Integrating field images and microclimate data to realize multi-day ahead forecasting of maize crop coverage using CNN-LSTM

Xin Wang[1,2], Yu Yang[1], Xin Zhao[1], Min Huang[1], Qibing Zhu[1*]

(1. *Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University,*
Wuxi 214122, Jiangsu, China;

2. *School of Physics and Electronic Electrical Engineering, Huaiyin Normal University, Huai'an, 223300, Jiangsu, China*)

Abstract: Crop coverage (CC) is an important parameter to represent crop growth characteristics, and the ahead forecasting of CC is helpful to track crop growth trends and guide agricultural management decisions. In this study, a novel CNN-LSTM model that combined the advantages of convolutional neural network (CNN) in feature extraction and long short-term memory (LSTM) in time series processing was proposed for multi-day ahead forecasting of maize CC. Considering the influence of climate change on maize growth, five microclimatic factors were combined with historical maize CC estimated from field images as the input variables of the forecasting model. The field experimental data of four observation points for more than three years were used to evaluate the performance of CNN-LSTM at the forecasting horizon of three to seven days ahead and compared the forecasting results to CNN and LSTM. The results demonstrated that CNN-LSTM obtained the lowest RMSE and the highest R^2 at all forecasting horizons. Subsequently, the performance of CNN-LSTM under univariate (historical maize CC) and multivariate (historical maize CC+microclimatic factors) input was compared, and the results indicated that additional microclimatic factors were effective in improving the forecasting performance. Furthermore, the 3-day ahead forecasting results of CNN-LSTM in different growth stages of maize were also analyzed, and the results showed that the highest forecasting accuracy was obtained in the seven leaves stage. Therefore, CNN-LSTM can be considered a useful tool to forecast maize CC.
Keywords: maize crop coverage, multi-day ahead forecasting, CNN-LSTM, field images, microclimatic factors
DOI: 10.25165/j.ijabe.20231602.7020

1　Introduction

Precision agriculture (PA) is the new trend of agricultural development in the world and the rudiment of agricultural development in the future[1,2]. The key goal of PA is to improve crop yield and quality while optimizing inputs, improving efficiency, and reducing environmental pollution[3]. Real-time observation of the changes in soil nutrient factors, climatic conditions, and crop status during the crop growth stages is the prerequisite of PA[1]. Therefore, accurately tracking and even forecasting the changes in these variables is the key to realizing PA.

Crop coverage (CC) is usually defined as the vertical projection of green parts of the crop (including leaves, stems, and branches) to the ground surface, and expressed as a percentage of the reference area[4,5]. CC is not only an important trait to represent crop growth and development but also has shown sufficient correlations with other crop parameters that are difficult to measure, such as leaf area index (LAI)[6] and canopy light interception[7]. In different growth stages of crops, CC can directly or indirectly provide the basis for agricultural managers to carry out corresponding agricultural activities. In the seedling stages, CC can be used to judge whether the plants are too dense, which provides help for thinning and fixing seedlings. In the rapid growth stages (such as from the seven leaves to the jointing stage of maize), the stems and leaves of the crops grow rapidly, and the soil moisture content and nitrogen status of crops can be inverted according to CC, which provides useful information for precision irrigation and fertilization[8,9]. In the reproductive growth stages (such as from the tasselling to the maturity stage of maize), CC has a strong correlation with crop yield, which provides a basis for yield prediction[10]. Therefore, continuous CC monitoring and forecasting in the whole growth stages of crops can explain the causes of differences in crop growth and provide a decision-making basis for agricultural producers[11].

Digital photograph taken in the field is an effective approach for estimating CC because of the advantages of high spatial resolution, immediate utility, and low price[12]. Chianucci et al.[13] verified the feasibility of using digital photographs to estimate CC and applied it to crops, including aromatic plants. Coy et al.[14] introduced an unsupervised, threshold-based segmentation method to estimate CC from digital photographs, and achieved good results in the test of four different crops (maize, oat, flax, and rapeseed). In summary, the estimation of CC from digital photographs is mainly achieved by dividing the number of pixels of target crops by the total number of pixels in the entire digital photographs. However, all these studies mentioned above mainly focus on the real-time estimation of CC from digital photographs, while the ahead forecasting of CC has received little attention. Notably, the ahead forecasting of CC can make agricultural managers master the

growth trends of crops, which is of great significance for guiding agricultural management decisions and realizing PA.

The performance of the forecasting model can be improved by using the influencing factors related to the forecasted quantity as additional inputs[15]. For example, Matsumura et al.[16] used multiple linear regression and ANN to forecast the maize yield of Jilin Province, with climatic conditions and the historical yield as input of the forecasting models, and achieved good forecasting results. Ferreira and da Cunha[17] assessed the potential of five models to forecast daily ET_O up to seven days ahead using three input combinations, and the results showed that the best-performing models were obtained by using input containing additional data. Meanwhile, many studies have shown that climatic factors such as temperature and precipitation will have an impact on crop growth[18,19]. Therefore, the maize CC forecasting model established in this study fully considered the microclimate data affecting maize growth and adopted multivariate inputs.

Recently, deep learning algorithms have attracted much attention and have been applied in many fields, which are superior to traditional machine learning algorithms[17]. Long short-term memory (LSTM)[20] and Convolutional neural network (CNN)[21] are the most effective and widely used deep learning methods. For time series forecasting, LSTM can effectively capture series pattern information, while CNN can automatically extract more valuable features. However, standard LSTM is difficult to filter the noise in the input data, while standard CNN is difficult to solve the problem of long-term dependence[22]. Therefore, a time series forecasting model combining the advantages of CNN and LSTM could obtain better forecasting performance.

Considering the importance of maize CC forecasting, this study proposed a novel forecasting model to realize the multi-day ahead forecasting of maize CC by integrating field images and microclimate data. The main contributions of this study were summarized as follows: 1) the multivariate input was constructed by introducing additional microclimatic factors related to maize growth, which can provide more abundant information for the forecasting model; 2) a hybrid CNN-LSTM model that combined the advantages of CNN and LSTM was proposed to realize multi-day ahead forecasting of maize CC; 3) the field experimental data collected at four observation points for more than three years provided sufficient data for the evaluation of the forecasting model.

2 Study region and observation data

2.1 Study region

In this study, four observation points (A1, A2, A3, and A4) were set up in three regions, of which A1 was located in Zhengzhou, Henan Province, China (34.46°N, 113.40°E), A2 and A3 were located in Tai'an, Shandong Province, China (36.11°N, 117.08°E), and A4 was located in Gucheng, Hebei Province, China (39.13°N, 115.67°E). These three regions are typical maize-planting regions in China. During the experiment, wheat-maize intercropping technology was adopted, and the sowing time and cultivation mode of crops were consistent with those of the local agricultural practice. Maize was sown with a row spacing of about 90 cm and a plant spacing of about 19 cm. Maize growth stages are from June to early October, which takes about 100 d from sowing to harvesting.

2.2 Observation data

2.2.1 Collection of field images

At each observation point, a digital camera (E450 Olympus) installed on a bracket about 5 m above the ground was used to capture the field images. The fixed-focus method with a focal length

of 16 mm was used for shooting. The resolution of the captured images was 3648×2736 pixels and the corresponding actual area at the time of sowing was 30 m², and the area would decrease accordingly with the growth of maize plants.

Field images of maize from sowing to harvesting were collected at observation points A2, A3, and A4 from 2011 to 2013, respectively, while at A1 from 2010 to 2013. On each experimental day, the field images were taken every hour from 9:00 to 16:00, so eight field images could be captured every day. The field images of maize from sowing to harvesting at one observation point were taken as one image series. Therefore, thirteen maize field image series were collected in this study. Detailed information on each maize field image series is listed in Table 1.

Table 1   Corresponding observation point, year, date range, days, and maize variety of each maize field image series

Field image series	Observation point	Year	Date range	Days	Maize variety
1	A1	2010	Jun. 28-Oct. 1	96	Jundan no.20
2		2011	Jun. 4-Sept. 20	109	Jundan no.20
3		2012	Jun. 9-Sept. 30	114	Zhengdan no.958
4		2013	Jun. 6-Sept. 14	101	Zhengdan no.958
5	A2	2011	Jun. 16-Sept. 26	103	Nongda no.108
6		2012	Jun. 16-Sept. 26	103	Zhengdan no.958
7		2013	Jun. 19-Sept. 27	101	Zhengdan no.958
8	A3	2011	Jun. 28-Sept. 26	91	Nongda no.108
9		2012	Jun. 16-Sept. 26	103	Zhengdan no.958
10		2013	Jun. 19-Sept. 27	101	Zhengdan no.958
11	A4	2011	Jun. 27-Oct. 9	105	Jidan no.32
12		2012	Jun. 24-Oct. 8	107	Jidan no.32
13		2013	Jun. 26-Oct. 7	104	Zhengdan no.958

Note: A1, A2, A3, and A4 represent four observation points that were set up in three regions, A1 was located in Zhengzhou, Henan, China, A2 and A3 were located in Tai'an, Shandong, China, and A4 was located in Gucheng, Hebei, China.

2.2.2 Estimating maize CC from field images

The core of estimating maize CC from field images is to segment maize accurately. There are two most common approaches: threshold-based approaches and machine learning-based approaches[14]. Many researchers used digital images to estimate CC without considering the problems caused by natural light changes or light reflection on crop leaves[11].

Owing to the field images in this study being mainly captured in summer, the effects of natural light changes and crop leaves reflection must be considered. The maize images from the field were segmented using a method as in Ye et al.[23], who proposed a novel crop segmentation method by using a probabilistic superpixel Markov random field. This segmentation method can improve crop segmentation performance under strong illumination or shadow. After segmenting, the percentage of the number of pixels in the green part of maize to the total number of pixels is expressed as maize CC (0-100%). A typical maize field image under strong natural light is shown in Figure 1a along with the corresponding segmentation result shown in Figure 1b. By calculating the percentage of green pixels in Figure 1b, the maize CC is 59.68% could be obtained.

As previously mentioned, eight field images were taken at each observation point every day. The average value of maize CC estimated from eight field images was taken as daily maize CC on that day.

2.2.3 Collection of microclimate data

Each study region had an agricultural microclimate observation

a. Maize field image example was taken at A1   b. Segmentation result of image a

Figure 1   Example of maize field images taken at 10 o'clock on July 29, 2010, at observation point A1 and the corresponding segmentation result

system, which was used to collect microclimate data of maize from sowing to harvesting. Five microclimatic factors were obtained, including cumulative precipitation, maximum temperature, relative humidity, dew point temperature, and maximum air pressure (Table 2). The microclimate data, collected on an hourly timescale, were converted to a daily timescale. The daily cumulative precipitation was the sum of the hourly cumulative precipitation, the daily maximum temperature and air pressure were the maximum values of the hourly maximum temperature and air pressure, respectively, and the daily relative humidity and dew point temperature were the average values of the hourly relative humidity and dew point temperature, respectively.

Table 2   Details of daily microclimatic factors used in this study

Microclimatic factors	Value ranges
Cumulative precipitation/mm	0-115
Maximum temperature/°C	14.4-41.9
Relative humidity/%	18.08-90.83
Dew point temperature/°C	1.17-26.51
Maximum air pressure/hPa	982.2-1031.1

2.2.4   Data management

In this study, the daily maize CC and its corresponding microclimate data were taken as observation data, so 1338 observation data were collected at four observation points for more than 3 years. All these observation data were divided into two parts: a training set and a testing set. Particularly, the data from 2010 to 2012 were used for model training and parameter optimization, including 931 data, and the data from 2013 were used to test the forecasting performance of the models, including 407 data.

3   Methods

3.1   Construction of multivariate input matrix

The collected observation data is used to form a time series dataset X, $X=[x_1, x_2, x_3, \ldots, x_t, \ldots, x_D]^T$, where D is the total number of days of observation data and x_t is a row vector representing maize CC and microclimate data at time t, $x_t = [x_t^1, x_t^2, \ldots, x_t^M] \in \mathbb{R}^{1 \times M}$. Among them, x_t^1 is the maize CC, and x_t^2 to x_t^M are microclimate data, and M is the number of variables of observation data.

Notice that the time series dataset should be converted into an input matrix suitable for the forecasting model. Suppose that the "look back window", which is the number of days of previous historical data that are taken into consideration by a forecasting model to make forecasting[24], is set to L, and the forecasting horizon is set to d days ahead, which means that the maize CC on the next d-th day would be forecasted. As shown in Figure 2, the input matrices of the forecasting model at time t and $t+1$ are X_t and X_{t+1},

respectively, and the dimension of the input matrix is $L \times M$. The goal of this study was to train the forecasting model so that when the input matrix is X_t, the output is x_{t+d}^1. That is to say, the maize CC on the next d-th day is forecasted by using the maize CC and microclimate data of previous L days.
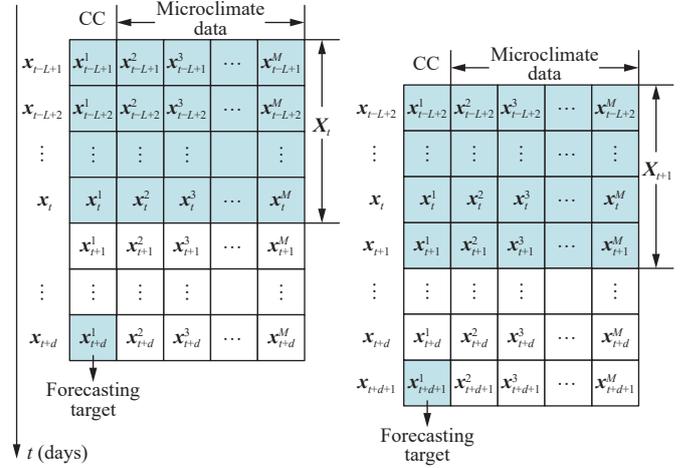

Figure 2   Input matrices of the forecasting model at time t and $t+1$

3.2   Proposed model

CNN[25] and LSTM[26] are the two main branches in the field of deep learning. CNN is good at extracting features from input data, while LSTM is effective for processing time series data. Inspired by this, a hybrid deep learning model (CNN-LSTM) is proposed to combine the advantages of CNN and LSTM.

3.2.1   CNN

The main difference between CNN and traditional neural networks is that CNN has the characteristics of local connection and weight sharing, so the training parameters of CNN are less than those of traditional neural networks[27]. The convolutional and pooling layers are the most important layer structures in CNN. The role of the convolutional layer is to automatically extract features from input data through convolutional kernels, and the role of the pooling layer is to reduce the size of feature maps. The convolutional kernel used in this study was one-dimensional, which was determined according to the dimension of the input data, and was called 1D CNN[15].

3.2.2   LSTM

LSTM is a variant of Recurrent Neural Network (RNN). It not only has the advantages of traditional RNN, but also solves the problem of gradient explosion or vanishing of traditional RNN in the process of back-propagation[28]. In particular, LSTM introduces a memory cell to store historical information, and combines three control gates (input, output, and forget gates) to read and write the memory cell. The decisions of the three control gates are all dependent on the previous output h_{t-1} and the current input x_t. The internal structure of LSTM unit is shown in Figure 3. More specifically, the equations of LSTM can be described as Equations (1)-(6):
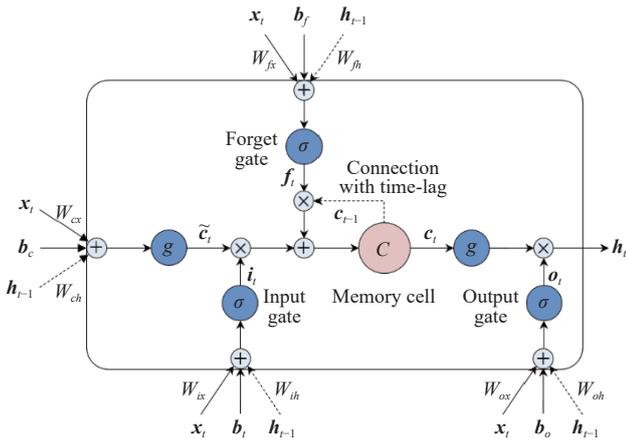
$$i_t = (W_{ix}x_t + W_{ih}h_{t1} + b_i) \tag{1}$$

$$f_t = (W_{fx}x_t + W_{fh}h_{t1} + b_f) \tag{2}$$

$$o_t = (W_{ox}x_t + W_{oh}h_{t1} + b_o) \tag{3}$$

$$\tilde{c}_t = g(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{5}$$

Note: LSTM: Long short-term memory.

Figure 3   Internal structure of LSTM unit

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot g(\boldsymbol{c}_t) \tag{6}$$

where, σ represents the sigmoid function, g represents the hyperbolic tangent function, $\boldsymbol{W}_{ix}$, $\boldsymbol{W}_{fx}$, $\boldsymbol{W}_{ox}$, and $\boldsymbol{W}_{cx}$ are input weight matrices, $\boldsymbol{W}_{ih}$, $\boldsymbol{W}_{fh}$, $\boldsymbol{W}_{oh}$, and $\boldsymbol{W}_{ch}$ are recurrent weight matrices, $\boldsymbol{b}_i$, $\boldsymbol{b}_f$, $\boldsymbol{b}_o$, and $\boldsymbol{b}_c$ are bias vectors, $\tilde{c}_t$ is the candidate state of the memory cell, and $\odot$ stands for the element-wise multiplication operation.

In order to handle time series data, L LSTM units are cascaded in turn to form an LSTM layer (Figure 4), in which each LSTM unit corresponds to a time slot. All LSTM units share weights and biases. Each row of the input matrix $\boldsymbol{X}_t$ is fed to the corresponding LSTM unit in the LSTM layer. The outputs of the j-th LSTM unit at time t, i.e., $\boldsymbol{h}_{t-L+j}$ and $\boldsymbol{c}_{t-L+j}$, are part of the inputs of the next LSTM unit. Therefore, the outputs of each LSTM unit are not only dependent on its inputs but also dependent on the inputs of the previous LSTM units.
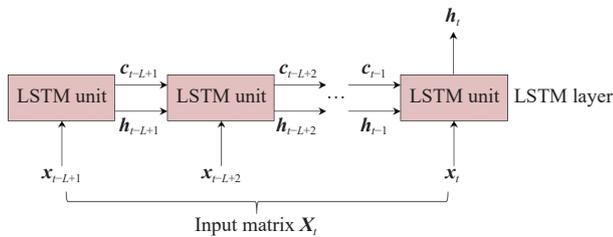

Figure 4   The unfolded LSTM layer structure

3.2.3   CNN-LSTM model

In order to combine the advantages of CNN and LSTM, a hybrid CNN-LSTM model was proposed. Firstly, the convolutional layer was used to automatically extract useful local features from fixed-length input data, and the max pooling layer was used to reduce the size of the extracted features. Then, the output was used as the input of the LSTM layer to learn the long short-term dependencies in time series data. Finally, the output of LSTM was used as the input of the full connection layer to obtain the final forecasting result. The number of convolution layers, LSTM layers, and full connection layers in the built network was obtained through many rigorous tests so that the proposed model can obtain the best forecasting performance. The structure of the hybrid CNN-LSTM framework is depicted in Figure 5, including one input layer, one 1D convolutional layer, one max pooling layer, one LSTM layer, and two fully connected layers.

The parameters of the convolutional layer were set as follows: the number of convolutional kernels equaled to 32, the size of
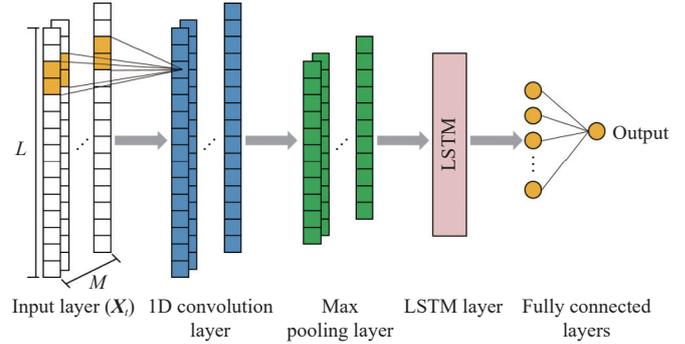
Note: L means "Look back window" which is the number of days of previous historical data that are taken into consideration by a forecasting model to make forecasting; M is the number of variables of observation data.

Figure 5   Network structure of the forecasting model based on CNN-LSTM

convolutional kernel equaled to 2, the activation function was set to Rectified Linear Unit (ReLU), and the padding was set to "same". In the max pooling layer, the following settings were applied: the size of pooling kernel equaled to 1, the stride equaled to 2, and the padding was set to "same". The number of neurons in the LSTM layer was set to 60. The number of neurons in the first fully connected layer and the second fully connected layer (output layer) was set to 10 and 1, respectively. In the training process of CNN-LSTM model, we employed mean absolute error as the loss function and used Adam[29] optimizer to update the network weights and biases by minimizing the loss function. The learning rate was set to 0.001, and the training epochs were set to 300. In addition, the look-back window was set to 5 d, and considering that sufficient time should be reserved for agricultural producers to make management decisions, the forecasting horizon was set to three to seven days ahead.

3.3   Design of experiments

In order to evaluate the forecasting performance of the CNN-LSTM model, we compared it with CNN and LSTM models (Figure 6). All three models were implemented on Python 3.5 by using Keras 2.2.2.

CNN model consisted of one input layer, one 1D convolutional layer, one max pooling layer, one flatten layer and two fully connected layers. The parameters of the convolutional layer, max pooling layer, and fully connected layers of the CNN model were the same as those of the CNN-LSTM model. The flatten layer had no parameter to set, and its role was to convert multi-dimensional features into one-dimensional data. The LSTM model included one input layer, one LSTM layer, and two fully connected layers. The parameters of LSTM layer and fully connected layers of LSTM model were the same as those of CNN-LSTM model. Furthermore, the loss function, optimizer, learning rate, training epochs, look-back window, and forecasting horizon of CNN and LSTM models were the same as those of CNN-LSTM model.

3.4   Data preprocessing

To eliminate the negative effects of variable differences on the performance of the forecasting model, all six variables (maize CC and five microclimatic factors) were normalized into [0, 1] according to Equation (7). It is worth noting that the minimum and maximum values of each variable were calculated only by the training data, not by the testing data.

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{7}$$

where, x and x_{norm} represent the original and normalized values,
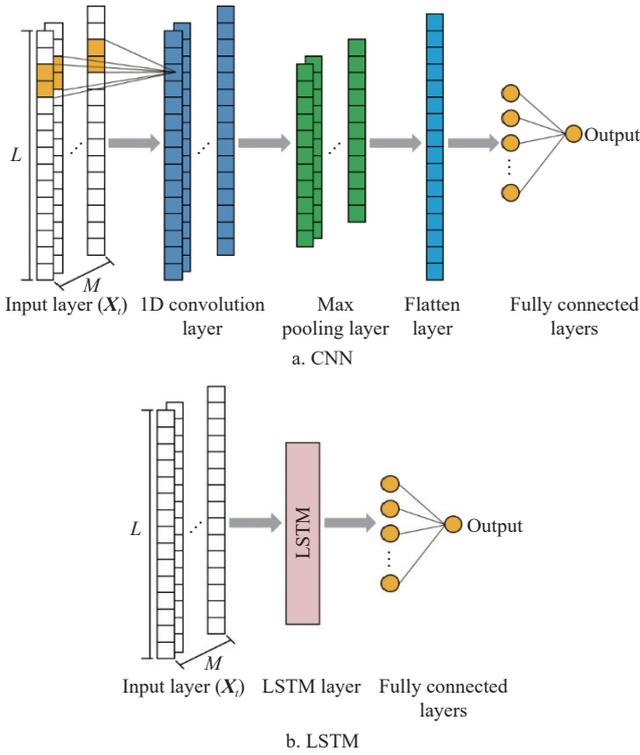
Figure 6   Network structures of the forecasting models based on CNN and LSTM

respectively; x_{min} and x_{max} represent minimum and maximum values, respectively.

3.5   Model evaluation

The performance of the forecasting model is evaluated by using Root Mean Square Error (RMSE) and coefficient of determination (R^2)[30,31]. Obviously, for a good forecasting model, RMSE should be close to 0 and R^2 should be close to 1. They are defined as Equations (8) and (9).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{8}$$

$$R^2 = \frac{\sum_{i=1}^{N}(y_i - \bar{y})^2 - \sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{9}$$

where, y_i is the observed maize CC, $\hat{y}_i$ is the forecasted maize CC, $\bar{y}$ is the mean of the observed maize CC, and N is the number of samples in the testing set.

4   Results and discussion

4.1   Analysis of the changing trend of maize CC

In this study, 13 field image series of maize from sowing to harvesting were collected. After estimating maize CC from field images, the change curves of maize CC can be obtained. Taking 2013 as an example, the change curves of maize CC at four observation points are shown in Figure 7. Overall, maize CC shows a slow rise at first, then a rapid rise, then relatively stable, and finally a slow decline in the trend. However, the change trends of maize CC at different observation points are also different. For example, the data fluctuation at observation point A1 is the most obvious, followed by A4, while A2 and A3 are relatively stable. The difference in change curves of maize CC at four observation

points is mainly related to climatic conditions. Therefore, in the ahead forecasting of maize CC, the traditional strategy of using only historical maize CC is difficult to obtain accurate results, while the strategy of combining additional microclimate factors can solve this problem.
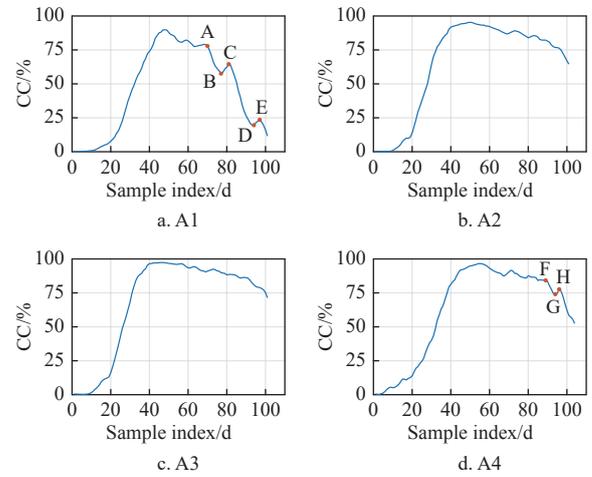

Figure 7   Change curves of maize CC at four observation points A1, A2, A3, and A4 in 2013

According to the microclimate data of maize during the total growth stage at four observation points in 2013, it was observed that the average maximum temperature and cumulative precipitation at A1 were 32.7°C and 125.8 mm, respectively. The temperature was high and the precipitation was low, so drought often occurred, causing maize leaves to turn yellow and wilt, and after rain, maize leaves returned to green and unfold, which led to the fluctuation of the changing curve of maize CC. A2 and A3 shared an agricultural microclimate observation system, the average maximum temperature and cumulative precipitation at A1 were 30.8°C and 339.1 mm, respectively. The temperature was suitable, and the precipitation was moderate and evenly distributed. Few maize leaves showed drought and wilt, so the change curve of maize CC was relatively stable. The average maximum temperature and cumulative precipitation at A4 were 28.8°C and 373.3 mm, respectively. The temperature was low, and the precipitation was abundant but unevenly distributed, which would also lead to slight fluctuations in maize CC.

Due to the accumulation or lag effect of microclimatic factors, there are some obvious data fluctuation points (A, B, C, D, E, F, G) in the change curves of maize CC (Figure 7). The average maximum temperature and cumulative precipitation between data fluctuation points are listed in Table 3. It can be seen that in the range of A-B, C-D, and F-G, there was basically no precipitation and the temperature was relatively high, which made maize leaves wilt and led to a decrease in maize CC. In the range of B-C, D-E, and G-H, there was sufficient precipitation and the temperature was suitable, which made the wilted maize leaves unfold again, and led to the increase of maize CC.

4.2   Comparison of forecasting performance of different models

The performance of three models at the forecasting horizon of three-seven days ahead is listed in Table 4 and Figure 8. The RMSE is ranked in the descending order of CNN, LSTM, and CNN-LSTM, which is completely consistent with the ascending order of the R^2. In addition, the performance of all models suffers from a loss with the increase of the forecasting horizons. However, even in the 7 d ahead forecasting, CNN-LSTM still obtains acceptable forecasting accuracy with the RMSE of 8.260% and the R^2 of 0.922.

Table 3   Average maximum temperature and cumulative precipitation between different data fluctuation points

Datapoints range	MaximumTemperature/°C	CumulativePrecipitation/mm
A-B	35.7	0.2
B-C	31.7	5.5
C-D	31.3	0.1
D-E	25.5	3.8
F-G	23.8	0.0
G-H	22.9	7.6

Furthermore, the advantages of CNN-LSTM are more obvious with the increase in the number of ahead forecasting days. For example, compared with the second-ranked LSTM, the RMSE of CNN-LSTM decreases by 0.48% and the R^2 increases by 0.10% in the 3-day ahead forecasting, while the RMSE decreases by 9.53% and the R^2 increases by 1.99% in the 7 d ahead forecasting. This indicates

that CNN-LSTM has better performance for long-term forecasting. In practical application, the forecasting accuracy for more distant days ahead is what we are most concerned about, because the more distant days ahead, the earlier the decision-making that agricultural producers can make. Therefore, CNN-LSTM is considered to be the optimal model for multi-day ahead forecasting of maize CC.

Table 4   Comparison of multi-day ahead forecasting performance based on RMSE and R^2 of different models

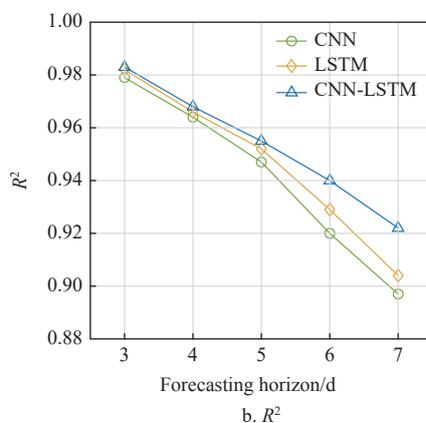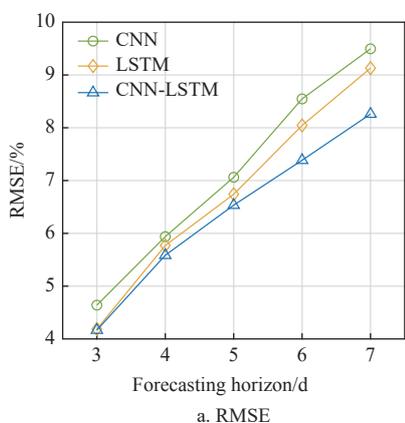| Horizon/day	RMSE/%			R^2		
	CNN	LSTM	CNN-LSTM	CNN	LSTM	CNN-LSTM
3	4.640	4.187	4.167	0.979	0.982	0.983
4	5.938	5.770	5.585	0.964	0.966	0.968
5	7.062	6.741	6.534	0.947	0.952	0.955
6	8.547	8.044	7.385	0.920	0.929	0.940
7	9.496	9.130	8.260	0.897	0.904	0.922

Figure 8   Comparison of RMSE and R^2 obtained by different models at the forecasting horizon of three to seven days ahead

In order to further illustrate the difference in the forecasting performance of the three models, the comparison of forecasting results of the three models at the forecasting horizon of three, five, and seven days ahead is given in Figure 9. In general, all three models can track the change of maize CC at three forecasting horizons, but their forecasting accuracy is slightly different. In terms of different forecasting horizons, the forecasting accuracy of all three models decreases with the increase of the forecasting horizons. In terms of different observation points, the forecasting accuracy of all three models at A2 and A3 is higher than that at A1 and A4 because the change of maize CC at A2 and A3 is relatively stable and that at A1 and A4 fluctuates obviously. However, the forecasted maize CC values using CNN-LSTM are closer to the observed maize CC values than using CNN and LSTM at the same forecasting horizon and observation point. This indicates that the combination of CNN and LSTM achieves additional performance gains over the use individually of CNN and LSTM.

4.3   Comparison of the forecasting performance of CNN-LSTM under different input types

To better illustrate the contribution of the introduced microclimate data to the forecasting performance of CNN-LSTM, the RMSE, and R^2 of different forecasting horizons obtained by CNN-LSTM under univariate and multivariate inputs are listed in Table 5. Among them, univariate refers to only one variable, maize CC, which is contained as the input of CNN-LSTM, while multivariate refers to five microclimatic variables that are also contained besides maize CC. Meanwhile, CNN-LSTM used for two different input types has the same network structure and parameters.

Compared with univariate input, the mean RMSE of multivariate input decreases by 3.36% and the mean R2 increases by 0.34%, and
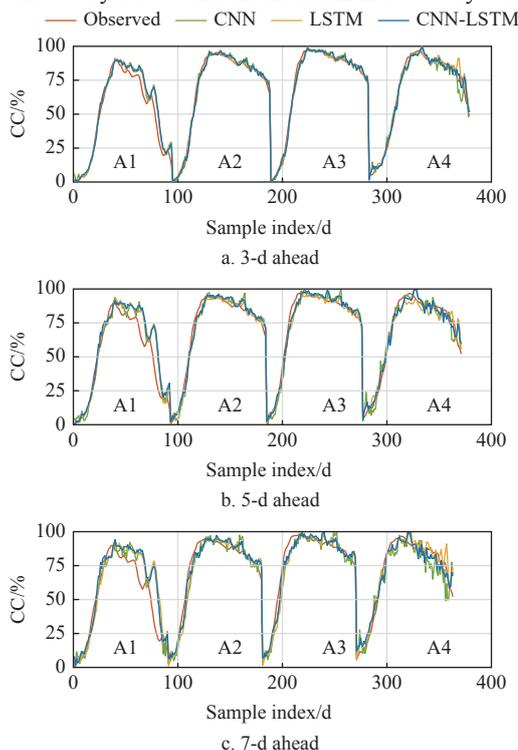

Figure 9   Comparison of forecasting results of different models at the horizon of 3-d, 5-d, and 7-d ahead

Table 5　RMSE and R² of different forecasting horizons obtained by CNN-LSTM under different input types

Input type	Metrics	Horizon/day				
		3	4	5	6	7
Univariate	RMSE/%	4.359	5.788	6.717	7.641	8.499
	R^2	0.981	0.966	0.952	0.936	0.917
Multivariate	RMSE/%	4.167	5.585	6.534	7.385	8.260
	R^2	0.983	0.968	0.955	0.940	0.922

the forecasting accuracy of multivariate input is better than that of univariate input at all forecasting horizons. In addition, existing studies have shown that the stability of the forecasting model will be better if the influencing factors are related to the forecasted quantity as an additional input of the forecasting model[32]. Therefore, from the perspective of forecasting accuracy and model stability, CNN-LSTM under multivariate input is considered the optimal input type.

4.4　Analysis of forecasting results of CNN-LSTM in different

growth stages of maize

From sowing to harvesting, maize includes the following growth stages: emergence, three leaves, seven leaves, jointing, tasselling, flowering, spinning, and maturity stage[33]. Taking observation point A3 in 2013 as an example, the dates of maize entering different growth stages and the corresponding days after sowing obtained by the field observation of the observers are listed in Table 6. In different growth stages, the change trends of maize CC are quite different, so the corresponding forecasting accuracy is also various. According to the characteristics of maize in different growth stages, it can be divided into four periods: emergence-three leaves stage, seven leaves the stage, jointing-flowering stage, and spinning-maturity stage. The 3-day ahead forecasting results of CNN-LSTM in different growth stages are analyzed.

As shown in Figure 10, the forecasted values are basically consistent with the observed values during the total growth stage of maize. From the emergence to the flowering stage, maize CC

Table 6　Dates of maize entering different growth stages and the corresponding days after sowing at observation point A3 in 2013

Growth stages	Sowing	Emergence	Three leaves	Seven leaves	Jointing	Tasselling	Flowering	Spinning	Maturity
Entering dates	Jun. 19	Jun. 25	Jun. 27	Jul. 6	Jul. 19	Aug. 7	Aug. 8	Aug. 9	Sept. 27
Days after sowing	0	6	8	17	30	49	50	51	100

increases continuously and tends to be stable in the later periods, while from the spinning to the maturity stage, maize CC begins to decrease, accompanies by fluctuations. As shown in Figure 10a, the forecasting errors in several days are large, which may be due to the maize CC being relatively small and easy to be disturbed by weeding and other agricultural activities. As shown in Figures 10b, maize CC increases approximately linearly, and CNN-LSTM can well capture this linear change relationship, so the highest forecasting accuracy is obtained, and the fitted line almost coincides with the 1:1 line. As shown in Figure 10c, the early part of this period is the jointing stage. Maize grows most vigorously, and the

ability of CNN-LSTM to capture rapid changes in maize CC is insufficient, which leads to low forecasting accuracy. However, in the latter part of this period, maize CC tends to be stable, and the coincidence degree between forecasted values and observed values is high. As shown in Figure 10d, the maize CC curve has fluctuations due to climatic conditions, but CNN-LSTM can still accurately track and forecast maize CC. This is because we fully considered the microclimatic factors affecting maize CC when establishing the forecasting model.

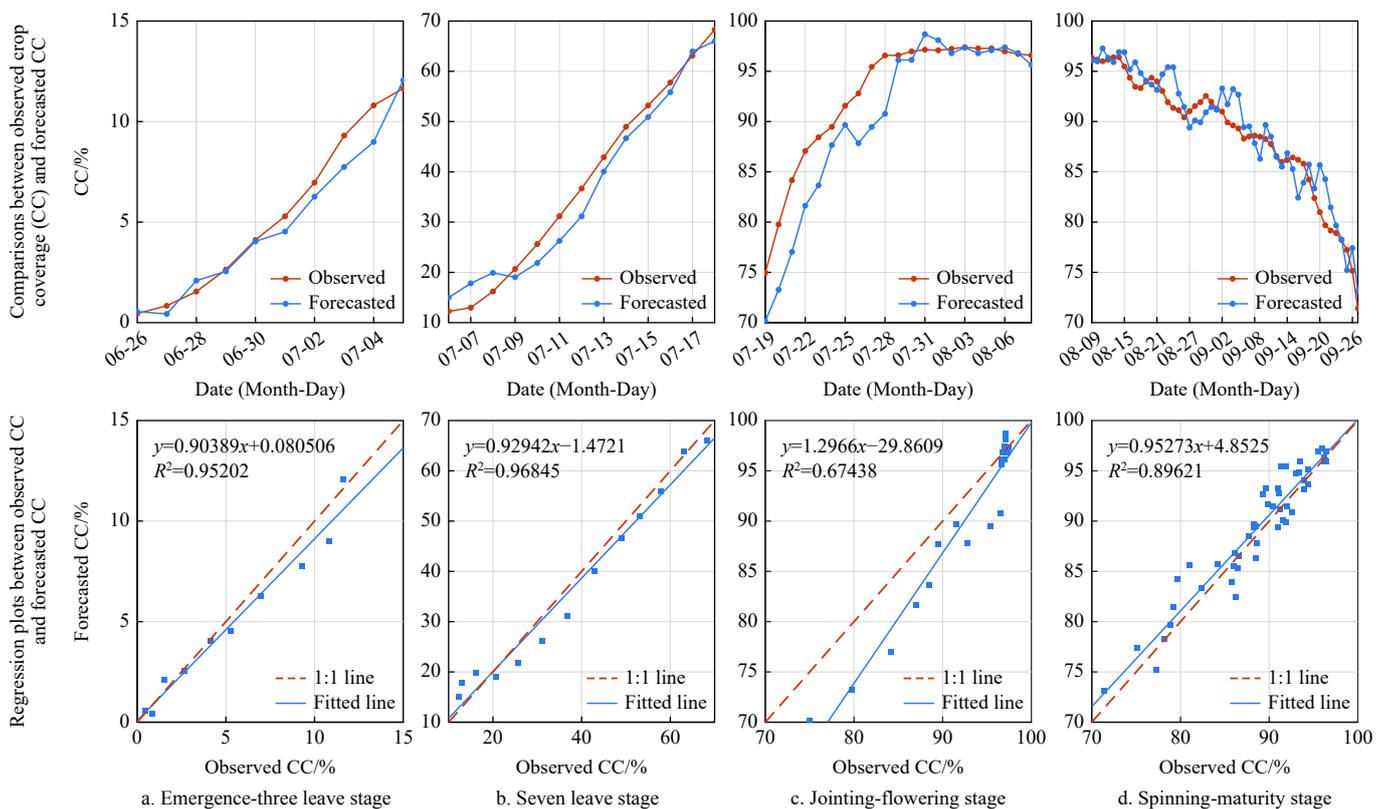Comparisons between observed crop coverage (CC) and forecasted CC

Figure 10　Comparisons and regression plots between observed CC and forecasted CC, respectively

5   Conclusions

In this study, a novel forecasting model CNN-LSTM was proposed to forecast maize CC, which combined the advantages of CNN in feature extraction and LSTM in time series processing. The forecasting performance of CNN-LSTM was compared with that of CNN and LSTM at the horizon of three to seven days ahead, the results showed that CNN-LSTM had the lowest RMSE and the highest R^2, followed by LSTM, and the worst was CNN. Furthermore, the forecasting performance of CNN-LSTM under univariate and multivariate inputs was compared, and the results showed that multivariate input performed better than univariate input, which indicated that the additional microclimatic variables had a positive effect on the improvement of the forecasting performance. Finally, taking observation point A3 in 2013 as an example, the 3 d ahead forecasting results of CNN-LSTM in different growth stages of maize were analyzed, and the results indicated that the forecasting accuracy of CNN-LSTM in different growth stages of maize was different, and the highest R^2 was obtained in the seven leaves stage. Overall, this study showed the potential of using CNN-LSTM to forecast maize CC by integrating field images and microclimate data.

Acknowledgements

[References]

[1]   Maes W H, Steppe K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. Trends in Plant Science, 2019; 24(2): 152–164.

[2]   Sui R X, Thomasson J A, Ge Y F. Development of sensor systems for precision agriculture in cotton. Int J Agric & Biol Eng, 2012; 5(4): 1–14.

[3]   Chlingaryan A, Sukkarieh S, Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and electronics in agriculture, 2018; 151: 61–69.

[4]   Purevdorj T S, Tateishi R, Ishiyama T, Honda Y. Relationships between percent vegetation cover and vegetation indices. International Journal of Remote Sensing, 1998; 19(18): 3519–3535.

[5]   Guevara-Escobar A, Tellez J, Gonzalez-Sosa E. Use of digital photography for analysis of canopy closure. Agroforestry Systems, 2005; 65(3): 175–185.

[6]   Tao Z Q, Bagum S A, Ma W, Zhou B Y, Fu J D, Cui R X, et al. Establishment of the crop growth and nitrogen nutrition state model using spectral parameters canopy cover. Spectroscopy and Spectral Analysis, 2016; 36(1): 231–236. (in Chinese)

[7]   Purcell L C. Soybean canopy coverage and light interception measurements using digital imagery. Crop Science, 2000; 40(3): 834–837.

[8]   Cheng M H, Jiao X Y, Liu Y D, Shao M C, Yu X, Bai Y, et al. Estimation of soil moisture content under high maize canopy coverage from UAV multimodal data and machine learning. Agricultural Water Management, 2022; 264: 107530.

[9]   Lee K J, Lee B W. Estimation of rice growth and nitrogen nutrition status using color digital camera image analysis. European Journal of Agronomy, 2013; 48: 57–65.

[10]   García-Martínez H, Flores-Magdaleno H, Ascencio-Hernández R, Khalil-Gardezi A, Tijerina-Chávez L, Mancilla-Villa O R, et al. Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and RGB images acquired with unmanned aerial vehicless. Agriculture, 2020; 10(7): 277.

[11]   De la Casa A, Ovando G, Bressanini L, Martínez J, Díaz G, Miranda C. Soybean crop coverage estimation from NDVI images with different spatial resolution to evaluate yield variability in a plot. ISPRS Journal of Photogrammetry and Remote Sensing, 2018; 146: 531–547.

[12]   Liu Y K, Mu X H, Wang H X, Yan G J. A novel method for extracting green fractional vegetation cover from digital images. Journal of Vegetation Science, 2012; 23(3): 406–418.

[13]   Chianucci F, Lucibelli A, Dell'Abate M T. Estimation of ground canopy cover in agricultural crops using downward-looking photography. Biosystems Engineering, 2018; 169: 209–216.

[14]   Coy A, Rankine D, Taylor M, Nielsen D C, Cohen J. Increasing the accuracy and automation of fractional vegetation cover estimation from digital photographs. Remote Sensing, 2016; 8(7): 474.

[15]   Li T Y, Hua M, Wu X. A hybrid CNN-LSTM model for forecasting particulate matter (PM2.5). IEEE Access, 2020; 8: 26933–26940.

[16]   Matsumura K, Gaitan C F, Sugimoto K, Cannon A J, Hsieh W W. Maize yield forecasting by linear regression and artificial neural networks in Jilin, China. The Journal of Agricultural Science, 2015; 153(3): 399–410.

[17]   Ferreira L B, da Cunha F F. Multi-step ahead forecasting of daily reference evapotranspiration using deep learning. Computers and Electronics in Agriculture, 2020; 178: 105728.

[18]   Butler E E, Huybers P. Adaptation of US maize to temperature variations. Nature Climate Change, 2013; 3(1): 68–72.

[19]   Mistry P, Bora G. Development of yield forecast model using multiple regression analysis and impact of climatic parameters on spring wheat. Int J Agric & Biol Eng, 2019; 12(4): 110–115.

[20]   Long Z H, Qin J M, Zhang T Y, Xu W. Prediction of continuous time series leaf area index based on long short-term memory network: A case study of winter wheat. Spectroscopy and Spectral Analysis, 2020; 40(3): 898–904. (in Chinese)

[21]   Sayeed A, Choi Y, Eslami E, Lops Y, Roy A, Jung J. Using a deep convolutional neural network to predict 2017 ozone concentrations, 24 hours in advance. Neural Networks, 2020; 121: 396–408.

[22]   Livieris I E, Pintelas E, Pintelas P. A CNN–LSTM model for gold price time-series forecasting. Neural Computing and Applications, 2020; 32(23): 17351–17360.

[23]   Ye M N, Cao Z G, Yu Z H, Bai X D. Crop feature extraction from images with probabilistic superpixel Markov random field. Computers and Electronics in Agriculture, 2015; 114: 247–260.

[24]   Wang Q L, Guo Y F, Yu L X, Li P. Earthquake prediction based on spatio-temporal data mining: an LSTM network approach. IEEE Transactions on Emerging Topics in Computing, 2017; 8(1): 148–158.

[25]   LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998; 86(11): 2278–2324.

[26]   Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997; 9(8): 1735–1780.

[27]   Zhang X L, Lin T, Xu J F, Luo X, Ying Y B. DeepSpectra: An end-to-end deep learning approach for quantitative spectral analysis. Analytica Chimica Acta, 2019; 1058: 48–57.

[28]   Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 1998; 6(2): 107–116.

[29]   Kingma D P, Ba J L. Adam: A method for stochastic optimization. In: Proceedings of the International Conference on Learning Representations (ICLR), 2015; In press. arXiv: 1412.6980.

[30]   Zhang J F, Zhu Y, Zhang X P, Ye M, Yang J Z. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. Journal of Hydrology, 2018; 561: 918–929.

[31]   Jung D H, Kim H S, Jhin C, Kim H J, Park S H. Time-serial analysis of deep neural network models for prediction of climatic conditions inside a greenhouse. Computers and Electronics in Agriculture, 2020; 173: 105402.

[32]   Yan R, Liao J, Yang J Q, Sun W, Nong M Y, Li F P. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Systems with Applications, 2021; 169: 114513.

[33]   Yue Y, Li J H, Fan L F, Zhang L L, Zhao P F, Zhou Q, et al. Prediction of maize growth stages based on deep learning. Computers and Electronics in Agriculture, 2020; 172: 105351.