

DOI:10.11705/j.issn.1672-643X.2021.01.30

# IOE模型在延安市宝塔区碾庄沟流域 滑坡易发性分区中的应用

张庭瑜<sup>1,2,3,4</sup>, 孙增慧<sup>1,2,3,4</sup>, 程杰<sup>1,2,3,4</sup>, 刘金宝<sup>1,2,3,4</sup>,  
石磊<sup>1,2,3,4</sup>, 孔辉<sup>1,2,3,4</sup>, 杨亮彦<sup>1,2,3,4</sup>, 罗丹<sup>5</sup>

(1. 陕西地建土地工程技术研究院有限责任公司, 陕西 西安 710075; 2. 陕西省土地工程建设集团有限责任公司, 陕西 西安 710075; 3. 自然资源部退化及未利用土地整治工程重点实验室, 陕西 西安 710075; 4. 陕西省土地整治工程技术研究中心, 陕西 西安 710075; 5. 陕西地建土地勘测规划设计研究院有限责任公司, 陕西 西安 710075)

**摘要:** 将延安市宝塔区碾庄沟流域作为研究区, 在野外调查以及遥感解译的基础上, 得到了73个滑坡点数据, 其中70%的滑坡点被当作训练样本, 剩余的30%的滑坡点被当作测试样本。选取坡度、坡向、高程、归一化植被指数(*NDVI*)、岩土体类型、土地利用类型、平面曲率和剖面曲率作为滑坡易发性分区建模的解释变量。利用熵指数模型(*IOE*)计算研究区的滑坡易发性指数(*LSI*), 得到了研究区滑坡易发性分区图(*LSM*)。最后利用准确率和接受者操作特征曲线下的面积(*AUC*)对分区结果进行评价。结果表明: 训练样本集和测试样本集的准确率均大于0.8, 且测试样本集的*AUC*值为0.964 1, 说明研究区的滑坡易发性分区结果可信度高, 且*IOE*模型具有较强的泛化能力。研究结果也可以为当地的滑坡防治工作提供参考。

**关键词:** 滑坡; 易发性分区; 熵指数模型(*IOE*); 黄土沟壑区; 碾庄沟流域; 延安市

中图分类号: P642.22

文献标识码: A

文章编号: 1672-643X(2021)01-0205-08

## Application of IOE model for landslide susceptibility mapping in Nianzhuangou Watershed, Yan'an City

ZHANG Tingyu<sup>1,2,3,4</sup>, SUN Zenghui<sup>1,2,3,4</sup>, CHENG Jie<sup>1,2,3,4</sup>, LIU Jinbao<sup>1,2,3,4</sup>,  
SHI Lei<sup>1,2,3,4</sup>, KONG Hui<sup>1,2,3,4</sup>, YANG Liangyan<sup>1,2,3,4</sup>, LUO Dan<sup>5</sup>

(1. Institute of Land Engineering and Technology, Shaanxi Provincial Land Engineering Construction Group Co., Ltd., Xi'an 710075, China; 2. Shaanxi Provincial Land Engineering Construction Group Co., Ltd., Xi'an 710075, China; 3. Key Laboratory of Degraded and Unused Land Consolidation Engineering, the Ministry of Natural Resources, Xi'an 710075, China; 4. Shaanxi Provincial Land Consolidation Engineering Technology Research Center, Xi'an 710075, China; 5. Shaanxi Construction, Land Survey, Planning, Design and Research Institute Co., Ltd., Xi'an 710075, China)

**Abstract:** The Nianzhuangou watershed of Yan'an City was regarded as the study area. Based on the field survey and remote sensing image, the data of 73 landslide points were obtained. 70% of the points were used as training dataset, and the remaining were used for validation purpose. Slope, aspect, elevation, normalized difference vegetation index (*NDVI*), lithology, land use, plan curvature and profile curvature were selected as explanatory variables for landslide susceptibility modeling. The landslide susceptibility index (*LSI*) of the study area was calculated using the index of entropy model (*IOE*), and a landslide susceptibility map (*LSM*) was generated. Finally, the results were evaluated using the accuracy and area under the receiver operating characteristic curve (*AUC*). The results show that the accuracy of the training dataset and the validation dataset are both greater than 0.8, and the *AUC* value of validation dataset is 0.964 1, indicating that the *LSM* is highly reliable, and the *IOE* model has strong generalization ability. The results can also provide some reference for the local prevention and control of landslides.

收稿日期: 2019-05-19; 修回日期: 2020-04-23

基金项目: 陕西省自然科学基金基础研究计划项目(2019JQ-945); 陕西省土地工程建设集团有限责任公司内部科研项目(DJNY2021-10); 长安大学中央高校基本科研业务费专项资金项目(300102351502)

作者简介: 张庭瑜(1990-), 男, 陕西汉中, 博士, 工程师, 研究方向为地质灾害与土地工程。

通讯作者: 孙增慧(1987-), 男, 山西忻州, 博士, 高级工程师, 研究方向为地质灾害与土地工程。

**Key words:** landslides; susceptibility mapping; index of entropy model (IOE); loess gully area; Nianzhuangou Watershed; Yan'an City

## 1 研究背景

自21世纪以来,随着人口和人类活动空间的不断增长,滑坡灾害爆发的频率也随之增高。据不完全统计,全世界每年因滑坡导致的人员伤亡约10 000人,经济损失约 $40 \times 10^8$  美元<sup>[1]</sup>。滑坡防治的第一步就是预测容易发生滑坡的范围,随着GIS技术在近年来的不断发展,许多统计学理论的分类模型被引入到滑坡易发性分区当中来,如信息量模型(SI)<sup>[2]</sup>、证据权模型(WOE)<sup>[3]</sup>、确定性系数模型(CF)<sup>[4]</sup>、熵指数模型(IOE)<sup>[5]</sup>、层次分析法(AHP)<sup>[6]</sup>、逻辑回归模型(LR)<sup>[7-8]</sup>和频率密度模型(FR)<sup>[9]</sup>等。随着区域滑坡易发性分区研究的逐渐深入,一些机器学习算法也被应用在滑坡预测领域,如决策树模型(DT)<sup>[10]</sup>、随机森林算法(RF)<sup>[11]</sup>、神经网络模型(ANN)<sup>[12-13]</sup>以及支持向量机模型(SVM)<sup>[14]</sup>等。由于这些分类模型的引入,使得区域滑坡易发性分区的结果变得较为可信,同时也使得滑坡易发性分区的方法变得更加多元化。

水土流失现象在中国西北部的黄土丘陵沟壑区十分严重,并且滑坡灾害频发。陕西省延安市宝塔区位于黄土丘陵沟壑区的核心区域,高陡边坡分布广泛,加之人口较为集中,所以在这一地区滑坡灾害对人民的安全威胁巨大。因此,将延安市宝塔区碾庄沟流域作为本文的研究区,通过构建IOE(index of entropy)模型对研究区进行滑坡易发性分区,所得到的结果可以为当地的滑坡防治工作提供对策和意见。

## 2 研究区概况与数据来源

### 2.1 研究区概况

研究区位于延安市宝塔区东北部,地理坐标为东经 $109^{\circ}28'30'' \sim 109^{\circ}35'00''$ ,北纬 $36^{\circ}38'00'' \sim 36^{\circ}43'30''$ 。研究区覆盖面积 $96 \text{ km}^2$ ,平均高程 $1100 \text{ m}$ ,相对高差 $24.58 \text{ m}$ 。半湿润半干旱大陆性季风气候是研究区的主要气候类型,年平均气温 $7^{\circ}\text{C}$ ,且干旱少雨,年平均降雨量约 $360 \text{ mm}$ ,降雨时段多集中在8、9月之间,且暴雨是主要的降雨形式。

研究区沟壑纵横,冲沟发育密集,所以短时暴雨会引发强烈的地表径流,冲刷坡脚,时常引发滑坡灾害。研究区内发育的地层主要为中生代-新生代地层,地表主要由全新统黄土层和侏罗系砂泥岩互层覆盖,其中黄土层厚度最高可达 $20 \text{ m}$ ,且黄土层被雨水浸润后结构会变得松散,具有一定的湿陷性,这

也是导致滑坡灾害发生的主要因素之一。

此外,地壳的间歇性抬升是新构造运动在研究区的主要表现,但稳定的沉积盆地的特征始终被保持着,所以构造作用对研究区地貌的改造没有显著的痕迹。《陕西省区域地质志》<sup>[15]</sup>显示区内无隐伏断层发育,地壳变形速率介于 $1 \sim 2 \text{ mm/a}$ 之间,在历史上发生过的地震震级均不超过4级<sup>[15]</sup>,所以本文所使用的滑坡数据为降雨型滑坡数据,而不涉及地震型滑坡。

### 2.2 数据来源

数据的选择在滑坡易发性分区建模过程中是十分重要的一个环节。本文所用到的数据源有:(1) $8 \text{ m} \times 8 \text{ m}$ 分辨率的GF-1多光谱遥感影像(传感器:PMS1;轨道圈号:27978;云量:6%;采集日期:2018-07-05);(2) $30 \text{ m} \times 30 \text{ m}$ 分辨率的AsterGDEM数据;(3)基于野外调查以及遥感解译的73个滑坡点数据。

### 2.3 滑坡编录

要建立滑坡易发性分区模型,首先需要进行滑坡编录,通过对研究区进行滑坡编录从而生成滑坡编录图(landslide inventory map)。滑坡编录图主要包含滑坡的位置信息、滑坡的规模和平面面积以及研究区的高程数据等<sup>[16]</sup>。为了提升后续计算的效率,本文利用质心法将73个滑坡图斑转换为73个滑坡点,结合DEM制作了研究区的滑坡编录图,如图1所示。

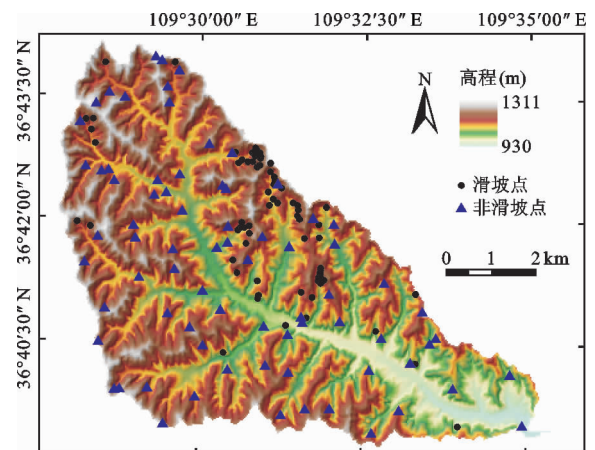


图1 碾庄沟流域滑坡编录图

为了检测分类模型的泛化能力和结果的精确程度,在滑坡编录图幅范围内又随机生成了等量的73个非滑坡点,同时从滑坡点和非滑坡点数据中按照70%的比例,分别选取51个滑坡点和51个非滑坡点用于建立IOE模型,剩余的22个滑坡点和22个

非滑坡点用于测试模型的精度和泛化能力。

### 2.4 滑坡影响因素

通过分析文献资料,并结合野外实际调查,本文选取了坡度、坡向、高程、归一化植被指数(normalized difference vegetation index, *NDVI*)、岩土体类型、土地利用类型、平面曲率和剖面曲率共 8 个滑坡影响因素作为滑坡易发性分区建模的解释变量,研究区各滑坡影响因素的相关数据分布见图 2。其中坡度数据、坡向数据、高程数据、平面曲率和剖面曲率数据基于数字高程模型(digital elevation model,

DEM)提取(图 2(a)–2(e)),并利用自然间断点法(natural break method)分别将这 5 种解释变量分为 5 级;*NDVI* 数据通过 GF-1 遥感影像提取(图 2(f)),同样利用自然间断点法将其分为 5 级;基于 1:200 000 地质图提取岩土体类型数据(图 2(g)),并将研究区的岩土体类型分为黄土、砂岩、泥岩和砂泥岩互层;利用 ENVI 软件中的支持向量机分类方法,基于 GF-1 遥感影像将研究区的土地利用类型分为建设用地、耕地、林草地、水域和裸地,从而获得土地利用类型数据(图 2(h))。

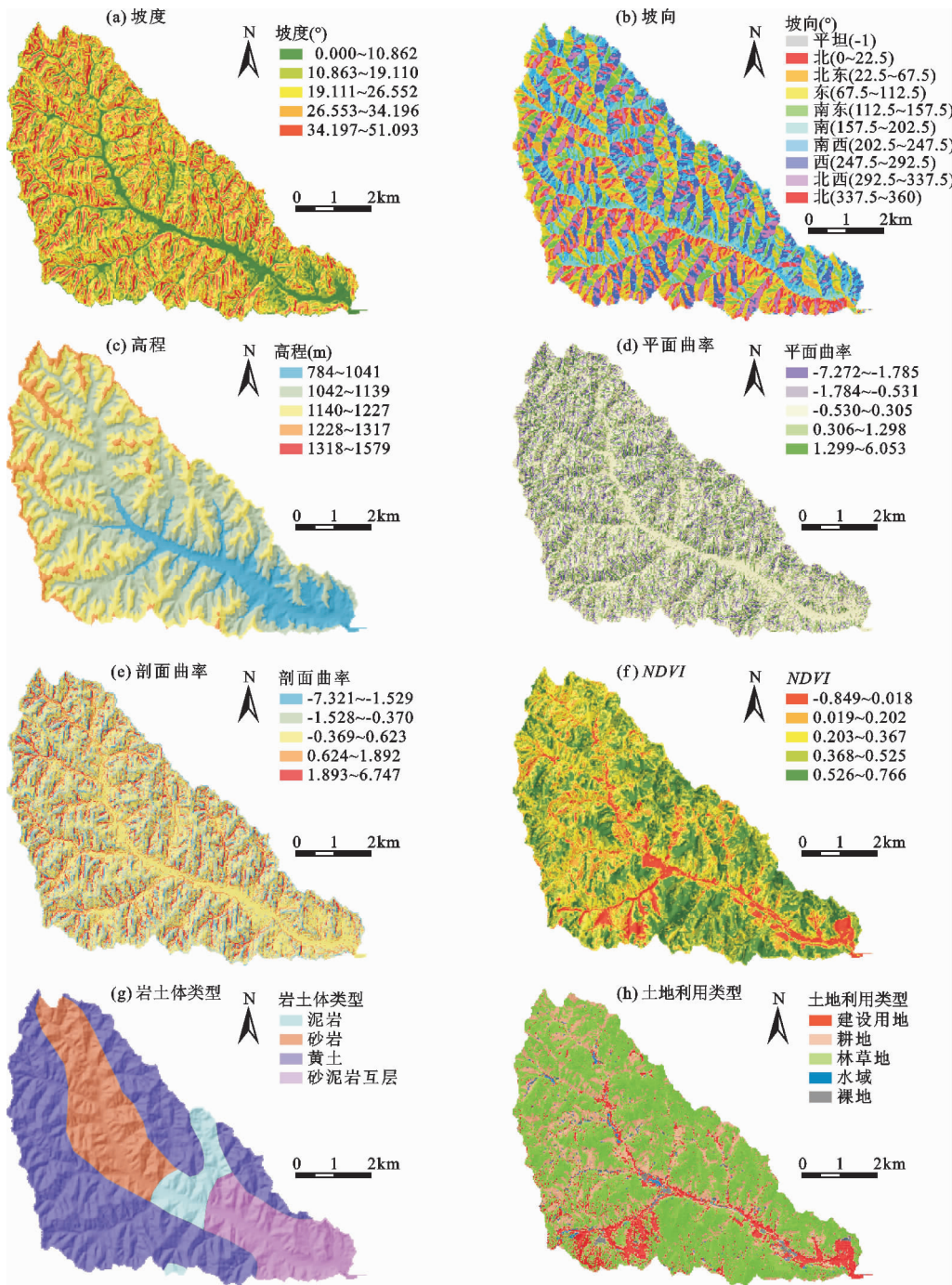


图 2 研究区各滑坡影响因素的相关数据分布

### 3 研究方法

#### 3.1 影响因子优选

滑坡易发性分区建立模型的前提是假设每一类解释变量均为相互独立,如果有强烈相关性存在于解释变量之间,那么过拟合的现象可能会出现在结果当中,导致精度下降。为了防止过拟合现象的发生,提高分类模型的泛化能力和精度,需要对分类模型的输入数据进行多重共线性检测<sup>[17]</sup>。一般来讲,检测多解释变量之间潜在的多重共线性的方法有相关性分析法、方差膨胀系数(variance inflation factor, *VIF*)检测法和条件系数检验法3种。本研究在对8种解释变量进行线性回归建模的基础上,利用方差膨胀系数检测法来检测解释变量之间可能存在的多重共线性问题。在线性回归分析时,每一个解释变量都会得到一个膨胀系数,一个对应的容差(tolerance, *TOL*)和一个Pearson系数,当 $VIF > 10$ ,  $TOL < 0.1$ 且Pearson系数大于0.7时,则认为该解释变量会对其他的变量造成干扰<sup>[18]</sup>,因此,需要将其排除,不参与后续的建模。

#### 3.2 IOE 模型

衡量事物混乱程度的指标可以用熵来表示,其具有平均信息量的含义。对于滑坡的空间分布而言,熵可以表征滑坡在一定区域内的分布状态,熵值越大,表示滑坡的空间分布越复杂,反之,熵值越小,表示滑坡的空间分布越简单<sup>[19]</sup>。此外,每一种解释变量对滑坡的影响程度同样可以用熵来体现。因此本文以熵为基础,通过建立熵指数模型对研究区进行滑坡易发性分区,该模型的计算过程如下:

$$FR_{ij} = \frac{y_{ij}}{x_{ij}} \quad (1)$$

$$S_{ij} = \frac{FR_{ij}}{\sum_{j=1}^{N_j} FR_{ij}} \quad (2)$$

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

$$M_{j \max} = \log_2 N_j \quad (4)$$

$$I_j = \frac{M_{j \max} - M_j}{M_{j \max}} \quad (5)$$

$$W_j = I_j \cdot FR_{ij} \quad (6)$$

式中: $i$ 和 $j$ 分别为解释变量矩阵中的行和列; $x$ 和 $y$ 分别为解释变量分级所占的比例和解释变量每一级中滑坡点所占的比例; $FR_{ij}$ 为研究区滑坡的频率密度; $S_{ij}$ 为滑坡发生可能性的密度; $M_j$ 和 $M_{j \max}$ 分别

为熵和熵的最大值; $N_j$ 为每一种解释变量所分的等级数; $W_j$ 为每一种解释变量的综合得分,利用公式(7)可以得到研究区的滑坡易发性指数(landslide susceptibility index, *LSI*),最终在ArcGIS软件中将*LSI*可视化得到滑坡易发性分区图(landslide susceptibility map, *LSM*)。

$$LSI_{IOE} = \sum_{j=1}^n W_j \quad (7)$$

#### 3.3 精度检测

由于影响滑坡发生的因素十分多样,且分区结果的准确性也未知,因此,有必要利用一种客观的评价方法对分区结果的准确性和模型的泛化能力进行评价。常见的评价指标有精确率(*Precision*)、召回率(*Recall*)和准确率(*Accuracy*)等,而在滑坡预测领域,准确率是评价分区结果准确性的最常用的指标<sup>[20]</sup>。同时,在医学领域广泛使用的接受者操作特征曲线(ROC)也经常用于评价滑坡易发性分区结果,所得到的结果较为客观。鉴于此,本文利用准确率和ROC曲线对分区结果进行评价。

在滑坡易发性分区结果中,用*TP*和*TN*分别表示分类正确的滑坡点和非滑坡点,用*FP*和*FN*分别表示分类错误的滑坡点和非滑坡点,则准确率的计算方法如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

准确率的取值范围为0~1,越接近于1,表明分类的准确率越高。

本文分别以灵敏度(*Sensitivity*)和1-特异度(*Specificity*)作为纵坐标和横坐标绘制ROC曲线。利用ROC曲线下的面积(*AUC*)来定量评价模型分类效果,*AUC*的取值范围为0~1,如果*AUC*=0,说明模型对滑坡没有分类能力,如果*AUC*介于0.5~0.7之间,表明模型分类能力较弱,如果*AUC*介于0.7~0.9之间,说明模型分类能力较强,当*AUC*的值介于0.9~1之间时,表明模型分类能力很强,所得出的结果有着很强的可信度<sup>[21]</sup>。*AUC*的计算方法如下:

$$AUC = \frac{(\sum TP + \sum TN)}{P + N} \quad (9)$$

式中: $P$ 为滑坡点的数量; $N$ 为非滑坡点的数量。

## 4 结果与分析

#### 4.1 影响因子优选结果

本文利用*VIF*、*TOL*和Pearson系数对所选取的

滑坡影响因子进行多重共线性检测,检测结果如表 1 所示。由表 1 可以看出,土地利用类型的 *VIF* 值最大 (1.412),其次为坡度 (1.335),岩土体类型的 *VIF* 值最小 (1.052)。从 *Pearson* 系数的计算结果来看,平面曲率的 *Pearson* 系数最大 (0.411),其次为坡度 (0.306),高程的 *Pearson* 系数最小 (-0.050)。所有影响因子的 *VIF* 值均小于 10,*TOL* 值均大于 0.1,且 *Pearson* 系数均小于 0.7,因此,多重共线性问题不存在于本文所选择的影响因子之间,所有的因子都可以参与后续建模。

4.2 滑坡易发性分区

根据各影响因子对应的分级,统计研究区滑坡的频率密度,最终建立起 IOE 模型,模型中的各指标参数计算结果见表 2。由表 2 可以看出,坡向的综合得分最高 (0.8401),且滑坡点主要集中在 26.553 ~

51.093 范围内,其次为平面曲率 (0.2748),然后为 *NDVI* (0.2491),岩土体类型的综合得分最低,为 0.0459。结果表明,影响因子分级的频率密度并不具有线性关系,与野外实际观察的结果相符合。

表 1 滑坡影响因子多重共线性问题检测表

影响因子	<i>VIF</i>	<i>TOL</i>	<i>Pearson</i>
坡度	1.335	0.749	0.306
坡向	1.177	0.849	0.169
高程	1.206	0.829	-0.050
<i>NDVI</i>	1.182	0.966	-0.021
岩土体类型	1.052	0.950	-0.004
土地利用类型	1.412	0.708	0.009
平面曲率	1.290	0.775	0.411
剖面曲率	1.254	0.797	0.217

表 2 各滑坡影响因子的指标参数计算结果

影响因子	分级	像元数	像元密度	滑坡点	滑坡点占比	$FR_{ij}$	$S_{ij}$	$M_j$	$M_{jmax}$	$I_j$	$W_j$
坡度/(°)	0.000 ~ 10.862	80906	0.1005	3	0.0411	0.4089	0.0904	2.1124	2.5850	0.1828	0.1379
	10.863 ~ 19.110	129054	0.1603	7	0.0959	0.5982	0.1147				
	19.111 ~ 26.552	201973	0.2509	17	0.2329	0.9283	0.1780				
	26.553 ~ 34.196	212681	0.2642	24	0.3288	1.2445	0.2386				
	34.196 ~ 51.093	180483	0.2242	22	0.3014	1.3443	0.2578				
坡向	平坦	3037	0.0038	2	0.0274	7.2629	0.4626	1.6434	3.1699	0.4816	0.8401
	北	83621	0.1039	3	0.0411	0.3957	0.0224				
	北东	103855	0.1290	1	0.0137	0.1062	0.0060				
	东	125221	0.1555	14	0.1918	1.2330	0.0697				
	南东	89216	0.1108	22	0.3014	2.7196	0.1536				
	南	74976	0.0931	17	0.2329	2.5006	0.1413				
	南西	97195	0.1207	9	0.1233	1.0212	0.0577				
	西	124830	0.1550	4	0.0548	0.3534	0.0200				
	北西	103146	0.1281	1	0.0137	0.1069	0.0060				
	784 ~ 1041	23805	0.0296	1	0.0137	0.4633	0.1103	2.1906	2.3219	0.0565	0.0475
高程/m	1042 ~ 1139	185051	0.2298	5	0.0685	0.2980	0.0826				
	1140 ~ 1227	339649	0.4219	46	0.6301	1.4937	0.4138				
	1128 ~ 1317	225478	0.2801	18	0.2466	0.8804	0.2439				
	1318 ~ 1579	31114	0.0386	3	0.0411	1.0634	0.2946				
	-7.727 ~ -1.785	37039	0.0460	4	0.0548	1.1910	0.2184	1.7370	2.3219	0.2519	0.2748
平面曲率	-1.784 ~ -0.531	147489	0.1832	11	0.1507	0.8225	0.0814				
	-0.530 ~ 0.305	393296	0.4885	45	0.6164	1.2619	0.1249				
	0.306 ~ 1.298	175772	0.2183	4	0.0548	0.2510	0.0248				
1.299 ~ 6.053	51501	0.0640	9	0.1233	1.9273	0.1907					

续表 2

影响因子	分级	像元数	像元密度	滑坡点	滑坡点占比	$FR_j$	$S_j$	$M_j$	$M_{j\max}$	$I_j$	$W_j$
NDVI	-0.849 ~ 0.018	90025	0.1118	2	0.0274	0.2450	0.0545	1.6782	2.3219	0.2772	0.2491
	0.019 ~ 0.202	114691	0.1425	5	0.0685	0.4808	0.1106				
	0.203 ~ 0.367	165209	0.2052	27	0.3699	1.8024	0.4146				
	0.368 ~ 0.525	219431	0.2726	33	0.4521	1.6586	0.3816				
	0.526 ~ 0.766	215741	0.2680	6	0.0822	0.3067	0.0706				
土地利用类型	耕地	110460	0.1372	5	0.0685	0.4992	0.1106	1.7179	2.3219	0.2602	0.2349
	建设用地	176294	0.2190	14	0.1918	0.8758	0.3273				
	林草地	492681	0.6120	51	0.6986	1.1416	0.4266				
	水域	9099	0.0113	0	0.0000	0.0000	0.0000				
	裸地	16563	0.0206	3	0.0411	1.9976	0.7465				
岩土体类型	黄土	524619	0.6516	49	0.6712	1.0301	0.2918	1.8961	2.0000	0.0520	0.0459
	砂岩	123931	0.1539	15	0.2055	1.3349	0.3819				
	泥岩	50120	0.0623	2	0.0274	0.4401	0.1259				
	砂质泥岩	106427	0.1322	7	0.0959	0.7254	0.2075				
剖面曲率	-7.321 ~ -1.529	39606	0.0492	1	0.0137	0.2785	0.0543	1.8177	2.3219	0.2172	0.2226
	-1.528 ~ -0.370	134139	0.1666	3	0.0411	0.2467	0.0394				
	-0.369 ~ 0.623	390750	0.4853	39	0.5342	1.1008	0.1758				
	0.624 ~ 1.892	186223	0.2313	18	0.2466	1.0660	0.1702				
	1.893 ~ 6.747	54379	0.0675	12	0.1644	2.4337	0.3886				

在 ArcGIS 平台中,按照表 2 的计算结果对影响因子重新分类,然后利用计算出的熵 ( $M_j$ ) 对重新分类后的因子图层赋值,最后按照公式 (7) 叠加各因子图层,最终计算出研究区的  $LSI_{IOE}$ ,得到研究区滑坡易发性分区图,如图 3 所示。基于  $LSI_{IOE}$  值,按照自然间断点法将滑坡易发性分区图分为 5 个区间,分别为极低易发区 (2.272 3 ~ 4.449 5),低易发区 (4.449 6 ~ 4.896 2),中易发区 (4.896 3 ~ 5.305 6),高易发区 (5.305 7 ~ 6.301 3) 和极高易发区 (6.301 4 ~ 7.017 6)。

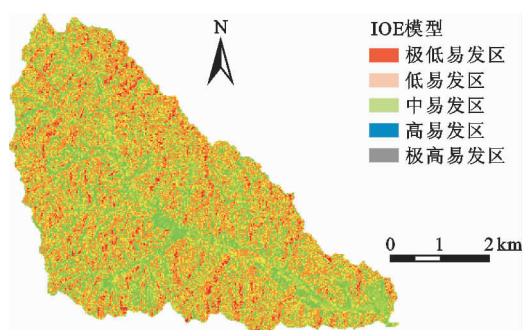


图 3 研究区滑坡易发性分区图

### 4.3 分区结果评价

本文分别利用训练样本集和测试样本集对滑坡易发性分区结果进行评价,见表 3。从表 3 可以看出,基于训练样本集检测出的研究区滑坡易发性分区结果的精确度为 86.27%,说明基于 IOE 模型得到的 LSM 准确率较高,分类结果的可信度较高;基于测试样本集检测出的研究区滑坡易发性分区结果的精确度为 90.91%,说明 IOE 模型在研究区进行滑坡易发性分区时的适用性较强。

表 3 滑坡易发性分区结果评价

统计学指标	IOE 模型	
	训练样本集	测试样本集
TP	42	21
TN	46	19
FP	4	1
FN	10	3
Accuracy / %	86.27	90.91

为了进一步客观地评价研究区滑坡易发性分区

的结果,本文使用 ROC 曲线下的面积( $AUC$ )来定量反映模型的分类效果,IOE 模型的 ROC 曲线见图 4。由图 4 可以看出,训练样本集的  $AUC$  值为 0.828 3 (图 4(a)),测试样本集的  $AUC$  值为

0.964 1 (图 4(b))。训练样本集和测试样本集的  $AUC$  值均大于 0.8,说明 IOE 模型的分类能力较强,在测试样本集中, $AUC$  的值介于 0.9~1.0 之间,也说明模型具有良好的泛化性能。

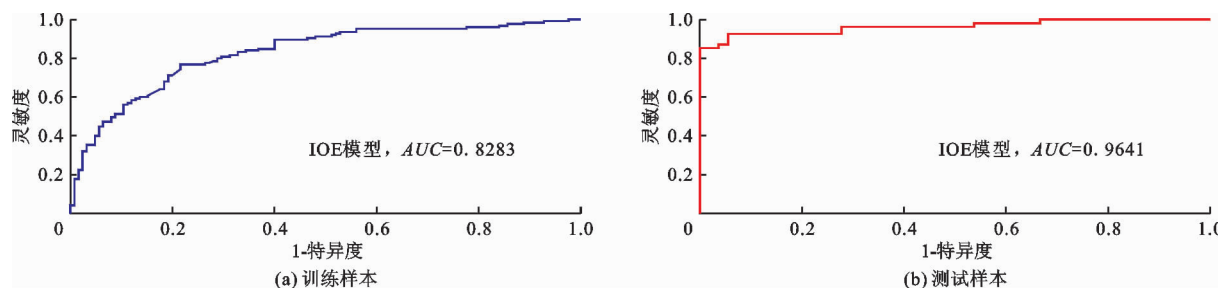


图 4 IOE 模型的 ROC 曲线图

## 5 讨论

数据的分辨率对滑坡易发性分区建模十分关键,由于现阶段可获得的地质资料的图幅的比例尺最大为 1:200 000,所以为了将低分辨率数据所引起数据的稀释性最小化,本文所选取的 8 种影响因子最终都被重采样成  $30\text{ m} \times 30\text{ m}$  分辨率的栅格图像。在影响因子分级方面,例如  $NDVI$  数据自身有较强的规律性,所以利用自然间断点法分级,数据不会产生突兀的现象。

本文通过建立 IOE 模型对研究区进行了滑坡易发性分区,IOE 模型作为常用的统计学分类模型,不需要专家经验,可以得到客观的分类结果。虽然利用 IOE 模型得到的研究区滑坡易发性分区的结果准确率大于 80%,并且测试样本集的  $AUC$  值为 0.964 1,但该模型能否在不同区域得到良好的分类效果,还需要增加更多的研究案例。

## 6 结论

人类工程活动在中国西北部黄土地区不断发展且日益强烈,随之而来的是滑坡灾害的频繁发生,对黄土地区人民的生产和生活造成了极大的威胁。本文以陕西省延安市宝塔区碾庄沟流域为研究区,在已有滑坡数据的基础上,利用 IOE 模型进行滑坡易发性分区,最后利用准确率和 ROC 曲线对分类结果进行评价。最终得到的主要结论如下:

(1) 利用  $VIF$ 、 $TOL$  和  $Pearson$  系数对 8 种影响因子检测,结果表明各影响因子之间不存在多重共线性问题。坡向在 IOE 模型中的综合得分最高 ( $W_j = 0.840 1$ ),是影响研究区滑坡发生的最关键因素。

(2) 训练样本集和测试样本集的准确率均大于 0.8,说明研究区的滑坡易发性分区结果可信度高。训练样本集和测试样本集的  $AUC$  值均大于 0.8,说明 IOE 模型的分类能力较强,且测试样本集的  $AUC$  值为 0.964 1,说明 IOE 模型具有良好的泛化性能。

### 参考文献:

- [1] PETLEY D. Global patterns of loss of life from landslides [J]. *Geology*, 2012, 40(10): 927–930.
- [2] 唐绪波. 基于信息量模型的神农架林区滑坡风险评价[J]. *水利与建筑工程学报*, 2018, 16(4): 115–119+169.
- [3] AGHDAM I N, PRADHAN B, PANABI M. Landslide susceptibility assessment using a novel hybrid model of statistical bivariate methods (FR and WOE) and adaptive neuro-fuzzy inference system (ANFIS) at southern Zagros mountains in Iran [J]. *Environmental Earth Sciences*, 2017, 76(6): 237.
- [4] 许冲,戴福初,姚鑫,等. 基于 GIS 的汶川地震滑坡灾害影响因子确定性系数分析[J]. *岩石力学与工程学报*, 2010, 29(S1): 2972–2981.
- [5] ZHANG Tingyu, HAN Ling, CHEN Wei, et al. Hybrid integration approach of entropy with logistic regression and support vector machine for landslide susceptibility modeling [J]. *Entropy*, 2018, 20(11): 884.
- [6] 樊晓一,乔建平,陈永波. 层次分析法在典型滑坡危险度评价中的应用[J]. *自然灾害学报*, 2004, 13(1): 72–76.
- [7] 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): 2150–2167.
- [8] 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.
- [9] MONDAL S, MAITI R. Integrating the Analytical hierarchy process (AHP) and the frequency ratio (FR) model in landslide susceptibility mapping of Shiv-khola watershed, Darjeeling Himalaya [J]. *International Journal of Disaster Risk Science*, 2013, 4(4): 200–212.
- [10] PRADHAN B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS [J]. *Computers & Geosciences*, 2013, 51: 350–365.
- [11] YOUSSEF A M, POURGHASEMI H R, POURTAGHI Z S, et al. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia [J]. *Landslides*, 2016, 13(5): 839–856.
- [12] PRABU S, RAMAKRISHNAN S S. Combined use of socio economic analysis, remote sensing and GIS data for landslide hazard mapping using ANN [J]. *Journal of the Indian Society of Remote Sensing*, 2009, 37(3): 409–421.
- [13] 黄 赠, 王 锐, 赵 宇, 等. 隐伏断层地震诱发滑坡易发性评价 [J]. *浙江大学学报(工学版)*, 2017(11): 2136–2143.
- [14] PANDEY V K, POURGHASEMI H R, SHARMA M C. Landslide susceptibility mapping using maximum entropy and support vector machine models along the Highway Corridor, Garhwal Himalaya [J]. *Geocarto International*, 2020, 35(2): 168–187.
- [15] 陕西省地质矿产局. 陕西省区域地质志[M]. 北京:地质出版社, 1989.
- [16] ZHANG Tingyu, HAN Ling, HAN Jichang, et al. Assessment of landslide susceptibility using integrated ensemble fractal dimension with kernel logistic regression model [J]. *Entropy*, 2019, 21(2): 218.
- [17] CHEN Wei, POURGHASEMI H R, NAGHIBI S A. A comparative study of landslide susceptibility maps produced using support vector machine with different kernel functions and entropy data mining models in China [J]. *Bulletin of Engineering Geology & the Environment*, 2017, 77(2): 1–18.
- [18] PHAM B T, BUI D T, PRAKASH I. Bagging based support vector machines for spatial prediction of landslides [J]. *Environmental Earth Sciences*, 2018, 77(4): 146.
- [19] DEVKOTA K C, REGMI A D, POURGHASEMI H R, et al. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya [J]. *Natural Hazards*, 2013, 65(1): 135–165.
- [20] POURGHASEMI H R, MOHAMMADY M, PRADHAN B. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran [J]. *Catena*, 2012, 97: 71–84.
- [21] 张桂荣, 殷坤龙, 刘传正, 等. 基于GIS的陕西省旬阳地区滑坡灾害危险性区划 [J]. *中国地质灾害与防治学报*, 2003, 14(4): 39–43.