Neural Network Implementation on Bias Correction of Sea Surface Temperature Forecast

Zhibo Li*, Yike Liu

TianJin YingHua experimental School, Tianjin, China
**Corresponding author. Email: 2254039500@qq.com*

ABSTRACT

This article introduces the implementation of neural network on making sea surface temperature forecasts more accurate. We manage to utilize LSTM neural network and intended to make comparisons with other networks such as Backward propagation network. We improve the accuracy of prediction by doing bias correction, the detailed solution will be illustrated as follows.

Keywords: *Neural Network, sea surface temperature forecast, bias correction*

1. INTRODUCTION

Air temperatures on Earth have been rising since the Industrial Revolution. According to an ongoing temperature analysis led by scientists at NASA's Goddard Institute for Space Studies (GISS), the average global temperature on Earth has increased by at least 1.1° Celsius (1.9° Fahrenheit) since 1880 [1]. Global warming does not mean temperatures rise everywhere at every time by same rate. Temperatures might rise 5 degrees in one region and drop 2 degrees in another. In this case, predicting temperature change and make proper preparations seems inevitable. In the meantime , information technology is developing in a unexpected fast speed. Deep learning--,one important branch of Artificial intelligence, is particularly useful in predicting data and processing data analysis. Deep Learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so, it allow us to build neural networks in a way that mimics the working of a biological brain [2]. With the help of deep learning we then, should be able to predicting temperature change .So in this essay, we'll be focusing on talking about using long short-term memory, LSTM, which is a type of an artificial recurrent neural network (RNN) architecture, to make data analysis and make predictions. On the next paragraph, will explain the dataset we chose to use to train our algorithm, and on the third paragraph, we'll be explaining why we use LSTM structure instead of other structures like RNN, GRU. Then on the fourth paragraph, we'll start talk about our infrastructure of the neuron network and the exact steps and approaches we use to analyze data. Then on the final paragraph, we 'll give a short conclusion on this algorithm we trained.

1.1.Data set introduction

Date set which we choose consists of observed and predicted values of sea surface temperature which we will call it SST instead in later paragraphs. These temperature values are extracted from massive data set of the South China Sea, we picked the data set at the Beibu Gulf only, in a time period from January 2003 to December 2008, for that we do not intend to use too much data which will in all aspects apply a heavy but unnecessary load upon the algorithm. The set of data which we eventually make use of is called bias, which is obtained by using simple arithmetic calculation, that is we subtract the observation value from the prediction value at a certain time spot to get the bias value at that specific time spot, and we do the same thing to all corresponding predicted and observed value to get bias value that we need at every given time spot. This magnitude of the bias value, with due consideration of their standard deviation and mean value, shows the integrity and accuracy of our prediction, the original data set is showing in the file named bias.csv for the convenience of use. Since it might not look as intuitive as we wish it to be, by looking at massive of figures, we can't really tell whether our prediction is good enough. So, our first step when we start writing a program is plotting a visualized graph which enables us to have a perceptual intuition. The graph is shown below as Fig 1a. We can see that the graph reflects perfectly the bias distribution, at this before utilizing our algorithm, the bias is not only highly dense but also very unevenly balanced, hence we do not expect a good result from our current prediction. In Fig 1b we stretch the x axis to gain a clearer view of the bias distribution with more detailed time line. We decided to use data from January 2003 to December 2007 for training our algorithm and the rest of them which is the whole year of 2008 for testing our algorithm.

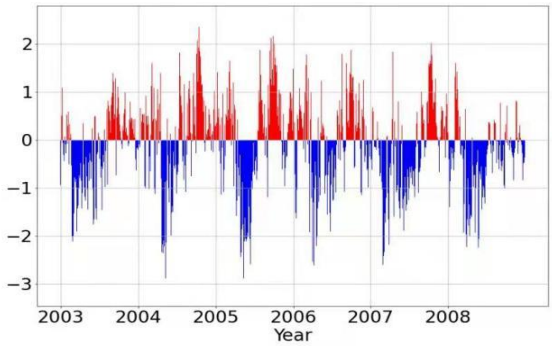

Figure 1a. bias values before bias correction.

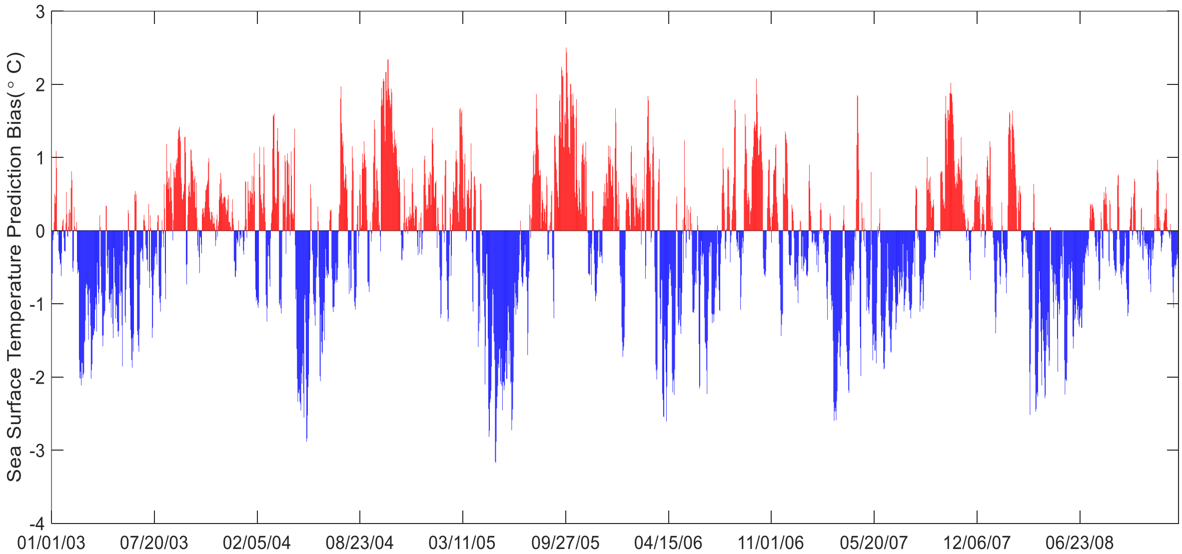

Figure 1b. bias values before bias correction with more detailed x axis values.

2. WHY LSTM

CNN is a type of recursive neural network that receives sequence data as input, performs recursion in the sequence's growth direction, and all nodes (cyclic units) are connected by chain [3]. RNN differs from ordinary BP neural networks in that the hidden layer of the ordinary BP neural network just accepts the output of the neuron from the previous layer, whereas RNN receives the output of the neuron from the previous layer as well as the information from the preceding time point. CNN can play a more effective forecast impact for time series with a particular correlation between the time nodes.

However, for the recurrent neural network, receiving the information of the last time point can improve the prediction effect of time series, but it will also cause the problems of gradient vanishing and exploding gradient. Therefore, recurrent neural networks are not suitable for analyzing long-term dependent time series. The LSTM neural network is an enhanced version of the recurrent neural network [4]. The LSTM neural network uses a gating method to handle the problem of gradient explosion and gradient disappearance while conserving the memory of the last time point, making it more appropriate for the analysis of long-term dependent time series. Before introducing our algorithm, first I ‘ll give you a brief idea on what is a LSTM so you may understand our algorithm more properly. LSTM has one most important component --memory cell, and 3 smaller component to control the memory cell, first one forget gate, it’s designed to judge whether to let the contents in memory cell affect the value in memory cell in this step; second one is the input gate, it’s used to read data into memory cell, it affects the value of cell mostly; the third gate is called output gate, it read out the entries from the cell.

The process starts when input X and Hidden state from last time is logged in, after the calculation of three gates is finished, we van then calculate candidate memory cell, which uses tanh function as its activation function to control it’s range from -1to 1, then we can conclude from last memory cell and new candidate cell to get our new memory cell, and finally, use output gate to control the flow of information to hidden state.

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \quad (1)$$

$$F_t = \sigma(X_t W_{xf} + H_{t-1}\ W_{hf} + b_f) \quad (2)$$

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \quad (3)$$

Our equation of calculating current values for three gates are provided above, I_t, F_t, O_t represent Input gate, forget gate, and output gate respectively; W_{xi}, W_{xf}, W_{xo} represents the weight parameter for each value, H_{t-1} is the hidden state of last time, b_i, b_f, b_o represents the bias parameter.

Then the value for new candidate memory cell is calculated in this case, it uses the tanh function to control the range of the output value between negative 1 and 1

$$C_c = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \quad (4)$$

The final step would be calculating immediate memory cell value, it makes use of the information from last time state and current time state, using this equation calculating current memory cell state, then output gate could be used to get the eventual output value.

$$C_t = F_t * C_{t-1} + I_t * C_c \quad (5)$$

2.1.Method used

Step 1 Problem Decomposition

To create this neural network, we must first have a clear understanding of the problem we're attempting to address. To do so, we employ decomposition, which is a strategy for collecting vital information from a problem and breaking it down into smaller chunks. In this scenario, we'll use the bias of the predicted value and the true value of the temperature, and then perform a training process related to LSTM. If it is successful, we will be able to anticipate temperature changes more accurately.

Step 2 LSTM Network for Regression

After we load the database, the data should be normalized before it becomes the input of the LSTM network, here we choose to use the default sigmoid function, to reshape the magnitude of data between negative 1 and 1. Then, we split the data set up into training set, which ranged from 0 to the 1826th data group, and the test set, which ranged from the 1826th to the final data group. we can now design and fit the LSTM to our problem. The network consists of a visible layer with one input, a hidden layer with four LSTM blocks or neurons, and an output layer that predicts a single value. For the LSTM blocks, the default sigmoid activation function is employed. The network is trained across 100 epochs with a batch size of one. After we've fitted the model, we can begin training our software and making predictions [5].

Step 3

The bias correction result is produced and output as figures, but as we have said before, figures are complicated and not straight forward enough, we then programs the algorithm to plot another corresponding graph which will give us a clear view of what happened as shown in Fig 2a. However, this graph seems to be squeezed as well, so we stretch the x axis to pull out a detailed time line in Fig2b. As we can tell, the density of the bias distribution graph has dramatically decreased since training, and the amplitude value of the bias has been well under a magnitude of 2 instead of the previous 3. After training our algorithm, we use our current prediction value at each time spot to minus the corrected bias value to strengthen the accuracy of our prediction. Thus we call this bias correction of sea surface temperature forecasts [6].

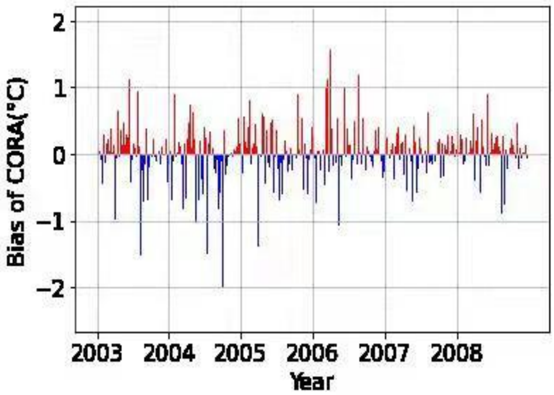

Figure 2a. bias values after bias correction

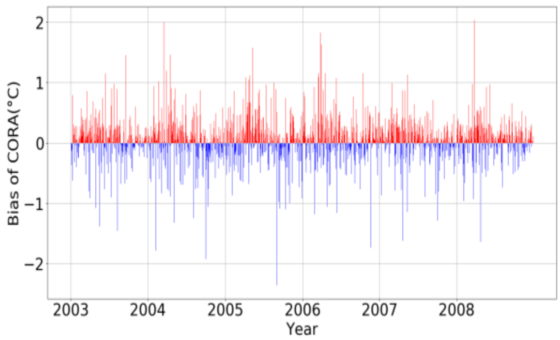

Figure 2b. bias values after bias correction with more detailed x axis values.

3. CONCLUSION

Nowadays, the accuracy of forecast of sea surface temperature is becoming more and more vital to satisfy all sorts of needs, including guidance for weather forecast and early warning of natural disasters. The only right course is to improve the forecast as much as possible, in this article we have made use of the LSTM deep learning which belongs to recurrent neural network and successfully diminishes part of the prediction biases using bias correction and the result of prediction of 2018 has been more accurate than before. However, there still leaves question to be asked and problem to be solved [7]. Some parameters in the program could be changed to seek for improvements, for instance, the use of an alternative batch size and as well as increase the number of look backs. We shall introduce a larger data set to the LSTM network so that it can produce further more accurate results, we might try some previously mentioned networks like BP to see whether our decision is right practically. For a vastly changing global climate, only more data could be implemented to foresee all sorts of unforeseeable problems, which could be done in due course.

REFERENCES

[1] NasserRostamAfsharHedayetFahmi (2018)Impact of climate change on water resources in Iran.

[2] ShraddhaSinghaSaroj KumarPandeybUrjaPawarcRekh RamJangheld (2019)Classification of ECG Arrhythmia using Recurrent Neural Networks

[3] XueJiangaceRuijieZhangabCongZhangaHaiqingYinaceXuanhuiQubde(2019)Fast prediction of the quasi phase equilibrium in phase field model for multicomponent alloys based on machine learning method

[4] Di Wu; Zhongkai Jiang; Xiaofeng Xie; Xuetao Wei; Weiren Yu; Renfa Li(2020)LSTM Learning With Bayesian and Gaussian Processing for Anomaly Detection in Industrial IoT

[5] S.Abhilash,A.K.Sahai,N.Borah,R.Chattopadhyay,S.Joseph,S.Sharmila,S.De,B.N.Goswami(2014)Does bias correction in the forecasted SST improve the extended range prediction skill of active-break spells of Indian summer monsoon rainfall

[6] LongyueLiangXuanyeCai(2020)Forecasting peer-to-peer platform default rate with LSTM neural network

[7] Qianchen Xianghua Lvng Mang Gaoa Wang(2020)A New Information-Theoretic Method for Advertisement Conversion Rate Prediction for Large-Scale Sparse Data Based on Deep Learning.