

基于机器学习算法的洪涝灾害风险评估 — 以宜昌市为例¹

王德运^{1,2}, 张露丹¹, 吴 祈¹, 郭海湘^{1,2}, 柯小玲¹, 吕新彪³

(1. 中国地质大学(武汉)经济管理学院, 湖北 武汉 430074; 2. 中国地质大学(武汉)自然灾害风险防控与应急管理实验室, 湖北 武汉 430074; 3. 中国地质大学(武汉)高等研究院, 湖北 武汉 430074)

【摘要】: 近年来, 城市洪涝事件频发, 对人民的生命和财产安全造成了严重影响。客观、准确的风险评估对于提升城市洪涝灾害防控水平至关重要, 以宜昌市 2020 年 6 月末遭受的洪涝灾害为例进行城市洪涝灾害影响因子分析及风险评估。首先, 基于 RS 遥感技术利用哨兵二号雷达影像提取和对比灾前、灾后水体形成淹没区范围并通过随机采样获取淹没点数据; 然后, 从致涝、孕涝、承涝和恢复能力 4 个方面选取 16 个指标, 使用随机森林模型计算各指标贡献率并依据排序结果优化指标体系; 最后, 使用 XGBoost 模型对优化后的指标体系赋权并对宜昌市的洪涝灾害风险进行评估。结果显示: (1) 在宜昌市洪涝灾害影响因素中, 地形地貌及河流分布的影响>社会经济因素>气象因素; (2) 高风险地区的范围与长江、清江、沮漳河、黄柏河、渔洋河等主要河系的分布关系十分密切, 对于以上主要河系的水位线监测应保持高度敏感并制定针对性的应急管理措施; (3) 低风险、较低风险和中风险地区占研究区域总面积的 71.8%, 但只包含 8% 的淹没点; 而高风险地区仅占研究区域总面积的 7.32%, 但却包含 81.33% 的淹没点, 表明高风险区域内洪涝灾害事件集中; (4) 使用小尺度历史灾害事件对评估模型的验证结果显示, 72% 的验证点落在高风险及较高风险区, 高达 92% 的验证点落在中高风险区。上述验证结果显示了该评估模型的有效性, 其评估结果与宜昌市实际情况相符, 研究解决了城市洪涝灾害风险难以精细化、量化评估的部分问题并为城市洪涝风险管理、防灾减灾和区域规划提供科学参考。

【关键词】: 洪涝灾害; 风险评估; 机器学习; 宜昌市

【中图分类号】: P426.616 **【文献标识码】**: A **【文章编号】**: 1004-8227(2023)08-1710-14

DOI: 10.11870/cjlyzyyhj202308014

随着全球气候变化加剧, 极端天气事件频发, 城市极端降雨事件不断引发城市内涝[1,2]。习近平总书记多次对防汛抢险救灾工作做出重要指示, 强调“要科学认识致灾规律, 坚持以防为主, 防灾、抗灾、救灾相结合, 全面提升综合防灾能力”。据中国应急管理部公开数据显示, 2021 年全年洪涝灾害共造成 5 901 万人次受灾, 因灾死亡失踪 590 人, 倒塌房屋 15.2 万间, 直接经济损失高达 2 458.9 亿元。在城市洪涝灾害日益严峻的背景下, 风险评估与区划可有效提高城市的防灾减灾和应急管理能力, 保障城市的可持续发展[3]。

国内外学者在洪涝风险评估领域已提出较多的模型和方法。在指标体系构建方面, 城市洪涝灾害的发生和危害程度既受自

¹ 收稿日期: 2022-11-04; 修回日期: 2023-01-03

基金项目: 国家自然科学基金项目(72274186); 湖北省自然科学基金项目(2022CFD128); 国家社会科学基金项目(23AZD072)

作者简介: 王德运(1981~), 男, 副教授, 主要研究方向为灾害风险评估与治理. E-mail: wangdy@cug.edu.cn

然因素的制约,例如,降水、地貌和植被等[4],又受社会因素的影响,例如,土地利用、人口分布和经济水平等[5],各因素在不同区域中对灾害的作用方式和程度有较大不同,指标体系的构建应全面、客观。但已有研究中的指标体系大都以地理特征和气象因素等自然因素为主进行指标选取[4,5,6,7,8,9],如地形起伏度、坡度和降雨等,缺少对社会因素致灾作用的考虑。此外,已有研究在指标赋权时,对不同因素致灾作用程度的考虑不足,缺少对指标体系的优化,限制了评估结果的客观性。从评估方法上来看,可将现有研究大致分为如下3种类型:第一种,以系统动力学[10,11]或水文学模型展开洪涝情景模拟;如叶陈雷等[12]使用InfoWorks ICM建立水文水动力耦合模型对福州市白马河区进行洪涝模拟及风险评价;吴海春等[13]基于PCSWMM模型以不同强度的降雨组合对海甸岛展开洪涝风险评估;第二种,以多元回归[14]、数据包络分析[15]等数理方法对洪涝灾情数据展开研究,如曾忠平等[16]利用爬虫方法获取积水点信息并选取8个洪水影响因素采用Logistic回归模型对江西省吉安市洪涝灾害敏感性进行评估;刘远等[17]利用投入产出模型对河南省2021年极端暴雨灾害进行经济损失的间接评估;第三种,使用层次分析法[18]、熵权法[19]和主成分分析法[20]等构建指标体系进而赋权评估,如赵佳慧等[21]提出了基于改进熵权-TOPSIS-灰色关联分析的城市洪涝风险评估方法,并以郑州市为例进行了实证研究;谢捷等[22]基于地理、社会经济和气象因素等13个指标,通过AHP-熵权组合法对西宁5区2县进行了洪涝风险评估。由于指标体系法可以显示宏观区域或较大尺度区域的洪涝风险情况,且计算方法相对简单,因而,应用非常广泛。在指标体系法中,指标权重的分配是决定风险评估准确性的关键步骤,目前常用的指标赋权方法可分为主观和客观赋权法两种类型[23,24]。以层次分析法和序关系法为代表的主观赋权方法[25]更多从定性角度分析且过于依赖专家知识,限制了评估结果的客观性;以熵权法为代表的客观赋权法易受原始数据分布的影响,评估结果的稳定性难以保证。为了解决上述问题,学者们开始将随机森林[26]、神经网络[27]等机器学习方法引入洪涝风险评估领域,且通过诸多实证研究证实了其有效性。此外,数据的质量对评估结果也有着至关重要的影响,由于统计数据的时效性较差且空间尺度较大,无法与小范围灾域精准匹配,难以做到精细化地评估[28];随着遥感解译技术的不断发展,遥感数据的空间分辨率持续提高,遥感数据在洪涝灾害风险评估领域得到了广泛应用[29]。

本文在构建指标体系时时效性较强的遥感数据代替传统统计数据,对宜昌市洪涝灾害进行了小尺度风险评估。在指标赋权时,先借助随机森林计算各指标贡献率并依据排序结果对指标体系进行优化,然后,基于XGBoost模型对优化后的指标体系进行赋权,减少了指标赋权的主观性。评估结果显示了XGBoost模型在洪涝风险评估领域的有效性。本文以宜昌市为研究区进行城市洪涝影响因子分析及风险评估,可为该地区防灾减灾及应急管理工作提供有益参考。

1 研究区域与数据资料

1.1 研究区域概况

宜昌市位于中国中部、湖北西南部,地处长江上游与中游的分界处,也是长江流域范围内遭遇洪灾影响较为严重的城市。宜昌境内水资源丰富,河流均属长江流域的长江上游干流水系、长江中游干流水系、清江水系、洞庭湖区的四口水系及澧水水系等五大水系,以长江干流为主脉,河流多、密度大、水量丰富。同时,宜昌市地处北亚热带向中亚热带的过渡区域,是亚热带季风性湿润气候,多年平均降水量超过1200mm,境内雨量充沛,降水分布不均,加之地形复杂、高低起伏颇大,导致暴雨洪涝灾害事件频发。

1.2 数据资料及来源

本文所使用到的数据包括如下2类:(1)城市洪涝灾害淹没点数据;(2)城市洪涝灾害影响因素数据,上述数据来源如下。

1.2.1 城市洪涝灾害淹没点数据

在应用机器学习模型进行洪涝灾害风险评估时需要大量的淹没点数据对其进行训练和测试。城市洪涝灾害淹没点数据的提取过程包括如下2步:(1)淹没区水体提取;(2)淹没点采样。

淹没区水体提取过程如下：通过欧洲航天局(<https://scihub.copernicus.eu/dhus/#/home>)获取宜昌市 2020 年 4 月 1 日至 4 月 30 日及 6 月 20 日至 7 月 15 日哨兵二号高分辨率多光谱成像产品，上述两个时间段影像可覆盖整个研究区域且能够形成灾前与灾后的水体对比图。在本研究，将从 2020 年 4 月 1 日至 4 月 30 日的影像中提取的水体记为灾前水体，从 6 月 20 日至 7 月 15 日的影像中提取的水体记为灾后水体，通过灾前与灾后的水体对比获取淹没区与非淹没区，并用二值结果进行区别。由于获取的数据影像级别为 level-1C,因而，所获取数据需要首先通过 Sen2cor 插件进行辐射定标和大气校正等预处理[31],然后，利用 Snap 拼接裁剪后基于 MNDWI 指数设置阈值进行水陆分离并提取水体，其流程及提取结果分别如图 1 和 2 所示。

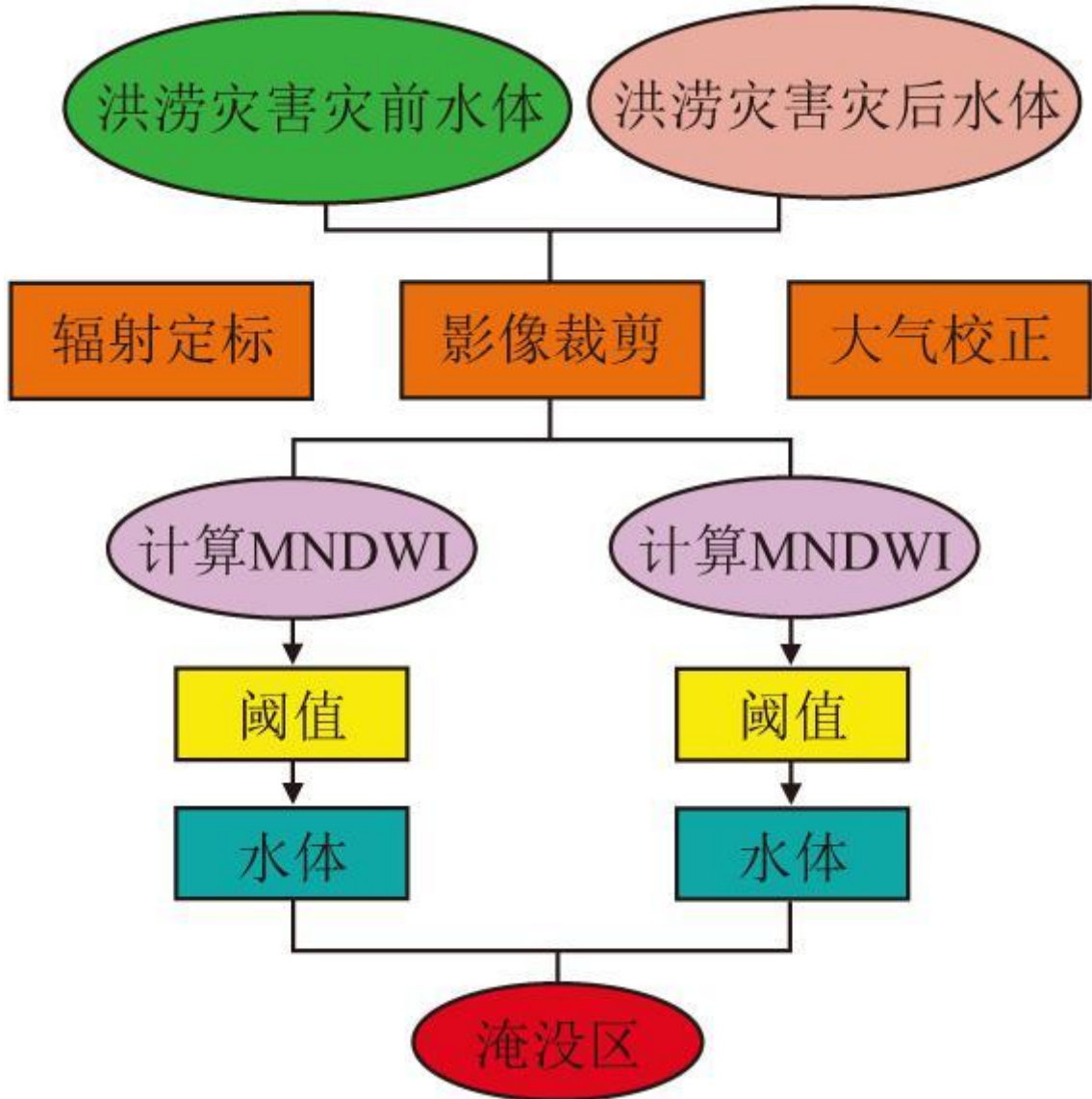


图 1 淹没区提取流程

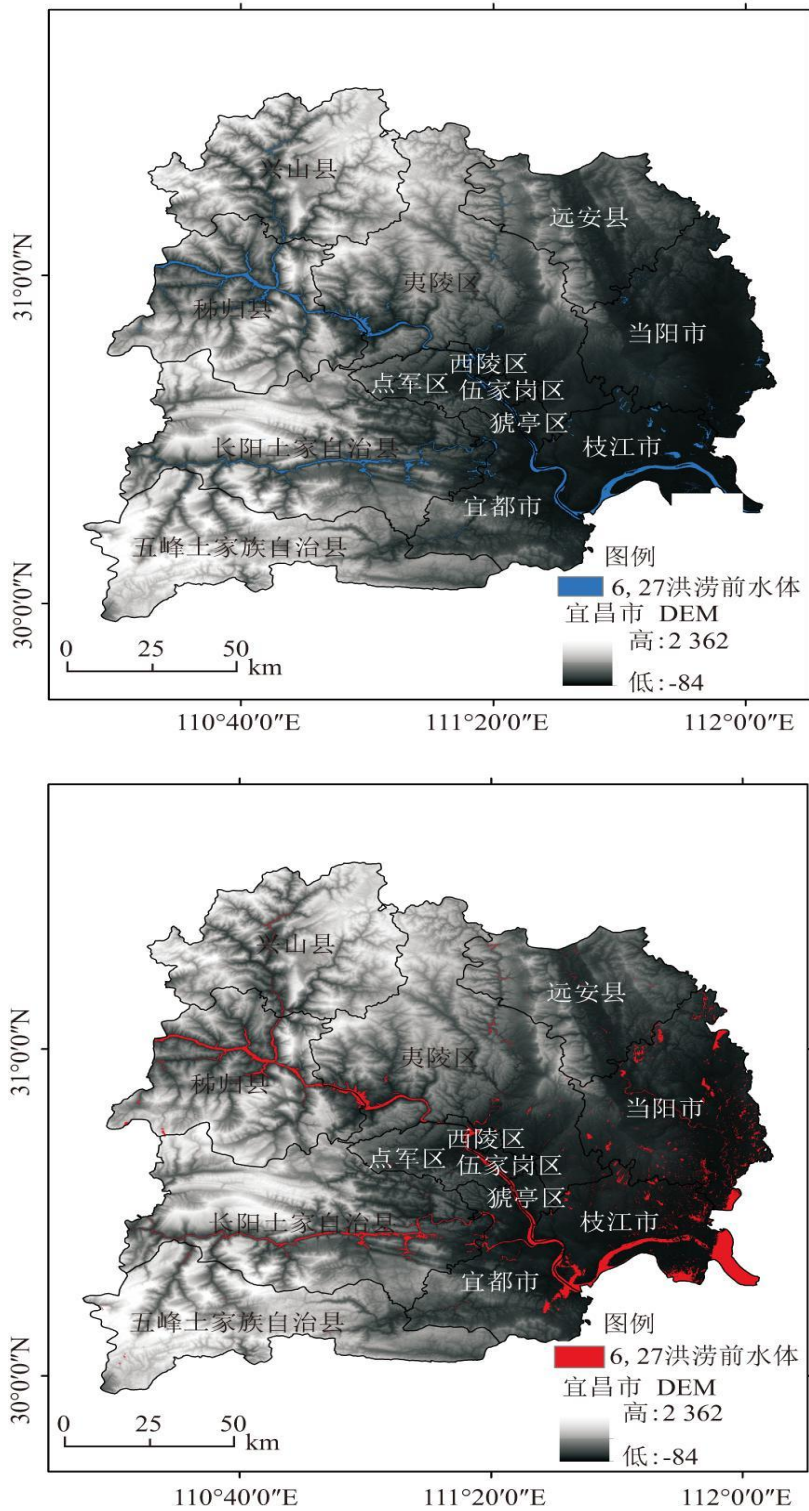


图2 宜昌市洪涝前后水体提取结果

淹没点采样过程如下：将洪涝水体分为淹没区及非淹没区，淹没区指该区域内有洪涝灾害事件的发生，将该区域内的采样点赋值为1,即淹没点；非淹没区指该区域内没有洪涝灾害事件，将该区域内的采样点赋值为0,即非淹没点。在淹没区和非淹没区范围内通过随机方式分别选取150个淹没点及非淹没点，共计300个采样点，其空间分布如图3所示。

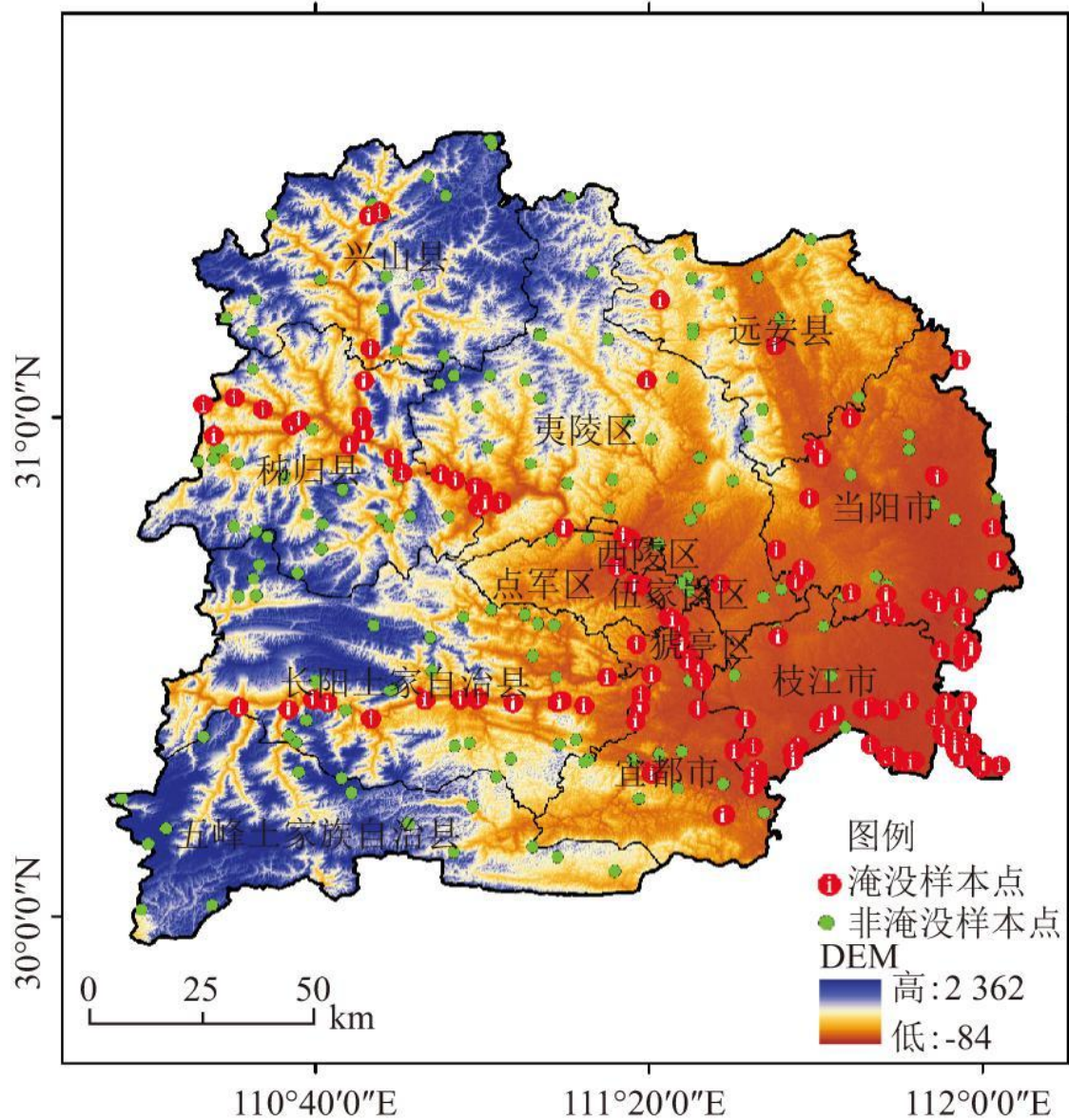


图3 样本点提取

1.2.2 城市洪涝灾害影响因素数据

城市洪涝灾害的发生是多因素共同作用的结果，既有自然因素的作用，又有社会经济因素的影响，其风险水平的大小主要由致灾因子危险性、孕灾环境敏感性、易损性和防灾减灾能力等因素决定[32]。本文结合宜昌市实际情况，从致涝、孕涝、承涝和恢复能力4个方面选取以下16个因素构建评估指标体系，如图4所示。评价指标体系中的16个指标具有不同的正负向属性，其中，高程、坡度、归一化植被指数和距主要河流缓冲区的距离为负向指标，其余均为正向指标。高程和坡度越小，越容易出现积水区，洪涝风险更高；植被覆盖程度低的区域缺乏涵养水的能力，洪涝风险也随之增大；在极端降雨的情况下，河网可能发生漫溢，此时离河系距离越小，洪涝风险越高。

上述指标体系所使用的数据来源如下表1所示：

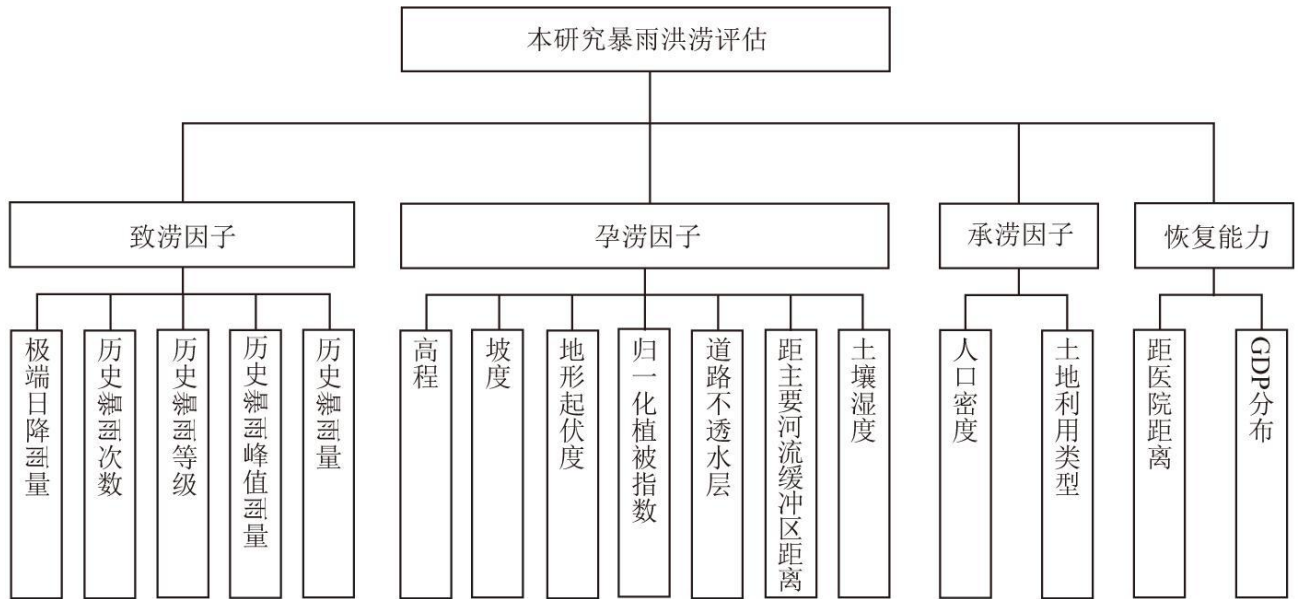


图 4 洪涝风险评估指标体系

表 1 指标数据的来源

因子层	指标层	数据来源
致涝因子	极端日(2020-06-27)降雨量(mm)	国家气象科学数据中心地面日值数据集(V3.0)
	历史暴雨次数(次)	基于 GPM 卫星产品的 2001-2019 中国暴雨数据集
	历史暴雨等级	
	历史暴雨峰值雨量(mm)	

	历史暴雨量(mm)	
孕 涝 因子	高程(m)	中国科学院计算机网络信息中心地理空间数据云平台
	坡度(°)	利用 GIS 的梯度分析工具对 DEM 高程数据进行处理得到
	地形起伏度(m)	利用 GIS 邻域分析工具对 DEM 高程数据进行处理得到
	归一化植被指数	地理空间数据云的 Landsat 8 OLI _ TIRS 卫星数字产品
	道路不透水层	道路分布来源国家基础地理信息中心
	距主要河流缓冲区距离(m)	河系分布来源于国家基础地理信息中心
	土壤湿度(kg/m ²)	全球高分辨率的陆面模拟系统(GLDAS)数字产品
承 涝 因子	人口密度(人/ km ²)	WorldPop (https://www.worldpop.org/)
	土地利用类型	武汉大学 Landsat 的年度中国土地覆盖数据集[33]

恢 复 能 力	距医院距离(m)	高德地图获取宜昌市二乙及以上等级医院分布
	GDP 分布(万元)	在宜昌市统计年鉴 GDP 数据的基础上, 将与经济活动密切相关的土地利用类型、夜间灯光亮度、居民点密度等多因素数据结合, 利用多因子权重分配法将 GDP 数据分布到栅格单元

2 研究方法

2.1 归一化差异水体指数

淹没区水体范围的提取是获取淹没点数据的前提。基于卫星遥感数据的水体指数法在水体提取领域有着广泛应用[34],其中,改进归一化差异水体指数法[35](Modified Normalized Difference Water Index, MNDWI)是在归一化差异水体指数法(Normalized Difference Water Index, NDWI)的基础上进行改进得到的,其在城镇范围内的提取精度更高,能够更加有效地区分水体边界。MNDWI 的计算公式如下:

$$MNDWI = \frac{Green - MIR}{Green + MIR} \quad (1)$$

式中: Green 是绿光波段, MIR 是中红外波段。

2.2 随机森林算法

随机森林是利用集成学习的思想将多棵决策树进行集成的一种算法,它属于机器学习中的集成学习方法。如何进行最优的节点分裂是构造决策树的关键步骤。一个最优的节点分裂应该使节点的不纯度水平减少得最多。Gini 指标是一种描述节点不纯度水平的函数。进行特征值筛选时,其重要性权重便是通过 Gini 指数得出。通过计算每个特征在每棵决策树节点分割时 Gini 指数的改变量的平均值,比较不同特征的平均 Gini 指数改变量占所有特征平均 Gini 指数改变量总和的百分比来判断特征重要程度,重要性评分的计算如下:

$$Gini(m) = 1 - \sum_{j=1}^k [p(j|m)]^2 \quad (2)$$

$$VIM_k = Gini(m) - Gini(r) - Gini(l) \quad (3)$$

式中: $p(j|m)$ 表示节点 m 上类别为 j 的概率; VIM_k 表示特征 X_k 分裂时 Gini 变化量; r 和 l 是分裂后的新节点。若随机森林共有 n 棵树,那么特征 X_k 在所有决策树中的 Gini 变化量总和可通过下式计算,其中, VIM 表示特征重要性评分:

$$VIM_k^{Gini} = \sum_{i=1}^n VIM_{ik} \quad (4)$$

最后，得到的特征重要性如下：

$$VIM_k = \frac{VIM_k}{\sum_{i=1}^c VIM_i} \quad (5)$$

2.3 XGBoost 算法

极端梯度提升模型(XGBoost)是以 CART 回归树为基础，对数据集进行分类和预测的一种 boosting 算法。其基本原理是基于迭代提升的思想，即后一个模型在前一个模型的基础上进行预测，不断地添加树，不断地进行特征分裂来生长一棵树，每添加一个树，相当于学习一个新函数，并用该函数去拟合上次预测的残差。预测值可用下式计算：

$$\hat{y} = \sum_{i=1}^n f_k(x_i), f_k \in F \quad (6)$$

式中： \hat{y} 表示第 i 个样本预测值； $f_k(x_i)$ 表示第 i 个变量的第 k 个回归方程； F 表示回归树集合。XGBoost 作为优化后的机器学习模型，在评估过程中能够结合多个学习器提升预测的准确率及效率；并且该模型能通过对正负样本的学习减小因原始多源数据异质性而导致的偏误问题，同时，引入正则项能够有效提升模型的稳健性，防止过拟合的情况。

2.4 正负指标的归一化处理

由于数据来源和量纲的差异，正负指标的性质对于风险水平的作用不同，为消除上述差异，需要首先对不同正负属性的指标进行归一化处理[5]。

$$T_{ij} = \frac{X_{ij} - X_{\min}}{X_{\max} - X_{\min}} \quad (7)$$

正向指标归一化：

$$T_{ij} = \frac{X_{\max} - X_{ij}}{X_{\max} - X_{\min}} \quad (8)$$

负向指标归一化：

式中： X_{ij} 为指标原始值； X_{max} 为指标原始值的最大值； X_{min} 为指标原始值的最小值。

3 洪涝风险评估指标体系构建

3.1 指标贡献率分析

本文分别从淹没区和非淹没区内随机选取了 150 个样本点作为研究数据，共计 300 个样本点。将其中 70% 的样本点划分为训练集，30% 的样本点划分为测试集，并用随机森林算法进行特征值筛选。由于数据来源、格式和量纲存在较大差异，因此，需首先依据公式(7)和(8)进行归一化处理并将归一化后的数据划入栅格网中，然后，将上述数据统一到投影坐标系 WGS84-UTM 下，并统一各栅格图层单元大小为 $30\text{ m} \times 30\text{ m}$ 。由随机森林模型计算得到各指标的贡献率排序如图 5 所示。

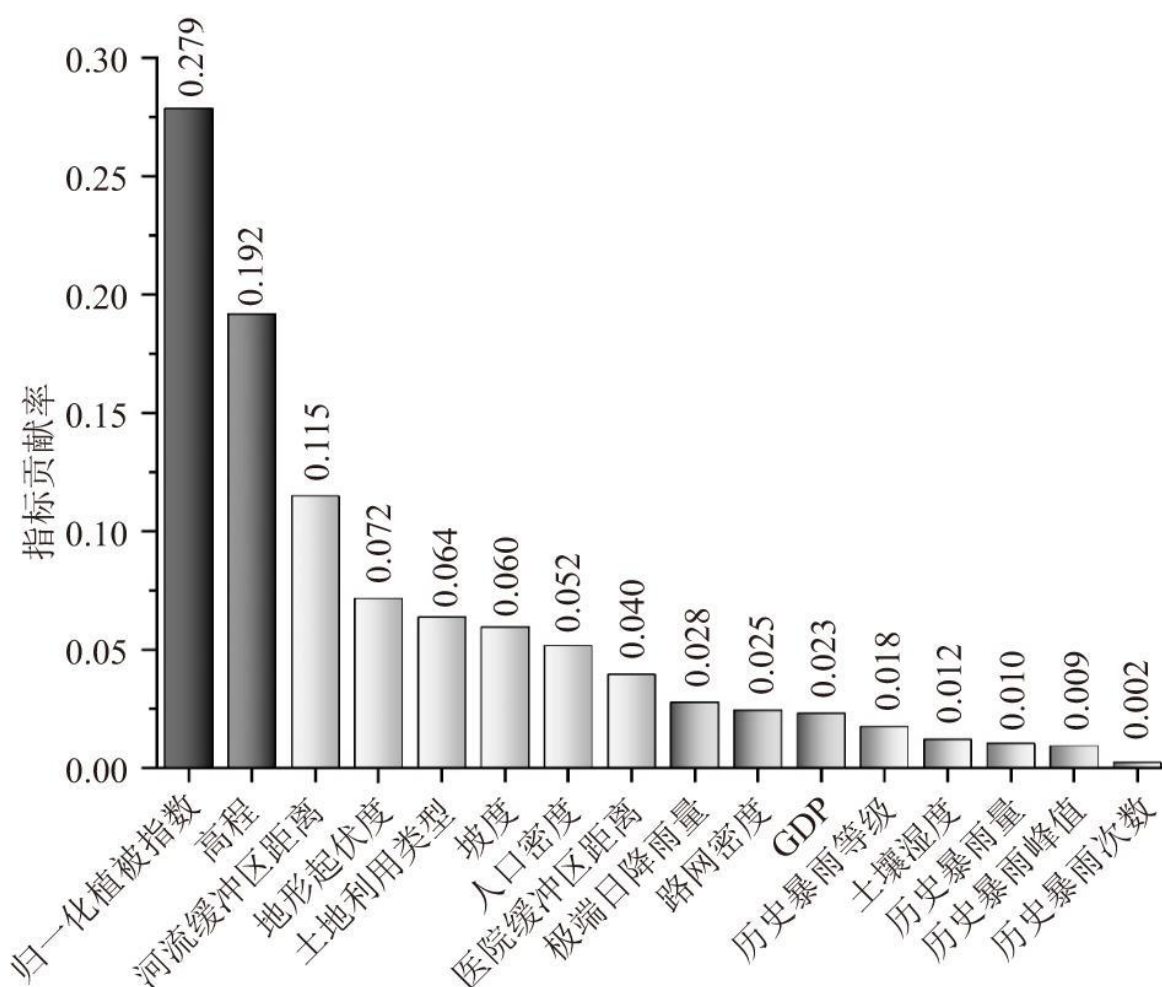


图 5 指标贡献率排序

由图 5 可以看出：

(1)16 项指标的贡献率总和为 1,其中，3 个指标贡献率大于 0.1,分别是植被指数、高程和距主要河流缓冲区的距离。植被指数

贡献率排在第一, 高达 0.279, 说明洪涝灾害与植被覆盖程度关联密切, 洪涝灾害发生区域与植被分布具有较强耦合性; 高程数据在贡献率中位列第二, 体现出地理特征对洪涝风险评估的重要性; 而距主要河流缓冲区的距离贡献率位居第三, 说明宜昌市的洪涝灾害与河系分布高度相关, 河系分布的周边地区是洪涝灾害的高发区。

(2) 贡献率在 0.02~0.1 之间的指标有 8 项, 涉及地理特征、社会经济及气象因素 3 个方面。地形起伏度、土地利用类型及坡度表明地形地质对于洪涝灾害具有较强的助推作用; 人口密度和距医院距离反映出社会经济层面的活动与洪涝灾害的风险也有较强的关联; 而极端日降雨则能够反映气象条件的致灾作用。

(3) 贡献率低于 0.02 的指标有 5 项, 包括 4 个历史暴雨指标及土壤湿度, 表明在宜昌市, 历史暴雨指标和土壤湿度与洪涝灾害相关程度不高。

3.2 最优指标组合筛选

为衡量模型在不同指标组合下的评估效果, 本文采用 ACC(Accuracy)和 AUC(Area Under Curve)共同评估模型结果的合理性。ACC 表示模型结果的准确率, AUC 是 ROC(Receiver Operating Characteristic curve)曲线下的面积, 通常取值在 0.5 和 1 之间, 越接近 1 表明模型真实性越高, 经常用于评判二值分类器的优劣; 以 ACC 和 AUC 为评价标准进行指标优化兼顾了模型的稳健性和准确性, 避免了数据偏斜造成的误差。在指标体系优化过程中, 首先基于随机森林计算各指标的贡献率并进行排序, 然后, 逐一从模型中剔除贡献率较低的指标形成新的指标体系, 以此类推, 直到模型效果达到最优, 即 ACC 和 AUC 的值最大, 此时的指标体系即为优化后的指标体系, 具体优化过程如表 2 所示。从实验 1 到 16, 将贡献率较低的指标逐一剔除, 模型的 ACC 和 AUC 值均呈现出先升后降的趋势。从实验 1 到 5, 模型的 ACC 和 AUC 值不断升高, 其原因为将贡献率较低的指标剔除后, 数据噪声减少, 模型性能也相应提高。在整个实验过程中, 实验 5 中模型的 ACC 和 AUC 值同时达到最大, 表明该指标体系下的模型性能最优。而从实验 6 到 16, 由于贡献率较大的指标被剔除, 数据重要信息也随之减少, 导致模型性能下降。

表 2 基于 ACC 和 AUC 值的最优指标组合筛选

	剔除指标	ACC	AUC
1	无	0.955 555 6	0.954 105 0
2	历史暴雨次数	0.922 222 2	0.921 723 6
3	历史暴雨次数、历史暴雨峰值	0.944 444 4	0.944 671 1
4	历史暴雨次数、历史暴雨峰值、历史暴雨量	0.955 555 6	0.958 184 6

5	历史暴雨次数、历史暴雨峰值、历史暴雨量、土壤湿度	0.955 555 6	0.958 184 6
6	历史暴雨次数、历史暴雨峰值、历史暴雨量、土壤湿度、历史暴雨等级	0.944 444 4	0.944 671 1
		
13	历史暴雨次数、历史暴雨峰值、历史暴雨量、土壤湿度、历史暴雨等级、GDP、道路密度、极端日降雨量、距医院距离、人口密度、坡度、土地利用类型	0.944 444 4	0.940 591 5
14	历史暴雨次数、历史暴雨峰值、历史暴雨量、土壤湿度、历史暴雨等级、GDP、道路密度、极端日降雨量、距医院距离、人口密度、坡度、土地利用类型、地形起伏度	0.933 333 3	0.927 078 0
15	历史暴雨次数、历史暴雨峰值、历史暴雨量、土壤湿度、历史暴雨等级、GDP、道路密度、极端日降雨量、距医院距离、人口密度、坡度、土地利用类型、地形起伏度、距离河网距离	0.933 333 3	0.931 157 6
16	历史暴雨次数、历史暴雨峰值、历史暴雨量、土壤湿度、历史暴雨等级、GDP、道路密度、极端日降雨量、距医院距离、人口密度、坡度、土地利用类型、地形起伏度、距离河网距离、高程	0.855 555 6	0.869 199 4

3.3 基于 XGBoost 算法的指标赋权

利用随机森林对上述 16 项指标进行优化后，剔除了历史暴雨次数、历史暴雨峰值、历史暴雨量和土壤湿度四项指标，保留了其余 12 个指标用于评估宜昌市洪涝灾害风险。为进一步提高指标赋权的客观性，本文使用 XGBoost 算法对指标进行赋权，得出各指标权重如表 3 所示。

4 宜昌市洪涝灾害风险评估结果分析

4.1 单项风险评估结果

洪涝灾害致灾因子危险性、孕灾因子敏感性、承涝因子脆弱性以及恢复能力的单项评估结果计算公式如下：

表 3 宜昌市洪涝风险评估指标权重测算结果

	因子类型	子指标	权重
城市洪涝灾害风险评估指标	致涝因子	极端日降雨量	0.034
		历史暴雨等级	0.010
	孕涝因子	高程	0.265
		坡度	0.015
		地形起伏度	0.034
		归一化植被指数	0.279
		道路密度	0.113
		距主要河系缓冲区距离	0.113
	承涝因子	人口密度	0.005
土地利用类型		0.010	

	恢复能力因子	距医院缓冲区距离	0.162
		GDP	0.020

$$L(X) = \sum_{i=1}^n [W_i \times L_{in}(X)] \quad (9)$$

致灾因子危险性:

$$B(X) = \sum_{i=1}^n [W_i \times B_{in}(X)] \quad (10)$$

孕灾因子敏感性:

$$S(X) = \sum_{i=1}^n [W_i \times S_{in}(X)] \quad (11)$$

承涝因子脆弱性:

$$P(X) = \sum_{i=1}^n [W_i \times P_{in}(X)] \quad (12)$$

恢复能力:

式中: $L(X)$ 、 $B(X)$ 、 $S(X)$ 和 $P(X)$ 为各单项评价指数(危险性、敏感性、脆弱性和恢复能力); W_i 是各指标的权重; $L_{in}(X)$ 、 $B_{in}(X)$ 、 $S_{in}(X)$ 和 $P_{in}(X)$ 为各指标归一化后的指标值。在得出各单项评价指数后, 利用幂指数加权评价法将 4 个单项评价指数集成叠加, 可以得到研究区域内的洪涝灾害综合风险, $R(X)$ 为区域洪涝综合风险值, f 为幂指数模型[8]。

$$R(X) = f(L(X), B(X), S(X), P(X)) \quad (13)$$

借助 GIS 分析可以得到宜昌市洪涝灾害危险性、敏感性、脆弱性和恢复能力的单项评价价值, 值域范围分别为[0.002,0.407]、[0.132,0.758]、[0,0.015]和[0.101,0.03]。采用自然断点法将各单项评价结果划分为 5 个不同的等级, 具体等级划分如表 4 所示。

表 4 单项风险等级划分

评价等级	低	较低	中	较高	高
洪涝危险性	[0.002,0.0110)	[0.011,0.016)	[0.016,0.022)	[0.022,0.297)	[0.297,0.407]
洪涝敏感性	[0.132,0.309)	[0.309,0.391)	[0.391,0.462)	[0.462,0.556)	[0.556,0.758]
洪涝脆弱性	[0,0.004)	[0.004,0.006)	[0.006,0.008)	[0.008,0.1)	[0.1,0.015]
恢复能力	[0.101,0.162]	[0.069,0.101)	[0.047,0.069)	[0.030,0.047)	[0,0.030)

4.1.1 洪涝灾害危险性评估

危险性是衡量洪涝灾害风险的关键指标之一，危险性与致灾因子有密切关系。本文在对 16 个指标进行优化后保留了极端日降雨量及 2001~2019 年历史暴雨等级两个致灾因子指标，并结合指标权重形成了宜昌市洪涝灾害危险性等级区划图，如图 6 所示。

由图 6 可以看出，宜昌市极端降雨在纵向上呈现由东北部向南梯度递减的趋势，整体趋势与历史暴雨等级并无明显的关联，历史暴雨等级分布较为分散。宜昌市洪涝灾害危险程度较高的区域位于远安县、当阳市西北部、夷陵区东部以及西陵区、伍家岗区 and 点军区东北区域，其中，由于远安县中部及南部地区以及夷陵区、西陵区和伍家岗区的三区交界地带极端日降雨量和历史暴雨等级均较高，因此叠加后危险程度最高，是洪涝灾害重危区域，在应急管理时需重点关注。

4.1.2 洪涝灾害敏感性评估

洪涝灾害敏感性体现在当洪涝来临时，研究区域内的地形地质、植被覆盖和河流分布等地理特征对于洪涝灾害发生的敏感程度。本文选取孕涝指标进行敏感性分析，形成的宜昌市洪涝灾害敏感性等级区划如图 7 所示。

图 7 中的(a)(b)(d)和(f)为高程、坡度、植被指数及距主要河流缓冲区距离进行负向归一化处理后的数据，其余为正向归一化处理后的数据。从图 7 可以看出，宜昌市整体地理环境较为复杂，高程及坡度均呈现从西向东递减，河流分布周边地区植被覆盖程度较低。从图(g)中的敏感性区划结果可以看出，宜昌市洪涝灾害高敏感地区主要分布在高程及坡度较低且植被覆盖程度不高的区域，此外，道路密度较大的地区由于不透水层比例高导致洪涝灾害敏感程度也相对较高，在极端降雨的情况下河流分布

周围地区的洪涝灾害敏感性相比其他区域更高。

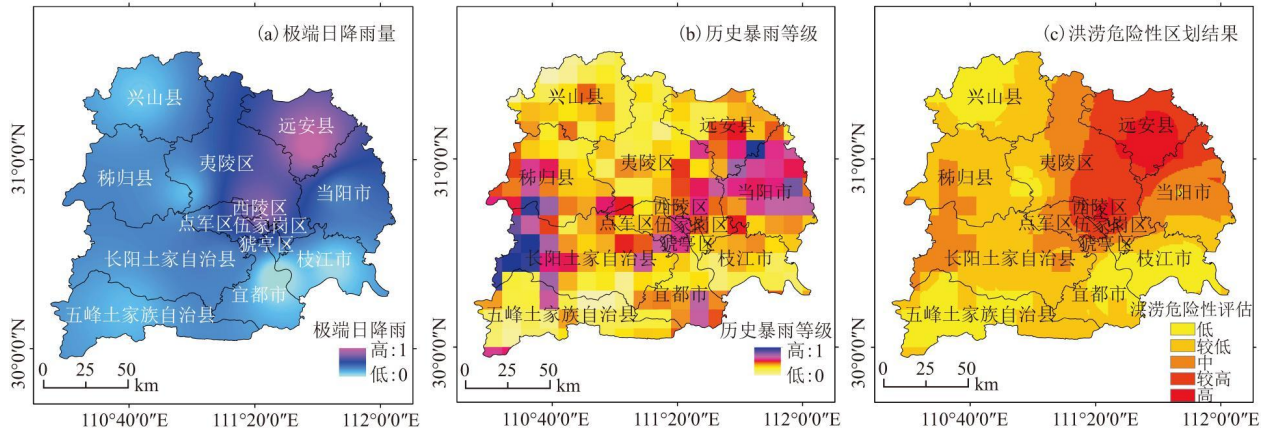
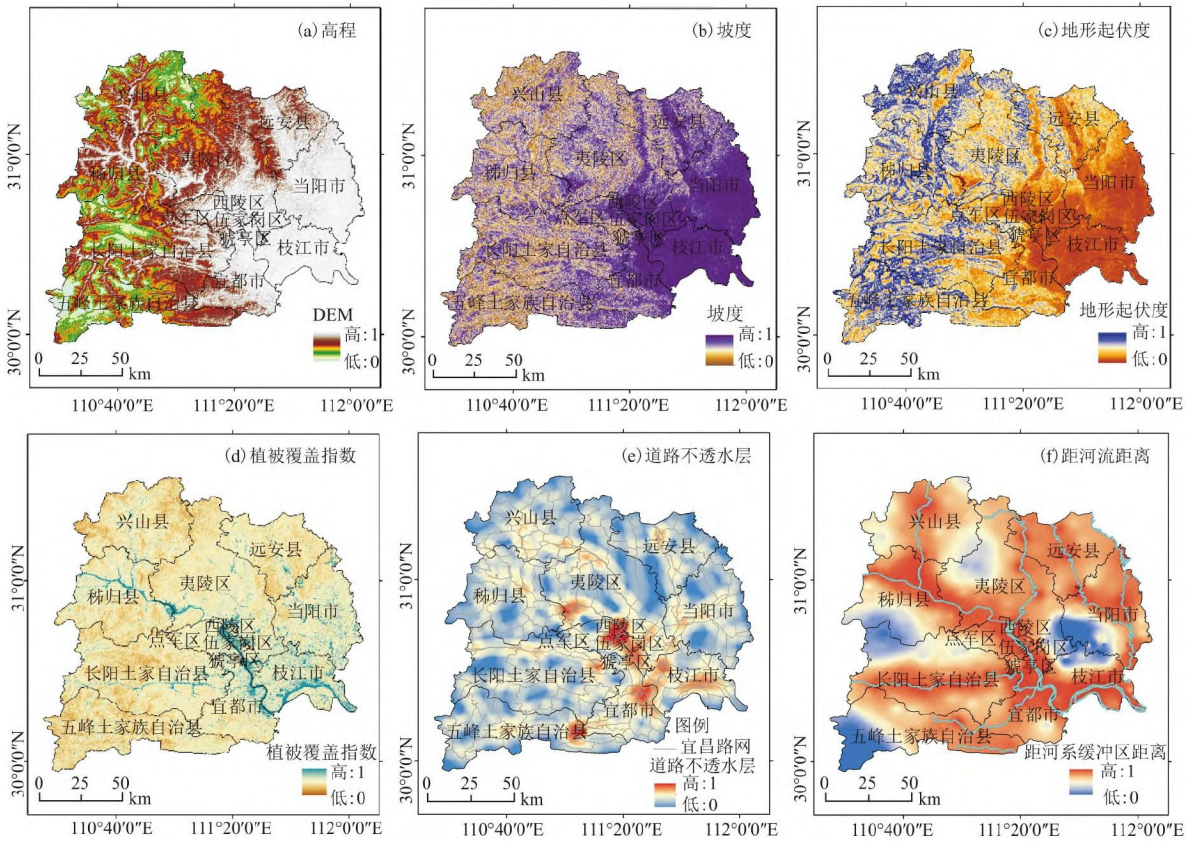


图6 洪涝灾害危险性评估结果



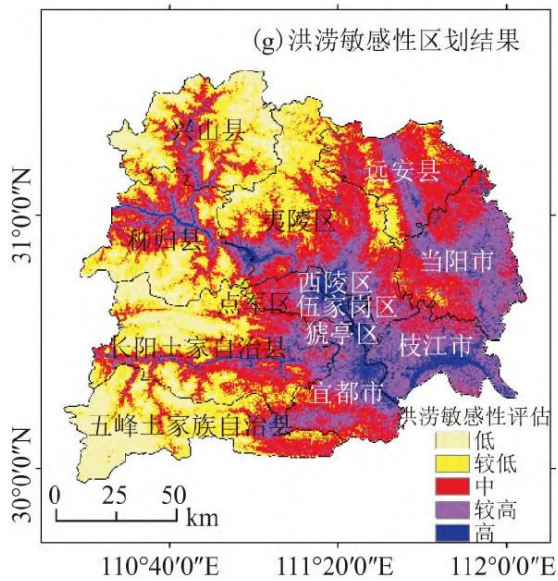


图 7 洪涝灾害敏感性评估结果

4.1.3 洪涝灾害脆弱性评估

洪涝灾害脆弱性体现在当洪涝来临时，承涝体遭受损失的可能性大小。本文选取承涝指标进行脆弱性分析，包括人口密度和土地类型，形成的宜昌市洪涝灾害脆弱性等级区划如图 8 所示。从图 8 中可以看出，在人口密度大的地区，其经济发展水平也较高，承涝体表现出的脆弱性程度也更高；宜昌市范围内的土地类型包括农田、灌木、水域和不透水地表 4 种，分别用归一化后的数值从低到高表示，如图 8(b)所示。当农田和灌木地区发生洪涝时，由于土壤具有较好的涵养水的能力，洪涝风险相对于其他两种用地类型也较低；在水域和不透水地表区域，当极端降雨发生时，水位线会在短时间内高涨，从而导致较高洪涝风险；在不透水层区域，由于不透水层比例大，地表下渗能力差，城市洪涝风险脆弱性也相应更高。

4.1.4 恢复能力评估

恢复能力主要来源于研究区域的医疗水平及经济实力的综合作用，恢复能力越强，其应急和抵御灾害的能力也越强，区域发生洪涝风险的可能性就越低。本文以社会经济指标衡量恢复能力，形成的宜昌市恢复能力等级区划如图 9 所示。由图 9 可知，宜昌市中心西陵区及周边夷陵区、点军区、伍家岗区及猗亭区医疗和经济发展水平较高，故恢复能力最强，而长阳土家族自治县和五峰土家族自治县由于医疗和经济水平落后，其恢复能力较差，在洪涝发生时需着重关注。

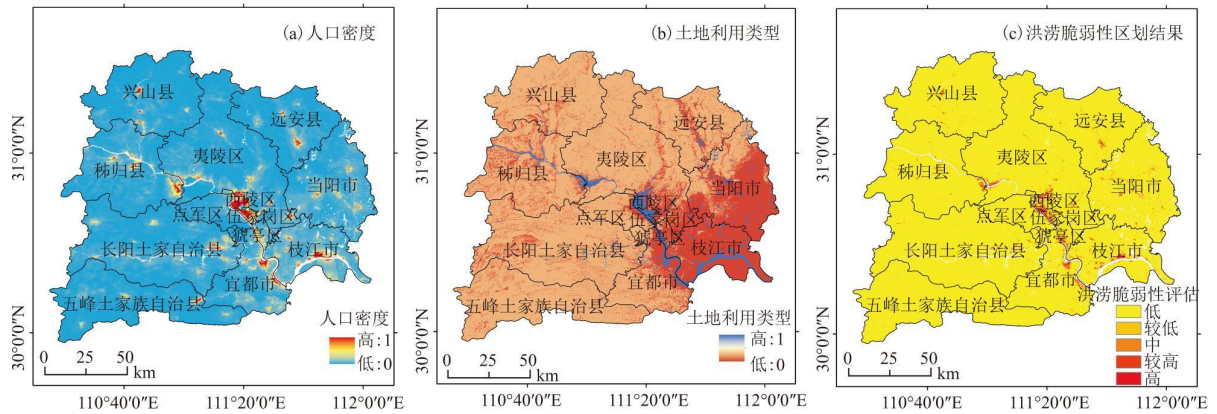


图8 洪涝灾害脆弱性评估结果

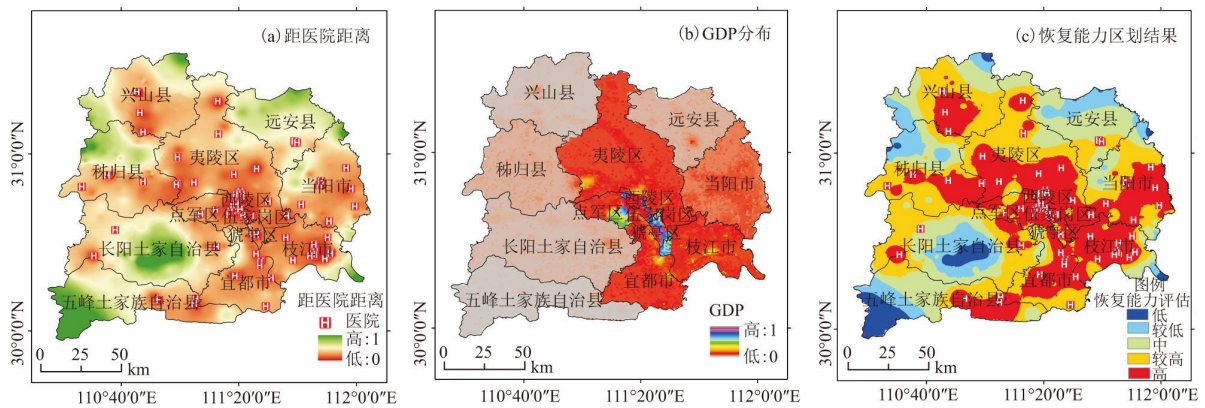


图9 恢复能力评估结果

4.2 洪涝灾害综合风险评估

洪涝灾害综合风险评估是在洪涝灾害危险性、敏感性、脆弱性和恢复能力评估结果的基础上，依据 XGBoost 得到的指标权重利用公式(13)得到的，宜昌市洪涝灾害综合风险评估结果如图 10 所示。

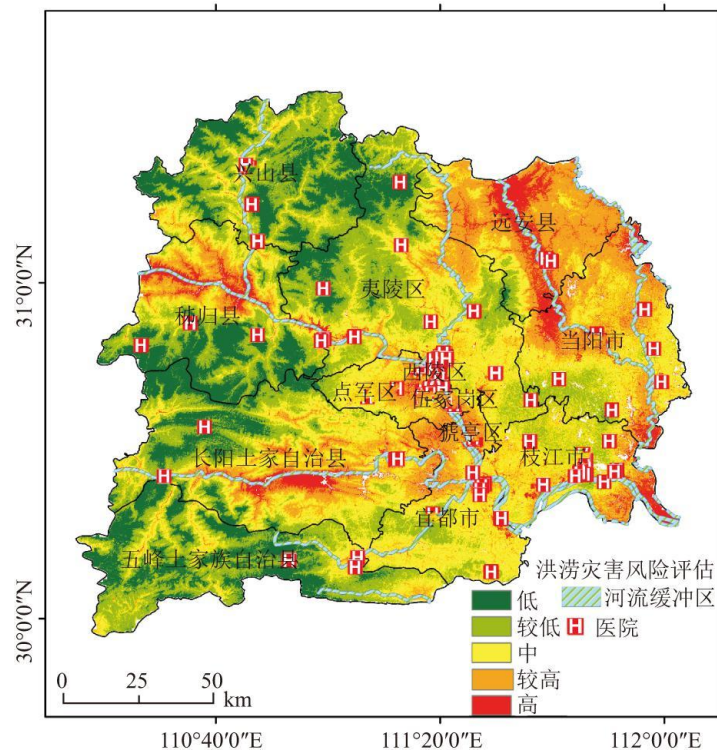


图 10 宜昌市洪涝灾害综合风险评估结果

由图 10 可以看出，宜昌市洪涝灾害综合风险呈现西部地区高于东部地区的态势，且与河流分布密切相关。宜昌市西部多山地，东部多平原，枝江市是长江沿岸城市，土地类型以耕地为主，但此地植被覆盖程度并不高，因此形成了洪涝高风险区域；西陵区及其周边区域位于宜昌市中心地带，人口密度大、路网密集且不透水层占比较大，当洪涝发生时损失也会更大，属于高风险地区；远安县为极端日降雨量最大的区域且历史降雨等级较高、距河网距离较近，洪涝风险较高；秭归县和长阳土家族自治县均为经济和医疗水平较低的区域，且这部分区域多为山地，恢复能力较差，属于高风险地区。从整体上看，宜昌市长江、清江、沮漳河、黄柏河和渔洋河等沿岸地区植被覆盖较少、距离河网较近，河系分布周围均是洪涝高风险区。

表 5 在图(10)的基础上进一步展示了各风险等级区域内淹没点的分布，从表 5 可以看出，低风险、较低风险和低风险地区占研究区域总面积的 71.8%,但只包含 8%的淹没点；高风险地区占研究区域总面积的 7.32%,却包含了 81.33%的淹没点，上述分析表明高风险区域内洪涝事件密集，符合宜昌市洪涝灾害的实际情况。

表 5 各风险等级区域内淹没点的分布

风险等级	栅格数目	分区面积 (km ²)	洪灾 数量	分区比例 (%)	灾害比例 (%)
低风险	3 301 295	2 971.165 5	0	14.5	0.00%
较低风险	6 453 018	5 807.716 2	3	28.34	2.00%
中风险	6 594 997	5 935.497 3	9	28.96	6.00%
较高风险	4 755 231	4 279.707 9	16	20.88	10.67%
高风险	1 666 339	1 499.705 1	122	7.32	81.33%

4.3 评估结果验证

4.3.1 基于人工检视法的实际淹没区与风险区划结果对比

本文使用人工视检法及历史数据分析法对评估结果的合理性和准确性进行验证。将 2020 年 6 月宜昌市实际淹没区与本文的风险区域结果进行对比，如图 11 所示，实际淹没区与本文给出的高风险区域重合程度较高，即大部分实际淹没区均落在了评估结果的较高风险及高风险区域，说明了本文评估结果的合理性。

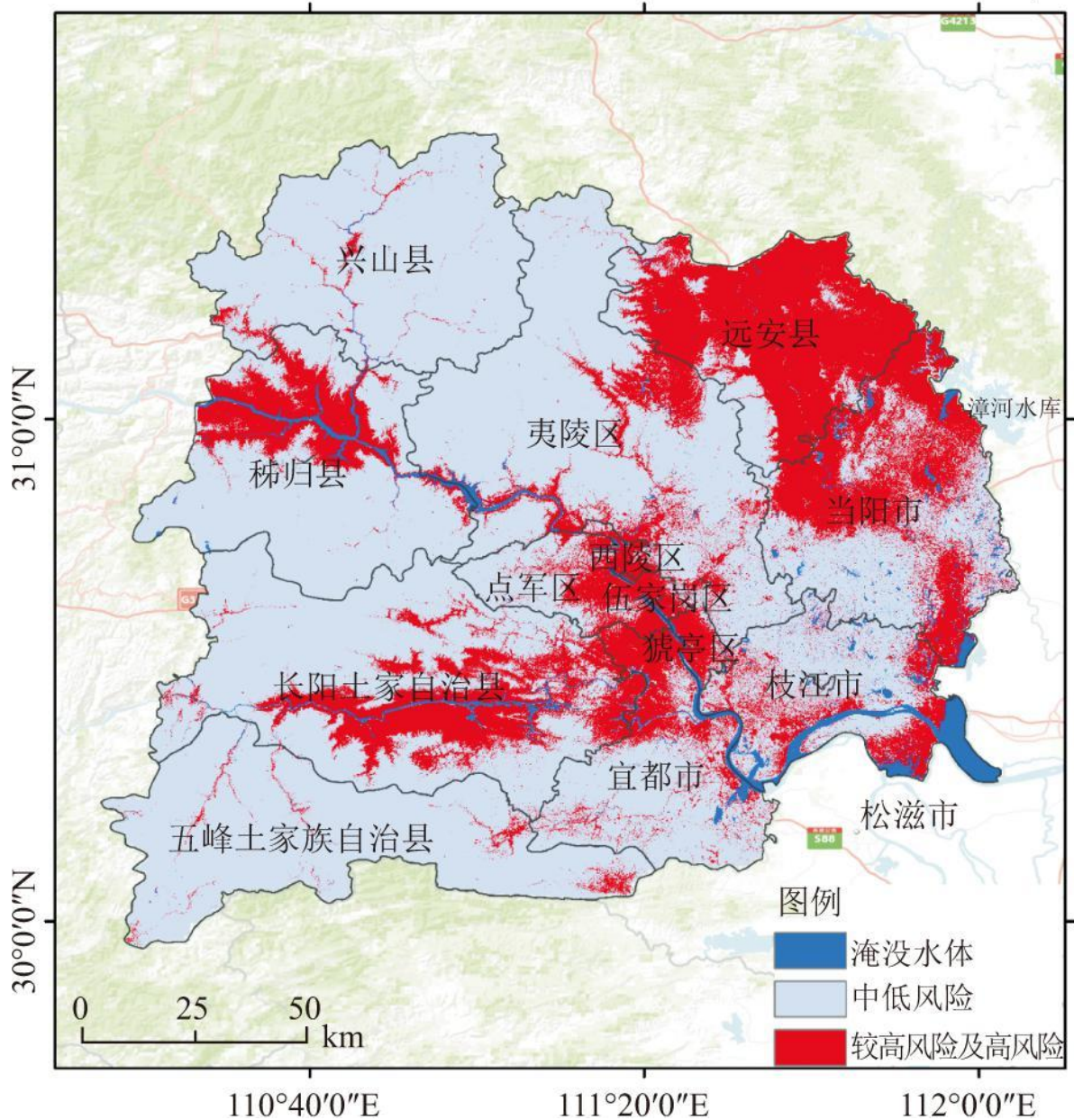


图 11 实际淹没区与风险区划对果对比图

4.3.2 基于历史灾害事件的模型合理性验证

为了进一步验证模型的合理性和泛化能力，本文结合近年来宜昌市发生的历史洪涝灾害事件通过宜昌民政局、宜昌发布以及各类新闻报道中获取了精确至乡镇、街道的 75 个历史洪涝灾害点，对上述数据进行可视化处理并对照到模型评估结果中，得到历史洪涝灾害点与不同风险等级区域的对应关系，如图 12 所示。其中，位于较高风险和高风险区域的灾害点占 72%；中高风险区内的灾害点占比高达 92%，上述结果表明，本文的评估结果合理、准确，符合宜昌市洪涝灾害实际情况。

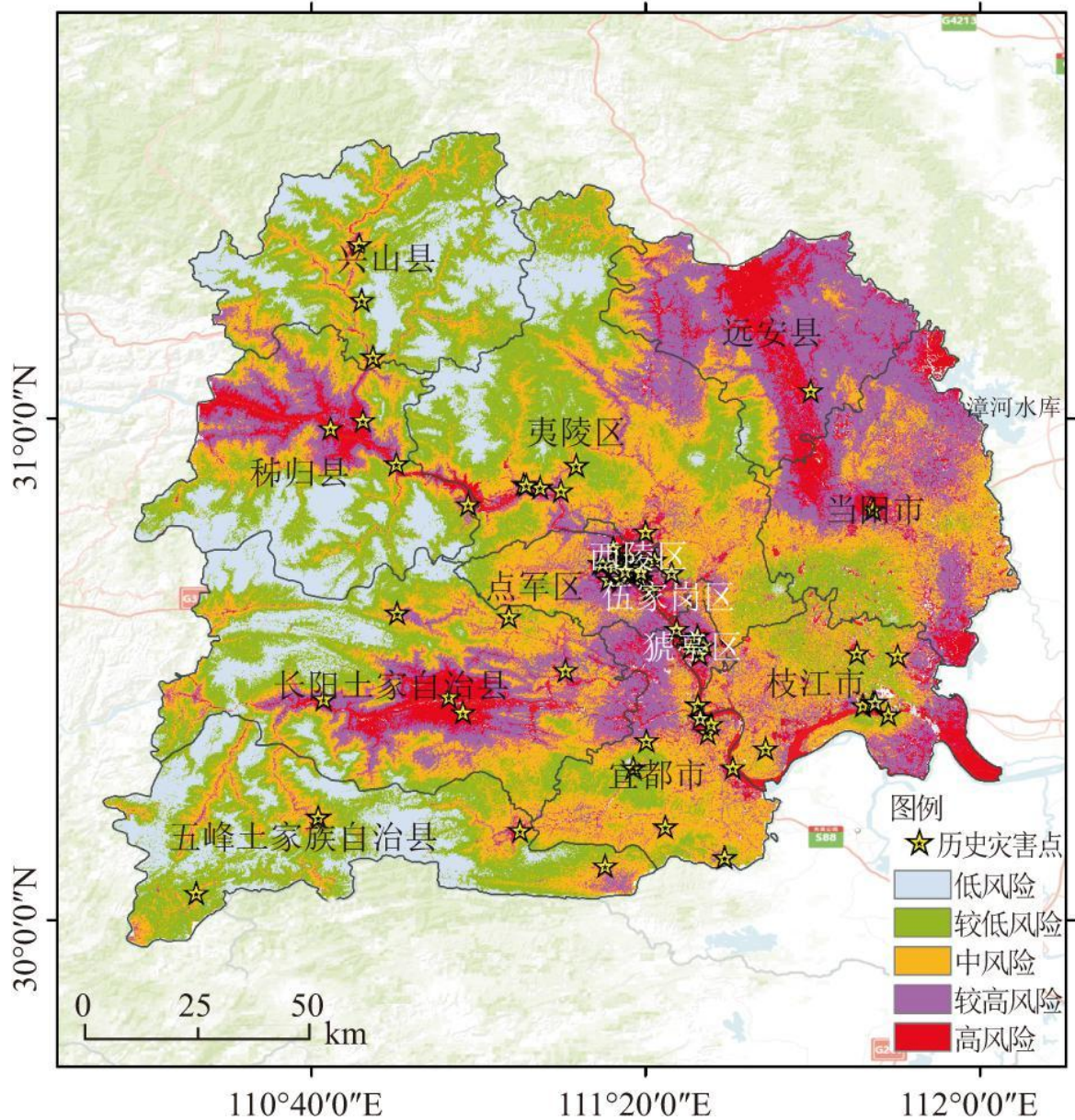


图 12 历史灾害点与不同风险等级区域的对应关系图

5 结论与讨论

5.1 结论

本文以宜昌市为研究区域，基于 RS 技术与 GIS 手段，结合随机森林和 XGBoost 模型对其洪涝灾害风险进行了评估和区划，得出如下主要结论：

(1) 本文利用 RS 技术提取淹没区范围并选取采样点，从四个层面选取了极端日降雨、历史暴雨等级、高程和植被覆盖等 16 个指标，使用随机森林模型根据指标贡献率对指标体系进行了优化。优化后的指标体系包括了 12 个指标，其中，植被指数、高程、主要河流缓冲区距离、地形起伏度、土地利用类型、坡度及人口密度对宜昌市洪涝风险水平影响较大。

(2)使用 XGBoost 模型对指标体系赋权后,对宜昌市洪涝灾害危险性、敏感性、脆弱性和恢复能力 4 个方面进行了单项评估,然后,利用幂指数模型形成综合风险评估结果。结果显示,低风险、较低风险和中风险地区占研究区域总面积的 71.8%,但只包含 8%的淹没点,高风险地区占研究区域总面积的 7.32%,却包含了 81.33%的淹没点,表明高风险区域内洪涝事件密集。

(3)宜昌市洪涝灾害影响因素作用程度的分析结果显示,地形地貌及河流分布的影响>社会经济因素>气象因素,高风险地区的范围与长江、清江、沮漳河、黄柏河和渔洋河等主要河系的分布关系密切,因此,对于以上主要河系的水位线监测应保持高度敏感并制定针对性的应急管理措施。

(4)本文通过人工视检法及历史数据分析法对评估结果进行了验证。小尺度历史灾害数据验证结果显示,72%的验证点落在高风险及较高风险区,92%的验证点落在中高风险区。上述结果表明了本文评估结果的合理性。

5.2 讨论与展望

本文基于小尺度数据进行研究,但传统社会经济统计数据的空间尺度通常为城市及以上尺度,无法与本研究区域较好地匹配,因此,部分指标难以纳入评估体系。但是其中部分数据可以用来进一步验证评估结果的合理性。例如,将宜昌市 2020 年旱涝保收面积与本文的评估结果进行对照,如图 13 所示。

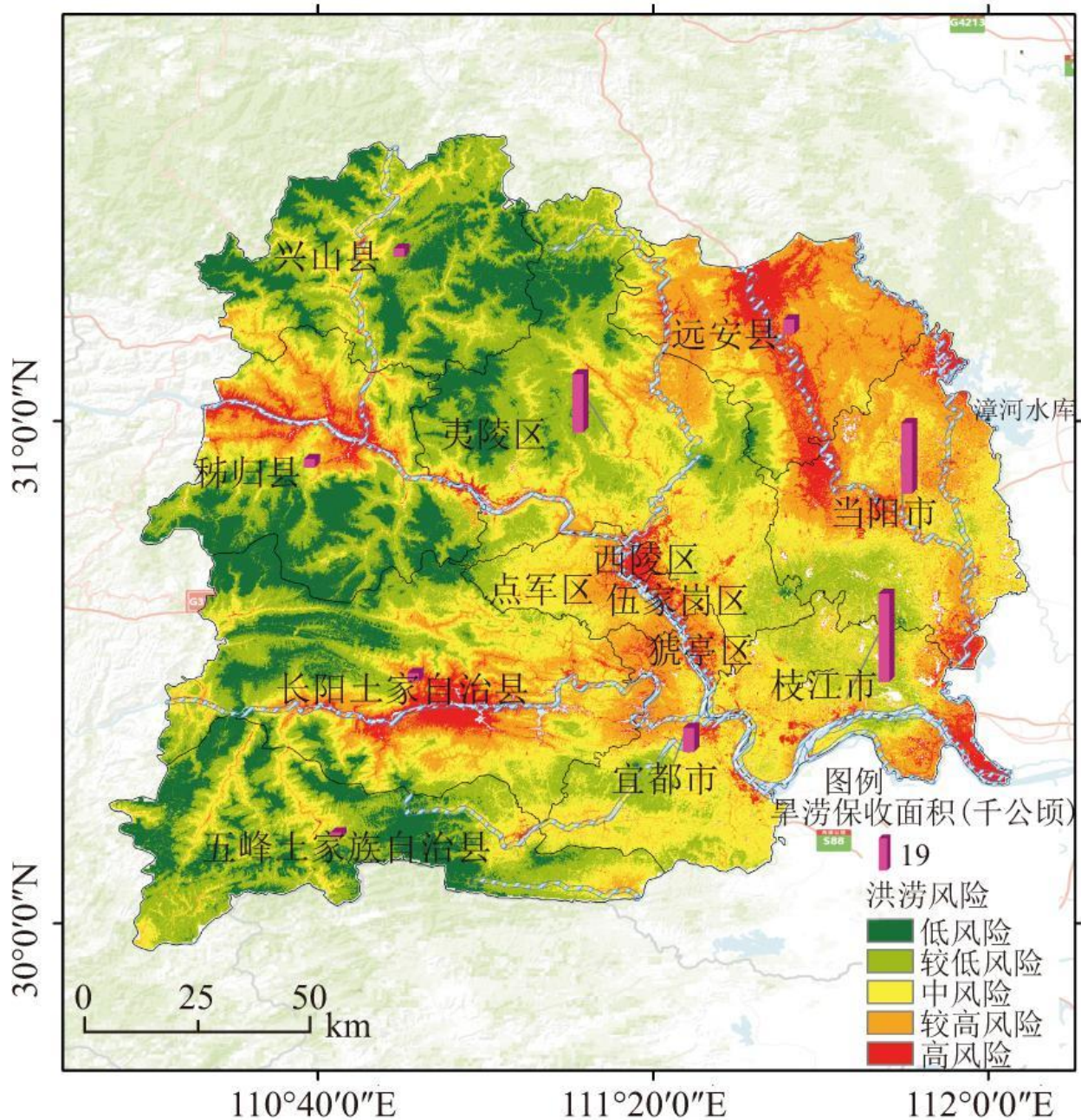


图 13 旱涝保守面积与评估结果对比图

从图 13 中可以发现，在高风险分布较多的区县，保收面积较低，如远安县、长阳县、秭归县；位于低风险地区的区县，保收面积相对较高，如夷陵区、枝江市。但也有兴山县、五峰县等地整体高风险区域占比不大，但旱涝保收面积较低，这可能是由于高风险面积虽占比不高但风险性较强，由于社会经济和地形等其他因素导致旱涝保收工作无法顺利展开。由于洪涝灾害事件的复杂性，其风险的评估与结果的验证相对困难，对数据质量和空间尺度的匹配度要求较高，在今后的风险评估中以期有更全面、更匹配的数据能加入评估体系。

参考文献：

[1]:何锦屏, 李双双, 段克勤, 等.面向事件过程的长江流域极端降水时空变化特征[J].长江流域资源与环境, 2022, 31(6): 1381-1392.

HE J P, LI S S, DUAN K Q, et al. Spatiotemporal variability of event - based extreme precipitation over Yangtze River Basin during 1961-2019[J]. Resources and Environment in the Yangtze Basin, 2022, 31(6): 1381-1392.

[2]孔 锋, 史培军, 方 建, 等.全球变化背景下极端降水时空格局变化及其影响因素研究进展和展望[J].灾害学, 2017, 32(2): 165-174.

KONG F, SHI P J, FANG J, et al. Advances and prospects of spatiotemporal pattern variation of extreme precipitation and its affecting factors under the background of global climate change [J]. Journal of Catastrophology, 2017, 32(2): 165- 174.

[3]傅春, 黄金燕, 付耀宗, 等.不同城市化程度对内涝弹性 的影响研究[J].长江流域资源与环境, 2022, 31(6): 1402- 1412.

FU C, HUANG J Y, FU Y Z, et al. Research on the influence of different urbanization degrees on waterlogging elasticity [J]. Resources and Environment in the Yangtze Basin, 2022, 31(6): 1402- 1412.

[4]胡博亭, 柳 江, 王文玲, 等.基于洪旱灾害的雅鲁藏布江 流域水资源脆弱性时空差异分析[J].长江流域资源与环 境, 2019, 28(5): 1092-1101.

HU B T, LIU J, WANG W L, et al. Temporal - spatial differences of water resources vulnerability in the Yarlung Zangbo River Basin based on the characteristics of flood and drought disasters[J]. Resources and Environment in the Yangtze Basin, 2019, 28(5): 1092-1101.

[5]孙殿臣, 王慧敏, 黄 晶, 等.鄱阳湖流域城市洪涝灾害风 险及土地类型调整策略研究——以景德镇市为例[J].长 江流域资源与环境,2018, 27(12): 2856-2866.

SUN D C, WANG H M, HUANG J, et al. Analysis of urban flood disaster risk in the Poyang Lake basin and land type ad justment strategy study: A case study of Jingdezhen city [J]. Resources and Environment in the Yangtze Basin, 2018, 27(12): 2856-2866.

[6]黄国如, 李碧琦.基于模糊综合评价的深圳市暴雨洪涝风 险评估[J].水资源与水工程学报,2021, 32(1): 1-6.

HUANG G R, LI B Q. Risk assessment of storm water flooding in Shenzhen City based on fuzzy comprehensive evaluation meth-od[J]. Journal of Water Resources and Water Engineering, 2021, 32(1): 1-6.

[7]刘 娜, 金有杰, 李 辉, 等.基于 GIS 的长江中下游城市 洪涝易发性评价[J].地理空间信息, 2020, 18(9): 57-61, 65, 7.

LIU N, JIN Y J, LI H, et al. Urban flood susceptibility assessment in the middle and lower reaches of the Yangtze River based on GIS [J]. Geospatial Information, 2020, 18(9): 57 - 61, 65, 7.

[8]石 涛, 谢五三, 张 丽, 等.暴雨洪涝风险评估的 GIS 和 空间化应用——以芜湖市为例[J].自然灾害学报, 2015, 24(5): 169- 176.

SHI T, XIE W S, ZHANG L, et al. Application of GIS and spatialization of rainstorm and flood risk assessment: A case study of Wuhu City[J]. Journal of Natural Disasters, 2015, 24(5): 169-176.

[9] 姜蓝齐, 马艳敏, 张丽娟, 等. 基于 GIS 的黑龙江省洪涝灾害 风险评估与区划[J]. 自然灾害学报, 2013, 22(5): 238-246.

JIANG L Q, MA Y M, ZHANG L J, et al. Assessment and zoning of flood disaster risk in Heilongjiang P rovince based on GIS [J]. Journal of Natural Disasters, 2013, 22(5): 238 - 246.

[10] 黄 晶, 余靖雯, 袁晓梅, 等. 基于系统动力学的城市洪涝 韧性仿真研究——以南京市为例[J]. 长江流域资源与环 境, 2020, 29(11): 2519-2529.

HUANG J, SHE J W, YUAN X M, et al. Simulation of urban flood resilience based on A system dynamic model: A case study in Nanjing[J]. Resources and Environment in the Yangtze Basin, 2020, 29(11): 2519-2529.

[11] 李国一, 刘家宏. 基于 TELEMAC-2D 模型的深圳洪涝风险 评估[J]. 水资源保护, 2022, 38(5): 58-64.

LI G Y, LIU J H. Flood risk assessment of Shenzhen City based on TELEMAC - 2D model [J] . Water Resources Protection, 2022, 38(5): 58-64.

[12] 叶陈雷, 徐宗学, 雷晓辉, 等. 基于 InfoWorks ICM 的城市 洪涝模拟及其风险分析——以福州市白马河片区为例[J]. 北京师范大学学报(自然科学版), 2021, 57(6): 784-793.

YE C L, XU Z X, LEI X H, et al. InfoWorks ICM flood simulation and risk analysis: Case of Baima River district, Fuzhou [J] . Journal of Beijing Normal University (Natural Science), 2021, 57(6): 784-793.

[13] 吴海春, 黄国如. 基于 PCSWMM 模型的城市内涝风险评估 [J]. 水资源保护, 2016, 32(5): 11-16.

WU H C, HUANG G R. Risk assessment of urban waterlogging based on PCSWMM model [J] . Water Resources Protection, 2016, 32(5): 11-16.

[14] 徐艺扬, 李 昆, 谢玉静, 等. 基于 GIS 的城市内涝影响因 素及多元回归模型研究——以上海为例[J]. 复旦学报(自 然科学版), 2018, 57(2): 182-198.

XU Y Y, LI K, XIE Y J, et al. Studies on the influencing factors and multiple regression model of urban waterlogging based on GIS : A case of Shanghai, China [J] . Journal of Fudan University (Natural Science), 2018, 57(2): 182- 198.

[15] 黄 晶, 余靖雯. 长江三角洲城市群洪涝灾害脆弱性评估 及影响因素分析[J]. 河海大学学报(哲学社会科学版), 2020, 22(6): 39-45, 110-111.

HUANG J, SHE J W. Vulnerability assessment and influencing factors analysis of urban flood disaster in Yangtze River Delta city cluster [J] . Journal of Hohai University (Philosophy and Social Sciences), 2020, 22(6): 39-45, 110-111.

[16] 曾忠平, 王江炜, 邹尚君. 基于 GIS 和逻辑回归分析的山地 城市洪涝灾害敏感性评估——以江西省吉安市为例[J]. 长

江流域资源与环境,2020, 29(9) : 2090-2100.

ZENG Z P, WANG J W, ZOU S J. Susceptibility assessment of flood disaster in mountain cities based on GIS and logistic regression analysis : A case study of Ji' an city, Jiangxi Province [J]. Resources and Environment in the Yangtze Basin, 2020, 29(9) : 2090-2100.

[17] 刘远, 李莹, 王国复, 等.河南省“21·7”极端暴雨灾害多区域间接经济损失评估[J].灾害学, 2022, 37(4): 45-51.

LIU Y, LI Y, WANG G F, et al. Multi-regional indirect economic loss assessment of “21 · 7” extreme rainstorm in Henan Province [J] . Journal of Catastrophology, 2022, 37(4) : 45- 51.

[18] 程朋根, 黄毅, 郭福生, 等.基于多源数据的城市洪涝灾害风险评估[J].灾害学, 2022, 37(3) : 69-76.

CHENG P G, HUANG Y, GUO F S, et al. Urban flooding risk assessment based on multi- source data[J] . Journal of Catastro-phology, 2022, 37(3) : 69-76.

[19] 程朋根, 黄毅.基于 AHP-熵权法的南昌市洪涝风险评估[J].人民长江, 2021, 52(10) : 18-25.

CHENG P G, HUANG Y. Flood risk assessment of Nanchang City based on AHP and Entropy Method [J] . Yangtze River,2021, 52(10) : 18-25.

[20] 叶瑞峰, 潘耀忠, 唐钰嫣, 等.基于主成分分析的城市内涝灾害风险评估——以杭州市临安区为例[J].杭州师范大学学报(自然科学版), 2022, 21(3) : 320-326.

YE R F, PAN Y Z, TANG Y Y, et al. Risk assessment of urban waterlogging disaster based on principal component analysis : An example of Lin' an district of Hangzhou [J]. Journal of Hangzhou Normal University (Natural Science Edition), 2022 , 21(3) : 320-326.

[21] 赵佳慧, 许红师, 王田野, 等.基于改进熵权-TOPSIS-灰色关联方法的城市洪涝风险评估[J].水利水电技术(中英文),2022, 53(10) : 58-73.

ZHAO J H, XU H S, WANG T Y, et al. Improved entropy weight- TOPSIS - grey correlation method - based urban flood -waterlogging risk assessment [J] . Water Resources and Hydropower Engineering, 2022 , 53(10) : 58-73.

[22] 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 J, LIU W, XU Y S, 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.

[23] 周姝天, 翟国方, 施益军, 等.城市自然灾害风险评估研究综述[J].灾害学, 2020, 35(4) : 180-186.

ZHOU S T, ZHAI G F, SHI Y J, et al. A literature review of urban natural disaster risk assessment[J] . Journal of Catastrop-hology, 2020, 35(4) : 180-186.

[24] 黄国如, 罗海婉, 卢鑫祥, 等.城市洪涝灾害风险分析与区划方法综述[J].水资源保护, 2020, 36(6) : 1-6, 17.

HUANG G R , LUO H W, LU X X, et al. Study on risk analysis and zoning method of urban flood disaster [J] . Water Resources Protection, 2020, 36(6) : 1-6, 17.

[25] 汪 怡, 梁 龙, 孙延忠, 等.山西省不可移动文物洪涝灾 害风险研究[J].自然灾害学报,2022, 31(3): 35-47.

WANG Y, LIANG L, SUN Y Z, et al. Research on flood disaster risk of immovable cultural relics in Shanxi Province [J] . Journal of Natural Disasters, 2022, 31 (3) : 35-47.

[26] 陈军飞, 李 倩, 邓梦华, 等.基于随机森林与可变模糊集 的城市洪涝脆弱性评估[J].长江流域资源与环境, 2020, 29(11) : 2551-2562.

CHEN J F, LI Q, DENG M H, et al. Urban flood vulnerability assessment based on random forests and variable fuzzy sets[J] . Resources and Environment in the Yangtze Basin, 2020, 29 (11) : 2551-2562.

[27] 刘媛媛, 刘业森, 郑敬伟, 等.BP 神经网络和数值模型相 结合的城市内涝预测方法研究[J].水利学报, 2022, 53 (3) : 284-295.

LIU Y Y, LIU Y S, ZHENG J W, et al. Intelligent rapid prediction method of urban flooding based on BP neural network and numerical simulation model[J] . Journal of Hydraulic Engineering, 2022, 53(3) : 284-295.

[28] 程先富, 郝丹丹, 韩 平, 等.基于格网数据的巢湖流域 洪涝灾害损失评估[J].长江流域资源与环境, 2014, 23 (10): 1479-1484.

CHENG X F, HAO D D, HAN P , et al. Flood loss assessment in Chaohu Basin based on grid data [J] . Resources and Environment in the Yangtze Basin, 2014, 23 (10) : 1479 - 1484.

[29] 徐永清, 陈 莉, 刘艳华, 等.基于不同空间尺度资料的自 然灾害风险评估对比分析——以暴雨洪涝灾害为例[J]. 灾害学, 2022, 37(3) : 61-68.

XU Y Q, CHEN L, LIU Y H, et al. Comparative analysis of natural disaster risk assessment based on different spatial resolution data: Taking rainstorm and flood disaster as an example [J] . Journal of Catastrophology, 2022 , 37 (3) : 61 -68.

[30] 林奕晨, 周鹏, 潘 悦, 等.荆州市洪涝灾害影响因子探究及风险评估——基于随机森林和 XGBoost 算法[J].中国 农村水利水电, 2022(6): 125-132.

LIN Y C, ZHOU P, PAN Y, et al. Influencing factor research and risk assessment of flood disasters in Jingzhou city: Based on random forest and XGBoost algorithm [J]. China Rural Water and Hydropower, 2022 , (6): 125- 132.

[31] 吕希希, 张春菊, 黄建伟, 等.暴雨洪涝敏感性影响因子分 析及评估——以江西安义县为例[J].灾害学, 2021, 36 (1) : 223-229.

LYU X X, ZHANG C J, HUANG J W, et al. Influencing factors analysis and assessment of storm and flood susceptibility: A case study of Anyi County [J] . Journal of Catastrophology, 2021, 36(1) : 223-229.

[32] 张会, 李钺, 程炯, 等. 基于“H-E-V”框架的城市洪涝风险评估研究进展[J]. 地理科学进展, 2019, 38(2): 175-190.

ZHANG H, LI C, CHENG J, et al. A review of urban flood risk assessment based on the framework of hazard-exposure-vulnerability [J]. Progress in Geography, 2019, 38(2): 175-190.

[33] YANG J, HUANG X. The 30m annual land cover dataset and its dynamics in China from 1990 to 2019[J] - Earth System Science Data, 2021, 13(8): 3907-3925.

[34] 王敬明, 王世新, 王福涛, 等. 基于 Sentinel-1 SAR 数据洪水淹没提取方法研究[J]. 灾害学, 2021, 36(4): 214-220.

WANG J M, WANG S X, WANG F T, et al. Flood inundation region extraction method based on sentinel - 1 SAR data [J] . Journal of Catastrophology, 2021, 36(4): 214-220.

[35] 吴庆双, 汪明秀, 申茜, 等. Sentinel-2 遥感图像的细小水体提取[J]. 遥感学报, 2022, 26(4): 781-794.

WU Q S, WANG M X, SHEN Q, et al. Small water body extraction method based on Sentinel- 2 satellite multi-spectral remote sensing image [J] . National Remote Sensing Bulletin, 2022, 26(4): 781-794.