DOI:10.12170/20190908001

胡江. 特高拱坝运行初期变形监测预报模型及构建方法 [J]. 水利水运工程学报, 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))

特高拱坝运行初期变形监测预报模型及构建方法

胡 江

(南京水利科学研究院水文水资源与水利工程科学国家重点实验室,江苏南京210029)

摘要:针对特高拱坝运行初期温度场的非稳定性和时效的非单调增长性,发展变形监测和预报模型,并提出构 建方法。通过主成分上的分层聚类法选取代表性温度测点,将其实测值作为温度变量,引入包含徐变及其恢复 项的时效变量表达式,论证其表达谷幅变形的能力。进而考虑库水位、实测温度、组合时效等变量,应用增强回 归树方法提出特高拱坝运行初期变形监测和预报模型,并通过后向消减变量建立优化模型。分析各变量对变形 的边际效应,得出相对影响,借助部分依赖图,辨识变量间相关关系及其对坝体变形的影响规律,揭示变形机 制。将该模型应用于某特高拱坝,验证该模型的可行性和有效性;并将结果与支持向量机、多元回归模型进行 对比分析,得出该模型具有显著的优越性。

关 键 词: 特高拱坝;变形;运行初期;安全监测;增强回归树;主成分分析;支持向量机 中图分类号: TV642 **文献标志码:** A **文章编号:** 1009-640X(2020)05-0063-09

多座 200~300 m 级特高拱坝蓄水后短期内均监测到上下游谷幅持续收缩,超出了已有工程经验和规律 认知^[1-2]。如当水库蓄水并水位长时间处于 540~560 m 高程时,溪洛渡拱坝向上游变形,并表现出非线性和 显著时效特征,蓄水完成后谷幅并未收敛,影响拱坝运行初期变形性态。除谷幅变形外,运行初期库水温垂 直分层、水泥持续水化导致坝体温度场仍处于非稳定阶段。可见,运行初期高拱坝处于非稳定、非线性的 环境和条件下。同时,根据国际大坝委员会的统计资料,运行初期失事大坝超总失事大坝数的一半。如 Malpasset 拱坝蓄水造成坝基岩体非均匀大变形、坝体溃决,Vajont 拱坝蓄水诱发库岸滑坡,库水翻坝而过、 大坝失事,两座拱坝从蓄水到发生事故分别只经过了 5 年和 3 年^[3-4]。可见,建立科学的运行初期安全监测 和预报模型,对掌握大坝安全状态、防患于未然意义重大。

变形是监测大坝安全状况的最主要效应量^[5-6]。传统统计模型(Hydrostatic Season Time, HST)应用最广 泛,但只适用于时效不显著、温度场稳定的运行期^[5]。基于实测温度值的模型(Hydrostatic Temperature Time, HTT)采用代表性的温度测点温变作为温度变量,能在一定程度上弥补 HST 模型的不足^[6]。多元回归 模型假设变量间独立,效应量是各变量分量的线性叠加,通过变量分离判断各变量对效应量的影响。当变 量间多重共线性时,模型精度差。基于数据的非参数方法如神经网络、支持向量机(Support Vector Machines, SVM)能克服变量间共线性的影响,可建立高精度的非线性模型,在近 30 年得到了大量应 用^[7-8]。但这些方法也存在难以解释大坝性态的缺陷。增强回归树(Boosted Regression Trees, BRT)方法除 有非参数方法的优点外,还能在运算过程中随机抽取数据分析变量对效应量的影响,从而为安全监测和预 报模型提供了新途径^[9]。

本文研究特高拱坝运行初期变形监测预报模型。通过基于主成分的分层聚类法(Hierarchical Clustering

收稿日期: 2019-09-08

基金项目: 国家重点研发计划资助项目(2018YFC0406705);国家自然科学基金资助项目(51879169,51779155)

作者简介: 胡 江(1983—), 男, 湖南衡阳人, 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 on Principal Components, HCPC)选取代表性温度测点,将其测值作为温度变量;引入包含徐变及其恢复项的 时效变量表达式,论证其表达谷幅变形的能力;进而考虑库水位、实测温度、组合时效等变量,应用增强回 归树方法提出特高拱坝运行初期变形监测和预报模型,并通过后向消减变量建立优化模型(Simple BRT, SBRT);采用相对影响(Relative Influence, *I*_R)表示各变量对变形的影响,并借助部分依赖图(Partial Dependence Plot, PDP)探寻变量间相关关系及其对坝体变形的影响规律。将该模型应用于某特高拱坝,并 将结果与 SVM, HST 及 HTT 模型对比分析,验证模型的可靠性和优越性。

1 监测预报模型

坝体温度变形在拱坝总变形中占较大比重。从力学观点看,温度变形与温度变化呈线性关系,温度变 形分量应选择温度计测值作为因子^[5]。当坝体和边界设置足够数量的温度计并连续观测时,实测值足以描 绘坝体变温场。此时,温度变形分量 δ_T 可表示为:

$$\delta_T = b_i \sum_{i=1}^N T_i \tag{1}$$

式中: bi 为待定系数; Ti 为测点的温度值; N 为选用温度计的数量。

时效是运行初期变形的显著特征,其特点是蓄水后急剧变化,而后随运行时间的增长逐步稳定。时效 包括刚固地基上坝体徐变、完整岩体徐变及岩体裂隙节理塑性变形等。同时,相关研究也表明,坝体和岩体 的徐变在卸荷后有一定的恢复^[5,10]。

徐变和塑性变形引发的时效变形 δ_{θ} 是随时间单调增加的函数,用一阶衰减微分方程的解来描述^[5]:

$$\delta_{\theta} = c \left(1 - \exp(-\gamma \theta) \right) \tag{2}$$

式中: c, y 均是待定常数; θ=t/100, t 为观测日至始测日的天数。

当库水位呈周期性变化时,徐变恢复项δ,可表示为^[5]:

$$\delta_{\rm r} = \sum_{i=1}^{2} \left(d_i \sin \frac{2\pi i t}{365} + f_i \cos \frac{2\pi i t}{365} \right) \tag{3}$$

式中: di, fi 均是待定常数, i=1~2, 这里取 1。

综合式(1)~(3),并考虑水压变形分量可得到高拱坝蓄水后运行初期的变形预报模型为:

$$\delta = \sum_{i=1}^{4} a_i H_i + \sum_{i=1}^{N} b_i T_i + c \exp(-\gamma t) + d \sin \frac{2\pi t}{365} + f \cos \frac{2\pi t}{365} + g$$
(4)

式中:H为水库水深;g为待定常数。

2 SBRT 的构建及实现方法

2.1 基于 HCPC 的温度因子选取

HTT 模型可显著提高复杂温度条件时的模型精度⁶⁰。但对于特高拱坝,埋设的温度计数量众多,如直接 选取全部的库水温、坝体温度等实测数据,会引起过多冗余变量和多重共线性,降低模型的鲁棒性。采用 HCPC 法,依据各点测值的变化规律,将测点分为不同的组,并选取各组中典型测点的温度测值作为模型的 温度变量。 PCA 的基本原理可参考文献 [11]。当 PCA 和分层聚类均采用欧氏距离度量时,可融合得到 HCPC 法, 从而更好地描述变量间的关系。对于数据集 *X_{KI}(K* 为变量个数,*I* 为观测数), PCA 的核心思想是采用 *S* 个 (*S*<*K*)不相关的主成分表征原始数据集,这些主成分反映了原始数据集中多变量蕴含信息的本质,其余的主 成分则可认为是噪音。在前 *S* 个主成分上进行分层聚类,可以获得比在原始变量上聚类更稳健的结果。可 见, PCA 也是一种去噪方法,可作为聚类分析的预处理步骤。分层聚类基于 PCA 和多维方差,采用 Ward 准 则作为聚集方法。该准则基于 Huygens 定理,分解组间、组内方差得到总方差:

$$\sum_{k=1}^{K} \sum_{q=1}^{Q} \sum_{i=1}^{I_q} \left(x_{iqk} - \bar{x}_k \right)^2 = \sum_{k=1}^{K} \sum_{q=1}^{Q} I_q \left(\bar{x}_{qk} - \bar{x}_k \right)^2 + \sum_{k=1}^{K} \sum_{q=1}^{Q} \sum_{i=1}^{I_q} \left(x_{iqk} - \bar{x}_{qk} \right)^2 \tag{5}$$

式中: x_{iqk} 为第 q 组的变量 i 在观测点 k 的值; \bar{x}_k 为各变量在观测点 k 的平均值; \bar{x}_{qk} 为第 q 组内各变量在观测 点 k 的平均值; I_q 为第 q 组内的观测数。

组内方差表征了组内变量的同质性, Ward 准则在聚类时使每个步骤中组内方差增长最小,即组间方差 减少最小。确定分组数是聚类分析的核心问题。分层聚类本质是一种嵌套分区,最底层上每个变量均是一 个小组,最顶层上所有变量都归属同一个大组。可根据组内方差的增长情况判断最优的聚类分组结果。当 分组数从 (*Q*-1) 到 *Q* 时的组间方差的增加值 Δ*Q* 远大于从 *Q* 到 (*Q*+1) 时的增加值 Δ(*Q*+1) 时,最优分组数 为 *Q* 组。可采用两种方法来获得最终的聚类分区,一是保持分层聚类得到的 *Q* 个分组;一是借助 Kmeans 算法,将分层聚类结果作为初始分区,通过若干次迭代获得改进的分区结果,迭代过程中通过组间方 差的比值判断。通常,初始分区不会被完全替换,而是得以改进。

变量的分层聚类分区结果可表示在主成分映射图上或树形图上。

2.2 基于 SBRT 的模型构建

BRT 综合了回归树和 Boosting 增强算法。通过回归树拟合一组单模型,并使用增强算法组合回归树的输出以计算总体预测值。BRT 的核心在于,每一棵树是从之前所有树的残差中学习的,利用残差梯度来优化回归树的集成过程。

回归树基于类似样本组中训练数据的递归划分,将特征空间划分为不同的区域,给每个区域以相应的 常数,通过将数据划分至不同区域进而得到预测值,其输出是每个组内观察的输出变量的平均值。回归树 每次生成树的子节点只有2个,即递归地二分每个特征,采取平方误差作为评价指标,在每一步选择一个最 好的特征来分裂。这样,将输入空间即特征空间划分为有限个单元,并对应以相应的数值。当考虑多个变 量时,计算每个变量的最佳节点,并选择导致误差减小最大的节点。因为一个节点中弱相关变量的误差减 少一般低于强相关变量,所以算法自动舍弃不相关的变量。Boosting 算法基于训练数据生成的多个简单模 型,通过集合中的每个模型的输出的加权和实现整体预测。算法的基本思想是让每个学习者适应前一个集 合的残差。

回归树和平方误差损失函数时,原始增强算法的主要步骤可总结如下。 用观测的平均值进行预测:

$$F_0\left(X^j\right) = f_0\left(X^j\right) = \bar{y}_i \tag{6}$$

对于 *m*=1, 2, …, *M*, 计算训练集的误差:

$$\tilde{y}_i = y_i - F_{m-1} \left(x_i^i \right) \tag{7}$$

绘制训练集的随机子样 Sm;考虑 Sm,基于前一个集合的残差拟合新的回归树:

$$\tilde{y}_i \approx f_m(X^j), \quad i \in S_m$$
(8)

更新集合:

$$F_m(X^j) \leftarrow F_{m-1}(X^j) + f_m(X^j) \tag{9}$$

 F_m 是最终的模型。通常认为该过程易于过拟合,因为训练误差随迭代而减小。为了克服这一问题,添加正则化参数 $v \in (0,1)$,从而上一步可以变换为:

$$F_m(X^j) \leftarrow F_{m-1}(X^j) + \upsilon \cdot f_m(X^j) \tag{10}$$

已有研究表明,相对较小的正则化参数(v<0.1)可极大提高泛化能力¹⁹。通常设较小的正则化参数并考虑多个树,使得误差稳定。随后,使用交叉验证来优化参数。本研究中设定 v=0.001,树的数量上限为 10 000。应用十重交叉验证来确定最终集合中树的数量,构建 BRT 模型。在此基础上,通过后向消减影响 较小的变量,优化建立更为稳健的 SBRT 模型。

2.3 变量相对影响大小的确定

训练过程主要变量被频繁地选取,次要变量被舍弃,变量 x_i的 I_R 与它们出现的频率成比例。

$$I_j^2 = \sum_{x_j} I_t^2 \tag{11}$$

式中: I_j 为变量 x_j 的相对影响。 I_i^2 是在这一节点上 x_j 的改进。变量 x_j 的 I_R 是 Boosting 算法所有生成树的平均值。根据 I_R 值来确定变量和效应量的关联性。

2.4 重要变量的部分依赖图

使用 PDP 通过边际效应识别变量-效应量间的关系。 X^{j} 为变量,在其范围内定义一组等间距值,即 $X^{j} = x_{k}^{j}$ 。对于这些值,模型效应量的平均值计算如下:

$$\bar{F}\left(x_{k}^{j}\right) = \frac{1}{N} \sum_{i=1}^{N} F\left(x_{i}^{j}, x_{i}^{jc}\right) \tag{12}$$

式中:x^{jc}是所有变量而非X^j的值。

2.5 监测预报模型的实现流程和性能评价

基于 SBRT 方法的特高拱坝运行初期变形监测和预报模型的实现步骤如下:①监测数据预处理;②基 于 HCPC 选取典型实测温度变量,确定合适的时效变形表达式;③构建 BRT 模型,通过后向消减建立 SBRT 模型;④计算各变量的 *I*_R,生成部分 PDP,判定变量对效应量的影响。

通过平均绝对误差(Mean Absolute Error, eMAE)分析模型的拟合效果,定义如下:

$$e_{\text{MAE}} = \frac{\sum_{i=1}^{N} \left| y_i - F\left(x_k^j\right) \right|}{N} \tag{13}$$

式中:N为训练或者预测样本集的大小;yi为实测的监测效应量值;F(x)为预测值。

3 工程应用

3.1 工程和监测概况

某混凝土双曲拱坝坝顶高程 610.0 m, 建基面高程 324.5 m, 共 31 个坝段。水库的正常蓄水位、死水位 和汛限水位分别为 600.0, 540.0 和 560.0 m, 水库具有年调节能力。坝体变形布设坝体垂线监测系统和坝后 桥外观大地测量两个监测系统。在坝体典型高程上、下游面各布置 1 支温度计观测环境温度, 内部则采用 应变计组的测温传感器观测内部温度。15#坝段垂线和 16#坝段的温度测点布置如图 1, 其中, T 为温度计, S(*k-i*) 表示 *k* 向应变计第 *i* 个测点。





Fig. 1 Layout of embedded pendulums and temperature sensors in Monolith 16

2013 年 5 月 4 日,导流底孔下闸水库开始蓄水 (图 2)。2014 年 3 月 6 日,大坝完成浇筑。2014 年 9 月 28 日,第一次蓄至正常蓄水位。当前谷幅出现 较大变形且仍未稳定(图 2)。通过比较 15#坝段相 同高程两套变形监测系统测值,发现两者在顺河向 变位观测值存在 3 mm 的差值,但变化规律一致^[12]。

3.2 变量选取

库水位、温度为每日平均测值, 垂线测点的监测 Fig. 2 Time histories of 频次为 2~3 d/次。分析的 PL15-1 的时间序列为 displacment and valle 2014 年 6 月—2018 年 6 月, 在此期间坝体经历了 4 次完整的加载和卸荷过程。

温度变量采用 16#坝段实测温度作为相邻 15#坝段的数据。典型测点实测温度过程如图 3。可看出,高高程内部测点温度仍有上升的趋势,而坝基内部测点温度变化跟位置有关。应用 HCPC 法,以 PL15-5 段为例,根据 Ward 准则对温度测点聚类,最终确定为 4 类(图 4),选取的该垂线段的温度计列于表 1,同理,其余垂线段的结果也列于表 1。高高程垂线测点模型构建还应考虑其下所有典型温度测点。







图 2 蓄水、坝顶径向位移和谷幅变形过程线

2 Time histories of reservoir water level, dam crest displacment and valley contraction deformation

图 4 PL15-5 库水温和典型段坝体温度测点 HCPC 结果

自蓄水以来,在坝址上下游侧均观测到了显著的、持续收缩的谷幅变形。如图 2,坝肩 VDL04 测线获得的谷幅变形呈指数函数变化,且存在小的周期性波动。因此,由式(2)和(3)组合表示的时效变量拟合谷幅变形,采用非线性最小二乘法得 $\gamma=0.15$, $R^2=0.993$, 拟合效果如图 5。可见,时效变量 δ_{θ} 、 δ_{r} 还可以较好地表达谷幅变形。最终考虑的变量包括库水位、选取的实测温度(表 1)、时效变量 δ_{θ} 和 δ_{r} 。

表 1 模型构建考虑的实测温度变量

Tab. 1 Measured temperature variables considered in model construction

垂线段	PL15-5	PL15-4	PL15-3	PL15-2	PL15-1
选取的温度计	T6, S3-1, S6-4, S6-9	T14, T15	S5-4, T18	S5-5, T28	T29

3.3 模型构建

一般地,选取样本总数的 10%~20% 作为预测 集^[13],考虑到运行初期的非线性,选取 15% 作为预测 集。SBRT 模型构建时的参数按 3.2 节方法选取。 十重交叉验证得到的最优的树的数量为 7 050。为 对比分析,还构建了 SVM,HTT,HST 等 3 种模型, 其中,SVM 的因子与 BRT 完全一致。SVM 也采用 了十重交叉验证优化参数,选择最佳的惩罚因子、不 敏感系数取值范围分别为 [10,1 000], [0.000 1, 1],最 优值分别为 1 000 和 0.001。HTT 和 HST 采用最小

二乘法拟合。HTT 模型的实测温度因子与 SBRT 模型相同。各模型的训练和预测的精度列于表 2, SBRT 模型最终选取的变量为 H, δ_{θ} , δ_{r} , S6-9, S6-4 和 T_{a} 等(其中 T_{a} 为空气温度)。各模型的训练和预测效 果如图 6。

表 2 各模型的训练集和预测集的平均绝对误差 单位: mm Tab. 2 Comparisons of mean absolute errors among training and prediction sets of constructed models SVM HST SBRT HTT 模型 训练 预测 训练 预测 训练 预测 训练 预测 平均绝对误差 0.42 0.84 0.95 1.49 3.07 5.96 1.83 6.62

3.4 部分依赖图

PDP 基于模型学习得到每个变量对效应量的边际效应,为此,可通过 PDP 识别变量-效应量间的关系。 PDP 可以排除其他变量,仅直观显示一个或一组变量与效应量响应间的关系。因篇幅限制,仅列出了 *H* 和 *δ*_θ, *H* 和 *T*_a 与 PL15-1 的三维 PDP 关系, 如图 7。

4 结果讨论

4.1 变量选取

从图 3、表 1 可看出, 388.7 m 高程以下库水温基本稳定, 温度变幅较小, 选取了 T6, S3-1, S6-4, S6-9 等 下游和内部共 4 个测点; 400~480 m 高程库水温存在年变化, 年变幅小于气温, 且与气温存在较大的相位差, 内部温度主要受环境温度影响, 选取了 T14 和 T15 等 2 个环境温度测点; 500~540 m 高程库水温存在明显 的年变化, 但坝体内部水化热温升明显, 选取了 S5-4, T18, S5-5, T28 等下游和内部共 4 个测点; 540 m 高程 库水温受上游来水温度和气温双重影响, 但主要受气温影响, 选取了上游测点 T29。可见, 上游面在温跃 层、表温层均选取了典型测点(T15 和 T29)。各区坝体温度从封拱后都存在温升现象, 内部测点主要受水 化热控制, 以单调温升为主。底高程温度渐趋稳定, 高高程受水化热影响仍持续温升, 选取的 S5-4 和 S5-5 测点(图 3)尤为典型。BRT 初步选定的变量为 19 个, SBRT 削减了影响较小的变量, 仅选取 *H*, δ_θ, δ_r, S6-9, S6-4 和 *T*_a 等 6 个变量。

4.2 模型对比分析

由表 2 可看出, SBRT 模型的训练集和测试集精度最高, 其次为 SVM 模型。HTT 和 HST 模型的精度 相对较差, HST 模型最差。

从 SBRT 模型及其 *I*_R 分析结果可知, 最主要的影响因素是 δ_θ, *H* 和 δ_r, 其值分别为 47.15%, 27.88% 和 9.80%。可见, 坝体和坝基的徐变及其恢复项对运行初期的变形影响较大, 水压荷载次之。对比 HTT 和 HST 模型, 分析 HTT 模型能否在一定程度上反映坝体的非稳定温度场, 从而改善 HST 模型的缺陷。

通过非参数方法和多元线性模型对比可知,受非线性时效和非稳定温度场影响,坝体变形表现出显著的非线性,HST和HTT等多元线性模型不适应,其精度不足以准确预测预报。SBRT和SVM均具备较高的精度,但SVM不能较好地联系和分析坝体变形机制,SBRT根据*I*_R判断变量对效应量的影响大小,并使用PDP通过边际效应识别变量对效应量的影响规律。

4.3 变形机制分析

由图 7 可知,库水位上升时,坝顶变形与库水位之间近似呈线性关系,且变形与库水位变化之间具有同步性、连续性的特征。库水位的多次式影响不显著。

由于坝体、坝基徐变的影响,运行初期坝体整体表现出向上游的变形趋势,2015年6月后变形速率有 所减小,但未收敛。加载卸荷引起的弹性位移也有相对较强的影响。考虑到时效能很好地表示谷幅变形, 可将坝体总体出现相对向上游变形的特征归结于谷幅变形。此外,坝体内部温度的缓慢回升也使得坝体产 生了轻微的向上游变形^[14-15]。可见,蓄水初期特高拱坝的变形是一个复杂的非线性过程,其影响因素主要包 括时效、库水位、温度回升引起的回弹变形等。

5 结 语

本文提出了特高拱坝运行初期变形监测和预报模型,得到以下主要结论:

(1)通过 HCPC 方法选取典型温度测点,以其温度测值作为温度变量,减少了实测温度变量的冗余,温度变量选取结果也反映了环境和坝体内部温度非稳定特征。

(2)引入包含徐变及其恢复项的运行初期时效变量,并对谷幅变形进行了拟合,结果表明时效变量能较 好地表征谷幅变形。

(3)提出了基于 SBRT 方法的运行初期变形监测和预报模型。工程应用表明, SBRT 模型具有高拟合和 预测精度, 并能解释各变量的影响大小; SVM 模型次之; HTT 和 HST 模型精度差, 不适用于运行初期变形 预测。借助变量相对影响和部分依赖图, SBRT 模型可直观体现主要变量对变形的影响, 有利于特高拱坝运 行初期的安全监控和管理。可见, 基于 SBRT 的特高拱坝运行初期变形监测和预警模型具有显著优越性。

参考文献:

- [1] 张冲, 王仁坤, 汤雪娟. 溪洛渡特高拱坝蓄水初期工作状态评价[J]. 水利学报, 2016, 47(1): 85-93. (ZHANG Chong, WANG Renkun, TANG Xuejuan. Safety evaluation of Xiluodu ultra-high arch dam during the initial impoundment period[J]. Journal of Hydraulic Engineering, 2016, 47(1): 85-93. (in Chinese))
- [2] 何柱, 刘耀儒, 杨强, 等. 溪洛渡拱坝谷幅变形机制及变形反演和长期稳定性分析[J]. 岩石力学与工程学报, 2018, 37(增刊2): 4198-4206. (HE Zhu, LIU Yaoru, YANG Qiang, et al. Mechanism of valley deformation of Xiluodu arch dam and back analysis and long-term stability analysis[J]. Chinese Journal of Rock Mechanics and Engineering, 2018, 37(Suppl2): 4198-4206. (in Chinese))
- [3] 刘宁. 国内外大坝失事分析研究[M]. 武汉: 湖北科学技术出版社, 2002. (LIU Ning. Analysis and research of damfailure[M]. Wuhan: Hubei Science and Technology Press, 2002. (in Chinese))
- [4] 顾冲时,苏怀智,王少伟. 高混凝土坝长期变形特性计算模型及监控方法研究进展[J]. 水力发电学报, 2016, 35(5): 1-14.
 (GU Chongshi, SU Huaizhi, WANG Shaowei. Advances in calculation models and monitoring methods for long-term deformation behavior of concrete dams[J]. Journal of Hydroelectric Engineering, 2016, 35(5): 1-14. (in Chinese))
- [5] 顾冲时, 吴中如. 大坝与坝基安全监控理论和方法及其应用[M]. 南京: 河海大学出版社, 2006. (GU Chongshi, WU Zhongru. Theory and method of dam and dam foundation safety monitoring and their applications[M]. Nanjing: Hohai University Press, 2006. (in Chinese))
- [6] HU J, WU S H. Statistical modeling for deformation analysis of concrete arch dams with influential horizontal cracks[J]. Structural Health Monitoring, 2019, 18(2): 546-562.
- [7] 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.
- [8] KANG F, LI J J, DAI J H. Prediction of long-term temperature effect in structural health monitoring of concrete dams using support vector machines with Jaya optimizer and salp swarm algorithms[J]. Advances in Engineering Software, 2019, 131: 60-76.
- [9] SALAZAR F, TOLEDO M Á, GONZÁLEZ J M, et al. Early detection of anomalies in dam performance: A methodology based on boosted regression trees [J]. Structural Control and Health Monitoring, 2017, 24(11): e2012.
- [10] 彭卫, 杜时贵. 混凝土徐变效应的计算机分析[J]. 西安公路交通大学学报, 1999, 19(2): 4-5, 9. (PENG Wei, DU Shigui. Computational analysis of concrete creep effect[J]. Journal of Xi'an Highway University, 1999, 19(2): 4-5, 9. (in Chinese))
- [11] ABDI H, WILLIAMS L J. Principal component analysis [J]. WIREs Computational Statistics, 2010, 2(4): 433-459.
- [12] LI Q B, ZUO Z, HU Y, et al. Smart monitoring of a super high arch dam during the first reservoir-filling phase[J]. Journal of

Aerospace Engineering, 2017, 30(2): B4016001.

- [13] SALAZAR F, TOLEDO M A, OÑATE E, et al. An empirical comparison of machine learning techniques for dam behaviour modelling[J]. Structural Safety, 2015, 56: 9-17.
- [14] 张国新, 周秋景. 特高拱坝封拱后温度回升及影响研究[J]. 水利学报, 2015, 46(9): 1009-1018. (ZHANG Guoxin, ZHOU Qiujing. Study on temperature recovery after arch sealing of super high arch dams and its influence[J]. Journal of Hydraulic Engineering, 2015, 46(9): 1009-1018. (in Chinese))
- [15] 高俊, 黄耀英, 万智勇, 等. 含冷却水管混凝土坝温度计埋设位置优选[J]. 水利水运工程学报, 2017(3): 93-98. (GAO Jun, HUANG Yaoying, WAN Zhiyong, et al. Optimization analysis of setting location of thermometer in concrete dam with cooling water pipes[J]. Hydro-Science and Engineering, 2017(3): 93-98. (in Chinese))

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

HU Jiang

(State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China)

Abstract: The dam temperature field has not been stabilized, and the time dependent effect does not increase nonmonotonously for super high arch dams during initial operation period. Therefore a special deformation monitoring and forecasting model was developed, and its modeling method was proposed in the study. The key temperature measurement points were chosen by the hierarchical clustering on principal component method, and the corresponding time series were inputted as thermal predictors. The combined time dependent effect, including creep and its restoration, was introduced as time dependent deformation. This time dependent effect was employed to validate its characterization for the valley contraction deformation. Considering the predictor variables such as reservoir water level, adopted measured temperatures, estimated time dependent effect, the simple boosted regression tree (BRT) based dam deformation prediction model was constructed. Through the backward elimination method, a simplified BRT (SBRT) model only including major predictors was obtained. The marginal effects of variables on deformation were analyzed, and the relative influences can be quantitatively analyzed. With the help of partial dependence plot, the correlations among variables and the influences of variables on deformation can be explored, and the deformation mechanism can be revealed. The model was applied to a super high arch dam, and the case study verified the feasibility of the model. The results were compared with those by the support vector machine model and the traditional multiple regression models, which shows the superiority of the developed model.

Key words: super high arch dam; deformation; initial impoundment period; safety monitoring; boosted regression tree; principal component analysis; support vector machine