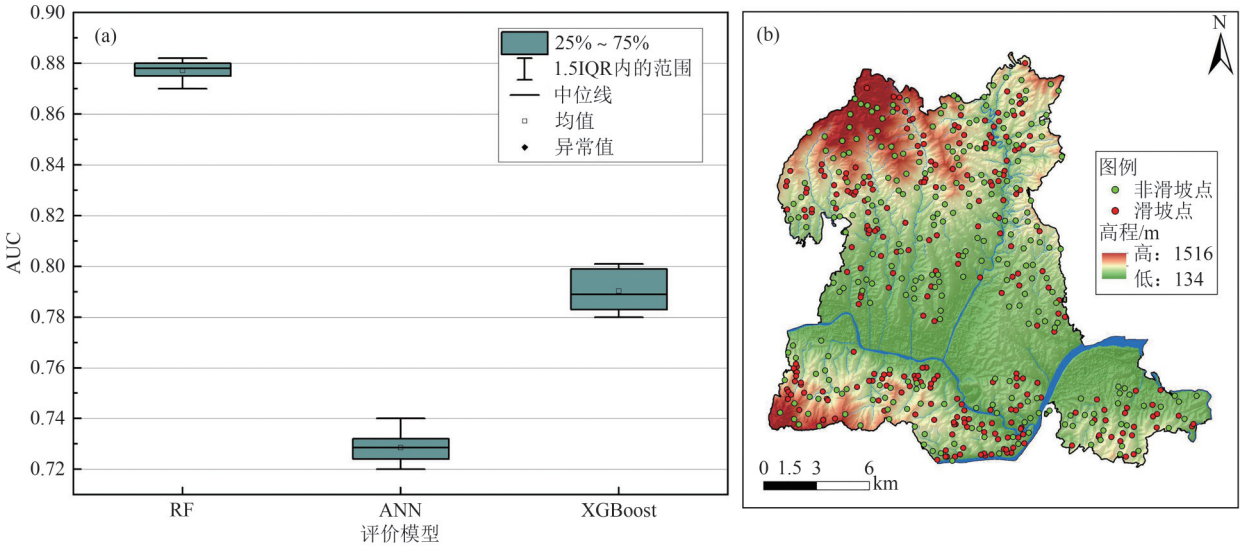


图8 负样本采样稳健性检验

Fig. 8 Robustness test of non-lanslide sampling

(a) 模型AUC值分布箱型图;(b) 正负样本空间分布图

(2) ANN模型结果。

ANN模型的易发性分区结果见图7 (c1—c5)。其极高与高易发区空间分布与统计学模型及RF基本一致，但空间异质性更强，分区更为合理。不同于RF，ANN的分区结果受联接方式影响较为明显。

(3) XGBoost模型结果。

XGBoost模型的分区结果见图7 (d1—d2)。由于除原始输入外，其余输入方式的易发性分区结果一致，故这里仅展示原始输入和LD联接的结果。XGBoost模型的空间格局与RF、ANN及统计学模型相似，但相较于统计学模型更为合理。

4 滑坡易发性精度检验

4.1 分区统计

对18种模型的易发性分区进行面积与滑坡发育统计，并计算滑坡密度以评估分区合理性（表3）。为直观展示分区特征，选取FR、IV、CF、RF-FR、ANN-IV与XGBoost等6种代表性模型绘制了各模型滑坡密度堆叠图（图9）、面积与滑坡占比直方图（图10）。

表3 安康市汉滨区主城区不同评价模型易发性分区结果统计

Table 3 Statistics of susceptibility zoning results from different evaluation models for the main urban area of Hanbin District in Ankang

模型	分区	分区面积/km²	滑坡数量	滑坡密度/(处/km²)	模型	分区	分区面积 /km²	滑坡数量	滑坡密度/(处/km²)
	极高	60.91	145	2.38		极高	40.18	96	2.39
	高	72.14	66	0.91		高	59.86	88	1.47
FR	中	74.33	38	0.51	ANN-LD	中	60.46	52	0.86
	低	42.47	11	0.26		低	58.18	21	0.36
	极低	16.37	0	0.00		极低	47.52	3	0.06
	极高	70.73	156	2.21		极高	42.50	102	2.40
	高	84.36	70	0.83		高	56.08	85	1.52
IV	中	70.00	31	0.44	ANN-FR	中	63.43	45	0.71
	低	28.72	3	0.10		低	61.07	26	0.43
	极低	12.39	0	0.00		极低	43.12	2	0.05
	极高	59.58	143	2.40		极高	49.73	109	2.19
	高	78.26	73	0.93		高	51.14	85	1.66
CF	中	70.87	34	0.48	ANN-IV	中	45.80	43	0.94
	低	41.24	10	0.24		低	40.86	18	0.44
	极低	16.25	0	0.00		极低	78.68	5	0.06

（续）表 3　安康市汉滨区主城区不同评价模型易发性分区结果统计

（continued）Table 3　Statistics of susceptibility zoning results from different evaluation models for the main urban area of Hanbin District in Ankang

模型	分区	分区面积/km²	滑坡数量	滑坡密度/(处/km²)	模型	分区	分区面积 /km²	滑坡数量	滑坡密度/(处/km²)
	极高	36.22	204	5.63		极高	46.66	93	1.99
	高	55.26	27	0.49		高	55.17	84	1.52
RF	中	63.85	19	0.30	ANN-CF	中	49.51	57	1.15
	低	60.58	10	0.17		低	39.96	20	0.50
	极低	50.29	0	0.00		极低	74.91	6	0.08
	极高	38.13	203	5.32		极高	44.18	191	4.32
	高	59.42	30	0.50		高	43.38	37	0.85
RF-LD	中	59.65	15	0.25	XGBoost	中	43.68	15	0.34
	低	61.67	10	0.16		低	49.58	8	0.16
	极低	47.34	2	0.04		极低	85.39	9	0.11
	极高	38.13	203	5.32		极高	43.89	140	3.19
	高	59.42	30	0.50		高	45.39	72	1.59
RF-FR	中	59.65	15	0.25	XGBoost-LD	中	43.68	20	0.46
	低	61.67	10	0.16		低	47.60	14	0.29
	极低	47.34	2	0.04		极低	85.65	14	0.16
	极高	39.34	202	5.14		极高	43.89	140	3.19
	高	55.92	32	0.57		高	45.39	72	1.59
RF-IV	中	66.68	14	0.21	XGBoost-FR	中	43.68	20	0.46
	低	61.23	11	0.18		低	47.60	14	0.29
	极低	43.03	1	0.02		极低	85.65	14	0.16
	极高	41.17	203	4.93		极高	42.89	133	3.10
	高	57.05	31	0.54		高	43.59	77	1.77
RF-CF	中	60.04	14	0.23	XGBoost-IV	中	44.01	22	0.50
	低	61.39	10	0.16		低	48.68	13	0.27
	极低	46.56	2	0.04		极低	87.04	15	0.17
	极高	46.12	117	2.54		极高	43.89	140	3.19
	高	56.27	73	1.30		高	45.39	72	1.59
ANN	中	61.79	49	0.79	XGBoost-CF	中	43.68	20	0.46
	低	57.71	16	0.28		低	47.60	14	0.29
	极低	44.31	5	0.11		极低	85.65	14	0.16

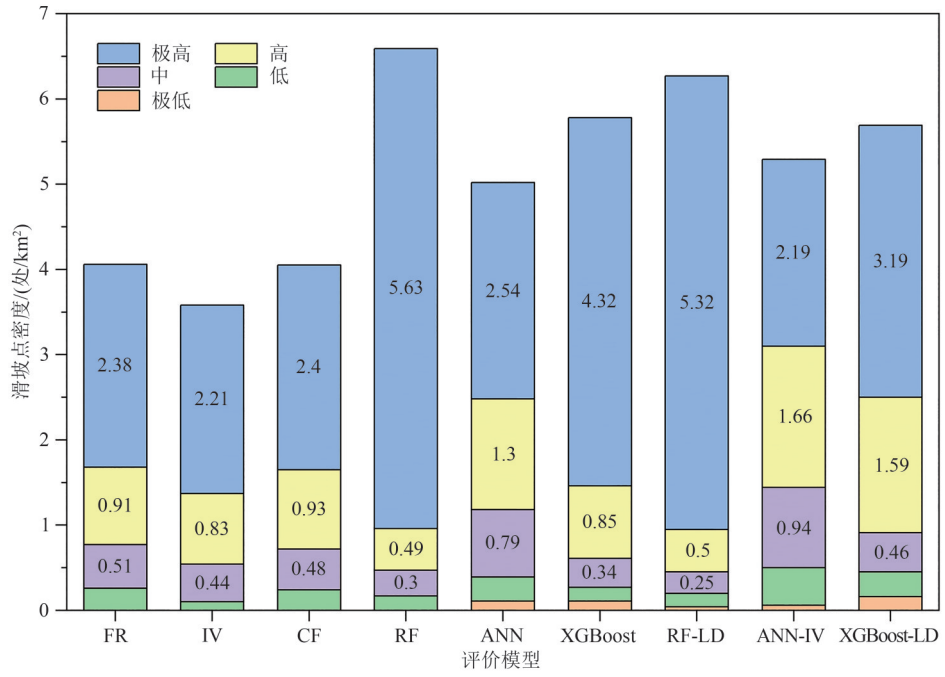


图 9　不同模型易发性分区滑坡点密度计算结果

Fig. 9　Calculated density of landslide points within susceptibility zones derived from different models

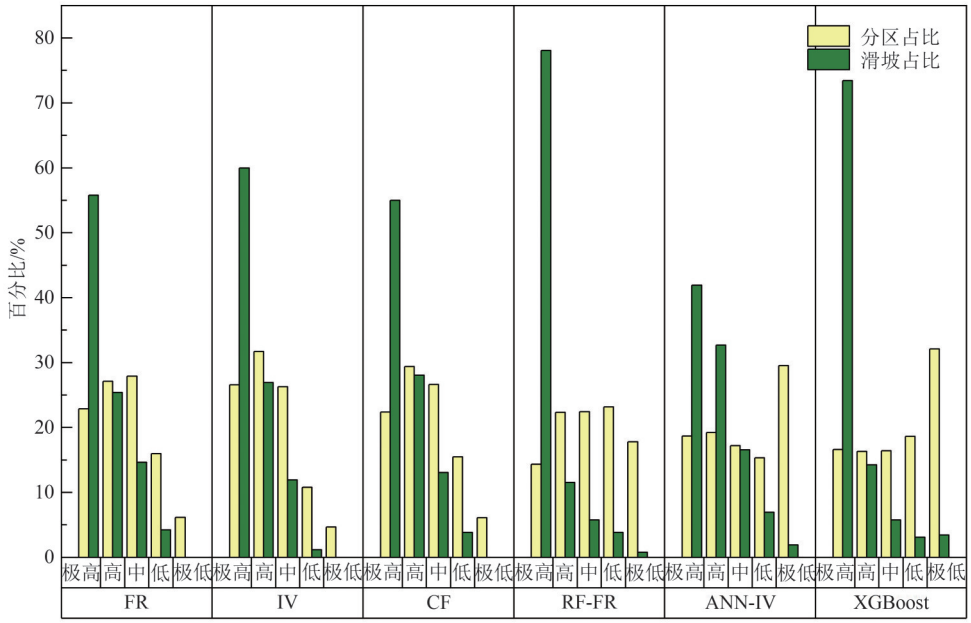


图 10 不同评价模型易发性分区结果

Fig. 10 Susceptibility zoning results from different evaluation models

统计表明，所有模型均表现出随易发等级降低，滑坡数量与密度逐步递减的合理趋势。对比发现，RF 模型在极高易发区的滑坡密度稳定在约 5 处/km^2，显著高于其他模型，体现出对高风险区域更优的识别能力。相比之下，统计学模型（FR，IV，CF）普遍呈现极高与高易发区面积偏大、低易发区偏小的特点，存在高估高风险区域的范围倾向。同时，在不同联接方式下，RF 与 XGBoost 的分区结果保持稳定，而 ANN 则表现出对输入方式的敏感性，尽管其总体分区趋势一致。

4.2 ROC 曲线

上述 18 种模型的 ROC 曲线与 AUC 值见图 11，其中机器学习模型的结果汇总于表 4。结果表明，ANN 模型的 AUC 值与统计学模型（FR、IV、CF）相近，甚至略低，但其易发性分区结果在空间合理性上优于统计学模型（图 10）。

综合比较，模型整体性能排序为：RF（AUC 最高 0.896）> XGBoost（最高 0.853）> ANN ≈ 统计学模型。此结果与分区统计结论相互印证，确认了 RF 模型的优越性。

在联接方式影响方面，除原始输入外，其余联接方式对 RF 与 XGBoost 的预测性能影响较小；而 ANN 的表现则对联接方式较为敏感，其 AUC 值随输入方式不同而呈现明显变化。

4.3 野外实例验证

基于 22 处未参与建模的滑坡灾情数据对 6 个代表性模型进行验证（表 5）。结果显示，就滑坡捕获数量而言，统计学模型（FR、IV、CF）表现优异，极高和高易发区涵盖 20 处以上的滑坡；机器学习模型中：RF（19 处）> XGBoost（16 处）> ANN-IV（11 处）。

从滑坡密度计算结果看，RF 在极高易发区的滑坡密度最大（0.33 处/km^2），优于 FR、CF（0.25 处/km^2）。此外，机器学习模型在高易发区内的平均滑坡密度为 0.14 处/km^2，高于统计学模型结果（0.06 处/km^2），表明其具有更高的滑坡空间预测效果。

以上结果说明，机器学习模型（尤其是 RF）并非简单地追求更高的滑坡点捕获数量，而是致力于以较小的面积准确预测更多的滑坡。该特征使得机器学习模型在滑坡风险防控中具有重要的优势，为区域地质灾害防控提供更为准确、科学的参考依据。

5 讨 论

5.1 环境因子联接方式对滑坡易发性的影响

本研究通过系统耦合 5 种联接方式与 3 种机器学习模型，得到 15 个工况组合结果。旨在揭示因子联接方式对滑坡易发性的影响机制。结果表明，联接方式的作用效果与评价模型本身具有密切的相关性。

对于 RF 与 XGBoost 这类“基于树”的模型，

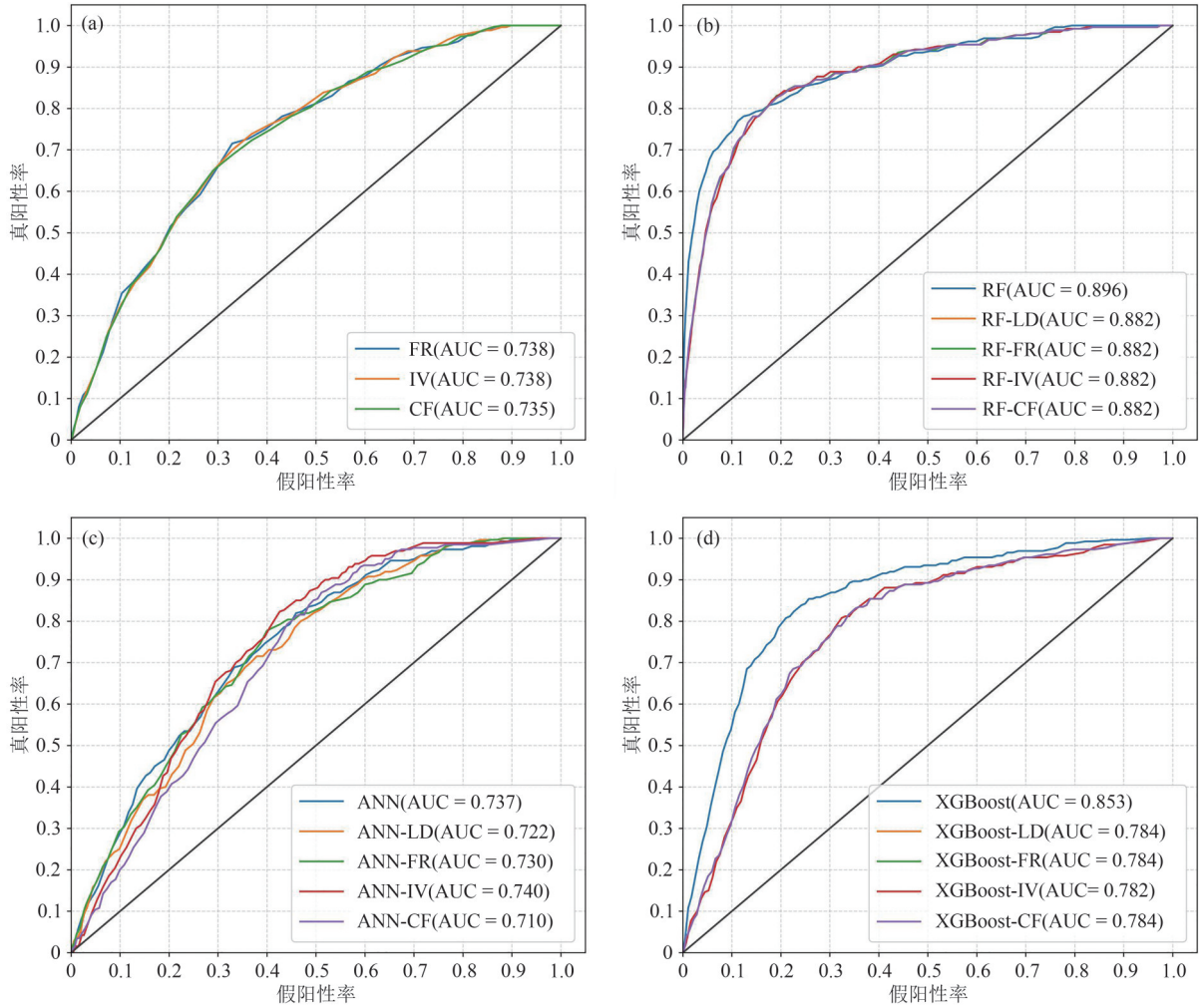


图 11　不同评价模型 ROC 曲线

Fig. 11　ROC curves of different evaluation models

表 4　不同环境因子联接方式下机器学习模型的 AUC 值

Table 4　AUC values of machine learning models under different environmental factor connection modes

模型	环境因子联接方式					平均值
	原始输入	LD	FR	IV	CF	
RF	0.896	0.882	0.882	0.882	0.882	0.885
ANN	0.737	0.722	0.730	0.740	0.710	0.728
XGBoost	0.853	0.784	0.784	0.782	0.784	0.797

除原始输入外，其余四种联接方式（LD，FR，IV，CF）对模型性能影响甚微。具体表现为：采用不同统计学联接方式时，RF 模型的 AUC 值稳定在 0.882，XGBoost 模型稳定在 0.784，且其易发性分区结果在空间格局和面积占比上高度一致。这表明，复杂的因子预联接处理对于提升此类本身具备强特征学习能力模型的精度贡献有限。

相比之下，ANN 这类“函数拟合型”模型对联接方式则表现出显著的敏感性。其性能优劣依赖于输入特征的质量，表现排序为：ANN-IV（AUC=0.740）> ANN - 原始输入（0.737）> ANN-FR（0.730）> ANN-LD（0.722）> ANN-CF（0.710）。这一规律与分区统计中 ANN-CF 模型在中易发区仍出现高滑坡密度（> 1 处/km^2）的不合理现象相互印证。该发现与大量将 FR、IV 用于 LR、SVM 等参数模型并取得性能提升的前人研究结论一致[17-19, 23]。同时，李文斌等的研究也表明 RF 模型对联接方式不敏感[6]，这与本研究的发现是吻合的。

笔者认为，导致上述差异的原因在于模型本身。“函数拟合型”模型属于参数模型，其性能受到输入数据分布和特征的影响，因此通过 FR、IV 等预处理方法可以强化有利因子区间的信号，从而提高模型性能。而“基于树”的模型属于非参数模型，处理数据时不需要显式建模特征之间的

表5 基于野外验证的各模型易发性分区滑坡发育统计

Table 5 Statistics of landslide development in susceptibility zones based on field validation for each model

模型	分区	分区面积/km²	滑坡数量	滑坡密度/(处/km²)	模型	分区	分区面积/km²	滑坡数量	滑坡密度/(处/km²)
FR	极高	60.91	15	0.25	RF	极高	36.22	12	**0.33**
	高	72.14	5	0.07		高	55.26	7	0.13
	中	74.33	2	0.03		中	63.85	2	0.03
	低	42.47	0	0		低	60.58	0	0
	极低	16.37	0	0		极低	50.29	1	0.02
IV	极高	70.73	**16**	0.23	ANN-IV	极高	49.73	6	0.12
	高	84.36	5	0.06		高	51.14	7	0.14
	中	70.00	1	0.01		中	45.8	4	0.09
	低	28.72	0	0		低	40.86	4	0.10
	极低	12.39	0	0		极低	78.68	1	0.01
CF	极高	59.58	15	0.25	XGBoost	极高	44.18	10	0.23
	高	78.26	5	0.06		高	43.38	6	0.14
	中	70.87	2	0.03		中	43.68	4	0.09
	低	41.24	0	0		低	49.58	1	0.02
	极低	16.25	0	0		极低	85.39	1	0.01

关系，故而环境因子联接方式对其精度影响较小。

一个值得深入探讨的发现是，原始输入在RF和XGBoost中反而取得了最高精度（AUC值分别为0.896和0.853）。这很可能是因为原始输入避免了因子分级造成的信息损失或引入的偏差，保留了环境因子最原始、最连续的变化信息。这与Huang等[54]提出的“增加因子分级数量可在一定程度上提升模型性能”的结论在逻辑上相通，因为原始输入本质上可视为不断增加分级数的情况，信息保留最为完整。

综上所述，本研究揭示了环境因子联接方式对于滑坡易发性评价的影响规律：

（1）环境因子联接方式对“基于树”的模型（RF，XGBoost）无显著影响，不会明显提高或降低模型预测精度。

（2）“函数拟合型”模型（LR，ANN，SVM）对环境因子联接方式较为敏感，模型预测结果精度取决于联接方式的优劣。

（3）原始输入对模型的影响主要在于增加了因子分级数量，一定程度上提高模型的精度。

5.2 评价模型对滑坡易发性的影响

评价模型的选择是影响滑坡易发性结果的重要因素，本研究系统对比三个机器学习模型外（RF，ANN，XGBoost）与三种统计学模型（FR，IV，CF），并通过分区统计、ROC曲线和野外实例验证对其性能进行了综合评估。结果表明，不同模型的性能存在显著差异。

在预测精度和分区合理性方面，机器学习模型整体上优于传统的统计学模型。其中，RF模型表现最为突出，其AUC值最高（原始输入达0.896），并且在分区统计中展现出极高的空间靶向性：其划定的极高易发区面积占比最小，却能预测最大数量的滑坡（204处），滑坡密度稳定在5处/km²。XGBoost模型性能次之，其AUC值（原始输入为0.853）和极高易发区内识别的滑坡数量（均值148处）均优于ANN模型。ANN模型的AUC值（0.710～0.740）虽然与统计学模型相近，甚至部分情况下略低，但其易发性分区结果在空间上更具异质性，分区更为合理，避免了统计学模型普遍存在的“片状”高估现象。

从实际防灾减灾的效能来看，机器学习模型，特别是RF模型，展现出更高的实用价值。基于22处未参与建模的滑坡灾情点验证发现，RF模型在极高易发区的滑坡密度达到0.33处/km²，相较于统计学模型（FR、CF为0.25处/km²）提升了32%（表5）。这意味着在相同的单位预警面积内，RF模型能够识别更多的隐患点，从而极大地优化了防灾资源的配置效率，为滑坡风险防控提供了更为可靠的依据。

综上所述，本研究通过三种独立的检验方法，得出模型性能的综合排序为：RF > XGBoost > ANN > FR = IV > CF。

值得注意的是，统计学模型与机器学习模型在结果上产生显著差异的根源在于其建模原理的

不同。统计学模型（如FR、IV、CF）仅基于滑坡点本身与环境因子的关系进行计算，并未引入非滑坡样本进行学习，因此在预测滑坡时，模型通常会对其进行“模糊处理”，将更多区域划定为高易发区[55-56]。

5.3　对区域滑坡风险防控的启示

基于上述研究结果，本文从实际应用角度出发，为区域滑坡风险防控提出以下启示。

（1）统计学模型适用于大范围风险初步筛查。在市级/区县级等大尺度（如1∶5万）地质灾害风险识别与规划中，信息量法等统计学模型操作简单、结果稳定，能有效圈定“重点防控区”，为早期防灾部署提供高效的底图支撑。

（2）机器学习模型支撑精细化防控。在集镇等更小尺度（如1∶1万）的精细化风险防控中，推荐采用随机森林（RF）等机器学习模型。其能以更小的预警面积实现更高的隐患识别效率，为“点状”精准投入提供科学依据，显著优化防灾资源配置，避免因预警范围过大带来的管控成本浪费。

（3）推荐RF模型作为优先选择的稳健方案。在实际工作中，若难以确定最佳建模策略，本研究结果表明随机森林（RF）模型是一项稳健且高效的优先选择。其在不同环境因子联接方式下均表现稳定、性能优异，能有效降低因方法选择不当带来的决策风险，尤其适用于模型技术积累初期的地区。

6　结　论

本研究系统评价了5种环境因子联接方式（原始输入、滑坡密度LD、频率比FR、信息量IV、确定性系数CF）与3种机器学习模型（随机森林RF、人工神经网络ANN、极端梯度提升XGBoost）在陕西省汉滨区滑坡易发性评价中的表现，主要结论如下。

（1）揭示了研究区滑坡发育的关键环境特征组合：高程490～628 m、坡度23.87～31.82°、东北与东南坡向、NDVI范围0.669～0.759、距河流<300 m、距构造<1000 m、较坚硬片状变质岩出露区、距道路<100 m，以及土地利用类型为耕地的区域。

（2）环境因子联接方式对“基于树”的模型（RF、XGBoost）影响不显著，不会明显改变其预测精度；而“函数拟合型”模型（ANN）对联接方式较为敏感，其性能优劣取决于所选联接方式的适用性。

（3）机器学习模型整体优于统计学模型，其中RF表现最优（AUC最高达0.896），模型性能排序为：RF > XGBoost > ANN > FR = IV > CF。

综上，建议将RF模型作为汉滨区及类似地质条件区域滑坡风险防控的首选评价工具，该模型能够精准识别滑坡隐患范围，为防控资源的优化配置提供有效支撑。

参考文献：

[1] 许强. 对滑坡监测预警相关问题的认识与思考[J]. 工程地质学报，2020，28(2)：360－374.

[2] 严天笑，张建通，朱月琴，等. 增量学习在滑坡易发性评价中的应用——以甘肃省天水市为例[J]. 地质通报，2024，43(4)：630－640.

[3] 黄发明，杨阳，蒋水华，等. 滑坡易发性关键问题综述及其基于半监督非对称理论的解决方案[J/OL]. 岩石力学与工程学报，1－30[2025-10-27]. https://link.cnki.net/urlid/42.1397.O3.20250917.1732.002.

[4] 吴益平，张秋霞，唐辉明，等. 基于有效降雨强度的滑坡灾害危险性预警[J]. 地球科学(中国地质大学学报)，2014，39(7)：889－895.

[5] REICHENBACH P，ROSSI M，MALAMUD B D，et al. A review of statistically-based landslide susceptibility models[J]. Earth-science Reviews，2018，180：60－91.

[6] 李文彬，范宣梅，黄发明，等. 不同环境因子联接和预测模型的滑坡易发性建模不确定性[J]. 地球科学，2021，46(10)：3777－3795.

[7] 屈飞行，张菁，汤明高. 基于GIS的区域地质灾害危险性评价流程与关键问题探讨[J]. 西华师范大学学报(自然科学版)，2018，39(4)：420－428.

[8] HUANG F M，YIN K L，HUANG J S，et al. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine [J]. Engineering Geology，2017，223：11－22.

[9] 刘帅，王涛，曹佳文，等. 基于优化随机森林模型的降雨群发滑坡易发性评价——以西秦岭极端降雨事件为例[J]. 地质通报，2024，43(6)：958－970.

[10] 陈航，刘惠军，王韬，等. 基于频率比-深度神经网络耦合模型的滑坡易发性评价——以盐源县为例[J]. 水文地质工程地质，2024，51(5)：161－171.

[11] 穆成林，周海鹏，阳博，等. 基于信息量-地理逻辑回归耦合模型的滑坡灾害易发性评价——以四川省德昌县为例[J]. 地理科学，2025，45(7)：1575－1589.

[12] LI M X，WANG H Y，CHEN J L，et al. Assessing landslide susceptibility based on the random forest model and multi-source heterogeneous data [J]. Ecological Indicators，2024，

158: 111600.
[13] 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.
[14] 赵晓东, 徐振涛,刘福,等. 基于极端梯度提升算法的滑坡易发性评价模型[J]. 科学技术与工程, 2022, 22(23): 10347 - 10354.
[15] 黄发明, 胡松雁, 闫学涯, 等. 基于机器学习的滑坡易发性预测建模及其主控因子识别[J]. 地质科技通报, 2022, 41(2): 79 - 90.
[16] 盛明强, 刘梓轩, 张晓晴, 等. 基于频率比联接法和支持向量机的滑坡易发性预测[J]. 科学技术与工程, 2021, 21(25): 10620 - 10628. .
[17] 刘山东, 李军, 江兴元, 等. 基于数学统计与机器学习模型耦合的滑坡易发性评价方法优化[J]. 科学技术与工程, 2025, 25(5): 1827 - 1839.
[18] 李正, 冷亮, 孙永鑫, 等. 基于信息量-机器学习耦合模型的水电梯级开发流域滑坡易发性评价[J]. 测绘通报, 2024(S1): 237 - 241.
[19] 栗泽桐, 王涛, 周杨, 等. 基于信息量、逻辑回归及其耦合模型的滑坡易发性评估研究:以青海沙塘川流域为例[J]. 现代地质, 2019, 33(1): 235 - 245.
[20] 赵振山. 汉阴县地质灾害危险性评价及群测群防体系建议[D]. 西安:长安大学, 2013.
[21] YANG J, HUANG X. 30 m annual land cover and its dynamics in China from 1990 to 2019[J]. Earth System Science Data Discussions, 2021, 2021: 1 - 29.
[22] 吴赛儿. 强震区泥石流易发性研究[D]. 北京:中国地质大学(北京), 2018.
[23] 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.
[24] XU C, XU X W, YAO X, et al. Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis[J]. Landslides, 2014, 11: 441 - 461.
[25] 肖婷, 刘庆丽, 邓敏, 等. 三峡库区万州区滑坡易发性演化规律[J]. 地球科学, 2025, 50(4): 1625 - 1637.
[26] 王小浩 ,唐亚明, 冯卫, 等. 基于加权信息量法的成家庄地质灾害敏感性评价研究[J]. 科技创新与应用, 2020, (5): 63 - 66 + 68.
[27] 伍敏婷, 孙岳, 周子腾. 基于层次分析 - 信息量法的滑坡易发性评价——以江西省浮梁县为例[J]. 东华理工大学学报(自然科学版), 2023, 46(2): 157 - 166.
[28] 陈芯宇, 师芸, 温永啸, 等. 基于确定性系数与支持向量机的滑坡易发性评价[J]. 科学技术与工程, 2023, 23(2): 518 - 527.
[29] BAGGING B L. predictors [J]. Machine Learning, 1996, 24(2): 123 - 140.
[30] CUTLER A, CUTLER D R, STEVENS J R. Random forests [J]. Ensemble machine learning: Methods and applications, 2012: 157 - 175.
[31] 李贞贵. 随机森林改进的若干研究[D]. 厦门:厦门大学, 2013.
[32] SAHA S, MAJUMDAR P, BERA B. Deep learning and benchmark machine learning based landslide susceptibility investigation, Garhwal Himalaya (India) [J]. Quaternary Science Advances, 2023, 10: 100075.
[33] LIU Y R, MENG Z Q, ZHU L, et al. Optimizing the sample selection of machine learning models for landslide susceptibility prediction using information value models in the Dabie mountain area of Anhui, China[J]. Sustainability, 2023, 15(3): 1971.
[34] 郭齐韵, 吴华, 李佳潼, 等. 基于机器学习与样本优化的滑坡易发性评价:以西藏自治区昌都市为例[J]. 地质科学, 2025, 60(1): 308 - 322.
[35] 林报嘉, 刘晓东, 杨川, 等. XGBoost机器学习模型与GIS技术结合的公路崩塌灾害易发性研究[J]. 公路, 2020, 65(7): 20 - 26.
[36] 许冲. 实际滑坡是滑坡危险性评价结果合理性检验最好的标准[J]. 工程地质学报, 2013, 21(6): 908 - 911.
[37] 张晓敏. 基于GIS的陕西省滑坡灾害危险性评价及分区研究[D]. 西安:长安大学, 2019.
[38] 麻晟, 陈剑, 李强, 等. 负样本选择和评价模型对滑坡易发性预测不确定性的影响[J]. 现代地质, 2025, 39(5): 1373 - 1391.
[39] CANTARINO I, CARRION M A, GOERLICH F, et al. A ROC analysis-based classification method for landslide susceptibility maps[J]. Landslides, 2019, 16: 265 - 282.
[40] 安雪莲, 密长林, 孙德亮, 等. 基于不同评价单元的三峡库区滑坡易发性对比——以重庆市云阳县为例[J]. 吉林大学学报(地球科学版), 2024, 54(5): 1629 - 1644.
[41] 麻晟, 陈剑, 吴赛儿. 陕西省安康市瀛湖镇滑坡灾害分布特征与易发性分析[J]. 现代地质, 2024, 38(2): 437 - 450.
[42] 刘增源. 安康市汉滨区滑坡敏感性评价研究[D]. 西安:西安科技大学, 2021.
[43] AYALEW L, YAMAGISHI H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan [J]. Geomorphology, 2005, 65(1/2): 15 - 31.
[44] 兰中孝, 张才兵, 杨雲兵, 等. 贵州省三都县滑坡特征及分布规律研究[J]. 地质与勘探, 2024, 60(4): 830 - 837.
[45] 殷翔, 齐普荣, 王滔. 陕西省石泉县地质灾害发育特征与形成条件分析[J]. 地质灾害与环境保护, 2014, 25(2): 16 - 18.
[46] 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.
[47] 朱路路. 基于多种负样本取样方法的滑坡易发性评价[D]. 淮南:安徽理工大学, 2024.
[48] 吴俭俭, 武震, 陈安安. 小尺度下基于斜坡和栅格单元的滑坡易发性评价对比研究[J]. 自然灾害学报, 2025, 34(5): 105 - 118.
[49] 张杰, 康晓波, 鲁慧, 等. 哀牢山中段地质灾害空间分布规律研究——以元江县为例[J]. 工程地质学报, 2018, 26(3): 720 - 731.
[50] KE C Y, HE S, QIN Y G. Comparison of natural breaks method

and frequency ratio dividing attribute intervals for landslide susceptibility mapping[J]. Bulletin of Engineering geology and the Environment, 2023, 82(10): 384.

[51] 王小浩, 王梦谣, 王滔, 等. 基于加权综合指数法集镇地质灾害易发性评价研究——以袁家庄街道为例[J]. 地下水, 2025, 47(5): 137 - 140.

[52] ABDELKADER M M, DAOUD A M A, ABDEL-GAWAD A G, et al. Improving Landslide Susceptibility Mapping: The Role of Non-Landslide Sample Selection in Machine Learning Models[J]. Remote Sensing Applications: Society and Environment, 2025: 101759.

[53] LIU Q, TANG A P, HUANG D L. Exploring the uncertainty of landslide susceptibility assessment caused by the number of non-landslides[J]. Catena, 2023, 227: 107109.

[54] HUANG F M, YE Z, JIANG S H, et al. Uncertainty study of landslide susceptibility prediction considering the different attribute interval numbers of environmental factors and different data-based models[J]. Catena, 2021, 202: 105250.

[55] MICHELETTI N, FORESTI L, ROBERT S, et al. Machine learning feature selection methods for landslide susceptibility mapping [J]. Mathematical geosciences, 2014, 46: 33 - 57.

[56] XIAO T, YIN K L, YAO T L, et al. Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, Three Gorges Reservoir, China[J]. Acta Geochimica, 2019, 38: 654 - 669.