International Journal of Social Science and Education Research Volume 3 Issue 4, 2020
ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

Research on Short-term Tide Forecast Based on Bi-LSTM Recurrent Neural Network

Guiqiang Zhu[1, a]

[1]Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China.

[a]zhgqwnx@163.com

Abstract

In order to improve the accuracy of tidal forecasting, this paper applies the bi-directional long short-term memory (Bi-LSTM) recurrent neural model to tidal forecasting for the first time. A high-precision Bi-LSTM recurrent neural network tide forecasting model is established, and the tidal water level corresponding to subsequent time points is predicted based on the tidal water level in the first few hours. Taking the Isabel port as an experimental object, the real-time tide forecast simulation was performed using the measured tide data of the Isabel port. Simulation results show: Compared with the traditional harmonic analysis method, BP neural network model, RBF neural network model, traditional RNN model, and LSTM recurrent neural network model, the accuracy of using Bi-LSTM recurrent neural network tide prediction model to predict the tides of the port increased by 99.0%, 89.9%, 93.8%, 79.8%, 29.5% respectively.

Keywords

Tide forecast,Deep learning,Bi-LSTM recurrent neural network,Harmonic analysis.

1. Introduction

Ocean tides are long-period fluctuations caused by celestial tidal forces. Tide forecasting is of great significance in national defense construction, transportation and shipping, marine resource development, energy utilization, environmental protection, harbour construction and coast protection, etc. The high requirements for safety also put forward higher requirements for the accuracy of real-time ocean tide numerical prediction.

The traditional tide prediction model mainly uses the harmonic analysis method for tide prediction. Darwin (Darwin G.H., 1883-1886) performed tide harmonic analysis on one month's tidal data, calculated the tide harmonic constant, and used it for tide prediction; Dudson (1928) proposed a harmonic analysis method to analyze 60 tides using one year's tide data; Dudson (1954) also proposed a harmonic analysis method for analyzing 60 tides using 29-day tide data. Since the 1960s, computers have replaced complicated manual calculations, which has resulted in a series of rigorous scientific analysis methods. In the tide forecast, the results of one year's tide data analysis have been used instead. Horn W. published a method of tide analysis using an electronic computer in 1960. Cartwright et al. Proposed the Fourier analysis method in 1963. In 1966 Van Ette proposed the least square method of tidal analysis. Franco (1993) proposed a method of tidal analysis using fast Fourier transform. He first performed Fourier tide expansion on the tidal data. After correction, he calculated the harmonic constant of the tidal tide development and carried out tide prediction. After hundreds of years of analysis and research, the harmonic analysis method has been widely used in tide forecasting, but this method only considers the astronomical tide affected by the gravitational influence of the sun and the moon, and ignores the non-effects caused by environmental factors such as weather, climate, and bottom terrain Astronomical tide parts are prone to large errors.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

With the rapid development of artificial intelligence, some neural network models have also been applied to tidal prediction, and have been well applied in the field of tidal prediction. TSAI[1] used artificial neural network models for the first time to predict whole-day and half-day tide; LIN [2] proposed an adaptive neural fuzzy inference system for sea level prediction; JAIN [3] used a neural network to establish a tide prediction model with a prediction step of 24 h, and applied this to the New Mangalore tidal station; Zhang Anran [4], Zhang Zeguo [5] and others have used BP neural network for tide prediction, and obtained better models than traditional methods; Zhou Tao [6] and others used the RBF neural network for tide prediction, and also obtained relatively ideal results. In recent years' research, BP neural network and RBF neural network are widely used in prediction because of their advantages such as easy implementation and strong non-linearity. However, their models are poorly global and easily fall into the defect of local optimization.

In order to improve the accuracy of tide forecast, F. Salcedo [7] first applied LSTM [8] recurrent neural network to the simulation of the coast conditions of San Francisco port in 2019 to supplement or replace the traditional hydrodynamic modeling work. To predict the tide of the San Francisco port in the United States, good results have been achieved. But whether it is BP neural network, RBF neural network or LSTM recurrent neural network, it is one-way propagation. It can only encode one-way information from front to back or back to front. It cannot encode information from front to back and back to front. , resulting in unsatisfactory model accuracy[9].

Based on this, based on previous research, this paper applies the bi-directional long-term memory (Bi-LSTM) recurrent neural network [9] to the tide forecast for the first time, and establishes a Bi-LSTM recurrent neural network tide model. The prediction model uses the actual tide data of the Isabel port in the United States for prediction simulation. The results show that the model is feasible and effective, and the prediction accuracy is greatly improved.

2. Recurrent Neural Network Principles

2.1. Recurrent Neural Network (RNN) Basic Principles

The reason why RNN is called recurrent neural network is that the current output of a sequence is also related to the previous output. The specific manifestation is that the network memorizes the previous information and applies it to the current output calculation, that is, the nodes between the hidden layers are no longer connected but connected, and the input of the hidden layer includes not only the output of the input layer It also includes the output of the hidden layer from the previous moment. It mainly consists of an input layer, a hidden layer, and an output layer, and a loop unit is introduced into the hidden layer. The structure of RNN is shown in Figure 1.

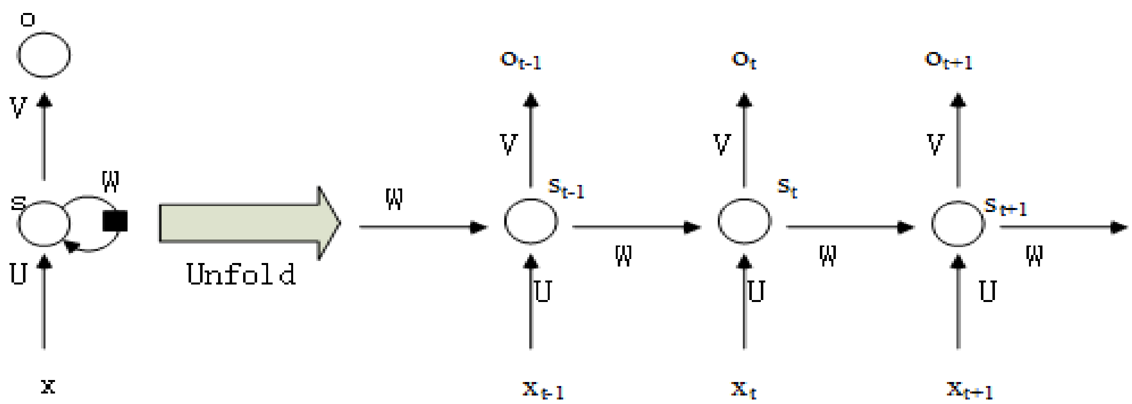

Fig 1. Circulating neural network development map

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

In Figure 1, x, s, and o are all vectors, represent the values of the input layer, the hidden layer, and the output layer, respectively; U and V respectively represent the weight matrix from the input layer to the hidden layer and the weight matrix from the hidden layer to the output layer; The biggest difference between a recurrent neural network and an ordinary neural network is that there are more recurrent units (the square arrow at the left letter W in the expanded view). The value of the hidden layer of the recurrent neural network at time t is not only related to the input x_t at time t but also The last time value of the hidden layer is related to St-1, W is the value of the hidden layer at the previous time as the weight of the hidden layer input at this time. The mathematical expression of the recurrent neural network is as follows:

$$o_t = g(Vs_t) \tag{1}$$

$$s_t = f(Ux_t + Ws_{t-1}) \tag{2}$$

g and f are activation functions; o_t is the calculation formulas for the output layer; s_t is the calculation formulas for the hidden layer.

If equation (2) is repeatedly substituted into equation (1), then:

$$\begin{aligned} o_t &= g(Vs_t) \\ &= g(Vf(Ux_t + Ws_{t-1})) \\ &= g(Vf(Ux_t + Wf(Ux_{t-1} + Ws_{t-2}))) \\ &= g\big(Vf(Ux_t + Wf(Ux_{t-1} + Wf(Ux_{t-2} + Wf(x_{t-3} + \ldots))))\big) \end{aligned} \tag{3}$$

It can be seen from the formula (3) that the output value o_t of the RNN is affected by the historical input information x_t, x_{t-1}, x_{t-2}, ... Therefore, when using recurrent neural network prediction, the correlation and dependence between time series data can be described.

2.2. Long Short-Term Memory (LSTM) Network Basic Principles

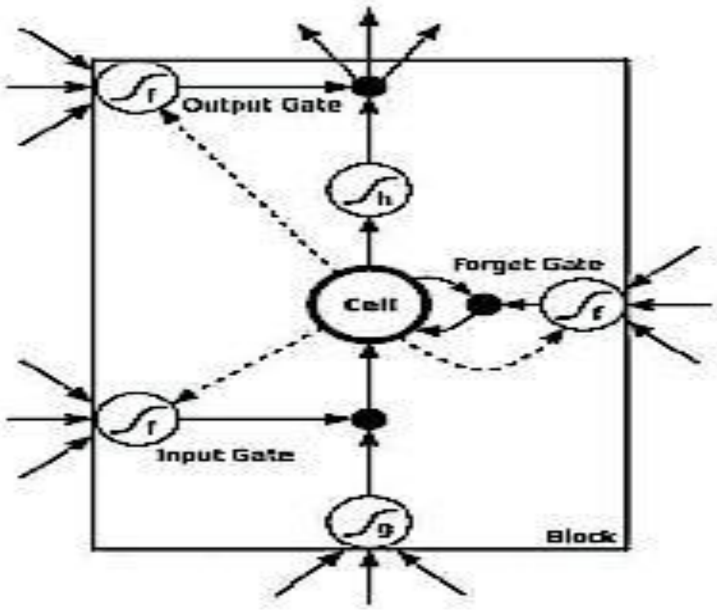

Fig 2. LSTM neural network structure

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

The LSTM network is an improved RNN network, which adds input gate, output gate, forget gate and cell state. Among them, the input gate and output gate control the inflow and outflow of information flow. The forget gate is used to select the previous moment. How much of the unit state is saved up to this moment. Figure 2 shows the network structure of LSTM.

Compared with the traditional RNN network, at time t, the LSTM network propagates forward in addition to the state h^t of the hidden layer, and there is more unit state C^t. The role of the forget gate is to control the state of the hidden layer at the previous moment with a certain probability matrix. The output f^t of the forgetting gate is determined by the hidden layer state h^{t-1} at the previous moment and the input x_t at the current moment:

$$f^t = \sigma(W_f h^{t-1} + U_f x^t + b_f) \tag{4}$$

Among them, W_f, U_f and b_f are weights and bias matrices related to forgetting gate, and σ is sigmod function.The function of the input gate is to control the input at the current moment. It consists of two parts:

$$i^t = \sigma(W_i h^{t-1} + U_i x^t + b_i) \tag{5}$$

$$a^t = \tanh(W_a h^{t-1} + U_a x^t + b_a) \tag{6}$$

Among them, W_i, W_a, U_i, U_a, b_i and b_a are weights and bias matrices related to the input gate.

The update of the cell state Ct in the LSTM consists of the results of the input gate and the forget gate:

$$C^t = C^{t-1} \odot f^t + i^t \odot a^t \tag{7}$$

⊙is the product of Hadamama.

Finally, the output of the output gate in LSTM is:

$$o^t = \sigma(W_o h^{t-1} + U_o x^t + b_o) \tag{8}$$

The update of the hidden layer h^t is:

$$h^t = o^t \odot \tanh(C^t) \tag{9}$$

2.3. Bidirectional Long Short-Term Memory Network (Bi-LSTM) Basic Principles

In order to overcome the limitations of the traditional recurrent neural network (RNN), a bi-directional long short-term memory (Bi-LSTM) recurrent neural network was first proposed in 1997 by Schuster, M [9]. Bi-LSTM recurrent neural network can process the input sequence in forward and backward directions, each direction has independent parameters, and its prediction accuracy for time series is better than traditional LSTM recurrent neural network. The structure of Bi-LSTM recurrent neural network is shown in Figure 3:

ISSN: 2637-6067 | DOI: 10.6918/IJOSSER.202004_3(4).0004

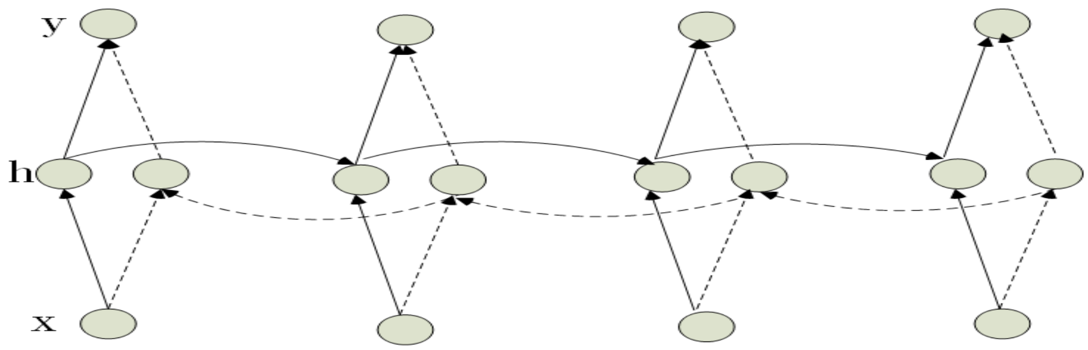

Fig 3. Structure of Bi-LSTM recurrent neural network

In the figure, x, h, and y represent the input layer, the hidden layer, and the output layer, respectively.

$$\overrightarrow{h}_t = f\left(\overrightarrow{W} x_t + \overrightarrow{V}\,\overrightarrow{h}_{t-1} + \overrightarrow{b}\right) \tag{10}$$

$$\overleftarrow{h}_t = f\left(\overleftarrow{W} x_t + \overleftarrow{V}\,\overleftarrow{h}_{t+1} + \overleftarrow{b}\right) \tag{11}$$

$$y_t = g\left(U\left[\overrightarrow{h_t}\;;\overleftarrow{h_t}\right] + c\right) \tag{12}$$

$\overrightarrow{h}_t$ is the forward calculation of the hidden vector, $\overleftarrow{h}_t$ is the backward calculation of the hidden vector, and y_t is the final output result.

3. Construction and prediction of Bi-LSTM-RNN model for short-term tide forecast

3.1. Data Extraction and Preprocessing

All raw tide level data used in this paper are downloaded from the National Oceanic and Atmospheric Administration (NOAA) website http://co-ops.nos.noaa.gov. In this paper, a total of 1,416 sets of actual observed tidal water level data from the GMT0000 of January 1, 2018 to GMT2400 of February 28, 2018 in the Port of Isabel, USA were used for model learning training and simulation fitting test. The model uses the first 1200 sets of measured tidal water level data for learning and training, and then uses the remaining 216 sets of measured tidal water level data as verification data for simulation prediction experiments.

3.2. Bi-LSTM Recurrent Neural Network Model Construction and Model Parameter Setting

This experiment uses the first 1416 sets of data in all the data as the training set. The neural network has 24 input layers, 20 hidden layers, and one output layer. Time_step is 24, which corresponds to the past 24 time points. The actual measured tidal water level data is used to predict the tidal water level at the next time point. The number of trainings is 5000. The

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

batch_size is 72. That is, 72 sets of data are taken for training each time. A neural network model is obtained through training.

In order to enhance the nonlinearity of the model, the activation function of the Bi-LSTM recurrent neural network selects relu. In the selection of the loss function and the optimization method, it is necessary to consider both the model convergence speed and the accuracy of the prediction. Time series prediction models generally choose Mean Squared Error (MSE) as the loss function, and choose adaptive learning optimization algorithm (adam) [10] as the model optimization method, which can achieve excellent results quickly.

During the prediction phase, according to the current time point, the tidal height value corresponding to the nearest whole point forward is obtained, and the real value of the berth demand corresponding to the first n hours is X_{T-n}, X_{T-n-1}, ... X_{T-2}, X_{T-1}. According to X_{T-n}, X_{T-n-1}, ..., X_{T-2}, X_{T-1}, X_T, use the trained Bi-LSTM recurrent neural network model to predict the next 1 hour, and obtain the corresponding tidal height value X_{T+1} in the next hour. Then, based on the measured tidal water level data corresponding to the first n-1 hours and the predicted tidal water level data corresponding to the T + 1 time, the tidal water level data corresponding to the T + 2 time is predicted to obtain X_{T+2}. The tidal height value corresponding to the hour is X_{T+1}, X_{T+2}, ... X_{T+m-1}, X_{T+m}. In many experiments, it was found that when n is 24, a better result can be obtained.

3.3. Determination of Evaluation Indicators

In order to better show the performance of the Bi-LSTM recurrent neural network in predicting the tide level, this article uses the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean squared error (RMSE) error) These three evaluation indicators perform comprehensive analysis and evaluation of the prediction performance of different neural networks [11] [12]. Among them, the mean absolute error represents the mean of the absolute value of the prediction error, which reflects the accuracy; the mean absolute percentage error is an indicator that reflects the overall effectiveness of the prediction method; and the root mean square error reflects the degree of error between the predicted value and the actual value. These three evaluation indicators are as follows (13), (14), (15):

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |t_{fi} - t_{mi}| \tag{13}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} |\frac{t_{fi} - t_{mi}}{t_{mi}}| \tag{14}$$

$$RMSE = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n} (t_{fi} - t_{mi})^2\right)} \tag{15}$$

t_f is the measured value of the tidal water level; t_m is the predicted value of the tidal water level.

3.4. Forecast Results and Analysis

After the neural network structure is determined, this paper uses Python simulation to implement the Bi-LSTM recurrent neural network. At the same time, it uses the 1,200 sets of tidal water level measured data of the Isabel port in the United States to complete the network training. The tidal water level of Port GMT0000 on February 20, 2018-GMT2400 on February

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

28, 2018 is predicted, and the actual data is used for error calculation. At the same time, in order to compare the prediction performance of the Bi-LSTM recurrent neural network and different methods, this paper uses the same training data and prediction data to use the harmonic analysis method, BP neural network, RBF neural network, traditional RNN neural network, and LSTM recurrent neural network for tide. Forecast of water level. The prediction results of the six models and the corresponding prediction errors are shown in Figures 4, 5, 6, 7, 8, 10, and 10.

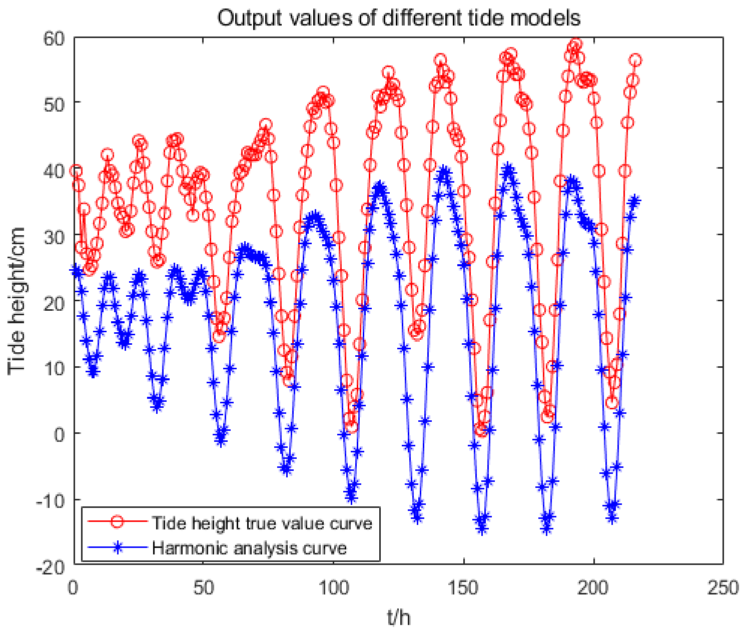

Fig 4 Comparison of Tide Height Predicted by Harmonic Analysis and Measured Tide Height

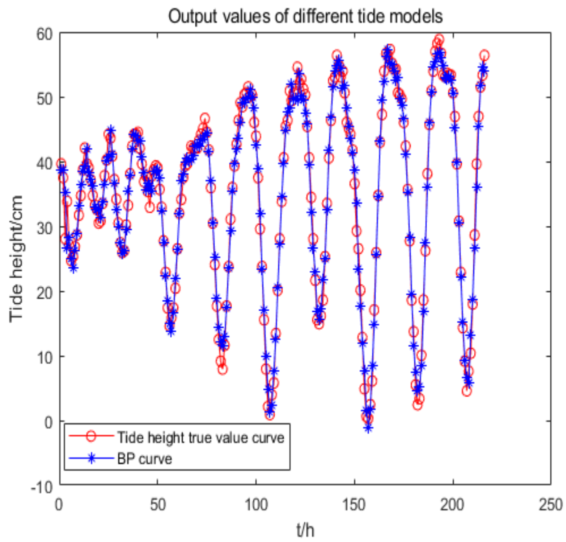

Fig 5. Comparison of tidal height predicted by BP neural network and measured tidal height

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

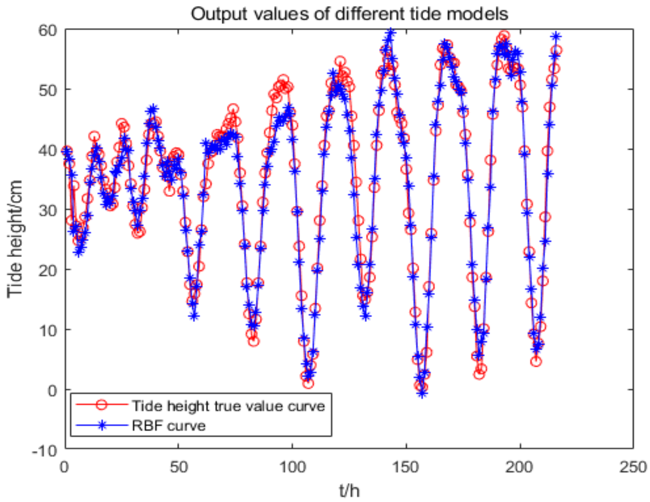

Fig 6. Comparison of tidal height predicted by RBF neural network and measured tidal height

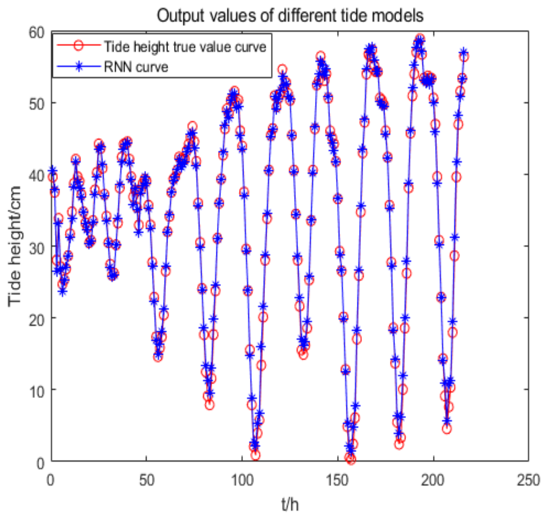

Fig 7. Comparison of tidal height predicted by RNN and measured tidal height

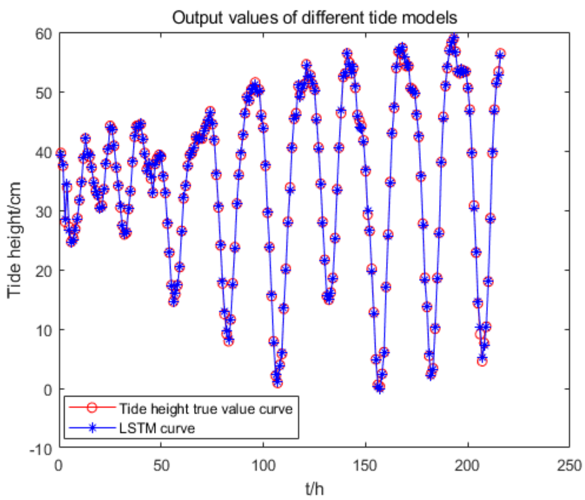

Fig 8. Comparison of tide height predicted by LSTM recurrent neural network and measured tide height

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

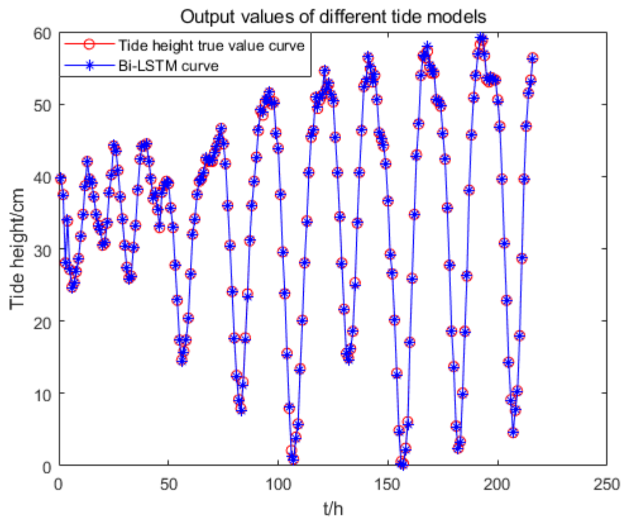

Fig 9. Comparison of tide height predicted by Bi-LSTM recurrent neural network and measured tide height

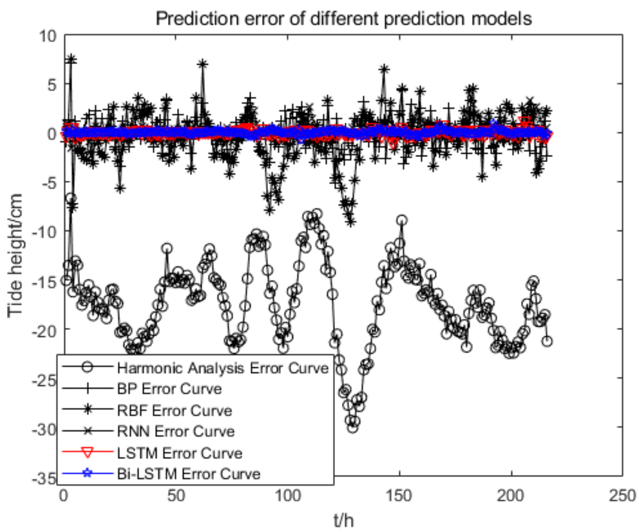

Fig 10. Comparison of prediction errors of different prediction models

It can be seen from Figures 4, 5, 6, 7, 8, 9, 10: harmonic analysis method, BP neural network, RBF neural network, RNN, LSTM cycle Compared with the neural network and Bi-LSTM recurrent neural network, the tidal water level prediction result of Bi-LSTM recurrent neural network is obviously closer to the actual value; meanwhile, the prediction result has less fluctuation, which proves that it has better stability.

In order to compare the prediction performance of different prediction models more deeply, based on the tidal water level prediction values and prediction errors of these six prediction models, this paper calculates these three evaluation indicators. The specific data is shown in Table 1.

ISSN: 2637-6067
DOI: 10.6918/IJOSSER.202004_3(4).0004

Evaluation index	Five prediction models					
	Harmonic analysis	BP	RBF	RNN	LSTM	Bi-LSTM
MAE	17.421	1.293	2.061	0.630	0.180	0.117
MAPE	1.112	0.098	0.119	0.078	0.017	0.014
RMSE	17.919	1.689	2.743	0.841	0.241	0.170

Table 1. Comparison of evaluation indicators of different prediction models

Table 1 shows that the root mean square errors of the harmonic analysis model, BP model, RBF model, RNN model, LSTM model, and Bi-LSTM recurrent neural network are: 17.919, 1.689, 2.743, 0.841, 0.241, 0.170. The prediction accuracy of the Bi-LSTM model was improved by 99.0%, 89.9%, 93.8%, 79.8% and 29.5% over the harmonic analysis model, BP model, RBF model, RNN model, and LSTM model respectively. The results show that the Bi-LSTM recurrent neural network tide forecasting model has better forecasting ability.

4. Conclusion

In this paper, Bi-LSTM recurrent neural network is used for the first time to predict tidal height. Based on the historical port data, the data is first reconstructed in time series, and then a Bi-LSTM recurrent neural network with a certain structure is established, and then the network is trained, predicted, and tested. Modeling is based on data from the Port of Isabel, USA Simulation and results analysis. The analysis results show that, compared with the traditional harmonic analysis method, BP neural network model, RBF neural network model, traditional RNN model, and LSTM recurrent neural network model, the accuracy of using Bi-LSTM recurrent neural network tide prediction model to predict the tides of the port increased by 99.0%, 89.9%, 93.8%, 79.8%, 29.5% respectively, the tidal prediction model based on the Bi-LSTM recurrent neural network has higher accuracy and can well solve the problem of low prediction accuracy of the current tide model, and has a wide application prospect in tide prediction. However, the model only extracts temporal features for modeling, but does not extract spatial features. Considering that the tide is affected by both time and space, it is the main research direction to extract the temporal and spatial features of the sequence for tide modeling and prediction.

Acknowledgements

Natural Science Foundation.

References

[1] TSAI Ching-Piao,LIN Chang,SHEN Jia-N.Neural network for wave forecasting among multi-stations[J].Ocean Engineering,2002,29(13):1683-1695.DOI:10.1016/S0029-8018(01)00112-3.

[2] LIN Li-Ching,CHANG Hsien-Kuo.An adaptive neuro-fuzzy inference system for sea level prediction considering tide-generating forces and oceanic thermal expansion[J].Terrestrial Atmospheric & Oceanic Sciences,2008,19(1/2):163-172.DOI:10.3319/TAO.2008.19.1-2.163(SA).

[3] JAIN P,DEO M C.Real-time wave forecasts off the western Indian coast[J].Applied Ocean Research,2007,29(1/2):72-79.DOI:10.1016/j.apor-2007.05.003.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202004_3(4).0004

[4] Zhang Anran. Modular tide forecast based on BP neural network [D] .Dalian Maritime University, 2015.

[5] Zhang Zeguo, Yin Jianchuan, Liu Cheng, et al.Real-time port tide forecast based on SAPSO-BP network model [J] .Water Transport Engineering,2017(1):34-40.DOI:10.16233/j.cnki.issn1002-4972.2017.01.006.

[6] Zhou Tao. Design and implementation of tide forecasting system based on improved RBF network [D] .Dalian Maritime University, 2016.

[7] F. Salcedo, C. Harter, S. Fenical. Predicting Long-Term Coastal Conditions in San Francisco Bay and Other Estuaries with the Use of Supervised Neural Networks[J]. Ports 2019: Port Planning and Development - Papers from Sessions of the 15th Triennial International Conference, p 101-111, 2019.

[8] Hochreiter S, Schmidhuber J. Long short-term memory.[J]. Neural Computation, 1997,9(8):1735-1780.

[9] Schuster, M., Paliwal, K. K.. Bidirectional recurrent neural networks[J]. IEEE Trans. Signal Process. 45 (11), 2673–2681 (1997).

[10] KINGMA, DIEDERIK P. and BA, JIMMY. Adam:A Methodfor Stochastic Optimization [EB/OL]. arXiv: 1412.6980, December 2014.

[11] MADSEN H, PINSON P, KARINIOTAKIS G. Standardizing the performance evaluation of short term wind power prediction models[J].Wind Engineering, 2005, 29(6): 475-489.

[12] DAI Lang, HUANG Shoudao, HUANG Keyuan, et al. Combination forecasting model based on neural networks for wind speed in wind farm[J]. Proceedings of the CSU-EPSA, 2011, 23(4): 27-31.