

• 水工岩石力学 •

DOI:10.12454/j.jsuese.202400517



本刊网刊

基于样本熵的大坝变形自适应预测模型

郭金婷¹, 李娅姣², 邓永东², 卢祥^{2*}, 周靖人², 陈建康^{2,3}

(1. 国能大渡河流域水电开发有限公司, 四川 成都 610041; 2. 四川大学 水利水电学院, 四川 成都 610065;
3. 四川大学 山区河流保护与治理全国重点实验室, 四川 成都 610065)

摘要: 外观变形监测是反馈大坝安全最直接、最重要的手段之一。高坝大库赋存与运行环境复杂, 准确预测其变形有助于及时识别大坝运行过程中的潜在风险。针对当前预测模型因大坝变形数据类型多样、非线性强而存在适用性差、预测精度波动大、泛化能力弱等问题, 首先, 基于“分解—优化—预测—重构”的思路, 采用自适应噪声完全集合经验模态分解算法将大坝变形数据时间序列分解为若干个不同频率的子模式序列; 然后, 基于各子模式序列的样本熵, 提出了样本熵—优化算法—模型自适应准则, 构建了与样本熵相匹配的超参数优化算法与预测模型, 通过不同子序列预测结果的再重构即可得到变形预测结果, 进而形成了基于样本熵的大坝变形自适应预测模型。以 PB 大坝为案例, 对坝顶位移测点的监测数据进行建模分析, 验证模型的可行性和有效性。工程应用表明, 与传统统计回归模型相比, 所提模型的均方根误差、平均绝对误差、平均绝对百分比误差在训练集和测试集上降低幅度分别超过了 60% 和 90%, 且精度高、适用性强。研究成果对准确预测大坝变形、掌握大坝安全运行性态具有重要的理论与工程应用价值。

关键词: 大坝; 变形; 样本熵; 优化; 自适应预测模型

中图分类号: TV698.1

文献标志码: A

文章编号: 2096-3246(2026)01-0101-10

中国已建水库大坝 9.5 万座, 其安全运行事关国民经济命脉、人民群众安全与社会稳定。大坝安全监测是保障大坝运行安全的重要手段之一, 常通过对大坝结构及环境参数的实时监测、分析监测数据的演变规律来诊断大坝安全运行性态、发现大坝潜在隐患, 以及预测大坝结构发展趋势^[1-2]。因此, 构建大坝安全监测数据预测模型对掌握大坝运行性态具有重要意义^[3-5]。

目前, 基于统计回归模型的传统监测数据预测模型应用较为广泛, 包括线性回归、稳健回归、随机森林回归、阈值回归等模型^[6-8]。传统统计回归模型结构简单, 易于实现, 但对数据非线性关系适应性较差、数据质量要求高、模型泛化能力不稳定。随着人工智能和大数据技术的迅猛发展, 许多高精度的时间序列智能预测模型不断涌现, 为大坝安全监测数据预测提供了新方法。例如: Zhang 等^[9]结合密集连

接卷积网络 (DenseNet) 的特征提取能力与长短期记忆网络 (LSTM) 的时间序列建模能力提出了重力坝位移的 DenseNet-LSTM 预测模型, 工程应用表明该模型预测精度高、泛化能力强。Wen 等^[10]提出一种基于支持向量机和自适应增强算法的大坝变形预测模型, 结果表明该算法比改进的支持向量机模型的预测精度更高, 且解决了数据中存在的非线性突变与离散性问题。岑威钧等^[11]基于集成经验模态分解方法、LSTM 和自回归滑动平均模型 (ARIMA) 构建土石坝渗压预测模型, 并以某深厚覆盖层上的土石坝工程为例进行验证, 结果表明, 相较于单一的 LSTM 模型和 ARIMA 模型, 该模型具有较高的预测性能和稳定性。徐丛等^[12]以模型误差指标最小和不同测点变形相关性最强分别构建目标函数, 通过关联向量机训练学习、模型参数优化构建拱坝变形监控模型, 提升了拱坝变形预测的精度与合理性。然而, 大坝变形数据类型多、受环境

收稿日期: 2024-07-05 修回日期: 2024-08-30 网络出版日期: 2024-09-10

基金项目: 国家自然科学基金项目 (52309162)

作者简介: 郭金婷 (1985—), 女, 高级工程师。研究方向: 水利水电建设工程管理及科技管理。E-mail: 12016399@ceic.com

* 通信作者: 卢祥, 副研究员, E-mail: scu-lx@scu.edu.cn

及荷载影响大,具有显著的非线性,同一种模型难以适用于各种类型变形数据^[13]。样本熵是一种衡量时间序列复杂性的指标,能够有效地捕捉数据中的随机性和不确定性,其在处理非线性、非平稳时间序列数据方面具有很大优势^[14-15]。因此,若能根据样本熵,将监测数据划分为若干数据子集,训练并匹配最优的预测模型,将有助于提高预测模型的精度和泛化能力。

当前,用来预测的智能预测模型越来越多,其中,LSTM^[16-17]、门控循环单元神经网络(GRU)^[18-19]和时序卷积神经网络(TCN)^[20]因在大坝变形、渗流等监测数据预测及检测中应用广泛,虽然存在长期依赖衰减、对多变量耦合能力有限、对异常数据敏感度不足等问题,但是预测精度可以满足工程要求。

因此,本文主要采用上述3种模型来进行研究,构建一种基于样本熵的大坝变形预测自适应模型。即通过对监测数据序列进行模态分解、计算各子模式序列的样本熵,采用不同预测模型组合对各子模式序列进行建模,提出样本熵-优化算法-模型自适应准则,构建基于“分解-优化-预测-重构”的大坝变形预测自适应模型;最后,以某高心墙堆石坝为案例,选取坝顶9个位移测点,对所提模型的可行性和有效性进行分析,并与其他模型进行对比验证。

1 大坝变形自适应预测模型

1.1 监测数据的模态分解

经验模态分解算法(EMD)^[21]可将监测数据时间序列分解为多个不同频率的子模态序列,每个子模态序列保留原始频率特征,但易产生数据频率模态混叠现象。此时,集成经验模态分解算法(EEMD)^[22]应运而生。EEMD算法先在监测数据时间序列中添加高斯白噪声,再采用EMD算法进行分解,进一步改善了EMD算法的分解效果,但是,EEMD算法较难彻底消灭残余噪声。针对上述问题,相关学者提出了自适应噪声完全集合经验模态分解(ICEEMDAN)算法^[23-25],其计算过程如下:

首先,依据EMD算法将各组高斯噪声进行分解,并分别加入到大坝安全监测数据时间序列的原始值 x 中,得到不同高斯噪声影响下的时间序列:

$$x^{(i)} = x + \varepsilon_0 E_1(w^{(i)}) \quad (1)$$

式中, $x^{(i)}$ 为第 i 个高斯噪声影响下的时间序列, $E_1(\cdot)$ 为分解得到的第1个子模态分量, $w^{(i)}$ 为第 i 个高斯噪声, ε_0 为第1次循环中添加噪声的标准差。

其次,用 $x^{(i)}$ 的局部均值计算平均值,并计算监测数据时间序列的第1个残余分量 r_1 :

$$r_1 = \langle M(x^{(i)}) \rangle \quad (2)$$

式中, $M(\cdot)$ 为局部均值, $\langle \cdot \rangle$ 为根据频率模态计算的均值。

此时,大坝安全监测数据时间序列第1个周期分量 f_1 的计算为:

$$f_1 = x - r_1 \quad (3)$$

对残差的第1个值由式(1)~(3)重复处理,从而得到可靠指标的第2个残余分量 r_2 与第2个周期分量 f_2 分别为:

$$r_2 = \langle M(r_1 + \varepsilon_1 E_2(w^{(i)})) \rangle \quad (4)$$

$$f_2 = r_1 - r_2 \quad (5)$$

式(4)中, ε_1 为第2次循环中添加噪声的标准差。

最后,可推导出监测数据的第 k 个残余分量与第 k 个周期分量分别:

$$r_k = \langle M(r_{k-1} + \varepsilon_{k-1} E_k(w^{(i)})) \rangle \quad (6)$$

$$f_k = r_{k-1} - r_k \quad (7)$$

式(6)、(7)中, ε_{k-1} 为第 k 次循环中添加噪声的标准差, r_{k-1} 为第 $k-1$ 个残余分量。

1.2 样本熵

样本熵(SampEn)是一种衡量时间序列复杂度的新方法,通过度量信号中产生新模式的概率大小来衡量时间序列的复杂度。时间序列的样本熵越大,表示新模式产生的概率越大,则时间序列的复杂度越大^[26]。对于大坝安全监测数据时间序列,样本熵的具体计算过程如下:

1)假设有一长度为 N 的时间序列 $\{x(1), x(2), \dots, x(N)\}$,按顺序组成一组维数为 m 的序列,即

$$X_m(i) = \{x(i), x(i+1), \dots, x(i+m-1)\}, \quad 1 \leq i \leq N-m+1 \quad (8)$$

2)定义 $X_m(i)$ 与 $X_m(j)$ 之间的距离 $d[X_m(i), X_m(j)]$ 为两者对应位置的元素中最大差值的绝对值,即:

$$d[X_m(i), X_m(j)] = \max_{k=0,1,\dots,m-1} (|x(i+k) - x(j+k)|) \quad (9)$$

式中, $1 \leq j \leq N-m+1, i \neq j$ 。

3)对于给定的 $X_m(i)$,统计 $X_m(i)$ 与 $X_m(j)$ 之间距离小于相似容限 l 的 j 的数目,记作 B_i ,并与距离个数 $N-m$ 做比值,记为 $B_i^m(l)$,即:

$$B_i^m(l) = \frac{1}{N-m} B_i \quad (10)$$

4)定义 $B^m(l)$ 为:

$$B^m(l) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(l) \quad (11)$$

5)增加维数到 $m+1$,重复式(8)~(10),计算得到 $B^{m+1}(l)$:

$$B^{m+1}(l) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^{m+1}(l) \quad (12)$$

6)此时,时间序列的样本熵可表达为:

$$S(m, l) = \lim_{N \rightarrow \infty} \left(-\ln \frac{B^{m+1}(l)}{B^m(l)} \right) \quad (13)$$

但实际中时间序列长度 N 不可能接近无穷大,故当 N 有限时,样本熵可估算为:

$$S(m, l, N) = -\ln \frac{B^{m+1}(l)}{B^m(l)} \quad (14)$$

1.3 基于样本熵的大坝变形自适应预测模型构建

1.3.1 模型构建思路

鉴于大坝变形监测数据非线性较强,直接对原始变形序列进行预测,精度往往得不到保障。因此,可先采用 ICEEMDAN 算法将大坝变形监测数据时间序列分解为多个不同频率的子模态序列,再采用智能模型分别预测各子模态序列。由于各子模态序列具有不同的数据特征,采用同样的智能模型可以保障预测精度的一致性。在此基础上,通过计算各子模态序列样本熵,并根据样本熵的大小及其分类,从常用的 LSTM、GRU 和 TCN 模型中匹配精度最高的模型,提出基于“分解—预测—重构”的预测模型。

此外,由于影响大坝变形的因素众多,在智能模型训练过程中,超参数的确定常常存在较大的不确定性和主观性。合理设定超参数可降低算法的循环次数,进而可优化变形预测模型的计算速度和精度。因此,为提升模型性能,常采用参数优化算法对超参数寻优。蜻蜓优化算法(DA)、鲸鱼优化算法(WOA)和灰狼优化算法(GWO)等优化算法^[27-29]在时序序列预测中应用广泛,其中,DA算法寻优能力强,WOA算法操作简单、需要设置的参数少、寻优性能强,GWO算法具有较强的收敛性能、参数少、易实现等特点。在本文基于“分解—预测—重构”的预测模型上,通过引入超参数优化,形成了本文基于样本熵的大坝安全监测数据自适应预测模型。

1.3.2 样本熵-优化算法-模型自适应准则

为确定样本熵与优化算法的匹配准则,本文以某高心墙堆石坝沉降位移监测数据作为数据集,经模态分解成9个子模态序列($f_1 \sim f_9$),计算各子序列的样本熵,并将3种优化算法(DA、WOA和GWO)和3种预测模型(GRU、LSTM和TCN)一一组合,对各子序列进行训练并预测,选用均方根误差 RMSE(记为 E_{RMSE})、平均绝对误差 MAE(记为 E_{MAE})、平均绝对百分比误差 MAPE(记为 E_{MAPE})及复相关系数 R 这4个指标来评价各模型的预测效果^[30-31],其计算如式(15)~(18)所示。

$$E_{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (15)$$

$$E_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (16)$$

$$E_{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (17)$$

$$R = \frac{\sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}}{\sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}} \quad (18)$$

式(15)~(18)中, y_i 和 \hat{y}_i 分别为第*i*个监测数据的实测值和预测值, \bar{y} 为监测数据的平均值, n 为监测数据个数。

优化算法与预测模型的组合情况如表1所示。

表1 预测模型组合

Tab. 1 Combination of prediction models

优化算法	预测模型		
	GRU	LSTM	TCN
DA	DA-GRU	DA-LSTM	DA-TCN
WOA	WOA-GRU	WOA-LSTM	WOA-TCN
GWO	GWO-GRU	GWO-LSTM	GWO-TCN

经不同优化算法和预测模型组合预测后,其预测精度如表2所示。由表2可知:3种优化算法条件下,TCN模型预测的精度优于其他两种模型,且受样本熵影响较小,因此,构建预测模型时应以TCN作为预测模型。优化算法受样本熵影响较大,具体如下:当 $S \geq 0.3$ 时,基于GWO优化算法的预测精度最高;当 $0.1 \leq S < 0.3$ 时,基于WOA优化算法的预测精度最高;当 $S < 0.1$ 时,基于DA优化算法的预测精度最高。

1.3.3 模型构建流程

基于上述模型构建思路、样本熵-优化算法-模型的自适应准则,得到基于样本熵的大坝安全监测数据自适应预测模型的构建流程,如图1所示,具体步骤如下:

1)监测数据模态分解。针对大坝安全监测数据时间序列,采用ICEEMDAN算法将监测数据时间序列分解成若干不同频率的子模态序列。

2)样本熵分析。对分解后的各子模态序列,按照第1.2节方法计算各子序列的样本熵,并根据样本熵的范围依据准则进行匹配。

3)子模态序列预测。基于匹配的子序列预测模型,对每个子序列逐一预测,形成若干个子序列预测模型。

4)预测结果重构。将每个子序列的预测值进行合并重构,得到最终监测数据预测结果,并计算 E_{RMSE} 、 E_{MAE} 、 E_{MAPE} 及 R 。

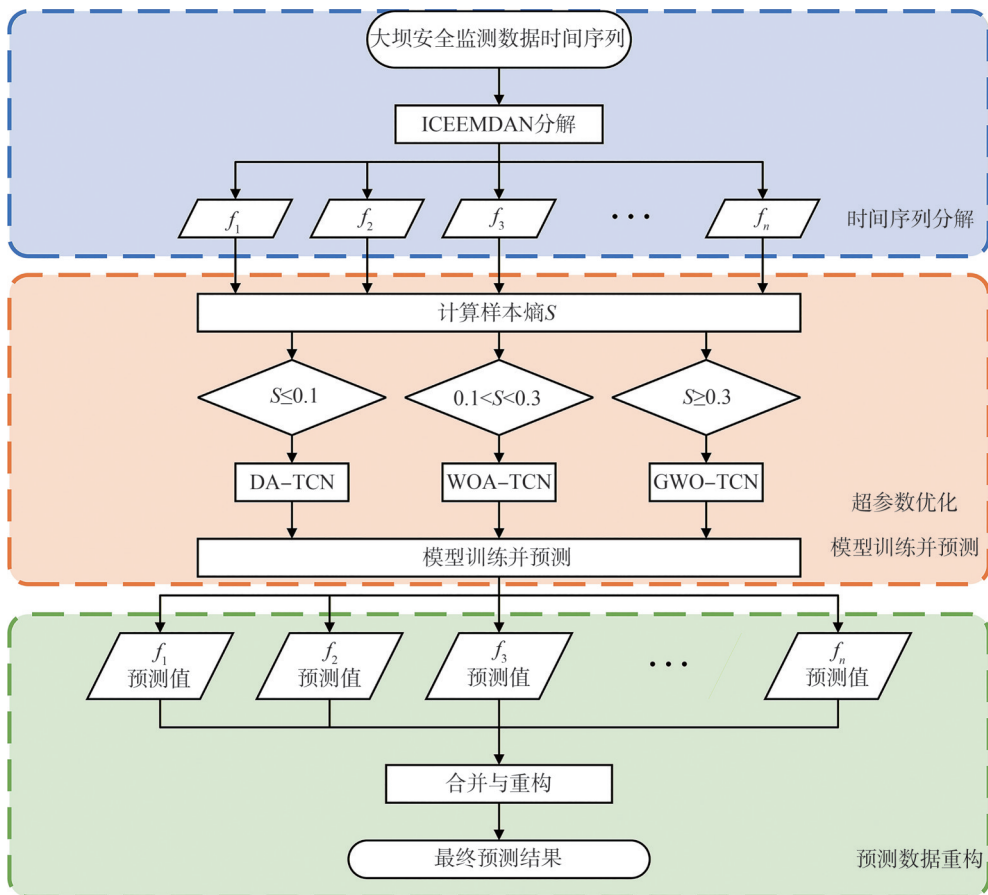


图 1 模型构建流程图

Fig. 1 Process diagram of model construction

2 模型精度验证

2.1 工程简介

PB水电站位于四川省大渡河中游汉源县和甘洛县接壤处,是大渡河干流水电规划22级开发方案的第17个梯级电站。电站水库正常蓄水位为850.00 m,汛期运行限制水位为841.00 m,死水位为790.00 m,设计总库容为53.37亿m³,其中,防洪库容为7.27亿m³,调节库容为38.94亿m³,是不完全年调节水库。工程属一等大(1)型,大坝等主要水工建筑物为I级建筑物,次要水工建筑物为3级建筑物。坝型为砾石土直心墙堆石坝,坝顶高程为856.00 m,最大坝高为186.00 m,坝顶

宽为14.00 m、大坝长为540.50 m。

2.2 监测布置及监测数据

大坝表面外观变形监测测点分别布置在大坝坝顶、下游坝面、上游795.00 m高程马道。本文以坝顶下游侧的所有9个测点(TP10~TP18)沉降位移时间序列作为实验数据集,数据的时间跨度为从2013-06-22至2024-03-17。其中,从2013-06-22至2023-08-07的数据序列作为训练集,从2023-08-08至2024-03-17的数据序列作为测试集,训练集与测试集的比例为9:1。坝顶外观变形监测测点布置如图2所示,测点TP10~TP18的沉降位移的时间序列如图3所示。

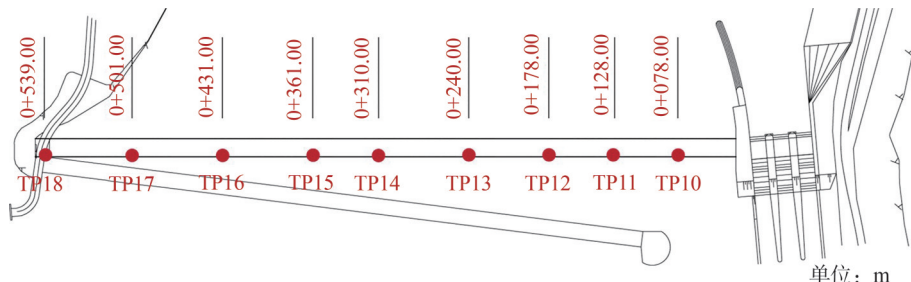


图 2 坝顶外观变形监测布置

Fig. 2 Monitoring layout diagram for deformation in the dam crest

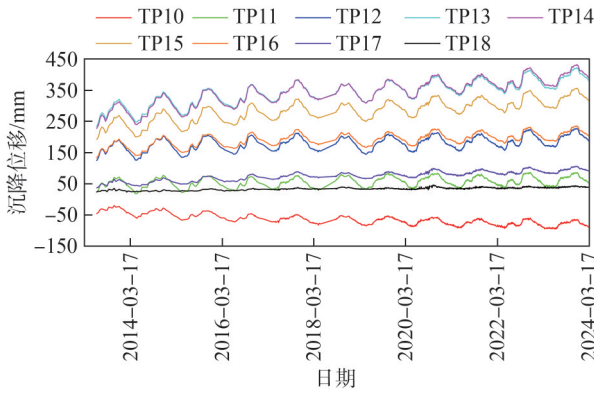


图3 坝顶测点沉降位移时间序列

Fig. 3 Time series of settlement displacement of measurement points in the dam crest

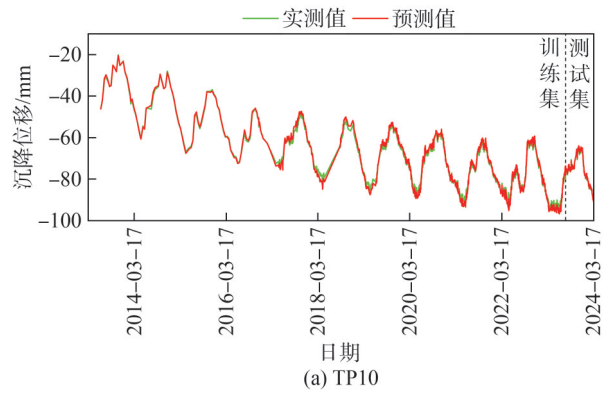
2.3 大坝变形自适应预测模型精度分析

为了评估本文所提自适应预测模型的精度,以坝顶下游侧的所有9个测点(TP10~TP18)沉降位移时间序列作为实验数据集,用于训练该模型并做预测。为了衡量模型的性能,选用均方根误差 E_{RMSE} 、平均绝对误差 E_{MAE} 、平均绝对百分比误差 E_{MAPE} 及复相关系数 R 这4个指标来评价所提模型的预测效果。各测点的评价指标如表3所示,典型测点TP10、TP13、TP15、TP18的预测结果如图4所示。

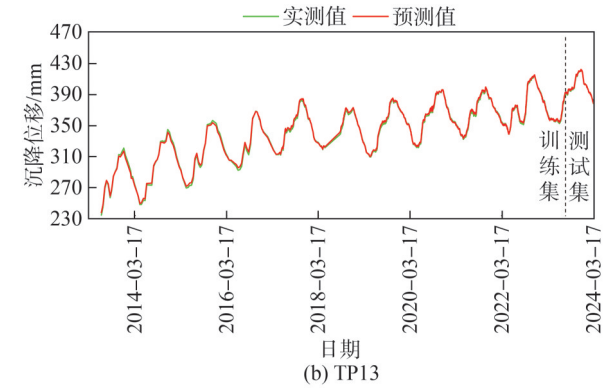
表3 本文自适应模型的不同测点预测精度

Tab. 3 Prediction accuracy of the proposed a adaptive model at different measurement points

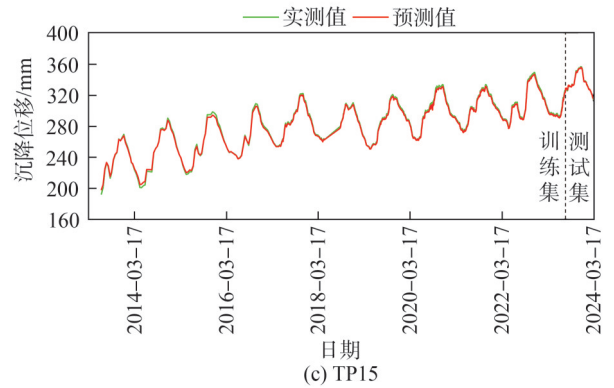
分类	测点	评价指标			
		E_{RMSE}	E_{MAE}	E_{MAPE}	R
训练集	TP10	1.19	1.01	0.010	0.991
	TP11	1.55	1.34	0.020	0.988
	TP12	1.80	1.53	0.009	0.990
	TP13	1.52	1.26	0.004	0.997
	TP14	1.05	0.87	0.002	0.999
	TP15	1.69	1.56	0.005	0.994
	TP16	0.71	0.58	0.003	0.998
	TP17	0.38	0.27	0.004	0.999
	TP18	0.12	0.10	0.003	0.998
测试集	TP10	0.34	0.29	0.004	0.996
	TP11	0.71	0.65	0.009	0.994
	TP12	0.58	0.50	0.002	0.997
	TP13	0.57	0.40	0.001	0.998
	TP14	0.36	0.26	0.001	0.999
	TP15	0.61	0.50	0.002	0.997
	TP16	0.22	0.18	0.001	0.999
	TP17	0.33	0.24	0.003	0.994
	TP18	0.08	0.06	0.002	0.996



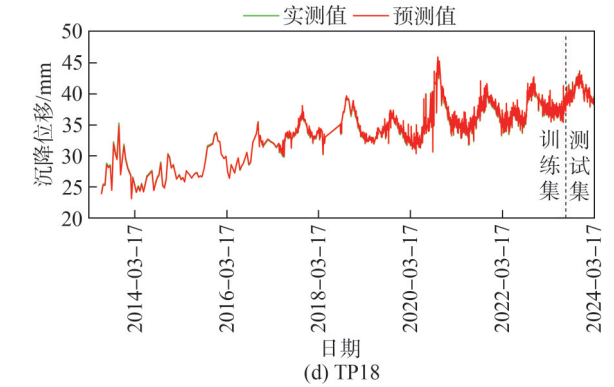
(a) TP10



(b) TP13



(c) TP15



(d) TP18

图4 本文自适应模型的典型测点预测结果

Fig. 4 Prediction results of typical measurement points for the proposed adaptive models

由表3和图4可知,本文所提模型的 E_{RMSE} 在0.08~1.86之间, E_{MAE} 在0.06~1.56之间, E_{MAPE} 小于0.02, R 均大于0.988,表明所提模型在训练集和测试集上都具有很好的预测精度。

为进一步验证本文所提自适应预测模型的性能,采用统计回归、GRU、LSTM和TCN4种模型与其进行精度对比。以TP13沉降位移监测数据为实验数据集,分别采用统计回归、GRU、LSTM、TCN和所提的自适应模型进行训练并预测,结果如表4和图5所示。

表4 不同模型的预测精度对比(TP13)

Tab. 4 Comparison of prediction accuracy among different models (TP13)

分类	模型	评价指标			
		E_{RMSE}	E_{MAE}	E_{MAPE}	R
训练集	统计回归	5.19	3.92	0.010	0.96
	GRU	7.55	6.05	0.020	0.91
	LSTM	5.37	4.17	0.010	0.96
	TCN	3.82	2.95	0.010	0.98
	所提自适应模型	1.52	1.26	0.004	1.00
测试集	统计回归	5.94	5.31	0.010	0.74
	GRU	9.83	7.83	0.020	0.28
	LSTM	10.27	8.15	0.020	0.21
	TCN	4.99	4.06	0.010	0.81
	所提自适应模型	0.57	0.40	0.001	1.00

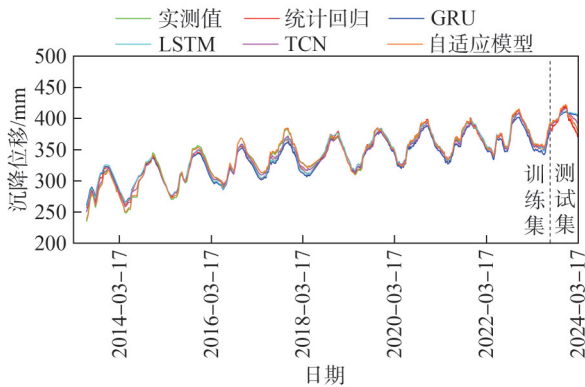


图5 不同模型的预测结果对比(TP13)

Fig. 5 Comparison of prediction results of different models (TP13)

由表4和图5可知:以统计回归模型精度为基准,在训练集上,所提自适应模型的 E_{RMSE} 、 E_{MAE} 、 E_{MAPE} 显著低于其他模型,分别降低了70.7%、67.9%和60.0%;在测试集上,所提自适应模型的精度提升更为明显,其 E_{RMSE} 、 E_{MAE} 、 E_{MAPE} 分别降低了90.4%、92.5%和90.0%。此外,尽管在训练集上各模型的复相关系数 R 均较高且接近,但是在测试集上差异明显。所提自适应模型的复相关系数 R 比TCN模型提高了23.5%,比统计回归模型提高了35.1%;GRU和LSTM模型的复相关系数 R 较低,其原因可能是在训练集中存在过拟合现象,导致在测试集中其精度显著降低,不适合长序列监测数据的预测。

3 结论

本文基于“分解—优化—预测—重构”的思想,构建了一种基于样本熵的大坝变形自适应预测模型。主要结论如下:

1)通过ICEEMDAN算法将大坝变形数据分解为不同频率的子序列,并根据子序列的样本熵,提出了样本熵—优化算法—模型自适应准则,构建了基于样本熵的大坝变形自适应预测模型。

2)模型精度验证表明,本文所提模型预测精度高,与常用的预测模型的精度对比分析表明,所提模型的 E_{RMSE} 、 E_{MAE} 、 E_{MAPE} 显著低于其他模型,训练集上其误差降低了60%以上,测试集上其误差降低了90%以上。

3)相较于训练集,本文所提模型在测试集的精度比其他模型提升得更为明显,表明所提模型在模型训练过程中能有效提取监测数据的关键特征,模型适应性和泛化能力强。

4)本文仅考虑了常用的3种智能模型和3种优化算法,而当前存在的模型及优化算法众多,未来亦可能产生新的模型和优化算法,可按照本文模型构建思路进行模型更新与修正,进一步提高模型的适应性。

参考文献:

- [1] Liu Hepeng, Li Denghua, Ding Yong, et al. MHA-ConvLSTM dam deformation prediction model considering environmental volume lag effect[J]. Applied Sciences, 2023, 13(14): 8538.
- [2] Song Sihan, Zhou Qiuqing, Zhang Tao, et al. Automatic concrete dam deformation prediction model based on TPE-STL-LSTM[J]. Water, 2023, 15(11): 2090.
- [3] Cheng Lin, Chen Jiamin, Ma Chunhui, et al. Multipoint deformation safety monitoring model for concrete arch dams based on Bayesian model selection and averaging[J]. Structural Control and Health Monitoring, 2023, 2023: 5042882.
- [4] Tian Jichen, Luo Yonghua, Lu Xiang, et al. Physical data-driven modeling of deformation mechanism constraints on earth-rock dams based on deep feature knowledge distillation and finite element method[J]. Engineering Structures, 2024, 307: 117899.
- [5] Wei Yilun, Li Qingbin, Hu Yu, et al. Deformation prediction model based on an improved CNN + LSTM model for the first impoundment of super-high arch dams[J]. Journal of Civil Structural Health Monitoring, 2023, 13(2): 431-442.
- [6] Chen Xudong, Hou Zhenzhen, Guo Jinjun. Hybrid model for

- deformation prediction of concrete dams based on long-term temperature simulation[J]. *Yellow River*, 2023, 45(9): 141–146. [陈旭东, 侯阵阵, 郭进军. 基于长期气温模拟的混凝土坝变形预测混合模型[J]. *人民黄河*, 2023, 45(9): 141–146.]
- [7] Han Zhang, Chen Jiankang, Zhang Fang, et al. An efficient online outlier recognition method of dam monitoring data based on improved M-robust regression[J]. *Structural Health Monitoring*, 2023, 22(1): 581–599.
- [8] Li Yanlong, Zhang Haiyang, Wen Lifeng, et al. A prediction model for deformation behavior of concrete face rockfill dams based on the threshold regression method[J]. *Arabian Journal for Science and Engineering*, 2021, 46(6): 5801–5816.
- [9] Zhang Ye, Zhong Wen, Li Yanlong, et al. A deep learning prediction model of DenseNet-LSTM for concrete gravity dam deformation based on feature selection[J]. *Engineering Structures*, 2023, 295: 116827.
- [10] Wen Lifeng, Li Yanlong, Zhao Weibo, et al. Predicting the deformation behaviour of concrete face rockfill dams by combining support vector machine and AdaBoost ensemble algorithm[J]. *Computers and Geotechnics*, 2023, 161: 105611.
- [11] Cen Weijun, Wang Xiaoxin, Jiang Minghuan. Seepage pore water pressure prediction model of earth-rock dams based on EEMD-LSTM-ARIMA[J]. *Journal of Water Resources and Water Engineering*, 2023, 34(2): 180–185. [岑威钧, 王肖鑫, 蒋明欢. 基于 EEMD-LSTM-ARIMA 的土石坝渗压预测模型研究[J]. *水资源与水工程学报*, 2023, 34(2): 180–185.]
- [12] Xu Cong, Wang Shaowei, Gu Chongshi, et al. A probabilistic prediction model for displacement of super high arch dams considering deformation spatial association[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(3): 433–442. [徐丛, 王少伟, 顾冲时, 等. 融合空间关联性的特高拱坝位移概率性预测模型[J]. *武汉大学学报(信息科学版)*, 2023, 48(3): 433–442.]
- [13] Zhang Fang, Lu Xiang, Li Yanling, et al. A self-matching model for online anomaly recognition of safety monitoring data in dams[J]. *Structural Health Monitoring*, 2023, 22(1): 746–773.
- [14] Lin Chuan, Wang Xiangyu, Su Yan, et al. Predictions of concrete dam deformation using clustering method and deep learning[J]. *Journal of Hydroelectric Engineering*, 2022, 41(10): 112–127. [林川, 王翔宇, 苏燕, 等. 联合聚类方法和深度学习的混凝土坝变形预测[J]. *水力发电学报*, 2022, 41(10): 112–127.]
- [15] Zhang Fangqing, Yuan Fang, He Yu, et al. Vibration signal de-noising of hydropower units based on NLM-CEEMDAN and sample entropy[J]. *China Rural Water and Hydropower*, 2023(6): 286–294. [章芳情, 袁方, 贺玉, 等. 基于 NLM-CEEMDAN 和样本熵的水电机组振动信号去噪[J]. *中国农村水利水电*, 2023(6): 286–294.]
- [16] Liu Changwei, Pan Jianwen, Wang Jinting. An LSTM-based anomaly detection model for the deformation of concrete dams[J]. *Structural Health Monitoring*, 2024, 23(3): 1914–1925.
- [17] Wang Shuming, Yang Bing, Chen Huimin, et al. LSTM-based deformation prediction model of the embankment dam of the Danjiangkou hydropower station[J]. *Water*, 2022, 14(16): 2464.
- [18] Lin Chuan, Weng Kailiang, Lin Youlong, et al. Time series prediction of dam deformation using a hybrid STL-CNN-GRU model based on sparrow search algorithm optimization[J]. *Applied Sciences*, 2022, 12(23): 11951.
- [19] Jiang Jiatong, Li Mingwei, Shang Xianchao, et al. Concrete dam deformation prediction based on convolutional and recurrent neural network[J]. *Journal of Harbin Engineering University*, 2023, 44(8): 1270–1274. [蒋佳彤, 李明伟, 尚宪朝, 等. 基于卷积循环神经网络的混凝土坝变形预报[J]. *哈尔滨工程大学学报*, 2023, 44(8): 1270–1274.]
- [20] Wang Yu, Liu Guohua. MLA-TCN: Multioutput prediction of dam displacement based on temporal convolutional network with attention mechanism[J]. *Structural Control and Health Monitoring*, 2023, 2023(1): 2189912.
- [21] Jamei M, Ali M, Karbasi M, et al. Monthly sodium adsorption ratio forecasting in rivers using a dual interpretable glass-box complementary intelligent system: Hybridization of ensemble TVF-EMD-VMD, Boruta-SHAP, and eXplainable GPR[J]. *Expert Systems with Applications*, 2024, 237: 121512.
- [22] Wu Zhaohua, Huang N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method[J]. *Advances in Adaptive Data Analysis*, 2009, 1(1): 1–41.
- [23] Mao Qinnan, Liu Zhao, Li Jie, et al. Runoff prediction and comparative study by combining ICEEMDAN and multiple intelligent optimization algorithms[J]. *Water Resources and Power*, 2023, 41(10): 23–26. [毛钦男, 刘招, 李杰, 等. 联合 ICEEMDAN 与多种智能优化算法的径流预测及对比

- 研究[J].水电能源科学,2023,41(10):23–26.]
- [24] Colominas M A, Schlotthauer G, Torres M E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing[J]. Biomedical Signal Processing and Control, 2014, 14: 19–29.
- [25] Liu Bo, Cen Weijun, Zheng Changhai, et al. A combined optimization prediction model for earth-rock dam seepage pressure using multi-machine learning fusion with decomposition data-driven[J]. Expert Systems with Applications, 2024, 242: 122798.
- [26] Fu Luying, Qi Huijun, Li Tongchun, et al. Dam displacement prediction based on CEEMDAN and correlation analysis[J]. Journal of China Three Gorges University (Natural Sciences), 2024, 46(1): 1–6. [傅露莹, 齐慧君, 李同春, 等. 基于 CEEMDAN 和相关性分析的大坝位移预测[J]. 三峡大学学报(自然科学版), 2024, 46(1): 1–6.]
- [27] Shah M, Borade H, Sanghavi V, et al. Enhancing tool wear prediction accuracy using Walsh – Hadamard transform, DCGAN and dragonfly algorithm-based feature selection [J]. Sensors, 2023, 23(8): 3833.
- [28] Su Kanhua, Da Wenhao, Li Meng, et al. Research on a drilling rate of penetration prediction model based on the improved chaos whale optimization and back propagation algorithm[J]. Geoenergy Science and Engineering, 2024, 240: 213017.
- [29] Liang Jing, Du Yukun, Xu Yipeng, et al. Using Adaptive Chaotic Grey Wolf Optimization for the daily streamflow prediction[J]. Expert Systems with Applications, 2024, 237: 121113.
- [30] Zhang Bo, Liu Jian, Wu Zhenyu, et al. R–OC criterion for dam deformation monitoring model identification[J]. Advanced Engineering Sciences, 2023, 55(3): 175–185. [张博, 刘健, 吴震宇, 等. 大坝变形监控模型识别的 R–OC 准则[J]. 工程科学与技术, 2023, 55(3): 175–185.]
- [31] Li Weiye, Wu Zhenyu. A methodology for dam parameter identification combining machine learning, multi-objective optimization and multiple decision criteria[J]. Applied Soft Computing, 2022, 128: 109476.

Adaptive Prediction Model for the Dam Deformation Based on the Sample Entropy

GUO Jinting¹, LI Yajiao², DENG Yongdong², LU Xiang^{2*}, ZHOU Jingren², CHEN Jiankang^{2,3}

(1.State Energy Dadu River Basin Hydropower Development Company, Limited, Chengdu 610041, China;

2.State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, China;

3.College of Water Resources & Hydropower, Sichuan University, Chengdu 610065, China)

Abstract:

Objective Deformation monitoring is one of the most direct and critical methods for assessing dam safety, and numerous studies have shown that deformation data typically exhibit significant changes prior to structural damage. The operating environment of high dams and large reservoirs is complex, and accurately predicting dam deformation is essential for timely identification of potential risks. However, the accuracy of deformation prediction is influenced by several factors, including data type, quantity, quality, patterns, and the prediction model itself. Research indicates that different models perform differently depending on the characteristics of the deformation sequence, making it difficult to ensure reliable predictions using a single model. To address the limitations of current models, such as poor applicability, large fluctuations in prediction accuracy, and weak generalization ability, this paper proposes an adaptive prediction model for dam deformation based on the “decomposition–optimization–prediction–reconstruction” framework.

Methods The dam deformation time series is first decomposed into several sub-mode sequences of different frequencies using the intrinsic computing expressive empirical mode decomposition with adaptive noise (ICEEMDAN) algorithm. Because high accuracy prediction is closely related to the extraction of key features from deformation data, sample entropy is introduced to quantify the complexity of each sub-mode sequence and to classify them according to entropy size. With prediction accuracy as the evaluation criterion, different hyperparameter optimization algorithms (e.g., dragonfly algorithm, whale optimization algorithm, and grey wolf optimizer) are combined with appropriate prediction models (e.g., long short-term memory network, gated recurrent unit network, and temporal convolutional neural network) for the classified sub-mode sequences. An adaptive criterion linking sample entropy, the optimization algorithm, and the prediction model is proposed. Then, the best combination of hyperparameter optimization algorithm and prediction model is determined. Finally, the deformation prediction results are obtained by reconstructing the predicted results of different sub-mode sequences, thereby forming an adaptive prediction model for dam deformation based on sample entropy.

Results and Discussions Engineering applications demonstrate that, compared with traditional statistical regression models, the proposed model reduces the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) by more than 60% and 90% in the training and testing sets, respectively, indicating high prediction accuracy and strong applicability. While the prediction of individual models and the proposed model is comparable on the training set, significant differences are observed in the testing set. For example, using the multi-correlation coefficient R as an indicator, the R values of all models are relatively high and similar in the training set, whereas significant differences occur in the testing set. The minimum increase in R achieved by the adaptive model is 23.5% compared with the TCN model and 35.1% compared with commonly used statistical regression models. These results indicate that the proposed model effectively extracts key features from deformation monitoring data and exhibits strong adaptability and generalization capability. With the increasing automation of dam monitoring systems, the growing number of monitoring points and increasingly complex data types place higher demands on prediction timeliness and accuracy.

Conclusions The proposed prediction model not only improves prediction accuracy but supports real-time calculation, enabling integration into modern online dam safety monitoring systems. By effectively utilizing deformation prediction data to evaluate the safety status of dams, this model demonstrates substantial theoretical value and practical engineering significance.

Key words: dam; deformation; sample entropy; optimization; adaptive prediction model

(编辑 赵 婧)

引用格式: Guo Jinting, Li Yajiao, Deng Yongdong, et al. Adaptive prediction model for the dam deformation based on the sample entropy[J]. *Advanced Engineering Sciences*, 2026, 58(1): 101–110. [郭金婷, 李娅姣, 邓永东, 等. 基于样本熵的大坝变形自适应预测模型[J]. *工程科学与技术*, 2026, 58(1): 101–110.]