# A hybrid CNN-LSTM model for typhoon formation forecasting

**Rui Chen[1] · Xiang Wang[1] · Weimin Zhang[2] · Xiaoyu Zhu[3] · Aiping Li[3] · Chao Yang[3]**

## Abstract

A typhoon is an extreme weather event that can cause huge loss of life and economic damage in coastal areas and beyond. As a consequence, the search for more accurate predictive models of typhoon formation; and, intensity have become imperative as meteorologists, governments, and other agencies seek to mitigate the impact of these catastrophic events. While work in this field has progressed diligently, this paper argues, that the existing models are deficient. Traditional numerical forecast models based on fluid mechanics have difficulty in predicting the intensity of typhoons. Forecasts based on statistics and machine learning fail to take into account the spatial and temporal relationships among typhoon formation variables leading to weaknesses in the predictive power of this model. Therefore, we propose a hybrid model, which we argue, can produce a more realist and accurate account of typhoon 'behavior' as it focuses on both the spatio-temporal correlations of atmospheric and oceanographic variables. Our CNN-LSTM model introduces 3D convolutional neural networks (3DCNN) and 2D convolutional neural networks (2DCNN) as a method to better understand the spatial relationships of the features of typhoon formation. We also use LSTM to examine the temporal sequence of relations in typhoon progression. Extensive experiments based on three datasets show that our hybrid CNN-LSTM model is superior to existing methods, including numerical forecast models used by many official organizations; and, statistical forecast and machine learning based methods.

✉ Xiang Wang
  xiangwangcn@nudt.edu.cn

  Rui Chen
  chenrui16@nudt.edu.cn

  Weimin Zhang
  wmzhang104@139.com

[1]  College of Meteorology and Oceanology, College of Computer, National University of Defense Technology, Changsha, China

[2]  College of Meteorology, National University of Defense Technology, Laboratory of Software Engineering for Complex Systems, Changsha, China

[3]  College of Computer, National University of Defense Technology, Changsha, China

# 1 Introduction

Tropical cyclones(TC) are synoptic-scale storms with a closed circulation around a center of low pressure and a warmed core, driven by the heat energy of the underlying warm ocean. TCs, in the Western Pacic Ocean, that reach the threshold speed at the center of the circulation of 64kt (1kt=0.514 m/s) are categorised as typhoons, while those originating in the North Atlantic and Eastern Pacic Oceans are called hurricanes. In this paper, all 'typhoons' or 'hurricanes' are designated as 'typhoons'. A Typhoon is an extreme weather event that not only affects maritime activities, but also can cause heavy loss of life and costs to the urban economy in coastal areas and beyond. Therefore, research into the origins of typhoons is a major concern for countries in risk areas.

Given the gravity of typhoon destruction, the onus is on meteorological and other interested agencies through the research and development of improved models of forecasting typhoon formation to ameliorate their impact through timely advanced notice. However, the complexity of typhoon phenomena and our current dearth of knowledge in the discipline make a predicative model of typhoon formation problematical. At the moment, the discipline is dominated by three forecast models. The first, is a numerical forecast model based on a set of equations using fluid mechanics, for example, the regional model - the Hurricane Weather and Research Forecasting Model (HWRF); the global model - such as the European center for Medium-Range Weather Forecasts global model (EMX); and, - the ensemble model - Florida State Super Ensemble (FSSE). The second, the statistical forecast model uses traditional statistical methods, for instance, the Climatology and Persistence model (CLIPER5); and, the statistical dynamical model - Statistical Hurricane Intensity Prediction Scheme (SHIPS)". These models are the basis of many typhoon early warning notices for such agencies as the Joint Typhoon Warning center (JTWC), National Hurricane center(NHC), and China Meteorological Administration (CMA). The third, is based on machine learning methods, such as the Logistic Regression (LR) and Linear Discriminant Analysis (LDA) [1–4]. Unfortunately, current methods based on machine learning, have weaknesses as they do not include data related the spatio-temporal relationships among variables of typhoon which are crucial for understanding the formation of typhoon and development.

In this paper, we propose a hybrid deep learning model which we argue is superior to existing models as a predictive model of typhoon formation and intensity with its emphasis on the various spatial and temporal features of typhoons. This approach is inspired by the outstanding performance of the Convolutional Neural Network (CNN) in capturing the spatial relationships of images and the Long Short-Term Memory (LSTM) in time-series prediction. Combining these novel and innovative methodologies together will create a more effective predictive model to forecast typhoon formation and intensity as it will create a greater synergy than the individual models on their own.

Some research has already been accomplished using these approaches in areas related to our research, for example, Xingjian Shi et al., proposed the Convolutional Long Short-Term Memory (ConvLSTM) model for precipitation forecast [5]. Yuting Yang et al. constructed the CFCC-LSTM model to predict sea surface temperature (SST) based on historical SST data. The CFCC-LSTM model uses both the Convolution layer and FC-LSTM [6]. Donahue et al. designed the Longterm Recurrent Convolutional Network (LRCN) to learn spatio-temporal characteristics for visual recognition and description [7]. Liang Zhang et al. combined the 3DCNN, ConvLSTM and 2DCNN to establish a deep learning model for gesture recognition [8]. Whilst our hybrid deep model uses and combines 3DCNN and, 2DCNN and LSTM to extract spatial incorporate spatio-temporal correlations, but also the framework developed here is different from existing models and takes full

advantage of the temporal relationships learnt by LSTM among the variables of typhoon activity [9–12].

Our hybrid CNN-LSTM model can be used to predict more accurately than existing methods whether a TC will form into a typhoon and its intensity. 3DCNN is used to determine the spatial relations of various atmospheric variables in 3-dimensional space (The dimensions are latitude, longitude and atmospheric pressure at differing altitudes). Simultaneously, 2DCNN is used to determine the spatial features of sea surface variables (this time without pressure levels at altitude). LSTM is used to examine the spatio-temporal relationships of the points in the path of a 'prospective' typhoon to reveal its status and intensity.

The contributions of this paper include:

1. An outline of the hybrid CNN-LSTM model for forecasting typhoon formation and intensity which uses data from spatio-temporal correlations of atmospheric and sea surface variables. To the best of our knowledge, this is the first time in the field a spatio-temporal approach will be applied to typhoon forecasting.
2. The model could also be eventually used as a general tool for teaching a unified experimental protocol for future researchers interested in spatio-temporal problems related to meteorological and oceanographic phenomena.
3. Extensive experiments on three datasets including the Western Pacific (WP), Eastern Pacific (EP), and North Atlantic (NA). These show our hybrid CNN-LSTM model to be more effective than existing methods, including the official numerical forecast models used by many official organizations, statistical forecast methods and classical machine learning methods.

The paper is structured in the following manner: Section 2 takes a look at a cross-section of the current state of forecasting methods and deep learning in the discipline; Section 3, elucidates the central problems of the paper; Section 4, explains the detailed methodological and operational principles underlying the CNN-LSTM model; Section 5, examines the validating metrics to analyse the effectiveness of the data for the model's hypotheses; and, finally, Section 6, deals with conclusions and recommendations for future work.

## 2 Related work

Our research focuses specifically on forecasting typhoon formation and intensity as distinguished from other research which analyses the genesis of TC's. The two traditional models of TC genesis and intensity are numerical models based on fluid mechanics and statistical models [13–16]. The numerical models are divided into global and regional models and so on. These models are themselves differentiated by initial fields and differing parameterization schemes which result in obvious and distinctive outcomes in performance and characteristics [17]. The statistical model is calibrated to deliver long and short term forecasts of TC genesis using statistical relationships based on TC genesis related variables. This data is used to create to forecast a seasonal frequency of tropical cyclone genesis by constructing a typhoon potential Generation Index (GPI) [18–20]. The short-term forecasting method aims to predict whether these 'embryonic' tropical cloud clusters or disturbances will develop into tropical cyclones by analyzing the environmental variables surrounding the 'embryo'. For example, Hennon et al. and Fu B et al. used Linear Discriminant Analysis (LDA) to classify the tropical clouds in the North Atlantic and the Northwest Pacific [2, 3].

Chand S et al. utilized Bayesian probability regression models to predict tropical cyclones genesis in the Fiji region [21]. However, TCs are complex and nonlinear phenomena and using the statistical model to predict genesis from this data is fraught with problems.

The rapid development of machine learning has opened another avenue to meteorologists in their efforts to improve the prediction of typhoon genesis and intensity [22–25]. Wijnands et al., established a set of screening mechanisms for tropical cyclone genesis related variables by using a probabilistic graphical model, and logistic regression algorithm to run a classification prediction experiment of tropical clouds [23]. Their predicted results showed this method can achieve a good Auc value (the area under the ROC curve). Zhang W used decision tree algorithms to predict whether tropical cloud clusters will form into tropical cyclones which not only achieved better predicted results, but also established threshold conditions for related variables [24]. To our knowledge, no evaluation of machine learning based methodology has taken place to test its efficacy in typhoon formation forecasting.

While machine learning with its attendant deep learning methodology has improved aspects of typhoon formation forecasting, it lacks the advantages of neural networks in capturing spatial and temporal features of weather phenomena. The typical Convolutional Neural Network (CNN), was initially modeled on the Neocognitron, which was used to extract spatial information. Lecunet al., added the Back Propagation (BP) algorithm to the Neocognitron, and then developed the CNN network structure, LeNet-5 to identify handwritten numbers [26]. CNNs have now developed spatial recognition to the point that they are widely recognised as a proven technology in image recognition, natural language processing and text classification and so on [26, 27]. This is complemented for our purposes of the typhoon prediction model by the development of temporal data of Recurrent Neural Networks (RNNs) which are used to capture temporal relationships. They have already been used successfully in other applications such as speech recognition, language modelling, machine translation, picture description, etc. However, there are two problems of gradient disappearance and gradient explosion which occur when errors are back propagated for RNN. Schmidhuber et al. improved the traditional RNN structure and then proposed the LSTM (Long Short Term Memory) network [28]. Compared with the previous RNN, the newer LSTM can deal with longterm dependencies in the sequences more effectively. Subsequent developments of the improved model have made the RNN more widely used for sequence prediction. These improved models include Gated Recurrent Unit (GRU), Multidimensional LSTM, and Grid LSTM, etc [29].

Deep learning can also be applied to several problems in the meteorological field relevant to out investigations in precipitation and the prediction of typhoon formation and intensity [5, 30, 31] . The complexity of atmospheric phenomena with its temporal and spatial dimensions necessitates both CNN and RNN are combined to overcome the problems of forecasting using their specialised advantages of spatial and temporal gathering data together. Shi Xingjian et al. proposed the ConvLSTM model for the spatial-temporal correlation of atmospheric variables, and transformed precipitation forecasting into spatio-temporal sequences and constructed an end-to-end precipitation forecasting model. The experimental results showed the model is better than the FC-LSTM and the ROVER algorithm [5]. Similarly, computer vision field also have to face spatio-temporal problems, as well as meteorological field. In order to solve the spatio-temporal dependence of different lengths of videos, Zhang L et al. combined 3DCNN, ConvLSTM and 2DCNN to establish a deep learning model for gesture recognition. In this model, 3DCNN is used to learn short-term spatio-temporal features, and Bidirectional ConvLSTM is used for learning long-term spatio-temporal features. They consider the deep learning framework to be an effective method to analyse spatio-temporal problems [8]. Researches using 3DCNN for

target recognition and behavior recognition have been found frequently in many international conferences in the last several years [8, 32–34].

## 3 Problem definitions

Our goal is to construct a predictive model for typhoon formation and intensity combining historic typhoon data, atmospheric and oceanic data. The transition from the genesis of TC to a typhoon is a process which occurs along a spectrum of related meteorological phenomena. For example, a transitional feature along this spectrum and precursor of a typhoon is the Tropical Storm (TS) ($>= 34$kt, $< 64$kt). The movement from the genesis of the tropical cyclone, to TS to Typhoon are moments fall under the definition of the tropical cyclone activity.

As this paper has argued these 'moments' constitute a temporal-spatial sequence of events or process which can be observed measured and interpreted through multi-moment analysis to produce more accurate forecasts of typhoon formation and intensity. An obvious difficulty in our model is finding a way to integrate the spatial and temporal aspects of typhoon formation and intensity in our forecasting model. The problem arises from the fact, whereas typhoons are a holistic phenomenon occurring in real time and space the limitations of current methodology separate these distinct but related data into atmospheric and oceanographic categories(based on a methodological division of the 3D spatial structure of typhoon and atmospheric phenomena; and, the 2D of the oceanographic). Our model seeks to overcome this division, and combine the data from both atmospheric and oceanographic variables to create a more accurate predictive model of typhoon formation and intensity. Because the surrounding variables of the TS center at any time is a spatial structure, the atmospheric variables are in a 3D space that includes longitude, latitude, level, while the oceanic variables are in a 2D space that only include longitude and latitude. Therefore, considering the existence of spatio-temporal relationships among various variables, the problem can be defined as a problem of spatio-temporal sequences forecasting.

A visual representation of the methodology used to define the research problem of the predictive CNN-LSTM model for typhoon formation and intensity is shown in Fig. 1. As Fig. 1 illustrates, the model uses as its starting point a 2-dimensional spatial grid point system which can be used to map and sequence the transition from TS to typhoon development variables related to atmospheric or sea level phenomena. The grid map around the
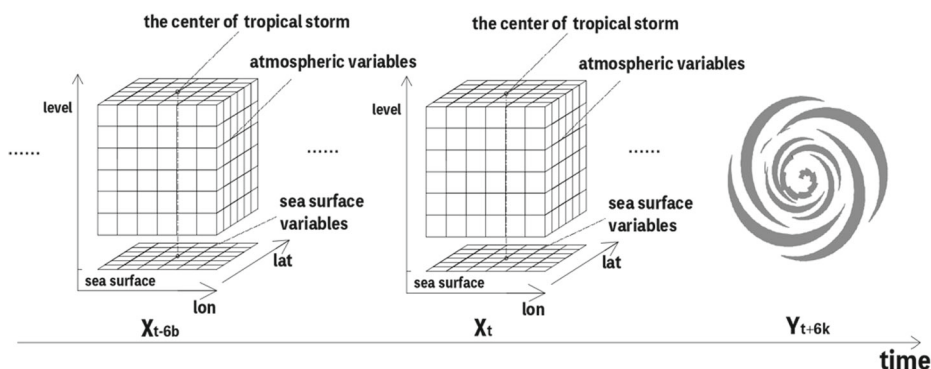


**Fig. 1** The structure of spatio-temporal variables sequence

TS center is shown as $M \times M$ grid points. $L$, and $R$ indicate atmospheric pressure levels and variables, whilst $O$, the sea-surface variables. The atmospheric variables $X^P$ around the TS center is shown by the $M \times M \times L \times R$ grid dataset. The sea surface variables $X^S$ around the TS center are denoted by the $M \times M \times O$ grid dataset. Therefore, the spatial variables including atmospheric and oceanic can be represented as $X = [X^P, X^S]$. The forecasting moment is $t$, and the variables at this moment $X$, combining to make $X_t$. $Y$ is the forecasting object that indicates whether the storm will develop into a typhoon or not; and the intensity of a typhoon; and $k$, is the forecasted timesteps, $b$ is the earlier lookback timesteps before the forecasting moment(each time sequence or step is six hours apart).

Therefore, the problem to be solved in this paper for the forecasting of typhoon formation and intensity involves the sequencing and inclusion of all the relevant atmospheric and oceanic grid data sets information around the TS and typhoon center, including, utilizing earlier synoptic time steps to predict whether the storm will form into a typhoon or not and the intensity of the storm in the following hours.

The central problems in our model of forecasting typhoon formation and intensity can likewise be represented in the following terms. The problem of typhoon formation forecasting can be defined using the spatio-temporal variable sequence of $X_{t-6b}(b = 0, 1, 2, 3, ...)$, to predict whether a TS will form a typhoon ($Y_{t+6k} = 1$) or not ($Y_{t+6k} = 0$). Similarly, using the spatio-temporal variables sequence $X_{t-6b}(b = 0, 1, 2, 3, ...)$, to predict the strength of the typhoon($Y_{t+6k} = intensity$) in the future.

The model can be also denoted by

$$Y_{t+6k} = f(X_t, X_{t-6b}), k, b = 1, 2, 3..., \tag{1}$$

where $Y_{t+6k}$ is the forecasted object in next $6k$ hours; $f$ is the final model learnt by the historical data, $X_t$ is the grid datasets of the environmental variables around the TS center at the forecasting moments; and , $X_{t-6b}$ is the grid datasets of the variables in $6b$ hours before the forecasting moments.

## 4 Methods

The section introduces the details of hybrid CNN-LSTM model which is capable of forecasting typhoon information and intensity. How to learn the spatial information of sea surface will be shown in the first subsection. Likewise, the mechanism of learning spatial information of atmosphere and temporal information of typhoon activity will be represented in the second and third subsetions seperately. Subsequently, the whole CNN-LSTM model and its loss function will be displayed in the last two subsections.

### 4.1 Learn the spatial feature of sea surface variables by 2DCNN

The sea surface variables are 2D in space. The reason why is 2DCNN be used is that sea surface variables are also 2D. Usually, a CNN is consisted of input layer, convolution layer, pooling layer, fully connected layer, and activation layer. After inputting sea surface variables, spatial features are learnt by 2D convolution layer. As shown in Fig. 2, it works by extracting features from the local neighborhood of the previous feature map, and then passing the activation function to learn.
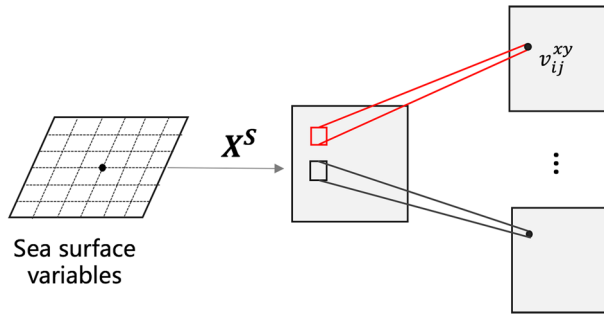
**Fig. 2** The 2DCNN layer for sea surface variables

According to reference [32], the value of unit at position $(x, y)$ in the $j$th feature map in the $i$th layer, denoted as $v_{ij}{}^{xy}$, is given by

$$v_{ij}{}^{xy} = f(\sum_{k=0}^{P_{i-1}} \sum_{p=0}^{Q_{i-1}} \sum_{q=0} w_{ijk}^{pq} v_{(i-1)k}^{(x+p)(y+q)} + b_{ij}), \tag{2}$$

where $f$ is an activation function, $m$ indexes over the set of feature maps in the $(i-1)$th layer connected to the current feature map, $w_{ijk}^{pq}$ shows the value at the $ijk$ position $(p, q)$ of the kernel connected to the $k$th feature map, and $P_i$ and $Q_i$ represent the height and width of the kernel, $b_{ij}$ denotes the bias for this feature map. Multiple layers built for extracting complex features, and then the feature vectors of sea surface variables could be got for forecasting.

## 4.2 Learn the spatial feature of atmospheric variables by 3DCNN

Typhoon is a strong convective weather processes. Additionally, the atmospheric variables are 3D in space and have strong correlation. Inspired by traditional physical features such as wind shear and vorticity etc., the CNN is used to effectively perform more complex feature learning. The correlation among atmospheric variables in the grid points can be learnt by CNN with 3D filters, which is 3DCNN in the paper. As shown in Fig. 3, it works by learning features from whole spatial structure.

According to reference [32], the value at position $(x, y, z)$ on the $j$th feature map in the $i$th layer is given by

$$v_{ij}{}^{xyz} = f(\sum_{k=0}^{P_{i-1}} \sum_{p=0}^{Q_{i-1}} \sum_{q=0}^{R_{i-1}} \sum_{r=0} w_{ijk}^{pqr} v_{(i-1)k}^{(x+p)(y+q)(z+r)} + b_{ij}), \tag{3}$$

where $R_i$ represents the size of the 3D kernel along the temporal dimension, $w_{ijk}^{pqr}$ is the $(p, q, r)$ value of the kernel connected to the $k$th feature map in the previous layer. The parameters of CNN, such as the bias $b_{ij}$ and the kernel weight $w_{ijk}^{pq}$, are usually trained using either supervisedly or unsupervisedly [27]. Multiple layers also needed to be built for extracting complex features, and then the feature vectors of atmospheric variables could be got for forecasting.
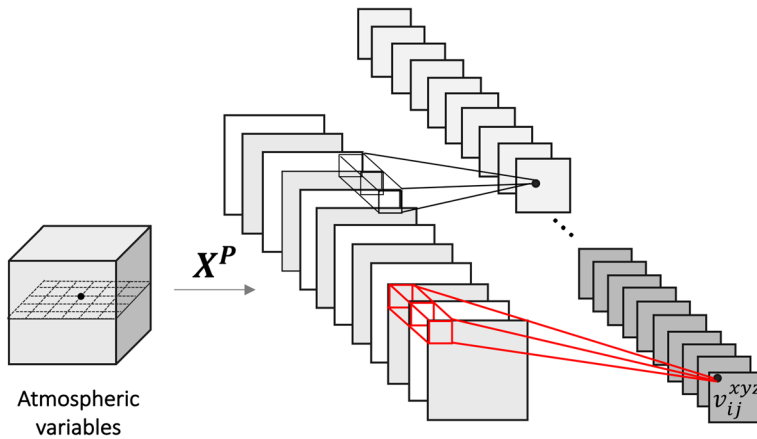
**Fig. 3** The 3DCNN layer for atmospheric variables

## 4.3 Learn the temporal realtionship of variables by LSTM

To learn temporal sequence relationships in the paths of typhoon, the LSTM networks are used. LSTM is an improved model of Recurrent Neural Networks(RNN). It can keep the error at a constant level, and enhance the robustness. The current input is $X_t$, the state of last hidden layer $h_{t-6}$ and the last memory $C_{t-6}$. LSTM works by forgetting and remembering the incoming information through three gates. The forget gate determines what information should be discarded from the $C_{t-6}$, the input gate determines what new information should be stored in the LSTM cell, $\widetilde{C}_t$ indicates the information that can be stored in the LSTM cell, and the output gate determines what kind of information can be passed to the next LSTM cell. These gates are all decided by the current input information and the last hidden layer state $h_{t-6}$ (Fig 4).
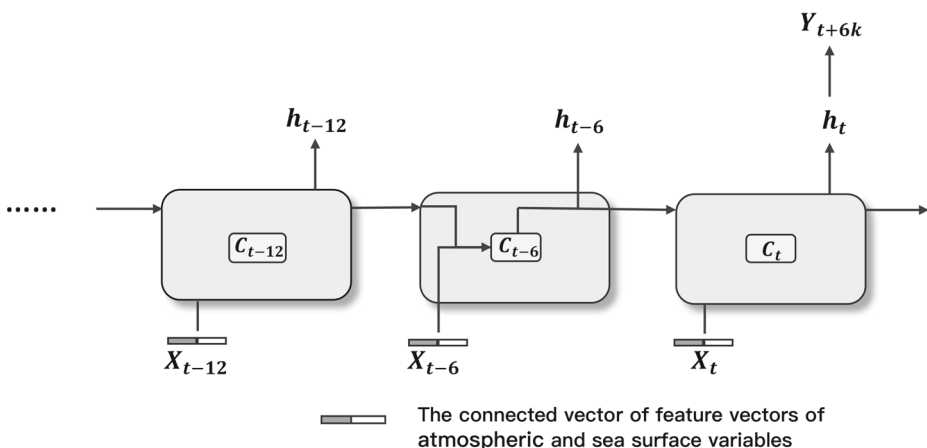


**Fig. 4** The LSTM layer for typhoon path

It can be concluded as

$$Y_{t+6k} = \sigma(Wh_t + b)$$
$$(h_t, C_t) = LSTM([h_{t-6}, X_t], C_{t-6}), \tag{4}$$

where $W$ is weight matrix and $b$ is bias, $\sigma$ denotes the 'sigmoid' function [28].

Inputting the features that learnt by atmospheric variables using 3DCNN and sea surface variables using 2DCNN into the LSTM, then get the output $Y_{t+6k}$ through the learning of temporal relationships at multi timesteps. $Y_{t+6k}$ is also the objective of hybrid CNN-LSTM model.

## 4.4 The framework of the hybrid CNN-LSTM model

The overall framework of our model is showed in Fig. 5. The hybrid CNN-LSTM model proposed to learn the spatio-temporal correlation of atmospheric and oceanic variables for typhoon formation and intensity forecasting. The model uses the 3DCNN to learn the spatial relationships of atmosphere variables, 2DCNN to learn the spatial relations of sea surface variables and the LSTM to learn the temporal relationships of variables.

The atmospheric variables $X^P$ around a TS center at any time is a $M \times M \times L \times N$ grid dataset, which is also a 4D tensor as the input of 3DCNN. In the experiments part, the 3DCNN of the model includes an input layer, three 3D convolution layers, a flatten layer, and a fully connected layer. The size of filters in the first convolution layer is $5 \times 5 \times 3$, the stride is $2 \times 2 \times 1$, and the number of filters is 32. The size of filters in the second convolution layer is $5 \times 5 \times 1$, the stride is $2 \times 2 \times 1$, and the number of filters is 64. The size of filters in the third convolution layer is $5 \times 5 \times 1$, the stride is $2 \times 2 \times 1$, and the number of filters is 128. The fully connected layer contains 100 neurons. The activation function of all layers is 'relu', which can keep the convergence speed of the model in a steady state.

The sea surface variables $X^S$ around the TS center at any time is a $M \times M \times O$ grid dataset, which is also a 3D tensor as the input of 2DCNN. In the experiments, the 2DCNN includes an input layer, three 2D convolution layers, a flatten layer, and a fully connected layer. The filters in different 2D convolution layers have different numbers, but the same
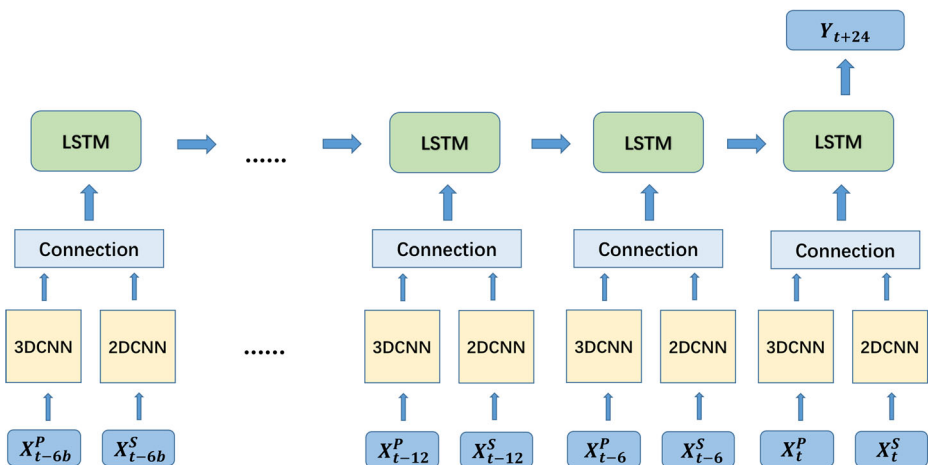


**Fig. 5** The framework of the hybrid model

size and stride. The size of all the filters is 5 × 5, stride is 2 × 2, and the numbers of three convolution layers are 32, 64, 128. The fully connected layer contains 100 neurons. The activation function of all 2D convolution layers and fully connected layers is 'relu'.

After extracting the features of atmospheric variables by 3DCNN and the features of sea surface variables by 2DCNN, the output features of the 3DCNN and 2DCNN are connected in the 'Connection' part, and as an input with the length $K$ to the LSTM. In the later experiments, only single-layer LSTM is considered since typhoon paths are short, and the number of neurons in hidden layer is 100; and, the activation function is 'relu'. The activation function chosen in output layer is 'sigmoid'.

The Hybrid CNN-LSTM model can be defined simply as

$$Y_{t+6k} = LSTM(3DCNN(X^P_{t-6b}), 2DCNN(X^S_{t-6b})), \tag{5}$$

where $k = 1, 2, 3...$ and $b = 0, 1, 2, 3...$

## 4.5 The loss function of the hybrid CNN-LSTM model

For the formation forecasting, as a classification problem in machine learning, the cross-entropy function which can measure the difference between the predicted value and actual value and presents the distances between two probability distributions is selected as the loss function. For the intensity forecasting, as a regression problem in machine learning, the Root-Mean-Square Error function(RMSE) is selected. RMSE which which denotes the sum of the squares of the difference between the predicted value and the actual value is commonly used for the loss function of regression problems.

The loss function of formation forecsting can be defined as

$$J = -\frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}((1 - y_{t,i}) \times log(1 - \hat{y}_{t,i}) + y_{t,i} \times log(\hat{y}_{t,i})). \tag{6}$$

And the loss function of intensity forecasting can be difined as

$$J = \sqrt{\frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{t,i} - \hat{y}_{t,i})^2}, \tag{7}$$

where T denotes the timestep, N shows the number of samples, $y_{t,i}$ is the true value of $i$th sample at $t (t \in T)$ moment, $\hat{y}_{t,i}$ represents the predictive value of $i$th sample at $t$ moment, $\hat{y}_{t,i} = \sigma(Wh_t)$, $h_t = \sigma[h_{t-1}, x_t]$, $x_t$ denotes the input of LSTM, which is also the concatenation of the output of 3DCNN and 2DCNN, $x_t = [O_{3D}, O_{2D}]$, $O_{3D}$ is the output vector of 3DCNN after inputing $X^P$, $O_{2D}$ is the output vector of 2DCNN after inputting $X^S$.

## 5 Experiments

This section detailedly presents the experiments of the proposed model. In terms of the typhoon and environmental datasets, evaluation metrics of the model, implementation and specific parameters analysis.

## 5.1 Datasets

The World Meteorological Organization (WMO) version of the International Best Track Archive for Climate Stewardship (IBTrACS) global tropical cyclone best track dataset, and ERA-Interim reanalysis datasets are used for our research. The global tropical cyclones best track dataset are divided by different regions. The best track data of a tropical cyclone recorded starts from the formation of a TS, and time interval of records is 6 hours. With the development of TSs, some could develop to typhoons, and the others began to decay after reaching the maximum wind speed near the center. Based on the maximum wind speed near the center in the typhoon records, tropical cyclones can be categorized into two types: one is TSs that will develop into typhoons, and the other is TSs that will not develop into typhoons. The ERA-Interim's global reanalysis data has been available since 1979. The types of level include pressure level, surface level, and model level. Variables in each level contain multi resolutions and multi types. Due to the dramatic changes in the atmosphere and sea surface around the TS center, the data of high-resolution and multi layers needed to be chosen.

Generally, different forecasted time is considered in operational typhoon forecasting. The longer the model could predict, the more significant it is. However, typhoon samples are rare unlike other common datasets in deep learning. In order to ensure enough samples for the model, forecast time selected is 24 hours. For typhoons, the forecasting moment is 24h before the typhoon is formed. For TS, the forecasting moment is the moment that reached maximum wind speed near the center. Additionally, to carry out the forecasting experiment, the spatial variables at the forecasting moment as well as the spatial variables at the lookback moments are consisdered. Here, the lookback moments selected is 6h, 12h, 18h, 24h before typhoon formed. These samples in the three typical sea areas that are the Western Pacific (WP), Eastern Pacific (EP), and North Atlantic (NA) are used for our experiments. The number of typhoons and TS are listed in Table 1.

As for the selection of environmental variables of ERA-Interim reanalysis datasets. Since the horizontal scale of typhoon or TS may reach about 1000km, the surrounding region of center selected ranges in $4° × 4° \sim 20° × 20°$, which can cover the entire typhoon zone. Besides, typhoon is a spatial structure in the vertical direction and it can be divided into three parts: the inflow layer, the intermediate layer and the outflow layer. From the sea surface to the height of 3 km is the inflow layer, the height of 3-8 km is the middle layer, and the height from 8 km to the typhoon roof is the outflow layer. So, atmospheric pressure level of 1000/975/925/850/800/700/600/500/400/300/200 /100hpa are selected. In addition, the sea surface has a crucial impact on the activities of the typhoon. Therefore, sea surface level is selected. For the type of atmospheric variables, the basic variables such as u-component of wind(u), v-component of wind(v), temperature(t), relative humidity(rh) and geopotential height(z) are chosen. For the Sea surface variables, the most critical sea surface temperature(sst) is enough for the experiments, and the other physical variables found in previous studies that are favorable to the development of TS, such as the vertical wind shear and max potential index [18] are also calculated from these basic variables. The model

| **Table 1** The number of Typhoons and TS | WP | EP | NA |
|---|---|---|---|
| Typhoon | 239 | 219 | 181 |
| TS | 211 | 220 | 186 |

proposed in this paper hopes to obtain better prediction effects by the basic atmospheric and sea surface variables, which can be more common.

The typhoon formation and intensity forecast is a problem of spatio-temporal sequence forecasting, so it is necessary to organize input and output datasets for the model based on the above selected spatio-temporal data. For the 3DCNN component of the model, the form of all the input datasets is $N \times M \times M \times L \times R$. For the 2DCNN of our model, the form is $N \times M \times M \times O$. For the LSTM of the model, the form is $N \times T \times K$. In the later experiments, in WP, $N = 450$. In EP, $N = 439$. In NA, $N = 367$. M ranges in $33 \sim 161$. $L = 12$. $R = 5$. $O = 1$. $T$ ranges in $1 \sim 5$. K is the length of the input vector, and $K = 200$. These parameters can be adjusted according to different objects. When conducting the experiments of typhoon formation forecasting, typhoons are labeled as 1 and TS are labeled as 0. When conducting the experiments of typhoon intensity forecasting, the intensity of a typhoon is the label. Among all of the experiments, 70% of all samples are used for training, 30% else for testing.

## 5.2 Evaluation metrics

In order to measure the prediction performance of the model, the Acc(Accuary) and Auc is used to evaluate the effect for formation forecasting (Table 2). The MAE(Mean Absolute Error) is used to evaluate the error of intensity forecasting.

Acc is the proportion of correctly predicted samples to the total number of samples, which can be defined as

$$Acc = \frac{TP + TN}{TP + FN + FP + TN}. \tag{8}$$

Auc is the area under the ROC curvecan, which can effectively describe the overall performance of the model. The abscissa of ROC curve is false positive rate ($FPR = \frac{FP}{FP+TN}$), and the ordinate is True positive rate ($TPR = \frac{TP}{TP+FN}$). Auc can be defined as

$$Auc = \int ROC. \tag{9}$$

The Auc value is a probability that is between 0.1 and 1. The larger the Auc value, the current classification model more likely to place positive samples in front of negative samples, so that these samples can be better classified.

MAE is the mean value of the absolute errors, which can reflect the forecasting error well. MAE can be defined as

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| (y_i - \hat{y}_i) \right|, \tag{10}$$

where $N$ is the number of samples, $\hat{y}_i$ indicates the predicted intensity value and $y_i$ indicates the true intensity.

**Table 2** The confusion matrix [35]

| Recorded | Forecasting | |
|---|---|---|
| | Typhoon | TS |
| Typhoon | TP | FN |
| TS | FP | TN |

### 5.3 Implementation

The model is built and implemented by keras that uses tensorflow as the backend. Two NVIDIA GTX 1080Ti GPU are used for training the model. In order to build the hybrid model, Conv3D, Conv2D, LSTM, Dense, Merge and TimeDistributed layer among keras layers are combined. The TimeDistributed is a layer wrapper which applies a layer to every temporal slice of an input. The convolution layer is chosen for extracting spatial environment features at each time step, so it is necessaty to ignore the effects of time first when training the model. The Merge layer can combine several tensors in a tensor list into a single tensor, which can concentrate the outputs of 3DCNN and 2DCNN. Because of the presence of missing values on some grid and a big difference between different variables, data preprocessing is necessary. Filling the missing values and normalize the data, which can accelerate the model. After fixing the network structure of the model, the next step is to adjust hyper-parameters such as learning rate, epochs, batchsize and so on, to get the best forecasting results. The sklearn's GridSearchCV module is also a good choice to find the most optimal hyper-parameters.

### 5.4 Analysis

#### 5.4.1 Results analysis for formation forecasting

There are four comparing traditional machine learning based methods for typhoon formation forecasting problem , that is Logistic Regression (LR), Random forest (RF), Linear Discriminant Analysis (LDA), and Decision Tree (DT) [4, 23, 24, 36]. In the references, these methods are used for the typhoon genesis, not exactly for typhoon formation forecasting. They regard the atmospheric and oceanic variables as independent features, without considering the spatio-temporal relations between variables. Since the samples of tropical cyclone genesis problem are extremely uneven, so these methods can achieve a high accuracy. Additionally, the original LSTM which doesn't consider the spatial relations and the existing improved methods ConvLSTM and other CNN methods combined with LSTM, such as 2DCNN+LSTM, 3DCNN+LSTM also are compared with our model. The experimental results of typhoon formation forecasting are showed in Table 3.

**Table 3** The experimental results of typhoon formation forecasting. Bold entries show the best results in different experimental regions

|  | WP | | EP | | NA | |
|---|---|---|---|---|---|---|
|  | Acc | Auc | Acc | Auc | Acc | Auc |
| LR | 0.718 | 0.720 | 0.727 | 0.728 | 0.730 | 0.730 |
| DT | 0.637 | 0.637 | 0.652 | 0.652 | 0.613 | 0.610 |
| RF | 0.778 | 0.780 | 0.682 | 0.681 | 0.649 | 0.651 |
| LDA | 0.741 | 0.741 | 0.735 | 0.763 | 0.613 | 0.612 |
| LSTM | 0.667 | 0.774 | 0.667 | 0.694 | 0.595 | 0.610 |
| ConvLSTM | 0.756 | 0.799 | 0.742 | 0.807 | 0.685 | 0.759 |
| 2DCNN+LSTM | 0.766 | 0.833 | 0.734 | 0.846 | 0.721 | 0.764 |
| 3DCNN+LSTM | 0.812 | 0.897 | 0.763 | 0.847 | 0.729 | 0.773 |
| **Hybrid CNN-LSTM** | **0.852** | **0.922** | **0.780** | **0.858** | **0.759** | **0.794** |

**Table 4** The MAE value of intensity forecasting compared with existing models (kt). Bold entries show the best results in different experimental regions

| | WP | | EP | | NA | |
| --- | --- | --- | --- | --- | --- | --- |
| | Model | MAE | Model | MAE | Model | MAE |
| Statistical models | PLS | 11.5 | SHIFOR5 | 10.2 | SHIFOR5 | 12.3 |
| | WIPS | 10.9 | SHIPS | 11.5 | SHIPS | **9.1** |
| Numerical models | GFS | 12.9 | GFS | 10.8 | GFS | 11.9 |
| | IFS | 14.3 | EMXI | 16.8 | EMXI | 10.6 |
| | TCM | 13.7 | HWRF | 11.6 | HWRF | 9.3 |
| | Hybrid CNN-LSTM | **7.4** | | **9.4** | | 9.4 |

For the traditional methods based on machine learning in WP, the optimal accuracy is 0.7778 and the best Auc value is 0.7800. Among other models based on deep learning, 3DCNN+LSTM can achieve the best accuracy of 0.812 and Auc of 0.897. ConvLSTM designed for precipitation (a typical meteorological phenomenon) prediction is not suitable for typhoon prediction unless three-dimensional atmospheric variables are projected into two-dimensional space, and then two-dimensional oceanic variables are combined into the forecasting model, resulting in performance degradation. This is due to the lack of capturing spatial characteristics of atmospheric information, 2DCNN+LSTM also has similar weaknesses. In addition, 3DCNN+LSTM, which can capture three-dimensional features has certain advantages shown in Table 4. Compared with the Hybrid CNN-LSTM model, the reason why 3DCNN+LSTM can achieve an approaching result but not higher accuracy is that the difference between atmosphere and ocean might be neglected while extracting the characteristics of variables. Therefore, it can be shown in Table 4. that the Hybrid CNN-LSTM model for 24h typhoon formation forecasting can achieve the best Acc and the best Auc, which is 0.852 and 0.922 respectively. To ensure the model is robust in different regions, the experiments in EP and NA region are also conducted. In EP region, the traditional methods achieve the best accuracy of 0.735, the best Auc value of 0.763 by LDA, the other models based on deep learning achieve the best accuracy at 0.763 and the best Auc value at 0.847 by using 3DCNN+LSTM model. The Hybrid CNN-LSTM model can achieve the best accuracy of 0.780 and the best Auc value of 0.858 which is better than existing methods. In NA region, the best Acc and Auc are computed by our hybrid model. The results in EP and NA have a lower accuracy than in WP, which is possibly caused by different numbers of the samples and the reason of differences between environment in different regions.

### 5.4.2 Parameters analysis for formation forecasting

**Lr Analysis** Gradient descent algorithm is the most common optimization method used to train neural networks. The gradient can be calculated by partial derivation. The learning rate (lr) is an important hyperparameter, which can control the update speed of parameters. It was found that at the beginning of the training process, the model was in a non-converging state when using the adaptive optimization algorithm such as 'Adam' algorithm. Then adjusting the lr in the fixed model framework and other hyper-parameters. The first step is dropping it from 0.01 to 0.0001, and the decrease rate is 10. Then, the model training and validation loss will be in a steadily declining state while the learning rate was on the level of $10^{-4}$. Finally, the experiments are conducted on different datasets, and the results are shown in
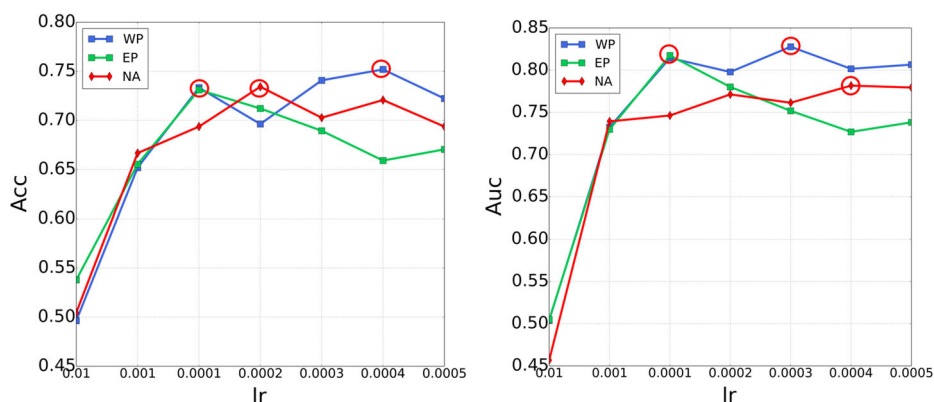
**Fig. 6** Auc and Acc changes with learning rate (lr)

Fig. 6. The left figure denotes the Acc changes with lr varying, the right figure denotes the Acc changes with lr varying. According to the Acc, in WP region, the best lr is 0.0004. In EP region, the best lr is 0.0001. In NA region is 0.0002. According to the Auc, in WP region, the best lr is 0.0003. In EP region, the best lr is 0.0001. In NA region is 0.0003. Therefore, the best lr is better to be on the level of $10^{-4}$ in our datasets.

**Epochs analysis** To determine the best epochs of datasets, running the model end with 25 epochs or 30 epochs in different regions. As is shown in Fig. 7, the Auc is in a steady state with 25 epochs in WP, the Auc begins to decrease after 15 epochs in EP, and the Auc starts reach a steady state at 30 epochs in NA region. Therefore, in our experiments, 25 epochs suitable for WP region, 15 epochs suitable for EP region and 40 epochs suitable for NA region.

**Region range analysis** Apart from the hyper parameters of the model play a crucial role in fitting model, the input dataset also has a critical influence for forecasting. We conduct test experiments for selecting the surrounding area. The surrounding area, that is region range, varies from $4° × 4°$ to $20° × 20°$. Figure 8 shows the Auc and Acc changes with different region range of environmental grid data. We can find that it's better to choose region range of environmental grid data to be between $9° × 9°$ and $13° × 13°$ in our datasets. It's not the
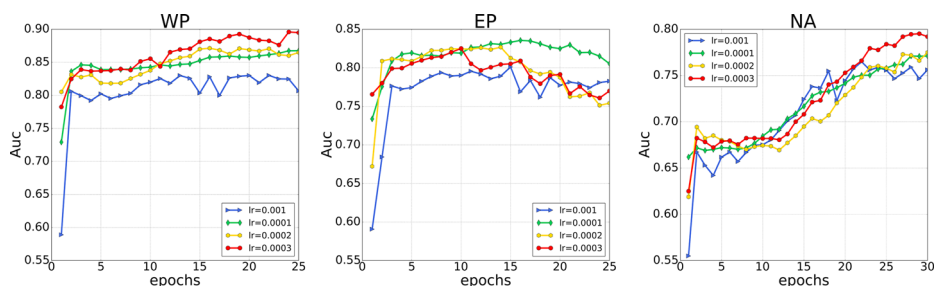


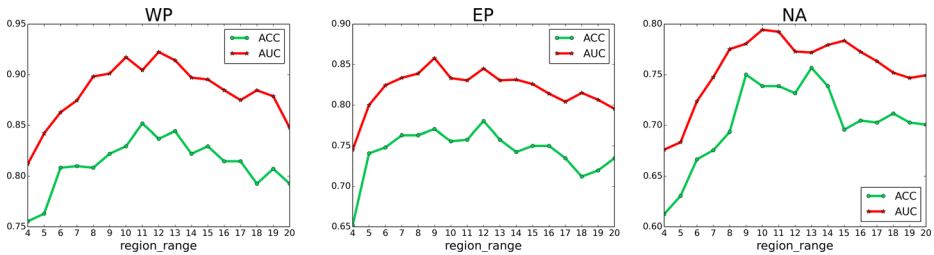**Fig. 7** The Auc results with the change of epochs

**Fig. 8** The Auc and Acc results with different region range of environmental grid data

region range bigger, the performance better. In WP region, the best region_range is $12° × 12°$ by evaluating Auc, and the best region_range is $11° × 11°$ by evaluating Acc. In EP region, the best region_range is $9° × 9°$ by evaluating Auc, and the best region_range is $12° × 12°$ by evaluating Acc. In NA region, the best region_range is $10° × 10°$ by evaluating Auc, and the best region_range is $13° × 13°$ by evaluating Acc.

**Timestep analysis** Timestep is an important parameter for the model to learn temporal information. Usually, the longer timestep is, the more past information can be used for forecasting. But there is a critical problem that whether the accumulation of errors will be greater than the accumulation of valid information or not. Therefore, the datasets in three regions are tested separately. As shown in Figs. 9, 10 and 11, timestep=2 is better than other timestep in most cases. To ensure the analysis has the statistical significance, the t-test is carried out for validation and a p-vlaue is calculated. If the p-value is lower than 0.05, the difference between two evaluation results can be considered statistically significant. The lower the p-value, the more significant the difference[13]. For the t-test between timestep=2 and timestep=1 in WP, the p-value of acc is 0.021. For the t-test between timestep=2 and timestep=5 in WP, the p-value of acc is $2.09 × 10^{-7}$. In EP and NA regions, the p-values are 0.002, $6.35 × 10^{-10}$, 0.1309, 0.0147 respectively.

## 5.5 Results analysis for intensity forecasting

In order to ensure the model is reliable in operational forecast. The intensity forecasting of 24 hours also is included in the paper, and the intensity errors of all points in the path of
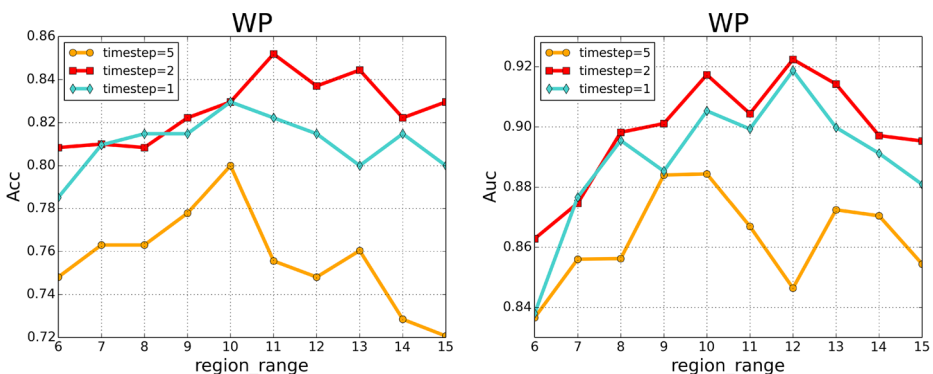


**Fig. 9** The Auc and Acc results of different timestep of LSTM in WP
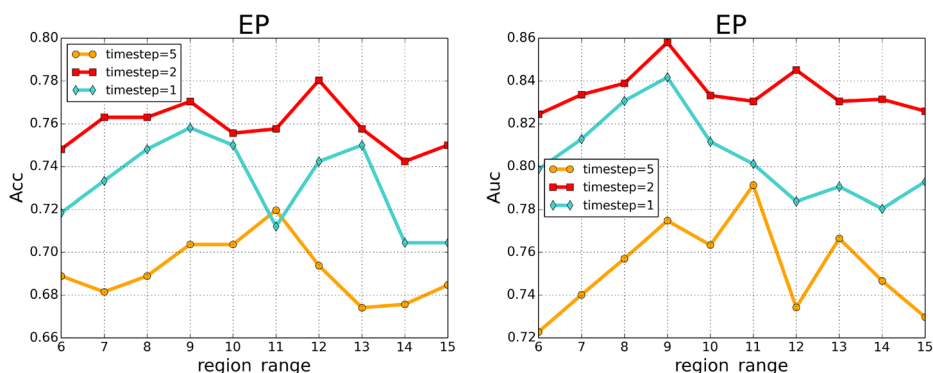
**Fig. 10** The Auc and Acc results of different timestep of LSTM in EP

typhoon are caculated. The error data comes from the 'Verification on Forecasts of Tropical Cyclones over Western North Pacific' of CMA [13] and 'National Hurricane center Forecast Verification Report' [15]. The selected numerical forecast model, including the regional model "Hurricane Weather and Research Forecasting Model (HWRF)", "Tropical Cyclone Model based on Global Regional Assimilation Prediction System(GRAPES-TCM)", "ShangHai typhoon Model"(SHTM), the global model "ECMWF global model (EMX)", the ensemble model "Florida State Super Ensemble (FSSE)", and the consensus model "Global Forecast System(GFS)". The statistical forecast methods selected include "Climatology and Persistence model (CLIPER5)", the Partial Least Square Regression Scheme (PLS), "WIPS(Western North Pacific Topical Cyclone Intensity Prediction Scheme (WIPS)" and the statistical-dynamical model "Statistical Hurricane Intensity Prediction Scheme (SHIPS)". They have been used by a lot of typhoon warming centers like China Meteorological Administration (CMA), National Hurricane center(NHC) and so on. The averaged intensity errors in past 5 years from CMA and NHC reports are used for the comparison.

Table 4 shows the MAE of intensity compared with existing model. The MAE of the hybrid CNN-LSTM model is potently better than existing methods in WP and EP regions.
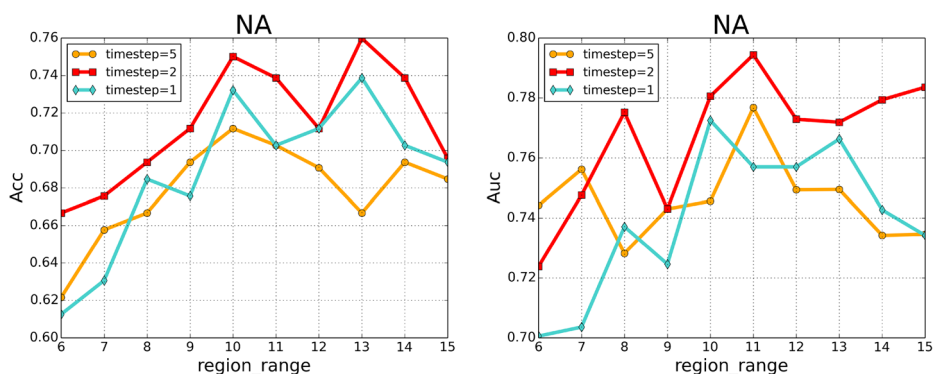


**Fig. 11** The Auc and Acc results of different timestep of LSTM in NA

The MAE of our model in WP is 7.4kt while the lowest MAE of existing models is 10.9kt, the MAE of our model in EP is 9.4kt while the lowest MAE of existing models is 10.2kt. But the MAE has a near result in NA region. Although the hybrid model is not the best in NA, the SHIPS is better than our model and its MAE is only 9.1kt while our MAE is 9.4kt. Our MAE is still lower than most of MAE from statistical models and Numerical models. In general, statistical models have advantages in existing models, but our model has a great improvement compared to the existing models on typhoon intensity forecasting.

## 6 Conclusions and future works

This paper has been concerned with the propagation of a hybrid CNN-LSTM model with a higher predictive forecasting potential for typhoon formation and intensity than existing systems. In this model, the forecasting of typhoon formation is defined as a problem of the classification of spatio-temporal sequences prediction. The problem of typhoon intensity forecasting is defined as the problem of spatio-temporal sequences regression prediction. The three components of the hybrid model, the 3DCNN (is used to analyse atmospheric variables in 3-dimensional space); 2DCNN(is used to analyse those at the sea surface); and, LSTM (is used to capture the temporal correlations) are combined to collate data which is used to analyse the relationship between spatio-temporal phenomena and the variables of typhoon formation and intensity.

The hybrid CNN-LSTM model can be trained, validated, and tested using historical meteorological records. The atmospheric and sea surface variables related to typhoon activity in the West Pacific, East Pacific, and North Atlantic Oceans goes back decades, and these can be used as scenarios for test and training purposes. Extensive experiments in all three areas show the model is better than existing methods. The model has an accuracy rate of 85.2%, in predicting typhoon formation; an Auc value of 92.2%; and, an error of 7.4kt for the intensity forecasting of typhoons in the West Pacific Ocean, (which is lower than pre-existing official forecast errors). These figures show the model's effectiveness in predicting typhoon formation and intensity compared with other current methods.

Additional analysis, experimentation and work on the parameters of the model enable model training successfully. The optimal choices are as follows for differing aspects of model operational parameters: the optimal learning rate (rate for refreshing of the parameters of the model) is on the level of $10^{-4}$; the optimal epochs for adjusting iterations is in the range in $15 \sim 30$; the best region range of environmental grid data surrounding the typhoon center is in $9° \times 9° \sim 13° \times 13°$; and, finally, the optimal timestep is 2 in our datasets.

As for future development, further refinements to the hybrid CNN-LSTM model will be undertaken. This will involve the use of more and varied data such as high resolution satellite information. Moreover, as the model is in its infancy, experimentation, experiential learning and integration will undoubtedly improve its efficacy for forecasting typhoon formation and intensity and maybe other meteorological and oceanographic phenomena. The model could also be eventually used as a didactic tool for teaching a unified experimental protocol for future meteorologists interested in typhoon phenomena.

# References

1. Peng MS et al (2012) Developing versus nondeveloping disturbances for tropical cyclone formation. Part I: North Atlantic. Mon Weather Rev 140.4:1047–1066
2. Fu B et al (2012) Developing versus nondeveloping disturbances for tropical cyclone formation. Part II: Western North Pacific. Mon Weather Rev 140(4):1067–1080
3. Hennon CC (2017) Forecasting tropical cyclogenesis over the atlantic basin using Large-Scale data. Monthly weather review. 1–14
4. Kerns BW, Chen SS (2013) Cloud Clusters and tropical cyclogenesis: Developing and nondeveloping systems and their Large-Scale environment. Mon Weather Rev 141(1):192–210
5. Xingjian S et al (2015) Convolutional LSTM network: A machine learning approach for precipitation nowcasting. 802–810
6. Yang Y et al (2018) A CFCC-LSTM model for sea surface temperature prediction. IEEE Geosci Remote Sens Lett 15(2):207–211
7. Donahue J et al (2015) Long-term recurrent convolutional networks for visual recognition and description. 2625–2634
8. Zhang L et al (2017) Learning Spatiotemporal Features Using 3DCNN and Convolutional LSTM for Gesture Recognition. 3120–3128
9. Qu Q, Liu S, Zhu F, Jensen CS (2016) Efficient online summarization of Large-Scale dynamic networks. IEEE Trans Knowl Data Eng 28(12):3231–3245
10. Liu S, Qu Q (2016) Dynamic collective routing using crowdsourcing data. Transp Res B Methodol 93:450–469
11. Liu S, Qu Q, Wang S (2018) Heterogeneous anomaly detection in social diffusion with discriminative feature discovery. Inf Sci 439-440:1–18
12. Liu S, Qu Q, Wang S (2015) Rationality analytics from trajectories. TKDD 10(1):10:1–10:22
13. CMA Verification on Forecasts of Tropical Cyclones over Western North Pacific (2012 ∼ 2016)
14. Chen P, Yu H, Chan JCL (2011) A Western North Pacific tropical cyclone intensity prediction scheme[J]. Acta meteorologica sinica 25(5):611–624
15. Franklin JL (2008) National Hurricane center forecast verification report. National Hurricane center Rep. (2012 ∼ 2016)
16. Song J, Wang Y, Chen P (2011) A statistical prediction scheme of tropical cyclone intensity over the western North Pacific based on the partial least square regression [J]. Acta Meteorol Sinica 69(5):745–756
17. Leiming M (2014) Research progress on China typhoon numerical prediction models and associated major techniques. Progress in Geophys.(in Chinese) 29(3):1013–1022DOI
18. Camargo SJ et al (2007) Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. J Climate 20(19):4819–4834
19. Murakami H, Sugi M (2010) Effect of model resolution on tropical cyclone climate projections. Sola 6:73–76
20. Zhang M et al (2016) A genesis potential index for Western North Pacific tropical cyclones by oceanic parameters. J Geophys Res Oceans 121(9):7176–7191
21. Chand SS, Walsh KJE (2011) Forecasting Tropical cyclone formation in the Fiji region: A Probit Regression Approach Using Bayesian Fitting. Weather Forecast 26(2):150–165
22. Su Y et al (2018) A New Data Mining Model for Hurricane Intensity Prediction. 98–105
23. Wijnands JS et al (2016) Variable selection for tropical cyclogenesis predictive modeling. Mon Weather Rev 144(12):4605–4619
24. Zhang W et al (2015) Discriminating developing versus nondeveloping tropical disturbances in the Western North Pacific through decision tree analysis. Weather Forecast 30(2):446–454
25. Zhang W et al (2013) The application of decision tree to intensity change classification of tropical cyclones in western North Pacific. Geophys Res Lett 40(9):1883–1887
26. LeCun Y et al (1989) Backpropagation applied to handwritten zip code recognition. Neural Comput 1(4):541–551
27. LeCun Y et al (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
28. Hochreiter S, Schmidhuber J (1997) Long Short-term memory. Neural Comput 9(8):1735–1780
29. Kalchbrenner N et al (2015) Grid long short-term memory. arXiv:1507.01526
30. Li Y et al (2017) Leveraging LSTM for rapid intensifications prediction of tropical cyclones. IV-4/W2, 101–105

31. Qiu M et al (2017) A Short-Term Rainfall Prediction Model using Multi-Task Convolutional Neural Networks. 395–404
32. Ji S et al (2013) 3D convolutional neural networks for human action recognition. IEEE Trans Pattern Anal Mach Intell 35(1):221–231
33. Li Y et al (2016) Large-scale gesture recognition with a fusion of rgb-d data based on the c3d model. 25–30
34. Tran D et al (2015) Learning spatiotemporal features with 3d convolutional networks. 4489–4497
35. Zhou Z (2016) Machine Learning, Tsinghua University press
36. Schumacher AB et al (2009) Objective estimation of the 24-h probability of tropical cyclone formation. Weather Forecast 24(2):456–471

**Rui Chen** is a master student at College of Computer, College of Meteorology and Oceanology, National University of Defense Technology, Changsha, China. She got her bachelor degree in College of Oceanography, Hohai University. She is major in computer technology, and her research interests include machine learning, meteorology and oceanography.

**Xiang Wang** is an assistant professor at College of Computer and College of Meteorology and Oceanology, National University of Defense Technology. He got his Ph.D. degree in National University of Defense Technology. His main research interests are Data Mining and Machine Leaning.

**Weimin Zhang** is a researcher at College of Computer and College of Meteo- rology and Oceanology, National University of Defense Technology. He is also a doctoral supervisor. He got the bachelor degree in department of mathematics from Sichuan University, Chengdu, China, in 1986 and Ph.D. in college of computer from National University of Defense Technology, Changsha, China, in 2006. His research interests include data assimilation, parallel computation and numerical weather prediction.



**Xiaoyu Zhu** is a master student at College of Computer, National University of Defense Technology. He got his bachelor degree in College of Atmospheric Sciences, Lanzhou University. His research interests include deep learning, numerical weather prediction and remote sensing. He has published several papers in international journals and conferences, such as IEEE Geoscience and Remote Sensing Letters.



**Aiping Li** is a professor and doctoral supervisor at College of Computer, National University of Defense Technology. He got his Ph.D. degree in National University of Defense Technology. His main research interests are Data Mining and Information Security.

**Chao Yang** is a master student at College of Computer, National University of Defense Technology. He got his bachelor degree in School of Mathematical Sciences, Zhejiang University. His research interests include deep learning, remote sensing. He has published several papers in conferences, such as IEEE 2018 International Geoscience and Remote Sensing Symposium.