

# 基于 EMD-LSTM 模型的河流水流量水位预测

王亦斌<sup>1</sup>, 孙涛<sup>1</sup>, 梁雪春<sup>2</sup>, 谢海洋<sup>1</sup>

(1. 南水北调东线江苏水源有限责任公司, 江苏 南京 210019;  
2. 南京工业大学电气工程与控制科学学院, 江苏 南京 211816)

**摘要:** 基于经验模式分解方法和长短期记忆网络 (empirical model decomposition and long short-term memory network, EMD-LSTM) 模型对水位数据进行预测。先采用中值滤波对数据序列进行预处理, 然后对数据序列进行 EMD 分解, 并对 EMD 分解的每个特征序列使用 LSTM 模型进行预测, 最后叠加各个序列预测值, 得到最终的预测结果。以南水北调工程某河流每隔 1 h 的瞬时流量、流速和水深监测数据为研究对象, 采用 EMD-LSTM 模型进行建模, 试验结果表明, 该模型能够实现水位、流速和瞬时流量连续 12 h 和 6 h 的准确预测, 且比 LSTM 模型具有更高的预测精度, 可为水位判别和水资源的实时调度提供决策依据。

**关键词:** 水量预测; 水位预测; 中值滤波; 经验模式分解方法; 长短期记忆神经网络

**中图分类号:** TV125      **文献标志码:** A      **文章编号:** 1006-7647(2020)06-0040-08

**Prediction of river water flow and water level based on EMD-LSTM model//**WANG Yibin<sup>1</sup>, SUN Tao<sup>1</sup>, LIANG Xuechun<sup>2</sup>, XIE Haiyang<sup>1</sup> (1. Eastern Route of South-to-North Water Diversion Project Jiangsu Water Resources Co., Ltd., Nanjing 210019, China; 2. College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211816, China)

**Abstract:** The combination of empirical model decomposition and long short-term memory network (EMD-LSTM) model was used to forecast the hydrological time series data. The median filter for data preprocessing was used first and then EMD was applied to decompose the time sequence. The LSTM model was used to prediction for each characteristic series from EMD, and each prediction sequence was superposed to obtain the final prediction result. Based on the data of the instantaneous flow, water velocity and water level per hour of a certain river in the South-to-North Water Diversion Project, the EMD-LSTM was used for modeling. The experimental results show that this method can realize accurate prediction of water level, water velocity and instantaneous flow for 12 h and 6 h continuously, has higher accuracy compared with the LSTM model and provides decision basis for the forecast of hydrological time series and the real-time dispatching of water resources.

**Key words:** water flow prediction; water level prediction; median filter; empirical mode decomposition; long short-term memory network

准确的水位和水量等水文时间序列预测是水资源管理的重要依据,对于调水监测有着重要的作用。传统的水文时间序列预测方法有回归分析、时间序列分析等,还有模糊识别法、灰色模型、小波神经网络分析<sup>[1-5]</sup>。随着机器学习方法的兴起,很多学者使用机器学习方法进行建模,相对于传统的统计预测模型,机器学习方法具有更好的适应性。面对数据量化不确定性问题,Porter 等<sup>[6]</sup>提出一种人工神经网络 (artificial neural networks, ANN) 方法来预测地下水水位;针对在台风期间受潮汐影响水位波动较大问题,Wei<sup>[7]</sup>提出了小波-支持向量机算法来预测每小时的水位数据,为台风袭击期间的水位预测问题提供了实用的解决方案;Moosavi 等<sup>[8]</sup>提出了小波-自适应神经模糊推理系统 (wavelet-adaptive neuro-fuzzy inference system, wavelet-ANIS) 方法,比较了 ANN、Wavelet-ANN、ANIS 和 Wavelet-ANIS 在地下水位预报上的表现;Zhong 等<sup>[9]</sup>分析了长江潮汐水位预测的复杂性,在获取了长江下游 (安庆、芜湖、南

京)的每日水位的历史数据后,提出使用人工智能-Kalman 滤波方法进行水位预报;Wang 等<sup>[10]</sup>分析南水北调中线陶岔渠首的水位数据情况,提出了一种改进 Cuckoo-搜索梯度增强树算法(CCS-gradient boosting decision tree)进行水位预报。近年来,随着深度学习技术的发展,深度神经网络也逐渐被应用于时间序列预测分析中,尤其是循环神经网络(recurrent neural networks, RNN)在时序数据处理上表现出很强的适应性。Chang 等<sup>[11]</sup>使用循环神经网络作城市水位预测和洪水预报。张轩等<sup>[12]</sup>为提高秦淮河流域东山站水位预报的精度,基于 BP 神经网络算法建立经验预报模型,分别根据降雨历时、起涨水位两种模式对水位涨幅进行预报。常见的 RNN 等算法存在梯度消失或者梯度爆炸的问题,而长短期记忆神经网络(long short-term memory networks, LSTM)是一种引入了遗忘门、输入门和输出门 3 种门控单元和记忆细胞的特殊形式的 RNN,这种特殊的结构可以缓解梯度消失或者梯度爆炸问题<sup>[13]</sup>。LSTM 通过门控结构调整隐藏状态中信息的流动,输入门可以识别需要添加到长期项的状态,遗忘门可以控制需要消除的长期项状态,可以使得模型能学习何时忘记以前的隐藏状态,何时根据新的信息更新隐藏状态,从而能够成功地捕获时间序列中的长期模式,因此,被广泛应用于交通、金融和经济序列预测中<sup>[14-16]</sup>。LSTM 模型在水文时间序列预报和模拟方面也受到了关注。为了有效提高流域内站点径流量预测的精度,冯锐<sup>[17]</sup>运用 LSTM 技术建立九龙江流域站点径流序列预测模型和数据模拟模型。刘亚新等<sup>[18]</sup>综合使用梯度下降与 Broyden-Fletcher-Goldfarb-Shanno 算法来训练 LSTM 模型,并将模型用于葛洲坝水电站水位预测。以入湖流量和长江干流流量作为输入条件,郭燕等<sup>[19]</sup>采用 LSTM 模型预测鄱阳湖的水位变化。闫佰忠等<sup>[20]</sup>基于地下水位、蒸发量、降水量、气温、气压、相对湿度、日照时长、开采量等多个指标构建多变量的 LSTM 地下水位预测模型。通常水文时间序列是非平稳且非线性的,为提高预测精度,陈旭等<sup>[21]</sup>提出了基于经验模式分解(empirical mode decomposition, EMD)的 AR 预测模型用于水文站的年径流量预测。基于上述研究,采用 EMD 和 LSTM 相结合的算法(EMD-LSTM)进行河流水流量水位预测,首先采用信号处理中常用的经验模式分解方法将水文时间序列数据分解为不同时间尺度的固有模态函数,然后对 EMD 分解的每个特征序列使用 LSTM 模型进行预测,最后叠加每个特征序列预测值得到预测结果。

南水北调工程对改善水环境、促进区域经济和

社会的可持续发展具有重要的战略意义。在我国南水北调工程的重要河段中,准确预测河流水位水量,对水库调度、防洪、发电、灌溉等具有重要意义。本文拟使用 EMD 和 LSTM 模型对水位、流速和水量进行预测。以南水北调工程某河流 2017 年 11 月到 2018 年 6 月中每隔 1 h 瞬时流量、流速和水深监测数据为研究对象,采用中值滤波对数据进行预处理,基于 EMD-LSTM 进行时间序列建模。

## 1 研究方法

### 1.1 EMD

EMD 是 Huang 等<sup>[22]</sup>提出的自适应信号时频处理方法,它不需要预先设置任何基函数,根据数据本身的时间尺度特征将信号分解成有限数目的固有模态函数。对时间序列 $\{X_t\}$ 的 EMD 分解过程如下:

**步骤 1:** 求出时间序列 $\{X_t\}$ 的极大值和极小值,由三次样条插值函数拟合所有极值点得到上包络线 $\{U_t\}$ 和下包络线 $\{L_t\}$ 。

**步骤 2:** 计算上包络线和下包络线的均值,记为

$$m_t = \frac{U_t + L_t}{2} \quad (1)$$

**步骤 3:** 令

$$\xi_t^{(1)} = X_t - m_t \quad (2)$$

**步骤 4:** 检验 $\xi_t^{(1)}$ 是否为本征模函数(intrinsic mode function, IMF),即是否满足极值点的个数和零点的个数相等或个数之差最多不能超过一个,且在任意时刻上下包络线的平均值为零,若是将 $\xi_t^{(1)}$ 取为第一个 IMF 记为 $I_t^{(1)}$ ;否则,将 $\xi_t^{(1)}$ 作为一个新的输入序列,重复步骤 1~4 直到得到 IMF 为止。

**步骤 5:** 令残差

$$r_t^{(1)} = X_t - I_t^{(1)} \quad (3)$$

**步骤 6:** 以 $r_t^{(1)}$ 作为新的输入序列重复步骤 1~5 直到满足终止条件(通常最后一个残差满足单调性)。通过以上步骤,筛选出一系列 IMF( $I_t^{(i)}, i = 1, 2, \dots, N$ ),原始信号 $\{X_t\}$ 由这些固有模态函数重建,即

$$X_t = \sum_{i=1}^N I_t^{(i)} + R_t \quad (4)$$

式中: $R_t$ 为 EMD 分解的残差。

### 1.2 LSTM

LSTM 是一种引入了遗忘门、输入门和输出门 3 种门控单元和记忆细胞(可看作一种特殊的隐藏状态)的特殊形式的循环神经网络,这种精心设计的结构缓解了 RNN 的梯度消失问题, LSTM 神经元结构如图 1 所示,其中, $X_t$ 为当前时刻输入; $H_{t-1}, H_t$ 分别为前一和当前时刻的隐藏状态; $I_t, F_t, O_t$ 分别为当前时

刻的输入门、遗忘门和输出门; $G_t$ 为当前时刻的候选记忆细胞; $c_{t-1}, c_t$ 分别为前一和当前时刻的记忆细胞; $Y_t$ 为输出。关于 LSTM 的更多细节可参考文献[15]。

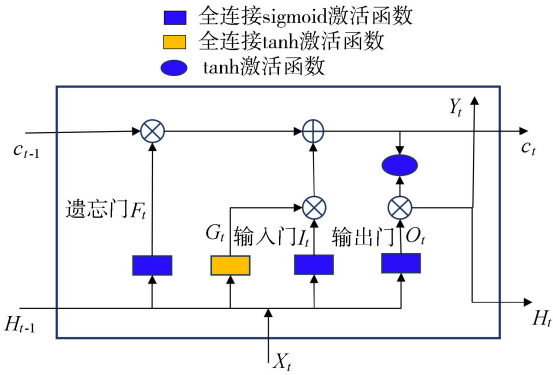


图1 LSTM 细胞单元结构

设隐藏单元数为  $h$ ，在  $t$  时刻输入  $X_t \in \mathbf{R}^{n \times d}$ （其中  $n$  为样本数， $d$  为输入长度），前一时刻的隐藏状态  $H_{t-1} \in \mathbf{R}^{n \times h}$ （可视爲短期状态），则在  $t$  时刻输入门  $I_t \in \mathbf{R}^{n \times h}$ ，遗忘门  $F_t \in \mathbf{R}^{n \times h}$  和输出门  $O_t \in \mathbf{R}^{n \times h}$  表达式为

$$I_t = \sigma(X_t W_{XI} + H_{t-1} W_{HI} + b_I) \tag{5}$$

$$F_t = \sigma(X_t W_{XF} + H_{t-1} W_{HF} + b_F) \tag{6}$$

$$O_t = \sigma(X_t W_{XO} + H_{t-1} W_{HO} + b_O) \tag{7}$$

式中： $I_t, F_t, O_t$  都由 sigmoid 激活函数控制，它们的输出值都在 0 ~ 1 之间。记忆细胞  $c_t$ （可视爲长期状态）可由式(9)计算：

$$G_t = \tanh(X_t W_{XG} + H_{t-1} W_{HG} + b_G) \tag{8}$$

$$c_t = F_t \otimes c_{t-1} + I_t \otimes G_t \tag{9}$$

式中： $W_{XI}, W_{XF}, W_{XO}, W_{XG} \in \mathbf{R}^{d \times h}$  为 4 个全连接层关于它们的输入向量的权重矩阵； $W_{HI}, W_{HF}, W_{HO}, W_{HG} \in \mathbf{R}^{h \times h}$  为 4 个全连接层关于它们的短期状态  $H_{t-1}$  的权重矩阵； $b_I, b_F, b_O, b_G \in \mathbf{R}^{1 \times h}$  为 4 个全连接层的线性偏倚； $\sigma(\cdot)$  和  $\tanh(\cdot)$  为 sigmoid 函数和双曲正切函数； $\otimes$  表示向量间的逐元素相乘。隐藏状态和输出的表达式为

$$Y_t = H_t = O_t \otimes \tanh(c_t) \tag{10}$$

### 1.3 经验模式分解方法和长短期记忆网络

本文使用 EMD-LSTM 方法进行建模，即先将原始数据序列  $\{X_t\}$  进行 EMD 分解为本征模函数  $I_t^{(i)}$  和残差部分  $R_t$ ，即  $X_t = R_t + \sum_{i=1}^N I_t^{(i)}$ ，对  $I_t^{(i)}$  和  $R_t$  序列分别使用 LSTM 方法进行预测，并将预测结果叠加得到原始数据序列的预测。

## 2 试验步骤

### 2.1 数据

以南水北调工程某河流 2017 年 11 月 3 日到

2018 年 6 月 30 日每隔 1 h 的中水文数据为试验对象，所记录的水文数据包括水深、流速和瞬时流量，共有 5 727 条记录，水深、流速和瞬时流量的时间序列曲线见图 2。

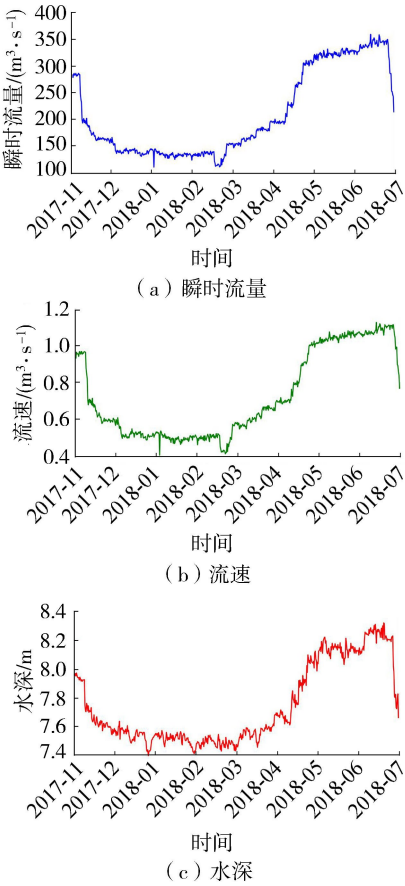


图2 瞬时流量、流速和水深数据序列

### 2.2 数据预处理和 EMD 分解

水流变化会受到天气、测量等随机因素的影响，原始数据包含一些噪声数据，噪声数据达到一定程度会影响预报的准确性，为此采用中值滤波来消除噪声数据。记  $\{x_1, x_2, \dots, x_T\}$  为观测时间序列，其中  $x_t$  为  $t$  时刻观测值，令

$$\hat{x}_i = \begin{cases} x_i & i \leq w \text{ 或 } i > T - w \\ \text{median} \{x_{i-w}, x_{i-w+1}, \dots, x_i, x_{i+1}, \dots, x_{i+w}\} & w < i \leq T - w \end{cases} \tag{11}$$

式中： $\text{median}$  为序列中位数； $\{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_T\}$  为预处理后的时间序列，然后对预处理后的时间序列进行 EMD 分解； $w$  为中值滤波的窗口长度。

### 2.3 LSTM 模型构建

为记号方便起见，记  $\{y_1, y_2, \dots, y_T\}$  为 LSTM 模型需要处理的时间序列，输入变量的时间窗口为  $K$ ，即选择当前时刻的前  $K$  个时间观测值作为输入变量来预测，LSTM 模型的输入为

$$\mathbf{Y}_{\text{input}} = \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_{T-K} \end{bmatrix} = \begin{bmatrix} y_1 & \cdots & y_K \\ \vdots & & \vdots \\ y_{T-K} & \cdots & y_{T-1} \end{bmatrix} \quad (12)$$

输出为

$$\mathbf{Y}_{\text{output}} = \begin{bmatrix} \hat{\mathbf{Y}}_1 \\ \vdots \\ \hat{\mathbf{Y}}_{T-K} \end{bmatrix} = \begin{bmatrix} \hat{y}_{K+1} & \cdots & \hat{y}_{K+L} \\ \vdots & & \vdots \\ \hat{y}_{T-K} & \cdots & \hat{y}_{T-K+L} \end{bmatrix} \quad (13)$$

式中: $L$ 为预测窗口长度,也就是说用前面的 $K$ 个时刻的观测值来预测接下来 $L$ 个时刻的数据。使用均方根误差 $E_R$ :

$$E_R = \sqrt{\frac{1}{T} \sum_{i=1}^T (\hat{y}_i - y_i)^2} \quad (14)$$

作为损失函数来确定隐藏状态的维数和式(5)~(10)的权重矩阵,采用均方根误差 $E_R$ 和平均相对误差 $E_M$ 来评价模型的预测精度, $E_M$ 定义为

$$E_M = 100 \sum_{i=1}^T \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (15)$$

式中: $\hat{y}_i$ 为 $t$ 时刻的预测值; $y_i$ 为 $t$ 时刻的真实值。

## 2.4 预测结果

把以南水北调工程某河流 2017 年 11 月 3 日到 2018 年 6 月 30 日每隔 1 h 瞬时流量、流速和水深

3 个时间序列分别按照 3 : 1 作为训练集数据和测试集数据,即每个序列的前 4 259 个数据作为训练集,后 1 432 个数据作为测试集。

首先对均值滤波预处理后的瞬时流量数据进行 EMD 分解为本征模函数部分和残差部分,对 EMD 分解的各个本征模函数和残差部分分别训练 LSTM 方法进行预报,并将预报结果进行叠加构建瞬时流量的预测结果。LSTM 实验在 PyTorch 环境下进行,在数据预处理阶段均值滤波的窗口长度设置为 2,LSTM 模型训练中以均方根误差作为损失函数,使用的梯度下降优化器为 Adam 优化器,输入变量的时间窗口  $K=24$  h,即每次预测需要前 24 h 的观测数据,epoch(指所有的训练样本输入网络中完成一次前向和反向传播的过程)设为 200,学习率为 0.000 1,隐藏单元数为 64。图 3 为预测窗口长度  $L=12$  h 的 LSTM 模型对瞬时流量的 EMD 分量( $I^{(1)}, I^{(2)}, \dots, I^{(7)}$ 和残差 $R_t$ )预测结果。

从图 3 可以看出,对 $I_t^{(1)}, I_t^{(2)}, I_t^{(3)}, I_t^{(4)}$ 预测准确性相对较低,这是由于高频分量类似噪声,规律性较差,而残差反映了指标数据的长期趋势,残差的预测值与实际值比较接近。将预报结果进行叠加构建瞬时流量的 EMD-LSTM 预测结果,图 4(a)(b)为预测

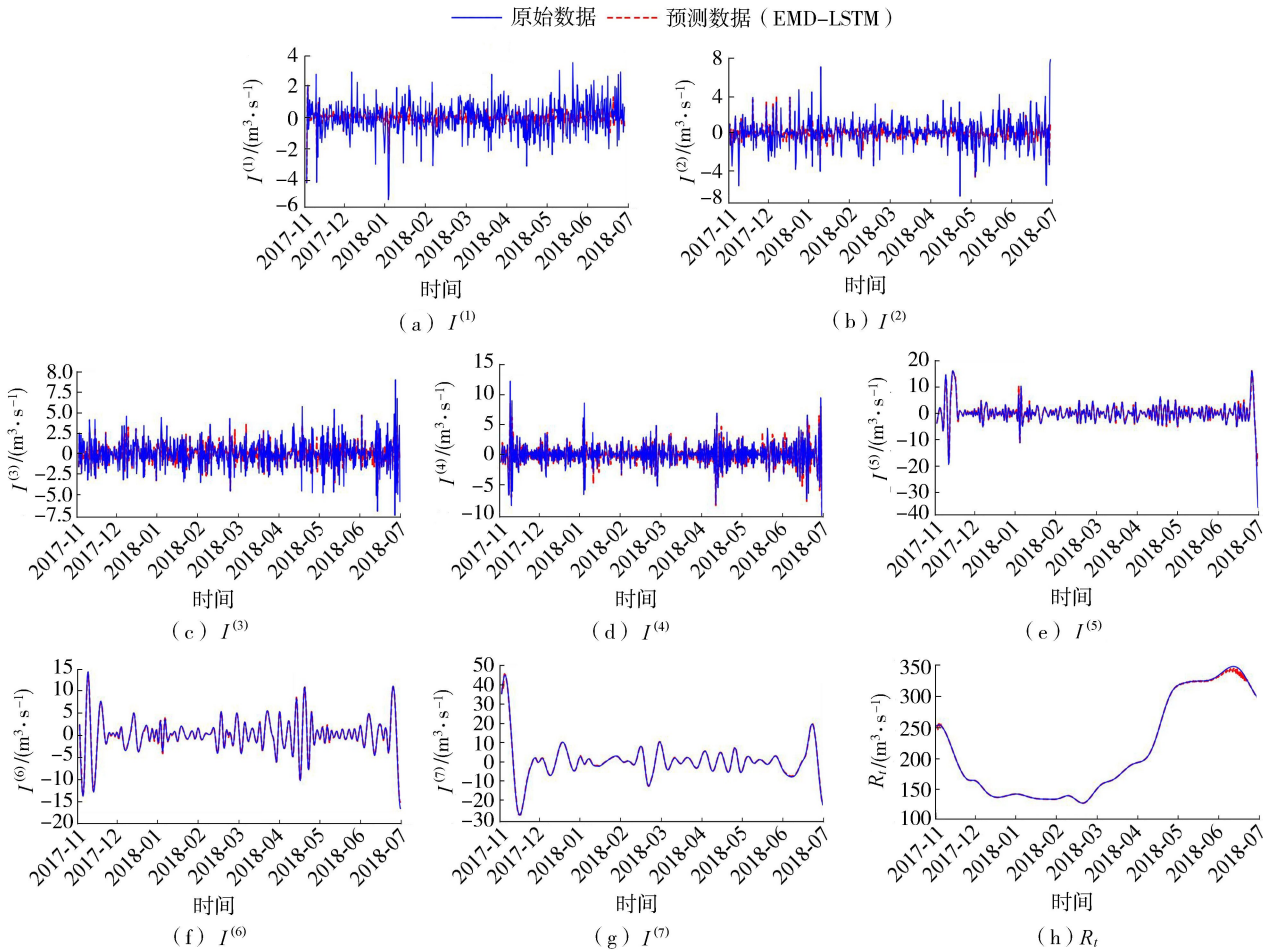


图 3 预测窗口长度为 12 h 的 LSTM 模型对瞬时流量的 EMD 分量预测结果

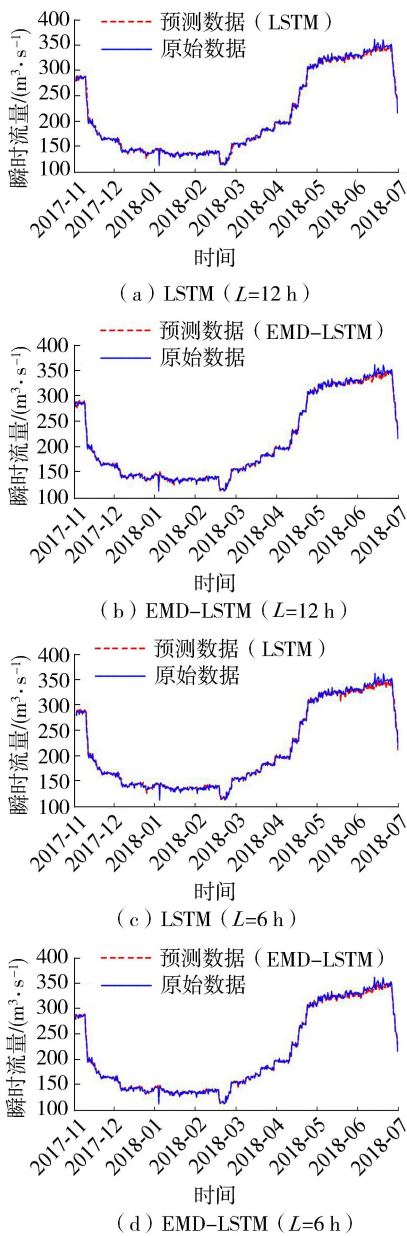


图4 不同预测窗口长度两种方法对瞬时流量的预测结果

窗口长度为 12 h 关于瞬时流量 LSTM 预测结果和 EMD-LSTM 预测的结果比较。改变预测时间窗口长度为 6 h,对瞬时流量进行预测,两种模型的预测效果见图 4(c)(d)。由图 4 可知,LSTM 方法和 EMD-LSTM 方法都能取得较好的预测效果,特别当预测时间较短时,预测效果更优。为了更为显著比较两种方法优劣,截取 2018 年 6 月 11 日 0 点到 6 月 25 日 24 点共 360 h 数据(该时段内的数据变换形式较多,

因此更能衡量预测算法优劣。)在预测窗口长度为 12 h 和 6 h 两种情况下的 LSTM 预测方法和 EMD-LSTM 预测方法的结果进行比较,如图 5 所示。由图 5 可知,对瞬时流量而言,EMD-LSTM 预测方法优于 LSTM 方法,特别对幅度变化较大时段,EMD-LSTM 方法优势更为显著。同时,短时段的预测效果优于长时段,原因在于预测窗口长度太长,数据信号会产生新的变化,而这些变化很难从历史数据中学习。

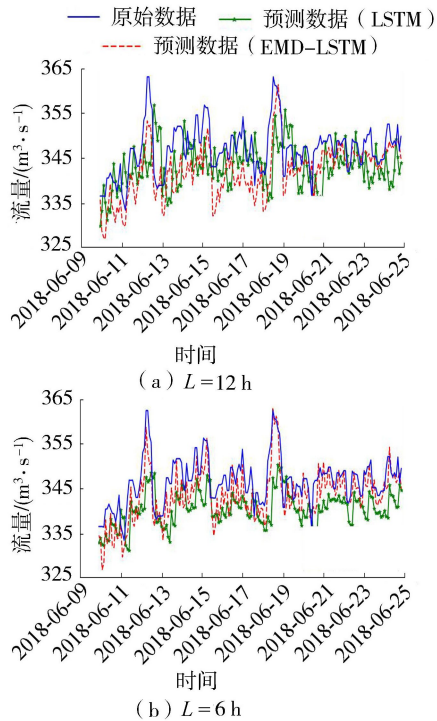


图5 不同预测窗口长度瞬时流量预测结果比较

对于两种不同预测窗口长度,对瞬时流量序列在训练集和测试集上的两种模型下瞬时流量预测误差比较见表 1,其中  $E_R^{To}$ 、 $E_R^{Tr}$  和  $E_R^{Te}$  分别为在整个数据集、训练集和测试集上的均方根误差,  $E_M^{To}$ 、 $E_M^{Tr}$  和  $E_M^{Te}$  分别为在整个数据集、训练集和测试集上的平均相对误差。从表 1 可以看出,两种不同预测窗口下,EMD-LSTM 模型的均方根误差和平均相对误差均小于 LSTM 模型,EMD-LSTM 模型的预测效果优于 LSTM;随着预测窗口长度变小,LSTM 和 EMD-LSTM 模型的均方根误差和平均相对误差都变小,预测精度提高。

进一步对流速和水深分别使用 LSTM 模型和 EMD-LSTM 模型进行预测,预测效果见图 6~9,预测误差见表 2 和表 3。从对瞬时流量、流速和水深

表1 两种模型下瞬时流量预测误差比较

| 模型       | L/h | $E_R^{To}/(\text{m}^3 \cdot \text{s}^{-1})$ | $E_R^{Tr}/(\text{m}^3 \cdot \text{s}^{-1})$ | $E_R^{Te}/(\text{m}^3 \cdot \text{s}^{-1})$ | $E_M^{To}/\%$ | $E_M^{Tr}/\%$ | $E_M^{Te}/\%$ |
|----------|-----|---------------------------------------------|---------------------------------------------|---------------------------------------------|---------------|---------------|---------------|
| LSTM     | 12  | 5.5801                                      | 4.4064                                      | 8.1443                                      | 1.6609        | 1.7065        | 1.5241        |
| EMD-LSTM | 12  | 2.9277                                      | 2.1229                                      | 4.5183                                      | 0.9292        | 0.9018        | 1.0115        |
| LSTM     | 6   | 4.6284                                      | 3.5371                                      | 6.9390                                      | 1.3659        | 1.3823        | 1.3168        |
| EMD-LSTM | 6   | 2.4473                                      | 1.6032                                      | 3.9897                                      | 0.7477        | 0.6734        | 0.9703        |

这 3 种指标的预测结果可以看出,两种模型都可以很好地预测水文数据的变化,但 EMD-LSTM 模型的预测精度更高,特别对波动较大时段,EMD-LSTM 方

法优势更为显著,这是由于 LSTM 方法适合于预测相对比较规律的信号,高频噪声的干扰会极大地降低 LSTM 方法的预测精度,使用 EMD-LSTM 可以把

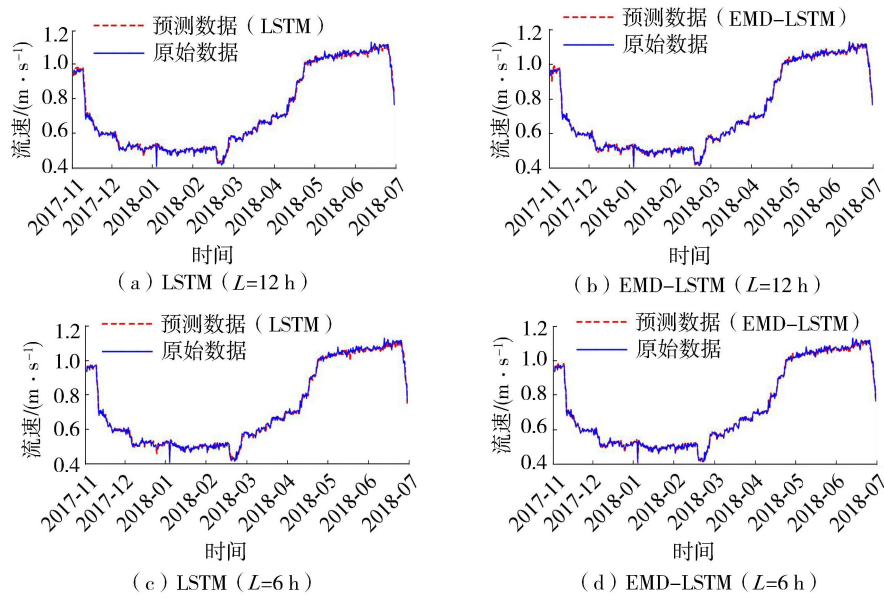


图 6 不同预测窗口长度两种方法对流速的预测结果

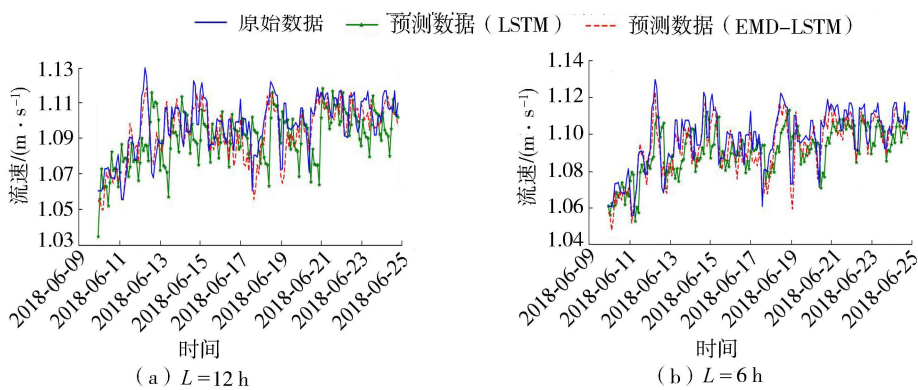


图 7 不同预测窗口长度流速预测结果比较

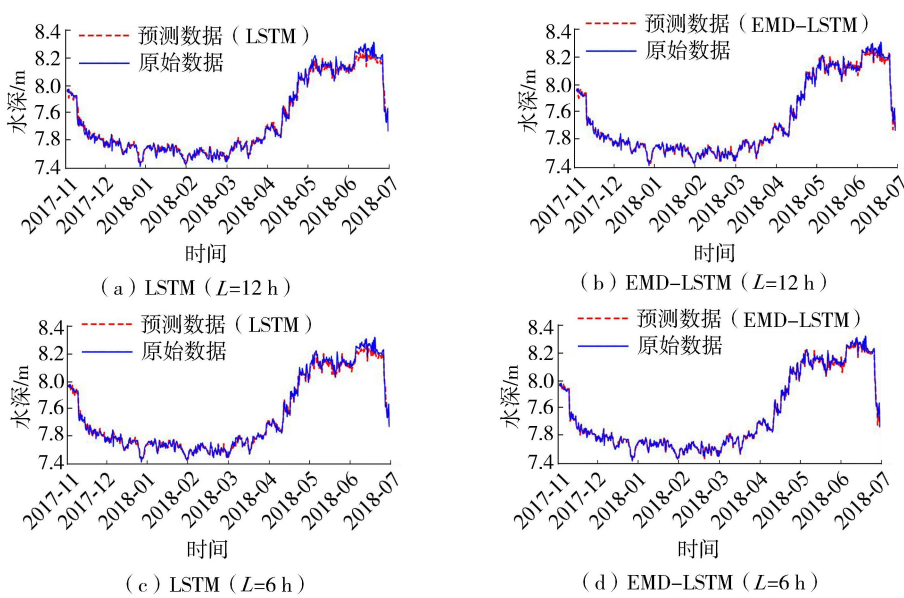


图 8 不同预测窗口长度两种方法对水深的预测结果

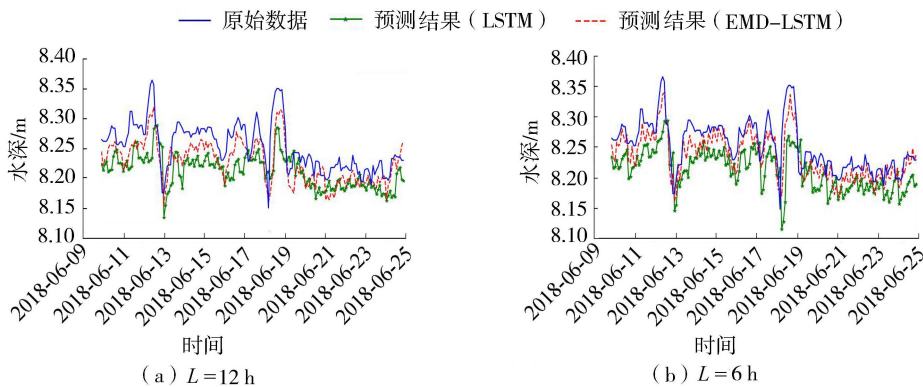


图9 不同预测窗口长度水深预测结果比较

表2 两种模型流速预测误差比较

| 模型       | $L/h$ | $E_R^{To}/(m^3 \cdot s^{-1})$ | $E_R^{Tr}/(m^3 \cdot s^{-1})$ | $E_R^{Te}/(m^3 \cdot s^{-1})$ | $E_M^{To}/\%$ | $E_M^{Tr}/\%$ | $E_M^{Te}/\%$ |
|----------|-------|-------------------------------|-------------------------------|-------------------------------|---------------|---------------|---------------|
| LSTM     | 12    | 0.0153                        | 0.0149                        | 0.0163                        | 1.5597        | 1.6755        | 1.2122        |
| EMD-LSTM | 12    | 0.0079                        | 0.0077                        | 0.0083                        | 0.8047        | 0.8745        | 0.5954        |
| LSTM     | 6     | 0.0127                        | 0.0118                        | 0.0151                        | 1.2224        | 1.3272        | 0.9081        |
| EMD-LSTM | 6     | 0.0060                        | 0.0056                        | 0.0069                        | 0.6221        | 0.6585        | 0.5129        |

表3 两种模型水深预测误差比较

| 模型       | $L/h$ | $E_R^{To}/m$ | $E_R^{Tr}/m$ | $E_R^{Te}/m$ | $E_M^{To}/\%$ | $E_M^{Tr}/\%$ | $E_M^{Te}/\%$ |
|----------|-------|--------------|--------------|--------------|---------------|---------------|---------------|
| LSTM     | 12    | 0.0314       | 0.0239       | 0.0474       | 0.2701        | 0.2212        | 0.4166        |
| EMD-LSTM | 12    | 0.0162       | 0.0115       | 0.0253       | 0.1360        | 0.1047        | 0.2298        |
| LSTM     | 6     | 0.0233       | 0.0181       | 0.0342       | 0.2164        | 0.1588        | 0.3891        |
| EMD-LSTM | 6     | 0.0122       | 0.0085       | 0.0192       | 0.0985        | 0.0743        | 0.1711        |

信号中的高频部分分离出来预测,极大地减少了高频噪声的干扰。另外,两种模型下,预测窗口长度越小,均方根误差和平均相对误差越小,预测越准确,曲线拟合越好,即短时段的预测效果优于长时段,原因在于预测窗口长度太长,受气候、环境和人类活动影响,水文数据信号会产生新的变化,而这些变化很难从历史数据中学习。

### 3 结 论

本文建立了基于 EMD-LSTM 神经网络的预测模型,并分别对水深、流速和瞬时流量序列预报,对每个序列预报的主要步骤是:①使用中值滤波消除噪声数据;②对预处理后的数据进行 EMD 分解为 IMF 部分和残差部分;③对步骤②中的每一部分使用 LSTM 方法进行预报,并将预报结果进行叠加。利用 LSTM 和 EMD-LSTM 两种模型分别对 3 个水文指标数据序列在不同预测窗口下进行预测,结果表明:①两种模型下,预测窗口长度越小,均方根误差和平均相对误差越小,预测越准确,曲线拟合越好,即短时段的预测效果优于长时段;②两种模型都可以很好地预测水文数据的变化,但 EMD-LSTM 模型的预测精度更高,特别对波动较大时段,EMD-LSTM 方法优势更为显著。EMD-LSTM 方法更适合于含有高频噪声干扰的水文时间序列。

本文提供的方法可以很好地对水文数据进行短期预报,为水位预判和水资源的实时调度提供决策依据和技术支撑。本研究数据不足 1 a,模型未考虑季节因素,后面随着数据的积累,可以进一步选取季节特征进行研究。

### 参考文献:

[1] KOUTROUMANIDIS T, SYLIAIOS G, ZAFEIRIOU E, et al. Genetic modeling for the optimal forecasting of hydrologic time-series: application in Nestos River [J]. Journal of Hydrology, 2009, 368(1/2/3/4): 156-164.

[2] 徐强,束龙仓,杨桂莲,等. 基于遗传算法优化的波小神经网络在地下水位预测中的应用[J]. 水文, 2010, 30(1): 27-30. (XU Qiang, SHU Longcang, YANG Guilian, et al. Application of optimized wavelet neural network based on genetic algorithm in groundwater level prediction [J]. Journal of China Hydrology, 2010, 30(1): 27-30. (in Chinese))

[3] 张展羽,梁振华,冯宝平,等. 基于主成分-时间序列模型的地下水位预测[J]. 水科学进展, 2017, 28(3): 97-102. (ZHANG Zhanyu, LIANG Zhenhua, FENG Baoping, et al. Groundwater level forecast based on principal component analysis and multivariate time series model [J]. Advances in Water Science, 2017, 28(3): 97-102. (in Chinese))

- [4] YU Z, LEI G, JIANG Z L. ARIMA modelling and forecasting of water level in the middle reach of the Yangtze River [C]//International Conference on Transportation Information & Safety. New York: IEEE, ICTIS, 2017:172-177.
- [5] 李宁,刘瑜,王大为. 基于 ARIMA 和 EEMD 的东江流域季节降水预报研究[J]. 人民珠江, 2019, 40(3): 52-58. (LI Ning, LIU Yu, WANG Dagang. Seasonal precipitation forecast study based on ARIMA and EEMD in the Dongjiang River Basin[J]. Pearl River, 2019, 40(3): 52-58. (in Chinese))
- [6] PORTER D W, GIBBS B P, JONES W F. Data fusion modeling for groundwater systems [J]. Journal of Contaminant Hydrology, 2000, 42: 303-335.
- [7] WEI Chinchang. Wavelet kernel support vector machines forecasting techniques: case study on water-level predictions during typhoons [J]. Expert Systems with Applications, 2012, 39:5189-5199.
- [8] MOOSAVI V, VAFAKHAH M, SHIRMOHAM M, et al. A Wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods [J]. Water Resources Management, 2013, 27:1301-1321.
- [9] ZHONG C, JIANG Z L, CHU X M, et al. Water level forecasting using a hybrid algorithm of artificial neural networks and local Kalman filtering [J]. Journal of Engineering for the Maritime Environment, 2019, 233(1):174-185.
- [10] WANG Y B, SUN T, SU J P, et al. Water level prediction of taocha based on CCS-GBDT model [C]//Computer Supported Cooperative Work and Social Computing. Guilin: Chinese CSCW, 2019:696-708.
- [11] CHANG F J, CHEN P A, LU Y R, et al. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control [J]. Journal of Hydrology, 2014, 517: 836-846.
- [12] 张轩,张行南,江唯佳,等. 秦淮河流域东山站水位预报研究[J]. 水资源保护, 2020, 36(2):41-46. (ZHANG Xuan, ZHANG Xingnan, JIANG Weijia, et al. Study on water level forecast of Dongshan Station in Qinhuai River Basin[J]. Water Resources Protection, 2020, 36(2):41-46. (in Chinese))
- [13] AURÉLIENGÉRON. Hands-on machine learning with scikit-learn and tensor flow [M]. New York: O'Reilly Media, 2017.
- [14] CAO J, LI Z, LI J. Financial time series forecasting model based on CEEMDAN and LSTM[J]. Physica A, 2019, 519:127-139.
- [15] YANG B, SUN S, LI J. Traffic flow prediction using LSTM with feature enhancement [J]. Neurocomputing, 2019, 332: 320-327.
- [16] KIM T Y, CHO S B. Predicting residential energy consumption using CNN-LSTM neural networks [J]. Energy, 2019, 182: 72-81.
- [17] 冯锐. 基于 LSTM 模型的九龙江流域径流序列预测研究[D]. 西安:长安大学, 2019.
- [18] 刘亚新,樊启祥,尚毅梓,等. 基于 LSTM 神经网络的水电站短期水位预测方法[J]. 水利水电科技进展, 2019, 39(2): 56-60. (LIU Yaxin, FAN Qixiang, SHANG Yizi, et al. Short-term water level prediction method for hydropower station based on LSTM neural network [J]. Advances in Science and Technology of Water Resources, 2019, 39(2): 56-60. (in Chinese))
- [19] 郭燕,赖锡军. 基于长短期记忆神经网络的鄱阳湖水位预测[J]. 湖泊科学, 2020, 32(3):865-876. (GUO Yan, LAI Xijun. Water level prediction of Lake Poyang based on long short-term memory neural network [J]. Journal of Lake Sciences, 2020, 32(3): 865-876. (in Chinese))
- [20] 闫佰忠,孙剑,王昕洲,等. 基于多变量 LSTM 神经网络的地下水水位预测[J]. 吉林大学学报(地球科学版), 2020, 50(1): 208-216. (YAN Baizhong, SUN Jian, WANG Xinzhou, et al. Multivariable LSTM neural network model for groundwater levels prediction [J]. Journal of Jilin University (Earth Science Edition), 2020, 50(1):208-216. (in Chinese))
- [21] 陈旭,赵雪花. 基于 EMD 分解的 AR 模型在年径流预测中的应用[J]. 水电能源科学, 2014(7): 14-18. (CHEN Xu, ZHAO Xuehua. Application of autoregressive model to annual runoff forecasting based on empirical model decomposition [J]. Water Resource and Power, 2014(7):14-18. (in Chinese))
- [22] HUANG N E, SHEN Z, LONG S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis [J]. Proceedings Mathematical Physical and Engineering Sciences, 1998, 454: 903-995.

(收稿日期:2019-04-26 编辑:彭桃英)

