

# Predicting the 25th solar cycle using deep learning methods based on sunspot area data

Qiang Li<sup>1</sup>, Miao Wan<sup>1</sup>, Shu-Guang Zeng<sup>1</sup>, Sheng Zheng<sup>1</sup> and Lin-Hua Deng<sup>2</sup>

<sup>1</sup> College of Science, China Three Gorges University, Yichang 443002, China; [zengshuguang19@163.com](mailto:zengshuguang19@163.com)

<sup>2</sup> Yunnan Observatories, Chinese Academy of Sciences, Kunming 650216, China

Received 2020 December 22; accepted 2021 March 22

**Abstract** It is a significant task to predict the solar activity for space weather and solar physics. All kinds of approaches have been used to forecast solar activities, and they have been applied to many areas such as the solar dynamo of simulation and space mission planning. In this paper, we employ the long-short-term memory (LSTM) and neural network autoregression (NNAR) deep learning methods to predict the upcoming 25th solar cycle using the sunspot area (SSA) data during the period of May 1874 to December 2020. Our results show that the 25th solar cycle will be 55% stronger than Solar Cycle 24 with a maximum sunspot area of  $3115 \pm 401$  and the cycle reaching its peak in October 2022 by using the LSTM method. It also shows that deep learning algorithms perform better than the other commonly used methods and have high application value.

**Key words:** Sun: activity — Sun: solar cycle prediction — Sun: sunspot area — Method: deep neural network

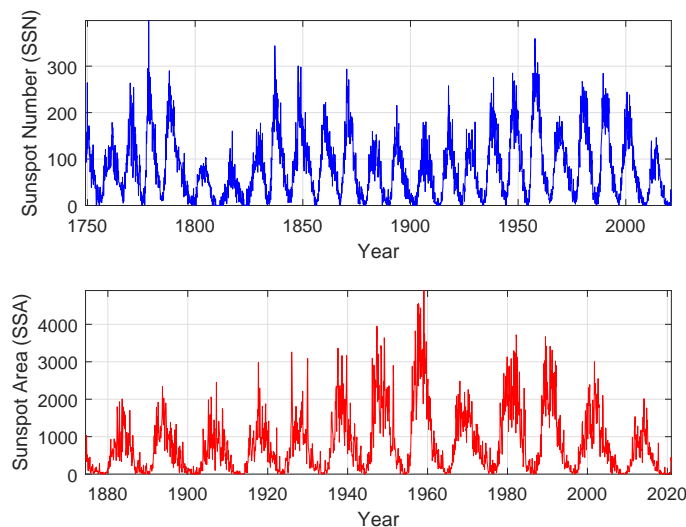
## 1 INTRODUCTION

Solar activity is closely related to human activities and many other phenomena on Earth. With the 11-year solar cycle rises and falls, such as space climate, space navigation and high-frequency radio communications will be affected by such changes (Pala & Atici 2019). When the solar activity is strong, the ultraviolet and X-ray radiation from the Sun will be enhanced, which will have a serious impact on the solar-terrestrial space and the upper atmosphere of the Earth. Furthermore, the intense solar activity may also lead to serious solar storms, which can bring about enormous damage to satellites around the Earth (Pulkkinen 2007). Serious solar storms can even bring great problems to our daily life because of the impact on the communication system or power grid. Therefore, predicting solar activity is of great significance not only for social technologies, but also for the understanding of the mechanism of solar activity.

The sunspot area (SSA) is a good indicator of magnetic activities of the Sun in long time series, which can be applied to predict the sunspot cycle like the same as the sunspot number (SSN) (Hathaway 2015). So far, it is one of the longest observed activity index, which has more physical significance than the SSN. Besides, compared

with SSN, the SSA has appended information about the position of the characteristic disk. Hence, SSA records play an important role in our understanding of the long-term behavior of solar magnetic activity and variability (Mandal et al. 2020).

Traditionally, there are three main prediction methods (Petrovay 2020). Many authors have used these methods to predict the upcoming solar cycle. The first method is the precursor method, which forecasts the maximum amplitude of the next solar activity based on the measured values of solar activity or magnetic field at a specific period. For example, Dabas & Sharma (2010) used geomagnetic precursors to predict Solar Cycle 24. Muñoz-Jaramillo et al. (2013) improved the solar cycle prediction on the basis of the polar magnetic fields by using the dipolar and quadrupolar moments. The second method is model-based prediction, which not only analyzes the observed data, but also uses various physically dynamo models to forecast solar activity (Petrovay 2020). For instance, Choudhuri et al. (2007) input the solar polar magnetic field data into the solar dynamo model to simulate the last few solar cycles, and they forecasted that the 24th solar cycle will be about 35% weaker than cycle 23. Jiang & Cao (2018) applied a surface flux transport (SFT) model to calculate the correlation of some important



**Fig. 1** Monthly distribution of the the sunspot number (SSN) (*top panel*) and sunspot area (SSA) data (*bottom panel*) from the Royal Observatory, Greenwich (RGO) USAF/NOAA during the period of May 1874 –December 2020.

properties of solar cycles to forecast the general trends of the next cycle. Upton & Hathaway (2018) predicted the solar cycle 25 based on the Advective Flux Transport (AFT) model. The last one is the extrapolation method, contrary to the precursor method, it only uses sunspot numbers or other solar activity indices time series data, but it generally depends on more than one former point to determine trends that can be applied to infer future data: such as Brajša et al. (2009) applied the ARMA model in R language to predict the trend of annual value sunspot number series. Rigozo et al. (2011) used the extrapolation of spectral components to evaluate the intensity of solar cycle 25. Noble & Wheatland (2012) used a Bayesian method for forecasting solar cycles, and Sarp & Kılıçık (2018) predicted Solar Cycle 25 using a nonlinear approach.

In the last few years, deep learning methods have developed rapidly and used extensively in various fields due to their powerful capabilities and flexibility. Therefore, deep learning methods are also applied to forecast time series except some classical methods. The advantage of deep learning methods is that they do not need to make any assumptions about any distribution information, and they have the ability to quickly simulate complex problems (Benson et al. 2020). A large number of studies have shown that deep learning algorithms have better performance than some classical algorithms in dealing with time series prediction, mainly because they have a stronger ability to deal with nonlinear problems. For example, Ajabshirizadeh et al. (2011) applied the Feed Forward Neural Network (FFNN) to predict the Solar Cycle 24. Pesnell (2012) used neural networks trained

on sunspot numbers to predict a solar cycle. Attia et al. (2013) employed the neural fuzzy approach to evaluate the geo-magnetic activity of the 25th solar cycle. Lately, Benson et al. (2020) combined the WaveNet and LSTM neural networks to forecast the 25th solar cycle.

In this paper, we make use of a monthly SSA time series during the period of May 1874 to December 2020 to make an estimation of SSA for Solar Cycle 25 with the help of deep learning algorithms. The structure of this paper is as follows: Section 2 introduces datasets and methods; the experimental analysis is shown in Section 3; Section 4 is a summary of this paper.

## 2 DATASETS AND METHODS

### 2.1 Datasets

In this work, we use the monthly SSA data from the Royal Observatory, Greenwich (RGO) USAF/NOAA during the period of May 1874 to December 2020. This time series is publicly available on the website (<http://solarcyclescience.com/activerregions.html>). The dataset is composed of a period of 146 years which includes 1760 months. According to the data features, there are some cyclic patterns, that is, the observed values rise and fall in a certain period. Figure 1 shows the monthly distribution of the SSN (top panel) and SSA data (bottom panel) from the Royal Observatory, Greenwich (RGO) USAF/NOAA during the period of May 1874 –December 2020.

**Table 1** The Cross-validation Scheme of SSA Datasets

Piece number of dataset	Piece length	Rolling-piece length	Training length	Validation length	Testing length	Skip-span length
4	1200	186	80[year]× 12=960	960/8=120	10[year]× 12=120	15.5[year]× 12=186
9	840	115	50[year]× 12=600	600/5=120	10[year]× 12=120	9.6[year]× 12=115

## 2.2 Methods

### 2.2.1 Long-Short-Term Memory (LSTM) Method

The Long-Short-Term Memory (LSTM) method is firstly proposed by Hochreiter & Schmidhuber (1997), which belongs to one of the Recurrent Neural Networks (RNNs), overcoming the vanishing gradient problem exhibited of RRNs. The special structural design allows it to avoid long-term dependency problems, remember information longer than RNNs, and present good capability for time series forecasting (Goodfellow et al. 2016). The key to LSTM is the cell state, which is used to store the current LSTM state information and transfer it to the next LSTM at the next moment. In addition, LSTM mainly consists of three different gate structures, which are the input, output and forget gate. These three gates are used to control the LSTM retention and transfer of information, ultimately reflected in the cell state and output signals. Firstly, the input gate determines which information needs to be added to the cell state to generate a new state. Then, the forget gate decides what information we discard from the cell state. Finally, the output gate determines the output value.

### 2.2.2 Neural Network Auto-Regressive (NNAR) method

The NNAR model is a non-linear parametric prediction model. There are mainly two steps in predicting by this model. The first step is to determine the order of autoregression. The order of auto-regression is the number of previous values on which the current value of a time series depends. Then, in a second step, training neural network by using training set which considers the order of autoregression. The order of auto-regression then decides the input node's number, and the inputs of neural network model is the former, lagged observation values. The neural network model's output value is the predicted value. Because of the lack of a theoretical basis for selection, the hidden node's number is usually decided by trial and experiment. To avoid the problem of over fitting, the number of iterations should be selected properly (Sena & Nagwani 2016).

## 2.3 Evaluation Index

The root mean square error (RMSE) is generally applied to examine the error of the real and predicted values, which makes an excellent general purpose error metric for numerical prediction (Adhikari & Agrawal 2013). It is defined as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (1)$$

where  $N$  is the observation numbers,  $y_i$  is the real value and  $\hat{y}_i$  is the predicted value. Compared to the similar Mean Absolute Error (MAE), RMSE amplifies and severely punishes large errors.

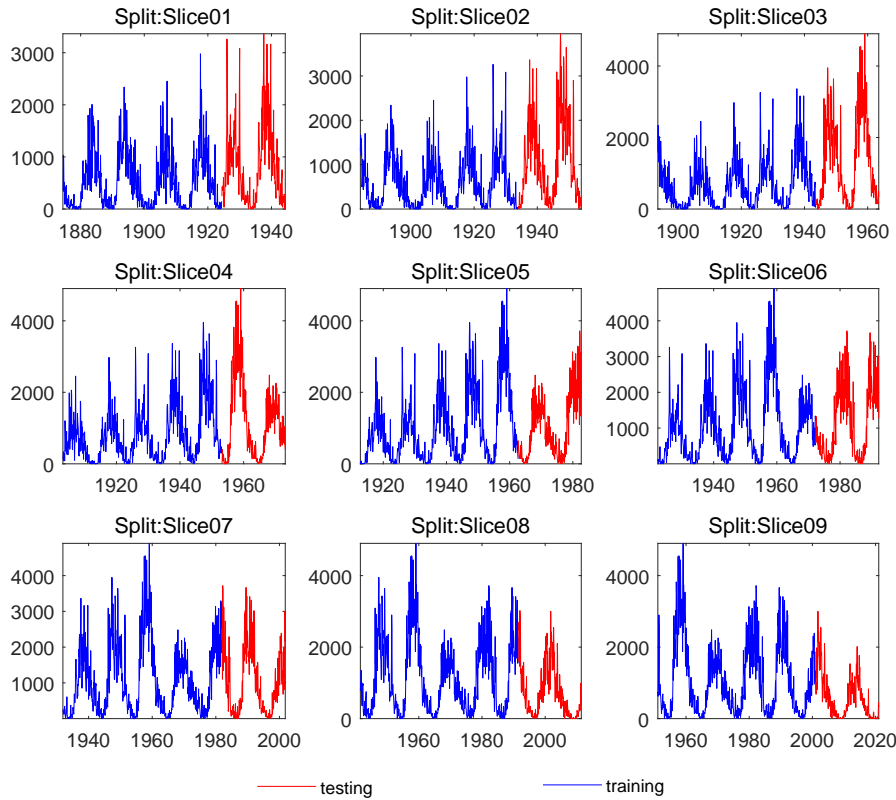
## 3 EXPERIMENTAL ANALYSIS

### 3.1 Data Preprocessing

To analyze the model errors and explore the influence of sample size on the prediction results, we chose the cross validation (CV) approach to improve the prediction accuracy. Using this approach, we analyzed 1760 observational data by using the resampling method of rolling origin prediction. Table 1 displays the 4-piece and 9-piece CV evaluation schemes of the SSA dataset, including a total of 1760 months of observation data.

From Table 1, we can easily see that every part of the sunspot area dataset is set up as 9-piece-based, and is divided into two parts as training and testing. Apart from that, one eighth or fifth of the dataset is arranged for training, which is reserved to the validation process in the 9-slice-based model and 4-slice-based model, respectively. After these processes, every part of the dataset is made up of the training, validation, and testing set. Their functions are to train the LSTM model, set the model hyperparameters and check the real model performance, respectively. Figure 2 shows the sampling size of the training and test set for each slice.

Combining Table 1 and Figure 2, we can see that every slice length of the nine separate pieces of data is 1200 records. Therefore, the first slice01 is composed of data from May 1874 to May 1944. According to the rolling-slice lengths, the second slice02 includes the data during



**Fig. 2** Sunspot area (SSA) dataset divided into nine parts, as training and testing.

**Table 2** The LSTM Model Performance of Different Slice Numbers on SSA Data Set

Slice number of data set	Length of Slice	RMSE value
2	1344	315.94
3	1272	300.22
4	1200	288.61
5	1128	297.38
6	1056	306.35
7	984	312.16
8	912	317.63
9	840	328.67

the period of May 1883 - May 1953. That is, the next piece begins 108 records earlier than the former one. The length of the skip span for the 9-piece-based model and the 4-piece-based schemes is 115 and 186 records, respectively.

### 3.2 LSTM Model Parameters

More often than not, all models need to be trained on a complete dataset to make accurate predictions on the basis of historical observations. However, it is difficult to judge the performance of the model. The reason is that there is no actual situation to verify the prediction. Thus, it needs to divide the dataset into training and validation parts, and

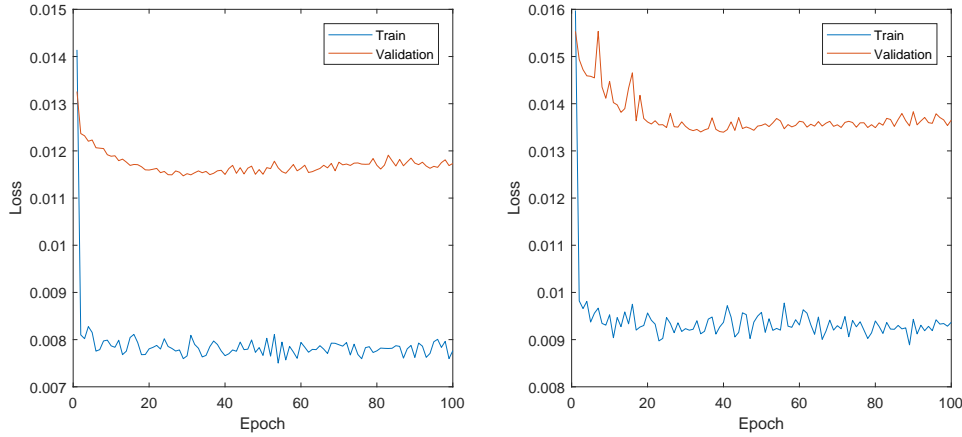
select the model with the best validation performance to be trained on the whole data to make our prediction.

As mentioned above, we split the sunspot area dataset to get a better performance. As shown in the Table 2, we divided the data set into eight different slice numbers. According to the RSME value, we find that the RMSE value decreases initially and increases afterwards as the increase of slice number, and reached the minimum at the 4-slice plan. Therefore, the 4-slice plan revealed greater predictive performance. We selected it as the proper data set, and correspondingly, the proper length of it is 1200.

After these operations, we selected the 4-slice type data set to make the following experiment. Table 3 shows the sunspot area dataset which is segmented into four slices in the second line, the number of training, test is 960, 240, respectively. The batch-size of them is all set to 10. Hence, the number of training and test iterations is 96 and 24, respectively. To get more accurate estimation, the gradient descent algorithm is usually applied to optimize the deep learning model by constantly updating weights through the training dataset. Thus, the number of epochs is set to 100.

**Table 3** The Training and Test Parameters of the LSTM Model

Model cross-validation strategy	Training/test numbers	Batch size	Training /test iteration numbers	Epoch numbers
4 piece-based model	960/240	10	96/24	100
9 piece-based model	600/240	10	60/24	100



**Fig. 3** The loss-epoch graphs of training and validation for 4-piece (left) and 9-piece (right) schemes, respectively.

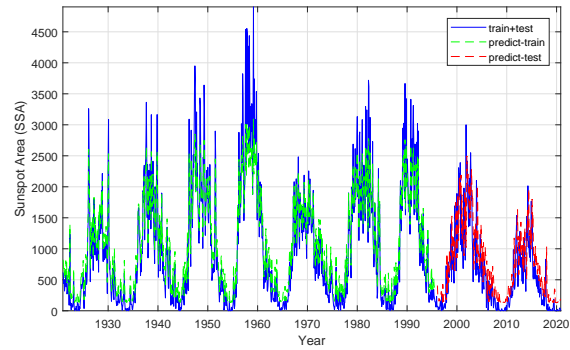
### 3.3 Analysis Results

#### 3.3.1 Results from deep learning methods

As noted above, the sunspot area dataset is split into two different types: 4-piece and 9-piece datasets. Therefore, these two different types of datasets in the LSTM model performed differently. First of all, LSTM model is used to train and test the last part of the 4-piece dataset and 9-piece dataset, respectively. Figure 3 shows the training and validation loss-epoch graphics for 4-slice (left) and 9-slice (right) plans, respectively. It is clear that the loss of the 4-slice plan is smaller than the 9-slice plan. Besides, the curves show that the model stopped the training process at about 20 iterations and remains constant for both plans. Meanwhile, it is easy to find that the training performance of the two schemes is better than verification performance.

Figures 4 and 5 show the training and test predict graphs of the last slices of the 4-slice and 9-slice dataset, respectively. It can be seen that the train predicted value, the train and test value are in good agreement, when the train and test value does not change much. However, the train predicting value is not as accurate when the train and test value change greatly, such as near the peak amplitude of the solar cycle. We can also see that the test predicting result has a very high fitting precision with the train and test dataset in both two types. Comparing Figures 4 and 5, it is clear that the estimation results using the last part of the 4-slice are better than the 9-slice by LSTM model.

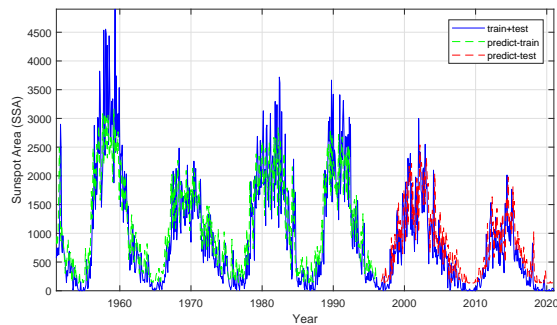
In addition, the RMSE is the standard deviation of the residuals or prediction errors, which offers a measure of



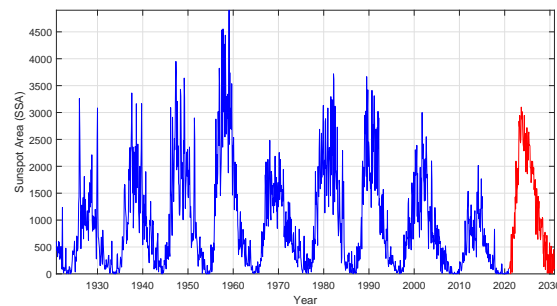
**Fig. 4** The predicting graphs for training and test of the last part of the 4-piece SSA dataset using the LSTM method (train and test: blue, train predict: green, test predict: red).

the distance between prediction and real data. Therefore, almost all experiments will choose it to estimate the performance of every model. From Table 4, it is not hard to find that the RMSE test values of the last slices of the 4-slice and 9-slice types of the sunspot area data were very similar to each other, but the 4-slice type revealed greater predictive performance. The RMSE test values for both of them were 288.61 and 328.67, respectively.

Through the above analysis, the prediction results of Solar Cycles 23 and 24 showed that the LSTM model can not only predict the trend of solar cycle but also accurately predict the strength of it for both types of dataset. After a comprehensive consideration, we select the 4-slice sunspot area dataset based on the LSTM model to predict the 25th Solar Cycle. After training, verification, and testing



**Fig. 5** The predicting graphs for training and test of the last part of the 9-piece SSA dataset using the LSTM method (train and test: blue, train predict: green, test predict: red).



**Fig. 6** The 25th solar cycle predict using the LSTM method (actual: blue, prediction: red).

operation, the predicted results for the next ten years (2020–2030) are displayed in Figure 6. In other words, we applied the LSTM model to predict the time period of 2020 to 2030 for the SSA dataset which includes 1760 months of observation data during the period of 1874 to 2020, with a time span of 146 years.

In the LSTM modeling, the selection of hyper parameters, such as the number of hidden units, input lags and hidden layers, is very important for model performance and data segmentation. In this paper, we use two stacked LSTM layers with 256 hidden units each. To select an appropriate hyperparameter, we have tested different numbers of hidden units. The result shows that approximately 256 hidden units can achieve good performance for the LSTM model. Adding more hidden units may increase the value of the RMSE, it is useless to enhance the performance of the model. The first LSTM layer was used to input data and return 3D shapes and the second LSTM layer was used to return the 2D data shapes. At the same time, we adopted the Adam optimization algorithm to optimize network parameters and minimize the training error (Kingma & Ba 2014). It is a step-descent algorithm with an adaptive learning rate. In an epoch, each batch is calculated once, and the network parameters are updated several times. When one epoch is over, the

**Table 4** The Performance of Different Methods on SSA Datasets

Method	RMSE value
LSTM 4-piece method	288.61
LSTM 9-piece method	328.67
NNAR 4-piece method	360.65
NNAR 9-piece method	398.38
ARIMA method	549.35

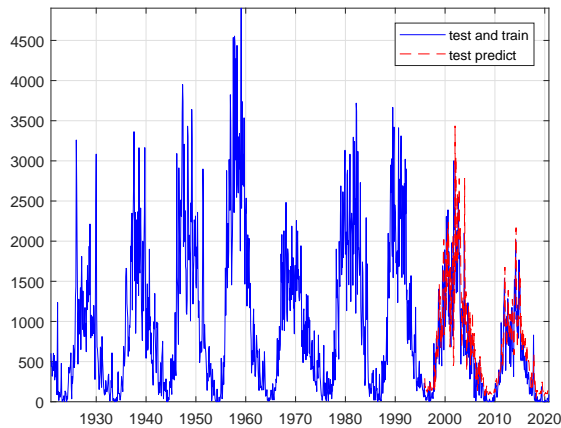
validation set will evaluate the updated network. In this model, the training batch size is 10 with a learning rate of  $10^{-4}$  and a decay rate of  $10^{-6}$  for 96 iterations.

The same strategy for the sunspot area dataset partition was also applied to the Neural Network Auto-Regressive (NNAR) method. Figure 7 and Figure 8 show the training and test predicted graphs of the last slices of the 4-slice and 9-slice plan using the NNAR method, respectively. We can easily see that the 4-slice plan test prediction has a higher fitting precision than the 9-slice plan. Moreover, the value of the RMSE is quite different for them. The RMSE test value for the 4-slice plan are lower, as shown in Table 4. Hence, we select the 4-slice plan to predict the next ten years (2020–2030) trend and strength of the sunspot area, as shown in Figure 9.

### 3.3.2 Comparing with the classical method

The Auto-Regressive Integrated Moving Average (ARIMA) model is a classical statistical method for time series prediction, which has dominated many areas of time series predicting (Box et al. 2013). In this paper, we used an ARIMA (p,d,q) model to predict the the upcoming Solar Cycle 25. To test the reliability of the ARIMA model, we also checked before the 22th cycle using data cycles. Through multiple experiments and analysis, we found that the ARIMA (2,0,5) is the optimal model for predicting the trend and strength of sunspot area with five MA terms and two AR terms. The results for predicting Solar Cycles 23 and 24 are presented in Figure 10. However, we find that the test predicting result has a slight deviation with the train and test dataset. Meanwhile, one can notice that the result for predicting the next ten years (2020–2030) trends and strength of the sunspot area is also different from the two previous deep learning methods, as shown in Figure 11.

In a similar study, Siami-Namini & Siami Namin (2018) used different datasets to measure the LSTM and ARIMA model's performance. By comparing the RMSE values of ARIMA and LSTM, they concluded that the performance of LSTM model is better than ARIMA. Actually, we approve of their view that the LSTM model has outperformed these classical methods for time series forecasting problems. Moreover, Maleki et al.



**Fig. 7** The predicting graphs for training and test of the last part of the 4-piece SSA dataset using the NNAR method (train and test: blue, test predict: red).

**Table 5** Predicting Result of the 25th Solar Cycle of Different Algorithms

Model/method name	Date of the peak occur	Maximum amplitude
LSTM method	2022.10	$3115 \pm 401$
NNAR method	2023.03	$3078 \pm 421$
ARIMA model	2021.09	$1241 \pm 616$

(2018) suggested that the performance of NNAR model is much better than ARIMA. From these studies, it is easy to conclude that the performance of LSTM model is best in relation to the NNAR and ARIMA model. In this paper, Table 4 is the RMSE values of these algorithm test procedures, which also proves this conclusion.

Based on the above analysis, we can conclude that deep learning algorithms perform better than the other commonly used classical methods. Table 5 summarizes the predicting of the 25th solar cycle trend and strength of the three algorithms we used in this study. From Table 5, predicting for the sunspot area dataset suggests that the maximum amplitude of the 25th solar cycle by the LSTM model is  $3115 \pm 401$  with the peak occurring in October 2022. Compared with the 24th solar cycle, the result shows that the 25th solar cycle will be stronger.

### 3.4 Comparison with Earlier Works

Many researchers have predicted the upcoming Solar Cycle 25 using all kinds of methods and data, their results were also notably different. The reason may be the extremely complex variability of solar activity itself. Table 6 lists early prediction results of Solar Cycle 25. For example, Covas et al. (2019) predicted that the 25th solar cycle would be the weakest cycle on record by using

the spatial-temporal data with neural networks, and the peak occurring around 2022–2023. Okoh et al. (2018) used a method called the Hybrid Regression Neural Network to evaluate the SSN. The prediction result showed that the maximum SSN amplitude of the 25th solar cycle was  $122.1 (\pm 18.2)$  with a peak in January 2025 ( $\pm$ six months). Compared with previous cycles, the strength of the 25th solar cycle will be moderate. Furthermore, an international team co-hosted by NOAA/NASA published a preliminary prediction in April 2019, which unanimously forecasted that the size of the 25th solar cycle would be similar with Solar Cycle 24.

However, in this paper we use the monthly sunspot area (SSA) data from the Royal Observatory, Greenwich (RGO) USAF/NOAA during the period of May 1874 to December 2020 to predict the upcoming Solar Cycle 25 with the help of the LSTM model, which shows the maximum amplitude of the 25th solar cycle is  $3115 \pm 401$  and the strength of it will be 55% stronger than Solar Cycle 24. Our result is different with theirs. The most likely reason is that the origin and length of the sunspot area data we used are different, and the selection of the model also has some influence. There are also some previous prediction results that are in agreement with us.

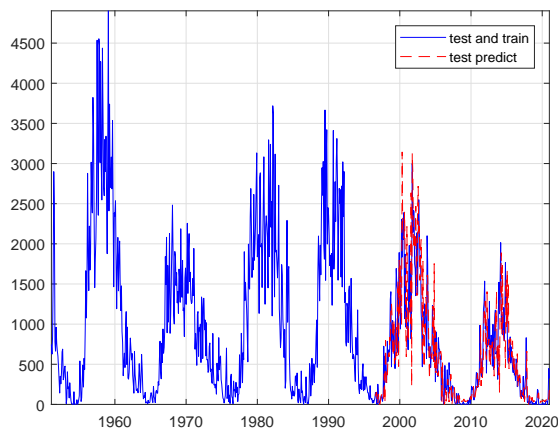
From Table 6 we can easily see that Kane (2007) suggested that the 25th solar cycle will be 29% stronger than the prior solar cycle, and the maximum amplitude of SSN is in the range of 112 to 127 with a peak occurring between 2022 and 2023. Rigozo et al. (2011) used the extrapolation of spectral components to evaluate the strength of the 25th solar cycle. They estimated that Solar Cycle 25 is about 17% stronger than Solar Cycle 24 and has its maximum sunspot number amplitude of 132.1 in April 2023. Dani & Sulistiani (2019) applied the linear regression technique to forecast Solar Cycle 25, and they found the maximum sunspot number amplitude for Solar Cycle 25 is of  $159.4 \pm 22$  in September 2023, about 10% stronger than cycle 24. According to the logarithmic relationship of the solar cycle and geomagnetic indices, Du (2020) found that the peak value of the 25th solar cycle is about  $151.1 \pm 16.9$ , about 30% stronger than cycle 24, and the cycle reaching its peak in October 2024.

## 4 CONCLUSIONS AND DISCUSSION

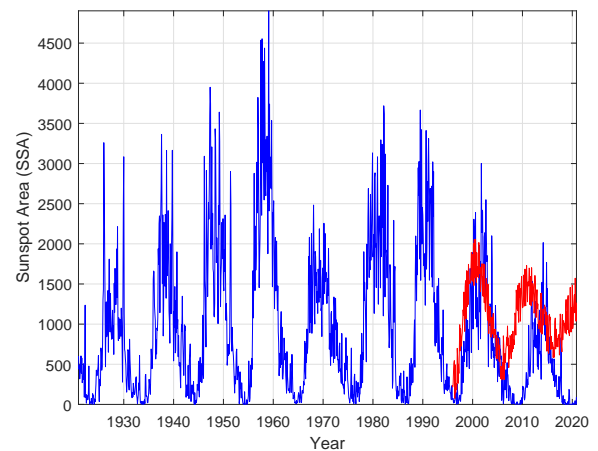
In this study, we present two deep neural network models and a classical method to forecast the 25th solar cycle using the monthly sunspot area (SSA) data from the Royal Observatory, Greenwich (RGO) USAF/NOAA during the period of May 1874 to December 2020. We can easily conclude that the performance of the LSTM model is best in relation to the NNAR and ARIMA model. The ARIMA model requires time series data to be stable. Its

**Table 6** A Selection of Early Prediction for the 25th Solar Cycle

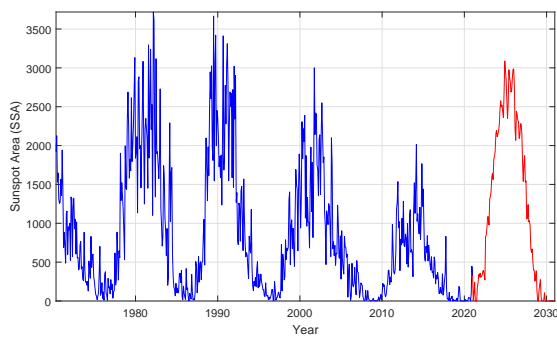
Method	Strength comparing to Cycle 24	Time of reaching the peak amplitude	Reference
LSTM	55% stronger	October 2022	Our works
Feed-Forward Neural Network	weaker	around 2022–2023	Covas <a href="#">Covas et al. (2019)</a>
Hybrid Regression Neural Network	similar	January 2025(±six months)	Okoh <a href="#">Okoh et al. (2018)</a>
Spectral wavelet decomposition tree	17% stronger	April 2023	Rigoz and Echer ( <a href="#">Rigoz et al. 2011</a> )
Similar-cycle method	30% stronger	October 2024	Du <a href="#">Du (2020)</a>
Extrapolation of spectral components	29% stronger	between 2022 and 2023	Kane ( <a href="#">Kane 2007</a> )
Linear Regression (LR)	10% stronger	September 2023	Dani and Sulistiani ( <a href="#">Dani &amp; Sulistiani 2019</a> )



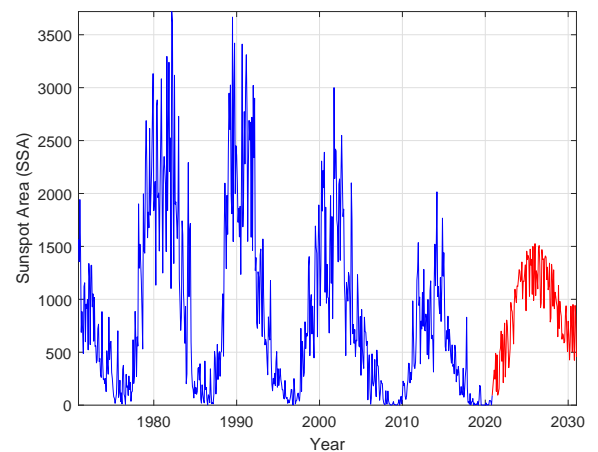
**Fig. 8** The predicting graphs for training and test of the last part of the 9-pieces SSA dataset using the NNAR method (train and test: blue, test predict: red).



**Fig. 10** Training and test predict graphs using the ARIMA model (train and test: blue, test predict: red).



**Fig. 9** The 25th solar cycle predict using the NNAR method (actual: blue, prediction: red).



**Fig. 11** The 25th solar cycle predict using the ARIMA model (actual: blue, prediction: red).

nonlinearity ability is poor, and it can only carry out short-term prediction. The NNAR model has better nonlinearity ability than the ARIMA model, but it does not have the ability of long-term learning. However, the LSTM model can learn and save information for a long time. Besides, the LSTM model tends to do better in unstable time series with more fixed components, which makes the prediction length longer and the result more reliable. Therefore, it is a good forecasting method.

By using the LSTM model, we predict that the maximum SSA amplitude of the 25th solar cycle is  $3115 \pm 401$  and the strength of it will be 55% stronger than Solar Cycle 24 with a peak occurring in October 2022. Our prediction is consistent with some previous predictions, such as [Kane \(2007\)](#), [Rigoz et al. \(2011\)](#),



Dani & Sulistiani (2019) and Du (2020). All of us used different methods and data to obtain the same conclusion that the 25th solar cycle would be stronger than the prior cycle.

Our proposed method also has some disadvantages that we cannot predict time series data for too long future time. Maybe, we can combine LSTM model with other linear and non-linear models. It may improve the length of predicting time. Of course, it can be explored in the future scope of this work. As mentioned above, predicting solar activity is of great significance. We hope that this study would attract more authors to use the deep learning methods widely in predicting tasks in heliophysics.

**Acknowledgements** This work is supported by the National Natural Science Foundation of China under Grant numbers U2031202, U1731124 and U1531247, the special foundation work of the Ministry of Science and Technology of the People's Republic of China under Grant number 2014FY120300, the 13th Five-year Informatization Plan of Chinese Academy of Sciences under Grant number XXH13505-04.

## References

- Adhikari, R., & Agrawal, R. K. 2013, arXiv e-prints, arXiv:1302.6613
- 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
- Benson, B., Pan, W. D., Prasad, A., Gary, G. A., & Hu, Q. 2020, Sol. Phys., 295, 65
- Box, G. E., Jenkins Gwilym, M., & Reinsel, G. C. 2013, Time Analysis Analysis (Wiley)
- Brajša, R., Wöhl, H., Hanslmeier, A., et al. 2009, A&A, 496, 855
- Choudhuri, A. R., Chatterjee, P., & Jiang, J. 2007, Phys. Rev. Lett., 98, 131103
- Covas, E., Peixinho, N., & Fernandes, J. 2019, Sol. Phys., 294, 24
- Dabas, R. S., & Sharma, K. 2010, Sol. Phys., 266, 391
- Dani, T., & Sulistiani, S. 2019, Journal of Physics: Conference Series, 1231, 012022
- Du, Z. L. 2020, Ap&SS, 365, 104
- Goodfellow, I., Bengio, Y., & Courville, A. 2016, Deep Learning 775 (MIT Press, Cambridge)
- Hathaway, D. H. 2015, Living Reviews in Solar Physics, 12, 4
- Hochreiter, S., & Schmidhuber, J. 1997, Neural Computation, 9, 1735
- Jiang, J., & Cao, J. 2018, Journal of Atmospheric and Solar-Terrestrial Physics, 176, 34
- Kane, R. P. 2007, Sol. Phys., 246, 487
- Kingma, D. P., & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980
- Maleki, Afshin, Nasser, et al. 2018, Ksce Journal of Civil Engineering, 22, 1
- Mandal, S., Krivova, N. A., Solanki, S. K., Sinha, N., & Banerjee, D. 2020, A&A, 640, A78
- Muñoz-Jaramillo, A., Balmaceda, L. A., & DeLuca, E. E. 2013, Phys. Rev. Lett., 111, 041106
- Noble, P. L., & Wheatland, M. S. 2012, Sol. Phys., 276, 363
- Okoh, D. I., Seemala, G. K., Rabi, A. B., et al. 2018, Space Weather, 16, 1424
- Pala, Z., & Atici, R. 2019, Sol. Phys., 294, 50
- Pesnell, W. D. 2012, Sol. Phys., 281, 507
- Petrovay, K. 2020, Living Reviews in Solar Physics, 17, 2
- Pulkkinen, T. 2007, Living Reviews in Solar Physics, 4, 1
- 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
- Sarp, V., & Kılıçık, A. 2018, in IAU Symposium, 340, eds. D. Banerjee, J. Jiang, K. Kusano, & S. Solanki, 321
- Sena, D., & Nagwani, N. K. 2016, ARPN Journal of Engineering and Applied Sciences, 11, 13123
- Siemi-Namini, S., & Siemi Namin, A. 2018, arXiv e-prints, arXiv:1803.06386
- Upton, L. A., & Hathaway, D. H. 2018, Geophys. Res. Lett., 45, 8091