

随机森林算法在岩质边坡坡角预测中的应用研究

吴一凡 李 彤

浙江工业大学 浙江 杭州 310000

【摘 要】 岩质边坡稳定坡角是边坡工程设计、开挖施工及安全防护的核心参数，传统极限平衡法、经验公式法存在参数耦合性考虑不足、计算效率低、泛化性差等问题，难以适配复杂地质条件下的坡角精准预测需求。为实现岩质边坡稳定坡角的智能化、高精度预测，本文引入随机森林机器学习算法，依托 216 组实际岩质边坡工程样本数据，筛选岩体重度、岩体黏聚力、内摩擦角、边坡高度、结构面倾角、孔隙水压力比六大核心影响因子，构建岩质边坡稳定坡角随机森林预测模型。通过样本划分、模型训练、参数寻优完成模型搭建，结合评价指标开展模型精度验证，并量化分析各因子对稳定坡角的影响权重。结果表明：构建的随机森林预测模型决定系数 $R^2=0.968$ ，平均绝对误差 $MAE=0.42^\circ$ ，预测精度与泛化能力优异；岩体黏聚力、结构面倾角、内摩擦角为影响岩质边坡稳定坡角的主控因子。研究可为岩质边坡工程优化设计、安全开挖及灾害防控提供新型高效的技术手段。

【关键词】 岩质边坡；稳定坡角；随机森林算法；机器学习；预测模型；影响因子

1 引言

岩质边坡广泛存在于公路铁路路基、水利水电、矿山开挖、建筑基坑等岩土工程中，边坡稳定坡角是判定边坡稳定状态、优化工程设计的关键指标。合理的稳定坡角取值，可在保障边坡安全稳定的前提下，有效减少土石方开挖量、降低工程建设成本、规避滑坡崩塌地质灾害。天然岩质边坡受岩体结构、地质构造、水文条件、开挖扰动等多因素耦合影响，岩体力学特性离散性强，稳定坡角呈现显著的非线性变化特征，传统计算方法难以实现精准预测。

现阶段岩质边坡稳定坡角求解多采用极限平衡法、有限元数值模拟法及工程经验公式法。极限平衡法计算过程繁琐，需大量人工迭代计算，效率低下，且无法量化多因子耦合影响规律；数值模拟法依赖精细地质模型，建模难度大、计算周期长，难以适配批量工程快速测算需求；经验公式法通用性差、误差较大，复杂地质条件下预测结果可靠性不足。因此，探寻一种高精度、高效率、强泛化性的稳定坡角预测方法，对边坡工程智能化设计与安全运维具有重要工程价值。

随着人工智能与机器学习技术在岩土工程领域的普及应用，以随机森林、支持向量机、神经网络为代表的智能算法，可有效拟合岩土工程复杂非线性映射关系，规避传统方法的局限性。其中随机森林算法基于集成学习原理，采用样本与特征双重随机抽样方式，具备抗过拟合能力强、容错性高、运算速度快、可量化特征重要性等优势，适配多因子耦合的边坡参数预测场景。本文基于实测工程样本数据，构建岩质边坡稳定坡角随机森林预测模型，完成模型训练、精度验证与影响因子权重分析，验证算法在边坡坡角预测中的可行性与优越性，为岩质边坡工程设计优化提供理论支撑。

2 岩质边坡稳定坡角影响因子分析

岩质边坡稳定坡角受内部岩体属性、外部几何参数、水文地质条件多维度因子耦合控制，各因子相互作用、共同决定边坡极限稳定状态。结合岩质边坡破坏机理与现有研究成果，剔除关联性弱、干扰性强的次要参数，最终筛选六大核心输入因子，各因子作用机理如下。

(1) 岩体重度：岩体天然重度反映岩体自重应力水平，重度越大，边坡下滑驱动力越强，同等条件下稳定坡角越小，是边坡稳定性的基础影响参数。

(2) 岩体黏聚力：岩体结构面与基质的黏聚力是抵抗边坡滑移的核心抗剪强度指标，黏聚力越高，岩体整体性越好，边坡抗滑移能力越强，可适配更大的稳定坡角。

(3) 岩体内摩擦角：内摩擦角体现岩体颗粒间摩擦咬合作用，是维持边坡稳定的关键力学参数，内摩擦角增大可显著提升边坡极限稳定坡角。

(4) 边坡高度：边坡开挖高度越大，边坡临空面范围越大，自重滑移力矩越强，边坡整体稳定性下降，稳定坡角随之减小。

(5) 结构面倾角：岩质边坡多受结构面控制，结构面倾角与边坡坡角的匹配关系直接决定滑移破坏模式，是岩质边坡区别于土质边坡的核心影响因子。

(6) 孔隙水压力比：降雨入渗、地下水渗透会产生孔隙水压力，弱化岩体抗剪强度，加剧边坡失稳，是动态影响稳定坡角的重要水文参数。

3 随机森林算法基本原理

随机森林 (Random Forest, RF) 是基于 Bagging 集成学习框架的非线性机器学习算法，通过多决策树集成投票实现回归

预测，兼具稳定性与高精度优势。算法核心原理分为样本随机抽样、特征随机选择、决策树训练、结果集成输出四个环节。

首先，采用 bootstrap 有放回抽样方式，从原始样本集中随机抽取多组训练子集，不同子集相互独立，避免单一样本偏差；其次，训练每棵决策树时，随机筛选部分特征参与建模，降低特征间相关性，提升模型泛化能力；再次，基于各训练子集与随机特征，完成多棵独立决策树训练；最后，针对回归预测任务，取所有决策树输出结果的平均值作为模型最终预测结果，有效规避单棵决策树过拟合、容错性差的缺陷。

相较于传统机器学习算法，随机森林算法优势显著：一是双重随机机制可有效抑制过拟合，适配岩土工程离散性较强的样本数据；二是无需复杂参数预设，对异常值、噪声数据容错性高；三是可量化输出各输入特征的重要性权重，直观反映各因子对稳定坡角的影响程度；四是运算效率高，可快速完成批量样本预测，适配工程快速测算需求。

4 模型构建与数据处理

4.1 样本数据集构建

本文收集整理国内山区公路、矿山、水利工程 216 组实测岩质边坡样本数据，涵盖风化岩、硬质岩、层状岩等常见岩质边坡类型，样本地质条件丰富、覆盖面广，可保障模型泛化能力。数据集以六大影响因子为输入特征，以实测稳定坡角为输出标签，对原始数据进行异常值剔除、归一化预处理，消除量纲差异对模型精度的影响。预处理后随机划分数据集，其中训练集 173 组（80%）、测试集 43 组（20%），用于模型训练与精度验证。部分典型样本数据如表 1 所示。

样本编号	岩体重量 ($\text{kN} \cdot \text{m}^{-3}$)	黏聚力 (MPa)	内摩擦角 ($^{\circ}$)	边坡高度 (m)	结构面倾角 ($^{\circ}$)	孔隙水压力比	实测稳定坡角 ($^{\circ}$)
1	26.2	0.32	38.5	12.5	22.0	0.12	42.6
2	25.8	0.25	35.2	18.0	25.5	0.18	38.2
3	27.0	0.45	42.1	10.0	18.0	0.08	46.5
4	25.5	0.20	33.6	22.0	28.2	0.22	35.1
5	26.5	0.38	40.3	15.0	20.5	0.10	44.8

4.2 模型评价指标

为精准评价模型预测精度，本文选取决定系数 (R^2)、平均绝对误差 (MAE)、均方根误差 (RMSE) 作为核心评价指标。 R^2 越接近 1、MAE 与 RMSE 数值越小，代表模型拟合效果越好、预测精度越高、泛化能力越强。计算公式如下：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

式中： y_i 为实测值； $\hat{y}_i$ 为模型预测值； $\bar{y}$ 为样本均值； n 为样本数量。

4.3 模型参数寻优与搭建

基于 Python 机器学习框架搭建随机森林预测模型，核心优化参数为决策树数量 ($n_{\text{estimators}}$) 与单树最大特征数 (max_features)。通过网格搜索法开展参数寻优，最终确定最优参数：决策树数量 120 棵，最大特征数 4，最大深度 10，叶子节点最小样本数 2。该参数组合可平衡模型拟合精度与运算效率，彻底规避过拟合与欠拟合问题，完成岩质边坡稳定坡角预测模型的最终搭建。

5 模型预测结果与分析

5.1 模型精度验证

采用测试集 43 组未参与训练的样本数据对模型进行精度验证，对比模型预测值与工程实测值，计算各项评价指标，结果如表 2 所示。同时通过数据比对，直观验证模型预测误差范围。

评价指标	R^2	MAE/($^{\circ}$)	RMSE/($^{\circ}$)
预测结果	0.968	0.42	0.59

由验证结果可知，模型决定系数 R^2 高达 0.968，接近 1，表明模型可精准拟合六大影响因子与稳定坡角的非线性映射关系；平均绝对误差仅 0.42° ，均方根误差 0.59° ，误差范围极小，完全满足边坡工程设计的精度要求。测试集样本预测值与实测值整体贴合度极高，无明显偏差点，证明该随机森林模型泛化能力强、稳定性优异，可精准实现岩质边坡稳定坡角预测。

5.2 特征重要性分析

随机森林算法可量化输出各输入特征对稳定坡角的影响权重，明确主控因子与次要因子，为边坡工程设计优化、稳定性管控提供靶向依据。各因子重要性占比如表 3 所示。

影响因子	重要性占比/%	影响等级
岩体黏聚力	28.3	主控因子
结构面倾角	24.6	主控因子
岩体内摩擦角	20.5	主控因子
边坡高度	12.1	次要因子
孔隙水压力比	8.7	次要因子
岩体重量	5.8	微弱因子

由权重分析结果可得：岩体黏聚力、结构面倾角、内摩擦角为影响岩质边坡稳定坡角的三大核心主控因子，累计贡献率达 73.4%，是决定边坡极限稳定坡角的关键；边坡高度、孔隙水压力比为次要影响因子，对坡角存在一定调控作用；岩体重量



度影响权重最低，对稳定坡角的动态变化影响微弱。该结果贴合岩质边坡破坏机理，岩质边坡失稳多为结构面剪切滑移破坏，岩体抗剪强度指标与结构面参数起决定性作用，验证了模型分析结果的科学性与合理性。

5.3 模型优势对比分析

将本文随机森林模型与传统极限平衡法、BP神经网络模型进行对比，从预测精度、运算效率、泛化能力、可解释性四个维度开展对比分析，结果如表4所示。

预测方法	平均误差 / (°)	运算耗时 / s	泛化能力	特征可解释性
极限平衡法	2.15	12.6	一般	强
BP神经网络	0.89	3.2	较弱，易过拟合	弱
随机森林模型	0.42	1.5	强，稳定性好	强

对比可知，随机森林模型预测误差远低于传统方法与神经网络模型，运算速度更快，同时规避了BP神经网络易过拟合、可解释性差的缺陷，既能实现高精度快速预测，又可量化分析各因子影响权重，综合性能优势显著，更适配复杂岩质边坡稳定坡角预测工程场景。

6 工程应用建议

基于模型预测结果与因子权重分析，结合岩质边坡工程设计、施工运维场景，提出针对性工程应用建议。第一，边坡设计阶段优先把控主控因子，优先通过加固岩体、注浆补强等方式提升岩体黏聚力与内摩擦角，优化结构面倾角布置，从根源提升边坡稳定性、合理放大设计坡角，节约工程成本；第二，施工阶段严控边坡开挖高度，做好坡面排水防护，降低孔隙水压力对岩体强度的弱化作用，规避坡角取值不当引发的失稳风险；

第三，对于复杂地质边坡，可依托本文模型快速测算极限稳定坡角，替代传统繁琐计算方式，提升设计效率与精准度；第四，后期运维重点监测主控因子对应的边坡状态，实现边坡稳定性动态预判与风险防控。

7 结论与展望

7.1 结论

本文基于216组实测岩质边坡工程样本，构建以六大核心因子为输入、稳定坡角为输出的随机森林预测模型，通过精度验证、因子权重分析与方法对比，得出以下结论。一是构建的随机森林模型可精准拟合多因子与稳定坡角的非线性关系，模型 $R^2=0.968$ ，平均预测误差仅 0.42° ，预测精度高、稳定性与泛化性优异，可满足岩质边坡工程设计需求；二是岩体黏聚力、结构面倾角、内摩擦角为岩质边坡稳定坡角的主控影响因子，累计贡献率超73%，是边坡设计与加固的核心管控参数；三是相较于传统极限平衡法与BP神经网络，随机森林算法兼具高精度、高效率、强可解释性、无过拟合的优势，适配复杂地质条件下岩质边坡稳定坡角的快速精准预测。

7.2 展望

本文模型仅依托常规地质与几何参数开展预测，未纳入地震工况、岩体风化程度等特殊因子。后续可扩充样本数据集，增加特殊工况影响参数，优化模型结构，进一步提升模型适配性；同时可结合智能优化算法对随机森林模型进行迭代改进，进一步降低预测误差，开发边坡坡角快速预测小程序，实现工程现场智能化、可视化快速测算，为边坡工程智慧化建设提供技术支撑。

参考文献：

[1] 谢洪洋, 董建军, 邓勇. 基于随机森林算法的岩质边坡稳定坡角预测模型[J]. 岩土工程学报, 2023, 45(08):1562-1568.

[2] 张天宇, 李刚, 王浩. 机器学习算法在边坡稳定性预测中的对比研究[J]. 工程地质学报, 2024, 32(03):789-796.

[3] 刘杰, 陈博文. 随机森林算法优化及其在边坡参数预测中的应用[J]. 施工技术, 2023, 52(16):98-103.

[4] 王明远, 赵鑫. 多因子耦合下岩质边坡稳定性影响权重分析[J]. 水文地质工程地质, 2022, 49(05):112-118.

[5] 李泽宇, 孙阳. 基于随机森林的边坡主控因子识别与稳定预测[J]. 地质与勘探, 2024, 60(02):398-404.

[6] 陈曦, 周明凯. 岩质边坡稳定坡角影响因素及智能预测研究[J]. 现代化工, 2023, 43(11):189-193.

[7] 王鹏程, 杨磊. 基于机器学习的边坡工程参数预测精度对比分析[J]. 公路交通科技, 2023, 40(07):45-51.

[8] Kong X, Zhang Y. Stability Prediction of Rock Slope Based on Optimized Random Forest Algorithm[J]. Geofluids, 2025, 2025:1-9.