The Impact of LSTM Network Structure on Flood Process Simulation and Forecast Period: A Case Study of the Middle Yellow River

*Caihong Hu[1]

1. school of water conservancy science and engineering, zhengzhou university, Zhengzhou city, Henan Province, P.R.China, 450001

Flood forecasting is an important non-engineering measure for flood prevention and disaster reduction. Recently, owing to the breakthrough in the field of computational science, neural network technology based on deep learning have received a growing interest from scientists. The Long Short-Term Memory (LSTM) neural network, one of the most advanced applications of deep learning technology, is applied to rainfall runoff simulation in the hydrological field. Most previous studies focused on verifying the feasibility of LSTM in hydrological forecasting. This paper explores the effects of LSTM network structure on flood process simulation and forecast period by constructing LSTM models with different network structures. The hydrological situation in the middle reaches of the Yellow River has changed significantly in the past 50 years. Jingle station control basin in the upper reaches of the Fenhe River is selected as the case study, which including 98 flood events from 1956 to 2014. The datasets of 1954-2003 was selected as the training sets and the dataset of 2003-2014 was selected as the verified set. The results show that, on the one hand, the prediction effect decreases with the increase of the foresight period. The LSTM simulation effect is the best when the foresight period is 1h. The number of cells and the cycle index are 120 and 200 respectively. The Nash-Sutcliffe efficiency (NSE) coefficient reaches 0.9958, Root Mean Square Error (RMSE) reaches 10.1595, BIAS reaches 0.5575%, and the allowable errors of flood peak forecast and peak time forecast were within 5% and 1h, respectively. At this time, LSTM has a perfect performance in flood forecasting; when the forecast period is 1-6h, the number of cells and the cycle index are 140 and 300 respectively, NSE reaches 0.8857, RMSE reaches 39.0867, BIAS reaches 10.5886%, and the allowable errors of flood peak forecast and peak time forecast are within 10% and 6h, respectively, which meet the allowable errors of flood forecast. When the forecast period exceeds 6h, NSE is lower than 0.8, RMSE is higher than 50, BIAS is higher than 15%, and the forecast error of peak time exceeds the length of a calculation period, and it does not have the ability to forecast floods. On the other hand, as far as the forecast period is selected as 1h, both the number of cells and the cycle index in the network structure are positively correlated with simulation accuracy. At first, when the network structure is simple which has 20 cells and 50 cycles index, the NSE is not higher than 0.5. As the number of cells and cycle index increase at the same time, the simulation accuracy increases rapidly. The NSE stabilizes until the number of neurons and the number of iterations reach 120 and 200, respectively. The reason is that the more cells cause the model more complex which can learn more features of data. By increasing the cycle index, the more the model updates the parameters, the better the simulation effect. In summary, LSTM can be used for short-term flood forecasting in the middle reaches of the Yellow River. It is recommended to use LSTM for flood process simulation when the forecast period is selected as 1h. The effect is best when the number of cells and the cycle index are selected as 120 and 200, respectively.

Keywords: LSTM;, network structure; , rainfall-runoff simulation; , middle Yellow River

