

DOI:10.19322/j.cnki.issn.1006-4710.2024.01.012

<https://xuebao.xaut.edu.cn>

引文格式:陈芯宇,师芸,赵侃,温永啸.基于CF融入SSA优化SVM和RF模型的滑坡易发性评价[J].西安理工大学学报,2024,40(1):121-131,142.

CHEN Xinyu, SHI Yun, ZHAO Kan, WEN Yongxiao. Landslide susceptibility evaluation based on CF integrated with SSA to optimize SVM and RF models [J]. Journal of Xi'an University of Technology, 2024, 40(1): 121-131, 142.

# 基于CF融入SSA优化SVM和RF模型的 滑坡易发性评价

陈芯宇<sup>1,2</sup>, 师芸<sup>1,2</sup>, 赵侃<sup>1,2</sup>, 温永啸<sup>1,2</sup>

(1. 西安科技大学 测绘科学与技术学院, 陕西 西安 710054;

2. 自然资源部煤炭资源勘查与综合利用重点实验室, 陕西 西安 710021)

**摘要:** 针对传统的区域滑坡易发性评价建模过程可能存在的样本数据量纲不统一以及模型参数选取误差等问题,本文以陕西省留坝县为研究区,选取高程、坡度、水系、降雨量、地层岩性等10个评价因子,采用确定性系数模型(CF)计算各评价因子的敏感值作为支持向量机模型(SVM)和随机森林模型(RF)的输入样本属性值,引入麻雀搜索算法(SSA)分别对SVM模型和RF模型的参数进行优化,获取最优参数对两种模型进行训练,最终构建了CF-SSA-SVM和CF-SSA-RF模型,从而对整个研究区进行预测,完成滑坡易发性评价,并通过受试者工作特征曲线(ROC)对两种模型进行精度验证。结果表明,两种模型的评价结果均有较多滑坡点落在极高易发区,无滑坡点落在极低易发区,评价结果均有较高的准确率。其中,CF-SSA-RF模型的成功率和预测率曲线AUC值分别为0.994和0.940,高于CF-SSA-SVM模型;并以三处典型滑坡为例进行验证,结果显示易发性分区与历史滑坡点分布较为吻合。进一步表明CF-SSA-RF模型更适用于留坝县的滑坡易发性评价,为当地滑坡灾害风险评估提供了指导依据。

**关键词:** 易发性评价; 麻雀搜索算法; 随机森林模型; 支持向量机模型; ROC曲线

**中图分类号:** P642

**文献标志码:** A

**文章编号:** 1006-4710(2024)01-0121-11

## Landslide susceptibility evaluation based on CF integrated with SSA to optimize SVM and RF models

CHEN Xinyu<sup>1,2</sup>, SHI Yun<sup>1,2</sup>, ZHAO Kan<sup>1,2</sup>, WEN Yongxiao<sup>1,2</sup>

(1. College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China;

2. Key Laboratory of Coal Resources Exploration and Comprehensive Utilization,  
Ministry of Natural Resources, Xi'an 710021, China)

**Abstract:** For the traditional modeling process of intra-regional landslide susceptibility evaluation, there may be problems such as non-uniformity of sample data outline and errors in the selection of model parameters. This paper takes Liuba County of Shaanxi Province as the research area, selects 10 evaluation factors such as elevation, slope, water system, rainfall, stratigraphic lithology, etc., and uses the certainty factor model (CF) to calculate the sensitivity of each evaluation factor as a support vector machine model (SVM) and random forest model (RF) input sample at-

收稿日期: 2022-09-27; 网络首发日期: 2023-04-24

网络首发地址: <https://kns.cnki.net/kcms/detail/61.1294.N.20230424.1518.014.html>

基金项目: 国家自然科学基金资助项目(41674013, 41874012)

第一作者: 陈芯宇, 女, 硕士生, 研究方向为地质灾害评价。E-mail: 930404815@qq.com

通信作者: 师芸, 女, 博士, 教授, 研究方向为环境遥感与防灾减灾。E-mail: shiyun0908@hotmail.com

tribute values; it introduces the sparrow search algorithm (SSA) to optimize the parameters of SVM model and RF model respectively, obtains the optimal parameters to train the two models, and finally constructs CF-SSA-SVM and CF-SSA-RF models, which can predict the entire study area, complete the landslide susceptibility evaluation, and verify the accuracy of the two models through the receiver operating characteristic curve(ROC). The results show that the evaluation results by the two models have more landslide points in the extremely high-prone areas, and no landslide points in the extremely low-prone areas, and that the evaluation results are of high accuracy. Among them, the AUC values at the success rate and prediction rate curves of the CF-SSA-RF model are 0.994 and 0.940, respectively, which are higher than those by the CF-SSA-SVM model;verified by three typical landslides, the results show that the prone zones and historical landslide points are relatively consistent. It further shows that the CF-SSA-RF model is more suitable for the landslide susceptibility evaluation research in Liuba County, providing a guiding basis for the local landslide disaster risk assessment.

**Key words:** ease of occurrence evaluation; sparrow search algorithm; random forest model; support vector machine model; ROC curve

滑坡是一种严重的自然地质灾害,因其分布广、危害大等特征,一直以来都是国民经济和人民生命安全的巨大威胁<sup>[1]</sup>。据资料统计,2019年全国共发生地质灾害6181起,其中滑坡4220起,占总地质灾害的68%,造成了巨大的经济损失<sup>[2]</sup>。为了减少滑坡灾害带来的人员和经济损失,开展滑坡易发性评价对当地防灾减灾和灾害预警有着重要意义。目前,易发性评价方法分为定性和定量两类,定性分析主要有专家打分法、层次分析法,这类方法在评价过程中主要依靠专家经验,具有较强的主观性<sup>[3-4]</sup>。随着3S技术的日益发展,定性分析逐步发展为定量分析(统计模型和机器学习模型)。统计模型可以明确反映影响因子与滑坡间的关系,原理简单,易于实现,如频率比法、信息量法、确定性系数法等<sup>[5-9]</sup>。以确定性系数模型(certainty factor, CF)为代表,它能够反映出各因子在不同分级类别下对滑坡发育的影响,可以消除不同量纲评价因子赋值引起的误差,但在因子的选取及分级过程中易受人为主观因素的影响。机器学习模型可以避免主观因素的干扰,它是通过较少的样本数据进行建模预测,从而获得较高的准确率,但存在模型输入样本量纲不统一等问题,从而影响评价结果,如逻辑回归、支持向量机、随机森林、神经网络等<sup>[10-13]</sup>。为了弥补单一模型的不足,多模型耦合已成为当前研究的热点话题,并已取得了较好的预测效果。如张钟远等<sup>[14]</sup>将频率比、信息量、确定性系数等单一模型与逻辑回归模型耦合来开展滑坡易发性评价,结果表明,耦合模型比单一模型评价效果更好。李远远等<sup>[15]</sup>采用CF方法计算各因子属性值作为SVM的分类数据,构建CF-

SVM模型对整个研究区进行地质灾害易发性评价,结果表明,CF方法可以消除复杂因子间同区间量化不统一等问题,相较于SVM模型预测效果更好。郑迎凯等<sup>[16]</sup>将CF模型与RF模型耦合,结果显示CF-RF模型具有比RF模型更高的预测率与成功率。

综上所述,在诸多耦合模型中,CF-SVM与CF-RF模型已被广泛应用于滑坡易发性评价中,但该过程往往忽略了模型训练过程中参数的选取问题,多数学者采用默认参数或网格搜索算法寻找最优参数<sup>[17]</sup>,此种方法带有一定的主观性,难以选取最优参数,而参数选取是否合理将直接影响滑坡易发性评价结果。为了合理设置模型参数,本文将引入麻雀搜索算法(sparrow search algorithm, SSA),该算法相较于其他智能优化算法具有较强的探索能力,在短时间内能找到全局最优点。利用该算法的这一优势,对常规支持向量机(support vector machine, SVM)模型与随机森林(random forest, RF)模型进行参数优化,并将CF模型融入SSA优化SVM和RF模型,最终构建成CF-SSA-SVM模型与CF-SSA-RF模型,并用于留坝县滑坡易发性评价中,评价结果可为当地防灾减灾工作提供参考。

## 1 研究区概况及数据来源

### 1.1 研究区概况

研究区留坝县位于陕西省汉中市北部,北纬 $33^{\circ}17' \sim 33^{\circ}52'$ ,东经 $106^{\circ}38' \sim 107^{\circ}18'$ ,地处秦岭南麓腹地,东、西、北三面环山。研究区中南部地势较低,区内地形起伏较大,海拔最低为618 m,最高

为 2 628 m。地貌类型按照形态及物质可划分为山地和河谷两大地貌单元。留坝县属大陆性季风气候区,多年平均气温 11.5℃,该区水系较为发达,褒河自北向南纵贯县境东西两岭之间,支流东疏西密,黑

河次之。区内地层以古生界分布最广,地处秦岭褶皱系,断裂构造复杂,大小和方向不同的断裂使地层遭到严重破坏,区内滑坡共 160 处。研究区概况及滑坡点分布如图 1 所示。

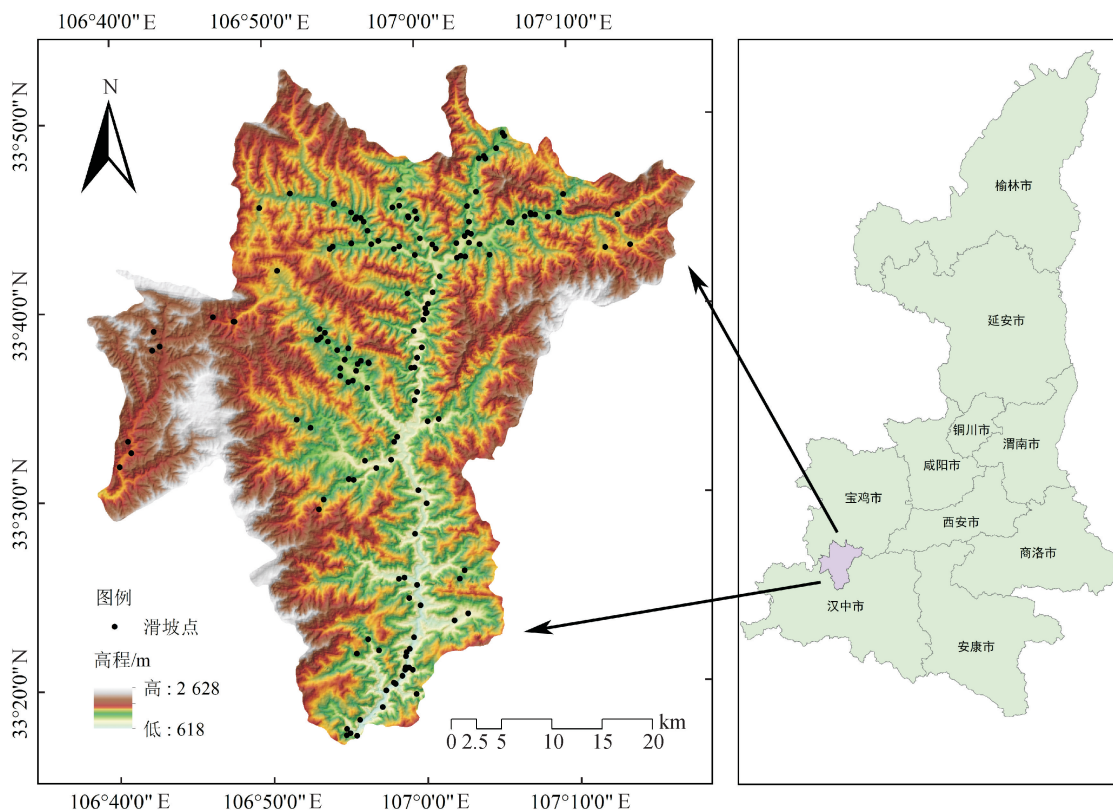


图 1 研究区概况图

Fig. 1 Overview map of the study area

注:此图基于国家自然资源部标准地图服务网站审图号为 GS(2022)1873 的标准地图制作,底图无修改。

## 1.2 数据来源

滑坡易发性评价因子包括地形地貌、基础地质、水文环境、人类工程活动 4 类,其主要数据来源及数据类型见表 1。

表 1 数据来源及类型  
Tab. 1 Data sources and types

数据名称	数据来源	数据类型
滑坡灾害点	中国科学院资源环境科学数据中心	点数据
高程、坡度、坡向、水系	地理空间数据云 ASTER GDEM	30 m×30 m 栅格
工程地质岩组、断层	地质图	1 : 50000 栅格
降雨量	中国气象数据网	点数据
植被指数	Landsat 8 OLI	30 m×30 m 栅格
道路	OpenStreetMap	矢量数据
土地利用类型	FROM-GLC	30 m×30 m 栅格

## 2 研究方法

### 2.1 CF 融入 SVM 和 RF 模型

近年来,随着人工智能的发展,机器学习逐渐被引入到滑坡易发性评价中,此类方法能更准确地计算滑坡与评价因子之间的关系,应用效果较好。目前应用较为广泛的主要有支持向量机模型和随机森林模型。支持向量机(SVM)模型最初被用来解决二分类问题,它是在统计学习理论上将结构风险最小化原则应用于分类中,能够适应小样本分类,分类速度快,性能不逊于人工神经网络<sup>[18]</sup>。其主要思想是通过寻找一个分离超平面,使得距离超平面最近的点至超平面距离最大,从而为分类问题提供较好的泛化能力。随机森林(RF)模型是以决策树为基础的学习器,它是基于集成学习思想将多颗决策树模型组合为一个更加稳定的分类模型<sup>[19]</sup>。针对分类问题,每颗决策树相当于一个分类器输出一个分类结果, $N$  棵树会有  $N$  个分类结果,通过投票



方式来决定最终的结果。这两种模型相较于神经网络模型均不需要大量参数,短时间内均可直接计算每个栅格单元的易发性指数值,能够较好地处理非线性关系,避免了人为主观因素的干扰。

单一机器学习模型主要使用影响因子原始数据作为模型的输入样本属性值,存在量纲不统一,易产生评价误差等问题,从而影响易发性评价结果。对此,考虑采用CF模型计算出各因子分级下的CF值,提取样本点各评价因子的CF值作为SVM和RF的输入值,以消除各因子之间量纲不统一以及误差等方面的影响,同时初步判断各因子在不同分级类别下与滑坡之间的关系。基于上述理论,可将CF模型分别与SVM模型和RF模型进行耦合。现有研究表明,CF-SVM模型与CF-RF模型的精度均高于单一SVM模型和单一RF模型<sup>[15-16]</sup>。CF模型是根据历史滑坡点在各因子分级类别下的分布情况,计算各因子分级类别下滑坡发生的概率,CF值计算公式为:

$$CF = \begin{cases} \frac{PP_a - PP_s}{PP_a(1 - PP_a)}, PP_a \geq PP_s \\ \frac{PP_a - PP_s}{PP_s(1 - PP_a)}, PP_a < PP_s \end{cases} \quad (1)$$

式中: $PP_a$ 为评价因子分级 $a$ 中滑坡点个数与该级单元面积的比值; $PP_s$ 为研究区滑坡总个数与研究区总面积的比值。 $CF$ 取值为 $[-1, 1]$ ,当 $CF$ 为正值时,值越接近1表明发生滑坡的可能性越大;当 $CF$ 为负值时,值越接近-1表明发生滑坡的可能性越小;当 $CF=0$ 时,无法确定评价因子在该分级下是否有利于滑坡的发生。

## 2.2 麻雀搜索算法(SSA)

尽管耦合模型具有以上优点,但机器学习模型在训练过程中,参数的选取会直接影响评价结果的准确性,故引入麻雀搜索算法进行参数优化。麻雀搜索算法于2020年由Xue等<sup>[20]</sup>首次提出,是基于麻雀种群的觅食和反捕食行为开发的一种新型智能优化算法,麻雀觅食的过程就是算法寻优的过程,该算法寻优能力强、收敛速度快。在麻雀觅食的过程中,发现者主要负责寻找食物以及提供觅食的区域和方向,具有较好的适应度;加入者会不断地监视发现者,在发现者周围觅食;警戒者主要负责侦察预警,当麻雀种群受到捕食者威胁或者遭遇危险时,将会进行反捕行为。

发现者位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \exp\left(\frac{-i}{\alpha I_{\max}}\right), \text{ if } R_2 < S_T \\ X_{i,j}^t + \mathbf{QL}, \text{ if } R_2 \geq S_T \end{cases} \quad (2)$$

式中: $t$ 为当前迭代次数; $\alpha$ 为 $(0, 1)$ 之间的均匀随机数; $Q$ 为服从正态分布的随机数; $\mathbf{L}$ 为一行多维矩阵,矩阵元素全为1; $I_{\max}$ 为最大迭代次数; $X_{i,j}^{t+1}$ 表示在 $t+1$ 次迭代中第 $i$ 只麻雀在第 $j$ 维中的位置; $R_2$ 和 $S_T$ 为预警值、安全值。

加入者的位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} Q \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right), \text{ if } i < \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \mathbf{A}^+ \mathbf{L}, \text{ if 其他} \end{cases} \quad (3)$$

式中: $X_p$ 为最优加入者的位置; $X_{\text{worst}}$ 为当前全局最差位置; $n$ 为种群数量; $\mathbf{A}$ 为一行多维矩阵,矩阵元素为1或者-1, $\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}$ 。

警戒者位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta |X_{i,j}^t - X_{\text{best}}^t|, \text{ if } f_i > f_g \\ X_{i,j}^t + K \left( \frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_w) + \epsilon} \right), \text{ if } f_i = f_g \end{cases} \quad (4)$$

式中: $X_{\text{best}}$ 为当前最优位置; $\beta$ 为步长控制参数; $f_i$ 为适应度值; $f_g$ 、 $f_w$ 分别为当前最优、最差适应度值; $\epsilon$ 为常数; $K$ 为 $[-1, 1]$ 内随机数。

## 2.3 CF融入SSA优化SVM和RF模型算法流程

为消除SVM模型和RF模型自身存在的各因子之间量纲不统一以及误差等方面的影响,采用CF法计算得到各评价因子的敏感值作为SVM模型和RF模型的样本属性值,构建CF-SVM模型与CF-RF模型。在SVM模型训练过程中,关键参数惩罚因子 $c$ 和径向基核参数 $g$ 的取值会直接影响预测结果的精度,而创建RF模型之前,确定决策树棵数 $n_{\text{tree}}$ 和最大特征数 $m_{\text{try}}$ 也至关重要。故引入SSA对SVM模型与RF模型进行参数优化,利用该算法搜索能力强、收敛速度快、能在短时间内寻找到最优参数组合的特点,优化SVM参数 $c$ 和 $g$ 以及RF参数 $n_{\text{tree}}$ 和 $m_{\text{try}}$ ,以最优参数 $c$ 、 $g$ 、 $n_{\text{tree}}$ 、 $m_{\text{try}}$ 作为SVM和RF的参数,将样本数据集带入SVM模型和RF模型进行训练,构建CF-SSA-SVM模型与CF-SSA-RF模型,然后对整个研究区进行滑坡发生的概率预测,计算区域滑坡易发性指数值,再利用ArcGIS软件绘制滑坡易发性分区图。

CF融入SSA优化SVM和RF模型的实现流程如图2所示。

## 3 评价因子的选取与分析

### 3.1 评价因子选取

滑坡的发生受到多个因素的影响,在滑坡易发



性评价过程中,合适的评价因子至关重要。根据前人经验以及研究区实际情况,本文综合选取了高程、坡度、坡向、距水系距离、距道路距离、工程地质岩组、距断层距离、年均降雨量、植被指数(NDVI)、土地利用类型等评价因子。采用 ArcGIS 提取评价因子图层(图 3),将研究区工程地质岩组以及土地利用类型按照野外调查划分标准进行划分,距水系距离、距道路距离、距断层距离按等间距划分,其余因子按照自然间断法进行划分。

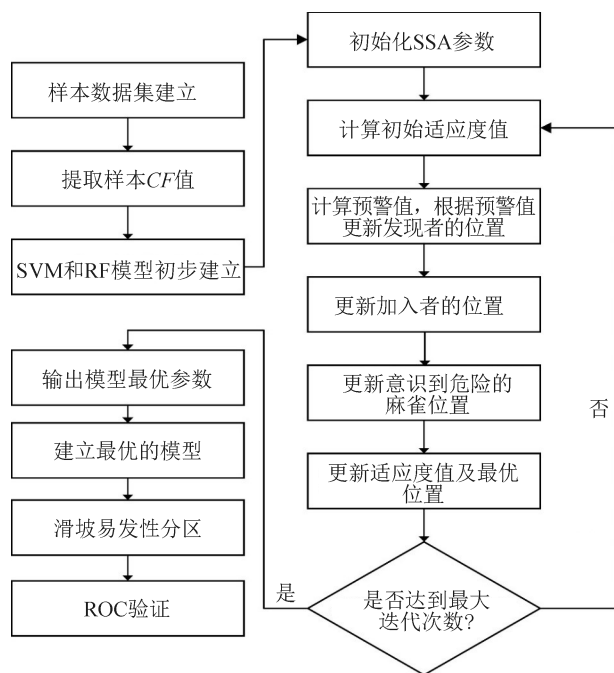


图 2 CF 融入 SSA 优化 SVM 和 RF 模型流程图

Fig. 2 CF into SSA optimization SVM and RF model flow chart

### 3.1.1 地形地貌因子

研究区中地形地貌因子由 DEM 生成。高程是一个重要的影响因子,不同的高程会使得区域人类活动和地形地貌存在明显差异。研究区的高程主要分布于 618~1 069 m,频率比值大于 1,且在 2 100 m 以上没有灾害发生,说明高程低的地方有利于滑坡的产生。坡度是影响滑坡的一个重要因素,地表土质松散发生位移会影响坡度的变化从而导致灾害发生,研究区坡度集中在  $0^{\circ}\sim 22^{\circ}$  时频率比大于 1,说明对灾害的发生会产生较大的影响。不同的坡向也会导致地表蒸发、植被覆盖等差异,将坡向分为 9 类,当坡向为南西、南东方向时,频率比值较大,对滑坡具有正向作用,有利于灾害的发生。剖面曲率反映不同的坡面形态,当剖面曲率大于等于 0 时,频率比值大于 1,有利于灾害的发生。

### 3.1.2 基础地质因子

基础地质因子分为工程地质岩组和断层构造两

个部分。岩土体的风化能力是根据岩石类型、坚硬程度进行划分,岩性对滑坡的分布具有重要的控制作用。黏土质单层土体岩组主要以粉质黏土为主,结构较为松散,不稳定,易发生滑坡。断层构造以 500 m 为步长划分为 6 类,整体趋势为离断层越远,发生滑坡的可能性越小。

### 3.1.3 水文环境因子

降雨会造成水流冲刷,使土质会变得疏松、软化,降低岩土体的强度,研究区年均降雨量在 767~996 mm,主要集中在 846~929 mm,对灾害发生具有较大影响。水系是影响滑坡发育的重要因素,水系分布密度和发育程度越高,对地表的侵蚀能力就越强,该区域水系是由 DEM 提取,以 200 m 为步长进行缓冲,主要集中在 200 m 以内,且离水系越远频率比越小,说明离水系越近,灾害产生的影响越大。植被指数可以反映地表植被的覆盖情况,取值在  $-1\sim 1$  之间,研究区植被指数低于 0.27 时,频率比大于 1,表明在此范围内发生灾害的概率较大。

### 3.1.4 人类工程活动因子

道路开挖或者周边施工引起的扰动都会诱导滑坡的发生,导致斜坡不稳定,将研究区道路按照 500 m 为步长进行缓冲,离道路越远频率比越小,说明离道路越近,越有利于滑坡的发育。土地利用类型不同,对滑坡发育的影响也不同,不合理的土地利用会对滑坡造成较大的影响,将研究区土地利用分为 8 类,其中耕地和灌木频率比较大,对滑坡的发育具有较大的促进作用。

## 3.2 评价因子共线性检验

多重共线性检验是评估各因子之间是否线性相关的一种方法,由于因子之间很容易存在高度共线相关关系,而共线性较强会影响模型的精度,故将评价因子代入 SPSS 26.0 软件进行多重共线性检验,统计 10 个评价因子的容忍度(tolerance, TOL)和方差膨胀因子(variance inflation factor, VIF)<sup>[21]</sup>。当容忍度小于 0.5 或者方差膨胀因子大于 2 时,表明因子之间相关性较强。由表 2 可知,各因子之间容忍度大于 0.5,方差膨胀因子小于 2,说明各个因子之间不存在多重共线性,可以用于模型。

## 4 滑坡易发性评价

### 4.1 评价因子确定性系数值的计算

统计 160 个滑坡点在各因子分级类别下的分布情况,采用式(1)计算出各评价因子分级类别下的确定性系数值。各分级的 CF 值如表 3 所示。

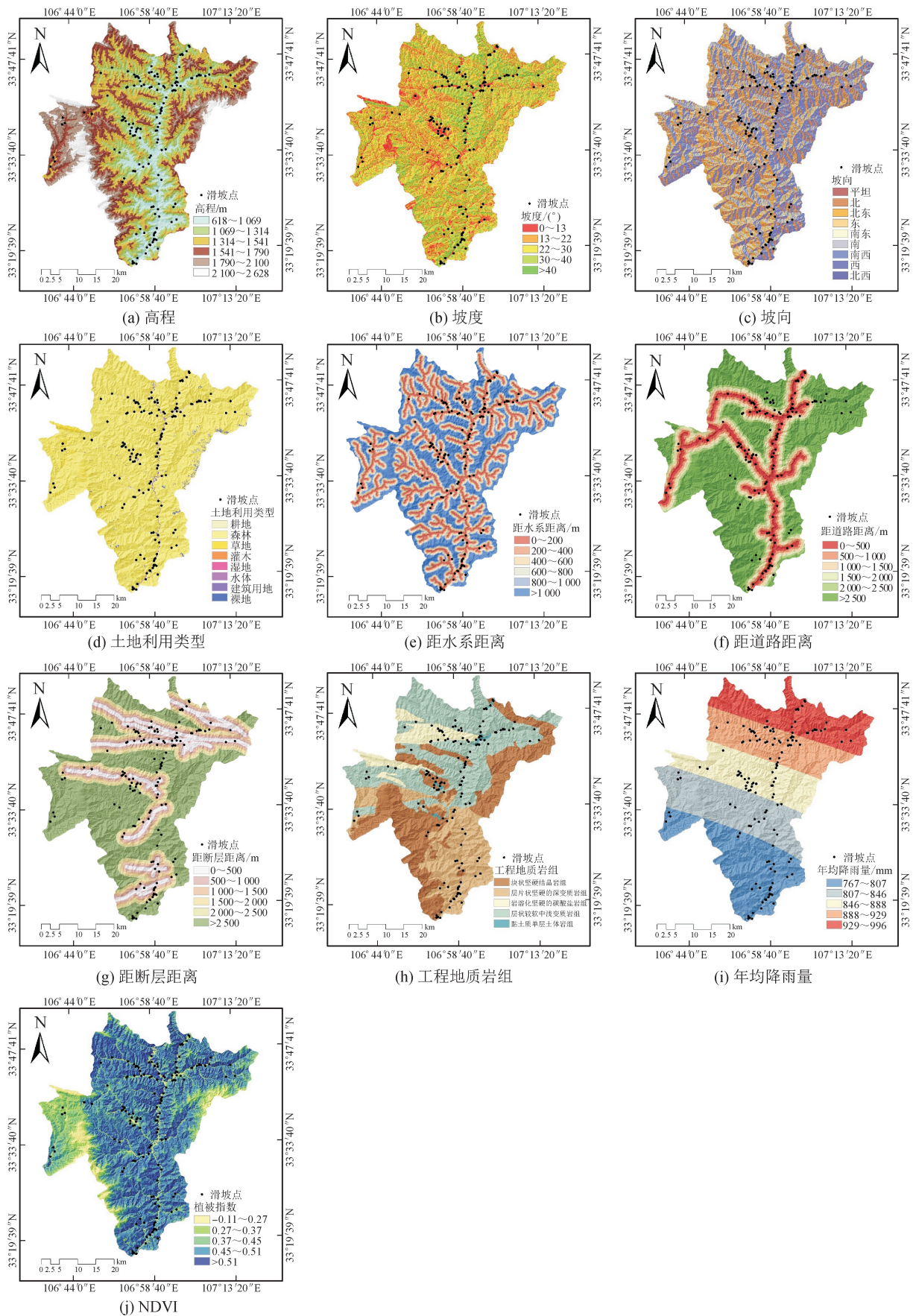


图 3 评价因子图层

Fig. 3 Evaluation factor layer

注:此图基于国家自然资源部标准地图服务网站审图号为 GS(2022)1873 的标准地图制作,底图无修改。

表2 多重共线性检验  
Tab.2 Multiple covariance check

检验指标	高程	坡度	坡向	土地利用类型	距水系距离	距道路距离	距断层距离	NDVI	年均降雨量	工程地质岩组
容忍度	0.544	0.816	0.917	0.691	0.551	0.501	0.804	0.693	0.898	0.766
方差膨胀因子	1.837	1.226	1.090	1.447	1.816	1.996	1.244	1.444	1.114	1.306

表3 评价因子分级  
Tab.3 Impact factor classification

评价因子	分级	频率比值	CF	评价因子	分级	频率比值	CF
高程/m	618~1 069	5.630 49	0.822 456	土地利用类型	耕地	12.954 75	0.922 876
	1 069~1 314	0.935 06	-0.064 940		森林	0.730 77	-0.269 249
	1 314~1 541	0.123 74	-0.876 264		草地	1.902 67	0.474 457
	1 541~1 790	0.090 73	-0.909 272		灌木	10.367 31	0.903 610
	1 790~2 100	0.090 86	-0.909 147		湿地	0.000 00	-1.000 000
	2 100~2 628	0.000 00	-1.000 000		水体	5.353 78	0.813 276
坡度/(°)	0~13	2.276 67	0.068	建筑用地	0.000 00	-1.000 000	
	13~22	1.117 36	0.560 803	裸地	1.668 89	0.400 829	
	22~30	0.823 55	0.105 042	NDVI	-0.11~0.27	3.661 04	0.726 907
	30~40	0.438 67	-0.176 457		0.27~0.37	2.175 17	0.540 306
	>40	0.224 90	-0.561 345		0.37~0.45	1.625 68	0.384 903
坡向	平坦	0.000 00	-1.000 000	0.45~0.51	0.610 52	-0.389 497	
	北	1.019 41	0.019 038	>0.51	0.289 80	-0.710 217	
	北东	0.927 59	-0.072 414	年均降雨量/mm	767~807	0.947 32	-0.052 680
	东	0.919 42	-0.080 586		807~846	0.425 78	-0.574 237
	南东	1.203 96	0.169 421		846~888	1.309 71	0.236 487
	南	0.864 19	-0.135 822		888~929	1.718 87	0.418 254
	南西	1.491 85	0.329 715		929~996	0.764 58	-0.235 429
	西	0.860 33	-0.139 680	工程地质岩组	块状坚硬结晶岩组	0.524 33	-0.475 693
北西	0.713 78	-0.286 234	层片状坚硬的深变质岩组		0.968 49	-0.031 513	
距水系距离/m	0~200	4.566 99	0.781 095		岩溶化坚硬的碳酸盐岩组	1.872 28	0.465 927
	200~400	0.918 11	-0.081 895		层状较软中浅变质岩组	1.064 30	0.060 422
	400~600	0.270 62	-0.729 393	黏土质单层土体岩组	12.228 23	0.918 290	
	600~800	0.203 13	-0.796 883	距断层距离/m	0~500	1.999 73	0.499 970
	800~1 000	0.163 49	-0.836 518		500~1 000	1.706 84	0.414 152
	>1 000	0.040 91	-0.959 097		1 000~1 500	1.930 26	0.481 969
距道路距离/m	0~500	5.680 66	0.824 025		1 500~2 000	1.236 99	0.191 599
	500~1 000	0.571 79	-0.428 229		2 000~2 500	1.029 98	0.029 108
	1 000~1 500	0.973 67	-0.026 328	>2 500	0.448 07	-0.551 949	
	1 500~2 000	0.524 77	-0.475 244				
	2 000~2 500	0.330 27	-0.669 745				
>2 500	0.266 51	-0.733 500					



## 4.2 基于 CF 融入 SSA 优化 SVM 和 RF 模型的易发性评价

本文将采用栅格单元作为滑坡易发性评价的基本单元(栅格单元计算方便、易操作)。将研究区划分为  $30\text{ m} \times 30\text{ m}$  大小的栅格单元,共计 2 166 353 个栅格单元,各评价因子均统一为 30 m 的栅格单元。CF 融入 SSA 优化 SVM 和 RF 模型的易发性评价具体算法如下。

1) 样本数据集的建立。根据留坝县已有的 160 个滑坡点的数据分布以及各因子的分级情况,采用确定性系数模型计算出各评价因子的 CF 值,在 160 个滑坡点的 500 m 缓冲区外随机选取等量的非滑坡点组成样本点,提取样本点的属性值(CF 值)作为 SVM 模型和 RF 模型的输入样本数据集,随机选取滑坡点与非滑坡点的 70% 作为训练样本点,剩余 30% 作为测试样本点。

2) 构建 CF 融入 SSA 优化 SVM 和 RF 模型。首先采用 Matlab 软件读取两种模型的训练样本数

据集,初始化 SSA 的参数,确定 SVM 参数  $c$  和  $g$  以及 RF 参数  $ntree$  和  $mtry$  的取值范围,通过不断更新麻雀位置,寻找最优适应度值,确定最佳麻雀个体的位置;然后根据最佳位置得到 SVM 模型和 RF 模型的最优参数,即  $c=0.940$ 、 $g=0.913$ 、 $ntree=68$ 、 $mtry=7$ ,将最优参数代入 SVM 模型和 RF 模型进行训练,最终构建 CF-SSA-SVM 模型与 CF-SSA-RF 模型。

3) 滑坡易发性分区。通过 ArcGIS 栅格转点工具获取整个研究区的属性值,通过上述训练好的两种模型对整个研究区进行预测,计算得到研究区滑坡易发性指数,采用 ArcGIS 软件中点转栅格工具绘制易发性分区图,将其按照自然间断法分为 5 个易发性等级:极低易发区、低易发区、中易发区、高易发区、极高易发区,如图 4 所示。

4) ROC 验证。采用 ROC 曲线对两种耦合模型的精度进行验证,通过成功率和预测率曲线对比两种耦合模型的评价精度,得出更适用于留坝县滑坡易发性评价的模型。

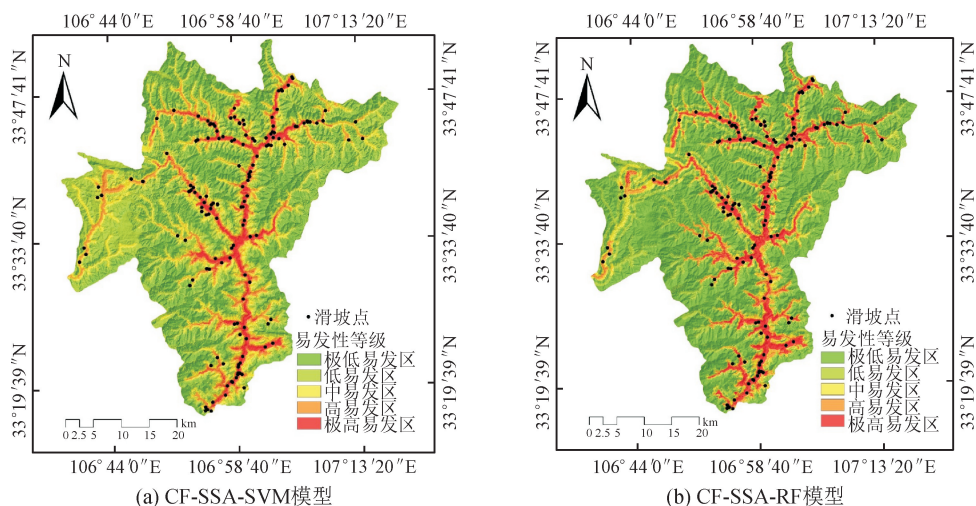


图 4 两种模型易发性分区图

Fig. 4 Susceptibility zoning map of the two models

注:此图基于国家自然资源部标准地图服务网站审图号为 GS(2022)1873 的标准地图制作,底图无修改。

## 5 结果与检验

### 5.1 易发性评价结果分析

1) 从易发性分区统计表得到(见表 4),两种模型评价结果较为一致,其中 CF-SSA-RF 模型的极高易发区包含了 82.5% 的灾害点,极低易发区面积占研究区总面积的 57.9%,相比 CF-SSA-SVM 模型,CF-SSA-RF 模型在极高易发区涵盖了更多的滑坡点,极低易发区面积占比也更大,说明 CF-SSA-RF 模型比 CF-SSA-SVM 模型预测得更加合理。

2) 两种模型均无滑坡点落在极低易发区,但仍

有少量滑坡点落在低易发区,这可能由于滑坡灾害还受到了其他因素的影响,而本文仅考虑了几种主要因素对滑坡灾害的影响,其中灾害点密度随着易发性等级的增大而逐渐增大。

3) 从图 4 中可以看出,极高和高易发区主要集中在研究区中部,即青桥驿镇、马道镇、武关驿镇中部靠近水系和道路较为发育的区域以及以粉质黏土为主的土质较为疏松的区域。在强降雨后,斜坡容易受到雨水的冲刷,使得斜坡土质变得松散,导致斜坡不稳定。极低易发区主要分布在东、西、北部的三面环山区域,因地势较高、人类活动较少,发生滑坡的可能性较小。

表4 易发性分区统计  
Tab.4 Vulnerability zoning statistics

易发性等级	CF-SSA-SVM 模型			CF-SSA-RF 模型		
	分区面积 占比/%	灾害点占 比/%	灾害点密度/ (处·km <sup>-2</sup> )	分区面积 占比/%	灾害点占 比/%	灾害点密度/ (处·km <sup>-2</sup> )
极低易发区	37.18	0.00	0.000 000	57.90	0.00	0.000 000
低易发区	32.57	1.25	0.000 003	17.21	1.88	0.000 008
中易发区	14.98	8.75	0.000 043	7.77	1.25	0.000 012
高易发区	9.97	27.50	0.000 204	8.37	14.38	0.000 127
极高易发区	5.29	62.50	0.000 873	8.76	82.50	0.000 696

## 5.2 精度检验

成功率和预测率曲线是评价滑坡易发性模型精度较常用的方法,通常用 ROC 曲线下面积 AUC 值来进行精度验证<sup>[22]</sup>。ROC 曲线表示拟合数据与真实数据之间的关系,横轴是假阳性率(1-特异性),表示非灾害点被预测正确的概率;纵轴是真阳性率(敏感度),表示灾害点被正确预测的概率。ROC 曲线越靠近左上角,表示模型预测效果越好,ROC 曲线

下面积 AUC 值的范围为[0.5,1],可以作为衡量模型准确率的指标,其值越接近于 1,表示该模型精度越高。由图 5 可知,CF-SSA-RF 模型、CF-SSA-SVM 模型的成功率曲线的 AUC 值分别为 0.994、0.971,预测率曲线的 AUC 值分别为 0.940、0.913,CF-SSA-RF 模型的成功率和预测率都高于 CF-SSA-SVM 模型,说明 CF-SSA-RF 模型具有更好的预测性能。

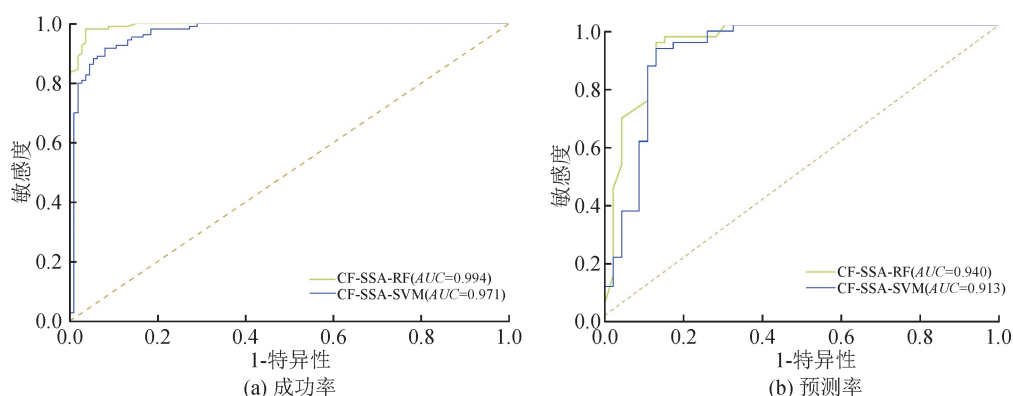


图5 ROC 曲线图  
Fig.5 ROC graph

## 5.3 典型滑坡验证

选择研究区典型滑坡对模型进行验证。根据已有的资料,老厂后山梁滑坡(图 6(a)),位于留侯镇庙台子村大坝沟组,海拔 1 458 m,规模 4.8 万 m<sup>3</sup>,受威胁群众 7 户 24 人,潜在威胁财产 90 万元。滑体为碎石土,滑面呈弧形。初次滑动时上部可见长约 100 m 的拉张裂缝,下伏基岩为千枚岩,产状 330°∠40°。地层岩性为志留系千枚岩,节理裂隙发育,风化强烈,坡体滑坡可能性较大,易发性等级为极高易发区。

郭家河坝滑坡(图 6(b)),位于武关驿镇五里铺村邓家沟组,海拔 742 m,规模 7.68 万 m<sup>3</sup>,受威胁群众 4 户 11 人,潜在威胁财产 70 万元。滑体滑动

方向 280°,滑面为半圆形,斜坡出现拉张裂缝三条,下伏基岩为片麻岩。前滑坡处于蠕变阶段,受降雨和河流冲刷等因素影响,可能进一步出现较大规模的滑动变形,稳定性差、危险性高,可能诱发滑坡灾害。

老房子滑坡(图 6(c)),位于留侯镇火烧关村,海拔 1 474 m,规模 7.68 万 m<sup>3</sup>,受威胁群众 13 户 43 人,潜在威胁财产 180 万元。地层岩性为泥盆系片岩,节理裂隙发育,风化强烈,滑体为碎石土,滑动方向 275°,斜坡体具一拉张裂缝,延伸 60 m,下伏基岩为片岩,产状 275°∠30°,斜坡稳定性较差,危险性高,属于极高易发区。

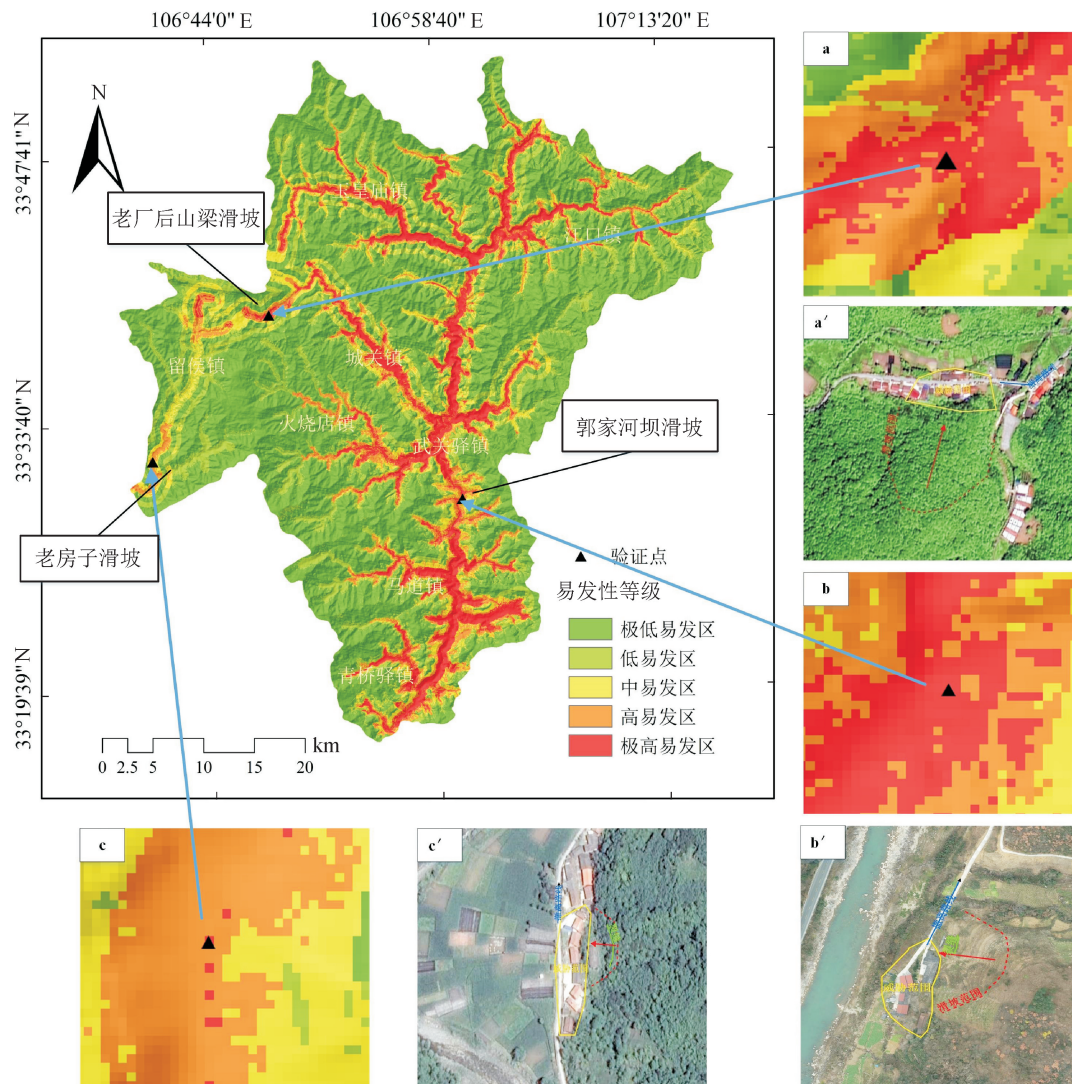


图 6 典型滑坡验证

Fig. 6 Typical landslide verification

注:此图基于国家自然资源部标准地图服务网站审图号为 GS(2022)1873 的标准地图制作,底图无修改。

### 6 结 论

本文以陕西省汉中市留坝县为研究区,分别建立了 CF-SSA-RF 模型与 CF-SSA-SVM 模型,并对该区进行了滑坡易发性评价。

1) CF-SSA-RF 模型相比 CF-SSA-SVM 模型在极高易发区涵盖了更多的滑坡点,极低易发区面积占比更大,且利用三处典型滑坡验证点对 CF-SSA-RF 模型进行了验证,滑坡点均落在极高易发区,评价结果较为合理。

2) 两种模型的滑坡易发性分区图与历史滑坡点的分布较为吻合,极高和高易发区主要集中分布在研究区中南部,即水系、道路较为发育的区域,主要沿着水系道路呈树枝状分布;极低易发区主要分布在东、西、北部的三面环山区域,由于地势较高、人

类活动较少,该区域无滑坡发生。

3) 将 CF 模型融入 SSA 优化 SVM 和 RF 模型能有效解决样本数据输入以及模型参数赋值问题,两种模型均具有较好的预测性能,评价结果基本保持一致,其中 CF-SSA-RF 模型的预测性能(AUC=0.940)略高于 CF-SSA-SVM 模型(AUC=0.913),评价结果更为合理,更适用于留坝县的滑坡易发性评价研究。

#### 参考文献:

[1] 张俊,殷坤龙,王佳佳,等. 三峡库区万州区滑坡灾害易发性评价研究[J]. 岩石力学与工程学报, 2016, 35 (2): 284-296.  
 ZHANG Jun, YIN Kunlong, WANG Jiajia, et al. Evaluation of landslide susceptibility for Wanzhou district of Three Gorges Reservoir[J]. Chinese Journal of Rock Mechanics and Engineering, 2016, 35 (2): 284-296.



- [2] 林荣福,刘纪平,徐胜华,等. 随机森林赋权信息量的滑坡易发性评价方法[J]. 测绘科学, 2020, 45(12): 131-138.  
LIN Rongfu, LIU Jiping, XU Shenghua, et al. Evaluation method of landslide susceptibility based on random forest weighted information[J]. Science of Surveying and Mapping, 2020, 45(12): 131-138.
- [3] 牛全福. 基于GIS的地质灾害风险评估方法研究——以“4.14”玉树地震为例[D]. 兰州:兰州大学, 2011.  
NIU Quanfu. Study on the method of geological disaster risk assessment based on GIS[D]. Lanzhou: Lanzhou University, 2011.
- [4] 王哲,易发成. 基于层次分析法的绵阳市地质灾害易发性评价[J]. 自然灾害学报, 2009, 18(1): 14-23.  
WANG Zhe, YI Facheng. AHP-based evaluation of occurrence easiness of geological disasters in Mianyang City[J]. Journal of Natural Disasters, 2009, 18(1): 14-23.
- [5] 李文彦,王喜乐. 频率比与信息量模型在黄土沟壑区滑坡易发性评价中的应用与比较[J]. 自然灾害学报, 2020, 29(4): 213-220.  
LI Wenyang, WANG Xile. Application and comparison of frequency ratio and information value model for evaluating landslide susceptibility of loess gully region [J]. Journal of Natural Disasters, 2020, 29(4): 213-220.
- [6] CHEN W, LI W, HOU E, et al. Landslide susceptibility mapping based on GIS and information value model for the Chencang District of Baoji, China[J]. Arabian Journal of Geosciences, 2014, 7(11): 4499-4511.
- [7] 魏江波,赵洲. 基于加权确定性系数法的地质灾害易发性分析[J]. 煤田地质与勘探, 2018, 46(6): 108-114.  
WEI Jiangbo, ZHAO Zhou. Analysis of geological hazard susceptibility based on the weighted certainty factor method[J]. Coal Geology & Exploration, 2018, 46(6): 108-114.
- [8] 胡燕,李德营,孟颂颂,等. 基于证据权法的巴东县城滑坡灾害易发性评价[J]. 地质科技通报, 2020, 39(3): 187-194.  
HU Yan, LI Deying, MENG Songsong, et al. Landslide susceptibility evaluation in Badong County based on weights of evidence method[J]. Geological Science and Technology Information, 2020, 39(3): 187-194.
- [9] 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): 367-372.  
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): 367-372.
- [10] SAHANA M, SAJJAD H. Evaluating effectiveness of frequency ratio, fuzzy logic and logistic regression models in assessing landslide susceptibility: a case from Rudraprayag District, India[J]. Journal of Mountain Science, 2017, 14(11): 2150-2172.
- [11] 杨康,薛喜成,李识博. 信息量融入GA优化SVM模型下的地质灾害易发性评价[J]. 安全与环境工程, 2022, 29(3): 109-118.  
YANG Kang, XUE Xicheng, LI Shibo. Geological hazard susceptibility assessment by incorporating information value into GA optimized SVM model[J]. Safety and Environmental Engineering, 2022, 29(3): 109-118.
- [12] 杨硕,李德营,严亮轩,等. 基于随机森林模型的乌江高陡岸坡滑坡地质灾害易发性评价[J]. 安全与环境工程, 2021, 28(4): 131-138.  
YANG Shuo, LI Deying, YAN Liangxuan, et al. Landslide susceptibility assessment in high and steep bank slopes along Wujiang River based on random forest model[J]. Safety and Environmental Engineering, 2021, 28(4): 131-138.
- [13] 武雪玲,杨经宇,牛瑞卿. 一种结合SMOTE和卷积神经网络的滑坡易发性评价方法[J]. 武汉大学学报:信息科学版, 2020, 45(8): 1223-1232.  
WU Xueling, YANG Jingyu, NIU Ruiqing. A landslide susceptibility assessment method using SMOTE and convolutional neural network[J]. Geomatics and Information Science of Wuhan University, 2020, 45(8): 1223-1232.
- [14] 张钟远,邓明国,徐世光,等. 镇康县滑坡易发性评价模型对比研究[J]. 岩石力学与工程学报, 2022, 41(1): 157-171.  
ZHANG Zhongyuan, DENG Mingguo, XU Shiguang, et al. Comparison of landslide susceptibility assessment models in Zhenkang County, Yunnan Province, China[J]. Chinese Journal of Rock Mechanics and Engineering, 2022, 41(1): 157-171.
- [15] 李远远,梅红波,任晓杰,等. 基于确定性系数和支持向量机的地质灾害易发性评价[J]. 地球信息科学学报, 2018, 20(12): 1699-1709.  
LI Yuanyuan, MEI Hongbo, REN Xiaojie, et al. Geological disaster susceptibility evaluation based on certainty factor and support vector machine[J]. Journal of Geo-Information Science, 2018, 20(12): 1699-1709.
- [16] 郑迎凯,陈建国,王成彬. 确定性系数与随机森林模型在云南芒市滑坡易发性评价中的应用[J]. 地质科技通报, 2020, 39(6): 131-144.  
ZHENG Yingkai, CHEN Jianguo, WANG Chengbin, et al. Application of certainty factor and random forest model in landslide susceptibility evaluation in Mangshi City, Yunnan Province [J]. Bulletin of Geological Science and Technology, 2020, 39(6): 131-144.
- [17] 李坤,赵俊三,林伊琳,等. 基于RF和SVM模型的东川泥石流易发性评价研究[J]. 云南大学学报:自然科学版, 2022, 44(1): 107-115.  
LI Kun, ZHAO Junsan, LIN Yilin, et al. Assessment of debris flow susceptibility in Dongchuan based on RF and SVM models[J]. Journal of Yunnan University: Natural Sciences Edition, 2022, 44(1): 107-115.
- [18] 傅文杰. GIS支持下基于支持向量机的滑坡危险性评价[J]. 地理科学, 2008, 28(6): 838-841.  
FU Wenjie. Landslide hazard evaluation based on GIS and SVM[J]. Scientia Geographica Sinica, 2008, 28(6): 838-841.

- [9] 李宁. 复合岩体穿层锚杆锚固力学机理及应用[D]. 徐州: 中国矿业大学, 2021.  
LI Ning. Study on anchorage mechanical mechanism and application of rock bolt in composite rock mass[D]. Xuzhou: China University of Mining and Technology, 2021.
- [10] 文竞舟. 隧道初期支护力学分析及参数优化研究[D]. 重庆: 重庆大学, 2012.  
WEN Jingzhou. Study on mechanical analysis of tunnel initial support and its parameters optimization [D]. Chongqing: College of Civil Engineering of Chongqing University, 2012.
- [11] 陶文斌, 吴平平, 陈铁林, 等. 基于锚杆拉拔试验优化锚固承载特性研究[J]. 煤炭科学技术, 2022, 50(9): 10-19.  
TAO Wenbin, WU Pingping, CHEN Tielin, et al. Experimental research on optimization of anchorage bearing characteristics based on bolt pull-out test. [J]. Coal Science and Technology, 2022, 50(9): 10-19.
- [12] 陈峰. 恒阻大变形锚杆支护机理数值试验研究[D]. 大连: 大连理工大学, 2020.  
CHEN Feng. Numerical experimental study on support mechanism of constant resistance and large deformation bolt [D]. Dalian: Dalian University of Technology, 2020.
- [13] 吴学震, 王刚, 蒋宇静, 等. 拉压耦合大变形锚杆作用机理及其试验研究[J]. 岩土工程学报, 2015, 37(1): 139-147.  
WU Xuezheng, WANG Gang, JIANG Yujing, et al. Mechanism of CTC-yield bolts and its experimental research[J]. Chinese Journal of Geotechnical Engineering, 2015, 37(1): 139-147.
- [14] ZHU Xunguo, CHEN Zhuoli, REN Yan. Numerical simulation study on the anchorage mechanism of yield supporting in deep tunnel[J]. Geotechnical and Geological Engineering, 2019, 37(3): 2091-2102.
- [15] XIA Qingyong, ZHU Xunguo, ZHANG Guofeng, et al. The improved theory of synergetic action between anchor support system and surrounding rock-I L 4S mechanism theory and its application in tunnel support engineering[J]. Geotechnical and Geological Engineering, 2021, 39(5): 3563-3572.
- [16] ZHU Xunguo, YANG Shuai, XIA Hongchun, et al. Joint support technology and its engineering application to deep soft rock tunnel with strong creep [J]. Geotechnical and Geological Engineering, 2020, 38(4): 3403-3414.
- [17] YANG Shuai, ZHU Xunguo, ZHANG Guofeng, et al. Research on the mechanics performance of the new tension-compression rock bolt through numerical simulation[J]. Geotechnical and Geological Engineering, 2022, 40(4): 2255-2266.
- [18] 黄炎杰. 基于双曲线模型的锚杆受力分析方法及其工程应用研究[D]. 长沙: 湖南大学, 2017.  
HUANG Yanjie. Analysis method of bolt force based on hyperbolic model and its engineering application [D]. Changsha: Hunan University, 2017.
- [19] 谭忠盛, 杨旻, 陈伟, 等. 中老铁路高地应力软岩隧道大变形控制技术[J]. 铁道学报, 2020, 42(12): 171-178.  
TAN Zhongsheng, YANG Yang, CHEN Wei, et al. Large deformation control technology of high geostress soft rock tunnel of China-Loas railway[J]. Journal of the China Railway Society, 2020, 42(12): 171-178.

(责任编辑 王绪迪)

(上接第131页)

- [19] 马啸, 王念秦, 李晓抗, 等. 基于 RF-FR 模型的滑坡易发性评价——以略阳县为例[J]. 西北地质, 2022, 55(3): 335-344.  
MA Xiao, WANG Nianqin, LI Xiaokang, et al. Assessment of landslide susceptibility based on RF-FR model: taking Lueyang County as an example [J]. Northwestern Geology, 2022, 55(3): 335-344.
- [20] XUE J, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm [J]. Systems Science & Control Engineering an Open Access Journal, 2020, 8(1): 22-34.
- [21] CHEN W, CHEN Y Z, TSANGARATOS P, et al. Combining evolutionary algorithms and machine learning models in landslide susceptibility assessments[J]. Remote Sensing, 2020, 12(23): 3854.
- [22] LEE S T, YU T T, PENG W F, et al. Incorporating the effects of topographic amplification in the analysis of earthquake-induced landslide hazards using logistic regression[J]. Natural Hazards and Earth System Sciences, 2010, 10(12): 2475-2488.

(责任编辑 周 蓓)