

靳春玲, 郭芮, 贡力, 等. 基于 RF 特征优选的 SSA-SVM 的铁路桥涵洪灾风险评价方法研究[J]. 安全与环境学报, 2026, 26(6): 2432-2440.

JIN C L, GUO R, GONG L, et al. Flood risk assessment method for railway bridges and culverts based on SSA-SVM with random forest feature selection[J]. Journal of Safety and Environment, 2026, 26(6): 2432-2440.

基于 RF 特征优选的 SSA-SVM 的铁路桥涵洪灾风险评价方法研究^{*}

靳春玲¹, 郭芮¹, 贡力^{1,2}, 陆浩伟¹, 袁鹏飞¹

(1 兰州交通大学土木工程学院, 兰州 730070; 2 甘肃民族师范学院甘肃省引调水工程行业技术中心, 甘肃合作 747000)

摘 要:随着我国西部铁路线路的建设,运营铁路线路的洪灾风险问题也不断突出,急需减小洪灾风险且提高防洪减灾效率。针对河西走廊漫流区铁路桥涵洪灾,研究基于麻雀搜索算法(Sparrow Search Algorithm, SSA)优化支持向量机模型(Support Vector Machine, SVM)以开展洪灾风险评价。首先,针对致灾因子、孕灾环境、承灾体(桥涵),搜集11个山洪影响因素;其次,采用随机森林(Random Forest, RF)算法进行重要度排序,并选出7个主要的影响因素作为模型特征值;再次,运用SSA算法优化SVM模型参数,建立基于麻雀搜索算法优化的支持向量机风险评价模型;最后,选取准确度、精确度、召回率和 F_1 -score共4个指标,将SSA-SVM模型与传统SVM模型、粒子群优化算法(Particle Swarm Optimization, PSO)优化SVM模型、鲸鱼优化算法(Whale Optimization Algorithm, WOA)优化SVM模型进行对比分析。研究结果显示:构建的SSA-SVM洪灾风险评价模型验证集的精确度、 F_1 -score分别为96.67%、96.91%,而且明显优于SVM、PSO-SVM、WOA-SVM模型,不仅能够实现铁路桥涵洪灾风险综合评价效能,还能为类似山洪风险评价提供决策支持。

关键词:安全工程;铁路桥涵;山洪灾害;风险评价;随机森林;麻雀搜索算法优化支持向量机模型

中图分类号:X43 **文献标志码:**A **文章编号:**1009-6094(2026)06-2432-09

DOI:10.13637/j.issn.1009-6094.2025.1125

0 引言

近年来,我国山洪灾害频发,山洪灾害点多面广、突发性强,严重危害人民群众生命财产安全^[1]。铁路运输承担着我国大部分的客运量和货运量,山洪灾害易引发路基损坏和桥梁垮塌的风险,尤其是西部地区铁路线路长、地质条件复杂,极端的气候条件与复杂的地理环境,导致突发洪水对铁路行车安全构成极大威胁。

目前,国内外针对山洪灾害的研究方法主要有:统计模型法、水文模型法和风险评价法。统计模型法^[2]需要以大量历史洪水数据为基础,通常用于特定历史数据充足地点的研究;水文模型法^[3-4]通过模拟洪水发生过程计算水位和流量,具有一套复杂的计算流程,且不考虑社会因素和人为因素,不适合进行综合分析;风险评价法综合考虑气象、水文、地形及人类活动等要素获取山洪发生的风险概率,通过对高风险位置提高防范,降低灾害发生风险。针对铁路洪灾的风险评价研究,Mark等^[5]从孕灾环境

和致灾因子出发,利用地理信息系统(Geographic Information System, GIS)技术对法国铁路网的洪水风险分布进行预测。王洁等^[6]首次以铁路线路为承灾体,收集了朔黄铁路40个站点的12个洪灾影响因子数据,并利用层次分析法进行风险等级评估。随着人工智能的快速发展,借助人工智能进行洪灾风险等级划分也逐渐成为洪灾风险识别的主要手段。王如锴等^[7]构建了基于粒子群算法优化的支持向量机(Particle Swarm Optimization-Support Vector Machine, PSO-SVM)的山洪风险评价模型,对温州市鹤盛流域山洪风险的空间分布特征进行了分析。韩帅兵等^[8]开创性地利用随机森林模型对山区铁路进行暴雨洪灾风险评价。这些研究都存在一些不足,传统的洪灾风险评估法的权重确定具有一定的主观性,而利用机器学习算法对铁路线路进行洪灾风险等级评价的研究较少,而考虑桥涵泄洪能力的研究更少。

因此,为提高铁路桥涵洪灾风险评价精度,本文以铁路桥涵为承灾体,结合山洪灾害致灾因子和孕

^{*} 收稿日期: 2025-06-30

作者简介: 靳春玲,教授,硕士,从事铁路洪灾风险管理方面的研究,522719061@qq.com。

基金项目: 国家自然科学基金项目(72261024);甘肃省科技重大专项(23ZDFA002);甘肃省科技重点研发项目(24YFGA041);甘肃省产业支持计划项目(2025CYZC-019)

灾环境,采用随机森林(Random Forest, RF)算法探究铁路桥涵洪灾主要影响因素,构建基于麻雀搜索算法优化支持向量机(SSA-SVM)的山洪灾害风险评价模型,并对河西走廊漫流区开展铁路桥涵洪灾风险等级预测,以期根据风险等级提出针对性的防范措施,为河西走廊地区铁路的洪灾风险管理和防灾减灾规划提供参考。

1 研究区与数据

1.1 研究区概况

河西走廊漫流区位于中国西北地区,地处东经 94° — 104° ,北纬 36° — 41° 。气候条件十分恶劣,年降雨量为 200 mm 左右,但降雨分布极不均匀,主要集中在 6—8 月,以暴雨为主。河西走廊存在一种特殊的山前流域——漫流区。漫流区没有固定的河道,一旦发生山洪灾害,洪水流动路径不固定,呈遍地漫流态势。这种情况会对现有的铁路路基造成不同程度的冲刷破坏,严重时甚至会危及列车的行车安全,进而导致巨大的经济和财产损失^[9]。

1.2 数据来源

历史受灾数据是确定洪水发生及其影响因素的关键数据基础,也是洪灾风险评价模型训练及验证的基本数据。研究随机选取河西走廊漫流区兰新铁路的 200 座桥涵为样本数据集,通过现场调研(图 1),获得运营期铁路部门定期复核标记的桥涵历史最高水位等数据。将桥涵历史最高水位与桥涵设计标高比作为桥涵历史受灾数据,这也是桥涵洪灾风险等级划分的重要依据。



图 1 现场调研照片

Fig. 1 On-site research photos

建立合理的指标体系是进行洪灾风险评价的必要条件。史培军^[10]提出的区域灾害系统理论认为区域灾害系统的结构体系由致灾因子(H)、孕灾环境(E)和承灾体(S)共同构成。致灾因子是指导致灾害发生的触发因素,铁路洪灾是自然因素(降雨量

是洪灾形成的动力条件,暴雨天数是反映洪灾强度的重要指标)和社会因素(人口密度增加导致人类活动增大,促使地质环境不稳定)相互作用导致的结果;孕灾环境是指能够影响致灾因子形成和发展的因素,主要与山洪灾害下垫面环境的稳定性有关;承灾体是灾害直接威胁和影响的对象,桥涵作为铁路线路主要的防洪工程,也是洪灾直接影响的对象。因此,本文以桥涵为铁路洪灾的直接承灾体进行评价分析,参照前人研究成果和研究区的气象、水文和地形特征,遴选出年平均降水量(H_1)、沿线人口密度(H_2)、暴雨天数(H_3)、河网密度(E_1)、植被覆盖率(E_2)、汇水面积(E_3)、坡度(E_4)、使用年限(S_1)、1 h 暴雨下桥涵满孔临界降雨量(S_2)、输沙抗淤能力(S_3)、养护效果(S_4)共 11 个评价指标,构成铁路桥涵洪灾风险评价指标体系。

本文数据来源:(1)年均降雨量通过甘肃省气象局 2019—2024 年资料获取;(2)沿线人口密度来自 2023 年甘肃省统计年鉴;(3)暴雨天数来自国家青藏高原科学数据中心提供的中国区域极端降水与月降水指数数据集;(4)河网密度、植被覆盖、汇水面积、坡度来自甘肃省 1:5 000 DEM 数据;(5)使用年限、输沙抗淤能力经实地调研获得;(6)1 h 暴雨下满孔临界降雨量、50 a 一遇水位达成灾水位比值来自 2023 年河西走廊某线路桥涵泄洪能力评估报告。

2 研究方法

铁路桥涵洪灾风险评价模型建模流程见图 2。

(1)铁路桥涵数据收集与处理。对收集的铁路桥涵数据进行清洗和整合,删除重复值、补全缺失值、剔除异常值,以保证样本数据的准确性。

(2)评价指标选取。计算各评价指标间的方差膨胀因子(VIF)和容差值(TOL),排除指标间可能存在的多重共线性问题,并采用随机森林算法进行特征重要性评估,找到主要因素,去除其他次要因素。

(3)洪灾风险评价模型建立与评估。建立基于麻雀算法优化支持向量机(SSA-SVM)算法的铁路桥涵洪灾风险评价模型,采用精确度、召回率、 $F_{1-score}$ 和准确度 4 个统计学评价指标,进行评价模型的准确性评估比较。

2.1 多重共线性分析

如果两个或两个以上自变量之间存在相关关系,则称这种相关关系为多重共线性。由于预测模

型中的各自变量存在较强的相互关系,可能会出现严重的多重共线性,从而导致模型训练的稳定性降低,造成回归系数、截距系数的估计准确性下降^[11]。为防止各指标间存在严重的多重共线性问题,本文对各评价指标采用方差膨胀因子(VIF)进行多重共线性检验。

$$F_i = \frac{1}{1 - R_i^2} \quad (1)$$

式中 F_i 表示变量 i 的方差膨胀因子; R_i 为第 i 个评价指标作为自变量与其他自变量的回归拟合系数。

2.2 特征重要性评估

随机森林算法(RF)由 Breiman^[12] 首次提出,其核心机制融合了装袋算法与随机特征选择策略。该算法通过可重复的随机抽样生成多个训练子集,并在每棵决策树构建过程中引入节点的随机分裂机制,最终通过集成多棵决策树的预测结果,采用多数表决机制形成综合决策。

在特征重要性评估方面,该算法提供了两种主流方法,一种是平均不纯度减少量法(Mean Decrease Impurity, MDI),即通过计算每个特征在所有决策树中的平均不纯度减少值,评估该特征对模型性能的贡献;另一种是平均精度下降量法(Mean Decrease Accuracy, MDA),即通过比较模型在原始数据集和打乱特征顺序后的数据集上的预测精确度来评估特征的重要性。特别值得注意的是,在抽样过程中,约有三分之一的原始样本未被选中,构成具有验证功能的袋外数据集(Out of Data, OOB)。基于该数据集,可通过量化特征缺失对模型精度的影响程度,建立 MDA 特征重要性评估体系,该方法为特征选择提供了可靠的量化依据^[13]。

本文基于随机森林算法确定洪灾评价指标权重。设有 N 个铁路桥涵样本。具体而言,假设随机森林中存在 n 棵决策树,随机抽取 N 个 Bootstrap 样本数据集,特征 X_i 的重要性程度计算按以下步骤。

Step 1: 初始令 $n = 1$;

Step 2: 根据第 n 个 Bootstrap 样本数据集训练构建决策树 T_k , 并计算出第 n 个 OOB data 的准确率之差 A_{ni} ;

Step 3: 对第 n 个 OOB data 中的特征 X_i 进行随机扰动,并再次计算进行随机扰动后的分类准确率 A_n ;

Step 4: 令 $n = 2, 3, \dots, N$, 重复 Step 2 ~ 3;

Step 5: 计算特征 X_i 的重要性程度 V_i 。

$$V_i = \frac{1}{N} \sum_{n=1}^N (A_n - A_{ni}) \quad (2)$$

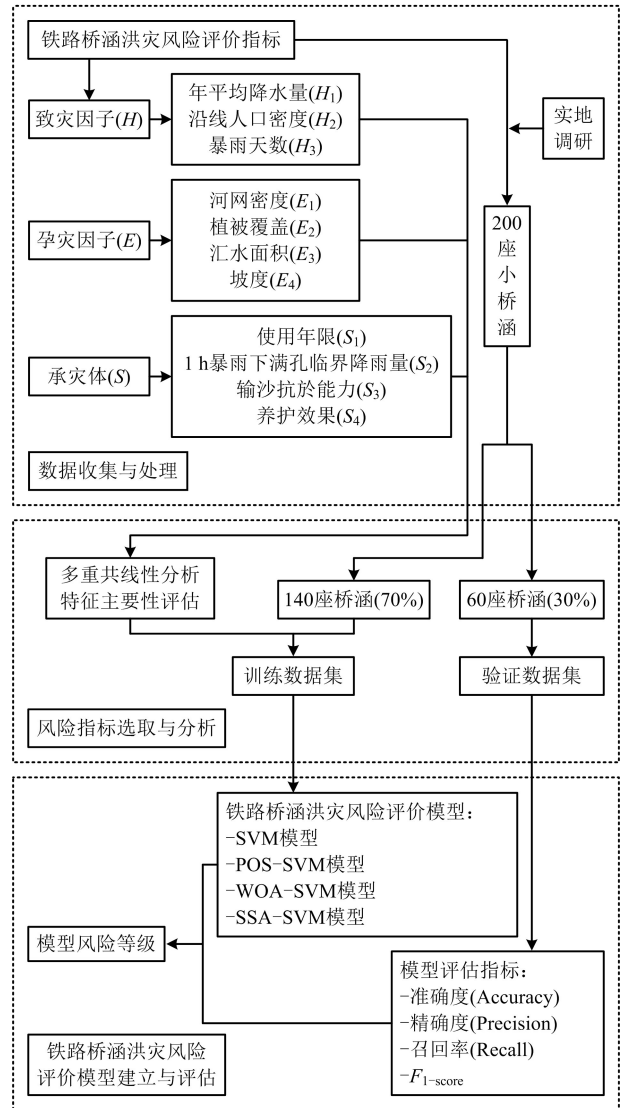


图2 铁路桥涵洪灾风险评价模型建模流程图

Fig. 2 Flowchart for modeling the flood risk evaluation model for railroad bridges and culverts

2.3 麻雀搜索算法优化支持向量机

支持向量机(SVM)是一种有监督的机器学习算法,最初用于解决二分类问题,后来经过扩展也可用于多分类和回归任务。SVM基于统计学习理论,旨在寻找一个最优的超平面来对不同类别的数据进行划分,具有坚实的理论基础和良好的泛化能力,可以将低维空间中复杂非线性问题映射转变为高维空间中的线性问题,并通过建立回归模型进行求解^[14]。其拟合函数为

$$f(x) = [\omega^T \varphi(x)] + b \quad (3)$$

式中 ω 为权重; $\varphi(x)$ 为非线性映射函数; b 为阈值。

基于结构风险最小化原则,可将支持向量机回归模型的目标函数转化为

$$\begin{cases} \min \frac{1}{2} \|\omega\|^2 + \frac{1}{2} C \sum_{i=1}^n \xi_i^2 \\ \text{s. t. } y_i(\omega^T x_i + b) \geq 1 - \xi_i \end{cases} \quad (4)$$

式中 C 为惩罚参数; ξ 为松弛变量。

通过构造拉格朗日函数,将原本问题转变为新的对偶问题,通过引入核函数,最终得出预测模型的决策函数。

$$f(x) = \sum_{i=1}^n \alpha_i K(x_i, x) + b \quad (5)$$

式中 n 为样本数; α_i 为拉格朗日乘子; $K(x_i, x)$ 为核函数。

麻雀搜索算法 (Sparrow Search Algorithm, SSA) 是一种新型的启发式优化算法,该算法具有寻优能力出色、收敛速度快的优势,可模拟麻雀群体的觅食和反捕食的行为。麻雀搜索算法的种群包含 3 种不同角色,分别是发现者、追随者和警戒者^[15]。

在这个算法模拟的生态场景中,发现者扮演着关键的引导角色,其主要任务是探寻并确定食物所在的大致范围以及具体方向,可为追随者提供觅食的指导。因此,发现者位置更新公式为

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \exp\left(\frac{-i}{\alpha t_{\max}}\right) & R_2 < S_i \\ X_{i,j}^t + QL' & R_2 \geq S_i \end{cases} \quad (6)$$

式中 $X_{i,j}^{t+1}$ 为经过 t 次迭代后维数为 j 的第 i 只麻雀的位置; α 为随机数, $\alpha \in (0, 1]$; $t_{\max}$ 为最大迭代次数; R_2 为预警值, $R_2 \in (0, 1]$; S_i 为安全阈值, $S_i \in (0.5, 1]$; Q 为随机数; L' 为元素全为 1 的 $1 \times d$ 矩阵。

追随者依靠发现者的成果来获取食物,会跟随发现者的指引,在发现者所确定的食物区域内进行捕食活动。追随者位置更新公式为

$$X_{i,j}^{t+1} = \begin{cases} Q \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{\alpha t_{\max}}\right) & i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L & i \leq \frac{n}{2} \end{cases} \quad (7)$$

式中 X_p^{t+1} 为目前发现者麻雀的最佳位置; X_{worst}^t 为全局最差位置; A^+ 为元素均为 ± 1 的 $1 \times d$ 矩阵。

警戒者承担着保障种群安全的重要职责。一旦群体中有麻雀察觉到周围存在捕食者威胁,警戒者会迅速发出警觉信号。当收到信号后,整个种群停止觅食行为,并开始转移。警戒者位置更新公式为

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta |X_{i,j}^t - X_{\text{best}}^t| & f_i > f_g \\ X_{i,j}^t + K \left[\frac{X_{i,j}^t - X_{\text{worst}}^t}{(f_i - f_w) + \varepsilon} \right] & f_i = f_g \end{cases} \quad (8)$$

式中 X_{best}^t 为目前麻雀所在的最佳位置; β 为符合正态分布的随机数; K 为飞行方向; ε 为不为零的常数; f_i 为适应度; f_g 为最佳适应度; f_w 为最差适应度。

支持向量机的性能在很大程度上依赖于核函数的选择和参数的设置,如惩罚参数 C 和核函数参数 g ,为提高 SVM 的性能,采用全局寻优算法 SSA 对 SVM 模型中的惩罚函数参数 C 与核函数参数 g 进行优化,以使模型的拟合效果显著提高^[16]。基于 SSA-SVM 的算法流程见图 3。

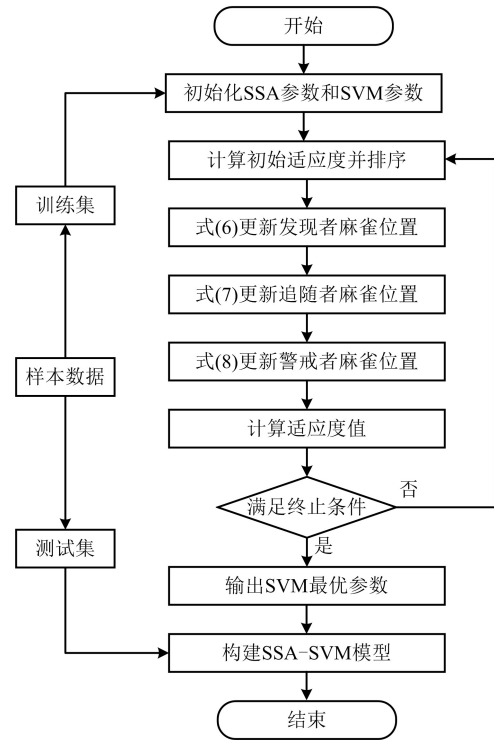


图 3 SSA-SVM 模型的算法流程图

Fig. 3 Algorithmic flowchart for SSA-SVM modeling

2.4 评价模型准确性评估

山洪灾害风险评价结果的准确性分析至关重要,混淆矩阵可将分类问题指标可视化,将评估类别与真实类别构造成列联表,并利用精确度 (P_{marco})、召回率 (R_{marco})、 $F_{1\text{-score}}$ 和准确度 (A_{cc}) 来度量不同模型进行洪灾风险等级评价的性能^[17]。准确度是模型正确分类样本占总样本的比例,可反映样本整体的正确率。精确度是衡量评价模型识别某类别的能力,而召回率是某类别被正确识别的能力,这两个指标之间一般存在此消彼长的关系,因此本文选取 $F_{1\text{-score}}$ 对二者进行平衡。研究属于多分类问题,精确度、召回率无法直接得出,在计算时须先对每一类统计指标值进行计算,然后对所有类别求算术平均值,进而得到宏平均值 (Macro-averaging)^[18]。

$$A_{cc} = \frac{\sum_{i=1}^n T_{Pi}}{N} \times 100\% \quad (9)$$

$$P_{marco} = \sum_{i=1}^n P_i = \sum_{i=1}^n \frac{T_{Pi}}{T_{Pi} + F_{Pi}} \times 100\% \quad (10)$$

$$R_{marco} = \sum_{i=1}^n R_i = \sum_{i=1}^n \frac{T_{Pi}}{T_{Pi} + F_{Ni}} \times 100\% \quad (11)$$

$$F_{1-score} = 2 \times \frac{P_{marco} R_{marco}}{P_{marco} + R_{marco}} \times 100\% \quad (12)$$

式中 n 为分类问题类别总数; N 为样本总数; T_{Pi} 为 i 类别被正确预测为 i 类别的样本数; T_{Ni} 为 i 类别被正确预测为其他类别的样本数; F_{Pi} 为其他类别被误判为 i 类别的样本数; F_{Ni} 为 i 类别被误判为其他类别的样本数。

3 结果分析

3.1 指标多重共线性检验

本文使用的 11 个铁路桥涵洪灾风险评价指标的多重共线性检验结果见表 1。由表 1 可知, 铁路洪灾风险评价指标的方差膨胀因子 (VIF) 均小于 5, 表明本文选取的 11 个铁路桥涵洪灾风险评价指标间不存在多重共线性, 可以用于构建铁路桥涵洪灾风险等级预测模型。

3.2 特征重要性评估

本文采用 RF 算法对数据进行特征重要性评估, 通过筛选主要影响的特征值, 以减少计算复杂度, 提高 SSA-SVM 模型的训练效果和准确率。利用 OOB 数据分别计算出 11 个评价指标特征的重要程度。为避免偶然性的发生, 进行 10 次 RF 重要度排序, 并取其重要度均值, 计算结果见图 4。

表 1 指标多重共线性检验结果

Table 1 Indicator multicollinearity test results

评价指标	VIF
H_1	1.197
H_2	1.053
H_3	1.238
E_1	1.209
E_2	1.030
E_3	1.089
E_4	1.066
S_1	1.927
S_2	1.232
S_3	1.152
S_4	1.477

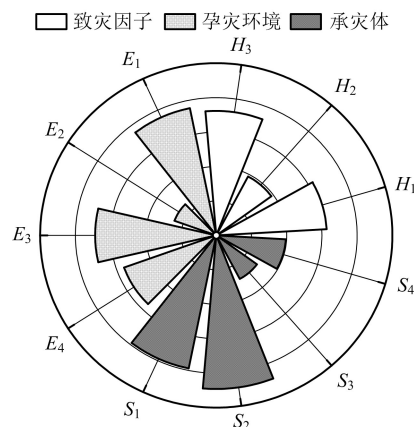


图 4 特征重要性分析

Fig. 4 Characteristic importance analysis

根据图 4 可知, 特征重要程度由大到小依次为: 1 h 暴雨下桥涵满孔临界降雨量 (S_2)、使用年限 (S_1)、河网密度 (E_1)、暴雨天数 (H_3)、汇水面积 (E_3)、年平均降水量 (H_1)、坡度 (E_4)、养护效果 (S_4)、沿线人口密度 (H_2)、输沙抗淤能力 (S_3)、植被覆盖率 (E_2)。以 0.5 为特征阈值, 筛选 1 h 暴雨下桥涵满孔临界降雨量 (S_2)、使用年限 (S_1)、河网密度 (E_1)、暴雨天数 (H_3)、汇水面积 (E_3)、年平均降水量 (H_1)、坡度 (E_4) 这 7 个特征作为机器学习模型的样本数据, 其余 4 个特征的重要程度较小可以忽略。

3.3 洪灾风险评价模型建立

(1) 输入样本。通过查阅 MZ/T 031—2012《自然灾害风险分级方法》、TB 10002—2017《铁路桥涵设计规范》等相关标准, 同时结合文献[19]将铁路桥涵洪灾风险划为 4 个等级: 历史最高水位比在 $[0, 0.25)$ 为低度风险 (I)、 $[0.25, 0.5)$ 为中度风险 (II)、 $[0.5, 0.75)$ 为高度风险 (III)、 $[0.75, 1)$ 为极高风险 (IV)。本文将样本数据集中的 140 座桥涵 (70%) 作为模型训练样本, 60 座桥涵 (30%) 作为模型验证样本。由于篇幅所限, 只列出部分验证集样本数据, 见表 2。

(2) 参数设计。对数据进行归一化处理, SVM 模型核函数选用径向基函数, 惩罚因子 C 的变化范围为 $[0, 100]$, 核函数参数 g 的变化范围为 $[0, 100]$ 。设置 SSA 算法的最大进化代数 50, 最大种群数量为 20。计算得到适应度曲线见图 5。由图 5 可知, 在迭代到 18 次后, 模型适应度平稳且达到最优, 最佳适应度为 0.057 14。经过 SSA 算法优化得到最优参数 $C=0.5659$, $g=5.1907$ 。

(3) 模型训练。图 6 为 SSA-SVM 模型的铁路

桥涵洪灾风险等级评价混淆矩阵。训练集[图 6(a)]总体准确度为 94.29%;验证集[图 6(b)]中低度风险(I)、中度风险(II)所有样本均被正确分类,准确度达到 100%,高度风险(III)、极高风险(IV)样本中仅有一个被误判,准确度为 94.7%,总体准确度为 96.67%。因此,利用 SSA 优化的 SVM 模型对铁路洪灾风险等级预测具有较高准确度,验证了其实用性以及准确性。

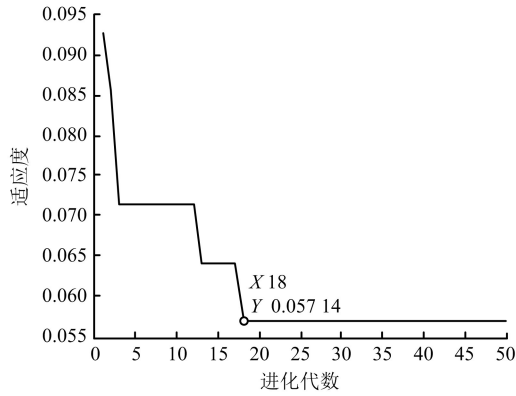


图 5 SSA-SVM 模型适应度曲线图

Fig. 5 SSA-SVM model fitness plot

3.4 洪灾风险评价结果对比分析

为了进一步验证 SSA-SVM 模型的有效性,本文引入传统 SVM 模型^[20]、PSO-SVM 模型^[21]、WOA-SVM 模型^[22]与 SSA-SVM 模型进行对比分析,且 3 个对比模型中种群数量与迭代次数也统一设置为 20 和 50。图 7 为不同模型下的预测结果。由图 7 可知,4 个模型对验证集 60 个漫流区铁路桥涵洪灾

风险等级预测准确度由大到小依次为:SSA-SVM 模型、WOA-SVM 模型、PSO-SVM 模型、SVM 模型。

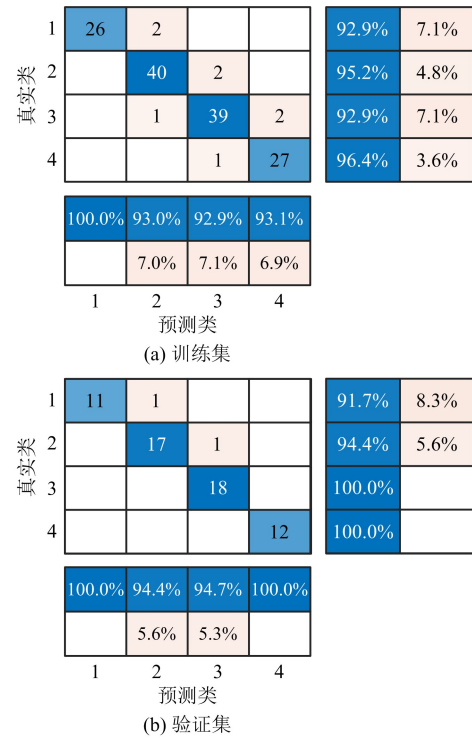


图 6 SSA-SVM 模型等级评价混淆矩阵

Fig. 6 Confusion matrix for SSA-SVM model level evaluation

由于试验目的是精准预测漫流区铁路桥涵洪灾风险等级,因此在所有评估指标中,识别准确率是最重要的指标之一, $F_{1-score}$ 是平衡精确度和召回率的指标, $F_{1-score}$ 越大,表明模型性能越好。表 3 为不同

表 2 实例样本数据

Table 2 Rockburst index data sheet

序号	H_1/mm	H_3/d	$E_1/(\text{m} \cdot \text{km}^{-2})$	E_3/km^2	$E_4/\%$	S_1/a	S_2/mm	风险等级
1	83.6	0.20	22.40	11.50	3.70	30	28.02	III
2	83.6	0.10	22.40	15.20	2.60	68	19.41	IV
3	85.6	0.22	22.40	28.50	5.20	68	42.50	III
4	85.3	0.23	116.52	1.03	3.70	55	34.75	II
5	76.9	0.05	116.52	42.50	2.76	10	72.65	I
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
56	85.3	0.03	116.52	19.62	1.00	13	36.78	II
57	152.4	0.04	22.40	0.18	20.15	68	42.33	II
58	178.5	1.84	26.20	0.88	4.80	64	55.28	I
59	178.5	2.28	26.20	0.29	4.90	54	33.69	IV
60	219.2	1.85	43.22	65.00	7.70	68	65.23	III

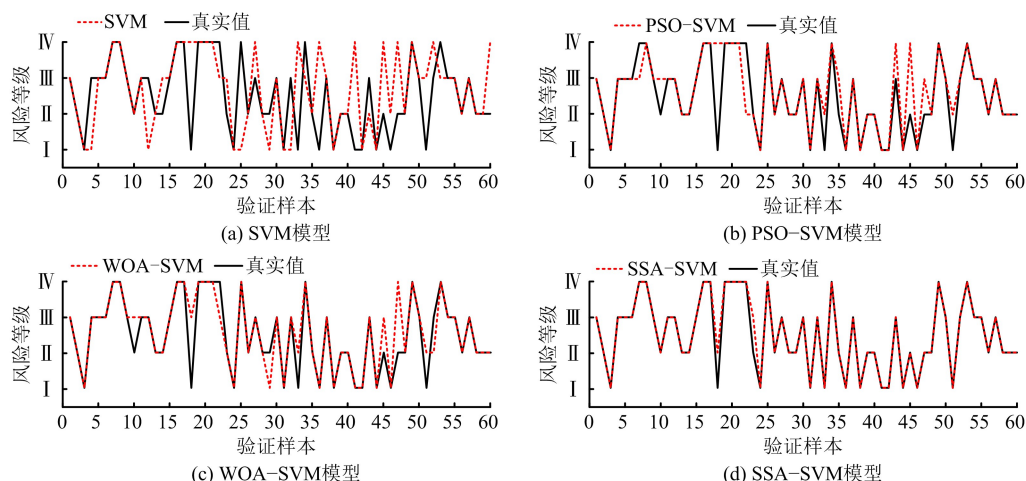


图7 不同模型预测结果对照图

Fig. 7 Comparison chart of prediction results from different models

模型的评估指标。由表3可以看出,传统的SVM模型随机生成 C 、 g 参数,因此各评估指标明显低于其他优化后的模型,其准确度和 $F_{1-score}$ 分别仅为66.67%和65.62%;PSO模型和WOA模型对参数 C 、 g 进行了优化,所有指标都显著提升,准确度分别达到了83.33%以及83.10%, $F_{1-score}$ 分别达到了86.67%和87.32%;而SSA-SVM模型进行风险等级评估的性能最好,所有评估指标都明显高于其他模型,识别准确度和 $F_{1-score}$ 分别为96.67%和96.91%。

表3 不同模型的评估指标对比

Table 3 Comparison of assessment indicators of different models

模型	准确度/%	精确度/%	召回率/%	$F_{1-score}/\%$
SVM	66.67	65.97	65.28	65.62
PSO-SVM	83.33	83.57	82.64	83.10
WOA-SVM	86.67	88.57	86.11	87.32
SSA-SVM	96.67	97.30	96.53	96.91

4 结论

(1)本文根据区域灾害理论搜集了11个河西走廊漫流区铁路桥涵洪灾影响因素,并利用RF算法对数据进行特征重要性评估,找出1h暴雨下桥涵满孔临界降雨量(S_2)、使用年限(S_1)、河网密度(E_1)、暴雨天数(H_3)、汇水面积(E_3)、年平均降水量(H_1)、坡度(E_4)这7个主要因素,去除其他次要因素,将主要因素输入模型中,提高了模型训练

效率。

(2)本文引入SSA算法优化SVM模型,得到最优参数 $C=0.5659$, $g=5.1907$ 。利用SSA-SVM模型将河西走廊漫流区铁路桥涵洪灾风险等级进行预测,总体准确度为96.67%,表明利用SSA优化的SVM模型对铁路洪灾风险等级预测具有较高的准确度,可为高风险桥涵的洪灾防护管理提供科学依据。

(3)将漫流区60座桥涵作为验证集代入SSA-SVM模型进行验证,得到准确度和 $F_{1-score}$ 分别是96.67%和96.91%,明显高于传统SVM模型、PSO-SVM模型和WOA-SVM模型,表明采用SSA-SVM模型进行河西走廊漫流区铁路桥涵洪灾风险等级预测不仅能够实现铁路桥涵洪灾风险综合评价效能,还能对类似山洪风险评价提供决策支持。

参考文献(References):

- [1] 方建,李梦婕,王静爱,等.全球暴雨洪水灾害风险评估与制图[J].自然灾害学报,2015,24(1):1-8.
FANG J, LI M J, WANG J A, et al. Global storm flood hazard risk assessment and mapping[J]. Journal of Natural Hazards, 2015, 24(1): 1-8.
- [2] 李磊,马梦格,折亚亚,等.复杂网络下雨洪灾害链风险分析及断链减灾研究[J].中国安全科学学报,2023,33(12):192-197.
LI L, MA M G, ZHE Y Y, et al. Risk analysis of rain-flood disaster chain in complex networks and disaster mitigation by breaking the chain[J]. China Journal of Safety Science, 2023, 33(12): 192-197.
- [3] DURGA RAO K H V, SHRAVYA A, VENKATESHWAR RAO V,

- et al. Flash flood disaster threat to indian rail bridges: a spatial simulation study of Machak river flood, Madhya Pradesh[J]. *Current Science*, 2017, 112(5): 1028 – 1028.
- [4] 张治中. 浅谈山前漫流区水文分析方法[J]. 铁道标准设计, 2006(11): 51 – 53.
ZHANG Z Z. Methods of hydrological analysis in mountain front diffuse flow area[J]. *Railway Standard Design*, 2006(11): 51 – 53.
- [5] MARK C, FRANCOIS C, LAURENT B. RISK VIP: evaluation of flood risk on the French railway network using an innovative GIS approach [J]. *E3S Web of Conferences*, 2016, 7: 10004.
- [6] 王洁, 张中杰, 曲晓黎, 等. 基于层次分析法的朔黄铁路暴雨灾害风险评估[J]. 气象科技, 2022, 50(6): 870 – 877.
WANG J, ZHANG Z J, QU X L, et al. Risk assessment of heavy rainfall disaster on Shuo Huang Railway based on hierarchical analysis [J]. *Meteorological Science and Technology*, 2022, 50(6): 870 – 877.
- [7] 王如锴, 练继建, 苑希民, 等. 基于 PSO-SVM 的鹤盛溪流域山洪风险评价[J]. 水资源保护, 2024, 40(2): 46 – 54.
WANG R K, LIAN J J, YUAN X M, et al. Evaluation of flash flood risk in Hesheng Creek watershed based on PSO-SVM[J]. *Water Resources Conservation*, 2024, 40(2): 46 – 54.
- [8] 韩帅兵, 苏成, 寇晓康, 等. 基于随机森林模型的山区铁路暴雨洪灾风险评估: 以朔黄铁路原平段为例[J]. 地理与地理信息科学, 2024, 40(2): 44 – 50.
HAN S B, SU C, KOU X K, et al. Risk assessment of heavy rainfall flooding on mountainous railroad lines based on random forest model: taking the Yuanping section of Shuo Huang Railway as an example[J]. *Geography and Geographic Information Science*, 2024, 40(2): 44 – 50.
- [9] 李盛, 秦军, 蔡磊, 等. 兰新铁路河西走廊漫流区桥涵水文沿线分布研究[J]. 中国铁道科学, 2022, 43(5): 30 – 41.
LI S, QIN J, CAI L, et al. Study on the distribution of bridge and culvert hydrology along the Hexi Corridor of Lanzhou-Xinjiang Railway [J]. *China Railway Science*, 2022, 43(5): 30 – 41.
- [10] 史培军. 三论灾害研究的理论与实践[J]. 自然灾害学报, 2002(3): 1 – 9.
SHI P J. Theory and practice of disaster research[J]. *Journal of Natural Hazards*, 2002(3): 1 – 9.
- [11] ALIREZA A, AMIR D S, SANTOSH M, et al. Flood susceptibility mapping using meta-heuristic algorithms [J]. *Geomatics, Natural Hazards and Risk*, 2022, 13(1): 949 – 974.
- [12] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5 – 32.
- [13] YUAN X, LIU S, FENG W, et al. Feature importance ranking of random forest-based end-to-end learning algorithm[J]. *Remote Sensing*, 2023, 15(21): 5203.
- [14] TIAN Y, LIU H, YANG Z, et al. A new background discrimination method using Support Vector Machine (SVM) with gaussian kernel in low-level 3H liquid scintillation measurement [J]. *Applied Radiation and Isotopes*, 2025, 220: 111754.
- [15] 何旭辉, 赵永帅, 蔡陈之. 基于 SSA-BP 神经网络的车-轨-桥系统随机振动分析[J]. 铁道科学与工程学报, 2024, 21(8): 3225 – 3236.
HE X H, ZHAO Y S, CAI C Z. Random vibration analysis of vehicle-rail-bridge system based on SSA-BP neural network [J]. *Journal of Railway Science and Engineering*, 2024, 21(8): 3225 – 3236.
- [16] 胡瑾秋, 张来斌, 李瑜环, 等. 基于支持向量机的油气生产复杂系统信息物理攻击识别方法[J]. 安全与环境学报, 2024, 24(8): 3053 – 3062.
HU J Q, ZHANG L B, LI Y H, et al. Cyber-physical attack identification method for complex system of oil and gas production based on support vector machine [J]. *Journal of Safety and Environment*, 2024, 24(8): 3053 – 3062.
- [17] 张振海, 李哲远, 鲍盛成, 等. 基于数字孪生的转辙机健康状态智能分析方法[J/OL]. 铁道科学与工程学报: 1 – 13 [2025 – 09 – 12]. <https://doi.org/10.19713/j.cnki.43-1423/u.T20242048>.
ZHANG Z H, LI Z Y, BAO S C, et al. Intelligent analysis method of rutting machine health state based on digital twin [J/OL]. *Journal of Railway Science and Engineering*: 1 – 13 [2025 – 09 – 12]. <https://doi.org/10.19713/j.cnki.43-1423/u.T20242048>.
- [18] WASHINGTON C, VITOR M, CHRISTIAN G, et al. On the cost-effectiveness of neural and non-neural approaches and representations for text classification: a comprehensive comparative study [J]. *Information Processing and Management*, 2021, 58(3): 102481.
- [19] 黄淇, 牟风云, 张用川, 等. 基于可拓云模型的重庆市中心城区公路洪灾风险评价[J]. 水土保持通报, 2022, 42(3): 157 – 165, 397.
HUANG Q, MOU F Y, ZHANG Y C, et al. Flood risk evaluation of highway in central Chongqing based on topological cloud model [J]. *Soil and Water Conservation Bulletin*, 2022, 42(3): 157 – 165, 397.
- [20] 吴瑕, 陈红环, 贾文龙, 等. 基于 SVM 的干线输气管道泄漏压降速率信号识别[J]. 中国安全科学学报,

- 2024, 34(6): 119–126.
- WU X, CHEN H H, JIA W L, et al. Signal identification of leakage pressure drop rate of trunk gas pipeline based on SVM [J]. China Safety Science Journal, 2024, 34(6): 119–126.
- [21] GEETHA K, MANU R, SUBBA R. Prediction of wave transmission over submerged reef of tandem breakwater using PSO–SVM and PSO–ANN techniques[J]. ISH Journal of Hydraulic Engineering, 2020, 26(3): 283–290.
- [22] 靳春玲, 姬照泰, 贡力, 等. 基于 WOA–SVM 的引水隧洞岩爆烈度评估模型[J]. 中国安全科学学报, 2023, 33(9): 41–48.
- JIN C L, JI Z T, GONG L, et al. Rock burst intensity evaluation model of diversion tunnel based on WOA–SVM[J]. China Safety Science Journal, 2023, 33(9): 41–48.

Flood risk assessment method for railway bridges and culverts based on SSA–SVM with random forest feature selection

JIN Chunling¹, GUO Rui¹, GONG Li^{1,2}, LU Haowei¹, YUAN Pengfei¹

(1 School of Civil Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;

2 Gansu Province Water Diversion and Transfer Engineering Industry Technology Center, Gansu Minzu Normal University, Hezuo 747000, Gansu, China)

Abstract: With the construction of railroad lines in western China, the issue of flood risk for operating railroads has become increasingly prominent. There is an urgent need to mitigate flood risks and enhance the efficiency of flood prevention and response strategies. This paper conducts a flood risk assessment for railroad bridges and culverts located in the diffuse flow area of the Hexi Corridor. The assessment employs the Sparrow Optimization Support Vector Machine Hybrid Algorithm (SSA–SVM). Initially, we identify 11 factors influencing flash flooding from three perspectives: disaster-causing factors, the disaster-containing environment, and the disaster-carrying infrastructure (bridges and culverts). Subsequently, we utilize the Random Forest (RF) algorithm to rank these factors by importance, selecting the seven most significant influencing factors as model features. The SSA algorithm was then employed to optimize the parameters of the Support Vector Machine (SVM) model, resulting in the establishment of a comprehensive risk assessment model based on the Sparrow Search Algorithm. To evaluate the performance of the SSA–SVM model, we compared it against traditional SVM, Particle Swarm Optimization (PSO) optimized SVM, and Whale Optimization Algorithm (WOA) optimized SVM models using four key indicators: Accuracy, Precision, Recall, and F_1 -score. The results indicate that the important features identified from the Random Forest model screening include critical rainfall (S_2), service life (S_1), river network density (E_1), rainstorm days (H_3), catchment area (E_3), annual average precipitation (H_1), and slope (E_4) within the 1-hour rainstorm model. The accuracy and F_1 -score of the SSA–SVM flood risk assessment model for the verification set are 96.67% and 96.91%, respectively. These figures represent improvements of 3 percentage points, 13.34 percentage points, and 10 percentage points over the combined algorithm models of SVM, PSO–SVM, and WOA–SVM, respectively, thereby confirming the efficacy of the proposed method. This study not only enhances the comprehensive assessment efficiency of flood risks for railway bridges and culverts but also provides valuable decision-making support for similar flash flood risk assessments.

Key words: safety engineering; railway bridges and culverts; flash floods; risk evaluation; random forests; the sparrow search algorithm optimized support vector machine model