

基于 GP-XGBoost 的大坝变形预测模型

徐 韧^{1,2}, 苏怀智^{1,2}, 杨立夫^{1,2}

(1. 河海大学水文水资源与水利工程科学国家重点实验室, 江苏 南京 210098;
2. 河海大学水利水电学院, 江苏 南京 210098)

摘要: 为了改善传统预测模型建模方法大样本训练效率低、容易过拟合、参数敏感性差的问题, 引入极端梯度提升(XGBoost)算法, 结合基于高斯过程(GP)的贝叶斯优化方法提升学习效率与预测精度, 构建了基于 GP-XGBoost 的大坝变形预测模型, 并通过工程实例与传统统计模型、神经网络模型的预测效果进行了比较。结果表明, 构建的大坝变形预测模型预测精度高, 迭代速度快, 通过调整正则项参数能有效避免过拟合。

关键词: 大坝变形; 预测模型; XGBoost 算法; 高斯过程; 贝叶斯优化方法

中图分类号: TV698.1 **文献标志码:** A **文章编号:** 1006-7647(2021)05-0041-06

Dam deformation prediction model based on GP-XGBoost//XU Ren^{1,2}, SU Huaizhi^{1,2}, YANG Lifu^{1,2} (1. *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China*; 2. *College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China*)

Abstract: In order to improve the problems of low training efficiency, easy over-fitting, and poor parameter sensitivity in traditional modeling methods, an extreme gradient boosting (XGBoost) algorithm is introduced, combined with a Bayesian optimization method based on Gaussian process (GP), to improve the learning efficiency and the prediction accuracy. The dam deformation prediction model based on the proposed method was established and the prediction effect was compared with the traditional statistical model and the neural network model. The results show that the monitoring model based on this method has high prediction accuracy and fast iteration speed. Overfitting can be effectively avoided by adjusting the regular term parameters.

Key words: dam deformation; prediction model; XGBoost algorithm; Gaussian process; Bayesian optimization

大坝作为重要的基础设施,其安全关乎民生,加强大坝性态监测是保障大坝安全的重要举措,而坝体的变形情况能够直观反映其性态,是判断大坝服役期间状态的重要依据^[1-2]。随着自动监测设备投入使用,坝体变形监测数据延续性更好、时效性更强、稳定性更佳,样本基数较人工观测时更大。而传统算法如灰色模型理论在大数据预测中效果欠佳^[3],BP神经网络在长周期数据预测中易过拟合^[4],基于时间序列的模型其参数在大时间跨度预测时敏感性较差^[5]。

近年来,随机森林、AdaBoost、Bootstrap等集成学习算法因其速度快、精度高、鲁棒性好的特点被学者运用在大坝变形预测领域。罗浩等^[6]通过随机森林算法,基于水压分量、温度分量、时效分量对拱坝位移进行了预测。沈晶鑫等^[7]将AdaBoost算法

与优化神经网络相结合,利用AdaBoost算法将粒子群优化后的神经网络模型集成,建立了大坝变形预测模型,并对向家坝第七坝段变形进行预测。王晓玲等^[8]基于Bootstrap和改进布谷鸟算法建立了混合模型,利用混沌理论和惯性权重理论以改善模型易陷入局部最优的问题,成功实现了大坝性态评价。极端梯度提升(extreme gradient boosting, XGBoost)算法由华盛顿大学陈天奇博士提出,是一种可进行并行计算的集成算法,通过加入正则项来控制模型复杂度,有效降低过拟合概率,能较好地解决传统模型存在的过学习、预测效率低、训练时间长、只适用于小集群样本等问题^[9],已在交通流量预测、商品销售量预测、疾病诊断等领域取得成功应用^[10-12],目前在大坝安全领域应用较少。

为提高大坝安全预测模型的预测精度及训练速

基金项目: 国家重点研发计划(2019YFC1510801, 2018YFC0407101); 国家自然科学基金(51979093)

作者简介: 徐韧(1996—),男,硕士研究生,主要从事水工结构与大坝安全监测研究。E-mail: 619235203@qq.com

通信作者: 苏怀智(1973—),男,教授,博士,主要从事涉水工程安全防控与提能延寿研究。E-mail: su_huaizhi@hhu.edu.cn

度,提升大坝安全监测系统对于数据噪音的响应能力,本文采用 XGBoost 算法对大坝变形数据进行预测,并引入贝叶斯优化,通过高斯过程(GP)对模型参数进行优化,建立基于 GP-XGBoost 的大坝变形预测模型,并结合原型观测数据对模型拟合预测能力进行评价。

1 建模原理

1.1 大坝变形预测统计模型基本原理

水压、温度等环境量的变化会导致坝体产生可逆变形,而时效因子则引起坝体材料劣化、力学性能下降从而形成不可逆影响^[13]。构建大坝变形预测模型时常选择水压因子、温度因子、时效因子作为影响因子^[14-15]。坝体位移公式为

$$\delta = \delta_H + \delta_T + \delta_\theta \tag{1}$$

式中: δ 为坝体位移; δ_H 、 δ_T 、 δ_θ 分别为位移的水压分量、温度分量和时效分量。

1.2 基于 XGBoost 算法的大坝变形预测模型构建原理

XGBoost 算法是梯度提升机器(gradient boosting machine,GBM)算法的扩展,是一种同时具备线性模型和树模型特点的优化模型,能够同时完成回归和分类任务。XGBoost 算法由多棵决策树(CART)组成,通过决策树集成实现机器学习,所有决策树的预测值相累加即为模型预测值,并由梯度提升决策树(gradient boost decision Tree,GBDT)算法进行模型训练^[16]。与 GBDT 算法不同的是,XGBoost 算法通过对损失函数进行二阶泰勒展开以快速逼近目标函数,并加入了正则项调节参数以避免过拟合,训练速度快,预测精度高^[17]。

在构建大坝变形预测模型时,XGBoost 算法通过在变形预测函数中不断加入关于各因子的新函数来逼近实测变形值,即:

$$\hat{y}_K = \sum_{i=1}^K F_i(x_i) = F_{K-1}(x_i) + f(x_i) \tag{2}$$

式中: K 为预测轮次; $\hat{y}_K$ 为第 K 轮变形预测函数的值; $f_K(x_i)$ 为关于水压因子、温度因子和时效因子的决策树函数。

XGBoost 算法目标函数定义如下:

$$F_0 = \sum_{i=1}^K L(y_i, \hat{y}_i) + \sum_{i=1}^K \Omega(f_K) \tag{3}$$

式中: L 为损失函数,用于评估变形预测值与真实值之间的损失; $\Omega(f_K)$ 为正则化函数,用于控制控制模型复杂度,避免过拟合。正则化函数定义如下:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \tag{4}$$

式中: γ 、 λ 为正则项惩罚系数; ω 为叶子节点对应的权重; T 为温度。对(2)式改写并作二阶泰勒展开:

$$F_{OK} = \sum_{i=1}^n (L(y_i, \hat{y}_{i,K-1} + f_K(x_i)) + \Omega(f_K)) \tag{5}$$

$$f(x + \Delta x) \cong f(x) + f'(x) \Delta x + \frac{1}{2} f''(x) \Delta x^2 \tag{6}$$

可以对公式(5)进行进一步简化得:

$$F_{OK} = \sum_{i=1}^n \left[g_i f_K(x_i) + \frac{1}{2} h_i f_K^2(x_i) \right] + \Omega(f_K) \tag{7}$$

式中: g_i 为损失函数的一阶梯度统计; h_i 为损失函数的二阶梯度统计。通过迭代找到 $f(x)$ 使目标函数最小则模型训练完成。

1.3 基于高斯过程(GP)的预测模型参数优化原理

贝叶斯优化算法(Bayesian optimization)是一种高效的优化工具^[18],其基于不同代理模型可分为3类:树型评估器 TPE、随机森林回归 AMAC 和高斯过程 GP。大坝原型观测数据集庞大,特选用迭代速度较快、质量较高的高斯过程作为优化模型内核。

以高斯过程为内核执行优化任务时,常对目标函数 $f(x)$ 假设一个满足高斯过程的先验分布模型,即模型各参数任何有限维度的组合仍满足高斯分布:

$$f(x) \sim \text{GP}(E(f(x)), C_{\text{ov}}(x, x')) \tag{8}$$

式中: E 为 $f(x)$ 的数学期望; $C_{\text{ov}}(x, x')$ 为 x 的协方差。每个参数组合 $\{x_1, x_2, \dots, x_i\}$ 都属于高斯分布,满足以下公式:

$$\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_i) \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} C_{\text{ov}}(x_1, x_1) & \cdots & C_{\text{ov}}(x_1, x_i) \\ \vdots & & \vdots \\ C_{\text{ov}}(x_i, x_1) & \cdots & C_{\text{ov}}(x_i, x_i) \end{bmatrix} \right) \tag{9}$$

在先验分布函数确定后,采样对模型进行修正。样本越多,模型越精确。为提升采样速度,优化算法通过定义采集函数确定下一个采样点^[19]。采集函数将权衡模型优化方向为探索(exploration)或开发(exploitation),探索将在新的区域采样以避免局部最优,开发则在现有局部最优解附近开展进一步搜索,尝试寻求全局最优解^[20]。优化目标则是在全集 A 中寻找使 $f(x)$ 值达到最大或最小的 X 集合如下式所示:

$$x^* = \arg \max_{x \in A} f(x) \tag{10}$$

1.4 K-折交叉验证

交叉验证值 C_v 常作为模型优劣评价指标。K-折交叉验证(K-fold cross-validation)是交叉验证的一种。调用交叉验证函数时,先将数据集均分为 K 份,选择其中一份作为验证集,其余 $K-1$ 份则用于训练模型,如此执行 K 次,将得到的 K 个评价值求

和平均即为交叉验证值 $C_v^{[21-22]}$ 。评价值的类型根据大坝观测数据特点选用均方根误差 RMSE (Root Mean Square Error)。

2 建模步骤

建模流程如图 1 所示,具体步骤如下。

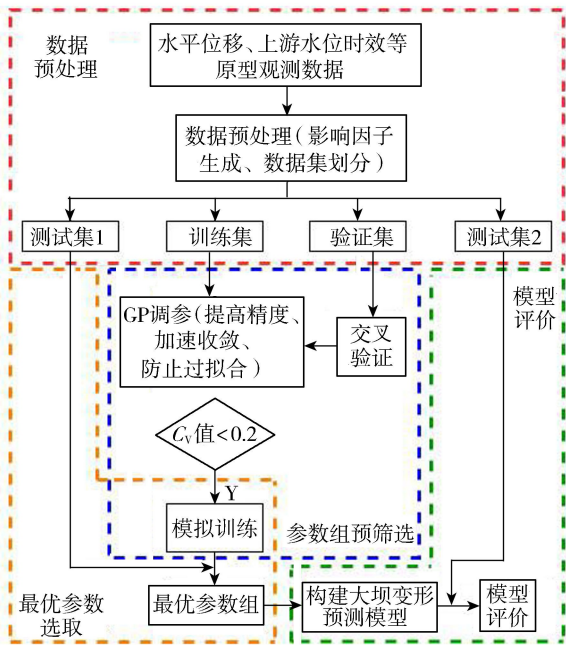


图 1 基于 GP-XGBoost 算法的建模流程

步骤 1 对上游水位、坝体上下游方向水平位移及时效数据进行处理,整理为合适影响因子作为输入样本集,并将其划分为训练集、验证集、测试集 1、测试集 2,其中验证集由交叉验证函数从训练集中划分。

步骤 2 划分好的训练集与验证集通过交叉验证产生评价指标 C_v ,确定各参数优化范围,带入贝叶斯优化算法进行优化迭代。由于验证集取自训练集内部,为避免局部最优和预测值突变的情况,选取 C_v 值小于最大 C_v 值 10% (本次选取 0.2) 的参数组构建预测模型^[23],对训练集 1 进行预测,若所选参数组均出现局部最优或预测值突变则适当放宽 C_v 阈值,综合选取最优参数组。综合评价各参数组训练速度与预测精度,选取最优参数组构建基于 GP-XGBoost 的大坝变形预测模型。

步骤 3 将测试集 2 中的数据带入步骤 2 中构建的模型,用于模型评价,通过与基于默认参数的 XGBoost 算法、线性回归方法、BP 神经网络所建立的模型进行比较,对 GP-XGBoost 模型做出评价。

3 工程实例验证

某水库为混凝土双支墩肋墩坝,控制流域面积 570 km²,总库容 3.47 亿 m³,是一座具有防洪、灌

溉、发电、养殖等综合功能的年调节的大(2)型水库。选取 2015 年 1 月至 2018 年 3 月共 1 174 组上游水位数据及 PL5 测点在上下游方向的水平位移测值作为初始数据集进行分析。其中 PL5 测点位于 5 号坝段坝顶,为自动测点,采用垂线法采集水平位移数据。

3.1 数据预处理

PL5 测点上下游方向水平位移、上游水位均为一天一测,数据较完整,自 2015 年 1 月至 2017 年 12 月,水平位移与上游水位一一对应,其中水平位移与水位过程线如图 2 所示。

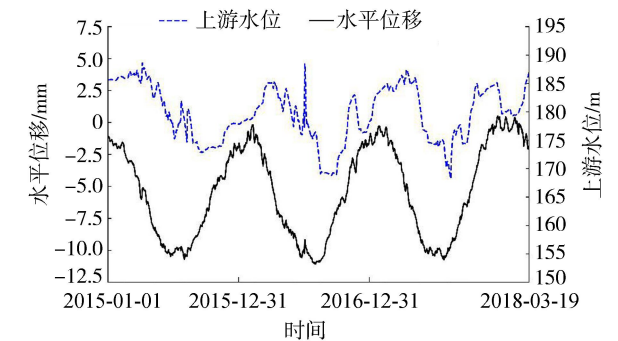


图 2 磨子潭水库 PL5 测点位移与上游水位过程线

结合目标大坝特性,选择 H 、 H^2 、 H^3 、 H^4 构成水压因子多项式,其中 H 为上游水深。温度的影响呈周期性作用于坝体,用正余弦函数可以很好地表示温度因子,取 $\sin\left(\frac{2\pi it}{365}\right) - \sin\left(\frac{2\pi i t_0}{365}\right)$ 、 $\cos\left(\frac{2\pi it}{365}\right) - \cos\left(\frac{2\pi i t_0}{365}\right)$ ($i=1,2$),其中 t 为监测日到实测日累计天数, t_0 为建模资料中第一监测日到始测日的累计天数。时效因子反映大坝劣化的程度,用 $\theta - \theta_0$ 、 $\ln \theta - \ln \theta_0$ 表示,其中 $\theta = \frac{t}{100}$ 。水压因子 4 项、温度因子 4 项、时效因子 2 项共计 10 项作为输入样本数据集。

将输入样本集前 600 项划分为为训练集,601 ~ 800 项划分为测试集 1,其余划分为测试集 2。验证集由内置交叉验证函数在训练集中自动生成,并在训练过程中生成模型评价指标 C_v 。

3.2 参数优化

对各关键参数进行初值定义,选定优化范围,具体参数含义、初值及取值范围如表 1 所示。

表 1 参数优化初值、范围及含义

参数	初始值	优化范围
训练步长 (learning_rate)	0.3	(0,0.5)
节点分裂折减系数 (gamma)	0	(0,10)
最大树深 (max_depth)	6	(0,10)
权重阈值 (min_child_weight)	1	(0.5,1)
样本采样率 (subsample)	1	(0.5,1)
对列采样率 (colsample_bytree)	1	(0,10)
正则系数 (alpha)	1	(0,10)

将表 1 中的参数组带入贝叶斯算法中迭代 2000 次,取出 2000 种参数组合中 C_v 值小于 0.2 的 5 组,如表 2 所示。

表 2 最优参数组

序号	C_v	learning_rate	gamma	max_depth
1	0.1853	0.1455	0.09267	6
2	0.0988	0.1500	0	2
3	0.1640	0.03315	0.02769	9
4	0.1790	0.5000	0	4
5	0.1040	0.1355	0.03450	5

序号	min_child_weight	subsample	colsample_bytree	alpha
1	5	0.7799	0.5562	0.0100
2	4	1	1	1
3	3	0.5071	0.8741	0.5734
4	8	0.6497	1	0.2514
5	8	0.9873	0.9877	0.9476

将 5 组参数分别带入 XGBoost 建模并对测试集 1 进行预测,迭代次数设置为 600。各参数组对测试集 1 预测情况如图 3 所示,其中第 2 参数组预测 RMSE 值最小为 0.90491。其中 11 月至 12 月变形预测数据出现小范围突变,其原因归结为此时段内坝体水压分量突然增大,模型参数 max_depth 对于突变数据较为敏感,当参数 max_depth 取值偏大时,预测值突变较明显,通过贝叶斯优化调整参数取值,选取合适的 max_depth 值可以缓解预测值突变的情况,如第 2 参数组所示。若根据条件 $C_v < 0.2$ 所选参数组建模预测均出现明显突变,则应适当扩大参数组选取范围,权衡各参数在合适范围做出选择。

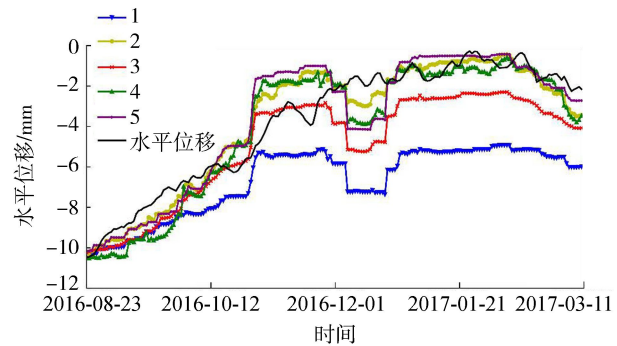


图 3 不同参数组预测结果

5 组参数组训练迭代收敛情况如图 4 所示。由图 4 可知,第 4 组参数建模收敛速度最快,但第 2 组参数建模在预测精度较高的条件下也满足了运算速度需求,综上所述,将选择第 2 参数组构建变形预测模型,用于精度评价。

图 4 说明预测模型最多经过 200 次迭代即可收敛,在后续的运算中将迭代次数缩减为 300 次以提升学习效率。

3.3 模型对比评价

针对回归模型选用均方根误差 (RMSE)、平均

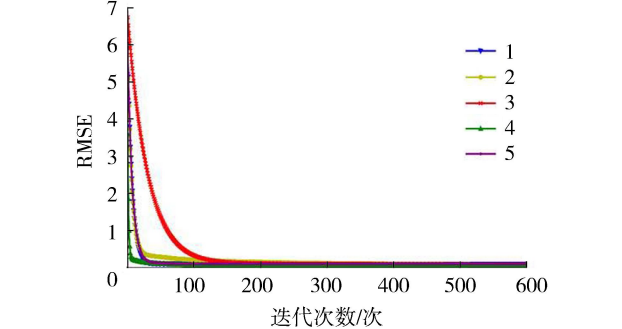


图 4 模型训练迭代过程

绝对误差 (MAE)、拟合优度 (R^2) 作为精度评价指标,分别用于评价模型稳定性、精确度及拟合效果。

将基于 GP-XGBoost 算法的预测模型与未经过参数优化的 XGBoost 预测模型、线性回归模型、BP 神经网络模型训练并对测试集 2 进行预测,对各模型的精度进行比较,评价模型性能。具体模型训练结果、预测结果及残差计算结果如图 5 所示,预测精度评价结果如表 3 所示。

表 3 各模型预测精度评价

模型	RMSE	MAE	R^2
GP-XGBoost	0.41023	0.19765	0.96832
XGBoost	0.76583	0.74567	0.91689
线性回归	0.82653	0.40213	0.89876
BP 神经网络	0.98763	0.87522	0.88654

由图 5 可知,基于 GP-XGBoost 建立的预测模型在水位突变的时段能够很好地对坝体水平位移进行预测,且整体精度高,拟合效果好,收敛速度快。BP 神经网络模型在预测阶段有时会出现过拟合的现象,而 XGBoost 模型加入了正则项参数,通过选择合适的参数值有效避免过拟合。由表 3 可以看出线性回归模型的拟合优度也较高,但线性回归模型优势在于反映线性趋势,呈现局部复杂变化较为困难,如 6 月底及 12 月初的线性回归模型预测数据难以体现坝体变形短期的波动,而基于 GP-XGBoost 所建立的大坝变形预测模型,预测值可以较好吻合坝体实际变形波动过程 (图 5),有助于大坝性态分析。实例中基于 GP-XGBoost 所建立的模型预测误差相较于传统模型明显较小,拟合更平滑、精度更高、模型更稳定,对于大变形预测将更具优势。参数优化能够进一步提升 XGBoost 算法模型的精确度,优化拟合效果,使模型能够很好地满足大坝变形预测的工作需求。

3.4 特征重要性分析

选择特征分裂率、特征平均收益率、特征平均覆盖率 3 个特征来衡量各分量在基于 GP-XGBoost 构建的大坝变形预测模型中占水平位移的比重。其中特征分裂率表示因子对大坝变形的参与度,通过统

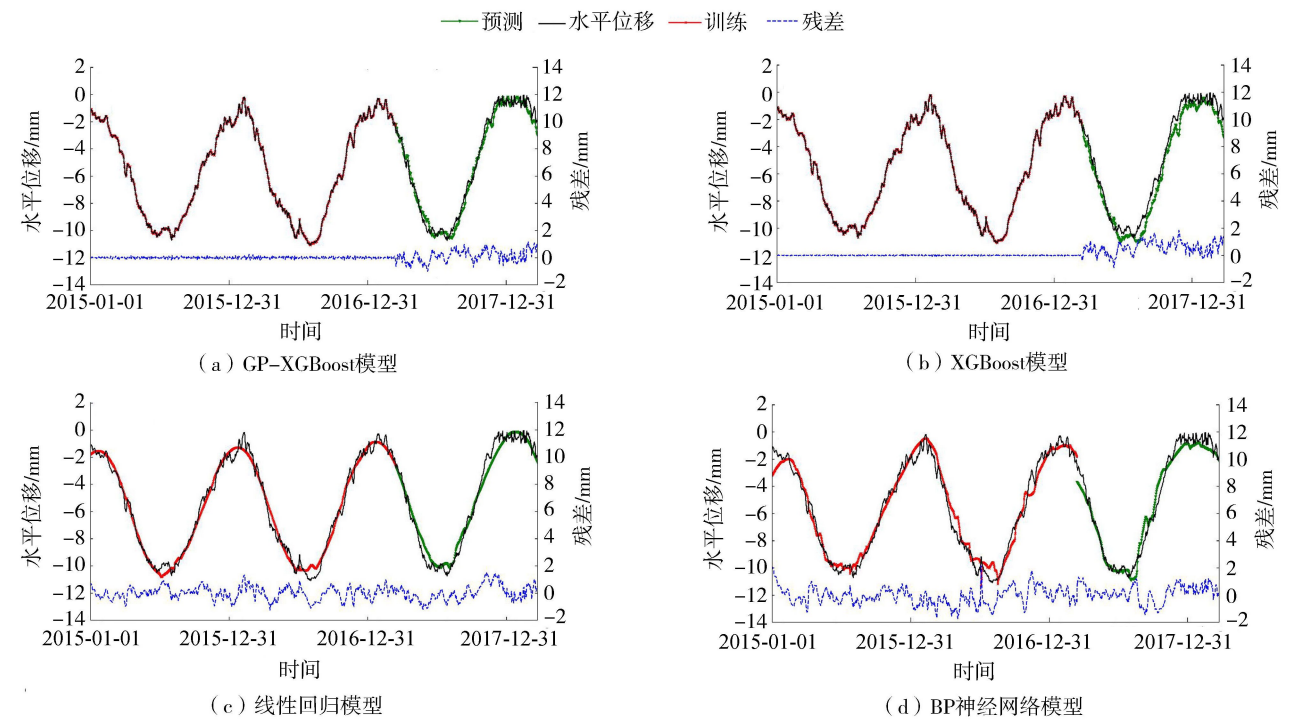


图5 各模型预测结果对比

计某个特征的分裂次数确定。特征平均收益率表示因子对大坝变形的影响率,通过损失函数的变化 ΔL 度量。特征平均覆盖率表示因子对大坝变形的贡献度,通过节点的二阶梯度 h_i 表征,特征平均收益率与特征平均覆盖率的计算方法如式(11)(12)所示。其中各因子互相独立,分别进行计算。

$$L = L(y_i, \hat{y}_{iK}) - L(y_i, \hat{y}_{i,K-1}) \tag{11}$$

$$h_i = \frac{\partial^2 L(y_i, \hat{y}_{i,K-1})}{\partial \hat{y}_{i,K-1}} \tag{12}$$

结果通过交叉验证互相比对,本次计算各因子交叉验证结果一致,具体结果如表4所示。由表4可知,温度分量占大坝变形最大比重,其特征平均收益率达98%,选择温度因子作为分裂结点可以使预测函数快速接近真实值,且温度分量对大坝变形的速率影响最大。同时温度因子特征分裂率、特征平均覆盖率都接近50%,表明温度分量与大坝变形联系较强,对变形的贡献较大。水压与时效分量对大坝变形速率影响较小,但对变形的大小有一定的影响。

表4 各因子特征重要性 单位:%

因素	特征平均覆盖率	特征平均收益率	特征分裂率
水压	30.60	0.68	20.94
温度	55.14	98.87	47.89
时效	14.26	0.45	31.17

结果表明,对坝体温度进行实时监测、控制可以有效帮助大坝健康服役,控制大坝变形速率,减少变形量,增加大坝维护工作的容错率,增强大坝变形的可控性。应对冷却管道等温控设施进行定期检查,对

于坝体温度较高部分重点防护,降低坝体开裂风险。

4 结 语

a. XGBoost 算法通过加入正则项参数,限制拟合程度,有效避免过拟合。且多决策树叠加保证了计算精度,对目标函数进行二阶泰勒展开大幅提升迭代速度,能够很好地适应大坝原型监测数据周期长、数据量大、精度要求高的特点。

b. 基于高斯过程的贝叶斯优化函数能够探索开发最优参数组,进一步提高模型的精度、稳定性,拟合更加平滑。基于磨石潭水库原型观测数据建立的大坝变形监测模型的成功应用证明,GP-XGBoost 为大坝安全监测提供了一种良好的思路。

c. XGBoost 是一种集成算法结合的树型算法,其包含了树型算法在预测过程中难以突破训练样本峰值的缺点。大坝上游水位在汛期上涨迅速,极有可能超过训练样本峰值,峰值限制问题将是今后的研究方向。

参考文献:

- [1] 毛宁宁,苏怀智,高建新. 基于FP-growth的大坝安全监测数据挖掘方法[J]. 水利水电科技进步,2019,39(5): 78-82. (MAO Ningning, SU Huaizhi, GAO Jianxin. Data mining method for dam safety monitoring based on FP-growth algorithm[J]. Advances in Science and Technology of Water Resources,2019, 39(5): 78-82. (in Chinese))
- [2] 谷艳昌,吴云星,黄海兵,庞琼. 基于遗传算法优化支持

- 向量机的大坝安全性态预测模型[J]. 河海大学学报(自然科学版), 2020, 48(5): 419-425. (GU Yanchang, WU Yunxing, HUANG Haibing, PANG Qiong. Prediction model of dam safety behavior based on genetic algorithm optimized support vector machine [J]. Journal of Hohai University (Natural Sciences), 2020, 48(5): 419-425. (in Chinese))
- [3] 胡德秀, 屈旭东, 杨杰, 等. 基于 M-ELM 的大坝变形安全监控模型[J]. 水利水电科技进展, 2019, 39(3): 75-80. (HU Dexiu, QU Xudong, YANG Jie, et al. A safety monitoring model of dam deformation based on M-ELM [J]. Advances in Science and Technology of Water Resources, 2019, 39(3): 75-80. (in Chinese))
- [4] 贾强强, 苏怀智, 郭芝韵, 等. 基于 NMEA-BP 大坝变形监测模型研究[J]. 水利水电技术, 2018, 49(1): 70-74. (JIA Qiangqing, SU Huaizhi, GUO Zhiyun, et al. NMEA-BP based study on dam deformation monitoring model [J]. Water Resources and Hydropower Engineering, 2018, 49(1): 70-74. (in Chinese))
- [5] 牛景太. 基于奇异谱分析与 PSO 优化 SVM 的混凝土坝变形监控模型[J]. 水利水电科技进展, 2020, 40(6): 60-65. (NIU Jingtai. Dam deformation monitoring model based on singular spectrum analysis and SVM optimized by PSO [J]. Advances in Science and Technology of Water Resources, 2020, 40(6): 60-65. (in Chinese))
- [6] 罗浩, 郭盛勇, 包为民. 拱坝变形监测预报的随机森林模型及应用[J]. 南水北调与水利科技, 2016, 14(6): 116-121. (LUO Hao, GUO Shengyong, BAO Weimin. Random forest model and application of arch dam's deformation monitoring and prediction [J]. South-to-North Water Transfers and Water Science & Technology, 2016, 14(6): 116-121. (in Chinese))
- [7] 沈晶鑫, 房彬, 郑东健, 等. 基于 AdaBoost 集成的 WPSO-RBF 大坝变形监控模型[J]. 长江科学院院报, 2018, 35(5): 57-62. (SHEN Jingxin, FANG Bin, ZHENG Dongjian, et al. dam deformation monitoring by radial basis function model optimized by particle swarm optimization with inertia weight and AdaBoost [J]. Journal of Yangtze River Scientific Research Institute, 2018, 35(5): 57-62. (in Chinese))
- [8] 王晓玲, 谢怀宇, 王佳俊, 等. 基于 Bootstrap 和 ICS-MKELM 算法的大坝变形预测[J]. 水力发电学报, 2020, 39(3): 106-120. (WANG Xiaoling, XIE Huaiyu, WANG Jiajun, et al. Prediction of dam deformation based on Bootstrap and ICS-MKELM algorithms [J]. Journal of Hydroelectric Engineering, 2020, 39(3): 106-120. (in Chinese))
- [9] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system [EB/OL]. [2020-06-28]. https://www.researchgate.net/publication/310824798_XGBoost_A_Scalable_Tree_Boosting_System.
- [10] 钟颖, 邵毅明, 吴文文, 等. 基于 XGBoost 的短时交通流预测模型[J]. 科学技术与工程, 2019, 19(30): 337-342. (ZHONG Ying, SHAO Yiming, WU Wenwen, et al. Short-term traffic flow prediction model based on XGBoost [J]. Science Technology and Engineering, 2019, 19(30): 106-120. (in Chinese))
- [11] 王国田, 戴筠一, 许少钧, 等. 基于 XGBoost 算法的实验室安全风险预测模型研究[J]. 实验技术与管理, 2019, 36(12): 245-251. (WANG Guotian, DAI Junyi, XU Shaojun, et al. Research on laboratory safety risk prediction model based on XGBoost algorithm [J]. Experimental Technology and Management, 2019, 36(12): 245-251. (in Chinese))
- [12] 李惠萍, 胡安民. 机器学习 DNN 和 XGBoost 算法对危重患者预后预测模型效能评估[J]. 实用医学杂志, 2020, 36(4): 466-469. (LI Huiping, HU Anmin. Machine learning for prediction of all-cause mortality in critically ill patients [J]. The Journal of Practical Medicine, 2020, 36(4): 466-469. (in Chinese))
- [13] SU H Z, HU J, WU Z R. A study of safty evaluation and early-warning method or dam global behavior [J]. Structural Health Monitoring, 2012, 11(3): 269-279.
- [14] 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.
- [15] SU H Z, WEN Z P, WANG F, et al. Dam safety prediction model considering chaotic characteristics in prototype monitoring data series [J]. Structural Control Health Monitoring, 2016, 15(6): 639-649.
- [16] 陈剑强, 杨俊杰, 楼志斌. 基于 XGBoost 算法的新型短期负荷预测模型研究[J]. 电测与仪表, 2019, 56(21): 23-29. (CHEN Jianqiang, YANG Junjie, LOU Zhibin. A new short-term load forecasting model based on XGBoost algorithm [J]. Electrical Measurement & Instrumentation, 2019, 56(21): 23-29. (in Chinese))
- [17] 韩涛, 尹伟石, 方明. 基于卷积神经网络和 XGBoost 的情感分析模型[J]. 复旦学报(自然科学版), 2019, 58(5): 560-564. (HAN Tao, YIN Weishi, FANG Ming. Sentiment analysis model based on convolutional neural network and XGBoost [J]. Journal of Fudan University (Natural Science), 2019, 58(5): 560-564. (in Chinese))
- [18] 冯钧, 王云峰, 邬炜, 等. 城市内涝事理图谱构建方法及应用[J]. 河海大学学报(自然科学版), 2020, 48(6): 479-487. (FENG Jun, WANG Yunfeng, WU Wei, et al. Construction method and application of event logic graph for urban waterlogging [J]. Journal of Hohai University (Natural Sciences), 2020, 48(6): 479-487. (in Chinese))
- [19] 徐锋, 方彦军. 基于贝叶斯优化 XGBoost 的现场校验仪误差预测[J]. 电测与仪表, 2019, 56(18): 120-125. (XU Feng, FANG Yanjun. Error prediction of field calibrator based on Bayesian optimization XGBoost [J]. Electrical Measurement & Instrumentation, 2019, 56(18): 120-125. (in Chinese))

(下转第 70 页)