

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 Xixi, ZHANG Chunju, HUANG Jianwei, 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. doi: 10.3969/j.issn.1000-811X.2021.01.041.]

暴雨洪涝敏感性影响因子分析及评估^{*}

——以江西安义县为例

吕希希¹, 张春菊^{1,2}, 黄建伟¹, 田宇¹, 王小丽¹

(1. 合肥工业大学 土木与水利工程学院, 安徽 合肥 230009; 2. 河南大学 黄河中下游数字地理技术教育部重点实验室, 河南 开封 475004)

摘要: 由于区域地理环境存在差异, 洪涝敏感性影响因子的选取主要根据区域地理特征及个人经验确定, 缺乏统一的标准、系统的分析和科学的评价。洪涝灾害与影响因子之间是复杂的、多变量的非线性关系, 与研究区的地质、地貌、土地覆盖等诸多因素密切相关。针对不同的研究区域, 搜集尽可能全面的影响因子并对其进行优化选取是实现洪涝敏感性准确评估的前提和保证。南昌市安义县位于潦河中下游, 是洪涝灾害的主要受灾区域。基于遥感影像数据和地理信息系统技术, 以安义县为例开展暴雨洪涝敏感性影响因子分析及评估研究。首先, 利用成灾前后哨兵一号雷达影像提取安义县2016年6月30日至7月5日暴雨洪涝的淹没范围, 选取高程、降水、用地类型、距河流距离、坡度等15个洪涝敏感性影响因子。然后, 基于随机森林模型对15个影响因子进行重要性排序, 按照排序结果对影响因子逐步精简, 并基于神经网络模型对影响因子进行优化选取。最后, 基于优化选取后的影响因子, 采用神经网络模型进行安义县洪涝敏感性评估, 并用实例验证洪涝敏感性评估结果的可靠性。研究结果表明, 在精简收敛指数、坡向、剖面曲率、地形位置指数和汇流动力指数5个最不重要的影响因子后, 神经网络模型的性能有一定提升; 敏感性等级为中等及中等以上区域主要分布在潦河两岸, 约占安义县总面积的1/3, 近70%的洪涝分布在敏感性中等及中等以上区域, 洪涝发生在洪涝敏感性等级为中等及中等以上区域的可能性非常大, 洪涝敏感性评估结果与安义县实际情况相符。

关键词: 暴雨洪涝; 敏感性评估; 哨兵一号; 随机森林; 神经网络

中图分类号: X43; X915.5; TV122; P642 **文献标志码:** A **文章编号:** 1000-811X(2021)01-0223-07

doi: 10.3969/j.issn.1000-811X.2021.01.041

洪涝灾害是我国最常见的自然灾害之一, 给我国社会经济带来严重的影响^[1]。暴雨洪涝灾害是由持续一定时间的强降水导致河道决口, 进而引发河水泛滥的洪涝灾害^[2]。洪涝敏感性评估通过分析与自然环境影响因子评估区域发生洪涝灾害的可能性, 是预测洪涝的重要手段和区域防灾减灾的基础^[3]。洪涝敏感性评估的方法主要有层次分析法^[4-5], 统计指标、频率比、证据权重等统计模型^[6-7], 以及支持向量机^[8]、随机森林^[9]、人工神经网络^[10]等机器学习模型。无论是基于统计模型还是机器学习模型的洪涝灾害预测, 均可以通过建立历史洪水淹没区域和导致洪涝发生的潜在的影响因子之间的关系来确定未

来暴雨时该地区可能发生洪涝的区域^[11]。例如, 文献[4]从地形、水系和植被因子出发, 基于层次分析法, 评估了广西洪涝灾害的敏感性; 文献[6]利用统计指标和频率比模型识别北京市房山区西曲沟的易涝区域; 文献[10]通过历史记录和遥感解译获得历史洪水位置, 结合证据权重和支持向量机、随机森林等数据挖掘技术, 评估了江西省鄱阳县的洪涝敏感性。在洪涝敏感性评估的指标体系研究方面, 高程、坡度、坡向、降水、距河流距离、归一化差异植被指数(Normalized Difference Vegetation Index, NDVI)、地形湿度指数(Topographic Wetness Index, TWI)、汇流动力

^{*} 收稿日期: 2020-05-14 修回日期: 2020-08-15

基金项目: 国家级大学生创新创业项目(201910359030); 黄河中下游数字地理技术教育部重点实验室开放基金(CTYR201907)

第一作者简介: 吕希希(1995-), 女, 汉族, 江西九江人, 硕士, 主要从事3S技术及应用研究. E-mail: 1466043734@qq.com

通讯作者: 张春菊(1984-), 女, 汉族, 安徽宿州人, 博士, 副教授, 主要从事地理信息智能处理与服务研究.

E-mail: zcjtzw@sina.com

指数 (Stream Power Index, SPI)、曲率、用地类型、地质和土壤这 12 个影响因子使用的频率较高^[6,10,12]。部分学者在上述影响因子的基础上,引入了平面曲率和剖面曲率^[10]、收敛指数和地形位置指数^[13]、地形耐用指数 (Terrain Ruggedness Index, TRI)^[14] 和 坡度 坡 长 (Length Slope, LS)^[15]。然而,由于区域地理环境存在差异,影响因子的选取主要根据区域地理特征及个人经验确定,尚未有统一的标准,且洪涝敏感性影响因子缺乏系统的分析。此外,洪涝灾害与影响因子之间是复杂的、多变量的非线性关系。神经网络模型具有强大的拟合多变量非线性函数的能力,部分学者初步探索了基于神经网络模型的洪涝预测。例如,通过历史洪水事件和遥感解译获得历史洪涝位置,基于神经网络模型对江西省赣州市定南县进行敏感性制图^[12],但是在洪涝敏感性评估方面忽略了影响因子的重要性差异。因此,为了使选取的影响因子与研究区的实际情况相符,搜集尽可能全面的影响因子并对其进行优化选取是实现洪涝敏感性准确评估的前提和保证。本文以南昌市安义县 2016 年 6 月 30 日至 7 月 5 日遭受的暴雨洪涝灾害为例,基于遥感影像数据和地理信息系统技术,开展暴雨洪涝敏感性影响因子分析及评估研究。首先,选取高程、降水、用地类型、距河流距离、坡度等 15 个暴雨洪涝影响因子;然后,基于随机森林和神经网络模型进行影响因子的优化选取;最后,基于神经网络模型评估安义县的暴雨洪涝敏感性,并通过实例验证洪涝敏感性评估结果的合理性。

1 淹没范围与影响因子提取

1.1 淹没范围提取

本文研究范围位于江西省北部的安义县,地处南昌市西北部,潦河中下游。在图 1 中,右下角是南昌市行政区划示意图,红色区域表示安义县,潦河贯穿全县。安义县属亚热带湿润季风气候,雨量充沛,降水分布不均,暴雨洪涝灾害发生较频繁。

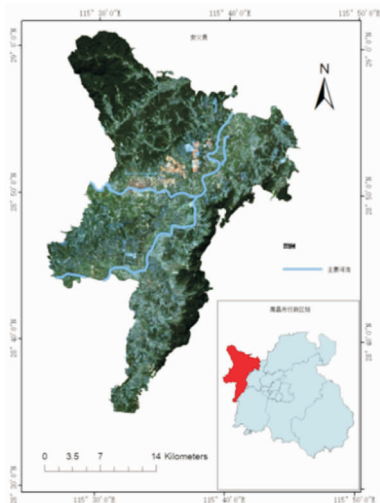


图 1 安义县地理位置

遥感数据被广泛地应用于洪涝水体的提取^[16-17],本文利用 2016 年 5 月 18 日和 2016 年 7 月 5 日的哨兵一号高分辨率双极化地距多视产品数据 (<https://urs.earthdata.nasa.gov/>) 提取洪水淹没范围,其影像数据的空间分辨率为 10 m × 10 m,波束模式是干涉宽幅,极化方式是 VV/VH 双极化。

洪水淹没范围提取技术流程如图 2 所示。首先,对数据进行预处理,包括配准、滤波、地理编码和辐射定标;然后,基于灾前和灾后数据,通过 ENVI 软件进行波段 RGB 彩色影像合成,其中, R 为灾后数据, G 和 B 为灾前数据。再者,以原有水体、水淹区域和陆地建立监督训练样本,采用监督分类法中的马氏距离法^[18]提取研究区域中的洪水淹没范围,并对分类结果进行聚类分析、人工目视校正等后处理,得到洪水淹没区域(参见图 3)。在图 3 中,蓝色是原有水体,红色是洪水淹没区域。

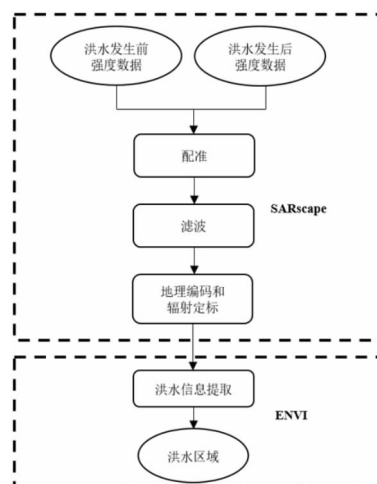


图 2 淹没范围提取技术流程

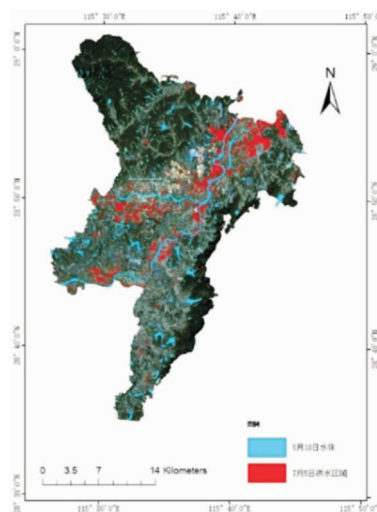


图 3 淹没范围提取结果

1.2 影响因子提取

基于安义县的地理特征以及数据的可获得性,同时参考文献[6, 10, 12-15],选取 15 个影响因子,包括降水、距河流的距离、用地类型、归一化差异植被指数、高程、收敛指数、平面曲率、

剖面曲率、坡度、坡向、地形湿度指数、汇流动力指数、地形位置指数、地形耐用指数、坡度坡长。其中,收敛指数是表示陆地水流收敛或发散的指标。平面曲率垂直于最大坡度的方向,影响流经表面的水流的汇聚和分散。剖面曲率与剖面平行,并指示最大坡度的方向,剖面曲率影响流经某表面水流的加速或减速,进而影响到侵蚀和沉积。地形湿度指数用于推断土壤湿润程度的空间分布和量化地形对水文过程控制作用的大小^[19]。汇流动力指数是描述水流侵蚀地表能力的指标^[20]。地形位置指数是一种局部高程指标,其定义为中心点高程与中心点邻域平均高程的差^[21]。地形耐用指数是一种量化地形异质性的指标,其定义为中心点与中心点邻域高程差的平方的和的平方根^[22]。坡度坡长是解释地形对水土流失和水文过程影响大小的指标^[23]。

从美国地质调查局(<https://earthexplorer.usgs.gov/>)获取空间分辨率为 30 m 的高程数据,使用自动化地球科学分析系统(System for Automated Geoscientific Analyses, SAGA)由数字高程模型提取收敛指数、地形耐用指数和地形位置指数^[24],平面曲率、剖面曲率、坡度、坡向、地形湿度指数、汇流动力指数和坡度坡长利用 ArcGIS 软件由数字高程模型提取。地形湿度指数 TWI ^[19]、汇流动力指数 SPI ^[20] 和坡度坡长 LS ^[23] 分别为

$$TWI = \ln(A_s / \tan \beta); \quad (1)$$

$$SPI = A_s \tan \beta; \quad (2)$$

$$LS = (n + 1) \left(\frac{A_s}{22.13} \right)^n \left(\frac{\sin \beta}{0.0896} \right)^m. \quad (3)$$

式中: n 、 m 为系数, $n = 0.4$, $m = 1.3$, A_s 表示单位等高线长度的集水面积 m^2/m , β 为以弧度表示的坡度^[25]。

从国家气象信息中心(<http://data.cma.cn/>)下载 2016 年 6 月 30 日至 7 月 5 日全国逐日网格降水量,首先利用 ArcGIS 栅格计算器求和得到 2016 年 6 月 30 日至 7 月 5 日研究区的总降水量,然后基于 ArcGIS 克里金插值法得到空间分辨率为 30 m 的降水数据。距河流的距离由主要河流基础数据使用 ArcGIS 欧式距离功能生成。用地类型来源于 2017 年全球 30 m 覆盖数据集^[26]。研究区有农田、森林、草地、灌丛、湿地、水体、不透水层和裸地这八种用地类型。

归一化差异植被指数利用哨兵二号遥感影像获得,哨兵二号数据来自美国地质调查局(<https://earthexplorer.usgs.gov/>)下载的 2018 年 10 月 Level-1C 级数据产品。Sen2Cor 是欧洲航天局发布的用于生成哨兵二号 Level-2A 级大气底层反射率数据产品的处理器,其输入数据是哨兵二号 Level-1C 级大气顶层反射率数据。使用 Sen2Cor 处理器对哨兵二号 Level-1C 级数据进行大气校正,归一化差异植被指数 $NDVI$ 为:

$$NDVI = \frac{NIR - R}{NIR + R}. \quad (4)$$

式中: NIR 和 R 分别代表近红外和红光波段的表面反射率。利用 ArcGIS 重采样方法,将所有数据的空间分辨率转换成与高程相同的空间分辨率(30 m),研究区由 778 841 个栅格组成。

2 暴雨洪涝敏感性影响因子优化选取

2.1 样本准备与精度指标

从所获取的洪水区域随机选取 130 个洪涝样本点,编码为“1”,考虑本节使用的随机森林模型属于机器学习的方法,为了避免由样本不均衡而导致的学习偏见^[27],从没有发生洪涝的区域随机选取相同数量的 130 个非洪涝样本点,编码为“0”,这 130 个洪涝样本点和 130 个非洪涝样本点便构成了总的样本,总样本数为 260 个。

接收者操作特征曲线(Receiver Operating Characteristic, ROC)被广泛地应用于评估模型的整体预测能力^[28],使用 ROC 评估模型的主要优点是可以由曲线下的面积(Area Under Curve, AUC)定量评价模型的预测性能^[9]。AUC 的值在 0.5 ~ 1 之间,一个完美模型的 AUC 值等于 1, AUC 值越接近 1,模型的预测性能越好^[12]。

2.2 基于随机森林的影响因子重要性排序

2.2.1 随机森林简介

随机森林是一种基于决策树的集成学习模型,由 Breiman 于 2001 年提出^[29]。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^[30]:

$$Gini(t) = 1 - \sum_{j=1}^k [p(j|t)]^2. \quad (5)$$

式中: $p(j|t)$ 表示节点 t 上类别 j 的概率,当 $Gini(t)$ 为 0 时,表示节点 t 上的样本属于同一类。根据 Gini 指标进行节点分裂的决策树,特征重要性可由单个特征使森林中所有决策树节点减少的 Gini 值的平均值衡量^[31]。

2.2.2 影响因子重要性分析

利用 python 语言的 scikit-learn 库实现随机森林模型,参数设置如下:决策树个数为 70,其他参数为模型默认参数。70% 的样本点(91 个洪涝样本点和 91 个非洪涝样本点)用于训练,剩下的 30% 的样本点(39 个洪涝样本点和 39 个非洪涝样本点)用于验证^[8]。

模型在验证集上的 AUC 为 0.994 083,说明该模型性能好,结果可靠。影响因子重要性排序如图 4 所示。由图 4 可知:

(1)重要性超过 0.1 的影响因子有 3 个,分别是高程、距河流距离和用地类型,重要性依次是 0.256、0.22 和 0.134,其中高程的重要性最大。

(2)重要性大于 0.05 且小于 0.1 的影响因子有 4 个,分别是坡度、降水、坡度坡长和地形耐用指数,重要性依次是 0.067、0.06、0.058 和 0.054。

(3)重要性小于 0.01 的影响因子有两个,即坡向和收敛指数,坡向的重要性为 0.009,收敛指数的重要性最低,为 0.007。

综合分析上述 15 种影响因子可以看出,高程、距河流距离、用地类型、坡度、降水、坡度坡长和地形耐用指数这七个因子的重要性之和超过 0.8,剩下的影响因子重要性之和小于 0.2。

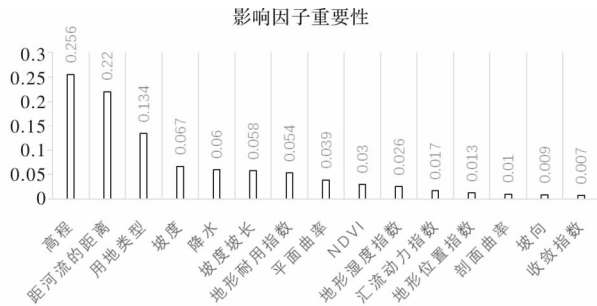


图4 影响因子重要性

2.3 基于神经网络的影响因子优化选取

2.3.1 神经网络简介

神经网络是一种基于真正的生物神经元特征的监督的机器学习算法,通常由输入层、隐藏层和输出层组成^[32]。本文设计的神经网络是一种简单的单隐藏层神经网络模型,其隐藏层只有一层,模型由接收影响因子的输入层、隐藏层和输出洪涝敏感性指标的输出层组成。

2.3.2 影响因子优化选取

根据影响因子重要性排序结果,从最不重要的影响因子开始,逐步精简影响因子,设计了如下9个实验:

实验1: 不精简任何影响因子;

实验2: 精简最不重要的收敛指数;

实验3: 精简收敛指数和坡向;

实验4: 精简收敛指数、坡向和剖面曲率;

实验5: 精简收敛指数、坡向、剖面曲率和地形位置指数;

实验6: 精简收敛指数、坡向、剖面曲率、地形位置指数和汇流动力指数;

实验7: 精简收敛指数、坡向、剖面曲率、地形位置指数、汇流动力指数和地形湿度指数;

实验8: 精简收敛指数、坡向、剖面曲率、地形位置指数、汇流动力指数、地形湿度指数和NDVI;

实验9: 精简收敛指数、坡向、剖面曲率、地形位置指数、汇流动力指数、地形湿度指数、NDVI和平面曲率。

神经网络模型结构及配置如下: 输入层神经元数量等于影响因子的个数,输出层神经元的数量为1。从输入层到隐藏层的激活函数为整流线性单元(Rectified Linear Unit, ReLU),从隐藏层到输出层的激活函数为 sigmoid,优化器为自适应矩估计(Adaptive Moment Estimation, Adam)^[33],损失函数为二分类交叉熵(binary_crossentropy),批大小(batch_size)为16,次数(epoch)为100。

由于隐藏层神经元的数量会影响神经网络模型的性能,为了尽量减小隐藏层神经元数量对神经网络模型性能的影响。首先,针对实验1到实验9依次设计五折交叉验证的网格搜索实验来确定相应的隐藏层神经元的最佳数量,假设输入层神经元的数量为 m ,隐藏层神经元数量网格搜索的范围为 $[m-5, m+5]$,间隔为1,最后比较实验1到

实验9的最优结果。实验1到实验9的最优结果及相应的隐藏层神经元数量见表1。由表1可以看到,从实验1到实验3,随着不重要影响因子的精简,神经网络模型的 AUC 的值先降低后升高,原因可能是前三个实验中精简的影响因子的重要性极小,这些重要性极小的影响因子对神经网络模型性能的影响不稳定,故 AUC 的值有升有降。从实验4到实验6,随着不重要影响因子的继续精简,神经网络模型的 AUC 的值持续升高,原因可能是随着越来越多不重要影响因子的精简,神经网络模型输入数据中包含的噪声越来越少,故神经网络模型的性能得到提高。从实验7到实验9,神经网络模型的 AUC 的值逐渐降低,原因可能是随着越来越多的影响因子被精简,重要性较大的影响因子也被精简,输入数据中的重要信息减少,故神经网络模型的性能降低。由表1可知,从实验7到实验9,即便神经网络模型的 AUC 的值在逐渐降低,然而实验9的 AUC 值依然比前4个实验的 AUC 值要高。本文推测,由于神经网络模型对数据中的噪声比较敏感,比起不加选择地使用影响因子,只使用部分重要性大的影响因子的神经网络模型的性能会更好。实验6的 AUC 值最高,高达0.997 04,即当精简收敛指数、坡向、剖面曲率、地形位置指数和汇流动力指数这5个不重要的影响因子,将高程、距河流距离、用地类型、坡度、降水、坡度坡长、地形耐用指数、平面曲率、NDVI、地形湿度指数这10个影响因子作为输入层的输入,且隐藏层神经元的数量等于14时,神经网络模型性能最优。

表1 实验1到实验9最优结果比较

实验	隐藏层神经元数量	AUC
1	13	0.992 28 ± 0.009 85
2	17	0.987 66 ± 0.019 77
3	13	0.995 85 ± 0.005 42
4	8	0.993 47 ± 0.008 82
5	10	0.996 14 ± 0.005 75
6	14	0.997 04 ± 0.005 22
7	12	0.997 03 ± 0.004 49
8	12	0.996 44 ± 0.005 75
9	8	0.996 14 ± 0.005 01

3 暴雨洪涝敏感性评估

3.1 基于神经网络的暴雨洪涝敏感性评估

由前文可知,实验6的神经网络模型性能最优,故采用该模型对安义县的洪涝敏感性进行评估。用于洪涝敏感性评估的神经网络模型的结构和配置如图5所示。该模型输入层有10个神经元,输入高程、距河流距离、用地类型、坡度、降水、坡度坡长、地形耐用指数、平面曲率、NDVI、地形湿度指数这10个影响因子,隐藏层有14个神经元,隐藏层的激活函数是 ReLU,输出层有1个神

经元,输出层的激活函数为 Sigmoid,用于将输出层的数值转化到区间(0,1),获得洪涝敏感性指标。此外,该模型使用的优化器是 Adam,模型使用的损失函数是二分类交叉熵。

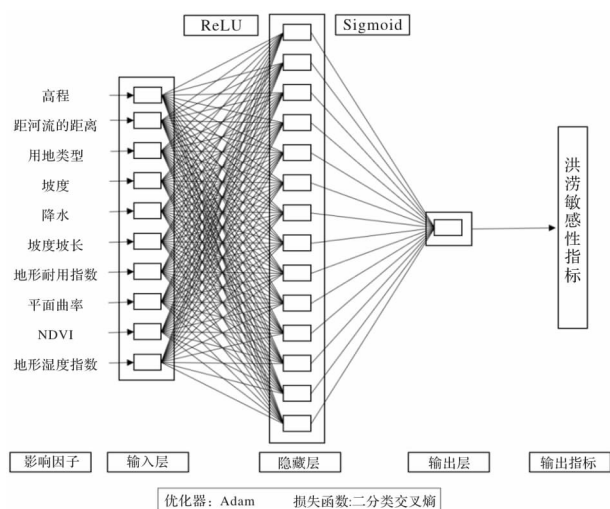


图5 洪涝敏感性评估的神经网络模型结构和配置

利用 ArcGIS 软件将模型输出的洪涝敏感性指标转化为栅格格式,洪涝敏感性指标的范围是(0,1),该指标代表在当前给定的影响因子下洪涝发生的概率,安义县洪涝敏感性指标的制图表达如图6所示。为了更加准确地表征安义县的洪涝敏感性区域划分,本文基于几何断点法,进一步将安义县的敏感性指标分为5个等级:非常低(0~0.0036)、低(0.0036~0.0602)、中等(0.0602~0.9397)、高(0.9397~0.9963)和非常高(0.9963~1)。安义县的洪涝敏感性等级制图如图7所示。由图可以看出,敏感性非常高的地区主要分布在潦河沿岸,且两岸分布的宽度大致相等,敏感性高的地区主要分布在安义县的北部、潦河的上游,敏感性中等和低的地区主要分布在潦河的两岸,非常高、高、中等和低敏感性地区由潦河沿岸向外扩展,敏感性非常低的地区主要分布在安义县的西北和东南部。

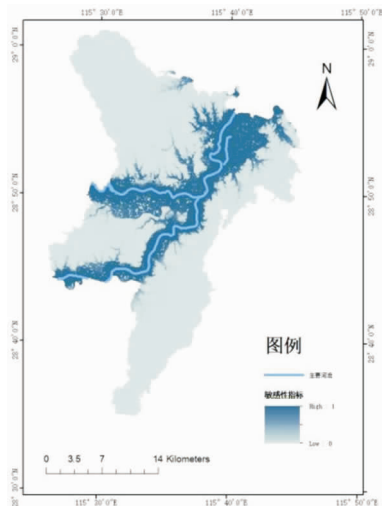


图6 洪涝敏感性指标的制图表达

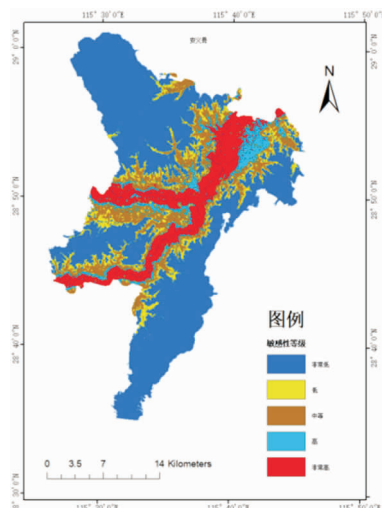


图7 安义县洪涝敏感性等级制图

本文同样定量统计了非常低、低、中等、高和非常高这五类敏感性地区占安义县总面积的百分比,统计结果见表2。其中,敏感性非常低的地区所占百分比最大,为58.60%,敏感性高的地区所占百分比最小,仅为6.24%,敏感性中等和非常高的地区所占百分比几乎相等,均在13%左右,敏感性低的地区所占百分比为8.96%,洪涝敏感性等级为中等及中等以上地区大约占安义县总面积的1/3。

表2 各敏感性等级占安义县百分比表

敏感性等级	面积百分比/%
非常低	58.60
低	8.96
中等	13.14
高	6.24
非常高	13.06

3.2 实例验证与分析

为了验证洪涝敏感性评估结果,利用2016年7月17日哨兵一号雷达影像,采用上文同样方法进行安义县2016年7月5日之后发生的洪涝灾害的洪水淹没范围提取,并将此洪水淹没范围与上文2016年7月5日洪涝敏感性制图进行叠加分析,如表3所示。第一列代表了洪涝敏感性等级,第二列是各敏感性等级面积占安义县总面积的百分比,第三列是位于各敏感性等级的2016年7月17日洪涝面积占总的洪水淹没面积的百分比,第四列是第三列与第二列的比值,该比值被称为频率比。该频率比可体现敏感性等级与淹没范围之间的相关性,如果频率比大于1,表示敏感性等级与洪涝发生之间的相关性较强;如果频率比小于1,则意味着较低的相关性^[34]。

表3 淹没范围与敏感性等级相关性分析表

敏感性等级	面积百分比/%	洪涝面积百分比/%	频率比/%
非常低	58.60	24.04	0.41
低	8.96	6.49	0.72
中等	13.14	21.00	1.60
高	6.24	16.28	2.61
非常高	13.06	32.19	2.47

从洪水区域的分布来看,敏感性非常高的区域分布的洪涝面积最大,达到洪水区域的 32.19%,敏感性中等和敏感性非常低的区域分布的洪涝面积百分比接近,均在 20%~25%之间,16.28%的洪水区域分布在敏感性高的区域,敏感性低的地区分布的洪涝面积最小,仅占洪水区域的 6.49%,近 70%的洪涝分布在敏感性中等及中等以上区域,敏感性非常低和低的区域只分布着大约 30%的洪涝。

从频率比模型的结果来看,敏感性高的区域的频率比最大,为 2.61,敏感性非常高的区域的频率比紧随其后,为 2.47,敏感性非常低的区域的频率比最小,为 0.41。敏感性非常低和低的区域频率比均小于 1,说明洪涝发生在敏感性非常低和低的区域概率较小,敏感性中等及中等以上区域的频率比均大于 1,说明洪涝更有可能发生在敏感性中等及中等以上区域。由洪水区域与安义县洪涝敏感性制图叠加分析的结果可知,洪涝敏感性评估结果是科学合理的,与安义县实际情况相符。

4 结论

本文以南昌市安义县为例,基于遥感影像数据和地理信息系统技术,开展暴雨洪涝敏感性影响因子分析及评估研究。主要结论如下:

首先,进行暴雨洪涝的淹没范围和高程、降水、用地类型和坡度等 15 个影响因子提取,基于随机森林模型对暴雨洪涝的 15 个影响因子进行重要性分析与排序,重要性最大的 7 个影响因子依次为高程、距河流距离、用地类型、坡度、降水、坡度坡长和地形耐用指数,这 7 个影响因子的重要性之和超过 0.8。

然后,基于随机森林和神经网络模型进行影响因子优化选取,在精简收敛指数、坡向、剖面曲率、地形位置指数和汇流动力指数这 5 个重要性靠后的影响因子后,神经网络模型的性能从 0.992 28 提高到 0.997 04。

最后,采用神经网络模型进行安义县洪涝敏感性评估,敏感性等级为中等及中等以上区域主要分布在潦河两岸,约占安义县总面积的 1/3,敏感性非常高地区在潦河两岸的分布宽度基本相同,但是高和中等敏感性地区在潦河两岸的分布范围具有一定的差异;中等及中等以上敏感性区域分布着近 70%的洪涝,因此安义县洪涝敏感性评估结果是科学合理的,应重点留意中等及中等以上敏感性地区,为区域制定适宜的防灾减灾措施和土地利用规划。

本文主要存在以下不足:尽管本文试图获取尽可能全面的影响因子,然而由于数据的限制,没有考虑安义县的土壤类型以及地质等影响因子,未来考虑更加全面的影响因子;由于数据的限制,本文使用的是研究区 30 m 空间分辨率的数字高程模型,基于数字高程模型提取的影响因子的空间分辨率也为 30 m,进一步的研究考虑空间分辨率更高的数字高程模型。

参考文献:

[1] 万金红,张葆蔚,刘建刚,等. 1950-2013 年我国洪涝灾情

- 时空特征分析[J]. 灾害学, 2016, 31(2): 63-68.
- [2] 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.
- [3] 田亚平,向清成,王鹏. 区域人地耦合系统脆弱性及其评价指标体系[J]. 地理研究, 2013, 32(1): 55-63.
- [4] 莫建飞,陆甲,李艳兰,等. 基于 GIS 的广西洪涝灾害孕灾环境敏感性评估[J]. 灾害学, 2010, 25(4): 33-37.
- [5] 方秀琴,王凯,任立良,等. 基于 GIS 的江西省山洪灾害风险评估与分区[J]. 灾害学, 2017, 32(1): 111-116.
- [6] CAO C, XU P, WANG Y, et al. Flash flood hazard susceptibility mapping using frequency ratio and statistical index methods in coalmine subsidence areas[J]. Sustainability, 2016, 8(9): 948-965.
- [7] RAHMATI O, POURGHASEMI H R, ZEINIVAND H. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran[J]. Geocarto International, 2016, 31(1): 42-70.
- [8] TEHRANY M S, PRADHAN B, MANSOR S, et al. Flood susceptibility assessment using GIS-based support vector machine model with different kernel types[J]. Catena, 2015, 125: 91-101.
- [9] CHAPI K, SINGH V P, SHIRZADI A, et al. A novel hybrid artificial intelligence approach for flood susceptibility assessment[J]. Environmental modelling & software, 2017, 95: 229-245.
- [10] HONG H, TSANGARATOS P, LLIA I, et al. Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China[J]. Science of the Total Environment, 2018, 625: 575-588.
- [11] BUI D T, HOANG N D. A Bayesian framework based on a Gaussian mixture model and radial-basis-function Fisher discriminant analysis (BayGmmKda V1. 1) for spatial prediction of floods[J]. Geoscientific Model Development, 2017, 10(9): 3391-3409.
- [12] WANG Y, HONG H, CHEN W, et al. Flood susceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm[J]. Journal of environmental management, 2019, 247: 712-729.
- [13] COSTACHE R D, 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.
- [14] MOJADDADI H, PRADHAN B, NAMPAK H, et al. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS[J]. Geomatics Natural Hazards & Risk, 2017, 8(2): 1080-1102.
- [15] SUNMIN L, SARO L, MOUNG-JIN L, et al. Spatial Assessment of Urban Flood Susceptibility Using Data Mining and Geographic Information System (GIS) Tools[J]. Sustainability, 2018, 10(3): 648.
- [16] 饶品增,蒋卫国,王晓雅,等. 基于 MODIS 数据的洪涝灾害分析研究——以 2017 年洞庭湖区洪水为例[J]. 灾害学, 2019, 34(1): 203-207.
- [17] 吴效勇,王晓青,丁玲,等. 基于光学和 SAR 影像的老挝溃坝洪涝灾害监测与评估[J]. 灾害学, 2020, 35(1): 211-215.
- [18] 成淑艳,曹生奎,曹广超,等. 基于高分辨率遥感影像的青海湖沙柳河流域土地覆盖监督分类方法对比[J]. 水土保持通报, 2018, 38(5): 261-268.
- [19] GRABS T, SEIBERT J, BISHOP K, et al. Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model[J]. Journal of Hydrology, 2009, 373(1/2): 15-23.
- [20] GOKCEOĞHU C, SONMEZ H, NEFESLIOĞLU H A, et al. The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity[J]. Engineering geology, 2005, 81(1): 65-83.
- [21] DE REU J, BOURGEOIS J, BATS M, et al. Application of the topographic position index to heterogeneous landscapes[J]. Geomorphology, 2013, 186: 39-49.
- [22] RILEY S J, DEGLORIA S D, ELLIOT R. Index that quantifies topographic heterogeneity[J]. intermountain Journal of sciences, 1999, 5(1/4): 23-27.
- [23] MOORE I D, GRAYSON R B, LADSON A R. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications[J]. Hydrological processes, 1991, 5(1): 3

- 30.
- [24] CONRAD O, BECHTEL B, BOCK M, et al. System for automated geoscientific analyses (SAGA) v. 2. 1. 4[J]. Geoscientific Model Development, 2015, 8(7): 1991–2007.
- [25] 梁倍瑜, 徐亚莉, 罗明良, 等. 三种重采样方法对冲沟坡度坡长因子的影响[J]. 土壤学报, 2018, 55(1): 64–74.
- [26] GONG P, LIU H, ZHANG M, et al. Stable classification with limited sample: Transferring a 30 – m resolution sample set collected in 2015 to mapping 10 – m resolution global land cover in 2017[J]. Science Bulletin, 2019, 64: 370–373.
- [27] NGO P T T, HOANG N D, PRADHAN B, et al. A novel hybrid swarm optimized multilayer neural network for spatial prediction of flash floods in tropical areas using Sentinel-1 SAR imagery and geospatial data[J]. Sensors, 2018, 18(11): 3704–3729.
- [28] BUI D T, PRADHAN B, NAMPAK H, et al. Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS[J]. Journal of Hydrology, 2016, 540: 317–330.
- [29] BREIMAN L. Random forests[J]. Machine learning, 2001, 45(1): 5–32.
- [30] WANG Z, LAI C, CHEN X, et al. Flood hazard risk assessment model based on random forest[J]. Journal of Hydrology, 2015, 527: 1130–1141.
- [31] VERIKAS A, GELZINIS A, BACAUSKIENE M. Mining data with random forests: A survey and results of new tests[J]. Pattern recognition, 2011, 44(2): 330–349.
- [32] RUMELHART D, WILLIAMS R, HINTON G R. Learning representations by back-propagating errors [J]. Nature, 1986, 323: 533–536.
- [33] KINGMA D P, BA J L. Adam: A method for stochastic optimization [C]//3rd International Conference on Learning Representations, ICLR 2015-Conference Track Proceedings, 2015.
- [34] AKGUN A, DAG S, BULUT F. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models[J]. Environmental Geology, 2008, 54(6): 1127–1143.

Influencing Factors Analysis and Assessment of Storm and Flood Susceptibility——A Case Study of Anyi County

LYU Xixi¹, ZHANG Chunji^{1,2}, HUANG Jianwei¹, TIAN Yu¹ and WANG Xiaoli¹

(1. College of Civil Engineering, Hefei University of Technology, Hefei 230009, China;

2. Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions Ministry of Education, Henan University, Kaifeng 475004, China)

Abstract: Due to differences in regional geographic environments, the selection of influencing factors of flood susceptibility is mainly determined based on regional geographic characteristics and personal experience, and lacks unified standards, systematic analysis and scientific evaluation. The relationship between flood disasters and influencing factors is a complex, multi-variable, non-linear relationship, which is closely related to the geology, landform, land cover and many other factors of the study area. For different research areas, collecting as comprehensive as possible influencing factors and optimizing them are the prerequisite and guarantee for realizing accurate assessment of flood susceptibility. Anyi County, Nanchang City is located in the middle and lower reaches of the Liaohe River and is the main area affected by flood disasters. Based on remote sensing image data and geographic information system technology, this article takes Anyi County, Nanchang City as an example to carry out the influencing factors analysis and susceptibility assessment of storm and flood. First, Sentinel-1 radar images before and after the disaster were used to extract the inundation range of storm and flood in Anyi County from June 30 to July 5, 2016, and 15 influencing factors of flood susceptibility were selected such as elevation, precipitation, land use, distance from the rivers and slope, etc. Then, based on the random forest, the importance of influencing factors was obtained, and the unimportant influencing factors were gradually reduced according to the ranking results of the importance, and the influencing factors were optimized and selected based on the neural network model. Finally, based on the optimized influencing factors, the neural network model was used to evaluate the flood susceptibility of Anyi County, and an example was used to verify the reliability of the flood susceptibility assessment results. The research results show that after reducing the five least important influencing factors of convergence index, aspect, profile curvature, topographical position index and stream power index, the performance of the neural network model has been improved to a certain extent; Areas with moderate or higher flood susceptibility levels are mainly distributed on both banks of the Liaohe River, occupying approximately one third Anyi County, and nearly 70% of the floods are distributed in areas with moderate or higher flood susceptibility levels, the possibility of floods occurring in areas with flood susceptibility levels of medium or above is very high, and the result of flood susceptibility assessment is consistent with the actual situation in Anyi County.

Key words: storm and flood; susceptibility assessment; sentinel-1; random forest; neural network