https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

Forecasting tropical cyclone tracks in the Northwest Pacific based on a deep-learning model

Liang Wang[1,2], Bingcheng Wan[3], Shaohui Zhou[3], Haofei Sun[1,2], Zhiqiu Gao[1,3*]

[1] State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China

[2] University of Chinese Academy of Sciences, Beijing, 100049, China

[3] School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, 210044, China

[*]*Correspondence to*: Dr. Zhiqiu Gao(zgao@mail.iap.ac.cn)

Abstract. Tropical cyclones (TCs) are one of the most severe meteorological disasters, making rapid and accurate track forecasts crucial for disaster prevention and mitigation. Because TC tracks are affected by various factors (the steering flow, thermal structure of the underlying surface, and atmospheric circulation), their trajectories present highly complex nonlinear behavior. Deep learning has many advantages in simulating nonlinear systems. In this paper, we explore the movement of TCs in the Northwest Pacific from 1979 to 2021 based on deep-learning technology, divided into training (1979–2014), validation (2015–2018), and test sets (2019–2021), and create 6–72 h TC track forecasts. Only historical trajectory data are used as input for evaluating the forecasts of the three recurrent neural networks utilized: recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU) models. The GRU approach performed best; to further improve forecast accuracy, a model combining GRU and a convolutional neural network (CNN) called GRU_CNN is proposed to capture the characteristics varying with time. By adding reanalysis data of the steering flow, sea-surface temperatures, and geopotential height around the cyclone, we can extract sufficient information on the historical trajectory features and three-dimensional spatial features. The results show that GRU_CNN outperforms other deep-learning models without CNN layers. Furthermore, by analyzing three additional environmental factors through control experiments, it can be concluded that the historical steering flow of TCs plays a key role, especially for short-term predictions within 24 h, while sea-surface temperatures and geopotential height can gradually improve the 24–72-h forecast.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

The average distance errors at 6 h and 12 h are 17.22 km and 43.90 km, respectively. Compared with the forecast results of the Central Meteorological Observatory, the model proposed herein is suitable for short-term forecasting of TC tracks.

1 Introduction

The Northwest Pacific is the most active basin for tropical cyclones (TCs) in the world, generating over one-third of the total number of TCs (Gray, 1968). China, located on the western side of the Pacific Ocean with a coastline longer than 18,000 km, is one of the countries most severely influenced by TCs. These storm systems are accompanied by strong winds, heavy precipitation, and storm surges, resulting in severe disasters that affect human lives and economic growth (Goldenberg et al., 2001). Studies have shown that global warming will progressively intensify TCs over time (Emanuel, 2017; Schulthess et al., 2019). Since disasters caused by TCs are unavoidable and potentially destructive, accurately predicting the movement of TCs can provide sufficient preparation time for people in affected areas to implement disaster mitigation strategies.

Given the uncertainty of TC movements, the complexity and nonlinearity inherent in the atmospheric system, and the scarcity of ocean-based observational data, accurately predicting the center positions and intensities of TCs is a challenge. Currently, forecasting methods for TCs are mainly divided into two categories, with the primary method being numerical weather prediction (NWP). NWP calculates the approximate solution of partial differential equations involving atmospheric state variables when the initial conditions and boundary conditions of the atmosphere are known. In this way, some elements, such as the tracks and intensities of the TCs, can be solved iteratively; GRAPES-TYM (CMA), GFS (NCEP), and IFS (ECWMF) are the main NWP models. Although these model forecasts can provide accurate results, there are limitations in methods requiring numerous calculations, accurate mathematical descriptions of physical atmospheric mechanisms, and precise initial conditions. At the same time, ensemble forecast methods (GRAPES-GEFS, ECMWF-EPS, NCEP-GEFS) have been used to reduce the influence of various uncertainties on the numerical prediction results (Goerss, 2000). The other forecasting method is a statistical model, which generally utilizes multiple regression. The statistical model is mainly based on the relationship between the movement of the TC and its specific historical characteristics, but it usually does not consider any

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

physical processes. The National Hurricane Center has successively adopted statistical models such as NHC64 (taking observational data and historical 12h movements as factors), NHC67 (increasing factors based on NHC64), CLIPER (climate persistence factors) (Neumann and Hope, 1972), and NHC72 (a combination of NHC67 and CLIPER). Most traditional TC statistical models adopt a linear regression model, and it is difficult for this approach to address the nonlinear problems in TC track forecasting (Roy and Kovordányi, 2012). At the same time, manual feature selection is unable to produce accurate predictions.

Deep learning is an emerging application of supercomputing that is continuously being developed; many researchers have tried to adopt this technology to forecast weather and meteorological elements, including visibility (Ortega et al., 2022), wind speeds (Liu et al., 2018), radar echoes (Klein et al., 2015), and precipitation nowcasting (Shi et al., 2015). Deep learning is a statistical model that solves nonlinear and complex relationships from historical sample data based on neural network algorithms. The weight factor between network nodes is automatically adjusted through repeated training; thus, neural network algorithms have the advantages of strong adaptability and fault tolerance. TCs have complex dynamic mechanisms and are easily affected by many factors, including environmental steering flow, Beta effects, underlying surface conditions, the asymmetric structure of the inner core, and mesoscale circulations (Chan and Kepert, 2010). Artificial neural networks (ANNs) have been applied to predict TC tracks due to their strong learning ability and advantages in simulating nonlinear systems. Until the 2010s, ANN and back propagation (BP) networks were the mainstream neural network methods for forecasting TC tracks (Ali et al., 2007; Li-Min et al., 2009; Wang et al., 2011). Since the mid-2010s, more new methods have been introduced into TC prediction due to the development of deep-learning technology. Recurrent neural networks (RNNs) are suitable for TC track forecasting owing to their ability to handle time series data of arbitrary lengths. Moradi Kordmahalleh et al. (2015)applied a sparse RNN to Atlantic hurricane trajectory prediction using the dynamic time warping (DTW) method to measure the hurricane most similar to the target hurricane for training. Gao et al. (2018) used long short-term memory (LSTM) to predict typhoon tracks in the Northwest Pacific Ocean; the ratio of the cyclone training set and test set was set at 8:2, and the 24-h prediction error could reach 105 km. Alemany et al. (2018) proposed an RNN based on a grid system to predict hurricanes in the Atlantic, potentially improving the 6-h prediction accuracy with a root mean square

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

error (RMSE) of 0.11 for the test set. Kim et al. (2018) performed a TC identification task based on ConvLSTM to train WRF-simulated data, and the results are significantly better than those of a convolutional neural network (CNN). These CNNs have attracted attention given their suitability for processing 2D image data; they maintain spatial correlations by implementing convolution layers and then pooling layers for feature extraction. Giffard-Roisin et al. (2020) combined historical trajectory data with wind field reanalysis data as input to a CNN and predicted Atlantic hurricane tracks since 1979, with an average error of 32.9 km for 6-h predictions.

Making full use of different types of data is essential for deep learning. TC-related data are mainly divided into the following three categories: observational trajectory data, remote sensing data, and meteorological reanalysis field data. A multi-modal approach enables more accurate predictions than an approach using a single data source does. Zhang et al. (2018) developed a matrix neural network (MNN) model that preserves the spatial information of the TC tracks, and it has demonstrated the ability to provide more accurate results compared with other models (GRU, LSTM, MLP, and RNN). Ruttgers et al. (2019) built generation adversarial networks (GANs) adding satellite images to predict the coordinates of the typhoon center and generate cloud maps of future typhoons. Liu et al. (2022) proposed a new deep learning–based model, DBFNet, to effectively fuse the inherent features of cyclones and reanalyze 2D pressure field data. Previous studies have shown that deep-learning models that incorporate multiple data types can improve the track forecast of TCs to a certain extent. Still, most of them have neglected to describe and analyze the meteorological factors that affect the movement of TCs, ignoring valuable features.

Therefore, in this paper, we propose a new method for TC track prediction based on a combination of CNN and GRU models that incorporate data regarding the trajectory, steering airflow, sea-surface temperatures, and geopotential height as input features, aiming to improve the accuracy of TC track forecasts by leveraging big data. The main contents of this paper are as follows: Section 2 introduces the necessary data and methodology principles. Section 3 describes the experimental design and the framework of the fusion model (GRU_CNN) proposed in this paper. Section 4 presents the experimental results and comparative analysis, and Section 5 provides a summary and discussion of shortcomings and directions for future work.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

2 Data and methods

2.1 Data

The data used in this paper are trajectory data and reanalysis environmental data. The TC track data come from the International Best Track Archive for Climate Stewardship (IBTrACS), which encompasses all TCs globally. For each TC, the latitude, longitude, central pressure, maximum wind speed, direction, moving speed, and other data are recorded at 3-h intervals. The IBTrACS dataset contains data from different basins where cyclones show different characteristics; thus, this paper only selects TCs that occur in the Northwest Pacific Ocean. To better mine the hidden information, the 2D cyclone track data were chosen according to the method proposed by Li-Min et al. (2009), and 19 movement characteristics were obtained, including the past 24-h longitude, latitude, central atmospheric pressure, maximum wind speed, meridional moving speed, zonal moving speed, moving direction/speed, the difference between those values and those at the current time, and the angle, zonal distance, and meridional distance formed between the data over the past 24 h and in the present moment. In addition, the Coriolis parameter corresponding to the latitude of the past 24 h influences the geostrophic deflection force on the TCs.

Both observational and theoretical studies have shown that TC movement is closely related to large-scale airflow fields (Holland, 1983), and TC movement is mainly affected by the steering flow (Brand et al., 1981; Chan, 1984). Because they are influenced by the earth's rotation, TCs will be biased to the northwest (Kitade, 1981). Interactions among weather systems, the subtropical anticyclone, Westerlies, and the Tibetan High will also affect the movement of cyclones (George and Gray, 1976; Chan et al., 1980).

The geopotential heights of 300 hpa, 500 hpa, and 700 hpa are selected as the locations for the high, middle, and low-level circulation data, respectively. In addition, the underlying surface conditions must be considered, and, in the case of a weak guidance environment, TCs tend to move toward warmer sea-surface temperatures (Sun et al., 2017; Katsube and Inatsu, 2016). Meteorological environmental data are obtained by downloading high-resolution ERA5 reanalysis data from the European Centre for Medium-Range Forecasting (ECMWF). Holland (1984) noted that the deep mean

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

circulation from 850 hpa to 300 hpa can better represent the direction of a TC. Therefore, the environmental data for the preceding 24 h were extracted as follows:

(1) The *u*- and *v*-component data of the wind field on the four isobaric surfaces (300 hpa, 500 hpa, 700 hpa, 850 hpa): A 10° radius is extended from the center of the TC. Since the resolution of the selected reanalysis data is 1° × 1°, a 21 × 21 grid can be formed.

(2) The sea-surface temperature (SST): A 10° radius is again extended from the TC center to form a 21 × 21 grid.

(3) The geopotential heights of 300 hpa, 500 hpa, and 700 hpa: A grid is extended +35° to the north, −10° to the south, −40° to the west, and +40° to the east from the center of the TC, forming a 46 × 81 grid.

Because the actual weather circulation is very complex and includes information about the TC itself, the surrounding airflow, and the interaction between the two, it is necessary to separate the cyclone vortex from the surrounding airflow to obtain the steering flow. The most commonly used method (Lownam, 2001; Galarneau and Davis, 2013) corrects the vorticity and divergence by solving the change in the velocity stream and potential functions, respectively, and then calculates the modified velocity field. The modified flow field can be interpreted as having a non-rotating wind and non-diverging wind. There must be potential velocity in the irrotational motion and a stream function in the non-divergent motion. The relationship between them can be expressed as follows:

$$\nabla^2\psi = \zeta \tag{1}$$

$$v_\psi = \hat{\mathrm{k}} \times \nabla\psi \tag{2}$$

Where ψ is the stream function without divergence, ζ is the relative vorticity, and v_ψ is the non-divergent wind (rotating wind). To define the rotating wind, the vorticity outside the vortex radius is set to zero, and $\psi = 0$ is specified on the horizontal boundary. The iterative relaxation method is used to solve the stream function of Eq. (1) at all layers and then to calculate v_ψ using Eq. (2). In the case of divergence, Eqs. (1) and (2) are replaced by:

$$\nabla^2\chi = \delta \tag{3}$$

$$v_\chi = \nabla\chi \tag{4}$$

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

Where χ is the potential velocity, δ is the divergence, and v_χ is the non-vorticity wind. The divergence outside the vortex radius is set to zero and the potential function $\chi = 0$ on the boundary of the region. The velocity potential can be solved in the same manner to calculate v_χ. The ambient wind field with the vortex removed can be obtained by subtracting the rotating wind and divergent wind from the original wind field, V:

$$v_{env}(x,y,p) = V(x,y,p) - v_\chi(x,y,p) - v_\psi(x,y,p) \quad (5)$$

2.2 Methods

2.2.1 Random forest

Selecting features based on importance is a fundamental step in machine learning that most efficiently directs variables for machine learning models (Díaz-Uriarte and Alvarez De Andrés, 2006; Genuer et al., 2010). Random forest is a supervised learning method that selects the best feature combination and reduces the input feature dimension. The random forest contains N decision trees, and N is generally set to 100. Since bootstrapping (random sampling with replacement) is used to generate the random decision tree, all samples are not in the generation process of a tree, and the unused samples are called "out-of-bag" (OOB) samples. Through OOB samples, the accuracy of this tree can be evaluated, and it is usually calculated via the following four steps:

(1) Select the OOB data and calculate the OOB data error:

$$\text{error} = \frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2 \quad (6)$$

f_i: predicted value, y_i: actual value, n: number of samples (1/3 of the total).

(2) Randomly add noise interference to the features of all samples of the OOB data and calculate the OOB data error ($error^m$) again.

$$\text{error}^m = \frac{1}{n}\sum_{i=1}^{n}(f_i^m - y_i)^2 \quad (7)$$

f_m: Predicted value after adding noise to feature m.

(3) Calculate the importance *I* of all features.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

$$I = \frac{1}{N}\sum_{j=1}^{N}(\mathrm{error}_j^m - \mathrm{error}_j) \tag{8}$$

N: the number of decision trees, $error_j$: OOB data error, $error_j^m$: OOB data error after adding noise.

(4) According to the order of feature importance, delete the features with less importance in turn, repeat the above three steps, calculate the OOB data error, and select the combination with the lowest error. The OOB score represents the score of the model performance:

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

$$OOB_{score} = \frac{1}{N}\sum_{j=1}^{N} R_j \tag{10}$$

Before model training, it is necessary to determine whether the 19 trajectory features all have an impact on the prediction results. Figure 1(b) shows the 19 features' order of importance calculated using the random forest. For forecasting the difference in longitude and latitude within the following 72 hours, characteristics like the historic longitude or the angle formed by the historical moment and the current moment are significant. The decision about whether to exclude some less important features, however, requires further consideration. Based on Eq. (9) and Eq. (10), the OOB scores under different input feature dimensions are computed, with variables input in the order of importance, as shown in Fig. 1(c). In the case in which the first 11 features are sorted by importance, the OOB score is the highest, and the features added later will no longer affect the result; in other words, the best combination is that of the first 11 features.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

(a)

Short_name	Long_name
ANG	Angle formed by past and current moments
LON	Longitude of past moments
DIR_DIF	Direction difference between past moments and previous moments
MER_DIS	Meridional distance between past and current moments
LON_DIF	Difference of longitude between past and current moments
LAT	Latitude of past moments
MOV_SP	Moving speed difference between past and current moments
COR	Coriolis parameters of past moments
ZON_DIS	Zonal distance between past and current moments
PRE	Central atmospheric pressure of past moments
WIND_SP	Maximum wind speed of past moments
LAT_DIF	Difference of latitude between past and current moments
DIS	Moving distance between past and current moments
ZON_SP	Zonal moving speed of the past moment
MER_SP	Meridiona moving speed of the past moment
MOV_DIR	Moving direction of past moments
PRE_DIF	Atmospheric pressure difference between past and current moments
WIND_SP_DIF	Maximum wind speed difference between past and current moments
MOV_SP	Moving speed of past moments

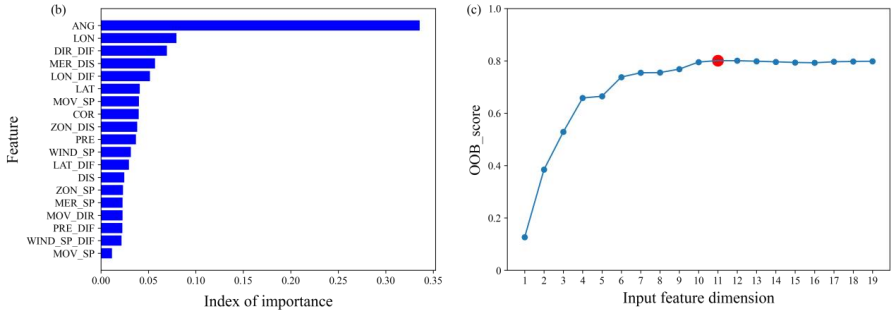


Figure 1: (a) Table displaying the short and long names of features, (b) the importance index of features, and (c) the OOB_score of different feature combinations based on the random forest (red dot indicates the maximum value).

2.2.2 Recurrent Neural Network

An RNN is an ANN characterized by architectural features intentionally designed to preserve historical information, showing a remarkable ability to process sequential data. RNNs can process sequences of any length using neurons with self-feedback, which has been successfully applied in many fields, such as speech recognition (Graves et al., 2013), stock market predictions (Bathla, 2020), and trajectory predictions (Wang and Fu, 2020). However, simple RNNs have difficulty in dealing with the long-term dependence of the sequence; when the sequence length exceeds a certain threshold, the information may disappear during the transmission process, resulting in large deviations in prediction accuracy. The LSTM network proposed by Hochreiter and Schmidhuber (1997) can avoid the gradient disappearance and explosion phenomena that occur in the standard RNN. While GRU (Cho et al., 2014) is an improved and optimized neural network based on LSTM, it has a faster convergence speed and maintains accuracy levels close to those of LSTM.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

An LSTM layer consists of multiple recurrently connected blocks called memory blocks. Each block has three multiplication units: an input gate, an output gate, and a forget gate. These three gates perform their duties in a manner similar to filters. The input gate updates the state of the memory cell, the forget gate selects relevant information from the previous memory cell, and the output gate controls the output of each LSTM unit. The value of the hidden layer h_t (Eq. 16) depends not only on the input variable x_t at the current moment but also on the hidden layer at the previous moment h_{t-1}. The forward propagation formula of LSTM is as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t]) \tag{11}$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t]) \tag{12}$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t]) \tag{13}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{14}$$

$$o_t = \sigma(W_i \cdot [h_{t-1}, x_t]) \tag{15}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{16}$$

Where σ and tanh are the sigmoid function and hyperbolic tangent activation function, respectively; i_t, f_t, and O_t are the input gate (Eq. 11), forget gate (Eq. 13), and output gate (Eq. 15), respectively; and C_t (Eq. 14) represents the cell state, the sizes of which are equal to the hidden vectors h_{t-1}. The weight matrices W_c, W_f, and W_i constitute the internal parameters of the neural network optimized by the back-propagation algorithm.

The principle of GRU is very similar to that of LSTM. GRU merges the forget gate and the input gate in LSTM to become the update gate z_t (Eq. 17), and the output gate is also named the reset gate r_t (Eq. 18). Intuitively, the reset gate controls the information from the current moment to the memory unit $\tilde{h}_t$ (Eq. 19), and the update gate determines the amount of previous memory information saved to the current time.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \tag{17}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \tag{18}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t \cdot h_{t-1}, x_t]) \tag{19}$$

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

$$h_t = (1 - z_t)h_{t-1} + z_t \cdot \tilde{h}_t \tag{20}$$

2.2.3 Convolutional Neural Network

A CNN is a hierarchical structure divided into convolution, pooling, and fully connected layers. CNN can extract features automatically by processing the input patterns and translating the same convolution kernel from top to bottom and from left to right. The spatial relationship is fixed with the distribution of neurons, and the local connection and weight sharing of neurons reduce the training complexity by reducing the number of parameters.

Lecun et al. (1998) first used CNN for handwritten character recognition with average pooling and the tanh activation function. Krizhevsky et al. (2012) proposed the AlexNet model in the ImageNet competition, using the ReLU function instead of the traditional tanh function to introduce nonlinearity and solve the gradient disappearance problem of the activation function when the network was relatively deep, employing maximum pooling to avoid the blurring effect of average pooling. Ioffe and Szegedy (2015) applied batch normalization to image classification models, which significantly accelerated the training of deep networks, and batch normalization helped alleviate the problem of gradient exploding or vanishing.

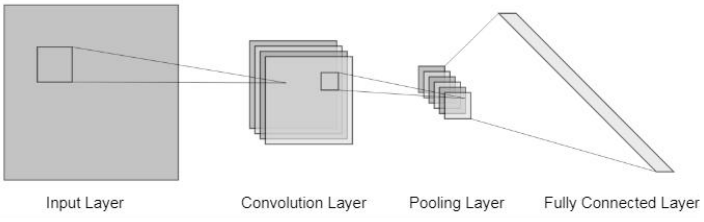


Figure 2: Diagram showing the typical structure of a CNN.

As shown in Fig. 2, all units in the same feature map share the same convolution kernel. The role of the pooling layer is to improve the operating efficiency of the algorithm and expand the perception range so that the subsequent convolution kernels can learn more global information. Typical operations include the average and maximum pooling values. The fully connected layer is usually located at the back end of the neural network, where all the features obtained from the hidden layer are expanded into a vector connected to the neurons of the output layer.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

3 Experiment

3.1 Experimental design

Our goal is to predict the TC movements for the following 6–72 h using the trajectory data and surrounding environmental field from the previous 24 h. We explore TC movement in the Northwest Pacific from 1979 to 2021 and consider the longitudinal and latitudinal changes in the following 6–72 h as the quantitative prediction variables, with the center of the TC at the current time as the reference point. Since the maximum forecast hour is 72 and the input sequence time length is 24 h, TCs that persist for longer than 96 hours are removed. All samples obtained based on the sliding window of the input-prediction sequence length are divided into three groups in chronological order: training set (1979–2014), validation set (2015–2018), and test set (2019–2021). There are 36473 samples, of which 90% are trained, and the remaining 10% are validated; 49 TCs from 2019 to 2021 are used for testing, and the number of test samples is 2095.

3.2 Model framework

Due to differences in the data sources, a new model must be developed to integrate the four information sources into the neural network using the Keras deep-learning framework. The specific model structure is shown in Fig. 3. For the 3D meteorological data, the data are superimposed on the geopotential height (pressure level), so the input data for the CNN consist of multiple two-dimensional matrices. The TimeDistributed layer is applied to a series of tensors in the processing of the time dimension. In addition, the CNN adopts a typical architecture with alternating convolution layers (Conv layers) and maximum pooling layers (Maxpool layers), converting two-dimensional data into one-dimensional vectors at the end of the CNN network and eventually creating a fully connected layer. All hidden layers are equipped with batch normalization, and this paper employs ReLU as the activation function.

For the two-dimensional trajectory data of the TCs, x_t^j represents the input value of the i^{th} feature at the j^{th} timestamp, $i \in (1, n)$, $j \in (1, t)$, and they are input into GRU. The model is based on the Adam optimizer and trained with the RMSE between the forecast and the actual value as a loss function. Due to the different properties among the wind field, pressure field, SSTs, and past trajectory data, different

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

learning rates are required for the neural network. Therefore, the parameters of each branch in the model can be trained with the same task, and then the branches can be fused into one network and stitched with a fully connected layer; thereafter, the parameters can be adjusted slightly.

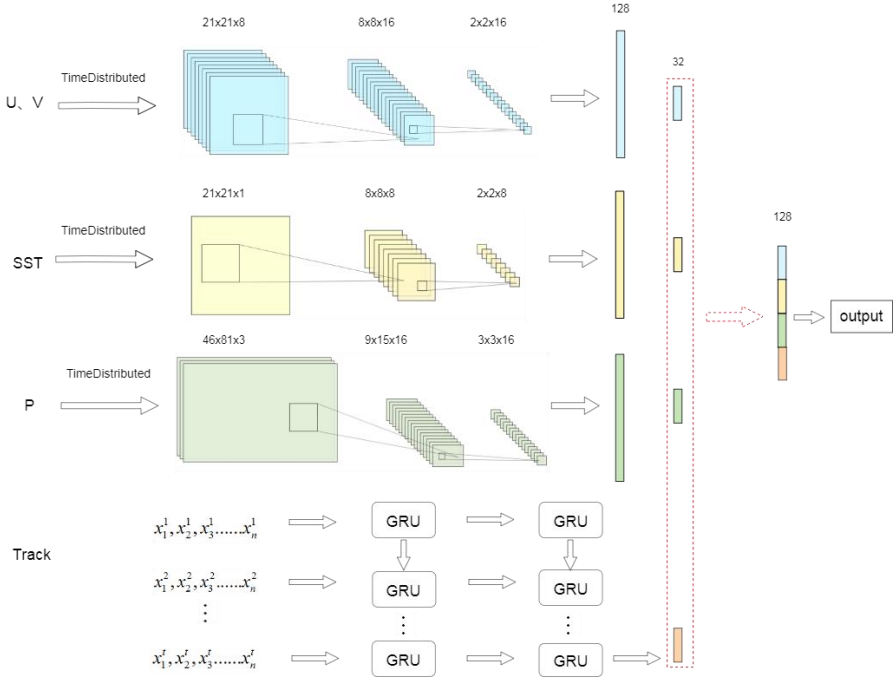


Figure 3: The model framework and network structure of GRU_CNN.

3.3 Data normalization

The differences in the dimensions of the input data will cause a variable with a larger value to have a more significant impact on the model. Therefore, it is necessary to normalize the data before model training to map the input data from 0 to 1. To eliminate the influence of different dimensions on the model, the original value x is normalized as x':

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{21}$$

Where x_{max} and x_{min} are the maximum and minimum values of the variable x, respectively.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

3.4 Evaluation criteria

RMSE is the root mean square error obtained by calculating the predicted value P_i and the observed value O_i. The formula is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{n}} \tag{22}$$

Since the latitude and longitude represent different spatial distances in kilometers, when comparing other models, the distance error (Dis) is calculated in kilometers. The formula is as follows:

$$Dis = R \times arccos(\cos(Lat_{pred})\cos(Lat_{obs}) \times cos(Lon_{pred} - Lon_{pred}) + \sin(Lat_{obs}) \times \sin(Lat_{pred})) \tag{23}$$

Where R is the radius of the earth, Lat_{obs} and Lon_{real} are the actual latitude and longitude, and Lat_{obs} and Lon_{pred} are the predicted latitude and longitude.

4 Results

Three types of recurrent neural networks (RNN, LSTM, GRU) are used to train samples with eight timestamps and 11 features selected by the random forest method, according to their importance; the results of analyzing 49 TCs in 2019–2021 are then evaluated. We set the value of the batch size to 64 and the epoch to 100 and found that the model performed best when the number of neurons in the hidden layer is set to 128; this was determined via experiments using different numbers of neurons in the hidden layer. Early stopping is used to prevent overfitting. When the performance of the model in the validation set begins to decline, training is stopped to avoid overfitting due to continued training. Figure 4 shows the predicted latitudes and longitudes based on three RNNs in the following 6 h, 12 h, 18 h, 24 h, 48 h, and 72 h, and the results of all three networks within 24 h are all approximately consistent with the real data. As the forecast time increases, the error accumulates.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

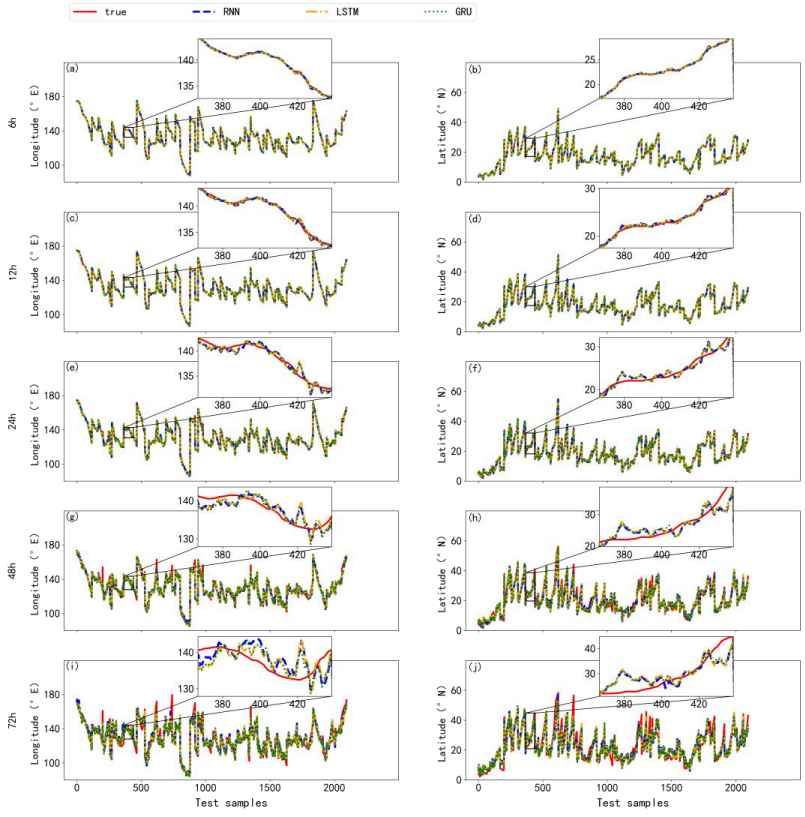


Figure 4: Comparing the longitudes and latitudes of all test samples predicted by the three recurrent neural networks (RNN, LSTM, GRU), including the longitude/latitude forecasts at (a–b) 6 h, (c–d) 12 h, (e–f) 24 h, (g–h) 48 h, and (i–j) 72 h.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

Table 1 Model performance evaluation (RMSE) for RNN, LSTM, and GRU

		Lat						Lon					
Forecast Hour		6 h	12 h	18 h	24 h	48 h	72 h	6 h	12 h	18 h	24 h	48 h	72 h
Training Sets	RNN	0.146	0.368	0.624	0.912	2.232	4.069	0.153	0.412	0.735	1.118	3.181	6.236
	LSTM	0.126	0.335	0.584	0.867	2.34	3.986	0.149	0.391	0.703	1.077	3.232	6.172
	GRU	**0.112**	**0.312**	**0.555**	**0.83**	**2.282**	**3.883**	**0.134**	**0.376**	**0.681**	**1.041**	**3.152**	**6.031**
Validation Sets	RNN	0.171	0.441	0.76	1.116	2.723	4.493	0.182	0.483	0.865	1.337	3.739	6.715
	LSTM	0.157	0.428	0.751	1.103	**2.675**	**4.349**	0.183	**0.480**	**0.855**	**1.305**	3.69	6.761
	GRU	**0.157**	**0.42**	**0.735**	**1.086**	2.699	4.434	**0.179**	0.484	0.868	1.33	**3.632**	**6.608**
Test Sets	RNN	0.166	0.411	0.685	0.988	2.481	4.207	0.176	0.461	0.797	1.178	3.534	6.157
	LSTM	0.149	0.389	0.661	0.965	2.473	4.154	0.169	**0.456**	0.812	1.215	3.346	5.989
	GRU	**0.149**	**0.387**	**0.653**	**0.951**	**2.446**	**4.143**	**0.167**	0.457	**0.8**	**1.185**	**3.325**	**5.969**

The performance evaluation of the three RNN models is displayed in Table 1 by calculating the RMSE values between the predicted longitude (latitude) and the actual longitude (latitude), including training, validation, and test sets; the best results are highlighted in bold font. It is clear that the GRU-based and LSTM-based models significantly outperformed the RNN-based model, which suggests that the RNN is inferior in handling the problem of long-term dependence. GRU is a variant of LSTM that combines the forget and input gates in LSTM into an update gate and also merges the cell and hidden states. Hence, the parameter amounts of GRU are less than those of LSTM, which results in the overall training speed of GRU being faster than that of LSTM. GRU is theoretically similar to LSTM and can achieve the same accuracy as LSTM (or even better), so the results of GRU and LSTM are close and their RMSE values are much lower than that of RNN. GRU achieves the best performance in all forecast hours, with the smallest RMSE in the test set. Therefore we use GRU as a part of the fusion network model called GRU_CNN, adding meteorological environment data processed with CNN.

Table 2 Comparison of the average absolute distance errors (km) predicted by multiple deep-learning models

	6 h	12 h	18 h	24 h	48 h	72 h
CLIPER (Demaraia, 1992)	—	—	—	213	442	659
BP	23.86	59.58	101.01	146.91	377.64	634.42
RNN	21.43	55.46	94.59	138.09	373.12	625.17
LSTM	19.65	52.38	91.76	136.05	360.32	614.76
GRU	19.51	52.6	91.21	134.73	357.25	607.44
NMSTN (Huang, 2022)	27.52	59.09	—	139.18	336.16	544.16
GRU_CNN	17.22	43.9	72.74	106.16	281.52	502.71

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

Table 2 compares the results between GRU_CNN and various deep-learning models, showing the forecast results in the form of the mean absolute distance error. It is evident that GRU_CNN presents an absolute advantage in long-term forecasting. Both LSTM and GRU retain important features through various gate functions, which ensures that they will not be lost during long-term propagation. They can better predict the medium and long-term tracks of the TCs, compared with standard RNNs and two traditional methods named CLIPER and BP. The GRU_CNN is more accurate than the models without CNN. The average distance errors at 6 h, 24 h, 48 h, and 72 h are 17.22 km, 106.16 km, 281.52 km, and 502.71 km, respectively. The error is also reduced compared with the NMSTN method proposed by Huang et al. (2022). In addition, although there is a big difference between the long-term forecast and the numerical prediction results, the average distance prediction results are better than the results provided by the Central Meteorological Observatory (CMO) in the short-term forecasts, including the 6 h (27.57 km) and 12 h (59.09 km) forecasts.

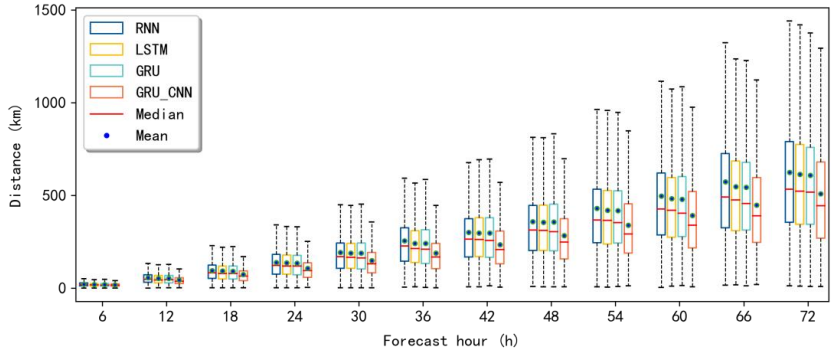


Figure 5: The absolute average distance boxplot of the three kinds of recurrent neural networks (RNN, LSTM, GRU) and the method in this paper (GRU_CNN) creating 6–72 h forecasts (interval 6 h).

As shown in Fig. 5, the maximum distance errors predicted by the three RNNs at 48 h and 72 h are over 500 km and 1000 km, respectively. Only considering the trajectory characteristics of the TCs in the RNN while ignoring the external atmospheric environmental characteristics will cause instability in the prediction of the TC tracks. The errors of the maximum and average values predicted by the GRU_CNN model are both significantly reduced. To illustrate GRU_CNN more comprehensively and intuitively, Fig. 6 shows a scatter plot of the predicted and actual values. The distance between the data points and the diagonal line represents the prediction error. The higher the wind speed, the stronger the

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

intensity of the TCs, and the closer the predicted value is to the actual value. In addition, with the increase in the forecast time, in high latitude and longitude forecasts when the TC is moving towards the northwest, the predicted value is often lower than the actual value.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

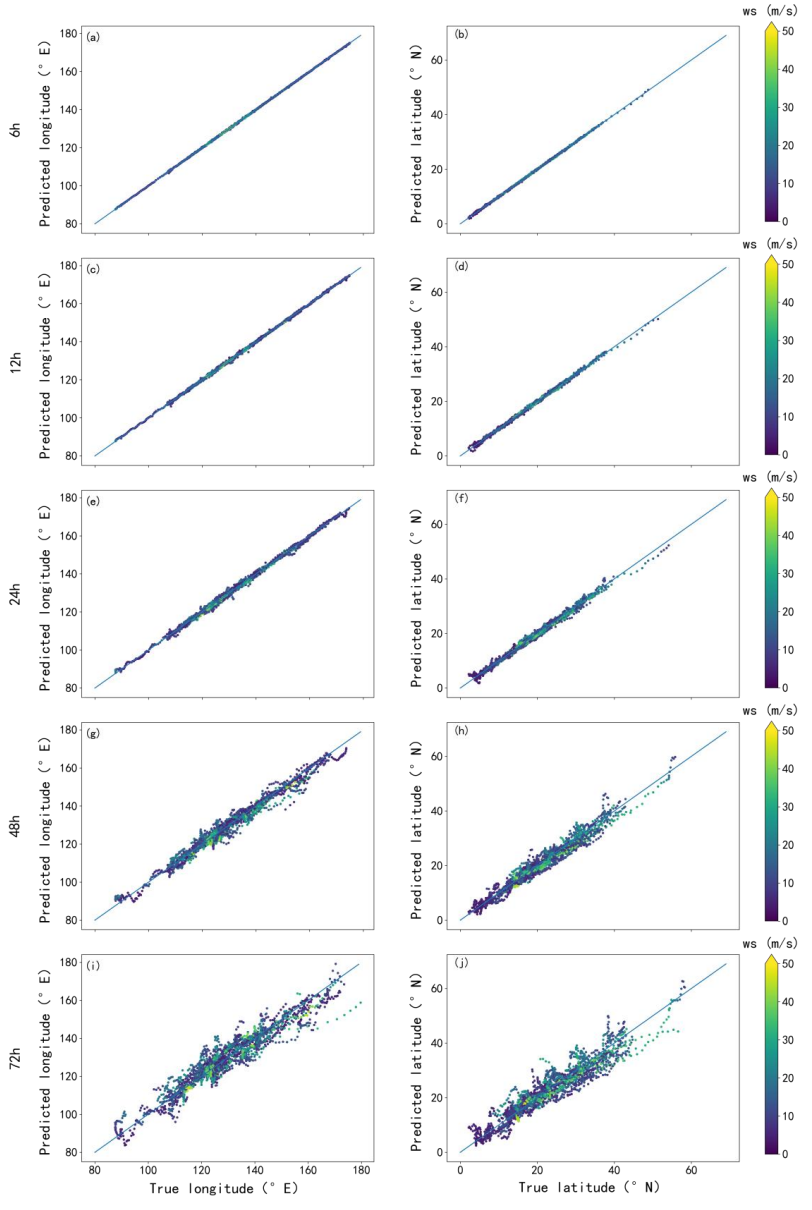


Figure 6: Scatter plot distributions of latitude predictions. The color bar represents the maximum wind speed, including the longitude and latitude forecasts at (a–b) 6 h, (c–d) 12 h, (e–f) 24 h, (g–h) 48 h, and (i–j) 72 h.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

Data from three environmental fields are used in this paper: SST, geopotential height (pressure), and wind field (*u*- and *v*-component) data. Different environmental input variables show different effects in the model (Table 3). GRU+SST (pressure, UV) represents only the combination of the trajectory characteristics and SST (geopotential height, wind field), while GRU+CNN is the result of the fusion of the three. The results in Table 3 indicate that GRU+UV performed best, followed by GRU+pressure and then GRU+SST, indicating that the steering flow plays a dominant role in TC forecasting, especially in the short-term < 24-h forecast. The forecasting results of adding only the steering flow are close to those of GRU_CNN, while the results at 48 h and 72 h illustrate that the influence of the SST and geopotential height on the long-term TC forecast track gradually increases.

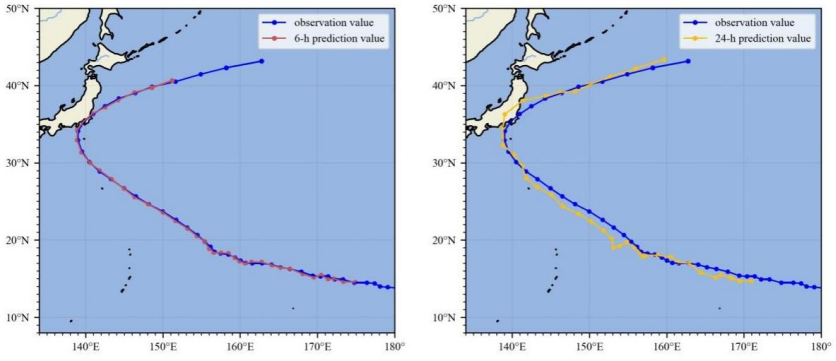
Table 3 Comparison of trajectory data combining different environmental features. RMSE is the root mean square error of latitude and longitude, and the distance is the average absolute distance error (km).

	GRU+SST		GRU+Pressure		GRU+UV		GRU+CNN	
	RMSE	Distance (km)	RMSE	Distance (km)	RMSE	Distance (km)	RMSE	Distance (km)
6 h	0.154	19.35	0.132	16.15	0.137	16.94	0.138	17.22
12 h	0.419	52.37	0.352	44.22	0.347	43.74	0.35	43.9
18 h	0.739	92.25	0.598	76.16	0.575	72.81	0.575	72.74
24 h	1.103	137.78	0.883	112.93	0.841	106.63	0.837	106.16
48 h	2.858	358.25	2.462	306.03	2.379	302.76	2.248	281.52
72 h	4.913	588.08	4.52	557.86	4.385	524.88	4.146	502.71

To better show the model forecast of GRU_CNN, Figures 7–9 present the observed and forecast tracks at 6 h and 24 h of TCs FAXAI, MITAG, and IN-FA, respectively, and the forecast tracks of other TCs in the test set are presented in Supplementary Fig. S1-51. The blue lines represent the observed tracks, while the red and yellow lines indicate the 6-h and 24-h forecast tracks. In general, it is particularly hard to forecast unexpected turns in the TC track. The three TCs shown all exhibit a sudden northward or northwestern turn in the TC track. For the 6-h forecast, the predicted path is approximately consistent with the actual track, while the 24-h forecast has some deviations. The average distance predicted near the northwest turn of FAXAI is 91.35 km; the error for MITAG's first turn to the north is 127.02 km, and the error for the second turn to the northwest is 121.91 km. The two average errors in the track forecast for In-fa are 84.27 km and 82.37 km. It can be seen that there is no significant deviation in the forecast around the steering point, but, for some abnormal track changes,

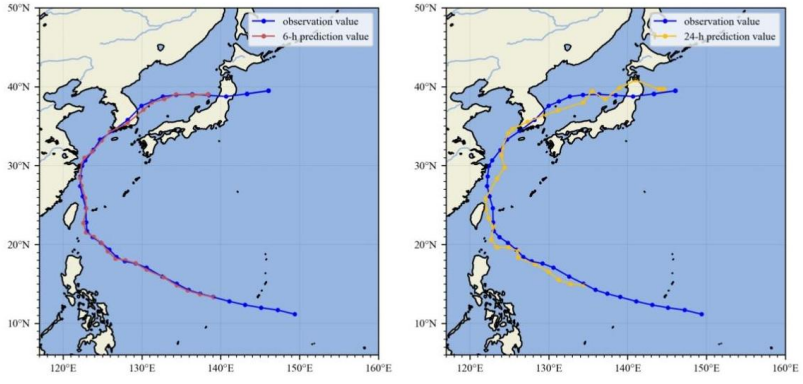
https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

407 such as crossing back over the same location, samples with more significant errors will be generated,
408 reducing the overall average absolute distance error.



409

410 **Figure 7: Forecast tracks of Tropical Cyclone FAXAI (1915) (left: 6 h, right: 24 h).**



411

412 **Figure 8: Forecast tracks of Tropical Cyclone MITAG (1918) (left: 6 h, right: 24 h).**

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

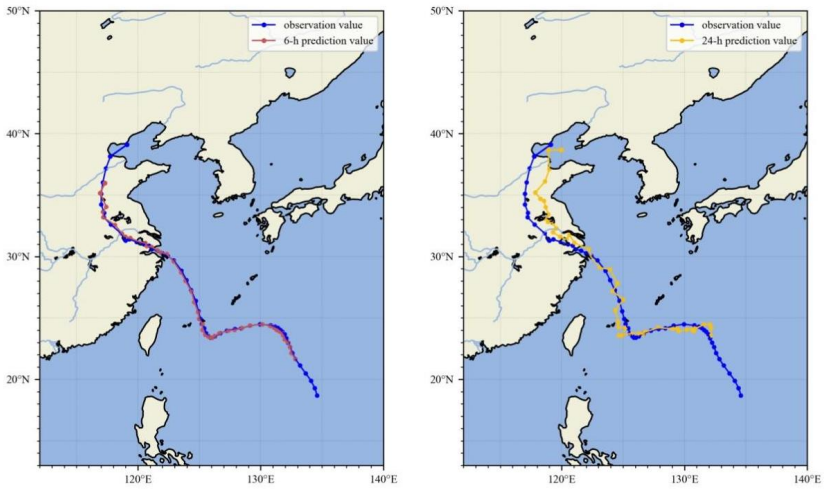


Figure 9: Forecast tracks of Tropical Cyclone IN-FA (2106) (left: 6 h, right: 24 h).

5 Conclusion

The past 24-h TC trajectory and meteorological field data have been used to forecast TC tracks in the Northwest Pacific from hours 6–72. First, in order to eliminate data redundancy and reduce the complexity of the prediction model, the random forest algorithm was used for feature extraction of the two-dimensional movement data. Second, three kinds of recurrent neural networks (RNN, LSTM, GRU) were used to evaluate and compare the models based on the input of trajectory features, and it was concluded that GRU performed relatively better in predicting TC tracks. Eventually, we combined GRU with CNN by adding the pre-processed meteorological environmental data around the cyclones (removing the vortex to obtain the steering flow); the CNN models the selected meteorological variables and extracts features, while GRU processes trajectory sequences. GRU_CNN has better prediction results than traditional single deep-learning methods do.

When a new TC generates in the ocean, the GRU_CNN model can quickly provide the forecast track within seconds. Short-term predictions within 12 h of initialization can provide better results than CMO can, and the average distance errors of the forecasts at 6 h and 12 h are 17.22 km and 43.9 km. When the forecast goes beyond 24 h, the model's accuracy declines. The historical steering flow of

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

cyclones has a significant effect on improving the accuracy of short-term forecasting, while, in long-term forecasting, the SST and geopotential height will have a particular impact, which is regarded as a crucial way to expand and improve the application of deep-learning models in TC track forecasting, In addition, the model can accurately predict TCs that suddenly turn to the north or northwest, but there will be a considerable distance error for abnormal trajectories, possibly due to a lack of synoptic analysis in our study.

Cyclone prediction has been a challenge in weather forecasting for a long time. With future scientific and technological advances, it is becoming increasingly convenient to obtain meteorological data, and the database has gradually expanded. At the same time, deep-learning models are flexible and can easily be expanded upon. In the future, more data can be integrated, and more valuable features can be extracted to improve the prediction accuracy of the deep-learning model. In addition, model predictor variables will be considered in future work, the inclusion of which can predict more useful information, such as cyclone intensity, rainfall, and wind speed.

Code availability. The code and model are available as a free access repository on GitHub at https://github.com/Hush980/TCs_DL_code.

Data availability. IBTrACS that we used in this study is publicly available. It can be down-loaded at https://www.ncei.noaa.gov/data/international-best-track-archive-for-Climate-stewardship-ibtracs.
ERA5 data can be obtained from Copernicus Climate Data Store (https://cds.climate.copernicus.eu).

Author contributions. Liang Wang wrote the paper and conducted most of the code implementation and data analysis. Bincheng Wan designed the research framework. Shaohui Zhou provided the code and revised the paper. Haofei Sun was involved in data collation and Zhiqiu Gao was responsible for supervision.

Competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

Acknowledgments. This study was supported by the National Key Research and Development Program of the Ministry of Science and Technology of China (2018YFC1506405), and by the National Natural Science Foundation of China (Grants 42175082 and 42222503).

References

Alemany, S., Beltran, J., Perez, A., and Ganzfried, S.: Predicting Hurricane Trajectories Using a Recurrent Neural Network, Proceedings of the AAAI Conference on Artificial Intelligence, 33, https://doi.org/10.1609/aaai.v33i01.3301468, 2018.

Ali, M. M., Kishtawal, C. M., and Jain, S.: Predicting cyclone tracks in the north Indian Ocean: An artificial neural network approach, Geophysical Research Letters, 34, https://doi.org/10.1029/2006gl028353, 2007.

Bathla, G.: Stock Price prediction using LSTM and SVR, 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), 6-8 Nov. 2020, 211-214, https://doi.org/10.1109/PDGC50313.2020.9315800,

Brand, S., Buenafe, C. A., and Hamilton, H. D.: Comparison of Tropical Cyclone Motion and Environmental Steering, Monthly Weather Review, 109, 908-909, https://doi.org/10.1175/1520-0493(1981)109<0908:cotcma>2.0.co;2, 1981.

Chan, J. and Kepert, J.: Global Perspectives on Tropical Cyclones: From Science to Mitigation, https://doi.org/10.1142/7597, 2010.

Chan, J. C.-L.: An Observational Study of the Physical Processes Responsible for Tropical Cyclone Motion, Journal of Atmospheric Sciences, 41, 1036-1048, https://doi.org/10.1175/1520-0469(1984)041<1036:aosotp>2.0.co;2, 1984.

Chan, J. C. L., Gray, W. M., and Kidder, S. Q.: Forecasting Tropical Cyclone Turning Motion from Surrounding Wind and Temperature Fields, Monthly Weather Review, 108, 778-792, https://doi.org/10.1175/1520-0493(1980)108<0778:FTCTMF>2.0.CO;2, 1980.

Cho, K., Merriënboer, B. v., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y.: Learning Phrase Representations using RNN Encoder--Decoder for Statistical Machine

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

480 Translation, In Proceedings of the 2014 Conference on Empirical Methods in Natural Language
481 Processing (EMNLP), https://doi.org/10.3115/v1/D14-1179, 2014.
482 Díaz-Uriarte, R. and Alvarez de Andrés, S.: Gene selection and classification of microarray data using
483 random forest, BMC bioinformatics, 7, 3, https://doi.org/10.1186/1471-2105-7-3, 2006.
484 Emanuel, K.: Will Global Warming Make Hurricane Forecasting More Difficult?, Bulletin of the
485 American Meteorological Society, 98, 495-501, https://doi.org/10.1175/bams-d-16-0134.1, 2017.
486 Galarneau, T. J. and Davis, C. A.: Diagnosing Forecast Errors in Tropical Cyclone Motion, Monthly
487 Weather Review, 141, 405-430, https://doi.org/10.1175/mwr-d-12-00071.1, 2013.
488 Gao, S., Zhao, P., Pan, B., Li, Y., Zhou, M., Xu, J., Zhong, S., and Shi, Z.: A nowcasting model for the
489 prediction of typhoon tracks based on a long short term memory neural network, Acta
490 Oceanologica Sinica, 37, 8-12, https://doi.org/10.1007/s13131-018-1219-z, 2018.
491 Genuer, R., Poggi, J.-M., and Tuleau-Malot, C.: Variable selection using random forests, Pattern
492 Recognition Letters, 31, 2225-2236, https://doi.org/10.1016/j.patrec.2010.03.014, 2010.
493 George, J. E. and Gray, W. M.: Tropical Cyclone Motion and Surrounding Parameter Relationships,
494 Journal of Applied Meteorology and Climatology, 15, 1252-1264,
495 https://doi.org/10.1175/1520-0450(1976)015<1252:TCMASP>2.0.CO;2, 1976.
496 Giffard-Roisin, S., Yang, M., Charpiat, G., Kumler Bonfanti, C., Kégl, B., and Monteleoni, C.:
497 Tropical Cyclone Track Forecasting Using Fused Deep Learning From Aligned Reanalysis Data,
498 Front Big Data, 3, 1, https://doi.org/10.3389/fdata.2020.00001, 2020.
499 Goerss, J. S.: Tropical Cyclone Track Forecasts Using an Ensemble of Dynamical Models, Monthly
500 Weather Review, 128, 1187-1193,
501 https://doi.org/10.1175/1520-0493(2000)128<1187:tctfua>2.0.co;2, 2000.
502 Goldenberg, S. B., Landsea, C. W., Mestas-Nuñez, A. M., and Gray, W. M.: The Recent Increase in
503 Atlantic Hurricane Activity: Causes and Implications, Science, 293, 474-479,
504 https://doi.org/doi:10.1126/science.1060040, 2001.
505 Graves, A., Jaitly, N., and Mohamed, A.: Hybrid speech recognition with Deep Bidirectional LSTM,
506 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, 8-12 Dec. 2013,
507 273-278, https://doi.org/10.1109/ASRU.2013.6707742,

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

508 Gray, W. M.: Global View of the Origin of Tropical Disturbances and Storms, Monthly Weather
509 Review, 96, 669-700, http://dx.doi.org/10.1175/1520-0493(1968)096<0669:GVOTOO>2.0.CO;2,
510 1968.
511 Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural computation, 9, 1735-1780,
512 https://doi.org/10.1162/neco.1997.9.8.1735, 1997.
513 Holland, G. J.: Tropical Cyclone Motion: Environmental Interaction Plus a Beta Effect, Journal of the
514 Atmospheric Sciences, 40, 328-342,
515 https://doi.org/10.1175/1520-0469(1983)040<0328:tcmeip>2.0.co;2, 1983.
516 Holland, G. J.: Tropical Cyclone Motion. A Comparison of Theory and Observation, Journal of
517 Atmospheric Sciences, 41, 68-75,
518 https://doi.org/10.1175/1520-0469(1984)041<0068:tcmaco>2.0.co;2, 1984.
519 Huang, C., Bai, C., Chan, S., and Zhang, J.: MMSTN: A Multi - Modal Spatial - Temporal Network
520 for Tropical Cyclone Short - Term Prediction, Geophysical Research Letters, 49,
521 https://doi.org/10.1029/2021gl096898, 2022.
522 Ioffe, S. and Szegedy, C.: Batch Normalization: Accelerating Deep Network Training by Reducing
523 Internal Covariate Shift, Proceedings of the 32nd International Conference on International
524 Conference on Machine Learning, 37, 448-456, https://doi.org/10.48550/arXiv.1502.03167, 2015.
525 Katsube, K. and Inatsu, M.: Response of Tropical Cyclone Tracks to Sea Surface Temperature in the
526 Western North Pacific, Journal of Climate, 29, 1955-1975,
527 https://doi.org/10.1175/jcli-d-15-0198.1, 2016.
528 Kim, S., Kang, J.-S., Lee, M., and Song, S.-k.: DeepTC: ConvLSTM Network for Trajectory Prediction
529 of Tropical Cyclone using Spatiotemporal Atmospheric Simulation Data, In Spatiotemporal
530 Workshop at 31st Conference on Neural Information Processing Systems, 2018.
531 Kitade, T.: A Numerical Study of the Vortex Motion with Barotropic Models, Journal of the
532 Meteorological Society of Japan. Ser. II, 59, 801-807, https://doi.org/10.2151/jmsj1965.59.6_801,
533 1981.
534 Klein, B., Wolf, L., and Afek, Y.: A Dynamic Convolutional Layer for short rangeweather prediction,
535 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7-12 June 2015,
536 4840-4848, https://doi.org/10.1109/CVPR.2015.7299117,

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

537 Krizhevsky, A., Sutskever, I., and Hinton, G.: ImageNet Classification with Deep Convolutional
538 Neural Networks, Neural Information Processing Systems, 25, https://doi.org/10.1145/3065386,
539 2012.
540 Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P.: Gradient-Based Learning Applied to Document
541 Recognition, Proceedings of the IEEE, 86, 2278-2324, https://doi.org/10.1109/5.726791, 1998.
542 Li-min, S., Gang, F. U., Xiang-chun, C., and Jian, Z.: Application of BP neural network to forecasting
543 typhoon tracks, Journal of Natural Disasters, 18, 104-111,
544 https://doi.org/0.3969/j.issn.1004-4574.2009.06.018, 2009.
545 Liu, H., Mi, X., and Li, Y.: Smart multi-step deep learning model for wind speed forecasting based on
546 variational mode decomposition, singular spectrum analysis, LSTM network and ELM, Energy
547 Conversion and Management, 159, 54-64, https://doi.org/10.1016/j.enconman.2018.01.010, 2018.
548 Liu, Z., Hao, K., Geng, X., Zou, Z., and Shi, Z.: Dual-Branched Spatio-Temporal Fusion Network for
549 Multihorizon Tropical Cyclone Track Forecast, IEEE Journal of Selected Topics in Applied Earth
550 Observations and Remote Sensing, 15, 3842-3852, https://doi.org/10.1109/JSTARS.2022.3170299,
551 2022.
552 Lownam, C.: The NCAR-AFWA tropical cyclone bogussing scheme. A report prepared for the Air
553 Force Weather Agency (AFWA), National Center for Atmospheric Research, 2001.
554 Moradi Kordmahalleh, M., Gorji Sefidmazgi, M., Homaifar, A., and Liess, S.: Hurricane Trajectory
555 Prediction Via a Sparse Recurrent Neural Network, 2015.
556 NEUMANN, C. J. and HOPE, J. R.: Performance Analysis of the HURRAN Tropical Cyclone
557 Forecast System, Monthly Weather Review, 100, 245-255,
558 https://doi.org/10.1175/1520-0493(1972)100<0245:paotht>2.3.co;2, 1972.
559 Ortega, L. C., Otero, L. D., Solomon, M., Otero, C. E., and Fabregas, A.: Deep learning models for
560 visibility forecasting using climatological data, International Journal of Forecasting,
561 https://doi.org/10.1016/j.ijforecast.2022.03.009, 2022.
562 Roy, C. and Kovordányi, R.: Tropical cyclone track forecasting techniques — A review, Atmospheric
563 Research, 104-105, 40-69, https://doi.org/10.1016/j.atmosres.2011.09.012, 2012.

https://doi.org/10.5194/egusphere-2022-1216
Preprint. Discussion started: 21 November 2022

564 Ruttgers, M., Lee, S., Jeon, S., and You, D.: Prediction of a typhoon track using a generative
565 adversarial network and satellite images, Sci Rep, 9, 6057,
566 https://doi.org/10.1038/s41598-019-42339-y, 2019.
567 Schulthess, T. C., Bauer, P., Wedi, N., Fuhrer, O., Hoefler, T., and Schär, C.: Reflecting on the Goal
568 and Baseline for Exascale Computing: A Roadmap Based on Weather and Climate Simulations,
569 Computing in Science & Engineering, 21, 30-41, https://doi.org/10.1109/MCSE.2018.2888788,
570 2019.
571 Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W. K., and Woo, W.-c.: Convolutional LSTM
572 Network: A Machine Learning Approach for Precipitation Nowcasting, Proceedings of the 28th
573 International Conference on Neural Information Processing Systems, 1, 802-810,
574 https://doi.org/10.48550/arXiv.1506.04214, 2015.
575 Sun, Y., Zhong, Z., Li, T., Yi, L., Camargo, S. J., Hu, Y., Liu, K., Chen, H., Liao, Q., and Shi, J.:
576 Impact of ocean warming on tropical cyclone track over the western north pacific: A numerical
577 investigation based on two case studies, Journal of Geophysical Research: Atmospheres, 122,
578 8617-8630, https://doi.org/10.1002/2017JD026959, 2017.
579 Wang, C. and Fu, Y.: Ship Trajectory Prediction Based on Attention in Bidirectional Recurrent Neural
580 Networks, 529-533, https://doi.org/10.1109/isctt51595.2020.00100, 2020.
581 Wang, Y., Zhang, W., and Fu, W.: Back Propogation(BP)-neural network for tropical cyclone track
582 forecast, Proceedings - 2011 19th International Conference on Geoinformatics, Geoinformatics
583 2011, 1-4, https://doi.org/10.1109/GeoInformatics.2011.5981095, 2011.
584 Zhang, Y., Chandra, R., and Gao, J.: Cyclone Track Prediction with Matrix Neural Networks, 2018
585 International Joint Conference on Neural Networks (IJCNN), 1-8,
586 https://doi.org/10.1109/IJCNN.2018.8489077, 2018.
587