

地质灾害易发性评价因子分级的AIFFC算法优化

陈 宾, 李颖懿, 张联志, 屈添强, 魏 娜, 刘 宁, 黄春林

Classification optimization of geological hazard susceptibility evaluation factors based on AIFFC algorithm

CHEN Bin, LI Yingyi, ZHANG Lianzhi, QU Tianqiang, WEI Na, LIU Ning, and HUANG Chunlin

在线阅读 View online: <https://doi.org/10.16031/j.cnki.issn.1003-8035.202210048>

您可能感兴趣的其他文章

Articles you may be interested in

基于RBF神经网络信息量耦合模型的滑坡易发性评价

Landslide susceptibility assessment by the coupling method of RBF neural network and information value: A case study in Min Xian, Gansu Province

黄立鑫, 郝君明, 李旺平, 周兆叶, 贾佩钱 中国地质灾害与防治学报. 2021, 32(6): 116–126

北京山区突发性地质灾害易发性评价

Assessment on the susceptibility of sudden geological hazards in mountainous areas of Beijing

罗守敬, 王珊珊, 付德荃 中国地质灾害与防治学报. 2021, 32(4): 126–133

基于GIS和加权信息量的湖北鄂州地质灾害易发性区划

`\${suggestArticle.titleEn}`

张波, 石长柏, 肖志勇, 张金朝, 邬磊, 习彪, 卢胜周 中国地质灾害与防治学报. 2018, 29(3): 101–107

基于快速聚类-信息量模型的汶川及周边两县滑坡易发性评价

Landslide susceptibility assessment based on K-means cluster information model in Wenchuan and two neighboring counties, China

周天伦, 曾超, 范晨, 毕鸿基, 龚恩慧, 刘晓 中国地质灾害与防治学报. 2021, 32(5): 137–150

基于遥感影像多尺度分割与地质因子评价的滑坡易发性区划

Landslide susceptibility assessment based on multi-scale segmentation of remote sensing and geological factor evaluation

李文娟, 邵海 中国地质灾害与防治学报. 2021, 32(2): 94–99

基于GIS和信息量模型的安徽池州地质灾害易发性评价

`\${suggestArticle.titleEn}`

王雷, 吴君平, 赵冰雪, 姚志强, 张乐勤 中国地质灾害与防治学报. 2020, 31(3): 96–102



关注微信公众号, 获得更多资讯信息

DOI: 10.16031/j.cnki.issn.1003-8035.202210048

陈宾, 李颖懿, 张联志, 等. 地质灾害易发性评价因子分级的 AIFFC 算法优化[J]. 中国地质灾害与防治学报, 2024, 35(1): 72-81.

CHEN Bin, LI Yingyi, ZHANG Lianzhi, et al. Classification optimization of geological hazard susceptibility evaluation factors based on AIFFC algorithm[J]. The Chinese Journal of Geological Hazard and Control, 2024, 35(1): 72-81.

地质灾害易发性评价因子分级的 AIFFC 算法优化

陈 宾^{1,2}, 李颖懿¹, 张联志³, 屈添强³, 魏 娜¹, 刘 宁¹, 黄春林³

(1. 湘潭大学土木工程学院, 湖南 湘潭 411105; 2. 湘潭大学岩土力学与工程安全湖南省重点实验室, 湖南 湘潭 411105; 3. 湖南省国土空间调查监测所, 湖南 长沙 421002)

摘要: 针对地质灾害易发性评价因子分级数不确定的问题, 引入自适应膨胀因子模糊覆盖分级方法 (fuzzy cover approach for clustering based on adaptive inflation factor, AIFFC) 对易发性评价因子分级进行优化。以湖南省湘乡市为研究区, 提取了坡度、坡向、高程、年平均降雨量、归一化植被指数、道路、断层、岩性和土地利用 9 类评价因子, 运用 AIFFC 及自然断点法 (natural breakpoint classification, NBC) 对连续型因子进行分级, 并分别代入加权信息量模型和随机森林模型, 获取研究区易发性区划图。采用单因子分级结果精度、灾积比分析和易发性分区结果对 AIFFC 分级法的优越性进行检验, 结果表明: 各因子采用 AIFFC 算法分级的 AUC 值均高于自然断点法; 基于 AIFFC 的随机森林模型及加权信息量模型的高易发区灾积比分别提升了 56.3%、74.6%, 低易发区灾积比分别降低了 48%、58.1%, AUC 值分别提升了 7.6%、2.7%。采用 AIFFC 分级方法优化了地质灾害易发性评价因子分级, 显著提高了地质灾害易发性评价的合理性。

关键词: 膨胀因子; 模糊覆盖; 加权信息量; 随机森林; 易发性评价; 地质灾害

中图分类号: P694

文献标志码: A

文章编号: 1003-8035(2024)01-0072-10

Classification optimization of geological hazard susceptibility evaluation factors based on AIFFC algorithm

CHEN Bin^{1,2}, LI Yingyi¹, ZHANG Lianzhi³, QU Tianqiang³, WEI Na¹, LIU Ning¹, HUANG Chunlin³

(1. School of Civil Engineering, Xiangtan University, Xiangtan, Hunan 411105, China; 2. Hunan Provincial Key Laboratory of Geomechanics and Engineering Safety, Xiangtan University, Xiangtan, Hunan 411105, China; 3. Hunan Provincial Territorial Space Survey and Monitoring Institute, Changsha, Hunan 421002, China)

Abstract: This paper addresses the issue of uncertainty in the grading of geological hazard susceptibility evaluation factors and introduces the adaptive expansion factor fuzzy coverage grading method (AIFFC) to optimize the grading of geological hazard susceptibility evaluation factors. Taking Xiangxiang City, Hunan Province as the research area, nine evaluation factors, including slope, slope direction, elevation and average annual rainfall, normalized difference vegetation index for land use, roads, faults, lithology, were extracted. The AIFFC method and the natural breakpoint method were used to grade continuous factors. These graded factors were then incorporated into a weighted information model and random forest model to obtain a susceptibility zoning map for the study area. The superiority of the AIFFC classification method was tested through the comparison of single-factor grading results, disaster product ratio analysis, and ROC curve comparison of susceptibility zoning

收稿日期: 2022-10-28; 修订日期: 2022-11-13

投稿网址: <https://www.zgdzhyfzxb.com/>

基金项目: 湖南省创新性省份建设专项(2019RS1059); 国家自然科学基金项目(51774131; 41972282)

第一作者: 陈 宾(1977—), 男, 河南驻马店人, 教授, 博士, 主要从事地质灾害防治方面的研究工作。

E-mail: 403021235@qq.com

通讯作者: 张联志(1988—), 男, 江苏沛县人, 硕士, 工程师, 主要研究方向为水工环地质。E-mail: 2149859375@qq.com

results. Based on AIFFC, the hazard accumulation ratio of the random forest model and the weighted information entropy model in the high susceptibility areas increased by 56.3% and 74.6%, respectively, while in the low susceptibility areas, it decreased by 48% and 58.1%, respectively. The *AUC* values increased by 7.6% and 2.7%, respectively. The AIFFC classification method is used to optimize the evaluation factor classification of geological disaster susceptibility, which significantly improves the rationality of the evaluation of geological disaster susceptibility.

Keywords: expansion factor; fuzzy covering; weighted information entropy; random forest; susceptibility assessment; geo-hazard

0 引言

我国是地质灾害频发的国家, 地质灾害易发性评价是我国重要的防灾减灾工作之一。地质灾害易发性评价因子厘定及其分级的合理性是易发性评价的工作基础和精度保障。目前, 常用的因子分级方法有自然断点法^[1]、专家经验法^[2]、等间距法^[3]、频率比法^[4]等。孙德亮等^[5]、杨得虎等^[6]、解明礼等^[7]通过对比不同分级方法获得的地质灾害易发性评价结果证明了灾害因子分级对评价精度的影响。鉴于此, 凌晓等^[8]、郭建华等^[9]、陈绪钰等^[10]、陈伟等^[11]分别采用对称分级法、方差分析法、迭代自组织聚类法、K-means 聚类算法对现有分级方法进行了改进, 但如何根据地质灾害分布特征客观确定各评价因子分级数的研究较为缺少。

自适应膨胀因子模糊覆盖分级 (fuzzy cover approach for clustering based on adaptive inflation factor, AIFFC) 算法是一种确定一维数据分级数的自适应膨胀因子的模糊覆盖分级方法, 最早被应用于地图制图领域^[12], 其能够根据数据分布特征, 动态生成分级数及分级区间, 有效的解决分级数确定受主观影响的问题。孙娟娟^[13]通过数据实例证明了 AIFFC 算法分级数的最优性及分级结果的精确性; 姚宇婕等^[14]基于 AIFFC 算法优化了引导型专题数据分级处理模式; 张涵斐^[15]将 AIFFC 算法应用于多尺度地理信息数据的分级处理, 实现了多尺度地理信息分级显示的效果。

本文以湖南省湘乡市为研究区, 分别采用自然断点法和 AIFFC 分级法对坡度、坡向、高程、年平均降雨量、归一化植被指数等评价因子进行分级赋值, 并分别代入加权信息量模型及随机森林模型对研究区进行地质灾害易发性区划评价。从单因子 ROC 曲线分析、易发性区划结果 ROC 曲线分析及灾积比对比三个方面对区划评价结果进行精度对比与分析, 从而获取最优分级方法。

1 基础研究理论

1.1 AIFFC 算法

AIFFC 其核心思想是对一个覆盖 ρ_i 定义相应的膨

胀因子 α_i 和覆盖中心 m , 对其进行扩张形成一个新的覆盖 ρ'_i ^[12, 16]。依次进行下去, 直到所有的数据被某个 ρ_i 覆盖, 则有 $X = \bigcup_i \rho_i, (\rho_i \cap \rho_j = \emptyset, i \neq j)$ 。AIFFC 算法具体计算步骤如下:

步骤一: 输入历史地质灾害点评价因子数据集 X , 并将各评价因子 X_i 从小到大排列, 取评价因子模糊覆盖半径 $\lambda, l=1, m=1$ 。

$$\lambda = \begin{cases} \lambda_1, & \lambda_1 < d_{\max} \\ d_{\max}, & \lambda_1 \geq d_{\max} \end{cases} \quad (1)$$

其中, $d_i = d(x_i, x_{i+1}), d(x, y) = |x - y|, d_{\max}$ 为 d_i 的最大值。各评价因子距离序列 $\{d_i\}$ 降序排序后仍记各元素为 d_i 。 λ_1 的取值分为以下 3 种情况:

1) 若 $[d_i/d_{i-1}] \in [1, 2]$, 取 $c=n$ 。

$$\lambda_1 = \sum_{i=1}^{n-1} \left[1 - \exp\left(-(d_i/d_{\max})^2/2\right) \right] d_i \quad (2)$$

2) 若 $d_{\max}$ 远远大于 d_i , 从 $d_{\max}$ 处将评价因子数据分为两部分, 若两部分数据分布均匀, λ_1 按 1) 中步骤分别运用算法分级。

3) 否则, 取 $d_{\text{mid}} = \frac{1}{n-1} \sum_{i=1}^{n-1} d_i, c=1$ 。

$$\lambda_1 = \sum_{i=1}^{n-1} \left\{ \exp\left[-(d_i/d_{\text{mid}})^2/2\right] / 2 \right\} d_i \quad (3)$$

步骤二: 从第 m 个历史地质灾害点元素起定义模糊覆盖, 设 $n_l = |\rho_l|$ 为 ρ_l 内历史地质灾害点个数, $k=1$ 。

$$\rho_l = \{y \mid d(x_m, y) < \lambda, y \in X\} \quad (4)$$

步骤三: $X = X / \rho_l$, 若 $X = \emptyset$, 转步骤四; 否则, 重新定义评价因子 ρ_l 向外膨胀的中心及半径为:

$$x_l^{(k)} = \left(\frac{n_l - 1}{n_l} \right)^{n_{l-1}} x_m + \left(\frac{n_l - 1}{n_l} \right)^{n_{l-2}} \frac{1}{n_l} x_{m+1} + \cdots + \frac{1}{n_l} x_{n_l} \quad (5)$$

取 $\lambda_l^{(k)} = \alpha_l^{(k)} \cdot \lambda$,

$$\alpha_l^{(k)} = \left[1 - \frac{n_l - 1}{n_l} \ln \left(1 + \frac{1}{c \cdot k} \right) \right]^k \quad (6)$$

其中, k 为覆盖膨胀的次数, c 为常数, 代表膨胀的大

小。再作 $x_l^{(k)}$ 的覆盖:

$$\rho_l^{(k)} = \{y | d(x_l^{(k)}, y) < \lambda_l^{(k)}, y \in X\} \quad (7)$$

若 $\rho_l^{(k)} \neq \emptyset$, $n_l = n_l + |\rho_l^{(k)}|$, $\rho_l = \rho_l \cup \rho_l^{(k)}$, $k = k + 1$, 继续步骤三; 否则, 若 $\rho_l^{(k)} = \emptyset$, 则令 $m = m + n_l$, $l = l + 1$, 转步骤二。

步骤四: 输出 l 及 ρ_l , l 为各评价因子的分级数, $\{\rho_l\}$ 为各评级因子的分级结果。

1.2 自然断点法(natural breakpoint classification, NBC)

NBC 法是由乔治·弗雷德里克·詹克斯提出的一种基于聚类思维的单变量分组方法^[17]。其核心思想是通过迭代对比各分组及分组中元素的均值与观测值之间的平方差之和确定每个数据在分组中的最佳排列。在确定的分级数下, 计算出数据分布中的中断点, 给出最佳分类区间, 使组间差异化最大, 组内差异化最小, 较好的保护数据的统计特性。NBC 计算原理如下^[18]:

$$DNB = \frac{1}{n} \sum_{i=1}^n DNB_i \quad (8)$$

$$SDAM = \sum_{i=1}^n \sum_{j=1}^m (DNB_i - DNB_j)^2 \quad (9)$$

其中, $SDAM$ 为研究区数组平均值的偏差平方和; DNB_i 为研究区内数据值; DNB 为平均数据值。迭代每个数据值范围组合, 计算类别均值的平方偏差平方和 $SDCM_ALL$, 找到最小值, 并计算方差拟合优度 GVF 。GVF 值的范围在 0~1, 其值代表分类结果是否类内差异最小, 类间差异最大, 1 表示拟合极好, 0 表示拟合极差。

$$GVF = (SDAM - SDCM)/SDAM \quad (10)$$

1.3 加权信息量模型(weighted information value based on entropy weight method, MIV)

通过熵权法得到各评价因子的权重, 再结合信息量法得到的各评价因子的信息量值, 两者相乘得到评价因子的最终信息量^[19-21]。MIV 计算方法如下:

$$I = \sum_{i=1}^n \omega_i I_i = \sum_{i=1}^n \omega_i \ln \frac{N_i/N}{S_i/S} \quad (11)$$

式中: I —— n 种评价因子组合下各地质灾害易发性栅格单元的加权信息量值;

I_i ——评价因子 i 的信息量值;

ω_i ——通过熵权法计算的评价因子 i 的权重值;

N ——研究区内历史灾害点总个数;

N_i ——各分级区间内历史灾害点个数;

S ——研究区总面积/ km^2 ;

S_i ——各分级区间面积/ km^2 。

1.4 随机森林模型(random forest, RF)

RF 是由 Breiman^[22]首次提出的一种集成方法。随机森林模型的基本思想是将 n 个独立的决策树组合建立一个模型, 森林中的 n 棵决策树具有相同的分布, 且每棵决策树都由独立采样的随机向量值决定, 运用模型中的决策树对输入的样本进行预测和判断, 通过机器学习形成不同的分类模型 $x_1(Y), x_2(Y), \dots, x_n(Y)$, 从而建立随机森林模型^[23]。计算公式如下:

$$\gamma(y) = \arg \max_Z \sum_{i=1}^k I(x_i(Y) = Z) \quad (12)$$

式中: $\gamma(y)$ ——随机森林模型;

$x_i(Y)$ ——单个决策树模型;

Z ——输出变量;

I ——显函数。

2 研究区概况

湘乡市总面积 1 967.4 km^2 , 居湘中偏东, 靠近北回归线, 属亚热带季风湿润气候, 年平均降雨量 1 408.6 mm; 境内主要由涟水、洧水和靳水组成, 水系较为发育。主要出露地层为三叠系、石炭系、泥盆系和震旦系等; 区域构造根据其力学性质、空间展布规律及组合关系分为东西向构造、北东向构造、弧形构造、北西向构造四类, 境内以北东向构造为主。湘乡市在气象水文、地形地貌、地质构造、新构造运动与地震等自然条件和采矿、修路等人为活动的共同作用下, 境内发生各类地质灾害共 270 处(图 1), 主要以滑坡灾害为主, 对湘乡市的人类生命财产安全造成了重大威胁, 有效的预测和治理地质灾害成为了相关部门有待解决的问题。

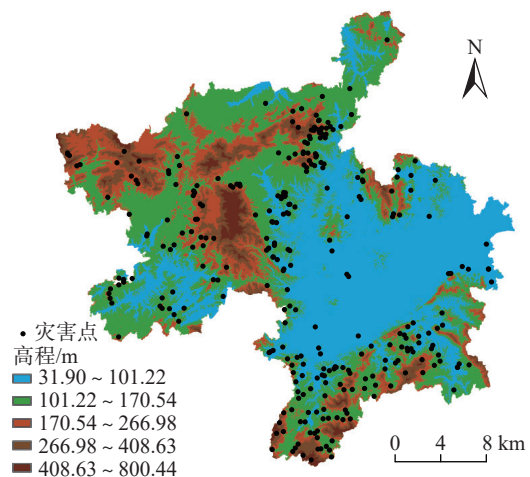


图 1 湘乡市地质灾害分布图

Fig. 1 Geological disaster distribution map of Xiangxiang City

3 评价因子选取与分级

以湘乡市地质灾害分布规律及孕灾环境分析^[24-26]为基础,选取坡度、坡向、高程、年平均降雨量、归一化植被指数、道路、断层、地层岩性和土地利用作为地质灾害易发性评价因子(图 2)。采用 30 m×30 m 的栅格单元作为评价研究和空间分析的基本单元,利用 ArcGIS 中栅格转点工具将栅格文件转化为点文件,然后用多值提取至点对所有评价因子图层进行值提取,

得到每个栅格的评价因子数据。其中,坡度、坡向和高程因子通过 ArcGIS 对研究区的 DEM 数据进行空间分析得到;年平均降雨量通过统计研究区各站点的年平均降雨量数据,利用反距离权重工具获取;地层岩性评价因子通过研究区工程地质图得到;归一化植被指数是利用 ENVI 软件提取遥感影像,然后在 ArcGIS 中赋值分类得到;道路、断层因子通过 ArcGIS 中的多环缓冲区工具获取;土地利用评价因子根据研究区土地利用类型

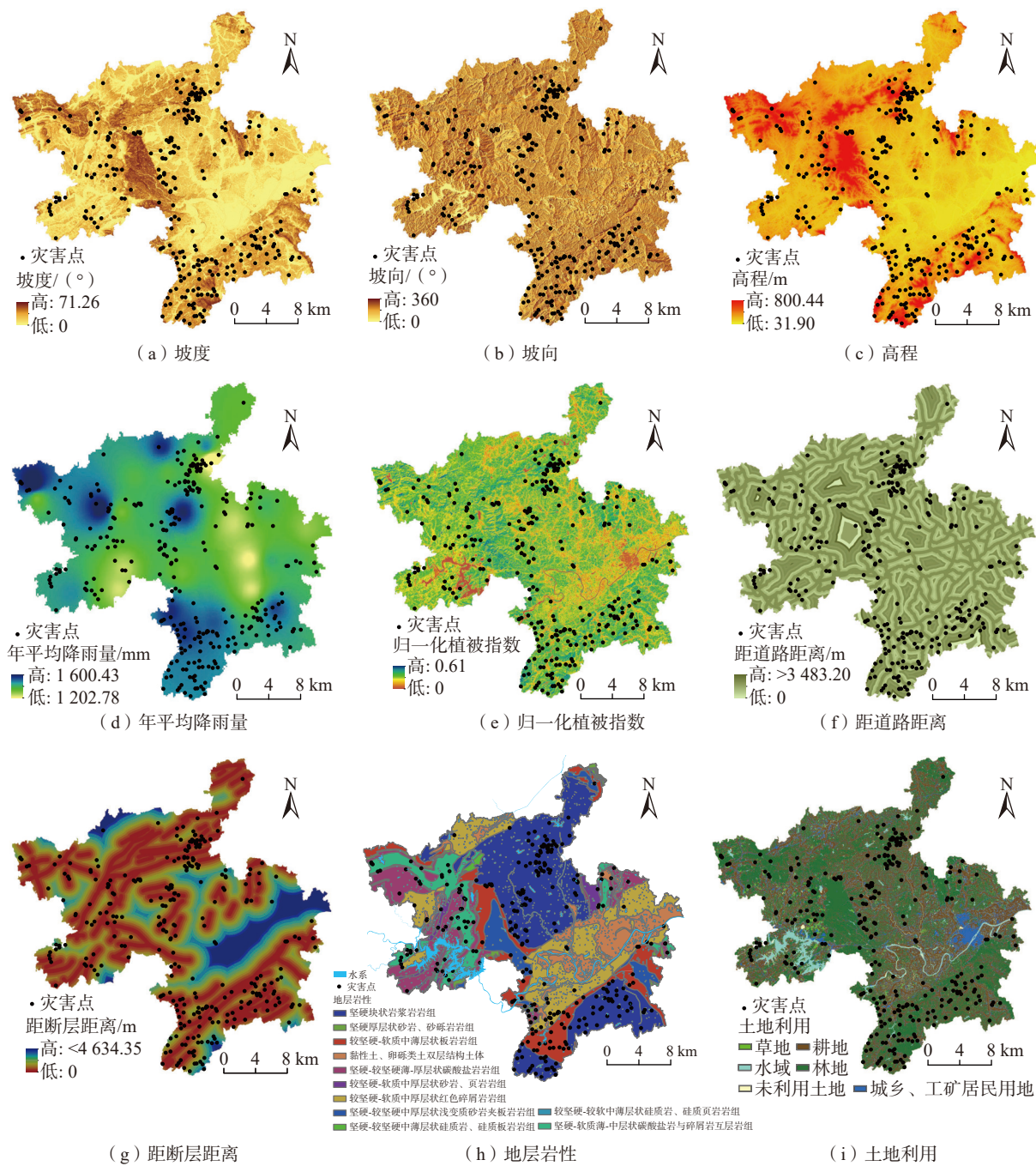


图 2 易发性各评价因子图层
Fig. 2 Layers of susceptibility evaluation factors

图提取分析得到。离散型评价因子地层岩性和土地利用根据其原有的类型等级划分如图 2(h)、(i), 其他 7 个连续型评价因子分别进行 AIFFC 分级划分及 NBC 分级划分。

3.1 AIFFC 分级划分

根据湘乡市 270 个已发生地质灾害点的原始数据, 统计各灾害点的坡度、坡向、高程、年平均降雨量、归一化植被指数、道路和断层 7 个连续型评价因子值。根据 1.1 节各步骤通过 Java 工具对 AIFFC 分级算法进行迭代编程, 利用程序对各评价因子进行分级计算, 各

评价因子 AIFFC 分级参数、分级数和分级区间见表 1。历史灾害点数据是离散型数据, 评价因子的范围是连续性数据, 通过历史灾害点数据分析, 研究区内坡度高于 70°, 灾害点高程高于 456.70 m, 年平均降雨量大于 1 554.89 mm, 归一化植被指数高于 0.417, 以及灾害点距离道路和断层在 3.48 km 及 4.63 km 以外, 未见地质灾害发生, 设定为地质灾害不易发范围, 统一为一级。对应评价因子分级数在原来分级数上增加一级, 分级区间采用大于各评价因子最大值的方法表示, 得到 AIFFC 最终分级结果见表 2。

表 1 各评价因子 AIFFC 分级参数
Table 1 AIFFC classification parameters for each evaluation factor

评价因子	研究区范围	灾害点分布范围	d_i	λ	分级数	分级区间($\{p_i\}$)
坡度/(°)	[0,71.26]	[2,70]	1.0~2.0	8	7	{1, 12, 125, 35, 14, 49, 34}
坡向/(°)	[0,360]	[0,360]	1.0~1.67	20	12	{10, 11, 19, 23, 38, 52, 25, 30, 25, 15, 9, 13}
高程/m	[31.90,800.44]	[113.12,456.70]	1.0~2.0	40.33	8	{112, 92, 34, 22, 5, 3, 1, 1}
年平均降雨量/mm	[1 202.78,1 600.43]	[1 221.03,1 554.89]	1.0~1.9	22.81	10	{1, 2, 14, 19, 59, 70, 60, 36, 8, 1}
归一化植被指数	[0,0.61]	[0.015,0.417]	1.0~1.5	0.04	6	{7, 27, 48, 89, 85, 14}
距道路距离/m	[0,>3 483.20]	[326.11,3 483.20]	1.0~1.79	230.84	10	{102, 55, 33, 37, 16, 10, 10, 5, 1, 1}
距断层距离/m	[0,>4 634.35]	[423.71,4 634.35]	1.0~1.5	236.17	11	{75, 57, 43, 38, 10, 16, 10, 6, 5, 5, 5}

表 2 AIFFC 分级结果
Table 2 AIFFC classification results

评价因子	分级数	分级结果
坡度/(°)	8	>0~2;>2~22;>22~35;>35~48;>48~55;>55~65;>65~70;>70
坡向/(°)	12	>0~30;>30~68;>68~90;>90~125;>125~162;>162~190;>190~225;>225~260;>260~280;>280~310;>310~338;>338~360
高程/m	9	>31.9~113.12;>113.12~165.57;>165.57~221.23;>221.23~278.54;>278.54~333.59;>333.59~367.50;>367.50~408.51;>408.51~456.70;>456.70
年平均降雨量/mm	11	>1 202.70~1 221.03;>1 221.03~1 283.85;>1 283.85~1 336.92;>1 336.92~1 368.54;>1 368.54~1 410.03;>1 410.03~1 446.28;>1 446.28~1 484.91;>1 484.91~1 516.62;>1 516.62~1 536.87;>1 536.87~1 554.89;>1 554.89
归一化植被指数	7	>0~0.015;>0.015~0.180;>0.180~0.246;>0.246~0.316;>0.316~0.377;>0.377~0.417;>0.417
距道路距离/m	11	0~326.11;>326.11~664.23;>664.23~989.03;>989.03~1 368.63;>1 368.63~1 688.75;>1 688.75~2 075.72;>2 075.72~2 497.47;>2 497.47~2 831.94;>2 831.94~3 065.30;>3 065.30~3 483.20;>3 483.20
距断层距离/m	12	0~423.71;>423.71~850.88;>850.88~1 272.26;>1 272.26~1 694.48;>1 694.48~2 150.42;>2 150.42~2 485.00;>2 485.00~3 005.74;>3 005.74~3 523.15;>3 523.15~4 107.21;>4 107.21~4 248.58;>4 248.58~4 634.35;>4 634.35

3.2 NBC 分级划分

为保证 AIFFC 分级法及 NBC 法分级数变量的统一, 以 AIFFC 算法计算的分级数为基础, 对 7 个连续型评价因子利用 ArcGIS 重分类工具中的自然断点分级法进行自然断点法分级, 结果见表 3。

4 易发性评价结果

将 AIFFC 和 NBC 法分级结果分别代入加权信息量模型及随机森林模型计算, 得到基于 AIFFC 的加权信息量模型(AIFFC-MIV)、基于 NBC 法的加权信息量模

型(NBC-MIV)、基于 AIFFC 的随机森林模型(AIFFC-RF)和基于自然断点法的随机森林模型(NBC-RF), 对其评价结果进行低、中、高易发性区间划分, 见图 3。

由四个模型易发性分区图可知: 各模型的分区结果存在一些差异, 但是高易发区、中易发区及低易发区的趋势范围基本相同, 分区结果符合湘乡市历史灾害点的分布特征。高易发区主要分布在西北及南部的构造侵蚀剥蚀花岗岩低山地区和构造侵蚀剥蚀变质岩-碎屑岩低山地区, 岩石风化强烈, 残坡积层发育, 地形起伏较大; 低易发区主要分布在中部的河流侵蚀堆积河谷平原

表 3 NBC 分级结果

Table 3 Natural break point method grading results

评价因子	分级数	分级结果
坡度/(°)	8	0 ~ 3.91; >3.91 ~ 9.50; >9.50 ~ 15.37; >15.37 ~ 20.95; >20.95 ~ 26.26; >26.26 ~ 31.85; >31.85 ~ 38.56; >38.56 ~ 71.26
坡向/(°)	12	0 ~ 24.48; >24.48 ~ 57.04; >57.04 ~ 88.18; >88.18 ~ 117.91; >117.91 ~ 147.64; >147.64 ~ 177.37; >177.37 ~ 208.52; >208.52 ~ 239.66; >239.66 ~ 269.39; >269.39 ~ 299.12; >299.12 ~ 328.85; >328.85 ~ 360.00
高程/m	9	31.90 ~ 79.93; >79.93 ~ 121.96; >121.96 ~ 166.99; >166.99 ~ 218.03; >218.03 ~ 275.07; >275.07 ~ 341.11; >341.11 ~ 425.17; >425.17 ~ 542.26; >542.26 ~ 800.44
年平均降雨量/mm	11	1 202.70 ~ 1 289.62; >1 289.62 ~ 1 333.91; >1 333.91 ~ 1 363.46; >1 363.46 ~ 1 386.40; >1 386.40 ~ 1 405.55; >1 405.55 ~ 1 425.18; >1 425.18 ~ 1 444.75; >1 444.75 ~ 1 463.71; >1 463.71 ~ 1 488.88; >1 488.88 ~ 1 523.98; >1 523.98 ~ 1 600.42
归一化植被指数	7	0 ~ 0.051; >0.051 ~ 0.152; >0.152 ~ 0.208; >0.208 ~ 0.259; >0.259 ~ 0.303; >0.303 ~ 0.347; >0.347 ~ 0.611
距道路距离/m	11	0 ~ 71.14; >71.14 ~ 167.93; >167.93 ~ 287.96; >287.96 ~ 398.67; >398.67 ~ 534.77; >534.77 ~ 757.79; >757.79 ~ 1 088.95; >1 088.95 ~ 1 532.08; >1 532.08 ~ 2 160.22; >2 160.22 ~ 3 483.20; >3 483.20
距断层距离/m	12	0 ~ 93.41; >93.41 ~ 195.93; >195.93 ~ 317.53; >317.53 ~ 464.40; >464.40 ~ 667.68; >667.68 ~ 850.88; >850.88 ~ 1 059.58; >1 059.58 ~ 1 363.72; >1 363.72 ~ 1 914.46; >1 914.46 ~ 2 795.16; >2 795.16 ~ 4 634.35; >4 634.35

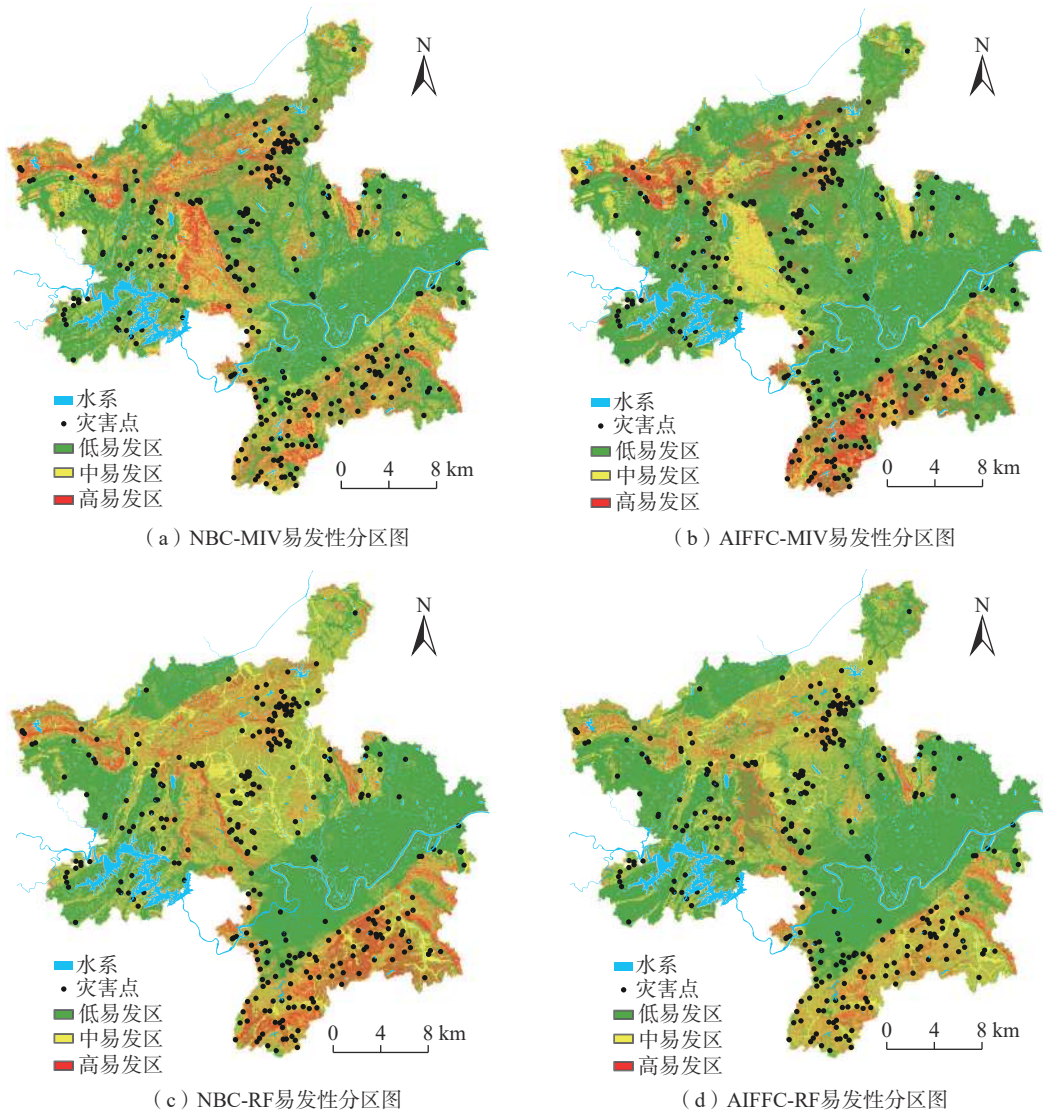


图 3 各模型易发性分区图

Fig. 3 Susceptibility zoning maps of each model

地区和侵蚀剥蚀碎屑岩-碳酸盐岩丘陵地区,地形相对平坦,植被及构造相对不发育。且高易发区人口密度相对集中,城乡建设频繁,矿山开采及人类工程经济活动较强烈,表明人类活动对地质灾害的发生具有较大的影响。

5 对比与分析

5.1 单因子分级结果精度对比

采用单因子 ROC 曲线法对每个因子的分级结果进行评价,根据各评价因子二级区间信息量值由高至低的面积比和对应的信息量值所处单元中的灾害点的百分比做单因子 ROC 曲线图,曲线下方面积(AUC 值)越大,说明该分级结果更合理^[27-29]。由 AIFFC 及 NBC 法的 7 个连续型因子分级结果 ROC 曲线图,见图 4(a)、(b),及 AUC 值对比图,见图 4(c),可知: AIFFC 分级曲线 AUC 值介于 0.546~0.911, NBC 法分级 AUC 值介于 0.541~0.895,均高于 0.5 具有较好的准确性。采用 AIFFC 法分级的各因子 AUC 值皆高于 NBC 法,提升精度介于 1.8%~6.3%,说明 AIFFC 分级结果更趋合理。

5.2 易发性评价结果分析

通过对四种区划评价结果的灾积比对比分析(表 4)可知, AIFFC-MIV 模型与 AIFFC-RF 模型区划的高易发区灾积比分别为 0.955 与 1.318,高于 NBC-MIV 模型与 NBC-RF 模型的 0.611 与 0.755,并分别提升了 56.3% 和 74.6%。AIFFC-MIV 模型与 AIFFC-RF 模型区划的低易发区灾积比分别为 0.026 与 0.013,低于 NBC-MIV 模型与 NBC-RF 模型的 0.050 与 0.031,并分别降低了 48.0% 和 58.1%。

因此,无论采用 MIV 模型或 RF 模型,基于 AIFFC 算法的灾害因子分级方法均大幅提升了高易发区的灾积比并降低了低易发区的灾积比,评价结果更具合理性。

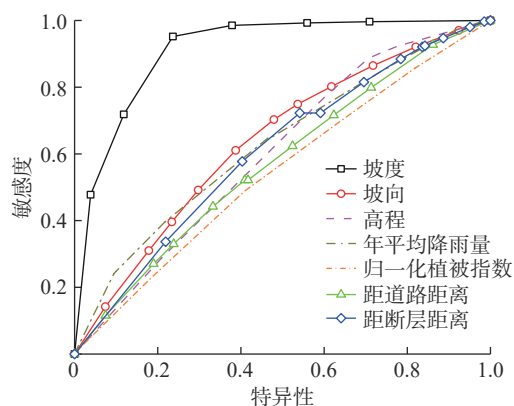
5.3 易发性评价精度对比

对易发性评价结果进行 ROC 曲线分析(图 5)可知:各模型 AUC 值均超过 0.5,评价结果预测精度满足要求^[30]。其中, AIFFC-MIV 模型与 AIFFC-RF 模型区划结果的 AUC 值分别为 0.835 与 0.905,高于 NBC-MIV 模型与 NBC-RF 模型区划结果的 0.776 与 0.880,并分别提升了 7.6% 和 2.7%。

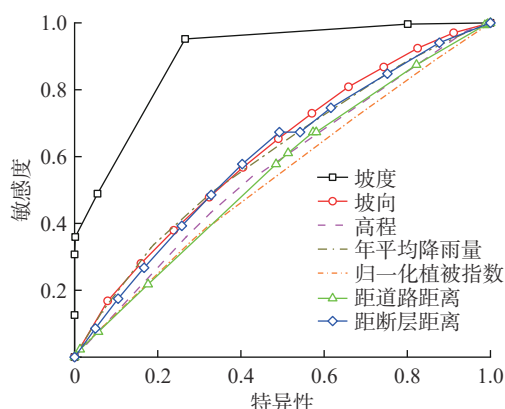
6 结论

(1)通过单因子评价精度对比可知,各因子采用 AIFFC 算法分级比 NBC 法分级的 AUC 值均有提高,提升幅度介于 1.8%~6.3%, AIFFC 分级方法更具合理性。

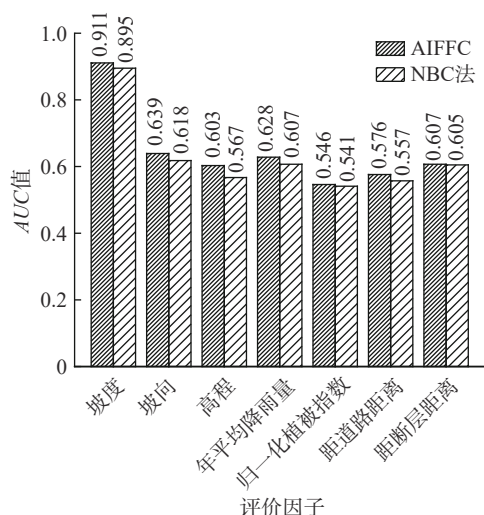
(2)4 种模型易发性区划评价结果中,基于 AIFFC



(a) AIFFC 单因子 ROC 曲线图



(b) NBC 法单因子 ROC 曲线图



(c) 单因子 AUC 值对比图

图 4 单因子 ROC 曲线分析图

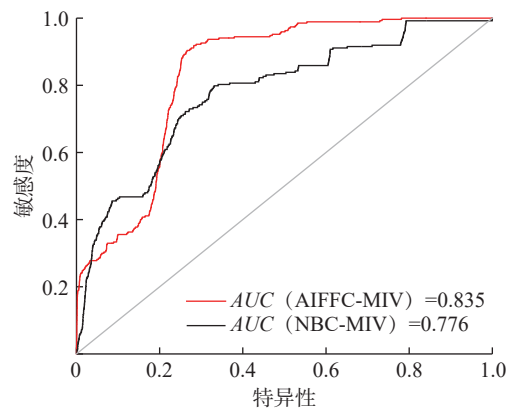
Fig. 4 Single-factor ROC curve analysis

的评价结果相比基于 NBC 法的评价结果高易发区灾积比分别提升了 56.3%、74.6%,低易发区灾积比分别降低了 48%、58.1%, AUC 值分别提高了 0.059(7.6%)、0.025(2.7%),基于 AIFFC 的易发性评价精确性和预测能力更佳。

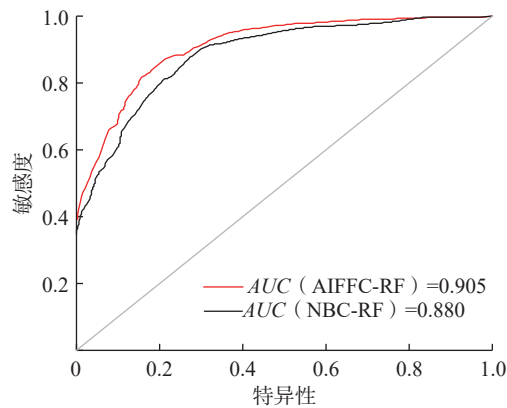
表 4 易发性评价结果灾积比统计表

Table 4 Statistical table of disaster accumulation ratio - product ratio for susceptibility evaluation results

评价模型	易发性分区	面积占比/%	灾积比	评价模型	易发性分区	面积占比/%	灾积比
NBC-MIV	高易发	13.62	0.611	AIFFC-MIV	高易发	11.73	0.955
	中易发	25.42	0.178		中易发	19.61	0.147
	低易发	60.96	0.050		低易发	68.76	0.026
NBC-RF	高易发	15.03	0.755	AIFFC-RF	高易发	9.68	1.318
	中易发	28.53	0.096		中易发	25.86	0.083
	低易发	56.44	0.031		低易发	64.45	0.013



(a) 信息量模型ROC曲线图



(b) 随机森林模型预测率ROC曲线图

图 5 各模型 ROC 曲线图

Fig. 5 ROC curve for various models

(3)AIFFC 分级方法在地质灾害易发性评价运用中,不仅能根据地质灾害数据特征与分布规律合理确定各因子的分级数,其分级区间划分的精度也具有一定的准确性。为今后确定易发性评价因子的合理分级数及分级方法提供了参考。

参考文献(References):

[1] 焦伟之,张明,谢鑫鹏,等.基于GIS与加权信息量模型的城镇地质灾害易发性评价——以大新镇为例[J].

安全与环境工程, 2022, 29(4): 119 – 128. [JIAO Weizhi, ZHANG Ming, XIE Xinpeng, et al. Susceptibility evaluation of urban geological disaster based on GIS and weighted information value model: A case study of Daxin Town [J] . Safety and Environmental Engineering, 2022, 29(4): 119 – 128. (in Chinese with English abstract)]

[2] 赵魁.基于ArcGIS的云安区地质灾害易发性分区评价 [J] .地质灾害与环境保护, 2020, 31(4): 38 – 42. [ZHAO Kui. The study of geological disaster susceptible division in Yun'an District based on ArcGIS [J] . Journal of Geological Hazards and Environment Preservation, 2020, 31(4): 38 – 42. (in Chinese with English abstract)]

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

[4] 刘福臻,王灵,肖东升.机器学习模型在滑坡易发性评价中的应用 [J] .中国地质灾害与防治学报, 2021, 32(6): 98 – 106. [LIU Fuzhen, WANG Ling, XIAO Dongsheng. Application of machine learning model in landslide susceptibility evaluation [J] . The Chinese Journal of Geological Hazard and Control, 2021, 32(6): 98 – 106. (in Chinese with English abstract)]

[5] 孙德亮,马祥龙,唐小娅,等.基于不同因子分级的滑坡易发性区划对比——以万州区为例 [J] .重庆师范大学学报(自然科学版), 2021, 38(5): 43 – 54. [SUN Deliang, MA Xianglong, TANG Xiaoya, et al. Comparison of landslide susceptibility mapping based on different factor classifications: Taking Wanzhou District as an example [J] . Journal of Chongqing Normal University (Natural Science), 2021, 38(5): 43 – 54. (in Chinese with English abstract)]

[6] 杨得虎,朱杰勇,刘帅,等.基于信息量、加权信息量与逻辑回归耦合模型的云南罗平县崩滑灾害易发性评价对比分析 [J] .中国地质灾害与防治学报, 2023, 34(5): 43 – 53. [YANG Dehu, ZHU Jieyong, LIU Shuai, et al. Comparative analyses of susceptibility assessment for landslide disasters based on information value, weighted information value and logistic regression coupled model in Luoping County, Yunnan Province [J] . The Chinese Journal of Geological Hazard and Control, 2023, 34(5): 43 – 53. (in Chinese with English abstract)]

[7] 解明礼,巨能攀,赵建军,等.区域地质灾害易发性分级方法对比分析研究 [J] .武汉大学学报(信息科学版), 2021, 46(7): 1003 – 1014. [XIE Mingli, JU Nengpan, ZHAO Jianjun, et al. Comparative analysis on classification methods of

- geological disaster susceptibility assessment [J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(7): 1003 – 1014. (in Chinese with English abstract)
- [8] 凌晓, 刘甲美, 王涛, 等. 基于致灾因子对称法分级的信息量模型在地震滑坡危险性评价中的应用 [J]. *国土资源遥感*, 2021, 33(2): 172 – 181. [LING Xiao, LIU Jiamei, WANG Tao, et al. Application of information value model based on symmetrical factors classification method in landslide hazard assessment [J]. *Remote Sensing for Land & Resources*, 2021, 33(2): 172 – 181. (in Chinese with English abstract)]
- [9] 郭建华, 刘初群, 刘翠. 基于遗传算法优化的城市标准循环工况构建 [J]. *科学技术与工程*, 2017, 17(15): 327 – 333. [GUO Jianhua, LIU Chuqun, LIU Cui. The construction of standard driving cycle based on genetic algorithm optimization [J]. *Science Technology and Engineering*, 2017, 17(15): 327 – 333. (in Chinese with English abstract)]
- [10] 陈绪钰, 倪化勇, 李明辉, 等. 基于加权信息量和迭代自组织聚类的地质灾害易发性评价 [J]. *灾害学*, 2021, 36(2): 71 – 78. [CHEN Xuyu, NI Huayong, LI Minghui, et al. Geo-hazard susceptibility evaluation based on weighted information value model and ISODATA cluster [J]. *Journal of Catastrophology*, 2021, 36(2): 71 – 78. (in Chinese with English abstract)]
- [11] 陈伟. 山区村镇滑坡灾害风险评估研究 [D]. 武汉: 武汉大学, 2019. [CHEN Wei. Study on landslide risk assessment in mountainous villages and towns [D]. Wuhan: Wuhan University, 2019. (in Chinese with English abstract)]
- [12] 孙娟娟, 江南, 刘纘武. 基于自适应膨胀因子的模糊覆盖分级方法 [J]. *测绘科学技术学报*, 2007, 24(5): 384 – 386. [SUN Juanjuan, JIANG Nan, LIU Zuanwu. Fuzzy cover approach for clustering based on adaptive inflation factor [J]. *Journal of Geomatics Science and Technology*, 2007, 24(5): 384 – 386. (in Chinese with English abstract)]
- [13] 孙娟娟. 专题地图数据分级模型的研究——现代数学在地图数据分级中的应用 [D]. 郑州: 解放军信息工程大学, 2007. [SUN Juanjuan. Research of classification model for thematic mapping data [D]. Zhengzhou: PLA Information Engineering University, 2007. (in Chinese with English abstract)]
- [14] 姚宇婕, 陈毓芬. 引导型专题数据分级处理研究 [J]. *测绘工程*, 2012, 21(1): 25 – 29. [YAO Yujie, CHEN Yufen. Research on guiding thematic data classification [J]. *Engineering of Surveying and Mapping*, 2012, 21(1): 25 – 29. (in Chinese with English abstract)]
- [15] 张涵斐. 多尺度地理信息显示处理与发布技术研究 [D]. 郑州: 解放军信息工程大学, 2012. [ZHANG Hanfei. Research on the display processing and publish technology of multi-scale geographic information [D]. Zhengzhou: PLA Information Engineering University, 2012. (in Chinese with English abstract)]
- [16] 牟廉明, 黄国兴. 一种基于自适应膨胀因子的聚类新方法 [J]. *计算机工程*, 2003, 29(9): 100 – 102. [MOU Lianming, HUANG Guoxing. New approach of clustering based on adaptive inflation factor [J]. *Computer Engineering*, 2003, 29(9): 100 – 102. (in Chinese with English abstract)]
- [17] 赖冠中, 陈文音. 基于自然断点法分析的城乡建设用地整理潜力分区研究——以汕头市濠江区为例 [J]. *广西城镇建设*, 2019(12): 123 – 127. [LAI Guanzhong, CHEN Wenyin. Study on potential zoning of urban and rural construction land consolidation based on natural breakpoint analysis: A case study of Haojiang District, Shantou City [J]. *Cities and Towns Construction in Guangxi*, 2019(12): 123 – 127. (in Chinese)]
- [18] 尹子虚. 基于夜光数据的城市空间特征提取尺度研究 [D]. 大连: 辽宁师范大学, 2021. [YIN Zixu. Study on the scale of urban spatial feature extraction based on nighttime light data [D]. Dalian: Liaoning Normal University, 2021. (in Chinese with English abstract)]
- [19] 孟晓捷, 张新社, 曾庆铭, 等. 基于加权信息量法的黄土滑坡易发性评价——以 1:5 万天水市麦积区为例 [J]. *西北地质*, 2022, 55(2): 249 – 259. [MENG Xiaojie, ZHANG Xinshe, ZENG Qingming, et al. The susceptibility evaluation of loess landslide based on weighted information value method: Taking 1:50 000 map of Maiji District of Tianshui City as an example [J]. *Northwestern Geology*, 2022, 55(2): 249 – 259. (in Chinese with English abstract)]
- [20] 孙滨, 祝传兵, 康晓波, 等. 基于信息量模型的云南东川泥石流易发性评价 [J]. *中国地质灾害与防治学报*, 2022, 33(5): 119 – 127. [SUN Bin, ZHU Chuanbing, KANG Xiaobo, et al. Susceptibility assessment of debris flows based on information model in Dongchuan, Yunnan Province [J]. *The Chinese Journal of Geological Hazard and Control*, 2022, 33(5): 119 – 127. (in Chinese with English abstract)]
- [21] 束龙仓, 黄蕾, 陈华伟, 等. 基于 AHP-EWM 的莱州市海岸带海水入侵灾害风险评价与区划 [J]. *吉林大学学报 (地球科学版)*, 2023, 53(6): 1864 – 1879. [SHU Longcang, HUANG Lei, CHEN Huawei, et al. Risk assessment and zoning of seawater intrusion hazard in coastal region of Laizhou city based on AHP-EWM method [J]. *Journal of Jilin University (Earth Science Edition)*, 2023, 53(6): 1864 – 1879. (in Chinese with English abstract)]
- [22] BREIMAN L. Bagging predictors [J]. *Machine Learning*, 1996, 24(2): 123 – 140.
- [23] 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. (in Chinese with English abstract)]
- [24] 温鑫, 范宣梅, 陈兰, 等. 基于信息量模型的地质灾害易发性评价——以川东南古蔺县为例 [J]. 地质科技通报, 2022, 41(2): 290 – 299. [WEN Xin, FAN Xuanmei, CHEN Lan, et al. Susceptibility assessment of geological disasters based on an information value model: A case of Gulin County in southeast Sichuan [J]. Bulletin of Geological Science and Technology, 2022, 41(2): 290 – 299. (in Chinese with English abstract)]
- [25] 李萍, 叶辉, 谈树成. 基于层次分析法的永德县地质灾害易发性评价 [J]. 水土保持研究, 2021, 28(5): 394 – 399. [LI Ping, YE Hui, TAN Shucheng. Evaluation of geological hazards in Yongde County based on analytic hierarchy process [J]. Research of Soil and Water Conservation, 2021, 28(5): 394 – 399. (in Chinese with English abstract)]
- [26] 贾雨霏, 魏文豪, 陈稳, 等. 基于 SOM-I-SVM 耦合模型的滑坡易发性评价 [J]. 水文地质工程地质, 2023, 50(3): 125 – 137. [JIA Yufei, WEI Wenhao, CHEN Wen, et al. Landslide susceptibility assessment based on the SOM-I-SVM model [J]. Hydrogeology & Engineering Geology, 2023, 50(3): 125 – 137. (in Chinese with English abstract)]
- [27] 范强, 巨能攀, 向喜琼, 等. 基于结果验证的信息量法地质灾害易发性评价——以贵州省开阳县为例 [J]. 人民长江, 2015, 46(15): 65 – 68. [FAN Qiang, JU Nengpan, XIANG Xiqiong, et al. Geohazard susceptibility assessment by using information quantity model with result validation: A case study in Kaiyang County, Guizhou Province [J]. Yangtze River, 2015, 46(15): 65 – 68. (in Chinese with English abstract)]
- [28] 陈前, 晏鄂川, 黄少平, 等. 基于样本与因子优化的黄冈南部地区地质灾害易发性评价 [J]. 地质科技通报, 2020, 39(2): 175 – 185. [CHEN Qian, YAN Echuan, HUANG Shaoping, et al. Susceptibility evaluation of geological disasters in southern Huanggang based on samples and factor optimization [J]. Bulletin of Geological Science and Technology, 2020, 39(2): 175 – 185. (in Chinese with English abstract)]
- [29] 刘月, 王宁涛, 周超, 等. 基于 ROC 曲线与确定性系数法集成模型的三峡库区奉节县滑坡易发性评价 [J]. 安全与环境工程, 2020, 27(4): 61 – 70. [LIU Yue, WANG Ningtao, ZHOU Chao, et al. Evaluation of landslide susceptibility based on ROC and certainty factor method in Fengjie County, Three Gorges Reservoir [J]. Safety and Environmental Engineering, 2020, 27(4): 61 – 70. (in Chinese with English abstract)]
- [30] WANG Liangjie, SAWADA K, MORIGUCHI S. Landslide susceptibility analysis with logistic regression model based on FCM sampling strategy [J]. Computers & Geosciences, 2013, 57: 81 – 92.