

引用格式: 尹超, 李仲波, 张凯, 王绍平, 郭兵. 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 Chao, LI Zhongbo, ZHANG Kai, WANG Shaoping, GUO Bing. Regional landslide susceptibility assessment based on Inter. iamb-Tabu algorithm [J]. Mountain Research, 2023, 41(4): 608-620.

基于 Inter. iamb-Tabu 算法的区域滑坡敏感性评价

尹超¹, 李仲波¹, 张凯², 王绍平³, 郭兵³

(1. 山东理工大学 建筑工程与空间信息学院, 山东 淄博 255049; 2. 山东金鼎智达集团有限公司, 山东 淄博 255049;

3. 日照城投建设集团有限公司, 山东 日照 276800)

摘要: 确定滑坡敏感性概率的空间分布, 可为滑坡防治政策制定和土地利用规划提供科学依据。由于成灾环境的差异和滑坡机理的复杂性, 基于不同逻辑架构(物理模型、理论模型等)的滑坡敏感性评价针对特定孕灾环境, 无法提前确定最优数学评价模型, 而基于深度学习法的、融合多种模型的混合算法能较好地解决这个问题。本文以山东省淄博市博山区为研究区域, 基于单因素 Logistic 回归法筛选了滑坡致灾因子, 通过频率比法对致灾因子进行了分级; 对 4 种基于贝叶斯网络的改进算法(MMHC、MMPC-Tabu、Fast. iamb-Tabu、Inter. iamb-Tabu)进行验证, 引入错误指数确定了建模效果最佳的算法; 计算了博山区 774 570 个栅格的滑坡敏感性概率, 绘制了博山区滑坡敏感性分布图; 基于 GIS 的空间叠加和栅格计算器功能, 对比了各模型的滑坡敏感性评价结果。研究结果表明: (1) 基于 Inter. iamb-Tabu 开展博山区滑坡敏感性建模时效果最佳。该模型包含 10 个节点(9 个致灾因子节点和 1 个结局节点)、14 条有向边。研究区域可划分为极高敏感区、高敏感区、中敏感区、低敏感区和极低敏感区, 分别占总面积的 7.30% (49.80 km²)、16.50% (112.56 km²)、26.10% (178.05 km²)、33.20% (226.49 km²) 和 16.90% (115.29 km²), 分别有 67 处、22 处、7 处、2 处和 1 处滑坡位于上述敏感区。(2) MMHC、MMPC-Tabu 和 Fast. iamb-Tabu 模型在训练时易舍弃样本中部分因子的特征信息, 以达到模型整体精度最优的目的。研究成果可为山东省淄博市博山区滑坡防治提供理论依据, 采用的研究方法同样适用于类似地区。

关键词: 滑坡敏感性评价; 致灾因子; 贝叶斯网络; Inter. iamb-Tabu 算法

中图分类号: P694

文献标志码: A

滑坡敏感性评价在选取致灾因子并分级的基础上基于数学模型计算滑坡发生概率并分析其空间分异状况, 为滑坡防治政策制定和土地利用规划提供数据依据^[1-2]。致灾因子选取与分级以及评价模型构建是滑坡敏感性评价的关键步骤^[3-5]。常用的致灾因子选取方法包括定性和定量两类, 其中, 定性方法是在充分揭示滑坡发生规律和孕灾环境特征的基础上, 结合专家意见选取致灾因子^[6]。定量方法是根据孕灾环境信息并基于特征提取算法筛选致灾因

子, 其本质是有监督特征提取问题^[7-10]。致灾因子分级通常采用经验判断法、等距划分法和统计分析法, 其中, 经验判断法和等距划分法主要依赖主观经验而忽视客观数据, 易导致评价结果准确率较低^[11]。统计分析法包括频率比法、信息量法和熵指数法等, 其中, 频率比法是目前广泛使用的致灾因子与评价模型的联结方法, 它通过引入区间面积(比)和滑坡面积(比)分析致灾因子的不同区间对滑坡发生的影响, 从而降低它与不同模型耦合造成的不

收稿日期(Received date): 2022-12-03; 改回日期(Accepted date): 2023-07-09

基金项目(Foundation item): 国家自然科学基金青年科学基金(51808327); 山东省自然科学基金(ZR2019PEE016, ZR2021MD011) [Young Scientists Fund of the National Natural Science Foundation of China (51808327); Natural Science Foundation of Shandong Province (ZR2019PEE016, ZR2021MD011)]

作者简介(Biography): 尹超(1987-), 男, 博士, 副教授, 主要研究方向: 公路自然灾害风险评价[YIN Chao, (1987-), male, Ph. D., associate professor, research on risk assessment of road natural disasters] E-mail: yinchao1987611@163.com

确定性^[12-13]。

滑坡敏感性评价常用的数学模型分为 3 类^[14]。第一类是数据统计模型,如 Logistic 回归法和二元/多元统计法等^[15]。第二类是机器学习法,如人工神经网络(Artificial Neural Network, ANN)和支持向量机(Support Vector Machines, SVM)等。近几年,深度学习法取得了长足进步,它可以避免浅层学习法易出现过拟合和拓扑结构不稳定等弊端,以随机森林(Random Forest, RF)、卷积神经网络(Convolutional Neural Networks, CNN)和贝叶斯网络应用最为广泛。第三,为克服某一种数学模型的缺点,有学者提出了融合多种模型的混合算法,由于混合算法的泛化能力和鲁棒性较高,其滑坡敏感性评价结果明显优于单一算法^[16]。不同地区的地质环境特点和滑坡成灾机理存在较大差异,各数学模型自身的逻辑结构也各不相同,因此,针对特定地区孕灾环境特征的最优评价模型往往无法提前获知,只能通过大量对比研究确定。贝叶斯网络的实质是利用一个联合概率分布描述变量间的概率依赖关系^[17]。最大最小爬山算法(Max-Min Hill-Climbing, MMHC)是一种广泛使用的贝叶斯算法,MMPC-Tabu、Fast. iamb-Tabu、Inter. iamb-Tabu 是通过禁忌搜索算法(Tabu Search)对 MMHC 进行改进后得到的混合算法。本文基于上述 3 种算法并以 MMHC 为对照,开展了山东省淄博市博山区滑坡敏感性评价研究,主要包括:(1)采用单因素 Logistic 回归法筛选滑坡致灾因子,通过频率比法对致灾因子进行分级;(2)验证 MMHC、MMPC-Tabu、Fast. iamb-Tabu 和 Inter. iamb-Tabu 模型的建模效果,通过错误指数确定效果最佳的模型;(3)开展博山区滑坡敏感性评价,对比各模型的评价结果。研究成果可为山东省淄博市博山区滑坡防治提供理论依据,采用的研究方法同样适用于类似地区。

1 研究区域和研究数据

1.1 研究区域概况

博山区(117°43′~118°42′E,36°16′~36°31′N)地处鲁中山区北部,海拔为 102~1066 m。地质构造复杂,主要存在两种地质构造类型,一是太古代的基底构造,二是中、新生代的盖层结构。南北向有姚家峪断裂,全长 60 km,走向 N5°W~N10°E,倾向

NW,倾角 55°~75°。北东向断裂有淄河断裂,全长 110 km,宽 400~1000 m,地表出露约 60 km,贯穿博山、源泉和池上等乡镇,长 19 km^[18-19]。

博山区地层有四界七系:太古界泰山群(Art)分布于岭西、乐疃、博山镇、池上、李家等地,厚 2700~15 000 m;寒武系(ε)分布于石门、博山镇、李家、池上、源泉、岭西等地,厚约 600 m;奥陶系(O)分布于石门、夏家庄、石马、乐疃、八陡、源泉、岳庄、博山镇等地,厚约 800 m;石炭系(C)分布于八陡、福山、山头、域城、白塔、夏家庄等,厚约 140 m;二迭系(P)分布于八陡、山头、福山、夏家庄、白塔、焦庄、域城等,厚约 430 m;中生界(J)由砂岩、页岩组成,属河湖沼泽沉积,含植物化石,厚约 180 m;第四系广泛分布于河谷、河川、低洼处,以粘质沙土为主,间夹砂砾层,厚 3~4 m^[20]。

1.2 研究数据

本文使用的数据包括:

(1) GDEM V2 30 m 分辨率数字高程模型,来自地理空间数据云(<http://www.gscloud.cn/>);

(2) 清华大学 10 m 分辨率土地利用数据,来自清华大学地理空间数据库(<http://data.ess.tsinghua.edu.cn/>);

(3) 山东省断层数据,来自地质专业知识服务系统(<http://geol.ckcest.cn/>);

(4) 博山区道路数据,来自 Openstreetmap 网(<https://www.openstreetmap.org/#map=5/34.574/113.247>);

(5) 博山区气象局提供的博山区降水数据;

(6) 淄博市自然资源和规划局编制的博山区地质灾害防治方案。

根据博山区地质灾害防治方案,确定了 99 处历史滑坡的发生时间和位置。现场踏勘滑坡,结果显示(图 1):博山区滑坡总体积 2 732 400 m³,总面积 1.03 km²;规模最大的滑坡位于博山镇,体积 74 160 m³,面积 0.02 km²;池上镇滑坡数量最多(19 处)。

2 滑坡致灾因子分析

2.1 滑坡致灾因子

选取 13 项滑坡致灾因子^[21-23],包括坡度、高程、坡向、土地利用类型、工程岩组、断层距离、河流

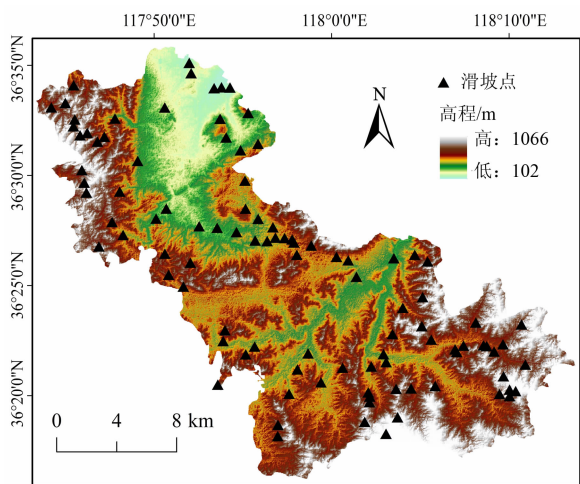


图 1 博山区滑坡分布图

Fig. 1 Landslide distribution in Boshan district of Zibo city, Shandong province of China

距离、道路距离、累积降水量、平面曲率、剖面曲率、归一化植被指数 (Normalized Difference Vegetation Index, NDVI)、地形湿度指数 (Topographic Wetness Index, TWI) 和河流动力指数 (Stream Power Index, SPI)。

其中,NDVI 是最常用的植被指数。当 $NDVI < 0$ 时,地表覆盖云、水、雪等,对可见光高反射;当 $NDVI$

$= 0$ 时,地表多为岩石或裸土;当 $NDVI > 0$ 时,地表有植被覆盖,且植被覆盖度随 $NDVI$ 的增大而增大^[24]。

TWI 考虑了地形和土壤特性对土壤水分分布的影响。基于 GIS 获取 TWI 值包括 5 个步骤:填洼、计算填洼后的坡度、计算流向、计算流量和计算单位面积的汇流量^[25]。

SPI 是反映地表水流侵蚀能力的指标,可用于确定水流汇集而形成的强水流路径和可能出现沟谷侵蚀的地点,地表水流的侵蚀能力随 SPI 值的增大而增强^[26]。

2.2 致灾因子筛选

在贝叶斯网络中,节点数量对结构学习的复杂程度有显著影响,即训练样本一定时,节点过多将导致模型精度下降,不利于反映主要节点与结局的关系^[27];另一方面,滑坡致灾因子并非完全独立,例如,河流距离、TWI 和 SPI 都是反映地表径流强度的指标,三者包含重复信息。为消除滑坡致灾因子的多重共线性并提高建模效率^[28],本文基于单因素 Logistic 回归法筛选致灾因子,计算结果如表 1 所示。

结合已有研究^[29-31]和表 1,选取 P 值为 0.1 ~ 0.2 的 9 项致灾因子用于博山区滑坡敏感性建模,包括坡度、高程、坡向、土地利用类型、断层距离、道

表 1 频率比法计算结果

Tab. 1 Calculation results of the frequency ratio method

致灾因子	分级	区间面积/m ²	滑坡面积/m ²	面积比/%	χ^2	P
坡度	[0, 10)	94 078 800	39 122	0.04	4.08	0.15
	[10, 25)	151 664 400	513 003	0.34		
	[25, 90]	89 056 800	78 875	0.09		
高程	[0, 300)	20 422 800	49 218	0.24	0.54	0.18
	[300, 450)	113 832 000	150 809	0.13		
	[450, 600)	182 466 000	413 936	0.23		
	[600, 1200]	18 079 200	17 037	0.09		
坡向	[0, 90)	85 039 200	195 610	0.23	0.47	0.12
	[90, 210)	108 810 000	164 060	0.15		
	[210, 270)	75 999 600	198 134	0.26		
	[270, 360]	64 951 200	73 196	0.11		
土地利用类型	建筑及道路用地	66 960 000	203 813	0.30	1.46	0.16
	耕地及草地	24 440 400	45 432	0.19		
	未开发用地及其他用地	243 399 600	381 755	0.16		
工程岩组	石灰岩层、泥岩层	79 012 800	34 705	0.04	10.55	0.28
	砂页岩互层	34 484 400	68 148	0.20		
	砂质土、粉质土	221 302 800	528 147	0.24		

续表 1

致灾因子	分级	区间面积/m ²	滑坡面积/m ²	面积比/%	χ^2	<i>P</i>
断层距离	[0, 2000)	188 492 400	468 833	0.25	11.48	0.12
	[2000, 4000)	62 272 800	123 045	0.20		
	[4000, 8000]	84 034 800	39 122	0.05		
河流距离	[0, 1200)	151 999 200	330 644	0.22	1.75	0.40
	[1200, 2400)	124 880 400	234 732	0.19		
	[2400, 3200]	57 920 400	65 624	0.11		
道路距离	[0, 600)	126 889 200	281 426	0.22	0.97	0.17
	[600, 1200)	75 999 600	141 975	0.19		
	[1200, +∞)	131 911 200	207 599	0.16		
累积降水量	[600, 700)	135 928 800	235 363	0.17	0.35	0.05
	[700, 800]	198 871 200	395 637	0.20		
平面曲率	(-∞, -0.5)	135 259 200	178 573	0.13	1.27	0.26
	[-0.5, 0.5)	106 131 600	198 765	0.19		
	[0.5, +∞)	93 409 200	253 662	0.27		
剖面曲率	(-∞, -0.5)	86 378 400	264 389	0.31	2.78	0.15
	[-0.5, 0.5)	112 158 000	209 492	0.19		
	[0.5, +∞)	136 263 600	157 119	0.11		
NDVI	[-1, -0.1)	158 695 200	329 382	0.21	1.47	0.17
	[-0.1, 0.05)	129 567 600	240 411	0.19		
	[0.05, 1]	46 537 200	61 207	0.13		
TWI	[0, 4)	80 686 800	111 056	0.14	1.06	0.04
	[4, 8)	107 470 800	203 813	0.19		
	[8, 12]	146 642 400	316 131	0.22		
SPI	[0, 30)	120 193 200	176 680	0.15	1.93	0.16
	[30, 70)	100 105 200	188 669	0.19		
	[70, +∞)	114 501 600	265 651	0.23		

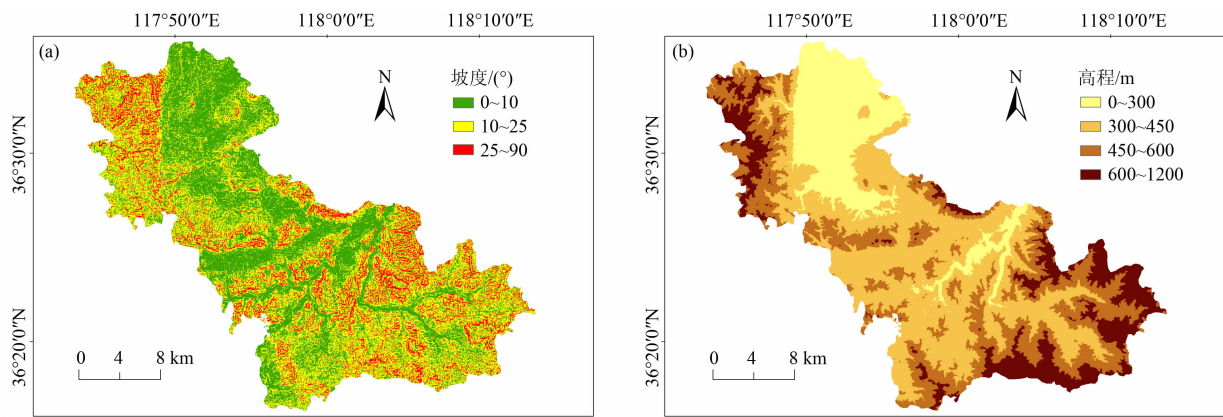
注: χ^2 为拟合优度; *P* 为逻辑概率。

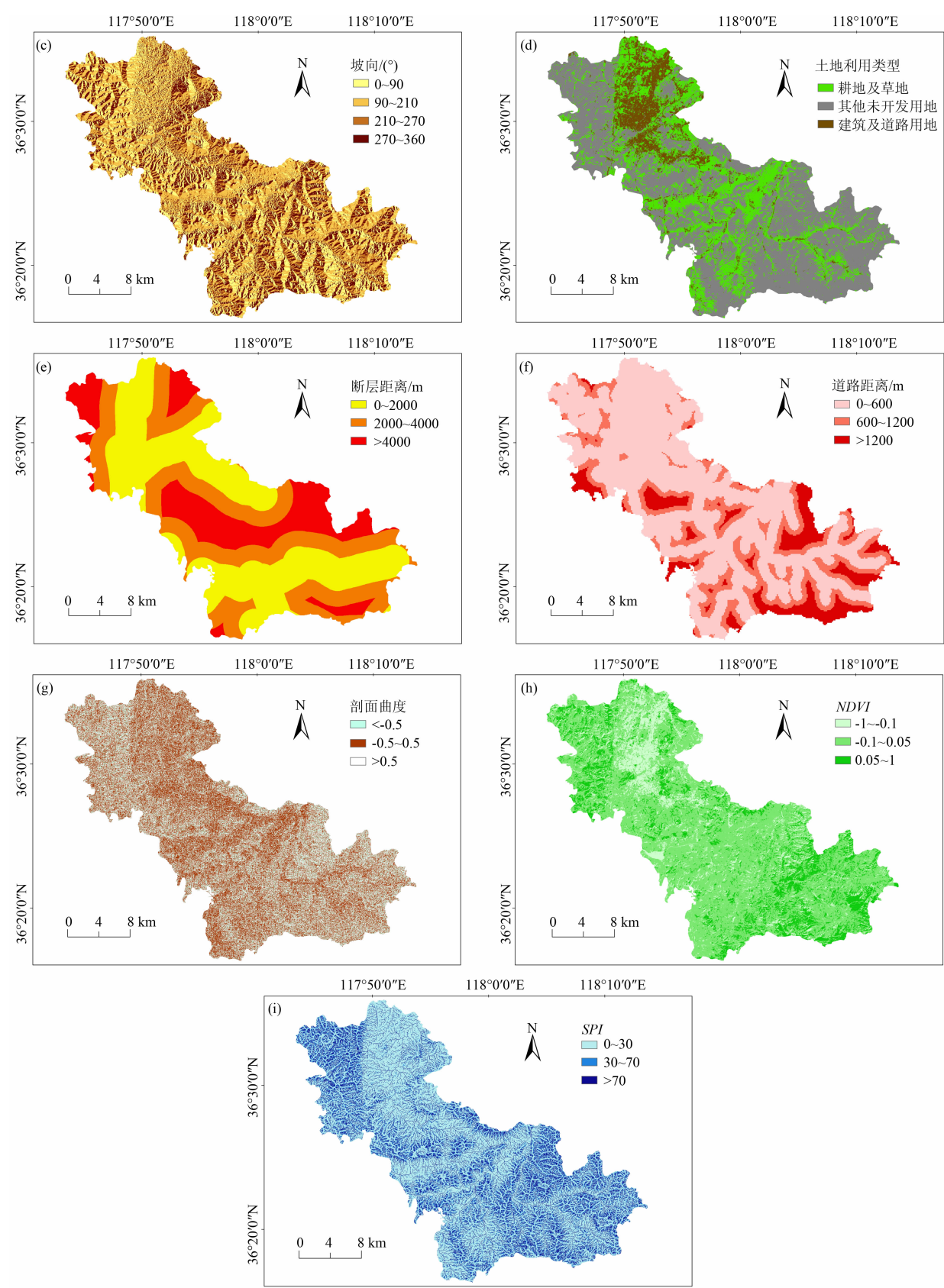
路距离、剖面曲率、NDVI 和 SPI。利用博山区 30 m 分辨率 DEM 提取坡度、高程、坡向、剖面曲率和 SPI 信息,利用 Landsat 8 OLI_TIRS 卫星的 Landsat TM 影像提取 NDVI 信息,利用清华大学 10 m 分辨率土地利用数据提取土地利用类型信息,利用山东省断

层数据提取断层距离信息,利用博山区道路数据提取道路距离信息,将研究单元重采样为 30 m × 30 m 栅格并绘制 9 项致灾因子的分布图(图 2)。

2.3 致灾因子分级

采用频率比法对滑坡致灾因子进行分级,频率





比的计算如下:

$$FR = \frac{N_i/N}{M_i/M}$$

(1)

式中, N_i 、 N 分别为某致灾因子第 i 区间的滑坡面积和滑坡总面积; M_i 、 M 分别为某致灾因子第 i 区间的面积和博山区总面积。 $FR < 1$ 说明该区间对滑坡具有限制作用; $FR > 1$ 说明该区间对滑坡具有促进作用; $FR = 1$ 说明该区间对滑坡的影响不明显。计算 9 项致灾因子各区间的 FR ,限于篇幅,仅列出高程的计算结果(表 2)。

表 2 高程频率比计算结果

Tab. 2 Calculation results of elevation frequency ratio

高程/m	区间 面积/km ²	区间 面积比	滑坡 面积/m ²	滑坡 面积比	FR
[0, 150)	4.69	0.01	10 727	0.020	1.21
[150, 300)	15.74	0.05	38 491	0.060	1.30
[300, 400)	54.24	0.16	71 303	0.110	0.70
[400, 450)	59.60	0.18	79 506	0.130	0.71
[450, 480)	75.00	0.22	165 953	0.260	1.17
[480, 520)	64.61	0.19	148 285	0.240	1.22
[520, 560)	25.44	0.08	60 576	0.100	1.26
[560, 600)	17.41	0.05	39 122	0.060	1.19
[600, 900)	11.05	0.03	15 775	0.030	0.76
[900, 1200]	7.03	0.02	1262	0.002	0.10

当高程为 0 ~ 300 m 和 450 ~ 600 m 时, $FR > 1$; 高程为 30 ~ 450 m 和 600 ~ 1200 m 时, $FR < 1$;据此将高程分为 0 ~ 300 m、300 ~ 450 m、450 ~ 600 m 和

600 ~ 1200 m 共 4 级,同理,根据表 2 对其他 8 项致灾因子进行分级(表 3)。

3 博山区滑坡敏感性评价

3.1 模型介绍

3.1.1 MMHC 算法

贝叶斯网络的实质是利用一个联合概率分布描述变量间的概率依赖关系。MMHC 算法是一种广泛使用的贝叶斯算法,包括两个阶段^[32-33],第一阶段采用最大最小启发式搜索策略,调用 MaxMinHeuristic() 函数获得每个变量的候选父子节点集 CPC(Candidate Parents and Children)。第二阶段是利用 Ind($X; T|Z$) 函数移除第一阶段不应进入 CPC 的节点,如果节点 X 与目标节点 T 在已知 Z 时具备独立性关系,则将 X 移出 CPC。Ind($X; T|Z$) 函数用于判断节点间的条件独立性,如果在给定 Z 时 X 和 T 相互条件独立,则 Ind() 的返回值为 true^[34]。

3.1.2 MMHC 算法的改进

MMPC-Tabu、Fast. iamb-Tabu、Inter. iamb-Tabu 是通过 Tabu Search 算法对 MMHC 进行改进后得到的混合算法。Iamb(Incremental association Markov blanket) 算法包括前进(Growing)和后退(Shirinking)2 个阶段^[35]。在一个贝叶斯网络中,马尔科夫毯节点 $MB(T)$ 为目标节点 T 的父节点、子节点及子节点的其他父节点。 $MB(T)$ 的所有估计值组成 $CMB(T)$,但在前进过程中易产生假阳性关系,即 $CMB(T)$ 的部分元素不是 $MB(T)$ 的估计值。对

表 3 致灾因子分级结果

Tab. 3 Classification results of hazard-inducing factors

分级	致灾因子分级标准								
	坡度/(°)	高程/m	坡向/(°)	土地利用类型	断层距离/m	道路距离/m	剖面曲率	NDVI	SPI
I	[0, 10)	[0, 300)	[0, 90)	建筑及道路用地	[0, 2000)	[0, 600)	(-∞, -0.5)	[-1, -0.1)	[0, 30)
II	[10, 25)	[300, 450)	[90, 210)	耕地及草地	[2000, 4000)	[600, 1200)	[-0.5, 0.5)	[-0.1, 0.05)	[30, 70)
III	[25, 90]	[450, 600)	[210, 270)	未开发用地及其他用地	[4000, 8000)	[1200, +∞)	[1200, +∞)	[0.05, 1]	[70, +∞)
IV		[600, 1200]	[270, 360]						

Iamb 算法进行优化得到 Inter. iamb (Interleaved Iamb)算法,即在 Iamb 的前进阶段加入独立性检验过程,以减少 $CMB(T)$ 中的假阳性关系和条件检验次数。Fast. iamb 算法是对 Iamb 算法的进一步优化,在前进阶段的每一次循环中, Fast. iamb 算法都会根据启发式搜索函数 h 值对每个变量从大到小进行排序,再根据依赖性大小将候选变量加入 $MB(T)$ 中。

3.2 模型性能比较

选取两个贝叶斯标准网络(Chest Clinic Network 和 TANK Network)验证 MMHC、MMPC-Tabu、Fast. iamb-Tabu、Inter. iamb-Tabu 的建模效果,并选取性能最优的算法开展博山区滑坡敏感性评价,其中, Chest Clinic Network 又称 Asia Network,包含 8 个节点和 8 条有向边, TANK Network 包含 14 个节点和 20 条有向边^[36-37]。

分别采用 Chest Clinic Network 和 TANK Network 生成样本数量为 1200、1500、1800 和 2100 的数据集,以数据集为训练样本,基于 MMHC、MMPC-Tabu、Fast. iamb-Tabu、Inter. iamb-Tabu 生成贝叶斯网络并与 Chest Clinic Network 和 TANK Network 的有向无环图进行对比,记录新生成网络的缺失边、错误边和反向边数量并计算错误指数。错

误指数 E_n 计算如下^[34-39]：

$$E_n = n_a + n_v + 0.5n_c \tag{2}$$

式中, n_a 为缺失边数量; n_b 为错误边数量; n_c 为反向边数量。

对比结果(表 4)可知,样本数量分别为 1800 和 2100 时, MMHC、MMPC-Tabu、Fast. iamb-Tabu、Inter. iamb-Tabu 对 Chest Clinic Network 和 TANK Network 的错误指数均相等,说明样本数量达到 1800 即可满足模型训练要求,继续增加样本数量对提高建模准确率没有显著影响;样本数量为 1800 时, Inter. iamb-Tabu 算法对 Chest Clinic Network 和 TANK Network 的错误指数均为 0,而其他算法对 Chest Clinic Network 和 TANK Network 的错误指数均不为 0,说明样本数量为 1800 时 Inter. iamb-Tabu 算法已经训练成熟且其性能明显优于其他算法。因此,本文选取 Inter. iamb-Tabu 开展博山区滑坡敏感性评价^[40]。

3.3 基于 Inter. iamb-Tabu 的博山区滑坡敏感性评价

3.3.1 模型建立

博山区 99 处滑坡共有栅格 1849 个,在非滑坡区域任意选取 1849 个栅格,建立滑坡栅格数据集和非滑坡栅格数据集。分别基于 900 个滑坡栅格和非滑坡栅格构建训练样本集,采用 Inter. iamb-Tabu 构

表 4 不同样本数量和算法的对比结果

Tab.4 Comparison results of different sample sizes and algorithms

样本数量	算法	Chest Clinic Network				TANK Network			
		n_a	n_b	n_c	E_n	n_a	n_b	n_c	E_n
1200	MMHC	2	2	1	4.5	5	4	4	11.0
	MMPC-Tabu	2	1	1	4.0	4	4	3	9.5
	Fast. iamb-Tabu	2	1	1	3.5	3	4	2	8.0
	Inter. iamb-Tabu	1	1	0	2.0	3	3	2	7.0
1500	MMHC	2	1	1	3.5	3	3	4	8.0
	MMPC-Tabu	1	1	1	2.5	3	3	3	7.5
	Fast. iamb-Tabu	1	0	1	1.5	2	2	3	5.5
	Inter. iamb-Tabu	1	0	0	1.0	1	2	2	4.0
1800	MMHC	1	1	1	2.5	2	2	2	5.0
	MMPC-Tabu	0	1	1	1.5	2	1	1	3.5
	Fast. iamb-Tabu	0	1	0	1.0	1	1	0	2.0
	Inter. iamb-Tabu	0	0	0	0	0	0	0	0
2100	MMHC	1	1	1	2.5	2	2	2	5.0
	MMPC-Tabu	0	1	1	1.5	2	1	1	3.5
	Fast. iamb-Tabu	0	1	0	1.0	1	1	0	2.0
	Inter. iamb-Tabu	0	0	0	0	0	0	0	0

建滑坡敏感性模型,并根据极大似然估计法计算各节点的条件概率,建模的硬件环境为 CPU i7-6700 处理器、8G 内存、GTX1050 Ti-8G 显卡,软件环境为 R3.5.0 软件中的贝叶斯网络学习程序包,构建的贝叶斯网络及各致灾因子的条件概率如图 3 所示。博山区滑坡敏感性模型包含 10 个节点(9 个致灾因子节点和 1 个结局节点)、14 条有向边;坡度、高程、土地利用类型和断层距离是滑坡的父节点,对滑坡具有直接诱发作用;剖面曲率和 NDVI 是滑坡的互馈节点,即剖面曲率和 NDVI 对滑坡具有直接诱发作用,且滑坡发生可能改变孕灾环境的剖面曲率和 NDVI;坡向、道路距离和 SPI 与其他致灾因子存在互馈关系,并通过这种互馈关系对滑坡发生起间接作用^[41-43]。

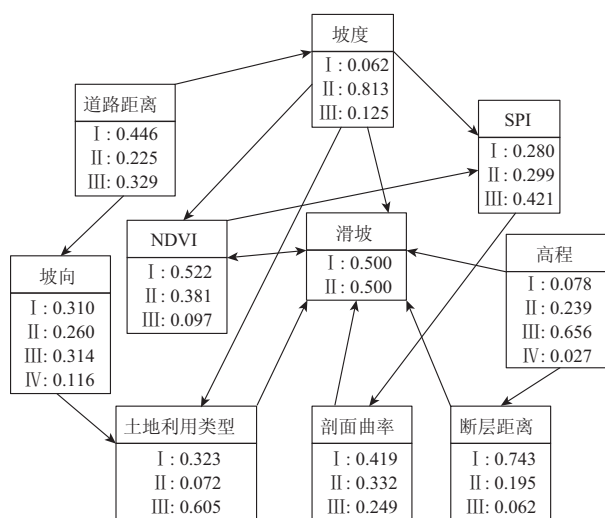


图 3 博山区滑坡敏感性模型

Fig. 3 Landslide susceptibility assessment model in Boshan district

3.3.2 模型验证

基于剩余 949 个滑坡栅格和 949 个非滑坡栅格构建验证样本集,采用受试者工作特征 (Receiver Operating Characteristic, ROC) 曲线下面积 (Area under Curve Roc, AUC) 法对博山区滑坡敏感性模型进行验证^[43-44]。图 4 为验证结果,由于 AUC 值达到 0.91,说明模型具有很高的准确率。

3.4 博山区滑坡敏感性区划

利用 ArcGIS10.2 的栅格计算器,将博山区滑坡敏感性模型按图 3 进行计算,经过 32.78 h 后得到 774 570 个栅格的滑坡敏感性概率,绘制了博山区滑坡敏感性概率分布图 (图 5)。博山区滑坡敏感性概

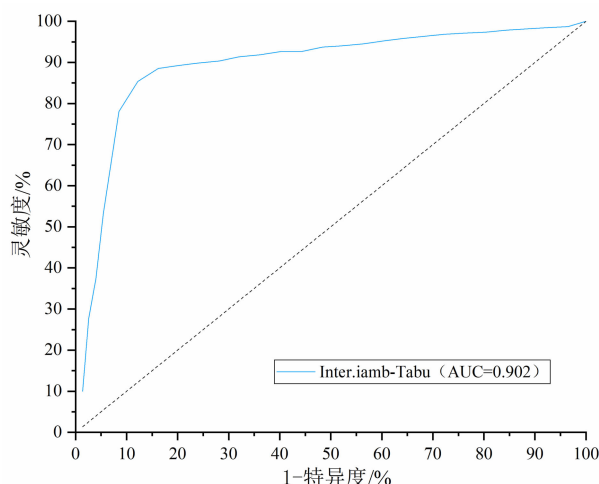


图 4 博山区滑坡敏感性模型验证结果

Fig. 4 Verification results of landslide susceptibility assessment model in Boshan district

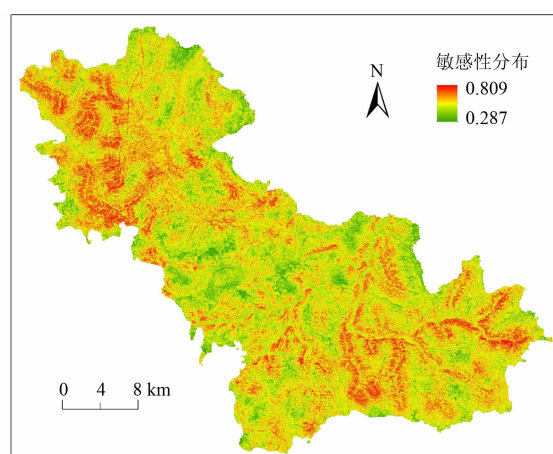


图 5 博山区滑坡敏感性概率分布图

Fig. 5 Landslide susceptibility probability distribution in Boshan district

率最高为 0.81、最小为 0.29;敏感性概率较高的区域是域城镇西部、池上镇和博山镇南部、石马镇的山地地带;较低的区域是源泉镇北部、白塔镇和博山城区、八陡镇和崮山镇的平原地带。

利用自然间断点法进行博山区滑坡敏感性区划 (间断点为 0.73、0.62、0.51 和 0.40^[45]),将博山区划分为极高敏感区、高敏感区、中敏感区、低敏感区和极低敏感区 (图 6),分别占博山区总面积的 7.30% (49.80 km²)、16.50% (112.56 km²)、26.10% (178.05 km²)、33.20% (226.49 km²) 和 16.90% (115.29 km²),各区分别有 67 处、22 处、7 处、2 处和 1 处滑坡,各区滑坡面积分别占滑坡总面积的

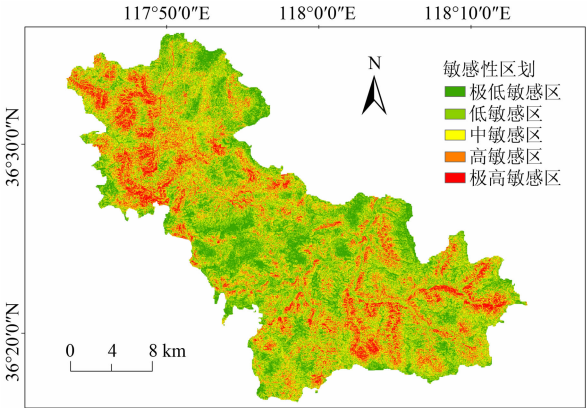


图 6 博山区滑坡敏感性区划图
Fig. 6 Landslide susceptibility regionalization in Boshan district

85.32% (1 391 569 m²)、10.58% (172 560 m²)、2.17% (35 393 m²)、1.09% (17 778 m²) 和 0.84% (13 700 m²)。

4 不同模型评价结果的讨论

以 Inter. iamb-Tabu 模型为基准模型, 基于 ArcGIS10.2 的空间叠加和栅格计算器功能, 将基准模型的滑坡敏感性评价结果与其他模型进行比较, 分别得到 MMHC、MMPC-Tabu、Fast. iamb-Tabu 模型较基准模型的过低估计区域、相等区域和过高估计区域 (图 7)。MMHC、MMPC-Tabu、Fast. iamb-Tabu 模型的滑坡敏感性评价结果与基准模型相等的区域面积远大于过低估计区域或过高估计区域面积^[46]。MMHC 模型错误估计区域最多, 多为过高估计, 呈散点式分布; Fast. iamb-Tabu 模型错误估计区域最少, 基本位于其他模型错误估计区域以内。将 Fast. iamb-Tabu 模型的过低估计区域与道路距离叠加, 过高估计区域同断层距离叠加 (图 8、图 9)。过低

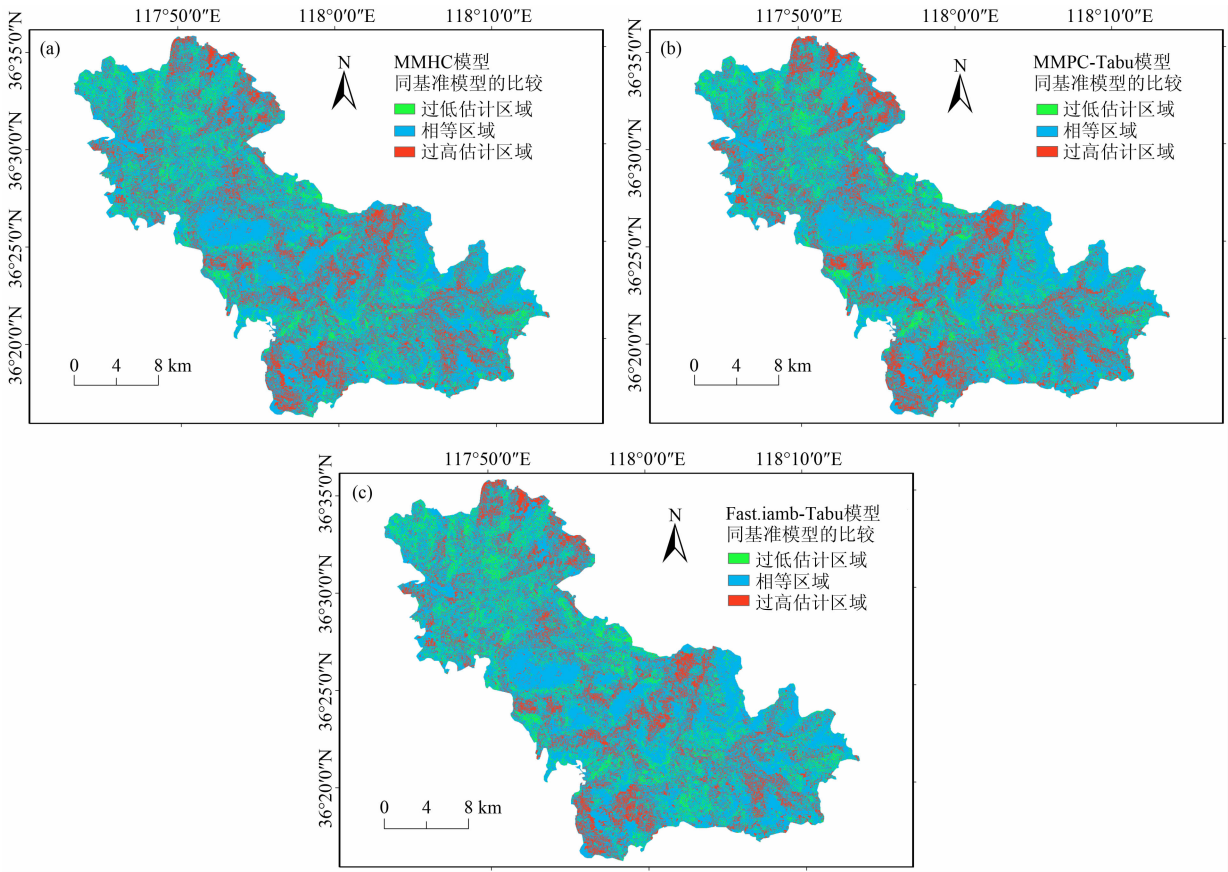


图 7 不同模型的滑坡敏感性评价结果比较:
(a) MMHC 模型同基准模型的比较; (b) MMPC-Tabu 模型同基准模型的比较; (c) Fast. iamb-Tabu 模型同基准模型的比较
Fig. 7 Comparison of landslide susceptibility assessment results of different models:
(a) comparison of the MMHC model with the benchmark model; (b) comparison of the MMPC-Tabu model with the benchmark model;
(c) comparison of the Fast. iamb-Tabu model with the benchmark model

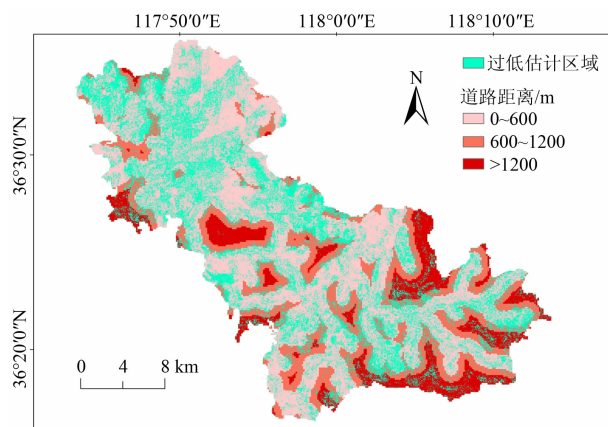


图8 过低估计区域与道路距离的叠加

Fig. 8 Overlay of underestimated area and road distance

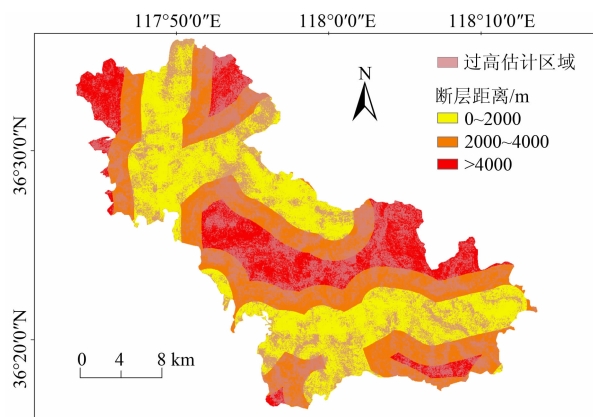


图9 过高估计区域与断层距离的叠加

Fig. 9 Superposition of overestimated areas and fault distances

估计区域分布在距离道路近的区域,过高估计区域的分布呈散点状,与距离断层近的区域重合率较高。结果表明,同 Inter. iamb-Tabu 模型相比,其他模型弱化了道路距离的作用,高估了断层距离的作用。综上所述,MMHC 模型、MMPC-Tabu 模型和 Fast. iamb-Tabu 模型在训练时,容易舍弃样本中部分因子的特征信息以达到模型整体精度最优的目的。

5 结论

(1)博山区滑坡致灾因子包括坡度、高程、坡向、土地利用类型、断层距离、道路距离、剖面曲率、NDVI 和 SPI;基于 Inter. iamb-Tabu 开展博山区滑坡敏感性建模时效果最佳,该模型包含 10 个节点(9 个致灾因子节点和 1 个结局节点)、14 条有向边;博

山区划分为极高敏感区、高敏感区、中敏感区、低敏感区和极低敏感区,分别占博山区总面积的 7.30% (49.80 km²)、16.50% (112.56 km²)、26.10% (178.05 km²)、33.20% (226.49 km²) 和 16.90% (115.29 km²),分别有 67 处、22 处、7 处、2 处和 1 处滑坡位于极高敏感区、高敏感区、中敏感区、低敏感区和极低敏感区。

(2)本文基于频率比法和 Inter. iamb-Tabu 算法开展了博山区滑坡敏感性评价,评价结果具有很高的准确率,MMHC 模型、MMPC-Tabu 模型和 Fast. iamb-Tabu 模型在训练时,容易舍弃样本中部分因子的特征信息以达到模型整体精度最优的目的。

本文未开展基于决策树、RF、CNN 和深度信念网络等深度学习算法及其混合算法的滑坡敏感性评价,相关算法的评价准确率需要在今后的研究中进行验证。

参考文献 (References)

- [1] CHEN Wei, PENG Jianbing, HONG Haoyuan, et al. Landslide susceptibility modeling using GIS-based machine learning techniques for Chongren county, Jiangxi province, China [J]. Science of the Total Environment, 2018, **626**: 1121 – 1135. DOI: 10.1016/j.scitotenv.2018.01.124
- [2] OUYANG Chaojun, WANG Zhongwen, AN Huicong, et al. An example of a hazard and risk assessment for debris flows; A case study of Niwan Gully, Wudu, China [J]. Engineering Geology, 2019, **263**: 105351. DOI: 10.1016/j.enggeo.2019.105351
- [3] YIN Chao, ZHANG Jinglei. Hazard regionalization of debris-flow disasters along highways in China [J]. Natural Hazards, 2018, **91** (2): S129 – S147. DOI: 10.1007/s11069-018-3229-8
- [4] MAO Yimin, MWAKAPESA D S, LI Yican, et al. Assessment of landslide susceptibility using DBSCAN-AHD and LD-EV methods [J]. Journal of Mountain Science, 2022, **19** (1): 184 – 197. DOI: 10.1007/s11629-020-6491-7
- [5] SHOJAEZADEH S A, NIKOO M R, MIRCHI A, et al. Probabilistic hazard assessment of contaminated sediment in rivers [J]. Science of the Total Environment, 2020, **703**: 134875. DOI: 10.1016/j.scitotenv.2019.134875
- [6] ZHANG Tingyu, MAO Zhongan, WANG Tao. GIS-based evaluation of landslide susceptibility using a novel hybrid computational intelligence model on different mapping units [J]. Journal of Mountain Science, 2020, **17** (12): 2929 – 2941. DOI: 10.1007/s11629-020-6393-8
- [7] ZHENG Hao, CHENG Gang, LI Yong, et al. A new fault diagnosis method for planetary gear based on image feature extraction and bag-of-words model [J]. Measurement, 2019, **145**: 1 – 13. DOI: 10.

- 1016/j. measurement. 2019. 05. 046
- [8] MOGHADDAM S H A, MOKHTARZADE M, BEIRAMI B A. A feature extraction method based on spectral segmentation and integration of hyperspectral images [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2020, **89**: 102097. DOI: 10. 1016/j. jag. 2020. 102097
- [9] LIN Qigen, LIMA P, STEGER S, et al. National-scale data-driven rainfall induced landslide susceptibility mapping for China by accounting for incomplete landslide data [J]. *Geoscience Frontiers*, 2021, **12**: 101248. DOI: 10. 1016/j. gsf. 2021. 101248
- [10] ZHOU Xinzhi, WEN Haijia, ZHANG Yalan, et al. Landslide susceptibility mapping using hybrid random forest with Geo-Detector and RFE for factor optimization [J]. *Geoscience Frontiers*, 2021, **12**: 101211. DOI: 10. 1016/j. gsf. 2021. 101211
- [11] 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] DOI: 10. 13722/j. cnki. jrm. 2015. 0318
- [12] TSANGARATOS P, ILIA I. Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size [J]. *Catena*, 2016, **145**: 164 – 179. DOI: 10. 1016/j. catena. 2016. 06. 004
- [13] WU Zening, SHEN Yanxia, WANG Huiliang, et al. Urban flood disaster risk evaluation based on ontology and Bayesian Network [J]. *Journal of Hydrology*, 2020, **583**: 124596. DOI: 10. 1016/j. jhydrol. 2020. 124596
- [14] MOHAMMADY M, POURGHASEMI H R, PRADHAN B. Landslide susceptibility mapping at Golestan Province, Iran: A comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models [J]. *Journal of Asian Earth Sciences*, 2012, **61**: 221 – 236. DOI: 10. 1016/j. jseas. 2012. 10. 005
- [15] SUN Deliang, WEN Haijia, WANG Danzhou, et al. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm [J]. *Geomorphology*, 2020, **362**: 107201. DOI: 10. 1016/j. geomorph. 2020. 107201
- [16] SONG Yiquan, GONG Jianhua, GAO Sheng, et al. Susceptibility assessment of earthquake-induced landslides using Bayesian network: A case study in Beichuan, China [J]. *Computers and Geosciences*, 2012, **42**: 189 – 199. DOI: 10. 1016/j. cageo. 2011. 09. 011
- [17] COSTACHE R, BUI D T. Spatial prediction of flood potential using new ensembles of bivariate statistics and artificial intelligence: A case study at the Putna river catchment of Romania [J]. *Science of the Total Environment*, 2019, **691**: 1098 – 1118. DOI: 10. 1016/j. scitotenv. 2019. 07. 197
- [18] 高晓梅, 秦增良, 王令军, 等. 鲁中山区地质灾害的气候特征分析[J]. *科技导报*, 2012, **30**(4): 55 – 60. [GAO Xiaomei, QIN Zengliang, WANG Lingjun, et al. The climatic characteristics of geological calamity in the mountainous area of the middle part of Shandong province [J]. *Science and Technology Review*, 2012, **30**(4): 55 – 60] DOI: 10. 3981/. issn. 1000 – 7857. 2012. 04. 008
- [19] 姜川, 曲宝杰, 董强, 等. 基于 GIS 的淄博市地质灾害预警预报系统建设[J]. *山东国土资源*, 2014, **30**(3): 89 – 91. [JIANG Chuan, QU Baojie, DONG Qiang, et al. Construction of geological disasters warning system based on GIS in Zibo city [J]. *Shandong Land and Resources*, 2014, **30**(3): 89 – 91]
- [20] 孙强, 史清明. 基于 GIS 的城市地震灾害风险区划研究——以淄博市为例[J]. *四川地震*, 2020(2): 19 – 24. [SUN Qiang, SHI Qingming. Study on the risk zoning of urban earthquake disaster based on GIS – take ZiBo city as an example [J]. *Earthquake Research in Sichuan*, 2020(2): 19 – 24] DOI: 10. 13716/j. cnki. 1001 – 8115. 2020. 02. 004
- [21] CHEN Wei, XIE Xiaoshen, PENG Jianbing, et al. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method [J]. *Catena*, 2018, **164**: 135 – 149. DOI: 10. 1016/j. catena. 2018. 01. 012
- [22] KHALAJ S, BAHOOTOROODY F, ABAEI M M, et al. A methodology for uncertainty analysis of landslides triggered by an earthquake [J]. *Computers and Geotechnics*, 2020, **117**: 103262. DOI: 10. 1016/j. compgeo. 2019. 103262
- [23] YILMAZ I. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey) [J]. *Computers and Geosciences*, 2009, **35**(6): 1125 – 1138. DOI: 10. 1016/j. cageo. 2008. 08. 007
- [24] ADITIAN A, KUBOTA T, SHINOHARA Y. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia [J]. *Geomorphology*, 2018, **318**: 101 – 111. DOI: 10. 1016/j. geomorph. 2018. 06. 006
- [25] BOLLMANN S, KRISTENSEN M H, LARSEN M S, et al. SHARQnet-Sophisticated harmonic artifact reduction in quantitative susceptibility mapping using a deep convolutional neural network [J]. *Zeitschrift für Medizinische Physik*, 2019, **29**(2): 139 – 149. DOI: 10. 1016/j. zemedi. 2019. 01. 001
- [26] WANG Yi, FANG Zhice, HONG Haoyuan. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan county, China [J]. *Science of the Total Environment*, 2019, **666**: 975 – 993. DOI: 10. 1016/j. scitotenv. 2019. 02. 263
- [27] MICHALOWSKI R L, PARK D. Stability assessment of slopes in rock governed by the Hoek-Brown strength criterion [J].

- International Journal of Rock Mechanics and Mining Sciences, 2020, **127**: 104217. DOI: 10.1016/j.ijrmms.2020.104217
- [28] NSEKA D, KAKEMBO V, BAMUTAZE Y, et al. Analysis of topographic parameters underpinning landslide occurrence in Kigezi highlands of southwestern Uganda [J]. *Natural Hazards*, 2019, **99** (2): 973 – 989. DOI: 10.1007/s11069-019-03787-x
- [29] EBID A H I M, ALI Z T, GHOBARY M A F. Blood pressure control in hypertensive patients: Impact of an Egyptian pharmaceutical care model [J]. *Journal of Applied Pharmaceutical Science*, 2014, **4** (9): 93 – 101. DOI: 10.7324/JAPS.2014.40917
- [30] KIM H, LEE J H, PARK H J, et al. Assessment of temporal probability for rainfall-induced landslides based on nonstationary extreme value analysis [J]. *Engineering Geology*, 2021, **294**: 106372. DOI: 10.1016/j.enggeo.2021.106372
- [31] CONFORTI M, IETTO F. Modeling shallow landslide susceptibility and assessment of the relative importance of predisposing factors, through a GIS-based statistical analysis [J]. *Geosciences*, 2021, **11** (8): 333. DOI: 10.3390/geosciences11080333
- [32] 钟坤华, 陈芋文, 秦小林. 基于子网融合的贝叶斯网络结构学习算法[J]. *计算机科学*, 2022, **49** (S2): 64 – 70. [ZHONG Kunhua, CHEN Yuwen, QIN Xiaolin. Sub-BN-Merge based bayesian network structure learning algorithm [J]. *Computer Science*, 2022, **49** (S2): 64 – 70] DOI: 10.11896/jsjcx.210800172
- [33] SONG Wenzhu, QIU Lixia, QING Jianbo, et al. Using Bayesian network model with MMHC algorithm to detect risk factors for stroke [J]. *Mathematical Biosciences and Engineering: MBE*, 2022, **19** (12): 13660 – 13674. DOI: 10.3934/MBE.2022637
- [34] 潘金花. 基于 Inter. iamb-Tabu 混合算法的贝叶斯网络效果评价及在高脂血症相关因素研究中的应用[D]. 太原: 山西医科大学, 2019: 1 – 58. [PAN Jinhua. Evaluation of Bayesian networks based on Inter. iamb-Tabu hybrid algorithm and its application in analyzing the relating factors of hyperlipidemia [D]. Taiyuan: Shanxi Medical University, 2019: 1 – 58]
- [35] YANG Xianglin, WANG Yujing, OU Yang, et al. Three-Fast-Inter Incremental Association Markov Blanket learning algorithm [J]. *Pattern Recognition Letters*, 2019, **122**: 73 – 78. DOI: 10.1016/j.patrec.2019.02.002
- [36] ERTUNCAY D, COSTA G. Determination of near-fault impulsive signals with multivariate Naïve Bayes method [J]. *Natural Hazards*, 2021, **108** (2): 1763 – 1780. DOI: 10.1007/s11069-021-04755-0
- [37] LEE Sunmin, LEE M J, JUNG H S, et al. Landslide susceptibility mapping using Naïve Bayes and Bayesian network models in Umyeonsan, Korea [J]. *Geocarto International*, 2020, **35** (15): 1665 – 1679. DOI: 10.1080/10106049.2019.1585482
- [38] ZHOU Yusheng, LI Xue, YUEN K F. Holistic risk assessment of container shipping service based on Bayesian Network Modelling [J]. *Reliability Engineering and System Safety*, 2022, **220**: 108305. DOI: 10.1016/j.res.2021.108305
- [39] LI Tingting, ZHOU Yangze, ZHAO Yang, et al. A hierarchical object oriented Bayesian network-based fault diagnosis method for building energy systems [J]. *Applied Energy*, 2022, **306** (PB): 118088. DOI: 10.1016/j.apenergy.2021.118088
- [40] MUKHAMMADZODA S, SHOHAFAZ F, ILHOMJON O, et al. Application of frequency ratio method for landslide susceptibility mapping in the Surkhob valley, Tajikistan [J]. *Journal of Geoscience and Environment Protection*, 2021, **9**: 168 – 189. DOI: 10.4236/gep.2021.912011
- [41] LI Ming, ZHANG Ren, LIU Kefeng. A new marine disaster assessment model combining Bayesian network with information diffusion [J]. *Journal of Marine Science and Engineering*, 2021, **9** (6): 640. DOI: 10.3390/jmse9060640
- [42] YIN Chao, LI Haoran, CHE Fa, et al. Susceptibility mapping and zoning of highway landslide disasters in China [J]. *PLoS ONE*, 2020, **15** (9): e0235780. DOI: 10.1371/journal.pone.0235780
- [43] ZÉZERE J L, PEREIRA S, MELO R, et al. Mapping landslide susceptibility using data-driven methods [J]. *Science of the Total Environment*, 2017, **589**: 250 – 267. DOI: 10.1016/j.scitotenv.2017.02.188
- [44] BERHANE G, KEBEDE M, ALFARAH N, et al. Landslide susceptibility zonation mapping using GIS-based frequency ratio model with multi-class spatial data-sets in the Adwa-Adigrat mountain chains, northern Ethiopia [J]. *Journal of African Earth Sciences*, 2020, **164**: 103795. DOI: 10.1016/j.jafears.2020.103795
- [45] DU Juan, GLADE T, WOLDAI T, et al. Landslide susceptibility assessment based on an incomplete landslide inventory in the Jilong Valley, Tibet, Chinese Himalayas [J]. *Engineering Geology*, 2020, **270**: 105572. DOI: 10.1016/j.enggeo.2020.105572
- [46] ZHANG Han, YIN Chao, WANG Shaoping, et al. Landslide susceptibility mapping based on landslide classification and improved convolutional neural networks [J]. *Natural Hazards*, 2023, **116**: 1931 – 1971. DOI: 10.1007/s11069-022-05748-3

Regional Landslide Susceptibility Assessment Based on Inter. Iamb-Tabu Algorithm

YIN Chao¹, LI Zhongbo¹, ZHANG Kai², WANG Shaoping³, GUO Bing³

(1. School of Civil Engineering and Geomatics, Shandong University of Technology, Zibo 255049, Shandong, China;

2. Shandong Jinding Zhida Group Co., Ltd., Zibo 255049, Shandong, China;

3. Rizhao City Construction Investment Group Co., Ltd., Rizhao 276800, Shandong, China)

Abstract: Determining the spatial distribution of landslide susceptibility probability can provide scientific basis for landslide prevention policy making and land use planning. Due to the spatial differentiation of geo-disaster-prone environment and the complexity of landslide formation, landslide susceptibility evaluation based on different logical frameworks (physical model, theoretical model, etc.) cannot determine the optimal mathematical evaluation model in advance for specific geo-environment, while the hybrid algorithm based on deep learning method and integrating of multiple models can better solve this problem.

In this study, it took Boshan district of Zibo city, Shandong province of China for a case study. It selected landslide-triggering factors by using univariate logistic regression method, and weighed them by frequency ratio method. Based on Bayesian Network, four improved algorithms, i. e., MMHC, MMPC-Tabu, Fast. iamb-Tabu and Inter. iamb-Tabu were verified for suitability, followed by the best algorithm to be determined for landslide susceptibility assessment by introducing the error exponent into all models. Landslide susceptibility probabilities of 774 570 grids in Boshan district were calculated and the resulting zoning map was plotted. Based on the spatial overlay and raster calculator function of GIS, the landslide susceptibility assessment results of each model were compared.

This research found (1) the Inter. iamb-Tabu model had the best performance in landslide susceptibility assessment. It contained 10 nodes (9 landslide-triggering factor nodes and 1 ending node) and 14 directed edges. Accordingly, Boshan district could be zoned into five areas, i. e., extremely high susceptible areas, high susceptible areas, medium susceptible areas, low susceptible areas and extremely low susceptible areas, accounting for 7.30% (49.80 km²), 16.50% (112.56 km²), 26.10% (178.05 km²), 33.20% (226.49 km²) and 16.90% (115.29 km²) of the whole areas of Boshan district, respectively. There were 67, 22, 7, 2 and 1 landslide identified in aforesaid five respective areas. (2) In MMHC, MMPC-Tabu and Fast. iamb-Tabu model, part of feature information of some landslide-triggering factor was likely to be skipped over in training samples for the best overall accuracy of the model.

The research can provide theoretical basis for landslide control in Boshan district, and the proposed landslide susceptibility assessment is also applicable to geohazard susceptibility delineation in similar areas.

Key words: landslide susceptibility assessment; landslide-triggering factor; Bayesian Network; Inter. iamb-Tabu algorithm

(责任编辑 朱颖彦 钟雨倩)