

# A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network

GAO Song<sup>1,2</sup>, ZHAO Peng<sup>1,2</sup>, PAN Bin<sup>3</sup>, LI Yaru<sup>1,2\*</sup>, ZHOU Min<sup>3</sup>, XU Jiangling<sup>1,2</sup>, ZHONG Shan<sup>1,2</sup>, SHI Zhenwei<sup>3</sup>

<sup>1</sup>North China Sea Marine Forecasting Center of State Oceanic Administration, State Oceanic Administration, Qingdao 266061, China

<sup>2</sup>Shandong Provincial Key Laboratory of Marine Ecological Environment and Disaster Prevention and Mitigation, Qingdao 266061, China

<sup>3</sup>Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, China

Received 16 July 2016; accepted 16 August 2017

© Chinese Society for Oceanography and Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

It is of vital importance to reduce injuries and economic losses by accurate forecasts of typhoon tracks. A huge amount of typhoon observations have been accumulated by the meteorological department, however, they are yet to be adequately utilized. It is an effective method to employ machine learning to perform forecasts. A long short term memory (LSTM) neural network is trained based on the typhoon observations during 1949–2011 in China's Mainland, combined with big data and data mining technologies, and a forecast model based on machine learning for the prediction of typhoon tracks is developed. The results show that the employed algorithm produces desirable 6–24 h nowcasting of typhoon tracks with an improved precision.

**Key words:** typhoon tracks, machine learning, LSTM, big data

**Citation:** Gao Song, Zhao Peng, Pan Bin, Li Yaru, Zhou Min, Xu Jiangling, Zhong Shan, Shi Zhenwei. 2018. A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network. *Acta Oceanologica Sinica*, 37(5): 8–12, doi: 10.1007/s13131-018-1219-z

## 1 Introduction

Typhoons are typical tropical weather systems, which impose severe threats to people's lives, property security and the development of the regional economics in the coastal areas. Timely prediction and warnings of a typhoon can provide effective information support to disaster prevention departments, and reduce injuries and economic losses effectively. Therefore, it is an important research topic to accurately predict typhoon tracks. However, typhoon tracks are influenced by the typhoon background fields, the thermodynamics and kinetics of the typhoon system, etc. (Huang and Jin, 2013). In addition, after the landing of a typhoon, its tracks are influenced by the complex bathymetry and coastline of the coastal region and the topography of the inland, etc. (Yu et al., 2012). It is thus a rather complex and integrated challenge to resolve the typhoon tracks.

The existing forecasting operations mainly rely on the subjective empirical forecasts in the early stage. Chen and Ding (1979) summarized the track types of the typhoon happened in the Northwest Pacific, which provided an empirical reference to the prediction of typhoon tracks. With the development of monitoring methods and computer technology, numerical forecasts of typhoon tracks have been widely developed in recent years. Currently, the forecasting of typhoon tracks has formed a comprehensive system utilizing a wide range of data and methods, which

is based on numerical forecasting and combined with human machine interface (Qian et al., 2012). Wang et al. (1996) proposed a scheme for the initial field for the numerical forecasting of typhoon based on artificial data and observations, and applied it to the forecast of typhoon tracks in the South China Sea. Qian et al. (2012) discussed the influence of different initial field and lateral boundary conditions on the accuracy of typhoon numerical forecasts. Li and Chen (2002) reviewed the operating application of the ensemble numerical forecast system in China. Xu et al. (2014) proposed a typhoon forecasting strategy based on GRAPES with an improved convective parameterization. However, although numerical forecasts have been widely applied, the accuracy is still lower than that of empirical forecasting methods (Chen et al., 2015). According to the published data by Shanghai Typhoon Institute of China Meteorological Administration, the current forecasting status of 24 h, 48 h and 72 h typhoon track errors based on subjective empirical methods are 84.2 km, 145.6 km and 205.4 km, respectively, and those based on regional numerical models are 97.4 km, 188.2 km and 302.7 km, respectively (Xu et al., 2010). It is still challenging to further improve the accuracy of the numerical forecasting models.

Currently, a three-dimensional typhoon observation system has been established preliminarily, including meteorological satellites, oceanic observation stations, and ground observation

Foundation item: The National Natural Science Foundation of China under contract Nos 61273245 and 41306028; the Beijing Natural Science Foundation under contract No. 4152031; the National Special Research Fund for Non-Profit Marine Sector under contract Nos 201405022-3 and 2013418026-4; the Ocean Science and Technology Program of North China Sea Branch of State Oceanic Administration under contract No. 2017A01; the Operational Marine Forecasting Program of State Oceanic Administration.

\*Corresponding author, E-mail: liyaru@bhj.gov.cn

posts, which provides numerous data to operating departments and research institutes (Hochreiter and Schmidhuber, 1997). In addition, an enormous number of typhoon observations have been accumulated by typhoon forecasting institutes since 1949. In the era of big data, it is an urgent task to improve the numerical forecasting accuracy by utilizing these data. With the development of machine learning algorithms, especially recurrent neural networks (RNNs), long and short term memory neural networks (LSTM), etc., new ways to solve prediction and regression issues have become available. Inspired by the outstanding performance of LSTM models in image recognition and visual descriptions (Xu et al., 2015), a forecasting model for typhoon tracks is developed based on a deep learning algorithm in this paper, with a large number of observation samples to train the model.

The essence of deep learning algorithms is using a large number of samples to train an end-to-end network and then performing forecasts using test data by this trained network. As there are many parameters involved in the network, sufficient training data are needed to train the model. Research shows that deep learning algorithms can improve the accuracy remarkably with sufficient samples to train the network. A reasonable employment of a deep neural network to deal with the large number of big data can improve the forecast precision of typhoon tracks. RNNs are one of the most widely used regression prediction neural network algorithms. However, the gradient in RNN may disappear, i.e., errors cannot propagate backward to a far-away previous neuron (Donahue et al., 2015). The LSTM is a classic solution to this issue. In LSTM model, the normal neurons containing activation functions are replaced by a memory cell to resolve the disappearance of the gradient. Meanwhile, values of information can be stored in the network within any length of time due to this scheme. Ranzato et al. (2014) proposed a video interpretation algorithm based on the LSTM. Sutskever et al. (2014) employed the LSTM algorithm to predict a time series, and obtained reasonable results. Studies show that deep learning algorithms, especially LSTM, can be employed in weather forecasting and ocean remote sensing areas. Yin et al. (2015) predicted the atmosphere pollution levels using a deep belief network. You et al. (2016) adopted a traditional error back-propagation neural network based on phase space reconstruction to predict storm surges. Shi et al. (2015) proposed a prediction model for precipitation based on a convolutional LSTM algorithm. The aforementioned studies indicate that deep learning algorithms are effective tools to establish models and perform predictions using big data.

A prediction model for typhoon tracks is proposed in this paper based on a deep learning algorithm. The observations of typhoon tracks from 1949 to 2012 provided by the North China Sea Marine Forecasting Center of State Oceanic Administration

of China are used to train the LSTM neural network. Six to twenty-four hours typhoon tracks are predicted and compared to observations. The contributions of this paper are threefold.

(1) The typhoon observations during 1949 and 2012 are analyzed comprehensively and a typhoon data set is established.

(2) A deep learning model based on the LSTM is established to predict typhoon tracks using typhoon big data. The applicability of the deep learning algorithm to typhoon tracks prediction is validated.

(3) Six to twenty-four hours typhoon tracks are predicted to provide information support to relevant personnel.

## 2 Prediction model for typhoon tracks based on LSTM algorithm

The LSTM algorithm is developed based on the RNNs. The recurrent and error back propagation processes are retained in LSTM, with long and short-term memory cells used to replace the hidden neurons in a traditional RNN. An introduction of the RNN is presented firstly in this section, which leads to the introduction to the LSTM algorithm for clarity. Then the LSTM algorithm is applied to establish a prediction model of typhoon tracks.

### 2.1 A brief introduction to RNN

Traditional feedforward neural networks (FNNs) have promising performance in many fields. However, their performance is poor in dealing with time series such as videos. This is mainly due to the following two aspects. First, the input format of an FNN is a vector of a fixed length. Second, samples in an FNN are assumed to be independent, i.e., they are not related to either temporally or spatially. RNNs have fixed this problem and achieved satisfying performance in speech recognition, machine translation, image interpretation, etc., because an RNN can process each element in the input series one by one, and pass the filtered information in a series in the neural network to retain the correlation among elements in a series.

Compared with an FNN, an RNN allows a series format for the input and output, and breaks the limit to only allow data propagate to a forward level one by one. The input and output of an FNN are normally a vector of a fixed length, while those of an RNN can be a time series like  $(x^1, x^2, \dots, x^t)$  and  $(y^1, y^2, \dots, y^t)$ , the length of which can be finite or infinite, where  $x$  and  $y$  are vectors, and the superscripts are index number which can be taken as time for a time series. Every sample in an RNN is a couple of an input series and an output series.

The output of a hidden neuron only can be directed to neurons on the next layer in an FNN, while in an RNN, that can be passed on to neurons on the same layer and also itself, therefore, information can be passed without a “time” limit. As shown in Fig. 1, the solid lines denote the information flow in an RNN, and

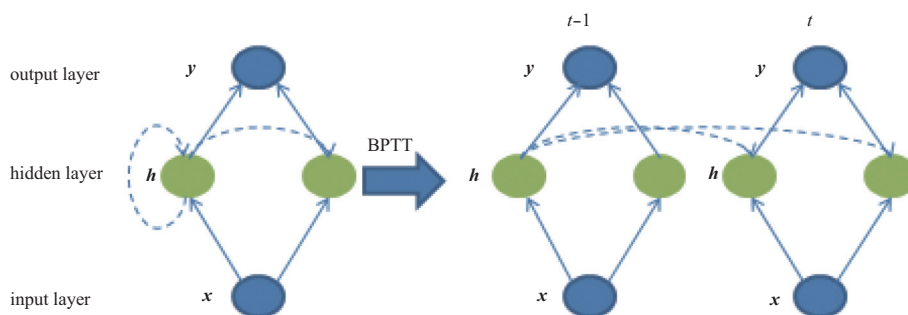


Fig. 1. Unfolding an RNN by the BPTT.

the dashed lines represent the information flow without a “time” limit. With this scheme, it is hard to train an RNN using error back propagation (BP), and thus the back propagation through time (BPTT) method is used to unfold the network by time. The left column in Fig. 1 is an illustration of an RNN and the right column is its unfolded structure by the BPTT. The hidden neurons at time step  $t$  can get information of the input at the current time step  $\mathbf{x}^t$ , as well as the information of the hidden layer at the previous time step  $\mathbf{h}^{t-1}$ , i.e.,  $\mathbf{h}^t = \sigma(W_{hx}\mathbf{x}^t + W_{hh}\mathbf{h}^{t-1} + b_h)$ . Unfolding an RNN by time makes it much easier to train the network as the temporally overlapped parts in an RNN is unfolded as a traditional FNN, which makes the RNN method widely applied.

## 2.2 LSTM models

Although time is introduced in an RNN, and the output of a current time step can be passed to the next time step, the information will be lost unless the value at the next time step is the same as the current one, i.e., the information in a series can only impact the neighboring elements, not any farther. Therefore, the impact is very short in time, the information cannot be stored in the network for a longer time during training, and disappearance or explosion of gradient can happen during error back propagation. Therefore, long and short term memory cells are introduced to store information for any time length to mitigate the disappearance or explosion of gradient remarkably, which is a temporally recurrent neural network called long-short term memory (LSTM).

A special cell structure is used to replace the original hidden neurons in the LSTM. An illustration of the LSTM is shown in Fig. 2, where  $\sigma$  represents the sigmoid function,  $\Pi$  represents multiplication, the subscript c means that it is a single structure, and all the solid arrows mean that the connection weight is 1.  $s_c$  is a memory cell, which is a linear element used to store information to guarantee that information can be stored for a long time to retain the correlation among elements in a sequence.  $g_c$  is the input node, which denotes the comprehensive interaction of the input at time step  $t$  and the information of previous network status. Its value can be passed on to a memory cell through the control of an input gate  $i_c$ . If  $W$  is the weight,  $b$  is the threshold, then  $g_c^t = \sigma(W_{hx}\mathbf{x}^t + W_{hh}\mathbf{h}^{t-1} + b)$ .  $i_c$  is the input gate, which receives the input at time step  $t$  and the network status information at previous time steps, and pass the input value of node  $g_c$  into a memory cell  $s_c$  after the control of the sigmoid function.  $f_c$  is the forget gate, which determines whether the value of  $s_c$  is stored or

not: if the weight is 1, it is stored as it was, and if it is 0, it is cleared.

$O_c$  is the output gate, which receives the input at the time step  $t$  and the network status information at previous time step. It controls the output of  $s_c$  after the sigmoid function.  $v_c$  is the output value.

Use  $\mathbf{x}$  and  $\mathbf{h}$  vectors to denote the values of each layer, then

$$\begin{aligned} g^t &= \tanh(W_{gx}\mathbf{x}^t + W_{gh}\mathbf{h}^{t-1} + b_g), \\ i^t &= \sigma(W_{ix}\mathbf{x}^t + W_{ih}\mathbf{h}^{t-1} + b_i), \\ f^t &= \sigma(W_{fx}\mathbf{x}^t + W_{fh}\mathbf{h}^{t-1} + b_f), \\ o^t &= \sigma(W_{ox}\mathbf{x}^t + W_{oh}\mathbf{h}^{t-1} + b_o), \\ s^t &= g^t \odot i^t + s^{t-1} \cdot f^t, \\ h^t &= \tanh(s^t) \odot o^t, \end{aligned}$$

where  $\odot$  is dot product.

## 2.3 Prediction model for typhoon tracks based on LSTM

Considering a two-dimensional coordinate representation of the typhoon location, for a typhoon which lasts for continuous  $t$  time steps ( $\Delta t$ ), its track can be represented by a sequence  $([x_1^1, x_2^1]^T, [x_1^2, x_2^2]^T, \dots, [x_1^t, x_2^t]^T)$ , where  $[x_1^1, x_2^1]^T$  is the coordinates of a two-dimensional vector, the subscript of  $x_1^1$  represents its dimension, the superscript of  $x_1^1$  represents its index in the series, and each two neighboring elements have a time difference of  $\Delta t$ . Considering that when performing predictions of the typhoon position after  $m \cdot \Delta t$ , the typhoon positions at the previous  $n$  time steps are necessary to show the typhoon track and subsequently obtain a more accurate prediction, the input series of the LSTM should be  $[x_1^1, x_2^1, x_1^2, x_2^2, \dots, x_1^n, x_2^n]^T, [x_1^2, x_2^2, x_1^3, x_2^3, x_1^{n+1}, x_2^{n+1}]^T, \dots, [x_1^{t-n-m+1}, x_2^{t-n-m+1}, x_1^{t-n-m+2}, x_2^{t-n-m+2}, \dots, x_1^{t-m}, x_2^{t-m}]^T$  and the corresponding output sequence is  $[x_1^{n+m}, x_2^{n+m}]^T, [x_1^{n+m+1}, x_2^{n+m+1}]^T, \dots, [x_1^t, x_2^t]^T$ . One input sequence and the corresponding output series form a sample.

An LSTM neural network for the prediction of typhoon tracks is designed in this way with three layers, i.e., an input layer, a hidden layer, and an output layer. There are  $2n$  neurons on the input layer, i.e., the dimension of each element in the input sequence is  $2n$ . The hidden layer consists of 20 long-short term memory cells. There are 2 neurons on the output layer, i.e., the dimension of each element in the output series is 2.

When training the network, all the network parameters are initialized as random numbers between 0 and 1, and then optimized using the training samples. For each training sample, each element in the series is read one by one by the LSTM neural network. An output vector is derived after the hidden layer and the output layer, which is then compared with labels and the errors are back propagated backward by the BPTT algorithm. The test process follows the same way as the training process.

## 3 Experiments

Two experiments were designed in this section. The first one was used to verify the feasibility of a typhoon-track-prediction LSTM network. In the second experiment, training data sets with different amount of samples were used to train the LSTM network, so as to discuss the effectiveness of big data. All the experimental data in this section were from CMA-STI optimal tracks data set of tropical cyclones, which recorded the tracks of the tropical cyclones happened in the Northwest Pacific during 1949

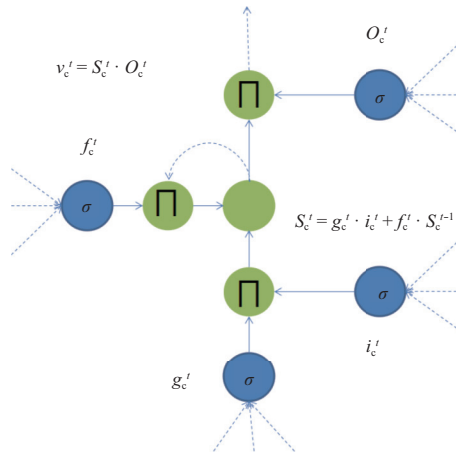


Fig. 2. A memory cell of the LSTM.

and 2011, where the typhoon tracks were recorded using longitude-latitude coordinates with an average precision of 0.1° at every 6 h. The test results were measured by the mean distance errors.

**3.1 Feasibility experiment**

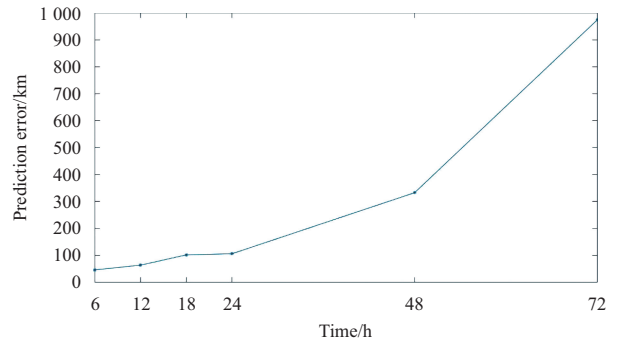
To prove the feasibility of the LSTM neural network to predict typhoon tracks, typhoon positions in 6 h, 12 h, 18 h, 24 h, 48 h, and 72 h were predicted in this experiment, and six data set couples were generated with training set to test set ratio being 8:1. Test results are shown in Fig. 3. The results demonstrate that with the prediction time goes on, the prediction error increases. The prediction results are compared with subjective empirical methods and regional numerical modeling in Table 1. The LSTM network obtained relatively accurate results in 6, 12, and 18 prediction. The prediction error of the LSTM network is comparable to those of the subjective empirical methods and the regional numerical modeling for 24 h predictions, while it soars and becomes much higher than the latter two for 48 h and 72 h predictions. Therefore, it is only meaningful to perform predictions within 24 h using the LSTM network, which also contributes a new method for the prediction of typhoon tracks.

**3.2 Experiment with data sets of various samples**

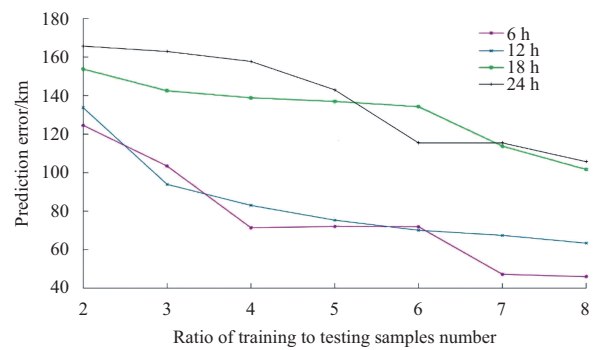
To examine whether the prediction accuracy can be improved with a larger training dataset, 6 h, 12 h, 18 h, and 24 h predictions were performed with the same dataset correspondingly, yet with the number ratio of samples in the training data set to test the data set being 2:1, 3:1, 4:1, 5:1, 6:1, 7:1 and 8:1, respectively. The experiment results are shown in Fig. 4 and Table 2. It can be seen that the prediction error decreases with a larger training data set, i.e., the prediction accuracy of a LSTM neural network will increase with a larger amount of typhoon tracks observations.

**4 Conclusions**

A prediction method for typhoon tracks based on a LSTM neural network is proposed in this paper. The CMA-STI optimal tracks data set of tropical cyclones are used for training and testing. The results show that the 24 h prediction errors are reasonable and comparable to those of the traditional subjective empir-



**Fig. 3.** Prediction errors of the LSTM neural network with time.



**Fig. 4.** Results of the experiment with various number of samples in the training data set.

ical method and the regional numerical modeling. Therefore, it provides a new and feasible method for the prediction of typhoon tracks. It also indicates that the typhoon tracks prediction error of the LSTM neural network can be reduced by using a larger training data set. Therefore, with increasingly larger amount of typhoon observations, the prediction accuracy of the LSTM neural network will increase.

In this paper, only the typhoon tracks observations are used as the input for the prediction. More information can be included in future studies, such as the central pressure, topo-

**Table 1.** Prediction errors (km) of different methods

Prediction time/h	LSTM	Subjective empirical methods	Regional numerical modeling
6	45.9540	-	-
12	63.3670	-	-
18	101.6243	-	-
24	105.6768	84.2	97.4
48	332.5407	145.6	188.2
72	974.4985	205.4	302.7

**Table 2.** Results of the experiment with various number of samples in the training data set

Ratio of sample numbers in the training data set to the test data set	6 h errors/km	12 h errors/km	18 h errors/km	24 h errors/km
2	124.5382	133.6872	153.7482	165.6832
3	103.4238	93.8955	142.4895	162.8953
4	71.3806	82.9868	138.8510	157.7417
5	72.0321	75.3128	136.9979	142.9761
6	71.9181	70.0683	134.2442	136.9106
7	47.2124	67.4215	113.7582	115.4255
8	45.9540	63.3670	101.6243	105.6768

graphy, etc., to improve the structure and accuracy of the model.

## References

- Chen Guomin, Cao Qing, Bai Lina. 2015. Verification on forecasts of tropical cyclones over western north Pacific in 2014. *Meteorological Monthly* (in Chinese), 41(12): 1554–1561
- Donahue J, Anne Hendricks L, Guadarrama S, et al. 2015. Long-term recurrent convolutional networks for visual recognition and description. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA: IEEE, 2625–2634, doi: 10.1109/CVPR.2015.7298878
- Hochreiter S, Schmidhuber J. 1997. Long short-term memory. *Neural Computation*, 9(8): 1735–1780
- Huang Xiaoyan, Jin Long. 2013. An artificial intelligence prediction model for typhoon tracks based on principal component analysis. *Chinese Journal of Atmospheric Sciences* (in Chinese), 37(5): 1154–1164
- Li Zechun, Chen Dehui. 2002. The development and application of the operational ensemble prediction system at national meteorological center. *Journal of Applied Meteorological Science* (in Chinese), 13(1): 1–15
- Qian Chuanhai, Duan Rihong, Ma Suhong, et al. 2012. The current status and future development of China operational typhoon forecasting and its key technologies. *Advances in Meteorological Science and Technology* (in Chinese), 2(5): 36–43
- Ranzato M A, Szelam A, Bruna J, et al. 2014. Video (language) modeling: a baseline for generative models of natural videos. *Eprint Arxiv*
- Shi Xingjian, Chen Z, Wang H, et al. 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Cambridge: MIT Press, 802–810
- Sutskever I, Vinyals O, Le Q V. 2014. Sequence to sequence learning with neural networks. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal, Canada: MIT Press, 3104–3112
- Wang Kangling, He Anguo, Xue Jishan. 1996. Preliminary test of typhoon trace numerical prediction for the South China Sea area. *Journal of Tropical Meteorology* (in Chinese), 12(2): 113–121
- Xu Daosheng, Chen Zitong, Dai Guangfeng, et al. 2014. The influence of an improved cumulus parameterization scheme on typhoon forecast from GRAPES model. *Journal of Tropical Meteorology* (in Chinese), 30(2): 210–218
- Xu K, Ba J, Kiros R, et al. 2015. Show, attend and tell: neural image caption generation with visual attention. *Computer Science*, 2048–2057
- Xu Yinglong, Zhang Ling, Gao Shuanzhu. 2010. The advances and discussions on China operational typhoon forecasting. *Meteorological Monthly* (in Chinese), 36(7): 43–49
- Yin Wenjun, Zhang Dawei, Yan Jinghai, et al. 2015. Deep learning based air pollutant forecasting with big data. *Chinese Journal of Environmental Management* (in Chinese), 7(6): 46–52
- You Cheng, Yu Fujiang, Yuan Ye. 2016. Storm surge prediction method of neural network based on phase space reconstruction. *Marine Forecasts* (in Chinese), 33(1): 59–64
- Yu Jinhua, Tang Jiexiang, Dai Yuhan, et al. 2012. Analyses in errors and their causes of Chinese typhoon track operational forecasts. *Meteorological Monthly* (in Chinese), 38(6): 695–700