Hindawi
Advances in Meteorology
Volume 2020, Article ID 5124274, 13 pages
https://doi.org/10.1155/2020/5124274

Research Article

A Convection Nowcasting Method Based on Machine Learning

Aifang Su,[1,2] **Han Li**,[1,2] **Liman Cui**,[1,2] **and Yungang Chen**[3]

[1]*Key Laboratory of Agrometeorological Safeguard Application Technique, CMA, Zhengzhou 450003, China*
[2]*Henan Provincial Meteorological Observatory, Zhengzhou 450003, China*
[3]*Beijing Presky Technology Co., Ltd., Beijing 100195, China*

Correspondence should be addressed to Aifang Su; suaifang68@126.com

Received 24 June 2019; Revised 6 November 2019; Accepted 20 December 2019; Published 27 January 2020

Academic Editor: Tomeu Rigo

In this study, a convection nowcasting method based on machine learning was proposed. First, the historical data were back-calculated using the pyramid optical flow method. Next, the generated optical flow field information of each pixel and the Red-Green-Blue (RGB) image information were input into the Convolutional Long Short-Term Memory (ConvLSTM) algorithm for training purposes. During the extrapolation process, dynamic characteristics such as the rotation, convergence, and divergence in the optical flow field were also used as predictors to form an optimal nowcasting model. The test analysis demonstrated that the algorithm combined the image feature extraction ability of the convolutional neural network (CNN) and the sequential learning ability of the long short-term memory network (LSTM) model to establish an end-to-end deep learning network, which could deeply extract high-order features of radar echoes such as structural texture, spatial correlation, and temporal evolution compared with the traditional algorithm. Based on learning through the above features, this algorithm can forecast the generation and dissipation trends of convective cells to some extent. The addition of the optical flow information can more accurately simulate nonlinear trends such as the rotation, or merging, or separation of radar echoes. The trajectories of radar echoes obtained through nowcasting are closer to their actual movements, which prolongs the valid forecasting period and improves forecast accuracy.

1. Introduction

Convection nowcasting refers to the short-term forecasting of the convective weather system and the catastrophic convective weather they may produce, up to 0–6 hours beyond the current observation time. At present, convection nowcasting based on radar data mainly involves thunderstorm identification, tracking, and automated extrapolation techniques [1], e.g., the centroid tracking method [2–4], the tracking radar echo by correlation (tREC) method [5–7], and the optical flow method [8–13].

The centroid tracking method can only be used for convective precipitation systems, while the TREC method can also track the layered cloud precipitation system. When faced with a fast-changing strong convective precipitation system, the TREC failure rate increased significantly [13]. Optical flow is a dense field of displacement vectors which defines the transition of each pixel in a region. It is computed using the brightness constraint, which assumes brightness constancy of corresponding pixels in consecutive frames. Thus, optical flow is commonly used as a feature in motion-based segmentation and tracking applications [10]. These advantages can solve the shortcomings of both the centroid tracking and TREC methods to some extent. However, the success of optical flow based methods is limited because it does not consider the physical meaning of radar echo development. As such, it is challenging to predict short-term local convection with rapid generation or extinction [14, 15].

The traditional optical flow method is only applicable to small movements of image features, although the pyramid delaminating technique can improve the calculation accuracy and convergence speed of this technique [16–18]. According to [19], even though the pyramid optical flow method has advantages in forecasting the convective radar echoes, it can only track echo characteristics that already exist in an image and cannot predict either the generation or dissipation of echoes. In addition, since atmospheric motion often exhibits highly nonlinear and random disturbance

behavior, the optical flow method is insufficient over a short valid forecasting period (0–6 h).

Traditional image recognition processing requires manually setting specific features such as the shape, length-width ratio, and area of a connected region in order to extract the desired information; occasionally all of the pixels in an image are used as the basic information for classifier training and classification. The former cannot guarantee the successful extraction of relevant or significant features, while the latter tends to introduce a great deal of redundant information.

Inspired by the cognitive mechanism of biological natural vision, Hubel and Wiesel [20] found that a unique network structure can effectively reduce the complexity of neural network feedback and proposed the concept of the convolutional neural network (CNN). The CNN effectively solved the forecasting problem of the traditional fully connected network algorithm via local connectivity, parameter sharing, and downsampling. Specifically, with local connections, each neuron is not connected to every neuron in the upper layer, effectively reducing the amount of parameters during training. Each group of connections shares the parameter of convolutional kernel, which reduces the amount of training required and accelerates the training speed. The pooling method used for downsampling greatly reduces the feature dimensions and amount of calculation and avoids overfitting. Compared with other neural networks, the CNN shows superior performance in automatically extracting the salient features of an image [21–23]. Liu et al. [24] used the CNN technology to extract extreme weather events (tropical cyclones, atmospheric rivers, and frontal activities) from climate datasets and obtained meaningful results.

Another neural network method is the recurrent neural network (RNN), which is applicable to data with time series features [25–27]. During the actual training process, gradient disappearance and explosion occurred sometimes. To solve these problems, a more complex RNN named long short-term memory network (LSTM) was proposed in [28]. Using LSTM, the medium- and long-term data information of a time series can be well-preserved. It is advantageous to train and model meteorological data with time series characteristics and to conduct forecast research using the LSTM. Akram and El [29] used 15 years of hourly weather data to train the multilayer LSTM model and to forecast meteorological conditions out to 24 and 72 hours. They found that the LSTM can forecast general weather variables with a better accuracy.

The CNN and LSTM are suitable for extracting spatial features and processing time series, respectively. A technique for integrating the two is worth exploring, in order to learn and train datasets with spatial and temporal features more efficiently. Since short-term nowcasting based on radar echoes is a forecasting problem that involves sequences of both spatial and temporal features, Shi et al. [14] proposed the ConvLSTM algorithm, based on both the CNN and LSTM. In the ConvLSTM, on the basis of the traditional LSTM, the convolutional structure was added to each LSTM unit. Compared with the traditional LSTM, this hybrid network has a greater ability to extract the spatial features of radar echoes. By using three-dimensional radar data, Kim et al. [30] applied ConvLSTM to predict the rainfall amount, and the results show that the ConvLSTM is better than FC-LSTM and linear regression.

Although the ConvLSTM algorithm performs well in convection nowcasting based on the training and learning of evolutionary patterns of the radar echoes using historical data, it lacks of dynamic field data (such as U,V wind field) in the input predictors that may be one of the potential factors to the generation or dissipation of echoes. In this paper, we simulate the nearly ideal background wind field (U,V) obtained from radar data using the pyramid optical flow method. Two components of background wind field (U,V) and three channels of Red-Blue-Green images of radar echoes are taken as 5 forecasting factors of ConvLSTM. As the results shown below, it will confirm that the new model performs better than the optical flow method. The present paper is organized as follows. In Section 2, a brief description of the machine learning system and the materials and methods are provided. Results of two cases are presented in Section 3, followed by a summary and conclusion in Section 4.

2. Data and Methodology

In this study, Henan Province in China is selected as the area of interest and the radar data from 2016 to 2017 are used. Based on the Red, Green, and Blue (RGB) image information of the radar images, the improved machine learning algorithm was used to study the nowcasting technique of convective echoes, by which we provide a technical support for nowcasting and early warning of severe convective weather.

2.1. Data Processing. The monitoring data collected by 8 radars (Figure 1) located in Zhengzhou City, Luoyang City, Puyang City, Sanmenxia City, Pingdingshan City, Shangqiu City, Nanyang City, and Zhumadian City in Henan Province were selected. After nonmeteorological clutter had been filtered, beam blockage compensated, and frequency attenuation corrected [31–33], the nearest-neighbor and vertical linear interpolations were applied in the radial and azimuthal directions [34, 35] in order to transform polar coordinates into grid points. Thereafter, the exponential weight function method was utilized to process the radar data from different scanning modes, different bands, and different generation times into data in a 3D Cartesian coordinate system of unified observation time and resolution, resulting in a CAPPI composite map consisting of 31 contour planes on the same base map. The vertical extension height is 18 km and the vertical grid spacing (Δz) is 0.25 km below a height of 3 km and stretches to 0.5 km up to 9 km, after which Δz remains constant at 1 km. The horizontal spacing and time resolution are, respectively, 0.01° and 6 min. The maximum of 31 levels of CAPPI is then mapped to the same layer to obtain the combined reflectivity, which is used as the input radar images to optical flow and ConvLSTM.

Figure 1: Distribution of 8 radars in Henan Province.

In order to make the model more generalized, data of nonsevere and severe weather were both included in the training dataset to learn the evolution characteristics of echoes under various weather conditions. This experiment selected volumetric data of the radars composite from May 2016 to April 2017 for each two consecutive hours. A total of 50,000 samples were used as data sets, 80% of which were used as training sets and 20% as test sets. The first 5 radar images (30 minutes) of radar echo image at each initial time were used as input data and the next 20 images (2 hours) as output. Due to the low proportion of strong convective weather in the sample set, the sample equalization (i.e., resample and data-augmentation technologies) was processed to make the proportion of samples in each batch more balanced. Then, about 5,000 samples selected from May to August 2017 were used as test sets to evaluate the forecasting ability of the model, and 7 strong weather events (Table 1) were selected from the test sets as case studies. Only radar echoes greater than 10 dBZ were considered for the evaluation of echo prediction bias, while echoes greater than 35 dBZ were considered for prediction accuracy and positional deviation.

2.2. Nowcasting Based on Machine Learning. In convection nowcasting, the pyramid delaminating technique [18] was used to classify and preprocess the mosaic radar data. During the process, the dense optical flow method was applied to calculate and track the evolution of each pixel in the network data in order to obtain the dense optical flow field. Feature points were selected from the strong convective cells, and the pyramid delaminating technique was utilized to reduce the dimensions in order to obtain the top-level wind field information satisfying the assumptions of the optical flow method. Through hierarchical iteration from the top to the bottom layers of the pyramid, the sparse optical flow field of the feature points of the severe convective cells in the network data was obtained and then the dense optical flow field was corrected. Finally, a corrected optical flow field, the background horizontal wind field (U,V), was generated.

The CNN structure in the ConvLSTM includes a convolutional layer, a downsampling layer, and a fully connected layer, among which the convolutional and downsampling layers may have a multilayer structure (deep convolutional). The convolutional and downsampling layers were not connected one by one. The next sampling layer could be connected after multiple convolutional layers in order to extract features of the output image in each dimension. With the output feature of the convolutional layer as the input, the fully connected layer acted as a classifier. Compared with the traditional recurrent neutral network (RNN), LSTM introduces three kinds of gate structures (forget gate layer, input gate layer, and output gate layer) to protect and control information. The first thing that LSTM has to solve is to decide which information to pass through this neuron, which is done by the forget gate layer. The input gate layer determines how much information needs to be saved to the current state at the current moment. After the selective memory and update steps, the output gate layer determines the information passed to the next neuron. The convolution kernel only needed to establish full-connection sampling for the local area of the image to extract the underlying information, including local information such as edges and corner points. Through the weight-sharing mechanism, a convolution kernel that had been trained based on a particular characteristic could complete the resampling of similar characteristics for the entire image [36]. Following convolution training, the image features were used as the information to be input into the LSTM. This information was protected and controlled through point-by-point multiplication, as well as the sigmoid activation function [28] of gate structure of the LSTM, so as to finally obtain the extrapolation nowcasting results of the radar reflectivity factor within 0–2 h.

Figure 2 shows the short-term nowcasting process based on pyramid optical flow and ConvLSTM machine learning. The main steps are summarized as follows:

Step 1. Optical flow field computation: the sparse and dense optical flow fields were obtained after preprocessing the mosaic radar data using the pyramid optical flow method. The dense optical flow field was calculated and tracked from every pixel of the radar data, while the sparse optical flow field was only obtained from pixels of the strong convective cells. Therefore, we correct the dense optical flow field with the sparse optical flow field to obtain a more accurate background wind field (U,V).

Step 2. Model configuration: we developed a ConvLSTM that has 3 ConvLSTM2D layers with 40 units in each layer. The ConvLSTM2D layer is an extension of the FC-LSTM layer, which replaced fully connected structures to convolutional structures in both the forget gate and input gate of LSTM.

Step 3. Model training: the input data of ConvLSTM, which is extracted from 5 radar images in the past 30

Table 1: Characteristics of the 7 cases.

Date (DD/MM/YYYY)	Severe weather phenomena	Mesoscale features
19/05/2017	Heavy rainfall, maximum wind gust (20 m/s), maximum diameter of hail (1 cm)	Shear line
22/05/2017	Heavy rainfall, maximum wind gust (25 m/s)	Trough and shear line
12/06/2017	Heavy rainfall, maximum wind gust (21 m/s)	Trough in northern China
20/06/2017-21/06/2017	Thundershower (light rainfall), maximum wind gust (23 m/s)	Cold air diffusing southward in northwesterly airflow
04/07/2017	Heavy rainfall, maximum wind gust (20 m/s)	Cyclonic vortex with shear line
14/07/2017	Heavy rainfall, maximum wind gust (20 m/s), maximum diameter of hail (1 cm)	Cyclonic vortex with shear line
01/08/2017	Continuous heavy rainfall with thunder and lightning	Typhoon trough

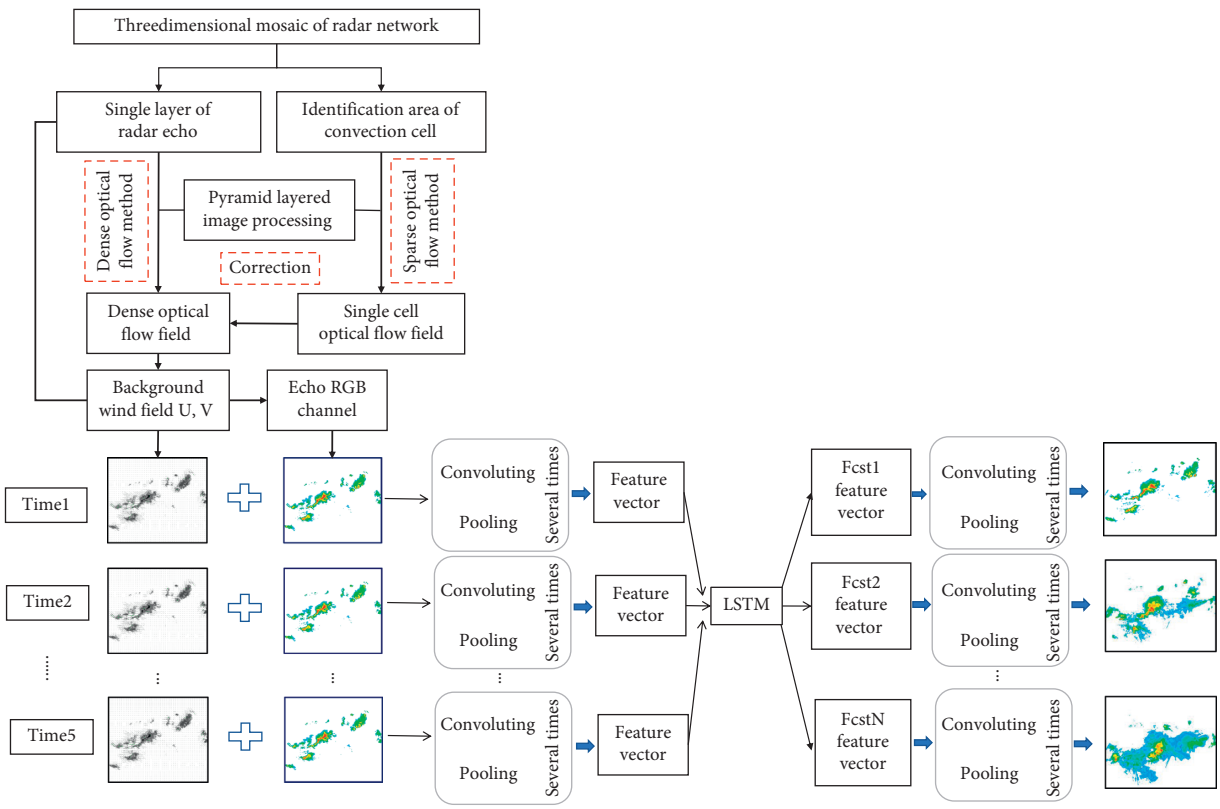

Figure 2: Flowchart of the convective echo extrapolation algorithm based on ConvLSTM machine learning.

minutes, contains 5 dimensions in each training sample: 2 components of background wind field and 3 channels of Red-Green-Blue (RGB) image information of radar echo. The output is 20 radar images in the next 2 hours.

2.3. Measures of Model Skill. In this study, the value of 35 dBZ is taken as the test threshold. To quantitatively describe the position, intensity, and valid period of the forecasted radar echoes, the resulting echoes were compared to the observed grid point by grid point. Considering the impact of the wind field, each grid point of the forecast field and its adjacent 3×3 grid point area were compared in the actual assessment. Both magnitudes, bias and forecast accuracy, were tested; ratio bias (BIAS), root-mean-square error (RMSE), and error of mean (EM) were used as magnitude bias indicators, and probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), and correlation coefficient (CC) were used as forecast accuracy indicators. The specific equations are as follows:

$$\mathrm{POD} = \frac{\mathrm{NA}}{\mathrm{NA} + \mathrm{NC}},$$
$$\mathrm{FAR} = \frac{\mathrm{NB}}{\mathrm{NB} + \mathrm{NC}},$$
$$\mathrm{CSI} = \frac{\mathrm{NA}}{\mathrm{NA} + \mathrm{NB} + \mathrm{NC}}, \tag{1}$$
$$\mathrm{CC} = \frac{\sum_j^n \left(g_{i,j} - \overline{g_i}\right)\left(r_{i,j} - \overline{r_i}\right)}{\sqrt{\sum_j^n \left(g_{i,j} - \overline{g_i}\right)^2 \left(r_{i,j} - \overline{r_i}\right)^2}}.$$

The explanations of verification index (NA, NB, and NC) are listed in Table 2. If the actual grid point value and the forecasted grid point value were both greater than the threshold value, then the grid point was considered to be a successful forecast (NA). If the actual grid point value was greater than the threshold, while the forecasted grid point value was less than the threshold, then the point was considered to be a missed forecast (NC). If the actual grid point value was less than the threshold, while the forecasted grid point value was greater than the threshold, then the point was considered to be a false forecast (NB).

2.4. Computational Platform. A C++ and Python mixed programming of optical flow method and ConvLSTM model was performed on a computer with 2 Intel Xeon E5-2650V3 CPUs (12 cores) and 2 NVIDIA GEFORCE GTX TITAN X GPUs, and the running efficiency of both algorithms is satisfied. For instance, the preprocessesing of 8 radars' data cost approximately 45 seconds. The calculation time of optical flow for each single radar composite image every 6 minutes is within 10 seconds. Although training the ConvLSTM model needs a relatively long time for about 30 hours, the extrapolation time for 2 hours took only 120 seconds after training. Thus, both two nowcasting algorithms in this paper can complete a single computation in 5 minutes.

3. Results

Seven severe convective weather events from May to August 2017 were selected for testing and analyzing nowcasting results based on the model in this study. The selected cases are listed in Table 1. A more in-depth analysis was done with the events on July 14 and August 1.

3.1. Comprehensive Assessment. Figures 3 and 4 show the test results of optical flow and ConvLSTM for the 2 forecasts starting at 1100UTC July 14 2017, and 0400UTC August 1 2017, respectively.

Figures 3(a)–3(c) show the comparison of average forecast bias of radar echoes greater than 10 dBZ within 2 h between optical flow and ConvLSTM. For ConvLSTM, the BIAS was maintained near 1, the absolute value of the EM was within 0.5, and the RMSE was kept within 7 dBZ, whereas the RMSE of optical flow was 1-2 dBZ higher than ConvLSTM on average. The similar results can also be seen from Figures 4(a) and 4(c), indicating a more accuracy in the extrapolation of radar echo location and strength in ConvLSTM.

Table 2: Verifying classification.

Actual value	Forecast value	
	Greater than the threshold	Less than the threshold
Greater than the threshold	NA	NC
Less than the threshold	NB	—

According to the average forecast bias of strong echoes greater than 35 dBZ in Figures 3(b) and 3(d), the forecast accuracy of two algorithms both decreased as the valid forecasting period increased. The POD, CSI, and CC of ConvLSTM show a change range of 0.73 to 0.28, 0.65 to 0.21, and 0.96 to 0.61, respectively, while the same indicators of optical flow show a range of 0.58 to 0.14, 0.52 to 0.13, and 0.91 to 0.43, respectively. From the comparison of Figures 4(c) and 4(d), the result of optical flow method shows a faster descent curve of POD, CSI, and CC than ConvLSTM as well. The results indicate that the average forecast bias of strong echoes of ConvLSTM was less than optical flow within the 2 h valid forecasting period.

Based on the 7 selected weather phenomena listed in Table 1, more than 20 forecast-initiation events were chosen for each process, resulting in a total of more than 150 forecast-initiation events. The TREC, optical flow, and ConvLSTM machine learning proposed in this study were used to conduct 0–2 h extrapolation forecasts on the selected forecast-initiation events. According to the average test results (Table 3), compared with the TREC extrapolation method, both the optical flow method and the ConvLSTM extrapolation method proposed in this study achieved qualitative improvements in forecast accuracy and correlation, as well as precision of forecast position and intensity. The average accuracy improved by more than 30%, illustrating the advantage of the nonlinear extrapolation method. The ConvLSTM method with pyramid optical flow field information further improved forecast accuracy compared with the optical flow method, reflecting the learning ability of the machine learning method.

3.2. Case Analysis

3.2.1. Case 1. Influenced by cyclonic vortex shear, there were thunderstorms in the northern and central parts of Henan Province on July 14-15, 2017, accompanied by strong convective weather such as lightning, strong wind, and heavy rainfall. During this event, echoes developed rapidly, lasted for an extended period, and exhibited a wide range of influence. However, the convective system was scattered and new echoes were constantly being generated, making it difficult to conduct an early-warning forecast of short-term heavy rainfall and thunderstorms.

Figures 5(a) and 5(c) show the convective echoes intensity and movement from 1100UTC–1300UTC. The convective echoes generally moved eastward, although the

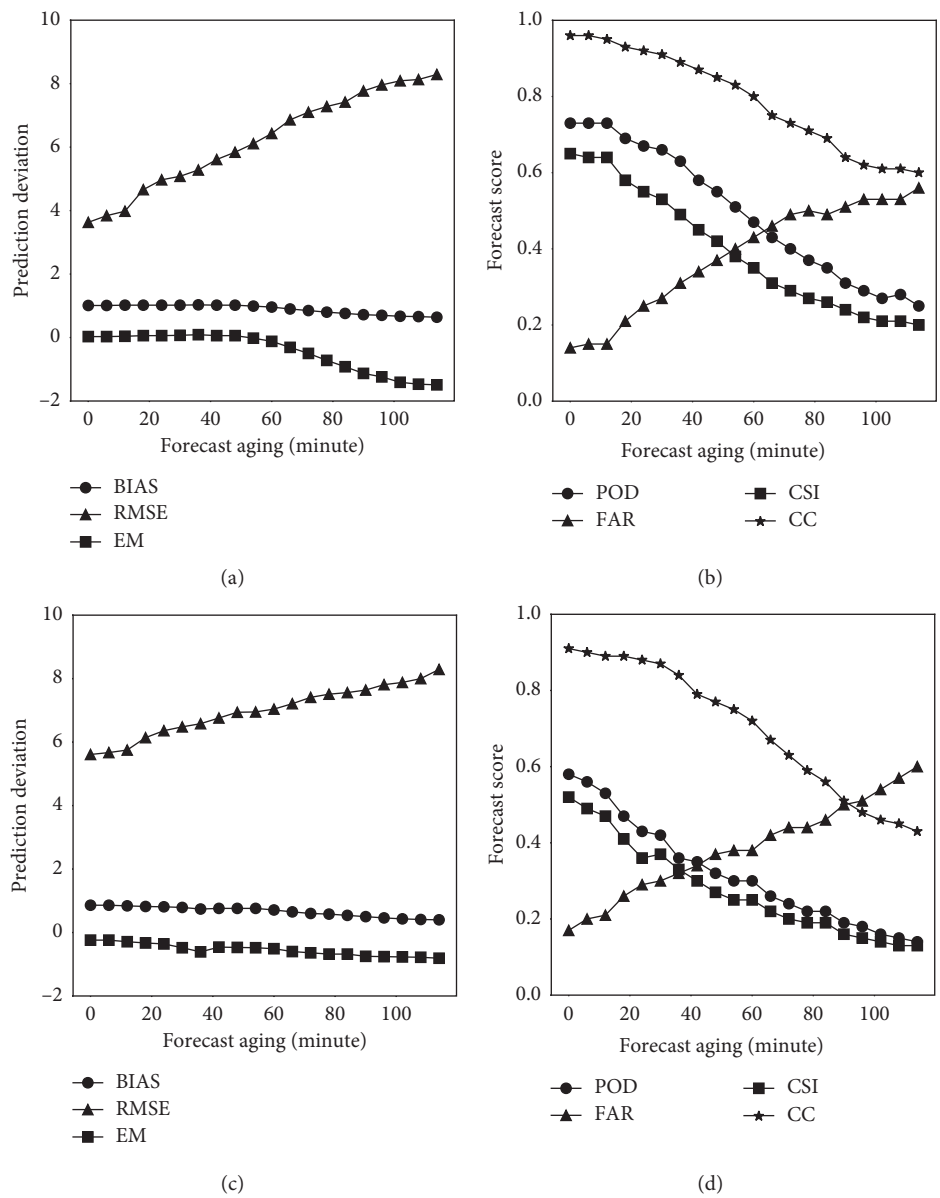

Figure 3: Forecast bias and forecast score for 2 hours starting at 1100UTC July 14 2017: (a, b) the results of the ConvLSTM method; (c, d) the results of the optical flow method.

speed of echo movement varied from region to region. The echoes in the central and southern areas moved relatively faster, whereas the echoes in the northern region propagated more slowly. The convective clouds in the northern region were generated at 1100UTC and tended to decay by 1254UTC, while the clouds in the south area evolved from stratified mixed clouds into strong convective clouds during this period. It can be seen from Figures 5(b) and 5(d) that based on the composited radar echo data of the 2.5 km height layer of the front and rear frames, the horizontal moving speed of the radar echo obtained using the pyramid optical flow technique has certain indication significance for the radar intensity change trend. The cloud system in the north had a low inversion wind speed and was divergent, whereas in the south, the northwest wind had a high speed and the wind field was characterized by obvious convergence, which was consistent with the variation pattern of the echoes, that is, moving quickly and being generated in the south while moving slowly and dissipating in the north.

The timing when numerous convective echoes initiated (i.e., 1100UTC July 14) was selected as the onset time for the 0–2 h extrapolation forecast. Figure 6 shows the echo forecast based on optical flow and machine learning. The forecasting start time was 1100UTC July 14 2017, and the valid periods of the forecast were the next 30, 60, 90, and 120 minutes. Comparison of the distribution of ConvLSTM forecasted echoes and the corresponding real-time radar echoes shows that the forecasted position and shape of the echoes over the 2 hour forecast period were in good agreement with the actual weather conditions. The forecasted echoes in southern Henan Province developed faster

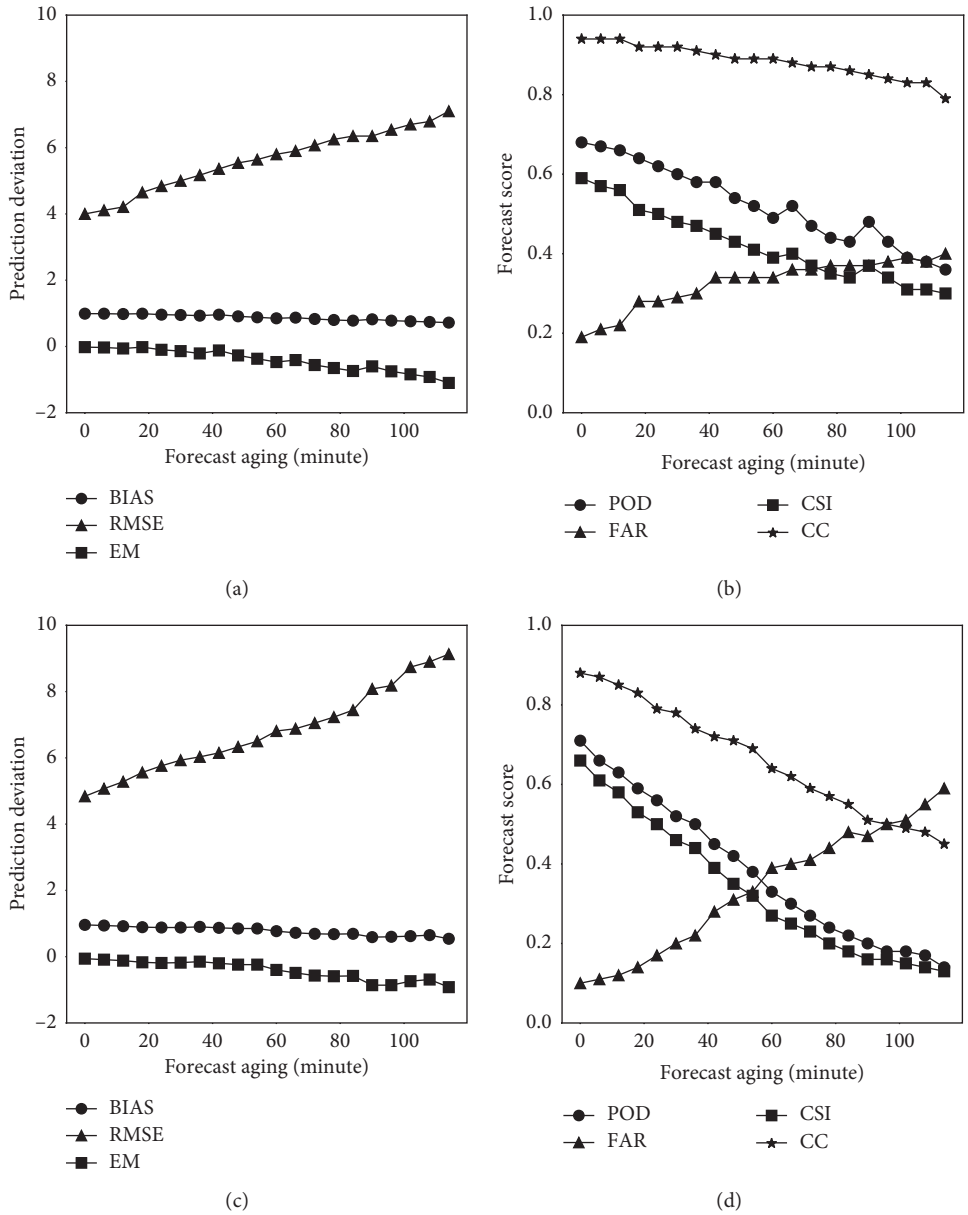

Figure 4: Forecast bias and forecast score for the 2 hours starting at 0400UTC August 1 2017: (a, b) the results of the ConvLSTM method; (c, d) the results of the optical flow method.

Table 3: Results of the comprehensive assessment.

Forecast method	Valid period of forecast (min)	POD	CSI	CC	Forecast centroid deviation (km)	RMSE (dBZ)
TREC		0.291	0.254	0.489	18.1	21.5
Optical flow method	60	0.482	0.428	0.728	12.8	15.3
Machine learning method		0.544	0.465	0.821	11.3	12.9
TREC		0.182	0.114	0.320	27.5	28.4
Optical flow method	120	0.216	0.195	0.466	17.8	20.3
Machine learning method		0.383	0.287	0.513	17.1	15.8

to the east with the increase of deviation degree and the valid forecasting period; the maximum amount of deviation can be controlled to a range of 0.03–0.05 (about 3–5 km). Based on the echo intensity, the characteristics indicating the evolution of the stratiform cloud system into the convective cloud system failed to be predicted, although they were reflected in the dissipation process of convective cells in the north, and there were similar forecasts of the evolution of

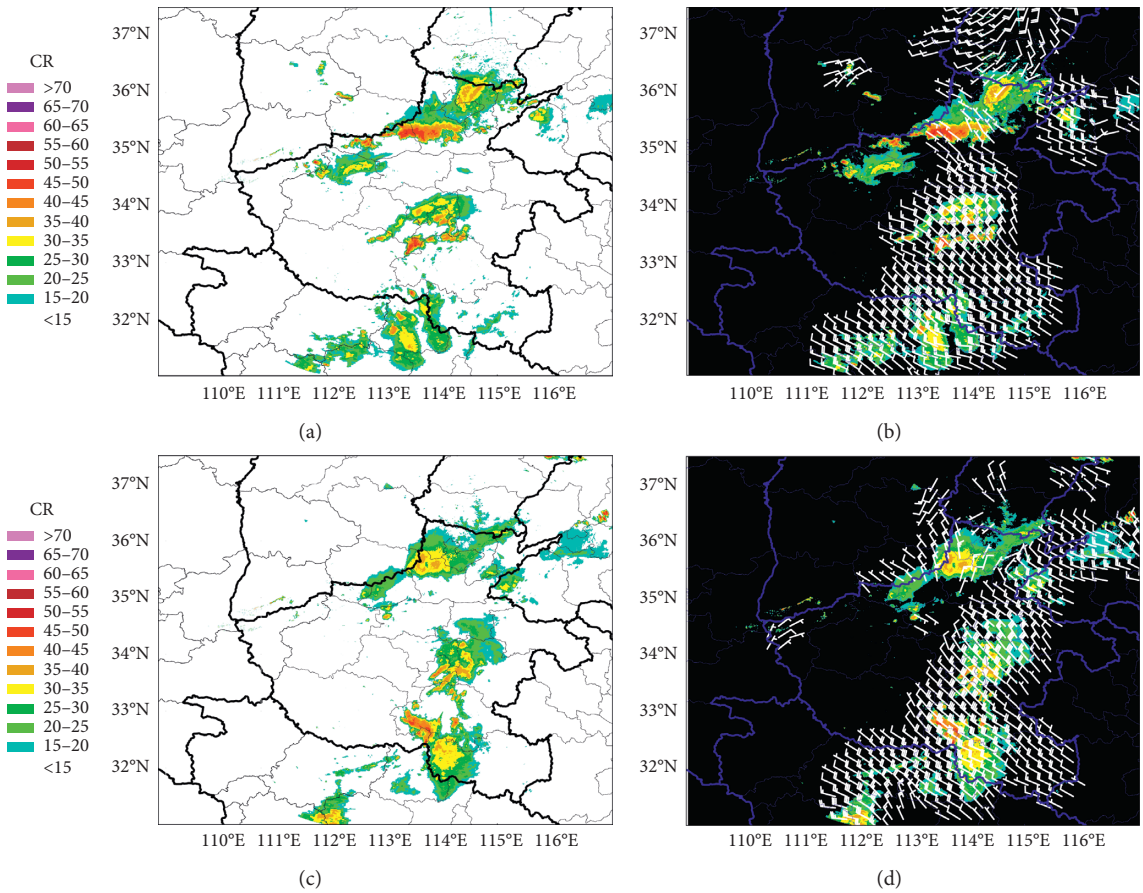

Figure 5: Evolutionary characteristics of the composite reflectivity (a, c) and wind field (b, d) at 1100UTC and 1254UTC July 14 2017.

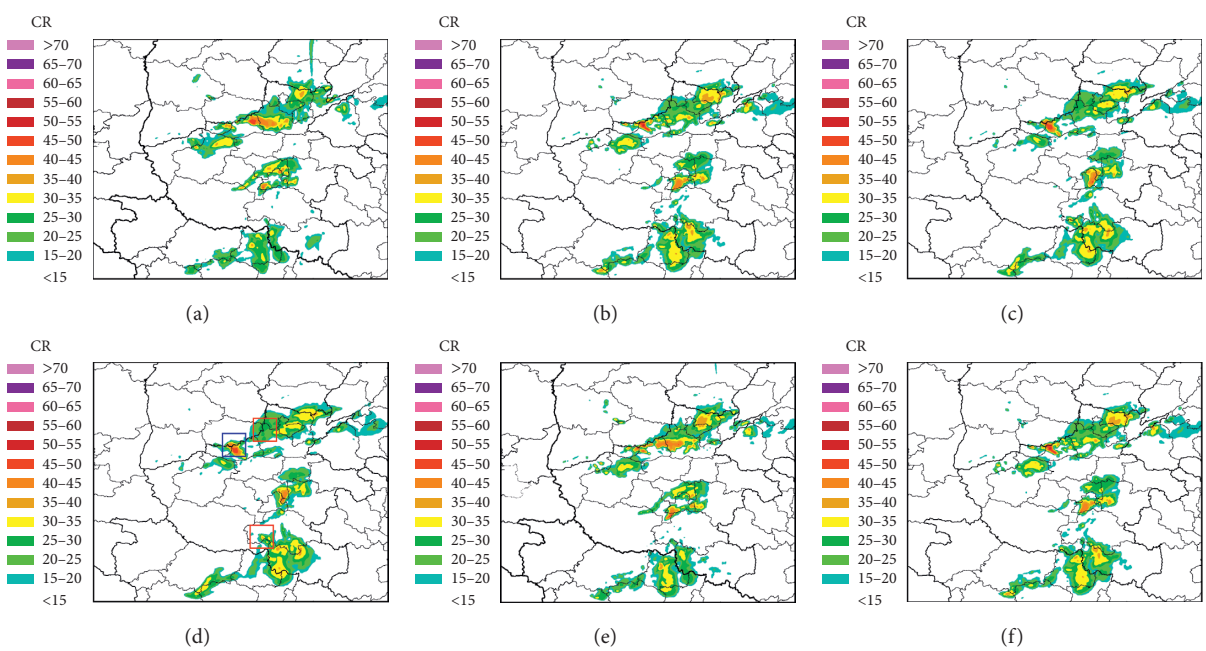

Figure 6: Continued.

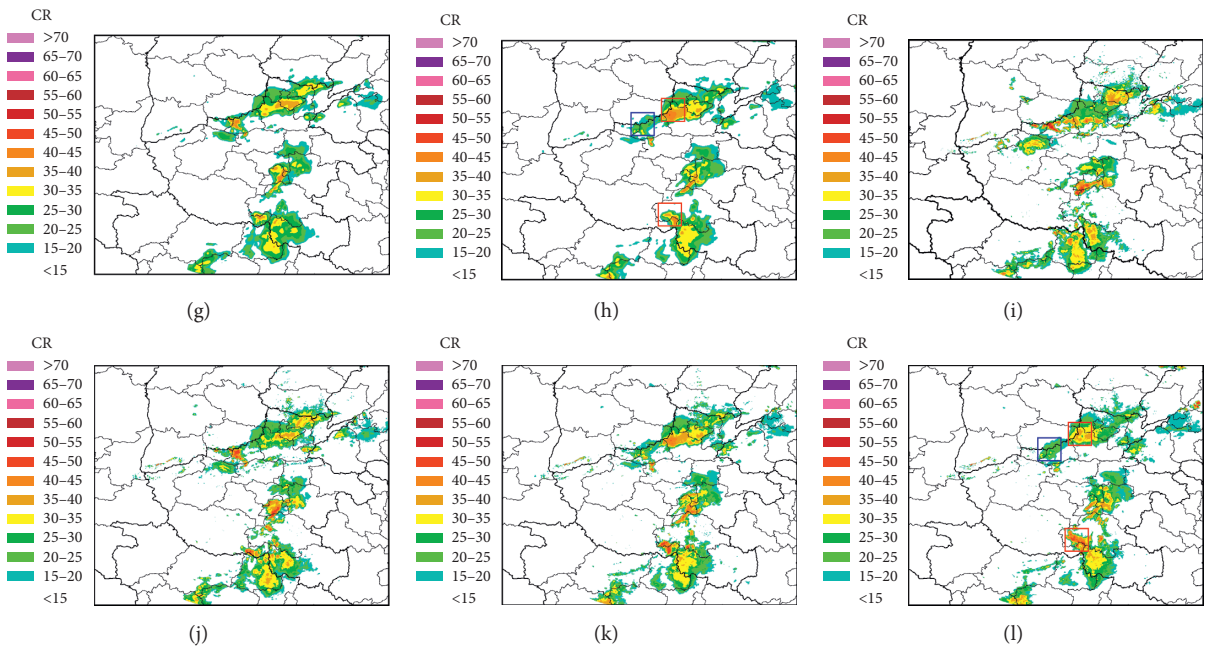

Figure 6: Comparison between the composite reflectivity factor forecast results and the actual weather conditions for the 2 h starting at 1100UTC July 14 2017. The forecast results of the optical flow method: (a) 30 min, (b) 60 min, (c) 90 min, and (d) 120 min after the start time of 1100UTC, respectively; the forecast results of the ConvLSTM method: (e) 30 min, (f) 60 min, (g) 90 min, and (h) 120 min after the start time of 1100UTC, respectively; the actual weather conditions: (i) 30 min, (j) 60 min, (k) 90 min, and (l) 120 min after the start time of 1100UTC, respectively; red square lines represents the enhancement area of observe radar echoes and blue square lines for the weakened area.

echo formation in central Henan Province. For this case, the traditional optical flow method only predicts the strong echo in northern Henan Province while failing to predict its weakening of strength and the evolution of radar echoes in southern and central Henan Province. During the development of this cloud system, in addition to the echo forms changing significantly, there was also the generation and dissipation of convective cells, occurrences that were reflected in the extrapolation algorithm proposed in this study. This also illustrates the improvement of this method compared with the traditional extrapolation method. With machine learning and the pyramid optical flow method, the convergence and divergence characteristics of the wind field can be accurately described, indirectly reflecting upward and downward air motions. Therefore, not only can this method accurately forecast the position and shape of radar echoes, but it can also forecast the generation and dissipation of local convection.

3.2.2. Case 2. Affected by the westward movement of the inverted trough of Typhoon Haitang, on August 1, a spiral-shaped strong convective system moving from east to west developed in Henan Province. Of strong intensity and long duration, this system brought a wide range of heavy rainfall to the eastern part of Henan Province, with a heavy rainstorm occurring in the northeastern area. As shown in Figure 7, the spiral cloud band of the typhoon was moving from east to west at 0400UTC on August 1, affecting the eastern and southeastern areas of Henan Province. There was an arc-shaped strong convection zone in the northern section of the cloud cluster, with strong echoes constantly rotating into Henan Province. As shown in Figures 7(a) and 7(c), the strong precipitation echoes in the southern section of the cloud cluster were separated from the main rain band during its movement towards the west, and the southern part of the system became an independent precipitation cloud cluster within 2 h. In addition, it can be seen from Figures 7(b) and 7(d) that there were obvious counterclockwise rotation characteristics in the wind field. The northerly wind speed was relatively high, but the wind speed in the south was low, and the wind was blowing westward, consistent with the separation process of the echoes in the southern part of the system.

At 0400UTC, when the convective system became mature, the 0–2 h extrapolation forecast started. Figure 8 shows the distribution of the forecasted radar echoes and the corresponding actual radar echoes for the next 30, 60, 90, and 120 min, starting at 0400UTC August 1 2017. Due to the concentrated nature of the convective system during this period, the forecasted shape and position of the echoes of ConvLSTM method in the 2 h forecast window were highly consistent with the actual radar echoes. The shape and movement trends of the convective cloud clusters in the stratiform cloud system were also accurately captured. In particular, the strong echoes rotating into the eastern part of the system could be predicted. However, the traditional optical flow method failed to predict the weakening process of echoes in southern and central Henan Province, and the strength of forecasted strong radar echo in northern Henan

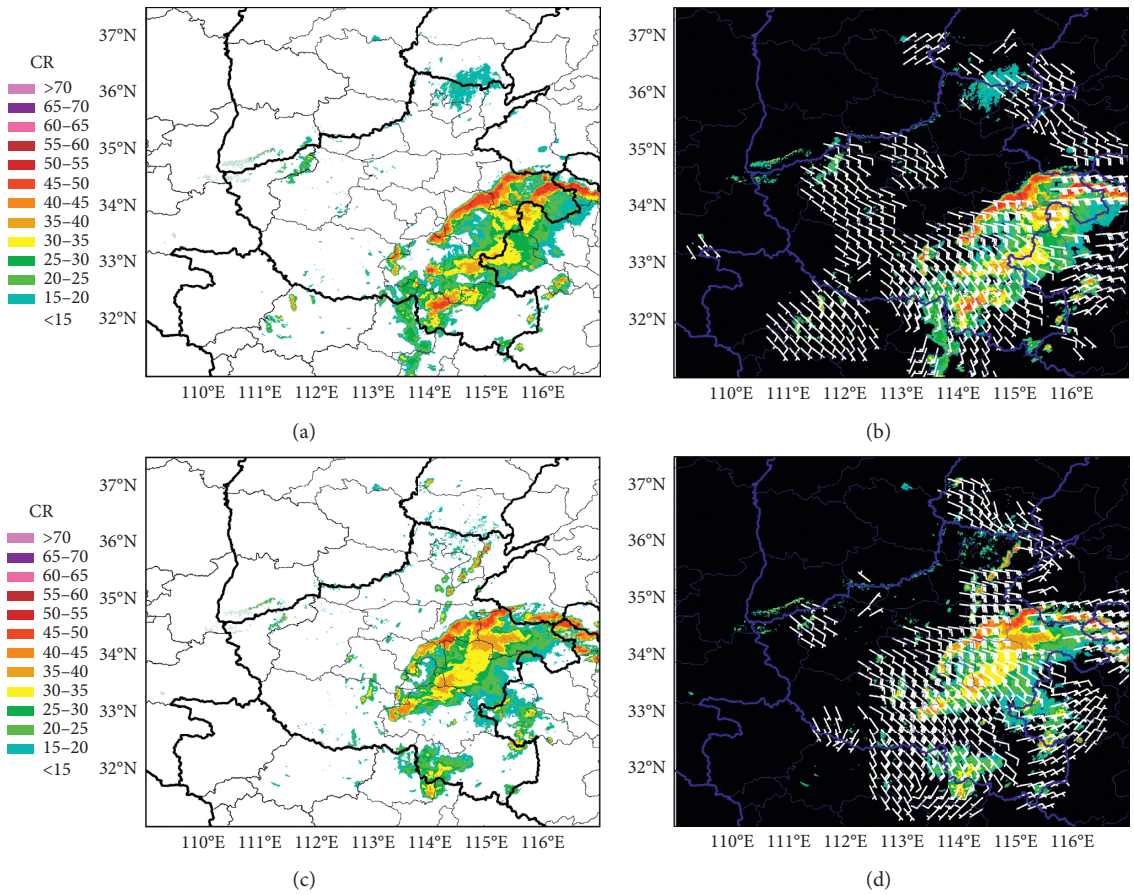

Figure 7: Evolutionary features of the composite reflectivity (a, c) and the wind field (b, d) at 0400UTC and 0554UTC August 1 2017.

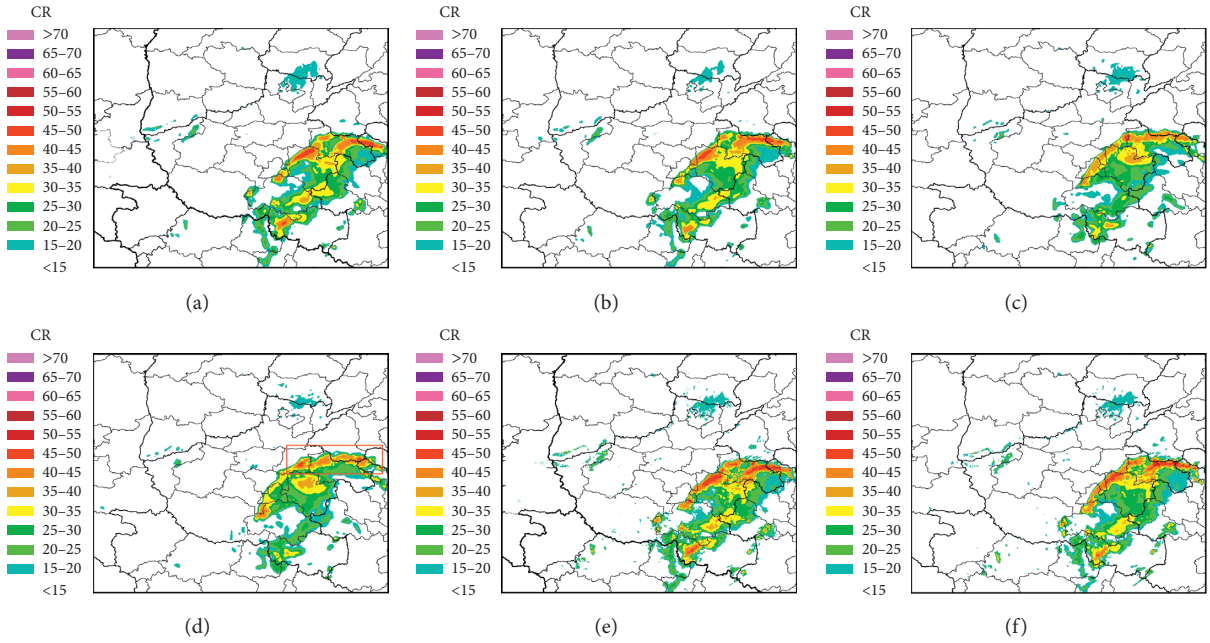

Figure 8: Continued.

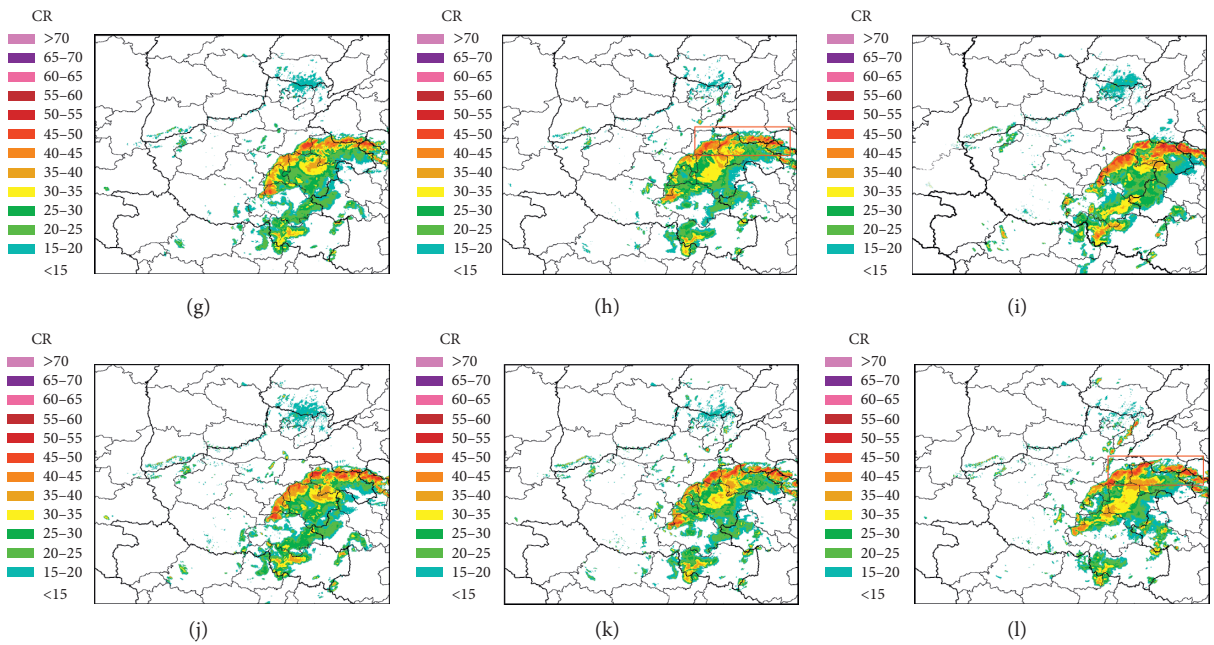

Figure 8: Comparison between the composite reflectivity factor forecast results and the actual weather conditions for the 2 h starting at 0400UTC August 1 2017. The forecast results of the optical flow method: (a) 30 min, (b) 60 min, (c) 90 min, and (d) 120 min after the start time of 1100UTC, respectively; the forecast results of the ConvLSTM method: (e) 30 min, (f) 60 min, (g) 90 min, and (h) 120 min after the start time of 1100UTC, respectively; the actual weather conditions: (i) 30 min, (j) 60 min, (k) 90 min, and (l) 120 min after the start time of 1100UTC, respectively; red square lines represents the area where the predicted radar echoes of ConvLSTM are closer to the observation than the prediction of the optical flow method.

Province was also weaker compared with observation. The above results indicate that the learning ability of the algorithm proposed in this paper has significant advantages over the traditional extrapolation algorithm. The central portion of the forecasted echoes gradually dissipated, and the southern section became separated, which agreed with the evolution of the actual radar echoes. However, the dissipation and separation speeds of the forecasted echoes were slower than those of the observed radar elements. Therefore, even with the optical flow field information added to the machine learning training set, the ability of the proposed method to separate, combine, and forecast convective systems still has a room for significant improvement.

4. Conclusions

Our experimental results demonstrated that the global optical flow field generated by pyramid delaminating technique can better reflect the movements of radar echoes, including nonlinear motions such as rotation, convergence, and divergence of local radar elements. By adding the background horizontal wind field (U,V) and Red-Green-Blue (RGB) image information to the ConvLSTM extrapolation nowcasting model, the forecast accuracy of echo position and intensity was further improved and the generation, dissipation, and merging of convective cells were also better identified.

Compared with the TREC method in business application, the machine learning method achieved qualitative improvements in forecast accuracy and correlation, as well as precision of forecast position and intensity. Moreover, the machine learning method also effectively eliminates the false merging of convective cells and improves the accuracy of convective cell recognition. Compared with the traditional optical flow method, the extrapolation of machine learning method is closer to the actual trajectory of radar echoes, especially in the case of cyclonic systems, for which the rotation vectors and curve movement trajectories can be generated.

As for scattered convective weather, the valid forecasting period of ConvLSTM for fast-developing convective cells is relatively short, ranging from 30–60 min. When it comes to systematic convective weather covering a broad expanse, with long system duration and complete echo structure, the ConvLSTM machine learning method has a greater learning capacity and higher forecasting accuracy than the traditional method [4, 37, 38].

Data Availability

Due to the Chinese Meteorological Law, the Doppler radar data provided by Chinese Meteorological Administration in our research are not open to the public.

Disclosure

Chen Yungang graduated from the University of Chinese Academy of Sciences with a major in meteorology, and now he is working in a meteorological company currently not affiliated with any academic institution.

Conflicts of Interest

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

Acknowledgments

This study was supported by the China Meteorological Administration 2017 Special Fund for the Development of Key Technologies in Weather Forecasting (YBGJXM(2017; 2018)02-08) and the National Key R&D Program of China (2017YFC1502000).

References

[1] X. D. Yu, X. G. Zhou, and X. M. Wang, "The advance of the nowcasting techniques," *Acta Meteor Sinica*, vol. 70, no. 3, pp. 311–337, 2012, in Chinese.

[2] R. Crane, "Automatic cell detection and tracking," *IEEE Transactions on Geoscience Electronics*, vol. 17, no. 4, pp. 250–262, 1979.

[3] M. Dixon and G. Wiener, "TITAN: thunderstorm identification, tracking, analysis, and nowcasting-a radar-based methodology," *Journal of Atmospheric and Oceanic Technology*, vol. 10, no. 6, pp. 785–797, 1993.

[4] J. T. Johnson, P. L. MacKeen, A Witt et al., "The storm cell identification and tracking algorithm: an enhanced WSR-88d algorithm," *Weather and Forecasting*, vol. 13, no. 2, pp. 263–276, 1998.

[5] R. E. Rinehart and E. T. Garvey, "Three-dimensional storm motion detection by conventional weather radar," *Nature*, vol. 273, no. 5660, pp. 287–289, 1978.

[6] J. W. Wilson, N. A. Crook, C. K. Mueller, J. Sun, and M. Sun, "Nowcasting thunderstorms: a status report," *Bulletin of the American Meteorological Society*, vol. 79, no. 10, pp. 2079–2099, 1998.

[7] M. X. Chen, Y. C. Wang, and X. D. Yu, "Improvement and application test of TRCE algorithm for convective storm nowcast," *Applied Meteorological Science*, vol. 18, no. 5, pp. 690–701, 2007, in Chinese.

[8] J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA, USA, 1950.

[9] U. Germann and I. Zawadzki, "Scale-dependence of the predictability of precipitation from continental radar images. Part I: description of the methodology," *Monthly Weather Review*, vol. 130, no. 12, pp. 2859–2873, 2002.

[10] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: a survey," *ACM Computing Surveys*, vol. 38, no. 4, pp. 1–45, 2006.

[11] L. Han, H. Q. Wang, and Y. J. Lin, "Application of optical flow method to nowcasting convective weather," *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 44, no. 5, pp. 751–755, 2008.

[12] H. Sakaino, "Spatio-temporal image pattern prediction method based on a physical model with time-varying optical flow," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 3023–3036, 2013.

[13] C. Y. Cao, Y. Z. Chen, D. H. Liu et al., "The optical flow method and its application to nowcasting," *Acta Meteor Sinica*, vol. 73, no. 3, pp. 471–480, 2015.

[14] X. Shi, Z. Chen, H. Wang et al., "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," *Advances in Neural Information Processing Systems*, vol. 9199, pp. 802–810, 2015.

[15] Z. B. Wang, Y. J. Xiao, and T. Wu, "Motion estimation for radar image based on improved optical flow method," *Computer Technology for Developing Areas*, vol. 27, no. 12, pp. 170–175, 2017.

[16] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," *RCA Engineer*, vol. 29, no. 6, 1984.

[17] J.-Y. Bouguet, *Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the Algorithm*, Microsoft Research Labs, Redmond, WA, USA, 1999.

[18] Z. J. Jiang and H. R. Yi, "An image pyramid based feature detection and tracking algorithm," *Geomatics and Information Science of Wuhan University*, vol. 32, no. 8, pp. 680–683, 2007.

[19] L. Zhang, M. Wei, R. Li et al., "Improved optical flow method application to extrapolate radar echo," *Science Technology and Engineering*, vol. 14, no. 32, pp. 133–137, 2014.

[20] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *The Journal of Physiology*, vol. 160, no. 1, pp. 106–154, 1962.

[21] S. Lawrence, C. L. Giles, A. C. Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.

[22] J. V. Back, K. C. Koch, E. Kleinhans et al., "Convolutional networks and applications in vision," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 253–256, IEEE, Paris, France, May 2010.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, pp. 1097–1105, Curran Associates Inc., Siem Reap, Cambodia, December 2018.

[24] Y. Liu, E. Racah, Prabhat et al., "Application of deep convolutional neural networks for detecting extreme weather in climate datasets," 2016, https://arxiv.org/abs/1605.01156.

[25] T. Mikolov, M. Karafiát, L. Burget et al., "Recurrent neural network based language model," in *Proceedings of the Conference of the International Speech Communication Association (INTERSPEECH 2010)*, pp. 1045–1048, DBLP, Chiba, Japan, September 2010.

[26] J. Donahue, L. A. Hendricks, M. Rohrbach et al., "Long-term recurrent convolutional networks for visual recognition and description," in *AB Initto Calculation of the Structures and Properties of Molecules*, pp. 85–91, Elsevier, Amsterdam, Netherlands, 2015.

[27] A. Karpathy and F. F. Li, "Deep visual-semantic alignments for generating image descriptions computer vision and pattern recognition," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3128–3137, IEEE, Boston, MA, USA, June 2015.

[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[29] M. Akram and C. El, "Sequence to sequence weather forecasting with long short-term memory recurrent neural networks," *International Journal of Computer Applications*, vol. 143, no. 11, pp. 7–11, 2016.

[30] S. Kim, S. Hong, M. Joh et al., "Deep rain: ConvLSTM network for precipitation prediction using multichannel radar data," 2017, https://arxiv.org/abs/1711.02316.

[31] M. Steiner and J. A. Smith, "Use of three-dimensional reflectivity structure for automated detection and removal of nonprecipitating echoes in radar data," *Journal of*

Atmospheric & Oceanic Technology, vol. 19, no. 19, pp. 673–686, 2002.

[32] C. Kessinger, S. Ellis, and J. VanAndel, "NEXRAD data quality: the AP clutter mitigation scheme," in *Proceedings of the 30th International Conference on Radar Meteorology*, pp. 707–709, American Meteor Society, Munich, Germany, 2001.

[33] R. A. Fulton, J. P. Breidenbach, D. J. Seo, D. A. Miller, and T. Miller, "The WSR-88d rainfall algorithm," *Weather and Forecasting*, Weather & Forecasting, vol. 13, no. 2, pp. 377–395, 1998.

[34] Y. J. Xiao and L. P. Liu, "Study of methods for interpolating data from weather radar network to 3-D grid and mosaics," *Acta Meteor Sinica*, vol. 64, no. 5, pp. 647–657, 2006, in Chinese.

[35] G. E. Klazura and D. A. Imy, "A description of the initial set of analysis products available from the NEXRAD WSR-88d system," *Bulletin of the American Meteorological Society*, vol. 74, no. 7, pp. 1293–1311, 1993.

[36] F. P. Liu, *Application of Neural Network and Deep Learning*, Publishing House of Electronics Industry, Beijing, China, 2018.

[37] X. Shi, Z. Gao, L. Lausen et al., "Deep learning for precipitation nowcasting: a benchmark and a new model," in *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, December 2017.

[38] Z. H. Zhou and J. Feng, "Deep forest: towards an alternative to deep neural networks," 2017, https://arxiv.org/abs/1702.08835.