DOI:10.14188/j.1671-8844.2025.0354 文章编号:1671-8844(XXXX)XX-0001-10

基于随机森林及SHAP驱动的拱坝坝体位移回归分析

梁海静[1,2],程井[1,2],杨军义[2],余雷[3],戴妙林[1]

(1. 河海大学水利水电学院,江苏 南京 210098; 2. 水利部黄河流域水治理与水安全重点实验室,河南 郑州 450003; 3. 贵州省大坝安全监测中心,贵州 贵阳 550002)

摘要: 为解决单一机器学习位移监控模型泛化能力不足与可解释性弱的问题,以拱坝为例构建基于随机森林算法的位移高精度监控模型,利用SHAP(Shapley Additive Explanations)框架量化水压、温度及时效三类模型因子对顺河向位移的贡献并对总位移进行解构,以揭示因子作用机理。算例结果表明:水压因子是引起大坝顺河向位移的主要因素,其影响随库水位的抬升而非线性增加,1288m以上水位贡献超70%;1285.1m被识别为YY拱坝位移响应的临界水位,超过该阈值时水压因子对位移影响由负向转正向;模型能有效识别并剔除存在共线性的冗余温度因子,显著提升关键温度因子的贡献度;时效因子对位移贡献无偏向性。

关键词: 拱坝;大坝位移监测;随机森林算法;SHAP理论;机器学习

中图分类号: TV698.1 **文献标志码:** A

Deformation regression analysis of arch dam based on mixed random forest and shap methods

LIANG Haijing[1,2], CHENG Jing[1,2], YANG Junyi[2], YU Lei[3], DAI Miaolin[1]

(1. College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China;
2. Key Laboratory of Water Management and Water Security for Yellow River Basin, Ministry of Water Resources, Zhengzhou 450003, China;
3. Guizhou Dam Safety Monitoring Center, Guiyang 550002, China)

Abstract: To address the limitations of insufficient generalization capability and weak interpretability in individual machine learning models for dam safety monitoring models, this study constructs a high-accuracy displacement monitoring model based on Random Forest algorithm, using the YY arch dam as a case study. The SHAP framework is employed to quantify the contributions of three model factors - water pressure, temperature, and time-dependent factors to the streamwise displacement. The total displacement is decomposed into three components based on feature attributes, thereby revealing the mechanism of environmental variables on dam displacement. Results demonstrate that water pressure is the primary factor, with its contribution increasing non-linearly with rising reservoir water levels, exceeding 70% at levels above 1288m. A critical water level of 1285.1m is identified for YY's displacement response, beyond which direct water pressure factors exert a positive influence on displacement. The model effectively identifies and eliminates redundant

收稿日期:XXXX-XX-XX

作者简介:梁海静(2001-),女,硕士研究生,研究方向为水工结构安全监控研究,E-mail:2714988569@qq.com

通信作者:程井(1982-),男,博士,副教授,研究方向为水工结构分析及库坝工程数字孪生模型开发研究,E-mail:cj042@126.com。

基金项目:国家重点研发计划项目(2022YFC3005501);水利部重大科技项目(SKS-2022153);贵州省水利厅科技专项(KT202537);水利部黄河流域水治理与水安全重点实验室研究基金(2024-SYSJJ-02)

本文引用格式:梁海静,程井,杨军义,等.基于随机森林及SHAP驱动的拱坝坝体位移回归分析[J].武汉大学学报(工学版),XXXX,XX(XX):1-10.

Citation:LIANG Haijing, CHENG Jing, YANG Junyi, et al.Deformation regression analysis of arch dam based on mixed random forest and shap methods[J].Engineering Journal of Wuhan University, XXXX, XX(XX):1-10.

temperature factors with collinearity, significantly enhancing the contribution of key temperature factors. The time-dependent factors exhibit no directional bias in its contribution to displacement.

Key words: arch dam; dam displacement monitoring; random forest algorithm; SHAP theory; machine learning

构建环境变量与结构响应内在关系的位移监控模型,是大坝运行管理的核心举措。位移监控统计模型利用历史监测数据,通过统计方法进行建模和分析。传统统计方法主要包括逐步回归、偏最小二乘回归等[1-4],因其物理含义明确、计算高效而被广泛应用。苏卫星[5]将相关性特征筛选算法(correlation-based feature selection, CFS)引入传统偏最小二乘回归,通过CFS提取最优因子子集,缓解变量间的多重共线性问题,展现了传统方法迭代升级的潜力。然而传统统计模型自动化程度低,对极端异常工况和非线性关系捕捉能力有限,且当监测数据维度大、时间序列较长时会出现过拟合等问题。

机器学习模型由数据驱动,在处理高维、非线性、复杂耦合关系时精度更优[6-9],因此近年来也被逐步应用于大坝位移监控。谷宇[10]等将多算法融合,实现自适应降噪-冗余因子筛选-时序预测的递进式建模,捕捉有效信息,但算法表征深度仍局限于因子重要性排序。柳聪聪[11]等构建最小绝对值收缩和选择算子(least absolute shrinkage and selection operation, LASSO)与长短期记忆(long short-term memory, LSTM)集成模型,实现环境变量物理意义的可解释筛选及多因素耦合影响下的变形高效预测。但机器学习算法内部决策机制不透明,难以解释各变量对位移的影响机理,从而限制了其在工程实践中的应用。

近年来,可解释人工智能的兴起为解读复杂机器学习提供了新思路。其中,基于合作博弈论的SHAP(SHapley Additive exPlanations)理论,因其能精准量化模型中变量的贡献,有研究尝试将该法用于大坝渗流、变形等关键指标的归因分析。余红玲[12]等在构建土石坝渗流IAO-XGBoost(Improved Aquila Optimization-eXtreme Gradient Boosting)模型的基础上引入SHAP,实现了对渗流性态影响因素的定量归因,然而其解释维度主要集中于静态的全局分析。马春辉[13]等将SHAP应用于高坝变形预测模型,有效量化了环境因子对坝体变形的全局及局部静态影响,但未能揭示多变量耦合作用随时间推移的演变规律。尽管SHAP具备在大坝位移监控模型中揭示变量作用机制的潜力,但目前相关研究的深入解析仍显不足。综上,本文基于随机森林机器学习方法建立拱坝位移监控模型,并引入SHAP原理来定量解析各类因子对位移时间序列的具体贡献及其影响规律,为复杂的"黑箱模型"提供了一种深入且直观的解释方式,推动数据-机理双驱动模型在大坝安全监控中的应用。

1 研究方法

1.1 安全监控模型

传统位移监控统计模型采用线性回归分析方法,将坝体位移响应拟合为水压、温度与时效三类分量的线性叠加。式(1)为拱坝顺河向位移δ的监控模型数学表达式:

$$\delta=\delta_H+\delta_T+\delta_\theta=\sum_i a_i x_H+\sum_j a_j x_T\sum_k a_k x_\theta \tag{1}$$

式中,δ_H、δ_T、δ_θ分别表示位移的水压、温度及时效分量;它们由对应的模型输入因子x_H、x_T、x_θ回归得到。

为了定量比较监控模型的回归精度,引入均值绝对误差(MAE)、拟合优度(R^2)这2个评价指标:

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \tag{2}$$

$$R^2=1-\sum_{i=1}^{n}\left(\hat{y}_i-y_i\right)^2\Big/\sum_{i=1}^{n}\left(\bar{y}_i-y_i\right)^2 \tag{3}$$

式中,n为样本数;y_i为监测值;$\hat{y}_i$为回归值;$\bar{y}_i$为监测值的平均值。

1.2 SHAP理论框架

SHAP是一种基于博弈论的模型事后解释方法[14-15]。它将模型预测值分解为模型输入因子的SHAP贡献值,通过加性方式量化每个因子对单个预测的正向或负向影响及强度,从而在全局和局部层面提供直观的因子归因分析。

单个样本预测值等于基线值与所有因子贡献求和,计算公式如下:

$$f(x)=\varphi_0(x)+\sum_{i=1}^{M}\phi_i(x) \tag{4}$$

式中,$f(x)$为单个样本模型预测值;x为标准化后的自变量;φ_0为基线值,取因子缺失时的期望值;M为因子数量即自变量个数;ϕ_i为该样本第i个因子的SHAP值。

单个样本第i个因子的贡献计算公式如下：

$$\phi_i(f,x)=\sum_{S\subseteq F\backslash\{i\}}\frac{|S|!(M-|S|-1)!}{M!}[f_x(S\cup\{i\})-f_x(S)] \tag{5}$$

式中，ϕ_i为第i个因子的SHAP值；F为所有因子的集合，$F=\{x_1,x_2,\ldots,x_M\}$，M为总因子数；$S\subseteq F\backslash\{i\}$为不包含当前计算因子$x_i$的任意因子子集，维度为$|S|$；$f_x(S)$为仅保留子集$S$中的因子时，模型的预测值；$f_x(S\cup\{i\})$为综合考虑$S$及因子$x_i$计算的模型预测值。

1.3 随机森林回归算法

随机森林是一种集成机器学习算法，通过构建多个决策树并将其结果进行汇总来提高统计与预测的准确性和稳健性[16-18]。其具体步骤如下：从N个样本的数据集中，随机且有放回地抽取N个训练样本形成一棵子决策树，并重复K次生成多棵相互独立的决策树，在每棵树的生成过程中，需预先设定超参数，如树的数量、每棵树分裂节点时考虑的最大因子数、以及树的最大深度；每棵树根据样本的取值，在树节点上进行分裂，直至达到叶节点满足停止条件，获取回归结果；最后通过加权均值的方式融合所有树结果。

随机森林(Random Forest RF)统计性能强，却不具有线性回归等模型的自解释性。为解释说明RF模型决策机理，研究对比了偏最小二乘回归、识别关键变量的逐步回归法和随机森林算法(RF)；采用SHAP原理解析RF位移监控模型，量化各因子的边际贡献，实现因子参数的智能筛选与解释分析，并分离分量，实现兼顾高精度拟合能力和解释能力的大坝位移监控模型。模型构建具体流程如图1：

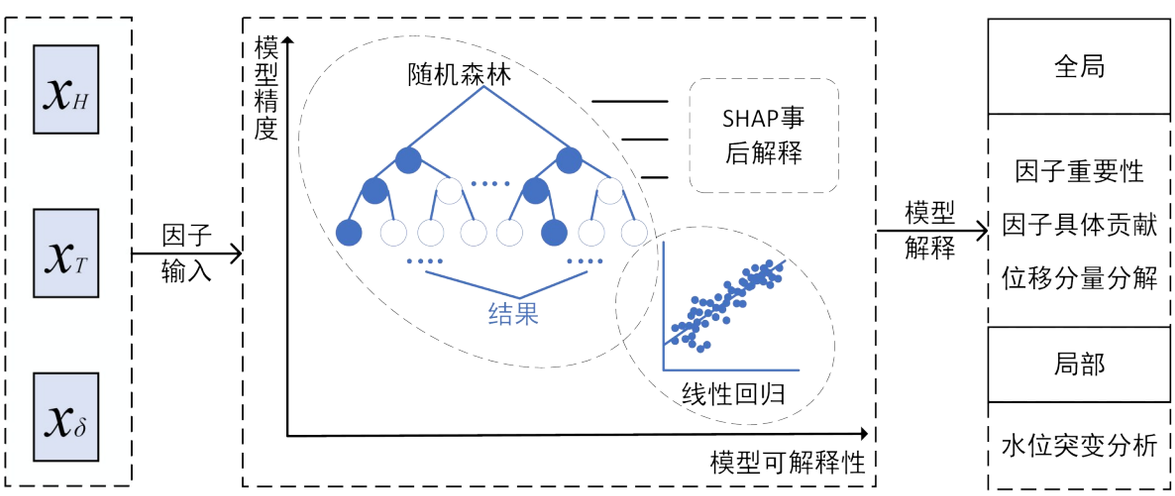


图1 基于随机森林及SHAP驱动的拱坝坝体位移回归分析

Fig.1 Flow chart of regression analysis of arch dam deformation integrating random forest and SHAP methods

2 实例分析

2.1 输入因子选取

YY拱坝为小(1)型水利工程，1982年12月竣工，2008年10月完成大坝除险加固。最大坝高51.9m，坝顶高程为1293.9m，正常蓄水位1290.0m。坝顶表面布设GNSS基站对坝体位移及沉降进行监测。本工程已实际运行44年，且坝内无有效温度计，参考文献[19-20]，拱坝的三类位移因子选取如下：(1)水压因子($x1$-$x6$)，静水压力作用采用水位H的1－3次方予以表征，并引入测前月平均水深的1－3次方，以反映长期荷载下应力重分布引起的可恢复变形；(2)温度因子($x7$-$x16$)采用多周期的谐波函数并结合坝面温度测值构建因子集合；(3)时效因子($x17$-$x18$)取时间一次式、时间对数式。具体见表1。

表中，H为观测日水位，$\overline{H}$为测前月平均水位，T观测当天温度，$\overline{T}$为测前月平均温度，t为位移观测日至建模资料起始日的累计天数，θ为观测日至基准日累计天数除以100；θ_0为观测日至资料起始日的累计天数除以100；a_i为回归系数。

2.2 监控模型回归

选取拱坝拱冠梁测点2024年1～10月共298例上下游向的水平位移数据作为目标变量，坝体水压及温度环境变量值如图2(a)。分别使用逐步回归、偏最小二乘回归及随机森林进行回归分析构建监

表 1 模型输入因子

Table 1 model input factors

$x1$	$x2$	$x3$	$x4$	$x5$	$x6$	$x7$	$x8$	$x9$
H	H^2	H^3	$\overline{H}$	$\overline{H}^2$	$\overline{H}^3$	T	$\overline{T}$	$\sin\frac{2\pi t}{365}$
$x10$	$x11$	$x12$	$x13$	$x14$	$x15$	$x16$	$x17$	$x18$
$\cos\frac{2\pi t}{365}$	$\sin\frac{4\pi t}{365}$	$\cos\frac{4\pi t}{365}$	$\sin\frac{6\pi t}{365}$	$\cos\frac{6\pi t}{365}$	$\sin\frac{8\pi t}{365}$	$\cos\frac{8\pi t}{365}$	$\theta-\theta_0$	$\ln\theta-\ln\theta_0$

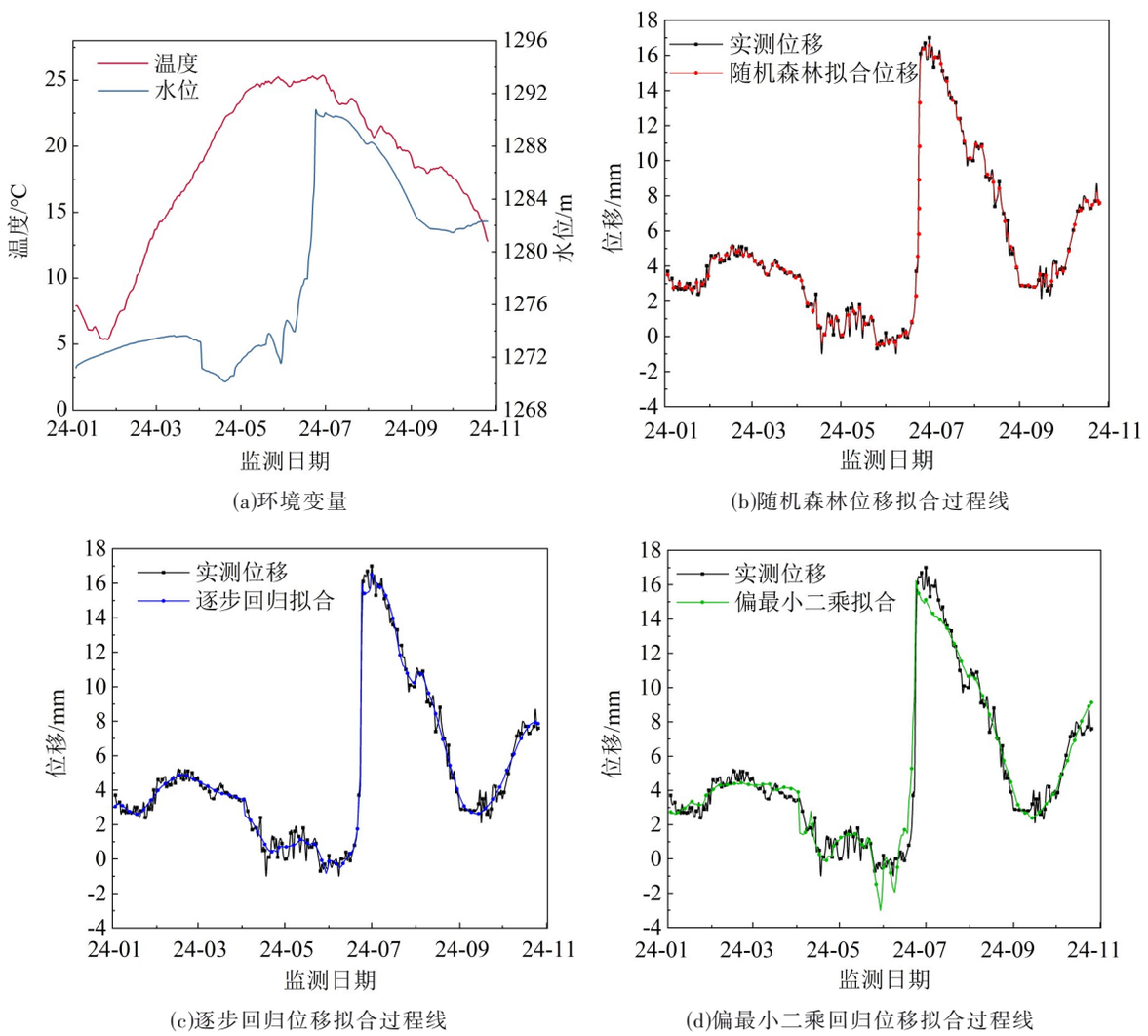
控模型。采用网格搜索结合交叉验证的方法来确定随机森林模型的最佳参数配置,参数搜索空间包括:树的数量(50,100,200)、最大因子数('sqrt','log2')、树的最大深度(None,2,4,6,7,8);交叉验证评估各参数组合的性能,以负均方误差作为评估指标。最终确定随机森林模型的最优参数为:树的数量100、最大因子数4与最大树深7,并据此构建最终模型。

三种方法拟合精度如表2,随机森林在拟合精度上显著优于偏最小二乘和逐步回归,R^2分别提高了4.1%、0.8%,MAE分别降低了72%及52%,具体拟合结果如图2所示。由于模型因子间的线性相关性与位移整体非线性,随机森林稳定性优于两种传统回归方法。

2.3 随机森林可视化

2.3.1 因子删减

采用SHAP框架进行全局因子重要性评估。如图3所示,拱坝水平位移主要受上游水位和环境



(a)环境变量 (b)随机森林位移拟合过程线

(c)逐步回归位移拟合过程线 (d)偏最小二乘回归位移拟合过程线

图2 环境变量与位移拟合值过程线

Fig.2 The time-process curve of environmental variables and fitted displacement values

表2 模型回归精度评价指标

Table 2 Statistical metrics for regression model evaluation

指标	偏最小二乘	逐步回归	随机森林
MAE/mm	0.633	0.373	0.178
R^2	0.956	0.987	0.995

气温影响，代表温度周期项的因子$x13$、$x16$影响最弱，考虑予以删减。进一步通过SHAP值相关性矩阵，计算$x13$和$x16$与其他因子的可替代率。如图4所示，两个目标能完全被其他相关温度因子替代。删除因子后，模型拟合精度提高，均值绝对误差和拟合优度变为0.172mm及0.996。

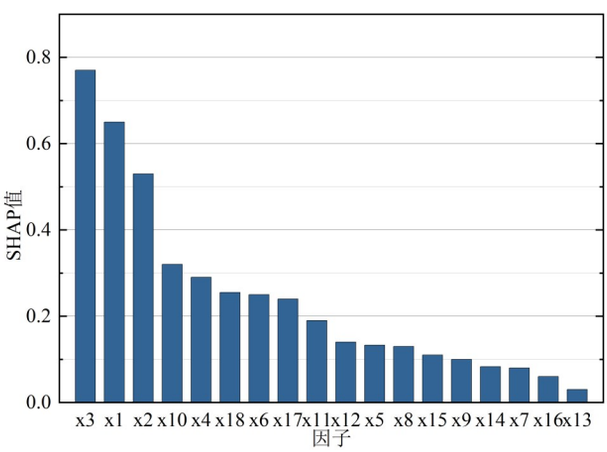


图3 因子重要性排序

Fig.3 Feature importance

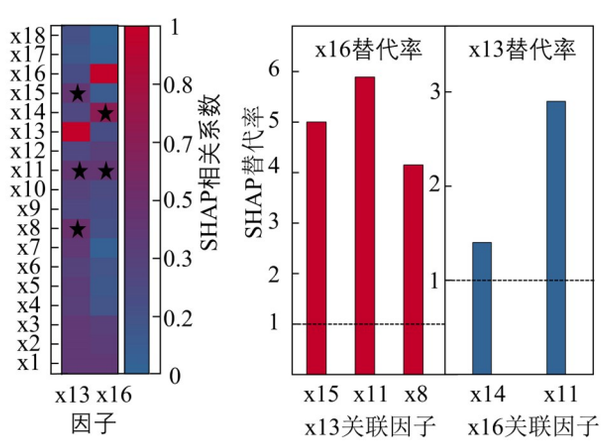


图4 因子可替代性图

Fig.4 Replaceability evaluation of factor

2.3.2 全局分析

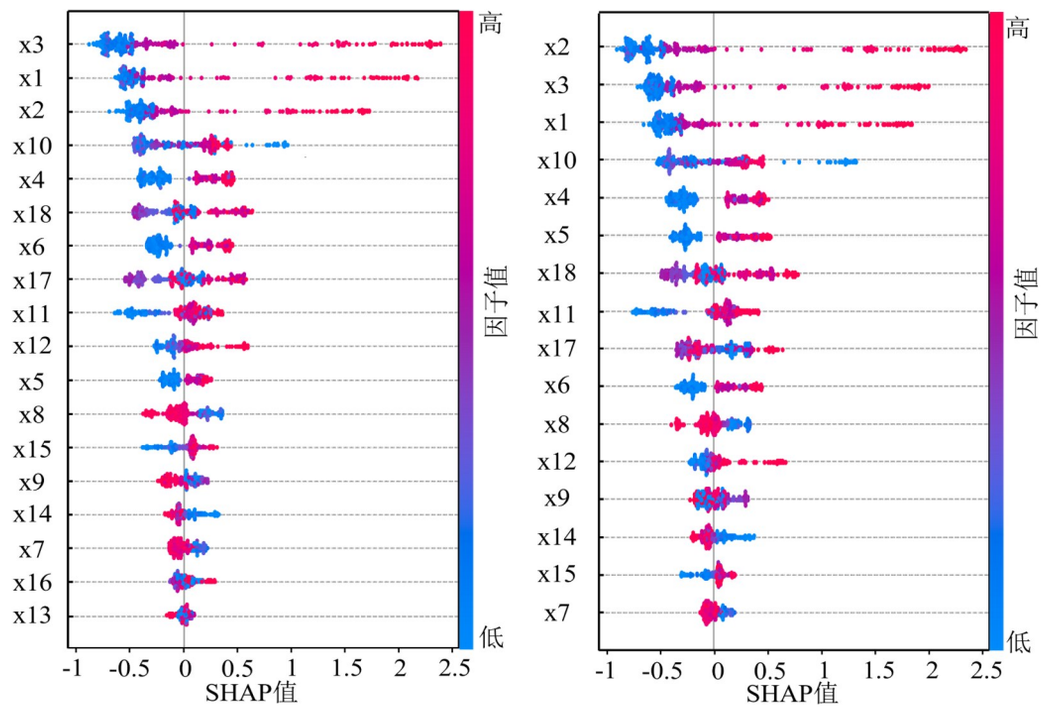
剔除冗余因子$x13$和$x16$后，重建随机森林模型，其最优因子数为3、最佳树深度6。分析新模型全局SHAP值，识别影响位移的核心驱动因子，结果见图5。横轴SHAP值为正表示因子对位移有正向影响，绝对值越大则影响越强；纵轴为因子重要性降序排列，颜色代表因子值大小。

1)温度因子。因子$x7$-$x8$直接温度项，因子$x9$-$x16$周期温度项。(1)图5(b)中直接温度项因子值与SHAP值呈负相关关系，从右至左因子值增大，对应SHAP值减小，且由正值转为负值，正负值范围基本相等。温度对位移的影响，低温驱动坝体产生正向(下游)水平位移，高温相反，符合材料热胀冷缩的物理规律；(2)分析$x9$-$x16$周期温度项因子。图5(a)中冗余因子$x13$和$x16$的SHAP值跨度小贡献有限，将其约减。约减后对应图5(b)，SHAP值重分配，最重要周期温度项$x10$的SHAP值提升0.5，增幅最大，$x11$和$x12$次之增加0.25，其余保留因子平均提升0.1，证明因子筛选能有效聚焦真实敏感因素，增强模型解释的可靠性。

2)水压因子。因子$x1$至$x6$。(1)图5(b)从左至右，因子值与SHAP值呈正相关关系，水位升高坝体产生正向(下游向)水平位移；(2)$x2$跨度最大、对位移影响最显著，与观测日水位相关的因子$x1$－$x3$的重要性排序高于与月平均水位相关的因子$x4$－$x6$；(3)因子$x2$右侧边界处的SHAP绝对值是其左侧边界处的2.5倍，推测高水位的贡献强于低水位。提取$x2$单因子全局影响如图6(a)，横轴为标准化后的因子值，结果显示当因子值超过1.2后，SHAP值的增长率提高至之前的3倍。为进一步验证，将$x2$值大小作为限制条件，对整体样本分组如图7虚线左侧，分组后$x2$值大于1.5、对应水位超1288.09m的44例的高水位样本，其SHAP均值是254例低水位样本的3倍。

综上，少数高水位样本是引发坝体位移的主导因素，且观测日水位影响大于月均水位作用。

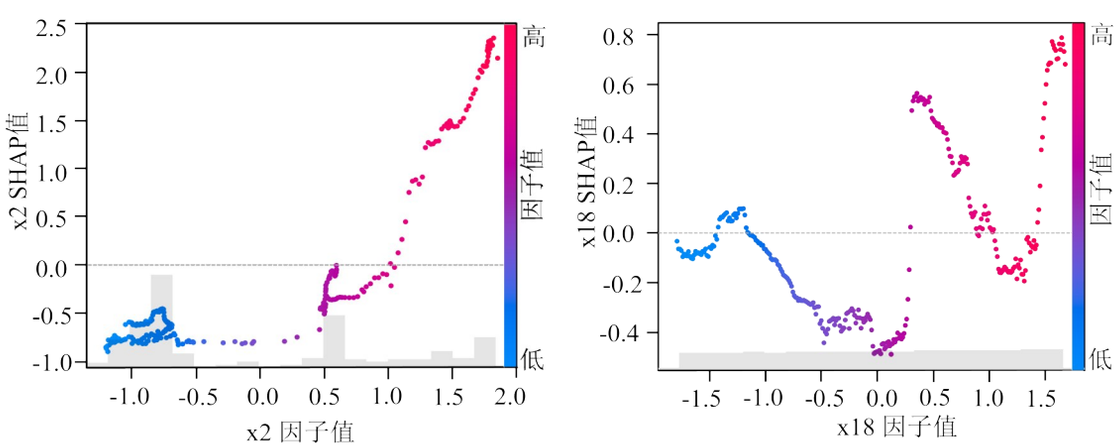
3)时效因子。因子$x17$与$x18$。(1)在图5(b)的整体统计中SHAP值与因子值相关性表现不明显，$x18$因子贡献稍大于$x17$；(2)提取$x18$进行单因子全样本分析如图6(b)，显示其与位移存在显著的非线性相关性。因子值小于0时，贡献值范围为[-0.6，0.1]，大于等于0时，范围为[-0.1，0.6]，两阶段变化幅度及速率基本一致，正负作用范围对称，该因子对位移方向的影响强度无显著偏向性。(3)将样本按时间顺序等分为前、后两组后，如图7虚线右侧结果所示。前50%样本与后50%样本的时效因子贡献度基本一致，表明其影响在时间维度上保持均衡。



(a)删减冗余因子前 (b)删减冗余因子后

图5 全局因子摘要图

Fig.5 Scatter plot of global characteristic density



(a)$x2$因子分布散点图 (b)$x18$因子分布散点图

图6 $x2$、$x18$因子分布散点图

Fig.6 Scatter plot of feature distribution for $x2$ and $x18$

2.3.3 *局部分析*

2024年6月末7月初,水库经历了一次蓄水与放水的过程。基于因子筛选后的监控模型,分别选取2024年5/30-6/28水位上升段和7/28-9/1水位下降段的样本,绘制SHAP决策图,以分析因子对单样本位移值的贡献。图8每一条彩色线代表一个样本的拟合累积过程,从底部基础值逐步叠加各因子贡献,最终到达顶部的模型输出值。

(1)图8(a)水位上升段,从左至右随水位上升顺河向位移持续上升。直接水压因子$x1$-$x3$(H^1-H^3)在低水位区间对位移呈负向影响,表现为虚线之前的样本线在$x1$-$x3$处向左偏移。随着水位升高,负向作用逐渐减弱。当水位超过1285.17m时即虚线位置,因子$x2$开始呈现正向贡献,虚线在$x2$处向右偏移;(2)图8(b)水位下降阶段,从右至左水位降低顺河向位移减小,$x1$-$x3$的正向作用逐步减弱,线条向右偏移量减小。当水位降至虚线位置1285.14m时,线条在$x1$和$x3$处向左偏移,其影响由

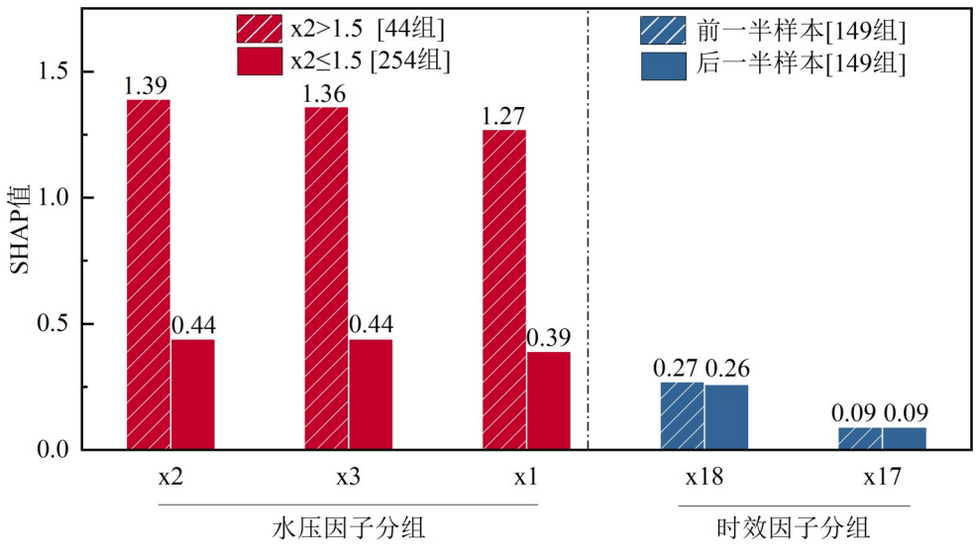
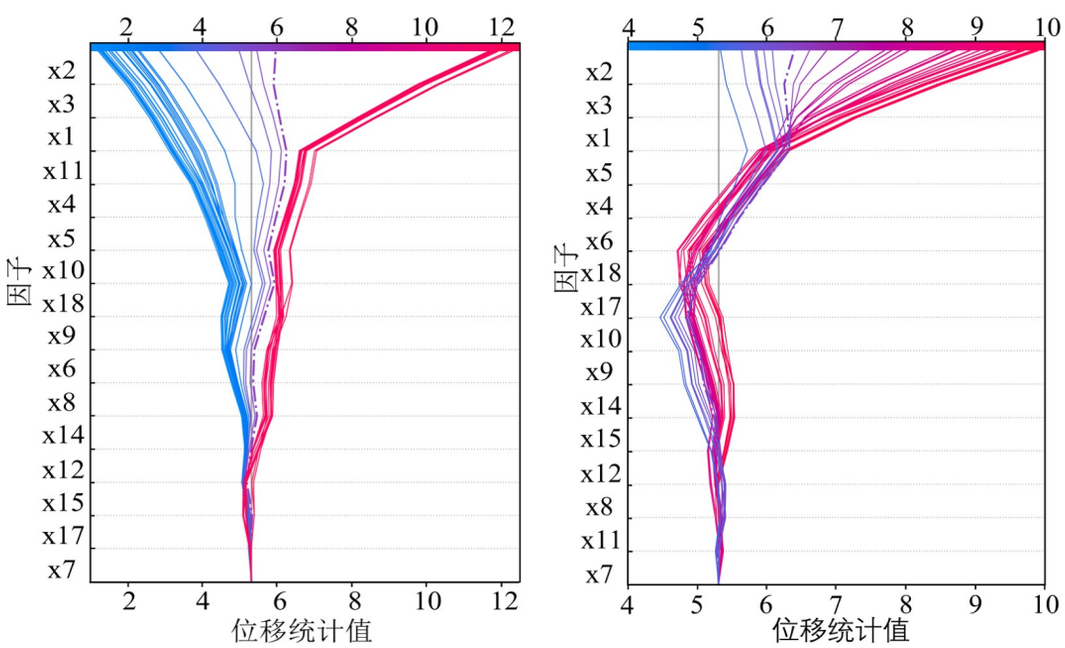


图7 水压因子、时效因子分组

Fig.7 Grouping of hydraulic pressure and time-dependent factors

正转负。因此可将水位1285.1m视为大坝位移响应转异的临界阈值水位。



(a)水位上升段决策图 (b)水位下降段决策图

图8 局部因子决策图

Fig.8 Interpretable local feature decision plot

具体分析，提取图(a)虚线及前后共三个样本。如图9所示，红色表示因子对位移正向影响，蓝色为负向。图9(a)到图9(b)，$x1$-$x3$的SHAP值由-0.21，-0.18，-0.26变为-0.18，+0.06，-0.16，直接水压分量的负向影响减弱，出现正向贡献。水位超过阈值后图9(c)，正向贡献主要由直接水压因子$x1$-$x3$提供，分别为1.59、1.91、1.69，是位移增量的主要驱动因素。

因此，在水位上升或下降至阈值时，建议控制水位变化的速率，减小水位波动对大坝位移的直接影响，缓解由于水压变化引发的应力集中，从而降低潜在的结构风险。

2.4 位移分解

据因子属性将SHAP值分别累加得到3个回归分量，见图10，其中水压分量变化最大，温度分量呈现季节性变化，时效分量相对稳定，且水压分量结

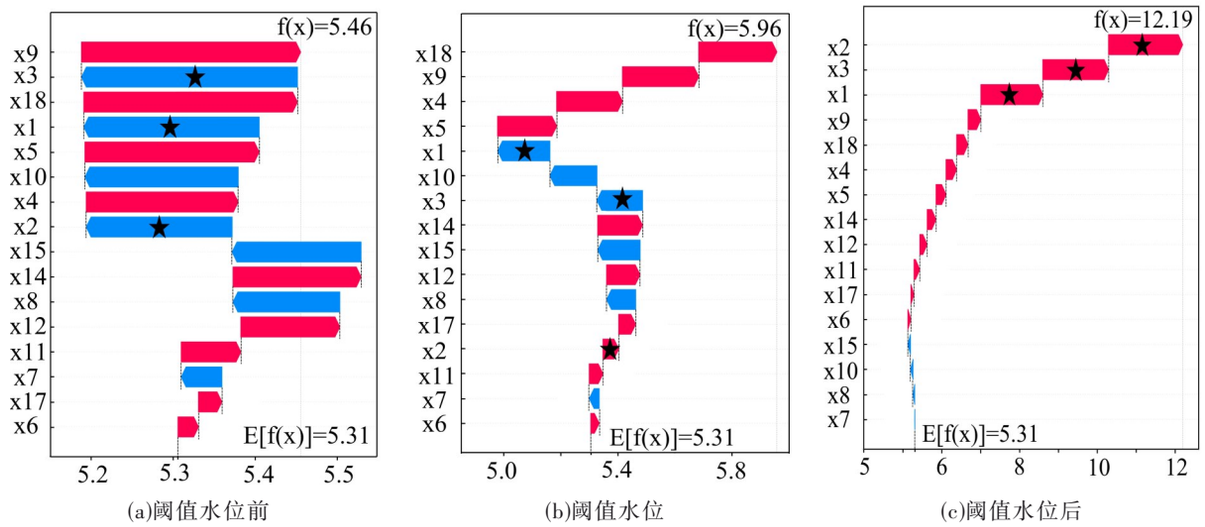


图9 单样本因子瀑布图

Fig.9 Additive feature contribution waterfall plot (single-sample)

果与有限元方法得到的大坝位移水压分量基本一致,模型分量结果能在一定程度上反映大坝的工作性态,具备良好定性解读价值。因为本工程目前数据量有限,待后续有更长期的数据时可采用置信区间法等进一步确定其分布特征,从而依据规范开展“正常-异常-险情”的定量安全评价[21]。

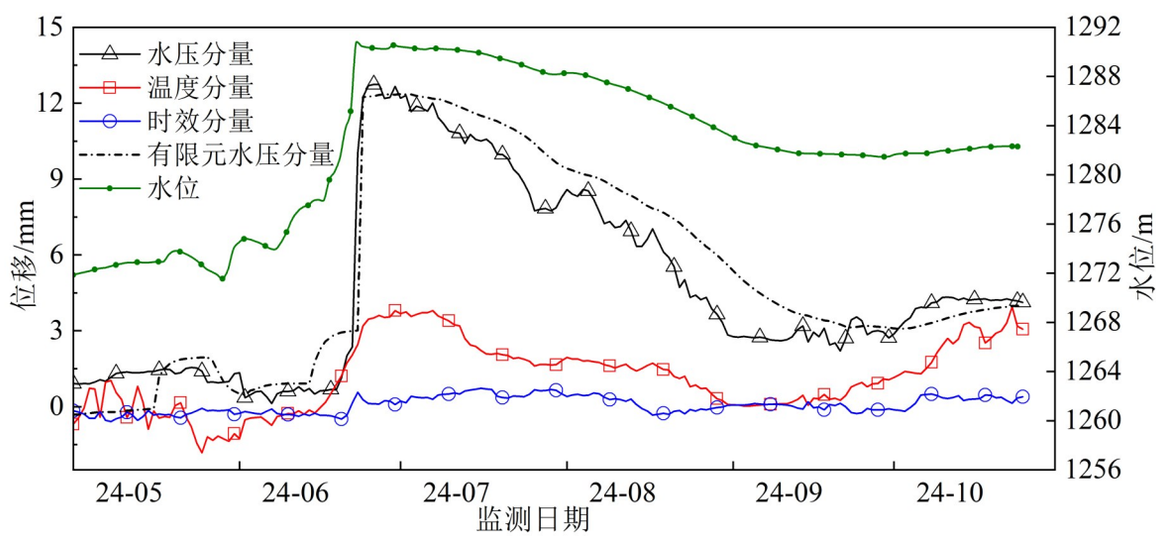


图10 位移回归值分离结果

Fig.10 Decomposed displacement regression outputs

3 结论

(1) 通过将SHAP理论与随机森林方法相结合构建的位移统计模型,能精准刻画复杂运行条件中的线性因子与非线性位移,从而实现了拟合精度与泛化能力的大幅提升。

(2) 定性分析各因子对大坝工作性态的影响规律:水压是主导因素,其贡献与因子值正相关,占总量17%的高水压样本贡献显著高于剩余样本贡献,且观测日水压因子影响大于测前月平均水压因子;时效因子与位移关系呈现高度非线性及无偏性;模型能有效识别并剔除冗余温度因子。

(3) YY拱坝位移响应在水位1285.1m左右发生转异,超过此临界点,观测日水压因子对位移影响由负转正,表现为影响位移开始向下游增长。

(4) 后续可结合更多工程实例,进一步揭示各分量的演化规律,并与有限元相结合实现数据-机理双驱动条件下的拱坝位移诊断与安全预警。

参考文献:

[1] Liu Xingyang, Li Zhanchao, Sun Linsong, et al. A critical review of statistical model of dam monitoring data[J]. Journal of Building Engineering, 2023, 80: 108-106.

[2] 苏荟,胡江,李星,等.改进温度分量的高海拔混凝土坝位移监控模型[J].水利水电科技进展,2025,45(02):106-112.
SU Hui,HU Jiang,LI Xing,et al.Displacement monitoring model of high-altitude concrete dam with improved temperature component[J]. Advances in Science and Technology of Water Resources,2025,45(02):106-112.

[3] Yang J, Su H, Liu H, et al. A multi-point combined prediction model for deformation monitoring of concrete dams[J].Engineering Applications of Artificial Intelligence,2025,159(PA):111652-111652.

[4] 王鑫鑫,凌薇,艾学轶.基于变分模态分解和集成深度学习的短期风速预测[J].武汉大学学报(工学版),2025,58(04):677-686.
WANG Xinxin,LING Wei,AI Xueyi.Short-term wind speed prediction based on variational mode decomposition and integrated deep learning[J].Engineering Journal of Wuhan University,2025,58(04):677-686.

[5] 苏卫星,冉顺义,刘芳,等.基于相关性变量筛选偏最小二乘回归的多维相关时间序列建模方法[J].信息与控制,2021,50(04):395-402.
SU Weixing, RAN Shunyi, LIU Fang, et al.Modeling method of multidimensional correlation time series based on correlation variable selection partial least squares regression[J].Information and Control,2021,50(04):395-402.

[6] 陈立秋,顾冲时,邵晨飞,等.基于相似测点对比的大坝变形监测数据粗差识别方法[J].水利水电科技进展,2024,44(04):72-77.
CHEN Liqiu, GU Chongshi, SHAO Chenfei, et al. Gross error identification method for dam deformation monitoring data based on similar measurement point comparison[J].Advances in Science and Technology of Water Resources,2024,44(04):72-77.

[7] B K, Wu Z Y.Data-based model with EMD and a new model selection criterion for dam health monitoring[J]. Engineering Structures,2022,260:114171.

[8] 程小龙,张斌,刘相杰,等.基于CEEMDAN-GMDH-ARIMA的大坝变形预测模型研究[J].人民黄河,2024,46(01):146-150.
CHENG Xiaolong, ZHANG Bin, LIU Xiangjie, et al. Research on dam deformation prediction model based on CEEMDAN-GMDH-ARIMA[J]. Yellow River, 2024,46(01):146-150.

[9] 黄海燕,艾星星,刘兴阳,等.基于可解释性分析的大坝变形监控模型对比研究[J].人民长江,2024,55(09):203-209.
HUANG Haiyan, AI Xingxing, LIU Xingyang, et al. Comparison of monitoring model for dam deformation based on interpretability analysis[J]. Yangtze River, 2024,55(09):203-209.

[10] 谷宇,苏怀智,张帅,等.基于自适应时序分解筛选的大坝变形预测模型[J].水利学报,2024,55(09):1045-1057+1070.
GU Yu, SU Huaizhi, ZHANG Shuai, et al.Dam deformation prediction model based on self-adaptive temporal decomposition screening[J].Journal of Hydraulic Engineering,2024,55(09):1045-1057+1070.

[11] 柳聪聪,张锋,胡超,等.基于可解释性因子选择的多模型耦合式大坝变形预测方法[J/OL].长江科学院院报,1-13[2025-09-26]. http://kns. cnki. net/kcms/detail/42.1171.TV.20250124.1454.032.html.
LIU Congcong,ZHANG Feng,HU Chao,et al.A multi-model coupled dam deformation prediction method based on interpretable factor selection[J/OL].Journal of Yangtze River Scientific Research Institute,1-13[2025-09-26]. http://kns. cnki. net/kcms/detail/42.1171. TV.20250124.1454.032.html.

[12] 余红玲,王晓玲,任炳昱,等.土石坝渗流性态分析的IAO-XGBoost集成学习模型与预测结果解释[J].水利学报,2023,54(10):1195-1209.
YU Hongling, WANG Xiaoling, REN Bingyu, et al. IAO-XGBoost ensemble learning model for seepage behavior analysis of earth-rock dam and interpretation of prediction results[J]. Journal of Hydraulic Engineering, 2023,54(10):1195-1209.

[13] 马春辉,余飞,程琳,等.特高拱坝变形机理可解释性智能预测模型[J].水力发电学报,2024,43(10):107-120.
MA Chunhui, YU Fei, CHENG Lin, et al. Explanatory intelligent prediction model for deformation mechanism of super-high arch dam[J].Journal of Hydroelectric Engineering,2024,43(10):107-120.

[14] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[J]. Arxiv Preprint Arxiv: 2017,1705:078-74.

[15] 向鑫,郭生练,崔震,等.考虑降水时空分布的洪水预报深度学习模型研究[J].武汉大学学报(工学版),2025,58(06):845-855.
XIANG Xin, GUO Shenglian, CUI Zheng, et al. Deep learning model for flood forecasting considering the spatio-temporal distribution of precipitation[J]. Engineering Journal of Wuhan University,2025,58(06):845-855.

[16] 田菊飞,苏怀智.基于随机森林算法的大坝应力预测模型的构建及其应用[J].水电能源科学,2018,36(5):54-56,5.
TIAN Jufei, SU Huaizhi. Development and application of dam stress prediction model based on random forest algorithm[J]. Water Resources and Power, 2018, 36(5): 54-56,5.

[17] Li K, Xu W Y, Wang Y, et al. Enhancing fuel cell system fault diagnosis: Using a bayesian optimized convolutional neural network-random forest hybrid model[J]. Sustainable Energy Technologies and Assessments,2025,82104479-104479.

[18] 黄佳健,王宝华,韩杰.基于LightGBM和可解释代理模型的风电场并网振荡稳定分析与控制[J].武汉大学学报(工学版),2024,57(10):1424-1433.
HUANG Jiajian, WANG Baohua, HAN Jie. Oscillation stability analysis and control of grid-connected wind farms based on LightGBM and interpretable surrogate model[J]. Engineering Journal of Wuhan University, 2024,57(10):1424-1433.

[19] 黄耀英,何一洋,沈振中,等.大坝监测量最佳统计模型优选方法[J].水利学报,2022,53(02):154-164.
HUANG Yaoying, HE Yiyang, SHEN Zhenzhong, et al. Optimization method of optimal statistical model of dam monitoring data[J]. Journal of Hydraulic Engineering,2022,53(02):154-164.

[20] 吴中如.水工建筑物安全监控理论及其应用[M].北京:高等教育出版社,2003.
WU Zhongru. Theory and Application of Safety Monitoring for Hydraulic Structures[M]. Beijing: Higher Education Press,2003.

[21] 中华人民共和国水利部.SL601-2013混凝土坝安全监测技术规范[S].北京:奔流电子音像出版(北京)有限公司,2013.
Ministry of Water Resources of the People's Republic of China. SL601-2013 Technical Specification for Safety Monitoring of Concrete Dams[S]. Beijing: Torrent Electronic & Audio-Video Publishing(Beijing)Co., Ltd., 2013.