

水利水运工程学报

HYDRO-SCIENCE AND ENGINEERING

水工结构变形预测模型构建与解释

胡江, 苏荟

Forecasting and analysis of deformation in hydraulic structures

HU Jiang, SU Hui

在线阅读 View online: <https://doi.org/10.12170/20230418001>

您可能感兴趣的其他文章

Articles you may be interested in

特高拱坝运行初期变形监测预报模型及构建方法

Deformation forecasting model and its modeling method of super high arch dams during initial operation periods

水利水运工程学报. 2020(5): 63 <https://doi.org/10.12170/20190908001>

湿度对薄壁渡槽受力变形的影响分析

Analysis of influences of humidity on stress deformation of thin-wall aqueduct

水利水运工程学报. 2019(4): 92 <https://doi.org/10.16198/j.cnki.1009-640X.2019.04.013>

取样卸荷对膨胀性泥岩强度与变形特性影响的试验研究

Experimental study of unloading impact on deformation behavior and strength of mudstone in sampling state

水利水运工程学报. 2017(5): 109 <https://doi.org/10.16198/j.cnki.1009-640X.2017.05.016>

框架式海堤变形模拟与稳定性影响因素分析

Deformation simulation and stability analysis of frame seawall

水利水运工程学报. 2020(5): 57 <https://doi.org/10.12170/20190806002>

基于应变评价边坡-地基失稳的过程控制

Process control of slope-foundation instability based on strain evaluation

水利水运工程学报. 2018(3): 71 <https://doi.org/10.16198/j.cnki.1009-640X.2018.03.010>

基于安全监测的水闸健康诊断体系研究

Analysis of sluice health diagnosis system based on safety monitoring

水利水运工程学报. 2018(5): 1 <https://doi.org/10.16198/j.cnki.1009-640X.2018.05.001>



扫码进入官网, 阅读更多精彩文章



关注微信公众号, 获得更多资讯信息

DOI:10.12170/20230418001

胡江, 苏荟. 水工结构变形预测模型构建与解释 [J]. 水利水运工程学报, 2024(2): 125-134. (HU Jiang, SU Hui. Forecasting and analysis of deformation in hydraulic structures[J]. Hydro-Science and Engineering, 2024(2): 125-134. (in Chinese))

水工结构变形预测模型构建与解释

胡江^{1,2}, 苏荟^{1,3}

(1. 南京水利科学研究院 水灾害防御全国重点实验室, 江苏 南京 210029; 2. 水利部大坝安全管理中心, 江苏 南京 210029; 3. 河海大学 水利水电学院, 江苏 南京 210098)

摘要: 现有基于机器学习算法的水工结构安全监控模型结果的可解释性较差。为提高安全监控模型的可解释性, 发展一种基于集成学习算法的水工结构变形预测模型构建与解释方法。简述改进统计模型及随机森林(RF)、极端梯度提升树(XGBoost)两种常用的集成学习算法, 引入沙普利值可加性解释(SHAP)方法实现集成学习算法模型结果的可解释性, 阐述 SHAP 方法的原理和推导过程。以某运行初期特高拱坝变形数据为例验证方法的有效性和实用性。结果表明, XGBoost 模型具有较高的预测精度, 预测集决定系数大于 0.982, 改进统计模型精度次之, RF 模型精度相对较低; SHAP 方法可以分离不同自变量对效应量的影响大小, 并能给出全局和局部的影响机制, 实现模型拟合和预测结果的可解释性。提出的方法综合了“机理驱动”和“数据驱动”模型的优势, 可为水工结构运行管理提供决策参考。

关键词: 水工结构; 变形; 安全监控; 机器学习; 集成学习算法; 可解释性

中图分类号: TV642

文献标志码: A

文章编号: 1009-640X(2024)02-0125-10

变形和渗流等效应量变化规律是水工结构安全性态的直观反映, 其安全监控是保障重大水工结构安全运行的重要手段。传统安全监控主要有统计、确定性和混合模型^[1]。对于处于稳定期的、自变量影响效应量机制明确的水工结构, 基于传统安全监控模型的效应量建模方法可获得较准确的拟合和预测结果, 得到了广泛应用。如将重力坝的变形分解为水压、温度和时效等自变量分量, 通过理论或数值模拟推导构建变形与自变量的关系式, 这类“机理驱动”模型得到了广泛应用^[2-3]。然而, 对于一些机理复杂、影响因素较多的水工结构的效应量, 如蓄水和运行初期特高拱坝的变形, 由于坝体温度场未稳定、时效非单调递增等影响, 导致传统安全监控模型的精度较低^[4]。

机器学习算法为解决复杂条件下安全监控模型构建提供了新思路。作为“数据驱动”方法, 机器学习算法直接挖掘数据样本中自变量(如水压、温度和时效)与效应量(如变形和渗流)间的映射关系, 从而建立精确的、鲁棒性强的效应量监控模型^[5-7]。如冷天培等^[8-9]利用贝叶斯方法优化超参数, 采用长短期记忆网络预测堆石坝和混凝土坝变形; 徐韧等^[10]引入极端梯度提升(eXtreme Gradient Boosting, XGBoost)算法, 结合基于高斯过程的贝叶斯优化方法, 构建了基于高斯过程的 XGBoost 算法的大坝变形预测模型; 胡江等^[11]构建了基于传统统计方法以及随机森林(Random Forest, RF)、最小二乘支持向量机和增强回归树等算法的高拱坝变形预测模型, 比较了模型的性能。上述应用表明了机器学习算法良好的精度和预测能力, 展现了其在水工结构效应量安全监控模型构建方面的优势, 但这些应用研究多为“黑盒子”模型, 即可以根据自变量较准确地拟合和预测效应量, 却无法给出模型结果的解释^[12-14]。这一缺陷极大地限制了机器学习算法的推广应用。

收稿日期: 2023-04-18

基金项目: 国家自然科学基金资助项目(52179138, 51879169, 52209165)

作者简介: 胡江(1983—), 男, 湖南衡阳人, 正高级工程师, 博士, 主要从事水工结构老化病害及安全监控研究。

E-mail: huj@nhri.cn

集成学习算法是一种重要的数据挖掘方法,可利用多个机器学习算法的组合和集成来解决问题,以提高方法的泛化能力^[15-18]。按照基分类器,集成学习算法分为同态和异态两类^[19]。其中,同态集成学习算法包括决策树和人工神经网络集成等,有统一、线性和堆融合等方式;异态集成学习包括叠加法(Stacking 算法)和元学习法。集成学习算法根据基分类器的生成顺序分为串行、并行和混合拓扑组合。经典的集成学习方法 Boosting 及其改进的 AdaBoosting、GDBT(Gradient Boosting Decision Tree)为串行组合, Bagging 及其改进 RF 则为并行组合, RF 基于套袋(Bootstrap Aggregation),对来自子样本的多个决策树结果进行平均。Boosting 按顺序生成预测值,反复利用残差中的模式,用弱预测增强模型。XGBoost 结合 RF 和 Gradient Boosting 的优点,预测误差比单独使用 RF 和 Gradient Boosting 低得多。然而,集成学习方法仍为“黑盒子”模型。

对机器学习算法模型结果进行解释,将结果与水工结构工作机理进行对比,可找出关键影响因素,并辨识出自变量与效应量间的逻辑关系。为了提高机器学习算法模型的可解释性,胡江^[20]通过分析自变量对特高拱坝变形的边际效应,借助部分依赖图,识别了自变量对变形的影响规律,阐明了变形机制;苏燕等^[21]利用 RF 评估各因子的重要性,提高了神经网络模型预测过程中对隐藏状态层级的解释性。总体看,模型结果解释性研究较少。Lundberg 等^[22]基于博弈论,提出了一类可解释机器学习方法即沙普利可加性解释(Shapley Additive Explanations, SHAP)方法,可用 SHAP 值量化每个影响因素的贡献,还可从单个观测日给出自变量的影响力和相关性,以反映自变量对效应量的影响规律。可见,SHAP 方法有效地统一了全局和局部可解释性,具有良好的应用前景。

为此,针对现有基于机器学习算法构建的安全监控模型结果可解释性较差这一不足,本文简述 RF、XGBoost 等常用的集成学习算法,以及 SHAP 方法的原理和推导过程,提出一种基于集成学习算法的水工结构变形安全监控模型构建与解释方法。基于已构建的统计模型^[23],以某运行初期特高拱坝变形数据为例,验证提出方法的有效性和实用性。

1 基于机理驱动的变形预测方法

以高拱坝变形为例进行说明,其变形主要由水压、温度和时效引起,统计模型可表示为:

$$\delta = \delta_H + \delta_T + \delta_\theta + g \quad (1)$$

式中: δ 为位移; δ_H 、 δ_T 、 δ_θ 分别为水压、温度和时效分量; g 为常数。

温度对高拱坝变形影响较大,温度因子的选取十分重要。采用实测温度可提高模型精度,此时 δ_T 可表示为:

$$\delta_T = \sum_{i=1}^M b_i T_i \quad (2)$$

式中: T_i 为实测温度; M 为温度计的数量; b_i 为系数。

采用式(2)表示温度因子的统计模型即为 HTT 模型(Hydrostatic Temperature Time, HTT)。运行初期的特高拱坝坝前水温垂直热分层显著,且由于封拱后水化热影响坝体混凝土温度回升,坝体温度场非稳定。采用基于主成分的分层聚类方法(Hierarchical Clustering on Principal Component, HCPC)可体现温度时间序列之间的相似性和差异,能从大量实测温度测点中选择具有代表性的温度时间序列^[20]。

考虑到主成分可以代表所有温度测点时间序列的变化,也可以采用下式来代表实测温度因子 δ'_T :

$$\delta'_T = \sum_{i=1}^N b_i T'_i \quad (3)$$

式中: T'_i 为实测温度的主成分(Principal Component, PC); N 为所选主成分的数量。

采用式(3)表示温度因子的统计模型即为 HT_{PC}T 模型(Hydrostatic-Principal Component based

Temperature-Time, HT_{PCT})。

对于运行初期的高拱坝, 时效变形受谷幅变形影响显著。Hu 等^[23] 根据大坝混凝土和围岩的蠕变机制推导了时效变形的解析式, 分析了谷幅和时效变形的关系, 对于溪洛渡、锦屏一级等谷幅变形幅度较大的特高拱坝, 蓄水和运行初期的谷幅和时效变形表现出良好的一致性。因此, 为更好反映变形机制, 采用谷幅变形来表示时效变形 δ_θ :

$$\delta_\theta = cv$$

(4)

式中: v 为谷幅变形; c 为系数。

通过式(2)、(4)或式(3)、(4)可更好地描述环境坝体温度的非稳定、谷幅变形等的影响, 提高机理驱动模型的精度^[20, 23]。

2 集成学习算法模型与结果解释方法

2.1 RF 算法

假设因变量 y 有 n 个观测值, 有 k 个自变量与之相关。如图 1, 在构建分类树的时候, RF 算法采用 Bootstrap 重新抽样的方法随机在原数据中选择 n 个观测值, 有的观测值被选择多次, 有的则未被选中。RF 算法随机从 k 个自变量选择部分变量确定分类树节点, 每次构建的分类树都可能不一样。一般情况下, RF 算法随机生成几百至几千个分类树, 然后选择重复程度最高的树作为最终模型。决策树的原始参数和树的数量是重要参数。

2.2 XGBoost 算法

Boosting 算法见图 2, 作为决策树算法, XGBoost 不受多重共线性的影响。给定具有 n 个样本的数据集, 存在自变量 x_i , 自变量有 m 个特征即 $x_i \in \mathbb{R}^m$ 。对于每一个变量, 都有相应的因变量 y_i , $y_i \in \mathbb{R}^m$ 。对于决策树集成学习算法, 模型使用 K 个自变量函数的和来预测因变量 y_i :

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F$$

(5)

式中: f_k 为具有叶分数的独立树结构; F 为树的空间。目标使得式(6)最小化:

$$\mathcal{L} = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k)$$

(6)

式中: $\mathcal{L}$ 为目标函数; l 为损失函数; Ω 为模型复杂度的惩罚项, 防止过拟合, 且:

$$\Omega(f) = \gamma G + \frac{1}{2} \lambda \|\omega_i\|^2$$

(7)

式中: G 为叶的数量; ω_i 为第 i 片叶的分数; γ 为节点切分的难度; λ 为 L2 正则化系数。

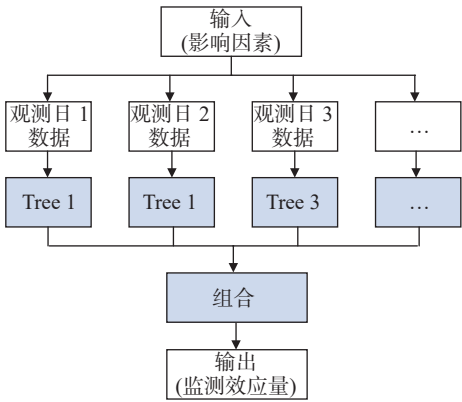


图 1 Bagging 算法

Fig. 1 Bagging learning algorithms

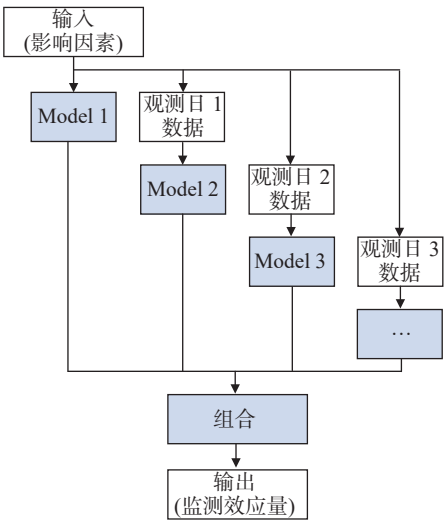


图 2 Boosting 算法

Fig. 2 Boosting learning algorithms

求解式(5)~(7),得到 ω_j^* 的最优值:

$$\omega_j^* = -\frac{\sum_{i \in I_j} \partial \hat{y}^{t-1} l(y_i, \hat{y}^{t-1})}{\sum_{i \in I_j} \partial^2 \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}) + \lambda} \quad (8)$$

$$\mathcal{L}^t(q) = -\frac{1}{2} \sum_{j=1}^G \frac{(\sum_{i \in I_j} \partial \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}))^2}{\sum_{i \in I_j} \partial^2 \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}) + \lambda} + \gamma G \quad (9)$$

由于很难计算所有可能树结构的最优值,因此使用下式:

$$\mathcal{L} = \frac{1}{2} \left[\frac{(\sum_{i \in I_L} \partial \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}))^2}{\sum_{i \in I_L} \partial^2 \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}) + \lambda} + \frac{(\sum_{i \in I_R} \partial \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}))^2}{\sum_{i \in I_R} \partial^2 \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}) + \lambda} - \frac{(\sum_{i \in I} \partial \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}))^2}{\sum_{i \in I} \partial^2 \hat{y}^{t-1} l(y_i, \hat{y}^{t-1}) + \lambda} \right] - \gamma \quad (10)$$

式中: $I = I_L + I_R$, I_L 和 I_R 分别是拆分后左、右节点的实例集。

2.3 基于 SHAP 的模型解释方法

SHAP 基于博弈论和局部解释,估计自变量贡献。假设集成学习算法模型中,具有 n 个特征的 N 组用于预测输出。SHAP 中,每个自变量对模型输出 $v(N)$ 的贡献 φ_i 是基于其边际贡献进行分配的。SHAP 量化了预测模型中每个自变量的贡献,将结果解释为各自变量的 SHAP 值之和:

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} [v(S \cup \{i\}) - v(S)] \quad (11)$$

式中: S 为 N 的特征子集。

3 基于集成学习算法的变形预测和解释流程

综合集成学习算法和 SHAP 方法,构建水工结构变形预测模型,并对模型的拟合和预测结果进行解释,流程如下:

(1)构造安全监测数据库,对效应量和自变量数据进行异常值处理与缺失值插值,获得较为平衡的数据集,对数据集进行归一化处理。

(2)划分数据集为训练集和测试集,划分比例一般可取 80% 和 20%。

(3)构建基于集成学习算法的变形预测模型。基于训练集进行模型训练;采用 10 阶交叉验证方法进行参数优化,寻求最优超参数。

(4)验证模型的真实性能。利用测试集评价模型性能,评价指标包含决定系数(R^2)、均方根差(Mean Square Error, E_{MSE})、均方根误差(Root Mean Square Error, E_{RMSE})以及平均绝对误差率(Mean Absolute Percentage Error, E_{MAPE})等数理统计指标。决定系数 R^2 反映了模型的精度,其值越接近于 1,模型拟合效果越好。

(5)利用 SHAP 方法进行解释。全局层面上利用 SHAP 分布来描述自变量的具体影响规律和相关性;局部层面上给出每个观测日中各自变量的量化贡献。

4 工程应用

为验证所提方法的有效性,以某运行初期特高拱坝变形数据为例进行验证。某水电站大坝为双曲拱坝(图 3(a)),坝顶高程 610.0 m,最大坝高 285.5 m,坝顶长 681.51 m,大坝共分 31 个坝段。正常、汛限和死水位分别为 600.0、560.0 和 540.0 m。2013 年 5 月开始下闸蓄水,2014 年 3 月完工。

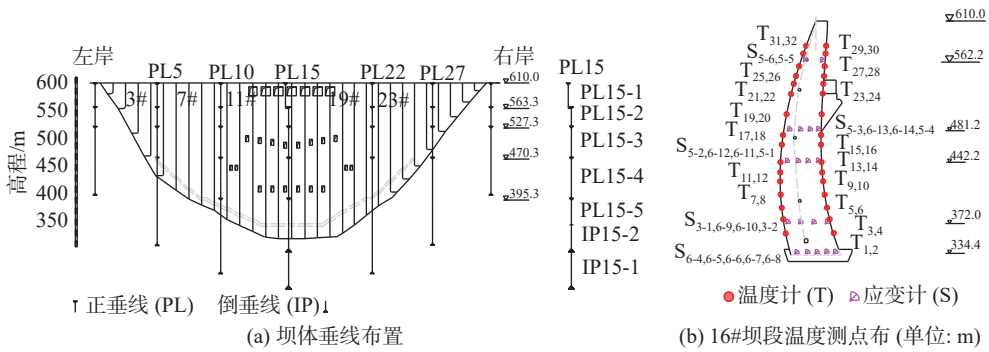


图3 坝体垂线和16#坝段温度测点布置(单位: m)

Fig. 3 Configuration of embedded pendulums and temperature sensors in monolith 16 (unit: m)

在5个坝段布置有垂线(图3(a)),正垂线(PL)测点布置在高程610.0、563.3、527.3、470.3、395.3和347.3 m,坝基布置倒垂线(IP),重点关注拱冠15#坝段坝顶的PL15-1测点的径向位移。在6个坝段布设环境和坝体温度测点,上、下游坝面各布置1支环境温度计(T),利用在多个高程均匀布置的应变计的测温传感器(S_{k-j} 为 k 向应变计第 j 个测点)观测混凝土温度。选取16#坝段温度测值分析,其温度测点布置见图3(b)。同一高程上的环境温度计、应变计测点编号从上游到下游用“,”号分隔。水库设有谷幅(Valley Deformation, S_{VD})监测系统,考虑到两岸坝肩611.0 m处的谷幅测线VDL4距大坝较近,故选取该测线测值分析,规定谷幅收缩为负,其测值变化过程见图4。

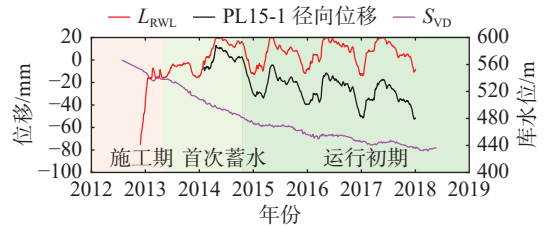


图4 库水位、PL15-1 径向位移和谷幅变形过程线

Fig. 4 Temporal profiles of reservoir water level, dam crest displacement, and valley contraction deformation

蓄水和运行初期库水位(Reservoir Water Level, L_{RWL})变化过程见图4,2013年5月—2018年6月共经历了4个加载和卸载过程。2013年5月4日开始蓄水,库水位从441.3 m逐渐升高,2014年9月28日首次蓄至600.0 m。随后库水位下降,2015年6月13日降至545.0 m。此后,库水位呈现加载-卸荷周期,库水位在5或6月达到最低值,汛期以相对较快的速度上升。除季节因素外,库水位还受到灌溉和航运等因素影响。

4.1 影响因子选取

由于库水温垂直分层、混凝土温度封拱回升等原因,坝体温度呈非稳定性。基于已开展的研究^[23],利用HCPC方法将库水温分为3类,选取的库水温主要为400.0~540.0 m范围内及540.0 m以上的(T_{15} 和 T_{31} 测点)测点温度,分别代表温变层和温跃层。混凝土内部温度的测点选取了 S_{5-1} 、 S_{6-4} 、 S_{6-9} 、 S_{5-4} 和 S_{5-5} ,下游面温度的测点选取了 T_6 、 T_{14} 、 T_{18} 和 T_{28} 。模型构建时考虑的自变量主要包括 L_{RWL} 、 T_i 、 S_{k-j} 、 T'_i 和 S_{VD} 等。构建了由式(2)表示温度因子的HTT模型。对库水温、坝体混凝土、下游面温度进行主成分分析,计算实测温度因子的 T'_i (图5),构建了由式(3)表示温度因子的HT_{PC}T模型。前4个PC(图5(a))解释了所有51个温度变量总方差的87.52%,前6个PC解释90.09%。具体地, T'_1 解释了总方差的40.75%,与库水位波动高度相关(图5(b)),解释了库水位波动引起的库水温变化。 T'_2 解释了总方差的28.23%,其值随时间增加,体现了以 S_{6-9} 测点为代表的因水化热引起的坝体混凝土温度封拱回升的特征(图5(c))。 T'_3 解释了总方差的14.97%,与气温波动相关。为此,使用前4个PC构建模型。

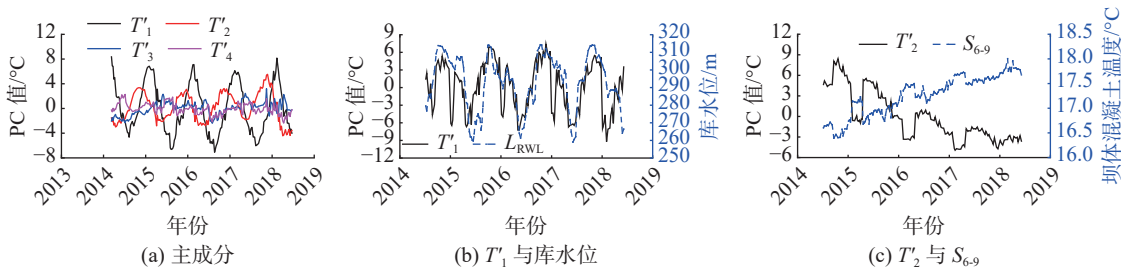


图 5 实测温度的主成分及其与环境因素的关系

Fig. 5 Principal components of recorded temperatures and their associations with environmental factors

4.2 模型构建

为对比分析, 分别基于实测温度及其主成分, 构建了 PL15-1 测点径向位移(图(3))的统计模型以及 RF、XGBoost 模型, 规定径向位移向下游为正。基于实测温度及其主成分的模型考虑的因子相同。RF 模型采用协变量数量默认值的 1/3 进行拟合, 决策树数为 500; XGBoost 模型叶的数量为 250, 点切分的难度 γ 为 0.3; 正则化系数 λ 为 4。构建的 6 个模型均通过了 10 阶交叉验证, 各模型的训练集和预测集的精度列于表 1, 各模型的训练和预测效果见图 6。

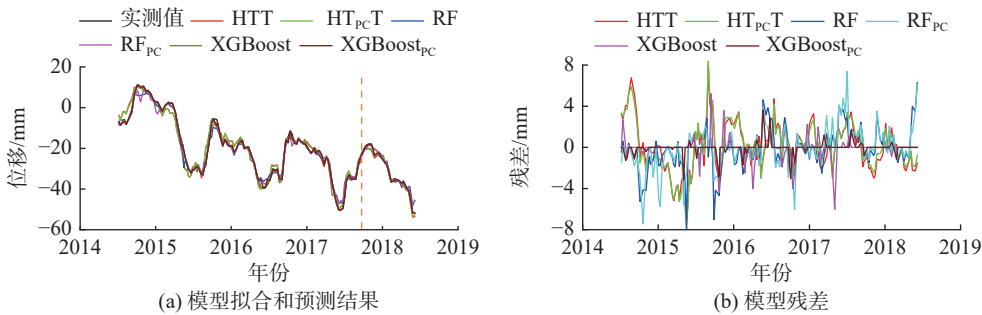


图 6 模型拟合、预测以及残差结果

Fig. 6 Results of model fitting, prediction and residual analysis

表 1 基于实测温度各模型及各主成分训练集和预测集的精度对比

Tab. 1 Comparative analysis of accuracy between training and prediction sets for various models based on the measured temperature data and temperature PCs						
项目	模型	模型训练或预测	R^2	E_{MAE}/mm	E_{MSE}/mm	E_{MAPE}
基于模型的精度对比	HTT	训练	0.973	1.753	4.417	0.315
		预测	0.961	2.511	8.573	0.258
	RF	训练	0.987	1.220	2.721	0.171
		预测	0.935	2.569	10.012	0.275
	XGBoost	训练	1.000	0.001	2.791×10^{-6}	1.680×10^{-4}
		预测	0.983	1.556	3.597	0.127
基于主成分的精度对比	HTT	训练	0.974	1.720	4.263	0.246
		预测	0.956	2.490	9.094	0.596
	RF	训练	0.973	1.569	5.275	0.189
		预测	0.987	1.074	3.063	0.096
	XGBoost	训练	1.000	0.005	3.803×10^{-5}	3.915×10^{-4}
		预测	0.982	1.276	3.830	0.171

从图 6 可看出, 各模型拟合和预测得到的全局变形、局部变化规律与实测值的吻合较好。由表 1 可见, 基于实测温度及其 PC 的模型预测性能相差不大。XGBoost 算法模型在 E_{MAE} 、 E_{MSE} 、 E_{MAPE} 这 3 个评价指标上都小于其他 2 种算法, 且决定系数 R^2 最高, 较好地反映了效应量与自变量之间的复杂关系, 预测效果最为理想。这表明运行初期特高拱坝的效应量表现出复杂的非线性关系, XGBoost 模型能较好地反映这种非线性关系, 获得了理想的预测效果。改进的统计模型也有较好的精度, RF 算法模型的精度相对较差。

4.3 基于 SHAP 的模型结果解释与分析

考虑到 XGBoost 算法模型精度较理想, 采用 SHAP 方法对该模型结果进行解释与分析。图 7(a)、(b) 分别为基于实测温度及其主成分的 SHAP 全局解释结果摘要, 其中, 横坐标表示 SHAP 值, 正、负 SHAP 值表示影响因素与变形的正、负相关, 散点颜色由黄至紫表示 SHAP 值由小至大, 纵坐标从上至下自变量的重要性逐渐降低。由图 7 可见, 谷幅变形、库水位和坝体温度等对变形的影响较显著。根据正、负值, 库水位为正面影响, SHAP 值越大, 对应变形越大; 谷幅变形、坝体温度为负面影响, SHAP 值越小, 对应变形越小。具体而言, 库水位值越高, SHAP 值越大, 对应的径向位移向下游发展。相反, 谷幅变形、坝体混凝土温度尤其是 S_{6-9} 、 T_{18} 等测点温度对其径向位移造成的负面影响较明显, 且随着特征值的增大, SHAP 值减小, 其负面影响程度增加。此外, 库水温 (T_{15})、坝体下游温度 (T_{18}) 等环境温度对坝体变形也有影响; T'_2 、 T'_3 也能得到同样的解释。

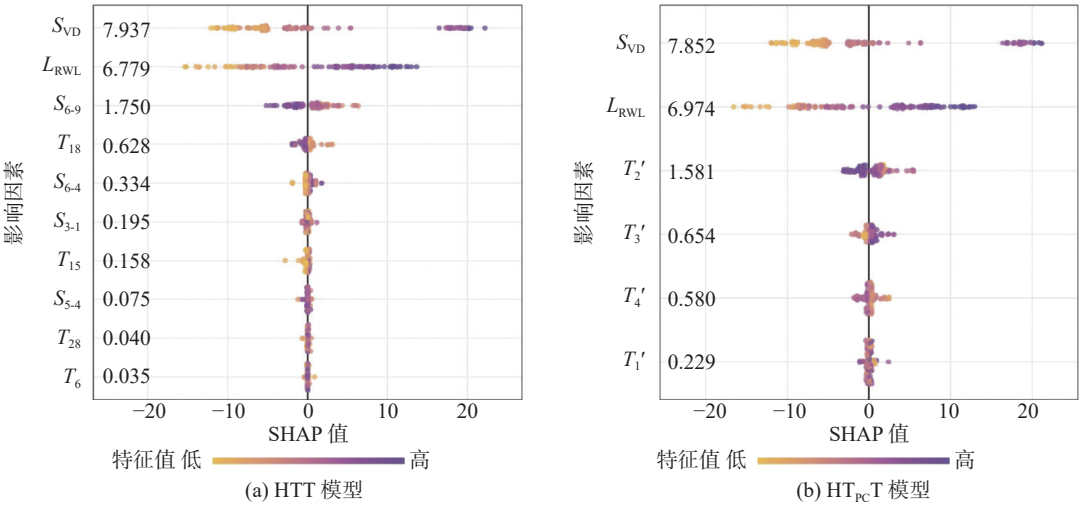


图 7 SHAP 全局模型解释结果

Fig. 7 Explanatory results for global model using SHAP

库水位、谷幅变形 SHAP 值的平均绝对值大, 数值分布方差也显著高于其他自变量, 对模型输出的贡献大, 便于模型对不同环境和荷载影响进行区分。

SHAP 将最终的预测分解为所有自变量贡献的总和, 可以展示所有自变量的定量影响。图 8 展示了如何利用 SHAP 解释模型对局部变形的预测结果。柱体越长表示该自变量 SHAP 值的绝对值越大, 也意味着对模型结果的影响越大。以 2015 年 9 月 28 日为例, 该日实测值、预测值分别为 -7.595 、 -7.591 mm, 谷幅变形、库水位、 S_{6-9} 测点温度等的贡献分别为 -1.465 、 -6.874 、 0.751 mm。

模型所得的结论与已有研究结果^[20] 吻合, 主要影响因素是谷幅变形(时效)、库水位和未稳定的环境与坝体温度, 其中蓄水引起的谷幅变形对运行初期特高拱坝的变形影响较大, 库水位次之, 坝体温度未稳定, 封拱后温度回升导致坝体也有向上游变形的趋势。为进一步了解自变量对变形的影响规律, 分析各变量对变形的边际效应, 辨识变量间相关关系及其对坝体变形的影响规律, 图 9 绘制了主要影响因素的特征依赖图。横坐标表示自变量大小, 纵坐标表示 SHAP 值。从图 9 可看出, 库水位的 SHAP 值随库水位升高而增

加,混凝土温度、谷幅变形的 SHAP 值随着特征参数的增大而减少。这说明库水位与径向位移呈正相关,混凝土温度、谷幅变形为负相关。同时还可以发现,谷幅收缩超过 60 mm、库水位在 290.0 m 以上时,对径向位移影响显著。

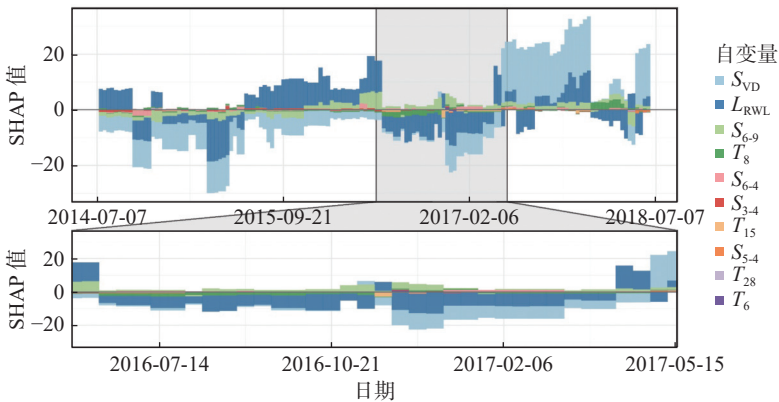


图 8 SHAP 对各观测日的解释结果
Fig. 8 Explanatory results for observation dates using SHAP

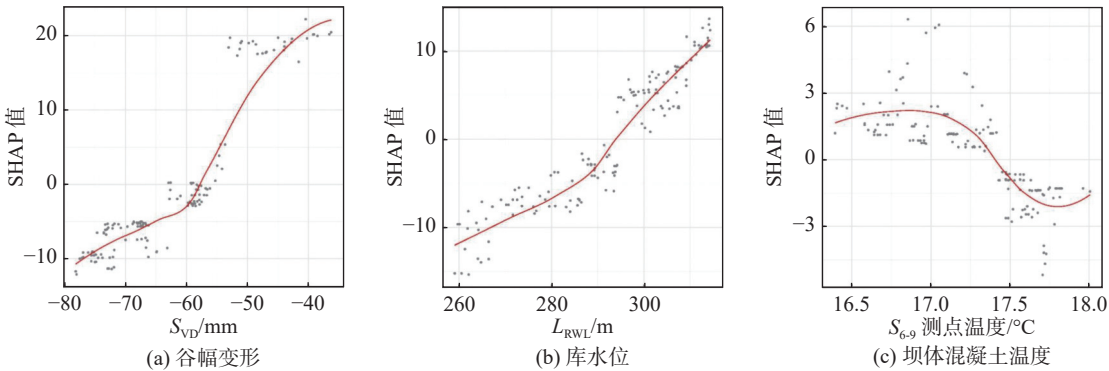


图 9 SHAP 特征依赖图
Fig. 9 SHAP dependency plots

5 结 语

为了提高水工结构变形预测模型的准确性和可解释性,提出了基于集成学习算法的水工结构变形预测模型构建与解释方法。以某运行初期特高拱坝变形为例,验证了方法的有效性和实用性,得到以下主要结论:

- (1)建立的基于集成学习算法的水工结构变形预测模型中,基于 XGBoost 变形预测模型测试集的决定系数 R^2 大于 0.982,模型预测效果好,改进统计模型的精度次之,RF 模型精度相对较低。
- (2)SHAP 方法可从全局角度解释自变量的重要性,挖掘影响水工结构效应量的主要自变量,并定性分析每一个自变量与效应量的正负相关性,较好地揭示了各因素影响变形的机制。
- (3)SHAP 方法可用于定量解释模型对每个观测日的计算结果,也可用于寻找变形的解析解,阐明模型对变形进行相应判断的原因,提高了模型拟合和预测结果的可信度和可解释性。
- (4)基于集成学习算法的水工结构变形预测与解释模型在推广泛化上具有显著优势,可为运行管理人员决策提供参考依据。

参 考 文 献:

- [1] 吴中如. 混凝土坝安全监控的确定性模型及混合模型[J]. 水利学报, 1989(5): 64-70. (WU Zhongru. Deterministic and hybrid models for safety monitoring of concrete dams[J]. Journal of Hydraulic Engineering, 1989(5): 64-70. (in Chinese))
- [2] 周仁练, 苏怀智, 韩彰, 等. 混凝土坝变形的长期预测模型与应用[J]. 水力发电学报, 2021, 40(9): 122-131. (ZHOU Renlian, SU Huaizhi, HAN Zhang, et al. Long-term deformation prediction model of concrete dams and its application[J]. Journal of Hydroelectric Engineering, 2021, 40(9): 122-131. (in Chinese))
- [3] 赵二峰, 顾冲时. 特高拱坝结构性态诊断与监控方法述评[J]. 水利水电工程学报, 2023(1): 16-26. (ZHAO Erfeng, GU Chongshi. Review on diagnosis and monitoring methods of structural behavior of superhigh arch dams[J]. Hydro-Science and Engineering, 2023(1): 16-26. (in Chinese))
- [4] 王仁超, 马钰明. 基于PCA和CS-KELM的重力坝变形预测模型[J]. 水利水电工程学报, 2021(4): 10-18. (WANG Renchao, MA Yuming. Prediction model of gravity dam deformation based on PCA and CS-KELM[J]. Hydro-Science and Engineering, 2021(4): 10-18. (in Chinese))
- [5] HU J, JIANG H, LI X. An optimized zonal deformation prediction model for super-high arch dams[J]. Structures, 2023, 50: 758-774.
- [6] FLAH M, NUNEZ I, BEN CHAABENE W, et al. Machine learning algorithms in civil structural health monitoring: a systematic review[J]. Archives of Computational Methods in Engineering, 2021, 28(4): 2621-2643.
- [7] SALAZAR F, TOLEDO M A, ONATE E, et al. An empirical comparison of machine learning techniques for dam behaviour modelling[J]. Structural Safety, 2015, 56: 9-17.
- [8] 冷天培, 马刚, 向正林, 等. 基于时序分解与深度学习的堆石坝变形预测[J]. 水力发电学报, 2021, 40(10): 147-159. (LENG Tianpei, MA Gang, XIANG Zhenglin, et al. Deformation prediction of rockfill dams based on time series decomposition and deep learning[J]. Journal of Hydroelectric Engineering, 2021, 40(10): 147-159. (in Chinese))
- [9] 欧斌, 吴邦彬, 袁杰, 等. 基于LSTM的混凝土坝变形预测模型[J]. 水利水电科技进展, 2022, 42(1): 21-26. (OU Bin, WU Bangbin, YUAN Jie, et al. LSTM-based deformation prediction model of concrete dams[J]. Advances in Science and Technology of Water Resources, 2022, 42(1): 21-26. (in Chinese))
- [10] 徐韧, 苏怀智, 杨立夫. 基于GP-XGBoost的大坝变形预测模型[J]. 水利水电科技进展, 2021, 41(5): 41-46, 70. (XU Ren, SU Huaizhi, YANG Lifu. Dam deformation prediction model based on GP-XGBoost[J]. Advances in Science and Technology of Water Resources, 2021, 41(5): 41-46, 70. (in Chinese))
- [11] 胡江, 马福恒, 王春红. 特高拱坝变形监测的分区分区及其模型构建方法[J]. 河海大学学报(自然科学版), 2021, 49(2): 148-154. (HU Jiang, MA Fuheng, WANG Chunhong. Zoned deformation monitoring and its model building method for super high arch dams[J]. Journal of Hohai University (Natural Sciences), 2021, 49(2): 148-154. (in Chinese))
- [12] 冯德成, 吴刚. 混凝土结构基本性能的可解释机器学习建模方法[J]. 建筑结构学报, 2022, 43(4): 228-238. (FENG Decheng, WU Gang. Interpretable machine learning-based modeling approach for fundamental properties of concrete structures[J]. Journal of Building Structures, 2022, 43(4): 228-238. (in Chinese))
- [13] SALAZAR F, MORÁN R, TOLEDO M Á, et al. Data-based models for the prediction of dam behaviour: a review and some methodological considerations[J]. Archives of Computational Methods in Engineering, 2017, 24(1): 1-21.
- [14] LAI D, DEMARTINO C, XIAO Y. Interpretable machine-learning models for maximum displacements of RC beams under impact loading predictions[J]. Engineering Structures, 2023, 281: 115723.
- [15] HU J, MA F H. Comparison of hierarchical clustering based deformation prediction models for high arch dams during the initial operation period[J]. Journal of Civil Structural Health Monitoring, 2021, 11: 897-914.
- [16] JIA J F, CHEN X Z, BAI Y L, et al. An interpretable ensemble learning method to predict the compressive strength of concrete[J]. Structures, 2022, 46: 201-213.
- [17] SUENAGA D, TAKASE Y, ABE T, et al. Prediction accuracy of Random Forest, XGBoost, LightGBM, and artificial neural network for shear resistance of post-installed anchors[J]. Structures, 2023, 50: 1252-1263.
- [18] FENG D C, WANG W J, MANGALATHU S, et al. Interpretable XGBoost-SHAP machine-learning model for shear strength prediction of squat RC walls[J]. Journal of Structural Engineering, 2021, 147(11): 04021173.
- [19] 周钢, 郭福亮. 集成学习方法研究[J]. 计算技术与自动化, 2018, 37(4): 148-153. (ZHOU Gang, GUO Fuliang. Research on ensemble learning[J]. Computing Technology and Automation, 2018, 37(4): 148-153. (in Chinese))
- [20] 胡江. 特高拱坝运行初期变形监测预报模型及构建方法[J]. 水利水电工程学报, 2020(5): 63-71. (HU Jiang. Deformation forecasting model and its modeling method of super high arch dams during initial operation periods[J]. Hydro-Science and

Engineering, 2020(5): 63-71. (in Chinese))

- [21] 苏燕, 付家源, 林川, 等. 基于时间注意力机制的大坝动态变形预测模型[J]. 水力发电学报, 2022, 41(7): 72-84. (SU Yan, FU Jiayuan, LIN Chuan, et al. Prediction model of dam structure dynamic deformation based on time attention mechanism[J]. Journal of Hydroelectric Engineering, 2022, 41(7): 72-84. (in Chinese))
- [22] LUNDBERG S M, NAIR B, VAVILALA M S, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery[J]. Nature Biomedical Engineering, 2018, 2(10): 749-760.
- [23] HU J, MA F H. Statistical modelling for high arch dam deformation during the initial impoundment period[J]. Structural Control and Health Monitoring, 2020, 27(12): e2638.

Forecasting and analysis of deformation in hydraulic structures

HU Jiang^{1,2}, SU Hui^{1,3}

(1. The National Key Laboratory of Water Disaster Prevention, Nanjing Hydraulic Research Institute, Nanjing 210029, China; 2. Dam Safety Management Center of MWR, Nanjing 210029, China; 3. College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China)

Abstract: The current machine learning-based safety monitoring models for hydraulic structures lack interpretability in their results. To address this issue, we propose a prediction and interpretation method for dependent variables based on ensemble learning algorithms. We provide a brief description of an improved statistical model and two commonly used ensemble learning algorithms, namely random forest (RF) and extreme gradient boosting tree (XGBoost). Additionally, we introduce the shapley additive explanation method (SHAP) to achieve interpretability in the results of ensemble learning algorithm models. We explain the principles and derivation process of the SHAP method. To verify the effectiveness and practicality of our approach, we utilize deformation data from a super high arch dam during its initial operation period as an example. The results demonstrate that the XGBoost model exhibits high prediction accuracy, with a decision coefficient greater than 0.982 in the prediction set. It is followed by the improved statistical model, while the RF model shows relatively poorer accuracy. The SHAP method effectively isolates the influence of different independent variables on the dependent variable, providing an impact mechanism and enhancing the interpretability of the fitting and prediction results. Our proposed method combines the strengths of both “mechanism driven” and “data driven” approaches, offering valuable insights for the operation and management of major hydraulic structures.

Key words: hydraulic structure; deformation; safety monitoring; machine learning; ensemble learning; interpretability