



引文格式: 付泉, 党光普, 李致博, 等. 基于分形维数耦合支持向量机和熵权模型的滑坡易发性研究[J]. 西北地质, 2024, 57(6): 255–267. DOI: 10.12401/j.nwg.2023196

Citation: FU Quan, DANG Guangpu, LI Zhibo, et al. Study of Landslide Susceptibility Mapping Based on Fractal Dimension Integrating Support Vector Machine with Index of Entropy Model[J]. Northwestern Geology, 2024, 57(6): 255–267. DOI: 10.12401/j.nwg.2023196

基于分形维数耦合支持向量机和熵权模型的 滑坡易发性研究

付泉¹, 党光普¹, 李致博^{1,*}, 田润青¹, 石琳¹, 赵鑫¹, 王昆², 石磊², 吕娜娜³

(1. 陕西地建土地勘测规划设计院有限责任公司, 陕西 西安 710075; 2. 陕西地建土地工程技术研究院有限责任公司, 陕西 西安 710075; 3. 陕西水务发展供水集团有限公司, 陕西 西安 710018)

摘要: 陕西省宝鸡市北部黄土高原滑坡灾害频发, 严重威胁当地人民的经济发展和生产生活。本研究基于分形维数, 分别利用熵权模型(IEO)、支持向量机模型(SVM)和两种混合模型即 F-IEO 和 F-SVM 对滑坡可能发生的范围进行定量预测。首先, 利用 179 个滑坡样本制作滑坡编录图, 将 70%(125 个)的滑坡样本用于训练, 其余 30%(54 个)用于测试。随后, 提取 12 种滑坡影响因子, 分别计算每个因子的信息增益率和分形维数, 并使用训练数据建立 4 种滑坡易发性分区模型。最后, 利用受试者工作特征曲线(ROC)和统计学指标包括阳性预测率(PPR)、阴性预测率(NPR)和准确率(ACC)测试模型的性能, 并比较模型的泛化性。结果表明, F-SVM 模型在训练和测试数据集上分别得到最高的 PPR、NPR、ACC 和 AUC 值, 其次是 F-IEO 模型。最终, F-SVM 模型在所有模型中表现最优, 因此, 基于分形维数构建的混合模型比原始模型更具优势, 可为当地滑坡防治决策提供参考。

关键词: GIS; 滑坡易发性研究; 混合模型; 分形维数

中图分类号: P694; P280

文献标志码: A

文章编号: 1009-6248(2024)06-0255-13

Study of Landslide Susceptibility Mapping Based on Fractal Dimension Integrating Support Vector Machine with Index of Entropy Model

FU Quan¹, DANG Guangpu¹, LI Zhibo^{1,*}, TIAN Runqing¹, SHI Lin¹, ZHAO Xin¹, WANG Kun², SHI Lei², LV Nana³

(1. Shaanxi Geological Construction Land Survey, Planning and Design Institute Co., Ltd, Xi'an 710075, Shaanxi, China; 2. Institute of Land Engineering and Technology, Shaanxi Provincial Land Engineering Construction Group Co., Ltd., Xi'an 710075, Shaanxi, China; 3. Shaanxi Water Development and Supply Group Co., Ltd., Xi'an 710018, Shaanxi, China)

Abstract: Landslides occur frequently on the Loess Plateau in the north of Baoji City, Shaanxi Province, which seriously threaten the economic development, production and life of the local people. Based on fractal dimension,

收稿日期: 2023-03-22; 修回日期: 2023-09-24; 责任编辑: 贾晓丹

基金项目: 陕西省重点研发计划项目(2024SF-YBXM-565), 陕西地建土地勘测规划设计院 2024 年度内部科研项目(KCNY2024-2, KCNY2024-4), 陕西省土地工程建设集团内部科研项目(DJNY-ZD-2023-1, DJNY-YB-2023-18, DJNY2024-18)和陕西地建—西安交大土地工程与人居环境技术创新中心开放基金(2024WHZ0240)联合资助。

作者简介: 付泉(1992-), 男, 硕士, 工程师, 主要从事地质灾害与土地工程研究。E-mail: 517043740@qq.com。

* 通讯作者: 李致博(1987-), 男, 硕士, 高级工程师, 主要从事地理信息与数据挖掘方面的研究。E-mail: 267001531@qq.com。

sion, entropy weight model (IOE), support vector machine model (SVM) and two hybrid models, namely F-IOE and F-SVM, are used to quantitatively predict the possible occurrence range of landslide. First of all, 179 landslide samples were used to make landslide cataloguing maps, 70% (125) of the landslide samples were used for training, and the remaining 30% (54) were used for testing. Then, 12 kinds of landslide influence factors are extracted, information gain rate and fractal dimension of each factor are calculated respectively, and four landslide vulnerability zoning models are established using training data. Finally, the performance of the model was tested using the receiver operating characteristic curve (ROC) and statistical indicators including positive predictive rate (PPR), negative predictive rate (NPR) and accuracy rate (ACC), and the generalization of the model was compared. The results show that F-SVM model has the highest PPR, NPR, ACC and AUC values in training and test data sets respectively, followed by F-IOE model. Finally, F-SVM model is the best among all models. Therefore, the hybrid model based on fractal dimension has more advantages than the original model, which can provide reference for local landslide control decisions.

Keywords: GIS; research on landslide susceptibility; mixed model; fractal dimension

滑坡是一种破坏性事件, 往往对人类生命、财产和生存环境构成严重威胁, 严重制约着人类的可持续发展。据统计, 发展中国家因滑坡造成的死亡率高达 95%, 造成的经济损失占国民生产总值的 1%~2% (Youssef et al., 2022)。作为发展中国家, 中国山区占国土面积的 70%, 地质环境条件复杂、构造活动强烈、工程建设规模大, 滑坡多发且分布较广 (刘光旭等, 2014)。

中国西北部的黄土覆盖区生态极其脆弱, 土壤结构松散, 降水集中, 水土流失严重, 已成为全国最容易发生滑坡的地区 (孙萍萍等, 2022)。此外, 滑坡对当地社会经济的发展 and 人类安全构成极大威胁 (韩玲等, 2019)。因此, 滑坡预测对该区域滑坡的防治具有重要意义。然而, 传统的野外调查难以满足滑坡预测的要求。随着地理信息技术 (GIS) 的发展, 滑坡易发性分区的研究结合了许多统计学方法, 即统计指数 (Abedini et al., 2018)、逻辑回归 (Thai et al., 2018)、确定性因素 (Li et al., 2017)、层次分析法 (张向营等, 2018)、频率比 (Abedini et al., 2018)、熵权法 (Jaafari et al., 2014) 和证据权法。虽然统计学方法和 GIS 技术可以很好地结合起来, 更适合大面积研究, 但滑坡易发性分区是一种典型的复杂非线性问题, 所以统计学方法得到的结果往往精度较低。因此, 越来越多的机器学习方法被引入, 如人工神经网络 (谢振华等, 2017)、核逻辑回归 (Chen et al., 2020)、随机森林 (Behnia et al., 2018)、支持向量机 (武雪玲等, 2016; 韩玲等, 2019)、朴素贝叶斯 (Sameen et al., 2020) 和决策树 (Kim et al., 2018)。机器学习模型的性能通常取决于训练数据的质量和数量, 并且对建模过程中的参数十分敏感。随机划分的数据、不同的研究规模和不

同的数据分辨率都会对训练数据产生影响 (Bazi et al., 2021)。此外, 对于哪种方法最适合研究区开展滑坡易发性分区, 目前还没有明确的结论。

近年来, 混合模型的出现为滑坡易发性分区研究提供了更多的选择, 也可探索出更加科学合理的结果。这些模型包括频率密度耦合核函数逻辑回归模型 (FR-KLR) (张庭瑜等, 2020)、确定性系数耦合径向基神经网络模型 (CF-RBFNN) (Li et al., 2017) 等。此外, 也有学者利用分形理论对滑坡进行研究, 但目前研究的主要方向是用分形维数来描述滑坡的空间分布, 几乎没有对滑坡易发性分区进行过研究。

由于评估方法会对滑坡易发性分区结果产生很大影响, 探索和研究分区模型是一项重要的工作, 并且将分形维数与 IOE 和 SVM 融合的研究鲜有报道。因此, 本研究基于分形维数和两种原始模型 (IOE、SVM), 分别构建两种混合模型 (F-IOE、F-SVM) 对宝鸡市北部典型黄土覆盖区进行滑坡易发性分区研究。研究结果可为研究区以及类似区域的灾害防治提供理论依据, 并且相关图件的发布也可当地的防灾减灾工作提供决策支持。

1 研究区域

研究区位于陕西省宝鸡市北部典型黄土覆盖区, 面积约 2 451 km², 地理坐标介于 E 107°00'01"~107°45'08" 和 N 34°29'22"~34°50'38"。由南向北横跨黄土台地和黄土高原两大地貌类型。黄土台地海拔 750~950 m, 相对高差小于 200 m, 黄土高原地区海

拔 900~1 700 m, 相对高差 100~300 m。研究区域内发育一条水系, 即千河, 水流方向为 WN—ES, 最终流入渭河。

研究区气候类型属暖温带大陆性气候, 年平均气温和月平均气温分别为 12 ℃ 和 25 ℃, 年平均降水量为 653 mm, 降雨期主要集中在 6~9 月, 近十年平均月最大降雨量为 103 mm。由于黄土遇水易塌陷, 故该区域容易发生地质灾害。研究区横跨黄土高原、鄂尔多斯盆地和秦岭造山带, 由南向北跨越 3 个构造单元, 区域内发育 7 条断裂带, 其中 F_1 — F_5 断裂走向为 NW—SE, F_6 — F_7 断裂走向为 W—E。根据《陕西省区域地质志》来看, 研究区未发生破坏性地震(陕西省地质矿产局, 1989)。

2 数据应用

2.1 滑坡编录图

包含滑坡数量、位置和规模信息的滑坡编录图是滑坡易发性分区研究的基础。根据野外调查、前期地质资料、1980~2017 年的历史滑坡数据以及航空遥感数据, 共提取地质灾害 179 处, 其中堆积层滑坡 166 处、基岩滑坡 13 处。滑坡周界面积的最大、最小和平均尺寸分别为 $2.2 \times 10^4 \text{ m}^2$ 、 $1.1 \times 10^3 \text{ m}^2$ 、 $8.3 \times 10^3 \text{ m}^2$ 。由于近 90% 的滑坡周界面积不超过 10 000 m^2 , 因此采用质心法将多边形转化为点, 然后将这些点映射到研究区的滑坡编录图中(图 1)。

2.2 训练和测试数据集的制备

为了建立滑坡易发性分区模型并获得滑坡易发性分区图(LSM), 训练和测试数据集的制备是重要且必要的步骤。由于建立滑坡易发性模型必须同时存

在正样本和负样本, 本研究将 179 个滑坡样本全部作为正样本。此外, 根据地质资料和野外观测数据分析, 研究区滑坡一般不会发生在坡度小于 5° 的区域(张茂省等, 2019), 因此将这些区域定义为非滑坡区(Pandey et al., 2020)。在非滑坡区随机生成相同数量的非滑坡点作为负样本, 然后将 179 个滑坡点以 70% 与 30% 的比例随机分为两组, 同时将 179 个非滑坡点按照相同的比例分为两组。最后, 将 70%(250 个) 的样本用来构建模型, 30%(108 个) 的样本用于模型测试。

由于单一的模型会因数据的随机性而出现拟合的现象(Mahdadi et al., 2018), 因此为了减少随机选择对滑坡易发性分区模型性能的影响, 将所有滑坡点以 7:3 的比例随机划分 30 次, 并对每组训练数据集进行 10 折交叉验证。保留最稳定的训练数据集检测多重共线性问题、确定滑坡影响因子和建立滑坡易发性分区模型。

2.3 滑坡影响因子分析

滑坡的发生受区域地质环境、水文、气象和构造运动等一系列条件因素的影响, 影响因子的选择和分类将直接影响滑坡易发性分区的准确性(Hong et al., 2017)。本研究在实地调查和以往研究的基础上, 共考虑了 12 个内部和外部的影响因子, 包括坡度、坡向、高程、降雨量、平面曲率、距河流的距离、土地利用类型、距道路的距离、距断层的距离、归一化植被指数(NDVI)、地层岩性和剖面曲率。

滑坡与坡度的关系密切, 坡度直接影响坡内应力分布及有效面。研究区的坡度信息由分辨率为 30 m 的 DEM 数据计算得出, 并以 5° 的等间隔划分为 6 类(图 2a)。坡向被认为是另一个重要的因素, 它通过影

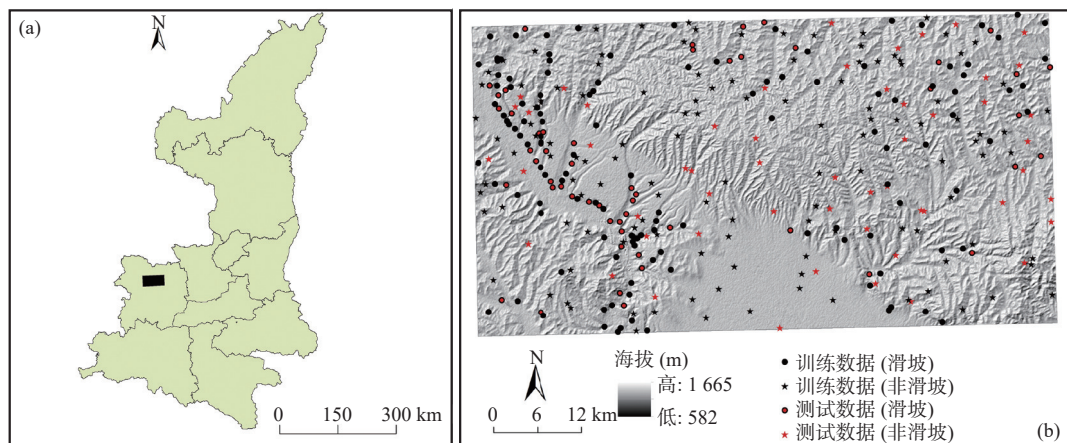
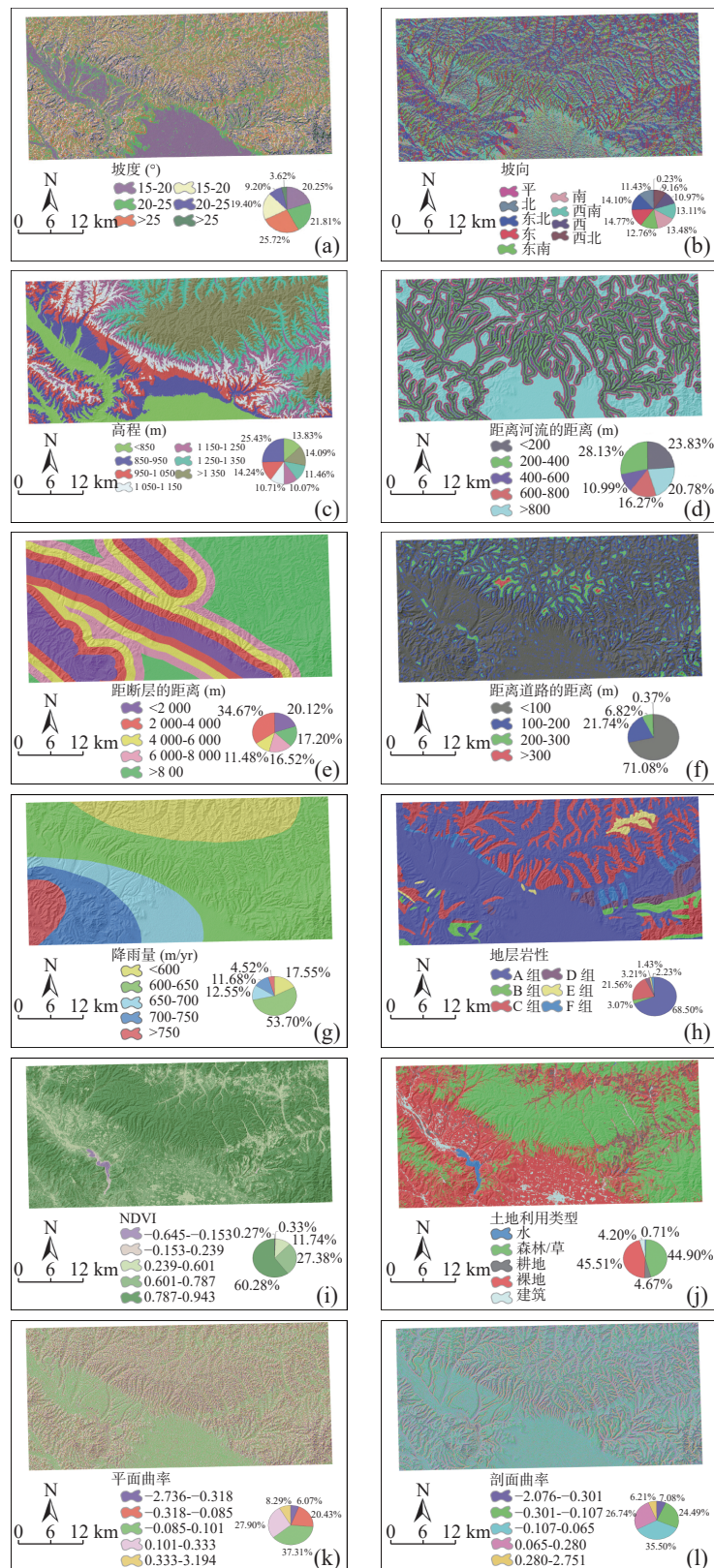


图1 研究区位置图(a)和滑坡编录图(b)

Fig. 1 (a) Location of study area and (b) landslide inventory map



(a).坡度; (b).坡向; (c).高程; (d).距河流的距离; (e).距断层的距离; (f).距道路的距离; (g).降雨量;
(h).地层岩性; (i).NDVI; (j).土地利用类型; (k).平面曲率; (l).剖面曲率

图2 各影响因子图

Fig. 2 Map of various Conditioning factors

响土壤的蒸发量、植被的生长和侵蚀来影响边坡的稳定性。研究区的坡向信息也从 DEM 中提取, 并重分为 9 类(图 2b)。同样地, 从 DEM 中提取高程数据, 按照 1 000 m 的间隔重分为 7 类(图 2c)。

许多学者发现河流侵蚀在滑坡发生中起重要作用(Kumar et al., 2017)。在本研究中, 距河流的距离被认为是重要的影响因素之一, 并以 100 m 的等间隔生成了 5 个不同的缓冲区(图 2d)。构造作用也是影响滑坡发生的重要因素。构造作用不仅是单一滑坡发生的必要条件, 也是区域滑坡发生的直接控制因素。因此, 以距断层的距离作为影响因子, 利用 ArcGIS 软件以 1 000 m 的相等间隔生成 5 个不同的缓冲区(图 2e)。此外, 研究发现道路建设需要开挖大量的边坡, 这很可能在研究区产生潜在的滑坡。因此, 滑坡易发性分区的另一个条件因素是距道路的距离, 以 100 m 的等间隔生成了 4 个不同的道路缓冲区(图 2f)。根据以往的文献, 降雨量对滑坡的稳定性有很大影响(Bui et al., 2017)。本研究使用的降雨数据提取自当地气象台提供的 2007~2017 年的年平均降雨量。将降雨数据按照 150 mm 的等间隔分为 5 类(图 2g)。地层岩性决定了机械强度、耐候性和应力分布, 进而影响边坡的稳定性和侵蚀强度。根据研究区(表 1)的内容, 将地层岩性重新划分为 6 组(图 2h)。

表 1 研究区地层岩性单元表
Tab. 1 Lithological units of study area

类别	地质年代	编码	主要地层岩性
A	新近纪	Q ₄	砂, 砾石, 黄土
	更新纪	Q ₃	黄土, 砾石
B	上新统	N ₂	砂土
	中新统	N ₁	石英砂, 黏土
C	晚白垩纪	K ₁	泥岩, 砂质泥岩, 泥质砂岩
D	早侏罗纪	J ₃	块状聚集物, 粘胶岩, 粉砂质泥岩
	中侏罗纪	J ₂	石质, 泥岩, 粉质泥岩
	晚侏罗纪	J ₁	粉砂岩, 煤层
E	早三叠纪	T ₃	泥岩, 页岩, 煤层
	中三叠纪	T ₂	中细砂岩, 粉砂岩, 泥岩
	晚三叠纪	T ₁	石质, 细砂岩, 粉砂岩, 砂质泥岩
F	二叠纪	P	砂质泥岩, 细砂岩, 粉砂岩

植被对滑坡的发生也有显著影响(Arabameri et al., 2018)。NDVI 一直被用作反映地表植被覆盖程度的重要指标, 本研究利用 ENVI 软件, 基于 GF-2 多光谱遥感影像计算 NDVI, 采用自然断点法(图 2i)将

NDVI 分为 5 类。土地利用类型通常被认为是滑坡发生的基本影响因素之一, 本研究利用 ArcGIS 软件将土地利用类型分为 5 组(图 2j)。此外, 基于 DEM 数据计算出平面曲率和剖面曲率, 并分别使用自然断点法(图 2k、图 2l)分为 5 个部分。最后, 为了保持数据分辨率的统一, 基于 ArcGIS 软件, 将这 12 个影响因子图全部重采样为 30 m 空间分辨率。

3 研究方法

本研究主要包括 4 个部分: 第 1 部分是基础数据的制备, 包括滑坡编录图的构建、滑坡影响因子的分析以及训练和测试数据集的准备; 第 2 部分涉及滑坡影响因子的潜在多重共线性问题的判断、滑坡影响因子的优化以及分形维数的计算; 第 3 部分主要建立了 4 种滑坡易发性分区模型, 即熵权模型(IOE)、基于分形维数的熵权模型(F-IOE)、支持向量机模型(SVM)和基于分形维数的支持向量机模型(F-SVM)。然后根据上述 4 种模型生成研究区的滑坡易发性分区图; 第 4 部分是利用统计学指标评价滑坡易发性模型的性能, 并根据 ROC 特征曲线对这 4 种模型的泛化性进行比较。接下来简要介绍本研究中采用的方法。

3.1 多重共线性

多重共线性是指由于影响因子之间存在高度相关性, 模型被扭曲或难以估计(Islam et al., 2011)。为了检测多重共线性的存在, 本研究引入了方差膨胀因子(VIF)。VIF 表示影响因子之间存在潜在多重共线性问题时的方差与不存在潜在多重共线性问题时的方差之比, VIF 的倒数是公差(TOL)。通常, VIF 值与影响因子之间的多重共线性问题的强度成正比。当 VIF>10 且 TOL<0.1 时, 认为数据集中存在潜在的多重共线性问题。

3.2 影响因子优选

在滑坡易发性分区制图中, 并非所有影响因子在建模中都具有相同的预测能力, 并且可能存在噪声, 这将使结果产生误差。因此, 可以计算信息增益率(IG)来定量反映每个影响因子对滑坡易发性模型的贡献, IG 值越高, 表明滑坡易发性映射的预测能力越高(Chen et al., 2017)。如果 IG 值等于 0, 这意味着影响因子没有预测能力, 应放弃。IG 值计算公式如下:

$$Entropy(D) = - \sum_{k=1}^{|D|} p_k \log_2 p_k \quad (1)$$

其中: D 是训练数据集; 熵(D)表示训练数据集的熵; y 是 D 中的物种数; p_k 表示 D 中 k 类的比例。然后使用表示影响因子之一的 s 将训练数据集划分为 $D_v (v=1, 2, 3, \dots, m)$, 并通过等式(2)计算增益(D, s)。

$$Gain(D, s) = Entropy(D) - \sum_{v=1}^m \frac{|D_v|}{D} Entropy(D_v) \quad (2)$$

调节因子的 IG 值计算为:

$$IG(D, s) = -\frac{Gain(D, s)}{IV(s)} \quad (3)$$

其中: $IV(s)$ 可通过等式(4)获得。

$$IV(s) = -\sum_{v=1}^m \frac{|D_v|}{|D|} \log_2 \frac{|D_v|}{|D|} \quad (4)$$

3.3 分形维数

分形维数是分形理论中用来定量描述分形特征和几何复杂性的参数(Shuren et al., 2005)。本研究引入分形维数来表示滑坡的聚类程度和滑坡系统的复杂性。在当前研究中, 使用 ArcGIS 软件将研究区域划分为边长为 r 的 n 个网格。然后将 r 值减少一半。在本研究中, r 值分别设定为 10 000 m、5 000 m、2 500 m、1 250 m、625 m 和 312.5 m。计算滑坡点位于不同边长的每个网格中的网格数(N_r), 分形维数计算示意图见图 3。然后, 使用最小二乘法拟合直线, 如下所示:

$$\ln N_r = a + f_j \ln r \quad (5)$$

其中: a 是 1 个参数; f_j 表示分形维数(Bouboulis et al., 2006)。最后, 通过盒计数技术测量每个影响因子的分形维数, 并将其作为滑坡易发性建模的输入数据。

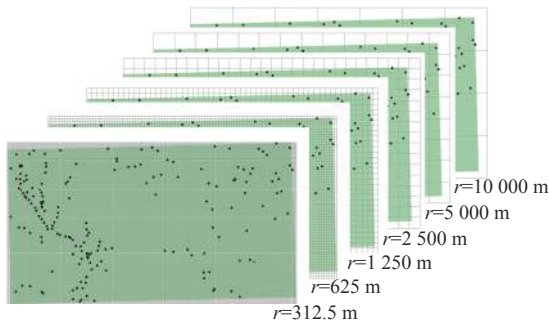


图3 分形维数计算示意图

Fig. 3 The schematic about the calculation of fractal dimensions

3.4 熵权模型(IOE)和基于分形维数的熵权模型(F-IOE)

IOE 作为二元线性统计分析模型被广泛应用(Li

et al., 2019)。原始 IOE 模型通过计算滑坡的频率比来计算每个影响因子的权重, 以建立滑坡易发性模型。具体计算过程如下:

$$FR_{ij} = \frac{b_{ij}}{a_{ij}} \quad (6)$$

$$P_{ij} = \frac{FR_{ij}}{\sum_{j=1}^{N_j} FR_{ij}} \quad (7)$$

$$H_j = -\sum_{i=1}^{N_j} P_{ij} \log_2 P_{ij}, j = 1, 2, 3, \dots, n \quad (8)$$

$$H_{j\max} = \log_2 N_j \quad (9)$$

$$I_j = \frac{H_{j\max} - H_j}{H_{j\max}} \quad (10)$$

$$W_j = I_j \times FR_{ij} \quad (11)$$

其中: a_{ij} 和 b_{ij} 分别代表像元百分比和滑坡百分比; N_j 是影响因子的分类数; FR_{ij} 是滑坡发生的频率; H_j 和 $H_{j\max}$ 是熵值, W_j 代表影响因子的综合评分。对于新的混合模型(F-IOE), 使用每个因子的分形维数 f_j 代替 H_j 作为建模数据。影响因子的最终权重由 $F-W_j$ 表示(贾俊等, 2023)。

3.5 支持向量机模型(SVM)和基于分形维数的支持向量机模型(F-SVM)

支持向量机(SVM)被认为是解决二值分类问题的有效监督学习算法(Wang et al., 2021)。支持向量机模型的基本理论是在特征空间中找到最佳分离超平面, 该超平面可以最大化训练数据集中滑坡和非滑坡数据之间的样本间隔(Balogun et al., 2021), 可以使用以下数学公式表示:

$$P = \frac{1}{2} \|w\|^2 \quad (12)$$

约束条件为:

$$y_i((w \times x_i) + b) \geq 1 \quad (13)$$

其中: b 为标量; $\|w\|^2$ 是超平面法向量的范数。引入拉格朗日乘数法求极值(λ_i), 生成辅助函数如下:

$$L = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \lambda_i (y_i((w \times x_i) + b) - 1) \quad (14)$$

对于不可分离的情况, 引入松弛变量 $\xi_i (i=1, 2, \dots, n)$, 等式(12)和等式(13)可以用如下代替:

$$y_i((w \times x_i) + b) \geq 1 - \xi_i \quad (15)$$

$$L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i, \left(C = -\frac{1}{vn} \right) \quad (16)$$

其中: C 为约束系数; $v \in (0, 1]$ 代表误分类罚值。此外, 非线性决策边界可以使用核函数 $K(x_i, x_j)$ 运算。

本研究采用径向基函数(RBF)进行滑坡易发性建模,其数学公式如下:

$$K(x_i, x_j) = \exp(-\delta \|x_i - x_j\|^2), \delta > 0 \quad (17)$$

其中: δ 说明了高斯核函数的宽度。

同样,使用分形维数 f_j 替换每个影响因子的原始数值数据作为输入数据,然后应用等式(12)在高维空间中使用超平面对输入数据进行分类。最后,构建一种新的混合模型(F-SVM)。

3.6 统计学指标

3.6.1 PPR, NPR 和 ACC

本研究采用3种统计指标,即阳性预测率(Positive Prediction Rate, PPR)、阴性预测率(Negative Prediction Rate, NPR)和准确性(Accuracy, ACC)来评估滑坡易发性制图的结果。采用PPR评价滑坡易发性模型对滑坡的预测能力,采用NPR评价滑坡易发性模型对非滑坡的预测能力。采用ACC测量滑坡易发性分区的准确性。这3个统计指标的计算过程如下:

$$PPR = \frac{TP}{TP + FP} \quad (18)$$

$$NPR = \frac{TN}{TN + FN} \quad (19)$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

其中: TP 全名为真阳性; TN 全名为真阴性,表示正确分类的像素数; FP 的全名为假阳性; FN 的全名为假阴性,表示未正确分类的像素数(杨光等, 2019)。一般来说,ACC值越大,滑坡易发性分区的精度越高(Pham et al., 2021)。

3.6.2 受试者工作特征曲线(ROC)

对4种模型(IOE、F-IOE、SVM和F-SVM)获得的分类结果进行比较评估是一个重要的部分(张文龙等, 2023)。在本研究中,引入了ROC曲线来比较这4种滑坡易发性模型,揭示了使用合成法的灵敏度和特异度值的关系。它为连续变量设置几个不同的临界值,以计算一系列灵敏度和特异度,然后绘制以灵敏度作为Y轴、1-特异度作为X轴的曲线。曲线下面积(简称AUC)用于比较本研究中每个模型的性能。AUC值范围为0~1。如果AUC=1,则表明滑坡和非滑坡的分类完全正确(Lombardo et al., 2020)。相反,如果AUC=0,则分类完全错误。计算公式如下:

$$\text{灵敏度} = \frac{TP}{TP + FN} \quad (21)$$

$$1 - \text{特异度} = 1 - \frac{TN}{TN + FP} \quad (22)$$

$$AUC = \frac{(\sum TP + \sum TN)}{P + N} \quad (23)$$

其中: P 和 N 分别表示研究区域内滑坡和非滑坡的总数。

4 结果

4.1 多重共线性诊断结果

在本研究中,VIF和TOL被用于检测影响因子之间潜在的多重共线性问题。从计算结果(表2)可以看出,VIF的最大值为NDVI(VIF=1.675),公差的最小值为土地利用类型(公差=0.650)。所有影响因子的VIF和TOL不在潜在的多重共线性问题阈值内(VIF>10, TOL<0.1),表明这12个影响因子中不存在多重共线性问题。

表2 影响因子的VIF和TOL表

Tab. 2 Variance inflation factors (VIF) and tolerances of each conditioning factor

影响因子	TOL	VIF
坡度	0.934	1.071
坡向	0.926	1.080
高程	0.656	1.525
距河流的距离	0.908	1.101
距道路的距离	0.877	1.141
距断层的距离	0.916	1.092
NDVI	0.597	1.675
土地利用类型	0.650	1.538
地层岩性	0.814	1.228
降雨量	0.814	1.229
平面曲率	0.912	1.096
剖面曲率	0.925	1.082

4.2 影响因子优选

本研究通过10折交叉验证计算平均信息增益率(AM)来选择影响因子。从图4可以看出,所有影响因子的AM中,高程的最高(0.598),其次是坡度和坡向的相同(0.299),距河流的距离(0.253),降雨量(0.230),土地利用类型(0.210),距道路的距离(0.184),距断层的距离和NDVI的相同(0.115),剖面曲率(0.046)和地层岩性(0.023)。然而,平面曲率的AM值为0,这表明平面曲率因子对建模的贡献为0,可能会造成干扰。因此,在后续建模和评估中排除平面曲率。

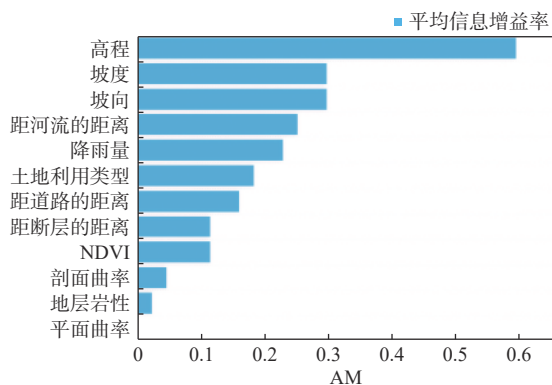


图4 各影响因子的平均信息增益率图

Fig. 4 Average Information Gain ratio of each conditioning factor

4.3 分形维数的计算

根据分形理论,为了计算影响因子的分形维数,研究区滑坡的空间分布必须满足分形特征(自相关性)。根据上述方法,使用公式(5)获得研究区内滑坡的空间分形维数。从图5可以看出,研究区滑坡的空间分形维数为1.903 2, R^2 为0.964 9,这表明滑坡的空间分布具有明显的分形特征和显著的自相关性。因此,继续计算每个影响因子的空间分维是有意义的,每个因子的分维数 f_j 见表3。

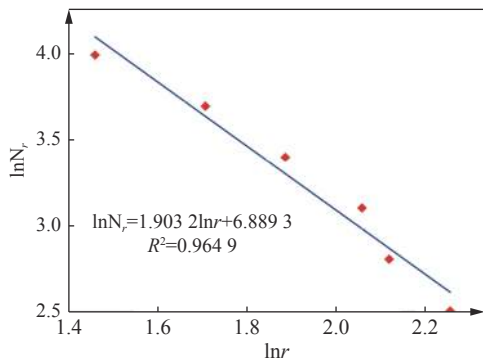


图5 对数图和线性方程图

Fig. 5 Logarithmic graphs and linear equations

4.4 基于 IOE 和 F-IOE 模型的滑坡易发性分区

使用 IOE 和 F-IOE 模型的最终计算数据见表3。如结果所示,每个影响因子的 IOE 值和分形维数具有基本相同的趋势。最大 W_j (0.364 6) 属于 NDVI, 其次是土地利用类型 (0.246 2), 距道路的距离 (0.169 8), 最小 W_j (0.015 5) 属于距断层的距离。最大 $F-W_j$ (0.138 8) 属于剖面曲率, 其次是 NDVI (0.133 7), 降雨量 (0.131 9), 最小 $F-W_j$ (0.094 6) 属于坡向。

从以上数据可以看出, NDVI 在 IOE 和 FIOE 模型

的建模过程中具有很高的权重。造成这种现象的原因是研究区水土流失严重, 降雨不足。基于 IOE 和 F-IOE 模型计算滑坡易发性指数 (LSI)。LSI 值在 0 到 1 之间, 越接近 1, 滑坡发生的概率越高, 反之亦然。最后, 将 LSI_{IOE} 和 LSI_{F-IOE} 划分为 5 个区间, 以生成滑坡易发性图 (LSM), 同时也通过自然断点法分为 5 类区域: 极低易发区 (0.094 0~0.208 4)、(0.104 7~0.194 5); 低易发区 (0.208 4~0.264 6)、(0.194 5~0.278 3); 中易发区 (0.264 6~0.322 8)、(0.278 3~0.383 4); 高易发区 (0.322 8~0.389 3)、(0.383 4~0.502 2); 极高易发区 (0.389 3~0.624 4)、(0.502 2~0.721 4) (图 6a、图 6b)。

4.5 基于支持向量机和 F-SVM 模型的滑坡易发性分析

在建立基于 RBF 核函数的支持向量机和 F-SVM 模型之前, C 和 δ 的确定对于建模至关重要。为了获得准确且稳定的结果, 基于 Python 环境, 利用 10 折交叉验证和网格搜索法, 确定了 SVM 和 F-SVM 的 C 和 δ 分别为 (256, 0.125 0) 和 (512, 0.016 5)。然后计算 SVM 和 F-SVM 模型的 LSI, 输出范围为 0~1, 越接近 1, 滑坡发生的概率越高, 反之亦然。最后, 将 LSI_{SVM} 和 LSI_{F-SVM} 划分为 5 个区间, 生成 LSM, 并通过自然断点法将其分为 5 类: 极低易发区 (0.014 5~0.245 9)、(0.074 0~0.306 1); 低易发区 (0.245 9~0.369 5)、(0.306 1~0.396 2); 中易发区 (0.369 5~0.516 1)、(0.396 2~0.496 7); 高易发区 (0.516 1~0.697 4)、(0.496 7~0.680 2); 极高易发区 (0.697 4~0.998 3)、(0.680 2~0.957 4) (图 6c、图 6d)。

4.6 结果的评估和比较

4.6.1 模型性能评估

基于训练数据集, 利用 3 种统计学指标 (PPR、NPR 和 ACC) 对滑坡易发性模型的性能进行评估, 结果见表 4。从 PPR 和 NPR 的计算结果可以看出, F-SVM 模型的最高值分别为 88.72% 和 94.02%。说明在本研究中, F-SVM 模型对滑坡是否发生的预测能力最强。从 ACC 计算结果来看, F-SVM 模型的值也最高 (91.20%), 其次是 SVM 模型 (89.20%)、F-IOE 模型 (87.60%) 和 IOE 模型 (84.80%)。

4.6.2 模型测试

使用测试数据集和 3 种统计指标完成了模型测试, 计算结果如表 5 所示。从 PPR 和 NPR 的计算结果可以看出, F-SVM 模型的值最高, 分别为 92.73% 和 94.34%。

这表明, 在目前的研究中, F-SVM 模型在这 4 种

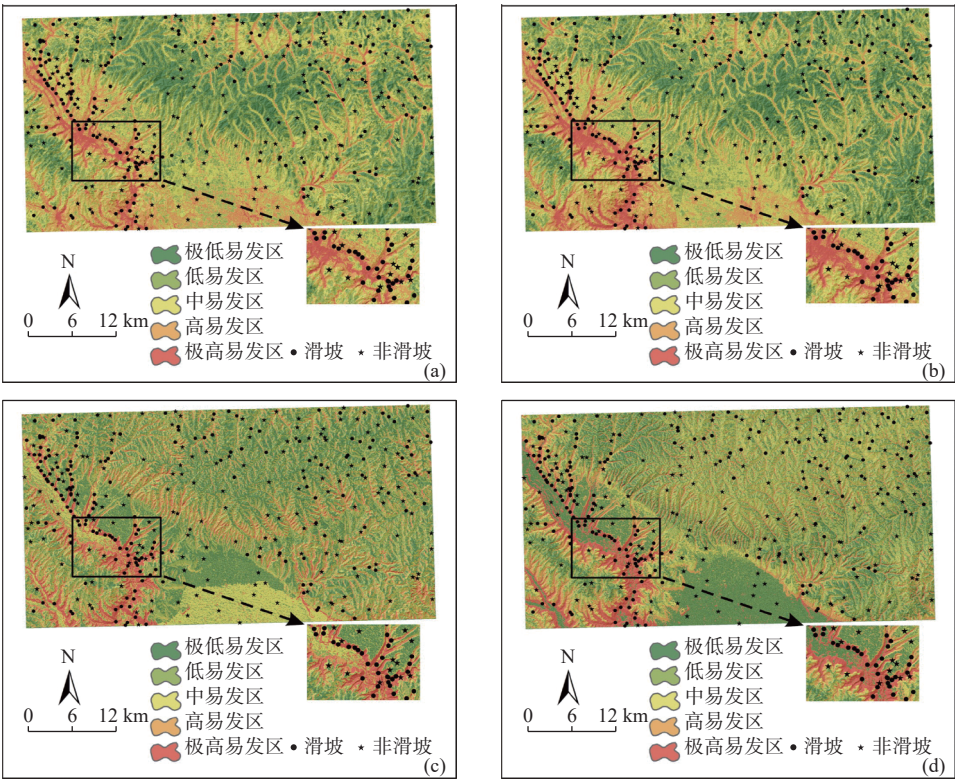
表 3 影响因子与滑坡的空间关系表

Tab. 3 Spatial relationship between influencing factors and landslides

影响因子	等级	f_j	FR_{ij}	P_{ij}	H_j	H_{jmax}	I_j	W_j	$F-W_j$
坡度 (°)	< 5	0.077 2	0.303 5	0.036 7					
	5~10	0.190 6	0.870 9	0.105 3					
	10~15	0.228 5	0.912 4	0.110 4					
	15~20	0.237 2	1.151 8	0.139 3	2.285 7	2.585 0	0.115 8	0.159 5	0.120 1
	20~25	0.168 2	1.942 7	0.235 0					
	> 25	0.228 4	3.086 4	0.373 3					
坡向	水平	0.000 0	0.000 0	0.000 0					
	北	0.045 9	0.609 7	0.078 7					
	东北	0.031 7	0.356 4	0.046 0					
	东	0.223 9	0.980 2	0.126 6					
	东南	0.156 7	1.036 4	0.133 8	2.900 4	3.169 9	0.085 0	0.073 2	0.094 6
	南	0.142 9	1.138 7	0.147 0					
	西南	0.199 0	1.021 5	0.131 9					
	西	0.226 0	1.624 8	0.209 8					
	西北	0.125 1	0.977 1	0.126 2					
高程 (m)	< 850	0.580 3	2.908 5	0.402 4					
	850~950	0.334 6	1.189 8	0.164 6					
	950~1 050	0.213 1	0.720 2	0.099 6					
	1 050~1 150	0.100 9	0.554 9	0.076 8	2.466 6	2.807 4	0.121 4	0.125 3	0.127 9
	1 150~1 250	0.094 2	0.573 8	0.079 4					
	1 250~1 350	0.160 6	0.863 0	0.119 4					
	> 1 350	0.209 2	0.417 4	0.057 8					
距河流的距离 (m)	< 200	0.399 8	2.273 8	0.470 3					
	200~400	0.222 0	0.833 6	0.172 4					
	400~600	0.053 9	0.446 4	0.092 3	2.015 4	2.321 9	0.132 0	0.127 7	0.100 9
	600~800	0.070 6	0.864 1	0.178 7					
	> 800	0.194 2	0.417 1	0.086 3					
距断层的距离 (m)	< 2000	0.435 8	1.221 6	0.227 5					
	2 000~4 000	0.272 5	1.234 3	0.229 9					
	4 000~6 000	0.328 6	1.149 6	0.214 1	2.288 3	2.321 9	0.014 5	0.015 5	0.126 3
	6 000~8 000	0.399 2	1.119 1	0.208 4					
	> 8 000	0.295 1	0.644 5	0.120 0					
距道路的距离 (m)	< 100	0.486 6	1.147 6	0.530 2					
	100~200	0.140 7	0.770 9	0.356 2					
	200~300	0.133 4	0.245 8	0.113 0	1.372 3	2.000 0	0.313 9	0.169 8	0.103 0
	> 300	0.000 0	0.000 0	0.000 0					

模型中对滑坡和非滑坡发生的分类性能最好。对于 ACC 指标, F-SVM 模型的值也最高 (93.52%), 其次是

SVM 模型 (92.59%)、F-IOE 模型 (88.89%) 和 IOE 模型 (87.96%)。总的来说, 这 4 种模型的性能对于研究区



(a).IOE 模型(LSM); (b).F-IOE 模型(LSM); (c).SVM 模型(LSM); (d).F-SVM 模型(LSM)

图6 滑坡易发性图

Fig. 6 Landslide susceptibility map using

表 4 模型性能评价统计指标计算结果表

Tab. 4 Calculation result of statistical indicators for model performance evaluation

指标	模型			
	IOE	SVM	F-IOE	F-SVM
真阳性	108	113	110	118
真阴性	104	110	109	110
假阳性	21	15	16	15
假阴性	17	12	15	7
PPR(%)	83.72	88.28	87.30	88.72
NPR(%)	85.95	90.16	87.90	94.02
ACC(%)	84.80	89.00	87.60	91.20

表 5 模型测试评价统计指标计算结果表

Tab. 5 Calculation result of statistical indicators for model validation evaluation

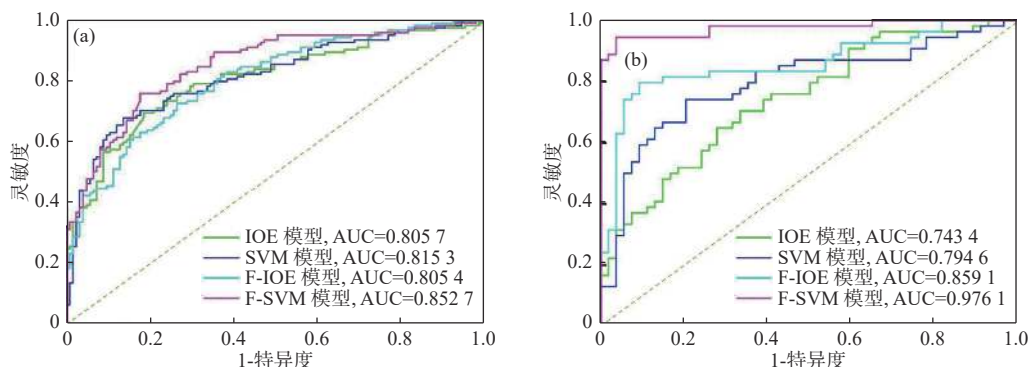
指标	模型			
	IOE	SVM	F-IOE	F-SVM
真阳性	48	52	110	118
真阴性	47	48	109	110
假阳性	7	6	16	15
假阴性	6	2	15	7
PPR(%)	87.27	89.66	86.21	92.73
NPR(%)	88.68	96.00	92.00	94.34
ACC(%)	87.96	92.59	88.89	93.52

域的滑坡易发性分区相对可靠。此外,根据统计测试方法的结果,F-SVM模型在所有指标上得分最高,具有更好的分类能力。在本研究中,两种混合模型的性能都优于其原始模型。

4.6.3 模型泛化性对比

本研究通过绘制 ROC 曲线,计算 AUC 值来对比滑坡易发性模型泛化性。基于训练和测试数据集计算的 AUC 值分别如图 7a 和图 7b 所示。从基于训练

数据集的 ROC 曲线来看,F-SVM 模型的 AUC 值最高(0.852 7),其次是 SVM 模型(0.815 3)、IOE 模型(0.805 7)和 F-IOE 模型(0.805 4)。从基于测试数据集的 ROC 曲线来看,F-SVM 模型的 AUC 值最高(0.976 1),其次是 F-IOE 模型(0.859 1)、SVM 模型(0.794 6)和 IOE 模型(0.743 4)。结果表明,F-SVM 在训练和测试数据集中表现出最强的泛化性,并且测试数据集中的 AUC 值明显高于其他模型。



(a) 训练数据集; (b) 测试数据集

图7 滑坡易发性模型的 ROC 曲线图

Fig. 7 ROC curves for landslide susceptibility models

5 讨论

在本研究中, 基于滑坡编录图, 在陕西省宝鸡市北部黄土覆盖区应用了4种滑坡易发性模型, 即 IOE、F-IOE、SVM、F-SVM 模型, 并进行了滑坡易发性分区。结果表明, 两种混合模型(F-IOE、F-SVM)对于滑坡具有更好的分类能力。虽然许多学者已经使用分形维数从微观和宏观上研究了滑坡, 但基本上没有使用分形维数作为输入数据建立混合模型的研究。由于分形维数反映了滑坡更真实的空间分布, 分形维数表示的不确定性不等于随机性, 分形维数与影响因子之间不存在线性关系。因此, 通过使用分形维数作为输入数据构建混合模型, 可以实现比普通混合模型更好的效果。另一方面, 分形维数可以作为未来研究中影响因子优选的指标。在滑坡易发性建模过程中, 尺度和数据分辨率的差异可能会产生一些不确定性。在区域滑坡易发性分区中, 通常需要选择一系列在时空尺度上具有不同变化率的影响因子。例如, 降雨量和坡度最有可能发生变化, 但地层岩性变化的可能性最小。在大尺度上, 影响因子的空间总体特征往往更受关注, 相反, 在小尺度下, 影响因子在空间中的整体特征通常被视为约束条件, 而细节信息则更受关注。随着空间尺度的增加, 变化率高的因子所带来的影响将减少甚至消除, 而变化率低的因子所带来的影响将保持甚至更突出。这些原因都将导致滑坡易发性分区的结果具有不确定性, 因此, 笔者将更加关注尺度变化对滑坡易发性分区的影响。

从模型性能的评估结果来看, 用分形维数构建的两种混合模型(F-IOE、F-SVM)的 PPR、NPR 和 ACC

值高于其原始模型(IOE、SVM)。同样, 从模型测试的结果来看, 两种混合模型的上述3个统计学指标也高于其原始模型。

从图7b可以看出, 两种混合模型的 AUC 值分别为 0.8591(F-IOE)和 0.9761(F-SVM), 显著高于原始模型。因此, 构建以分形维数为输入数据的混合模型来研究滑坡易发性比原始模型更有优势。然而, 研究区域不足以反映整个方法的普遍适用性, 需要结合更多的区域进行研究。

6 结论

(1)本研究基于179个滑坡样本点, 随机选择70%的滑坡点进行训练, 其余30%的滑坡点用于测试。选择了12个影响因子, 包括坡度、坡向、高程、距河流的距离、距道路的距离、距断层的距离、NDVI、土地利用类型、地层岩性、降雨量、平面曲率和剖面曲率。多重共线性诊断结果表明, 影响因子之间不存在潜在的多重共线性问题, 然后计算影响因子的信息增益率, 排除了平面曲率因素的影响。

(2)构建了两种原始分类模型, 即 IOE 和 SVM 模型。同时, 计算每个影响因子的分形维数, 并将其作为输入数据来构建两种混合模型, 即 F-IOE 和 F-SVM 模型。最终使用自然断点法将滑坡易发性分区图分为5类: 极低易发区、低易发区、中易发区、高易发区和极高易发区。最后, 使用3种统计学指标(PPR、NPR 和 ACC)来评估模型的性能, 并使用 ROC 曲线来比较这4种模型的泛化性。

(3)4种模型的 ACC 值都高于80%, 这表明所有模型的分类结果都可靠, 并且两种混合模型的 ACC 值都高于其原始模型。此外, 从基于测试数据集的

ROC 曲线来看,结果相似,F-IOE 模型和 F-SVM 模型的 AUC 值分别为 0.859 1 和 0.976 1,高于 IOE 模型(0.743 4)和 SVM 模型(0.794 6)。

(4)本研究的结果表明混合模型在滑坡易发性分区中的可靠性高,可以减少野外滑坡详细调查的工作量,为更精准、有效的滑坡灾害预测预警提供数据支撑,同时,相关图件的发布也可为当地的防灾减灾工作提供决策支持。

参考文献(References):

- 韩玲,张庭瑜,张恒.基于 IOE 和 SVM 模型的府谷镇滑坡易发性分区[J].水土保持研究,2019,26(3):373-378.
- HAN Ling, ZHANG Tingyu, ZHANG Heng. Landslide susceptibility mapping based on IOE and SVM model in Fugu town[J]. Research of Soil and Water Conservation, 2019, 26(3): 373-378.
- 刘光旭,席建超,戴尔阜,等.中国滑坡灾害承灾体损失风险定量评估[J].自然灾害学报,2014,33(2):39-46.
- LIU Guangxu, XI Jianchao, DAI Erfu, et al. Loss risk assessment of the hazard-affected body of landslides in China[J]. Journal of Natural Disasters, 2014, 33(2): 39-46.
- 孙萍萍,张茂省,贾俊,等.中国西部黄土区地质灾害调查研究进展[J].西北地质,2022,55(3):96-107.
- SUN Pingping, ZHANG Maoxing, JIA Jun, et al. Progress in Geological Hazard Investigation and Research in Loess Regions of Western China[J]. Northwestern Geology, 2022, 55(3): 96-107.
- 陕西省地质矿产局.陕西省区域地质志[M].西安:地质出版社,1989. Bureau of Geology and Mineral Resources of Shaanxi Province, 1989. Regional geology of Shaanxi Province[M]. Xi'an: Geological Publishing House, 1989.
- 张茂省,薛强,贾俊,等.山区城镇地质灾害调查与风险评价方法及实践[J].西北地质,2019,52(2):125-135.
- ZHANG Maoxing, XUE Qiang, JIA Jun, et al. Method And Practice of Geological Hazard Investigation and Risk Assessment in Mountainous Towns[J]. Northwestern Geology, 2019, 52(2): 125-135.
- 贾俊,李志忠,郭小鹏,等.多源遥感技术在降雨诱发勉县地质灾害调查中的应用[J].西北地质,2023,56(3):268-280.
- JIA Jun, LI Zhizhong, GUO Xiaopeng, et al. Application of Multi-Source Remote Sensing Technology in the Investigation of Geological Disasters Induced by Rainfall in Mianxian County[J]. Northwestern Geology, 2023, 56(3): 268-280.
- 武雪玲,沈少青,牛瑞卿. GIS 支持下应用 PSO-SVM 模型预测滑坡易发性[J]. 武汉大学学报(信息科学版),2016,41(5):665-671.
- WU Xueling, SHEN Shaoqing, NIU Ruiqin. PSO-SVM model is used to predict landslide susceptibility with GIS support[J]. Geomatics and Information Science of Wuhan University, 2016, 41(5): 665-671.
- 谢振华,窦培谦.基于 BP 神经网络的矿山排土场滑坡预警模型[J].金属矿山,2017,(6):166-169.
- XIE Zhenhua, DOU Peiqian. Landslide early warning model for mine dumping site based on BP neural network[J]. Metal Mine, 2017, (6): 166-169.
- 杨光,徐佩华,曹琛,等.基于确定性系数组合模型的区域滑坡敏感性评价[J].工程地质学报,2019,27(5):1153-1163.
- YANG Guang, XU Peihua, CAO Chen et. al. Assessment of regional landslide susceptibility based on combined model of certainty factor method[J]. Journal of Engineering Geology, 2019, 27(5): 1153-1163.
- 张文龙,张振凯,杨帅.勉略宁地区地质灾害危险性智能评价和区划研究[J].西北地质,2023,56(1):276-282.
- ZHANG Wenlong, ZHANG Zhenkai, YANG Shuai. Research on Intelligent Evaluation and Zoning of Geological Hazard Risk in Mian-lv-ning Area[J]. Northwestern Geology, 2023, 56(1): 276-282.
- 张庭瑜,韩玲,张恒,等.混合分类模型在滑坡易发性分区中的适用性研究——以延安市宝塔区为例[J].干旱区资源与环境,2020,34(1):192-201.
- ZHANG Tingyu, HAN Ling, ZHANG heng, et al. Feasibility of hybrid classification models for landslide susceptibility mapping—A case of Baota district, Yan'an City[J]. Journal of Arid Land Resources and Environment, 2020, 34(1): 192-201.
- 张向营,张春山,孟华君,等.基于 Random Forest 和 AHP 的贵德县北部山区滑坡危险性评价[J].水文地质工程地质,2018,45(4):142-149.
- ZHANG Xiangying, ZHANG Chunshan, MENG Huajun, et. al. Landslide hazard evaluation in the northern mountainous area of Guide County based on Random Forest and AHP[J]. Hydrogeology & Engineering Geology, 2018, 45(4): 142-149.
- Abedini M, Tulabi S. Assessing LNRF, FR, and AHP models in landslide susceptibility mapping index: a comparative study of Nojjan watershed in Lorestan province, Iran[J]. Environmental Earth Sciences, 2018, 77: 405-405.
- Arabameri A, Pradhan B, Rezaei K, et al. GIS-based landslide susceptibility mapping using numerical risk factor bivariate model and its ensemble with linear multivariate regression and boosted regression tree algorithms[J]. Journal of Mountain Science, 2018, 16: 595-618.
- Balogun A L, Rezaie F, Pham Q B, et al. Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with with GWO, BAT and COA algorithms[J]. Geoscience Frontiers, 2021, 12: 101-104.
- Bazi Y, Bashmal L, Rahhal M, et al. Vision Transformers for Remote Sensing Image Classification[J]. Remote Sensing, 2021, 13: 516.
- Behnia P, Blais-stevens A. Landslide susceptibility modelling using the quantitative random forest method along the northern portion of the Yukon Alaska Highway Corridor, Canada[J]. Natural Hazards, 2018, 90: 1407-1426.
- Bouboulis P, Dalla L, Drakopoulos V. Construction of recurrent bivariate fractal interpolation surfaces and computation of their

- box-counting dimension[J]. *Journal of Approximation Theory*, 2006, 141: 99–117.
- Bui D T, Tuan T A, Hoang N D, et al. Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization[J]. *Landslides*, 2017, 14: 447–458.
- Chen W, Li Y, Tsangaratos P, et al. Groundwater Spring Potential Mapping Using Artificial Intelligence Approach Based on Kernel Logistic Regression, Random Forest, and Alternating Decision Tree Models[J]. *Applied Sciences*, 2020, 10: 425–441.
- Chen W, Xie X, Wang J, et al. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility[J]. *Catena*, 2017, 151: 147–160.
- Hong H, Liu J, Zhu A X, et al. A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China)[J]. *Environmental Earth Sciences*, 2017, 76: 652–667.
- Islam M Z, Giggins H. Knowledge Discovery through SysFor - a Systematically Developed Forest of Multiple Decision Trees [C], Proceedings of the 9-th Australasian Data Mining Conference, December 15–16; Ballarat, Australia, 2011.
- Jaafari A, Najafi A, Pourghasemi H R, et al. GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran[J]. *International Journal of Environmental Science & Technology*, 2014, 11: 909–926.
- Kim J, Lee S, Jung H. Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea [J]. *Geocarto International*, 2018, 45: 1000–1015.
- Kumar D, Thakur M S, Dubey C, et al. Landslide susceptibility mapping & prediction using Support Vector Machine for Mandakini River Basin, Garhwal Himalaya, India[J]. *Geomorphology*, 2017, 295: 115–125.
- Li J, Zhang Y. GIS-supported certainty factor (CF) models for assessment of geothermal potential: A case study of Tengchong County, southwest China[J]. *Energy*, 2017, 140: 552–565.
- Li R, Wang N. Landslide Susceptibility Mapping for the Muchuan County (China): A Comparison Between Bivariate Statistical Models (WoE, EBF, and IoE) and Their Ensembles with Logistic Regression[J]. *Symmetry*, 2019, 11: 762–781.
- Lombardo L, Tanyas H, Nicu I C. Spatial modeling of multi-hazard threat to cultural heritage sites[J]. *Engineering Geology*, 2020, 277: 105776.
- Mahdadi F, Boumezeur A, Hadji R, et al. GIS-based landslide susceptibility assessment using statistical models: a case study from Souk Ahras province, N-E Algeria[J]. *Arabian Journal of Geosciences*, 2018, 11: 476–488.
- Pandey V K, Pourghasemi H R. Landslide susceptibility mapping using maximum entropy and support vector machine models along the highway corridor, Garhwal Himalaya[J]. *Geocarto International*, 2020, 35: 168–187.
- Pham Q B, Yacine A, Ali S A, et al. A comparison among fuzzy multi-criteria decision making, bivariate, multivariate and machine learning models in landslide susceptibility mapping[J]. *Geomatics, Natural Hazards and Risk*, 2021, 12: 1741–1777.
- Sameen M I, Pradhan B, Lee S. Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment[J]. *Catena*, 2020, 186: 104249.
- Shuren W, Han J, Shi J, et al. A Study of the Relationship between Landslide Stability and Fractal Dimension of the Major Landslide Trace Pattern near Badong in the Reservoir Region of the Yangtze Gorge Project[J]. *Acta Geoscientia Sinica*, 2005, 26: 455–460.
- Thai P B, Indra P. Machine Learning Methods of Kernel Logistic Regression and Classification and Regression Trees for Landslide Susceptibility Assessment at Part of Himalayan Area, India[J]. *Indian Journal of Science & Technology*, 2018, 11: 1–10.
- Wang Z, Brenning A. Active-Learning Approaches for Landslide Mapping Using Support Vector Machines[J]. *Remote Sensing*, 2021, 13: 2588–2607.
- Youssef A M, Pradhan B, Dikshit A, et al. Landslide susceptibility mapping using CNN-1D and 2D deep learning algorithms: comparison of their performance at Asir Region, KSA[J]. *Bulletin of Engineering Geology and the Environment*, 2022, 81: 110–135.