

基于改进 IGGⅢ-ELM 法的混凝土坝变形监测数据粗差识别方法

王岩博^{1,2}, 顾冲时^{1,2}, 石立³, 顾昊^{1,2}, 张建中^{1,2}, 陆希³, 吴艳⁴, 朱明远⁴

(1. 河海大学水灾害防御全国重点实验室, 江苏 南京 210098; 2. 河海大学水利水电学院, 江苏 南京 210098;
3. 中国电建集团西北勘测设计研究院有限公司, 陕西 西安 710065;
4. 新疆水利水电科学研究院, 新疆 乌鲁木齐 830049)

摘要:针对混凝土坝变形监测数据存在非线性强、粗差识别与剔除困难等问题,结合改进 IGGⅢ法对异常值的抗差能力强、极限学习机(ELM)对数据序列预测效率高和对非线性问题处理能力强、增量 ELM 寻找最优网络结构速度快等优势,提出了基于改进 IGGⅢ-ELM 法的混凝土坝变形监测数据粗差识别方法,并采用包含两个调和系数的四段权函数对 IGGⅢ-ELM 法进行改进,使得权函数的一阶导数处处光滑,增强了权函数突变范围内信息的可利用性。与 IGGⅢ-ELM 法、DBSCAN 聚类算法、罗曼诺夫斯基准则和拉依达准则的实例处理结果对比表明:基于改进 IGGⅢ-ELM 法的混凝土坝变形监测数据粗差识别方法较其他 4 种方法粗差识别率更高、泛化能力更强、预测效果更好。

关键词: 变形监测数据; 混凝土坝; 粗差识别; 机器学习; 变形预测

中图分类号: TV698.1 **文献标志码:** A **文章编号:** 1006-7647(2023)06-0089-07

Gross error identification method for deformation monitoring data of concrete dams based on improved IGGⅢ-ELM method//WANG Yanbo^{1,2}, GU Chongshi^{1,2}, SHI Li³, GU Hao^{1,2}, ZHANG Jianzhong^{1,2}, LU Xi³, WU Yan⁴, ZHU Mingyuan⁴ (1. *The national key Laboratory of Water Disaster Prevention, Hohai University, Nanjing 210098, China*; 2. *College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China*; 3. *Northwest Engineering Corporation Limited, Xi'an 710065, China*; 4. *Xinjiang Institute of Water Resources and Hydropower Research, Urumqi 830049, China*)

Abstract: Aiming at the problems of strong nonlinearity, difficulty in identifying and eliminating gross errors in concrete dam deformation monitoring data, an approach of gross error identification and a dam safety monitoring model for deformation monitoring data of concrete dams are proposed based on the improved IGGⅢ-ELM method, which combines the advantages of the improved IGGⅢ method with the good robustness to outliers, the extreme learning machine (ELM) method with high efficiency in predicting data sequences and strong ability to handle nonlinear problems and the incremental ELM method with rapid optimal network structure seeking, the IGGⅢ-ELM method is improved by using a four-segment weight function containing two harmonic coefficients, making the first derivative of the weight function smooth everywhere, and enhancing the availability of information in the mutation range of the weight function. In the study case, the processing result of the improved IGGⅢ-ELM method was compared with the IGGⅢ-ELM method, DBSCAN clustering algorithm, Romanovsky criterion and Pauta criterion. The results show that the gross error identification and prediction method of concrete dam deformation monitoring data based on the improved IGGⅢ-ELM method has a gross error identification rate, stronger generalization ability and better prediction effect than the other four methods.

Key words: monitoring data; concrete dams; gross error identification; machine learning; deformation prediction

混凝土坝是重要的国民基础设施,然而大型混凝土坝一旦失事,会给下游人民的生命财产带来巨大损失,并对周边城市、村庄和生态环境造成毁灭性的灾难^[1],因此,确保混凝土坝健康长效服役至关

基金项目:国家自然科学基金联合基金项目(U2243223);中央高校基本科研业务费专项资金资助项目(B230201011);江苏省水利科技项目(2022024)

作者简介:王岩博(1998—),男,博士研究生,主要从事水工结构安全监控研究。E-mail:a19617230662022@163.com

通信作者:顾冲时(1962—),男,教授,博士,主要从事水工结构安全监控研究。E-mail:csgr@hhu.edu.cn

重要。混凝土坝在设计施工中存在的不合理因素和在服役期间受到的动静循环荷载与恶劣环境侵蚀,会导致混凝土性能和大坝结构整体安全性降低,严重影响大坝的安全运行^[2],需要大坝安全监控系统实时掌握混凝土坝真实的工作性态,及时发现混凝土坝运行过程中的缺陷和隐患,并对病险坝进行除险加固^[3-5]。

在获取混凝土坝安全监测原始数据时,经常会因为读数错误、计算错误、检测仪器故障等产生一定的粗差。粗差的存在严重影响了大坝监测数据序列的准确性,只有采取有效措施予以剔除,才能取得真实可靠的大坝安全预测效果。目前应用广泛的粗差去除方法有过程线法、统计检验法和数学模型法^[6-7],这些传统方法存在粗差识别率低,经常出现误判、漏判等问题,因此,Zhang 等^[8]基于集成学习算法提出了用实时数据更新的异常指标矩阵来识别大坝监测数据异常值的方法;Song 等^[9]基于多变量面板数据和 K 均值聚类算法建立了大坝异常值监测的理论方法;Shao 等^[10]利用图像处理技术和布谷鸟搜索算法处理大坝监测数据中的异常值。同时,也有众多学者将稳健估计理论引入粗差识别模型中,从而增强模型的鲁棒性和泛化能力。赵泽鹏等^[11]提出了一种基于最小协方差行列式估计的拉依达准则改进算法来识别大坝监测数据中的残差异异常值;Li 等^[12]提出了结合稳健统计量和置信区间的早期预警模型来识别监测异常值;李兴等^[13]提出了基于 M 估计的改进 Pauta 准则去除粗差算法,有效地解决了异常值误判的问题;胡德秀等^[14]结合稳健估计抵抗粗差能力强和极限学习机 (extreme learning machine, ELM) 能够处理非线性问题的优势,建立了基于 M-ELM 的大坝安全监控模型。与传统方法相比,这些方法能够较好地解决大坝变形数据中的异常值问题,但当存在连续偏离正常值或低偏差度的监测序列时,仍存在误判漏判现象,导致可利用信息在稳健估计中不能得到充分利用。这一现象需要进一步研究,以提高对异常数据的识别能力。

目前传统的大坝安全监控模型有数理统计模型、确定性模型和混合模型。葡萄牙、意大利等国家早在 20 世纪 50 年代就开始采用统计回归法建立大坝变形监控模型^[15-16]。我国大坝安全监控模型起步相对较晚,20 世纪 80 年代中期吴中如^[17]采用逐步回归、加权回归等数理统计方法,建立了实测数据与效应量之间的统计模型,并成功地应用于实际工程中。近年来,随着计算机、大数据和人工智能的发展,随机森林算法、神经网络、粒子群算法等智能算法相继被引入大坝安全监控领域。Huang 等^[18]首

次将核偏最小二乘法引入超高坝的安全监测中;Xu^[19]构建了基于径向基函数神经网络的大坝变形监测模型;Gu 等^[20]基于优化随机森林方法建立了混凝土坝监测因子挖掘模型;Chen 等^[21]基于狄利克雷过程-高斯混合模型聚类算法提出了一种考虑时空分异的大坝综合位移预测模型。这些智能算法能够更好地解决监控模型中复杂的非线性问题和众多的不确定因素,提高大坝安全监控模型的鲁棒性和泛化能力,进而提高大坝安全监控的预测精度。

选权迭代法中的 IGGⅢ法是稳健估计中最常用的方法,其初值的选取一般采用最小二乘法。然而,IGGⅢ权函数的一阶导数存在不光滑的“尖点”,导致其附近范围内的信息不能被充分准确地利用,对于小偏差异常值易造成数据误判漏判的现象。因此,本文结合 IGGⅢ与 ELM 方法进行混凝土坝变形监测数据的粗差识别与回归分析,采用包含两个调和系数的四段权函数对 IGGⅢ-ELM 进行改进(以下简称“改进 IGGⅢ-ELM 法”),并结合工程案例讨论该方法在粗差识别和回归预测中的效果,以期更好地解决监测模型中复杂的非线性问题和众多的不确定因素,提高大坝安全监测模型的鲁棒性和泛化能力,进而提高大坝安全监测的预测精度。

1 改进 IGGⅢ-ELM 法

1.1 IGGⅢ-ELM 法

ELM 是一种简单高效的单隐含层前馈神经网络算法,相比于传统的神经网络算法具有学习速度快、泛化能力强、预测精度高等优点^[22-23]。神经网络隐含层节点输出权重^[14]为

$$\hat{\beta} = (H^T B H)^{-1} H^T B T^T \quad (1)$$

式中: H 为隐含层输出矩阵; B 为隐含层的随机偏置矩阵, $B = (b_1, b_2, \dots, b_l)^T$; T 为混凝土坝的变形实测值矩阵。

若大坝实际变形的权重为 P_i ,令 $\bar{P} = (\bar{P}_i) = (P_i b_i)$,其中 $\bar{P}_i$ 为 IGGⅢ等价权函数的一阶导数,则法方程组矩阵为

$$\hat{\beta} = (H^T \bar{P} H)^{-1} H^T \bar{P} T^T \quad (2)$$

稳健估计中的选权迭代法是根据残差大小确定权因子,其中最有效的方案是利用 IGGⅢ法逐步降低异常值的权重,达到准确识别粗差的目的。具体公式为

$$\bar{P}_i = \begin{cases} 1 & |\tilde{v}_i| \leq k_1 \\ \frac{k_1}{|\tilde{v}_i|} \left(\frac{k_2 - |\tilde{v}_i|}{k_2 - k_1} \right)^2 & k_1 < |\tilde{v}_i| \leq k_2 \\ 0 & |\tilde{v}_i| > k_2 \end{cases} \quad (3)$$

式中: $k_1=1.5$; $k_2=2.5$; $\tilde{v}_i$ 为标准化残差。处理方法如表 1 所示。

表 1 IGGⅢ法权重处理方法			
判别条件	监测值	措施	权重
$ \tilde{v}_i \leq k_1$	正常	保权	1
$k_1 < \tilde{v}_i \leq k_2$	可利用	降权	$>0 \sim <1$
$ \tilde{v}_i > k_2$	异常	淘汰	0

1.2 IGGⅢ-ELM 法的改进

稳健估计是一个保留有效数据信息、削弱可用数据信息、剔除干扰数据信息的反复迭代过程。IGGⅢ权函数的一阶导数在 k_1 处存在不光滑的“尖点”,使得 k_1 附近的数据信息不能得到充分利用,易出现误判漏判的情况。因此,本文采用包含两个调和系数的四段权函数一阶导数来替代式(3):

$$\bar{P}_i = \begin{cases} 1 & |\tilde{v}_i| \leq k_1 \\ 1 - 2 \left(\frac{|\tilde{v}_i| - k_1}{k_2 - |\tilde{v}_i|} \right)^2 & k_1 < |\tilde{v}_i| \leq \frac{k_1 + k_2}{2} \\ 2 \left(\frac{k_2 - |\tilde{v}_i|}{|\tilde{v}_i| - k_1} \right)^2 & \frac{k_1 + k_2}{2} < |\tilde{v}_i| \leq k_2 \\ 0 & |\tilde{v}_i| > k_2 \end{cases} \quad (4)$$

改进 IGGⅢ的权函数如图 1 所示。

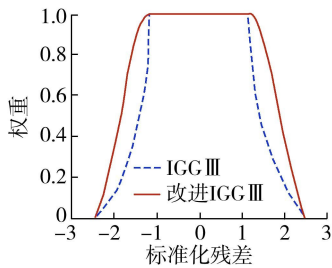


图 1 IGGⅢ及改进方案的权函数

1.3 增量 ELM 法确定隐含层节点数

在 ELM 中,隐含层节点数是影响网络泛化能力的一个重要参数,本文采取增量 ELM 法确定隐含层的节点个数。该方法的原理是初始化一个较小的 ELM 网络结构,逐渐增加隐含层的节点个数,使其结果达到期望精度为止。具体步骤如下:

假设 ELM 网络结构初始隐含层节点数为 $s-1$, 则 $\boldsymbol{\beta}_{s-1} = (\boldsymbol{H}_{s-1}^T \boldsymbol{H}_{s-1})^{-1} \boldsymbol{H}_{s-1}^T \boldsymbol{T}^T$, 其中 $\boldsymbol{\beta}_{s-1}$ 表示节点数为 $s-1$ 时隐含层与输出层之间的连接权重;当新的节点加入后,有 $\boldsymbol{\beta}_s = (\boldsymbol{H}_s^T \boldsymbol{H}_s)^{-1} \boldsymbol{H}_s^T \boldsymbol{T}^T$, 其中 $\boldsymbol{H}_s = (\boldsymbol{H}_{s-1}, \boldsymbol{v}_s)$, $\boldsymbol{v}_s$ 为与第 s 个隐含层节点连接的隐含层输出向量。

由于 $(\boldsymbol{H}_s^T \boldsymbol{H}_s)^{-1}$ 是对称的,因此有

$$(\boldsymbol{H}_s^T \boldsymbol{H}_s)^{-1} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^T & \boldsymbol{E} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{A}' & \boldsymbol{B}' \\ \boldsymbol{B}'^T & \boldsymbol{E}' \end{bmatrix} \quad (5)$$

$$\begin{aligned} \text{其中 } \boldsymbol{A} &= \boldsymbol{H}_{s-1}^T \boldsymbol{H}_{s-1} \quad \boldsymbol{B} = \boldsymbol{H}_{s-1}^T \boldsymbol{v}_s \quad \boldsymbol{C} = \boldsymbol{v}_s^T \boldsymbol{v}_s \\ \text{因此有} \\ \boldsymbol{I} &= \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^T & \boldsymbol{E} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}' & \boldsymbol{B}' \\ \boldsymbol{B}'^T & \boldsymbol{E}' \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}\boldsymbol{A}' + \boldsymbol{B}\boldsymbol{B}'^T & \boldsymbol{A}\boldsymbol{B}' + \boldsymbol{B}\boldsymbol{E}' \\ \boldsymbol{B}^T\boldsymbol{A}' + \boldsymbol{E}\boldsymbol{B}'^T & \boldsymbol{B}^T\boldsymbol{B}' + \boldsymbol{E}\boldsymbol{E}' \end{bmatrix} \end{aligned} \quad (6)$$

解得:

$$\begin{cases} \boldsymbol{A}' = \frac{\boldsymbol{A}^{-1} \boldsymbol{B} (\boldsymbol{A}^{-1} \boldsymbol{B})^T}{\boldsymbol{E} - \boldsymbol{B}^T (\boldsymbol{A}^{-1} \boldsymbol{B})} + \boldsymbol{A}^{-1} \\ \boldsymbol{B}' = \frac{\boldsymbol{A}^{-1} \boldsymbol{B}}{\boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B} - \boldsymbol{E}} \\ \boldsymbol{C}' = \frac{1}{\boldsymbol{E} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B}} \end{cases} \quad (7)$$

因此可得:

$$\boldsymbol{D}_s = \begin{bmatrix} \boldsymbol{A}' & \boldsymbol{B}' \\ \boldsymbol{B}'^T & \boldsymbol{E}' \end{bmatrix} \begin{bmatrix} \boldsymbol{H}_{s-1}^T \\ \boldsymbol{v}_s^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}' \boldsymbol{H}_{s-1}^T + \boldsymbol{B}' \boldsymbol{v}_s^T \\ \boldsymbol{B}'^T \boldsymbol{H}_{s-1}^T + \boldsymbol{E}' \boldsymbol{v}_s^T \end{bmatrix} \quad (8)$$

令 $\boldsymbol{D}_s = (\boldsymbol{M}, \boldsymbol{N})^T$, 把式(7)代入式(8)可以计算 $\boldsymbol{M}$ 和 $\boldsymbol{N}$, 即可算得 $\boldsymbol{\beta}_s$ 。

1.4 大坝监测数据中剔除粗差的具体步骤

对于给定的数据拟合段 N , 初始隐含层节点个数为 s , 最大隐含层节点数为 $s_{\max}$, 期望精度为 ε 。

步骤 1 在 ELM 网络结构中随机生成输入大坝环境量因子的权重矩阵 $\boldsymbol{\omega}$ 和偏置矩阵 $\boldsymbol{B} (i=1, 2, \dots, s)$ 。

步骤 2 令 $s=1$, 计算隐含层输出矩阵 $\boldsymbol{H}_0$ 和混凝土坝拟合效应量因子权重 $\boldsymbol{\beta}_0$, 并计算网络学习精度 $e_s, e_s = e_0$ 。

步骤 3 当 $s < s_{\max}$ 且 $e_s > \varepsilon$ 时, 令 $s=s+1$, 计算输出矩阵 $\boldsymbol{H}_s$ 和混凝土坝拟合效应量因子权重 $\boldsymbol{\beta}_s$, 直至满足条件为止, 得到最优隐含层节点数 s^* 。

步骤 4 利用改进 IGGⅢ法计算混凝土坝变形监测数据的初始等价权矩阵 $\bar{\boldsymbol{P}}$ 。

步骤 5 根据步骤 3 中 s^* 对应的 $\boldsymbol{H}_{s^*}$ 和步骤 4 中 $\bar{\boldsymbol{P}}$, 利用式(2)计算 $\hat{\boldsymbol{\beta}}$ 。

步骤 6 逐次迭代计算, 直至 $|\hat{\boldsymbol{\beta}}_{i-1} - \hat{\boldsymbol{\beta}}_i| \leq \delta$ 时停止。

步骤 7 输出混凝土坝在连续时间序列上的监测变形权重 P_i, P_i 为 0 的样本序列判定为粗差, 予以剔除。

改进 IGGⅢ-ELM 法粗差识别流程如图 2 所示。

2 工程案例

2.1 工程概况

某水电站是雅砻江干流梯级滚动开发的龙头水库电站, 水库正常蓄水位 1 880 m, 主要任务是发电, 兼有防洪、拦沙作用。该水电站的主要建筑物为

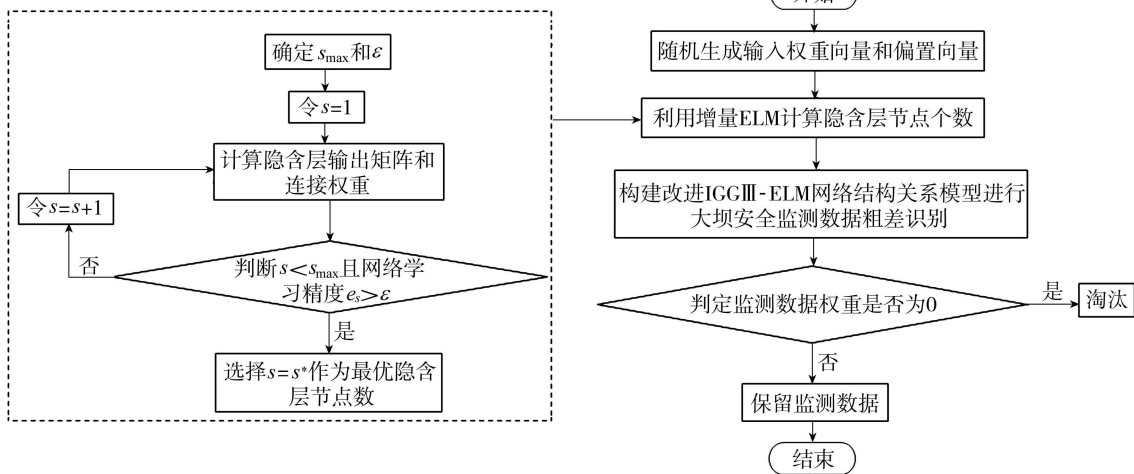


图2 改进 IGG III-ELM 法粗差识别流程

混凝土双曲拱坝,坝顶高程1885 m,最大坝高305 m。本文选取电站运行期2016年9月1日至2018年5月19日的位移监测数据,分析并验证所用方法的准确性和适用性。图3为该水电站双曲拱坝的三维模型。



图3 双曲拱坝的三维模型

2.2 模型影响因子的选取及数据集的划分

在混凝土坝运行过程中,大坝会受到水压力、扬压力、泥沙压力和温度等荷载的共同作用,选取水压分量 δ_H 、温度分量 δ_T 、时效分量 δ_θ 为自变量,大坝变形的位移 δ 为因变量,利用 ELM 建立自变量与因变量间的非线性映射,能够较好地构建大坝变形位移的预测模型,其表达式为

$$\delta = \delta_H + \delta_T + \delta_\theta \quad (9)$$

选取 $H-H_0$ 、 $H^2-H_0^2$ 、 $H^3-H_0^3$ 、 $H^4-H_0^4$ 作为水压因子, $\sin \frac{2\pi t}{365} - \sin \frac{2\pi t_0}{365}$ 、 $\cos \frac{2\pi t}{365} - \cos \frac{2\pi t_0}{365}$ 、 $\sin \frac{4\pi t}{365} - \sin \frac{4\pi t_0}{365}$ 、 $\cos \frac{4\pi t}{365} - \cos \frac{4\pi t_0}{365}$ 作为温度因子, $\theta - \theta_0$ 、 $\ln \theta - \ln \theta_0$ 作为时效因子,其中 H 为上游水位, H_0 为观测起始日的上游水位, t 为监测日至实测日的累计天数, $t_0 = 1$, $\theta = 0.01t$, $\theta_0 = 0.01t_0$ 。

拱冠梁的监测数据是馈控大坝变形安全的关键性数据,因此选取13号坝段坝中监测点 PL13-3 作为研究对象,将其径向位移监测数据划分为拟合段

(用于模型拟合的数据样本)和验证段(用于验证模型泛化能力的数据样本),拟合段为2016年9月1日至2018年3月22日(共计522个数据样本),验证段为2018年3月23日至2018年5月19日(共计58个数据样本)。

为了检验改进 IGG III-ELM 法识别粗差的效果,将震荡、台坎、突跳、台阶4种类型及偏差量较小的粗差引入到测点 PL13-3 中(粗差类型分别计为①、②、③、④、⑤),测点的位移过程线见图4。

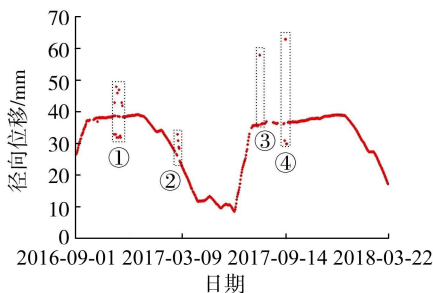


图4 PL13-3 测点加入粗差的位移过程

首先在 ELM 网络结构中随机产生权重 ω_i 和 b_i ,将隐含层节点数初始值取为1,将拟合段数据按式(5)处理后输入该网络结构。其次不断增加隐含层节点数 s ,当 $s=42$ 时,满足所需的学习精度,因此选定隐含层节点数为42个。取 $\delta=0.001$,利用改进 IGG III-ELM 法进行选权迭代,计算每个拟合段数据的权重。图4中所插入粗差的初次迭代权重如表2所示,最终迭代权重均降为0,根据最终迭代权重可识别出拟合段所有的异常值数据。

2.3 粗差去除效果对比分析

采用 IGG III-ELM 法、DBSCAN 聚类算法、罗曼诺夫斯基准则、拉依达准则与改进 IGG III-ELM 法进行对比,分析不同方法去除粗差的效果。表3为不同方法的粗差识别情况对比。由表3可知,传统的

表 2 粗差初次迭代权重

日期	初次迭代权重	日期	初次迭代权重
2016-01-11	0.972 3	2016-11-25	0.926 1
2016-10-12	0.986 7	2017-03-05	0
2016-10-13	1	2017-03-06	0.677 2
2016-11-10	0	2017-03-07	1
2016-11-11	0.841 5	2017-03-08	1
2016-11-13	0	2017-04-05	0.962 2
2016-11-14	0	2017-04-06	0
2016-11-15	0	2017-04-07	0
2016-11-16	0	2017-08-01	0
2016-11-18	0	2017-09-15	0
2016-11-19	0	2017-09-16	0
2016-11-21	0	2017-09-17	0
2016-11-22	0	2017-09-18	0
2016-11-24	0.696 7	2017-09-19	0

拉依达准则和罗曼诺夫斯基准则粗差识别能力较差,识别率分别为 7.14% 和 10.71%;DBSCAN 聚类算法能识别出 16 个粗差,识别率为 57.14%;IGGⅢ-ELM 法能识别出①、②、③、④类型的粗差,但没有识别出粗差⑤,表明该方法对于小偏差量的异常值没有抵抗能力;改进 IGGⅢ-ELM 法能识别出 28 个粗差,识别率为 100%。通过对比可知,改进 IGGⅢ-

ELM 法不仅具有更好的稳健能力,而且能够充分利用 k_1 附近的数据信息,精准地识别出大坝位移监测数据中偏差量较小的异常值。

2.4 剔除粗差后的验证效果

根据表 3 数据,使用改进 IGGⅢ-ELM 法对 5 种方法剔除异常值后的数据集建立大坝安全监测模型,并根据 2018 年 5 月 21 日至 2018 年 12 月 30 日的水压、温度、时效分量对大坝径向位移进行预测,从而评估模型最终的泛化能力,如图 5 所示。由图 5 可知,利用改进 IGGⅢ-ELM 法剔除异常值后,模型拟合段的拟合值与实测值基本重合,拟合效果很好;从预测段的散点排列中可以看出,预测段的大坝径向位移变化趋势符合拟合段的大坝径向位移变化规律,表明该方法具有准确的预测精度和良好的泛化能力。

选取平均绝对误差 (MAE)、均方根误差 (RMSE)、决定系数 (R^2) 作为评判指标进行对比,结果如表 4 所示。从表 4 可知,在剔除异常值后的大坝径向位移验证中,5 种方法的 R^2 均在 0.85 以上,验证段的 MAE、RMSE 均在 1 以内,表明利用改进

表 3 粗差插入位置 and 不同方法的粗差识别情况

日期	原始位移/mm	位移测值 粗差/mm	加入粗差后的 位移测值/mm	粗差识别情况				
				改进 IGGⅢ- ELM 法	IGGⅢ- ELM 法	DBSCAN 聚类算法	罗曼诺夫 斯基准则	拉依达准则
2016-10-11	38.14	-1.13	37.01	√				
2016-10-12	38.13	-1.23	36.90	√				
2016-10-13	38.15	-1.15	37.00	√				
2016-11-10	38.85	-5.85	33.00	√	√	√		
2016-11-11	38.75	4.25	43.00	√	√	√		
2016-11-13	38.82	-5.82	33.00	√	√			
2016-11-14	38.81	9.19	48.00	√	√	√		
2016-11-15	38.78	-6.78	32.00	√	√			
2016-11-16	38.70	7.30	46.00	√	√	√		
2016-11-18	38.67	-6.67	32.00	√	√			
2016-11-19	38.60	8.40	47.00	√	√	√		
2016-11-21	38.52	-6.02	32.50	√	√	√		
2016-11-22	38.48	-6.48	32.00	√	√	√		
2016-11-24	38.47	4.53	43.00	√	√	√		
2016-11-25	38.48	3.52	42.00	√	√	√		
2017-03-05	26.05	6.95	33.00	√	√	√		
2017-03-06	25.72	-12.22	13.50	√	√			
2017-03-07	25.36	3.64	29.00	√	√			
2017-03-08	25.00	3.50	28.50	√	√			
2017-04-05	13.87	1.30	15.17	√				
2017-04-06	13.57	1.40	14.97	√				
2017-04-07	13.14	1.50	14.64	√				
2017-08-01	28.24	29.76	58.00	√	√	√	√	
2017-09-15	36.56	-5.56	31.00	√	√	√		
2017-09-16	36.48	26.52	63.00	√	√	√	√	√
2017-09-17	36.73	26.27	63.00	√	√	√	√	√
2017-09-18	36.74	-6.74	30.00	√	√	√		
2017-09-19	36.63	-6.63	30.00	√	√	√		

注:“√”表示该方法可以识别粗差。

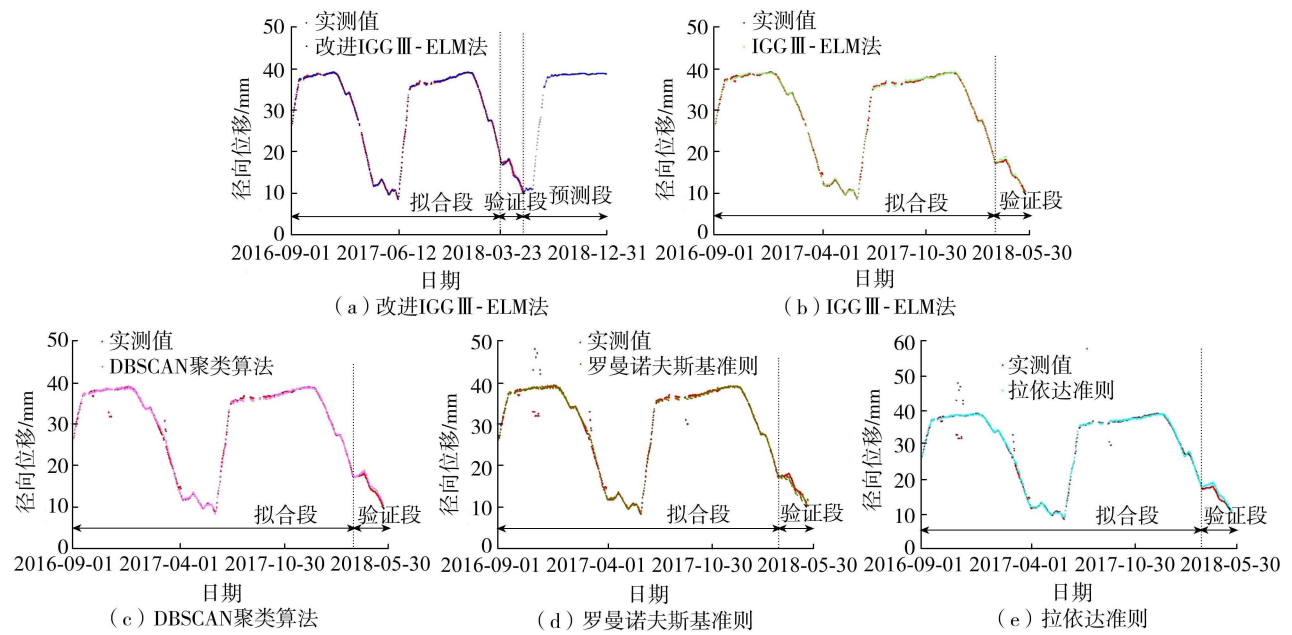


图 5 5 种方法剔除粗差后的大坝监测模型泛化能力

表 4 评判指标对比

粗差识别方法	训练段			验证段		
	MAE/mm	RMSE/mm	R^2	MAE/mm	RMSE/mm	R^2
改进 IGGⅢ-ELM 法	0.0890	0.126 1	0.999 8	0.230 1	0.345 7	0.981 5
IGGⅢ-ELM 法	0.1790	0.247 2	0.998 3	0.415 8	0.485 5	0.960 4
DBSCAN 聚类算法	0.3764	0.952 6	0.970 6	0.557 2	0.612 7	0.937 5
罗曼诺夫斯基准则	0.4684	1.535 5	0.976 2	0.866 8	0.910 7	0.861 4
拉依达准则	0.4104	1.158 2	0.986 3	0.594 7	0.731 6	0.910 6

IGGⅢ-ELM 法进行大坝安全监测建模是合理和准确的。其中利用改进 IGGⅢ-ELM 法剔除粗差后拟合段和验证数据是 5 种方法中误差最小、相关性最强的一组数据。综上可知,利用改进 IGGⅢ-ELM 法去除粗差的比例最高,对后续大坝径向位移的监测效果最好。

3 结 语

针对复杂因素影响下大坝安全监测数据中的粗差问题,结合增量 ELM 法寻找最优网络结构速度快、改进 IGGⅢ选权迭代法对异常值的抗差能力强、ELM 对数据序列的预测效率高和处理非线性问题能力强等方面的优势,提出了基于改进 IGGⅢ-ELM 法的混凝土坝变形监测数据粗差识别方法。改进 IGGⅢ-ELM 法能够很好地解决连续时间序列数据中的粗差识别问题。工程案例表明,在人为添加 28 个粗差的大坝监测数据序列中,粗差识别率为 100%。相比于其他 4 种粗差识别方法,该模型剔除粗差效果更好、识别精度更高、泛化能力更强,验证了模型的准确性与可靠性。通过 5 种方法剔除粗差后,利用改进 IGGⅢ-ELM 法对大坝安全进行监测,结果表明所用方法在剔除粗差后具有更好的拟合效

果和更高的预测精度。

参考文献:

[1] 顾冲时,苏怀智,刘何稚. 大坝服役风险分析与管理研究述评[J]. 水利学报, 2018, 49 (1) : 26-35. (GU Chongshi, SU Huaizhi, LIU Hezhi. Review on service risk analysis of dam engineering [J]. Journal of Hydraulic Engineering, 2018, 49 (1) : 26-35. (in Chinese))

[2] 袁冬阳,顾冲时,顾昊. 严寒地区混凝土重力坝变形行为分析与预测模型[J]. 水利学报, 2022, 53 (6) : 733-746. (YUAN Dongyang, GU Chongshi, GU Hao. Displacement behavior analysis and prediction model of concrete gravity dams in cold region [J]. Journal of Hydraulic Engineering, 2022, 53 (6) : 733-746. (in Chinese))

[3] 顾昊,曹文翰,汪程,等. 混凝土坝服役性态监测效应量安全监控指标拟定方法[J]. 水利水电科技进展, 2021, 41 (1) : 30-34. (GU Hao, CAO Wenhan, WANG Cheng, et al. Safety monitoring index determination for monitoring effect quantity of concrete dams in service [J]. Advances in Science and Technology of Water Resources, 2021, 41 (1) : 30-34. (in Chinese))

[4] 赵二峰,顾冲时. 混凝土坝长效服役性态健康诊断研究述评[J]. 水力发电学报, 2021, 40 (5) : 22-34. (ZHAO Erfeng, GU Chongshi. Review on health diagnosis of long-

- term service behaviors for concrete dams[J]. Journal of Hydroelectric Engineering, 2021, 40 (5): 22-34. (in Chinese))
- [5] 赵二峰,顾冲时.碾压混凝土坝安全服役关键技术研究进展[J]. 水利水电科技进展,2022,42(1):11-20. (ZHAO Erfeng, GU Chongshi. Review on key technologies for safe operation of roller compacted concrete dams[J]. Advances in Science and Technology of Water Resources,2022,42(1):11-20. (in Chinese))
- [6] 刘千驹,陈代明,陈少勇,等.小波理论在大坝安全监测数据粗差探测中的应用[J]. 西北水电,2020(增刊1):129-132. (LIU Qianju, CHEN Daiming, CHEN Shaoyong, et al. Application of wavelet theory in detection of gross errors in dam safety monitoring data [J]. Northwest Hydropower,2020(Sup1):129-132. (in Chinese))
- [7] 花胜强,李永红,高磊,等.基于离差和改进的大坝变形监测粗差判别方法[J]. 水电能源科学,2020,38(6):67-69. (HUA Shengqiang, LI Yonghong, GAO Lei, et al. A gross error discriminant method based on difference squares in dam displacement monitoring [J]. Water Resources and Power,2020,38(6):67-69. (in Chinese))
- [8] ZHANG Jun, XIE Jiemin, KOU Pangao. Abnormal diagnosis of dam safety monitoring data based on ensemble learning [J]. IOP Conference Series: Earth and Environmental Science,2019,267(6):062027.
- [9] SONG Jintao, ZHANG Shengfei, TONG Fei, et al. Outlier detection based on multivariable panel data and K-means clustering for dam deformation monitoring data [J]. Advances in Civil Engineering,2021,2021:3739551.
- [10] SHAO Chenfei, ZHENG Sen, GU Chongshi, et al. A novel outlier detection method for monitoring data in dam engineering[J]. Expert Systems with Applications,2022,193:116476.
- [11] 赵泽鹏,陈建康,张瀚,等.基于稳健估计的大坝监测数据粗差识别方法[J]. 水电能源科学,2018,36(12):68-71. (ZHAO Zepeng, CHEN Jiankang, ZHANG Han, et al. The method for gross error identification of dam monitoring data based on robust estimation[J]. Water Resources and Power,2018,36(12):68-71. (in Chinese))
- [12] LI Xing, LI Yanling, LU Xiang, et al. An online anomaly recognition and early warning model for dam safety monitoring data[J]. Structural Health Monitoring, 2020, 19(3):796-809.
- [13] 李兴,李艳玲,张鹏,等.基于M估计的改进Pauta准则在监测数据粗差识别中的研究及应用[J]. 中国农村水利水电, 2019 (8): 133-136. (LI Xing, LI Yanling, ZHANG Peng, et al. Research on an improved Pauta criterion based on M-estimation for gross error identification of monitoring data and its application[J]. China Rural Water and Hydropower,2019(8):133-136. (in Chinese))
- [14] 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))
- [15] 马佳佳,苏怀智,王颖慧.基于EEMD-LSTM-MLR的大坝变形组合预测模型[J]. 长江科学院院报,2021,38(5):47-54. (MA Jiajia, SU Huaizhi, WANG Yinghui. Combinatorial prediction model for dam deformation based on EEMD-LSTM-MLR [J]. Journal of Yangtze River Scientific Research Institute, 2021, 38 (5): 47-54. (in Chinese))
- [16] 侯回位,郑东健,刘永涛,等.基于EEMD-SE-LSTM的混凝土坝变形监测模型[J]. 水利水电科技进展,2022,42(1):61-66. (HOU Huiwei, ZHENG Dongjian, LIU Yongtao, et al. Deformation monitoring model of concrete dams based on EEMD-SE-LSTM [J]. Advances in Science and Technology of Water Resources, 2022, 42 (1): 61-66. (in Chinese))
- [17] 吴中如.论混凝土坝安全监控的确定性模型及混合模型[J]. 水利学报,1989(5):64-70. (WU Zhongru. Research on the deterministic model and hybrid model of concrete dam safety monitoring[J]. Journal of Hydraulic Engineering,1989(5):64-70. (in Chinese))
- [18] HUANG Hao, CHEN Bo, LIU Chungao. Safety monitoring of a super-high dam using optimal kernel partial least squares[J]. Mathematical Problems in Engineering,2015,2015:571594.
- [19] XU Guanghua. Application of RBF neural network in dam deformation prediction [J]. Applied Mechanics and Materials,2014,675-677:261-264.
- [20] GU Hao, YANG Meng, GU Chongshi, et al. A factor mining model with optimized random forest for concrete dam deformation monitoring [J]. Water Science and Engineering,2021,14(4):330-336.
- [21] CHEN Wenlong, WANG Xiaoling, CAI Zhijian, et al. DP-GMM clustering-based ensemble learning prediction methodology for dam deformation considering spatiotemporal differentiation [J]. Knowledge-Based Systems,2021,222:106964.
- [22] 董小刚,曹馨潼,秦喜文.基于改进LSTM-ELM模型的汇率预测研究[J]. 中国集体经济,2022(29):112-116. (DONG Xiaogang, CAO Xintong, QIN Xiwen. Exchange rate prediction study based on an improved LSTM-ELM model[J]. China Collective Economy, 2022 (29): 112-116. (in Chinese))
- [23] 岳兆新,艾萍,熊传圣,等.基于信息熵与改进极限学习机的中长期径流预测[J]. 水利水电科技进展,2021,41(4):7-14. (YUE Zhaoxin, AI Ping, XIONG Chuansheng, et al. Medium-long term runoff forecasting based on information entropy and improved extreme learning machine [J]. Advances in Science and Technology of Water Resources,2021,41(4):7-14. (in Chinese))

(收稿日期:2022-11-02 编辑:俞云利)