

基于数据驱动与可解释性分析的堤防管涌灾害预测模型

张茜^{1,2,3}, 张超^{1,2,3,4}, 夏洋洋^{1,2,3}, 王翠霞^{1,2,3,4}, 方宏远^{1,2,3,4}

(1. 郑州大学 水利与交通学院/黄河实验室(郑州大学)/地下工程研究院, 河南 郑州 450001;

2. 重大基础设施检测修复技术国家地方联合工程实验室, 河南 郑州 450001;

3. 地下工程灾变防控省部共建协同创新中心, 河南 郑州 450001; 4. 黄河实验室(河南), 河南 郑州 450001)

摘要:【目的】探析兼顾预测精度与可解释性的堤防管涌灾害预测方法, 破解传统识别方法精度与效率偏低、常规机器学习模型可解释性缺失的黑箱问题, 为堤防管涌风险预警与防灾管控提供理论借鉴。【方法】依托长江阳新堤段管涌实测数据集, 基于数据驱动的机器学习算法, 构建堤防管涌多层次可解释预测模型。对比 3 种单一学习模型(反向传播神经网络、多层感知机、支持向量机)与 4 种集成学习模型(随机森林、梯度提升决策树、极限梯度提升树、CatBoost)的预测效果, 筛选最优预测模型; 结合极限梯度提升树(XGBoost)与 Shapley 加性解释开展多层次可解释性分析, 识别管涌关键影响因子, 揭示各因子对管涌灾害的作用机制。【结果】①集成学习模型的预测精度优于单一学习模型的, 其中极限梯度提升树测试集的拟合优度(R^2)为 0.961 7, 均方根误差(RMSE)为 0.097 0, 预测精度最高, 且鲁棒性好; ②弱透水层上覆土层厚度 H 、孔隙比 e 和水位高度差 Δh 为管涌灾害的主要影响因子; ③依据个体条件期望曲线, 管涌灾害的 3 个主要影响因子触发管涌的临界值分别为 7.5 m、0.84、4.2 m。【结论】构建的数据驱动模型可以高精度预测堤防管涌, 多层次可解释性分析明确了关键致灾因子及其触发管涌的临界值。该研究成果为堤防管涌灾害超前预警、风险研判提供理论支撑与科学依据。

关键词:堤防管涌; 数据驱动; 机器学习; 单一学习模型; 集成学习模型; 多层次可解释性分析

中图分类号:TV641

文献标识码:A

文章编号:2096-6792(2026)04-0010-10

水利堤防是江河防洪体系的关键基础设施, 肩负径流拦蓄、防汛抗洪、水量调控等重要功能。截至 2023 年底, 我国已建成五级及以上江河堤防 33.2 万 km, 保护人口 6.4 亿、耕地面积 4.2×10^5 万 hm^2 。管涌是水利堤防工程中最常见的灾害之一。据统计, 1998 年长江特大洪水期间, 长江中下游堤防累计发生各类险情 73 825 处, 其中管涌占比 35.2%; 重大险情 1 702 处, 其中管涌占比高达 51.2%。管涌已成为堤防失稳溃决的主要诱因, 是堤防风险评估的核心研究方向^[1-2]。

国内外学者围绕管涌致灾机理与侵蚀演化过程开展了大量研究。在机理方面, 姚秋玲等^[3]将数学模型、物理模型试验和数值模型相结合, 研究了

单层、双层及三层堤基结构的管涌灾害发育的宏观机理, 发现整体破坏时的临界水力比降随结构层数的增加而逐步降低; 黄海均等^[4]研究了上覆粗粒土层对堤基管涌影响的微观机制, 采用三维颗粒流数值模拟和休止角标定法建立了材料宏观参数与颗粒细观参数之间的联系; 倪小东等^[5]基于透明土技术与颗粒流模拟, 从颗粒间接触力链破坏与重构的角度对管涌灾害的细观机理展开研究, 发现颗粒级配为不稳定级配, 且细料含量越高、颗粒摩擦系数越小时, 土体内部力链容易出现断裂与重构现象, 土体结构容易发生失稳。在数值模拟方面, El Shamy U 等^[6]建立了瞬态全耦合连续离散流体力学模型, 模拟极端流动条件下管涌诱发土体变形的

收稿日期/Received: 2024-10-21

修回日期/Revised: 2024-12-17

基金项目:国家自然科学基金项目(U2568213); 河南省自然科学基金项目(262300421398, 242300421001); 安徽省自然科学基金水科学联合基金项目(2308085US15); 河南省重点研发计划项目(241111322700); 中原基础研究领军人才计划项目(22ZYLJ005); 河南省科技攻关计划重点项目联合基金项目(235200810014)。

第一作者:张茜(2000—), 女, 硕士研究生, 从事堤坝工程灾害预测方面的研究。E-mail: zhangxi_8816@163.com。

通信作者:张超(1984—), 男, 教授, 博导, 博士, 从事工程修复材料的多尺度物理力学性能与提升、先进结构材料的设计与应用和数据驱动技术等方面的研究。E-mail: chao.zhang.zzu@outlook.com。

情况;Liang Yu 等^[7]基于固液相互作用和两相流理论,建立流固耦合模型,模拟管涌演化过程;Wewer M 等^[8]建立了一种将二维达西渗流方程解与 Exner 的一维输沙质量守恒方程相结合的可用于边界元泥沙输移的管涌数值模型。

传统机理模型求解复杂、时效性差,难以直接用于工程现场的快速风险评估^[9]。为此,诸多学者开展管涌灾害识别技术研究。刘胜等^[10]通过对管涌渗流场的数值模拟,分析堤防不同位置处的孔隙水压力变化规律,提出基于孔隙水压力变化特征的管涌快速识别方法;周华敏等^[1]依托磁电探测技术,通过堤防表层磁场空间变化反演地下土体电导率,根据电导率的变化识别管涌通道;房纯钢等^[11]基于地球物理瞬变电磁法(Geophysical Transient Electromagnetic Method, GTEM)研发了 SDC-2 型堤防渗漏检测仪,该仪器可有效识别堤防内部渗漏及管涌。上述探测方法依赖人工巡查与专业设备,普遍存在监测效率低、人力成本高、实时性不足等问题。

近年来,数据驱动技术得到快速发展,机器学习算法逐步被应用于堤防管涌灾害预测领域。陈亮等^[12]基于径向基函数(Radial Basis Function, RBF)人工神经网络模型预测堤防管涌,并证明了在预测性能方面,RBF 神经网络模型优于反向传播神经网络(Back-Propagation Neural Network, BPNN)

模型;陆公义等^[13]结合 YOLO 图像识别算法与多权值神经网络,依据管涌实景图片,实现对险情的智能分类和识别;杨端阳^[14]采用灰色关联法筛选致灾因子,耦合遗传算法与深度信念神经网络,预测管涌风险;Band S S 等^[15]基于随机森林(Random Forest, RF)、支持向量机(Support Vector Machine, SVM)和贝叶斯广义线性模型,建立了伊朗马尔卡齐省 Zarandieh 流域的管涌侵蚀模型。现有机器学习模型虽能实现对管涌发生的快速预测,但普遍存在黑箱缺陷,模型内部逻辑透明度低、可解释性差,制约了其在水利工程中的应用。

本文基于数据驱动技术和可解释性机器学习算法,以长江阳新堤防管涌灾害为研究对象,构建了该堤段管涌灾害数据库;对比单一学习模型与集成学习模型的预测性能,优选高精度预测模型;结合敏感性分析与 Shapley 加性解释(SHapley Additive exPlanations, SHAP)框架开展多层次可解释性分析,明确主要致灾因子,量化因子作用规律,求解各因子致灾的临界值,为堤防管涌风险评估、隐患管控提供科学参考。

1 基于数据驱动与可解释性分析的堤防管涌灾害预测模型

1.1 建模流程

建模流程如图 1 所示,建模步骤如下。

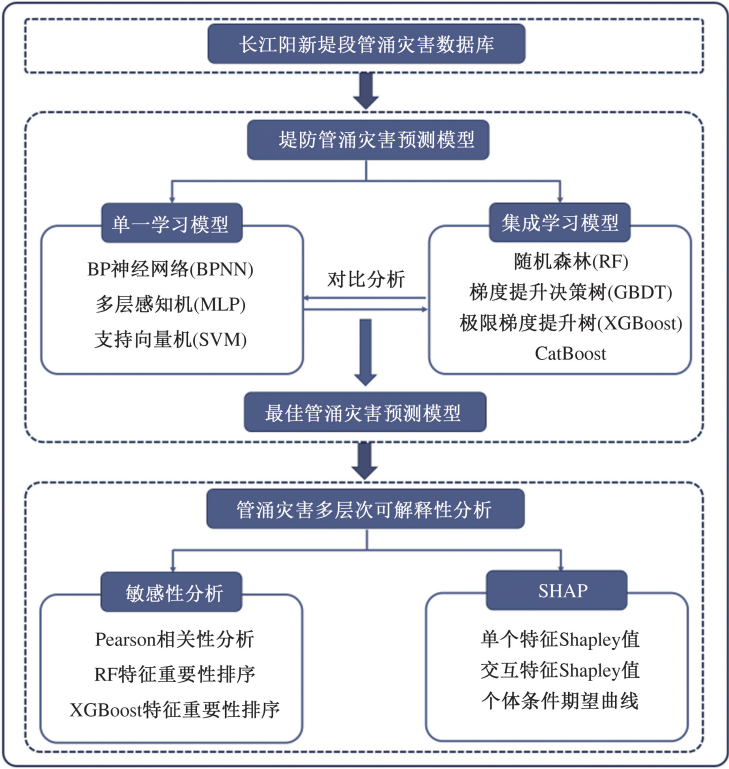


图 1 建模流程

Fig. 1 Modeling process

1)数据库构建。通过梳理堤防管涌灾害的成因和文献调查,筛选出尽可能丰富的影响管涌灾害发生的因子;基于筛选结果和长江阳新堤段管涌灾害的基础数据,以及室内试验测得的各影响因子值,整理并建立长江阳新堤段管涌灾害数据库。

2)模型筛选。将数据库中的管涌灾害样本按4:1划分出训练集与测试集,分别利用BPNN、多层感知机(Multi-Layer Perceptron, MLP)、支持向量机(Support Vector Machine, SVM)3种单一学习模型和随机森林(Random Forest, RF)、梯度提升决策树(Gradient Boosting Decision Tree, GBDT)、极限梯度提升机(eXtreme Gradient Boosting, XGBoost)和CatBoost 4种集成学习模型,建立堤防管涌灾害的预测模型,以均方根误差(Root Mean Square Error, RMSE)和拟合优度 R^2 为评价指标,评价各模型的预测精度和稳定性,综合确定最优模型。

3)多层次可解释性分析。采用Pearson相关性分析、RF及XGBoost内置特征重要性开展敏感性分析;结合Shapley加性解释(SHAP),对堤防管涌进行多层次可解释性分析,筛选影响管涌灾害的关键特征;根据个体条件期望(Individual Conditional Expectation, ICE)曲线得到导致管涌灾害的各因子临界值。

堤防管涌灾害本身是一个二分类问题。为量

化灾害风险等级、直观表征险情发育程度,本文统一采用回归模型计算管涌发生的概率,取值区间为0~1。

1.2 堤防管涌灾害数据库构建

文中的“管涌”是指水流通过土壤孔隙或裂缝形成水流通道,对土壤造成侵蚀,导致土壤被带走形成空洞的现象。管涌灾害的发生受土体自身条件、水力条件、水土耦合因素及外界破坏等影响^[3],堤防地质条件差异显著影响管涌发育特征。

长江阳新堤断面大多为二元堤基,上部为透水性弱的黏土层,下层为透水性强的砂土层^[16],其基本参数见表1。由表1可以看出:黏土层与砂土层的渗透系数差异显著。渗透系数与颗粒级配密切相关,是导致管涌发生的关键因素。上部弱透水层渗透系数 k 越大,土体抗管涌能力越弱,管涌发生的可能性越大^[5,17]。弱透水层上覆土层厚度 H 对管涌影响显著,当 $H>8$ m时,可有效抑制管涌发生^[14]。除此之外,有效黏聚力 c' 、有效内摩擦角 φ' 、干密度 ρ_d 、孔隙比 e 、压缩系数 a_v 等细观物理参数也会影响管涌的发生^[18-20]。水力条件方面,水力梯度是影响管涌过程的主要因子^[17],其为水位高差 Δh 与水力半径之比,为便于采集,使用 Δh 作为水力条件参数。

表1 长江阳新堤基断面参数

Tab. 1 Cross-section parameters of Yangxin Levee foundation along the Yangtze River

堤底宽度/m	黏土层深度/m	黏土层渗透系数/(cm/s)	砂土层深度/m	砂土层渗透系数/(cm/s)
60~70	1~3	1×10^{-7}	>10	1×10^{-5}

结合已有研究成果及研究区地质特征,最终选取 Δh 、 H 、 k 、 c' 、 φ' 、 ρ_d 、 e 、 a_v 等8个因子作为影响堤防管涌灾害的因子。其中, Δh 、 H 和 k 值通过实地测量获得,其余因子的值通过室内试验得到。 c' 和 φ' 的值通过快剪试验测得, a_v 的值通过侧限压缩试验获得,密度(如 ρ 和 ρ_d)值利用环刀法测得。 e 的值由式(1)计算得到,含水率 ω 和比重 G_s 的值分别通过击实试验和比重瓶法测得。最终获取193组数据作为堤防管涌灾害数据库的数据。由这些数据可知,管涌发生与不发生的比例为74:119,各因子的取值范围如下: Δh 为1.8~9.0 m, H 为1.0~13.2 m, k 为 $7.5 \times 10^{-7} \sim 5.0 \times 10^{-4}$ cm/s, c' 为5~100 kg/cm, φ' 为14.5~26.5°, ρ_d 为1.3~1.7 g/cm³, e 为0.6~1.3, a_v 为0.1~0.6 MPa⁻¹。

$$e = \frac{G_s \gamma_w (1 + \omega)}{\gamma} - 1, \quad (1)$$

$$\gamma = \rho g, \quad (2)$$

$$\gamma_w = \rho_w g. \quad (3)$$

式中: γ 为土的天然重度,N/m³; γ_w 为水的重度,N/m³; ρ 为土体天然密度,g/m³; ρ_w 为水体天然密度,g/m³; g 为重力加速度,取9.8 m/s²。

1.3 机器学习预测模型筛选

1.3.1 单一学习模型

1)BPNN由输入层、隐含层、输出层构成,每一层由若干个神经元(节点)构成,依靠神经元权重传递更新误差。该模型是目前最成熟、应用最广的神经网络模型^[21]。

2)MLP本质上是一种神经网络模型,拥有多

个隐含层。在其网络架构中,除输入层外,其余每一层均由激活函数进行非线性变换来增强输入与输出间的复杂映射能力。

3) SVM 基于核函数,将低维非线性数据映射至高维空间,求解最优分离超平面,最小化训练误差^[22]。

1.3.2 集成学习模型

1) RF 是基于 Bagging 思想的一种算法^[23],通过 Bootstrap 重采样将原始数据集划分为 N 个新数据集,创建 N 棵决策树,最终输出所有决策树的平均值。

2) GBDT 基于 Boosting 迭代优化思想,以分类回归树 (Classification And Regression Tree, CART) 作为弱学习器,在每次迭代过程中构建一个决策树,使用上一轮迭代所得残差来训练新学习器并更新权重,一直迭代到要求轮数为止。GBDT 的最终输出为全部 CART 预测值的加权和。

3) XGBoost 和 CatBoost 均为 GBDT 框架下的变体算法。与 GBDT 不同的是, XGBoost 的目标函数在 GBDT 泰勒一阶展开的基础上将其展开到了二阶,完整保留损失函数曲率信息; CatBoost 增加了类别特征处理和自适应学习率两个功能,可以处理多种类型的特征,并根据不同的特征自适应地调整学习率,从而加速收敛速度,提高预测精度。

1.3.3 交叉验证

为规避人为参数设置带来的主观性误差,本文采用 5 折交叉验证 (5-fold cross-validation, 5-CV) 完成集成学习模型超参数寻优,筛选最优组合,以便保障模型的泛化性能。

1.4 多层次可解释性分析

1.4.1 敏感性分析

敏感性分析用于量化输入因子对模型输出的影响程度。本文主要采用 Pearson 相关性、RF 特征重要性与 XGBoost 特征重要性开展分析。其中, RF 和 XGBoost 依托算法内置逻辑计算因子的贡献度。Pearson 相关性分析是一种参数统计方法,常用于分析两个变量间线性相关强弱的程度,通常包括显著性 p 和相关系数 r 两个参数。

1.4.2 SHAP 分析

SHAP 依托博弈论思想量化单因子预测的边际贡献,可以弥补机器学习黑箱的缺陷,同步实现对模型全局和局部的可解释性分析,能够直观表征各因子对管涌发生的正负驱动作用。

1.4.3 个体条件期望曲线

个体条件期望 (ICE) 曲线是可视化机器学习预测模型的常用工具。该方法以预测模型为基础,展示预测结果随单个样本特定特征因子数值变化而变化的情况,直观反映输入特征对模型预测结果的影响,能够有效提升模型的可解释性。

2 结果分析

2.1 不同机器学习模型预测管涌的精度

2.1.1 评价指标

评价指标均方根误差 (RMSE) 和拟合优度 (R^2) 的计算公式分别如下:

$$RMSE = \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{n}, \quad (4)$$

$$\begin{cases} R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \\ \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \end{cases} \quad (5)$$

式中: n 为样本数量; y_i 、 y_i^* 分别为第 i 个样本的实测值、预测值; $\bar{y}$ 为样本均值。

RMSE 数值越小,拟合精度越高; R^2 的值越趋近于 1,说明模型泛化能力越强。

2.1.2 模型的超参数

超参数的选取对于机器学习模型的性能至关重要。在本次研究中,对于单一学习模型,参考已有研究^[24-27]取值, BPNN 隐含层层数设为 5,迭代次数为 1 000 次,学习率为 0.01,训练目标最小误差为 0.001;对于 MLP,设置隐含层层数为 4,激活函数为 ReLU,优化器设置为 Adam,学习率为 0.001,迭代次数为 200;对于 SVM,采用 epsilon-SVR 模型,核函数类型为 RBF,损失系数取值 0.01。

集成学习模型通过 5 折交叉验证完成参数寻优,得到的最优超参数见表 2。表中: $n_estimators$ 为决策树的棵数,控制模型规模,寻优范围为 0~500; max_depth 为决策树的最大深度,控制模型的复杂程度,寻优范围为 1~10; $max_features$ 为每棵决策树在划分节点时所考虑的特征数量,寻优范围为 0.1~1.0; $learning_rate$ 为每棵决策树的权重缩减系数,控制每次迭代中参数更新的幅度,寻优范围为 0.1~1.0; $iterations$ 为迭代次数,控制模型的复杂度和精度,寻优范围为 0~500。

表 2 集成学习模型的超参数
Tab. 2 Hyperparameters of ensemble learning models

模型	n_estimators	max_depth	max_features	learning_rate	iterations
RF	100	8	0.7	—	—
GBDT	500	4	0.7	0.1	100
XGBoost	500	3	0.3	0.3	100
CatBoost	100	6	0.7	0.3	100

2.1.3 预测精度

各模型评价指标计算结果见表 3。其中,训练集和测试集是将数据集中数据按 4:1 的比例进行划分所得,所有模型基于相同的训练集和测试集进行训练。由表 3 可以看出:对于单一学习模型,MLP 模型的预测精度最低,其测试集的 RMSE 值和 R^2 值分别为 0.172 2 和 0.879 3,是 R^2 值唯一未达到 0.900 0 的预测模型;BPNN 和 SVM 模型的预测精度相近,其 R^2 值均大于 0.940 0, RMSE 值均小

于 0.020 0,说明模型的预测精度较高。对于集成学习模型,其预测性能整体优于单一学习模型,测试集 R^2 值排序为: XGBoost>CatBoost>GBDT>RF;除了 RF 外,其余集成学习模型的 R^2 值均大于 0.950 0,其中 XGBoost 模型拟合效果最优,测试集 R^2 值为 0.961 7, RMSE 值为 0.097 0。最优单一学习模型 BPNN 的 R^2 小于最差集成学习模型 RF 的,说明集成学习模型在堤防管涌预测中具备显著优势,该结论与文献[27]的研究成果一致。

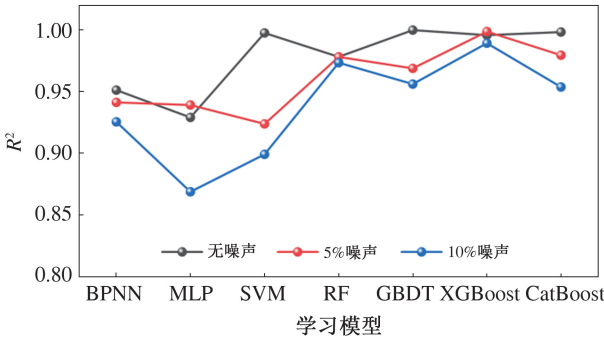
表 3 机器学习模型评价指标
Tab. 3 Evaluation indicators of machine learning models

机器学习模型		训练集		测试集	
		RMSE	R^2	RMSE	R^2
单一学习模型	BPNN	0.001 7	0.951 2	0.019 1	0.947 8
	MLP	0.126 0	0.929 0	0.172 2	0.879 3
	SVM	0.000 5	0.997 6	0.013 8	0.943 9
集成学习模型	RF	0.069 9	0.978 2	0.111 9	0.949 1
	GBDT	0.002 6	0.999 9	0.105 9	0.954 4
	XGBoost	0.030 8	0.995 8	0.097 0	0.961 7
	CatBoost	0.019 6	0.998 3	0.102 2	0.957 6

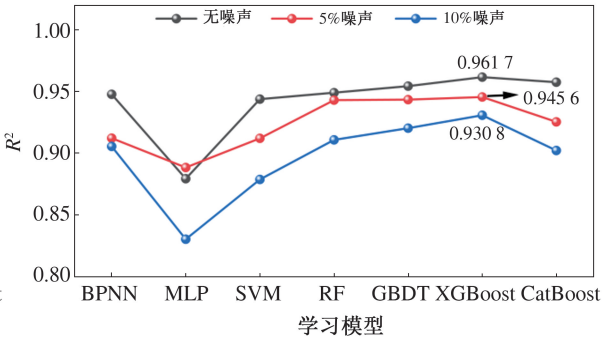
2.1.4 模型鲁棒性

为验证模型的抗干扰能力,分别向训练集中掺入 5%、10%噪声,测试集保持不变,各模型评价指标的变化规律如图 2 所示。由图 2 可知:掺入噪声后,所有模型 R^2 值小幅减小、RMSE 值略有增大,模型精度因受扰动影响而衰减;集成学习模型的稳

定性普遍优于单一学习模型,其中 RF、GBDT、XGBoost 指标的波动幅度最小。XGBoost 抗干扰能力最优,分别在 5%、10%噪声下,其测试集的 R^2 值依次为 0.945 6、0.930 8,在扰动环境下仍可维持预测结果的高精度,鲁棒性优异。



(a) 各模型在训练集上的 R^2 值



(b) 各模型在测试集上的 R^2 值

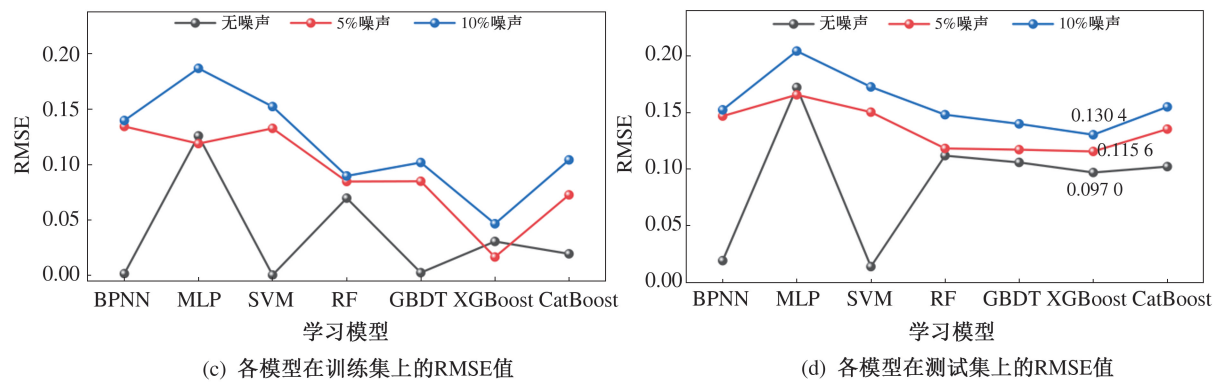


图 2 不同噪声比下各模型的预测效果

Fig. 2 Prediction performance of various models under different noise ratios

2.2 堤防管涌灾害的多层次可解释性分析

对比机器学习模型虽可获得高精度堤防管涌预测模型,但模型内在运行机理难以获取,无法明确各输入因子的作用机制,也难以阐释模型输出结果随输入因子变化的响应规律,这在一定程度上限制了模型的可信度。本文依托敏感性分析、SHAP、ICE 曲线开展多层次可解释性分析,从多个角度探究各因子与管涌灾害之间的关系及不同因子之间的相互依赖性。

2.2.1 敏感性分析

表 4 与图 3(a) 为 Pearson 相关性分析结果。其中,表 4 为各影响因子与堤防管涌灾害的显著性检验结果,图 3(a) 为依据 Pearson 相关系数绘制的相关性热力图。图 3(a) 内正、负数值分别对应变量间正、负相关关系,图例圆圈颜色越深、尺寸越大,代表变量间相关性程度越高。

由表 4 可以看出:管涌灾害与 H 、 Δh 、 e 和 a_v 在检验水平 $\alpha=0.01$ 下呈极显著性相关,说明 H 、 Δh 、 e 和 a_v 是堤防管涌灾害的主要影响因子,该结论与图 3(a) 的相符。在对堤防管涌影响较大的几个因子中,只有 H 与管涌灾害之间呈负相关,其原因在后文解释。

为了直观展示各因子对管涌灾害的影响程度,图 3(b) 和图 3(c) 分别给出了 RF 和 XGBoost 模型的特征重要性排序及对应数值。由图 3(b) 可以看出,RF 各特征因子重要性从大到小的排序为: H 、 e 、 Δh 、 a_v 、 ρ_d 、 k 、 φ' 、 c' ,前 5 个因子的重要性之和为 0.9518,累计贡献度达 95.18%。由图 3(c) 可以看出,XGBoost 各特征因子重要性从大到小的排序为: H 、 e 、 Δh 、 a_v 、 ρ_d 、 k 、 φ' 、 c' ,前 5 个因子重要性的排序结果与 RF 的一致。XGBoost 模型的前 3 个因子的重要性之和为 0.9806,累计贡献度为 98.06%。

表 4 显著性水平 p

Tab. 4 Significance level p

特征因子	Δh	H	k	c'	φ'	ρ_d	e	a_v
p	2.925E-67*	1.436E-70*	0.380	0.392	0.697	0.404	2.112E-50*	1.684E-33*

注: * 表示在 0.01 水平(双侧)上显著相关。

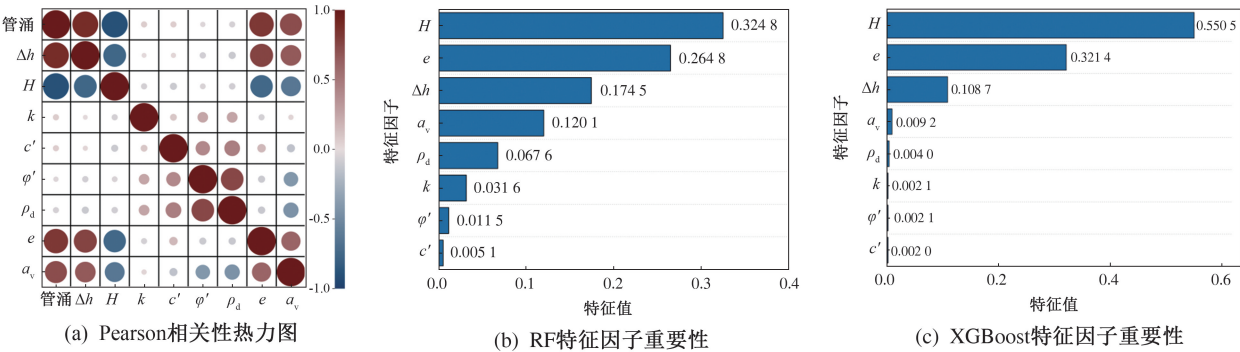


图 3 特征因子重要性分析结果

Fig. 3 Results of feature importance analysis

2.2.2 SHAP 分析

图 4(a) 是 Shapley 值的汇总图,图中的颜色表

示特征值的大小(颜色越红代表取值越大,颜色越蓝代表取值越小),Shapley 值的正负代表特征因子

与模型输出之间的正相关或负相关。图4(b)是各特征因子之间的交互分析图,图中的颜色表示左边列特征因子的特征值大小(越红代表值越大,越蓝

代表值越小),Shapley值的正或负表示着交互后的特征因子组合与管涌灾害的正相关或负相关。

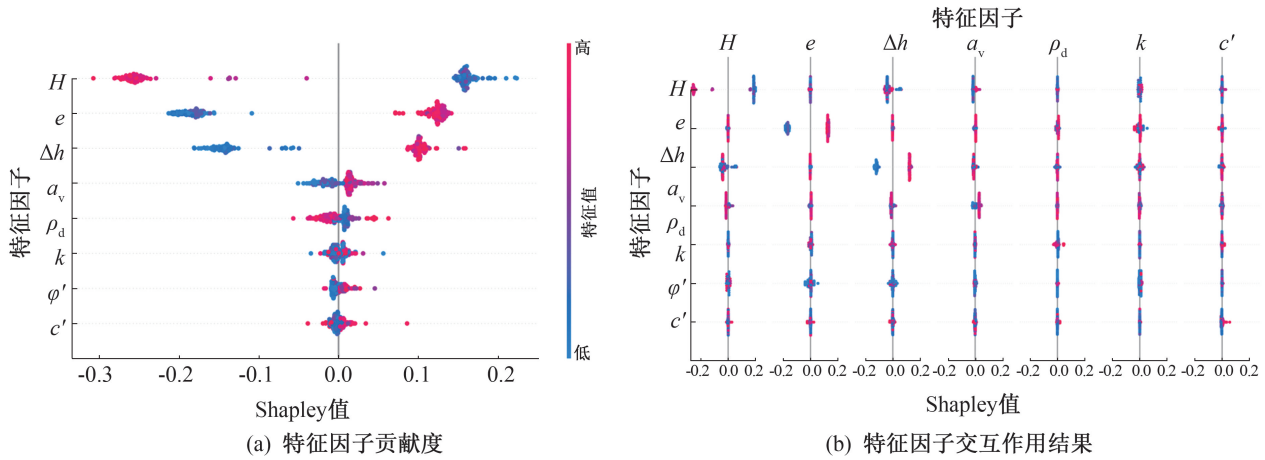


图4 SHAP分析结果

Fig. 4 SHAP analysis results

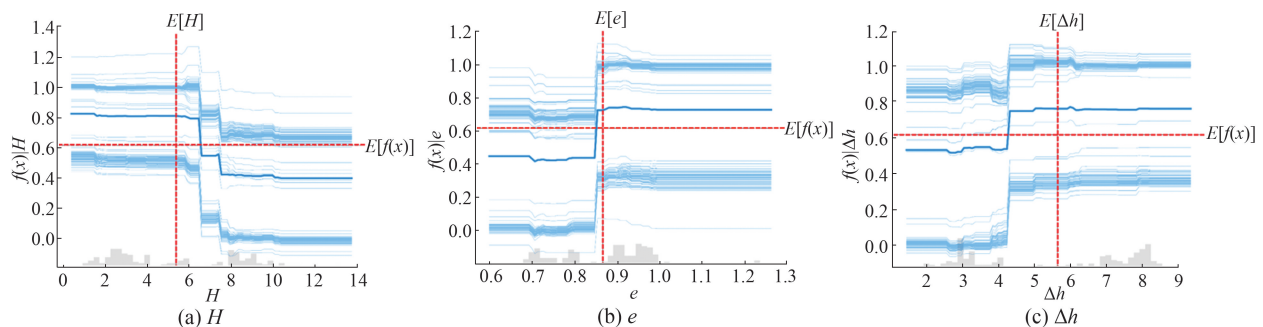
由图4(a)可知, H 、 e 和 Δh 的Shapley值的绝对值最大,表明它们与堤坝管涌灾害的相关性较强,这与敏感性分析结果一致。此外, e 、 Δh 、 a_v 、 k 、 ϕ' 和 c' 的高特征值对应的Shapley值均为正值,表明它们与堤防管涌灾害的发生呈正相关,即 e 、 Δh 、 a_v 、 k 、 ϕ' 和 c' 的值越大,管涌灾害发生的概率越大;相反, H 、 ρ_d 的高特征值对应的Shapley值为负值,表明它们与堤防管涌灾害的发生呈负相关,即 H 、 ρ_d 越大,管涌发生概率反而越低。 H 与堤防管涌灾害的发生呈现负相关,是因为江水位与地下水水位之间的水位差所产生的压力一般作用在上覆土层上,加上堤防强透水层的渗透系数远远大于弱透水层的,使得实际水位差产生的压力大部分作用在弱透水层上覆土层上,当汛期水位抬升时,水位差突然增大,若此时 H 值足够大,便可有效减少管涌的发生。 Δh 与管涌灾害发生呈现极强的正相关,这与文献[28-29]的结论一致。这是因为水力梯度与 Δh 成正比关系, Δh 越大,水力梯度越大,堤防内部水流对土体的侵蚀越严重,当侵蚀发展到一定程

度时会有管涌发生。 e 与管涌灾害呈现显著的正相关,是因为 e 与土体的临界水力梯度紧密相关, e 越大,土体的有效重力和土颗粒间的咬合力越小,土颗粒越容易被水流带走,使得土体的临界水力梯度变小,致使管涌灾害更易发生。

由图4(b)可知:非对角线因子组合中,仅 H 和 Δh 组合的特征值对应的Shapley值明显偏离了零值,且为负值,表明该特征因子组合对堤防管涌的影响较大,与管涌灾害呈负相关,而其中 H 与管涌灾害呈负相关,说明 H 与管涌的相关性比 Δh 与管涌的相关性强,这验证了特征重要性排序的有效性。

2.2.3 特征因子的临界值

基于XGBoost模型绘制各特征因子的个体条件期望曲线,如图5所示,纵坐标为管涌发生概率,横坐标为因子实测取值, $E[i]$ 为特征因子 i 的平均值, $E[f(x)]$ 为模型预测均值,单条曲线对应单个样本的预测值变化,蓝色实线为模型预测结果的平均值。



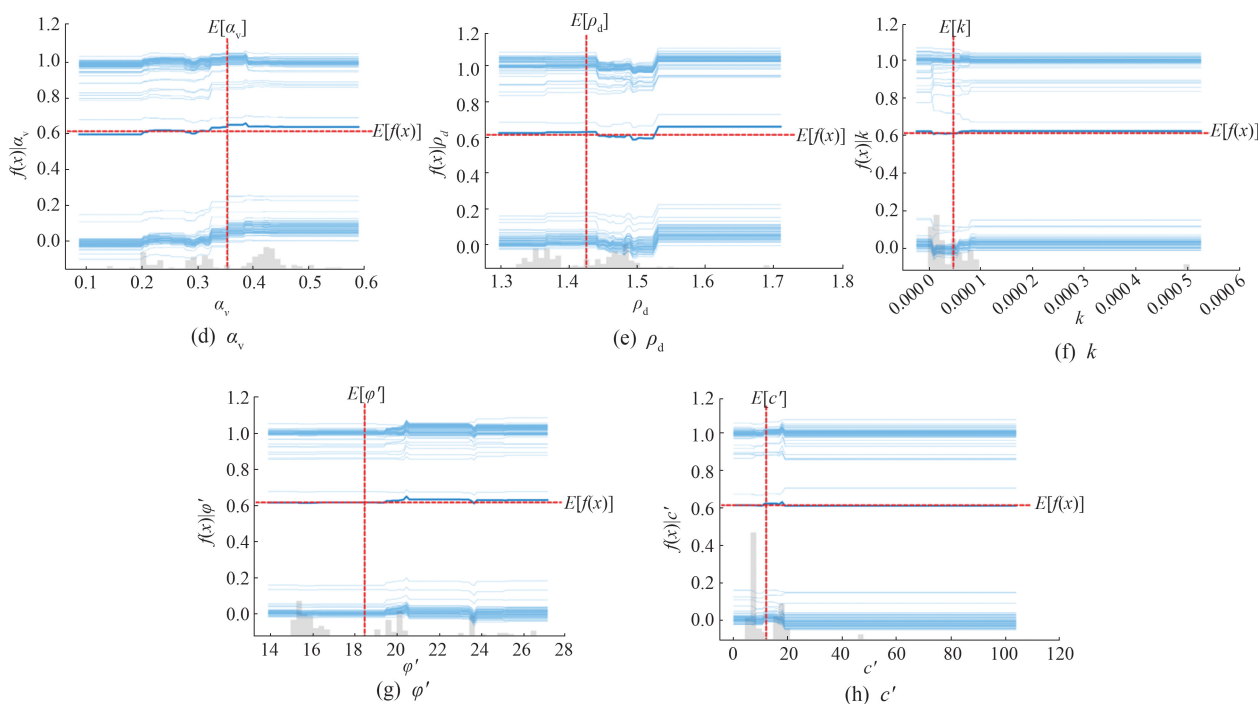


图 5 特征因子的 ICE 曲线

Fig. 5 ICE curves of feature factors

图 5 表明:所有因子均存在明显临界值突变特征,当因子数值突破临界值后,管涌发生概率快速增加;结合图中各曲线拐点确定各因子临界值,分别如下: H 约为 7.5 m, e 约为 0.84, Δh 约为 4.2 m, a_v 约为 0.35 MPa^{-1} , ρ_d 约为 1.5 g/cm^3 , φ' 约为 19° , c' 约为 15 kg/cm 。该成果可为实际工程中堤防稳定性评估提供参考。

3 结论

本文基于数据驱动的机器学习方法,建立了堤防管涌灾害的多层次可解释性机器学习预测模型,对比分析了单一学习(BPNN、MLP、SVM)模型与集成学习(RF、GBDT、XGBoost、CatBoost)模型的预测性能,并对模型的可解释性进行分析。所得主要结论如下:

1) 集成学习模型的预测性能显著优于单一学习模型。单一学习模型中 BPNN 的预测精度最优,其测试集的 R^2 值为 0.947 8, RMSE 值为 0.019 1;集成学习模型中的 XGBoost 预测性能最佳,并具有良好的鲁棒性,其测试集的 R^2 值为 0.961 7, RMSE 值为 0.097 0。

2) 多层次可解释性结果显示, H 、 e 和 Δh 是影响堤防管涌灾害发生的关键特征因子, H 与管涌灾害发生呈负相关, e 和 Δh 与管涌灾害发生呈正相关。

3) 对于管涌灾害,各关键因子的临界值分别

为: H 约为 7.5 m, e 约为 0.84, Δh 约为 4.2 m, a_v 约为 0.35 MPa^{-1} , ρ_d 约为 1.5 g/cm^3 , φ' 约为 19° , c' 约为 15 kg/cm 。

本文构建的可解释性机器学习模型解决了传统黑箱模型可信度低的问题,为堤防管涌灾害风险管控提供了技术支撑。

参 考 文 献

- [1] 周华敏, 郭爱清, 肖国强, 等. 堤防管涌通道微弱磁场信号检测技术[J]. 长江科学院院报, 2023, 40(8): 112–119. [Zhou Huamin, Wu Aiqing, Xiao Guoqiang, et al. Weak magnetic field signal detection technology for dike piping channel [J]. Journal of Yangtze River Scientific Research Institute, 2023, 40(8): 112–119.]
- [2] 康捷. 砂土内部侵蚀机理及其影响因素研究[D]. 西安: 西安理工大学, 2023. [Kang Jie. Study on internal erosion mechanism and influencing factors of sandy soil [D]. Xi'an: Xi'an University of Technology, 2023.]
- [3] 姚秋玲, 丁留谦, 刘昌军, 等. 堤基管涌机理及防治设计准则研究[J]. 中国防汛抗旱, 2022, 32(1): 75–79. [Yao Qiuling, Ding Liujian, Liu Changjun, et al. Study on piping mechanism and prevention design criteria of dike foundation [J]. China Flood & Drought Management, 2022, 32(1): 75–79.]
- [4] 黄海均, 毛海涛, 严新军, 等. 上覆粗粒土层对堤基管涌破坏的细观机制研究[J]. 水资源与水工程学报, 2020, 31(2): 147–154, 160. [Huang Haijun, Mao Haitao, Yan Xinjun, et al. Study on the meso-mechanism of piping damage to embankment foundation from overlying coarse grained soil

- layers[J]. *Journal of Water Resources and Water Engineering*, 2020, 31(2): 147–154, 160.]
- [5] 倪小东,寇恒琦,左翔宇,等. 基于透明土技术与颗粒流方法联合开展管涌细观机理研究[J]. *水利学报*, 2021, 52(12): 1482–1497. [Ni Xiaodong, Kou Hengqi, Zuo Xiangyu, et al. Research on meso mechanism of piping based on transparent soil technology and particle flow method[J]. *Journal of Hydraulic Engineering*, 2021, 52(12): 1482–1497.]
- [6] El Shamy U, Aydin F. Multiscale modeling of flood-induced piping in river levees[J]. *Journal of Geotechnical and Geoenvironmental Engineering*, 2008, 134(9): 1385–1398.
- [7] Liang Yu, Wang Junjie, Liu Mingwu. Two-flow model for piping erosion based on liquid-solid coupling[J]. *Journal of Central South University*, 2013, 20(8): 2299–2306.
- [8] Wewer M, Aguilar-López J P, Kok M, et al. A transient backward erosion piping model based on laminar flow transport equations[J]. *Computers and Geotechnics*, 2021, 132: 103992.
- [9] 刘潘. 土石堤坝管涌演化进程中的电场特征及识别方法研究[D]. 重庆: 重庆交通大学, 2024. [Liu Pan. Study on electric field characteristics and identification method in piping evolution process of earth-rock embankment dams[D]. Chongqing: Chongqing Jiaotong University, 2024.]
- [10] 刘胜,王媛,冯迪. 堤防管涌渗流场响应特征的数值模拟[J]. *水利水电科技进展*, 2023, 43(6): 10–16. [Liu Sheng, Wang Yuan, Feng Di. Numerical simulation of response characteristics of seepage field during embankment piping[J]. *Advances in Science and Technology of Water Resources*, 2023, 43(6): 10–16.]
- [11] 房纯纲,鲁英,葛怀光,等. 堤防管涌渗漏隐患探测新方法[J]. *水利水电技术*, 2001, 32(3): 66–69. [Fang Chungang, Lu Ying, Ge Huaiguang, et al. A new technique for detecting hidden piping defaults in levees[J]. *Water Resources and Hydropower Engineering*, 2001, 32(3): 66–69.]
- [12] 陈亮,赵正信,陈建生,等. 基于 RBF 神经网络的堤防管涌预测方法[J]. *探矿工程(岩土钻掘工程)*, 2007, 34(11): 5–8. [Chen Liang, Zhao Zhengxin, Chen Jiansheng, et al. Prediction of piping occurring in embankment based on RBF neural network[J]. *Drilling Engineering*, 2007, 34(11): 5–8.]
- [13] 陆公义,欧阳鹏,程赞,等. 基于 YOLO 模型的堤坝管涌监测智能识别方法[J]. *水利水电科技进展*, 2024, 44(1): 89–94. [Lu Gongyi, Ouyang Peng, Cheng Yun, et al. Intelligent identification method of dyke piping monitoring based on YOLO model[J]. *Advances in Science and Technology of Water Resources*, 2024, 44(1): 89–94.]
- [14] 杨端阳. 基于灰色关联和 GA-DBN 的长江二元堤防管涌险情预测分析[D]. 郑州: 郑州大学, 2021. [Yang Duanyang. Prediction and analysis of binary embankment piping danger in the Yangtze River based on grey correlation and GA-DBN[D]. Zhengzhou: Zhengzhou University, 2021.]
- [15] Band S S, Janizadeh S, Saha S, et al. Evaluating the efficiency of different regression, decision tree, and Bayesian machine learning algorithms in spatial piping erosion susceptibility using ALOS/PALSAR data[J]. *Land*, 2020, 9(10): 346.
- [16] 王天宇. 管涌发展与悬挂式防渗墙作用机制研究[D]. 北京: 清华大学, 2015. [Wang Dayu. Research on piping development and the function mechanism of suspended cut-off wall[D]. Beijing: Tsinghua University, 2015.]
- [17] 谷华. 堤基管涌险情评价研究及应用[D]. 扬州: 扬州大学, 2022. [Gu Hua. Study and application of piping risk assessment of embankment foundation[D]. Yangzhou: Yangzhou University, 2022.]
- [18] 吴梦喜,叶发明,张琦. 细颗粒流失对砂砾石土本构关系的影响研究[J]. *岩土力学*, 2017, 38(6): 1550–1556. [Wu Mengxi, Ye Faming, Zhang Qi. Effect of fine grain loss on the stress-strain relationship of sand and gravel soils[J]. *Rock and Soil Mechanics*, 2017, 38(6): 1550–1556.]
- [19] 刘孟适,罗强,郭建湖,等. 过渡型粗粒土渗透变形试验及破坏类型判别[J]. *岩石力学与工程学报*, 2017, 36(12): 3102–3110. [Liu Mengshi, Luo Qiang, Guo Jianhu, et al. Experiment on seepage deformation of the transitional coarse grained soil and criterion of failure type[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2017, 36(12): 3102–3110.]
- [20] 梁越,喻金桃,张强,等. 骨架颗粒组成对散粒土管涌规律影响的试验研究[J]. *河海大学学报(自然科学版)*, 2024, 52(1): 63–69. [Liang Yue, Yu Jintao, Zhang Qiang, et al. Experimental study on the effect of skeleton particle composition on piping law of cohesionless soil[J]. *Journal of Hohai University(Natural Sciences)*, 2024, 52(1): 63–69.]
- [21] Zhang Li, Wang Fulin, Sun Ting, et al. A constrained optimization method based on BP neural network[J]. *Neural Computing and Applications*, 2018, 29: 413–421.
- [22] Zhang Chengyu, Brodylo D, Rahman M. Using an object-based machine learning ensemble approach to upscale evapotranspiration measured from eddy covariance towers in a subtropical wetland[J]. *Science of The Total Environment*, 2022, 831: 154969.
- [23] Zhang Xu, Xia Yuanyuan, Zhang Chao, et al. A prediction model for dyke-dam piping based on data augmentation and interpretable ensemble learning[J]. *Engineering Failure Analysis*, 2025, 182: 110174.
- [24] 袁静,王锐,喻国良. 基于人工神经网络的沿海地区底泥盐度计算模型[J]. *华北水利水电大学学报(自然科学版)*, 2024, 45(4): 102–108. [Yuan Jing, Wang Rui, Yu Guoliang. Calculation models for sediment salinity in coastal areas based on artificial neural networks[J]. *Journal of North China University of Water Resources and Electric Power (Natural Science Edition)*, 2024, 45(4): 102–108.]
- [25] 张鹏,梅书浩,石成春,等. 基于机器学习的水库溶解氧预测模型比较研究[J]. *华北水利水电大学学报(自然科学版)*, 2025, 46(1): 87–95. [Zhang Peng, Mei Shuhao,

- Shi Chengchun, et al. Comparative study of dissolved oxygen prediction models for reservoirs based on machine learning [J]. Journal of North China University of Water Resources and Electric Power (Natural Science Edition), 2025, 46 (1): 87–95.]
- [26] 陈存宇, 陈爱喜, 戚晓秋, 等. 基于 MLP 神经网络优化改进的 BW 模型 [J]. 物理学报, 2025, 74 (1): 36–46. [Chen Cunyu, Chen Aixi, Qi Xiaoli, et al. Improved BW model based on MLP neural network optimization [J]. Acta Physica Sinica, 2025, 74 (1): 36–46.]
- [27] 成佩, 孟勇. 集成学习方法的应用与比较 [J]. 统计与决策, 2023, 39 (23): 28–33. [Cheng Pei, Meng Yong. Application and comparison of ensemble learning methods [J]. Statistics and Decision, 2023, 39 (23): 28–33.]
- [28] Xie Qingyu, Liu Jing, Han Bin, et al. Experimental investigation of interfacial erosion on culvert-soil interface in earth dams [J]. Soils and Foundations, 2019, 59 (3): 671–686.
- [29] Schmertmann J H. A method for assessing the relative likelihood of failure of embankment dams by piping; discussion [J]. Canadian Geotechnical Journal, 2002, 39 (2): 495–496.

Prediction Model for Levee Piping Disasters Based on Data-Driven and Interpretability Analysis

ZHANG Xi^{1,2,3}, ZHANG Chao^{1,2,3,4}, XIA Yangyang^{1,2,3}, WANG Cuixia^{1,2,3,4}, FANG Hongyuan^{1,2,3,4}

(1. School of Water Conservancy and Transportation/Yellow River Laboratory/Underground Engineering Research Institute, Zhengzhou University, Zhengzhou 450001, China; 2. National Local Joint Engineering Laboratory of Major Infrastructure Testing and Rehabilitation Technology, Zhengzhou 450001, China; 3. Collaborative Innovation Center for Disaster Prevention and Control of Underground Engineering Jointly Built by Provinces and Ministries, Zhengzhou 450001, China; 4. Yellow River Laboratory (Henan), Zhengzhou 450001, China)

Abstract: [Objective] This study aims to explore a levee piping disaster prediction method that considers both high prediction accuracy and interpretability, address the low accuracy and efficiency of traditional identification approaches and the black-box problem caused by the lack of interpretability in conventional machine learning models, and provide theoretical references for risk early warning of levee piping and disaster prevention and control. **[Methods]** Using a field-measured dataset on piping from the Yangxin levee section of the Yangtze River and data-driven machine learning algorithms, a multi-level interpretable prediction model for levee piping was constructed. The prediction performance of three single learning models (back-propagation neural network, multi-layer perceptron, and support vector machine) and four ensemble learning models (random forest, gradient boosting decision tree, extreme gradient boosting, and CatBoost) was compared to select the optimal prediction model. Extreme gradient boosting (XGBoost) and SHapley Additive exPlanations (SHAP) were combined for multi-level interpretability analysis to identify the key factors influencing piping and reveal the mechanisms by which each factor affects piping disasters. **[Results]** (1) Ensemble learning models had higher prediction accuracy than single learning models. Among them, the XGBoost model had the highest prediction accuracy and good robustness, with a test-set coefficient of determination (R^2) of 0.961 7 and a root mean square error (RMSE) of 0.097 0. (2) The thickness of aquitard overburden (H), void ratio (e), and water level difference (Δh) were the key influencing factors of piping disasters. (3) According to individual conditional expectation curves, the critical values of the three main factors triggering piping were 7.5 m, 0.84, and 4.2 m, respectively. **[Conclusions]** The constructed data-driven model can predict levee piping with high accuracy, and multi-level interpretability analysis clarifies the key causative factors and their critical thresholds for triggering piping. The results provide theoretical reference and a scientific basis for early warning and risk assessment of levee piping disasters.

Keywords: levee piping; data-driven; machine learning; single learning model; ensemble learning model; multi-level interpretable analysis

(编辑:杜明侠)