Agricultural Sciences, 2020, 11, 777-792
https://www.scirp.org/journal/as
ISSN Online: 2156-8561
ISSN Print: 2156-8553

Research on Dynamic Forecast of Flowering Period Based on Multivariable LSTM and Ensemble Learning Classification Task

Chao Chen, Xingwei Zhang, Shan Tian

College of Communication Engineering, Chengdu University of Information Technology, Chengdu, China
Email: 804928397@qq.com

How to cite this paper: Chen, C., Zhang, X.W. and Tian, S. (2020) Research on Dynamic Forecast of Flowering Period Based on Multivariable LSTM and Ensemble Learning Classification Task. *Agricultural Sciences*, **11**, 777-792.
https://doi.org/10.4236/as.2020.119050

Received: July 6, 2020
Accepted: September 21, 2020
Published: September 24, 2020

Abstract

The flowering forecast provides recommendations for orchard cleaning, pest control, field management and fertilization, which can help increase tree vigor and resistance. Flowering forecast is not only an important part of the construction of agro-meteorological index system, but also an important part of the meteorological service system. In this paper, by analyzing local meteorological data and phenological data of "Red Fuji" apples in Fen County, Linfen City, Shanxi Province, with the help of machine learning and neural networks, we proposed a method based on the combination of time series forecasting and classification forecasting is proposed to complete the dynamic forecasting model of local flowering in Ji County. Then, we evaluated the effectiveness of the model based on the number of error days and the number of days in advance. The implementation shows that the proposed multivariable LSTM network has a good effect on the prediction of meteorological factors. The model loss is less than 0.2. In the two-category task of flowering judgment, the idea of combining strategies in ensemble learning improves the effect of flowering judgment, and its AUC value increases from 0.81 and 0.80 of single model RF and AdaBoost to 0.82. The proposed model has high applicability and accuracy for flowering forecast. At the same time, the model solves the problem of rounding decimals in the prediction of flowering dates by the regression method.

Keywords

Multivariable LSTM, Ensemble Learning Combination Strategy, Random Forest, Adaboost

1. Introduction

Flowering forecast is an important part in the construction of agro-meteorological

index system and meteorological service system. The phenology is used to indicate the seasonal changes of the seasons, the response and adaptation process of the ecosystem to changes in the external environment [1]. The change in the phenology of plants or animals reflects the response of living systems to climate change [2]. Phenology is no longer a pastime of little scientific value [3]. Climate change and crop management measures affect crop growth [4]. Among the agrometeorological elements, the meteorological elements closely related to the production activities of crops are solar radiation, temperature, precipitation, humidity and wind *et al*. These meteorological elements not only provide basic materials and energy for organisms, but also constitute the external environmental conditions that are the growth and development of organisms, with determined yield and quality. Climatic factors influence the phenology of many animal and plant species, rendering them susceptible to the effects of climate change [5]. During the growth and development of the plant in each year, with the seasonal changes of weather and climate, it changes from sprouting, branching and leafing, flowering and fruiting stages to deciduous dormancy and other phenological phenomena with regular changes. The plant phenology model is based on the reaction mechanism of the plant growth and development process to the factors that constitute the climatic factors. It is a mathematical expression used to simulate the plant growth and development process.

According to the number of meteorological elements, the flowering forecast model can be divided into single factor forecast model and multi-factor forecast model. Temperature is a common meteorological factor of flowering and can influence a variety of stages in floral development [6]. Alcalá and Barranco [7] determined the heat storage period based on 10-year phenology and temperature data, and used the heat accumulation threshold to predict flowering time. The method of the heat storage period is the same as the accumulated temperature method. When the accumulation of temperature reaches a certain degree Celsius, it is the initial flowering period of the plant. Temperature-based models include the developmental rate (DVR) model, the chill day model and the new chill day model. The developmental rate (DVR) has been described as a functions of temperature, with the relationships between DVR and temperature being considered linear or exponential [8] [9] [10]. Surgiura and Honjo [11] founded that the DVR model accurately predicted the flowering dates with Root Mean Square Error (RMSE) of 1.23 days. Jong Ahn Chun *et al.* [12] developed the prediction models of full blooming dates for the five peach cultivars at the six major peach cultivation sites using the DVR model, chill day model and new chill day models. Jina Hura and Joong Bae Ahn [13] used the dynamically downscaled daily temperature to research the effect of global warming on the first-flowering data (FFD) of cherry, peach and pear in Northeast Asia. Adnane *et al.* [14] considers the successive and the independent effect of temperature on the dormancy and the induction of ecodormancy in flower buds of fruit trees, using a phenology model based on a sequential model to simulate flowering

dates. Soil moisture is also used to predict the flowering period. Chauhan *et al.* [15] used the APSIM model to predict the effect of soil moisture on the flowering time of chickpea and wheat.

Schneemilch *et al.* [16] used multivariate logistic regression to determine which environmental variables were influential on flowering timing. The resultant models described a large amount of variation in flowering presence or absence, with R^2 ranging from 0.72 to 0.79. Cenci and Ceschia [17] used daily maximum and minimum temperatures, daily rainfall, relative humidity and daily global insolation to establish flowering prediction models $z = \alpha x^{\beta} + \gamma y$. Among them, z represents flowering time that the equation has been simply expressed as the power relationship of the variable x and the linear correlation of the variable y. Park, *et al.* [18] used Elasticnet regression techniques with the quantity of winter and spring precipitation that fell as snow in a given year, the number of frost-free days that occurred in a given winter and spring, the date of the beginning of the frost-free period five climatic parameters to establish a predictable 2320 Species-specific model of angiosperm flowering date. Finding the best meteorological elements (golden feature) is a key step in improving the accuracy of the model for regression models. The flowering prediction model is also a time series prediction model, which is a special regression problem. With the popularization of artificial intelligence big data, the prediction of flowering period has also begun to use neural networks for prediction. Elizondo *et al.* [19] used the back propagation training algorithm of BP neural network to complete the prediction of soybean flowering and physiological maturity. The meteorological factors used include daily maximum, minimum temperature and photoperiod.

Since the prediction result of the regression model contains decimals, and the flowering period is an integer in days, different decimal rounding methods will cause prediction errors. In order to predict as far as possible in the process of regression model prediction, it will cause a long blank period. If meteorological disaster such as low-temperature freezing damage occurs during the blank period, the flowering period will be delayed, thereby affecting the accuracy of the forecast. This paper proposes a flowering period prediction model based on multi-variable LSTM (Long Term short Memory) and ensemble learning classifier to solve the problem between long blank period and Decimal rounding on the prediction process.

2. Materials and Methods

2.1. Experimental Data

In this study, meteorological and phenological data were provided by the Ji county of Linfen City, Shanxi Province. Among them, the meteorological data includes 12 meteorological elements, such as temperature (maximum, minimum, average), precipitation, sunshine time, ground temperature (5 cm, 10 cm, 15 cm) and humidity from 2005 to 2019. The phenological data is the phenolog-

ical data of local apples named "Red Fushi" in Ji county Prefecture from 2010 to 2019.

To determine that a certain day is a flowering day, it is necessary to combine meteorological data and phenological data. In the classification task, the more data dimensions, the better the model, so the two features of the sum of air temperature (SAT) and The sum of geothermal temperature (SGT) are added. The formula is:

$$SAT = MaxT + MinT + AT \tag{1}$$

Among them, MaxT, MinT, AT represent the maximum temperature, minimum temperature, average temperature.

$$SGT = GT5 + GT10 + GT15 \tag{2}$$

Among them, GT5, GT10, GT15 respectively represent 5 cm, 10 cm, 15 cm ground temperature.

Extract the meteorological data from March 25 to April 30 every year from 2010 to 2019, and add tags according to the phenological data of the corresponding year. The flowering observations for each date are converted to binary data (1 = flowering, −1 = non-flowering). In this decade, there are 370 data, including 187 data with label 1 and 183 data with label-1. Basically meet the balance of positive and negative sample data.

2.2. Meteorological Data Analysis

Data quality directly affects the performance indicators of the model. Statistical description helps to discover some obvious problems related to data quality (e.g., missing values, duplicate values, outliers), deepen the understanding of the relationship between data and variables, and provide useful information for subsequent data preprocessing and model selection. Descriptive statistical analysis of meteorological data is shown in **Table 1**. Observe the number, distribution and quartiles of the data. It can be found that from January 1, 2015 to December 31, 2019, there should be 5478 data. However, there are only 2039 precipitation data. There are 5472 data in 5 cm ground temperature and 10 cm ground temperature, but 6 data are missing. There is no missing value for the 15 cm ground temperature data. The average wind speed has 5 missing values. The minimum value of the maximum wind speed during this period was a positive value, and no abnormal value occurred.

Periodicity is a prerequisite for time series forecasting. Before the periodic analysis, the missing data has been filled. The methods are as follows, 1) Fill in the missing values directly as zero according to the reason for the lack of precipitation. 2) For meteorological elements other than precipitation, first find the year corresponding to the missing value, and then infer the corresponding missing value based on the annual data. When there are multiple missing values in the same year, the missing values are treated according to averaging. **Figure 1** shows the periodicity of each meteorological element.

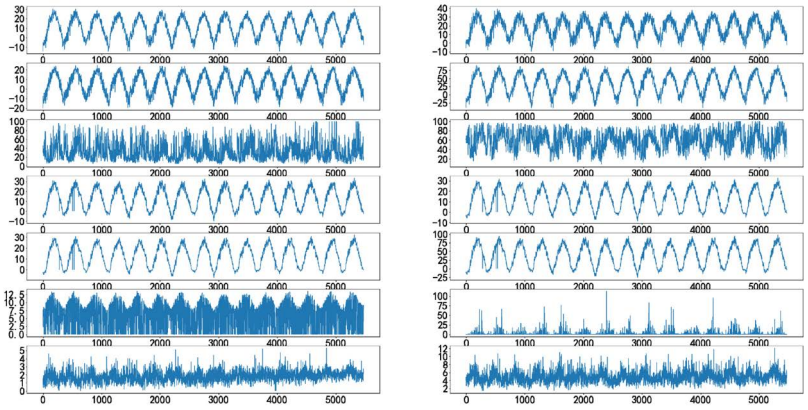

Figure 1. The periodicity of meteorological elements. From left to right, from top to bottom are: average temperature, maximum temperature, minimum temperature, the sum of temperature, minimum humidity, average humidity, 5 cm ground temperature, 10 cm ground temperature, 15 cm ground temperature, sum of ground temperature, sunshine hours, precipitation, average wind speed, maximum wind speed.

Table 1. Meteorological data quality.

Describe	Count	Mean	Std	Min	Max	Q1	Q2	Q3
Average Temperature	5478	11.05	10.45	−14.40	30.6	1.80	12.40	20.20
Maximum Temperature	5478	18.24	10.68	9.70	39.70	9.50	19.70	27.40
Minimum Temperature	5478	5.58	10.40	−19.80	24.60	−3.20	6.60	14.80
Precipitation	2039	3.86	8.70	0	111.90	0	0.40	3.80
5 cm Ground Temperature	5472	12.75	10.67	−10.10	33.20	2.10	13.80	22.30
10 cm Ground Temperature	5472	12.78	10.38	−8.70	32.40	2.20	13.90	22.10
15 cm Ground Temperature	5472	12.79	10.13	−7.60	32.10	2.30	13.90	22.00
Average Humidity	5478	58.21	18.78	8.00	99.00	44.00	59.00	73.00
Minimum Humidity	5478	31.92	19.41	3.00	99.00	16.25	27.00	44.00
Sunshine Hours	5478	5.94	3.95	0	13.40	2.00	7.00	9.00
Average Wind speed	5473	1.78	0.63	0.10	5.30	1.40	1.80	2.20
Maximum Wind speed	5478	4.81	1.33	1.40	12.10	3.90	4.70	5.60

Where count represents the number of data, and men, std, min, max, Q1, Q2, and Q3 represent the mean, variance, minimum, maximum, first, second, and third quartiles of different meteorological factors.

Selection of predictive variables. As show in Figure 1, it can be found that the average temperature, the maximum temperature, the minimum temperature, the sum of temperature, the 5 cm ground temperature, the 10 cm ground temperature, the 15 cm ground temperature, and the sum of ground temperature have obvious periodicity. At the same time, it can be seen that the upper edge of the sunshine hours is also periodic, while the periodicity of other meteorological elements is not obvious. Therefore, these nine meteorological elements (maximum temperature, minimum temperature, the sum of temperature, 5 cm ground temperature, 10 cm ground temperature, 15 cm ground temperature, the sum of ground temperature and sunshine hours) with observable periodicity are

selected as the input of multivariable LSTM network and binary classification task.

2.3. Multivariable LSTM Network

The neural network used to process serialized data is called recurrent neural network. When folded according to time, it can be regarded as a deep neural network with infinite layers [20]. In the traditional neural network model, the layers are fully connected, and the nodes of each layer are not connected. Theoretically, the function of each layer of the recurrent neural network is used to memorize the data, rather than hierarchical processing. There may be problems of gradient disappearance and gradient explosion. Long-term and short-term memory networks have designed new computing nodes based on maintaining the original recurrent neural network structure [21] [22] [23]. In the long-term and short-term memory network, storage units are used to replace conventional neurons. Each storage unit is composed of input gate, output gate, forget gate, and free state, as shown in **Figure 2** LSTM network.

2.4. Basic Binary Classifier

Ensemble learning refers to constructing multiple weak learners first, and then using a certain integration strategy to combine to obtain a "strong learner" with better performance indicators [24]. Logistic regression, Naive Bayes classification, Support vector machine, Random forest, Bagging classification, Decision tree classification, AdaBoost classification and Extra Trees classification are used as weak learners.

2.5. Experimental Process

The machine learning method makes the prediction results of the flowering prediction model dynamic. The proposed method is based on the combination of multivariate LSTM prediction and combined strategy binary classification prediction. In this way, it can solve the vacuum period caused by the early prediction of the regression prediction model. It can also solve the problem of decimal rounding in the prediction process. The multivariable LSTM prediction model and the combination strategy binary classification model mainly include three aspects, namely data processing, Multivariable LSTM and binary classification ensemble learning model and model evaluation, as show in **Figure 3**.

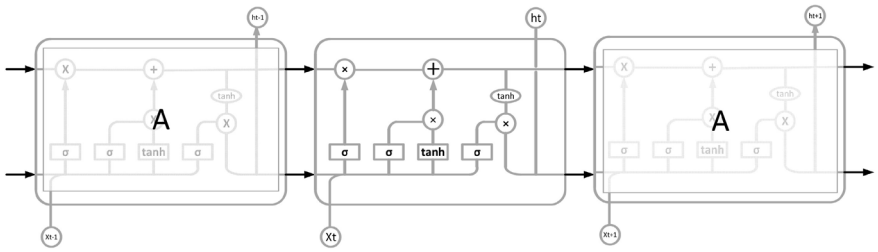

Figure 2. LSTM Network.

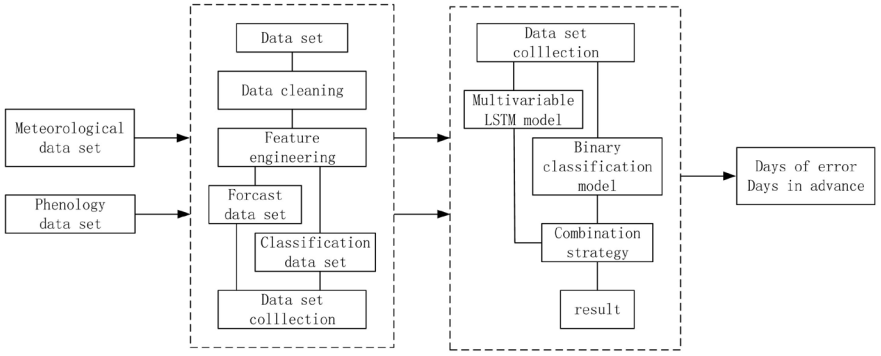

Figure 3. Multivariable LSTM and binary classification ensemble learning model Flow Chart.

The steps of flowering forecast are as follows:

Step 1: Data set partition

1) Forecast data set

For LSTM, the input data must be sequential data. The nine meteorological element data from January 1, 2005 to December 31, 2016 were used as the training set, and the nine meteorological element data from January 1, 2017 to December 31, 2019 were used as the test set.

2) Classification data set

For the classification model, input x (data of nine meteorological elements) and output y (1 or −1) are required. The data from March 25 to April 30 of each year from 2010 to 2019 is extracted, and the data is randomly divided into train set and test set according to a ratio of 7 - 3.

Step 2: Multi-LSTM network

1) Data Normalization

The LSTM network is particularly sensitive to the size of the input value, so the Min-Max normalization method is used to process the data. The common method of Data Normalization is Min-Max normalization. Through the linear transformation of the data, the result falls within the range of [0, 1]. This makes it easier and faster to transform dimensional data into pure values without dimensions to ensure comparability between data. The formula is:

$$x^* = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{3}$$

Among them, x is the observed value, $\min(x)$, $\max(x)$ corresponding to the x minimum and maximum values.

2) Window method

The window method is to use multiple recent time items to predict the next time item. Use the data in the first 90 days of t to predict the data in the last 7 days of t.

3) Model building

Define the model. It is to create a sequential model and add a configuration layer. Sequential model is a linear stacking of multiple network layers, that is

"one road goes to black". The layers used are LSTM layer, Repeat Vector layer, Dropout layer, Time Distributed layer and Dense layer. Among them, the activation function used for the LSTM layer is ReLU.

Compilation model. It is to select the parameters of the loss function and optimizer. The model is compiled with Adam as the optimizer and MSE as the loss function. The parameters of the model are shown in Table 2.

4) Model validation

Take the meteorological elements that can be observed from January 1, 2019 to April 17, 2019. After normalizing these data, a windowing process that predicts the next 3 days in 90 days is performed, and a total of 18 windows are generated by windowing. Among them, the first window is based on 90 days of data from January 1 to March 31 to predict the three days of data on April 1, April 2 and April 3. Extract the data of the next three days of the 18 windows predicted by the multivariate LSTM model, and divide them according to the first day of the future, the second day of the future, and the third day of the future into Dataset 1, Dataset 2, and Dataset 3. As shown in Table 3, it is the detailed information of the three data sets.

Step 3: Binary classification

1) Data standardization

Data standardization helps to remove the unit restrictions of the data and converts the data into pure values without dimensional constraints, ensuring the comparability between the data. The formula is:

$$x^* = \frac{x - \mu}{\sigma} \tag{4}$$

Among them, x is the observed value, μ is the overall mean, and σ is the overall standard deviation.

Table 2. LSTM model parameters.

Layer	Type	Output Shape	Params
lstm_1	LSTM	(None, 64)	18,944
repeat vector_1	Repeat Vector	(None, 7, 64)	0
dropout_1	Dropout	(None, 7, 64)	0
Lstm_2	LSTM	(None, 7, 32)	12,416
dropout_2	Dropout	(None, 7, 32)	0
time distribute_1	Time Distribute	(None, 3, 9)	297

Table 3. Information about the data set.

Dataset	Starting time	Termination time	Length
Dataset 1	2019-4-01	2019-4-18	18
Dataset 2	2019-4-02	2019-4-19	18
Dataset 3	2019-4-03	2019-4-20	18

2) Basic classifier

With Logistic regression, Naive Bayes classification, Support vector machine, Random forest classification, Bagging classification, Decision tree classification, AdaBoost classification and Extra Trees classification, eight classification learners are used as weak learning to complete the selection of the learner.

3) Combination strategy

The binary classifier combination strategy makes each classifier to solve the same original task, and combine the results of each model through a specific strategy to obtain a better global model. Using the arithmetic average combination strategy in the ensemble learning idea, when multiple classification learners judge all to 1, then judge to 1. When a classification learner judges that the result is not 1, the model judges that it is −1. The formula is:

$$\text{result} = \frac{1}{n}\sum_{i}^{n} pred_i = \begin{cases} 1 \\ -1 \end{cases} \tag{5}$$

Among them, $pred_i$ represents the predicted value of the i learner.

Step 4: Judging the initial flowering period

When the result value is 1 and it appears for the first time, the corresponding date is the initial flowering period.

Step 5: Model evaluation

Use the number of error days (actual value-predicted value) and the number of days in advance as evaluation indicators to complete the evaluation of model performance.

The input and output of the multivariable LSTM prediction model and binary classification model are shown in Figure 4. Both the input and output of the prediction model are meteorological element data. The input of the binary classification model is a meteorological element, and the output is the binary classification result of flowering. Finally, the classifier outputs whether it is the flowering period through the combination strategy.

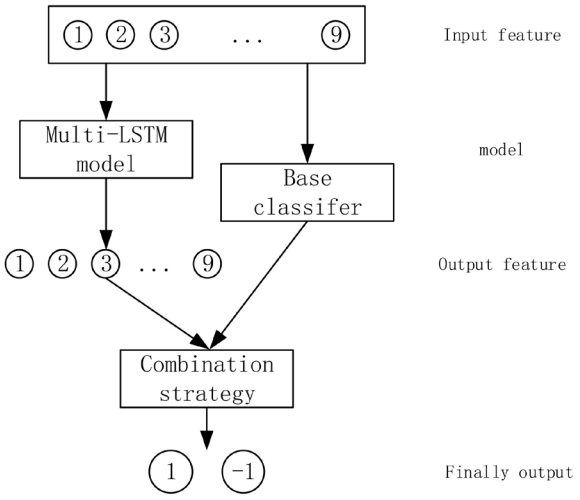

Figure 4. Model input and output.

3. Results

3.1. Multivariable LSTM Prediction Effect

Training the multivariable LSTM model and evaluating the model with MSE as the loss function. The loss of the model is shown in Figure 5. It can be seen that after 50 iterations, the loss of the model has stabilized and converged.

Figure 6 shows the RMSE value when the predicted and actual values of each window are normalized. It can be found that although the RMSE value is fluctuating, the overall value is less than 0.25. In order to see the prediction effect more intuitively, the prediction value of each window is separated to separate the data of the first day in the future, the second day in the future and the third day in the future. The predicted and actual values are evaluated by RMSE, and the results are shown in Figure 7.

3.2. Binary Classifier Selection

Use F1 score, accuracy and recall rate as evaluation indicators to complete the screening of weak learners. As shown in Table 4, the accuracy of different classification task learners is greater than 70% in both the training set and the test set. Among them, the accuracy of logistic regression, naive Bayes classification and support vector machine learner in the test set is better than that in the training set. In general, the performance in the training set is always better than the test

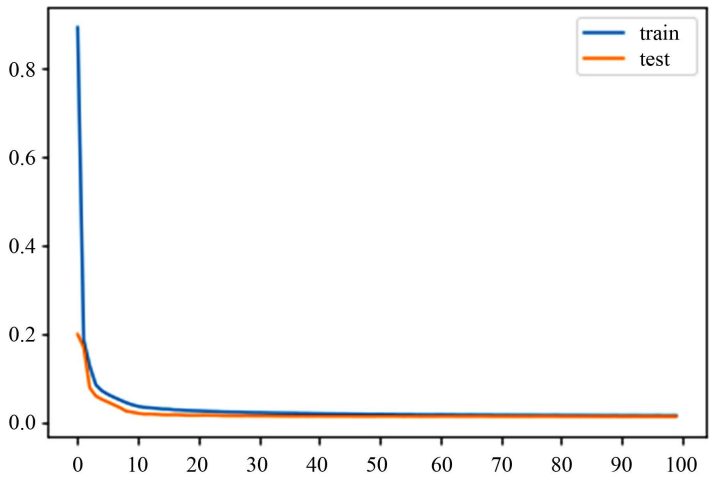

Figure 5. Model loss.

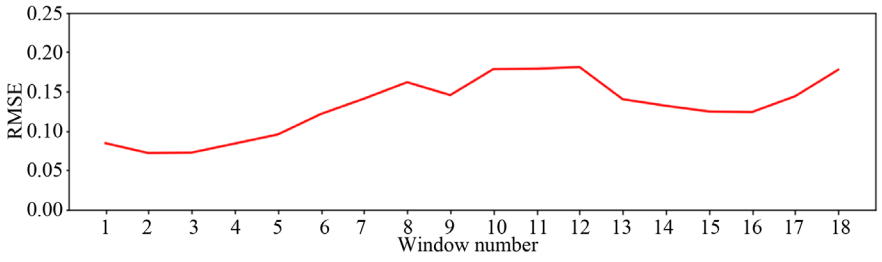

Figure 6. RMSE values of 18 windows.

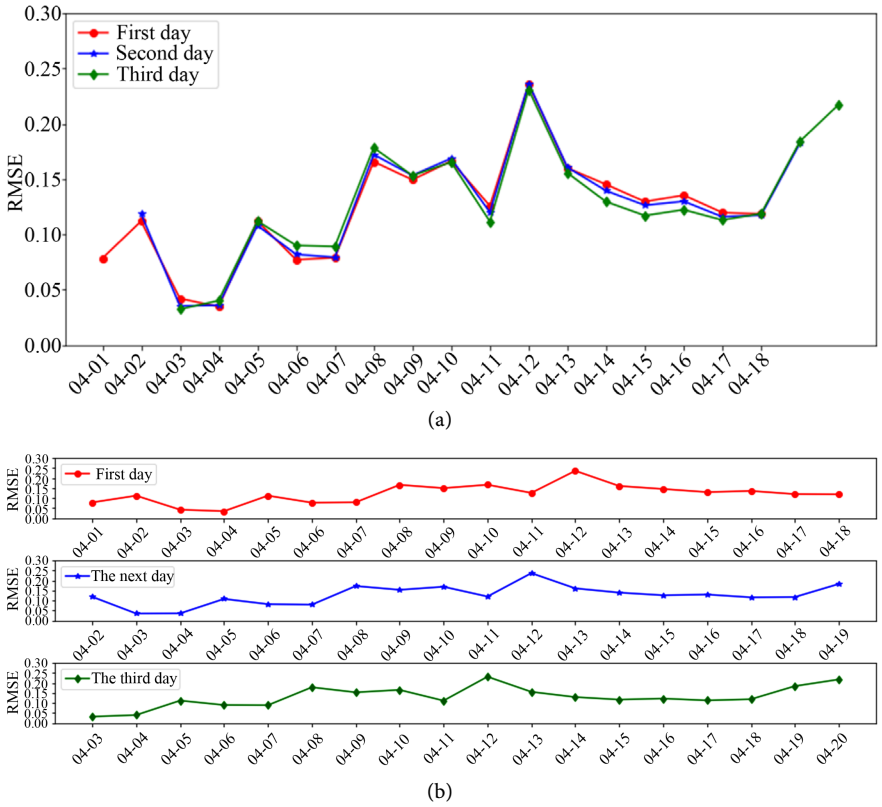

Figure 7. RMSE value for each day, (a) overall; (b) alone.

Table 4. Classifier performance.

Classifier	Accuracy score		F1 score	Recall score
	Train data	Test data		
Logistic Regression	0.78	0.79	0.79	0.78
Native Bayestion	0.73	0.77	0.78	0.80
Support Vector Machine (SVM)	0.79	0.81	0.80	0.75
Random Forest (RF)	0.98	0.81	0.80	0.76
Bagging	0.98	0.77	0.74	0.65
Decision Tree	1.00	0.76	0.72	0.62
Adaboost	0.96	0.80	0.79	0.76
Extra Trees	1.00	0.77	0.75	0.67

set. Therefore, when selecting a learner, first exclude the Logistic Regression, Native Bayestion, and support vector machine whose test set is more accurate than the training set. The accuracy of the remaining five classifiers in the training set is higher than 95%, and the accuracy in the test set is also higher than 75%. There are RF classifier and Adaboost classifier with F1 score greater than 0.75, and the corresponding Recall score is 0.76, which is the largest among the remaining five classifiers.

Draw ROC-AUC curve for the remaining 5 learners. The results are shown in Figure 8. The AUC values are ranked RF, AdaBoost, Bagging, Extra Trees and Decision Tree in descending order. Area Under the ROC Cure (AUC) is the area under the ROC curve and is used to measure the stability of the model. It is a curve with True Positive Rate (TPR) as the vertical axis and False Positive Rate (FPR) as the horizontal axis.

Combining Table 4 and Figure 8, select RF and Adaboost as weak learners to complete the binary-class ensemble learning combination strategy. As shown in Figure 9, the AUC value corresponding to the model using the integrated learning combination strategy is 0.82, which is greater than 0.81 and 0.80 of RF and AdaBoost.

3.3. Forecast Result

Dataset 1, Dataset 2, Dataset 3 are used to the ensemble learning model. Finding the date corresponding to the first occurrence of 1 is the initial flowering period. The results are shown in Table 5. Using the number of error days (actual value-predicted value) and the number of days ahead as evaluation indicators, the model is evaluated.

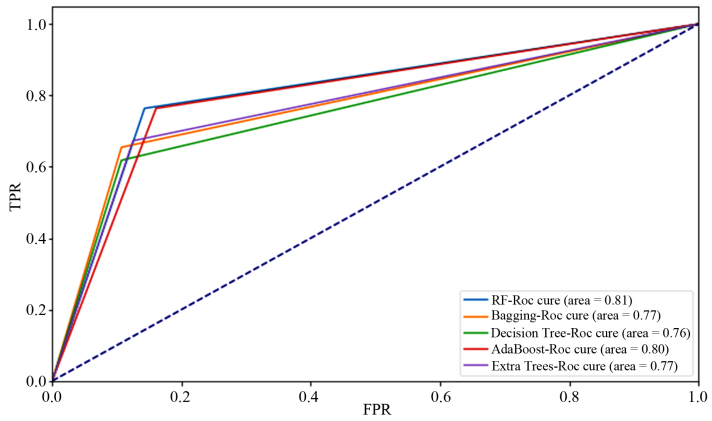

Figure 8. Classifier Roc-AUC curve.

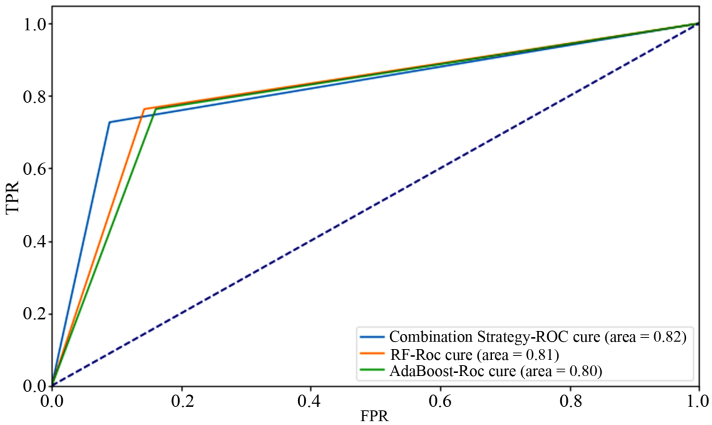

Figure 9. Combination strategy Roc-AUC curve.

Table 5. Model evaluation.

Dataset	predicted value	actual value	Days of error	Days in advance
Dataset 1	2019-4-07	2019-4-08	+1	1
Dataset 2	2019-4-07	2019-4-08	+1	2
Dataset 3	2019-4-08	2019-4-08	0	3

4. Discussion

Crop phenology is highly dictated by weather variables such as radiation, precipitation and temperature [25] [26]. Thus, the accuracy of predicting weather inputs is critical for crop simulation model [27]. Select meteorological factors with obvious periodicity helps to improve the accuracy of LSTM model prediction. Since LSTM adds a cell to judge whether the information is useful, an input gate, a forget gate, and an output gate are placed in a cell. Therefore, it is particularly suitable for the prediction of time series. Meanwhile, the advantage of the multivariable LSTM network is that it can complete the prediction of multiple meteorological factors simultaneously. In this study, the LSTM network simultaneously completes the prediction of maximum temperature, minimum temperature, average temperature, 5 cm ground temperature, 10 cm ground temperature, 15 cm ground temperature, sunshine hours, the sum of air temperature (SAT) and the sum of geothermal temperature (SGT).

For the binary-class classification task of ensemble learning. Firstly, it screens different classifiers to find classifiers that have no underfitting or overfitting. Underfitting is usually due to insufficient learning ability of the learner, and overfitting is usually due to too strong learning ability. Both will affect the generalization ability of the model. Secondly, complet the judgment of flowering period with the idea of combination strategy (1 = flowering, −1 = non-flowering). The advantages of multiple classifiers are combined to enhance the classification effect. In addition, the parameters of the selected classifier can be further optimized.

In this paper, a machine learning technique that combines time series prediction (special regression prediction) and classification prediction to complete flowering prediction is proposed. By analyzing the quality of the data and the periodicity of the data, seven feature variables with no missing values and obvious periodicity were extracted, and two features of SAT and SGT were added. Secondly, the weather data and phenology data are combined to divide the data into forecast data sets and classification data sets. Then the prediction results of the multivariable LSTM network are passed into the trained combined strategy binary classification learner to complete the prediction of flowering. Finally, the date corresponding to the first occurrence of the classification label 1 is the initial flowering period. The model solves the problem of decimal rounding in the regression prediction process, realizes the dynamic prediction of the flowering period, and the model error is within the range of one day.

In addition, in order to further improve the accuracy of the model, several problems need to be solved. First, by improving the LSTM network model, the RMSE value is further reduced. Secondly, further adjust the classification learner. Finally, increase the scope of the data. The LSTM network can predict time series prediction problems as special regression problems. Neural network can complete not only regression prediction but also classification prediction. Our future work will focus on using one network to complete flowering forecast.

5. Conclusion

This research combines neural networks with integrated learning, and proposes a method to dynamically predict whether the next three days will be flowering dates, effectively solving the problems of decimal rounding and long-term blank periods brought by regression prediction. This method utilizes the long-term storage characteristics of the LSTM network and the classification functions of Random Forest (RF) and Adaboost. The loss of the multivariable LSTM model is below 0.2, and the RMSE value is below 0.3. The AUC value of the combined classification model based on RF and AdaBoost is 0.82. In short, the error of the prediction model is 1 day.

Acknowledgements

This research was funded by the Chengdu Science and Technology Bureau Fund (2018-YF05-01217-SN). We would like to thank the Ji County Meteorological Bureau for the data provided.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ahas, R., Jaagus, J. and Aasa, A. (2000) The Phenological Calendar of Estonia and Its Correlation with Mean Air Temperature. *International Journal of Biometeorology*, **44**, 159-166. https://doi.org/10.1007/s004840000069

[2] Ge, Q.S., Wang, H.J., Rutishauser, T. and Dai, J.H. (2015) Phenological Response to Climate Change in China: A Meta-Analysis. *Global Change Biology*, **21**, 265-274. https://doi.org/10.1111/gcb.12648

[3] Menzel, A. (2003) Plant Phenological Anomalies in Germany and Their Relation to Air Temperature and NAO. *Climatic Change*, **57**, 243-263. https://doi.org/10.1023/A:1022880418362

[4] Liu, Y.J. and Dai, L. (2020) Modelling the Impacts of Climate Change and Crop Management Measures on Soybean Phenology in China. *Journal of Cleaner Production*, **262**, Article ID: 121271. https://doi.org/10.1016/j.jclepro.2020.121271

[5] Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.M., Guldberg, O.H. and Bairlein, F. (2002) Ecological Responses to Recent Climate Change. *Nature*, **416**, 389-395. https://doi.org/10.1038/416389a

[6] Rohde, A. and Bhalerao, R.P. (2007) Plant Dormancy in the Perennial Context. *Trends in Plant Science*, **12**, 217-223. https://doi.org/10.1016/j.tplants.2007.03.012

[7] Alcalá, A.R. and Barranco, D. (1992) Prediction of Flowering Time in Olive for the Cordoba Olive Collection. *American Society for Horticultural Science*, **27**, 1205-1207. https://doi.org/10.21273/HORTSCI.27.11.1205

[8] Behdani, M.A., Koocheki, A., Nassiri, M. and Rezvani, P. (2008) Models to Predict Flowering Time in the Main Saffron Production Regions of Khorasan Province. *Journal of Applied Sciences*, **8**, 907-909. https://doi.org/10.3923/jas.2008.907.909

[9] Yasuyuki, A.O.N.O. and Yukio, O.M.O.T.O. (1990) A Simplified Method for Estimation of Blooming Date for the Cherry by Means of DTS. *Journal of Agricultural Meteorology*, **46**, 147-151.

[10] Aono, Y. (1993) Climatological Studies on Blooming of Cherry Tree (*Prunus yedoensis*) by Means of DTS Method. B*ulletin of the University of Osaka Prefecture. Ser. B, Agriculture & Life Sciences*, **45**, 155-192.

[11] Sugiura, T. and Honjo, H. (1997) A Dynamic Model for Predicting E Flowering Date Developed Using an Endodormancy Break Model and A Flower Bud Development Model in Japanese Pear. *Journal of Agricultural Meteorology*, **52**, 897-900.

[12] Chun, J.A., Kang, K., Kim, D., Han, H.-H. and Son, I.-C. (2017) Prediction of Full Blooming Dates of Five Peach Cultivars (*Prunus persica*) Using Temperature-Based Models. *Scientia Horticulturae*, **220**, 250-258. https://doi.org/10.1016/j.scienta.2017.04.007

[13] Hur, J. and Ahn, J.B. (2017) Assessment and Prediction of the First-Flowering Dates for the Major Fruit Trees in Korea Using a Multi-RCM Ensemble. *International Journal of Climatology*, **37**, 1603-1618. https://doi.org/10.1002/joc.4800

[14] El Yaacoubi, A., Oukabli, A., Hafidi, M., Farrera, I., Ainane, T., Cherkaoui, S.I. and Legave, J.-M. (2019) Validated Model for Apple Flowering Prediction in the Mediterranean Area in Response to Temperature Variation. *Scientia Horticulturae*, **249**, 59-64. https://doi.org/10.1016/j.scienta.2019.01.036

[15] Chauhan, Y.S., Ryan, M., Chandra, S. and Sadras, V.O. (2019) Accounting for Soil Moisture Improves Prediction of Flowering Time in Chickpea and Wheat. *Scientific Reports*, **9,** Article No. 7510. https://doi.org/10.1038/s41598-019-43848-6

[16] Schneemilch, M., Michael, K. and Williams, C.R. (2012) Flowering Timing Prediction in Australian Native Understorey Species (*Acrotriche* R.Br Ericaceae) Using Meteorological Data. *International Journal of Biometeorology*, **56**, 95-105. https://doi.org/10.1007/s00484-010-0400-7

[17] Cenci, C.A. and Ceschia, M. (2000) Forecasting of the Flowering Time for Wild Species Observed at Guidonia, Central Italy. *International Journal of Biometeorology*, **44**, 88-96. https://doi.org/10.1007/s004840000065

[18] Park, I., Jones, A. and Mazer, S.J. (2019) Phenoforecaster: A Software Package for the Prediction of Flowering Phenology. *Applications in Plant Sciences*, **7**, 1230-1236. https://doi.org/10.1002/aps3.1230

[19] Elizondo, D.A., Mcclendon, R.W. and Hoogenboom, G. (1994) Neural Network Models for Predicting Flowering and Physiological Maturity of Soybean. *Transactions of the ASAE*, **37**, 981-988. https://doi.org/10.13031/2013.28168

[20] Hermans, M. and Schrauwen, B. (2013) Training and Analysis Deep Recurrent Neural Networks. *Advances in Neural Information Processing Systems*, 190-198.

[21] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[22] Gers, F.A., Schmidhuber, J. and Cummins, F. (2000) Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, **12**, 2451-2471. https://doi.org/10.1162/089976600300015015

[23] Gers, F.A., Schraudolph, N.N. and Schmidhuber, J. (2003) Learning Precise Timing with LSTM Recurrent Networks. Journal of Machine Learning Research, **3**, 115-143.

[24] Vega-Pons, S. and Ruiz-Shulcloper, J. (2011) A Survey of Clusering Ensemble Algorithms. *International Journal of Pattern Recognition and Artificial Intelligence*, **25**, 337-372. https://doi.org/10.1142/S0218001411008683

[25] Barnett, T.L. and Thompson, D.R. (1982) The Use of Large-Area Spectral Data in Wheat Yield Estimation. *Remote Sensing of Environment*, **12**, 509-518. https://doi.org/10.1016/0034-4257(82)90025-6

[26] Tollenaar, M., Fridgen, J., Tyagi, P., Stackhouse, P.W. and Kumudini, S. (2017) The Contribution of Solar Brightening to the US Maize Yield Trend. *Nature Climate Change*, **7**, 275-278. https://doi.org/10.1038/nclimate3234

[27] Togliatti, K., Archontoulis, S.V., Dietzel, R., Dietzel, R., Puntel, L. and Vanloocke, A. (2017) How Does Inclusion of Weather Forecasting Impact In-Season Crop Model Predictions? *Field Crops Research*, **214**, 261-272. https://doi.org/10.1016/j.fcr.2017.09.008