RAA 2021 Vol. **21** No. **1**, 12(8pp) doi: 10.1088/1674-4527/21/1/12

http://www.raa-journal.org *http://iopscience.iop.org/raa*

***R**esearch in*
***A**stronomy and*
***A**strophysics*

Solar cycle prediction using a long short-term memory deep learning model

Qi-Jie Wang, Jia-Chen Li and Liang-Qi Guo

School of Geosciences and Info-Physics, Central South University, Changsha 410083, China; *guolqch@csu.edu.cn*

Abstract In this paper, we propose a long short-term memory (LSTM) deep learning model to deal with the smoothed monthly sunspot number (SSN), aiming to address the problem whereby the prediction results of the existing sunspot prediction methods are not uniform and have large deviations. Our method optimizes the number of hidden nodes and batch sizes of the LSTM network structures to 19 and 20, respectively. The best length of time series and the value of the timesteps were then determined for the network training, and one-step and multi-step predictions for Cycle 22 to Cycle 24 were made using the well-established network. The results showed that the maximum root-mean-square error (RMSE) of the one-step prediction model was 6.12 and the minimum was only 2.45. The maximum amplitude prediction error of the multi-step prediction was 17.2% and the minimum was only 3.0%. Finally, the next solar cycles (Cycle 25) peak amplitude was predicted to occur around 2023, with a peak value of about 114.3. The accuracy of this prediction method is better than that of the other commonly used methods, and the method has high applicability.

Key words: Sun: solar activity — Sun: sunspot number — techniques: deep learning — techniques: long short-term memory

1 INTRODUCTION

Changes in solar activity affect activities in space, the Earth's magnetic field, the Earth's climate, and human activities (Usoskin 2008). Sunspots are one of the most basic and obvious features of solar activity. The number of sunspots is also related to the intensity of solar radiation (Solanki 2003).

Forecasting sunspots is not only important for the study of changes in solar activity and for an understanding of the mechanism of solar activity, but it is also important for space navigation and solar-terrestrial relations (Prabhakaran Nayar et al. 2002; Gholipour et al. 2005; Ahluwalia & Jackiewicz 2012). However, the change process and mechanism of solar activity are extremely complicated. The prediction of sunspots is also very difficult using general techniques (Kitiashvili 2016).

Spectral analysis, neural networks, climatological prediction, dynamo models, and precursor methods are the main methods for solar cycle prediction. Spectral analysis is an analytical method for calculating structural response, which combines modal analysis results with known spectra. It examines a Fourier analysis of sunspot time series for invariant quantities. Neural networks are nonlinear statistical algorithms, which can determine complex relationships between input and output (Attia et al. 2013). Climatological prediction assumes that the future of a system can be determined by the statistical characteristics of the past behavior of the system. It is a general term for a class of statistical methods. Dynamo models are based on physics. The conservation equation of a physical model can be integrated by the dynamo model, which in turn can predict solar activity (Nandy 2002). Precursor methods are the most common prediction method at present; these methods use the solar polar magnetic field and a geomagnetic activity index to forecast solar activity.

Many scholars have used these methods to predict the peak value of the smoothed monthly sunspot number (SSN) for Cycle 24. Among them, Noble and Noble & Wheatland (2012) (66+5) and Rigozo et al. (2011) (113.3) used spectral analysis methods; Ajabshirizadeh et al. (2011) (65) and Attia et al. (2013) (101+8) used neural network methods; Han et al. (2018) (134.1), Sabarinath & Anilkumar (2018) (78+25) used climatological methods; dynamo models were used by Choudhuri et al. (2007) (80); and Dabas & Sharma (2010) (131+20), and Muñoz-Jaramillo et al. (2013) (78) used precursor methods. These forecast results are combined with the results collected by Pesnell (2008), Pesnell (2016) and Han et al. (2018) and are listed in Table 1, according to the different methods used. The observed peak value of Cycle 24 is 81.9 (Version 1.0).

The methods listed in Table 1 include both original and improved methods, where $N_{\rm pre}$ is the number of

Table 1 Cycle 24 Amplitude

Method	N_{pre}	RMp	DRM	PDRM	N(10%)
Spectral	24	43-180	19.8	17%	5
Neural network	6	65-145	13.1	11.3%	0
Climatological	33	40-185	11.6	10.0%	10
Dynamo model	5	80-167	11.9	10.2%	3
Precursor	36	53-180	8.3	7.1%	5

forecast cases, RMp is the range of the predicted Cycle 24 amplitude, DRM is the absolute error of the forecast value, PDRM is the relative error of the forecast value, and N is the forecast number with a prediction error of less than 10%. DRM and PDRM can be expressed as:

$$\mathrm{DRM} = |\mathrm{Predictedvalue} - \mathrm{Observedvalue}| \quad (1)$$

$$\mathrm{PDRM} = \frac{|\mathrm{Predictedvalue} - \mathrm{Observedvalue}|}{\mathrm{Observedvalue}} * 100\% \quad (2)$$

It can be seen that the prediction values obtained not only have large distribution ranges, but also low prediction accuracies. The prediction errors of most of the prediction methods (we usually use the relative error (PDRM) to compare the accuracy of prediction results, so the prediction errors in the following are all relative errors) are above 10%. In general, obtaining an accurate prediction of future solar activity amplitude is still very difficult (Pesnell 2014).

Deep learning, which was developed from ordinary neural networks, has developed rapidly in various fields due to its powerful capabilities and flexibility. However, to date, there has been little research on solar cycle forecasting using deep learning techniques. In this study, the long short-term memory (LSTM) deep learning model was used to predict the SSN. The number of hidden nodes and the batch size in the LSTM model were also optimized, Finally, appropriate training samples were selected experimentally.

2 DEEP LEARNING AND NETWORK OPTIMIZATION

2.1 Introduction to deep learning and recurrent neural networks

Deep learning, a type of machine learning, originated from image recognition. Deep learning algorithms are both powerful and flexible. In deep learning, there are two typical neural network structures — the convolutional neural network (CNN) and the recurrent neural network (RNN) — among which, the RNN is more suitable for dealing with time-series analysis.

The RNN model was derived from the Hopfield network model by Jordan in 1986 (Saratha & Tajuddin 2008). However, the RNN model was not widely used

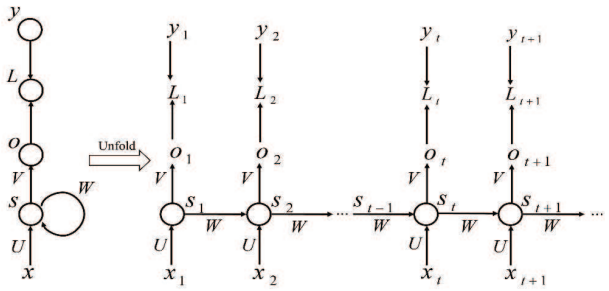

Fig. 1 Structure of a recurrent neural network.

at that time. As more efficient RNN structures have continually been proposed, the ability of RNNs to explore the temporal and semantic information in data has been fully utilized. Language models (Yamamoto et al. 2001), machine translation (Saha & Raghava 2010), speech recognition (Graves et al. 2013), and timing analysis (Lukosevicius & Jaeger 2009) have all been achieved using RNNs.

RNNs contain input units, where the corresponding input unit is labeled $\{x_1,x_2,\cdots,x_t,x_{t+1},\cdots\}$, $\{o_1,o_2,\cdots,o_t,o_{t+1},\cdots\}$ are recorded as the output units, the corresponding output unit is denoted as $\{y_1,y_2,\cdots,y_t,y_{t+1},\cdots\}$, and $\{s_1,s_2,\cdots,s_t,s_{t+1},\cdots\}$ is a hidden unit. U, V, and W are parameters from the input layer to the hidden layer, the hidden layer to the output layer, and the hidden layer to the hidden layer, respectively. When the network calculates the output unit $\{o_1,o_2,\cdots,o_t,o_{t+1},\cdots\}$, the loss $\{L_1,L_2,\cdots,L_t,L_{t+1},\cdots\}$ can be calculated from each o and the corresponding training target y. A back propagation algorithm is then used to update the network parameters and minimize the loss L, finally obtaining the trained network model. The loss L can be calculated by the following formula:

$$L = -(y \cdot \log \hat{y} + (1-y)\log(1-\hat{y})) \quad (3)$$

where $\hat{y} = \mathrm{softmax}(o)$, which is the value of the output value o after normalization. The loss function can be used to predict the probability of a certain situation for different independent variables. Only the loss function is minimized to determine the parameters of the model when LSTM is used for prediction (Graves 2012). A typical RNN and its expanded structure are shown in Figure 1.

In an RNN, the main work is done by the hidden units. One flow of information is from the input unit to the hidden unit, and a second flow of information is from the hidden unit to the output unit, both of which flow in the same direction. In some specific cases, the network will direct information from the output unit directly to the hidden unit. This information is called “back projection”, and the input of the hidden layer includes not only the information from the output unit, but also the state of the previous hidden layer.

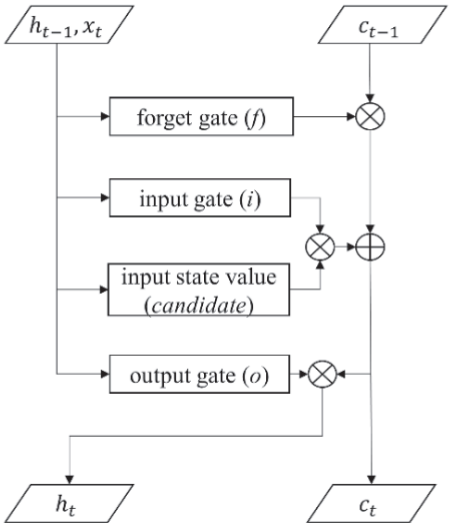

Fig. 2 LSTM unit details.

The hidden unit can contain important information about the past, and the former hidden unit is indirectly connected to the current part of the network only through the generated prediction. More parallelism and a shorter training time can be achieved during training because the time steps can be separated. The variables and operations in the hidden layer are the same at different times, so the RNN can process arbitrarily long sequences with limited parameters.

2.2 The long short-term memory network structure

The LSTM network structure was proposed by Hochreiter and Schmidhuber in 1997 to solve the long-term dependence problem in the previous RNNs (Hochreiter & Schmidhuber 1997).

The "forget gate" and "input gate" are the core of the LSTM structure. The forget gate calculates a vector f whose dimension is n according to the current input x_t and the output $h_{(t-1)}$ of the previous moment. Its value is in the range of (0,1) in each dimension. The state $c_{(t-1)}$ of the previous moment is then multiplied by the f vector. Information in dimensions whose values are close to 0 is forgotten and information in dimensions whose values are close to 1 is retained (Graves 2012). The input gate is based on x_t. $h_{(t-1)}$ decides which information needs to be added to the state $c_{(t-1)}$ to generate a new state c_t. The output gate determines the output h_t based on the latest state c_t, the output of the last moment $h_{(t-1)}$, and the current input x_t at that moment. The specific structure of the LSTM unit is shown in Figure 2.

The functions used by each of these nodes are defined as follows:

$$\text{candidate} = \tanh(W_z[h_{t-1}, x_t]) \tag{4}$$

$$i = \text{sigmoid}(W_i[h_{t-1}, x_t]) \tag{5}$$

$$f = \text{sigmoid}(W_f[h_{t-1}, x_t]) \tag{6}$$

$$o = \text{sigmoid}(W_o[h_{t-1}, x_t]) \tag{7}$$

$$c_t = f \cdot c_{t-1} + i \cdot z \tag{8}$$

$$h_t = o \cdot \tanh(c_t) \tag{9}$$

where Equations (4)–(9) are the input state value (candidate), input gate (i), forget gate (f), output gate (o), new state (c_t), and output (h_t) in the nodes, respectively (Graves et al. 2013). W_z, W_i, W_f and W_o are four parameter matrices with dimensions $[2n, n]$, and the dimension of n is the dimension of state c.

Differing from the ordinary linear and nonlinear methods, LSTM networks do not simply impose an element-by-element nonlinearity after affine transformation of the input and recurrent elements. Instead, they incorporate not only external recurrent neural networks, but also internal self-cycling. Self-cycling exists in the state unit, whose weight is controlled by the forget gate. The accumulated time scale can be dynamically changed by the input sequence, and it uses different functions to calculate the state of the hidden layer (Graves, 2012). Therefore, considering the long intervals in sunspot time series, the use of the LSTM network structure is appropriate for the prediction of future solar activity.

2.3 Optimization of LSTM Parameters

The parameters of the LSTM network structure include the number of hidden nodes, the batch size, and the network layers. Hidden nodes are located in the hidden units of the LSTM network. These nodes are called hidden nodes because the training data do not specify the outputs needed by these nodes. The number of hidden nodes directly affects the performance of the network. The network may fail to train or the network performance will be poor when the number of hidden nodes is too small. In addition, if the number of hidden nodes is too large, the training time of the network will be prolonged, and the network can easily fall into local minima, although the error of the network will be reduced.

The batch size is the amount of data that the network traverses for each training pass. Some data can be trained at the same time by setting a certain batch size. However, when the batch size is too small the multi-layer neurons and nonlinear networks in LSTM will modify the parameters in the gradient direction of their samples each time, which means that the model will not easily reach convergence. Increasing the batch size within a certain range can improve memory utilization and parallelization efficiency, but the number of training rounds required to

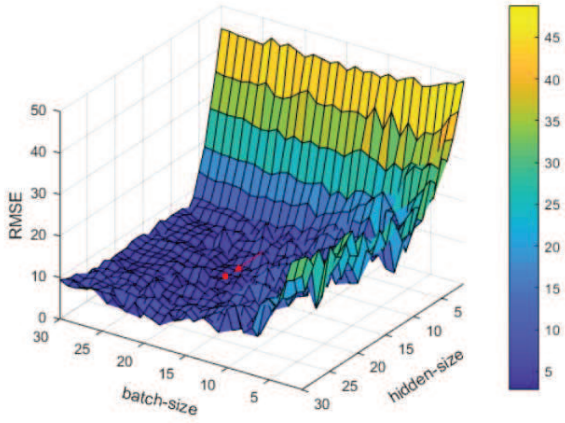

Fig. 3 Parameter determination of the network model.

Table 2 Parameter Setting of the Network Model

HID_S	NUM_L	TIM_S	TRA_S	BAT_S	SAM_G
19	2	10	10 000	20	0.01

HID_S is the number of hidden nodes in LSTM, NUM_L is the number of layers in LSTM, TIM_S is the length of the training sequence of the RNN, TRA_S is the number of training rounds, BAT_S is the size of the batch, and SAM_G is the sampling interval.

achieve the same accuracy will increase with an increase in the batch size.

The number of LSTM network layers refers to the number of layers between the input and output at each time step. This differs from a deep feedforward neural network, as a deep RNN does not stack too many layers between input and output. In the sequential processing of RNNs, an increase in the numbers of layers results in an exponential growth of the time and memory overheads. The disappearance of the gradient becomes very obvious, and the convergence effect and efficiency drop sharply. The local minima problem can also be encountered when the number of layers exceeds three. Therefore, a deep RNN with two to three layers is a large-scale network.

In order to train a better network model, we evaluated the impact of the hidden node number and batch size on the model's performance. The other parameters were left unchanged, and we changed the number of hidden nodes and the batch size in a range from 1 to 30. The results are shown in Figure 3.

From the results, the error after training the network model is minimized when the number of hidden nodes is 19 and the batch size is 20. Therefore, we chose 19 for the number of hidden nodes and 20 for the batch size for the subsequent experiment. Due to the number of nodes in each hidden layer being the same in LSTM (Graves 2012), the hidden nodes of each layer number are 19 in this application. These parameters are listed in Table 2, including the other network model parameters.

3 DATASETS AND METHOD

The data types of sunspot numbers include annual mean value, monthly mean value, and smoothed monthly mean value. The SSN is the average of the 13 months before and after the current month, and with a weighting of half for the first month and the 13th month, which is also called the 13-month smoothed sunspot number. Solar activities include sunspots, solar wind, solar proton events, etc. Sunspots are caused by the large difference between the temperature of some areas of the Sun's surface and the surrounding area. Their change is the most basic phenomenon in solar activity. The number of sunspots changes periodically, and the state of the solar cycle is usually described by the magnitude and change of the SSN.

The data we used were the version 2.0 SSN data from the year of 1750 onwards, downloaded from $sidc.oma.be/silso/datafiles$. In order to analyze the model errors and explore the influence of sample size on the prediction results, we first selected the data of different periods to train the LSTM network to predict the 24th solar cycle, and obtained four different prediction models. We then analyzed the model errors and selected the appropriate data as training sets to model and forecast Cycles 22 and 23 according to the forecast results, and we analyzed the root-mean-square error (RMSE) of the established model and the relative error of the predicted amplitude. Finally, we forecast the future via one-step and multi-step predictions using the LSTM model, and predicted the peak amplitude value and occurrence time of Cycle 25.

We took the observation sequence $O(i-9)$, $O(i-8)$,$\cdots$,$O(i)$ in the first 10 months as the input of the network, and the value $O(i+1)$ in the next month was trained as the output of the network, for the one-step prediction. We then took the observation sequence $O(i-719)$,$O(i-718)$,$\cdots$,$O(i)$ in the first 720 months as the input of the network, and the values $O(i+1)$ in the next 72 months were trained as the output of the network, for the multi-step prediction. A prediction model was then established. The known data were selected as the test set of the model, and the RMSE of the test set data and the relative error of the predicted amplitude were used as the evaluation indicators. The RNN model takes a very short time to process the time series of pure numbers, so the time complexity is generally not considered in the prediction of sunspot numbers.

Similar to the training process, the prediction took the observation sequence of the first some months as the input of the network and the value of some future months as the output. For example, we used the data series of 10 months' observations from March 2008 to December 2008 (i.e. $O(i-9)$ to $O(i)$) as input of the model, and January 2009

Table 3 RMSE and Amplitude Prediction Error for the Cycle 24 One-step Prediction Using the Data from Different Time Series

SAMPLE	1750C-2008	1770–2008	1790–2008	1810–2008
RMSE	2.45	2.96	3.68	3.86
PDRM	2.10%	2.42%	4.78%	5.28%

(i.e. $P(i+1)$) as output of the model in one-step prediction. The output of the model was treated as the correct output, which was fed back to the hidden unit of the model (Graves A, 2012). The information passed back to the hidden unit can be saved as a specific value to the hidden unit and propagated into the future. However, it is impossible to predict the next month's SSN based only on the first few months of SSN data using the trained LSTM model. The model needs to combine all the previous SSN data to make the next prediction (Graves 2012). Since the initial parameters of each network training pass were random, each result is the average of 10 runs, to minimize the influence of the initial parameters on the forecast results.

4 RESULTS

4.1 One-step prediction of SSN

Simulated forecasts for the Cycle 24 SSN using the data from 1750–2008, 1770–2008, 1790–2008, and 1810–2008 were obtained using the LSTM model. The results for the different time series are shown in Figure 4(a)-(d), and their RMSEs are listed in Table 3. It can be seen that the RMSE and amplitude prediction error of the established model increase as the number of samples decreases. This is the feature of deep learning, where the larger the training sample set, the lower the error of the established prediction model. The RMSE obtained using the samples from 1750 to 2008 is 2.45, and the amplitude prediction error is 2.1%, which represent the best prediction results. Therefore, the data before the forecast solar cycle were selected as the training samples in the next forecasting experiment. We used the SSN data from 1750 to 1985, 1750 to 1996, 1750 to 2008, and 1750 to 2018 as training data sets to forecast Cycle 22 (1986–1996), Cycle 23 (1997–2007), Cycle 24 (2009–2019), and Cycle 25 (2019–2029), respectively.

It can be seen that the predicted value and the observed value are in good agreement when the observed value does not change much. However, the predicted value is not as accurate when the observed value changes significantly (e.g. near the peak amplitude of the solar cycle). There is also an inevitable problem of postponement near the peak amplitude of the solar cycle, with a postponement period of about 1.5 months.

Table 4 Statistical Results for the One-step Amplitude Predictions for Cycles 22–24 Using the LSTM Model

SC	RM	RMp	DRM	PDRM	RMSE
22	212.5	119.3	13.2	6.2%	6.12
23	180.3	178.8	1.5	0.8%	4.28
24	116.4	118.8	2.2	2.1%	2.45

SC is the solar cycle number, RM is the actual observation value of the SC, RMp is the forecast value of the solar cycle, DRM is the absolute error of the forecast value, PDRM is the relative error of the forecast value, and RMSE is the LSTM model root-mean-square error.

Table 5 RMSE and Amplitude Prediction Error for the Cycle 24 Multi-step Prediction by Setting

TIMESTEPS (MON_PRED)	1200 (120)	1080 (108)	960 (96)	840 (84)	720 (72)
RMSE	46.2	28.3	18.4	11.2	6.1
PDRM	18.6%	10.7%	8.9%	6.1%	3%

Table 6 Statistical Results of the Amplitude Multi-step Prediction for Cycles 22–25 Using the LSTM Model

SC	RM	RMp	DRM	PDRM	RMSE
22	212.5	175.9	36.6	17.2%	35.3
23	180.3	167.8	12.5	6.9%	28.8
24	116.4	112.8	3.6	3.0%	12.1
25	—	114.3	—	—	—

Figure 5(a)-(c) are the results of the one-step prediction predictions for Cycle 22 to Cycle 24 using the LSTM model. The amplitude prediction error and RMSE are listed in Table 4. The forecast results are generally low when forecasting Cycle 22, where the final amplitude forecast error is 6.2% and the RMSE reaches 6.12. On the one hand, this may be due to the use of fewer training samples or, on the other hand, it could be due to the complexity of Cycle 22 itself. The amplitude deviation of Cycle 23 is the smallest, at only 1.5 difference from the observed value, the relative error is 0.8%, and the RMSE is also reduced to 4.28. With an increase of training samples, although the amplitude forecast error of the 24th solar activity week is slightly higher than that of the 23rd solar activity week, reaching 2.1%, the RMSE is reduced to 2.45 compared with Cycles 22 and 23.

4.2 Multi-step prediction of SSN

The RMSE and PDRM results in Table 5 were obtained by setting the timesteps to 1200, 1080, 960, 840, and 720 to predict the SSN of Cycle24. The timesteps and prediction timestep are generally in a ratio of 10:1 when setting timesteps and prediction timesteps (Graves 2012). Therefore, the numbers of predicted months corresponding to the different timesteps are 120, 108, 96, 84, and 72, respectively.

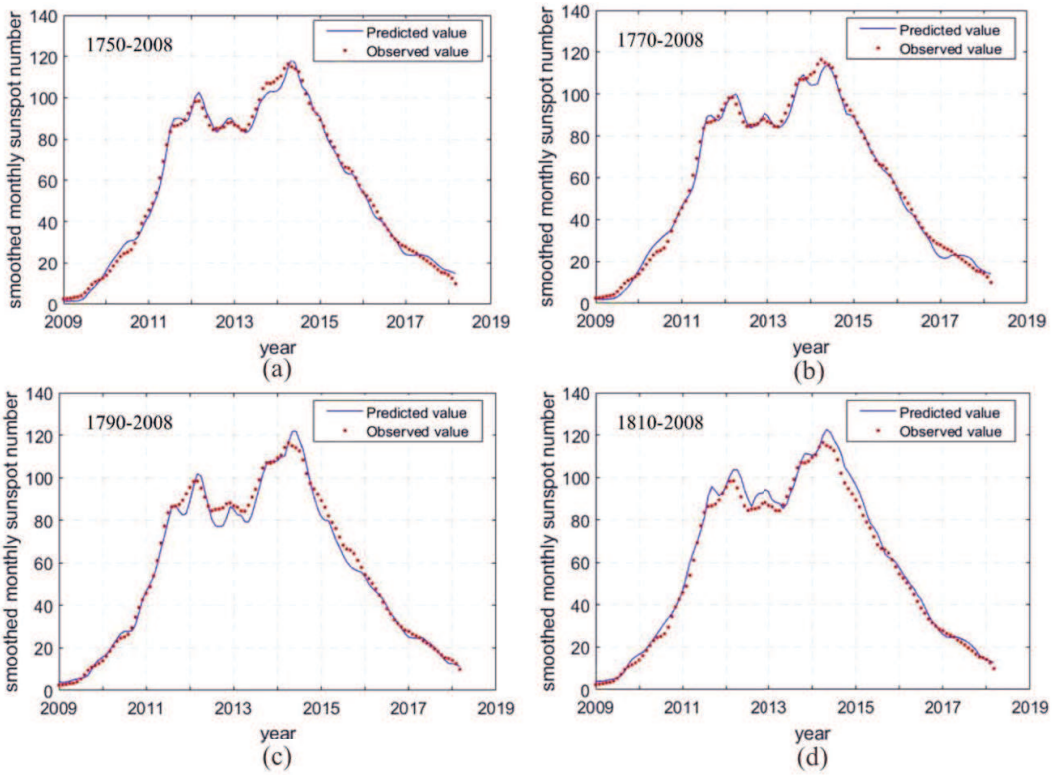

Fig. 4 One-step prediction results for Cycle 24 using the data from different time series.

Table 7 Predicted Peak Year and Peak Value of Cycle 25

	Rigozo et al. (2011)	Helal & Galal (2013)	Miao et al. (2015)	Our forecasting
Peak year	2023	2023	2024	2023
Peak value	132.1	118.2	119.2±5.5	114.3

It can be seen that the RMSE and amplitude prediction error of the established model are reduced with the decrease of the timesteps. The prediction effect is better when the timestep number is 720. The RMSE is then 6.1 and the amplitude prediction error is 3.0%. It would be impossible to predict the peak value of the next solar cycle if the value of the timesteps continues to decrease, because the peak value of the next solar cycle is predicted to be at least 72 months. Therefore, we set the value of the timesteps to 720 in the subsequent experiments.

Figure 6(a)-(d) are the results of the multi-step predictions for Cycle 22 to Cycle 25 obtained using the LSTM model. The amplitude prediction error and RMSE are listed in Table 6. It can be seen from the results that there are some deviations between the predicted values and the observed values. The predicted values for the first two years are generally larger than the observed values, and the predicted values for the last three years also deviate to some extent. This is because the principle of establishing a prediction model with the LSTM model is to minimize the overall deviation between the output of the network and the observed value (Graves 2012). This reflects the overall trend of the SSN. In the first six years of Cycle 22, the amplitude prediction error is 17.2% and the RMSE is 35.3. The amplitude deviation for Cycle 24 is the smallest, at only 3.6 difference from the observed value. The relative error is 3.0% and the RMSE is also reduced to 12.1. Finally, the next solar cycle's (Cycle 25) peak amplitude is predicted to occur around 2023, with a peak value of about 114.3. A number of scholars have forecast Cycle 25, such as Rigozo et al. (2011) (113.3), Helal & Galal (2013) (118.2), and Miao et al. (2015) (119.2±5.5). Our results for the peak value are consistent with those of these researchers. The different results are listed in Table 7.

5 CONCLUSIONS AND DISCUSSION

The solar cycle has been predicted by many researchers; the results of long-term forecasting (especially for Cycle 24) are highly biased, and the results obtained by different methods are notably different. The reason for this may be the extremely complex and variable character of the solar activity itself.

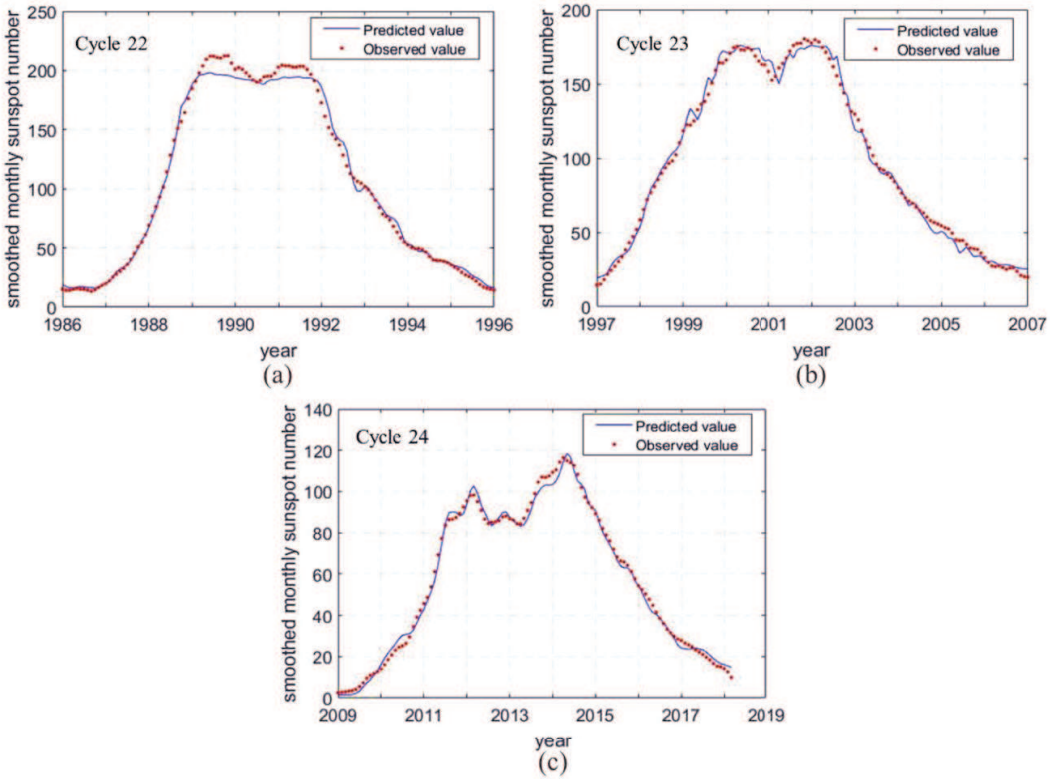

Fig. 5 One-step prediction results for Cycles 22–24 using the LSTM model.

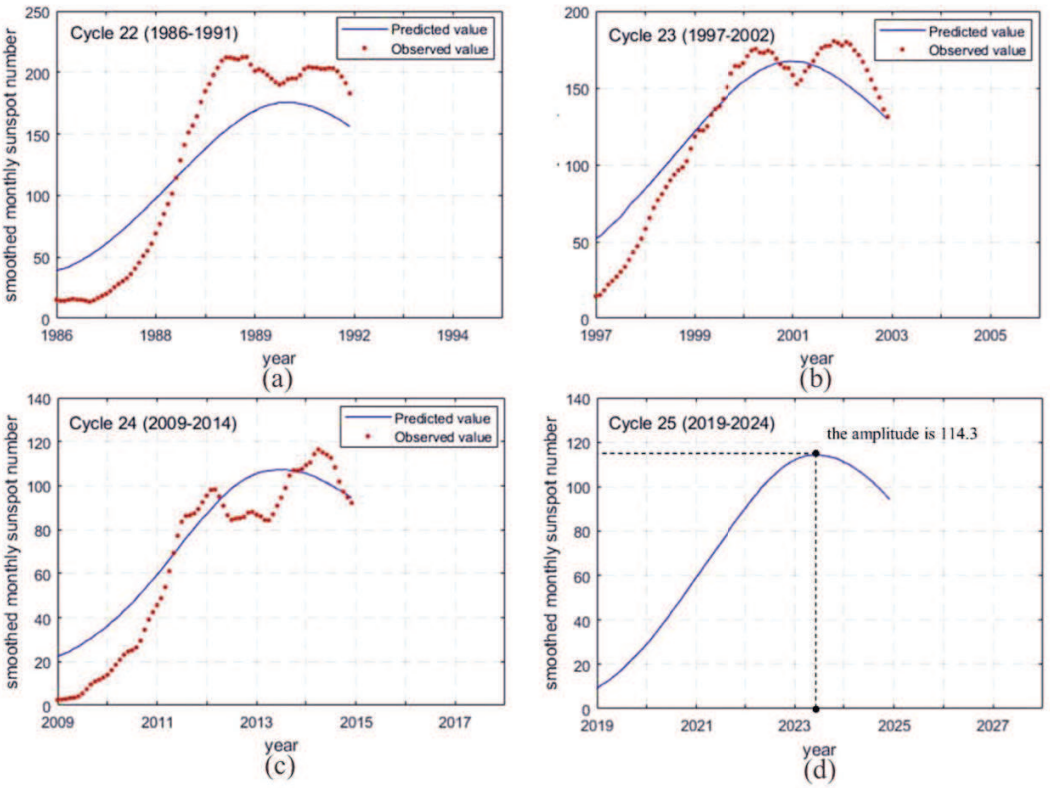

Fig. 6 Multi-step prediction results for Cycles 22–25 using the LSTM model.

In this study, an optimized LSTM network was trained and used to forecast the SSN. For Cycles 22–24, the maximum RMSE of the one-step prediction model was 6.12 and the minimum was only 2.45. The maximum prediction error of the multi-step prediction of the amplitude was 17.2% and the minimum was only 3.0%. Finally, the next solar cycle's (Cycle 25) peak amplitude was predicted to occur around 2023, with a peak value of about 114.3.

The proposed LSTM deep learning model can not only predict the peak amplitude value of the solar cycle, but also the time when the solar cycle reaches its peak amplitude; previous years of solar cycle data are not required for the forecast. However, the disadvantage is that sufficient training samples are required to adjust the network parameters before training the network, and the network does take some time to train.

Acknowledgements The sunspot number data were provided by SILSO data/image, Royal Observatory of Belgium, Brussels. This work was supported by the National Natural Science Foundation of China (Grant No. U1531128).

References

Ahluwalia, H. S., & Jackiewicz, J. 2012, Advances in Space Research, 50, 662

Ajabshirizadeh, A., Masoumzadeh Jouzdani, N., & Abbassi, S. 2011, RAA (Research in Astronomy and Astrophysics), 11, 491

Attia, A.-F., Ismail, H. A., & Basurah, H. M. 2013, Ap&SS, 344, 5

Choudhuri, A. R., Chatterjee, P., & Jiang, J. 2007, Phys. Rev. Lett., 98, 131103

Dabas, R. S., & Sharma, K. 2010, Sol. Phys., 266, 391

Gholipour, A., Lucas, C., Araabi, B. N., & Shafiee, M. 2005, Journal of Atmospheric and Solar-Terrestrial Physics, 67, 595

Graves, A. 2012, Supervised Sequence Labelling (Springer Berlin Heidelberg)

Graves, A., Mohamed, A.-r., & Hinton, G. 2013, arXiv e-prints, arXiv:1303.5778

Han, Y.-B., Yin, Z.-Q., & Wang, B. 2018, Chinese Science Bulletin, 63, 311

Helal, H. R., & Galal, A. A. 2013, Journal of Advanced Research, 4, 275

Hochreiter, S., & Schmidhuber, J. 1997, Neural Computation, 9, 1735

Kitiashvili, I. N. 2016, ApJ, 831, 15

Lukosevicius, M., & Jaeger, H. 2009, Computer ence Review, 3, 127

Miao, J., Gong, J., Li, Z., & Ren, T. 2015, Scientia Sinica Physica, Mechanica & Astronomica, 45, 099601

Muñoz-Jaramillo, A., Balmaceda, L. A., & DeLuca, E. E. 2013, Phys. Rev. Lett., 111, 041106

Nandy, D. 2002, Ap&SS, 282, 209

Noble, P. L., & Wheatland, M. S. 2012, Sol. Phys., 276, 363

Pesnell, W. D. 2008, Sol. Phys., 252, 209

Pesnell, W. D. 2014, Sol. Phys., 289, 2317

Pesnell, W. D. 2016, Space Weather, 14, 10

Prabhakaran Nayar, S. R., Radhika, V. N., Revathy, K., & Ramadas, V. 2002, Sol. Phys., 208, 359

Rigozo, N. R., Souza Echer, M. P., Evangelista, H., Nordemann, D. J. R., & Echer, E. 2011, Journal of Atmospheric and Solar-Terrestrial Physics, 73, 1294

Sabarinath, A., & Anilkumar, A. K. 2018, Journal of Earth System Science, 127, 84

Saha, S., & Raghava, G. P. S. 2010, Proteins-structure Function & Bioinformatics, 65, 40

Saratha, S., & Tajuddin, W. A. W. A. 2008, Modern Applied Science, 2, 57

Solanki, S. K. 2003, A&A Rev., 11, 153

Usoskin, I. G. 2008, Living Reviews in Solar Physics, 5, 3

Yamamoto, H., Isogai, S., & Sagisaka, Y. 2001, in ACL '01: Proceedings of the 39th Annual Meeting on Association for Computational Linguistics, 531