DOI：10.16285/j.rsm.2024.1298 CSTR:32223.14.j.rsm.2024.1298

基于刃脚土压力的超大锚碇沉井基础下沉智能预测

蔡启航[1,2]，董学超[2,3]，郭明伟[2,3]，卢 正[2,3]，徐 安[4]，蒋 凡[4]

（1. 华北水利水电大学 地球科学与工程学院，河南 郑州 450046；2. 中国科学院武汉岩土力学研究所 岩土力学与工程安全全国重点实验室，湖北 武汉 430071；3. 中国科学院大学，北京 100049；4. 中铁桥隧技术有限公司，江苏 南京 210061）

摘 要：沉井基础因整体刚度大，承载能力强已广泛应用于大型桥梁工程，其下沉施工的关键在于控制下沉状态的安全平稳，准确预测取土施工过程中沉井基础的下沉速率与倾斜程度对下沉控制至关重要。沉井基础在下沉过程中获取了大量的刃脚土压力实时监测数据，刃脚土压力实测数据的数据维度较高，与沉井下沉速率和倾斜程度的作用机制复杂，采用传统方法难以处理，故采用机器学习中的极限树算法建立下沉状态预测模型，模型可提取刃脚土压力监测数据的时间和空间特征，捕捉刃脚土压力与沉井基础下沉速率和倾斜程度之间的复杂关系，智能化预测沉井的下沉速率和倾斜程度，并将预测模型应用于张靖皋长江大桥北锚碇沉井基础工程中，计算模型评估参数、验证模型预测精度。另外，将所选取的极限树算法与其他常用机器学习算法进行对比，并分析模型参数对所提出模型预测精度的影响规律。结果表明：所建立的分析模型预测精度高，工程应用时决定系数 R^2 均大于 0.9，可以满足工程需要，且极限树算法预测效果优于其他机器学习方法，基决策树个数的增加与最大树深度的增大有助于提高模型的预测精度。研究结果可为类似沉井工程控制下沉速率及倾斜程度提供参考。

关 键 词：沉井基础；下沉速率预测；倾斜程度预测；刃脚土压力；机器学习；极限树；张靖皋长江大桥

中图分类号：U446.2 **文献标识码：**A **文章编号：**1000－7598 (2025) 增刊－0377－12

Intelligent prediction of sinking of super-large anchorage caisson foundation based on soil pressure at cutting edges

CAI Qi-hang[1,2], DONG Xue-chao[2,3], GUO Ming-wei[2,3], LU Zheng[2,3], XU An[4], JIANG Fan[4]

(1. School of Earth Science and Engineering, North China University of Water Resources and Electric Power, Zhengzhou, Henan 450046, China; 2. State Key Laboratory of Geomechanics and Geotechnical Engineering safty, Institute of Rock and Soil Mechanics, Chinese Academy of Sciences, Wuhan, Hubei 430071, China; 3. University of Chinese Academy of Sciences, Beijing 100049, China; 4. China Railway Bridge and Tunnel Technology Co. Ltd., Nanjing, Jiangsu 210061, China)

Abstract: The caisson foundation is widely used in large bridge projects due to its high overall stiffness and strong bearing capacity. The key to successful sinking construction is to maintain a safe and stable sinking process. Accurately predicting the sinking rate and tilt degree of the caisson foundation during soil excavation is crucial for effective sinking control. During the sinking process, extensive real-time monitoring data of soil pressure at the cutting edge are collected. These data exhibit high dimensionality, and the underlying mechanisms linking soil pressure to sinking rate and tilt degree are complex, posing challenges for traditional analytical methods. Therefore, the extra trees algorithm from machine learning is employed to establish a sinking state prediction model. This model extracts temporal and spatial features from the monitoring data, captures the complex relationships between cutting edge soil pressure and sinking behavior, and enables intelligent prediction of sinking rate and tilt degree. The model was applied to the northern anchor caisson project of the Zhangjinggao Yangtze River Bridge, with model evaluation parameters calculated to verify prediction accuracy. Additionally, the extra trees algorithm was compared with other common machine learning methods, and the influence of model parameters on prediction accuracy was analyzed. Results show that the established model achieves high prediction accuracy, with R^2 values consistently greater than 0.9 in engineering applications, meeting project requirements. The extra trees algorithm outperforms other machine learning methods, and prediction accuracy improves with an increased number of individual decision trees and greater maximum tree depth. These findings provide valuable reference for controlling sinking rates and tilt degrees in similar caisson foundation projects.

Keywords: caisson foundation; sinking rate prediction; tilt degree prediction; soil pressure under cutting edges; machine learning; extra tree; Zhang-Jing-Gao Yangtze River Bridge

收稿日期：2024-10-22 录用日期：2025-01-24

基金项目：国家自然科学基金面上项目（No.42077262）。

This work was supported by the General Program of National Natural Science Foundation of China (42077262).

第一作者简介：蔡启航，男，2000 年生，硕士研究生，主要从事计算岩土力学方面的研究工作。E-mail: z202210020234@stu.ncwu.edu.cn

通讯作者：郭明伟，男，1981 年生，博士，副研究员，硕士生导师，主要从事岩土稳定性分析方面的研究工作。E-mail: mwguo@whrsm.ac.cn

1 引 言

沉井是一种常见的深基础类型[1]，因其具有埋深较大、无深度约束、抗震性能好、占地面积小、承载能力强和稳定性好等优点而被广泛应用于各种大型桥梁工程中[2-3]。沉井施工的重点是保证沉井顺利平稳的下沉，下沉速率与倾斜程度是反映沉井下沉状态和控制下沉的关键指标，沉井是否平稳下沉可通过实时监测下沉速率与倾斜程度进行判断[4]。在下沉过程中，下沉速率过大可能会导致突沉、沉井结构破坏等施工风险，倾斜程度过大可能会导致沉井颠覆风险[5-6]。因此，为提高沉井下沉的安全性与稳定性、有效预防施工过程中的潜在风险，预测和分析沉井下沉的下沉速率与倾斜程度具有重要意义[7-8]。

许多学者通过反演分析、构建模型和推导方程等多种方法对沉井下沉速率和倾斜情况展开了研究[9]：马振兴等[10]通过对沉井下沉过程中下沉速率与沉井倾斜角度的变化规律进行监测，将沉井下沉过程中的各个阶段分开进行对比，发现在下沉各阶段沉井基本展示出一类阻尼振荡衰减演变特征。管政霖等[11]通过对常泰长江大桥 5 号墩沉井基础下沉阶段实测数据反演分析，建立了下沉系数与下沉速率的关系，验证了通过下沉系数对下沉速率进行控制的可行性；陈保国等[12]通过建立力学模型和推导运动学方程对沉井下沉的运动学特性进行研究，发现了影响沉井下沉的重要因素；Li 等[13]提出了经验公式来预测矩形基础穿透引起的沉井最大倾角，可为沉井的设计和施工提供有用的指导；刘锦军[14]使用主动控沉方案对南通市李港水厂沉井基础初沉阶段下沉井的下沉姿态进行监测与分析，发现沉井在初沉阶段下易发生下沉困难、倾斜等情况，对类似的沉井工程具有参考意义。现有研究多关注于下沉速率与倾斜角度的特点与控制，虽然部分研究涉及沉井下沉预测，但是研究仍不全面，难以应用于实际工程中，因此，结合大型沉井实际工程，准确预测沉井基础下沉状态具有重要意义。

沉井下沉阻力的分布能够反映沉井的整体受力情况，刃脚土压力作为沉井下沉阻力的重要特征，可以反映沉井底部的受力状态，是沉井受力以及沉井结构与土体相互作用的综合表现，对沉井下沉的分析与预测具有重要意义[15]。研究表明，当取土方案[16]、地质条件[17]、下沉深度[3]等与沉井下沉姿态密切相关的物理因素发生变化时，将直接体现于刃脚土压力特征的变化，而下沉速率和倾斜程度变化伴随着刃脚土压力特征的变化，因此可以通过沉井基础刃脚土压力的特征对下沉速率和倾斜程度进行预测。

机器学习方法作为人工智能核心，具有速度快、精度高等优点，多应用于大数据的分类与回归，在工程领域具有重要作用[18]。李振阳等[19]使用粒子群优化算法（particle swarm optimization，简称 PSO）构建了 PSO-XGBoost（极端梯度提升 extreme gradient boosting，简称 XGBoost）预测模型，基于露天铜钴矿的实测数据预测爆破振动的质点峰值速度（peak particle velocity，简称 PPV），并与传统 XGBoost 预测模型进行对比，发现该模型准确预测了矿山爆破时的振动峰值质点速度，且预测精度最优；Thangam 等[20]提出了一种将虚拟蜜罐技术与机器学习相结合、具有动态和自适应功能的系统，可以实现实时威胁检测，并随着时间的推移不断适应、学习和改进；伍圣超等[21]提出了基于随机森林算法的孔压静力触探（piezocone penetration testing，简称 CPTU）土类识别模型，准确识别土类并结合多种算法进行预测分析，其中随机森林算法性能最优，预测精度较归一化土壤行为类型指数（the normalized soil behaviour type，简称 SBTn）图表法有显著提升；基于 325 组岩爆案例数据集，汤志立等[22]基于机器学习算法建立了岩爆预测模型，将预测模型应用于西藏多雄拉隧道工程中进行验证，结果表明，模型准确率较高，预测结果与实际情况基本一致；龙潇等[23]针对常规液化判别方法和现有机器学习方法的不足建立了一种基于卷积神经网络的预测模型；蒋明镜等[24]根据结构性土破损演化规律传统方法的局限性提出了一种基于符号回归的损伤规律预测模型，并结合不同损伤数据库对模型预测精度进行了验证。

由于下沉过程中沉井刃脚土压力数据维度大、与沉井下沉速率以及倾斜程度的内在关系复杂，传统方法无法进行下沉状态的预测分析，而机器学习方法非常适用于分析大维度数据之间的复杂关系，因此，本文利用机器学习方法中效果优越的极限树算法建立基于刃脚土压力监测数据的沉井下沉速率和倾斜程度预测模型，依托张靖皋长桥大桥北锚碇沉井基础工程，采用 5 折交叉验证方法对下沉速率与横桥向高差、顺桥向高差进行合理预测，并结合实测数据验证了模型预测效果。对比分析了极限树算法与随机森林算法、随机梯度下降与贝叶斯岭回归等算法的预测精度，验证了本文预测模型的实用性，对模型参数的影响程度进行了研究与分析。本

文实现了对沉井下沉速率与倾斜程度的预测，可为控制类似沉井工程的下沉提供参考。

2　智能预测算法

智能预测算法是一种用于预测未来值的方法，算法核心是机器学习算法，包括线性回归、梯度提升机、决策树和极限树等。张靖皋长江大桥北锚碇沉井基础在下沉过程中获取了大量的刃脚土压力实时监测数据，刃脚土压力实测数据的数据维度较高，与沉井下沉速率和倾斜程度的作用机制复杂，且在监测期间由于环境因素等受到了噪声的影响，采用传统方法难以处理。极限树算法是一种以决策树为基本单元的集成学习算法，具有训练速度快、泛化能力出色、鲁棒性好、预测精度高等优点，能够提取刃脚土压力的时间和空间特征，捕捉刃脚土压力与沉井基础下沉速率和倾斜程度之间的复杂关系，适用于处理含有噪声的大规模多维度数据集，故选取极限树智能预测算法构建模型预测沉井下沉速率与倾斜程度。限于篇幅，本文只简要介绍极限树算法。

2.1　集成学习

集成学习是目前最常用的智能预测算法之一，采用构建与结合多个学习器的方法来完成学习任务，可以通过多个模型训练取平均值的方法减少偏差，一些模型的错误也可以被其他模型纠正，因此，相较于单一学习器，集成学习的泛化能力与鲁棒性更优越[25]。而且集成学习具有多样性，在集成中既可以选择同种类型的个体学习器（“同质”集成），也可以选择不同类型的个体学习器（“异质”集成），“同质”集成下个体学习器被称为“基学习器”，而“异质”集成下的个体学习器则被称为“组件学习器”。这种多样性使得集成模型可以捕捉更多的数据特征，从而能够在未见过的数据上有着更好的表现。

2.2　决策树

决策树是集成学习算法中最常用的基学习器类型，是一种基于树形结构进行预测和决策的常见机器学习算法，可应用于分类与回归任务[26]。决策树的基本组成部分包括根节点、内部节点和叶节点。根节点是决策树的起始点，包含整个数据集；叶节点对应于决策结果，在分类树中是类别，在回归树中是预测值；内部节点用于分割数据，代表一个特征或属性判断条件。每一个判定序列都与从根节点到叶节点的一个路径相对应。

决策树常见的特征准则包含信息增益、增益率和基尼指数，其生成是一个递归过程，通过连续的分裂构建一颗树，其中每次分裂都是通过选择最优的特征和分支阈值进行确定[27]。该过程便于理解，可以通过树状图直观地展示出来。

2.3　极限树

极限树是集成学习算法的一种，采用决策树作为基学习算法。极限树中每棵决策树都是一个独立的学习器（个体学习器）。当保持个体学习器的数量不变时，集成学习的准确率随个体学习器之间的独立性增强而提高[28]。相较于传统的决策树算法，极限树在每棵决策树构建时会使用训练集中全部的样本数据，且在传统决策树算法的基础上引入了随机性[29]。极限树会随机选择特征和分割点，即无需对特征值进行排序，也无需在各个节点上寻找最优分割点，从而减少了计算量，提高了模型的训练速度，同时，因为单个决策树的随机性有利于减少个别错误的影响，使得极限树具有一定的鲁棒性。

3　工程实例分析

3.1　工程概况

张靖皋长江大桥位于长江下游澄通河段如皋沙群段，距离上游江阴大桥 28 km，距离下游沪苏大桥 16 km，线路由 3 个部分组成，包括跨江段大桥、南接线、北接线，总长度约为 29.849 km，是一座世界上最大跨度的悬索桥（见图 1）。张靖皋长江大桥北航道桥北锚碇沉井共分 11 节，其中第 1 节是钢壳混凝土结构，平面尺寸为 75.4 m×70.4 m，第 2 节～11 节是钢筋混凝土结构，平面尺寸均为 75.0 m×70.0 m，是平面尺寸仅次于五峰山长江大桥锚碇沉井的世界第二大陆地沉井。沉井形状为矩形，高为 57.0 m，底部设计标高为−53.5 m，中间设置 36 个标准尺寸为 10.5 m×9.7 m 的矩形井孔，具体结构如图 2 所示。沉井基础所在区域地形平缓，地面标高为+3.5 m，表层覆有杂填土，中层主要由粉质黏土、粒径不同的粉砂和粉土组成，下层主要由中砂组成，该地区工程地质具体情况如表 1 所示。

3.2　数据获取

沉井倾斜程度可由横桥向高差与顺桥向高差来确定。在沉井下沉期间，通过布置传感器对横桥向高差、顺桥向高差与刃脚土压力进行实时监测。如图 3 所示，沉井底部布置了 60 个刃脚土压力传感器，其中外井壁底部布置了 24 个传感器，隔墙底部布置了 36 个传感器，每 30 min 测量一次数据。沉井顶部布置了 4 个 GPS 传感器，每 1 min 测量一次数据，其中沉井横桥向高差由 1 号和 3 号传感器的竖向高差进行衡量，高差为正代表下游侧高于上游

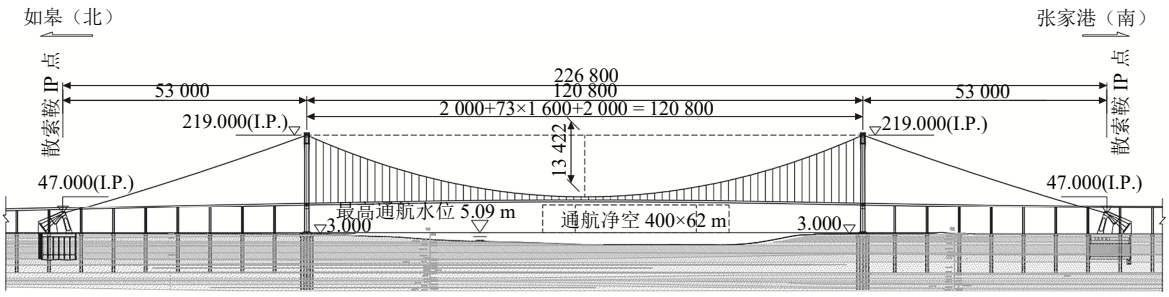
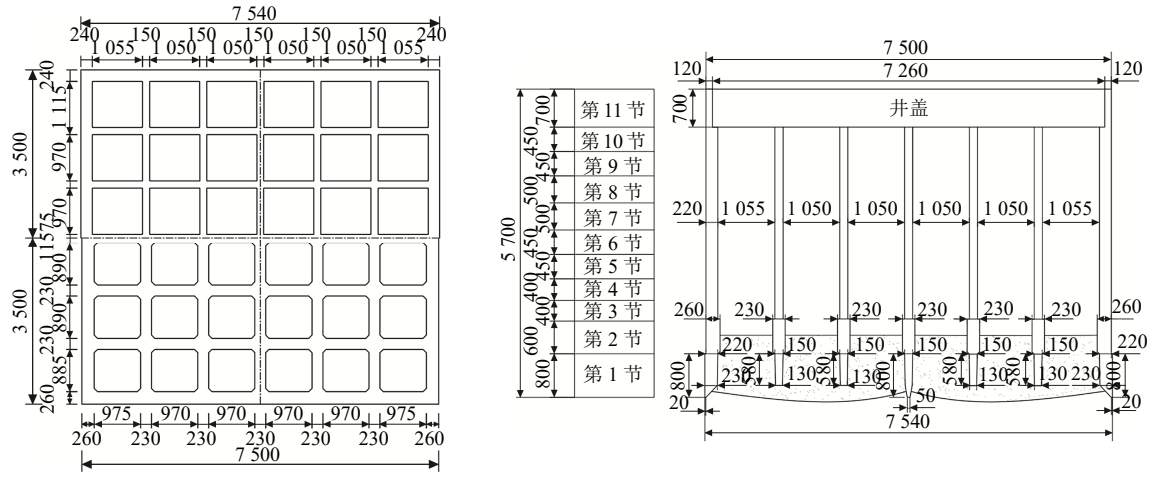


图 1 北航道桥立面布置（尺寸单位：cm，高程单位：m）

Fig.1 Elevation layout of the north channel bridge (size unit: cm, elevation unit: m)



(a) 平面图 (b) 立面图

图 2 沉井结构图（单位：cm）

Fig.2 Structure of caisson foundation (unit: cm)

表 1 土体分层情况

Table 1 Soil layers

序号	土层	层顶高程/m	层底高程/m	厚度/m
1	换填土	3.50	−1.00	4.50
2	淤泥质粉质黏土	−1.00	−8.10	7.10
3	粉砂+粉土	−8.10	−16.98	8.88
4	粉砂①	−16.98	−24.72	7.74
5	粉砂②	−24.72	−41.31	16.59
6	粉质黏土+粉土	−41.31	−47.70	6.39
7	粉砂③	−47.70	−62.99	15.29
8	中砂	−62.99	−84.04	21.05

侧，顺桥向高差由 2 号和 4 号传感器的竖向高差进行衡量，高差为正代表张家港侧高于如皋侧。

4 沉井基础下沉状态预测模型

4.1 数据预处理

数据预处理是数据分析中的一个关键步骤，包括数据清洗和数据滤波，若输入数据与输出数据的获取频率不同，则还需在处理之前进行数据重采样，即使用相同的时间间隔 t 重新获取输入数据与输出数据，输入数据和输出数据的数量应保持一致，将经过处理后的输入数据与输出数据一一对应，生成数据样本集。

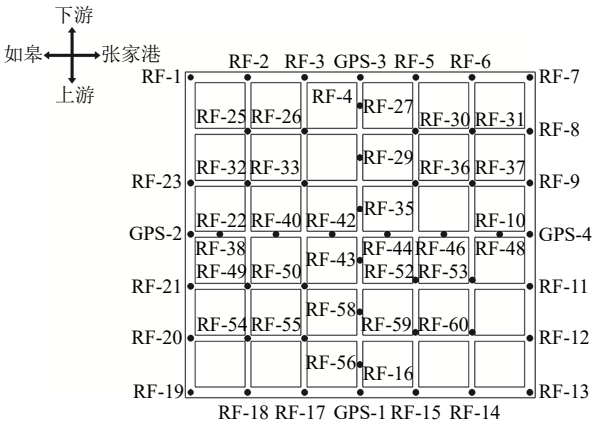


图 3 沉井传感器布置

Fig.3 Sensor layout in the open caisson

基于极限树算法构建沉井基础下沉速率和倾斜程度预测模型，在极限树模型中，输入的数据是模型用来做出预测的数据，刃脚土压力的变化与影

响沉井基础下沉速率和倾斜程度的取土方法[16]、土层条件[17]、下沉深度[3]等因素密切相关，当刃脚土压力出现明显变化时，沉井下沉速率和倾斜程度也会发生改变，沉井底部所安置的刃脚土压力传感器相互独立运行，互不影响，因此选择提取刃脚土压力的特征作为模型的输入层；输出是模型做出的预测结果，沉井的倾斜程度可由横桥向高差与顺桥向高差进行确定，因此选择沉井基础下沉速率和横桥向高差、顺桥向高差作为模型的输出层。

根据沉井下沉过程中实时测量的原始数据，传感器测量输入数据与输出数据的频率不同，选择30 min为测量时间间隔，重新采集数据样本。因横桥向高差、顺桥向高差和下沉速率的测量时间间隔为30 min，故仅需要重新采集刃脚土压力的监测数据，最终得到的数据应确保每一个测量时刻都有一组刃脚土压力数据对应着1个下沉速率数据和2个高差数据。

在上述工作完成后，将处理后的输入数据和输出数据进行匹配，组成数据样本。在匹配时应满足相应的匹配规则，每个数据样本应都有一组输入数据和对应的输出数据[9]，输出数据应比输入数据晚一个时间间隔，匹配结束后即可得到完整的数据样本。

4.1.1 数据清洗

由机器学习算法所构建模型的预测效果会受到输入数据及输出数据的影响，而在实际工程中，传感器所测量的数据可能会因传感器性能或外部影响而导致出现数据缺失或异常值[30]，因此在训练模型前对原始数据进行数据清洗是必要的。设置缺失比例的值为 p，对每个传感器所测量数据的缺失值依次进行处理：缺失比例大于 p，删除该传感器所测量的全部数据；缺失比例小于 p，采用线性插值方法补全该传感器的缺失数据[9]。异常值通常指与其他数据差异过大的值，在构建模型进行预测时会对预测效果产生不良影响。异常值可使用数据分布方法进行筛选，分别计算数据样本中每个特征的均值μ和标准差σ，认定在$\mu+3\sigma$与$\mu-3\sigma$之间的监测值为正常值，否则为异常值，将数据样本中的异常值全部去除，并使用线性插值方法进行补全。

设置缺失比例为20%，对输入数据与输出数据进行缺失值处理。沉井底部所设置的60个刃脚土压力传感器中，缺失比例小于20%的传感器有51个，将缺失值按照线性插值方法进行补全，剩下9个传感器的数据全部删除，下沉速率与横桥向高差、顺桥向高差监测数据的缺失比例均小于20%，故只需要将缺失值也采取线性插值方法补全。经过筛选后，将输入数据和输出数据中存在的异常值去除，并使用线性插值方法补全。

4.1.2 数据滤波

数据滤波是指从数据中去除噪声和异常值改善数据质量的过程。沉井在取土下沉过程中，传感器所监测的实时数据可能会因为环境因素、姿态改变、人工因素等影响而存在噪声，它们可能会干扰到预测模型的性能，因此为提高模型的预测效果，在数据样本进行清洗后需进行降噪处理。数据滤波的方法包括均值滤波、高斯滤波与中值滤波等方法，选取滑动均值算法作为本次数据的降噪方法，具体计算方法为

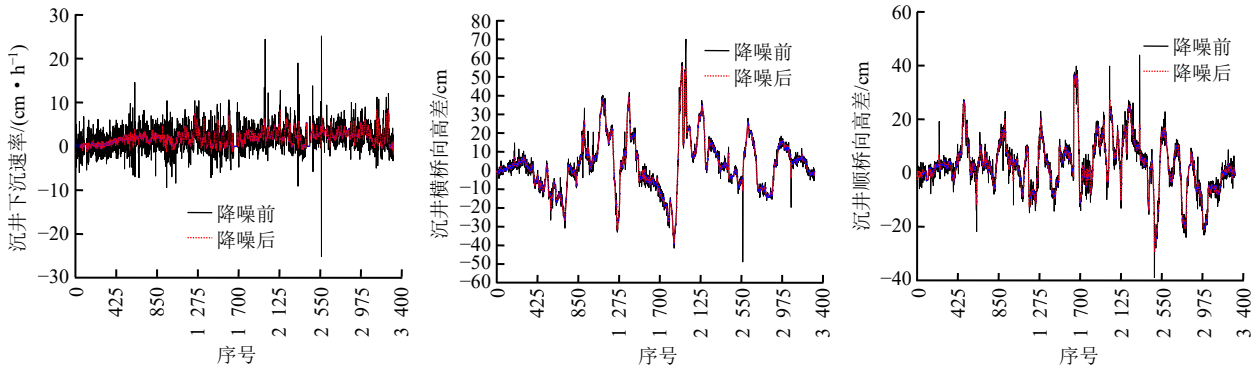
$$h_i^{\text{out}}=\frac{1}{N}\sum_{i=-m}^{m}h_i^{\text{in}}(i+j) \tag{1}$$

式中：h_i^{out}为滤波后的数据；h_i^{in}为滤波前的数据；m为滑动窗口的半径；N为滑动窗口的大小，通常$N=2m+1$；i为数据序号；$j=\{-(N-1)/2\ldots,-1,0,1,\ldots,(N-1)/2\}$。

根据沉井下沉过程中的实际监测数据，刃脚土压力监测值的数据质量较好，不需要进行数据滤波处理，而下沉速率与横桥向高差、顺桥向高差的数据质量较差，数据受到的干扰较为严重，需要进行滤波处理。设置滑动均值窗口为7[9]，数据滤波处理效果如图4所示，图中曲线分别展示了下沉速率与横桥向高差、顺桥向高差在降噪前后的数据，图中横坐标“序号”是对数据监测时间的排序，相较降噪前，降噪后下沉速率和横桥向高差、顺桥向高差的数据曲线均更加平稳。

4.2 模型训练与测试

得到全部的数据样本后，采用极限树算法建立沉井下沉速率和倾斜程度预测模型，然后使用K折交叉验证进行测试。K折交叉验证是一种模型评估方法，在使用前需要将数据样本划分为训练集和测试集，分别用于训练模型和评估模型性能，其具体步骤为：将整个数据集分成K个子集，每个子集所包含数据的数量基本一致；选择1个子集作为模型的测试集，其余的子集作为模型的训练集；预测模型需要进行K轮训练，每个子集都需要作为一次测试集；训练完成后输入测试集对下沉速率与横桥向高差、顺桥向高差进行预测，将预测结果与实测数据进行对比，选择并计算合适的评估性能指标；最后记录每轮训练的评估性能指标并计算评估指标的均值，以此作为模型性能最终评估的结果。



(a) 下沉速率 (b) 横桥向 (c) 顺桥向

图 4 沉井下沉速率与高差滤波效果

Fig.4 Filtering effects on sinking rate and height difference during open caisson sinking

4.3 模型评价指标

采用拟合优度 R^2 和均方误差 MSE（mean squared error）作为模型评价指标，R^2 能够衡量回归模型对数据的拟合程度，一般在 0～1 之间，R^2 越接近于 1，表明模型的预测精度越高，MSE 能够反映模型预测值与实际值之间的差异，MSE 越小，表明模型的预测精度越高。评估指标具体的计算公式为

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-y_i'\right)^2 \tag{2}$$

$$R^2=1-\frac{\sum_{i=1}^{n}\left(y_i'-y_i\right)^2}{\sum_{i=1}^{n}\left(y_i-\overline{y_i}\right)^2} \tag{3}$$

式中：y_i 为第 i 个样本的实际值；y_i' 为第 i 个样本的预测值；$\overline{y_i}$ 为 n 个样本实际值的平均值；n 为样本个数。

4.4 模型参数设置

设置 $K=5$，即采用 5 折交叉验证方法（见图 5），预测模型共需要进行 5 轮训练与测试。模型训练完成后，获得横桥向高差、顺桥向高差两个高差的预测值与下沉速率的预测值，然后利用决定系数 R^2 和均方误差 MSE 分别对 3 个预测值的预测精度进行评估，每一折横桥向高差、顺桥向高差与下沉速率均会得到两个模型评估指标，计算 5 轮训练与测试中所计算模型评估指标的均值，将其作为最终评估结果。

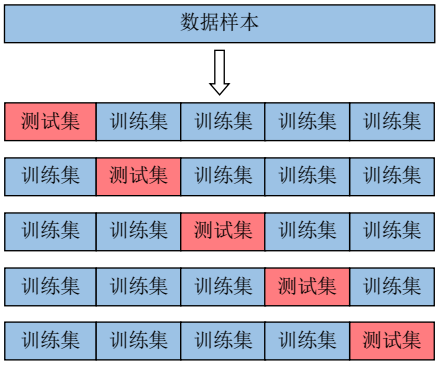


图 5 5 折交叉验证示意图

Fig.5 Schematic diagram of 5-fold cross validation

基于极限树算法所构建的预测模型参数设置，如表 2 所示，设置模型的基学习器个数为 200 个，最大树深度为“None”，即该模型的最大树深度不受限制，其余参数都按算法默认值设置。

表 2 模型参数设置

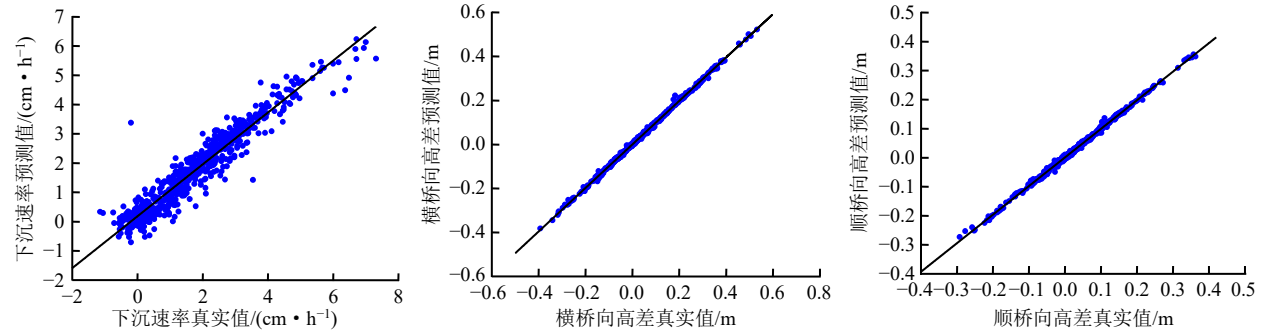
Table 2 Model parameter setting

模型名称	最大树深度	分裂所需最小样本数/个	叶节点最小样本数/个	基学习器个数/个
基于极限树预测模型	None	2	1	200

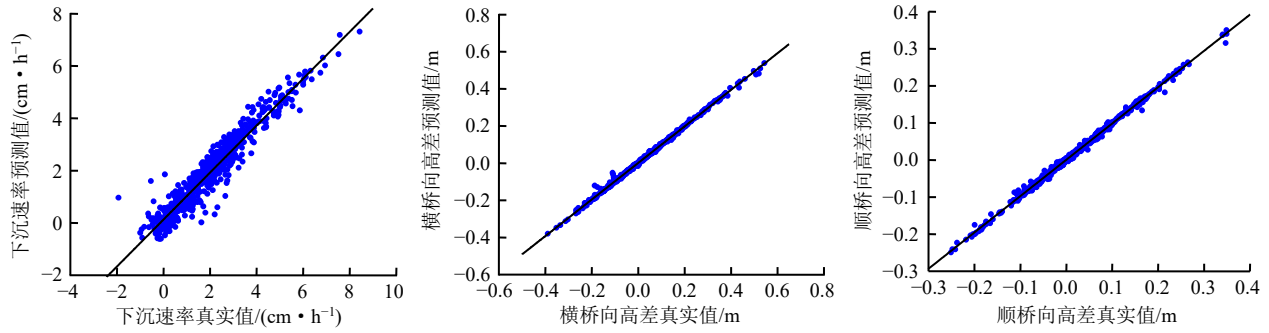
5 预测结果分析

5.1 预测结果

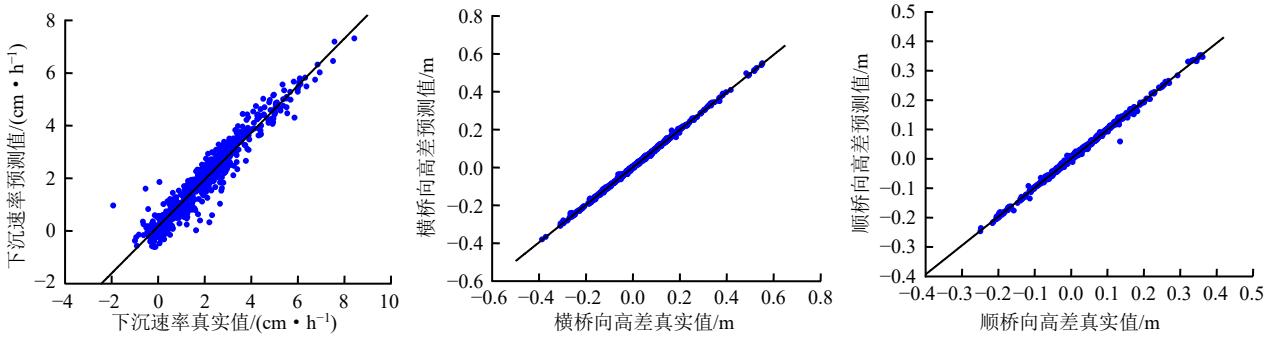
根据预处理后的北锚碇沉井下沉监测数据，采用极限树算法构建沉井下沉速率及倾斜程度预测模型，使用 5 折交叉验证方法对预测模型进行训练，每一轮训练结束后，模型输出当前预测的下沉速率、横桥向高差与顺桥向高差，为直观展示模型的预测效果，将预测值与实际值进行对比并绘制散点图。将全部的对比散点图排列成 5 行 3 列的形式（见图 6），每一行表示 5 折交叉验证一轮训练与测试的结果，第 1 列为沉井下沉速率的对比散点图，第 2 列、第 3 列分别为沉井横桥向高差、顺桥向高差的对比散点图。预测值与实际值越接近，则图上的数据散点与虚线越接近，观察图 6 可知，图中数据散点基本落在虚线附近，仅有少数散点分布在距虚线



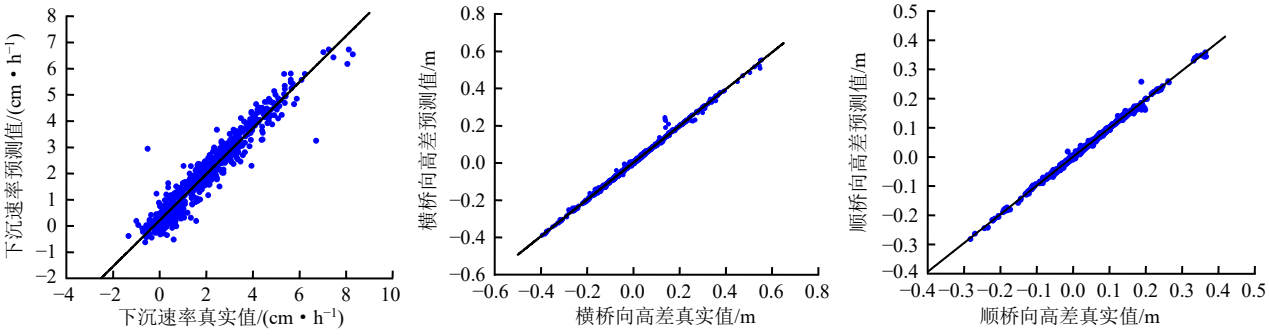
(a) 第 1 轮



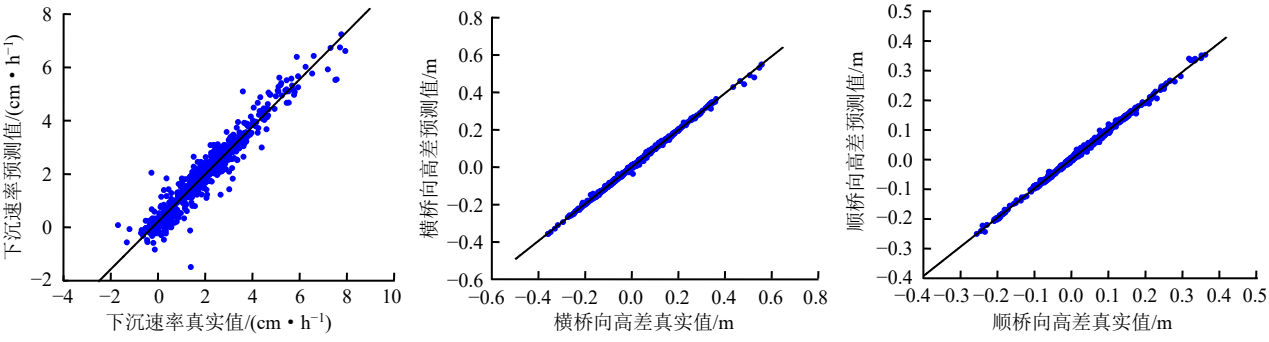
(b) 第 2 轮



(c) 第 3 轮



(d) 第 4 轮



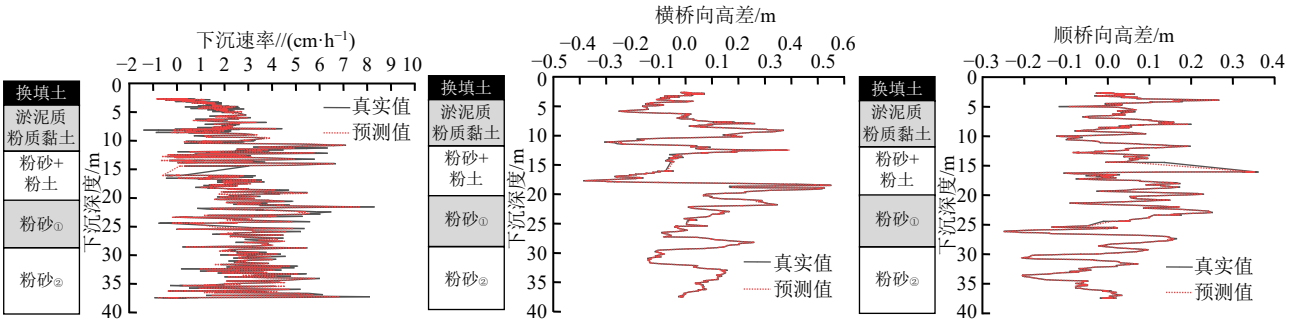
(e) 第 5 轮

图 6　预测值与实际值对比

Fig.6　Comparison between predicted and actual values

较远的地方，表明下沉速率与横桥向高差、顺桥向高差的预测值和实际值相近，基于极限树算法所够构建的模型预测效果较好。

绘制沉井在下沉过程中沿土层分布真实值与预测值的变化曲线图。如图 7 所示，左侧图例为沉井基础所在区域的土层分布情况，与右侧 y 轴下沉深度相对应，右侧为沉井在下沉过程中沿土层分布真实值与预测值的变化曲线。根据变化曲线可知，沉井在下沉过程中下沉速度和下沉姿态均在允许范围内，未出现突沉、滞沉或倾斜过大的现象，且真实值与预测值的变化曲线基本重合，故模型的预测精度高，在该工程上具有良好的应用效果。



(a) 下沉速率 (b) 横桥向高差 (c) 顺桥向高差

图 7 沉井下沉过程中真实值与预测值的变化曲线

Fig.7 Variations of actual vs. predicted values during caisson sinking

计算模型在 5 折交叉验证方法中每一轮次的模型评价指标，然后计算 5 轮训练与测试中决定系数和均方误差的均值。如表 3 所示，模型在 5 折交叉验证每一轮所计算的模型评价指标：下沉速率预测结果的决定系数均大于 0.9，横桥向高差与顺桥向高差预测结果的决定系数 R^2 均大于 0.99；下沉速率预测结果的均方误差 MSE 均在 0.2 左右，横桥向高差与顺桥向高差预测结果的均方误差 MSE 均小于 0.1。结合决定系数 R^2 和均方误差 MSE 的均值可知，极限树算法预测模型对横桥向高差与顺桥向高差之间的预测精度相近，二者差异较小，虽然对下沉速率的预测精度低于对横桥向高差与顺桥向高差的预测精度，但是 3 个参数预测精度均可以满足工程需求。

表 3 极限树算法模型预测精度

Table 3 Prediction accuracy of the extra trees model

轮次	下沉速率		横桥向		顺桥向	
	MSE	R^2	MSE/10^{-5}	R^2	MSE/10^{-5}	R^2
1	0.190 7	0.917 9	2.641	0.998 7	2.208	0.997 4
2	0.192 7	0.924 9	4.798	0.997 7	2.854	0.996 8
3	0.182 7	0.932 4	2.453	0.998 9	3.142	0.996 8
4	0.205 3	0.923 0	8.881	0.996 0	3.446	0.996 2
5	0.189 4	0.924 7	3.370	0.998 3	2.723	0.997 0
均值	0.192 2	0.924 6	4.429	0.997 9	2.875	0.996 8

5.2 预测精度分析

箱形图是一种形状与箱子相似的图表，能够显示一组数据的分布情况，因箱形图法所采用的实际数据不受数据偏态的影响，故能够更加直观地显示出数据的实际情况[31]。箱形图由箱体、须线和异常值组成：箱体通常指第 1 四分位数 Q_1 与第 3 四分位数 Q_3 之间，第 1 四分位数表示数据样本中 25%的数据小于该值，第 3 四分位数表示数据样本中 75%的数据小于该值，故箱体表示为数据中间 50%的范围；须线是指从箱体的两端延伸出去的线，分别表示数据样本的最大值和最小值，统计最大值和最小值时通常会排除异常值，须线的长度反映了数据离散程度；异常值指位于须线之外的数值。

为了进一步验证模型的预测精度，设置随机种子 1～50，采用 5 折交叉验证方法，对模型进行 250 轮不重复的训练与测试，计算每一轮的模型评估指标并绘制相应的箱形图，由于预测模型对下沉速率的预测精度较低，而对两个高差的预测精度相近，故采用双坐标系可视化方案—左轴显示下沉速率计算结果，右轴同步显示横桥向与顺桥向高差数据。

如图 8 所示，模型在下沉速率方面的预测精度小于在横桥向高差、顺桥向高差方面的预测精度，但均取得了良好的预测效果。

5.3 算法对比分析

为验证基于极限树算法模型的预测效果，使用线性回归、岭回归、贝叶斯岭回归、支持向量回归、决策树算法 5 种单一机器学习算法与梯度提升决策

树、Bagging 算法、自适应提升、极端梯度提升、随机森林算法、LightGBM 算法 6 种集成学习算法以及随机梯度下降构建沉井下沉速率与倾斜程度预测模型，选取相同的数据集作为输入数据，采用 5 折交叉验证方法对沉井下沉速率与横桥向高差、顺桥向高差进行预测并计算训练与测试中的模型评估指标及其均值，计算结果见表 4。极限树算法预测模型的预测精度最优，其次是随机森林算法、LightGBM 算法与极端梯度提升，下沉速率的决定系数 R^2 在 0.90 左右，高差的决定系数 R^2 均大于 0.98；决策树算法与梯度提升决策树效果一般，虽然高差的 R^2 均大于 0.96，但是下沉速率的 R^2 仅在 0.80 左右；其他算法预测模型的预测效果较差，与上述模型预测效果具有较大差距。

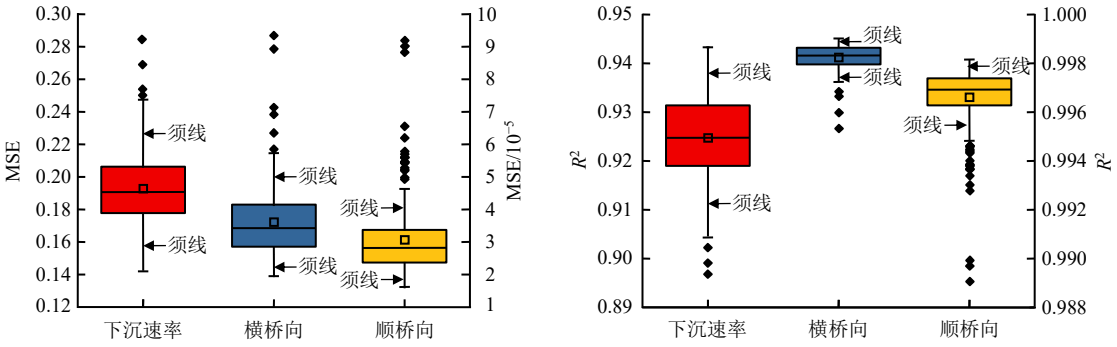


图 8　不同模型评价指标箱形图

Fig.8　Box plots of different model evaluation indicators

表 4　基于不同算法建立模型的预测精度

Table 4　Model prediction accuracies by different algorithms

算法名称	下沉速率		横桥向		顺桥向	
	MSE	R^2	MSE	R^2	MSE	R^2
极限树	0.192 2	0.924 6	$0.442\ 9\times10^{-4}$	0.997 9	$0.287\ 5\times10^{-4}$	0.996 8
线性回归	1.488 7	0.417 3	0.005 2	0.747 1	0.003 1	0.657 0
岭回归	1.513 0	0.407 9	0.005 8	0.720 4	0.003 5	0.609 5
贝叶斯岭回归	1.524 0	0.403 5	0.005 3	0.746 0	0.003 1	0.656 0
随机梯度下降	1.706 5	0.333 2	0.009 0	0.568 5	0.006 1	0.321 8
支持向量回归	1.528 6	0.401 9	0.004 6	0.779 9	0.004 1	0.540 9
决策树	0.547 3	0.785 7	0.000 7	0.967 5	0.000 3	0.967 5
Bagging 算法	1.526 0	0.403 0	0.004 6	0.779 4	0.004 1	0.540 7
随机森林	0.253 6	0.900 7	0.000 2	0.992 3	0.000 1	0.986 8
梯度提升决策树	0.470 5	0.815 7	0.000 4	0.982 4	0.000 3	0.970 0
自适应提升	1.012 3	0.603 8	0.003 8	0.815 2	0.002 1	0.765 8
极端梯度提升	0.258 4	0.898 8	0.000 2	0.989 3	0.000 1	0.983 9
LightGBM	0.234 8	0.907 9	0.000 2	0.991 0	0.000 1	0.985 0

5.4　模型参数分析

模型的预测精度可能会受到模型参数设置的影响，基决策树个数和最大树深度是影响以决策树为基学习算法的极限树算法所构建模型预测精度的最重要的参数。为探究模型参数对极限树算法模型的影响程度，通过分别更改基决策树个数与最大树深度对采用 5 折交叉验证方法对模型进行训练与测试，并计算模型评估指标均值。

在表 2 模型参数设置的基础上，将基决策树个数分别设置为 10、20、30、50、80、120、200、300、500 个，其他参数保持不变。如图 9 所示，随着基决策树个数的增加，基于极限树算法的预测模型对下沉速率与横桥向高差、顺桥向高差进行预测时，均方误差 MSE 逐渐减小，决定系数 R^2 逐渐增大；基决策树个数较少时，模型评估指标值变化幅度较大，直到基决策树个数达到一定数量，模型评估指标值变化幅度逐渐变缓直至趋于稳定；当基决策树个数达到 120 个时，模型的各个评估指标值基本达到最优，曲线拐点处于[50，120]区间。

同样在表 2 模型参数设置的基础上，设置最大树深度分别为 3、5、7、10、15、20 和“None”，其中“None”表示最大树深度不受限制，其它参数保持不变。如图 10 所示，随着最大树深度的增加，基于极限树算法的预测模型对下沉速率与横桥向高

差、顺桥向高差进行预测时，均方误差 MSE 逐渐减小，决定系数 R^2 逐渐增大；与基决策树个数相同，模型评估指标根据最大树深度的变化同样呈先快后慢的规律，指标的变化拐点处于[10，15]区间，最大树深度为“None”时，模型各个评估指标达到最优。

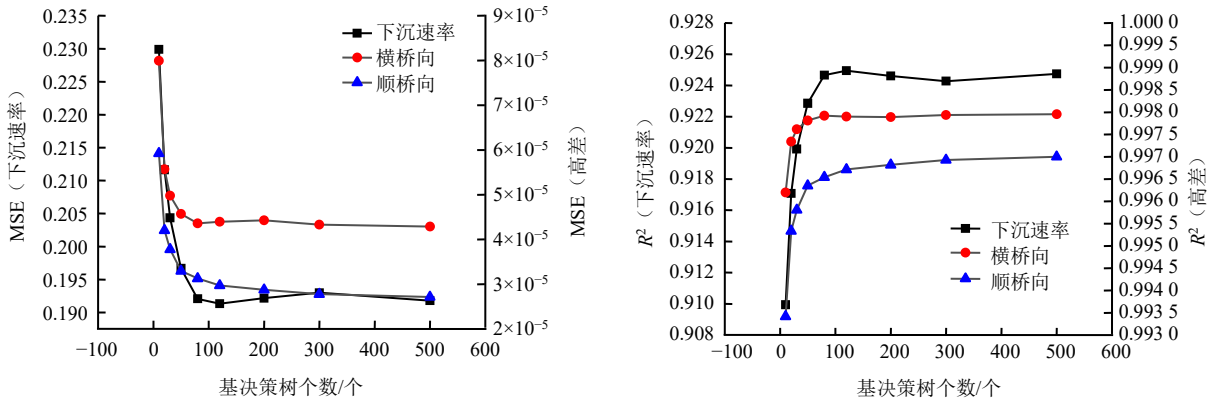


图 9 模型评价指标值受基决策树个数的影响效果

Fig.9 Impact of base decision tree count on model evaluation metrics

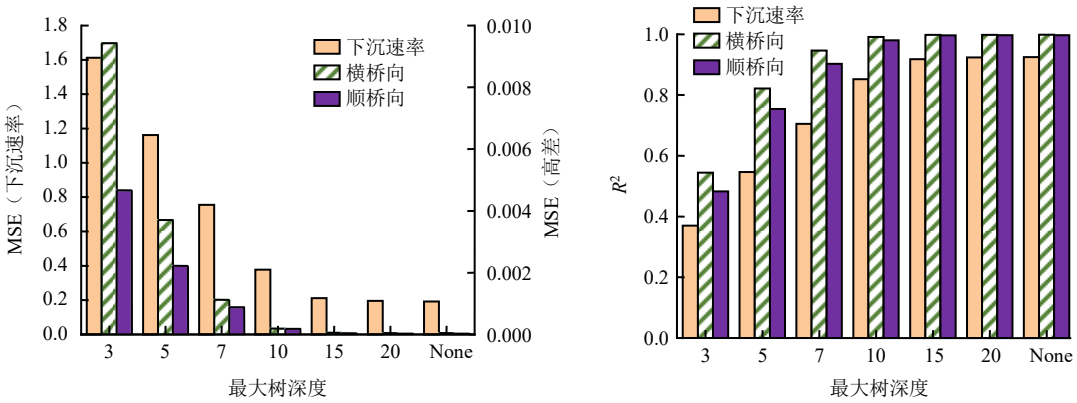


图 10 模型评价指标值受最大树深度的影响效果

Fig.10 Impact of maximum tree depth on model evaluation metrics

6 结 论

依托张靖皋长江大桥北锚碇沉井项目，获取并处理沉井下沉过程中实际监测数据，基于极限树算法构建沉井下沉速率与倾斜程度预测模型，采用刃脚土压力作为输入数据，下沉速率和倾斜程度作为输出数据，通过模型提取刃脚土压力监测值的数据特征，利用 5 折交叉验证方法进行训练与测试，对沉井基础下沉速率和倾斜程度进行预测，并将预测值与实测值进行对比，计算模型评估指标均值，验证模型的预测效果。对比分析极限树算法模型与其他算法模型的预测精度并探究模型重要参数对预测效果的影响，主要结论如下：

（1）结合沉井下沉过程中刃脚土压力实测数据，极限树算法预测模型可以合理预测沉井下沉速率与横桥向高差、顺桥向高差，可为类似沉井工程控制下沉速率及倾斜程度提供参考。

（2）张靖皋长江大桥北锚碇沉井基础工程下沉状态预测结果表明：沉井下沉速率与高差的预测值与实际值相近，3 个参数预测精度均可以满足工程需求；下沉速率预测的决定系数 R^2 均大于 0.90，均方误差 MSE 均在 0.2 左右；横桥向高差与顺桥向高差预测的决定系数 R^2 均大于 0.99，均方误差 MSE 均小于 0.1；对横桥向高差与顺桥向高差的预测精度高于对下沉速率的预测精度。

（3）本文提出的模型预测效果好，实用性强。极限树算法预测模型的预测精度最优，其次是随机森林算法、LightGBM 算法与极端梯度提升，下沉速率的决定系数 R^2 在 0.90 左右，高差的决定系数 R^2 均大于 0.98；决策树算法与梯度提升决策树效果一般，虽然高差的 R^2 均大于 0.96，但是下沉速率的 R^2 仅在 0.80 左右；其他算法预测模型的预测效果较差，与上述模型预测效果具有较大差距。

（4）保持其他参数不变，模型预测精度会随着

基决策树个数与最大树深度的增加而提高，当基决策树个数达到 120 个时，最大树深度为“None”时，模型各个评估指标达到最优。

参 考 文 献

[1] 白明智, 莫祥昆, 乔鑫, 等. 我国桥梁基础的主要形式及其特点分析[J]. 科学技术创新, 2012(34): 280.
BAI Ming-zhi, MO Xiang-kun, QIAO Xin, et al. Analysis of the main forms and characteristics of bridge foundations in China[J]. **Scientific and Technological Innovation**, 2012(34): 280.

[2] 郭明伟, 董学超, 沈孔健, 等. 超大沉井基础取土下沉端阻力变化规律研究[J]. 岩石力学与工程学报, 2021, 40(增刊 1): 2976-2985.
GUO Ming-wei, DONG Xue-chao, SHEN Kong-jian, et al. Study on the variation of the bottom resistance during sinking stage of super large caisson foundation[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2021, 40(Suppl.1): 2976-2985.

[3] 周和祥, 马建林, 张凯, 等. 沉井下沉阻力离心模型试验研究[J]. 岩土力学, 2019, 40(10): 3969-3976.
ZHOU He-xiang, MA Jian-lin, ZHANG Kai, et al. Study of sinking resistance of large and deep caisson based on centrifugal model test[J]. **Rock and Soil Mechanics**, 2019, 40(10): 3969-3976.

[4] 唐庆. 自动测量技术在沉井下沉姿态测量中的研究及应用[J]. 山西建筑, 2024, 50(9): 161-166.
TANG Qing. Research and application of automatic measurement technology in open caisson sinking attitude measurement[J]. **Shanxi Architecture**, 2024, 50(9): 161-166.

[5] DONG X C, GUO M W, WANG S L. GBDT-based multivariate structural stress data analysis for predicting the sinking speed of an open caisson foundation[J]. **Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards**, 2024, 18(2): 333-345.

[6] 陈佳培. 基于 MJS 工法的大角度倾斜沉井纠偏处理技术研究[J]. 水利建设与管理, 2024, 44(9): 59-63.
CHEN Jia-pei. Research on correction technology for large-angle tilted caissons based on MJS method[J]. **Water Conservancy Construction and Management**, 2024, 44(9): 59-63.

[7] 董学超, 郭明伟, 王水林. 基于多元结构应力特征的沉井基础下沉速度预测[J]. 岩石力学与工程学报, 2022, 41(增刊 2): 3476-3487.
DONG Xue-chao, GUO Ming-wei, WANG Shui-lin. Sinking speed prediction of an open caisson foundation based on the characteristics of multivariate structural stress data[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2022, 41(Suppl.2): 3476-3487.

[8] LI J, CHEN S X, YU F, et al. Mechanics and deformation characteristics of an oversized inclined caisson foundation when being reused[J]. **Ocean Engineering**, 2022, 248: 110780.

[9] 董学超, 郭明伟, 王水林, 等. 基于集成学习的超大型沉井基础下沉倾斜程度预测[J]. 岩石力学与工程学报, 2023, 42(增刊 1): 3812-3822.
DONG Xue-chao, GUO Ming-wei, WANG Shui-lin, et al. Inclination prediction of a super-sized open caisson foundation during sinking process based on ensemble learning[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2023, 42(Suppl.1): 3812-3822.

[10] 马振兴, 刘建, 要康凯, 等. 沉井下沉过程倾斜度演变特征研究[J]. 黑龙江科学, 2024, 15(8): 85-87.
MA Zhen-xing, LIU Jian, YAO Kang-kai, et al. Study on slope evolution characteristics of the sinking process of caisson[J]. **Heilongjiang Science**, 2024, 15(8): 85-87.

[11] 管政霖, 郭万中, 刘修成. 常泰长江大桥 5 号墩沉井取土下沉施工技术研究[J]. 施工技术(中英文), 2024, 53(8): 66-71, 80.
GUAN Zheng-lin, GUO Wan-zhong, LIU Xiu-cheng. Soil taking and sinking construction technology of open caisson for Changtai Yangtze River Bridge No. 5 pier[J]. **Construction Technology**, 2024, 53(8): 66-71, 80.

[12] 陈保国, 贺洁星, 骆瑞萍, 等. 沉井下沉过程的运动学特性与下沉控制研究[J]. 岩土力学, 2022, 43(增刊 2): 425-430, 453.
CHEN Bao-guo, HE Jie-xing, LUO Rui-ping, et al. Kinematic characteristics and sinking control of open caisson during sinking process[J]. **Rock and Soil Mechanics**, 2022, 43(Suppl.2): 425-430, 453.

[13] LI S Y, QI Y L, ZHOU M, et al. Response of an installed suction caisson induced by rectangular footing penetration in nonhomogeneous clay[J]. **International Journal of Geomechanics**, 2024, 24(9): 04024195.

[14] 刘锦军. 邻江涉铁砂层地区超深沉井初沉阶段实测研究——以南通市李港水厂项目沉井工程为例[J]. 中国水运(下半月), 2024, 24(8): 153-155.
LIU Jin-jun. Field measurement study on the initial settlement stage of ultra-deep caisson in the iron-bearing sand layer area adjacent to the river—taking the caisson project of Ligang waterworks in Nantong city as an example[J]. **China Water Transport**, 2024, 24(8): 153-155.

[15] 蒋凡, 刘华, 岳青, 等. 超大沉井基础取土下沉刃脚土压力变化规律研究[J]. 岩土力学, 2022, 43(增刊 2): 431-442.
JIANG Fan, LIU Hua, YUE Qing, et al. Variation trend of soil pressure under cutting edges of the super large caisson during sinking stage[J]. **Rock and Soil Mechanics**, 2022, 43(Suppl.2): 431-442.

[16] 董学超, 郭明伟, 蒋振雄, 等. 基于土压力监测值的超大沉井下沉端阻力计算与分析[J]. 桥梁建设, 2022, 52(5): 78-84.
DONG Xue-chao, GUO Ming-wei, JIANG Zhen-xiong, et al. End resistance calculation and analysis of super-large open caisson during soil excavation sinking based on monitored soil pressure data[J]. **Bridge Construction**, 2022, 52(5): 78-84.

[17] 邵风行, 叶嘉庭, 池奥博, 等. 软土地区大型沉井下沉阻力分布特征研究[J]. 地下空间与工程学报, 2024, 20(增刊 1): 128-133.
SHAO Feng-xing, YE Jia-ting, CHI Ao-bo, et al. Study on distribution characteristics of sinking resistance of open caisson foundation in soft soil area[J]. **Chinese Journal of Underground Space and Engineering**, 2024, 20(Suppl.1): 128-133.

[18] 陈进帆, 尚德磊, 赵志宏, 等. 基于机器学习的含充填裂隙岩样断裂行为预测方法[J]. 岩石力学与工程学报, 2023, 42(增刊 1): 3458-3472.
CHEN Jin-fan, SHANG De-lei, ZHAO Zhi-hong, et al. Prediction method of fracture behaviors of rock samples with an infilled fracture based on machine learning[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2023, 42(Suppl.1): 3458-3472.

[19] 李振阳, 张宝岗, 熊信, 等. 基于 PSO-XGBoost 的露天矿山 PPV 预测模型研究[J]. 黄金科学技术, 2024, 32(4): 620-630.
LI Zhen-yang, ZHANG Bao-gang, XIONG Xin, et al. Research on PPV prediction model of open-pit mine based on PSO-XGBoost[J]. **Gold Science and Technology**, 2024, 32(4): 620-630.

[20] THANGAM S, CHAKKARAVARTHY S S. An edge-enabled virtual honeypot based intrusion detection system for vehicle-to-everything (V2X) security using machine learning[J]. **IAENG International Journal of Computer Science**, 2024, 51(9): 1374-1384.

[21] 伍圣超, 王睿, 张建民. 基于随机森林算法的 CPTu 土类识别模型研究及其在不同区域的应用[J]. 中南大学学报(自然科学版), 2023, 54(11): 4391-4402.
WU Sheng-chao, WANG Rui, ZHANG Jian-min. Research on CPTu-based soil classification model using random forest algorithm and its application in different regions[J]. **Journal of Central South University (Science and Technology)**, 2023, 54(11): 4391-4402.

[22] 汤志立, 徐千军. 基于 9 种机器学习算法的岩爆预测研究[J]. 岩石力学与工程学报, 2020, 39(4): 773-781.
TANG Zhi-li, XU Qian-jun. Rockburst prediction based on nine machine learning algorithms[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2020, 39(4): 773-781.

[23] 龙潇, 孙锐, 郑桐. 基于卷积神经网络的液化预测模型及可解释性分析[J]. 岩土力学, 2024, 45(9): 2741-2753.
LONG Xiao, SUN Rui, ZHENG Tong. Convolutional neural network-based liquefaction prediction model and interpretability analysis[J]. **Rock and Soil Mechanics**, 2024, 45(9): 2741-2753.

[24] 蒋明镜, 张卢丰, 韩亮, 等. 基于符号回归算法的结构性砂土损伤规律研究[J]. 岩土力学, 2024, 45(12): 3768-3778.
JIANG Ming-jing, ZHANG Lu-feng, HAN Liang, et al. Damage law of structured sand using symbolic regression algorithm[J]. **Rock and Soil Mechanics**, 2024, 45(12): 3768-3778.

[25] 李博, 张婧婧, 雷嘉诚, 等. 基于集成学习算法和 WOFOST 模型的小麦生长模拟分析与产量预测[J]. 湖北农业科学, 2024, 63(8): 85-91.
LI Bo, ZHANG Jing-jing, LEI Jia-cheng, et al. Simulation analysis and yield prediction of wheat growth based on ensemble learning algorithm and WOFOST model[J]. **Hubei Agricultural Sciences**, 2024, 63(8): 85-91.

[26] 王冠杰, 刘盛咸, 周健, 等. 材料研究中的可解释机器学习[J]. 金属学报, 2024, 60(10): 1345-1361.
WANG Guan-jie, LIU Sheng-xian, ZHOU Jian, et al. Explainable machine learning in the research of materials science[J]. **Acta Metallurgica Sinica**, 2024, 60(10): 1345-1361.

[27] PRAGATI A, GNANAPRAKASH R, H. S D. Prediction of biodiesel yield employing machine learning: interpretability analysis via Shapley additive explanations[J]. **Fuel**, 2024, 359: 130516.

[28] GEURTS P, ERNST D, WEHENKEL L. Extremely randomized trees[J]. **Machine Learning**, 2006, 63(1): 3-42.

[29] 朱子龙, 张立臣. 基于堆叠极限树集成算法的信息物理系统入侵检测方法[J]. 计算机应用与软件, 2021, 38(11): 314-321.
ZHU Zi-long, ZHANG Li-chen. Intrusion detection method of cyber-physical system based on stacking extra tree integration algorithm[J]. **Computer Applications and Software**, 2021, 38(11): 314-321.

[30] 周小雄, 龚秋明, 殷丽君, 等. 基于 BLSTM-AM 模型的 TBM 稳定段掘进参数预测[J]. 岩石力学与工程学报, 2020, 39(增刊 2): 3505-3515.
ZHOU Xiao-xiong, GONG Qiu-ming, YIN Li-jun, et al. Predicting boring parameters of TBM stable stage based on BLSTM networks combined with attention mechanism[J]. **Chinese Journal of Rock Mechanics and Engineering**, 2020, 39(Suppl.2): 3505-3515.

[31] 谢志炜, 温锐刚, 孟安波, 等. 基于箱形图和隔离森林的施工人次数据处理与预测研究[J]. 工程管理学报, 2018, 32(5): 92-96.
XIE Zhi-wei, WEN Rui-gang, MENG An-bo, et al. Data processing and forecast analysis of construction workers based on boxplot method and isolation forest[J]. **Journal of Engineering Management**, 2018, 32(5): 92-96.