

DOI:10.11784/tdxbz202502005

# 考虑空间分异性的土石坝变形安全分区评价指标拟定模型

王 雷, 王晓玲, 张 君, 王佳俊

(天津大学水利工程智能建设与运维全国重点实验室, 天津 300072)

**摘要:** 针对传统大坝安全评价指标仅表征整体结构安全状态而忽略了变形测点空间分布差异性与类聚性, 由此导致其无法揭示坝体局部安全状况的问题, 本研究提出了基于改进算数优化算法优化注意力双向门控循环单元自编码器、凝聚型层次聚类(AHC)与超阈值(POT)理论的土石坝变形安全分区评价指标拟定模型。首先, 本文建立一种有效的序列数据降维模型, 通过引入注意力机制耦合双向门控循环单元自编码器, 解决了传统单向门控循环单元序列信息提取不完整及数据降维过程中的信息丢失问题。其次, 设计了一种混沌搜索策略改进的算数优化算法, 显著提升了网络超参数的优化效率, 有效避免了深度学习模型超参数优化易陷入局部最优解的问题。随后, 采用基于曼哈顿距离的 AHC 方法, 有效实现了坝体变形监测数据的空间分区, 并在空间分区基础上结合 POT 理论进行安全诊断指标拟定。实际土石坝工程案例分析结果表明: 本文所提方法的聚类性能优异, 聚类评价轮廓系数高达 0.886, 戴维斯-鲍丁指数低至 0.151, 显著优于现有方法; 安全诊断指标考虑了坝体结构的时空分异性, 合理性显著提高。本研究提升了大坝变形监测数据挖掘的深度与精度, 为大坝安全监测与评价指标研究提供了新思路。

**关键词:** 土石坝; 安全评价指标拟定; 算数优化算法; 双向门控循环单元; 自编码器; 注意力机制; 凝聚型层次聚类; 超阈值模型

中图分类号: TV641

文献标志码: A

文章编号: 0493-2137(2025)11-1205-13

## Formulation Model for Safety Zoning Evaluation Indexes of Rockfill Dam Deformation Considering Spatial Differentiation

Wang Lei, Wang Xiaoling, Zhang Jun, Wang Jiajun

(State Key Laboratory of Hydraulic Engineering Intelligent Construction and Operation, Tianjin University, Tianjin 300072, China)

**Abstract:** The traditional dam safety evaluation indexes only characterize the overall structural safety state and neglect the spatial distribution variability and clustering characteristics of deformation monitoring points, thereby failing to accurately reveal the local safety conditions within the dam structure. To address this issue, a safety zoning evaluation index formulation model for rockfill dam deformation is proposed in this paper, which integrates an improved arithmetic optimization algorithm-optimized attention-based bi-directional gated recurrent unit autoencoder, agglomerative hierarchical clustering(AHC) and the peak over threshold(POT) theory. First, an efficient dimensionality reduction model for sequential data is established by coupling an attention mechanism with a bidirectional gated recurrent unit autoencoder, thus solving the problems associated with incomplete sequence information extraction inherent in the traditional unidirectional gated recurrent unit models and information loss during dimensionality reduction. Second, a chaos search strategy-enhanced arithmetic optimization algorithm is designed, significantly improving the optimization efficiency for network hyperparameters and effectively avoiding the common problem of local optima encountered during deep learning model hyperparameter optimization. Third, the AHC method based on Manhattan distance is employed to achieve effective spatial zoning of dam deformation monitoring data, upon which

收稿日期: 2025-02-11; 修回日期: 2025-05-09.

作者简介: 王 雷(1992—), 男, 博士研究生, wlayp\_123@tju.edu.cn.

通信作者: 张 君, zhangdajun@tju.edu.cn.

基金项目: 国家自然科学基金青年科学基金资助项目(52309165).

Supported by the Youth Program of the National Natural Science Foundation of China (No. 52309165).

the safety diagnostic indexes are formulated using the POT theory. The application to a practical rockfill dam engineering case demonstrates that the proposed method exhibits superior clustering performance, achieving a silhouette coefficient as high as 0.886 and a Davies-Bouldin index as low as 0.151, obviously surpassing the existing methods. Furthermore, the formulated safety diagnostic indexes account for spatial variability within the dam structure, thereby significantly enhancing their rationality. The approach developed in this paper improves the depth and accuracy of dam deformation monitoring data mining and provides novel insights into dam safety monitoring and research on evaluation indexes.

**Keywords:** rockfill dam; formulation of safety evaluation indexes; arithmetic optimization algorithm; bidirectional gated recurrent unit; autoencoder; attention mechanism; agglomerative hierarchical clustering (AHC); peak over threshold (POT) model

在大坝建设与运行阶段,科学有效的大坝安全监测反映了大坝整体结构的安全性态,而安全监测效应量是否异常需要通过安全诊断指标来判断<sup>[1-2]</sup>。通过收集大坝设计、施工、运维各阶段的安全监测资料,采用概率统计<sup>[3]</sup>、结构计算<sup>[4]</sup>等分析方法确定各监测效应量的界限值即为大坝安全诊断指标的拟定过程,其本质为根据大坝和周边结构历史上已经历的外来荷载确定其监测效应量的安全范围,因此合理的指标拟定对大坝的风险控制和安全稳定具有重要的意义<sup>[5]</sup>。

大坝在其全生命周期中面临的外来荷载具有复杂性、动态性、叠加性<sup>[6]</sup>等特点。同时,由于材料特性和人为因素,大坝本身的结构性态亦随时间发生变化,从而导致大坝安全诊断指标的拟定较为困难。因此,在大坝安全监控领域,该问题得到了相关科研工作者的广泛研究。现有关于大坝安全诊断指标拟定的研究大多基于两大类方法<sup>[7]</sup>:一类是通过数理统计的方法分析和挖掘历史安全监测数据,简称概率统计法,包括小概率法<sup>[8-9]</sup>、云模型<sup>[10-11]</sup>、置信区间<sup>[12]</sup>等;另一类是采用结构计算的思路,通过数值模拟确定各种荷载引起的监测效应量变化,简称结构计算法<sup>[13-15]</sup>,包括极限状态、有限元仿真计算等。还有一些研究混合了两类方法来确定诊断指标<sup>[16-17]</sup>。

已有研究方法能取得较为合理的安全诊断指标,亦各有优势:概率统计法简单易行,结构计算法有明确的物理意义且可以解释坝体安全失效的机理,而混合法能够综合二者优势。随着大坝安全监测技术的快速发展,监测点在空间中的分布更加密集<sup>[18]</sup>,各测点在空间中既有分异性,也有类聚性。在大坝变形预测领域,已有研究考虑测点之间的空间相关性以提高预测精度<sup>[19-21]</sup>。在现有的大坝安全诊断指标拟定中,大多数研究拟定的指标是反映大坝整体安全状态的界限值,而实际上大坝的安全性态亦有空间分布特点,因此在安全诊断指标拟定中,应考虑到测点的空间分布特征以提高所拟定指标的合理性,进而实现空

间分区安全诊断。

大坝变形分区能够将坝体变形的完整复杂性态在空间上实现分解,为准确评估大坝安全状态提供基础。目前关于大坝变形分区的研究较少,大多采用  $k$  均值、层次聚类等传统无监督机器学习聚类算法<sup>[22-24]</sup>,主要聚焦于变形监测数据相似性度量上的改进<sup>[24-26]</sup>。然而,大坝变形监测数据具有时间依赖性和复杂的非线性关系,传统的聚类方法无法处理这些问题。相较于传统的机器学习模型,深度学习模型能够全面地深度挖掘数据的隐藏特征,实现精确建模<sup>[27]</sup>。利用深度学习方法进行序列聚类的基本思路为:首先实现准确序列建模,而后提取高维数据的低维特征表示,最终基于原始数据的低维特征进行聚类分析<sup>[28]</sup>。

在序列建模领域,递归神经网络尤其是其变体门控循环单元(gated recurrent unit, GRU)既能基于其门控机制有效捕捉序列数据中的长、短期依赖特征,又能通过非线性激活函数处理序列数据的非线性关系<sup>[29-30]</sup>。然而,单向 GRU 只能按照过去到未来的顺序进行序列建模,这在某些任务中可能导致提取信息不完整,例如序列重建<sup>[31]</sup>。双向门控循环单元(bidirectional GRU, BiGRU)则能从前向和后向两个角度全面地提取序列信息<sup>[32]</sup>。学习率、 $L_2$  正则化系数等超参数的设置对 BiGRU 建模精度有重要的影响,而且该模型处理长序列数据仍然存在信息丢失的问题。因此,一方面需要对 BiGRU 的超参数进行高效率的优化以提升模型性能表现,另一方需要解决其信息丢失的缺陷。

在数据降维方面,自动编码器(autoencoder, AE)是一种无监督的确定性前馈人工神经网络结构,也是一种数据降维技术,由编码器和解码器组成<sup>[33]</sup>。AE能够在去除冗余信息的基础上,提取高维输入数据主要特征,实现有效降维。

综上所述,为了在大坝变形安全诊断指标拟定考虑变形的空间类聚性,本文提出基于改进算数优化算

法 (improved arithmetic optimization algorithm, IAOA) 优化注意力双向门控循环单元自编码器 (attention-based BiGRU AE, ABiGRUAE) 土石坝变形分区方法. 首先, 一方面于 BiGRUAE 中耦合注意力机制解决信息丢失问题, 另一方面基于混沌搜索增强算数优化算法 (arithmetic optimization algorithm, AOA) 跳出局部最优的能力, 对 ABiGRUAE 的超参数进行自适应寻优, 从而实现变形监测数据的有效降维. 其次, 基于曼哈顿距离的凝聚型层次聚类 (agglomerative hierarchical clustering, AHC) 算法实现土石坝变形分区. 最后, 通过超阈值 (peak over threshold, POT) 理论拟定各分区的变形安全诊断指标.

AHC-POT 模型应用于中国西南某土石坝工程的变形分区安全诊断指标拟定研究中, 对比了不同的优化算法、序列降维方法和聚类方法的性能, 验证了本文所提方法的有效性.

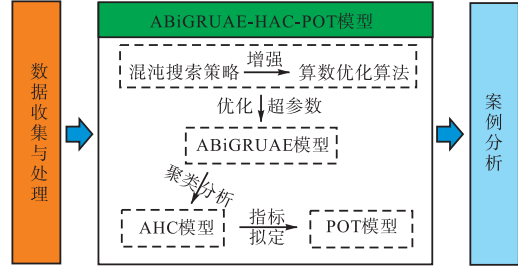


图1 研究框架

Fig.1 Research framework

### 1 研究框架

图1为本文所提模型的研究框架, 主要包括3个部分.

(1) 数据的收集与处理. 基于变形安全监测系统采集坝体表面变形监测数据, 并通过插值补全缺失数据.

(2) 建立土石变形分区安全诊断指标拟定 ABiGRUAE-AHC-POT 模型. 首先, 构建 ABiGRUAE 序列降维模型, 其中利用混沌搜索策略改进的 AOA 优化网络的超参数, 提高模型的重构精度. 其次, 通过基于曼哈顿距离的 AHC 对降维序列进行聚类分析, 获得土石坝变形分区结果. 最后, 基于 POT 模型拟定分区变形安全诊断指标.

(3) 案例分析. 将本研究提出的 ABiGRUAE-

### 2 ABiGRUAE-AHC-POT 土石坝变形分区指标拟定模型

#### 2.1 耦合注意力机制的 BiGRUAE 神经网络模型

双向门控循环单元由前向 GRU 和后向 GRU 组成, 如图2所示. 前向 GRU 从序列的开始到结束处理数据, 后向 GRU 则从序列的结束到开始处理数据. 两个 GRU 的隐藏状态在每个时间步上进行连接, 形成最终的输出, 具体为: 前向 GRU 从时间步  $t=1$  到  $t=T$  处理输入序列  $x$ , 得到最终的前向隐藏状态  $h^p$ ; 后向 GRU 从时间步  $t=T$  到  $t=1$  逐步处理输入序列  $x$ , 得到最终的后向隐藏状态  $h^n$ ; 最后将  $h^p$  和  $h^n$  拼接得到双向 GRU 的隐藏状态  $[h^p; h^n]$ .

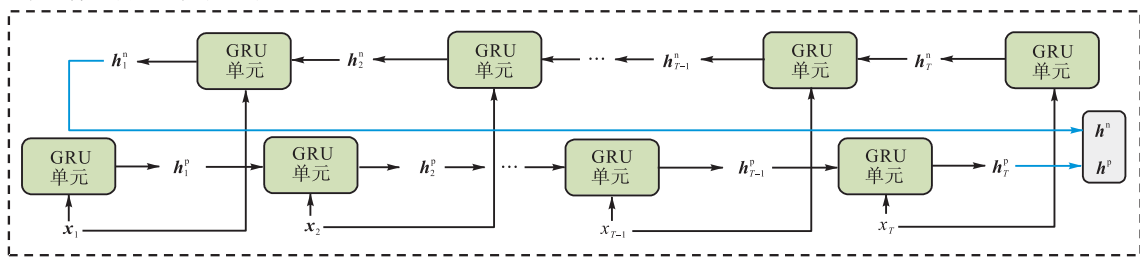


图2 BiGRU网络结构

Fig.2 Structure of BiGRU network

AE 是一种无监督的前馈神经网络结构<sup>[8-9]</sup>, 也是一种数据降维技术, 由编码器和解码器组成. 图3为 AE 的网络结构, 其由输入层、隐藏层和输出层组成. 编码器是一种神经网络模型, 它是将输入数据向量从高维低级映射到低维高级潜在空间向量  $z$ .  $z$  包含描述输入数据分布的主要信息, 其计算式为.

$$z = f_{enc}(W_{enc}x + b_{enc}) \quad (1)$$

式中:  $W_{enc}$  和  $b_{enc}$  分别为权重矩阵和偏置向量;  $f_{enc}$  为

一个激活函数, 它对线性映射  $W_{enc}x + b_{enc}$  进行非线性变换. 一般而言,  $f_{enc}$  是 Sigmoid、tanh、Relu 或线性函数. 解码器部分也是一个神经网络模型, 其通过将潜在向量  $z$  作为输入来重建输入数据向量的估计值, 可表示为

$$z_r = f_{dec}(W_{dec}z + b_{dec}) \quad (2)$$

式中:  $W_{dec}$  和  $b_{dec}$  分别为解码器的权重矩阵和偏差向量;  $z_r$  为重构向量;  $f_{dec}$  为一个激活函数, 一般为

Sigmoid、tanh 或 Relu 函数。

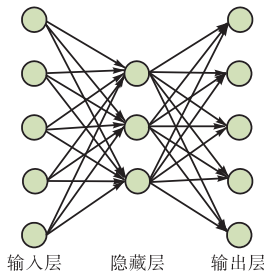


图 3 AE 网络结构

Fig.3 Structure of AE network

本文基于 BiGRU 和 AE 提出的 ABiGRUAE 模型如图 4 所示,在编码器中,  $X_t = [x_1, x_2, \dots, x_T]$  是作为输入的原始变形监测序列,通过双向 GRU 得到编码器输出的隐藏状态  $[h^p; h^n]$  (下文用  $h_E$  表示),该隐藏状态作为解码器中双向 GRU 的输入。  $h_E$  和编码器

的输出向量  $H_E$  是注意力模块的输入,通过计算得到各个时间步的权重向量  $A_{IN}$ 。为了表达重要时间步影响,编码器的输出向量  $H_E$  与  $A_{IN}$  乘法运算后得到注意力输出向量  $A_{OU}$ 。在解码器中  $h_E$  通过双向 GRU 得到各个时间步的隐藏状态,与  $A_{OU}$  向量拼接能够获得更加充分的序列特征信息,最后通过解码得到重构变形监测序列  $X'_t = [x'_1, x'_2, \dots, x'_T]$ 。

在模型的训练过程中,通过最小化输入数据和重构数据之间的误差或者准确的低维特征表示。将均方误差 (mean square error, MSE) 作为模型的损失函数,其表达式为

$$E_{MS} = \frac{1}{T} \sum_{t=1}^T \|X_t - X'_t\|_2^2 \quad (3)$$

式中  $\|\cdot\|_2^2$  为二平方范数。模型训练结束后,将原始序列降维后的向量  $h_E$  作为聚类算法的输入。

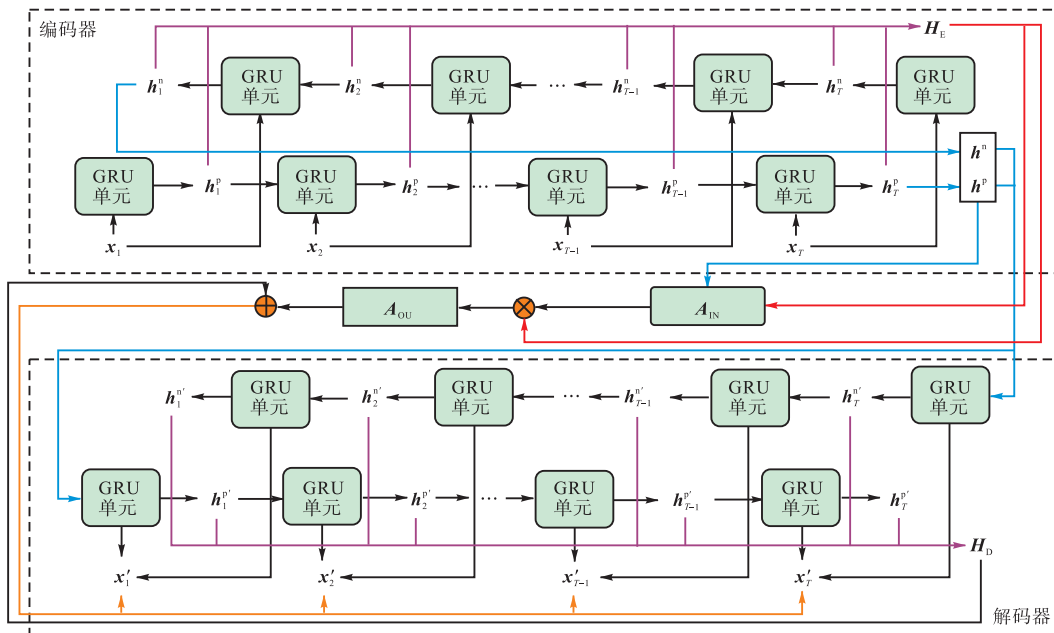


图 4 ABiGRUAE 模型结构

Fig.4 Structure of ABiGRUAE model

## 2.2 混沌搜索策略改进的 AOA

### 2.2.1 AOA

AOA 是 Abualigah 等<sup>[34]</sup>受到加减乘除四则运算符号的启发而提出的一种智能优化算法。该算法使用四则运算符作为数学优化,从一组候选解中确定符合特定标准的最佳解。算法分为 2 个搜索阶段,即探索阶段和开发阶段。

在 AOA 开始优化工作之前,首先应选择算法所处的搜索阶段。通过定义数学最优化加速函数 (S) 来确定第 k 次迭代时的算法处于的阶段,即

$$S^{(k)} = S_{min} + k(S_{max} - S_{min}) / I_{max} \quad (4)$$

式中:  $S^{(k)}$  为第 k 次迭代的函数值;  $S_{min}$  和  $S_{max}$  分别为加速器的提前设定的最小值和最大值;  $I_{max}$  为算法最大迭代次数; k 为当前的迭代次数,其范围为  $1 \sim I_{max}$ 。探索阶段是利用乘法和除法这 2 种运算法则进行高分散性地、大范围地全局搜索。开发阶段是利用加法和减法这 2 种运算法则进行低分散性地局部搜索。乘法和除法搜索策略见式 (5),加法和减法搜索策略见式 (6)<sup>[34]</sup>。如果  $d_1 > S^{(k)}$  ( $d_1$  为一个在区间  $[0, 1]$  上服从均匀分布的随机数),则执行探索阶段;反

之,执行开发阶段.

$$v_{i,j}^{(k+1)} = \begin{cases} v_j^* \div (W^{(k)} + \omega) \times [(v_j^U - v_j^L) \times \eta + v_j^U] & d_2 < 0.5 \\ v_j^* \times W^{(k)} \times [(v_j^U - v_j^L) \times \eta + v_j^L] & d_2 \geq 0.5 \end{cases} \quad (5)$$

$$v_{i,j}^{(k+1)} = \begin{cases} v_j^* - W^{(k)} \times [(v_j^U - v_j^L) \times \eta + v_j^L] & d_3 < 0.5 \\ v_j^* + W^{(k)} \times [(v_j^U - v_j^L) \times \eta + v_j^L] & d_3 \geq 0.5 \end{cases} \quad (6)$$

式中:  $d_2$  为另一个在区间 $[0, 1]$ 上服从均匀分布的随机数,当  $d_2 \geq 0.5$  时,执行乘法搜索策略,当  $d_2 < 0.5$  时,则执行除法搜索策略;  $v_{i,j}^{(k+1)}$  为第  $i$  个解在第  $k+1$  次迭代中的第  $j$  个维度的值;  $d_3$  也为一个在区间 $[0, 1]$ 上服从均匀分布的随机数;  $v_j^*$  为当前迭代次数下第  $j$  个维度的最优解;  $\omega$  为一个小整数;  $v_j^U$  和  $v_j^L$  分别为第  $j$  个维度的搜索上、下限;  $\eta$  为调整搜索过程的控制参数,其固定为 0.5;  $W^{(k)}$  为数学优化器概率<sup>[34]</sup>,其定义为

$$W^{(k)} = 1 - k^{1/\theta} / I_{\max}^{1/\theta} \quad (7)$$

式中  $\theta$  为控制开发精度的敏感参数,固定为 5.

### 2.2.2 混沌搜索策略改进 AOA

混沌现象的遍历性质也可以作为一种优化机制,以避免在搜索过程中局部收敛.对于可能涉及相当多局部最优的深度神经网络模型的训练, AOA 可能在解空间中不进一步搜索就终止.双混沌优化机制被集成到 AOA 的局部搜索策略中形成 IA OA,以提高最优化结果,其中两种完全不同的混沌最优化机制,即 logistic 映射和三次映射,用于独立搜索<sup>[35-37]</sup>.研究表明:在相同的初始值和相同的迭代次数下,两种映射的混沌变量值在大多数情况下是不同的,这有助于优化过程收敛到更好的解.双混沌优化机制使搜索过程同时在多个位置独立进行探索,显著提高了搜索空间遍历能力,有效避免了优化算法过早陷入局部最优解.这种机制在超参数优化过程中(尤其是深度学习模型训练时)能够更高效地寻找全局最优解,从而提高模型重构精度和泛化性能.

假设种群大小为  $p$ , 每个个体为  $n$  维向量,考虑迭代  $k$  处的一个最优个体  $v^{(k)}$  变换为两个混沌值  $v_{i,j}^{(k)}$  和  $v_{c,i}^{(k)}$ , 分别作为 logistic 映射和三次映射的输入<sup>[38]</sup>, 即

$$v_{i,j}^{(k)} = (v_i^{(k)} - l_i) / (u_i - l_i) \quad (8)$$

$$v_{c,i}^{(k)} = 2(v_i^{(k)} - l_i) / (u_i - l_i) - 1 \quad (9)$$

式中:  $u$  和  $l$  分别为搜索域的上限和下限; 索引  $i = 1, 2, \dots, n$ . logistic 映射和三次映射分别由式(10)和式

(11)定义<sup>[38]</sup>, 即

$$c_{i,i}^{(k+1)} = \varphi v_{i,i}^{(k)} (1 - v_{i,i}^{(k)}) \quad (10)$$

$$c_{c,i}^{(k+1)} = \sigma (v_{c,i}^{(k)})^3 + (1 - \sigma) v_{c,i}^{(k)} \quad (11)$$

当  $\varphi = 4$  且  $v_{i,i}^{(k)} \in (0, 1)$  时, logistic 映射被认为是混沌的; 当  $\sigma \in [3.3, 4.0]$  且  $v_{c,i}^{(k)} \in (-1, 1)$  时, 三次映射是混沌的.

### 2.3 基于曼哈顿距离的 AHC 聚类变形分区

AHC 是一种自底向上的聚类算法<sup>[39]</sup>, 其基本原理是从最底层(即单个数据对象)开始, 在全部数据合并到一个簇或满足预设的终止条件之前, 不断迭代合并满足相似性度量准则的簇, 以形成上一层中的簇, 如图 5 所示. 分裂型层次聚类是一种自顶向下的聚类算法, 其基本原理是先将所有数据看作一个簇, 然后根据相似性度量准则不断进行分裂, 生成的子节点再以相同的方式递归向下分裂, 直到只包含一个数据对象的子簇出现或达到指定条件.

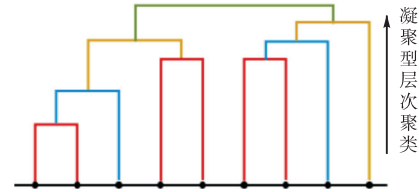


图 5 AHC 结构

Fig.5 Structure of AHC

聚类相似性度量方法可用于衡量数据点之间的相似性或差异性, 以便将相似的数据点分组到同一簇中. 欧氏距离是常用的方法, 其适用于需要衡量直线距离的场景, 尤其是当数据点的坐标是连续值时, 而曼哈顿距离更适用于衡量向量之间的相似性. 土石坝变形监测数据经第 2.1 节所提方法降维后成为多组低维向量, 故选用曼哈顿距离为相似度量方法.

曼哈顿距离用于计算两个点在标准坐标系上的绝对距离之和<sup>[40]</sup>. 在二维空间中, 这个距离相当于沿着网格路径走的距离. 令  $\mathbf{X} = [x_1, x_2, \dots, x_T]$  和  $\mathbf{Y} = [y_1, y_2, \dots, y_T]$  为两个需要度量相似性的向量, 则二者的曼哈顿距离为

$$M_{\text{DIS}}(\mathbf{X}, \mathbf{Y}) = \left( \sum_{i=1}^T |x_i - y_i| \right) / T \quad (12)$$

$$m_{\text{DIS}}(\mathbf{X}, \mathbf{Y}) = 1 / [1 + M_{\text{DIS}}(\mathbf{X}, \mathbf{Y})] \quad (13)$$

通过式(13)归一化后可将该距离转换为 $[0, 1]$ 区间内的某个值, 该值越接近 1, 代表二者的相似度越高.

### 2.4 基于 POT 模型的变形安全分区诊断指标拟定

#### 2.4.1 变形极值的定义

狭义上的极值定义是指随机变量中的极大值与

极小值,而广义上的极值定义是指在某一固定阈值控制下的极值集合.假设  $\{x_1, x_2, \dots, x_n\}$  为一组独立且同分布的顺序排列的变形监测数据.存在阈值  $\mu$ ,使得  $x_n \geq x_{n-1} \geq x_{n-2} \geq \dots \geq x_2 > \mu > x_1$ ,则变形极值序列为  $\{x_2, \dots, x_{n-1}, x_n\}$ .

2.4.2 POT 模型

广义帕累托分布 (generalized Pareto distribution, GPD) 是一种连续概率分布,其可描述数据中超过某一阈值的部分,用于 POT 极值建模<sup>[41]</sup>.其概率密度函数 (probability density function, PDF) 可表示为

$$g(x|\gamma, \sigma', u') = 1/\sigma' [1 + \gamma(x-u')/\sigma']^{-1/\gamma} \quad (14)$$

式中:  $\gamma, \sigma'$  和  $u'$  分别为形状参数、尺度参数和阈值,  $\gamma \neq 0, \sigma' > 0, 1 + \gamma(x-u')/\sigma' > 0$ . 当  $\gamma > 0$  时,  $g(x|\gamma, \sigma', u')$  为厚尾分布. 当  $\gamma < 0$ ,  $g(x|\gamma, \sigma', u')$  为薄尾分布. 当  $\gamma = 0$  时,  $g(x|\gamma, \sigma', u')$  为指数分布.

设  $\{x_1, x_2, \dots, x_n\}$  为一组独立且同分布的顺序排列的样本数据,存在阈值  $u'$ ,当  $x_i > u'$ ,  $x_i$  为超阈值,超过阈值的观测数据个数为  $n_i$ .  $y_i = x_i - u'$  的分布函数可表示为

$$F_{u'}(y) = P(x - u' \leq y | x > u') = \frac{P(u' \leq x \leq u' + y)}{P(x > u')} = \frac{F(u' + y) - F(u')}{1 - F(u')} \quad y \geq 0 \quad (15)$$

当阈值  $u'$  足够大时,可用  $(1 - n_{u'}/n)$  近似估计  $F(u')$ ,其中  $n_{u'}$  为超阈值的样本数量,可得  $y_i = x_i - u'$  的总体分布函数,其可表示为

$$F(x) = \begin{cases} 1 - n/n_{u'} [1 + \gamma/\sigma' (x - u')]^{-1/\gamma} & \gamma \neq 0 \\ 1 - n/n_{u'} [1 - e^{-(x-u')/\sigma'}] & \gamma = 0 \end{cases} \quad (16)$$

在确定失效概率的基础上,经过总体分布函数的反函数可得安全诊断指标,即

$$E_m = \begin{cases} u' + \gamma/\sigma' (P_\alpha n/n_{u'} - 1) & \gamma \neq 0 \\ u' - \sigma' \ln(P_\alpha n/n_{u'}) & \gamma = 0 \end{cases} \quad (17)$$

式中  $P_\alpha$  为失效概率.

2.4.3 阈值的确定

不同阈值的选择会显著影响诊断指标的敏感性,阈值过高可能忽略部分关键风险,而阈值过低则可能引入不必要的报警.采用 Hill 图法<sup>[42]</sup>通过统计量的变化趋势直观地确定合理的阈值,以确保样本数据能合理符合 GPD.考虑到变形监测序列  $\{x_1, x_2, \dots, x_n\}$  为独立而且同分布的随机变量,其顺序统计量为  $x_{(n)} \geq x_{(n-1)} \geq x_{(n-2)} \geq \dots \geq x_{(1)}$ .以式(18)计算 Hill 统计量,即

$$H_{k,n} = \frac{1}{k} \sum_{j=1}^k j (\ln x_{(n-j+1)} - \ln x_{(n-j)}) \quad (18)$$

式中  $1 \leq k \leq n-1$ .以  $k$  为横轴、 $H_{k,n}$  为纵轴的曲线图为 Hill 图,图形处于稳定的初始点对应的样本值  $x_{(j)}$  为阈值  $u'$ .

2.5 所提指标拟定方法流程

本文所提模型的实现如图 6 所示<sup>[38]</sup>,具体包括 5 个步骤.

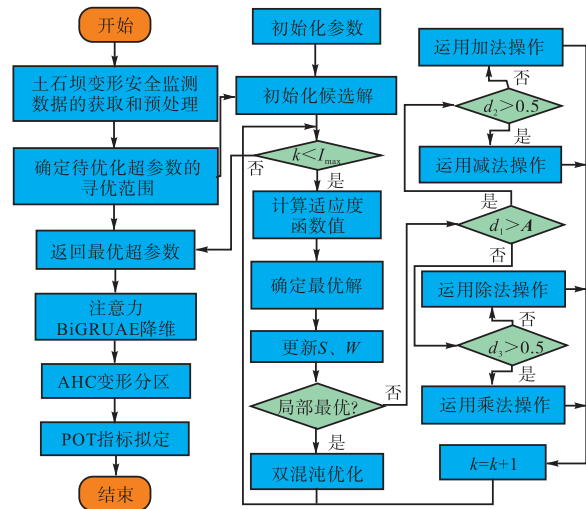


图 6 所提模型建立流程

Fig.6 Flow chart of establishment of proposed model

步骤 1 首先将获取的变形监测数据进行缺失值填补、归一化处理.

步骤 2 通过所提出的 IAOA 算法对 ABiGRUAE 模型的学习率、正则化系数  $L_2$ 、隐藏层节点数和最大训练周期这 4 项超参数进行寻优.

步骤 3 基于寻优所得到的最优超参数建立 ABiGRUAE 模型,而后对变形监测序列数据进行特征提取降维.

步骤 4 通过曼哈顿距离相似性度量的 AHC 算法对降维后的监测序列特征聚类分析,获得土石坝变形分区结果.

步骤 5 基于 POT 模型拟定各变形分区安全诊断指标.

3 案例

3.1 工程概况

某高心墙堆石坝位于云南省澜沧江下游河段,为 I 等大(1)型工程,最大坝高 261.5 m.在堆石坝施工期间布置了丰富的安全监测仪器,具有完备的变形安全监测系统.为验证所提方法的有效性,本文选取坝

体下游侧 2015-01-11—2018-11-10 的表观变形监测数据进行研究. 图 7 为监测点布置, 部分变形监测数据见图 8<sup>[38]</sup>.

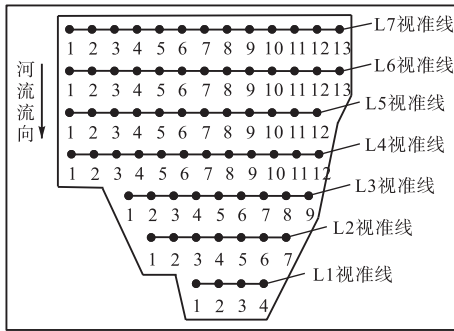


图 7 下游侧表观变形监测点布置

Fig.7 Layout of apparent deformation monitoring points on downstream side

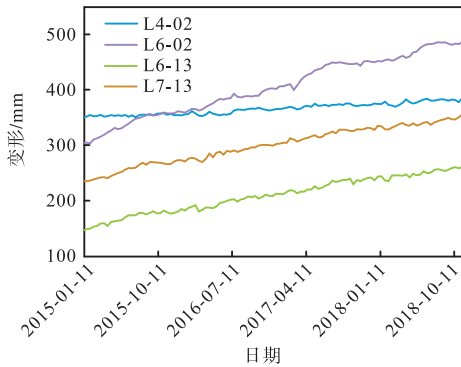


图 8 部分变形监测点的监测数据

Fig.8 Monitoring data at some deformation monitoring points

### 3.2 数据预处理

由于工程所在地区的自然环境条件恶劣、设备易故障损坏以及人为失误等因素, 时常会发生数据缺失和数据异常的情况. 通过数据预处理(数据插值和数据归一化)可提高数据的质量和可靠性, 优化特征选择和提取, 消除特征之间的差异, 为数据分析和建模提供良好的基础.

数据插值采用三次 Hermite 插值, 该方法较好地逼近原始数据点, 在相邻数据点之间的曲线是光滑连续的, 而且对数据点之间的分布没有特殊要求, 适用于各种不规则的数据分布情况. 通过该方法可以获得平稳且连续变化的土石坝变形安全监测数据, 案例工程原始监测数据的缺失比例为 2.57%, 数据插值结果如图 9 所示.

数据归一化采用 Min-Max 标准化, 可将数据缩放到一个固定的范围并且使得数据的取值范围一致, 具体见式(19), 即

$$M' = (M - \min M) / (\max M - \min M) \quad (19)$$

式中:  $M$  为实际监测数据;  $\min M$  为实际监测数据中的最小值;  $\max M$  为实际监测数据中的最大值;  $M'$  为归一化后的数据.

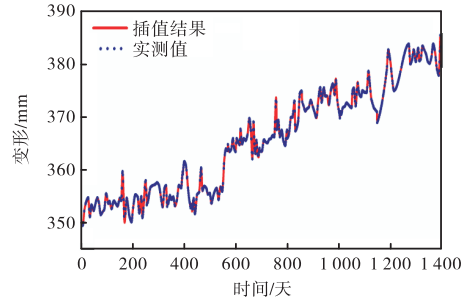


图 9 监测点 L4-02 数据插值结果

Fig.9 Interpolation results at monitoring point L4-02

### 3.3 变形分区聚类模型的建立

将所选取的图 7 中 70 个测点的变形监测数据经过插值和归一化预处理后, 得到 70 条步长为 1400 步的序列作为聚类分析数据. ABiGRUAE 降维模型基于深度学习框架 Pytorch 3.7 构建, 模型损失函数为 MSE 函数, 优化器为 Adam. 为提高降维模型的重构精度, 基于 IAOA 寻优 ABiGRUAE 模型的 4 个超参数, 即隐藏层节点数量  $N_u$ 、学习率  $L_r$ 、正则化系数  $L_2$  和最大训练周期  $M_t$ . 这 4 个参数的优化范围分别为:  $N_u \in [10, 100]$ ,  $L_r \in [1 \times 10^{-3}, 1 \times 10^{-1}]$ ,  $L_2 \in [1 \times 10^{-4}, 1 \times 10^{-1}]$ ,  $M_t \in [50, 200]$ . IAOA 的种群数量为 20, 最大迭代次数为 100. 最优参数结果如表 1 所示.

表 1 ABiGRUAE 模型的最优超参数

Tab.1 Optimal hyperparameters of ABiGRUAE model

超参数	数值
$N_u$	64
$L_r$	0.01
$L_2$	0.042 7
$M_t$	150

为了比较不同方法的聚类分区性能表现. 本文使用轮廓系数 ( $A_s$ ) 和戴维斯-鲍丁指数 (Davies-Bouldin index, DBI) 作为评价指标.

轮廓系数评价指标的定义为

$$s(x_i) = \frac{b(x_i) - a(x_i)}{\max[a(x_i), b(x_i)]} \quad (20)$$

式中:  $a(x_i)$  为数据点  $x_i$  的簇内距离;  $b(x_i)$  为数据点  $x_i$  的簇间距离. 整体聚类结果的轮廓系数可以通过所有数据点的轮廓系数评价指标的平均值来衡量, 见式(21). 平均轮廓系数越接近 1, 聚类效果就越好.

$$A_s(x_i) = \frac{1}{N} \sum_{i=1}^N s(x_i) \quad (21)$$

戴维斯-鲍丁指数的计算式为

$$I_{DB} = \frac{1}{M} \sum_{i=1}^M \max_{i \neq j} \frac{k_i + k_j}{d_{ij}} \quad (22)$$

式中:  $k_i$  为簇  $i$  中每个数据点到簇中心的平均距离;  $d_{ij}$  为簇  $i$  和簇  $j$  中心间距离.  $I_{DB}$  越接近 0, 聚类效果就越好.

### 3.4 变形分区指标拟定结果

#### 3.4.1 变形分区聚类结果

通过比较不同聚类簇数量下的聚类平均距离(各类的样本与类中心的平均距离之和)的下降程度,同时参考 DBI, 可以确定最佳的聚类簇数量, 见图 10.

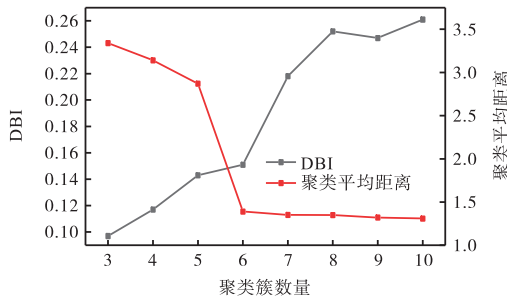
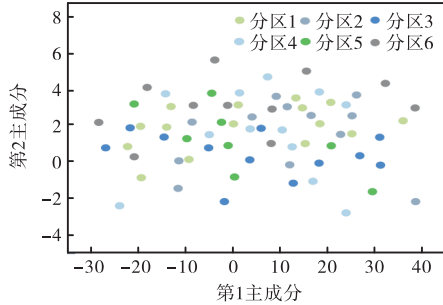


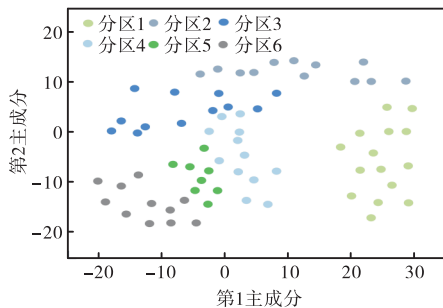
图 10 聚类簇数量与 DBI 及聚类平均距离的关系

Fig.10 Relationship among the number of clusters, DBI and clustering average distance

在图 10 中, 当聚类簇数量为 6 时, 聚类平均距离下降程度最大, 并且在附近戴维斯-鲍丁指数较



(a) 10 个训练周期



(c) 120 个训练周期

小, 即 0.151. 因此, 确定聚类簇数量为 6, 即变形分区数量为 6. 变形分区结果如图 11 所示, 可以看出土石坝变形整体呈现左右岸对称分布, 每一个分区是聚类分析产生的一个簇: 分区 1 包含 15 个测点; 分区 2、4 和 5 都包含 12 个测点; 分区 3 包含 11 个测点; 分区 6 包含 8 个测点.

图 12 对不同训练周期下的聚类情况进行了二维化展示, 可以看出随着训练特征相似的测点序列被逐渐地聚类为一簇, 在达到最大训练周期时, 即所有全部训练集数据对模型完整训练 150 次后, 聚类效果达到了最佳.

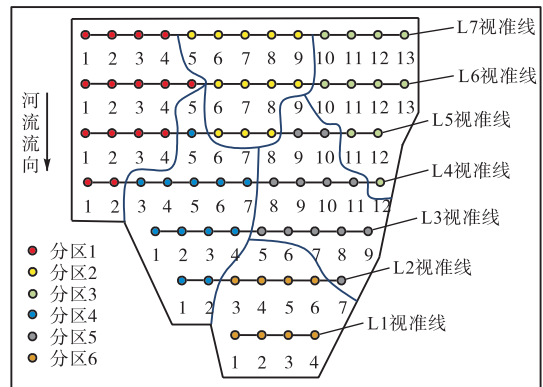
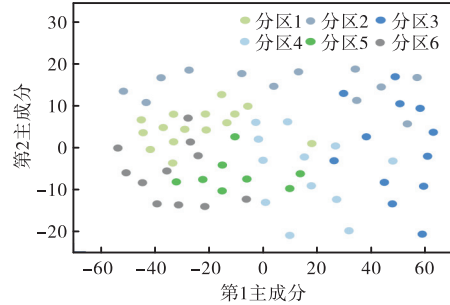
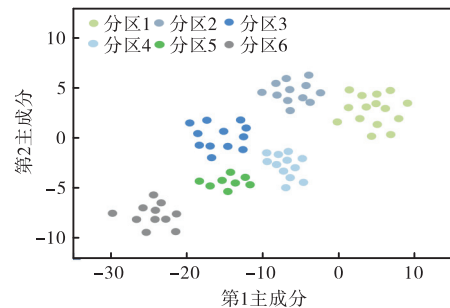


图 11 变形分区结果

Fig.11 Zoning results of deformation



(b) 60 个训练周期



(d) 150 个训练周期

图 12 不同训练周期下的聚类可视化结果

Fig.12 Results of clustering visualization under different training cycles

#### 3.4.2 分区指标拟定结果

首先, 通过 Hills 图对 6 个变形分区各自的样本

数据阈值进行选取; 然后, 利用矩法估计 POT 序列样本服从的 GPD 的参数; 最后, 基于工程失事概率确

定分区安全诊断指标. 表 2 为各分区的超阈值模型计算参数.

表 2 各分区 POT 模型计算参数

Tab.2 Calculation parameters of POT model for each zone

分区编号	$n$	$n_u'$	$u'$	$\gamma$	$\sigma'$
1	200	35	1 335.0	8.215	-0.281
2	200	26	1 200.0	4.239	-0.135
3	200	23	974.0	6.790	-0.164
4	200	61	786.9	5.989	-1.095
5	200	44	559.0	41.848	-1.065
6	200	39	222.8	0.956	-0.349

参考相关工程经验<sup>[10]</sup>, 通过式 (17) 分别计算显著性水平 1% 和 10% 对应的安全诊断指标值. 各分区在不同显著性水平下的指标见表 3, 可见各分区在同一显著性水平下的诊断指标值差异较大, 相较于不考虑空间分异和类聚性的传统方法得到的整体性指标, 分区指标明显更具合理性.

表 3 各分区的安全诊断指标

Tab.3 Safety diagnostic indexes for each zone

分区编号	显著性水平 1%	显著性水平 10%
1	1 362.566 0	1 347.530 0
2	1 229.080 0	1 207.270 0
3	1 011.782 0	979.397 5
4	792.190 5	790.576 4
5	596.506 6	580.432 3
6	225.397 4	224.133 8

### 3.5 对比分析和讨论

#### 3.5.1 超参数优化效果对比

为了验证本文所提 IAOA 在寻优性能上的优越性, 将其与遗传算法 (genetic algorithm, GA)、粒子群优化 (particle swarm optimization, PSO) 算法、鲸鱼优化算法 (whale optimization algorithm, WOA) 和 AOA 的适应度值收敛过程以及测点序列的重构误差进行对比, 结果如图 13 和图 14 所示.

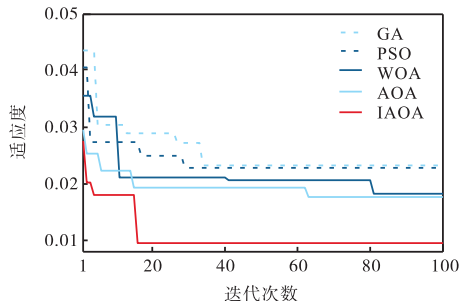


图 13 不同优化算法的适应度曲线

Fig.13 Fitness curves of different optimization algorithms

各优化算法的种群数量均设置为 20, 最大迭代次数为 100. 从图 13 可以看出, IAOA 在 20 次迭代之前就收敛至最优适应度值, 且该适应度值优于其他

方法. AOA 与 WOA 寻优性能接近, 前者略优于后者. GA 和 PSO 则是在这个案例中寻优性能最差的 2 种算法. 从图 14 展示的重构误差结果来看, IAOA 的重构误差为 0.009 53, 是所有方法中最小的; 其次是 AOA 和 WOA, 二者的重构误差分别为 0.012 68 和 0.015 24; GA 和 PSO 最差, 其重构误差分别为 0.020 25 和 0.019 81. 以上对比分析说明本文所提的 IAOA 具有较好的寻优能力.

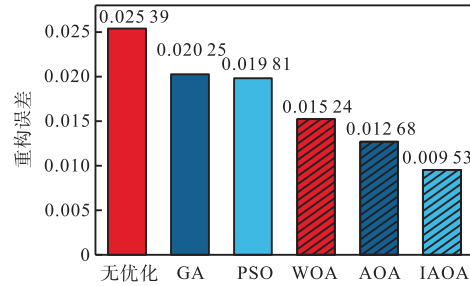


图 14 不同优化算法的序列重构误差

Fig.14 Sequence reconstruction errors of different optimization algorithms

#### 3.5.2 降维方法性能对比

为了验证所提 ABiGRUAE 变形监测序列降维模型的优越性, 本文选择门控循环单元自编码器 (gated recurrent unit autoencoder, GRUAE)、长短期记忆自编码器 (long short-term memory autoencoder, LSTM AE)、双向 GRUAE (bi-directional GRUAE, BiGRUAE)、双向 LSTM AE (bi-directional LSTM AE, BiLSTM AE)、注意力 BiLSTM AE (attention-based BiLSTM AE, ABiLSTM AE) 与本文模型进行对比分析.

由图 15 可见, 相较于其他对比方法, 所提方法的序列重构误差最小, 即 0.009 53, 说明本文方法能够精确重构变形监测序列, 其网络编码器输出的上、下文语义向量能够更全面地包含原始数据的特征. 图 16 所示为各方法在降维性能上的相对提升比率, 可见 GRU 相比 LSTM 的性能优势以及双向结构和注意力机制对网络特征提取能力的提升是十分明显的.

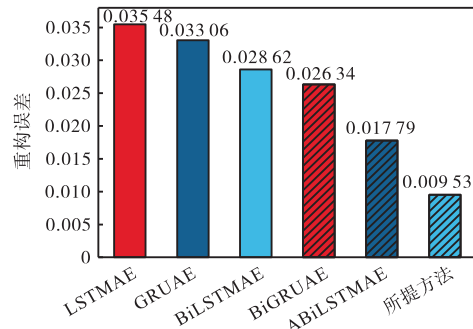


图 15 不同降维方法的序列重构误差

Fig.15 Sequence reconstruction errors of different dimensionality reduction methods

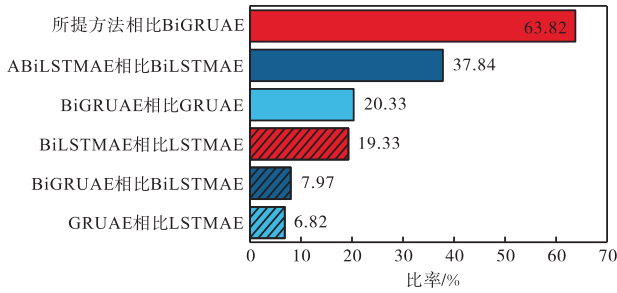


图 16 不同降维方法的性能相对提升

Fig.16 Relative performance enhancement of different dimensionality reduction methods

表 4 为基于各降维方法的 AHC 评价指标,可以看出本文方法的  $A_s$  和  $I_{DB}$  是最优的,分别为 0.886 和 0.151,说明相比于其他方法本文所提的降维方法能够实现序列的有效降维,为聚类提供良好基础。

表 4 不同降维方法的聚类评价指标

Tab.4 Clustering evaluation indexes for different dimensionality reduction methods

降维方法	$A_s$	$I_{DB}$
LSTMAE+AHC	0.701	0.336
GRUAE+AHC	0.749	0.317
BiLSTMAE+AHC	0.769	0.284
BiGRUAE+AHC	0.781	0.235
ABiLSTMAE+AHC	0.854	0.189
本文方法	0.886	0.151

### 3.5.3 聚类相似性度量方法性能对比

为验证所提基于曼哈顿距离的 AHC 聚类方法在聚类性能上的优越性,本文选取基于欧氏距离和余弦相似度的 AHC 聚类方法与所提方法进行对比。评价指标选取  $A_s$  和  $I_{DB}$ ,各方法的评价指标值见表 5。可以看出所提方法(基于曼哈顿距离)的  $A_s$  为 0.886,  $I_{DB}$  为 0.151,均优于对比方法,表明所提方法对于本研究中的序列特征具有更佳的区别能力。

表 5 不同相似性度量方法的评价指标

Tab.5 Evaluation indexes for different similarity measurement methods

距离度量方法	$A_s$	$I_{DB}$
基于欧氏距离	0.743	0.299
基于余弦相似度	0.769	0.287
基于曼哈顿距离(所提方法)	0.886	0.151

### 3.5.4 聚类分区方法对比

为了验证所提方法的在聚类上的有效性,本文选取 K-means、PCA+K-means、AHC 和 PCA+AHC 与所提方法进行对比分析。评价指标选取  $A_s$  和  $I_{DB}$ ,各方法的评价指标值展示于表 6。首先,可以看出 AHC 相较于 K-means 各指标表现都更好,其中  $A_s$  提升了 0.106,  $I_{DB}$  降低了 0.054,说明 AHC 聚类效果更

优。其次,分别对比 K-means 和 PCA+K-means 以及 AHC 和 PCA+AHC,可以看出降维后聚类的效果有明显提升。最后,对比本文方法和 PCA+AHC,发现  $A_s$  提升了 0.245,  $I_{DB}$  降低了 0.427,说明本文提出的混沌算数优化的 BiGRUAE 方法能够在降维后尽可能地保留序列数据的原始特征,从而获得最好的聚类结果。

表 6 不同聚类方法的评价指标

Tab.6 Evaluation indexes for different clustering methods

聚类方法	$A_s$	$I_{DB}$
K-means	0.497	0.732
PCA+ K-means	0.578	0.598
AHC	0.603	0.678
PCA+ AHC	0.641	0.578
所提方法	0.886	0.151

## 4 结 论

合理的指标拟定对大坝的风险控制和安全稳定具有重要的意义。本文考虑了变形监测点在空间中的分异性和类聚性,提出了混沌算数优化的 BiGRUAE-AHC 土石坝变形分区方法,基于 POT 理论实现了土石坝空间分区安全诊断指标拟定,提高了所拟定指标的合理性,得到了如下结论。

(1) 针对序列降维中存在的信息丢失问题,提出了 IAOA-ABiGRUAE 方法,采用混沌搜索策略优化 ABiGRUAE 网络超参数,从而提高了变形监测序列的重构精度,实现了序列的有效降维。

(2) 相较于仅拟定单一指标表征整体安全状态,本文充分考虑测点在空间的分异性和类聚性,在变形监测序列有效降维的基础上,采用 AHC 和 POT 实现了土石坝变形安全分区指标拟定。

(3) 将所提方法应用于中国西南某高心墙堆石坝工程实例,实现了变形安全诊断分区指标拟定。研究结果表明,所提出模型的序列重构误差仅为 0.00953,聚类性能评价指标  $A_s$  为 0.886,  $I_{DB}$  为 0.151,远优于其他对比方法,证明了所提模型的优越性,为大坝安全状态诊断提供了新的思路。

未来研究中应做进一步探索:一方面,应充分考虑不同类型土石坝的材料特性、结构型式及地质环境对模型性能的影响,以进一步增强模型的泛化能力与适用范围;另一方面,应结合结构分析理论与方法对所提出的数据驱动型指标进行更深入的物理机理验证与定性解释,以更清晰地揭示监测数据与实际坝体变形及力学响应之间的内在关联。

## 参考文献:

- [1] 姜振翔, 陈 辉, 陈柏全. 基于云理论的大坝整体性态评价模型[J]. 工程科学学报, 2022, 44(3): 464-473.  
Jiang Zhenxiang, Chen Hui, Chen Baiquan. Evaluation model of overall dam behavior based on cloud theory[J]. Chinese Journal of Engineering, 2022, 44(3): 464-473(in Chinese).
- [2] 李宗坤, 姜景山, 王广印. 土石坝实测性态综合评价方法研究[J]. 岩土工程学报, 2007, 29(2): 255-259.  
Li Zongkun, Jiang Jingshan, Wang Guangyin. Comprehensive assessment of observed behavior of earth-rockfill dams[J]. Chinese Journal of Geotechnical Engineering, 2007, 29(2): 255-259(in Chinese).
- [3] 李家田, 苏怀智. 重力坝失稳破坏变形表征指标警戒值的拟定[J]. 水利水运工程学报, 2023(1): 34-42.  
Li Jiatian, Su Huaizhi. Anti-sliding stability reliability-based approach determining the early-warning value of deformation safety for gravity dam[J]. Hydro-Science and Engineering, 2023(1): 34-42(in Chinese).
- [4] 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. Safety Monitoring of Dams and Dam Foundations—Theories & Methods and Their Application[M]. Nanjing: Hohai University Press, 2006(in Chinese).
- [5] 何金平, 施玉群, 吴雯娴. 大坝安全监测系统综合评价指标体系研究[J]. 水力发电学报, 2011, 30(4): 175-180.  
He Jinping, Shi Yuqun, Wu Wenxian. Index system of comprehensive evaluation on the dam safety monitoring system[J]. Journal of Hydroelectric Engineering, 2011, 30(4): 175-180(in Chinese).
- [6] 马 睿, 尹 韬, 李浩欣, 等. 大坝机理-数据融合模型的基本结构与特征[J]. 水力发电学报, 2022, 41(5): 59-74.  
Ma Rui, Yin Tao, Li Haoxin, et al. Basic structure and characteristics of dam mechanism-data-driven fusion models[J]. Journal of Hydroelectric Engineering, 2022, 41(5): 59-74(in Chinese).
- [7] 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]. 水利水电科技进步, 2021, 41(1): 30-34, 48.  
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, 48(in Chinese).
- [8] 余正源, 黄耀英, 高大水, 等. 基于典型小概率法的土石坝倾度安全监控指标拟定[J]. 水电能源科学, 2022, 40(11): 103-106, 54.  
Yu Zhengyuan, Huang Yaoying, Gao Dashui, et al. Typical low probability method based determination of inclination security monitoring index of earth-rock dam[J]. Water Resources and Power, 2022, 40(11): 103-106, 54(in Chinese).
- [9] 周稳忠, 谷艳昌, 黄海兵, 等. 基于改进典型小概率法的大坝变形安全监控指标拟定[J]. 人民珠江, 2020, 41(6): 39-43.  
Zhou Wenzhong, Gu Yanchang, Huang Haibing, et al. Determination of dam deformation safety monitoring index based on improved typical small probability method[J]. Pearl River, 2020, 41(6): 39-43(in Chinese).
- [10] 肖 磊, 万智勇, 黄耀英, 等. 基于最大熵和云模型的RCC坝变形监控指标拟定[J]. 水利水运工程学报, 2018(4): 24-29.  
Xiao Lei, Wan Zhiyong, Huang Yaoying, et al. RCC dam deformation monitoring index based on maximum entropy and cloud model[J]. Hydro-Science and Engineering, 2018(4): 24-29(in Chinese).
- [11] 袁友翠, 胡少华, 吴 浩, 等. 基于CM-GCT的尾矿坝在线监测位移分级预警阈值研究[J]. 武汉理工大学学报(信息与管理工程版), 2022, 44(5): 758-765.  
Yuan Youcui, Hu Shaohua, Wu Hao, et al. Research on graded early-warning threshold of tailings dam online monitoring displacement based on CM-GCT[J]. Journal of Wuhan University of Technology(Information & Management Engineering), 2022, 44(5): 758-765(in Chinese).
- [12] 汪俊波, 李 黎, 程 帅, 等. 功果桥大坝坝顶水平位移安全监控指标拟定[J]. 水利与建筑工程学报, 2020, 18(1): 256-260.  
Wang Junbo, Li Li, Cheng Shuai, et al. Safety monitoring index determination of horizontal displacement for Gongguoqiao Dam crest[J]. Journal of Water Resources and Architectural Engineering, 2020, 18(1): 256-260(in Chinese).
- [13] 苏 龙, 孙建生, 王 刚. 基于分项系数极限状态法与突变理论的重力坝深层抗滑稳定分析[J]. 水电能源科学, 2015, 33(12): 78-80, 100.  
Su Long, Sun Jiansheng, Wang Gang. Deep anti-sliding stability analysis of gravity dam based on partial coefficient limit state method and catastrophe theory[J]. Water Resources and Power, 2015, 33(12): 78-80,

- 100 (in Chinese).
- [14] 徐小蓉, 何涛洪, 雷峥琦, 等. 超长坝段堆石混凝土重力坝蓄水运行安全评价[J]. 清华大学学报(自然科学版), 2022, 62(9): 1375-1387.
- Xu Xiaorong, He Taohong, Lei Zhengqi, et al. Safety evaluation of rock-filled concrete gravity dam with long section during the impoundment operation period[J]. Journal of Tsinghua University (Science and Technology), 2022, 62(9): 1375-1387 (in Chinese).
- [15] 吴相豪, 刘俊汝. 基于安全系数法拟定面板堆石坝变形监控指标[J]. 中国安全科学学报, 2020, 30(2): 127-132.
- Wu Xianghao, Liu Junru. Monitoring index for deformation of face rockfill dam based on safety factor method[J]. China Safety Science Journal, 2020, 30(2): 127-132 (in Chinese).
- [16] 刘西, 康飞, 李俊杰. 基于智能参数反演的混凝土坝位移监控指标拟定[J]. 水电能源科学, 2022, 40(7): 111-114, 97.
- Liu Xi, Kang Fei, Li Junjie. Determination of concrete dam displacement monitoring index based on intelligent parameter inverse[J]. Water Resources and Power, 2022, 40(7): 111-114, 97 (in Chinese).
- [17] 徐小枫, 黄耀英, 朱赵辉, 等. 混凝土大坝变形监控指标拟定的混合法[J]. 中国水利水电科学研究院学报(中英文), 2022, 20(4): 334-342.
- Xu Xiaofeng, Huang Yaoying, Zhu Zhaohui, et al. Hybrid method for deformation monitoring index of concrete dam[J]. Journal of China Institute of Water Resources and Hydropower Research, 2022, 20(4): 334-342 (in Chinese).
- [18] 黄海兵, 吴云星, 谷艳昌. 应用聚类算法的大坝安全监控方法[J]. 水利规划与设计, 2021(8): 122-126.
- Huang Haibing, Wu Yunxing, Gu Yanchang. Overview of application research on dam safety monitoring based on clustering algorithm[J]. Water Resources Planning and Design, 2021(8): 122-126 (in Chinese).
- [19] Jiedeerbieke M, Li T C, Chao Y, et al. Gravity dam deformation prediction model based on I-KShape and ZOA-BiLSTM[J]. IEEE Access, 2024, 12: 50710-50722.
- [20] Lin C, Zou Y, Lai X H, et al. Variation trend prediction of dam displacement in the short-term using a hybrid model based on clustering methods[J]. Applied Sciences, 2023, 13(19): 10827.
- [21] Chen W L, Wang X L, Cai Z J, et al. DP-GMM clustering-based ensemble learning prediction methodology for dam deformation considering spatiotemporal differentiation[J]. Knowledge-Based Systems, 2021, 222: 106964.
- [22] 张 晓. 基于聚类集成的大坝异常数据检测研究[D]. 郑州: 郑州大学软件学院, 2021.
- Zhang Xiao. Research on Dam Abnormal Data Detection Based on Clustering Ensemble[D]. Zhengzhou: School of Software, Zhengzhou University, 2021 (in Chinese).
- [23] 胡 江, 王春红, 李 星. 改进的特高拱坝分区变形预测模型[J]. 水力发电学报, 2023, 42(7): 69-83.
- Hu Jiang, Wang Chunhong, Li Xing. Improved zonal deformation prediction model for super-high arch dams [J]. Journal of Hydroelectric Engineering, 2023, 42(7): 69-83 (in Chinese).
- [24] 刘伟琪, 陈 波, 葛盼猛, 等. 基于聚类分区和 MO-LSSVR 的高拱坝变形预测模型[J]. 水利水电科技进展, 2023, 43(2): 102-108.
- Liu Weiqi, Chen Bo, Ge Panmeng, et al. Deformation prediction model of a high arch dam based on clustering and MO-LSSVR[J]. Advances in Science and Technology of Water Resources, 2023, 43(2): 102-108 (in Chinese).
- [25] 胡添翼. 基于面板数据分析方法的混凝土拱坝变形数据时空聚类模型[J]. 长江科学院院报, 2021, 38(2): 39-45.
- Hu Tianyi. Spatial and temporal clustering model of concrete arch dam deformation data based on panel data analysis method[J]. Journal of Yangtze River Scientific Research Institute, 2021, 38(2): 39-45 (in Chinese).
- [26] 崔欣然, 石 立, 陆 希, 等. 基于面板数据模型的拱坝缺失数据填补方法[J]. 水力发电学报, 2024, 43(3): 94-107.
- Cui Xinran, Shi Li, Lu Xi, et al. Panel data model-based method to fill in missing data for arch dams[J]. Journal of Hydroelectric Engineering, 2024, 43(3): 94-107 (in Chinese).
- [27] Wei X X, Zhang Z H, Huang H J, et al. An overview on deep clustering[J]. Neurocomputing, 2024, 590: 127761.
- [28] Cortés D G, Onieva E, López I P, et al. Autoencoder-enhanced clustering: A dimensionality reduction approach to financial time series[J]. IEEE Access, 2024, 12: 16999-17009.
- [29] 董靖川, 武晓鑫, 高宇博, 等. 融合注意力机制的刀具磨损预测方法[J]. 天津大学学报(自然科学与工程技术版), 2024, 57(4): 362-373.
- Dong Jingchuan, Wu Xiaoxin, Gao Yubo, et al. Tool wear prediction method based on attention mechanism[J]. Journal of Tianjin University (Science and

- Technology), 2024, 57(4): 362-373 (in Chinese).
- [30] Yuan D Y, Gu C S, Wei B W, et al. Displacement behavior interpretation and prediction model of concrete gravity dams located in cold area[J]. Structural Health Monitoring, 2023, 22(4): 2384-2401.
- [31] Chen C B, Tang L Q, Lu Y H, et al. Reconstruction of long-term strain data for structural health monitoring with a hybrid deep-learning and autoregressive model considering thermal effects[J]. Engineering Structures, 2023, 285: 116063.
- [32] Zhou G Y, Hu G F, Zhang D X, et al. A novel algorithm system for wind power prediction based on RANSAC data screening and Seq2Seq-Attention-BiGRU model[J]. Energy, 2023, 283: 128986.
- [33] Zheng Z, Zhang Z J. A temporal convolutional recurrent autoencoder based framework for compressing time series data[J]. Applied Soft Computing, 2023, 147: 110797.
- [34] Abualigah L, Diabat A, Mirjalili S, et al. The arithmetic optimization algorithm[J]. Computer Methods in Applied Mechanics and Engineering, 2021, 376: 113609.
- [35] Feng J H, Zhang J, Zhu X S, et al. A novel chaos optimization algorithm[J]. Multimedia Tools and Applications, 2017, 76: 17405-17436.
- [36] Luo Y Q, Yu J, Lai W R, et al. A novel chaotic image encryption algorithm based on improved baker map and logistic map[J]. Multimedia Tools and Applications, 2019, 78: 22023-22043.
- [37] 王晓玲, 梁羽翎, 王佳俊, 等. 耦合注意力机制大坝变形改进 LSTM 序列到序列预测模型[J]. 天津大学学报(自然科学与工程技术版), 2023, 56(7): 702-712.  
Wang Xiaoling, Liang Yuling, Wang Jiajun, et al. Improved LSTM sequence-to-sequence prediction model for dam deformation coupled with attention mechanism[J]. Journal of Tianjin University(Science and Technology), 2023, 56(7): 702-712 (in Chinese).
- [38] Wang L, Wang J J, Tong D W, et al. A novel long short-term memory Seq2Seq model with chaos-based optimization and attention mechanism for enhanced dam deformation prediction[J]. Buildings, 2024, 14(11): 3675.
- [39] 章玉, 张婷婷, 姚成北, 等. 考虑交通状态影响的服务区凝聚层次聚类入区判别模型[J]. 重庆交通大学学报(自然科学版), 2024, 43(3): 84-91.  
Zhang Yu, Zhang Tingting, Yao Chengbei, et al. An identification model of entering the service area of agglomerative hierarchical clustering considering the influence of traffic status[J]. Journal of Chongqing Jiaotong University(Natural Science), 2024, 43(3): 84-91 (in Chinese).
- [40] Wang C D, Yang J L, Zhang B Q. A fault diagnosis method using improved prototypical network and weighting similarity-Manhattan distance with insufficient noisy data[J]. Measurement, 2024, 226: 114171.
- [41] 张耀, 黄耀英, 何一洋, 等. 基于 POT 模型的泄水闸检修门库裂缝开合度安全监控指标拟定[J]. 三峡大学学报(自然科学版), 2023, 45(1): 1-5.  
Zhang Yao, Huang Yaoying, He Yiyang, et al. Formulation of safety monitoring index of crack opening and closing degree of sluice maintenance door library based on POT model[J]. Journal of China Three Gorges University(Natural Sciences), 2023, 45(1): 1-5 (in Chinese).
- [42] 周志杰, 沈振中, 徐力群, 等. 改进 POT 模型及其在边坡安全监测预警中的应用[J]. 南水北调与水利科技, 2016, 14(4): 192-197.  
Zhou Zhijie, Shen Zhenzhong, Xu Liqun, et al. Modified POT model and its application in early-warning methods for slope safety monitoring[J]. South-to-North Water Transfers and Water Science & Technology, 2016, 14(4): 192-197 (in Chinese).

(责任编辑:武立有)