

# 基于熵权-随机森林优化赋权法的新源县 地质灾害风险评估

刘京会, 李鑫旭, 袁旭山, 李艳敏

(防灾科技学院应急技术与管理学院, 河北 三河 065201)

**摘要:** 伊犁河流域地质灾害频发, 新源县位于伊犁河谷腹地, 是流域内发生最频繁的县市之一。以新源县为研究区, 构建包含致灾因子危险性、孕灾环境敏感性与承灾体脆弱性的三维度指标体系, 并提出一种结合熵权法与随机森林模型的权重优化方法, 结合 ArcGIS 平台对新源县地质灾害风险进行综合评估。结果表明: (1) 权重优化后, 危险性主要由水动力因子主导, 敏感性受地形因素影响显著, 脆弱性以交通与人口经济特征为主, 同时优化后权重在受试者工作特征曲线下的面积较熵权法更高。 (2) 危险性、敏感性与脆弱性在空间上呈现明显差异, 分别对应地质条件复杂区域、阴坡缓坡区域和人口经济集聚区域。 (3) 整体风险空间格局与脆弱性高度一致, 高风险区主要集中在镇区与主要道路沿线, 中高风险区多位于灾害环境与承灾体叠加区, 低风险区分布于生态稳态区或无显著承灾体区。 (4) 风险高值区与历史灾害点及典型承灾体空间分布高度重合, 验证了评估结果的合理性与可靠性。研究结果揭示了新源县地质灾害风险空间分布格局, 为防灾减灾工作提供一定科学依据与理论参考。

**关键词:** 风险评估; 地质灾害; 熵权法; 随机森林; 新源县

**文章编号:** 1000-6060(2026)04-0778-13(0778~0790)

近年来, 地质灾害呈现频发、多发、高发趋势, 对社会经济造成严重影响, 也使地质灾害风险评价成为国内外学术研究的前沿领域。伊犁河谷地处天山构造带, 是我国地质灾害发育最为活跃的地区之一, 其中新源县因构造活动频繁而成为典型代表<sup>[1]</sup>。新源县受矿山开采、植被破坏等复合因素影响, 地质灾害频次逐渐上升。同时, 新源县是伊犁州人口与经济密度最高的县市之一, 灾害频发对社会经济与群众生产生活造成严重影响。现有研究虽对新源县或伊犁河谷地质灾害分布开展调查与识别, 但对地质灾害风险的系统评价与指标权重的优化研究仍存在不足。因此, 开展新源县地质灾害系统风险评估具备现实的防治意义。

地质灾害风险评估体系随着发展已经逐步完善。史培军<sup>[2]</sup>在 2002 年提出的理论中将灾害风险

评估体系分为广义评估和狭义评估, 奠定了自然灾害风险评估的方法基础。其中广义风险评估强调在致灾因子、孕灾环境与承灾体基础上对灾害风险进行综合评估<sup>[3-4]</sup>, 许多学者围绕此框架在构建指标与确定权重方面开展大量研究。

传统权重确定方法大致可以分为主观赋权与客观赋权。层次分析法 (Analytic hierarchy process, AHP) 作为典型的主观赋权法, 在赋权过程中能够体现专家的经验知识, 但其结果也容易受到专家的主观影响<sup>[5]</sup>; 熵权法则是一种客观赋权法, 它依据数据自身信息熵确定权重, 但会忽视数据之间的相关性, 做出孤立的判断<sup>[6]</sup>。为了克服不同方法的局限, 广大学者提出了如 AHP-熵权法、AHP-灰色关联法等主客观综合赋权法<sup>[7]</sup>, 考虑兼顾主客观信息进行赋权。然而该类方法仍存在局限, 一方面难以消除

收稿日期: 2025-05-22; 修订日期: 2025-08-20

基金项目: 第三次新疆综合科学考察 (2022xjkk0600) 资助

作者简介: 刘京会 (1975-), 女, 博士, 副教授, 主要从事灾害监测与评估等方面的研究. E-mail: liujh@cidp.edu.cn

通讯作者: 李鑫旭 (2000-), 男, 硕士研究生, 主要从事灾害风险评估等方面的研究. E-mail: lxxmail824@163.com

主观性,另一方面客观方法虽依赖数据特性,但难以刻画数据间的非线性关系,无法得到更全面的权重。因此,如何避免赋权过程中的局限,仍是热门的研究话题。

近年来,机器学习方法逐渐被引入灾害风险评估领域,展现出能够直接构建指标与风险的关系,不易受人为干扰的优势<sup>[8-9]</sup>。典型的机器学习方法适用于不同场景,如逻辑回归仅适合处理线性问题,难以应对复杂的非线性关系;支持向量机在小样本上表现出色,但在样本量较大的情景中出现效率不降的问题。与之相比,随机森林具备稳健性高、抗噪性强等优势,并且能够通过内置机制直接量化指标贡献度。这一特点为解决传统赋权方法无法考虑数据之间关系的局限提供了可能,也令其在风险评估领域得到了广泛应用。

新源县作为伊犁河谷地区地质灾害最频发的县市之一,已有大量学者围绕其地质灾害问题开展研究。周平华等<sup>[10]</sup>对新源县地质灾害易发性做出了量化评价;常顺利等<sup>[11]</sup>通过证据权法为指标赋权,对危险性做出一定评价;程乙峰等<sup>[12]</sup>通过证据权法为滑坡影响因素赋权,提出了一种基于数据驱动的滑坡危险性评价方法;赵忠国等<sup>[13]</sup>分别使用逻辑回归法与多元自适应回归样条法构建新源县滑坡敏感性模型,发现高程、坡度及降雨量对滑坡敏感性影响较大。综合来看,已有新源县地质灾害研究聚焦于易发性或危险性评估,而风险评估研究略显不足,缺少对评估结果的验证工作。

本研究从灾害系统角度综合考虑致灾因子危

险性、孕灾环境敏感性与承灾体脆弱性三维度风险评估体系,使用熵权-随机森林优化赋权法得到指标权重,对新源县地质灾害风险做出综合评估,并通过历史灾害事件和3种典型承灾体对风险评估结果进行验证,以更准确地揭示新源县地质灾害风险空间分布特征,为区域灾害防治提供参考。

## 1 数据与方法

### 1.1 研究区概况

新源县隶属于新疆维吾尔自治区伊犁哈萨克自治州,处在伊犁河谷东端,全县总面积7581 km<sup>2</sup>,下辖9个镇、2个乡、2个国有农牧场。全境大致划分为山地、丘陵、河谷平原3种地形,分别占全县总面积的73.19%、5.70%、21.11%(图1)。研究区呈三面环山、东高西低、南北高中间低的特殊地形,年均降雨量自西向东为280~1200 mm。由于其特殊地形与高降水量特点,新源县地质灾害频发。经勘测境内共有地质灾害隐患点417个,这些隐患点的存在对新源县的未来发展有着极大的威胁。

### 1.2 数据来源

本研究共使用13类数据,其来源、时空分辨率与数据说明如表1所示,数据均来源于被广泛应用的数据集。数据处理过程中,对其进行严格预处理,将缺失值与异常值填补为邻域均值,同时统一数据坐标系、空间分辨率等,保证数据的一致性与完整性。后续矢量数据处理过程中同样进行严格的质量控制,以保障数据质量。总体而言,所有数据均经过质量控制,以保证后续分析的可靠性。

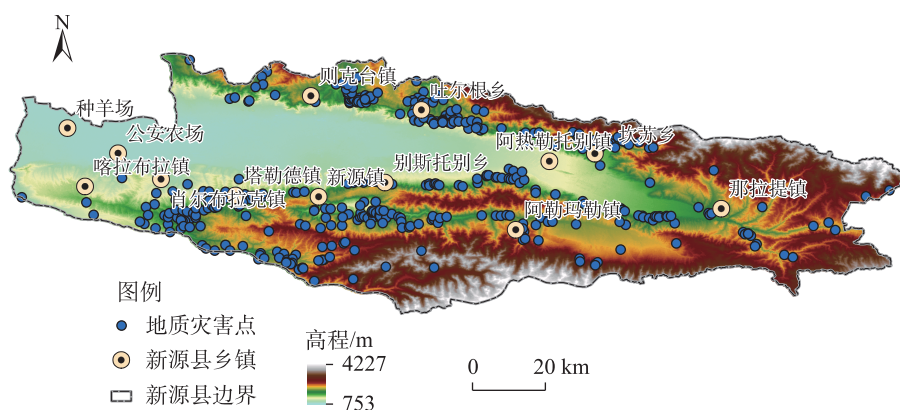


图1 研究区示意图

Fig. 1 Schematic diagram of the study area

表 1 数据来源

Tab. 1 Data source

数据类型	数据来源	时间分辨率	空间分辨率	数据说明
地质灾害点分布数据	地理遥感生态网( <a href="http://www.gisrs.cn/">http://www.gisrs.cn/</a> )	统计至2020年	矢量数据	导入 ArcGIS 平台使用
岩性分布数据	GLADA 数据集( <a href="https://data.isric.org/">https://data.isric.org/</a> )	长期不变	1000 m	长期稳定且不随时间变化
土壤类型数据	GLADA 数据集( <a href="https://data.isric.org/">https://data.isric.org/</a> )	长期不变	1000 m	长期稳定且不随时间变化
地貌类型数据	GLADA 数据集( <a href="https://data.isric.org/">https://data.isric.org/</a> )	长期不变	1000 m	长期稳定且不随时间变化
历史断层分布数据	全国地理信息资源目录服务系统( <a href="https://www.webmap.cn/">https://www.webmap.cn/</a> )	统计至2020年	矢量数据	通过缓存区工具处理为栅格数据使用
降雨数据	国家青藏高原科学数据中心( <a href="https://data.tpcd.ac.cn/">https://data.tpcd.ac.cn/</a> )	2020 年	矢量数据	通过克里金插值处理为栅格数据使用
高程数据	地理空间数据云( <a href="https://www.gscloud.cn/">https://www.gscloud.cn/</a> )	长期不变	100 m	长期稳定且不随时间变化
植被覆盖数据	全球土壤覆盖数据库( <a href="https://soilgrids.org/">https://soilgrids.org/</a> )	2020 年	100 m	—
人口数据	World Pop 官方网站( <a href="https://www.worldpop.org/">https://www.worldpop.org/</a> )	2020 年	100 m	—
GDP 数据	资源环境科学与数据平台( <a href="https://www.resdc.cn/">https://www.resdc.cn/</a> )	2020 年	100 m	—
兴趣点数据	高德地图( <a href="https://ditu.amap.com/">https://ditu.amap.com/</a> )	2020 年	矢量数据	通过核密度工具处理为兴趣点密度使用
基础地理数据	国家基础地理信息中心( <a href="https://www.ngcc.cn/">https://www.ngcc.cn/</a> )	2020 年	矢量数据	导入 ArcGIS 平台使用

### 1.3 研究方法

**1.3.1 风险评估体系** 本文选取广义灾害风险评估体系进行风险评估<sup>[14]</sup>,该体系按致灾因子危险性、孕灾环境敏感性、承灾体脆弱性3个维度表征灾害风险,通过逐一计算相应指标得到危险性、敏感性与脆弱性得分,最终综合得到风险指数。表达式如下:

$$R = W_h \times \sum_{i=1}^{m_h} (H_i \times w_i) + W_s \times \sum_{i=1}^{m_s} (S_i \times w_i) + W_v \times \sum_{i=1}^{m_v} (V_i \times w_i) \quad (1)$$

式中:  $R$  为风险;  $H_i$ 、 $S_i$ 、 $V_i$  分别为第  $i$  个危险性、敏感性与脆弱性指标;  $m_h$ 、 $m_s$ 、 $m_v$  为危险性、敏感性与脆弱性各自指标数量;  $W_h$ 、 $W_s$ 、 $W_v$  分别为危险性、敏感性、脆弱性权重;  $w_i$  为第  $i$  个指标的权重。

通过对指标栅格数据进行空间加权叠加得到具体风险结果,具体流程如图2所示。首先对数据进行统一预处理,包括栅格化与重采样,将数据统一至 WGS 1984 坐标系与 100 m 分辨率,并对异常值采用邻域均值填补的方式处理;为保证不同量纲数据的可比性,对各数据进行栅格归一化;将归一化的栅格数据导出为 GeoTIFF 文件,并通过 Python 的 rasterio 库将其读取生成数值矩阵,从而实现数据量化;后续对指标多重共线性进行检验并筛选冗余指标;根据余下的指标数据结合熵权法与随机森林

模型集成的熵权-随机森林赋权法,计算优化权重;在 ArcGIS 中的栅格计算器中计算指标栅格数据与优化权重的加权叠加结果,实现空间上的综合风险评估。

**1.3.2 赋权方法** 熵权法基于信息熵原理,是一种使用已获得的客观数据得到指标权重的客观赋权法<sup>[15]</sup>。根据熵权法基本理论,数据指标的数值差异度越大,则表示该指标离散程度较大,相应的权重较高。具体计算步骤如下:

(1) 指标数据标准化。使用最小-最大法对指标归一化,根据贡献方式分为正向与负向指标,表达式为:

$$\text{正向: } X_{ij} = \frac{x_{ij} - \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}}{\max\{x_{1j}, x_{2j}, \dots, x_{nj}\} - \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}} \quad (2)$$

$$\text{负向: } X_{ij} = \frac{\max\{x_{1j}, x_{2j}, \dots, x_{nj}\} - x_{ij}}{\max\{x_{1j}, x_{2j}, \dots, x_{nj}\} - \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}} \quad (3)$$

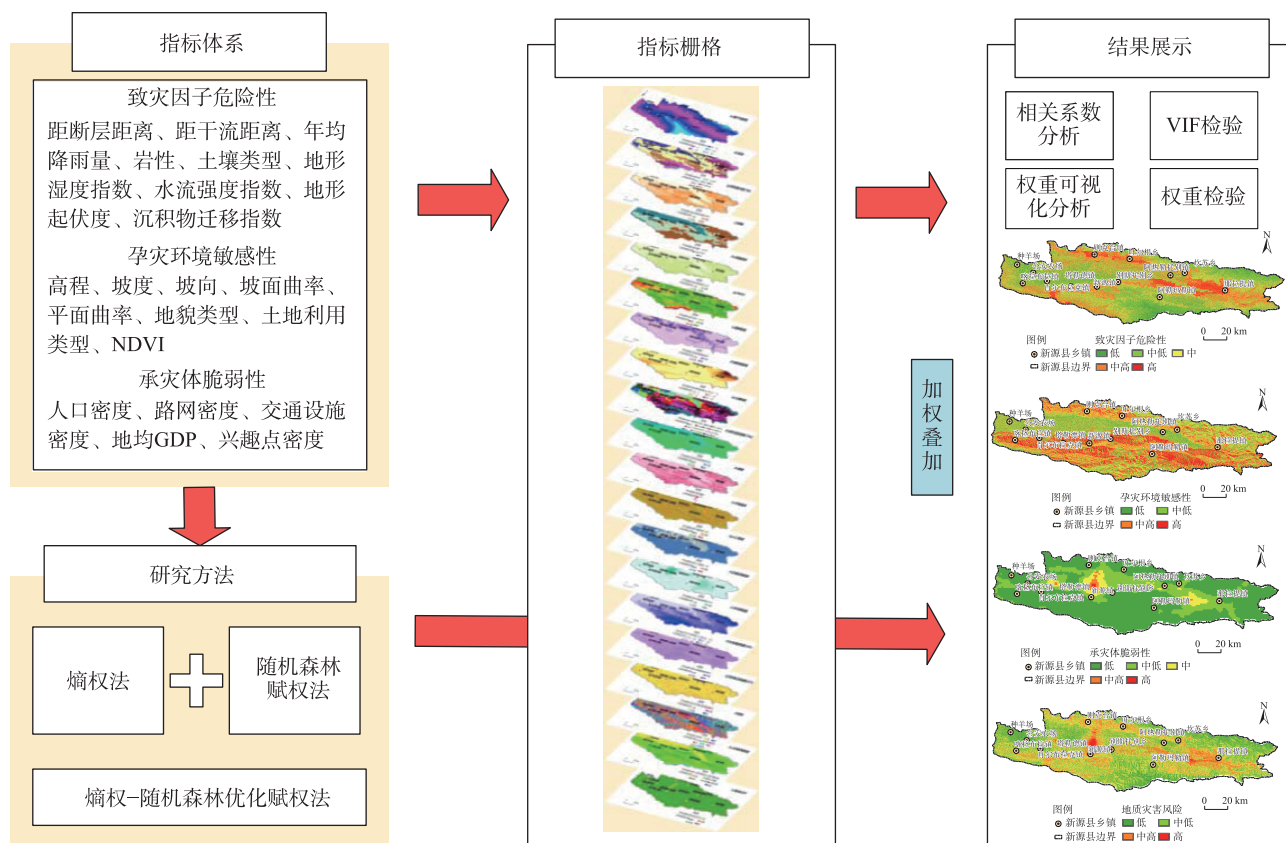
式中:  $X_{ij}$  为第  $i$  个区域第  $j$  个指标的标准化数据;  $x_{ij}$  为对应的原始数据;  $x_{nj}$  为  $i$  达到最大值  $n$  时的原始数据。

(2) 非负平移处理:

$$X'_{ij} = X_{ij} + \varepsilon \quad (4)$$

式中:  $X'_{ij}$  为非负平移处理后的数据;  $\varepsilon$  为极小值,





注:NDVI为归一化植被指数;VIF为方差膨胀因子。下同。

图2 风险评估流程图

Fig. 2 Flowchart of risk assessment

设置为0.0001,对实际结果不产生影响<sup>[16]</sup>。

(3) 第*i*区域第*j*指标占该指标和的比例( $P_{ij}$ ):

$$P_{ij} = \frac{X'_{ij}}{\sum_{i=1}^n X'_{ij}} \quad (5)$$

(4) 第*j*项指标的熵值( $E_j$ )、信息熵冗余度( $D_j$ )与权重( $W_j$ ):

$$k = 1/\ln(n) \quad (6)$$

$$E_j = -k \sum_{i=1}^n [P_{ij} \times \ln(P_{ij})] \quad (7)$$

$$D_j = 1 - E_j \quad (8)$$

$$W_j = \frac{D_j}{\sum_{j=1}^n D_j} \quad (9)$$

式中: $k$ 为保证 $E_j$ 归一化在[0, 1]区间的常数。

随机森林(Random forest, RF)是一种基于自助采样的集成学习算法,通过集成多棵决策树形成“森林”,并统计每棵树结果,通过投票或回归得到

最终结果<sup>[17]</sup>。使用随机森林方法进行指标重要性评估已经被应用在各个领域<sup>[18-19]</sup>,通过基于矩阵随机置换的残差均方减小量来评估指标重要性,以此得到各指标权重。

熵权法权重作为初始权重与栅格数据进行加权计算,得到初始结果;栅格数据与初始结果分别作为随机森林模型的输入输出数据进行训练,得到的特征重要性即为指标优化权重。随机森林训练采用5折交叉验证,每折内按8:2比例划分训练集与测试集,由网格搜索法确定最优超参数,同时结合随机森林自身特性来降低过拟合风险。训练完成后,对每折的特征重要性结果进行平均和归一化处理,以此提高优化权重的稳健性,流程图如图3所示。

## 2 结果与分析

### 2.1 风险评估指标体系

**2.1.1 指标选取** 地质灾害风险受到多方因素的影响,需考虑各类影响因素构建指标体系。通过参考



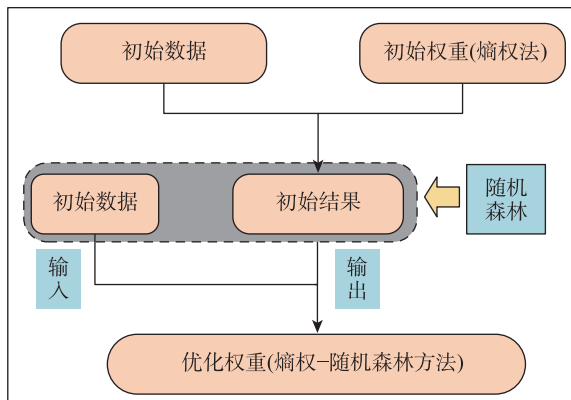


图3 熵权-随机森林优化赋权流程图

Fig. 3 Flowchart of entropy weight-random forest optimization weighting

相关文献<sup>[20-21]</sup>,在构建指标体系时以灾害形成机制为基础,结合地质条件、水文气象等地质灾害发育的必备条件,地形特征与土地类型等孕灾环境,人类、经济与基础设施等承灾体数据,在危险性-敏感性-脆弱性三维度下初步选取22个指标作为新源县地质灾害评价指标。

危险性受地质构造和水文气象2类因子共同影响,新源县具备复杂的地质环境与水文地质条件,且降雨分布不均,这些因素共同提供地质灾害发育条件;敏感性则受地形特征与土地类型2类因子影响,高程、坡度、坡向等地形因子,地貌类型、土地利用类型与归一化植被指数(Normalized difference vegetation index, NDVI)分别从土地的地形特点、自然地形特征、使用状况、植被覆盖程度3方面表征孕灾环境;承灾体以人口、经济与基础设施为主,以人口密度、地均国内生产总值(Gross domestic product, GDP)分别表征人口、经济脆弱性,路网密度、兴趣点(Point of interest, POI)密度与交通设施密度则表征基础设施脆弱性,并统一综合为承灾体脆弱性。

基于以上分析,按照危险性-敏感性-脆弱性三维度进行指标选取:距断层距离、距干流距离、岩性、土壤类型、地形湿度指数(Topographic wetness index, TWI)、水流强度指数(Stream power index, SPI)、沉积物迁移指数(Sediment transport index, STI)、地形起伏度(Terrain ruggedness index, TRI)和年均降雨量作为致灾因子危险性指标;以高程、坡度、坡向、平面曲率、坡面曲率、土地利用类型、地貌类型和NDVI作为孕灾环境敏感性指标;以人口密

度、路网密度、交通设施密度、地均GDP和POI密度作为承灾体脆弱性指标。为确保指标体系科学合理,后续会通过Pearson相关系数与方差膨胀因子(Variance inflation factor, VIF)检验提出冗余或多重共线指标,从而形成合理的指标体系。其中部分指标计算公式如下:

$$SPI = A_s \tan \beta \quad (10)$$

$$TWI = \ln \frac{A_s}{\tan \beta} \quad (11)$$

$$STI = \frac{A_s \tan \beta}{E} \quad (12)$$

$$TRI = |E_i - E_{i-1}| \quad (13)$$

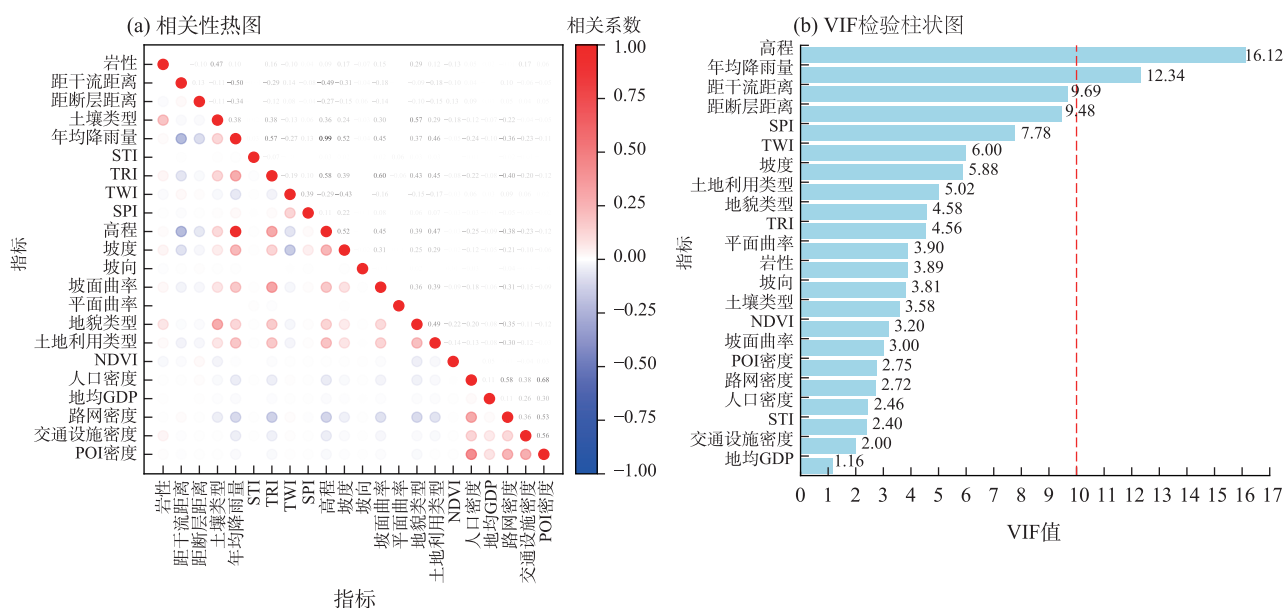
式中:  $A_s$  为汇水面积;  $\beta$  为坡度;  $E$  为高程值;  $|E_i - E_{i-1}|$  为相邻点的高程差值。

**2.1.2 相关性分析** 基于多重共线性的考虑,在构建指标体系后需评估指标之间的相关性强度<sup>[22]</sup>。本研究从Pearson相关系数与VIF检验2个角度进行多重共线性检验:Pearson相关系数被用于量化指标间的线性相关性,取值范围为 $[-1, 1]$ ,其中-1代表完全负相关,1代表完全正相关,0代表无相关;VIF检验则用于检验指标之间的多重共线性程度,VIF值高表明该指标能够被多个指标共同解释。图4a与图4b分别为相关性热图与VIF检验柱状图。

当相关系数或VIF值达到阈值时,即表明存在多重共线性问题。张向宁等<sup>[23]</sup>定义相关系数绝对值大于0.9为高度关联,而VIF值大于10时通常被认为存在多重共线性。经检验,高程与年均降雨量的相关系数大于0.9,且VIF值均超过10,说明二者存在严重的多重共线性。高程能够被派生指标间接反映,而降水是地质灾害的直接驱动因子,最终本文筛除高程指标。值得一提的是,随机森林由于训练过程中通过样本和特征的随机抽样,在面对多重共线性时仍能保持稳健性,已有文献证明了这一优势<sup>[24]</sup>。筛查后的风险评估指标体系栅格图如图5所示。

## 2.2 权重优化结果

优化赋权法得到的指标权重与熵权法、等权法权重进行了对比(图6)。结果显示,致灾因子中SPI、TWI、STI等权重明显提升,而TRI、岩性、土壤类型等下降,反映了水动力触发因素对地质灾害危险性的更直接影响。孕灾环境敏感性指标中,坡向和平面曲率权重提升,其他指标有所下降,说明坡向



注:STI为沉积物迁移指数;TRI为地形起伏度;TWI为地形湿度指数;SPI为水流强度指数;POI为兴趣点。下同。

图4 多重共线性检验

Fig. 4 Multicollinearity test

在积水及地质灾害孕育中的重要性。承灾体脆弱性指标中,路网密度、交通设施密度与POI密度权重有所增加,而人口密度和地均GDP权重虽略降但仍占重要比例,表明人口、经济和基础设施在脆弱性评估中同样关键。

为验证优化权重可靠性,本研究通过受试者工作特征曲线(Receiver operating characteristic, ROC)与曲线下面积(Area under the curve, AUC)对权重效果进行评估。使用新源县历史地质灾害点数据集,并按1:2比例随机生成非灾害点构建样本集,采集各样本指标值后结合3种权重方法,逐样本生成地质灾害风险值。将风险值作为预测分数,与实际灾害发生情况进行比较,计算不同阈值下的真正率与假正率,绘制曲线并计算AUC,以评估各赋权方法在识别高风险区域的可靠性(图7)。结果显示,3种方法的AUC均高于0.7,表明均能有效区分高风险区域,其中优化赋权法的AUC最高达到0.776,进一步验证优化权重的可靠性。

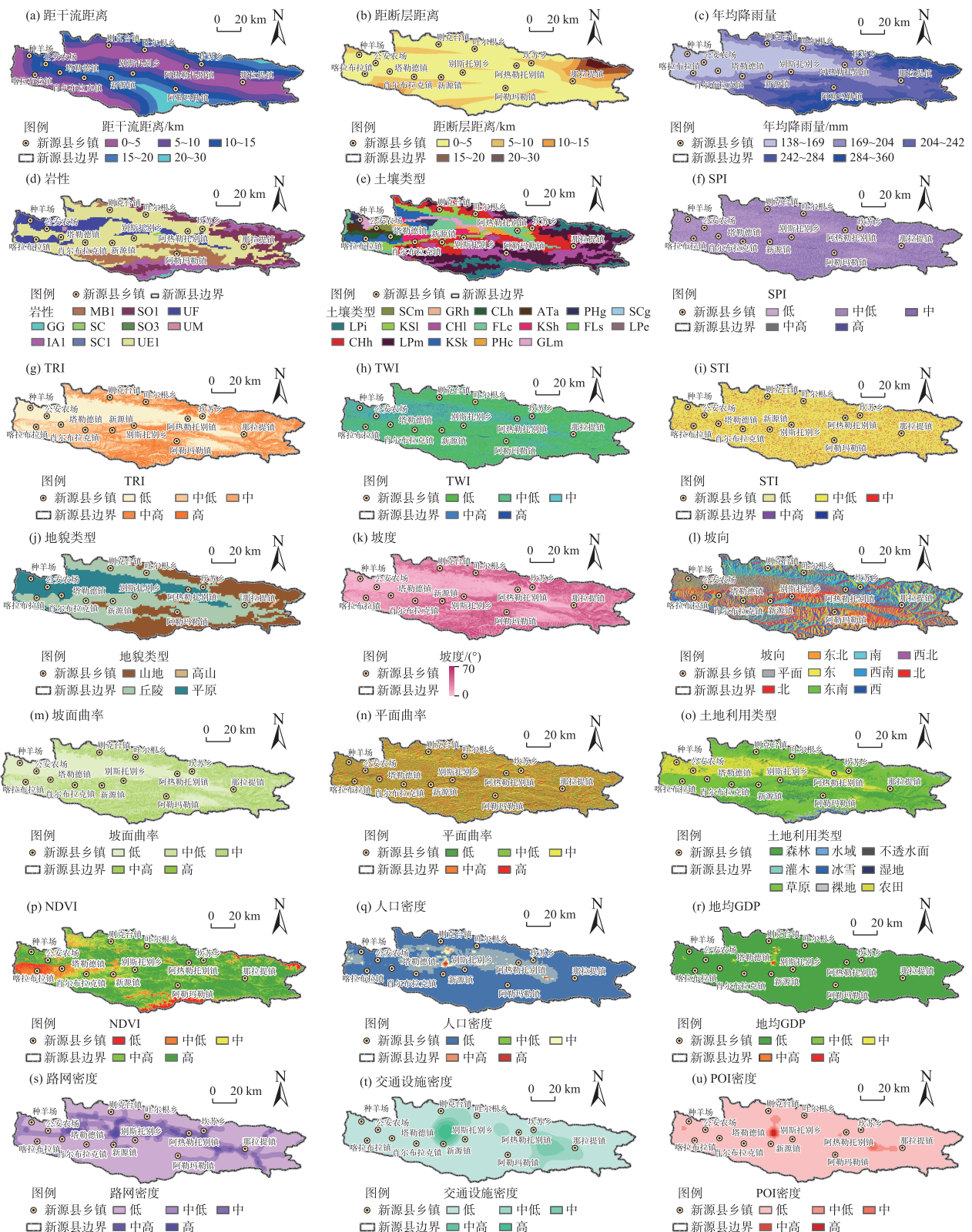
## 2.3 评估结果分析

**2.3.1 致灾因子危险性** 致灾因子危险性用于衡量地质灾害发生的潜在条件强度,是区域风险评估的重要组成部分之一。如图8a所示,新源县危险性分布整体呈显著的区域差异,有权重可知水文条件在灾害发育中占据主导地位。高危地区主要分布在

新源县东部与北部乡镇部分区域,这些区域地形汇水集中导致SPI和STI较高,形成了有利于松散物质滑移的地质水文条件。同时,这些地区岩性多为风成沉积岩,土壤类型以黑钙土为主,进一步增强了区域滑移的可能性。相比之下,阿勒玛勒乡南部和那拉提镇北部山区则由于地质构造活动较弱,同时地形封闭、汇水能力差,使得地表水动力过程不明显,危险性较低;新源县西北部受限于低降水量,难以形成有效的侵蚀与搬运过程,亦表现为低危险性区域。

**2.3.2 孕灾环境敏感性** 孕灾环境敏感性反映研究区环境与致灾因子形成发育之间的关系,同样是在区域风险评估中非常重要。如图8b所示,新源县敏感性空间分布主要受地形因素调控,其中坡向是敏感性水平的主控指标。高敏感区分布在矮山缓坡地带,这种区域易累积沉积物,同时以含水率较高的阴坡与半阴坡为主,导致其敏感性最高。低敏感区分布在新源县西部的河谷平原地形与南北两侧的高山地形,河谷平原地形坡度极低且坡向分布不均导致无法形成有效的孕灾结构,而高山地形由于不易累积沉积物使得临界土体难以形成,因此是敏感程度较低的孕灾环境。

**2.3.3 承灾体脆弱性** 承灾体脆弱性反映人口、经济等承灾体能够承受灾害影响的能力,是区域风险



注:MB1为板岩、千层岩;SO1、SO3为石灰岩;UF为冲沉岩;GG为基岩;SC为碎屑沉积岩;UM为洋沉积;IA1为花岗岩;SC1为砾岩、角砾岩;UE1为风成沉积物;SCm为暗沃盐土;GRh为普通灰色森林土;CLh为普通钙土;ATa为埃里克人为土;PHg为潜育黑土;SCg为潜育盐土;LPi为寒冻薄层土;KSI为淋溶栗钙土;CHI为淋溶黑钙土;FLc为石灰性冲积土;KSh为普通栗钙土;FLs为盐积冲积土;LPe为饱和薄层土;CHh为普通黑钙土;LPm为暗沃薄层土;KSk为钙积栗钙土;PHc为石灰性湿草原土;GLm为暗沃潜育土。

图5 指标栅格图

Fig. 5 Grid maps of indicators



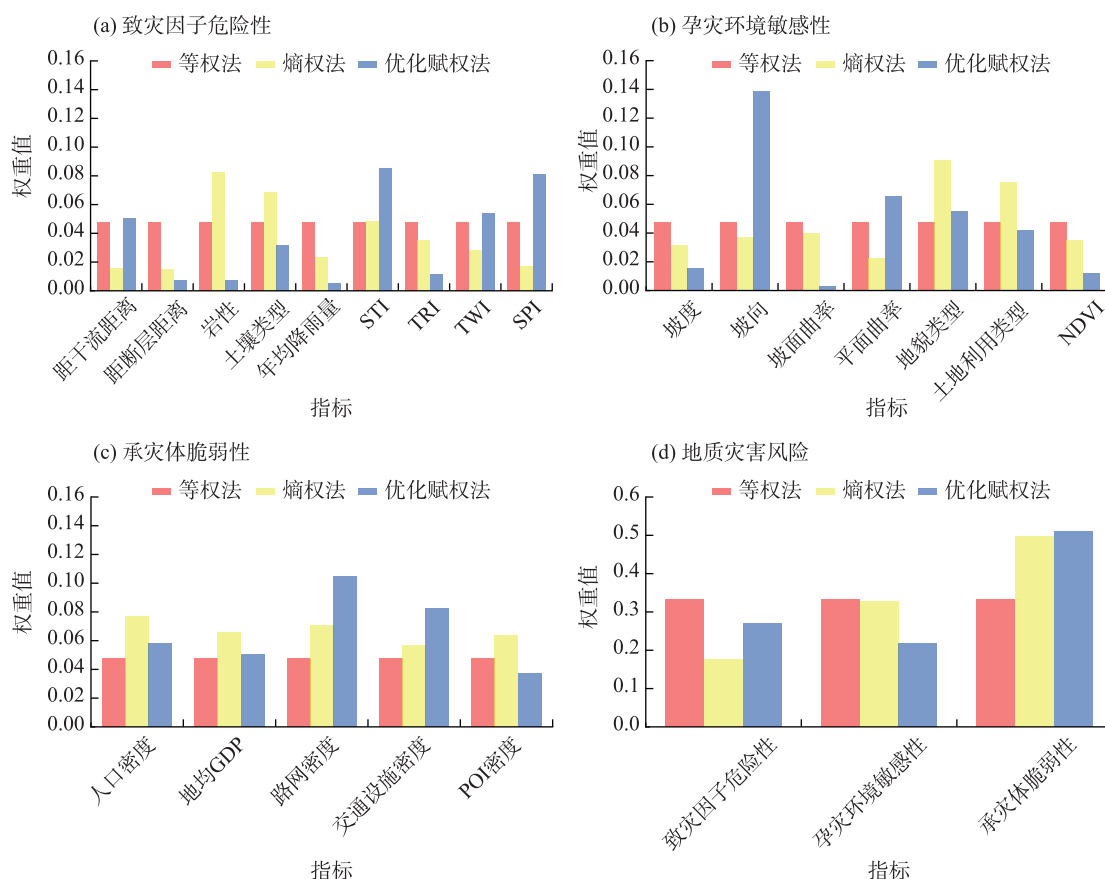


图6 权重对比

Fig. 6 Comparison of indicator weights

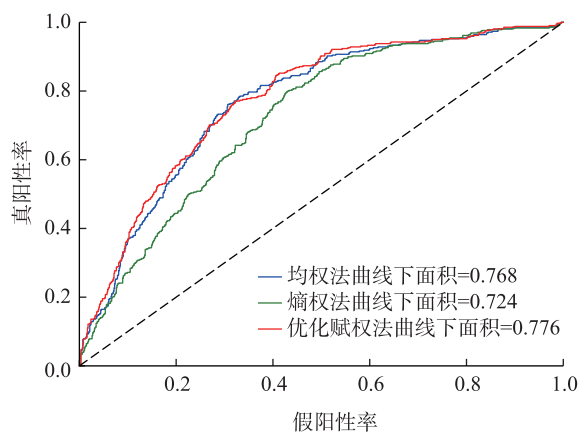


图7 不同赋权方法的受试者工作特征曲线

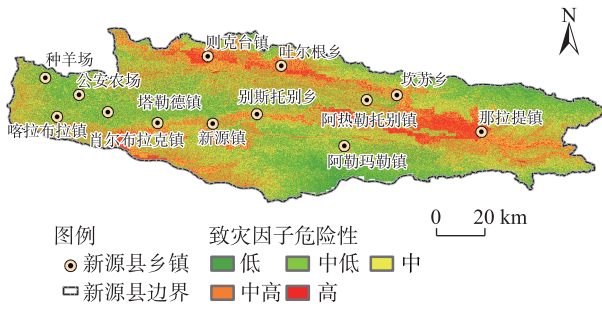
Fig. 7 Receiver operating characteristic curve of different weighting methods

组成要素之一。如图8c所示,空间上新源县脆弱性主要沿路网呈带状分布,部分区域呈环状聚集。脆弱性各指标贡献较为均衡,重要的承灾体都要被考虑。新源镇与别斯托别乡交界处和那拉提镇中心区域因其人口经济水平高的特点有着最高的脆弱

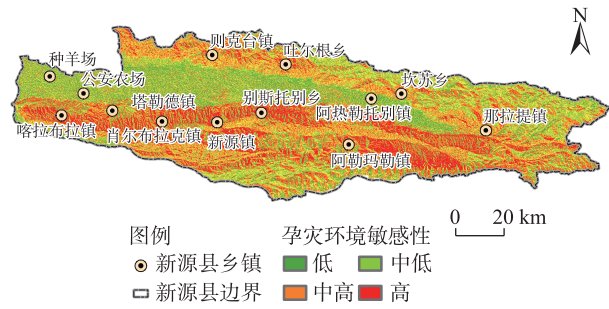
性。由新源镇、别斯托别乡、那拉提镇中心向外扩散,形成的外围区域人口经济水平略低于中心区域但显著高于其余乡镇,所以脆弱性水平同样偏高。道路作为重要的基础设施与承灾体,其存在显著影响脆弱性水平,所以仅有道路分布的区域脆弱性均高于中低水平。由于新源县特殊的地理环境,除去种羊场到那拉提镇的河谷平原区域,其余区域均为无明显承灾体,脆弱性水平最低。

**2.3.4 地质灾害风险** 风险结果由致灾因子危险性、孕灾环境敏感性、承灾体脆弱性共同决定,体现了区域范围内风险水平空间分布特征。如图8d所示,新源县地质灾害风险与脆弱性空间分布高度一致,表明脆弱性水平是影响高风险的关键因素。高风险区主要集中在新源镇、别斯托别乡和那拉提镇,这些区域因人口高度聚集、交通物流密集及旅游经济发达,显著提升了脆弱性,从而使风险水平达到最高。具体来说,新源镇作为行政中心拥有12处公共服务设施及万人居住区,别斯托别乡沿G218国道分布了32处工商业设施,那拉提镇则是5A景

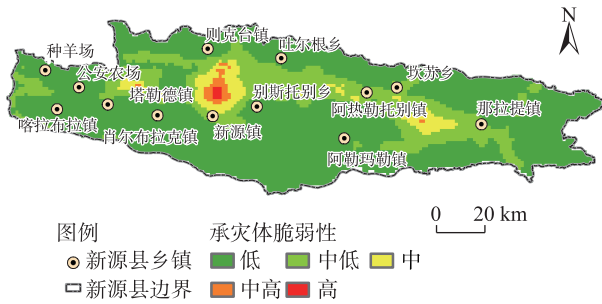
(a) 致灾因子危险性



(b) 孕灾环境敏感性



(c) 承灾体脆弱性



(d) 地质灾害风险

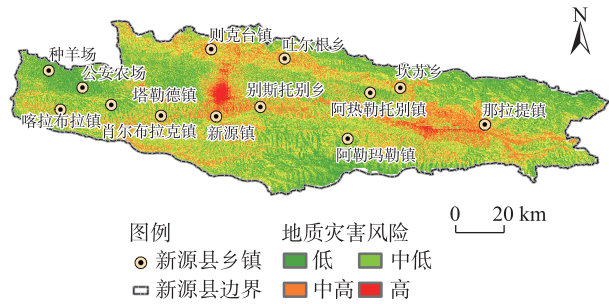


图8 地质灾害风险及风险组成要素地图

Fig. 8 Map of geological disaster risk and risk components

区,日均接待游客达 $1.2 \times 10^4$ 人次,均存在重要承灾体。高风险区体现了短板效应,即使危险性与敏感性未达到最高水平,但受脆弱性这一“短板因素”影响,其风险仍显著高于其他区域。中高风险区主要分布在那拉提镇中部滑雪场及民宿集群、阿勒玛勒镇北部矿区和则克台镇东部工业园区等脆弱性略低区域,在三要素共同作用下其风险相对较高。中风险区分布在喀拉布拉镇农产品带、吐尔根乡杏花谷等地,这些区域在特定季节人口和活动密度明显升高,导致其具有一定的风险水平。种羊场河漫滩区域及那拉提镇北部生态保护区等地因地质环境稳定且缺少明显承灾体,整体风险水平较低。总体来看,新源县地质灾害风险空间分布清晰,表明城镇功能集聚区和资源开发区是地质灾害防控的重点区域,应实施针对性的防控措施:高风险区应加强人口密集区、交通枢纽等地的地灾防控设施建设,并完善应急疏散系统;中高风险区可以通过加固关键基础设施、优化土地利用规划等措施降低风险;中及以下风险区则应实现常态防控,以日常巡查、公众教育等维持风险水平稳定的措施为主。

## 2.4 风险验证

为验证本文风险评估结果的合理性与准确性,

本文选取新源县历年来的地质灾害事件与3种典型承灾体(经济、人口、基础设施)作为验证样本与风险评估结果进行空间叠加分析。地质灾害事件选择地质灾害点进行表征;典型承灾体验证则选择POI点表示经济承灾体,人口与人口密度表示人口承灾体,道路长度与道路密度表示基础设施承灾体。基于以上数据分布展开风险结果验证,结果如图9所示。

如图9灾害点验证结果显示大多数地质灾害事件分布在中风险区与中高风险区,分别达到39.7%与31.0%。这些区域多分布在河谷平原周边的矮山缓坡地形,地质灾害易发。过往研究表明<sup>[1,10-11,25-27]</sup>,吐尔根乡、则克台镇和塔勒德镇等河谷周边乡镇地质灾害发生最频繁,与本研究结论相符。值得注意的是,高风险区虽脆弱性较高,但危险性与敏感性水平有限,导致历史灾害事件较少。

近年来研究表明,承灾体分布是影响灾害风险的重要原因。Bastos等<sup>[28]</sup>在研究中指出,大规模城市扩张导致承灾体脆弱性上升,与2001年相比建筑物受灾害损失提高了46%;Mosaffaie等<sup>[29]</sup>的研究得到84.35%的高风险区与91.21%的中高风险区位于人口居住区的结论,这些结论与本研究统计结果相

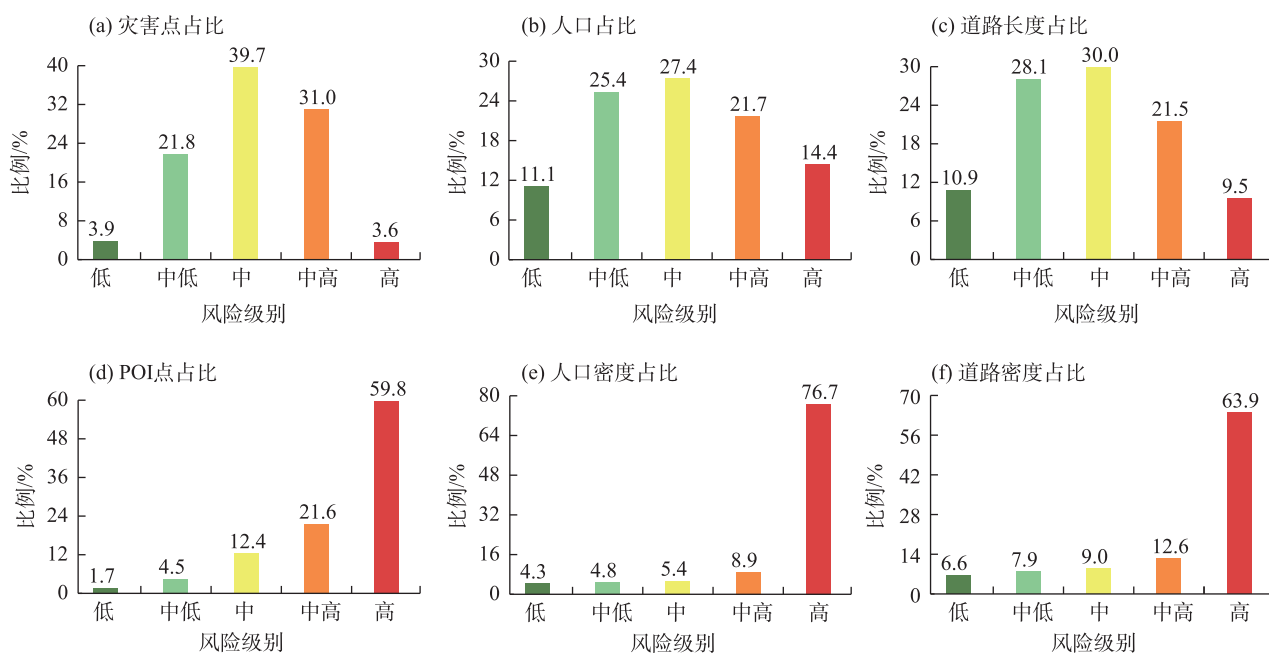


图9 风险评估结果验证

Fig. 9 Validation of risk assessment results

一致。承灾体验证显示:在经济承灾体方面,新源县共有3007个POI点,其中59.8%分布在高风险区,而中高级以上风险区占比高达93.8%,表明风险与经济承灾体在空间分布上高度一致;人口承灾体与基础设施承灾体分布规律类似,在高风险区呈现明显的聚集特性,分布其中的人口密度与道路密度分别达到76.7%与63.9%。综上所述,历史灾害事件及经济、人口、基础设施3类承灾体的空间分布与风险评估结果高度匹配,进一步支持了风险评估结果的科学性和准确性。

经过验证比较与历史研究相比,本文综合考虑了孕灾环境与承灾体的空间分布,使风险评估更加系统和全面。研究结果显示,高风险区主要集中在具有明显承灾体的区域,而易发性高但承灾体水平较低的区域风险则略低于高风险区。这表明承灾体分布在地质灾害风险空间格局中具有关键作用,也凸显了熵权-RF优化赋权法在综合风险评估中的优势。

### 3 结论

(1) 熵权-RF优化赋权法有效融合客观分布特征与灾害发生机理,使权重分布更符合实际。危险性受水动力因子(SPI、TWI、STI、距干流距离)主导,

坡向和平面曲率主控敏感性,路网与交通设施密度显著提升脆弱性解释力,人口密度和地均GDP仍保持较高影响力;同时绘制ROC曲线,可知优化赋权法AUC最高,达到0.776。

(2) 危险性、敏感性、脆弱性空间分布清晰,主控因子作用区域分异显著。高危险性区集中于地质结构复杂、水动力强的新源县东部与北部;高敏感性区多分布于阴坡缓坡地形;高脆弱性区对应乡镇区域、交通走廊和经济节点,呈现显著空间演变特征。

(3) 地质灾害风险空间格局与高脆弱性区高度相似。高风险区集中在人口密集、功能集聚的镇区核心,虽致灾条件一般,但因脆弱性极高导致整体风险升高;中高风险区多为孕灾环境与承灾体叠加区域;中低与低风险区则分布于无显著承灾体的山区或生态稳定区,整体风险水平较低。

(4) 历史灾害事件与典型承灾体分布在空间上与风险评估结果高度匹配,验证了其合理性及准确性。灾害点主要分布于中及中高风险区(70.7%),表明评估结果能有效反映历史灾害分布;POI点中有59.8%位于高风险区,人口密度在高风险区占比达76.7%,道路密度达63.9%,3类承灾体均呈现出与高风险区高度重合的空间格局,验证结果与现实



风险分布高度一致。

本研究综合考虑孕灾环境与承灾体空间分布,从灾害系统角度构建三维度风险评估体系,并对风险评估结果进行验证分析,使得评估结果更加准确、系统和全面。结果显示高风险区主要集中在承灾体比较集中的区域,而传统高易发区中承灾体水平较低的区域则整体风险略低。这表明承灾体分布在地质灾害风险空间格局中具有关键作用,也凸显了熵权-RF 优化赋权法在综合风险评估中的优势。同时,在许多方面仍存在一定的不足与未来发展空间,首先研究内容主要为静态风险评估,在后续研究可以尝试通过动态指标构建动态风险评估体系,以实现时地灾害风险的时序动态分析;其次未来会尝试将研究区从新源县扩展到伊犁河谷,从而提升区域适用性,为伊犁河谷地区地质灾害防控提供一定参考。

## 参考文献 (References)

- [1] 张策, 揭文辉, 付丽华, 等. 新疆新源县滑坡灾害遥感影像特征及分布规律[J]. 国土资源遥感, 2017, 29(增刊 1): 81-84. [Zhang Ce, Jie Wenhui, Fu Lihua, et al. Remote sensing image and distribution characteristics of landslide disasters in Xinyuan County, Xinjiang[J]. Remote Sensing for Land and Resources, 2017, 29 (Suppl. 1): 81-84. ]
- [2] 史培军. 三论灾害研究的理论与实践[J]. 自然灾害学报, 2002, 11(3): 1-9. [Shi Peijun. Theory on disaster science and disaster dynamics[J]. Journal of Natural Disasters, 2002, 11(3): 1-9. ]
- [3] 熊凡, 李沛鸿, 袁逸敏, 等. 基于 GIS 和组合赋权的城市洪涝灾害风险评估[J]. 人民长江, 2023, 54(8): 60-66. [Xiong Fan, Li Peihong, Yuan Yimin, et al. Urban flood risk assessment based on GIS and combined weights[J]. Yangtze River, 2023, 54(8): 60-66. ]
- [4] 郑德凤, 高敏, 李钰, 等. 基于 GIS 的大连市暴雨洪涝灾害综合风险评估[J]. 河海大学学报(自然科学版), 2022, 50(3): 1-8, 2. [Zheng Defeng, Gao Min, Li Yu, et al. Comprehensive risk assessment of rainstorm-flood disaster in Dalian City based on GIS[J]. Journal of Hohai University (Natural Sciences), 2022, 50(3): 1-8, 2. ]
- [5] 费杜秋, 刘峰贵, 周强, 等. 青藏铁路沿线滑坡泥石流灾害风险分析[J]. 干旱区地理, 2016, 39(2): 345-352. [Fei Duqiu, Liu Fenggui, Zhou Qiang, et al. Risk analysis of landslide and debris flow disasters along the Qinghai-Tibet Railway[J]. Arid Land Geography, 2016, 39(2): 345-352. ]
- [6] 罗紫元, 田健, 丁懿媛, 等. 基于熵权 TOPSIS 和神经网络复合方法的厦门市雨洪灾害风险评估[J]. 灾害学, 2022, 37(4): 184-192. [Luo Ziyuan, Tian Jian, Ding Siyuan, et al. Hybrid EWM-TOPSIS and neural network for flood risk evaluation of Xiamen City[J]. Journal of Catastrophology, 2022, 37(4): 184-192. ]
- [7] 谢捷, 刘玮, 徐月顺, 等. 基于 AHP-熵权法的西宁地区汛期暴雨灾害风险评估[J]. 自然灾害学报, 2022, 31(3): 60-74. [Xie Jie, Liu Wei, Xu Yueshun, et al. Rainstorm disaster risk assessment in Xining area in rainy season based on the AHP weight method and entropy weight method[J]. Journal of Natural Disasters, 2022, 31(3): 60-74. ]
- [8] 袁旭山, 刘京会, 宋珂. 基于 BP 神经网络的洪涝灾害承灾体脆弱性评估[J]. 人民长江, 2024, 55(2): 26-34. [Yuan Xushan, Liu Jinghui, Song Ke. Assessment on vulnerability of flood disaster bearing body based on BP neural network[J]. Yangtze River, 2024, 55(2): 26-34. ]
- [9] Al-Abadi A M. Mapping flood susceptibility in an arid region of southern Iraq using ensemble machine learning classifiers: A comparative study[J]. Arabian Journal of Geosciences, 2018, 11(9): 218, doi: 10.1007/s13201-015-0283-1.
- [10] 周平华, 李英成, 白洁, 等. 新源县汇水域单元滑坡灾害易发性评价[J]. 测绘科学, 2017, 42(9): 82-87. [Zhou Pinghua, Li Yingcheng, Bai Jie, et al. Susceptibility assessment of landslide hazard to catchment unit in Xinyuan County[J]. Science of Surveying and Mapping, 2017, 42(9): 82-87. ]
- [11] 常顺利, 张钟月, 孙志群, 等. 基于 GIS 的新源县滑坡灾害分析与区划[J]. 自然灾害学报, 2011, 20(5): 216-221. [Chang Shunli, Zhang Zhongyue, Sun Zhiqun, et al. GIS-based analysis and zoning of landslide hazard in Kunes County[J]. Journal of Natural Disasters, 2011, 20(5): 216-221. ]
- [12] 程乙峰, 刘志辉. 基于证据权模型的滑坡影响因子研究[J]. 新疆地质, 2016, 34(2): 275-279. [Cheng Yifeng, Liu Zhihui. A research based on weight of evidence model landslide impact factor [J]. Xinjiang Geology, 2016, 34(2): 275-279. ]
- [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 Zhongguo, Zhang Feng, Zheng Jianghua. Evaluation of landslide susceptibility by multiple adaptive regression spline method[J]. Geomatics and Information Science of Wuhan University, 2021, 46 (3): 442-450. ]
- [14] 史培军. 五论灾害系统研究的理论与实践[J]. 自然灾害学报, 2009, 18(5): 1-9. [Shi Peijun. Theory and practice on disaster system research in a fifth time[J]. Journal of Natural Disasters, 2009, 18(5): 1-9. ]
- [15] Zhu Y X, Tian D Z, Yan F. Effectiveness of entropy weight method in decision-making[J]. Mathematical Problems in Engineering, 2020 (1): 3564835, doi: 10.1155/2020/3564835.
- [16] 吴佳瑶, 张煜洲, 陈柯, 等. 基于 AHP-熵权法的城市内涝风险评估——以杭州市滨江区为例[J]. 浙江水利水电学院学报, 2024, 36(1): 38-45. [Wu Jiayao, Zhang Yuzhou, Chen Ke, et al. Urban waterlogging risk assessment based on AHP-entropy method: Case study of Binjiang District[J]. Journal of Zhejiang University

- ty of Water Resources and Electric Power, 2024, 36(1): 38–45. ]
- [17] Bersabe J T, Jun B W. The machine learning-based mapping of urban pluvial flood susceptibility in seoul integrating flood conditioning factors and drainage-related data[J]. ISPRS International Journal of Geo-Information, 2025, 14(2): 57, doi: 10.3390/ijgi14020057.
- [18] 郭桂祯, 刘乃山. 基于随机森林模型的区域洪涝灾害房屋倒塌评估方法研究[J]. 灾害学, 2018, 33(3): 48–51. [Guo Guizhen, Liu Naishan. A method research of house damage assessment based on random forest model of regional flood disaster[J]. Journal of Catastrophology, 2018, 33(3): 48–51. ]
- [19] 郭四代, 袁子寒, 雷高文. 企业碳信息披露质量评价及影响因素研究[J]. 地球环境学报, 2023, 14(6): 848–860. [Guo Sida, Yuan Zihan, Lei Gaowen. Research on quality evaluation and influencing factors of corporate carbon information disclosure[J]. Journal of Earth Environment, 2023, 14(6): 848–860. ]
- [20] 刘亚静, 刘红健. 基于信息量-随机森林模型的地震带地质灾害易发性评价: 以松潘-较场地震带为例[J]. 科学技术与工程, 2024, 24(1): 143–154. [Liu Yajing, Liu Hongjian. Evaluation of geological hazard susceptibility in seismic zone based on information data-RF model: A case study of Songpan-Jiaochang seismic zone[J]. Science Technology and Engineering, 2024, 24(1): 143–154. ]
- [21] 马红娜, 刘江, 冯卫, 等. 地质灾害风险评估在国土空间规划中的应用——以陕北榆林高西沟为例[J]. 西北地质, 2023, 56(3): 223–231. [Ma Hongna, Liu Jiang, Feng Wei, et al. Application of geological hazard risk assessment in territorial space planning: A case study of Gaixigou Village in Yulin City of northern Shaanxi Province[J]. Northwestern Geology, 2023, 56(3): 223–231. ]
- [22] Alin A. Multicollinearity[J]. Wiley Interdisciplinary Reviews: Computational Statistics, 2010, 2(3): 370–374.
- [23] 张向宁, 王小军, 齐广平, 等. 基于DPSR-信息敏感性的水安全评价指标体系构建——以庆阳市为例[J]. 干旱区资源与环境, 2022, 36(9): 44–53. [Zhang Xiangning, Wang Xiaojun, Qi Guangping, et al. Construction of water security evaluating index system based on driving force-pressure-state-response model and information sensitivity[J]. Journal of Arid Land Resources and Environment, 2022, 36(9): 44–53. ]
- [24] Strobl C, Boulesteix A L, Kneib T, et al. Conditional variable importance for random forests[J]. BMC Bioinformatics, 2008, 9(1): 307, doi: 10.1186/1471-2105-9-307.
- [25] 袁旭山, 刘京会, 黄龙生, 等. 基于机器学习模型的新源县滑坡易发性评估[J]. 科学技术与工程, 2025, 25(5): 1815–1826. [Yuan Xushan, Liu Jinghui, Huang Longsheng, et al. Assessment of landslide susceptibility in Xinyuan County based on machine learning models[J]. Science Technology and Engineering, 2025, 25(5): 1815–1826. ]
- [26] 梁世川, 乔华, 吕东, 等. 伊犁谷地地质灾害分布特征及主控因素分析[J]. 干旱区地理, 2023, 46(6): 880–888. [Liang Shichuan, Qiao Hua, Lü Dong, et al. Distribution characteristics and main controlling factors of geohazards in Ili Valley[J]. Arid Land Geography, 2023, 46(6): 880–888. ]
- [27] 陈世龙, 孟庆凯, 戴勇, 等. 基于CMIP6未来情景的伊犁河流域地质灾害危险性评估预测[J]. 干旱区地理, 2025, 48(4): 599–611. [Chen Shilong, Meng Qingkai, Dai Yong, et al. Geological disaster hazard assessment and prediction in the Ili River Basin based on CMIP6 future scenarios[J]. Arid Land Geography, 2025, 48(4): 599–611. ]
- [28] Bastos M C, Thieken A H. Urban growth and spatial segregation increase disaster risk: Lessons learned from the 2023 disaster on the north coast of São Paulo, Brazil[J]. Natural Hazards and Earth System Sciences, 2024, 24(9): 3299–3314.
- [29] Mosaffaie J, Salehpour Jam A, Sarfaraz F. Landslide risk assessment based on susceptibility and vulnerability[J]. Environment, Development & Sustainability, 2024, 26(4): 9285–9303.

## Geological disaster risk assessment of Xinyuan County based on entropy weight-random forest optimization weighting method

LIU Jinghui, LI Xinxu, YUAN Xushan, LI Yanmin

(School of Emergency Technology and Management, Institute of Disaster Prevention, Sanhe 065201, Hebei, China)

**Abstract:** Geological disasters frequently occur in the Ili River Basin, and Xinyuan County is among the most disaster-prone areas. This study focuses on Xinyuan County and establishes a three-dimensional indicator system encompassing the hazard of triggering factors, the susceptibility of disaster-prone environments, and the vulnerability of exposed elements. An optimized weighting scheme was developed by integrating the entropy weight method with the Random Forest model, and the comprehensive geological disaster risk of the county was evaluated using the ArcGIS platform. The results indicate that (1) After weight optimization, hazard was primarily governed by hydrodynamic factors, susceptibility was strongly controlled by topographic factors, and vulnerability was mainly determined by transportation networks and population-economic characteristics. The optimized weights yielded higher area under curve values than those obtained using the traditional entropy weight method. (2) The spatial distributions of hazard, susceptibility, and vulnerability differed significantly, corresponding to areas with complex geological conditions, gentle shaded slopes, and population-economic clusters, respectively. (3) The overall spatial pattern of geological disaster risk closely corresponded to the distribution of vulnerability, with high-risk areas concentrated in town centers and along major roads, medium-high-risk areas mainly located where disaster-prone environments overlapped with exposed elements, and low-risk areas distributed in ecologically stable zones or regions lacking significant exposed elements. (4) High-risk zones exhibited strong spatial consistency with historical disaster points and typical exposed elements, thereby confirming the rationality and reliability of the assessment results. The findings reveal the spatial distribution pattern of geological disaster risk in Xinyuan County and provide a scientific basis and theoretical reference for disaster prevention and mitigation.

**Keywords:** risk assessment; geological disasters; entropy weight method; random forest; Xinyuan County