

DOI:10.19651/j.cnki.emt.2415498

# 基于 ICEEMDAN 分解与 SE 重构和 DBO-LSTM 的滑坡位移预测\*

封青青 李丽敏 陈飞阳 张碧涵 余兵

(西安工程大学电子信息学院 西安 710600)

**摘要:** 滑坡位移预测是防灾减灾的一项重要工作,针对位移分解后趋势项和周期项重构的合理性问题以及周期项位移预测精度不高的问题,提出了一种改进的自适应噪声完备集合经验模态分解(ICEEMDAN)、样本熵(SE)以及蜣螂算法(DBO)优化的长短期记忆网络(LSTM)组合模型进行位移预测。以八字门滑坡为研究对象,利用 ICEEMDAN 方法将滑坡累计位移进行分解,并用样本熵值表征分解得到的子序列,将其重构为趋势项和周期项位移。之后利用 LSTM 模型预测趋势项和周期项位移;通过灰色关联度的方法确定周期项位移的影响因素。考虑到 LSTM 网络中超参数的随机性会影响模型预测精度,引入蜣螂优化算法获取 LSTM 最优超参数,最终将预测得到的趋势项和周期项位移叠加得到累计位移。本文所提的 ICEEMDAN-SE-DBO-LSTM 模型预测周期项位移的 RMSE、MAE、 $R^2$  3 项指标分别为 1.803 mm、1.584 mm、0.988,相较于 DBO-BP、LSTM、GRU 和 BP 模型预测效果更优,证明了模型的有效性。

**关键词:** 滑坡位移;改进的自适应噪声完备集合经验模态分解;样本熵;蜣螂优化算法

**中图分类号:** TN306; TP18; P642.22 **文献标识码:** A **国家标准学科分类代码:** 620.1030

## Landslide displacement prediction based on ICEEMDAN decomposition and SE reconstruction and DBO-LSTM

Feng Qingqing Li Limin Chen Feiyang Zhang Bihan Yu Bing

(School of Electronic Information, Xi'an Polytechnic University, Xi'an 710600, China)

**Abstract:** Landslide displacement prediction is an important task in disaster prevention and mitigation. Aiming at the rationality problem of trend term and period term reconstruction after displacement decomposition as well as the problem of low accuracy of period term displacement prediction, a combined model of improved adaptive noise complete ensemble empirical modal decomposition (ICEEMDAN), sample entropy (SE), and dung beetle optimization algorithm (DBO) optimization of the long- and short-term memory network (LSTM) is presented displacement prediction is performed. Taking the Bazimen landslide as the research object, the cumulative displacement of the landslide was decomposed using the ICEEMDAN method, and the subsequence obtained from the decomposition was characterized by the sample entropy value, which was reconstructed into the trend term and the period term displacements. After that, the LSTM model is used to predict the trend term and the period term displacements. The influence factors of the period term displacement are determined by the method of gray correlation. Considering that the randomness of hyperparameters in the LSTM network affects the model prediction accuracy, the dung beetle optimization algorithm is introduced to obtain the optimal hyperparameters of the LSTM, and finally the predicted trend term and period term displacements are superimposed to obtain the cumulative displacement. The ICEEMDAN-SE-DBO-LSTM model proposed in this paper predicts the period term displacement with the RMSE, MAE and  $R^2$  of 1.803 mm, 1.584 mm and 0.988, respectively, which is better than the DBO-BP, LSTM, GRU and BP models, and proves the effectiveness of the model.

**Keywords:** landslide displacement; improved adaptive noise complete ensemble empirical mode decomposition; sample entropy; dung beetle optimization algorithm

## 0 引言

长期以来,我国一直受到滑坡灾害问题的困扰和威胁,

每年都有大量的滑坡发生。它带来的危害是多方面的,不但会破坏生态环境,还会造成人员伤亡,给工农业生产带来巨大损失,因此很有必要加强对滑坡的预防工作。滑坡的

收稿日期:2024-02-13

\* 基金项目:国家自然科学基金(62203344)、陕西省自然科学基金基础研究计划(2022JM-322)、陕西省教育厅服务地方专项(22JC036)资助

发生是一个极其复杂的地质过程,受地质结构、降雨作用、土壤特性等因素的综合影响<sup>[1]</sup>,而位移可以直观反映滑坡的变形情况,可以通过分析历史监测数据对滑坡位移进行预测,以准确判断滑坡未来发生的可能性。

目前应用广泛的滑坡位移预测模型是基于数据驱动的模式,这类模型通过数据挖掘、机器学习算法对地面变形数据进行处理分析,可以全面、准确地预测地表的位移情况。具体实现方法是将滑坡位移分解为趋势项和周期项,结合滑坡影响因素,对分解项进行预测和累加,得到总位移的预测值。

针对滑坡位移预测,已有许多模型被提出。最初的模型结合经验公式实现预测。传统的机器学习模型如极限学习机(extreme learning machine, ELM)、支持向量回归(support vector regression, SVR)、反向传播(back propagation neural network, BPNN)等,因能够处理非线性数据而被广泛应用于位移预测中。曹博等<sup>[2]</sup>通过蚁群算法优化 ELM,取得了良好的位移预测效果。王晨辉等<sup>[3]</sup>用经验模态分解(ensemble empirical mode decomposition, EEMD)方法分解位移后结合 SVR 模型预测。之后,杨背背等<sup>[4]</sup>应用深度学习的方法,为滑坡位移预测提供了新思路。Duan 等<sup>[5]</sup>将自回归移动平均模型与 LSTM 模型进行对比,证明了 LSTM 在滑坡位移预测领域的优势。王朝阳等<sup>[6]</sup>采用卷积神经网络(convolutional neural network, CNN)耦合 LSTM 模型预测八字门滑坡周期项位移,张明岳等<sup>[7]</sup>首先应用变分模态分解方法(variational mode decomposition, VMD)分解位移后结合双向 LSTM(BiLSTM)预测周期项位移和随机项位移,Lin 等<sup>[8]</sup>通过自适应噪声完备集合经验模态分解(complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN)分解位移后用 CNN-BiLSTM 进行预测,这些单一 LSTM 和组合 LSTM 模型都能提高预测精度。

以上研究虽然取得一定的进展,但仍存在局限性。最初的模型主要依赖经验判据,无法实现高精度的预测。ELM、SVR、BP 等模型由于结构简单,无法充分挖掘数据中的复杂特征,限制了精度的提高。LSTM 模型虽在一定程度上能提升预测精度,但模型中超参数较多,随机组合的超参数也会影响模型性能。在位移分解上,移动平均法<sup>[9]</sup>会导致分解结果在时间上产生滞后,使趋势项无法反映滑坡实际变形情况。EEMD 改善了 EMD 分解存在的模态混叠问题,但引入随机噪声,使结果存在残余噪声。在此基础上,CEEMDAN<sup>[10]</sup>引入了自适应噪声,解决了 EEMD 的噪声残留问题,但分解早期仍存在伪模态。而改进的自适应噪声完备集合经验模态分解(improved complete ensemble empirical mode decomposition with adaptive noise, ICEEMDAN)<sup>[11]</sup>能同时减少伪模态的数量、降低噪声成分的影响,还能够保证分解得到的各个子序列之间相互独立。

综上所述,本文以八字门滑坡为研究对象,首先,将累计位

移通过 ICEEMDAN 方法分解后,为定量分析分解得到的固有模态函数(IMF),使周期项和趋势项更具合理性,采用样本熵(sample entropy, SE)方法重构 IMF。然后,使用蛭螂优化算法(dung beetle optimizer, DBO)获取 LSTM 模型最优超参数组合,采用 DBO-LSTM 模型预测趋势项位移;之后,通过灰色关联度值来分析影响因素与周期项位移之间的相关性以确定相关性强的影响因素,作为 DBO-LSTM 模型的输入,预测滑坡的周期项位移。最后,将两项位移相加得到滑坡累计位移预测值。

## 1 预测模型及方法

### 1.1 位移序列分解

在对滑坡影响因素综合分析的基础上,可以将滑坡累计位移分解为趋势项位移和周期项位移。通过该分解方式,滑坡位移的时间序列可表示为:

$$y(t) = p(t) + q(t) \quad (1)$$

式中: $y(t)$  表示累计位移, $p(t)$  表示趋势项位移, $q(t)$  表示周期项位移。

### 1.2 ICEEMDAN 算法

滑坡位移序列通常为非平稳序列<sup>[12]</sup>,ICEEMDAN 方法通过集成多个 CEEMDAN 分解结果,能更好地适应不同类型的序列特性,提高分解的可靠性。该方法也被应用于各个领域进行序列的分解,如短期电力负荷预测<sup>[13]</sup>、轴承故障诊断<sup>[14]</sup>。ICEEMDAN 算法实现流程如下:

1)对原始信号  $x$  添加白噪声得到  $x^{(i)}$ :

$$x^{(i)} = x + \beta_0 E_1(\omega^{(i)}), i = 1, 2, \dots, N \quad (2)$$

式中: $x^{(i)}$  是第  $i$  个构造信号; $\beta_0$  是第一次分解时的信噪比; $\omega^{(i)}$  是第  $i$  个被添加的零均值单位方差白噪声; $E_1(\omega^{(i)})$  是信号的首个 *imf* 分量。

2)重复以上步骤  $N$  次,计算第一个残差分量  $r_1$ :

$$r_1 = \frac{1}{N} \sum_{i=1}^N M(x^{(i)}) \quad (3)$$

式中: $M(x^{(i)})$  为局部均值函数。

3)将原始信号减去首个残余分量  $r_1$  得到首个模态分量 *imf*:

$$imf_1 = x - r_1 \quad (4)$$

4)经过  $k$  次分解后得到第  $k$  个残余分量:

$$r_k = \frac{1}{N} \sum_{i=1}^N M(r_{k-1} + \beta_{k-1} E_k(\omega^{(i)})) \quad (5)$$

5)用前一项残余分量减去  $r_k$  得到第  $k$  个 ( $k > 1$ ) 模态分量 *imf* <sub>$k$</sub> :

$$imf_k = r_{k-1} - r_k \quad (6)$$

6)重复步骤 4)和 5)即可获得所有分量。

### 1.3 样本熵

样本熵<sup>[15]</sup>是 Richman 等在 2000 年提出的,用于度量时间序列的复杂性和不规则性<sup>[16]</sup>,序列的熵值越大,复杂程度越高。其基本原理如下:

1) 假设分解得到的某个分量的时间序列表示为  $X = \{x(1), x(2), \dots, x(N)\}$ , 将时间序列  $X$  构造成  $m$  维向量, 即:

$$X(i) = \{x(i), x(i+1), \dots, x(i+m-1)\} \quad (7)$$

2) 计算  $X(i)$  与  $X(j)$  之间的距离:

$$d[X(i), X(j)] = \max_{k \in (0, m-1)} |x(i+k) - x(j+k)| \quad (8)$$

3) 将式(8)小于阈值  $r$  的数量与  $N-m$  作比值运算:

$$B_i^m(r) = \frac{1}{N-m} \text{num}\{d[X(i), X(j)] < r\} \quad (9)$$

4) 计算式(9)的平均值:

$$B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r) \quad (10)$$

5) 将维数  $m$  增加 1, 重复 2)~4) 步。计算该序列的样本熵值为:

$$\text{SampEn}(m, r) = \lim_{N \rightarrow \infty} \{-\ln[\frac{B^{m+1}(r)}{B^m(r)}]\} \quad (11)$$

嵌入维数  $m=1$  或 2; 相似容限  $r = (0.1 \sim 0.25) \times \text{std}$ , 其中  $\text{std}$  表示原序列的标准差。

#### 1.4 LSTM 神经网络

LSTM 是一种循环神经网络, 它通过引入 3 个门控制器即遗忘门、输入门和输出门来解决传统 RNN 存在的短时记忆问题。LSTM 的网络结构如图 1 所示。

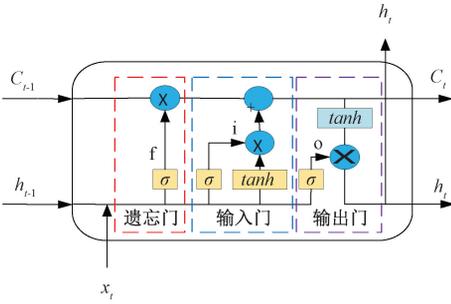


图 1 LSTM 网络结构图

图中  $x_t$  表示当前时刻的输入信息,  $h_{t-1}$  表示上一时刻的隐藏状态,  $h_t$  是当前时刻的隐藏状态。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (12)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (13)$$

$$\bar{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (14)$$

$$c_t = f_t \times c_{t-1} + i_t \times \bar{c}_t \quad (15)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (16)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (17)$$

式中:  $f_t, i_t, o_t$  分别为遗忘门、输入门、输出门,  $W$  和  $b$  分别为权重矩阵和偏置。

#### 1.5 蜣螂优化算法

蜣螂优化算法是 Xue 等<sup>[17]</sup>在 2022 年底提出的一种新的群智能算法, 具有寻优能力强、收敛效率高的特点<sup>[18]</sup>。

其思路来源于蜣螂的滚球、跳舞、觅食、偷窃和繁殖行为。该算法将蜣螂种群分为滚球蜣螂、产卵蜣螂、小蜣螂和小偷蜣螂 4 个部分。

1) 滚球蜣螂

滚球蜣螂以太阳为导航使粪球沿着直线滚动, 位置更新为:

$$x_i(t+1) = x_i(t) + \alpha \times k \times x_i(t-1) + b \times \Delta x \quad (18)$$

$$\Delta x = |x_i(t) - X^*| \quad (19)$$

式中:  $t$  代表当前迭代次数,  $x_i(t)$  表示第  $t$  次迭代时第  $i$  只蜣螂的位置,  $\alpha$  是自然系数, 取值为  $-1$  或  $1$ ,  $k$  是取值为  $(0, 0.2)$  的偏转系数,  $b$  为  $(0, 1)$  的常数,  $X^*$  表示全局最差位置,  $\Delta x$  用于模拟光强度变化。

当前进受阻会跳舞来调整自己的方向, 位置更新为:

$$x_i(t+1) = x_i(t) + \tan(\theta) |x_i(t) - x_i(t-1)| \quad (20)$$

式中:  $\theta$  取值范围为  $[0, \pi]$ 。

2) 产卵蜣螂

产卵蜣螂将粪球滚到安全区进行产卵, 产卵区域的定义为:

$$Lb^* = \max(X^* \times (1-R), Lb) \quad (21)$$

$$Ub^* = \min(X^* \times (1+R), Ub) \quad (22)$$

式中:  $X^*$  表示当前的局部最佳位置,  $Lb^*$  和  $Ub^*$  分别表示产卵区的上限和下限,  $R$  是惯性权重,  $Lb$  和  $Ub$  分别代表优化问题的下限和上限。产卵区的动态变化导致产卵蜣螂的位置也不断变化, 位置更新为:

$$B_i(t+1) = X^* + b_1 \times (B_i(t) - Lb^*) + b_2 \times (B_i(t) - Ub^*) \quad (23)$$

式中:  $B_i(t)$  是第  $t$  次迭代时第  $i$  个产卵球的位置信息,  $b_1$  和  $b_2$  代表两个大小为  $1 \times D$  的独立随机向量,  $D$  表示优化问题的维度。

3) 小蜣螂

小蜣螂在最佳觅食区寻找食物, 定义为:

$$Lb^b = \max(X^b \times (1-R), Lb) \quad (24)$$

$$Ub^b = \min(X^b \times (1+R), Ub) \quad (25)$$

式中:  $X^b$  表示当前的局部最佳位置,  $Lb^b$  和  $Ub^b$  分别表示最佳觅食区的下限和上限。小蜣螂的位置更新为:

$$x_i(t+1) = x_i(t) + C_1 \times (x_i(t) - Lb^b) + C_2 \times (x_i(t) - Ub^b) \quad (26)$$

式中:  $C_1$  为遵从正态分布的随机数,  $C_2$  是  $1 \times D$  随机向量。

4) 小偷蜣螂

小偷蜣螂会从其他蜣螂处偷取粪球, 位置更新为:

$$x_i(t+1) = X^b + S \times g \times (|x_i(t) - X^*| + |x_i(t) - X^b|) \quad (27)$$

式中:  $g$  是遵从正态分布的  $1 \times D$  随机向量,  $S$  是常数。

#### 1.6 DBO-LSTM 模型

在应用 LSTM 模型进行位移预测时, 超参数设置会对

模型预测结果产生较大影响,本文选用 DBO 算法对 LSTM 模型重要的超参数进行寻优,获得最优超参数组合以提高预测精度。DBO-LSTM 模型步骤如下:

- 1) 将 ZG111 监测点累计位移应用 ICEEMDAN 分解与 SE 重构后,得到趋势项和周期项位移;
- 2) 对 DBO 算法进行参数初始化设置,包括蛭螂种群 N、最大迭代次数 T、不同类蛭螂的比例等,适应度函数为训练集的均方误差。设置 LSTM 输入层、输出层神经元数量;
- 3) 计算蛭螂种群的初始适应度值;
- 4) 根据 DBO 算法原理更新蛭螂个体的位置;
- 5) 比较个体位置更新前后的适应度值,保留适应度值最优的位置;
- 6) 判断是否满足终止条件;
- 7) 根据以上步骤确定 LSTM 最优超参数组合,建立 DBO-LSTM 模型,实现位移预测。

## 2 研究区概况

### 2.1 八字门滑坡

八字门滑坡位于三峡库区湖北省秭归县归州镇香溪村,长江北岸支流香溪河右岸河口处,距三峡坝址 38 km。滑坡体整体西高东低,呈阶梯状起伏。八字门滑坡是三峡库区典型的呈“台阶状”的滑坡,因此对八字门滑坡进行分析预测,其方法对于三峡库区同类型的滑坡具有极强的借鉴意义。八字门滑坡地质平面图如图 2 所示。

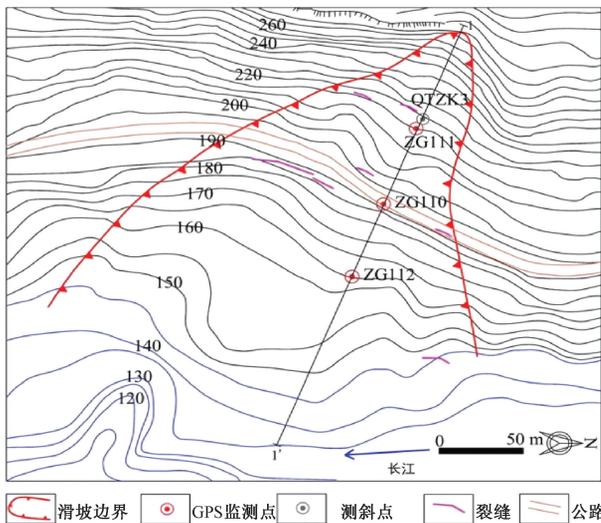


图 2 八字门滑坡平面及监测点布置图

八字门滑坡体上自 2003 年起布设有 ZG110、ZG111 监测点,从 2013 年 9 月又新增了 GSC1、GSC2、GSC3、GSC4、GSC5、GSC6、GSC7、GSC8、GSC9 等监测点。本文获取了八字门滑坡 2003 年 7 月~2018 年 12 月 ZG110 和 ZG111 监测点累计位移、降雨量、库水位的监测数据,数据变化情况如图 3 所示(数据来源:国家冰川冻土沙漠科学数据中心

<https://www.ncdc.ac.cn>)。

### 2.2 滑坡影响因素分析

降雨:从图 3 中可以看出,当每年的降雨量较少时,监测点 ZG110、ZG111 的累计变形位移总是缓慢增长,当降雨量处于每年的汛期(5~10 月)时,监测点累计位移总是呈“阶梯状”迅速增长,由此可见,降雨这一影响因素与滑坡变形之间存在着密切的联系。

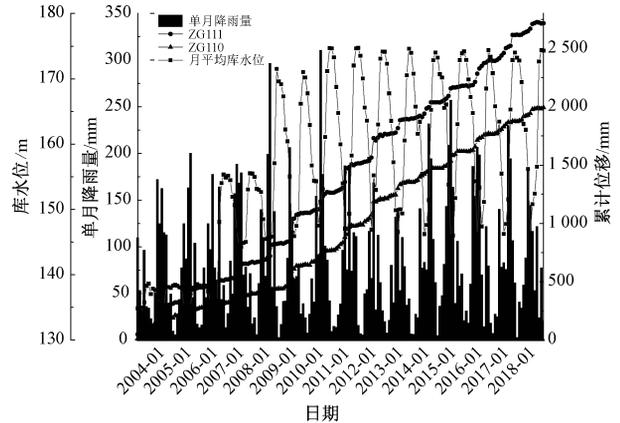


图 3 监测点累计位移、降雨量、库水位变化数据

库水位:库水位的变化与累计位移的关系也较为明显,在 2008 年 7 月三峡大坝首次开闸泄洪后,长江库水位开始出现更明显的周期变化,在 6~8 月之间,水库水位逐渐降至当年的最低点,滑坡累计位移出现了明显的增长。

## 3 滑坡位移预测实验验证分析

结合八字门滑坡两个监测点的监测数据分析,ZG111 监测点的监测数据更完整,位移变化特征更加明显,因此,本文选用了 ZG111 的监测数据,监测时间为 2004 年 1 月~2018 年 12 月,监测周期为每月一次,共 180 组监测数据。对该监测点的累计位移数据用 ICEEMDAN 方法分解为多个子序列,再通过样本熵的方法将子序列重构,得到周期项位移和趋势项位移。然后,应用 DBO-LSTM 模型预测趋势项位移和周期项位移,将预测结果相加得到累计位移预测值。

### 3.1 ICEEMDAN 分解位移

采用 ICEEMDAN 方法对八字门滑坡 ZG111 监测点 2004 年 1 月~2018 年 12 月的累计位移进行分解,设置参数值分别为:白噪声标准差为 0.1,最大迭代次数为 200,加入噪声的次数为 500,经过分解得到 IMF1~IMF5,分解结果如图 4 所示,其中 ZG111a 为原始序列数据。

### 3.2 样本熵重构

经过 ICEEMDAN 分解得到的子序列,计算各序列的样本熵值,根据样本熵值将子序列重构为两个序列,重构后的两个序列分别代表趋势项位移和周期项位移。

设置样本熵的嵌入维数  $m$  为 2,相似容限  $r$  为原位序序列标准差的 0.2 倍,各子序列样本熵值计算结果如表 1 所示。

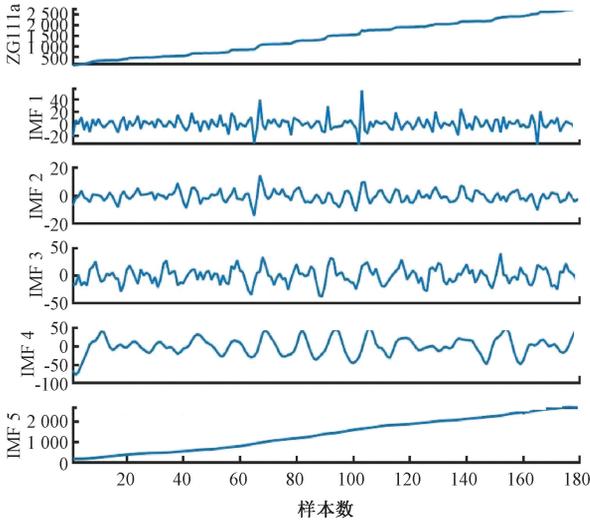


图 4 位移序列分解结果

表 1 子序列的样本熵值

序列	IMF1	IMF2	IMF3	IMF4	IMF5
样本熵值	1.504	1.666	1.806	0.924	0.012

根据表 1 可知,IMF1~IMF4 子序列的熵值均较高,说明复杂度较高,波动性较强。因此,将这 4 项子序列重构后作为周期项位移。

IMF5 子序列的熵值较小,由分解图也可以看出该序列相对平滑,几乎没有波动性,具有明显的平稳性,可以将其作为滑坡的趋势项位移表示滑坡长期的变形情况。重构后得到的周期项位移和趋势项位移如图 5 所示。

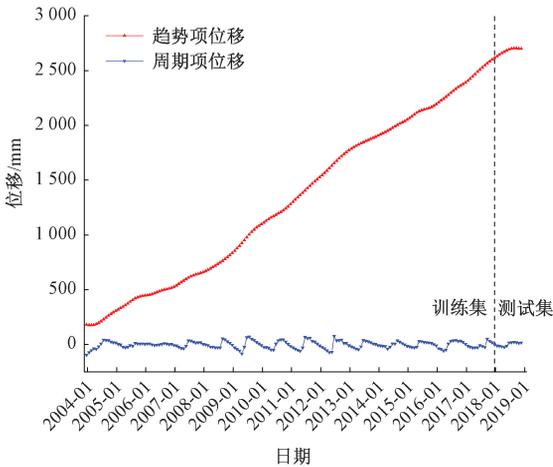


图 5 趋势项位移和周期项位移提取结果

由图 5 可以看出,周期项位移从 2008 年 7 月份开始周期性更加明显,这是由于三峡大坝首次开闸泄洪使长江库水位开始出现更明显的周期变化,而趋势项位移逐渐增加,变化平稳,符合滑坡受内部因素影响长期缓慢变形的特征。因此,ICEEMDAN 的位移分解和 SE 重构的方法对位移监测数据的处理效果具有合理性。

### 3.3 模型评价指标

为了进一步对模型的预测效果进行评价,选用了均方根误差(RMSE)、平均绝对值误差(MAE)和拟合优度( $R^2$ ) 3 项指标衡量,计算公式如下。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (28)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (29)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (30)$$

式中: $y_i$ 、 $\hat{y}_i$ 、 $\bar{y}$  分别表示真实值、预测值、真实值的平均值。

### 3.4 LSTM 模型超参数寻优

对 LSTM 模型的超参数进行调整,可以有效提升模型性能。本文选用 4 个需要优化的超参数,分别是隐含层单元数量、最大训练周期、学习率和正则化系数。隐含层单元数量决定了模型的复杂度,取值范围为[1,200],最大训练周期影响模型训练效果,取值范围为[10,1 000],学习率影响模型的收敛速度和性能,取值范围为[0.000 01,0.1],正则化系数用于防止模型过拟合。通过 DBO 算法改进的 LSTM 最优参数值如表 2 所示。

表 2 LSTM 网络最优参数值

隐含层单元数量	最大训练周期	学习率	正则化系数
101	100	0.009 9	0.000 1

### 3.5 趋势项位移预测

滑坡趋势项位移反映了滑坡长期的运动趋势和演化状态。如果呈现出不断变化或波动的趋势,则滑坡可能存在不稳定的迹象。根据上述提取出的趋势项位移图中可以看出,八字门滑坡的趋势项位移呈现逐渐增加的趋势,走势较为简单,较为稳定,应用 DBO-LSTM 网络模型对趋势项位移进行预测,预测结果如图 6 所示。

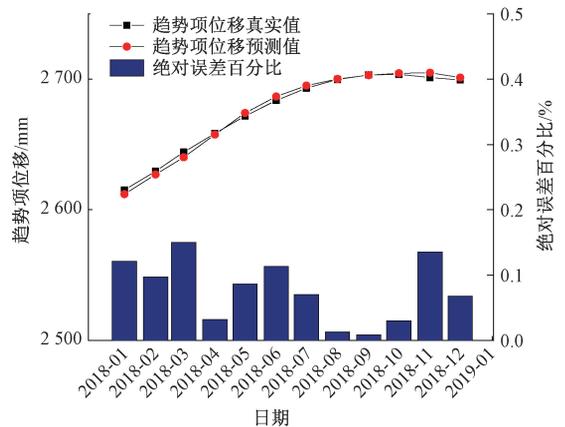


图 6 趋势项位移预测结果

采用 DBO-LSTM 网络模型预测趋势项位移,可以看出预测结果的趋势能较好跟随原趋势项位移的变化,预测结果的均方根误差为 5.73 mm,平均绝对误差百分比为 0.077%,拟合优度  $R^2$  为 0.993。

### 3.6 周期项位移预测

#### 1) 影响因素的确定

影响因素的合理选择决定了预测模型的可靠性。基于对外部影响因素的分析结果,降雨量,库水位都是影响八字门滑坡变形的关键因素。已有研究表明,滑坡过去的状态会在一定程度上影响未来的发展趋势。因此,将库水位高程、当月平均库水位、当月降雨量、双月降雨量以及前一月位移作为影响因素。

灰色关联度分析(grey relational analysis, GRA)方法适用于多因素的统计分析,是指在一个灰色系统中,已知某一指标以及影响该指标的相关因素,需要知道该指标与哪个因素相对来说关系强一点,用灰色关联度值的大小来描述因素间关系的强弱,保留关联性较强的影响因素。该方法也常用于选取滑坡周期项位移影响因素,杨玲等<sup>[19]</sup>、邓冬梅等<sup>[20]</sup>都应用灰色关联度分析方法来确定相关性强的影响因素。

根据上述所选影响因素,计算各因素与周期项位移的关联度值,结果如表 3 所示。

表 3 各影响因素与周期项位移的关联度值

影响因素	关联度
库水位高程	0.811 7
当月平均库水位	0.813 4
单月降雨量	0.706 2
双月降雨量	0.705 2
前一月位移	0.772 7

当影响因素与滑坡周期项位移的关联度超过 0.6,则认为该影响因素与滑坡变形有着较强的相关性。由灰色关联度结果可知,降雨量、库水位和前一月位移均与滑坡周期项位移有着较强相关性。因此,将以上几项因素作为滑坡周期项位移的影响因素。

#### 2) 数据处理

由于影响因素中降雨量和前一月位移的单位为 mm,而库水位的单位为 m,数值范围存在较大差异,并不适合直接输入模型进行训练,因此,需要对数据进行归一化确保数据在相同的尺度范围内,避免模型在训练过程中受到不同量级数据的影响。本文采用最小-最大归一化的方法,将数据统一映射到[0,1]范围内,归一化方法可用式(28)表示:

$$x' = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (31)$$

式中:  $x$  表示原始数据值,  $x_{\min}$  表示所在序列最小值,  $x_{\max}$  表示所在序列最大值,  $x'$  表示归一化后的值。

#### 3) 周期项位移预测结果

对周期项位移的预测在滑坡位移预测精度上起着决定

性作用,本文采用蛭螂优化算法优化的 LSTM 模型预测周期项位移,模型输入为库水位、当月平均库水位、单月降雨量、双月降雨量、前一个月位移,输出为当月的周期项位移。选取 ZG111 监测点 2004 年 1 月~2017 年 12 月共 168 组数据作为训练集,2018 年 1 月~2018 年 12 月共 12 组数据作为测试集,用来验证模型的效果。不同模型的周期项位移预测结果如图 7 所示。

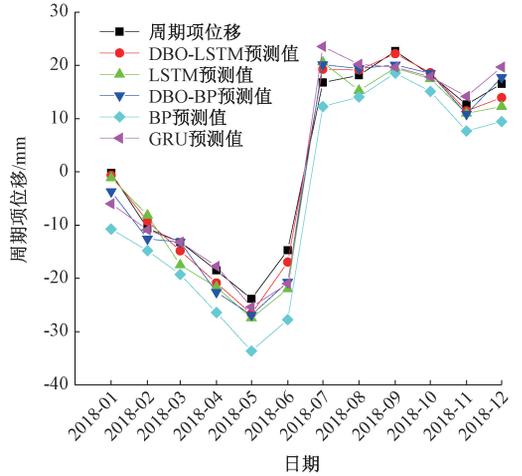


图 7 周期项位移预测结果

图 7 显示了不同模型的预测效果,5 种模型预测结果中只有未经优化的 BP 模型的预测效果较差,其他 4 种模型都能在各个阶段较好地跟随周期项位移真实值的变化趋势,预测误差相对较小。6~7 月间,数据急剧增长,导致各位移模型预测误差较大,BP 网络模型和 GRU 网络模型误差最大,分别为 4.51 mm 和 6.76 mm,而 DBO-LSTM 模型的误差最小,为 2.48 mm。

通过 RMSE、MAE 和  $R^2$  3 项指标来评价上述 5 种模型的预测效果,评价指标结果如表 4 所示。

表 4 各模型评价指标结果

模型	RMSE/mm	MAE/mm	$R^2$
DBO-LSTM	1.803	1.584	0.988
LSTM	3.609	3.193	0.952
DBO-BP	2.914	2.419	0.969
BP	7.267	6.623	0.805
GRU	3.450	2.566	0.956

结合表 4 结果可以得出,BP 网络模型的预测误差最大, $R^2$  仅为 0.805,经 DBO 算法优化的 BP 模型预测性能也有所提高。而其他模型预测误差相对较低, $R^2$  也都达到 0.95 以上,DBO-LSTM 模型的 3 项指标值均为最优,模型的预测效果最好。

### 3.7 累计位移预测

将 DBO-LSTM 模型预测的趋势项位移和周期项位移

相加后得到滑坡累计位移预测值。累计位移预测结果如图 8 所示。

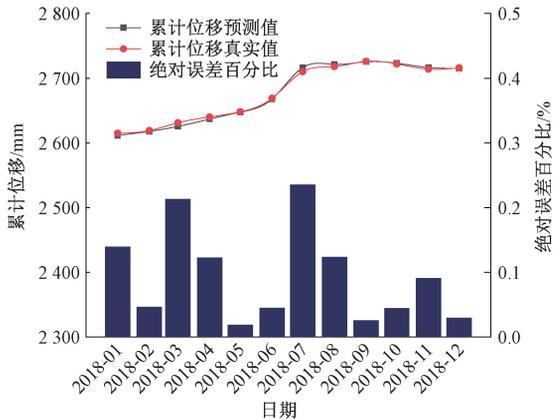


图 8 累计位移预测结果

由累计位移预测结果图可以看出,监测点 ZG111 的累计位移预测得到的结果值能很好地跟随实际值的变化,在 6~7 月的突然剧增时出现较大误差,导致 7 月份预测结果绝对误差百分比最高,为 0.24%,最低为 0.02%,平均绝对误差为 2.523 mm。由此可见,DBO-LSTM 位移模型能够反映出滑坡体的演变过程。

## 4 结 论

本文针对滑坡体位移时序变化特征,提出了一种基于 ICEEMDAN 分解与 SE 重构结合 DBO-LSTM 的滑坡位移预测模型,对八字门滑坡 ZG111 监测点的累计位移进行预测分析,得到以下结论:1)ICEEMDAN 方法是一种改进的自适应噪声完备集合经验模态分解方法,能够根据数据特点自适应地对噪声进行处理,还能够通过叠加模态函数重构原始数据,保证数据的完备性。2)周期项位移的预测结果对于滑坡位移预测的准确性有着决定性作用,因此要选取相关性较强的影响因素,本文中根据灰色关联度值来确定,降雨量、库水位以及前一个月位移 3 类因素都表现出较强的相关性,确定了合适的影响因素,有利于提高周期项位移预测精度。3)LSTM 神经网络具有强大的长序列预测能力,但其性能受到超参数的影响,合理的进行优化设置有助于发挥网络性能。因此引入了 DBO 优化算法进行超参数寻优,优化后的 LSTM 预测精度更高。4)采用 ICEEMDAN 分解并用样本熵重构后,基于 DBO-LSTM 模型对 ZG111 监测点位移进行分析预测,各项指标均表明预测结果较好,证明了所提模型的有效性。然而该模型对于其他类型的滑坡位移预测效果有待考证。

## 参考文献

[1] 吴益平,滕伟福,李亚伟. 灰色-神经网络模型在滑坡变形预测中的应用[J]. 岩石力学与工程学报,2007(3):

632-636.

- [2] 曹博,汪帅,宋丹青,等. 基于蚁群算法优化极限学习机模型的滑坡位移预测[J]. 水资源与水工程学报,2022,33(2):172-178.
- [3] 王晨辉,赵贻玖,郭伟,等. 滑坡位移 EEMD-SVR 预测模型[J]. 测绘学报,2022,51(10):2196-2204.
- [4] 杨背背,殷坤龙,杜娟. 基于时间序列与长短时记忆网络的滑坡位移动态预测模型[J]. 岩石力学与工程学报,2018,37(10):2334-2343.
- [5] DUAN G H, SU Y W, FU J. Landslide displacement prediction based on multivariate LSTM model [J]. International Journal of Environmental Research and Public Health, 2023, 20(2): 1167-1167.
- [6] 王朝阳,李丽敏,温宗周,等. 基于时间序列和 CNN-LSTM 的滑坡位移动态预测[J]. 国外电子测量技术,2022,41(3):1-8.
- [7] 张明岳,李丽敏,温宗周. 基于变分模态分解和双向长短时记忆神经网络模型的滑坡位移预测[J]. 山地学报,2021,39(6):855-866.
- [8] LIN Z, JI Y, SUN X. Landslide displacement prediction based on CEEMDAN method and CNN-BiLSTM model [J]. Sustainability, 2023, 15(13). Doi:10.3390/su151310071.
- [9] 唐菲菲,唐天俊,朱洪洲,等. 结合注意力机制和 Bi-LSTM 的降雨型滑坡位移预测[J]. 测绘通报,2022(9):74-79,104.
- [10] WANG C H, LI G C, GUO W, et al. Landslide displacement prediction based on CEEMDAN and grey wolf optimized-support vector regression model [J]. Frontiers in Earth Science, 2022, 10: 1-14.
- [11] 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.
- [12] 缪海波,殷坤龙,柴波,等. 基于非平稳时间序列分析的滑坡变形预测[J]. 地质科技情报,2009,28(4):107-111.
- [13] 高超,孙谊嫒,赵洪峰,等. ICEEMDAN-ISSA-LSTM 短期电力负荷预测[J]. 中国测试,2023,49(9):99-107.
- [14] 王朝兵,靳福涛,张龙,等. 基于 ICEEMDAN 与支持向量机的轴承故障诊断方法[J]. 机械设计与研究,2023,39(3):115-120.
- [15] RICHMAN J S, MOORMAN J R. Physiological time-

- series analysis using approximate entropy and sample entropy[J]. *American Journal of Physiology: Heart and Circulatory Physiology*, 2000, 278(6): H2039-49.
- [16] WANG Y T, WANG D. Investigations on sample entropy and fuzzy entropy for machine condition monitoring: revisited[J]. *Measurement Science and Technology*, 2023, 34(12), DOI:10.1088/1361-6501/aceb0d.
- [17] XUE J K, SHEN B. Dung beetle optimizer: A new meta-heuristic algorithm for global optimization[J]. *The Journal of Supercomputing*, 2022, 79(7): 7305-7336.
- [18] 李斌,高鹏,郭自强.改进蜣螂算法优化LSTM的光伏阵列故障诊断[J/OL]. *电力系统及其自动化学报*, 1-10 [2024-05-30]. <https://doi.org/10.19635/j.cnki.csu-epsa.001317>.
- [19] 杨玲,魏静,许子伏.基于平滑先验法-麻雀搜索算法-支持向量机回归模型的滑坡位移预测——以三峡库区八字门和白水河滑坡为例[J]. *地球科学与环境学报*, 2022, 44(6): 1096-1110.
- [20] 邓冬梅,梁焯,王亮清,等.基于集合经验模态分解与支持向量机回归的位移预测方法:以三峡库区滑坡为例[J]. *岩土力学*, 2017, 38(12): 3660-3669.

### 作者简介

**封青青**, 硕士研究生, 主要研究方向为智能算法及其在地质灾害监测领域的应用研究。

E-mail: fqqxjcn@163.com

**李丽敏**(通信作者), 博士, 副教授, 主要研究方向为智能算法及其在地质灾害监测领域的应用研究。

E-mail: 2364225096@qq.com