

DOI:10.13869/j.cnki.rswc.2024.06.028.

潘网生, 赵恬茵, 李鑫, 等. 基于多元自适应回归样条的黄土区滑坡敏感性评价[J]. 水土保持研究, 2024, 31(6): 271-280.

Pan Wangsheng, Zhao Tianyin, Li Xin, et al. Susceptibility Assessment of Landslide in the Loess Region Based on Multivariate Adaptive Regression Spline[J]. Research of Soil and Water Conservation, 2024, 31(6): 271-280.

基于多元自适应回归样条的黄土区滑坡敏感性评价

潘网生¹, 赵恬茵¹, 李鑫², 卢玉东³

(1. 黔南民族师范学院 旅游与资源环境学院, 贵州 都匀 558000;

2. 石家庄铁道大学 土木学院, 石家庄 050043; 3. 长安大学 水利与环境学院, 西安 710054)

摘要: [目的] 探讨和评价致灾因子对黄土滑坡敏感性区划的影响, 以辅助滑坡灾害的监测、预警和防治。[方法] 在因子筛选和独立性检验基础上, 基于多元自适应回归样条模型(MARS)自主分割数据, 正向拟合若干基函数, 同时构建了铰链函数在各基函数之间建立连接机制, 进而反向自我修正, 剔除或修改部分基函数, 实现自主选择致灾因子和给出因子权重, 最后引入了概率比(PR)模型做精度对比分析。[结果] (1) MARS 模型的基函数物理含义明确, 极大地降低了模型的复杂度, 易于 GIS 空间分析实现; (2) 评价因子的多重共线性和相关性检验有助于 MARS 模型优化; (3) MARS 模型对极高敏感区和高敏感区的划分较 PR 模型更为严谨、客观, 且滑坡敏感等级区划结果是所有参评因子综合作用的结果, 关键因子对评价结果的影响并不明显。[结论] MARS 模型在铜川市耀州区黄土滑坡敏感性评价研究中, ROC 曲线的 AUC 值为 0.879, 整体拟合效果好, 与野外实地勘察结果相吻合, 结果可靠。

关键词: 多元自适应回归样条; 黄土; 滑坡; 敏感性评价

中图分类号: P642.22

文献标识码: A

文章编号: 1005-3409(2024)06-0271-10

Susceptibility Assessment of Landslide in the Loess Region Based on Multivariate Adaptive Regression Spline

Pan Wangsheng¹, Zhao Tianyin¹, Li Xin², Lu Yudong³

(1. School of Tourism and Resources Environment, Qiannan Normal University for Nationalities,

Duyun, Guizhou 558000, China; 2. School of Civil Engineering, Shijiazhuang Tiedao University,

Shijiazhuang 050043, China; 3. School of Water and Environment, Chang'an University, Xi'an 710054, China)

Abstract: [Objective] The purpose of this study is to explore and evaluate the influence of disaster factors on the susceptibility zoning of Loess landslide, so as to assist the monitoring, warning and prevention of landslide disaster. [Methods] Based on factor screening and independence test, multiple adaptive regression spline model (MARS) was used to independently segment the data, to fit several basis functions, to construct hinge functions to establish a connection mechanism between the basis functions, and then to reverse self-correct, to eliminate or modify part of the basis functions to achieve independent selection of disaster factors and give factor weights. At last, probability ratio (PR) model was introduced to analyze the accuracy. [Results] (1) The physical meaning of the base function of MARS model is clear, which greatly reduces the complexity of the model and is easy to implement GIS spatial analysis. (2) The multicollinearity and correlation tests of assessment factors are helpful for MARS model optimization. (3) The classification of extremely sensitive and highly sensitive areas by MARS model is more rigorous and objective than that by PR

收稿日期: 2023-12-19

修回日期: 2024-02-23

资助项目: 黔南民族师范学院支持引进高层次人才研究专项项目(2021qnsyrc03); 贵州省自然科学基金重点项目(黔科合基础[2018]1416); 黄土高原土壤侵蚀与旱地农业国家重点实验室基金课题(A314021402-202113); 中央引导地方科技发展资金项目(236Z5405G); 河北省自然科学基金(E2021210092); 黔南民族师范学院 2024 年度重点研究科研能力提升项目(2024zdzk07)

第一作者: 潘网生(1975—), 男, 江苏镇江人, 博士, 副教授, 主要从事地质灾害与地质环境研究。E-mail: 450392404@qq.com

通信作者: 赵恬茵(1988—), 女, 四川巴中人, 博士, 副教授, 主要从事土壤侵蚀研究。E-mail: zhaotianyin1988@126.com

<http://stbcyj.paperonce.org>

model, and the result of landslide susceptibility classification is the result of the comprehensive effect of all the participating factors, and the influence of key factors on the assessment results is not obvious. [Conclusion] In the study on the susceptibility assessment of Loess landslide in Yaozhou District of Tongchuan City by MARS model, the AUC value of ROC curve is 0.879, the overall fitting effect is good, and the results are consistent with the field investigation results, and the results are reliable.

Keywords: multivariate adaptive regression spline; loess; landslide; susceptibility assessment

滑坡敏感性评价是以预测评价区域的地质环境条件为基础,构建致灾的敏感性评价因子体系,评估滑坡发生概率及其空间概率分布,并制作滑坡敏感性等级区划图,它是地质灾害危险性评价和风险性评价的重要基础,也是滑坡灾害预警和防灾减灾的必要前提^[1-2]。国外最早开展滑坡敏感性评价研究始于 1970s^[3],国内则稍晚,主要以惠振德等^[4]关于陕南山区滑坡研究一文为代表。滑坡敏感性研究主要集中在两个方向,一是评价因子的敏感性分析^[5],二是滑坡敏感性时空概率和区划研究。后者又分为局部地貌单元的滑坡敏感性评价^[6]和区域性滑坡敏感性评价^[7]。

目前,滑坡敏感性评价模型和方法较多。根据学者 Aghdam 等^[8]观点,滑坡敏感性评价模型可以划分为确定性和数据驱动两种类型。确定性模型即完全基于确定的函数关系,其在滑坡敏感性评价领域并不实际存在;而数据驱动模型是一种遵循建模—分析—反馈—调整—应用这一循环机制的应用模型,可以再细分为启发式模型(如层次分析法等)、常规数理统计模型(如逻辑回归模型、证据权重法、概率模型法等)和机器学习模型(如支持向量机、神经网络、随机森林等)^[9]。数据驱动模型在滑坡敏感性评价领域具有较好的评价效果,但因其评价因子权重、分级量化阈值等确定方面具有较强的人为主观性,给滑坡敏感性评价带来不确定性。为此,部分学者致力于多模型融合来试图减小这种不确定性,如苏刚等^[10]提出基于支持向量机、神经网络和随机森林的 Stacking 集成法,该方法思路主要来源于 Wolpert^[11]的多模型融合评价法。尽管此类模型取得了较好的评价效果,但因模型本身相对复杂,参数物理意义不明晰,以致在普及和推广方面受到一定的限制。

基于以上认识,本文引入多元自适应回归样条法(Multivariate Adaptive Regression Spline, MARS)开展相关研究。该模型相对简单、实用,且能充分考虑各变量之间的交互作用,对因变量与自变量之间函数关系进行显式表达,具有完善的自我修复、适应、调整和完善等功能,力求模型最大程度与目标现实相吻合,极具推广和应用价值^[12]。目前,该模型在地质灾

害领域的应用尚不多见,仅有赵忠国等^[13]开展了此类研究。其他可查文献主要是关于冻土空间分布研究^[14]、作物蒸散量模拟^[12]、矿井温度预测^[15]、电网输出功率预测^[16]、地下水资源动态变化模拟^[17]及岩土工程中多维非线性问题模拟等^[18]。本文以陕西省铜川市耀州区为例,运用 MARS 开展区域黄土滑坡敏感性评价研究,以期对黄土滑坡致灾因子敏感性评价和敏感性等级区划研究做一些有益探索和尝试,并辅助滑坡灾害的监测、预警和防治。

1 研究区与数据源

1.1 研究区概况

铜川市位于陕西省中部,地处鄂尔多斯台地与渭河断陷盆地的过渡地带,下辖耀州区、王益区、印台区和宜君县。耀州区地理位置在铜川市西南,行政区划面积 1 608.4 km²,区内梁峁相间,沟壑纵横,地势西北高东南低,中部多丘陵沟壑,南部较为平缓(图 1)。最低海拔约 550 m,最高海拔约 1 730 m,相对高差 1 180 m,区域地质构造为渭河断陷盆地的一部分,堆积有厚度较大的松散沉积层,表层为黄土覆盖。耀州区因其丰富的煤炭资源及近 60 年的开采历史,地质环境遭受严重破坏,生态极为脆弱^[19-20]。

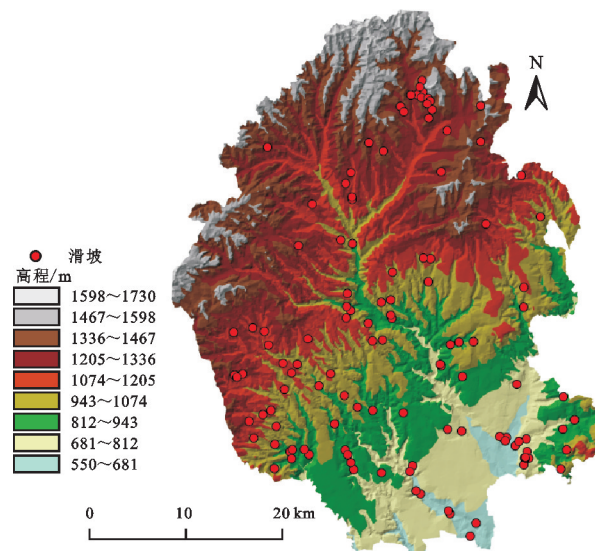


图 1 研究区地貌概况与滑坡灾害分布

Fig. 1 Geomorphic situation and landslide disaster distribution in the study area

1.2 数据源

1.2.1 滑坡编录数据 根据中国地质调查局项目“陕西省铜川市耀州区地质灾害详细调查报告”,研究区内分布典型滑坡 120 处,主要为黄土滑坡和堆积层滑坡^[21]。滑坡数据的采集和验证基于野外调查和 Bigmap 遥感影像。本文按照 1:1 比例采集非滑坡点数据用于 MARS 模型精度验证,120 个非滑坡点数据基于双缓冲区采样法获取^[22],即以滑坡点为中心,设内外两个缓冲区,内缓冲区半径为 1 km,外缓冲区半径为 10 km,非滑坡点数据在内外缓冲区之间采样产生,采样遵循比率模型低概率原则。

1.2.2 因子选取与分析 本文评价因子主要考虑地形地貌、气象水文、地质构造、地质岩组、植被覆盖、人类工程活动等。根据研究需要,上述基本因子可以进一步细化分解,如地形地貌因子可细分出坡度、坡向、地表剖面曲率、高程变异系数、地表粗糙度、地表切割深度等。本文地形地貌数据源于数字高程模型(DEM),DEM 栅格尺寸为 25 m×25 m,将研究区划分成 2 578 304 个栅格。基于 DEM 派生出坡度(°)、坡向(°)、剖面曲面、高程变异系数、沟壑密度(km/km²)、地表切割深度(m)等。以上地形地貌因子图层处理见图 2A—F。

气象水文因子细分为降水分布和水力侵蚀系数。研究区降水量较小,且降水空间分布并不均匀(图 2G),长期地表径流会侵蚀黄土地表,致使斜坡应力平衡被打破,进而引发黄土滑坡,本文采用水力侵蚀系数反映该影响。水力侵蚀系数经由 DEM 水文分析生成汇流累积量,并做归一化处理,然后乘以坡度,水力侵蚀系数因子见图 2H。岩层产状大致呈水平分布,岩土类型及其空间分布见图 2I。植被因子由归一化植被指数(Normalized Difference Vegetation Index, NDVI)体现(图 2J)。地质构造因子采用地层断裂数据,断裂构造空间分布及缓冲区见图 2K。人类工程活动以公路网线密度体现,单位:km/km²(图 2L)。

2 方法与模型

2.1 因子独立性检验

2.1.1 皮尔逊相关系数检验 皮尔逊相关系数法是检验两组数据是否相关的简单、实用工具。其数学公式详见纪德洋等^[23]研究成果。皮尔逊相关系数 ρ_{XY} 介于-1~1,越接近于 1,表明 X 和 Y 之间的相关性就越强。

2.1.2 多重共线性检验 多重共线性是关于两个或两个以上变量之间是否存在高度相关性以及如何影响因变量的问题^[24],如果存在高度相关性,且难以区分其对因变量的影响和作用,则将这种高度相关性所

产生的后果定义为多重共线性^[25]。自变量之间的高度相关性不一定导致共线性,但共线性必然说明自变量高度相关^[24]。因此,本文采用容忍度(tolerance, TOL)和方差膨胀因子(variance inflation factor, VIF)对滑坡敏感性因子做多重共线性检验,如果 $VIF > 10$ 或 $TOL < 0.1$,则表明滑坡敏感性因子之间存在多重共线性。

2.2 多元自适应回归样条(MARS)模型

MARS 最早由美国斯坦福大学的 Friedman^[26]提出,MARS 建模思想是在分段节点 t 处将整体数据划分为若干区块,分块建立线性回归拟合,由此创建若干个基函数(Basis Function, BF),从而实现整体数据域的非线性拟合。分段节点 t 两侧的样条曲线,属于截断型幂函数,可用式(1)和式(2)表示^[12,26-27]。

$$BF(t)_+ = \begin{cases} (x-t)^q & x > t \\ 0 & \text{其他} \end{cases} \quad (1)$$

$$BF(t)_- = \begin{cases} (x-t)^q & x < t \\ 0 & \text{其他} \end{cases} \quad (2)$$

式中: $BF(t)_+$ 为分段节点 t 右侧基函数; $BF(t)_-$ 为分段节点 t 左侧基函数; x 为基函数评价因子; q 为分段节点 t 两侧样条函数的幂次,当 $q=1$ 时,即为一阶基函数。

一阶 MARS 模型表示如下:

$$y = a_0 + \sum_{i=1}^n k_i \cdot BF_i(t) \quad (3)$$

式中: y 为模型预测值; a_0 为截距; $BF_i(t)$ 为第 i 段基函数; k_i 为第 i 段基函数的系数; n 为基函数的总段数。

该模型的创建主要有两个步骤,第一步,正向拟合:首先,对样本数据进行分割并分段拟合创建基函数,其次,以常数为基础项,逐步添加基函数,直至基函数数量达到约定的最大限值,仅需保证拟合残差的平方和最小,此时 1 个过度拟合的全局函数即被创建^[12];第二步,反向修正:即对第一步产生的过度拟合的全局函数做剪枝修正处理,通过广义交叉验证(Generalized Cross Validation, GCV)和预测性试验,找出贡献度最小的基函数,做删除或调整,使其拟合效果最佳。

广义交叉验证(GCV),其数学表达式^[12,16,26]如下:

$$GCV = \frac{\sum_{i=1}^n (y_i^* - y_i)^2}{n \left[1 - \frac{T(M)}{n} \right]^2} \quad (4)$$

式中: n 为参与拟合的样本总数,如果预设其最大值,即可作为模型正向拟合的终止条件; $T(M)$ 为模型复杂度的惩罚函数; y_i 为模型原始预测值; y_i^* 为剪枝修正之后模型预测值。

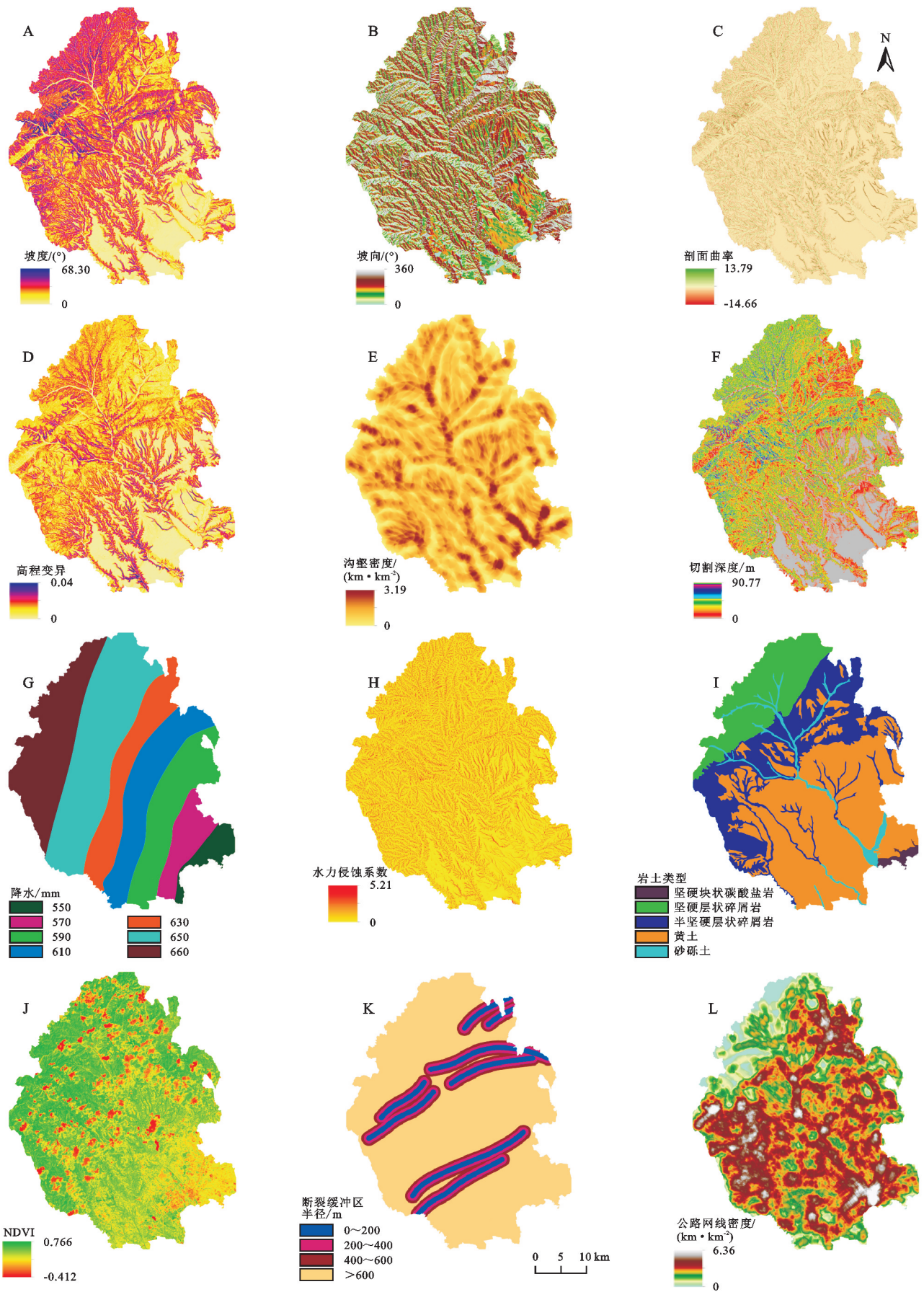


图 2 评价因子数据

Fig. 2 Data of assessment index

其中:

$$T(M) = M + 1 + b \cdot \frac{M}{2} \quad (5)$$

式中: M 为基函数总数; b 为惩罚系数, 通常介于 $2 \sim 4$ ^[16]。模型精度越高, GCV 值越小^[12]。

MARS 模型任意两个基函数由铰链函数 $h_i(x)$ 进行光滑链接, 且拟合建模前不需要对因变量和自变量做任何假设^[12], 这就降低了模型的复杂度。正向拟合和反向修正完全自主完成, 且预测效果良好。此外, 还可以用于缺失数据的插值拟合。当然, MARS 也有两大缺陷, 一是不能实现连续分割数据块, 二是尚不能自动识别如线性的、累积的或交互的这种简单又低阶的关系^[26]。尽管 MARS 不尽完美, 但这并不影响我们对其使用, 仅需消耗一些额外的时间成本即可克服其第一个缺陷, 基于程序的 MARS 函数高效处理技术及高素质的人为干预也可以很好弥补其第二个缺陷。本文基于 R 语言 earth 程序包, 高效实现陕西省铜川市耀州区的黄土滑坡敏感性评价研究。相关程序包资源可通过 <https://cran.r-project.org> 下载。

2.3 概率比 (PR) 模型

PR 模型是一种基于统计学原理、简单且实用的数学方法, 其与信息量模型同宗同源, 评价精度也不亚于逻辑回归模型^[28], 在不同比例尺的滑坡敏感性分析评价实例中已经取得比较理想的效果, 其结果更具有科学性和客观性^[13-14]。因此, PR 模型可以用来对 MARS 模型的评价结果做比较验证。PR 模型的评价结果采用滑坡敏感性评价指数来衡量, 区域敏感

性评价结果由各评价因子的概率比求和得到, 具体公式参考 Bharadwaj 等^[28]的成果。

2.4 模型精度评价

MARS 模型和 PR 模型的评价精度均可以由受试者特征曲线 (Receiver Operating Characteristic, ROC) 检验, ROC 曲线是一种简单、直观且准确的模型优劣评价方法^[29], 近年来在地质灾害评价领域得到推广和应用^[13]。ROC 曲线的线下面积 (Area Under Curve, AUC) 可以对 MARS 模型精度做深度评价, AUC 值越大, 表明模型精度越高, 一般要求该值大于 0.8。

此外, 可以通过绘制滑坡密度分布图对模型预测精度进行验证。该方法可以直观反映不同滑坡敏感性等级内的滑坡密度差异, 通常以实际滑坡是否尽可能多地分布于较高敏感区且较高敏感区的面积占比是否尽可能小作为该模型精度优劣的判断标准^[30]。

3 结果与分析

3.1 因子独立性分析

3.1.1 因子相关性 皮尔逊相关系数分析基于 SPSS.25。分析结果显示, 坡度、坡向、断裂、公路网线密度、剖面曲率、水力侵蚀系数、沟壑密度、岩组、植被归一化指数 (NDVI)、降水 10 个因子满足相关性要求 (表 1)。坡度与公路网线密度高度相关可能源于一个事实, 即地形坡度越大, 环山公路线密度就越大。由于坡度属自然因素, 公路网线密度属人为因子, 二者之间不具有相互替代性, 故不做因子剔除处理。

表 1 因子相关性

Table 1 Correlation of factors

因子	坡度	坡向	断裂	公路网线密度	剖面曲率	水力侵蚀系数	沟壑密度	岩组	NDVI	降水量
坡度	1.000	0.057	-0.091	0.832 **	0.162	-0.002	-0.177	0.059	0.206 *	0.368 **
坡向	0.057	1.000	-0.119	-0.018	0.030	0.010	-0.028	0.023	0.140	0.044
断裂	-0.091	-0.119	1.000	-0.070	-0.160	-0.019	-0.004	-0.143	-0.094	-0.134
公路网线密度	0.832 **	-0.018	-0.070	1.000	0.160	0.039	-0.119	0.095	0.163	0.456 **
剖面曲率	0.162	0.030	-0.160	0.160	1.000	-0.067	-0.036	-0.089	-0.075	-0.064
水力侵蚀系数	-0.002	0.010	-0.019	0.039	-0.067	1.000	-0.036	-0.051	0.038	-0.073
沟壑密度	-0.177	-0.028	-0.004	-0.119	-0.036	-0.036	1.000	-0.092	-0.284 **	-0.259 **
岩组	0.059	0.023	-0.143	0.095	-0.089	-0.051	-0.092	1.000	0.049	0.250 **
NDVI	0.206 *	0.140	-0.094	0.163	-0.075	0.038	-0.284 **	0.049	1.000	0.394 **
降水量	0.368 **	0.044	-0.134	0.456 **	-0.064	-0.073	-0.259 **	0.250 **	0.394 **	1.000

注: ** 代表在 0.01 级别 (双尾), 相关性显著; * 代表在 0.05 级别 (双尾), 相关性显著。

3.1.2 因子多重共线性 因子多重共线性检测也基于 SPSS.25, 将满足 $TOL < 0.1$ 或 $VIF > 10$ 的高程变异系数、地表粗糙度及地表切割深度 3 个因子剔除, 符合要求的评价因子见表 2。因此, 坡度、坡向、断裂、公路网线密度、剖面曲率、水力侵蚀系数、沟壑密度、岩组、植被归一化指数 (NDVI)、降水 10 个因子

将作为 MARS 模型和 PR 模型的参选因子。

3.2 MARS 模型应用

评价因子多重共线性和相关性检验有助于 MARS 模型的优化。120 个滑坡点的 MARS 模拟研究结果表明, degree 取值大小会影响广义交叉验证指数 GCV 和模型拟合决定系数 Rsq , degree = 2 时, GCV 值较

degree=1时降低约 6%，而模型拟合精度则提高约 1%。模型正向拟合阶段，整体数据被分割为 21 个区块，即产生 21 个初始基函数；后向修正阶段，经剪枝处理，有 14 个关键基函数(含截距)被保留，其中涉及 5 个重要的评

价因子，分别为坡度、公路网线密度、沟壑密度、降水量和剖面曲率(表 3)。此时，GCV 为 0.847，拟合决定系数 Rsq 为 0.918。根据陈宣全等^[12]观点，GCV 值减小，表明 MARS 模型精确度提高。

表 2 因子多重共线性

Table 2 Multicollinearity of factors

因子	坡度	坡向	断裂	公路网线密度	剖面曲率	水力侵蚀系数	沟壑密度	岩组	NDVI	降水量
VOL	0.291	0.967	0.917	0.264	0.899	0.965	0.899	0.911	0.773	0.673
VIF	3.434	1.034	1.091	3.783	1.113	1.036	1.113	1.098	1.294	1.485

表 3 MARS 模型基函数及系数

Table 3 Basis function and coefficients of MARS model

序号	基函数(BF)	系数
BF ₀	(截距)	12.87644
BF ₁	BF(17.5597—坡度)	0.09242
BF ₂	BF(坡度—17.5597)	0.32138
BF ₃	BF(坡度—24.554)	2.15075
BF ₄	BF(坡度—26.1967)	—15.43513
BF ₅	BF(0.204905—公路网线密度)	—5.73558
BF ₆	BF(沟壑密度—0.539335)	1.85861
BF ₇	BF(590—降水量)	0.05311
BF ₈	BF(坡度—26.1967)×降水量	0.01995
BF ₉	BF(坡度—17.5597)×BF(沟壑密度—0.726538)	—0.29793
BF ₁₀	BF(坡度—17.5597)×BF(0.726538—沟壑密度)	—0.95251
BF ₁₁	BF[剖面曲率—(—0.0795057)]×BF(沟壑密度—0.539335)	—1.04403
BF ₁₂	BF(沟壑密度—0.539335)×BF(降水量—650)	—0.21206
BF ₁₃	BF(沟壑密度—0.539335)×BF(650—降水量)	—0.01109

基于 GIS 技术对 MARS 模型的基函数进行空间运算和敏感等级区划，得到 MARS 模型的滑坡敏感性评价结果(图 3)。采用自然断点法将滑坡敏感区划分为极高敏感、高敏感、中敏感、低敏感和极低敏感 5 个等级。结果显示：极高敏感区面积 322.60 km²，面积占比 20.07%，主要分布于沟壑密度较大的黄土区和半坚硬层状碎屑岩区域；高敏感区面积 106.82 km²，面积占比 6.64%，集中分布于人类工程活动强度较大或沟壑密度较大的黄土区；中敏感区面积 783.50 km²，面积占比 48.74%，主要分布于沟壑密度相对较低的黄土区及河谷砂砾土区，河谷砂砾土虽结构松散，但因地处河谷平缓地势，滑坡敏感性相对较低；低敏感区面积 213.28 km²，面积占比 13.27%和极低敏感区面积 181.29 km²，面积占比 11.28%，主要分布于部分坚硬块状碳酸盐和坚硬层状碎屑岩区域，该区域受稳定岩土结构控制。

3.3 PR 模型比较研究

PR 模型的评价因子选择与 MARS 模型保持一致，基于该模型的滑坡敏感性评价结果见图 4。为区别各细分因子对滑坡灾害的影响，PR 模型对各因子做分级处理，分别统计各分级因子栅格数占比和各细分因子的滑

坡栅格占比，进而计算各细分因子的概率比值(表 4)。该模型的滑坡敏感性区划方法同 MARS 模型，极低敏感区面积 71.30 km²，面积占比 4.45%，主要分布于坚硬层状碎屑岩区域；低敏感区面积 339.25 km²，面积占比 21.19%，主要分布于半坚硬层状碎屑岩及黄土地势相对平缓的区域；中敏感区面积 500.29 km²，面积占比 31.25%，主要分布于半坚硬层状碎屑岩区的沟壑密度较大和黄土区的沟壑密度较小的区域；高敏感区面积 462.09 km²，面积占比 28.86%，集中分布于黄土区沟壑密度较大的区域；极高敏感区面积 228.22 km²，面积占比 14.25%，主要分布于坚硬块状碳酸盐区、人类工程活动强烈或沟壑密度较大的黄土区。

MARS 模型对极高敏感区和高敏感区的划分较 PR 模型更为严谨、客观，极低敏感区、低敏感区和中敏感区的划分在空间上更为粗放。滑坡敏感性区划结果是所有参评因子综合作用的结果，关键因子对评价结果的影响并不明显。尽管上述两个模型的评价结果存在一定的差异，但与我们对黄土滑坡灾害发育的一般性规律认识基本一致，也符合我们的预期，结果可信度较高。但哪一个模型预测精度更高，尚需引入精度模型做深入分析。

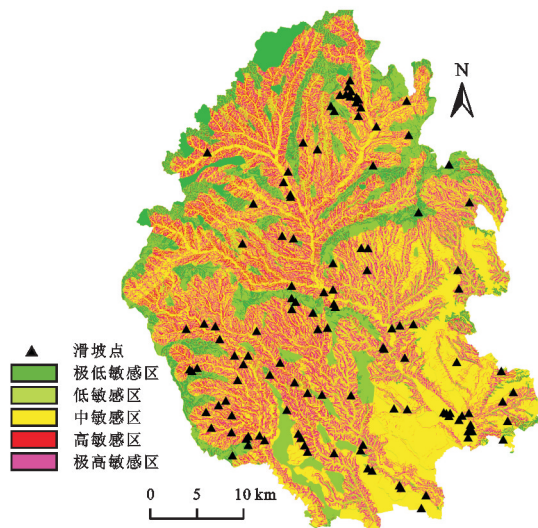


图 3 MARS 滑坡敏感性区划

Fig. 3 Susceptibility zoning of landslide based on MARS

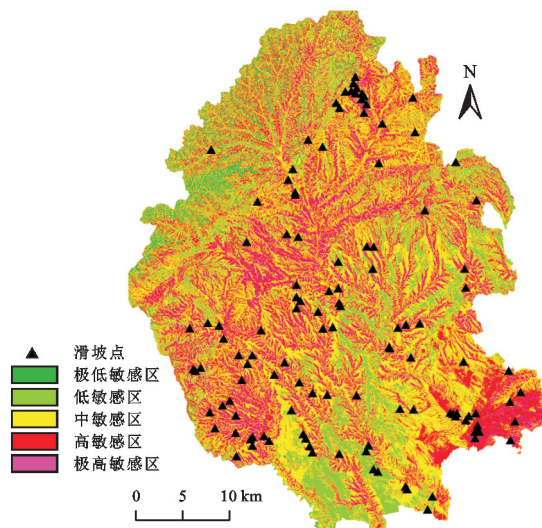


图 4 PR 滑坡敏感性区划

Fig. 4 Susceptibility zoning of landslide based on PR

表 4 PR 模型因子分级与概率比值计算

Table 4 Factor classification of PR model and probability ratio calculation

因子	因子分级	滑坡栅格数/个	滑坡栅格数占比/%	分级因子栅格数/个	分级因子栅格数占比/%	概率比值
坡度/(°)	[0~10]	51	0.425	933528	0.362	1.174
	(10~20]	30	0.250	684488	0.265	0.942
	(20~30]	34	0.283	671747	0.260	1.088
	(30~40]	5	0.042	251947	0.098	0.427
	(40~50]	0	0.000	35459	0.014	0.000
	(50~90]	0	0.000	2405	0.001	0.000
坡向/(°)	(315~45]	29	0.242	568775	0.220	1.096
	(45~135]	34	0.283	738040	0.286	0.990
	(135~225]	32	0.267	730225	0.283	0.942
	(225~315]	25	0.208	542534	0.210	0.991
断裂缓冲区半径/m	[0~200]	10	0.083	215448	0.084	0.987
	(200~400]	12	0.100	218385	0.085	1.173
	(400~600]	10	0.083	202181	0.079	1.052
	(600~+∞)	88	0.733	1925833	0.752	0.975
公路网线密度/ (km·km ⁻²)	[0~1]	3	0.025	274544	0.106	0.235
	(1~2]	24	0.200	774769	0.300	0.666
	(2~3]	58	0.483	1063339	0.412	1.173
	(3~+∞)	35	0.292	466408	0.181	1.613
剖面曲率	(-∞~-1]	4	0.033	192087	0.074	0.448
	(-1~0]	66	0.550	1244516	0.482	1.140
	(0~1]	43	0.358	929045	0.360	0.995
	(1~+∞)	7	0.058	213926	0.083	0.703
水力侵蚀系数	[0~5]	42	0.350	926689	0.359	0.974
	(5~10]	16	0.133	316777	0.123	1.086
	(10~15]	9	0.075	206449	0.080	0.937
	(15~+∞)	53	0.442	1129659	0.438	1.009

续表 4						
因子	因子分级	滑坡栅格数/个	滑坡栅格数占比/%	分级因子栅格数/个	分级因子栅格数占比/%	概率比值
沟壑密度/ ($\text{km} \cdot \text{km}^{-2}$)	[0~0.5]	14	0.117	563566	0.218	0.534
	(0.5~1]	34	0.283	990753	0.384	0.738
	(1~1.5]	51	0.425	748359	0.290	1.465
	(1.5~+∞)	21	0.175	268301	0.104	1.683
	坚硬块状碳酸盐	3	0.025	27427	0.011	2.351
岩组	坚硬层状碎屑岩	2	0.017	436592	0.169	0.098
	半坚硬层状碎屑岩	40	0.333	697038	0.270	1.234
	黄土	73	0.608	1306284	0.506	1.201
	砂砾土	2	0.017	113106	0.044	0.380
NDVI	[0~0.2]	5	0.042	178907	0.069	0.601
	(0.2~0.3]	17	0.142	277325	0.108	1.318
	(0.3~0.4]	29	0.242	598036	0.232	1.042
	(0.4~+∞)	69	0.575	1525303	0.591	0.972
降水量/mm	550	11	0.092	91326	0.035	2.589
	570	13	0.108	166201	0.064	1.681
	590	15	0.125	367596	0.143	0.877
	610	7	0.058	380752	0.148	0.395
	630	20	0.167	334691	0.130	1.285
	650	47	0.392	737340	0.286	1.370
	660	7	0.058	502243	0.195	0.300

3.4 模型精度分析

根据 ROC 曲线求得 MARS 模型和 PR 模型的 AUC 值分别为 0.879,0.875,表明 MARS 模型预测精度并不低于 PR 模型。滑坡密度统计结果(图 5)表明,MARS 模型和 PR 模型的评价结果均满足较高敏感等级所占面积尽可能小这一标准。PR 模型的滑坡密度随敏感等级提高而增加,而 MARS 模型则存在较大波动,在高敏感区达到极值,是 PR 模型的 4 倍,在极高敏感区呈现回落。总体而言,MARS 模型对高敏感区的预测较 PR 模型更优,表明该模型值得在黄土区做进一步探索研究。

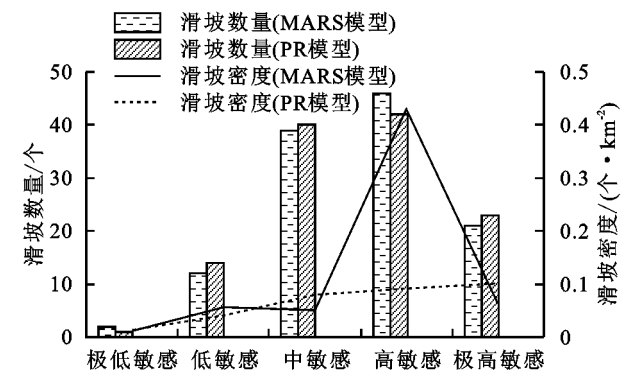


图 5 不同模型滑坡密度分布

Fig. 5 Landslide density distribution map with different models

4 讨论与结论

4.1 讨论

(1) 在理论层面,MARS 模型的广义交叉验证 GCV 随 degree 值增加而减小,拟合决定系数 Rsq 值随 degree 值增加而增加,但基于 degree 不同取值的实际预测效果(如 ROC 曲线的 AUC 值)的比较研究尚未开展,也即 MARS 模型的精度评价理论在实践层面是否具有正确的指示作用,仍有待实践检验。

(2) 虽然 MARS 模型对高敏感区的划分较 PR 模型更为严谨、客观,且符合我们的预期,结果的可信度也较高,但该模型是否适用于非黄土覆盖区,尚需结合岩土类型、强度及其他敏感性因子做进一步探讨和分析。

4.2 结论

(1) MARS 模型实现自主分割数据,正向拟合若干基函数,同时构建铰链函数在基函数之间建立连接机制,并反向做自我修正,剔除或修改部分基函数,使模型拟合效果最佳。MARS 模型的这种正向拟合与反向修正完全自主完成,极大地降低了模型的复杂度,且基函数物理含义明确,基于 GIS 的空间分析易于实现。其在铜川市耀州区黄土滑坡敏感性评价研究中,整体拟合效果好,与野外实地勘察结果相吻合,

评价结果可信度高。

(2) 评价因子的多重共线性和相关性检验有助于 MARS 模型优化。MARS 模型对极高敏感区和高敏感区的划分较 PR 模型更为严谨、客观。滑坡敏感等级区划结果是所有参评因子综合作用的结果,关键因子对评价结果的影响并不明显。

(3) MARS 模型在铜川市耀州区黄土滑坡敏感性评价研究中,ROC 曲线的 AUC 值为 0.879,整体拟合效果好,与野外实地勘察结果相吻合,结果可靠。极高敏感区主要分布于沟壑密度较大的黄土区和半坚硬层状碎屑岩区域;高敏感区集中分布于人类工程活动强度较大或沟壑密度较大的黄土区;中敏感区主要分布于沟壑密度相对较低的黄土区及河谷砂砾土区,河谷砂砾土虽结构松散,但因地处河谷平缓地势,滑坡敏感性相对较低;低敏感区和极低敏感区则分布于部分坚硬块状碳酸盐和坚硬层状碎屑岩区域,该区域受稳定岩土结构控制。

参考文献(References):

- [1] 段宇英,汤军,刘远刚,等.基于随机森林的山西省柳林县黄土滑坡空间敏感性评价[J].地理科学,2022,42(2):343-351.
Duan Y Y, Tang J, Liu Y G, et al. Spatial sensitivity evaluation of loess landslide in Liulin County, Shanxi based on random forest[J]. Scientia Geographica Sinica, 2022,42(2):343-351.
- [2] 贺倩,汪明,刘凯.基于 Logistic 回归和 MCMC 方法评价地震滑坡敏感性[J].水土保持研究,2022,29(3):396-403,410.
He Q, Wang M, Liu K. Assessment on earthquake-triggered landslide susceptibility based on logistic regression and MCMC method[J]. Research of Soil and Water Conservation, 2022,29(3):396-403,410.
- [3] Fell R, Corominas J, Bonnard C, et al. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning[J]. Engineering Geology, 2008,102(3/4):85-98.
- [4] 惠振德,李晓玲,孙虎.陕南山区滑坡发育特征与敏感性分析[J].陕西师范大学学报:自然科学版,1996,24(1):96-101.
Xi Z D, Li X L, Sun H. Characteristics and sensitivity of landslide in mountainous areas of Southern Shaanxi Province[J]. Journal of Shaanxi Normal University: Natural Science Edition, 1996,24(1):96-101.
- [5] 黄健,贺子城,黄祥,等.基于地貌特征的滑坡堰塞坝形成敏感性研究[J].地质科技通报,2021,40(5):253-262.
Huang J, He Z C, Huang X, et al. Formation sensitivity of landslide dam based on geomorphic characteristics[J]. Bulletin of Geological Science and Technology, 2021,40(5):253-262.
- [6] 唐川,马国超.基于地貌单元的小区域地质灾害易发性分区方法研究[J].地理科学,2015,35(1):91-98.
Tang C, Ma G C. Small regional geohazards susceptibility mapping based on geomorphic unit[J]. Scientia Geographica Sinica, 2015,35(1):91-98.
- [7] 尹超,李仲波,张凯,等.基于 Inter. iamb-Tabu 算法的区域滑坡敏感性评价[J].山地学报,2023,41(4):608-620.
Yin C, Li Z B, Zhang K, et al. Regional landslide susceptibility assessment based on Inter. iamb-Tabu algorithm[J]. Mountain Research, 2023,41(4):608-620.
- [8] Aghdam I N, Pradhan B, Panahi 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.
- [9] 黄发明,叶舟,姚池,等.滑坡易发性预测不确定性:环境因子不同属性区间划分和不同数据驱动模型的影响[J].地球科学,2020,45(12):4535-4549.
Huang F M, Ye Z, Yao C, et al. Uncertainties of landslide susceptibility prediction: Different attribute interval divisions of environmental factors and different data-based models[J]. Earth Science, 2020,45(12):4535-4549.
- [10] 苏刚,秦胜伍,乔双双,等.基于 Stacking 集成学习的泥石流易发性评价:以四川省雅江县为例[J].世界地质,2021,40(1):175-184.
Su G, Qin S W, Qiao S S, et al. Debris flow susceptibility evaluation based on Stacking ensemble learning: A case study in Yajiang, Sichuan Province[J]. Global Geology, 2021,40(1):175-184.
- [11] Wolpert D H. Stacked generalization[J]. Neural Networks, 1992,5(2):241-259.
- [12] 陈宣全,崔宁博,李继平,等.多元自适应回归样条算法模拟川中丘陵区参考作物蒸散量[J].农业工程学报,2019,35(16):152-160.
Chen X Q, Cui N B, Li J P, et al. Simulation of reference crop evapotranspiration in hilly area of central Sichuan based on MARS [J]. Transactions of the Chinese Society of Agricultural Engineering, 2019,35(16):152-160.
- [13] 赵忠国,张峰,郑江华.多元自适应回归样条法的滑坡敏感性评价[J].武汉大学学报:信息科学版,2021,46(3):442-450.
Zhao Z G, Zhang F, Zheng J H. Evaluation of landslide susceptibility by multiple adaptive regression spline method[J]. Geomatics and Information Science of Wuhan University, 2021,46(3):442-450.

- [14] 张秀敏,南卓铜,吴吉春,等.基于多元自适应回归样条的青藏高原温泉区域的冻土分布制图[J].冰川冻土, 2011,33(5):1088-1097.
- Zhang X M, Nan Z T, Wu J C, et al. Modeling permafrost distribution in Wenquan Area over Qinghai-Tibet Plateau by using multivariate adaptive regression splines[J]. Journal of Glaciology and Geocryology, 2011,33(5):1088-1097.
- [15] Krzemień A. Fire risk prevention in underground coal gasification (UCG) within active mines: Temperature forecast by means of MARS models[J]. Energy, 2019, 170:777-790.
- [16] 鲍长庚,闫贻鹏,黄一楠,等.基于多元自适应回归样条的光伏并网系统日输出功率预测[J].电力系统保护与控制, 2021,49(5):124-131.
- Bao C G, Yan Y P, Huang Y N, et al. Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines[J]. Power System Protection and Control, 2021,49(5):124-131.
- [17] Rezaie-balf M, Naganna S R, Ghaemi A, et al. Wavelet coupled MARS and M5 Model Tree approaches for groundwater level forecasting[J]. Journal of Hydrology, 2017,553:356-373.
- [18] 仇文岗,洪利,黎泳钦.基于多元自适应回归样条的高维岩土工程问题分析[J].河海大学学报:自然科学版, 2019,47(4):359-365.
- Zhang W G, Hong L, Li Y Q. Analysis of multi-dimensional geotechnical engineering problems based on multivariate adaptive regression splines[J]. Journal of Hohai University: Natural Sciences, 2019,47(4):359-365.
- [19] 穆柯,谢婉丽,刘琦琦,等.基于 LR-RF 模型的滑坡易发性评价:以铜川市耀州区为例[J].灾害学, 2022,37(3):212-218.
- Mu K, Xie W L, Liu Q Q, et al. Research on landslide susceptibility evaluation based on logistic regression and LR coupling model[J]. Journal of Catastrophology, 2022,37(3):212-218.
- [20] 延延,杜荣军,严芳.空间分析在铜川市耀州区地质灾害易发性分区中的应用[J].水土保持通报, 2010,30(1):146-150.
- Yan Y, Du R J, Yan F. Application of spatial analysis in susceptibility regionalization of geological disaster in Yaozhou District of Tongchuan City[J]. Bulletin of Soil and Water Conservation, 2010,30(1):146-150.
- [21] 孟文强.基于 GIS 和层次分析法的地质灾害危险性评价研究:以铜川市耀州区为例[D].西安:长安大学, 2012.
- Meng W Q. Study on Geo-Hazards Risk Evaluation Based on GIS and AHP: Taking Yaozhou District of Tongchuan City for example[D]. Xi'an: Chang'an University, 2012.
- [22] Zhu J, Baise L G, Thompson E M. An updated geospatial liquefaction model for global application[J]. Bulletin of the Seismological Society of America, 2017,107(3):1365-1385.
- [23] 纪德洋,金锋,冬雷,等.基于皮尔逊相关系数的光伏电站数据修复[J].中国电机工程学报, 2022,42(4):1514-1523.
- Ji D Y, Jin F, Dong L, et al. Data repairing of photovoltaic power plant based on Pearson correlation coefficient[J]. Proceedings of the CSEE, 2022,42(4):1514-1523.
- [24] 肖霞,伍兴国.线性回归中多重共线性的几何解释[J].统计与决策, 2021,37(21):46-51.
- Xiao X, Wu X G. Geometric interpretation of multicollinearity in linear regression[J]. Statistics & Decision, 2021,37(21):46-51.
- [25] Dar I S, Chand S. Improved heteroscedasticity-consistent ridge estimators for linear regression with multicollinearity[J]. Iranian Journal of Science, 2023,47:1593-1604.
- [26] Friedman J H. Multivariate adaptive regression splines[J]. Annals of Statistics, 1991,19(1):1-67.
- [27] 王洪礼,向建平,葛根.多元自适应样条回归预报浮游植物总量分析[J].海洋技术, 2006,25(3):7-9.
- Wang H L, Xiang J P, Ge G. The application of multivariate adaptive regression splines in forecast of total number of phytoplankton [J]. Ocean Technology, 2006,25(3):7-9.
- [28] Bharadwaj D, Sarkar R. Landslide susceptibility mapping using probabilistic frequency ratio and Shannon entropy for Chamoli, Uttarakhand Himalayas[J]. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 2024,48:377-395.
- [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 Y, Wang N T, Zhou C, 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.
- [30] 张帆宇,刘高,谌文武,等.基于多变量统计分析的大型滑坡敏感性评价:以汶川地震影响的陇南地区为例[J].中南大学学报:自然科学版, 2012,43(9):3595-3600.
- Zhang F Y, Liu G, Chen W W, et al. Large landslide susceptibility assessment by multivariate statistical analysis in the Longnan area affected by the Wenchuan earthquake[J]. Journal of Central South University: Science and Technology, 2012,43(9):3595-3600.