Sea surface temperature prediction model based on long and short-term memory neural network

Xiaojing Li*

Guangxi College for Preschool Education, Nanning, 530022, China
*Corresponding author’s e-mail: 50921626@qq.com; jjjin@cs.ecnu.edu.cn

Abstract. In response to the monitoring and forecasting of El Nino/La Nina phenomenon, this paper proposes a sea surface temperature prediction method based on long and short-term memory neural network for the average sea surface temperature in the NINO comprehensive area. This method uses the monthly anomaly sea surface temperature sequence for the mean sea surface temperature in the NINO Comprehensive area as the input of the long- and short-term memory neural network to establish a forecast model. The average sea surface temperature of the NINO comprehensive area is forecasted for the next 1 to 3 months. The results show that the method can better predict the average sea surface temperature of the NINO comprehensive area, which is useful for the monitoring and forecasting of El Nino/La Nina phenomenon. Provides a new approach.

1. Introduction

With the increase of global abnormal weather and climate and the frequent occurrence of severe weather, to study and analyze the causes of abnormal weather and climate, and to make scientific and effective estimates of them is the focus and hotspot of climate change research today. A large number of observational data and theoretical studies show that ENSO events have a significant impact on abnormal changes in global climate[1]. Its occurrence and development often cause severe heat, drought, floods and other disasters in many areas[2]. Therefore, predicting the sea surface temperature in the El Niño sea area is the key to studying ENSO events and making short-term climate predictions, and it is also an important part of studying global climate change.

The ocean and the atmosphere are a coupled system that affects each other. The change of sea surface temperature not only affects the ocean, but also greatly affects the state of the atmosphere. The sea surface temperature prediction methods generally include the dynamic model method of numerical forecasting and some traditional statistical methods. After more than ten years of development, the dynamic model based on the laws of physics has made great progress, but there are still great uncertainties in its prediction results[3-4], and the prediction results of different models are also quite different. Traditional statistical methods are generally based on solving linear equations, and these methods are used to predict sea surface temperature with nonlinear changes, which have great limitations.

In recent years, some nonlinear statistical forecasting methods have also been used in sea surface temperature forecasting[5-6]. For example, literature[5] used the integrated prediction method of wavelet decomposition and ANFIS model to predict the sea surface temperature of the equatorial East Pacific effectively, Literature[6] introduced chaos theory into sea surface temperature forecasting, combined phase space reconstruction theory with fuzzy neural network, and carried out prediction experiments on sea surface temperature in the sea east of Taiwan, and achieved good results.
In recent years, Long Short Term Memory Networks (LSTM) is a variant of many Recurrent Neural Networks (RNN). It makes up for the problems of RNN's gradient disappearance, gradient explosion, and insufficient long-term memory ability, so that the recurrent neural network can truly effectively use long-term timing information[7]. Unlike traditional RNNs, LSTM has a more complex memory unit, which can maintain good memory for long-term time series. Therefore, the model has outstanding performance in the problem of time series forecasting, and has been a research hotspot in the field of machine learning in recent years[8-11]. This paper attempts to use long and short-term memory neural network to establish a prediction model for the sea surface temperature anomaly in the next 1 to 3 months in the monthly average sea surface temperature anomaly time series of the NINO comprehensive area, and analyze and discuss the forecast effect of the model.

2. Research material
The research data adopts the reanalysis data of the monthly mean sea surface temperature field from January 1854 to December 2015 provided by NCEP/NCAR. At the same time, according to the definition and monitoring of El Nino events by the National Climate Center, the research scope of this article is defined as NINO comprehensive area (namely NINO1+2+3+4 area: 90°W-160°E, 5°S-5°N, 80°W-90°W, 10°S-0°).

3. LSTM neural network prediction model
LSTM Neural Network was first proposed by Sepp Hochreiter and Jurgen Schmidhuber in 1997[12], mainly improved the problem of the disappearance of gradient when the recurrent neural network is dealing with distant sequences. The architecture for an LSTM block is shown below. An LSTM block typically has a memory cell, input gate, output gate, and a forget gate in addition to the hidden state in traditional RNNs. As a recurrent neural network model, LSTM model has memory characteristics and is suitable for dealing with sequential problems. In addition, LSTM itself has a special structure, which is more suitable for dealing with delay and persistence characteristics than other neural network models.

![LSTM structure diagram](image)

The learning process of LSTM algorithm is controlled through forget gate, input gate and output gate. \( \sigma \) is the 'sigmoid' activation function, \( w_f \) is the weight matrix of the forgetting gate, \( w_i \) is the weight of the input gate, matrix, \( w_c \) is the weight matrix of the update gate, \( w_o \) is the weight matrix of the output gate, \( b_f \) is the bias of the forgetting gate, and \( b_i \) Represents the bias of the input gate, \( b_c \) represents the bias of the update gate, \( b_o \) represents the bias of the output gate, \( h_t \) represents the output at time t, and \( c_t \) represents the cell state updated at time t.
Forgotten door: The forget gate can determine which information is discarded from the unit information state. Using sigmoid as the activation function, it can output a value between 0-1 for each element in the unit state. The formula is:

\[ f_t = \sigma(w_f [h_{t-1}, x_t] + b_f) \] (1)

Input gate: The input gate is used to determine the information to be added to the unit state. One is the input gate using sigmoid as the activation function to determine the value to be updated; the other is the 'tanh' layer, which is used to create a new value to be added to the unit state. From the above results, the cell state update value \( C_t' \) at time \( t \) can be calculated, and the formula is:

\[ C_t' = \tanh(w_c [h_{t-1}, x_t] + b_c) \] (2)

Output gate: Determine the value of the output gate, and the formula is:

\[ h_t = O_t \text{tanh}(C_t') \] (3)

Output gate:

\[ O_t = \sigma(w_o [h_{t-1}, x_t] + b_o) \] (4)

\[ h_t = O_t \text{tanh}(C_t') \] (5)

4. Prediction experiment and result analysis of average sea surface temperature in NINO comprehensive area.

The paper divides the sample of data into two sections, among which the average sea surface temperature data from January 1854 to December 1994 is a 140-year modeling sample, a total of 1680 samples. Among them, the average sea surface temperature data from January 1995 to December 2015 are independent forecast samples for 20 years, with a total of 240 independent forecast samples. For the monthly mean sea surface temperature time series \{ \( X_i(t), i = 1,2,L,N \); where \( t \) represents the current time \}, this article will forecast the next month \{ \( X(t + j), j = 1,2,3 \) \}.

Use the LSTM neural network structure to train and predict the sea surface temperature anomaly sequence. The specific prediction steps are as follows:

1. Network initialization. Initialize the weight \( W \) and the bias vector \( b \), set the total number of memory blocks \( N \), the learning rate \( \eta = 0.005 \), the maximum number of iterations \( T = 250 \), the training target error \( \varepsilon = 0.0001 \), the initial memory \( c_0 = 0 \), the initial output \( s_0 = 0 \).

2. Data standardization. Perform min-max standardization on the data set \( x \) to obtain a standardized data set \( X^* = \{x_1^*, x_2^*, L, x_n^*\} \), among them:

\[ x_t^* = (x_t - x_{\text{min}}) / (x_{\text{max}} - x_{\text{min}}) \quad 1 \leq t \leq n \] (6)

3. Data division. Divide \( X^* \) into the training set \( X_{tr}^* = \{x_1^*, x_2^*, L, x_d^*\} \) and the test set \( X_{te}^* = \{x_{d+1}^*, x_{d+2}^*, L, x_n^*\} \).

4. Error calculation. Calculate the error term between the output of the output layer and the theoretical output \( E \).

\[ E = \frac{1}{2}(s_t - x_t)^2 \] (7)

5. Weight and threshold update. Use Adam gradient optimization algorithm to update \( W \) and \( b \) based on the error term \( E \).

6. Repeat (3)-(5), and end when the number of training reaches the maximum number of iterations or the absolute value of the difference between the updated value and the value before the update is less than \( \varepsilon \).
### Table 1  Forecast results of 240 independent samples

| NINO integrated average sea surface temperature | 1 month ahead | 2 month ahead | 3 month ahead |
|-----------------------------------------------|--------------|--------------|--------------|
| Correlation coefficient                       | 0.9443       | 0.9026       | 0.8832       |
| Mean absolute error                           | 0.2473       | 0.3981       | 0.4365       |
| Anomaly correlation coefficient( ACC)          | 0.9083       | 0.8847       | 0.8246       |
| Anomaly sign agreement rate (%)               | 90.83%       | 84.58%       | 78.08%       |

It can be seen from Table 1 that, firstly, the correlation coefficients of 240 independent samples have reached above 0.88, among which the correlation coefficient between the forecast results of the next month and the independent sample forecast will reach 0.94; secondly, the average absolute error of the forecast is the smallest It reached 0.25, and the maximum did not exceed 0.44. Third, after calculating the corresponding anomaly value using the forecast value in the next 1 to 3 months, the agreement rate of the anomaly signs also reached more than 78%.

Figure 2 is a comparison diagram of the anomaly between the actual value and the predicted value one month ahead. It can be seen from the figure that the forecast result of the forecast model can be better close to the actual situation in terms of overall trend and local details, and can be compared. The main trend of the sea surface temperature change is well predicted, and the forecast model performs well in the prediction of multiple break points. It can be seen that the prediction effect of the sea surface temperature prediction model established in this paper is relatively ideal.

---

5. Summary and discussion

Although the changing law of sea surface temperature is extremely complicated, its short- and medium-term changing trend is still predictable. This article attempts to use the LSTM with memory function to predict the average sea surface temperature of the NINO complex in the next 1-3 months, and obtain the following conclusions:
1) The LSTM network has a strong ability to fit training samples.
2) In the 20-year independent sample experiment, most of the forecasts of sea surface temperature anomaly trends are consistent. The correlation coefficients of the independent samples are all above 0.88, and the coincidence rate of the anomaly signs is above 78%.
3) The prediction of sea surface temperature by the LSTM network increases with the increase of the prediction step, and its prediction performance decreases relatively greatly.

It can be seen that using the LSTM neural network model to predict sea surface temperature is a new prediction method, and it has indeed achieved good results in the sea surface temperature prediction experiment. Therefore, it can be a better tool and method for predicting sea surface temperature changes.

Acknowledgments
This work was supported by the Guangxi Natural Science Foundation (Grants 2018GXNSFAA294128, 2018GXNSFAA281229, 2017GXNSFDA198030) and the National Natural Science Foundation of China (Grant 41765002).

References
[1] Duanjun Lu and Xiangong Zhang. The response of precipitation and temperature in china to ENSO[J]. Quarterly Journal of applied meteorology, 1995, 6( 1): 118—123.
[2] Wucheng Xu, Jinsong Ma, Wen Wang. A review of studies on the influence of enso events on the climate in china[J]. Scientia Meteorologica sinica. 2005,25(1):212-220.
[3] Luo J J, Masson S, Behera S K, et al. Extended ENSO predictions using a fully coupled ocean–atmosphere model [J]. J. Climate, 2008,21 (1):84–93.
[4] Jin E K, Kinter III J L, Wang B, et al. Current status of ENSO prediction skill in coupled ocean–atmosphere models [J]. Climate Dyn., 2008,31 (6): 647–664.
[5] Zhangren Wang, Jiguang Jiang, Guo-rong. A conpositive prediction of eastern pacific equatorial sst based on wavelet decomposition and anfis model. Journal of tropical oceanography. 2002 , 21:77- 84.
[6] Zhaojun Dong, Teng jun, Jipeng Wang. Application of phase space reconstruction and ANFIS model in SST forecasting. [J].Journal of tropical oceanography. 2008,27(4):73-76.
[7] Graves A. Supervised sequence labelling with recurrent neural networks[J]. Studies in Computational Intelli-gence, 2012,2(385):42-45.
[8] Ziming Wang and Shunhuai Chen. Real-time Forecast of Fuel Consumption of Ship Main Engine Based on LSTM Neural Network[J]. Journal of Wuhan University of Technology (Transportation Science & Engineering ), 2020,44(05):923-927.
[9] Liu Xin, Zhao Ning, Guo, Jinyun, Guo Bin, Prediction of Monthly Precipitation over the Tibetan Plateau based on LSTM Neural Network[J]. Journal of Geo-information Science, 2020,22(08):1617-1629.
[10] Zhao Mingzhu, Wang Dan, Li Yan, et al. Prediction of subway station temperature based on LSTM neural network [J].JOURNAL OF BEIJING JIAOTONGUNIVERSITY. 2020,44(04):94-101.
[11] Qing Yang and Chenwei Wang. A Study on Forecast of Global Stock Indices Based on Deep LSTM Neural Network[J]. Statistical Research. 2019, 36(03):67-79.
[12] Hochreiter S, Schmidhuber J. Long short- term memory[J]. Neural Computation,1997,9(8):1735-1780.