

# 基于 WMA-SVM 的边坡稳定性预测模型

孙华芬<sup>1,2</sup>, 饶辉<sup>1,2</sup>, 侯克鹏<sup>1,2</sup>, 汪洪林<sup>1,2</sup>, 王泽奇<sup>1,2</sup>

(1. 昆明理工大学国土资源工程学院, 云南 昆明 650093;  
2. 云南省中-德蓝色矿山与特殊地下空间开发利用重点实验室, 云南 昆明 650093)

**摘要:** 为了提升数据驱动模型在边坡稳定性分类任务中的预测精度, 提出了一种融合新型鲸鱼迁徙优化算法 (whale migrating algorithm, WMA) 与支持向量机 (SVM) 的混合智能模型 (WMA-SVM)。首先, 构建了一个涵盖多种工程背景的异构边坡案例数据集。针对其显著的类别不平衡问题, 采用合成少数类过采样技术 (SMOTE) 与局部异常因子 (LOF) 算法相结合的策略, 生成了高质量平衡数据集。随后, 利用经 8 个基准测试函数验证的具有优越寻优性能的 WMA 算法, 对 SVM 的超参数进行自适应寻优。模型评估结果表明, WMA-SVM 在各项性能指标上均显著优于对比模型。此外, 基于置换特征重要性 (permutation feature importance, PFI) 算法分析得出, 容重、边坡角和内摩擦角是影响该数据集分类结果的关键敏感性特征。最后, 通过 8 组独立的工程案例测试, 进一步验证了模型的泛化能力, 结果显示, 预测与实际状态高度吻合。研究结果为边坡稳定性的智能分析提供了一种具备较好泛化潜力的建模框架。

**关键词:** 边坡稳定性; 机器学习; 鲸鱼迁徙算法; 支持向量机; 不平衡数据; 泛化能力  
**中图分类号:** TU470; TP18; O521.9 **文献标志码:** A

近年来, 随着我国基础建设的迅速发展, 在露天矿开采、交通建设工程、水利水电工程以及建筑等工程领域中会产生大量边坡, 这些边坡的稳定性在地质构造、岩土特性、水文及人类工程活动等多种因素作用下, 直接影响到了工程安全与可持续发展。边坡失稳为世界三大地质灾害之一, 其稳定性的分析与精准预测预报始终是边坡防治中的关键问题<sup>[1]</sup>, 构建科学、有效且精准的边坡稳定性预测模型, 对岩土工程的发展与边坡灾害防治而言, 意义重大<sup>[2]</sup>。

目前, 边坡稳定性分析方法主要可划分为传统和人工智能两大类<sup>[3]</sup>, 其中, 基于极限平衡法或有限元法的数值模拟是当前工程界公认的分析手段, 其成果已在 GeoStudio、FLAC 等商业软件中得到广泛应用与验证<sup>[4]</sup>, 尽管这类基于物理力学的方法在精确性上久经考验, 但它们通常面临建模复杂、计算量大、耗时较长等问题, 尤其在区域性初步风险筛查或大量参数敏感性分析时效率不高。为此, 近年来, 随着人工智能和计算机技术的进步, 机器学习方法为边坡稳定性预测提供了新的研究方向<sup>Error! Reference source not found.</sup>, 诸多学者已广泛应用各类机器学习模型, 如反向传播神经网络 (BPNN)<sup>Error! Reference source not found.</sup>、随机森林 (RF)<sup>Error! Reference source not found.</sup>和支持向量机 (SVM)<sup>Error! Reference source not found.</sup>等, 支持向量机基于核函数理论将样本映射到高维空间, 在处理小样本、非线性及高维数据方面表现出显著优势。吕鹏<sup>Error! Reference source not found.</sup>通过对 6 种人工智能方法在边坡稳定性预测中的效果: 发现 SVM 的预测精度在很大程度上依赖惩罚因子与核函数参数的选取, 参数的优化方法尚未形成统一标准。

为优化 SVM 参数, 诸多群体智能优化算法被引入, 如粒子群算法 (PSO)<sup>Error! Reference source not found.</sup>、遗传算法 (GA)<sup>Error! Reference source not found.</sup>等, 其中, 王团辉等<sup>Error! Reference source not found.</sup>提出了一种多策略改

收稿日期: 2025-10-27; 修回日期: 2026-01-22

基金项目: 云南省基础研究计划面上项目 (202001AT070205); 云南省“兴滇英才支持计划”青年人才专项 (CG25111F287A)

作者简介: 孙华芬 (1984—), 女, 博士, 副教授, 主要从事地质灾害研究. E-mail: 154221644@qq.com.

通信作者: 侯克鹏, 男, 博士, 教授, 博士生导师, 主要从事安全工程、采矿工程和岩土工程研究. E-mail: 1912864323@qq.com.

进的麻雀搜索算法 (MISSA) 优化 SVM 的边坡稳定性预测模型, 通过 9 组边坡工程实例验证了其良好性能。胡军等<sup>[1]</sup>则采用灰狼算法 (GWO) 优化 SVM 实现尾矿坝地下水位高精度预测。

然而, 传统算法 (如 GA、PSO) 易陷入局部最优且全局寻优能力受限, 为此, 本文引入鲸鱼迁徙优化算法 (Whale Migrating Algorithm, WMA)<sup>[13]</sup>, 该算法通过模拟鲸鱼群体迁徙行为, 结合领导者-跟随者动态与自适应迁移策略, 在增强全局搜索能力的同时提升局部开发效率; 其在寻优性能上的突出表现, 恰好能弥补传统 SVM 模型在参数寻优上的短板。

值得注意的是, 现实中的边坡工程地质条件极其复杂, 涉及非均质土体、多层土体、复杂坡面形态、地下水动态变化及地震等多种内外因素的耦合作用。将这些复杂因素完全参数化并构建预测模型是一项极具挑战性的任务。因此, 作为一项探索性研究, 本文聚焦于在相对理想化的均质边坡假设下, 验证 WMA-SVM 融合算法在边坡稳定性分类问题上的有效性和可行性, 旨在为数据驱动方法在该领域的应用提供一种新的算法思路。

基于此, 本文提出融合 WMA 与 SVM 的边坡稳定性预测模型 (WMA-SVM)。利用边坡案例数据, 采用混淆矩阵、ROC 曲线及 AUC 值系统评估模型性能。通过几种对比模型验证及特征重要性分析, 验证模型在分类精度、泛化能力和寻优性能上的表现, 以期为该领域的算法研究提供参考。

## 1 数据选取与相关性分析

### 1.1 边坡数据库建立

本文选取边坡高度  $H$ 、边坡角  $\beta$ 、容重  $\gamma$ 、内聚力  $c$ 、内摩擦角  $\varphi$  和孔隙水压系数  $r_u$  作为特征参数, 并根据文献<sup>[14]</sup>构建了包含 284 个边坡案例的数据库。为确保模型的泛化能力, 该数据库整合了源于矿山、公路及水库等多种工程背景的案例, 在地质构成上形成了一个以岩质边坡为主体, 同时涵盖土质与岩土混合边坡的异构数据集。这种多样性旨在将模型训练成一个普适性的快速评估工具, 而非针对特定地质类型的精细化模型。同时, 该数据库存在显著的类别不平衡问题, 其中稳定状态边坡案例 185 个, 而失稳状态边坡案例仅 99 个。为解决此问题并提升后续模型的预测精度, 本文采用了一种结合过采样与数据清洗的两阶段预处理策略。

首先, 采用合成少数类过采样技术 (SMOTE)<sup>[17]</sup>对少数类 (失稳状态) 样本进行扩充。通过该方法将 99 个失稳样本过采样至与多数类样本数量相等的 185 个, 形成一个包含 370 个案例 (185 个稳定, 185 个失稳) 的初步平衡数据集。随后, 考虑到 SMOTE 可能生成位于类别边界或远离原始样本分布的噪声点, 采用局部异常因子 (LOF)<sup>[18]</sup>对初步平衡后的数据集进行数据清洗。通过剔除被识别为异常点的样本, 最终获得一个包含 341 个高质量案例的最终数据集, 其中, 稳定状态案例为 166 个, 失稳状态案例为 175 个。该处理流程不仅有效缓解了类别不平衡问题, 还通过剔除潜在噪声提升了数据质量。最终数据集的部分数据示例如表 1 所示。

表 1 边坡案例数据库  
Table 1 Slope case database

Number	Slope parameters						Slope state
	$H$ (m)	$\beta$ (°)	$\gamma$ (kN·m <sup>-3</sup> )	$c$ (kPa)	$\varphi$ (°)	$r_u$ /kPa	
1	10	45	22.4	10	35	0.4	F
2	110	41	27.3	14	31	0.25	S
3	115	40	16	10	35	0.35	F
4	115	40	16	70.9	20	0.4	F
5	135	41	27.3	31.5	29.7	0.25	S
6	220	30	26	150	45	0.29	S
7	289	42	27	32	33	0.25	S

8	305	47	31.3	68.6	37	0.25	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
339	320	42.6	27	32	33	0.29	F
340	359	42	27	35	35	0.25	S
341	420	43	27	40	35	0.25	F

### 1.2 相关性分析

边坡失稳受边坡角、内聚力、内摩擦角等多因素协同作用，需通过相关性分析揭示边坡系统协调机制。基于皮尔逊相关系数，可视化如图 1 所示，其中红色和蓝色分别代表边坡的失稳和稳定状态，分析得出，所有变量间相关性均低于 0.6，表明特征间无强线性关联，存在较为复杂的非线性关系。



图 1 特征参数相关性分析矩阵图

Fig. 1 Correlation matrix of the characteristic parameters

## 2 WMA-SVM 预测模型构建

### 2.1 支持向量机分类原理

支持向量机 (Support Vector Machine, SVM) 的核心思想在于利用非线性映射手段，将原始输入空间的样本数据向高维特征空间转换，从而将输入空间中复杂的非线性决策边界问题转化为高维特征空间中的线性超平面构造问题 [Error! Reference source not found.](#)。其分类问题的数学模型如图 2 所示：

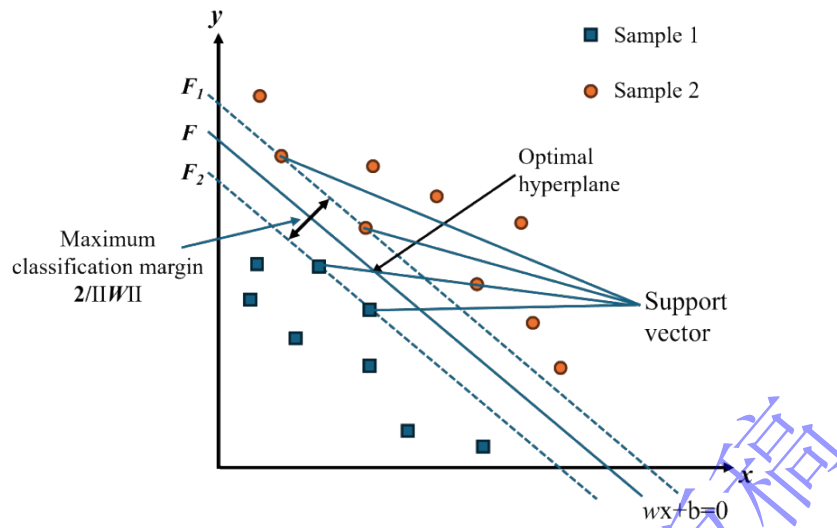


图 2 SVM 最优超平面示意图

Fig.2 Optimal hyperplane of the SVM

假定边坡的数据集为  $(x_i, y_i)$ ， $(i=1,2,\dots,n)$ ，其中， $x_i$  为边坡的第  $i$  个特征向量； $y_i \in \{-1,+1\}$  为类别标签，代表着边坡的稳定与失稳。

使距离最近的支持向量之间分类间隔最大化。

$$\min_{w,b,\xi} \left( \frac{1}{2} \|w\|^2 + C \sum_{n=1}^N \xi_n \right) \quad (1)$$

约束条件:

$$y_n (w^T x_n + b) \geq 1 - \xi_n, \xi_n \geq 0 (n = 1, 2, \dots, N). \quad (2)$$

式中： $w$  为权重向量 (weight vector)，决定了分离超平面的方向； $b$  为偏置项 (bias)，决定了超平面与原点的距离； $C$  为惩罚参数， $C > 0$ ； $\xi_n$  为松弛变量，通过拉格朗日乘子法求解式(2)得到决策超平面和分类模型的函数表达式分别为

$$\sum_{n=1}^N \alpha_n^* y_n x_n^T x + b^* = 0 \quad (3)$$

$$f(x) = \text{sgn} \left( \sum_{n=1}^N \alpha_n^* y_n x_n^T x + b^* \right) \quad (4)$$

式中： $\alpha_n^* \geq 0$  对应拉格朗日乘子最优解，非零值对应支持向量；引入核函数  $K(x_n, x_m)$  取代内积  $x_n^T x_m$ ：

$$K(x_n, x_m) = \exp(-\beta \|x_n - x_m\|^2) \quad (5)$$

式中， $K(x_n, x_m)$  为径向基核函数， $\beta$  为高斯核宽度参数，且  $\beta > 0$ 。可得该模型的函数表达式为

$$f(x) = \text{sgn} \left( \sum_{n=1}^N \alpha_n^* y_n K(x_n, x_m) + b^* \right) \quad (6)$$

高斯核函数参数  $\beta$  和惩罚参数  $C$ ，本文将采用 WMA 算法对二者开展全局寻优，构建 WMA-SVM 边坡稳定性预测模型。

## 2.2 WMA 算法原理

WMA (Whale Migration Algorithm) 是一种新型的元启发式优化算法<sup>[20]</sup>。其设计灵感来源于座头鲸在迁徙过程中所表现的协作与领导行为。该算法通过模拟鲸鱼群迁徙中“领导者-跟随者”机制，在优化搜索过程中既保持较快的收敛性能又避免陷入局部最优。其优化流程如下：

### (1) 种群初始化

首先，WMA 随机生成一个初始鲸鱼种群  $(W_1, W_2, \dots, W_i, \dots, W_{N_{pop}})$ ，其中， $N_{pop}$  表示种群规模，每个个体  $W_i = (w_{i,1}, w_{i,2}, \dots, w_{i,d}, \dots, w_{i,D})$ ，( $i=1, 2, \dots, N_{pop}$  和  $d=1, 2, \dots, D$ )，代表优化问题的一个潜在解。种群个体在搜索空间的初始位置由式 (7) 生成：

$$W_i = L + \text{rand}(1, D) \odot (U - L), i = 1, 2, \dots, N_{pop} \quad (7)$$

式中， $L$  和  $U$  分别为搜索空间的下界和上界向量，函数  $\text{rand}(1, D)$  是维度为  $D$  的随机向量，运算符  $\odot$  表示向量的 Hadamard 积；

### (2) 鲸鱼当前位置的表示

在算法中，设定前  $N_L$  个适应度较优的个体作为领导者，为表征鲸群的整体移动趋势，定义领导者位置的平均向量  $W_{Mean}$

$$W_{Mean} = \frac{1}{N_L} \sum_{j=1}^{N_L} W_j \quad (8)$$

该平均位置用于引导经验较少的个体移动，有助于提升收敛效率并维持种群多样性。

### (3) 幼鲸向邻近个体移动

将整个鲸群按适应度值降序排列：

$$W_1, \dots, W_{i-1}, W_i, W_{i+1}, \dots, W_{N_{pop}} \quad (9)$$

式 (9) 中， $W_1$  为最优成员， $W_{N_{pop}}$  为最差成员，经验较少的鲸鱼将向其邻近的更优个体方向移动。

### (4) 领导者引导机制

在迁徙过程中，若平均领导值位置  $W_{Mean}$  与当前最优位置  $W_{Best}$  的距离逐渐缩小，表明鲸群正朝有利区域移动，此时，经验不足的鲸鱼也开始沿  $D$  维区间  $[0, 1]$  随机向量方向向领导者靠拢，其位置更新公式如式 (10) 所示：

$$W_i^{new} = W_{Mean} + \text{rand}(1, D) \odot (W_{i-1} - W_i) + \text{rand}(1, D) \odot (W_{Best} - W_{Mean}), \quad (10)$$

$$i = N_L + 1, \dots, N_{pop}.$$

式 (10) 中，仅当  $f(W_i^{new}) < f(W_i)$  时，才替换原位置。

### (5) 领导者探索行为

经验丰富的领导者个体除引导群体外，还负责探索潜在更优区域，其位置更新策略如下：

$$W_i^{new} = W_i + r_1 \odot L + r_1 \odot r_2 \odot (U - L), i = 1, \dots, N_L, \quad (11)$$

式 (11) 中， $r_1$  和  $r_2$  为  $D$  的区间  $[0, 1]$  的随机数的向量， $L$  表示位置向量， $U-L$  是相对方向向量，仅当  $f(W_i^{new}) < f(W_i)$  时，更新位置可替换，且迭代结束进行排序后，选择所有  $N_L$  中最好的成员作为领导者。

WMA 算法优化过程如图 3 所示：

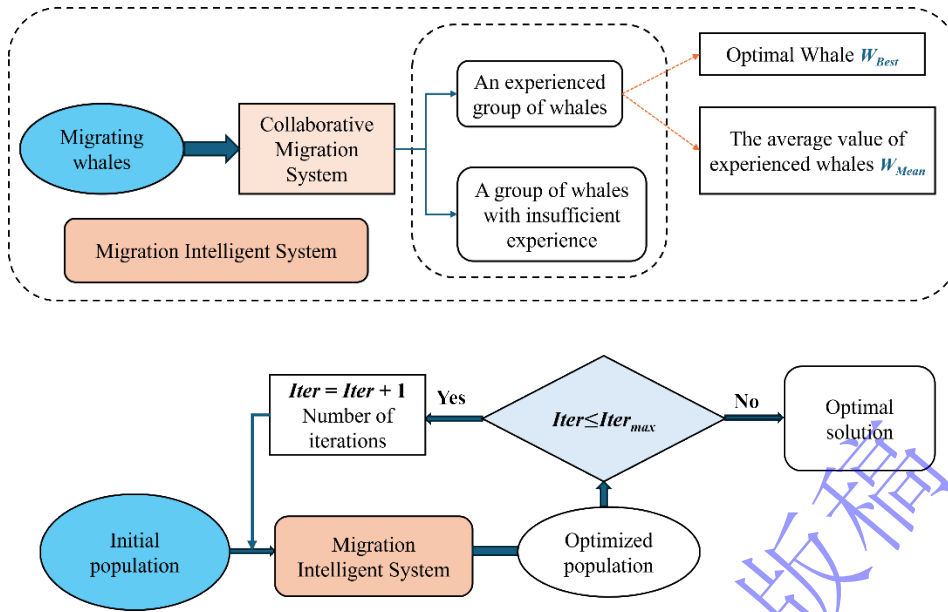


图3 合作迁徙系统和算法优化过程

Fig.3 The cooperative migration system and its optimization

2.2.1 WMA 优化算法性能对比分析

本研究选取了粒子群优化算法 (Particle Swarm Optimization, PSO)、蜣螂算法 (Dung Beetle Optimizer, DBO)、蝙蝠算法 (Bat Algorithm, BA) 等与 WMA 算法进行性能比较, 为公平对比, 各个算法的初始种群规模均设置为 50, 最大迭代次数设置为  $Z_{max} = 1000$ , 每组实验独立运行 30 次以消除随机性影响。

针对边坡稳定性预测这一复杂的非线性工程问题, 本文选取 CEC-2005 基准测试集中的 8 个代表性函数进行综合验证<sup>[21]</sup>。各函数具体信息见表 2。

表 2 相关测试函数信息

Table 2 Specifications of the related test functions

Function name	Type	Range	Dimension	Theoretical Optimum
$f_1$ : Shifted Sphere	Unimodal	[-100,100]	30	-450
$f_2$ : Shifted Schwefel's 1.2	Unimodal	[-100,100]	30	-450
$f_3$ : Shifted Rotated Elliptic	Unimodal	[-100,100]	30	-450
$f_4$ : Shifted Schwefel's 1.2 with Noise	Unimodal	[-100,100]	30	-450
$f_9$ : Shifted Rastrigin's	Multimodal	[-5.12,5.12]	30	-330
$f_{10}$ : Shifted Rotated Ackley's	Multimodal	[-32,32]	30	-140
$f_{11}$ : Shifted Rotated Griewank's	Multimodal	[-600,600]	30	-130
$f_{13}$ : Shifted Penalized Function 2	Multimodal	[-50,50]	30	-130

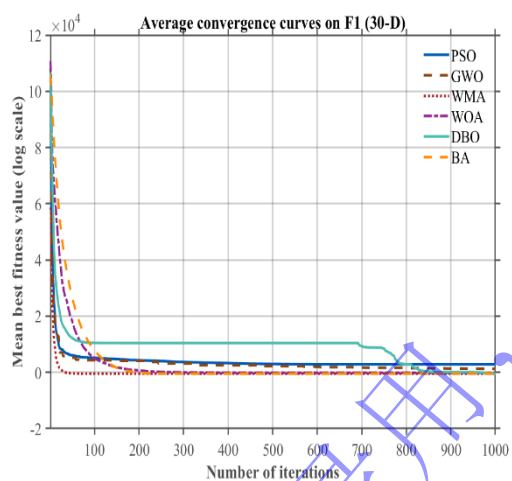
为全面考察算法的综合性能, 测试函数涵盖了两种类型, 单峰函数( $f_1 \sim f_4$ ): 用于检验算法在无局部最优干扰下的收敛速度与寻优精度; 多峰函数( $f_9 \sim f_{11}, f_{13}$ ): 用于测试算法的全局探索及勘探开发平衡能力。

结合两类函数的测试, 可从“探索”与“开发”两个关键维度系统性地评价算法性能, 这对于验证其解决复杂工程问题的适用性至关重要。算法性能的评估基于统计指标 (平均值 Mean 和标准差 Std, 见表 3) 和收敛曲线特性 (见图 4) 两个维度。平均值反映了算法的寻优精度, 标准差则反映了算法的鲁棒性。

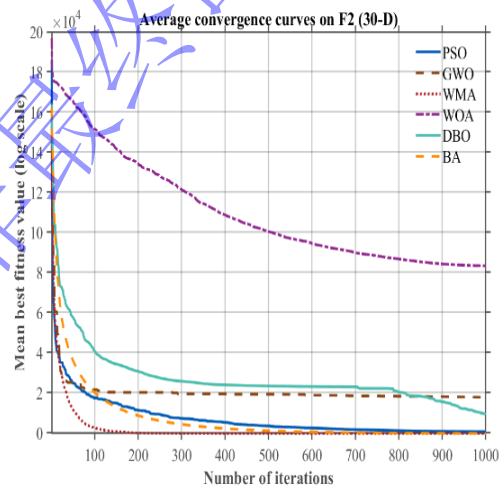
表 3 各算法性能参数对比

Table 3 Performance comparison of the algorithms

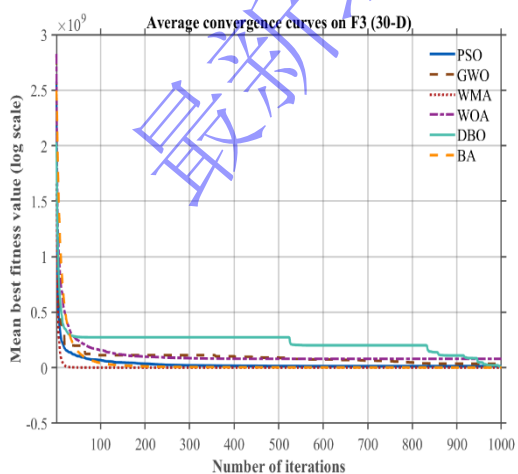
Function	Indicator	WMA	WOA	BA	GWO	DBO	PSO
$f_1$	Mean	-450	-202.68	-449.998	1.26E+03	-415.204	2.88E+03
	Std	5.99E-14	180.2488	2.79E-04	2.40E-13	15.1724	1.87E+03
$f_2$	Mean	-449.964	8.32E+04	-423.848	1.76E+04	9.08E+03	448.4377
	Std	0.0623	1.93E+04	44.1018	3.83E-12	8.35E+03	947.1713
$f_3$	Mean	-450	8.01E+07	1.07E+06	3.44E+07	1.08E+07	1.38E+07
	Std	5.99E-14	4.15E+07	7.13E+05	7.85E-09	6.50E+06	2.70E+07
$f_4$	Mean	4.02E+03	1.64E+05	1.38E+05	2.28E+04	2.25E+04	4.10E+03
	Std	1.21E+03	6.00E+04	3.71E+04	0	7.13E+03	2.16E+03
$f_9$	Mean	-258.493	-89.8118	-65.4021	-254.179	-167.004	-270.576
	Std	28.5664	59.8542	47.5989	0	28.3757	15.7052
$f_{10}$	Mean	-327.475	-311.632	-310.755	-324.633	-324.773	-327.277
	Std	1.2912	0.7667	0.2131	0	3.3295	4.4153
$f_{11}$	Mean	90.0511	91.9102	802.9144	92.8505	91.3767	133.5164
	Std	0.0572	0.4661	71.4287	1.50E-14	0.2318	45.7768
$f_{13}$	Mean	-129.535	-129.957	-58.94	-129.904	-129.979	-129.991
	Std	0.598	0.0481	5.5381	3.00E-14	0.0453	0.0146



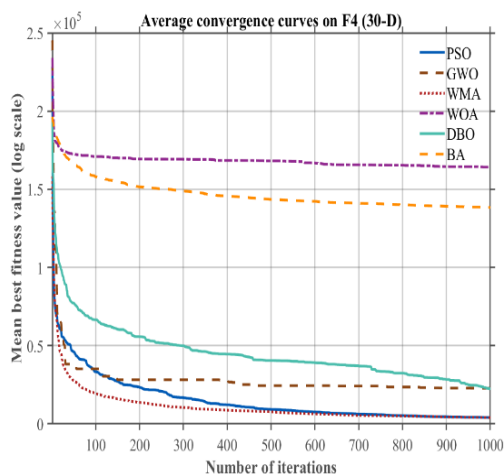
(a) Function  $f_1$



(b) Function  $f_2$



(c) Function  $f_3$



(d) Function  $f_4$

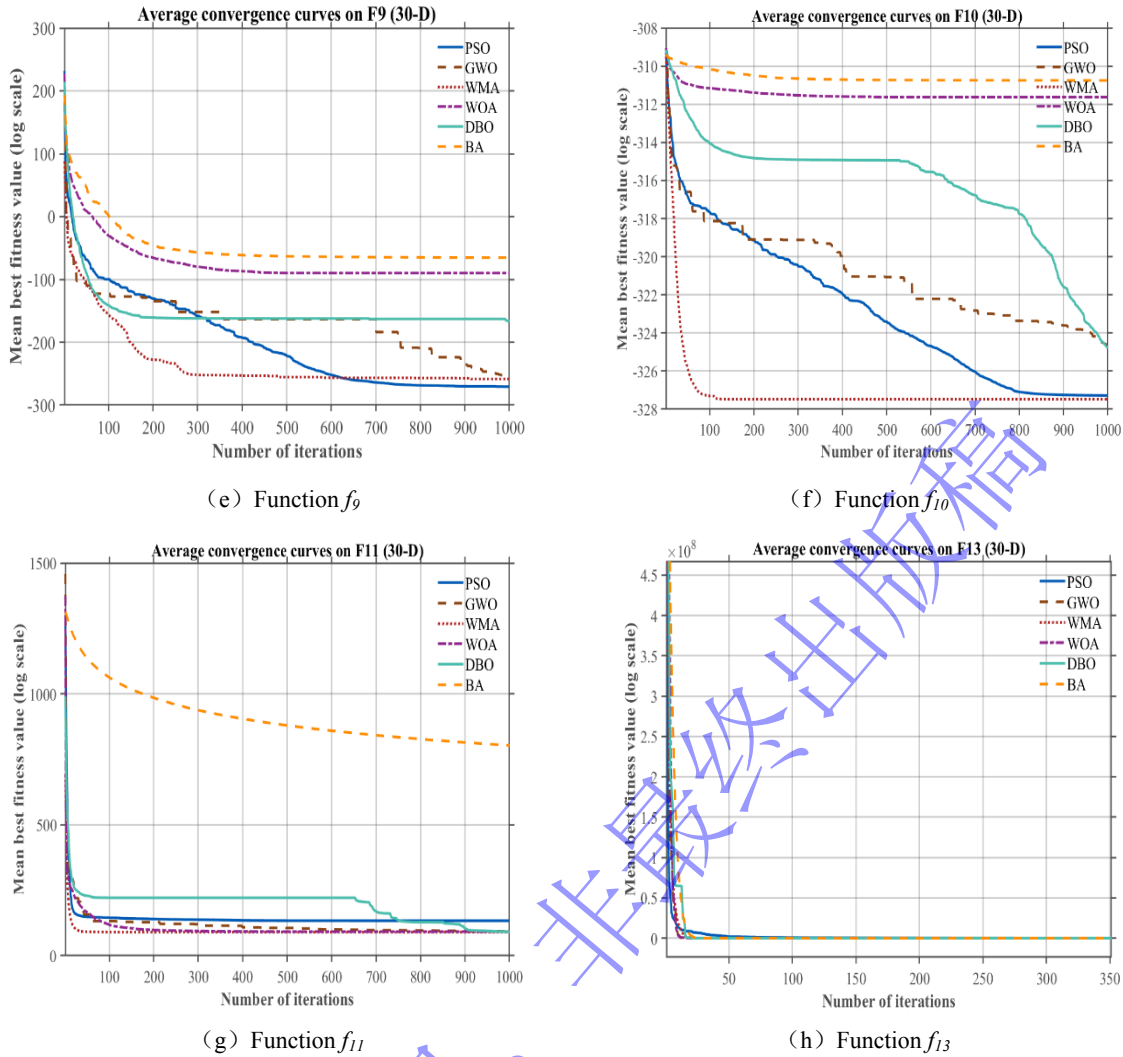


图4 算法收敛曲线

Fig.4 Convergence curve of the algorithm

平均值越小，表明算法在相应函数上的优化性能越优，标准差越小，则说明算法的稳定性越高，一致性越好。从表3中可以看出，WMA算法在绝大多数测试函数上表现出显著优势，在单峰函数( $f_1 \sim f_4$ )上表现出极强的局部开发能力，成功收敛至理论最优值；在多峰函数( $f_9 \sim f_{13}$ )上，WMA展现出优秀的全局探索能力，在 $f_{10}$ 和 $f_{11}$ 上取得了最佳寻优精度，并在其余函数上保持了较强的竞争力。

从图4所示的收敛曲线可以看出，WMA算法在所有测试函数的迭代初期均表现出最快的收敛速度，迅速逼近最优值；对于 $f_1$ ,  $f_3$ ,  $f_{10}$ , WMA的曲线呈持续下降趋势直至达到极高精度，未出现明显的早熟收敛现象；而在 $f_2$ 和 $f_4$ 和等复杂函数上，当其他算法陷入停滞时，WMA仍能保持寻优趋势。

综上，WMA算法兼顾了单峰函数的高精度开发能力和多峰函数的全局探索能力，这一特性与边坡稳定性预测模型超参数优化的需求高度契合。

### 2.2.2 基于WMA的SVM参数寻优策略

为提高支持向量机(SVM)模型的预测性能，本文将WMA算法的寻优机制与SVM的参数空间进行映射，对SVM模型中的惩罚因子 $C$ 和核函数参数 $g$ 进行优化；完整的寻优流程、参数设置及理论依据如下，(1)参数映射与搜索范围：WMA中每个个体的位置向量 $w_i = (w_{i,1}, w_{i,2})$ 被映射为一组SVM参数( $C, g$ )；综合考虑惩罚因子 $C$ 和核函数参数 $g$ 对模型的理论影响，同时借鉴现有智能优化算法优化SVM的广泛实践。例如，徐智超等将SVM模型 $C$ 和 $g$ 均设置为 $[0.1, 100]$ <sup>[22]</sup>，基于此，本研究设置SVM参数优化

范围分别为，惩罚因子  $C=[0.01,1000]$ ,核函数参数  $g=[0.0001,1000]$ 。该宽泛范围旨在增强算法的全局搜索能力，以寻得兼顾精度与泛化能力的最优参数组合。

(2) 适应度函数：在超参数寻优中引入适应度函数，公式为  $Fitness(W_i) = 1 - Accuracy$ ，即 WMA 算法的目标是找到适应度值最低时的个体，对应着分类准确率最高的参数组合。

(3) 寻优过程：WMA 算法设置初始种群为  $N=30$ ，最大迭代次数为 150。在迭代过程中，算法首先根据适应度值对种群排序，并将其分为“领导者”和“追随者”两部分：追随者种群向领导者平均位置和全局最优解方向移动，以加速收敛；领导者种群则通过增加随机扰动进行探索，以避免陷入局部最优。通过该双重更新机制，迭代结束后算法输出适应度值最低的个体坐标  $(C_{best}, g_{best})$  作为 SVM 模型的最终参数。

### 2.3 WMA-SVM 模型构建

WMA-SVM 预测模型具体构建流程如图 5 所示。

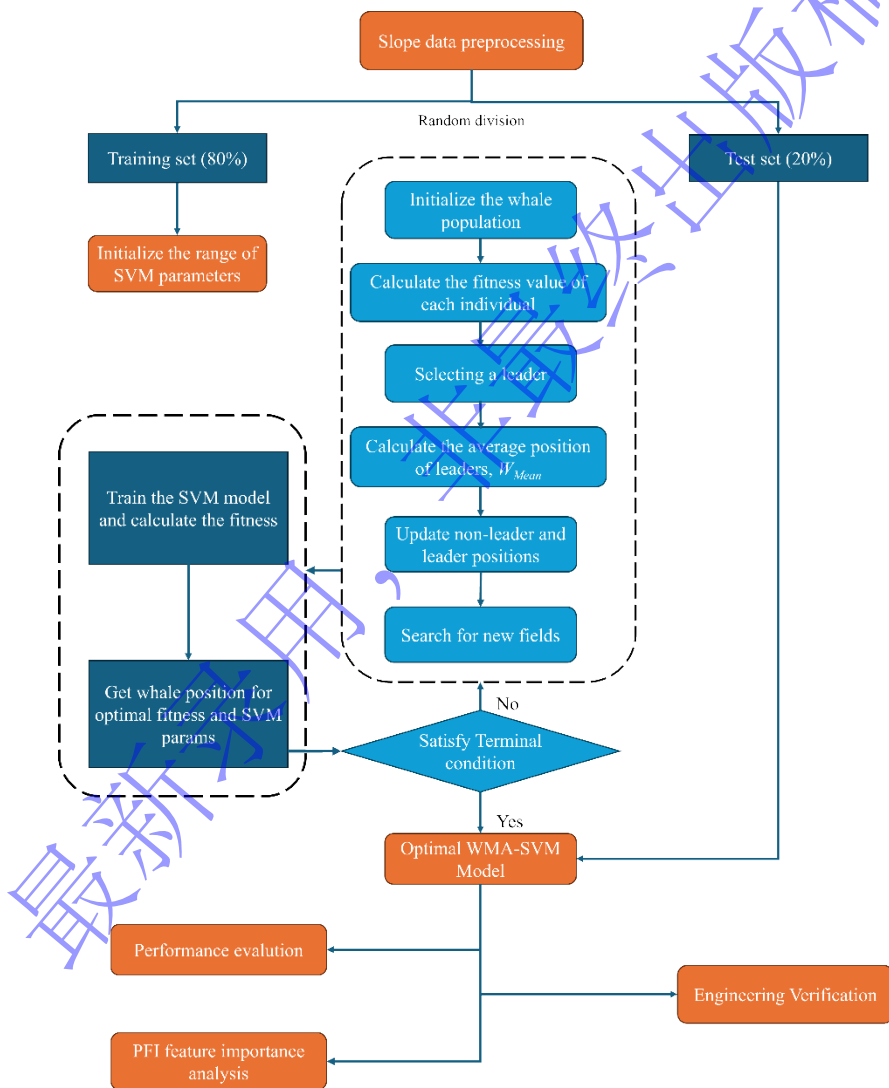


图 5 WMA-SVM 构建流程

Fig.5 Workflow of the WMA-SVM model

详细步骤如下：

- (1) 对所收集的边坡数据进行统计学分析和归一化预处理，随机划分为训练集（80%）和测试集（20%）；

- (2) 设定 SVM 超参数的取值范围（上下限）；
- (3) 对 WMA 算法进行参数初始化，设置鲸鱼种群数量；
- (4) 根据式（7）~（11）进行位置迭代更新；
- (5) 迭代终止后，得到最优超参数组合，基于此构建最优 WMA-SVM 模型，并利用测试集数据对模型精度进行验证；
- (6) 对 WMA-SVM 模型性能进行评估，并基于置换特征重要性算法对影响模型预测结果的关键参数进行敏感性分析；
- (7) 选取独立的工程案例对训练完成的模型进行泛化性能测试，以检验其在未见样本上的预测表现。

## 2.4 模型性能评估

为全面评估 WMA-SVM 模型的性能，选用了 5 个关键指标：准确率（ $A$ ）、精确率（ $P$ ）、召回率（ $R$ ）、F1 分数（ $F_{1-Score}$ ）以及受试者工作特征曲线下面积（AUC），这些指标的数值越高，表示模型的性能越优 Error! Reference source not found.。

各个指标对应公式如下：

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$R = \frac{TP}{TP + FN} \quad (14)$$

$$F_{1-Score} = \frac{2PR}{P + R} \quad (15)$$

$$TPR = \frac{TP}{TP + FN} \quad (16)$$

$$FPR = \frac{FP}{TN + FP} \quad (17)$$

式中，TN 表示实际失稳且预测失稳的样本数，TP 表示实际稳定且预测稳定的样本数，FN 表示实际稳定但预测失稳的样本数，FP 表示实际失稳但预测稳定的样本数。ROC 曲线以假阳性率（FPR）为横坐标、真阳性率（TPR）为纵坐标构建（二者取值范围均为[0,1]），AUC（曲线下面积）越接近 1，模型的预测性能越好。

## 3 结果预测与分析

### 3.1 模型训练结果分析

处理后的数据导入最优 WMA-SVM 模型并训练，其混淆矩阵可视化结果如图 6 所示。

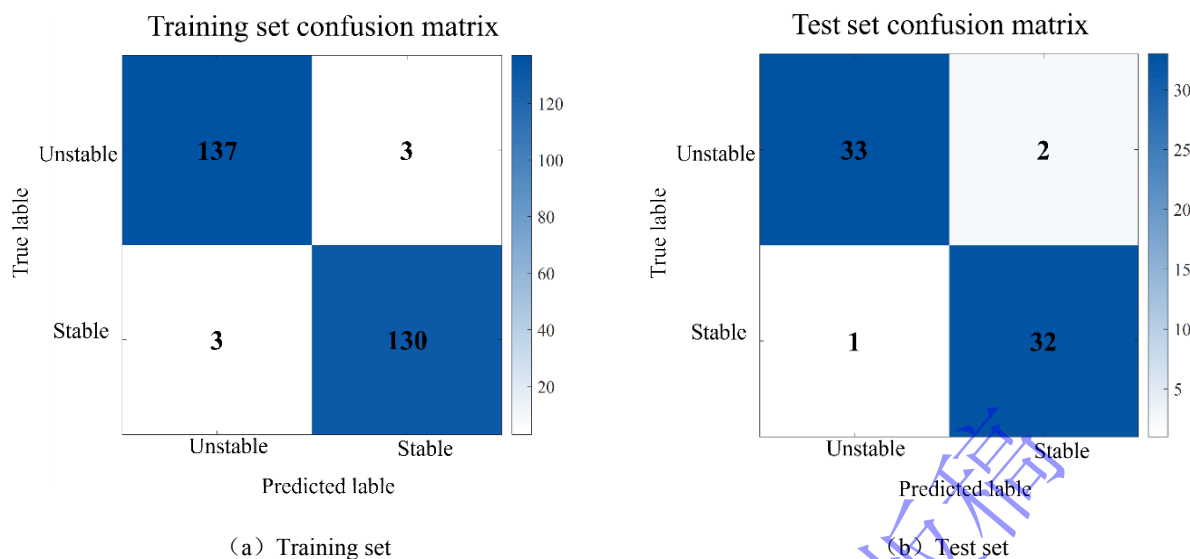


图 6 WMA-SVM 模型混淆矩阵

Fig.6 Confusion matrix of the WMA-SVM model

模型在训练集上的表现为：3 个稳定数据被误判为失稳，3 个失稳样本被预测为稳定；在测试集中，仅 1 个稳定数据被误判为失稳，2 个失稳数据被预测为稳定。结果表明，所构建的 WMA-SVM 模型在边坡稳定性预测方面具有较好的鲁棒性与可靠性。受试者工作特征 (ROC) 曲线如图 7 所示：

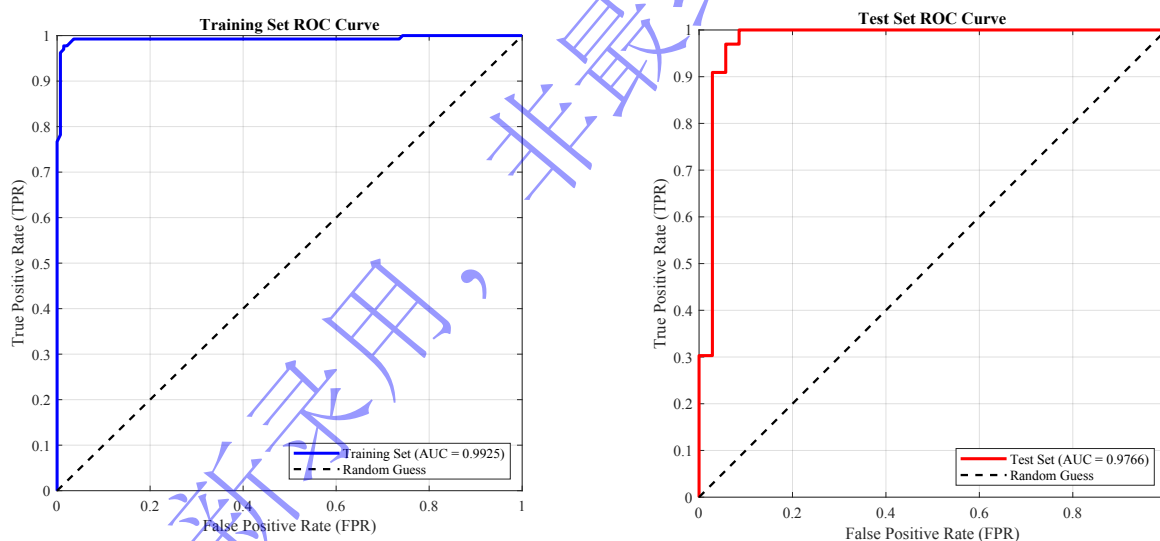


图 7 WMA-SVM 模型 ROC 曲线

Fig.7 ROC curves of the WMA-SVM model

由图 7 可得，WMA-SVM 模型在训练集和测试集中，AUC 值均大于 0.95，表明该模型具备卓越的泛化能力。

### 3.2 不同分类模型性能对比分析

为了验证 WMA 优化 SVM 超参数的可行性及其在边坡稳定性预测中的适用性，本文选取了 6 种经典算法对 SVM 进行超参数优化，采用了准确率、精确率、 $F_1$  分数及 AUC 值等多个指标对不同分类模型进行综合评价与分析。所有模型的训练集和测试集数据比例设置为 8:2，并进行了 30 次独立重复试验以减少随机性影响，不同模型的测试性能对比结果见表 4 所示。

表 4 不同分类模型的性能对比

Table 4 Performance comparison of different classification models

Model	Accuracy	Precision	R	F <sub>1</sub> -Score	AUC
WMA-SVM	0.9558	0.9559	0.9563	0.9559	0.9766
WOA-SVM	0.8971	0.8971	0.8974	0.8970	0.9310
SSA-SVM	0.9118	0.9137	0.9108	0.9115	0.9390
PSO-SVM	0.8676	0.8711	0.8662	0.8669	0.8860
GWO-SVM	0.8824	0.8930	0.8848	0.8819	0.9190
DBO-SVM	0.8382	0.8750	0.8429	0.8354	0.8690
BA-SVM	0.8529	0.7949	0.9394	0.8611	0.9230

从表 4 可知，WMA-SVM 模型在各项评估指标上，均优于其他模型，表明具有更优的分类预测性能，在边坡稳定性预测中具有更强的适用性和更高的可靠性。

### 3.3 基于 PFI 算法的 WMA-SVM 特征重要性分析

置换特征重要性 (Permutation Feature Importance, PFI) 是一种用于解释机器学习模型的方法，适用于深度学习模型。其核心思想是通过将输入特征进行排列 (置换)，以此来量化特征对模型预测性能贡献价值。

为了系统评估各边坡参数对 WMA-SVM 模型预测结果的影响程度，本研究采用 PFI 算法，并进行 100 次独立重复试验以确保结果的可靠性。图 8 展示了各特征在 100 次重复试验中的平均特征重要性值，其中横坐标为边坡特征参数，纵坐标为对应的特征重要性平均值；顶部的误差棒则表示得分的标准差，它反映了特征重要性在 100 次试验中的波动范围。通过结果分析各参数的平均重要性值及其波动范围 (误差棒)，将影响因素划分为三个层次。第一层次为核心敏感因素，包括重度 (0.4019)、边坡角 (0.3841) 和内摩擦角 (0.3729)，它们的重要性得分最高且彼此差异不具统计显著性。第二层次为次要影响因素，包括边坡高度 (0.3484) 和内聚力 (0.3073)。第三层次为孔隙水压力比 (0.2330)，其重要性明显偏低。

值得注意的是，本研究发现重度的重要性与边坡角及内摩擦角相当，这与传统基于极限平衡理论的认识存在一定差异。然而，该现象在基于数据驱动的边坡研究中并非孤例，例如，石峻峰等<sup>[24]</sup>与 QI 等<sup>[25]</sup>的研究均得出重度是影响稳定性的关键因素之一。其根本原因在于模型机理不同：机器学习模型不依赖预设物理公式，而是从数据本身的统计分布中学习模式。因此，本研究的结论科学地反映了当前数据集的内在统计规律。

综上，本研究的敏感性分析结果表明，在实际工程案例中，除了高度关注传统的剪切强度参数外，对岩土体自重 (重度) 的控制——例如对坡顶卸载以减小下滑力，同样是提升边坡稳定性的一项关键且直接有效的措施。

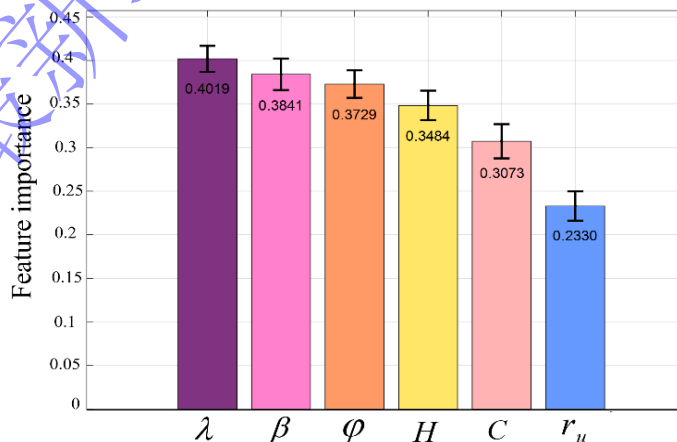


图 8 特征重要性分析图

Fig.8 Feature importance Analysis

### 4 工程实例验证

为进一步检验 WMA-SVM 模型的预测性能与泛化能力，本文选取文献 [Error! Reference source not found.](#) 中的苏家坪滑坡、三峡库区刘家屋场边坡等 8 组工程案例进行独立测试，其中，岩质边坡和土质边坡各四组；并选取 WOA-SVM，SVM 作为对比模型，工程实例详见表 5；

表 5 工程实例

Table 5 Engineering examples

Number	Project Name	Slope characteristic parameters						Slope type	Actual state of the slope
		$\gamma/(\text{kN}\cdot\text{m}^{-3})$	$c/(\text{kPa})$	$\phi/(\text{°})$	$\beta/(\text{°})$	$H/(\text{m})$	$r_u/\text{kPa}$		
1	Sujiaping Landslide	20	30	36	45	50	0.29	Soil slope	F
2	Jipaizi Landslide	27	35	35	42	359	0.29	Rock slope	F
3	Sichuan Kualiangzi Slope	21	10	30.34	30	30	0.29	Soil slope	S
4	Touzhai Slope in Yunnan	21.5	15	29	41.5	123.6	0.36	Rock slope	S
5	Zhongyan Village Landslide	12	0	30	45	8	0.29	Soil slope	F
6	Songshan Ancient Landslide	27	43	35	43	420	0.29	Rock slope	F
7	Huanglashi Landslide	22	0	40	33	8	0.30	Soil slope	S
8	Liu JiaWu Slope	27	32	33	42.2	239	0.29	Rock slope	S

将八组边坡案例数据作为外部测试集进行验证，各模型的预测混淆矩阵如图 9 所示，从图中可以看出，

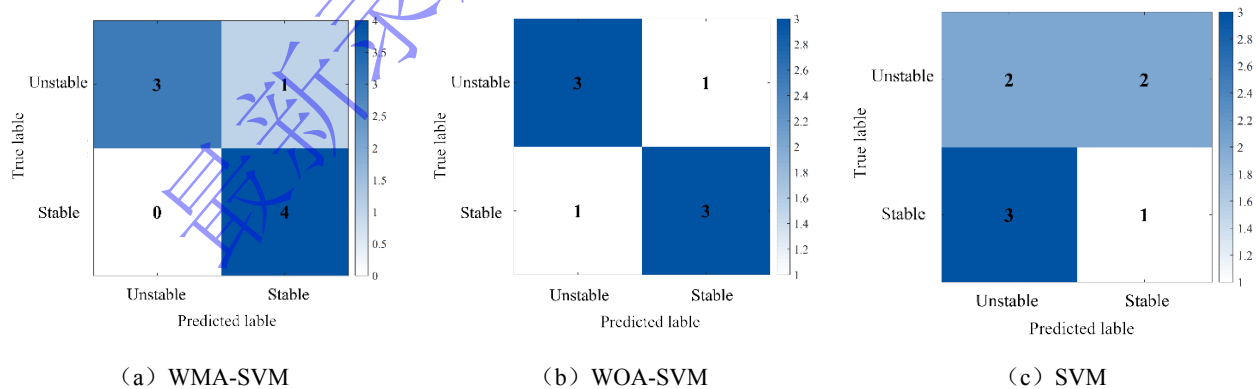


图 9 各模型混淆矩阵结果

Fig.9 Confusion matrices of the model

在边坡稳定性预测中，WMA-SVM 模型误分类了 1 个样本，WOA-SVM 模型误分类了 2 个样本，效果略差；而未优化的 SVM 模型误分类样本数最多，达 5 个。为深入分析 WMA-SVM 模型的性能，我们验证了其在不同边坡类型上的表现。结果显示，模型对与训练集特征差异较大的土质边坡实现 100% 的准确预

测, 唯一的误判案例发生在与训练集特征更相似的岩质高边坡中。这一结果清晰地表明, 模型具备良好的泛化能力, 其预测性能并未受建模样本与验证案例间的类型差异限制。

综上, WMA-SVM 模型在参数优化和特征相关性挖掘方面表现更强, 具有较高的预测精度, 为边坡稳定性预测提供了更可靠的参考。

## 5 结 论

(1) 本研究构建了一种融合鲸鱼迁徙算法(WMA)与支持向量机(SVM)的边坡稳定性混合预测模型。通过构建多特征边坡参数数据库并进行相关性分析, 发现特征参数之间存在较为复杂的非线性关系。

(2) 试验结果表明, 相较于 WOA-SVM、SSA-SVM 等六种对比模型, 本文提出的 WMA-SVM 模型在准确率(0.9558), AUC 值(0.9776)及  $F_1$ -Score(0.9559)等多项评估指标上均表现最佳。这证明了 WMA 算法在优化 SVM 超参数方面具有显著优势, 能有效提升模型在本数据集上的整体预测性能与泛化能力。

(3) 基于特征重要性分析得出, 在以本次数据集为基础的研究中, 重度、边坡角与内摩擦角是影响模型预测结果的核心敏感因素。这一发现揭示, 在数据驱动的稳定性的评估中, 除了传统力学参数外, 反映边坡几何与自重特性的相关参数同样重要。同时, 通过一个具体的工程实例验证, 模型展现了良好的应用潜力, 其预测结果可为类似地质条件下的工程实践提供有价值的参考。

(4) 本文提出的 WMA-SVM 模型并非旨在替代 GeoStudio 等数值模拟软件, 而是其高效补充。数值模拟的优势在于对特定边坡进行基于物理机理的分析; 而本模型作为一种数据驱动方法, 旨在实现区域性边坡的快速评估与风险评价。因此, 两者定位不同、优势互补, 本模型可为传统方法的重点详勘分析提供高效指引。

(5) 本研究的局限性在于, 模型构建基于相对理想化的均质边坡数据, 未能完全涵盖现实工程中岩土非均质性、复杂水文地质条件等关键因素。因此, 其普适性有待在更复杂的工程案例中进一步验证。未来研究将致力于融合多源数据, 以提升模型在复杂工况下的适用性。

## 参考文献

- [1] LIN Y, ZHOU K P, LI J L. Prediction of Slope Stability Using Four Supervised Learning Methods[J]. IEEE Access, 2018, 6: 31169-31179.
- [2] 张建涛, 刘志祥, 张双侠, 等. 基于 WOA-RF 的边坡稳定性预测模型[J]. 高压物理学报, 2024, 38(3): 194-205.  
ZHANG J T, LIU Z X, ZHANG S X, et al. Slope Stability Prediction Model Based on WOA-RF[J]. Chinese Journal of High Pressure Physics, 2024, 38(3): 194-205.
- [3] KURTZ W, LAPIN A, SCHILLING O S, et al. Integrating Hydrological Modelling, Data Assimilation and Cloud Computing for Real-Time Management of Water Resources[J]. Environmental Modelling & Software, 2017, 93: 418-435.
- [4] 杨杰, 马春辉, 程琳, 等. 高陡边坡变形及其对坝体安全稳定影响研究进展[J]. 岩土力学, 2019, 40(6): 2341-2353, 2368.  
YANG J, MA C H, CHENG L, et al. Research Progress on Deformation of High-Steep Slopes and Its Influence on Dam Safety and Stability[J]. Rock and Soil Mechanics, 2019, 40(6): 2341-2353, 2368.
- [5] 张化进, 吴顺川, 张中信, 等. 边坡稳定性自动机器学习预测方法研究[J]. 中国安全生产科学技术, 2023, 19(1): 35-40.  
ZHANG H J, WU S C, ZHANG Z X, et al. Research on Automatic Machine Learning Prediction Method for Slope Stability[J]. Journal of Safety Science and Technology, 2023, 19(1): 35-40.
- [6] GONG B. Study of PLSR-BP Model for Stability Assessment of Loess Slope Based on Particle Swarm Optimization[J]. Scientific Reports, 2021, 11(1): 1-10.
- [7] 汪伟, 崔欣超, 祁云, 等. 基于 SSA-RF 的采空区煤自燃温度回归分析模型[J]. 中国安全科学学报, 2023, 33(9).  
WANG W, CUI X C, QI Y, et al. Regression Analysis Model for Spontaneous Combustion Temperature of Coal in Goaf Based on SSA-RF[J]. China Safety Science Journal, 2023, 33(9).
- [8] 吕鹏. 基于信息技术的深凹露天矿高陡边坡稳定性综合分析研究[D]. 北京: 北京科技大学, 2019.

- LV P. Comprehensive Analysis of Stability of High-Steep Slopes in Deep Open-Pit Mines Based on Information Technology[D]. Beijing: University of Science and Technology Beijing, 2019.
- [9] 黄俊. 改进量子粒子群算法优化的 SVM 边坡变形预测研究[D]. 赣州: 江西理工大学, 2020.  
HUANG J. Research on Slope Deformation Prediction Using SVM Optimized by Improved Quantum Particle Swarm Algorithm[D]. Ganzhou: Jiangxi University of Science and Technology, 2020.
- [10] CAI Z L, XU W Y, MENG Y D, et al. Prediction of Landslide Displacement Based on GA-LSSVM with Multiple Factors[J]. Bulletin of Engineering Geology and the Environment, 2016, 75(2): 637-646.
- [11] 王团辉, 王超, 吴顺川, 等. 基于 MISSA-SVM 模型的边坡稳定性预测及应用[J]. 中国安全科学学报, 2024, 34(4): 135-144.  
WANG T H, WANG C, WU S C, et al. Prediction and Application of Slope Stability Based on MISSA-SVM Model[J]. China Safety Science Journal, 2024, 34(4): 135-144.
- [12] 胡军, 邱俊博. 多策略灰狼算法优化 SVM 的尾矿坝地下水位预测[J]. 矿冶工程, 2021, 41(3): 24-27.  
HU J, QIU J B. Groundwater Level Prediction of Tailings Dam Using SVM Optimized by Multi-Strategy Grey Wolf Algorithm[J]. Mining and Metallurgical Engineering, 2021, 41(3): 24-27.
- [13] GHASEMI M, DERICHE M, TROJOVSKÝ P, et al. An Efficient Bio-Inspired Algorithm Based on Humpback Whale Migration for Constrained Engineering Optimization[J]. Results in Engineering, 2025: 104215.
- [14] YAN X M, LI X B. Bayes Discriminant Analysis Method for Predicting the Stability of Open Pit Slope[C]// 2011 International Conference on Electric Technology and Civil Engineering (ICETCE). Lushan, 2011: 147-150.
- [15] WANG M, ZHAO G, WANG S. Hybrid Random Forest Models Optimized by Sparrow Search Algorithm (SSA) and Harris Hawk Optimization Algorithm (HHO) for Slope Stability Prediction[J]. Transportation Geotechnics, 2024, 48.
- [16] ZHOU J, LI E, YANG S, et al. Slope Stability Prediction for Circular Mode Failure Using Gradient Boosting Machine Approach Based on an Updated Database of Case Histories[J]. Safety Science, 2019, 118: 505-518.
- [17] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: Synthetic Minority Over-Sampling Technique[J]. Journal of Artificial Intelligence Research, 2002, 16: 321-357.
- [18] BREUNIG M M, KRIEGEL H P, NG R T, et al. LOF: Identifying Density-Based Local Outliers[C]// Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. 2000: 93-104.
- [19] 苏亮, 何海健. 基于 SVM 的 RC 框架结构地震易损性分析[J]. 华中科技大学学报(自然科学版), 2018, 46(5): 115-120.  
SU L, HE H J. Seismic Vulnerability Analysis of RC Frame Structures Based on SVM[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2018, 46(5): 115-120.
- [20] JIA H, WEN Q, WU D, et al. Modified Beluga Whale Optimization with Multi-Strategies for Solving Engineering Problems[J]. J. Comput. Des. Eng., 2023, 10(6): 2065-2093.
- [21] SUGANTHAN P N, HANSEN N, LIANG J J, et al. Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization[R]. Singapore: Nanyang Technological University, 2005.
- [22] 徐智超, 陈匀杉, 邓超, 等. 基于 PSO-SVM 模型的边坡稳定性预测研究[J]. 安全与环境学报, 2025, 25(9): 3531-3537.  
XU Z C, CHEN Y S, DENG C, et al. Slope Stability Prediction Research Based on PSO-SVM Model[J]. Journal of Safety and Environment, 2025, 25(9): 3531-3537.
- [23] 杨星雨, 王宁. 基于 SCA-GBDT 的边坡稳定性预测混合模型[J]. 黄金, 2025, 46(3): 59-65.  
YANG X Y, WANG N. Hybrid Model for Slope Stability Prediction Based on SCA-GBDT[J]. Gold, 2025, 46(3): 59-65.
- [24] 石峻峰, 周琳, 任宇联, 等. 基于 INGO-RF 的边坡稳定性预测模型[J]. 安全与环境学报, 2025, 25(4): 1380-1390.  
SHI J F, ZHOU L, REN Y L, et al. Slope Stability Prediction Model Based on INGO-RF[J]. Journal of Safety and Environment, 2025, 25(4): 1380-1390.
- [25] QI C C, TANG X L. Slope Stability Prediction Using Integrated Metaheuristic and Machine Learning Approaches: A Comparative Study[J]. Computers & Industrial Engineering, 2018, 118: 112-122.
- [26] 熊振涛. 基于机器学习的边坡稳定性预测模型优化研究[D]. 昆明: 昆明理工大学, 2023.  
XIONG Z T. Optimization Research on Slope Stability Prediction Model Based on Machine Learning[D]. Kunming: Kunming

## A WMA-SVM Model for Slope Stability Prediction

SUN Huafen<sup>1,2</sup>, RAO Hui<sup>1,2</sup>, HOU Kepeng<sup>1,2</sup>, WANG Honglin<sup>1,2</sup>, WANG Zeqi<sup>1,2</sup>

(1. Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650093, Yunnan, China;

2. Yunnan Key Laboratory of Sino-German Blue Mining and Special Underground Space Development and Utilization, Kunming 650093, Yunnan, China)

**Abstract:** To enhance the prediction accuracy of data-driven models in slope stability classification, this study proposes a hybrid intelligent model (WMA-SVM) that integrates a novel Whale Migration Algorithm (WMA) with a Support Vector Machine (SVM). First, a heterogeneous dataset of slope cases from diverse engineering backgrounds was constructed. To address its significant class imbalance, a combined strategy using the Synthetic Minority Over-sampling Technique (SMOTE) and the Local Outlier Factor (LOF) algorithm was adopted to generate a high-quality balanced dataset. Subsequently, the WMA algorithm, which demonstrated superior optimization performance on eight benchmark test functions, was employed to optimize the hyperparameters of the SVM adaptively. Evaluation results show that the proposed WMA-SVM model significantly outperforms all benchmark models across various performance metrics. Moreover, based on the Permutation Feature Importance (PFI) method, the unit weight ( $\gamma$ ), slope angle ( $\beta$ ), and internal friction angle ( $\phi$ ) were identified as the most critical features influencing the classification outcomes for this dataset. Finally, the model's generalization capability was further validated through eight independent engineering case studies, revealing a high consistency between the predictions and the actual stability states. This research provides a modeling framework with considerable generalization potential for the intelligent analysis of slope stability.

**Keywords:** slope stability; machine learning; whale migration algorithm (WMA); support vector machine (SVM); imbalanced data; generalization capability