

基于 PSO-SVM 的鹤盛流域山洪风险评价

王如锴^{1,2}, 练继建^{1,2,3}, 苑希民^{1,2}, 田福昌^{1,2}, 陈隆吉⁴, 马文豪^{1,2}

(1. 天津大学水利工程仿真与安全国家重点实验室, 天津 300350; 2. 天津大学建筑工程学院, 天津 300350;
3. 天津理工大学海洋能源与智能建设研究院, 天津 300384; 4. 温州市水文管理中心, 浙江 温州 325000)

摘要:为探究温州市鹤盛流域山洪风险空间分布,综合考虑山洪致灾因子、孕灾环境和承灾体3方面的山洪影响因子,建立基于粒子群优化-支持向量机(PSO-SVM)混合算法的山洪风险评价模型。选取准确度、灵敏度、特异性、F-score值、Kappa系数和受试者工作特征曲线等6个指标,采用学习矢量量化(LVQ)算法量化山洪影响因子对山洪灾害发生的影响程度,并将 PSO-SVM 混合算法模型与单一算法模型进行对比。结果表明:混合算法具有一定的迁移能力,能够更加准确地反映山洪风险的空间分布特征;验证集受试者工作特征曲线指标、Kappa系数和准确度分别为0.934、0.833、0.912,PSO-SVM混合算法模型能显著提高山洪风险评价精度。

关键词:山洪灾害;风险评价模型;山洪影响因子;PSO-SVM;鹤盛流域

中图分类号:TV877

文献标志码:A

文章编号:1004-6933(2024)02-0046-09

Risk assessment of torrential floods in the Heshengxi Watershed based on PSO-SVM//WANG Rukai^{1,2}, LIAN Jijian^{1,2,3}, YUAN Ximin^{1,2}, TIAN Fuchang^{1,2}, CHEN Longji⁴, MA Wenhao^{1,2} (1. State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300350, China; 2. School of Civil Engineering, Tianjin University, Tianjin 300350, China; 3. Ocean Energy and Intelligent Construction Research Institute, Tianjin University of Technology, Tianjin 300384, China; 4. Wenzhou Hydrological Management Center, Wenzhou 325000, China)

Abstract: To explore the spatial distribution of torrential flood risk in the Heshengxi Watershed of Wenzhou City, a risk assessment model of torrential flood based on particle swarm optimization-support vector machine (PSO-SVM) hybrid algorithm was established, taking into account the torrential flood causing factors, the disaster-prone environment, and the hazard bearing bodies. Six evaluation indicators were selected, including accuracy, sensitivity, specificity, F-score value, Kappa coefficient, and subject working characteristic curve. The learning vector quantization algorithm was used to quantify the impact of torrential flood impact factors on the occurrence of torrential flood disasters, and the PSO-SVM hybrid algorithm model was compared with single algorithm models. The results show that the hybrid algorithm has a certain transfer ability and can more accurately reflect the spatial distribution characteristics of torrential flood risk. The working characteristic curve indicators, Kappa coefficient, and accuracy of the validation set subjects were 0.934, 0.833, and 0.912, respectively. The PSO-SVM hybrid algorithm model can significantly improve the accuracy of torrential flood risk assessment.

Key words: torrential flood disaster; risk assessment model; torrential flood impact factors; PSO-SVM; Heshengxi Watershed

洪涝灾害是发生频繁且影响广泛的自然灾害之一^[1]。山丘地区暴雨洪水受地形影响较大,水土流失严重,山洪及其伴生的次生灾害可能给流域防洪安全带来更大威胁^[2]。开展山洪风险评价,准确评估山洪风险的空间分布和内在规律,对于指导山区小流域防洪规划和提高山洪防治能力具有重要意

义。目前山洪风险评价方法主要有水文学方法、统计学方法和数据驱动方法^[3]。水文学方法通过模拟流域洪水过程计算河道水位和流量,从而确定洪水风险区域^[4],对于形态不稳定、存在高侵蚀等特性的河道,无法对河道和淹没区进行综合分析^[5]。统计学方法需要大量历史资料,对于水文监测站点

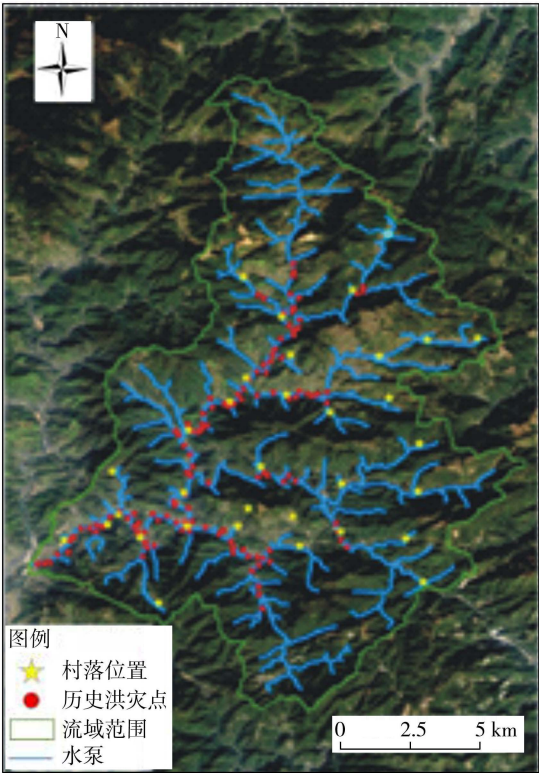
基金项目:国家重点研发计划项目(2022YFC3202501);水利部重大科技项目(SKS-2022002);科技部重点领域创新团队项目(2014RA4031);国家自然科学基金委创新团队项目(51621092)

作者简介:王如锴(1995—),男,博士研究生,主要从事防洪减灾及水利信息化研究。E-mail:rukaiwang@tju.edu.cn

通信作者:苑希民(1968—),男,教授,博士,主要从事防洪减灾及水利信息化研究。E-mail:yxm@tju.edu.cn

稀缺的山区小流域,难以反映山洪风险与流域地形水文变量间的统计学关系^[6]。近年来随着人工智能技术的发展,机器学习算法因其具有较强的非线性处理能力,已广泛应用于山洪风险评价^[7]。机器学习算法可分为两种建模类型:单一算法模型与混合模型,大量研究表明混合模型能有效降低输入数据的噪声、方差和过拟合问题,从而提高预测精度^[8-11]。目前利用机器学习算法进行山洪风险评价仍然存在一些挑战:①决策树需要进行大量数据预处理工作,同时容易陷入局部最优解的问题;②人工神经网络模型存在训练速度慢、参数多等问题^[12];③遗传算法初始种群及参数的选择依赖人为经验,且存在搜索速度较慢等问题^[13];④流域不同风险单元间存在相互作用和影响,但目前大多数山洪风险评价模型只考虑单个风险单元的特征和环境条件,难以全面解释流域的风险分布和变化趋势。

鉴于此,本文综合考虑山洪致灾因子、孕灾环境和承灾体 3 个方面,依托粒子群优化 (particle swarm optimization, PSO) 算法较强的全局搜索能力以及支持向量机 (support vector machine, SVM) 算法良好的分类能力,构建基于 PSO-SVM 混合算法的山洪灾害风险评价模型,对温州市永嘉县鹤盛溪流域开展山洪风险评价,探究鹤盛溪流域山洪灾害风险分布规律,以为山洪风险管理、防洪规划等提供科学依据。



(a) 历史洪灾分布

1 研究区概况与数据资料

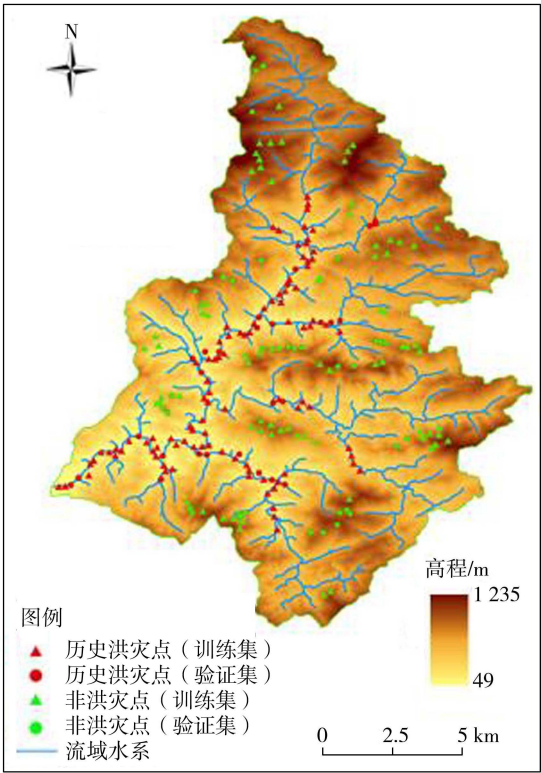
1.1 研究区概况

鹤盛溪流域位于浙江省温州市永嘉县,地处东经 120° 40' 34" ~ 121° 29' 19", 北纬 27° 58' 40" ~ 28° 36' 54", 发源于黄岩区富山乡后岙村,由北向南流经云岭乡、鹤盛镇,在鹤盛镇上日川村双溪口入楠溪江。鹤盛溪流域面积约 313 km²,主流长 40.3 km,平均比降 17.1‰,平均河宽 65m。鹤盛溪流域属亚热带海洋性季风气候,雨量充沛,流域内降水年际变化较大,主要集中在 4—9 月,占全年降水量的 76%。降水主要为春雨、梅雨和台风雨,后两者为形成洪水的主要因素。

1.2 数据资料

1.2.1 历史洪灾分布

历史洪灾分布是确定洪水发生及其驱动因素间关系的重要数据基础,也是山洪风险评价模型训练及验证的基本数据。近年来,遥感和地理信息系统已成为确定历史洪灾分布的重要方法,本文通过实地调研和对比研究区洪灾发生前后的卫星图像,结合 2010—2022 年水利年鉴记载的洪灾事件,确定研究区历史洪灾分布 (图 1(a))。在历史洪灾淹没区内随机选取 111 个洪灾点,为保证数据平衡,在没有任何历史洪水记录的区域内随机选择相同数量的非



(b) 模型样本数据集

图 1 研究区历史洪灾分布与模型样本数据集

Fig. 1 Historical flood distribution and model sample dataset in study area

洪灾点,组成模型样本数据集,见图 1(b)。为便于模型分类识别,将历史洪灾点风险等级量化为 1,非洪灾点风险等级量化为 0。同时将历史洪灾点和非洪灾点样本数据集随机分为两组,78 处(70%)用于模型训练,33 处(30%)用于模型验证。根据相关洪水调查报告和文献[14]的描述,鹤盛流域历史受灾村落及受灾情况如表 1 所示,受流域暴雨的影响,下游地区村落如东皋村和鹤盛村受灾最为严重,普遍淹没水深为 1~1.5 m,而中游地区的村落如梅坦村和上埠村,普遍淹没水深为 0.5 m 左右。

1.2.2 山洪影响因子

确定山洪影响因子是进行山洪风险评价的必要条件^[15]。Li 等^[16]研究指出山区流域经短历时强降雨后,陡坡地区径流下渗和植被截流时间较短,导致产汇流迅速进而引发下游河道周围地区洪水泛滥。因此孕灾环境因子选取距河道距离、坡度、高程和坡向。短历时强降雨是山洪灾害发生的驱动力^[17],同时郑德凤等^[18]研究表明河道及流域地形指数也是主要致灾因素。Fuller^[19]分析了新西兰洪水的地貌影响,指出高河流功率指数可能导致河道剧烈变化进而影响流域水系格局。地形湿度指数可量化地形对包括洪水在内的基本水文过程的控制,并反映流域湿度的空间变化。因此致灾因子选取年平均降水量、河流功率指数和地形湿度指数。人口密集区往往城镇化进程较快,但同时会降低地表渗透能力使其更易受洪灾影响。土地利用与土地覆盖对径流速度、拦截、渗透和蒸发输送等起着关键影响,不同土地利用影响流域径流范围^[20]。因此承灾体选择人口密度和土地利用。

本文采用的高程、土地利用、人口密度数据来源于温州市水文管理中心;年平均降水量数据来源于浙江省短历时暴雨图集。对上述指标进行极差标准化处理,生成 9 个栅格数据图层(图 2),在满足计算精度要求的基础上利用 GIS 技术生成全流域

12.5 m×12.5 m 共计 1 989 237 个计算单元,通过空间链接方式将栅格数据赋值到计算单元中,图 2(d)中坡向为-1°表示平坡。

2 研究方法

利用流域数字高程、遥感、气象和河道矢量等数据生成流域山洪影响因子矢量图层,并通过实地调研获取流域历史山洪灾害点;计算影响因子间方差膨胀因子与公差以消除可能存在多重共线性的问题,并采用学习矢量量化(learning vector quantization, LVQ)算法量化影响因子权重,最终确定鹤盛流域山洪影响因子;建立基于 PSO-SVM 混合算法的山洪风险评价模型,选取准确度、灵敏度、特异性、F-score 值和 Kappa 系数 5 个统计学评价指标^[10]和受试者工作特征曲线对 PSO-SVM 混合算法模型结果与单一算法模型(PSO、SVM)结果进行比较。山洪风险评价模型流程如图 3 所示。

2.1 多重共线性检验

在回归分析中,两个或两个以上自变量之间存在相关性称作多重共线性。由于预测模型中的各变量存在多重共线性会导致模型不稳定,造成回归系数、截距系数的估计不准确^[21],本文采用方差膨胀因子和公差进行模型变量的多重共线性检验,方差膨胀因子小于阈值 10,公差大于阈值 0.1 表明变量间不存在多重共线性。计算公式为

$$V = 1/T \tag{1}$$

其中

$$T = 1 - R_j^2$$

式中:V 为方差膨胀因子;T 为公差;R_j 为第 j 个山洪影响因子作为因变量,其山洪影响因子作为余变量的回归拟合系数。

2.2 权重量化分析

LVQ 算法是一种基于向量量化的有监督学习算法,可用于分类和回归。其算法思想是通过逐步调整聚类中心,使得不同类别之间的距离变大,同

表 1 鹤盛流域历史受灾情况

Table 1 Historical disaster situation of the Heshengxi Watershed

受灾村落	黄海高程/m	地理位置		最大淹没水深/m	
		东经	北纬	“黑格比”台风洪灾	“利奇马”台风洪灾
东皋村	73	120°48′14″	28°22′52″	1.1	1.4
鹤盛村	87	120°50′4″	28°22′32″	1.0	1.3
罗家寮村	260	120°52′57″	28°22′42″	0.6	0.8
霞岭根村	210	120°54′15″	28°22′17″	0.8	1.0
上埠村	168	120°52′8″	28°23′58″	0.4	0.5
西源村	113	120°50′12″	28°24′53″	0.9	1.2
新源村	122	120°51′49″	28°26′6″	1.2	1.4
梅坦村	143	120°52′53″	28°25′39″	0.9	1.1
腾家龙村	306	120°53′5″	28°26′30″	0.5	0.6
云栖村	420	120°54′47″	28°28′8″	0.3	0.4

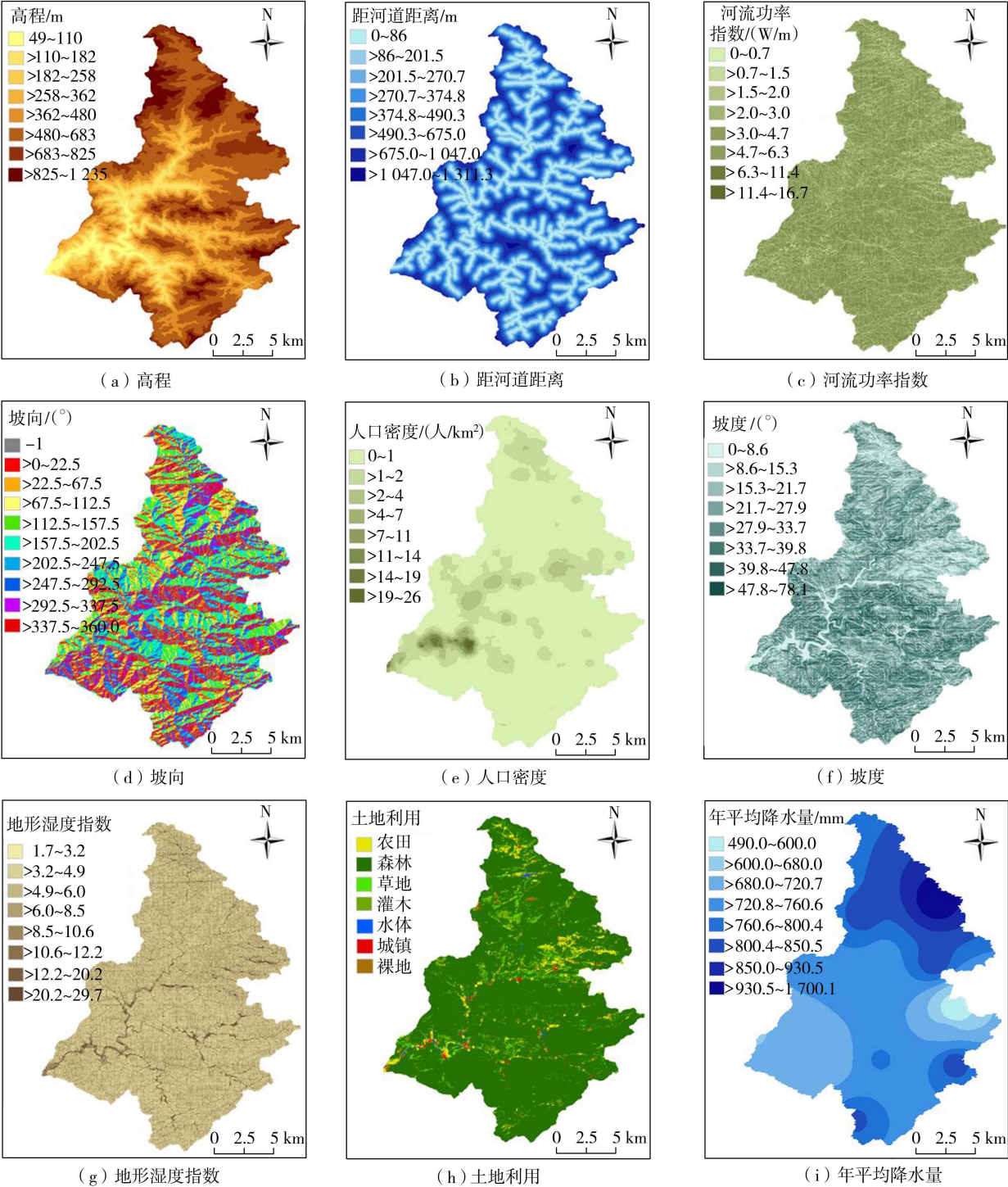


图2 山洪影响因子

Fig.2 Torrential flood conditioning factors

一类别之间的距离变小,从而实现分类任务^[22]。本文利用LVQ算法分析山洪风险分布和地形水文变量间的关系,确定山洪影响因子权重。距离计算公式为

$$D_i = \sqrt{\sum_{j=1}^R (x_j - w_{ij})^2} \quad (i = 1, 2, \dots, n) \quad (2)$$

式中: D_i 为原型向量与竞争层神经元*i*之间的欧氏距离; w_{ij} 为训练向量与竞争层神经元*i*之间的权重;

R 为竞争层神经元总数; n 为流域计算单元总数。

2.3 PSO-SVM 混合算法

PSO 是一种基于鸟类捕食行为研究的进化计算方法,该算法受鸟类觅食过程、启发式特征以及进化算法的随机搜索启发,可有效解决复杂的优化问题,其独特之处在于,PSO 算法通过模拟群体中粒子的随机运动,寻找全局最优解,具有快速收敛、高效搜索等特征^[23]。在该算法中,每个粒子视为 N 维搜索

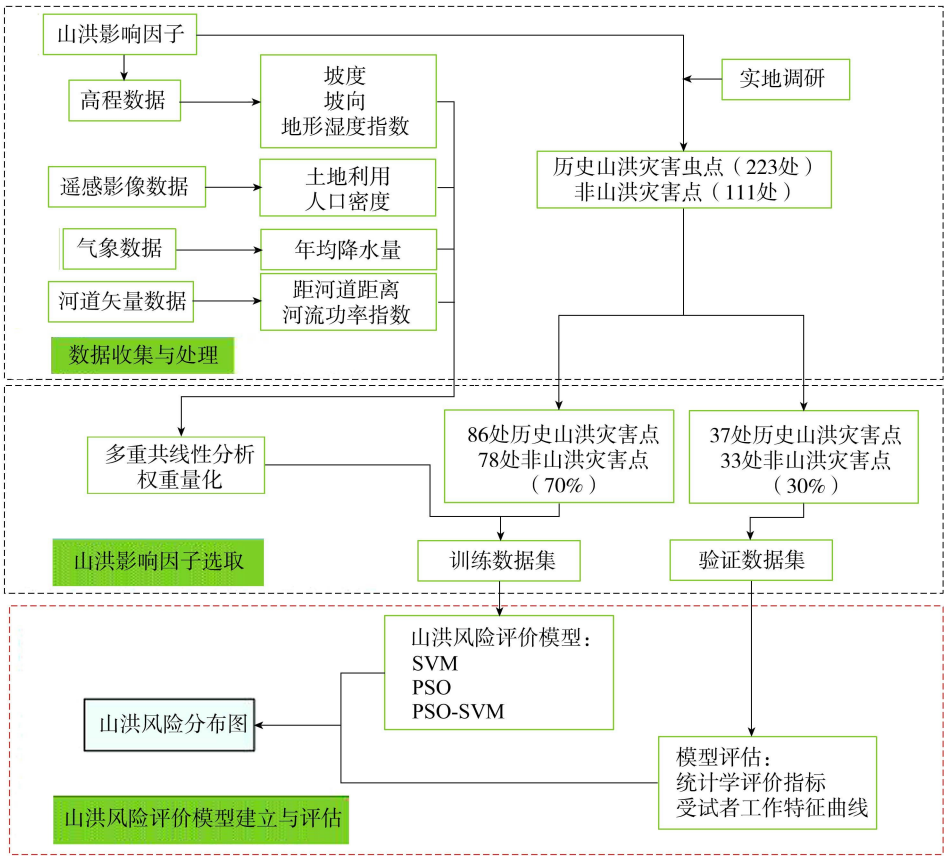


图3 山洪风险评价模型流程

Fig.3 Flow chart of torrential flood risk assessment model

空间中的搜索个体,粒子的最佳历史位置和种群的最佳历史定位有助于动态调整粒子的飞行速度。在洪水风险评价模型中,将流域内的风险单元视为粒子群,训练数据集的历史山洪灾害点和非山洪灾害点视为粒子最佳历史位置,PSO 算法可以充分利用群体智能的优势,综合考虑各风险单元之间相互作用,通过模拟多个粒子的运动,实现全局搜索,使得模型的结果更加准确。

SVM 的核心思想是对数据点进行有效划分,并确保分离超平面与其他数据点的距离尽可能大,惩罚函数 C 与核函数 σ 对其预测能力影响较大^[24-25]。因其能够处理输入和输出之间的非线性关系,且能够处理大量高维数据,SVM 可根据各种因素来估计某地区的洪水风险,如海拔、地形坡度、城市农业用地以及基础设施与资源分布等。基于 SVM 的山洪灾害风险评价模型的目标是识别出最优分离超平面,将训练数据集划分为洪水点与非洪水点两类。

尽管 SVM 和 PSO 算法均为高度结构化且能够解决多维问题,但 PSO-SVM 混合算法采用 PSO 算法对 SVM 中的惩罚函数 C 与径向基核函数 σ 进行全局寻优,能够克服 SVM 参数选择敏感等问题,显著提高模型的拟合效果^[26-27]。

在 PSO-SVM 混合中,首先随机生成 m 个粒子

$\mathbf{x} = (x_1, x_2, \dots, x_m)$ 形成初始种群, $\mathbf{x}$ 为 N 维欧氏空间中的向量,每个粒子对应一个类别表示标识 y_i , $y_i \in \{-1, 1\}$, 组成训练集 (x_i, y_i) 。 x_i 为输入空间向量元素,即高程、距河道距离、河流功率指数、坡度、人口密度、坡度、地形湿度指数、土地利用和年平均降水量; y_i 为与 x_i 相关的输出值,即历史山洪灾害点和非山洪灾害点。在高维特征空间中建立线性回归函数:

$$\mathbf{y}(\mathbf{x}) = \mathbf{W}^T \Phi(\mathbf{x}) + b \quad (3)$$

式中: $\mathbf{W}$ 为超平面权重向量; $\Phi(\mathbf{x})$ 为样本数据点; b 为偏置项。

将一组粒子初始化为随机粒子,在搜索迭代过程中,粒子通过个体极值和种群极值调整粒子状态,寻找全局最优解,迭代计算公式为

$$\begin{cases} V_{i,n+1} = w V_{i,n} + c_1 r_1 (p_{i,n} - x_{i,n}) + c_2 r_2 (G_{i,n} - x_{i,n}) \\ x_{i,n+1} = x_{i,n} + V_{i,n} \end{cases} \quad (i = 1, 2, \dots, m) \quad (4)$$

式中: m 为粒子群中粒子的个数; n 重复次数; w 为权重; c_1 、 c_2 为学习因子,一般为 2; r_1 、 r_2 为介于 (0,1) 之间的随机数; $p_{i,n}$ 为第 i 个粒子的个体最优解; $G_{i,n}$ 为第 n 次重复循环中全局最优解; $V_{i,n+1}$ 为第 $n+1$ 次重复时粒子第 i 个数的修正速度; $x_{i,n+1}$ 为第 $n+1$ 次重复时粒子第 i 个数的位置。

计算每个粒子对应的适应度 $S(x)$:

$$S(x) = \min(\frac{y_j - \bar{y}_j}{y_j}) \quad (j = 1, 2, \cdots, J) \quad (5)$$

式中: y_j 为训练样本风险等级实际值; $\bar{y}_j$ 为训练样本的风险等级评价价值; J 为训练样本的总数。

PSO-SVM 混合算法具体的训练步骤为: ①模型输入。以训练数据样本点对应的 9 个山洪影响因子数据及其山洪风险等级(洪灾点量化为 1、非洪灾点量化为 0) 作为模型输入进行训练; ②初始化粒子群。设置粒子数、惯性重量、加速度系数和迭代次数, 给出起始粒子位置 $x_{i,0}$ 和该粒子的起始速度 $V_{i,0}$; ③适应度评估。利用单个粒子相关核函数建立 SVM 模型, 核函数选用径向基函数, 初设径向基函数系数 $\gamma = 0.1$ 、惩罚函数 $C = 5$, 使用适应度函数 $S(x)$ 评估当前粒子的适应度; ④模型验证。重复式(5)直至粒子满足最佳粒子适应度, 输出最优解, 得到最佳径向基函数系数 $\gamma = 0.2$ 、惩罚函数 $C = 13$, 并与验证集样本点的风险等级进行对比; ⑤模型计算。输入流域共计 1989237 个计算单元对应的山洪影响因子数据, 调用 PSO-SVM 混合算法模型得到各计算单元风险等级。

3 结果与分析

3.1 山洪影响因子多重共线性检验结果

表 2 为山洪影响因子的多重共线性检验结果, 可见, 洪水影响因子的方差膨胀因子均小于阈值 10, 公差均大于阈值 0.1。因此, 本文选取的山洪影响因子间不存在多重共线性, 可用于构建山洪风险评价模型。

表 2 山洪影响因子多重共线性检验结果

Table 2 Multi-collinearity test results of torrential flood impact factors		
山洪影响因子	公差	方差膨胀因子
高程	0.519	1.925
坡度	0.767	1.304
人口密度	0.773	1.293
距河道距离	0.677	1.478
地形湿度指数	0.782	1.279
坡向	0.997	1.003
年平均降水量	0.812	1.232
土地利用	0.760	1.316
河流功率指数	0.806	1.241

3.2 山洪影响因子权重量化结果

采用 LVQ 算法计算输入向量(山洪影响因子的网格数据)与原型向量(山洪影响因子的历史洪灾点网格数据)间的欧式距离, 根据各因子与山洪发生的相关性计算山洪影响因子权重, 得到距河道距离、坡度和高程是鹤盛溪流域山洪灾害发生的主要

影响因素, 权重分别为 0.89、0.671 和 0.573。鹤盛溪流域地形起伏较大, 发生短历时强降雨后, 下渗时间较短导致大量降雨短时间内积聚在地表, 且由于山坡陡峭进而引发流速相对较快的山洪, 水流的强冲击力和破坏力对下游地区造成严重威胁。地形湿度指数、坡向和河流功率指数反映流域的水分分布与汇流方向, 当降水量超过地表渗透排水能力和河道容量时, 水体无法有效下渗且顺利排入河道进而引发洪涝灾害, 此过程受地形湿度指数、坡向和河流功率指数影响, 权重分别为 0.421、0.317 和 0.112。人口密度、土地利用和年平均降水量对山洪灾害发生影响程度较低, 权重分别为 0.097、0.217 和 0.246。

3.3 山洪风险评价结果

基于山洪影响因子多重共线性检验及权重量化结果, 采用 SVM、PSO 算法与 PSO-SVM 混合算法分别构建鹤盛溪流域山洪风险评价模型, 对全流域计算单元进行风险评价, 利用 GIS 技术生成栅格数据图层, 绘制流域山洪风险分布图(图 4)。根据风险等级评价价值, 采用自然断点分级法^[28-29]将山洪风险划分为 5 个风险等级: 0 ~ 0.40 为极低风险、>0.40 ~ 0.65 为低风险、>0.65 ~ 0.75 为中风险、>0.75 ~ 0.85 为高风险、>0.85 ~ 1 为极高风险。

由图 4 可见, SVM 算法模型计算结果中鹤盛溪流域极低风险、低风险、中风险、高风险和极高风险的空间覆盖率分别为 23.79% (72.23 km²)、17.11% (51.95 km²)、15.11% (45.87 km²)、24.78% (75.23 km²) 和 19.21% (58.32 km²)。PSO 算法模型计算结果中鹤盛溪流域极低风险、低风险、中风险、高风险和极高风险的空间覆盖率分别为 21.89% (66.46 km²)、19.63% (59.60 km²)、19.98% (60.66 km²)、25.79% (78.30 km²) 和 12.71% (38.58 km²)。PSO-SVM 混合算法模型计算结果中鹤盛溪流域极低风险、低风险、中风险、高风险和极高风险的空间覆盖率分别为 25.36% (76.99 km²)、20.36% (61.81 km²)、18.71% (56.80 km²)、21.05% (63.91 km²) 和 14.52% (44.09 km²)。3 种算法模型结果均表明鹤盛溪流域山洪极高风险区主要分布在中下游坡度较缓、地势较平坦区域, 包括东皋村、鹤盛村、西源村, 极低风险区主要分布在上游山高坡陡、人口稀疏且受台风影响较小区域, 包括锦里村、上港村。PSO-SVM 混合算法模型计算结果中山洪风险分布图中高、极高风险区与文献[14]描述的台风洪灾受淹地区基本重合, 且与相关文献[14,30-31]描述受淹地区基本重合, 模型评价结果更合理。

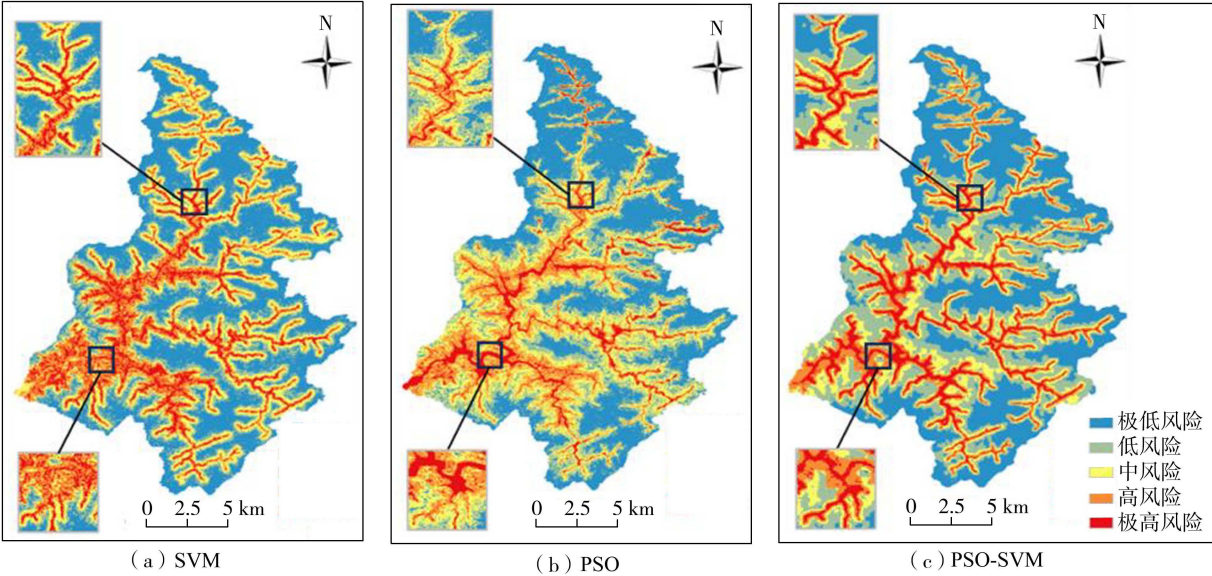


图 4 不同算法下鹤盛溪流域山洪风险分布

Fig.4 Distribution of torrential flood risk in the Heshengxi Watershed under different algorithms

3.4 混合算法模型与单一算法模型评价结果对比

采用准确度、灵敏度、特异性、F-score 值、Kappa 系数和受试者工作特征曲线对 3 种算法模型计算结果进行评估,表 3 为训练和验证数据集的各统计指标值。由表 3 可见,训练阶段 PSO-SVM 混合算法模型具有最高的准确度(0.932)、灵敏度(0.935)、特异性(0.922)、F-score 值(0.909)和 Kappa 系数(0.852)。结果表明该模型对 93.5%的洪灾单元和 92.2%的非洪灾单元进行分类,对93.2%的计算单元进行正确分类,对 90.9%的计算单元进行正确的风险等级评估。此外, Kappa 系数为 0.852 (>0.81),表明风险评价结果与实际分类结果几乎完全一致,达到了较高的一致性水平。由表 3 可见,验证集的评估指标表明模型性能与验证过程一致, PSO-SVM 混合算法模型具有最高的准确度(0.912)、灵敏度(0.931)、特异性(0.919)、F-score 值(0.896)和 Kappa 系数(0.833)。这说明 PSO-SVM 混合算法模型在山洪风险评价方面具有较好的性能,能够准确地分类洪水和非洪水单元,并对风险等级进行正确评估。

表 3 模型性能评估指标值

Table 3 Model performance evaluation index values						
数据集	算法模型	准确度	灵敏度	特异性	F-score 值	Kappa 系数
训练数据集	PSO	0.913	0.918	0.899	0.899	0.796
	SVM	0.891	0.899	0.893	0.891	0.759
	PSO-SVM	0.932	0.935	0.922	0.909	0.852
验证数据集	PSO	0.899	0.897	0.893	0.892	0.783
	SVM	0.883	0.892	0.879	0.863	0.832
	PSO-SVM	0.912	0.931	0.919	0.896	0.833

由图 5(a)可见,对于训练集,基于 PSO-SVM 混合算法模型的线下面积 (area under curve, AUC) 最高(0.953),高于 SVM 算法模型(0.898)和 PSO 算法模型(0.938)。由图 5(b)可见,对于验证集,基于 PSO-SVM 混合算法模型 AUC 值最高(0.934),高于 SVM 算法模型(0.876)和 PSO 算法模型(0.901)。从评估结果可知,3 种算法模型都具有较高的 AUC 值,但混合算法模型评价精度最高,这表明尽管

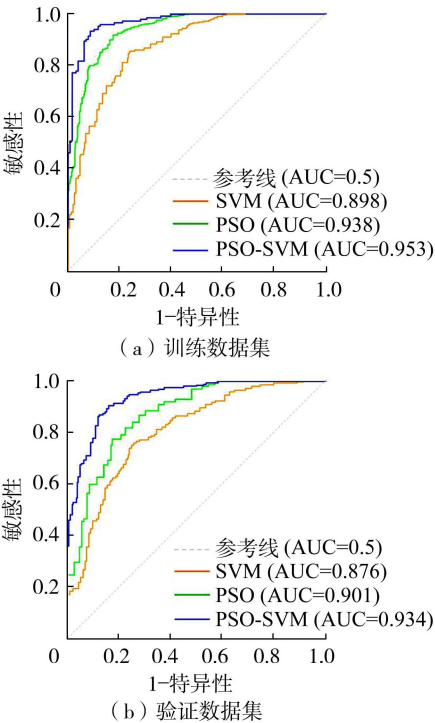


图 5 3 种算法模型受试者工作特征曲线

Fig.5 Models receiver operating characteristic curves of three algorithms

图 5 为 3 种算法模型的受试者工作特征曲线。

SVM 和 PSO 算法是高度结构化且能够解决多维问题,但采用 PSO 算法对 SVM 模型中的惩罚函数 C 与核函数 σ 进行全局寻优,能够克服 SVM 参数选择敏感等问题,显著提高模型的拟合效果。

4 结 论

a. 鹤盛溪流域山洪灾害受地形地貌、气候条件和人类活动等多种因素综合影响,山洪高风险区主要分布于流域中下游地区,确定山洪影响因子包括孕灾环境因子(距河道距离、坡度、高程和坡向)、致灾因子(年平均降水量、河流功率指数和地形湿度指数)和承灾体(人口密度和土地利用)。

b. 采用 LVQ 算法量化山洪影响因子对山洪灾害发生的影响程度,发现距河道距离、坡度和高程是鹤盛溪流域山洪灾害发生的主要影响因素,权重分别为 0.89、0.671 和 0.573。

c. 将 PSO 算法与 SVM 相结合,构建 PSO-SVM 混合算法山洪风险评价模型,其验证集受试者工作特征曲线指标、Kappa 系数和准确度分别为 0.934、0.833、0.912,能更加准确地反映山洪风险的空间分布特征。

参考文献:

[1] 王秀杰,王艳鹏,苑希民,等. 基于多重现期洪水情境模拟和综合风险度方法的入海河口地区洪水风险区划研究[J]. 自然灾害学报,2021,30(5):1-11. (WANG Xiujie,WANG Yanpeng,YUAN Ximin,et al. Research on flood risk zoning in the estuarine area based on multiple return periods flood situations simulation and comprehensive risk method [J]. Journal of Natural disasters,2021,30(5):1-11. (in Chinese))

[2] 赵刚,庞博,徐宗学,等. 中国山洪灾害危险性评价[J]. 水利学报,2016,47(9):1133-1142. (ZHAO Gang,PANG Bo,XU Zongxue,et al. Assessment on the hazard of flash flood disasters in China [J]. Journal of Hydraulic Engineering,2016,47(9):1133-1142. (in Chinese))

[3] PHAM T B,JAAFARI A,PHONG V T,et al. Improved flood susceptibility mapping using a best first decision tree integrated with ensemble learning techniques [J]. Geoscience Frontiers,2021,12(3):339-348.

[4] FABRIZIO F,DMITRI K,HUBERT H G S,et al. Catchment properties, function, and conceptual model representation: is there a correspondence [J]. Hydrological Processes,2014,28(4):2451-2467.

[5] MAHYAT S T,BISWAJEET P,MUSTAFA N J. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS[J]. Journal of Hydrology,2014,512:332-343.

[6] RAZAVITERMEH S V,SEO M,SADEGHINIARAKI A,et al. Application of genetic algorithm in optimization parallel ensemble-based machine learning algorithms to flood susceptibility mapping using radar satellite imagery [J]. The Science of the Total Environment,2023,873:162285-162285.

[7] 何昕宇,田文翀,张智宇,等. 基于数据驱动的洪涝风险评估方法研究进展[J]. 人民珠江,2022,43(5):60-67. (HE Xinyu,TIAN Wenchong,ZHANG Zhiyu,et al. Progress in flood risk assessment based on data-driven methods [J]. Pearl River,2022,43(5):60-67. (in Chinese))

[8] DIEU T B,BINH T P,QUOC P N,et al. Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of least-squares support vector machines and differential evolution optimization: a case study in Central Vietnam [J]. International Journal of Digital Earth,2016,9(11):1077-1097.

[9] BAHRAM C,AMIR M,ESMAIL H A,et al. Earth fissure hazard prediction using machine learning models [J]. Environmental Research,2019,179:108770.

[10] ARABAMERI A,SEYED D A,SANTOSH M,et al. Flood susceptibility mapping using meta-heuristic algorithms [J]. Geomatics,Natural Hazards and Risk,2022,13(1),949-974.

[11] HAMID R P,SEYED V R,NARGES K,et al. An assessment of metaheuristic approaches for flood assessment [J]. Journal of Hydrology,2020,582(C):124536.

[12] 赖成光,陈晓宏,赵仕威,等. 基于随机森林的洪灾风险评价模型及其应用[J]. 水利学报,2015,46(1):58-66. (LAI Chengguang,CHEN Xiaohong,ZHAO Shiwei,et al. A flood risk assessment model based on random forest and its application [J]. Journal of Hydraulic Engineering,2015,46(1):58-66. (in Chinese))

[13] NOROUZI H,BAZARGAN J. Calculation of water depth during flood in rivers using linear Muskingum method and particle swarm optimization (PSO) algorithm [J]. Water Resources Management,2022,36(11):4343-4361.

[14] 陈光跃,陈亚威. 2019 年“利奇马”台风期间温州暴雨洪水分析[J]. 河南水利与南水北调,2022,51(11):17-18. (CHEN Guangyue,CHEN Yawei. Analysis of rainstorm and flood in Wenzhou during Lichma Typhoon in 2019 [J]. Henan Water Resources and South-to-North,2022,51(11):17-18. (in Chinese))

[15] TANG X Z,TAKASHI M,LIU W,et al. A novel index to evaluate discretization methods: a case study of flood susceptibility assessment based on random forest [J]. Geoscience Frontiers,2021,12(6):319-331.

[16] LI Kaizhong,WU Shaohong,DAI Erfu,et al. Flood loss

- p>analysis and quantitative risk assessment in China [J]. Natural Hazards, 2012, 63(2): 737-760.
- [17] 张伟, 庄子孟, 孙慧超, 等. 我国城市内涝风险图编制关键问题和发展展望[J]. 水资源保护, 2023, 39(5): 58-68. (ZHANG Wei, ZHUANG Zimeng, SUN Huichao, et al. Key issues and development prospects for urban waterlogging risk maps in China [J]. Water Resources Protection, 2023, 39(5): 58-68. (in Chinese))
- [18] 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. (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. (in Chinese))
- [19] FULLER I C. Geomorphic impacts of a 100-year flood: Kiwitea Stream, Manawatu Catchment, New Zealand [J]. Geomorphology, 2007, 98(1): 84-95.
- [20] DAS S, GUPTA A. Multicriteria decision based geospatial mapping of flood susceptibility and temporal hydrogeomorphic changes in the Subarnarekha Basin, India [J]. Geoscience Frontiers, 2021, 12(5): 333-347.
- [21] MIND' J E R, LI L H, AMANAMBU A C, et al. Flood susceptibility modeling and hazard perception in Rwanda [J]. International Journal of Disaster Risk Reduction, 2018, 38: 101211.
- [22] RAVICHANDRAN J, KADEN M, VILLMANN T. Variants of recurrent learning vector quantization [J]. Neurocomputing, 2022, 502: 27-36.
- [23] SEYDI S T, KANANISADAT Y, HASANLOU M, et al. Comparison of machine learning algorithms for flood susceptibility mapping [J]. Remote Sensing, 2022, 15(1): 192-192.
- [24] 谷艳昌, 吴云星, 黄海兵, 等. 基于遗传算法优化支持向量机的大坝安全性能预测模型[J]. 河海大学学报(自然科学版), 2020, 48(5): 419-425. (GU Yanchang, WU Yunxing, HUANG Haibing, et al. Prediction model of dam safety behavior based on genetic algorithm optimized support vector machine [J]. Journal of Hohai University (Natural Sciences), 2020, 48(5): 419-425. (in Chinese))
- [25] 周婷, 汪炎, 邹俊, 等. 基于 PCA 和 SVM 的遥感影像水体提取方法及验证[J]. 水资源保护, 2023, 39(2): 180-189. (ZHOU Ting, WANG Yan, ZOU Jun, et al. PCA and SVM-based algorithm of water area extraction from remote sensing images and its verification [J]. Water Resources Protection, 2023, 39(2): 180-189. (in Chinese))
- [26] KUNT O G, RAO M, RAO S. Prediction of wave transmission over submerged reef of tandem breakwater using PSO-SVM and PSO-ANN techniques [J]. Journal of Hydraulic Engineering, 2020, 26(3): 283-290.
- [27] 张才溢, 傅蜀燕, 欧斌, 等. 基于 APSO 和 TWSVM 的特高拱坝变形预测模型[J]. 水利水电科技进展, 2023, 43(4): 46-51. (ZHANG Caiyi, FU Shuyan, OU Bin, et al. Deformation prediction model of ultra-high arch dams based on APSO and TWSVM [J]. Advances in Science and Technology of Water Resources, 2023, 43(4): 46-51. (in Chinese))
- [28] 王晟, 张珂, 晁丽君, 等. 基于集合模拟的汉江上游洪水与滑坡灾害风险评估[J]. 水资源保护, 2023, 39(6): 70-76. (WANG Sheng, ZHANG Ke, CHAO Lijun, et al. Assessment of flood and landslide hazard risk based on ensemble simulation over upstream of the Han River [J]. Water Resources Protection, 2023, 39(6): 70-76. (in Chinese))
- [29] 徐奎, 潘昊, 宾零陵, 等. 基于强化学习的海南岛山洪灾害易发性评估[J]. 水资源保护, 2023, 39(2): 95-100. (XU Kui, PAN Hao, BIN Linglin, et al. Research on flash flood susceptibility in Hainan Island based on reinforcement learning [J]. Water Resources Protection, 2023, 39(2): 95-100. (in Chinese))
- [30] 陈佩燕, 陆逸. 浙江超强台风“利奇马”灾害事件回顾及部门调查情况介绍[J]. 中国减灾, 2022(11): 36-39. (CHEN Peiyan, LU Yi. Zhejiang super Typhoon Lichma disaster event review and department investigation introduction [J]. Disaster Reduction in China, 2022(11): 36-39 (in Chinese))
- [31] 刘松波. 2020 年 4 号台风“黑格比”过程分析[J]. 黑龙江科学, 2021, 12(20): 36-37. (LIU Songbo. Analysis of the process of No. 4 Typhoon “Hagupit” in 2020 [J]. Heilongjiang Science, 2021, 12(20): 36-37. (in Chinese))

(收稿日期: 2023-04-29 编辑: 王芳)

