Forecasting Different Types of Convective Weather: A Deep Learning Approach

Kanghui ZHOU[1,2,3*], Yongguang ZHENG[3], Bo LI[4], Wansheng DONG[1], and Xiaoling ZHANG[3]

1 *Chinese Academy of Meteorological Sciences*, *China Meteorological Administration*, *Beijing* 100081, *China*
2 *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*
3 *National Meteorological Center*, *China Meteorological Administration*, *Beijing* 100081, *China*
4 *University of Illinois at Urbana-Champaign*, *Champaign*, *IL* 61820, *USA*

(Received October 24, 2018; in final form June 13, 2019)

ABSTRACT

A deep learning objective forecasting solution for severe convective weather (SCW) including short-duration heavy rain (HR), hail, convective gusts (CG), and thunderstorms based on numerical weather prediction (NWP) data was developed. We first established the training datasets as follows. Five years of severe weather observations were utilized to label the NCEP final (FNL) analysis data. A large number of labeled samples for each type of weather were then selected for model training. The local temperature, pressure, humidity, and winds from 1000 to 200 hPa, as well as dozens of convective physical parameters, were taken as predictors in our model. A six-layer convolutional neural network (CNN) model was then built and trained to obtain optimal model weights. After that, the trained model was used to predict SCW based on the Global Forecast System (GFS) forecast data as input. The performances of the CNN model and other traditional methods were compared. The results show that the deep learning algorithm had a higher classification accuracy on HR and hail than support vector machine, random forests, and other traditional machine learning algorithms. The objective forecasts by use of the deep learning algorithm also showed better forecasting skills than the subjective forecasts by the forecasters. The threat scores (TSs) of thunderstorm, HR, hail, and CG were increased by 16.1%, 33.2%, 178%, and 55.7%, respectively. The deep learning forecast model is currently used in the National Meteorological Center of China to provide guidance for the operational SCW forecasting over China.

Key words: deep learning, convolutional neural network, convective weather, forecasting

Citation: Zhou, K. H., Y. G. Zheng, B. Li, et al., 2019: Forecasting different types of convective weather: A deep learning approach. *J. Meteor. Res.*, **33**(5), 797–809, doi: 10.1007/s13351-019-8162-6.

1. Introduction

Severe convective weather (SCW), which includes thunderstorms and/or lightning, hail, convective gusts (CG), short-duration heavy rain (HR), and tornadoes, poses a serious threat to life and property in most areas of the world. Due to the rapid evolution of small scale convective systems and their complicated interaction with environmental features, forecasting SCW is still a challenging issue in operational meteorology today (Ray, 1986; Stensrud et al., 2009).

At present, the National Meteorological Center (NMC) of the China Meteorological Administration issues subjective SCW forecasts using the ingredients-based (IB) method, which was first proposed by Doswell III et al. (1996) and then further developed in China (Zhang et al., 2010; Yu, 2011). The IB method determines the basic ingredients of SCW events, which generally include relatively independent meteorological variables or parameters, such as potential instability, atmospheric moisture, lifting indices, vertical wind shear, etc. (Doswell III et al., 1996). Different synoptic situations require different thermodynamic and dynamic parameters of the convective environment, which are subjectively determined by the meteorologists based on their experience and knowledge of SCW. These ingredients can provide a clear idea of weather conditions for forecasters/meteorologists. The evaluation of subjective forecasts shows that the IB

Supported by the National Key Research and Development Program of China (2017YFC1502000 and 2018YFC1507504) and National Natural Science Foundation of China (41375051). Li's research is partially supported by AGS-1602845 and DMS-1830312 of the US National Science Foundation.

*Corresponding author: zhoukh@cma.gov.cn.

method is quite effective (Zhang et al., 2010; Yu, 2011). However, there are still some limitations in applying this method. First, due to vast extent and extremely complex terrains of China, climatological features in different regions appear to be significantly different. As a result, a variety of synoptic conditions, such as cold fronts and easterly waves, can lead to convective storms (Meng et al., 2013; Xia et al., 2015; Yang et al., 2017). Therefore, it is difficult to achieve accurate forecasts of strong convection in different regions of China using uniform thresholds of different ingredient variables. Second, with the rapid development of numerical weather prediction (NWP) and meteorological observation networks, the amount of available meteorological information has been exploding during recent years. It is almost beyond the capability of meteorologists to discover and synthesize useful and valuable information from the massive amount of data available without computer support. In addition, implementation of the IB method requires many meteorologists to reinforce their scientific understandings of convective systems, because meteorologists who lack sufficient knowledge and fail to keep up with new scientific developments may not be able to fully capture the valuable information for making optimal SCW forecasts.

Compared to the fact that the subjective extraction of physical features is often limited by the meteorologists' understanding about SCW, the machine learning (ML) method is less dependent on the experience and knowledge of users. Many attempts have already been made using traditional ML algorithms, such as artificial neural network (ANN), support vector machine (SVM), and random forest (RF), for weather forecasting (Gardner and Dorling, 1998). Manzato (2005, 2007) and Chaudhuri (2010) used indices derived from atmospheric sounding data to develop a short-term thunderstorm and rainfall forecasting tool based on ANN that can be applied to different regions. Their results showed that ANN can be a powerful statistical method for performing a multivariate data analysis. ML has also been applied to hail forecasting (Manzato, 2013; Gagne II et al., 2015, 2017), tornado prediction and detection (Marzban and Stumpf, 1996; Lakshmanan et al., 2005), damaging winds prediction (Marzban and Stumpf, 1998; Lagerquist et al., 2017), extreme precipitation forecasting (Gagne II et al., 2014; Herman and Schumacher, 2018), and thunderstorms nowcasting (Han et al., 2017), and the results all seem encouraging.

A deep neural network (DNN) is an ANN with multiple hidden layers between the input and output layers (Bengio, 2009; Schmidhuber, 2015). Similar to traditional machine learning algorithms like ANN and SVM, the DNN can model complex nonlinear systems. Moreover, compared to the traditional algorithms, DNN has been shown to perform better at extracting advanced features by using deeper layers. DNN has a wide range of applications in computer vision (Kubat et al., 1998; Beijbom et al., 2012; Simonyan and Zisserman, 2015), face recognition (Matsugu et al., 2003), and medical diagnosis (Mac Namee et al., 2002; Grzymala-Busse et al., 2004). It can yield results comparable to and in some cases superior to those produced by human experts (Krizhevsky et al., 2012; Ciregan et al., 2012).

There have been some preliminary applications of deep learning to meteorology. For example, a deep learning algorithm was employed to capture spatiotemporal correlations from the radar echo spatiotemporal sequences to obtain the extrapolation vectors, and then further used to predict the development and movement of the radar echoes. In particular, Klein et al. (2015) created a dynamic convolutional layer, Shi et al. (2015) created a convolutional long short-term memory (ConvLSTM) network, and Wang et al. (2017) created a predictive recurrent neural network (PredRNN). Evaluation of the predictions indicates that the deep learning solutions can provide better predictions than traditional algorithms such as the optical flow method. Zhang et al. (2017) showed that convective storm initiation, growth, and advection can be simultaneously better predicted with a deep learning framework when multi-source meteorological data are available. In their study, a five-layer convolutional neural network (CNN) was constructed to extract features from radar and reanalysis data created by variational Doppler radar analysis system (VDRAS). Their experimental results showed that deep learning methods achieve better performance than traditional extrapolation methods. Gope et al. (2016) created a storm forecasting model based on historical climatological data with stacked automatic encoder (SAE). This model is a type of DNN model, and has successfully forecasted heavy rainfall 6–48 h in advance in Mumbai and Calcutta, with less false alarms than conventional methods.

Complex physical processes and dynamic characteristics are often involved in convective systems at small spatial and temporal scales (Doswell III, 2001). Thus, in order to improve the prediction for SCW, it is critical to understand the mechanism of their occurrence and development under various conditions, to fully extract convective characteristics automatically for various types of SCW, and to comprehensively consider its geographical environments and climatological background. Deep learning provides a practical tool that can effectively improve the forecasts of SCW.

In this study, a deep CNN was constructed and trained for thunderstorms, HR, hail, and CG forecasting based on the data from Global Forecast System (GFS) of NCEP. This is essentially a variation of the perfect-prognosis method (Klein et al., 1959), which now is improved by replacing its trainer with more powerful deep learning algorithms. Our proposed forecast postprocessing method can be applied to provide real-time objective probabilistic forecasts over the entire China.

2. Data

2.1 *NWP data*

The data used for establishing different deep learning models are extracted from the global 1° × 1° NCEP final (FNL) analysis data during the period 2010–14. All data are available 4 times a day (0000, 0600, 1200, 1800 UTC), providing global scenarios of weather. After the deep learning model is established, the 1° × 1° forecast data from GFS are used for forecasting SCW.

In order to accelerate the training process and improve the predictive accuracy of the deep network, we first select a set of predictors among all variables in the FNL analysis data. The predictors contain all major environmental conditions that are favorable for SCW events, which include basic meteorological elements such as pressure, temperature, geopotential height, humidity, and wind, as well as a number of convective physical parameters that can reflect water vapor, atmosphere instability, and uplift conditions (Tian et al., 2015). For example, most unstable convective available potential energy (MUCAPE), precipitable water (PWAT), convective inhibition (CIN), convergence, and wind shear are such physical parameters. To account for the geographical differences between various regions, we also use elevation, longitude, and latitude in our model. In total, 144 predictors were selected to describe environmental characteristics of SCW (Table 1) and all those predictors were extracted from the FNL analysis data.

2.2 *SCW observations*

The SCW events, in particular the HR, hail, and CG, are very rare. On average, the maximum number of days

Table 1. Selected predictors for SCW forecasting with the deep learning model

	Feature	Level (hPa)
Multi-level variable	*T* (temperature) *H* (geopotential height) WS (wind speed) WD (wind direction) *W* (vertical wind speed) TDD (temperature dew point difference) *Q* (specific humidity) VAPFLUXDIV (water vapor flux divergence) PV (potential vorticity) TMPADV (temperature advection) SITASE (potential pseudo-equivalent temperature) DIV (divergence) VOR (vorticity) VORADV (vorticity advection)	1000, 925, 850, 700, 600, 500, 400, 300, 200
Single-level convective parameter	MUCAPE (most unstable convective available potential energy) BLI (best lift index) CIN (convective inhibition) DCAPE (downdraft convective available potential energy) K (K-index) LI (lift index) Z0 (altitude of 0°C) Z20 (altitude of 20°C) PWAT (precipitable water) SHIP (significant hail parameter) SHEAR1 (0–1-km wind shear) SHEAR3 (0–3-km wind shear) SHEAR6 (0–6-km wind shear) SI (Showalter index) TT (total index)	
Others	Elevation Longitude Latitude	

per year with thunderstorm, HR, hail, and CG is less than 110, 13, 5, and 13, respectively. Thunderstorm and HR seem to share somewhat similar spatial pattern. They both occur most frequently in South China. However, only thunderstorm while no HR is observed in West China. The spatial pattern of hail and CG also appears to be similar with hot-spots of both types of events concentrated in Tibet (Sun et al., 2014).

Observations of thunderstorms, HR, hail, and CG, which were used to label the predictors, were obtained from the severe weather observation dataset of NMC (Zheng et al., 2013). The thunderstorm observations consist of lightning location data collected by the National Lightning Location Network (NLLN) of China. It has been installed with ground-based advanced time of arrival and direction system of cloud-to-ground lightning detection sensors, reaching 394 in operation in 2016, covering most of China. According to relevant studies (Xia et al., 2015; Yang et al., 2015), the lightning location accuracy of the whole network is approximately 300 m, the detection rate is larger than 80%, and the average detected radius of a sensor is approximately 300 km. A thunderstorm is recorded if at least one lightning strike is observed by the NLLN. The HR data consist of observations of hourly rainfall no less than 20 mm. Rainfall is measured by automatic rain gauges at 2420 national-level weather stations (NWSs) and more than 20,000 automatic weather stations. The hail and CG observations are from observer reports, and are available 24 h a day at the 2420 NWSs in mainland China.

3. Deep learning method

The deep learning algorithm for SCW forecasting (Fig. 1) includes three major steps. First, the training and testing datasets are collected. Second, a deep learning network is constructed, trained, and tested. Third, the trained network is implemented for forecasting.

3.1 *Training/testing sets*

The weather forecasting can be regarded as a two-category classification problem, i.e., 0 indicates that the event will not happen, and 1 indicates that it will happen. To feed the deep learning network with sufficient spatial information of climate variables, our model input is set to be 144 observed climate variables over a square patch with dimension $L \times L$ centered at each SCW event grid. These $L \times L \times 144$ data arrays are labeled by either 1 or 0 depending on whether SCW occurs at its center grid or not. These labeled data arrays form either the training or testing samples. The choice of L serves as a balance of tradeoff between computational efficiency and model performance. Some preliminary experiments suggested that $L = 7$ is an optimal choice for this purpose.

Since the NWP system yields gridded fields and the SCW observations are site-based data, the observations need to be remapped to the NWP grids first. If an observed SCW event occurred within a radius, R, of the grid point, the grid point is marked by 1, indicating that the event occurred at this grid point. Otherwise, the grid point is marked by 0. Considering that the SCW often occurs on a meso-γ scale, R is set to be 20 km in this study. Note that if R is set too small, there will be too many missing forecasts; while if R is too large, there will be too many false alarms.

Compared to non-SCW events, the SCW is a high-impact and low-probability event. Therefore, positive samples (marked as "with SCW") are far fewer than negative samples (marked as "without SCW"). This reflects a typical sample set imbalance (Krawczyk, 2016). To remedy this issue, positive samples are replicated to balance the positive and negative samples in the training sets. This process is called over-sampling (Buda et al., 2018). Over-sampling is unnecessary for test sets though, as the test sets are mainly used to assess the performance of the trained models. We therefore constructed test sets without over-sampling to assess the performance of our trained models under the real positive–negative sample ratio.

Two independent datasets were constructed based on SCW observations and NCEP FNL analysis data for the period of March–October during 2010–14. One contains a sample of 50 days that were selected by randomly

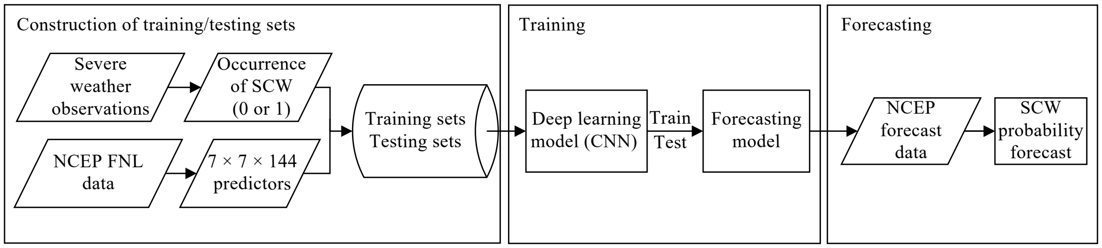

Construction of training/testing sets
Severe weather observations
Occurrence of SCW (0 or 1)
NCEP FNL data
7 × 7 × 144 predictors
Training sets Testing sets
Training
Deep learning model (CNN)
Train
Test
Forecasting model
Forecasting
NCEP forecast data
SCW probability forecast

Fig. 1. Flow chart of SCW forecasting with deep learning method.

choosing one day in each month from March to October of 2010–14, and is considered as the test set. The other constructed dataset contains all the remaining positive and negative samples (4,582,577 thunderstorm samples, 3,609,185 HR samples, 1,468,158 hail samples, and 1,488,531 CG samples) and is treated as the training set.

3.2 *CNN*

As mentioned above, the prediction of SCW can be regarded as a classification task with binary categories. A deep learning network for classification is therefore constructed for this purpose. Among various deep learning networks, CNN is a class of deep and feed-forward artificial neural networks, which has been successfully applied to many fields, especially image and video recognition (LeCun and Bengio, 1995; Krizhevsky et al., 2012). CNN algorithm can effectively extract two-dimensional (2D) features, reduce the number of model parameters, and accelerate the training speed by utilizing receptive fields and weights sharing (LeCun and Bengio, 1995). We constructed deep 2D CNN classification models for binary classification and trained them to predict thunderstorms, HR, hail, and CG.

Our CNN consists of convolution, fully connected layer, and the Softmax classifier. The input of a 2D CNN requires a data array with a format of height × width × channel (channel corresponds to predictors here). Due to the 144 predictors selected for each patch, our input is a three-dimensional (3D) array with dimensions 7 × 7 × 144. As mentioned above, these predictors represent the environmental conditions favorable for SCW events.

The core of our processing was carried out by a feed-forward stack of five convolutional layers (C1 to C5), followed by one fully connected layer that outputs class scores. Each channel of the five convolutional layers was obtained by convolving the channels of the previous layer with a bank of linear 2D filters such as summing, adding a bias term, and applying a pointwise nonlinearity, as follows:

$$\boldsymbol{X}_n^l = \mathrm{ReLU}\left(\boldsymbol{b}_n^{(l)} + \sum\nolimits_{k=1}^{K} \boldsymbol{W}_n^{(k,l)} * \boldsymbol{X}_{n-1}^{(k)}\right),\ l \in \{1,\cdots,5\}, \quad (1)$$

where ReLU(x) = max(0, x) is the rectified linear unit activation function. The symbol "*" denotes the two-dimensional convolution operation. The matrices $\boldsymbol{W}_n^{(k,l)}$ represent the filters of layer n, and $\boldsymbol{b}_n^{(l)}$ the bias for feature map l. Note that a feature map $\boldsymbol{X}_n^l$ is obtained by computing a sum of K convolutions of the feature maps from the previous layer.

The fifth layer was followed by a fully connected layer with 128 neurons, which transformed the 3D array ($\boldsymbol{X}^5$) into a one-dimensional (1D) array ($\bar{\boldsymbol{X}}^5$). Then $\bar{\boldsymbol{X}}^5$ was processed by a linear and fully connected layer to compute the class scores S_c with c = 0 or 1 as follows:

$$S_c = \sum\nolimits_{i=1}^{128} \bar{\boldsymbol{X}}_i^5 \cdot \boldsymbol{W}_{ci}^6 + b_c^6. \quad (2)$$

Finally, we applied the Softmax classifier function to class scores to obtain a properly normalized probability distribution (p), which could be interpreted as a posterior distribution of the two classes given the input X^0 and the network parameters $\boldsymbol{W}$ and $\boldsymbol{b}$:

$$p_c = \mathrm{P}\left(\text{class} = c|X^0, \boldsymbol{W}, \boldsymbol{b}\right) = \frac{\exp(S_c)}{\exp(X_0) + \exp(X_1)}, c = \{0, 1\}, \quad (3)$$

where $\boldsymbol{W}$ = {W^1, ⋯, W^6} is the set of weights, and $\boldsymbol{b}$ = {b^1, ⋯, b^6} is the set of biases.

The structure of the deep CNN is shown in Fig. 2. The input for the model was a 7 × 7 × 144 data array. After the data were imported, they were convolved by five convolution layers that have 256, 256, 512, 256, and 128 filters, respectively. The convolution kernel size was 2 × 2, and a valid mode was applied to every convolutional layer. Because the height and width of the input patch (7 × 7) were small, no pooling layer was utilized. The output was then passed through a fully connected layer with 128 neurons, and the 3D array was transformed into a 1D output array. Finally, the classified probability was calculated by the Softmax classifier function. The number of parameters used in this CNN model was 1,647,362.

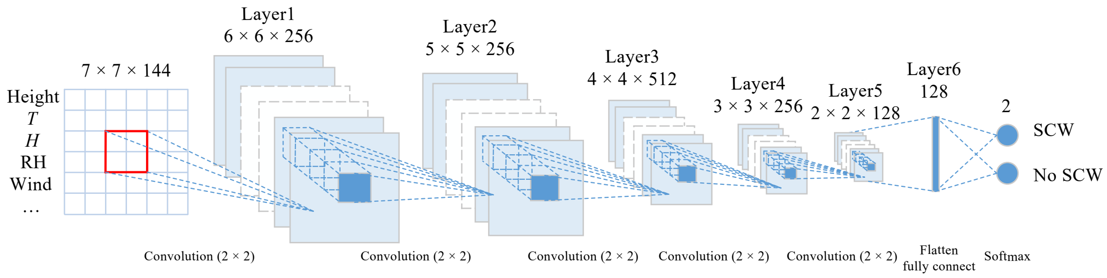


Fig. 2. Structure of the deep CNN algorithm for SCW forecasting, including five convolutional layers, a fully connected layer, and the Softmax classifier.

We optimized the network parameters by minimizing an L2-regularized cross-entropy loss function. For optimization, we used the ADAM algorithm (Kingma and Ba, 2015), an algorithm for first-order gradient-based optimization of stochastic objective functions, which kept track of the first- and second-order moments of the gradients and was invariant to any diagonal rescaling of the gradients. We set the learning rate at 10^{-4} and kept all other parameters to their default values recommended by Kingma (Perol et al., 2018).

3.3 *SCW forecasting*

After the optimal forecasting model was established, we now make forecast based on the 144 predictors obtained from GFS. To accommodate to the format of the input to the SCW forecasting model, the 144 predictors were normalized and transformed into an $M \times 7 \times 7 \times 144$ array (M is the number of forecast samples). Probabilistic forecasts were then produced again by the Softmax classifier. Here, probabilistic forecasts were preferred mainly because the meso-γ scale SCW is difficult to observe, especially for the hail and CG, and thus it is challenging to fully understand the phenomena (Cintineo et al., 2014).

Deep CNN training is computationally intensive. Compared to the usually small number of logical CPUs (central processing unit), the GPU (graphics processing unit) used in CNN training is a huge computational matrix with thousands of compute cores. GPUs are able to support parallel computing which is crucial for deep learning because it greatly accelerates the training process (Sanders and Kandrot, 2010). The NVIDIA CUDA (Compute Unified Device Architecture) library and NVIDIA GeForce 1080 Ti graphics chip were utilized in our training and forecasting processes. Tests showed that 0–72-h forecasts (at 6-h intervals) at 1° × 1° resolution over mainland China can be completed in 3 min, which makes the forecasts practical for operation.

4. Results

4.1 *Evaluation methods*

We chose four skill scores to measure the performance of the forecasts: threat score (TS), equitable threat score (ETS), probability of detection (POD), and false alarm rate (FAR), which are defined as follows:

$$\mathrm{POD} = \frac{h}{h+m}, \tag{4}$$

$$\mathrm{FAR} = \frac{f}{h+f}, \tag{5}$$

$$\mathrm{TS} = \frac{h}{h+m+f}, \tag{6}$$

$$\begin{aligned} \mathrm{ETS} &= \frac{h-h_{\mathrm{random}}}{h+m+f-h_{\mathrm{random}}}, \\ h_{\mathrm{random}} &= (h+f)\times(h+m)/(h+m+f+c), \end{aligned} \tag{7}$$

where h is the number of hits, m is the number of missing forecasts, f is the number of false forecasts, and c is the number of correct negatives.

Although the above scores are typically used for evaluating the deterministic forecasts, they can also be used for evaluating the probabilistic forecasts by thresholding the probabilistic forecasts and turning them into deterministic forecasts. Thus, we used those four scores to evaluate the deep CNN forecasting results. After exploring different thresholds for each prediction, we found that the most effective probabilistic threshold values for thunderstorms, HR, hail, and CG were 0.5, 0.5, 0.9, and 0.9, respectively.

4.2 *Evaluation of different algorithms*

In order to compare the performance of deep CNN to that of traditional algorithms, we report the classification performance of various algorithms in Table 2 using the HR test set (592 positive samples and 14,049 negative samples) as well as the hail test set (149 positive samples and 14,492 negative samples). The input of traditional ML algorithms is the 144 predictors at each individual SCW event grid. Note that the logistic regression (LR), RF, SVM, and multilayer perceptron (MP) algorithms are from the scikit-learn package (Pedregosa et al., 2011), and we used GridSearchCV function to conduct exhaustive search over specified parameter values for each classifier.

It can be seen from Table 2 that different algorithms have different classification skill. The performance of RF, SVM, and MP is similar in terms of ETS and TS, and is slightly better than that of LR.

Compared to traditional algorithms, the deep CNN model has a deeper network architecture and more model parameters. Owing to the increased complexity, deep CNN's training time is much longer than simpler algorithms. However, the added complexity can take into account the spatial features of SCW occurrence, leading to improved forecasting performance. As shown in Table 2, deep CNN achieved best performance among all the algorithms for HR and hail forecasting in terms of the four skill scores. Because the comparison results for thunderstorm and CG are quite similar to those for HR and hail respectively, we omitted the results for thunderstorm and CG in Table 2.

Table 2. Skill scores for different algorithms using an HR test set (592 positive samples and 14,049 negative samples) and a hail test set (149 positive samples and 14,492 negative samples). The input of CNN is 7 × 7 × 144 data arrays formed by 144 predictors over 7 × 7 patches centered at each SCW event grid, while the input of traditional ML algorithms is 144 predictors at each individual SCW grid

SCW	Algorithm	POD	FAR	ETS	TS
Heavy rain (HR)	LR	0.515	0.570	0.285	0.306
	RF	0.499	0.531	0.300	0.319
	SVM	0.509	0.543	0.297	0.317
	MP	0.526	0.562	0.294	0.314
	Deep CNN	0.536	0.504	0.328	0.347
Hail	LR	0.178	0.933	0.044	0.051
	RF	0.182	0.916	0.054	0.061
	SVM	0.185	0.922	0.051	0.058
	MP	0.192	0.925	0.050	0.057
	Deep CNN	0.213	0.892	0.070	0.077

In summary, the above results showed that deep CNN algorithm outperforms traditional machine learning algorithms in SCW forecasting over China.

4.3 *Case evaluation*

On 21 September 2017, thunderstorms, CG, and hail occurred over a large area in northern China. Meanwhile, a large area in southern China suffered from thunderstorms and HR. The SCW observations and forecasts for this case are shown in Fig. 3.

Figure 3 clearly shows that the deep CNN algorithm has a good forecasting skill for thunderstorms, hail, and CG in northern China. Most of the occurrences in the forecasting area were successfully identified based on our forecasts. Moreover, the deep CNN algorithm also appeared to be skillful on forecasting thunderstorms and HR in southern China.

The meteorologist's forecasts seemed quite different from the objective forecasts. We can see that with the meteorologist's forecasts there were a large number of false alarms for the thunderstorm, and lots of missing HR, hail, and CG events. The TSs of the deep CNN forecasts of thunderstorms, HR, hail, and CG are 0.48, 0.41, 0.13, and 0.46, respectively, while those of the meteorologist's forecasts are only 0.40, 0.25, 0.08, and 0.33, respectively.

In summary, for the above typical SCW case, deep CNN forecasts demonstrated much better performance than the forecasters' forecasts.

4.4 *False forecast case*

The false forecast of SCW that occurred in northeastern China on 2 August 2015 was selected to investigate the reason for false forecast. For this SCW, the forecasts indicated that HR would occur over a large area in the eastern Inner Mongolia, southwestern Heilongjiang, Jilin, and Liaoning provinces. However, observations showed that false alarms of HR were issued over most of the above regions except for eastern Inner Mongolia, central Heilongjiang, southern Liaoning, and some other areas. In order to identify the reason for the false forecasts, two important parameters for HR forecasting, i.e., PWAT and K-index, were selected to compare the forecasts and observations.

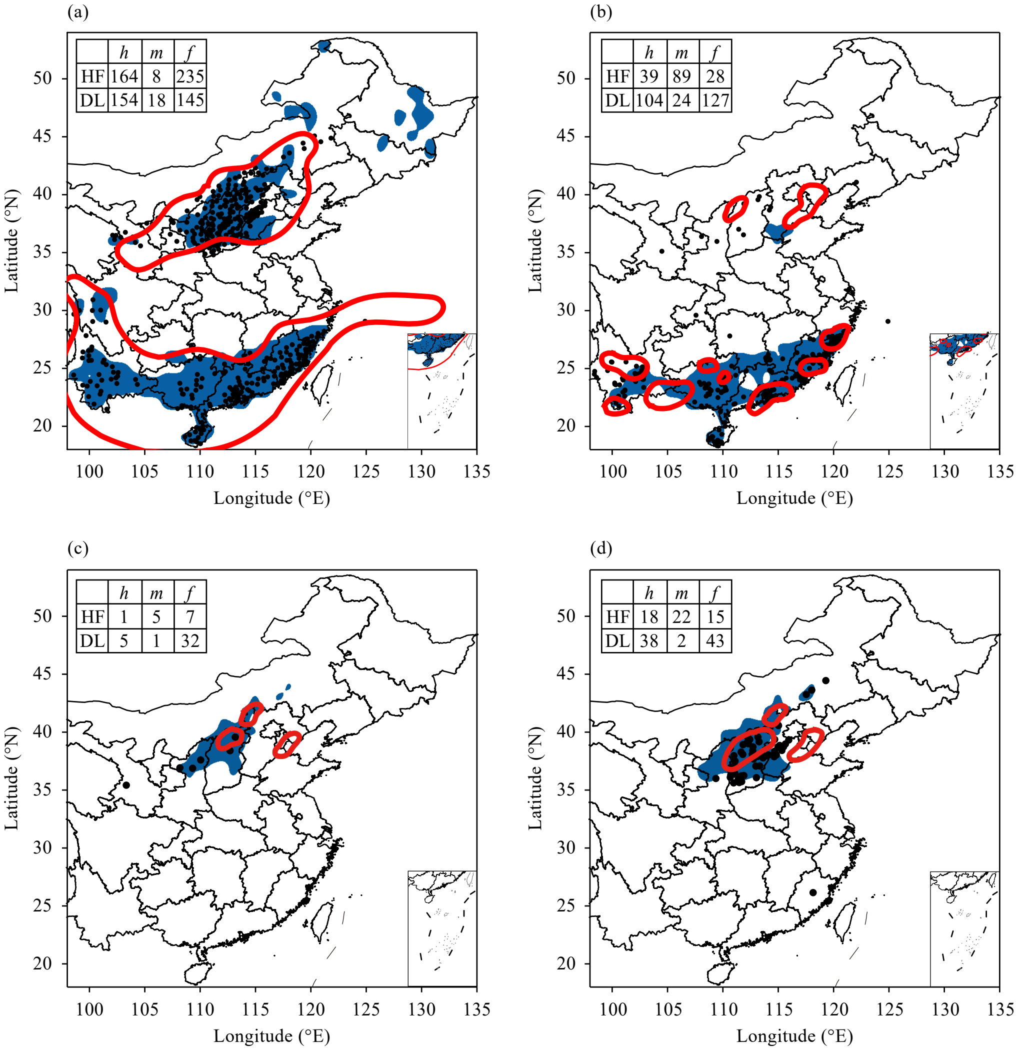
The GFS forecasts at 1400 Beijing Time (BJT) on 2 August 2015 (issued at 2000 BJT 1 August 2015) indicated that PWAT in Heilongjiang, Jilin, and Liaoning provinces would exceed 50 mm and the K-index would exceed 40°C, suggesting good environmental conditions for HR events. Meteorologists also predicted the occurrence of HR events in the same area and during the same time period. Based on the GFS forecasts, an HR warning was issued over the above regions.

However, a comparison between the NWP forecast fields and the analysis fields indicates big differences between the NWP predictions and observations. Figure 4a shows that the observed PWAT values were significantly lower than the forecasted values, and the predicted values of PWAT in central Jilin and Liaoning were about 8 mm higher than the observed values. In addition, the predicted K-indices were also 3–7°C higher than the observations as seen in Fig. 4b.

The above results indicate that the water vapor condition and the distribution of K-index in the NWP forecasts favored the occurrence of HR events. For this reason, a false alarm of SCW was issued. Since the deep CNN algorithm took the incorrect NWP forecasts as its input, it therefore also predicted that HR events would occur over a large area in Northeast China.

4.5 *Missed forecast case*

During 0200–1400 BJT on 9 July 2015, the deep CNN algorithm predicted that HR events would occur only in central Sichuan Province and its surrounding areas, while observations indicate that HR events actually occurred over a much larger area in eastern Sichuan. As a result, this was a case of missed forecasts.



(a)

	h	*m*	*f*
HF	164	8	235
DL	154	18	145

(b)

	h	*m*	*f*
HF	39	89	28
DL	104	24	127

(c)

	h	*m*	*f*
HF	1	5	7
DL	5	1	32

(d)

	h	*m*	*f*
HF	18	22	15
DL	38	2	43

Fig. 3. Forecasts and observations of (a) thunderstorm, (b) HR, (c) hail, and (d) CG on 21 September 2017. Blue shades are objective forecasts, black points are observations of SCW events, and red lines are meteorologist's subjective forecasts. Only forecasts on land were evaluated. The table in each image lists the number of hits, misses, and false alarms made by either human forecasters (HF) or deep learning (DL) algorithm.

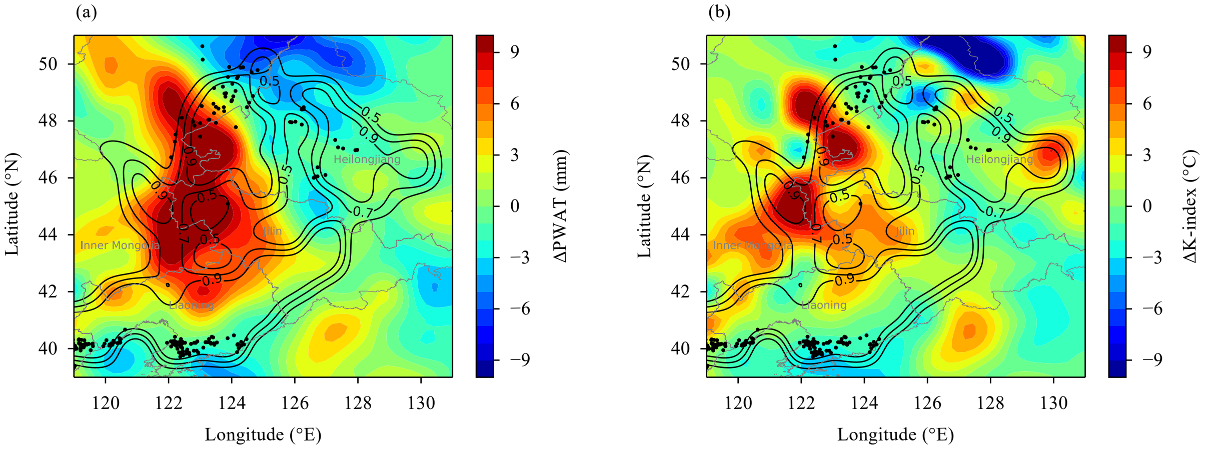


Fig. 4. Comparison between the NWP forecasts and FNL analyses from NCEP at 1400 BJT 2 August 2015 in terms of (a) PWAT and (b) K-index. Contours are HR probabilistic forecasts, black dots indicate HR observations, and shaded areas indicate differences between GFS forecasts and FNL analyses, i.e., GFS forecasts minus FNL analyses.

In order to analyze the causes for the missed forecasts, Fig. 5 shows a comparison between the GFS forecast fields and observations. At 0800 BJT 9 July 2015, one of the SCW conditions in Sichuan was reflected in PWAT, which was about 50 mm, implying an adequate water vapor condition. Meanwhile, the K-indices were 30–35°C, indicating an instability in the atmosphere; the best lift index (BLI) was within −1 to 1, and the energy condition was relatively weak; and the vertical velocity at 500 hPa was from −20 to 10 × 10^{-2} Pa s^{-1}, suggesting a weak dynamic lifting. Overall, the GFS forecasted fields suggested that the SCW environmental conditions were only favorable for the development of weak convection.

In general, there were obvious differences between the FNL analysis fields and the forecast fields. According to the FNL analysis data, at 0800 BJT 9 July 2015, the PWAT values exceeded the forecasted values by 2–8 mm while the K-indices exceeded the forecasted values by 1–6°C. Although the forecast missed the SCW, surprisingly other features of the forecasts demonstrated a more favorable condition for the development of SCW. For example, BLI values were less than −2, the convective available potential energy (CAPE) values were more than 600 J kg^{-1}, and the well-developed dynamic uplifting was found at 500 hPa and even higher levels. Significant differences between the forecasts and analyses are displayed in Fig. 5, which shows that the analysis values of PWAT and K-index are much larger than the forecasted values in eastern Sichuan where the forecast missed the HR event.

In summary, the GFS forecast fields suggested that the environmental conditions were only favorable for the development of weak convection. In contrast, the analyses indicated that the true environmental conditions were actually very favorable for the development of SCW. The meteorologists also failed to predict the HR events in the eastern part of Sichuan. We found the reason that the deep CNN algorithm missed forecasting the SCW event was because the NWP forecasts as input were only favorable for weak convection.

4.6 *Overall evaluation*

In order to overall evaluate the performance of the deep CNN algorithm, the forecasts of SCW from April to September in 2015, 2016, and 2017 by both the CNN algorithm and meteorologists were reported (Fig. 6).

Table 3 shows that the deep CNN algorithm significantly improved the forecasts of all kinds of SCW when compared to forecaster's forecasts.

The average TS of HR forecasts by deep CNN algorithm was 0.336, which showed an increase of 33.2% over the TS of 0.252 for subjective forecasts. It is worth mentioning that the miss rate (1 − POD) of forecaster's forecasts was larger than the FAR, which means that a large number of heavy rainfall events were missed. In contrast, the opposite was true for forecasts by the deep CNN algorithm; that is, fewer missing forecasts were made by CNN. Thus, the deep CNN algorithm provides valuable guidance for meteorologists in their operational forecasts.

The deep CNN algorithm has stable performance in thunderstorm forecasting. The average TS of forecasts by the algorithm exceeded 0.44 each year, which is on average 16.1% higher than the score of meteorologist's forecasts. The deep CNN algorithm had a higher POD and lower FAR, implying more reliable forecasting.

The performance diagrams are given in Fig. 6. Forecasts of hail and CG were greatly improved by the deep

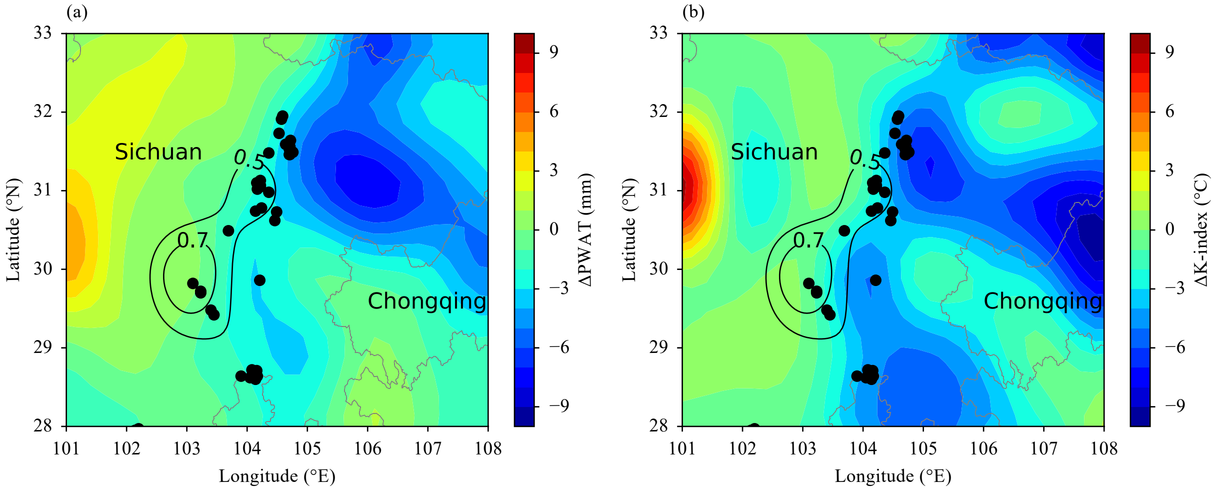


Fig. 5. As in Fig. 4, but over eastern Sichuan Province, China at 0800 BJT 9 July 2015.

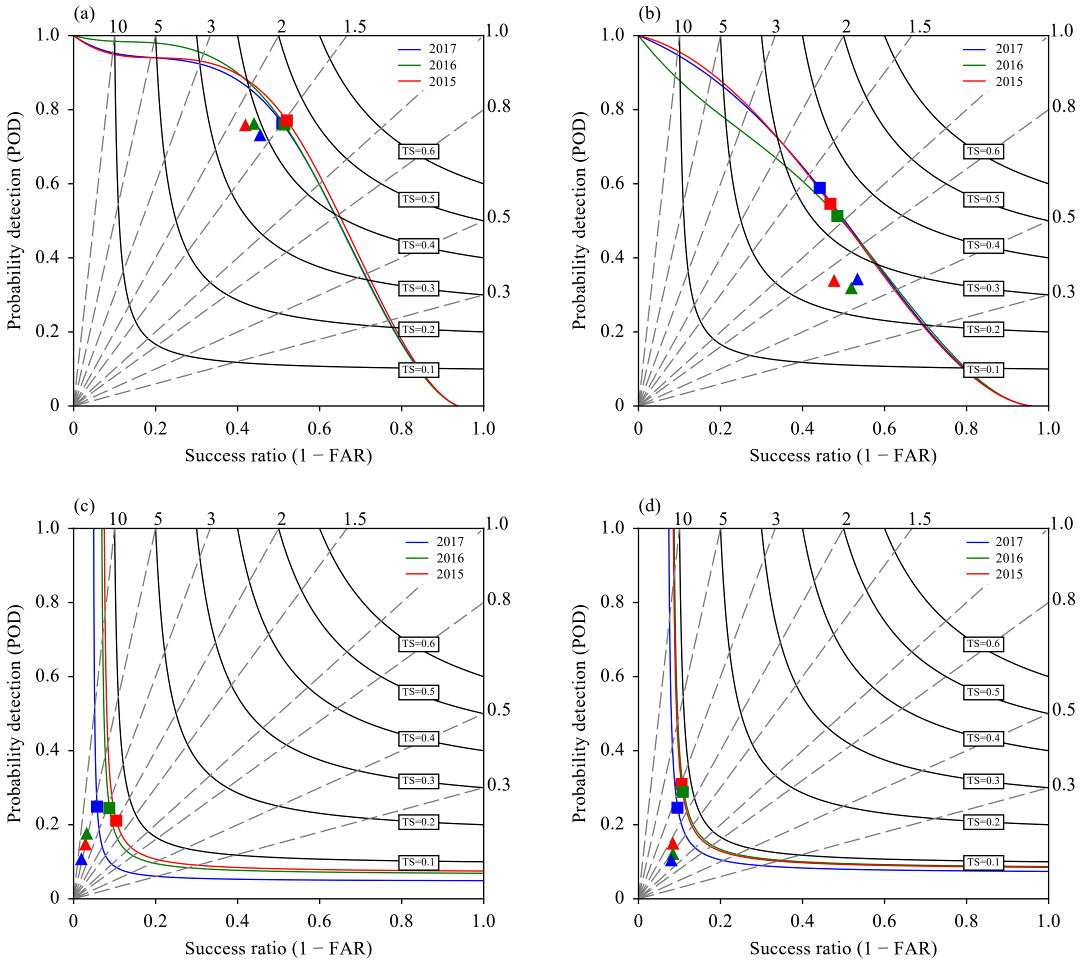


Fig. 6. Performance diagram of forecasts by the deep CNN algorithm and meteorologists from April to September of 2015, 2016, and 2017 for (a) thunderstorm, (b) HR, (c) hail, and (d) CG. Dashed lines represent bias scores with labels on the outward extension of the line, while labeled solid contours are TS. The red, green, and blue lines indicate the performance of CNN in 2015, 2016, and 2017, respectively. The red, green, and blue triangles indicate the forecast performance by meteorologists in 2015, 2016, and 2017, respectively. The squares indicate the performance of CNN with the determined probability thresholds, which were 0.5, 0.5, 0.9, and 0.9 for thunderstorms, HR, hail, and CG, respectively.

Table 3. Evaluation of deep CNN (DL) forecasts and human forecasts (HF) from April to September of 2015, 2016, and 2017 (for 12-h forecasts initialized at 0800 BJT)

SCW	Year	POD(DL)	POD(HF)	FAR(DL)	FAR(HF)	TS(DL)	TS(HF)	ETS(DL)	ETS(HF)
HR	2015	0.546	0.338	0.532	0.523	0.337	0.247	0.292	0.211
	2016	0.513	0.318	0.515	0.481	0.332	0.246	0.289	0.212
	2017	0.589	0.342	0.558	0.466	0.338	0.264	0.290	0.229
Thunderstorm	2015	0.771	0.758	0.480	0.581	0.451	0.370	0.372	0.277
	2016	0.760	0.762	0.486	0.560	0.442	0.387	0.363	0.297
	2017	0.763	0.731	0.491	0.545	0.440	0.390	0.360	0.303
Hail	2015	0.211	0.147	0.896	0.971	0.075	0.025	0.071	0.021
	2016	0.245	0.176	0.913	0.968	0.069	0.028	0.065	0.023
	2017	0.249	0.107	0.943	0.981	0.049	0.016	0.044	0.012
CG	2015	0.310	0.149	0.895	0.916	0.085	0.057	0.074	0.047
	2016	0.289	0.121	0.892	0.916	0.085	0.052	0.075	0.044
	2017	0.246	0.104	0.905	0.920	0.074	0.047	0.063	0.039

CNN algorithm in comparison to meteorologist’s forecasts. The average TS of hail and CG forecasts by the algorithm were 0.064 and 0.081, respectively. This indicates 178% and 55.7% improvements to the meteorologist’s forecast scores of 0.023 and 0.052. The deep CNN algorithm also had better performance in terms of POD and FAR. Because hail and CG are incompletely observed due to their local distribution, the forecasts from both the algorithm and meteorologists had high FAR value.

In summary, deep CNN algorithm showed higher capability in all four types of SCW forecasting. Compared to the forecaster's forecasts, the deep CNN forecasts showed notable improvements in both qualitative and quantitative evaluations, suggesting that the algorithm has much better overall performance. Nevertheless, there are still inadequacies in the algorithm as it issues too many false alarms of hail and CG, a problem also found in meteorologist's forecasts. Thus, a key goal of improving the algorithm in the future will be reducing the FAR of forecasts.

It is worth mentioning that we also pinned down to the performance of deep learning at hot-spots, and we found that the deep learning has better forecasting capability at hot-spots than in other areas. For example, the TS of HR is 0.30 in North China in 2017, while 0.39 in South China where hot-spots are observed. Similar pattern is also observed for other types of SCW events.

5. Conclusions and discussion

Based on the NCEP global FNL analysis and GFS forecast data and the SCW observations, we constructed a deep CNN algorithm to forecast thunderstorms, HR, hail, and CG. The performance of the deep CNN forecasts was evaluated for both three SCW cases and long-run SCW forecasts from April to September of 2015, 2016, and 2017. The major conclusions are summarized as follows.

(1) Compared with traditional machine learning algorithms, the deep CNN algorithm can automatically extract nonlinear features of SCW, and yield a better forecast performance.

(2) Compared with meteorologist's forecasts, the deep CNN forecasts of SCW show significant improvements. The TS of thunderstorm, HR, hail, and CG forecasts of the deep CNN algorithm were 16.1%, 33.2%, 178%, and 55.7% better than their respective scores from traditional methods.

(3) Incorrect GFS forecasts led to false or missed SCW forecasts by the deep CNN algorithm. Correcting NWP forecast errors or inadequacies is necessary prior to applying postprocessing techniques such as deep CNN algorithm since such techniques rely on reasonable input data.

The deep CNN algorithm not only can automatically extract physical characteristics of SCW from the massive historical data, but also can consider terrain features in different areas. Therefore, the deep CNN takes into account more comprehensive environmental conditions of SCW such as the dynamical processes, water vapor contents, instability, etc. Our results have demonstrated overall better forecasting skills of deep CNN than other forecasting methods.

Despite the already good performance of deep CNN algorithm demonstrated here, there are still many ways to improve the results. For example, a hyperparameter optimization, deeper network architectures, and deep learning model ensembles would likely lead to improved performance of the deep CNN model. Furthermore, only GFS data were used as the data source for training and forecasting in this study. However, to extract more optimal forecast results, we may design a new deep CNN model with input data from several NWP models, such as from GFS, ECMWF, and Global/Regional Assimilation and Prediction Enhanced System (GRAPES).

High FAR and low POD values for hail and CG forecasts are observed in Table 3. This could be caused by two main reasons. One is the small sample size of hail and CG observations due to their relatively low occurrence frequency compared to thunderstorms and HR as well as their incomplete record for omitting the events with small spatial sizes. The other possible reason is the limited capability of the current global NWP models as indicators of hail and CG events. Many studies (Gagne II et al., 2015; Sheridan, 2018) have shown that high-resolution NWP models possess greater capability of forecasting these events, and thus using high-resolution NWP models in the future is expected to further improve the forecasts.

In summary, the deep CNN algorithm can effectively extract the characteristics of SCW and has demonstrated great forecasting skills. Results from the algorithm can provide useful guidance for meteorologists in their operational weather forecasting.

Acknowledgments. The authors thank the editor, one anonymous reviewer, and Dr. David J. Gagne at NCAR for their constructive comments that have greatly improved the content and presentation of this article. One of the authors, Bo Li, acknowledges partial support from the US NSF grants AGS-1602845 and DMS-1830312.

REFERENCES

Beijbom, O., P. J. Edmunds, D. I. Kline, et al., 2012: Automated annotation of coral reef survey images. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Providence, RI, USA, 1170–1177, doi: 10.1109/CVPR.2012.6247798.

Bengio, Y., 2009: *Learning Deep Architectures for AI*. Foundations and Trends® in Machine Learning, Vol. 2, No. 1, 1–127, now Publishers Inc., Hanover, MA, USA, doi: 10.1561/2200000006.

Buda, M., A. Maki, and M. A. Mazurowski, 2018: A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.*, **106**, 249–259, doi: 10.1016/j.neunet.2018.07.011.

Chaudhuri, S., 2010: Convective energies in forecasting severe thunderstorms with one hidden layer neural net and variable learning rate back propagation algorithm. *Asia-Pacific J. Atmos. Sci.*, **46**, 173–183, doi: 10.1007/s13143-010-0016-1.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, et al., 2014: An empirical model for assessing the severe weather potential of developing convection. *Wea. Forecasting*, **29**, 639–653, doi: 10.1175/WAF-D-13-00113.1.

Ciregan, D., U. Meier, and J. Schmidhuber, 2012: Multi-column deep neural networks for image classification. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Providence, RI, USA, 3642–3649, doi: 10.1109/CVPR.2012.6248110.

Doswell III, C. A., 2001: *Severe Convective Storms*. American Meteorological Society, Boston, MA, USA, 561 pp, doi: 10.1007/978-1-935704-06-5.

Doswell III, C. A., H. E. Brooks, and R. A. Maddox, 1996: Flash flood forecasting: An ingredients-based methodology. *Wea. Forecasting*, **11**, 560–581, doi: 10.1175/1520-0434(1996)011<0560:FFFAIB>2.0.CO;2.

Gagne II, D. J., A. McGovern, and M. Xue, 2014: Machine learning enhancement of storm-scale ensemble probabilistic quantitative precipitation forecasts. *Wea. Forecasting*, **29**, 1024–1043, doi: 10.1175/WAF-D-13-00108.1.

Gagne II, D. J., A. McGovern, J. Brotzge, et al., 2015: Day-ahead hail prediction integrating machine learning with storm-scale numerical weather models. Proceedings of the 27th Conference on Innovative Applications of Artificial Intelligence, AAAI, Austin, TX, USA, 3954–3960.

Gagne II, D. J., A. McGovern, S. E. Haupt, et al., 2017: Storm-based probabilistic hail forecasting with machine learning applied to convection-allowing ensembles. *Wea. Forecasting*, **32**, 1819–1840, doi: 10.1175/WAF-D-17-0010.1.

Gardner, M. W., and S. R. Dorling, 1998: Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmos. Environ.*, **32**, 2627–2636, doi: 10.1016/S1352-2310(97)00447-0.

Gope, S., S. Sarkar, P. Mitra, et al., 2016: Early prediction of extreme rainfall events: A deep learning approach. Proceedings of the 16th Industrial Conference on Data Mining, Springer, New York, NY, USA, 154–167, doi: 10.1007/978-3-319-41561-1_12.

Grzymala-Busse, J. W., L. K. Goodwin, W. J. Grzymala-Busse, et al., 2004: An approach to imbalanced data sets based on changing rule strength. *Rough-Neural Computing: Techniques for Computing with Words*, S. K. Pal, L. Polkowski, and A. Skowron, Eds., Springer, Berlin Heidelberg, 543–553, doi: 10.1007/978-3-642-18859-6_21.

Han, L., J. Z. Sun, W. Zhang, et al., 2017: A machine learning nowcasting method based on real-time reanalysis data. *J. Geophys. Res. Atmos.*, **122**, 4038–4051, doi: 10.1002/2016JD025783.

Herman, G. R., and R. S. Schumacher, 2018: Money doesn't grow on trees, but forecasts do: Forecasting extreme precipitation with random forests. *Mon. Wea. Rev.*, **146**, 1571–1600, doi: 10.1175/MWR-D-17-0250.1.

Kingma, D. P., and J. Ba, 2015: Adam: A method for stochastic optimization. Proceedings of the 3rd International Conference on Learning Representations, IEEE, San Diego, USA, 3156–3165.

Klein, B., L. Wolf, and Y. Afek, 2015: A dynamic convolutional layer for short range weather prediction. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Boston, MA, USA, 4840–4848, doi: 10.1109/CVPR.2015.7299117.

Klein, W. H., B. M. Lewis, and I. Enger, 1959: Objective prediction of five-day mean temperatures during winter. *J. Meteor.*, **16**, 672–682, doi: 10.1175/1520-0469(1959)016<0672:OPOFDM>2.0.CO;2.

Krawczyk, B., 2016: Learning from imbalanced data: Open challenges and future directions. *Prog. Artif. Intell.*, **5**, 221–232, doi: 10.1007/s13748-016-0094-0.

Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012: ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems, ACM, Lake Tahoe, USA, 1097–1105.

Kubat, M., R. C. Holte, and S. Matwin, 1998: Machine learning for the detection of oil spills in satellite radar images. *Mach. Learn.*, **30**, 195–215, doi: 10.1023/A:1007452223027.

Lagerquist, R., A. McGovern, and T. Smith, 2017: Machine learning for real-time prediction of damaging straight-line convective wind. *Wea. Forecasting*, **32**, 2175–2193, doi: 10.1175/WAF-D-17-0038.1.

Lakshmanan, V., G. Stumpf, and A. Witt, 2005: A neural network for detecting and diagnosing tornadic circulations using the mesocyclone detection and near storm environment algorithms. Proceedings of AI Applications with a Nowcasting Flavor (Joint between the Fourth Conference on Artificial Intelligence and the 21st International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology), Amer. Meteor. Soc., San Diego, CA, USA, J5.2.

LeCun, Y., and Y. Bengio, 1995: Convolutional networks for images, speech, and time-series. *The Handbook of Brain Theory and Neural Networks*, M. A. Arbib, Ed., MIT Press, Cambridge, MA, USA, 10 pp.

Mac Namee, B., P. Cunningham, S. Byrne, et al., 2002: The problem of bias in training data in regression problems in medical decision support. *Artif. Intell. Med.*, **24**, 51–70, doi: 10.1016/S0933-3657(01)00092-6.

Manzato, A., 2005: The use of sounding-derived indices for a neural network short-term thunderstorm forecast. *Wea. Forecasting*, **20**, 896–917, doi: 10.1175/WAF898.1.

Manzato, A., 2007: Sounding-derived indices for neural network based short-term thunderstorm and rainfall forecasts. *Atmos. Res.*, **83**, 349–365, doi: 10.1016/j.atmosres.2005.10.021.

Manzato, A., 2013: Hail in northeast Italy: A neural network ensemble forecast using sounding-derived indices. *Wea. Forecasting*, **28**, 3–28, doi: 10.1175/WAF-D-12-00034.1.

Marzban, C., and G. J. Stumpf, 1996: A neural network for tornado prediction based on Doppler radar-derived attributes. *J. Appl. Meteor.*, **35**, 617–626, doi: 10.1175/1520-0450(1996)035<0617:ANNFTP>2.0.CO;2.

Marzban, C., and G. J. Stumpf, 1998: A neural network for damaging wind prediction. *Wea. Forecasting*, **13**, 151–163, doi: 10.1175/1520-0434(1998)013<0151:ANNFDW>2.0.CO;2.

Matsugu, M., K. Mori, Y. Mitari, et al., 2003: Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Netw.*, **16**, 555–559, doi: 10.1016/S0893-6080(03)00115-1.

Meng, Z. Y., D. C. Yan, and Y. J. Zhang, 2013: General features of squall lines in east China. *Mon. Wea. Rev.*, **141**, 1629–1647, doi: 10.1175/MWR-D-12-00208.1.

Pedregosa, F., G. Varoquaux, A. Gramfort, et al., 2011: Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.

Perol, T., M. Gharbi, and M. Denolle, 2018: Convolutional neural network for earthquake detection and location. *Sci. Adv.*, **4**, e1700578, doi: 10.1126/sciadv.1700578.

Ray, P. S., 1986: *Mesoscale Meteorology and Forecasting*. American Meteorological Society, Boston, USA, 793 pp, doi: 10.1007/978-1-935704-20-1_2.

Sanders, J., and E. Kandrot, 2010: *CUDA by Example: An Introduction to General-Purpose GPU Programming*. Addison-Wesley Educational Publishers Inc., Upper Saddle River, NJ, USA, 312 pp.

Schmidhuber, J., 2015: Deep learning in neural networks: An overview. *Neural Netw.*, **61**, 85–117, doi: 10.1016/j.neunet.2014.09.003.

Sheridan, P., 2018: Current gust forecasting techniques, developments and challenges. *Adv. Sci. Res.*, **15**, 159–172, doi: 10.5194/asr-15-159-2018.

Shi, X. J., Z. R. Chen, H. Wang, et al., 2015: Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Proceedings of the 28th International Conference on Neural Information Processing Systems, ACM, Montreal, Canada, 802–810.

Simonyan, K., and A. Zisserman, 2015: Very deep convolutional networks for large-scale image recognition. Proceedings of the 3rd International Conference on Learning Representations, IEEE, San Diego, USA, 313–318.

Stensrud, D. J., M. Xue, L. J. Wicker, et al., 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90**, 1487–1500, doi: 10.1175/2009BAMS2795.1.

Sun, J. S., J. H. Dai, L. F. He, et al., 2014: *The Basic Principle and Technical Method of Strong Convective Weather Forecast*. China Meteorological Press, Beijing, 1–21. (in Chinese)

Tian, F. Y., Y. G. Zheng, T. Zhang, et al., 2015: Statistical characteristics of environmental parameters for warm season short-duration heavy rainfall over central and eastern China. *J. Meteor. Res.*, **29**, 370–384, doi: 10.1007/s13351-014-4119-y.

Wang, Y. B., M. S. Long, J. M. Wang, et al., 2017: PredRNN : Recurrent neural networks for predictive learning using spatiotemporal LSTMs. Proceedings of the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, 879–888.

Xia, R. D., D.-L. Zhang, and B. L. Wang, 2015: A 6-yr cloud-to-ground lightning climatology and its relationship to rainfall over central and eastern China. *J. Appl. Meteor. Climatol.*, **54**, 2443–2460, doi: 10.1175/JAMC-D-15-0029.1.

Yang, X. L., J. H. Sun, and W. L. Li, 2015: An analysis of cloud-to-ground lightning in China during 2010–13. *Wea. Forecasting*, **30**, 1537–1550, doi: 10.1175/WAF-D-14-00132.1.

Yang, X. L., J. H. Sun, and Y. G. Zheng, 2017: A 5-yr climatology of severe convective wind events over China. *Wea. Forecasting*, **32**, 1289–1299, doi: 10.1175/WAF-D-16-0101.1.

Yu, X. D., 2011: Ingredients based forecasting methodology. *Meteor. Mon.*, **37**, 913–918. (in Chinese)

Zhang, W., L. Han, J. H. Sun, et al., 2017: Application of multi-channel 3D-cube successive convolution network for convective storm nowcasting. IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Honolulu, USA, 118–127.

Zhang, X. L., S. Y. Tao, and J. H. Sun, 2010: Ingredients-based heavy rainfall forecasting. *Chinese J. Atmos. Sci.*, **34**, 754–766, doi: 10.3878/j.issn.1006-9895.2010.04.08. (in Chinese)

Zheng, Y. G., Y. J. Lin, W. J. Zhu, et al., 2013: Operational system of severe convective weather comprehensive monitoring. *Meteor. Mon.*, **39**, 234–240. (in Chinese)

Tech & Copy Editor: Qi WANG