

温度变形效应认知不确定性影响下高拱坝位移置信区间的预测方法

隋旭鹏¹, 王少伟¹, 郅俊力²

(1. 常州大学城市建设学院, 江苏 常州 213164; 2. 江苏省常州高级中学, 江苏 常州 213004)

摘要:为解决传统预测模型未考虑拱坝变形因果关系的不确定性,导致所建立的位移静态置信区间缺乏合理的因果机理,提出了温度变形效应认知不确定性影响下高拱坝位移置信区间的预测方法。采用动态时间规整法衡量坝体温度测点之间的时间序列相似性,构建最小相似性实测温度变形因子筛选准则,基于支持向量机构建不放回采样和正交试验设计采样的机器学习模型,通过统计多模型预测值的分布规律来拟定动态变化的置信区间。以锦屏一级拱坝为例的预测结果表明该预测方法及筛选准则可有效实现高拱坝最具表征性温度测点的优选,基于最小相似性实测温度因子的正交试验设计误差小、建模效率高,所得位移预测置信区间更符合因果机理;高拱坝的最优实测温度因子组合是动态变化的。

关键词:高拱坝;位移预测;置信区间;温度变形效应;认知不确定;正交试验设计

中图分类号:TV698.1

文献标志码:A

文章编号:1006-7647(2024)03-0095-06

Prediction method of confidence interval for displacement of high arch dams under influence of cognitive uncertainty of temperature deformation effects//SUI Xupeng¹, WANG Shaowei¹, TAI Junli² (1. School of Urban Construction, Changzhou University, Changzhou 213164, China; 2. Changzhou Senior High School of Jiangsu Province, Changzhou 213004, China)

Abstract: Traditional prediction models are difficult to consider the uncertainty of causality of arch dam deformation, leading to the lack of causality mechanism in the established static displacement confidence interval. To solve the difficulty, a prediction method which considers the influence of cognitive uncertainty of temperature deformation effects, was proposed for the confidence interval of high arch dams' displacement. The dynamic time warping method was used to calculate the similarity between the measured time series of multiple temperature monitoring points on the dam body, and a criterion was established to select the measured temperature deformation factors with the minimum similarity. Multiple support vector machine models were then established, in which the orthogonal experimental sampling of selected temperature factor was used as inputs. The distribution of multi-model predicted displacements was analyzed and used to establish a dynamic changed confidence interval. Research results of the Jinping I Arch Dam indicate that the proposed methods and criteria can effectively optimize the most representative dam temperature monitoring points. The minimum similarity temperature factor-based orthogonal experiment design has small error and high modeling efficiency. The confidence interval of displacement prediction is more consistent with the causal mechanism and the optimal measured temperature factor combination of the high arch dam is dynamic.

Key words: high arch dam; displacement prediction; confidence interval; temperature deformation effect; cognitive uncertainty; orthogonal experimental design

全生命周期内的结构健康监测在保证高坝安全运行中起着不可或缺的作用。混凝土坝位移监控模型中,水压-谐波温度-时效(hydraulic, seasonal and

time, HST)模型应用广泛,其所需外因输入数据仅为库水位^[1]。然而,HST模型中周期性谐波温度因子难以反映极端气候影响和高拱坝坝体各部位温度

基金项目:国家自然科学基金项目(51709021);中国博士后科学基金项目(2020M670387);中国水利水电科学研究院水利部水工程建设与安全重点实验室开放研究基金(202107)

作者简介:隋旭鹏(2000—),男,硕士研究生,主要从事结构健康监测与加固改造研究。E-mail:sxp1010100@163.com

通信作者:王少伟(1988—),男,副教授,博士,主要从事水工结构安全监控及老化病害研究。E-mail:shaowei2006nanjing@163.com

演变规律的差异性^[2]。从力学观点上讲,构建水压-实测温度-时效(hydraulic, temperature and time, HTT)模型是最合理的^[1],但从海量温度监测数据中如何提取少量最具表征性的有效信息是仍需进一步解决的问题。对此,Tatin 等^[3]将网格化和分层等效处理后的坝体实测温度作为建模因子;Mata 等^[4]将从坝体多测点温度时间序列中提出的主成分作为温度变形因子;Wang 等^[5-6]提出了基于时间序列形状相似度的高拱坝实测温度场空间分区方法和基于灵敏度分析的实测温度变形因子优选方法。

建模因子优化后,如何提高预测性能是另一个焦点。机器学习模型具有良好的非线性数据挖掘能力,因而被广泛应用于大坝监测效应量的预测建模中,整体而言,其预测性能普遍优于线性回归模型^[7-8],但需谨慎处理过拟合问题。此外,现有数学监控模型主要是在固定输入因子组合下进行的单个时间节点的位移确定性预测,导致所建立的位移置信区间缺乏因果机理。高拱坝温度测点数以百计,甚至上千,这使得利用坝体实测温度模拟其温度变形效应时存在较大不确定性^[6-7,9]。对于温度变形效应认知不确定性的影响,Ren 等^[10]建议将建模数据集拆分成多份并建立多个模型,然后利用多模型输出之间的偏差来构建置信区间;黄耀英等^[11]基于拟合良好性、模型简单性和检验有效性等指标,提出了大坝监测量最佳统计模型的优选方法。

为考虑温度变形效应认知不确定性的影响,并使位移预测置信区间具备合理的因果机理,本文在缩减坝体多测点温度时间序列数量和合理采样的基础上,统计输入因子不确定性影响下多模型预测位移的分布特征,据此构建高拱坝位移的预测置信区间。

1 基于坝体实测温度的高拱坝位移预测置信区间

用于高拱坝位移因果解释和预测的 HTT 模型包含水压分量 δ_H 、实测温度分量 δ_T 和时效分量 δ_θ ,其数学表达式如下^[1]:

$$\delta = \delta_H + \delta_T + \delta_\theta = \sum_{i=1}^4 a_i H^i + \sum_{j=1}^M b_j T_j + c_1 \theta + c_2 \ln \theta \tag{1}$$

其中 $\theta = t/100$
式中: δ 为高拱坝位移; H 为当日上游库水深; T_j 为经动态时间规整法(dynamic time warping, DTW)相似性分析后,优选得到的第 j 个坝体温度测点的当日实测温度; M 为选用的温度测点总数; t 为位移监测日到起始监测日的累计天数; a_i 、 b_j 、 c_1 和 c_2 为拟合系数。

温度分量是由坝体混凝土和基岩温度变化引起的位移,对其如何合理选用,目前存在较大不确定性,因而需要考虑实测温度变形因子表征方式不确定性对高拱坝位移预测的影响。本文整体思路如下:首先,通过 DTW 量化坝体多测点温度时间序列之间的相似性,构建最小相似性实测温度变形因子筛选准则;其次,基于最小相似性温度因子构建不放回采样和正交试验设计采样的多组输入因子组合,并分别建立支持向量机(support vector machine, SVM)模型,然后统计各预测时间点多 SVM 模型的预测位移分布特征;最后,按照保证率要求,构建高拱坝位移的预测置信区间。

1.1 基于 DTW 的温度时间序列相似性

DTW 可按照非线性距离最小原则来衡量两个时间序列之间的相似性,其最大优势是可建立一对多或多对一的匹配关系,从而使两者之间的总距离最小化。高拱坝气温和上游库水温对坝体温度场的滞后影响机理异常复杂,采用 DTW 构建两温度时间序列之间的对应关系,可较好地反映上述局部差异性对温度时间序列整体相似性的影响。给定两个时间序列 $Z = \{z_1, z_2, \cdots, z_n\}$ 和 $O = \{o_1, o_2, \cdots, o_m\}$,基于 DTW 的相似距离计算步骤如下:

步骤 1 构造 $n \times m$ 的距离矩阵 D ,矩阵中 (i, j) 处的元素为序列值 z_i 和 o_j 间的平方距离,即 $(z_i - o_j)^2$ 。

步骤 2 DTW 计算目标是在距离矩阵 D 中找到一条从左上角到右下角的最佳路径,使得所经过的元素值之和最小。构造路径曲线时,相邻两点的行号和列号可同时加 1 或单独加 1,按照此规则,从矩阵的左上角至右下角,所有可能的路线均需构造为路径曲线。

1.2 最小相似性温度因子优选

坝体温度监测数据中包含大量具有一定相似性的温度时间序列,若直接对所有温度时间序列进行正交试验设计,所得单个温度因子组合中会包含功能基本重复的多个温度时间序列,导致正交试验设计得到的温度时间序列组合不具有最佳的表征能力,据此得到的位移分布规律也是错误的。本文拟定的温度因子优选步骤具体如下:

步骤 1 将坝体所有可用的 p 个温度时间序列作为输入量,基于 DTW 计算任意两温度时间序列之间的相似距离,从而得到一个 $p \times p$ 的相似距离矩阵。

步骤 2 对相似距离矩阵进行归一化处理,分级拟定相似距离阈值,即两时间序列之间的相似距离如果小于该阈值,则可以按照步骤 3 中的规则来执行实测温度因子的剔除流程。

步骤3 按照以下4种相似情况进行剔除:

①情况1. 多个温度时间序列与所选温度时间序列A的相似距离小于阈值,比如,A与B、C、D3个温度时间序列相似,则剔除A,其他3个则保留。②情况2. 所选时间序列A仅与1个温度时间序列B的相似距离小于阈值,若在先前或后续中B被剔除,则应保留A,反之保留B。③情况3. 所选时间序列A仅与1个温度时间序列B的相似距离小于阈值,且两者在先前或后续均不可剔除,则保留A,剔除B。④情况4. 该温度时间序列与其他时间序列的相似距离均大于阈值,则应全部保留。

步骤4 建立保留因子数量与相似距离阈值之间的定量关系,综合考虑保留因子的表征能力和所需建模数量,拟定最佳的相似距离阈值,最终选用与此对应的温度因子。

1.3 温度因子的不放回采样和正交试验设计采样

通过相似性分析,尽管已经剔除了功能基本重复的坝体温度时间序列,但仍具有不确定性:一是温度时间序列保留数量没有严格的理论意义,具有一定的经验性;二是保留的温度时间序列是为了全面表征坝体温度场,但在建立位移预测模型时,各位移测点对同一温度测点的敏感性不同^[9],因而保留的温度时间序列仅是可供选择的上限,各位移测点仍需从中选择最优的组合。针对温度变形效应认知不确定性的影响,常用策略是变量选择与模型平均。本文在此基础上,对具有不同输入因子组合的多模型预测值进行统计分析,然后按照大坝安全监控领域的保证率要求,拟定预测位移的置信区间。

按照最小相似性原则优选后,假设共保留有 N 个温度因子,分别标记为 $T_1、T_2、\cdots、T_N$,则考虑温度变形效应认知不确定影响下,式(1)所示HTT模型中的温度分量可表示为

$$\delta_r = \chi_1 b_1 T_1 + \chi_2 b_2 T_2 + \cdots + \chi_N b_N T_N$$

$$(\chi_1, \chi_2, \cdots, \chi_N \in \{0, 1\}) \tag{2}$$

式中 $\chi_1、\chi_2、\cdots、\chi_N$ 为选择指标,若 $\chi_p = 0(p = 1, 2, \cdots, N)$,则不选择此因子;反之若 $\chi_p = 1$,则选择此因子。因此,在整个模型空间中每个温度因子均存在选中与不选中两种情况,共需建立 2^N 个SVM模型,称为全模型法。全模型法所需建模数量非常大,为减少SVM建模数量,提高计算效率,本文采用正交试验设计来优化温度因子的合理组合,通过JMP软件来构造多因素正交表,其采用的构造方法是Taguchi正交阵列法^[12]。

2 工程实例分析

锦屏一级拱坝为坝高305 m的混凝土双曲拱

坝,位于四川省凉山州境内,是世界已建最高拱坝。坝顶、正常蓄水位和死水位高程分别为1885、1880、1800 m。本次以拱冠梁坝顶PL13-1测点为例,对其径向位移建立温度变形效应认知不确定下的预测置信区间。拱冠梁坝段的温度测点最具代表性,通常可用于表征拱坝的实测温度场。锦屏一级拱坝13号拱冠梁坝段共包含140个可有效使用的温度测点,采用的监测数据时段为2015年9月1日至2018年12月31日,以2018年5月28日为分界,前1000组数据用于模型训练,后218组数据用于模型预测性能测试。

2.1 坝体最小相似性实测温度因子优选

可选用的实测温度因子数量过于庞大(共计140个),直接进行建模所需计算工作量巨大,因此按照本文构建的优选准则,剔除相似性过大的因子,提高计算效率。由于混凝土坝位移中温度分量与温度时间序列的起伏形状密切相关,因此Wang等^[5]将140个温度测点分为六类。在基于DTW相似距离进行建模因子缩减时,要兼顾坝体实测温度场的分区特征。

通过DTW计算140个温度因子之间的相似距离,可发现同一类别内的温度时间序列因波动规律基本一致而相似距离较小,尤其是演变规律基本符合周期性谐波因子的第二和第三类,而第四类测点之间的相似距离明显偏大,原因在于第四类测点所在坝体区域的实测温度演变规律最为复杂^[5]。温度因子保留数随相似距离阈值的变化规律如图1所示,在保证各类别均有实测温度因子被选中的前提下,为使保留的温度因子数最小,本文拟定的相似距离阈值为0.0041,保留温度因子数为23,如表1所示。

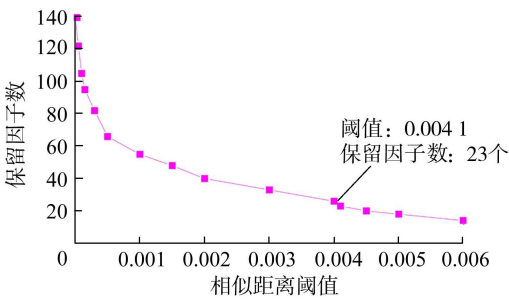


图1 温度因子保留数随相似距离阈值的变化规律

选用的23个温度测点主要位于坝体上、下游边界,尤其是受库水位变化影响较大部位,各类测点实测温度时间序列如图2所示。①第一类测点的实测温度时间序列基本呈下降趋势;②第二类和第三类测点的实测温度演变规律与年周期性谐波因子所示规律基本一致,温度年变幅相对气温已明显削弱,但

表 1 优选的最小相似性坝体温度测点

所属类别	选用温度测点
第一类	T13-5、T13-10、T13-43
第二类	T13-136
第三类	T13-53
第四类	T13-9、T13-21、T13-26、T13-82、T13-98、T13-102、 T13-114、T13-122、T13-127
第五类	T13-41、T13-71、T13-121、T13-152
第六类	T13-32、T13-62、T13-111、T13-149、T13-150

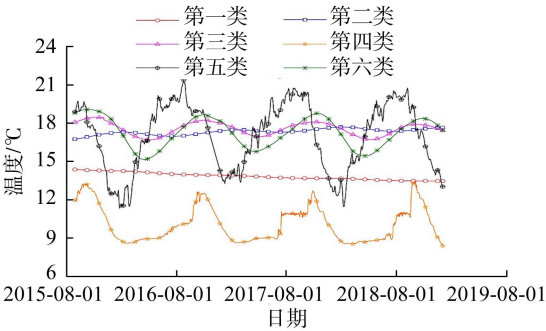


图 2 各类代表测点的实测温度时间序列

第二类测点的年变幅明显小于第三类;③第四类测点受库水位升降影响非常大,主要是因为该拱坝上游水库在气温年周期温升阶段时水位较低,因而坝体升温较快,导致不同温度时间序列在整体相似的情况下又存在细节上的较大差异;④第五类测点位于坝体下游面和坝顶,直接受太阳辐射影响;⑤第六类测点则较为均匀地出现在坝体温度场第六分区的 3 个子区域中。综上,经 DTW 相似距离计算后,按照本文构建的筛选准则,最终保留的最小相似性温度因子能反映该拱坝实测温度场的演变规律,且聚集程度低,可最大程度保证所选实测温度变形因子的表征能力和建模的高效性。

2.2 位移区间预测模型

在 HTT 模型中,不考虑水压分量和时效分量的不确定性,因此各模型中均包含式(1)所示的 4 个水压因子和 2 个时效因子,选用的 23 个实测温度因子则通过不放回采样和正交试验设计采样来考虑其不确定性的影响。不放回采样的全模型组需要建立的 SVM 模型数为 2^{23} ,即 8 388 608;对 23 个实测温度因子设计 23 因子 2 水平正交试验,总模型数为 32,大大缩减了建模时间。此外,为验证进行最小相似性温度因子优选的必要性,同时设置了 140 因子的正交试验设计组,对应的模型数为 148。

全模型组和正交试验设计组分别建立规定数目的 SVM 模型后,对多模型预测值进行统计分析,可发现各预测时间点预测位移的分布规律符合正态分布,在预测阶段中随机选取两个预测时间点,所得结果如图 3 所示,其他预测时间点的分布规律与此类似。从图 3 可以看出,23 因子的全模型组和正交试

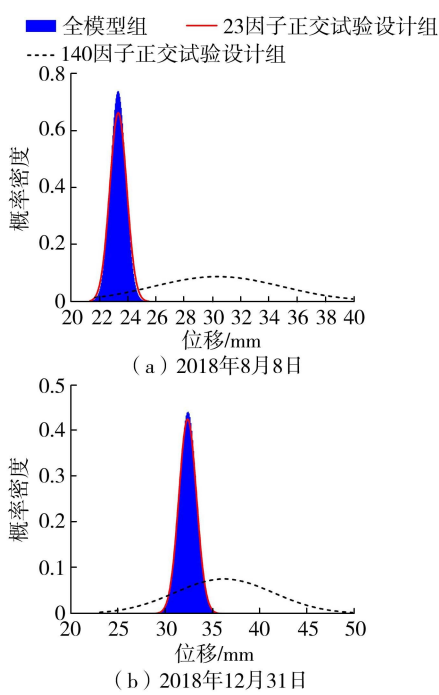


图 3 预测阶段全模型组和正交试验设计组的多模型位移预测值分布规律

验设计组所得位移分布规律和均值基本一致,由此验证了后者所用因子组合的合理性,但后者的模型数远少于前者,23 因子正交试验组在均值上的概率密度略小于全模型组;然而,不对实测温度因子进行最小相似性优选情况下,由于初始 140 个温度因子中包含了大量非常相近的实测温度时间序列,导致按照正交试验设计所建 148 个模型并不具备代表性,与 23 因子的全模型组和正交试验设计组在位移预测均值和分布区间上均出现了明显偏差,位移预测区间明显增大,其原因就在于模型数偏少,没有包含最优输入因子组合。

为进一步验证对最小相似性温度因子采用正交试验设计的合理性,给出预测阶段 23 因子的全模型组和正交试验设计组所得各预测时间点的预测均值及 95% 保证率时的最大值和最小值,如图 4 所示。由图 4 可知,两种方法所得各预测时间点的特征值基本一致,偏差较小,预测均值的最大偏差为 0.066 mm,仅为该预测时间点全模型组预测均值的 0.52%;偏差较大预测时间点主要出现在库水位下降至死水位 1 800 m 的短暂阶段以及库水位突变阶段,其原因就在于库水位的突变将直接导致坝体温度场边界的突变,由此引起实测温度时间序列的短期波动,而正交试验设计组由于模型数非常少,难以反映此细微影响。整体而言,正交试验设计组可反映全模型组的结果,可据此构建位移的预测置信区间。

将正交试验设计组中 32 个模型 2018 年 5 月 15 日至 2019 年 1 月 15 日的 218 个位移预测值分别

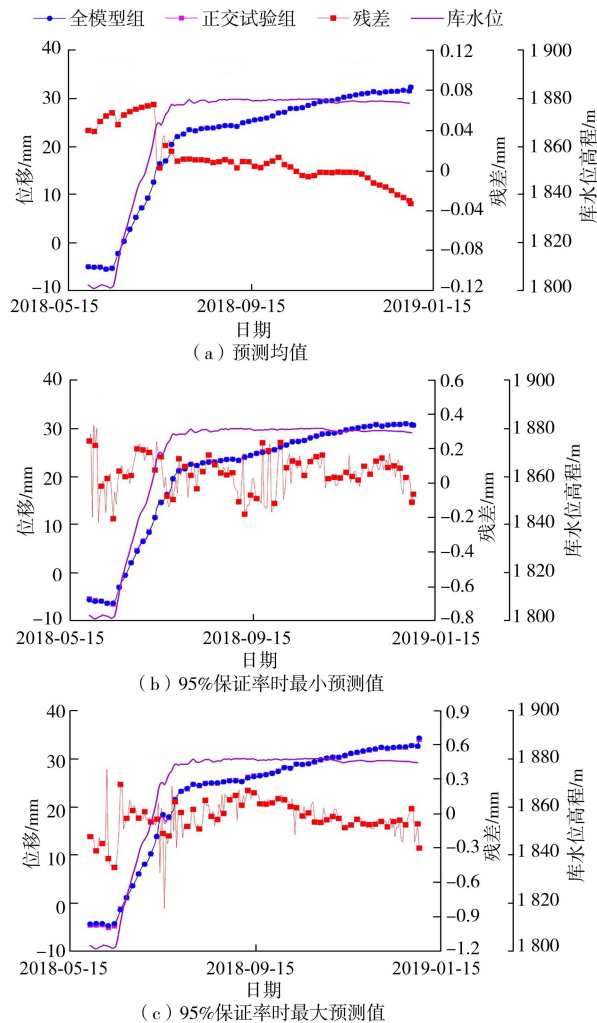
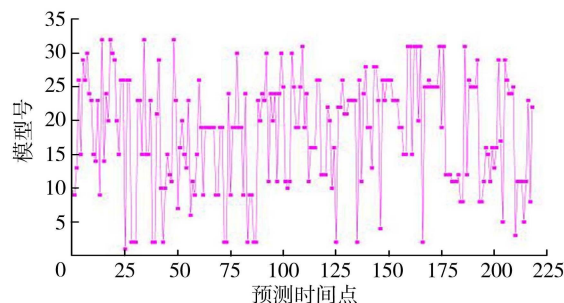


图4 全模型组和正交试验设计组的预测位移特征值

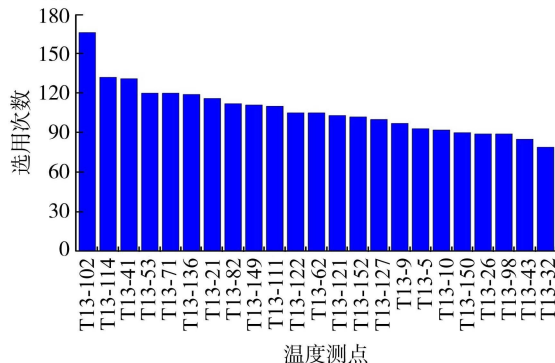
逐个与其预测均值进行对比,可寻找出各预测时间点最接近位移预测均值的模型,如图5(a)所示,可明显发现不同预测时间点的最接近模型是不同的,各预测时间点的最优温度因子也不完全相同,由此表明荷载周期性循环过程中高拱坝温度变形效应的因果关系是动态变化的,这也是本文采用不同实测温度因子组合来构建位移区间预测模型的初衷;此外,仅1个预测时间点的最接近模型同时包含了23个实测温度因子(预测时间点为第25天),由此表明最小相似性实测温度时间序列集是各位移测点建立预测模型时可供选择的温度变形因子的上限。结合图5(b)可发现,在218个预测值中,各温度因子选用次数介于79和166之间,其中T13-102、T13-114和T13-41的选用次数明显高于其他温度因子,这3个温度测点分别位于死水位1800 m高程附近的坝体上游面、坝体中部,以及坝体下游面,对坝体库水温边界和气温边界具有最强的代表性,因而也更适合表征坝体实测温度场的整体演变规律。

2.3 与传统HST模型的对比分析

模型预测精度主要通过均方根误差、最大绝对



(a) 正交试验设计模型



(b) 温度因子选用次数

图5 预测阶段各预测时间点最接近预测均值的正交试验设计模型及温度因子选用次数

误差和复相关系数来衡量,其中前两者越小、后者越大,代表模型精度越高^[1,6]。预测阶段的模型性能评价指标统计如表2所示,其中23因子的全模型组和正交试验设计组以预测均值为计算依据,HST模型采用相同的SVM建模方法。从表2可对比发现,23因子的正交试验设计组和全模型组的均方根误差和复相关系数相同,但前者的最大绝对误差明显大于后者,此结果也是因为正交试验设计组难以兼顾外界环境突变的影响;此外,全模型组预测模型的均方根误差明显小于HST模型,由此验证了本文方法具有较高的精度。

表2 预测阶段模型性能评价指标

建模方法	均方根误差	复相关系数	最大绝对误差
全模型组	0.463 6	0.996 8	1.722 6
正交试验设计组	0.463 6	0.996 8	2.477 6
HST模型	0.581 9	0.996 0	2.173 3

为考虑各种不确定性对位移诊断结果的影响,传统方法在基于数学监控模型拟定置信区间时,其在预测值的两侧分别加减固定倍数的模型标准差,比如,取保证率为95%时,可形成2S置信区间(实际为1.96S,其中S为建模阶段的均方根误差)^[1]。然而,该方法仅是从误差统计角度出发的,所拟定的置信区间并不具备严格的因果机理。本文方法与固定输入因子HST模型所得位移预测值及置信区间对比如图6所示,两种模型所得预测值与监测值均较为接近,本文方法得到的位移置信区间在不同预

测时间点其区间大小是不同的,处于动态变化过程中;而传统方法所建位移置信区间是静态的,没有考虑荷载周期性循环过程中因果关系的动态变化,这也是本文构建温度变形效应认知不确定下位移区间预测模型的原因之一。

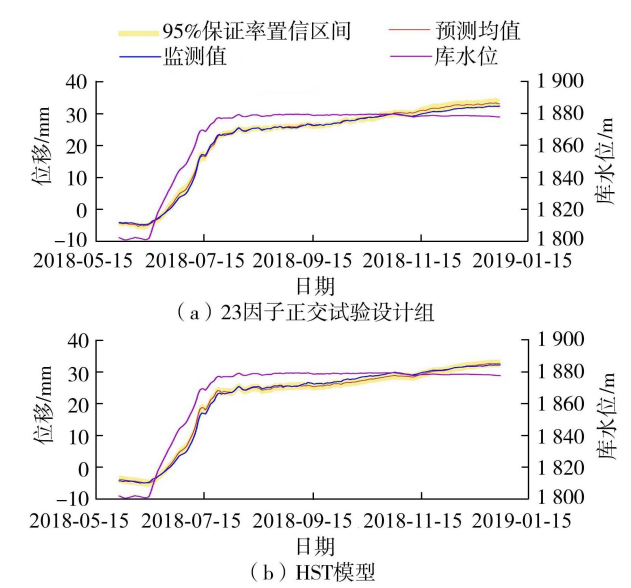


图6 PL13-1 测点径向位移区间预测值(均值)及置信区间

3 结 论

a. DTW 相似性分析可有效实现高拱坝温度测点的优选,从锦屏一级拱坝坝体 140 个温度测点中选用的 23 个测点具有较强的表征能力,可全面反映坝体实测温度场的演变规律,据此建立的位移区间预测模型具有良好的性能。

b. 正交试验设计可合理构建实测温度因子不确定下的输入因子组合,显著提高建模效率,多模型预测所得拱坝位移分布规律、均值及区间与全模型组基本一致,但必须剔除坝体实测温度时间序列中相似性较大者。

c. 对于高拱坝温度变形效应而言,最优实测温度变形因子组合是动态变化的,其中影响最大的是可表征库水温和坝址气温的坝体温度测点。基于最小相似性实测温度因子和正交试验设计建立的位移预测置信区间,可较好地表征温度变形效应认知不确定性对高拱坝位移预测的影响,所建位移置信区间具有明确的因果机理。

参考文献:

[1] 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.
[2] KANG Fei, LI Junjie, ZHAO Sizeng, et al. Structural health monitoring of concrete dams using long-term air temperature for thermal effect simulation [J].

Engineering Structures,2019,180:642-653.
[3] TATIN M, BRIFFAUT M, DUFOUR F, et al. Statistical modelling of thermal displacements for concrete dams: influence of water temperature profile and dam thickness profile [J]. Engineering Structures,2018,65: 63-75.
[4] MATA J, CASTRO A T D, COSTA J S D. Constructing statistical models for arch dam deformation[J]. Structural Control and Health Monitoring,2014,21(3):423-37.
[5] WANG Shaowei, XU Cong, GU Chongshi, et al. Displacement monitoring model of concrete dams using the shape feature clustering-based temperature principal component factor [J]. Structural Control and Health Monitoring, 2020, 27(10): e2603.
[6] WANG Shaowei, SUI Xupeng, LIU Yi, et al. Prediction and interpretation of the deformation behaviour of high arch dams based on a measured temperature field[J]. Journal of Civil Structural Health Monitoring,2023,13(2/3):661-675.
[7] 隋旭鹏,王少伟,朱圣辉,等. 不同机器学习模型对混凝土坝位移预测性能的影响[J]. 水电能源科学, 2022,40(1):107-111. (SUI Xupeng, WANG Shaowei, ZHU Shenghui, et al. Performance evaluation of different machine learning models on predicting displacement of concrete dams [J]. Water Resources and Power, 2022, 40(1):107-111. (in Chinese))
[8] 向镇洋,包腾飞,白妍丽,等. 基于混合注意力机制和深度学习的大坝变形预测模型[J]. 水利水电科技进展, 2023,43(2):96-101. (XIANG Zhenyang, BAO Tengfei, BAI Yanli, et al. Dam deformation prediction model based on mixed attention mechanism and deep learning [J]. Advances in Science and Technology of Water Resources, 2023,43(2):96-101. (in Chinese))
[9] 胡江,王春红,马福恒. 特高拱坝运行初期变形预测模型温度因子选取方法[J]. 长江科学院院报, 2021, 38(1): 59-65. (HU Jiang, WANG Chunhong, MA Fuheng. Selecting temperature factors for deformation prediction model for super-high arch dams during initial operation [J]. Journal of Yangtze River Scientific Research Institute,2021,38(1):59-65. (in Chinese))
[10] REN Qiubing, LI Mingchao, KONG Rui, et al. A hybrid approach for interval prediction of concrete dam displacements under uncertain conditions [J]. Engineering with Computers, 2023,39(2):1285-1303.
[11] 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(2):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(2):154-164. (in Chinese))
[12] 姬振豫. 正交设计的方法与理论[M]. 天津:世界科技出版社,2001.

(收稿日期:2023-05-05 编辑:骆超)