Long-term prediction of runoff of Heiher River in China based on EMD-LSTM networks

Heru Xue [1,*], and *Guodong* Wu[1,2]

[1]College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010018, P.R. China
[2]College of Science, Inner Mongolia Agricultural University, Hohhot 010018, P.R. China

Abstract. Due to the influence of global climate change and human activities, the time series data of river runoff become complex and non-stationary. These properties make sequence prediction difficult and with low accuracy. In order to improve the prediction accuracy, Empirical Mode Decomposition (EMD) was introduced to Long Short-Term Memory (LSTM) networks in this paper. The EMD decomposed a non-stationary time series into multiple components. We trained an LSTM network for each component, and added their predictions to obtain the final forecasting value of original sequence. At last, the EMD -LSTM model was applied to the annual runoff sequence in the upper Heihe River. By comparing with single LSTM, it shows that the EMD-LSTM has higher accuracy for the long-term prediction of river runoff.

1 Introduction

Accurate and timely forecast of river runoff is conducive to improving the utilization rate of water resources, to alleviate the shortage of regional water resources, and to prevent droughts and floods in advance. Influenced by meteorological changes such as temperature and precipitation, river runoff shows non-stationarity and complexity [1], which cause large deviations in traditional deterministic prediction processes. Therefore, it is considered the non-stationarity to be processed by signal decomposition technology before forecasting [2]. Artificial neural networks have better advantages for long-term time series forecasting [3,4]. Moreover, the forecast results of combined models are often more accurate than single model [5,6].

LSTM is a typical nonlinear prediction model with good long-term prediction ability and generalization ability [7]. Its application scope covers many fields, such as financial forecasting, traffic flow research, market price fluctuation analysis, etc. In recent years, LSTM has been widely used in the prediction of various hydrometeorological time series, such as precipitation, air temperature, sunshine, wind speed, evapotranspiration, and runoff [8,9]. Moreover, compared with traditional methods such as ARIMA, BP, and RNN, LSTM has great advantages in processing long-term sequences and is suitable for long-term prediction [7,10].

* Corresponding author: xuehr@126.com

When predicting non-stationary time series, it is a feasible way to use signal decomposition technology to decompose the time series into a series of sub-sequences, and then use a certain forecasting method for forecasting analysis. In recent years, the EMD method has been applied to the study of hydrometeorological non-stationarity [11].

In the study, the annual runoff data series from 1800 to 2008 in the upper reaches of Heihe River is taken as the research object. The first 90 % of the sequence (1800-1988) is used for training the model, and the last 10% of the data (1989-2008) is used for model validation. The complexity of the sequence is reduced by EMD decomposition, and then LSTM neural network is performed on the obtained IMFs components. Finally, the combined prediction results are compared with the prediction results of a single LSTM.

2 Methodology

2.1 Empirical mode decomposition

EMD is an adaptive data processing method proposed by Huang in 1998 [12], which is essentially the process of smoothing the data. EMD can decompose the components with different scales from the original signal one by one, to obtain a series of Intrinsic Mode Function (IMF) components containing local feature information. Denoting the original time series data as (t) , the process of the EMD method is as follows:

(1) Let the number of iterations $i = 1$, $d_{i=1}(t) = x(t)$;

(2) Find all maxima and minima of the time series data $d_i(t)$;

(3) Use the local maximum value to perform interpolation calculation to obtain the upper envelope $e_{max}(t)$, and use the local minimum value to perform the interpolation calculation to obtain the lower envelope $e_{min}(t)$;

(4) Calculate the average envelope: $m_i(t) = (e_{max}(t) + e_{min}(t))/2$;

(5) Let $h_i^1(t) = d_i(t) - m_i(t)$, determine whether $h_i^1(t)$ is an IMF. If true, then $IMF_i(t) = h_i^1(t)$, $d_{i+1}(t) = d_i(t) - IMF_i(t)$; else, $d_{i+1}(t) = h_i^1(t)$;

(6) Repeat (2) to (5) until the stopping condition is satisfied, that is, the residual $d_i(t)$ is a monotonic function or the number of extreme points in the envelope is less than or equal to 3.

After above all processes, the original data sequence $x(t)$ is finally decomposed.

$$x(t) = \sum_{j=1}^{N} IMF_j(t) + d_N(t) \tag{1}$$

where, $IMF_j(t)(j = 1,2,\cdots,N)$ is the jth *IMF*, which includes the components from high frequency to low frequency in turn; $d_N(t)$ is the residual term, including the main trend component of original sequence.

2.2 Long short-term memory neural network

LSTM is mainly used for forecasting time series data because of its ability to find more complex nonlinear relationships in the dataset. LSTM overcomes the vanishing and exploding gradient problems and is able to learn long-term dependencies. The LSTM neural network has 3 gating units, which are forget gate, input gate and output gate. Depending on the input gate, the model can decide how much to update itself. These parameters will adaptively adjust the weight of memory and update during the training process to achieve the best performance.

Assuming $\Gamma_f, \Gamma_i, \Gamma_o$ to indicate the forget gate, input gate and output gate respectively and $x^{<t>}$ be the input of the system, the expression corresponding to the LSTM structure is as follows:

$$\Gamma_f = \sigma(W_f.[a^{<t-1>}, x^{<t>}] + b_f) \tag{2}$$

$$\Gamma_i = \sigma(W_i.[a^{<t-1>}, x^{<t>}] + b_i) \tag{3}$$

$$\Gamma_o = \sigma(W_o.[a^{<t-1>}, x^{<t>}] + b_o) \tag{4}$$

$$\tilde{c}^{<t>} = tanh(W_c.[a^{<t-1>}, x^{<t-1>}] + b_c) \tag{5}$$

$$c^{<t>} = \Gamma_i * \tilde{c}^{<t>} + \Gamma_f * c^{<t>} \tag{6}$$

$$a^{<t>} = \Gamma_o * tanh(c^{<t>}) \tag{7}$$

where $c^{<t>}$ and $a^{<t>}$ are the output of the current module and the memory unit interconnected between the modules, and $\sigma(\cdot)$ represents the sigmoid function.

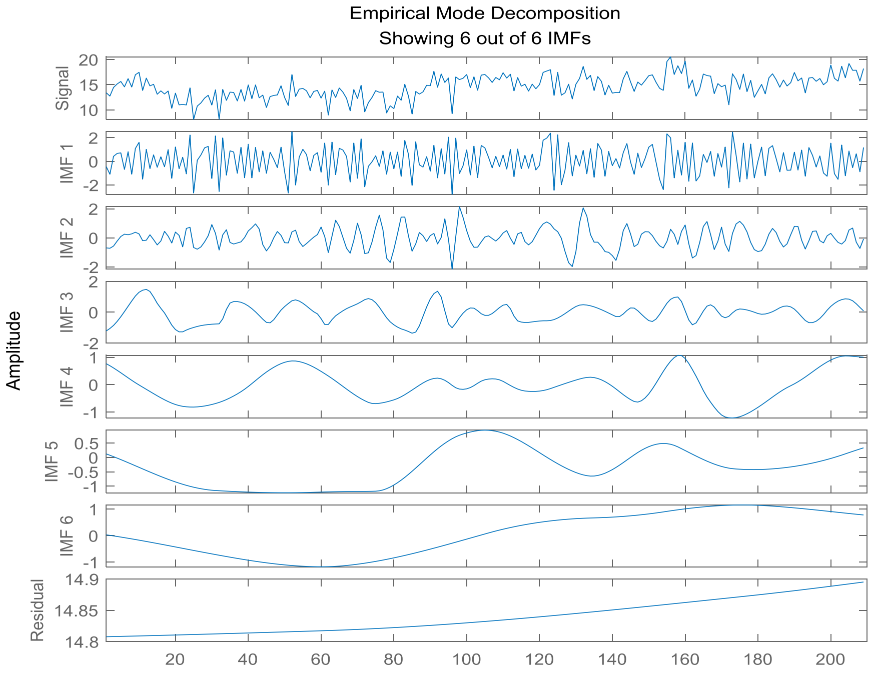

Fig. 1. EMD decomposition results of annual runoff.

3 Model establishment and analysis

3.1 Data

The data used in this paper are reconstructed annual runoff of the upper reaches of the Heihe River from China National Data Center for Glaciology and Geocryology and Desert Science (http://www.ncdc.ac.cn). The time interval of the reconstructed sequence was 1000-2008, and the collection site was in the upper reaches of Heihe River with elevation between 3200.0m and 3600.0m.

The first 190 data, i.e., from year 1800 to 1989, are used for training network, and the rest data from 1990 to 2008 are used for evaluation. The training data was normalized as mean 0 and variance 1.

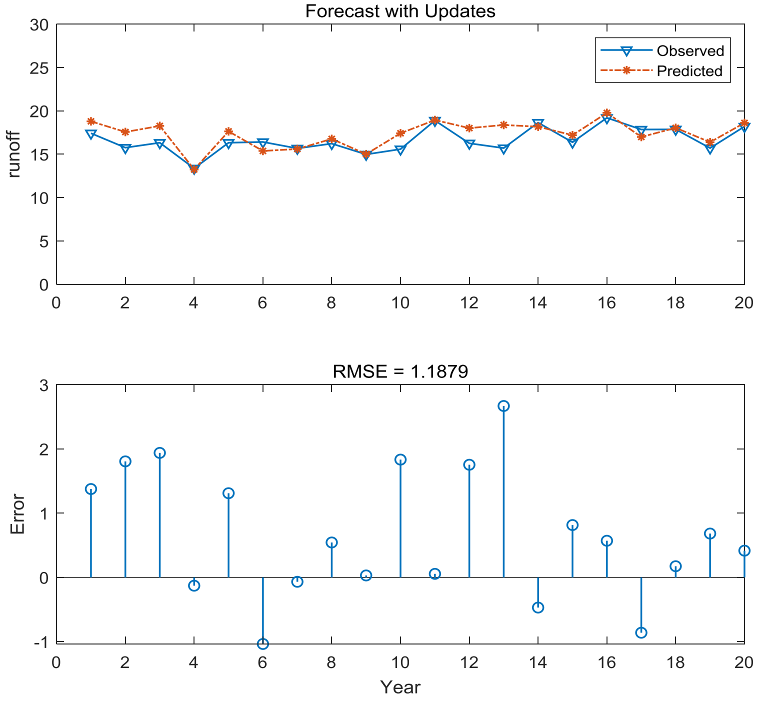

Fig. 2. LSTM prediction results of annual runoff based on EMD.

3.2 EMD-LSTM model

EMD decomposition decomposes the annual runoff time series into 6 IMFs and 1 trend series. The decomposition results are shown in Figure 1. Each decomposed sub-sequence is normalized to construct the LSTM network. The first 90% of every sequence is used for training and the last 10% for evaluation. After repeatedly training the LSTM by adjusting the parameters, the best model is obtained. The detailed parameter settings are recorded in Table 1. The final runoff prediction comes from the sum of all subsequence's predictions (Fig.2).

Table 1. Parameter table of LSTM model for annual runoff based on EMD.

Parameter name	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	Res
numFeatures	1	1	1	1	1	1	1
numResponses	1	1	1	1	1	1	1
numHiddenUnits	80	80	80	80	80	80	80
MaxEpochs	160	160	150	140	100	100	80
GradientThreshold	1	1	1	1	1	1	1
InitialLearnRate	0.02	0.02	0.02	0.02	0.02	0.02	0.02
LearnRateDropPeriod	80	80	70	70	50	50	40
LearnRateDropFactor	0.2	0.2	0.2	0.2	0.2	0.2	0.2

3.3 Evaluation

In order to evaluate the accuracy of EMD-LSTM, we use three evaluation indicators including root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{obs}(i)\text{-}Y_{pre}(i))^2} \tag{8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_{obs}(i) - Y_{pre}(i)\right| \tag{9}$$

$$MAPE = \frac{100\%}{n} \times \sum_{i=1}^{n}\left|\frac{Y_{obs}(i) - Y_{pre}(i)}{Y_{obs,\ i}}\right| \tag{10}$$

where $Y_{obs}(i)$ and $Y_{pre}(i)$ respectively represent the observation and prediction and n represents the data number of test set.

Table 2. Error comparison of annual runoff prediction models in the upper reaches of Heihe River.

Algorithm	RMSE	MAE	MAPE
LSTM	2.0449	1.5258	0.0890
EMD-LSTM	1.1879	0.9272	0.0566

It can be seen from Table 5 that the RMSE, MAE and MAPE of the EMD-LSTM prediction model are 1.1879, 0.9272 and 0.0566 respectively, while the RMSE, MAE and MAPE of the LSTM model are 2.0449, 1.5258 and 0.089 respectively. The results show that the EMD-LSTM is better than LSTM.

4 Results

In this paper, we construct an EMD-LSTM model for prediction of river runoff. Applied to the record of upper reaches of the Heihe River, the EMD-LSTM shows better performance than single LSTM. Therefore, the EMD method improves the prediction performance of LSTM. This method can effectively deal with the non-stationarity of nonlinear time series and improve its prediction accuracy.

This work was financially supported by Research Program of science and technology at Universities of Inner Mongolia Autonomous Region, P.R. China (grant number: NJSY21477)

References

1. H. Maier; A. Jain, G. Dandy, K.P. Sudheer, Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. Environmental Model. Softw. **25**, 891 (2010).
2. Q. Tao, Y.U. Chuanjin. Wind speed forecasting for bridge sites based on empirical mode decomposition and Elman neural network. Journal of Catastrophology, **32**, 85 (2017).
3. V. Nourani, G. Andalib, F. Sadikoglu, E. Sharghi, Cascade-based multi-scale AI approach for modeling rainfall-runoff process. Hydrology Research, **49**, 1191 (2018).
4. D.W. Cui. Application of multi-hidden layer BP neural network models in runoff prediction. Hydrology, **33**, 68 (2013).
5. Z. Qiang, B.D. Wang, B. He, et al. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. Water Resources Management, **25**, 2683 (2011).

6. C.T. Cheng, W.J. Niu, Z.K. Feng, et al. Daily reservoir runoff forecasting method using artificial neural network based on quantum-behaved particle swarm optimization. Water, **7**, 4232 (2015).
7. F. Kratzert, D. Klotz, C. Brenner, et al. Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrology and Earth System Sciences, **22**, 6005 (2018).
8. Z. Karevan, A.K. Johan. Transductive LSTM for time-series prediction: An application to weather forecasting. Neural Networks, **125**, 1 (2020).
9. Z.Xiang, J. Yan, I. Demir. A rainfall-runoff model with LSTM-based sequence to sequence learning. Water Resources Research, **56**, 1 (2020).
10. N.S. Siami, N. Tavakoli, A.S. Namin. A Comparison of ARIMA and LSTM in Forecasting Time Series. *17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. (IEEE, 2018).
11. Y. Zhang, Y. Li, M. Yang, et al. Multi-scale nonlinear runoff prediction model based on EMD. International Journal of Earth Sciences and Engineering, **8**, 552 (2015).
12. N.E. Huang, Z. Shen, S.R. Long, et al. The empirical mode decomposition method and the Hilbert spectrum for non-stationary time series analysis. *Proceedings of the Royal Society of London*. **454**, 903 (1998).