Ocean Modelling 192 (2024) 102444

Contents lists available at ScienceDirect

Ocean Modelling

journal homepage: www.elsevier.com/locate/ocemod

Deep learning approaches in predicting tropical cyclone tracks: An analysis focused on the Northwest Pacific Region

Peng Hao, Yaqi Zhao, Shuang Li *, Jinbao Song, Yu Gao

Institute of Physical Oceanography and Remote Sensing, Ocean College, Zhejiang University, Zhoushan, China

ARTICLE INFO

Keywords:
Tropical cyclone track prediction
Northwest pacific region
Deep learning
Time series prediction

ABSTRACT

In this study, we conducted a comprehensive and integrated test of tropical cyclone track prediction using deep learning technologies, aiming to enhance the efficiency and accuracy of the prediction methods. We employed the Best Track dataset from the China Meteorological Administration's Tropical Cyclone Data Center, which covers the Northwest Pacific region from 1949 to 2023. This dataset provides comprehensive coverage, encompassing critical tropical cyclone details like time, latitude, longitude, and wind speed. Our focus was on evaluating and comparing different deep learning models, including Recurrent Neural Networks (RNN), Long Short-Term Memory Networks (LSTM), and Gated Recurrent Units (GRU), for their effectiveness in handling complex time series data. Through detailed analysis of various model configurations, including factors such as input-output lengths, hidden size, the number of layers, the implementation of bi-directional networks, and attention mechanisms, we discovered that LSTM and GRU models significantly outperform traditional RNN models in dealing with long-term dependencies and enhancing prediction accuracy. Moreover, the LSTM model, used to forecast key tropical cyclones during the 2023 Pacific tropical cyclone season, achieved mean errors of 21.84 km, 37.56 km, and 26.12 km for Typhoons Mawar, Doksuri, and Saola, respectively. This method also demonstrated high efficiency in rapid response to extreme weather changes, processing each tropical cyclone's forecast in just about 8 s. The results not only illustrate the practical utility of deep learning in tropical cyclone track prediction but also provide new, effective tools for disaster prevention and mitigation efforts.

1. Introduction

Tropical Cyclones are among the most destructive natural disasters on Earth (Peduzzi et al., 2012), posing severe challenges annually to disaster prevention and mitigation efforts in many countries worldwide (Easterling et al., 2000). The Northwest Pacific region is the most active area for tropical cyclones globally, accounting for over one-third of the total number of tropical cyclones (Gray, 1968). With the intensification of global climate change, the behavior patterns of tropical cyclones are evolving, marked by an increase in their intensity, frequency, and impact range. This evolution makes accurate predictions of their tracks more urgent and significant. Therefore, developing and selecting effective prediction models is not only crucial for mitigating the potential impacts of these storms but also holds profound implications for long-term urban planning and the formulation of disaster response strategies.

Against this backdrop, traditional methods for forecasting tropical cyclone tracks are encountering new challenges. These methods primarily rely on statistical models (Neumann, 1972; Aberson, 1998; Wang et al., 2011) and numerical models, such as the European Centre for Medium-Range Weather Forecasts Integrated Forecasting System (ECMWF-IFS), the National Centers for Environmental Prediction Global Forecast System (NCEP-GFS), the United Kingdom Met-Office Unified Model (UKMO-MetUM), the China Meteorological Administration Global Forecast System (CMA-GFS), and the Japan Meteorological Administration Global Spectrum Model (JMA-GSM). The essence of statistical models lies in their use of historical data analysis, predicting cyclone tracks by comparing current cyclones with similar historical events. Although widely used for their simplicity and intuitiveness, these models often face limitations in accuracy, especially when dealing with anomalous or unprecedented cyclone behaviors. On the other hand, numerical models, which are more complex, rely on physical equations to simulate atmospheric movement and cyclone behavior, offering more precise predictions (Bauer et al., 2015). However, this increased accuracy comes at the cost of high computational demands and sensitivity to initial conditions and parameter settings. To compensate for the

* Corresponding author.
E-mail address: lshuang@zju.edu.cn (S. Li).

https://doi.org/10.1016/j.ocemod.2024.102444
Received 13 December 2023; Received in revised form 15 September 2024; Accepted 18 September 2024
Available online 18 September 2024
1463-5003/

limitations of individual models, ensemble forecasting systems have been introduced. These systems combine outputs from multiple models to increase the reliability of predictions by generating a range of possible tracks (Bishop and Toth, 1999; Buizza et al., 2008; Haiden et al., 2018; Chen et al., 2019a; Hodges and Klingaman, 2019). Despite recent advancements and improvements in dynamic models and integration methods, which have enhanced prediction accuracy, traditional methods still face limitations in accurately predicting under extreme weather conditions. Furthermore, these methods often require significant time for data processing and model updating, hindering their ability to rapidly respond to swift changes in meteorological conditions.

In light of the limitations of traditional forecasting methods, the rise of deep learning technologies offers new perspectives and approaches. Recently, in the field of meteorology, particularly in handling and analyzing large volumes of complex meteorological data, deep learning has shown significant potential. For instance, Convolutional Neural Networks (CNNs), with their exceptional performance in image recognition and feature extraction, have become key tools in processing and analyzing meteorological satellite data (LeCun et al., 1995; Pradhan et al., 2017; Jiang et al., 2018; Matsuoka et al., 2018). Furthermore, RNN, LSTM, and GRU play a vital role in the time series modeling of meteorological data, such as in the analysis of significant wave height, wind speed, and temperature changes (Hao et al.; Hochreiter and Schmidhuber, 1997; Zaremba et al., 2014; Dey and Salem, 2017; Yu et al., 2019; Hao et al., 2023). Moreover, in exploring the prediction of tropical cyclone tracks, deep learning models have shown potential to improve prediction accuracy and provide deeper insights into cyclone behavior. For instance, Gao et al. (2018) proposed an LSTM approach which demonstrated advantages in handling time series data for cyclone prediction, improving prediction accuracy by effectively capturing long-term dependencies in the data. Alemany et al. (2019) employed an RNN to predict hurricane trajectories, effectively capturing the complex temporal patterns of hurricane behavior and demonstrating the potential of deep learning in processing complex meteorological data. Chen et al. (2019b) introduced a hybrid CNN-LSTM model which improved prediction accuracy by combining spatial and temporal features, enhancing the model's ability to extract both local and global patterns in cyclone behavior. These studies, among others (Gao et al., 2018; Alemany et al., 2019; Chen et al., 2019b; Giffard-Roisin et al., 2020; Higa et al., 2021; Song et al., 2022; Bi et al., 2023; Wang et al., 2023; Zhang et al., 2023), suggest that deep learning has the potential to improve the accuracy of tropical cyclone track predictions and deepen our understanding of cyclone behavioral patterns.

The primary objective of this research is to explore effective methods for setting parameters in the prediction of tropical cyclone tracks, aiming to contribute to the existing research in this field. In pursuit of this objective, we utilized tropical cyclone latitude, longitude, and wind speed data from the China Meteorological Administration's Tropical Cyclone Data Center, covering the Northwest Pacific region from 1949 to 2023, to deeply explore the potential applications of deep learning technologies in predicting tropical cyclone tracks. We conducted thorough testing and analysis of various deep learning models, with a focus on maintaining high prediction accuracy while reducing reliance on high-performance computing resources. This work not only demonstrates the effectiveness of simplified feature input strategies in predicting extreme weather conditions but also provides new insights into the analysis of tropical cyclone behavior and emergency response strategies in the context of global climate change. Our findings underscore the crucial role of data-driven approaches in addressing the challenges of extreme weather events.

2. Data and methods

In this chapter, we clearly outline the core methodologies of our study on tropical cyclone track prediction, including the problem definition, data preparation steps, and model evaluation metrics. Additionally, we provide detailed descriptions of the experimental design and the chosen model structures, laying a solid foundation for further result analysis and discussion.

2.1. Definition of tropical cyclone track prediction problem

The problem of tropical cyclone track prediction is fundamentally a time series prediction issue, where the objective is to predict future tracks based on historical data. We use past latitude, longitude, and wind speed data as input to forecast the cyclone's track over a future period through deep learning models. Specifically, we utilize the observed sequence from the past T time points, $X_{t-T+1}, \ldots, X_t$, to predict the track sequence for the next F time points, $\widehat{Y}_{t+1}, \ldots, \widehat{Y}_{t+F}$. These predicted values are then compared with the actual observed values, $Y_{t+1}, \ldots, Y_{t+F}$, at those future F time points. This prediction task can be expressed by the following conditional probability formula:

$$\widehat{Y}_{t+1}, \ldots, \widehat{Y}_{t+F} = \underset{\widehat{Y}_{t+1}, \ldots, \widehat{Y}_{t+F}}{arg\ max}\ P(Y_{t+1}, \ldots, Y_{t+F} | X_{t-T+1}, \ldots, X_t) \tag{1}$$

In this study, we ensured the relative independence of the training and validation sets, with the validation dataset completely excluded during the model training phase and only utilized in the model evaluation stage. This approach helps to prevent data leakage, thereby enhancing the accuracy of the model assessment.

2.2. Best track dataset for tropical cyclones

We utilized the Best Track dataset from the Tropical Cyclone Data Center of the China Meteorological Administration (Ying et al., 2014; Lu et al., 2021). This dataset meticulously records the tropical cyclone activities in the Northwest Pacific region since 1949, including essential information such as latitude, longitude, and wind speed, thus offering high representativeness and completeness. To ensure continuity in time series analysis, our study uniformly adopted a 6-hour data recording interval, aligning with the data frequency before 2017. Additionally, to maintain the reliability of the dataset, we excluded tropical cyclone tracks with fewer than 20 recorded points to avoid the model learning incorrect or imprecise trends, thereby enhancing the accuracy of model training and prediction.

We focus on using four key parameters: time, longitude, latitude, and wind speed. Although these parameters are simple, they play a crucial role in the dynamic analysis of tropical cyclones. Time series help us understand the development and changing trends of tropical cyclones, while latitude and longitude data accurately depict the location and movement trajectory of tropical cyclones. Wind speed, as a direct measure of tropical cyclone intensity, is an indispensable element in forecasting. The comprehensive analysis of these parameters enables our model to more accurately capture key dynamic characteristics related to tropical cyclone behavior, thereby significantly improving the accuracy

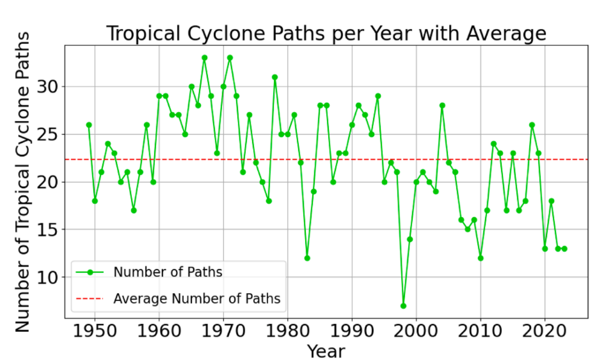

Fig. 1. Interannual changes in the number of tropical cyclones in the northwest Pacific region from 1949 to 2023.

of track predictions.

As shown in Figs. 1 and 2, a visual analysis of the annual number of tropical cyclone tracks shows that although there are some fluctuations from year to year, no sustained long-term upward or downward trends are observed. In terms of tropical cyclone intensity distribution, tropical depressions occur most frequently, while super typhoons occur relatively rarely, indicating that tropical cyclones of extreme intensity are not a common phenomenon. This visualization of trends and distributions provides deep insights into understanding the historical activity of tropical cyclones and lays the data foundation for further prediction of tropical cyclone tracks using deep learning models. In further studies, models must take into account the following: first, natural variability reflected in annual fluctuations, which may be associated with specific seasonal factors; second, data imbalance issues, given the high incidence of tropical depressions and The scarcity of super typhoons, requires corresponding strategies to deal with this imbalance during model training; in addition, the model also needs to be able to identify and adapt to abnormal climate patterns, especially when predicting rare extreme events.

2.3. Model introduction

The tropical cyclone data are time-series data with high seriality and complexity. Therefore, we specifically focus on RNN, LSTM, and GRU in model selection, which offer potential advantages in processing time series data (Graves, 2012; Chung et al., 2014; Zaremba et al., 2014). Fig. 3 shows the structures of these models and how they handle time-series information.

2.3.1. Recurrent neural network (RNN)

The Recurrent Neural Network (RNN) is a classic neural network architecture for processing time series data, capable of capturing temporal dynamic characteristics within a sequence. RNN have internal looping units that allow information to be passed from one time step to the next. This structure enables RNN to excel in handling sequences of variable lengths. At each time step, RNN produces a hidden state that carries not only the information of the current input but also information from previous time steps. The update of the RNN's hidden state can be represented by the following formula:

$$h_t = tanh\ (W_{hh}h_{t-1} + W_{xh}x_t + b_h) \tag{2}$$

In this context, h_t represents the hidden state at time step t and x_t is the input vector. W_{hh} and W_{xh} denote the weight matrices for the hidden-to-hidden connections and input-to-hidden connections, respectively. Bias terms is represented by b_h for the hidden layer.

2.3.2. Long short-term memory network (LSTM)

The Long Short-Term Memory Network (LSTM) is an enhancement of the RNN, designed to address the issues of vanishing or exploding gradients encountered by RNN in learning long sequences. The LSTM introduces three gate mechanisms: the forget gate, input gate, and output gate, which control the retention, updating, and output of information. This structure enables the LSTM to maintain stable learning and memory capabilities over long sequences. The updated formulas for the LSTM are as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

$$\widetilde{C}_t = tanh\ (W_C \cdot [h_{t-1}, x_t] + b_C) \tag{5}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{6}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = o_t *\ tanh\ (C_t) \tag{8}$$

Here, σ denotes the sigmoid function, C_t represents the cell state, $\widetilde{C}_t$ is the candidate cell state, and f_t, i_t, and o_t are the activation values of the forget gate, input gate, and output gate, respectively. W_f, W_i, W_C, and W_o denote the weight matrices associated with the forget gate, input gate, candidate cell state, and output gate, respectively. The bias terms b_f, b_i, b_C, and b_o represent the biases for the forget gate, input gate, candidate cell state, and output gate, respectively. The term $[h_{t-1}, x_t]$ indicates the concatenation of the previous hidden state h_{t-1} and the current input x_t. The operation * denotes pointwise multiplication.

2.3.3. Gated recurrent unit (GRU)

The Gated Recurrent Unit (GRU) is a variant of the LSTM, which simplifies the LSTM structure by merging the forget and input gates into a single update gate and introducing a reset gate. This streamlined structure reduces the number of parameters, often leading to faster model training speeds, and can provide performance comparable to LSTM in certain scenarios.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{9}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{10}$$

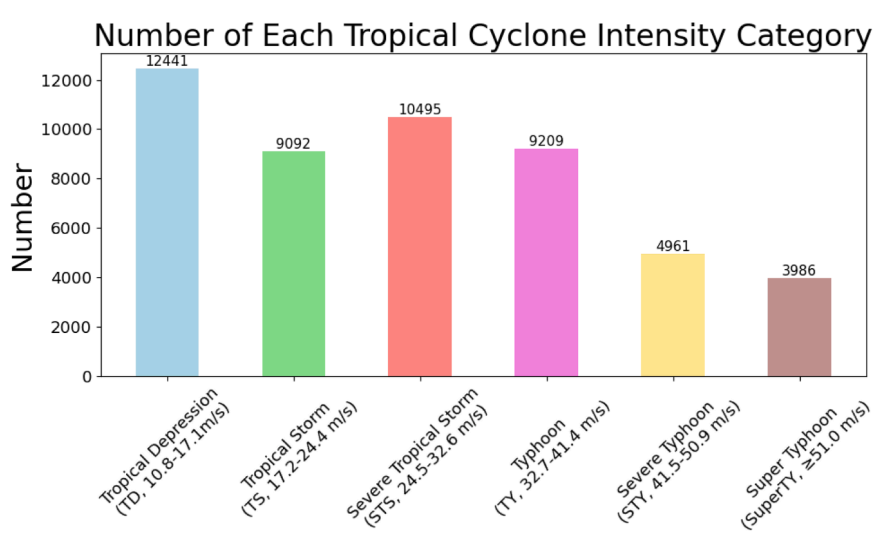

Fig. 2. Distribution of tropical cyclone intensity categories in the Northwest Pacific Region from 1949 to 2023.

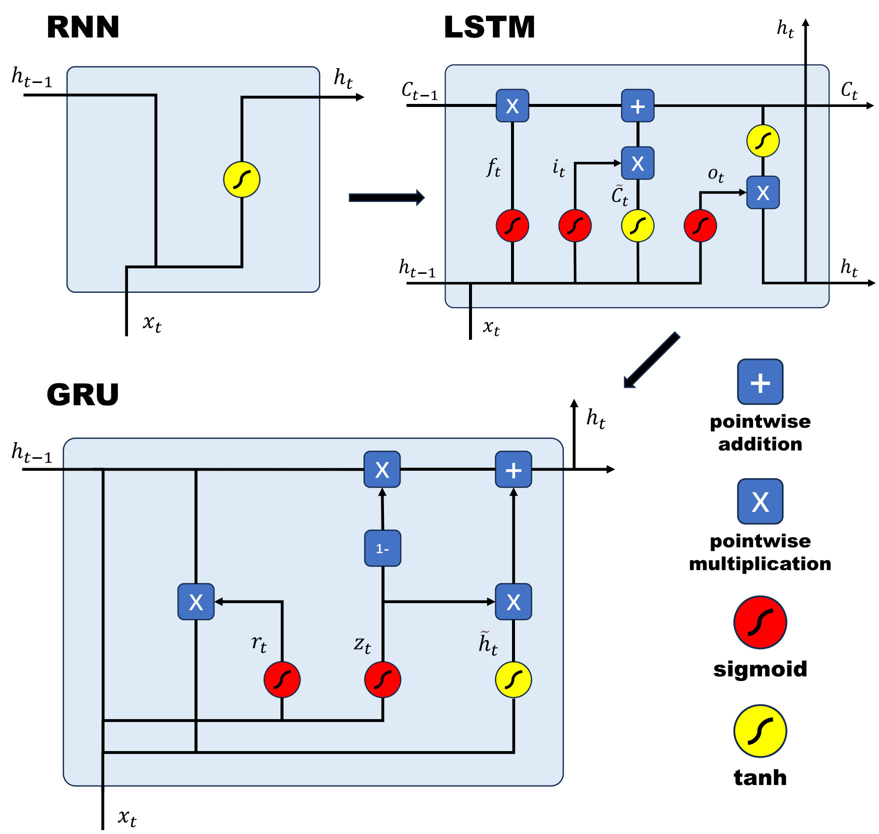

Fig. 3. Structural Illustration of RNN, LSTM, and GRU Models.

$$\widetilde{h}_t = tanh(W_h \cdot [r_t * h_{t-1}, x_t] + b_h) \tag{11}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \widetilde{h}_t \tag{12}$$

In this context, z_t and r_t represent the activation values of the update gate and reset gate, respectively, while $\widetilde{h}_t$ is the candidate hidden state. W_z, W_r, and W_h denote the weight matrices associated with the update gate, reset gate, and candidate hidden state, respectively. The bias terms b_z, b_r, and b_h represent the biases for each respective gate or component. The term $[h_{t-1}, x_t]$ indicates the concatenation of the previous hidden state h_{t-1} and the current input x_t. The operation * denotes pointwise multiplication.

2.3.4. Attention mechanism

The attention mechanism is a technique in deep learning designed to enhance a model's ability to recognize important parts of the input data. In the task of tropical cyclone track prediction, the mathematical expression of the attention mechanism is as follows:

$$c_t = \sum_{i=1}^{T} a_{t,i} h_i \tag{13}$$

$$a_{t,i} = \frac{\exp(score(s_{t-1}, h_i))}{\sum_{j=1}^{T} \exp(score(s_{t-1}, h_j))} \tag{14}$$

Here, c_t represents the context vector, $a_{t,i}$ denotes the attention weights, and h_i is the hidden state of the encoder at time step i. The attention weights $a_{t,i}$ are typically computed by a small feed-forward network that considers the relationship between the current decoder state s_{t-1} and the hidden states h_i of each encoder. We will compare the performance of models with and without integrated attention layers in the task of predicting tropical cyclone tracks. Such a comparison will aid in understanding and evaluating the practical application value of the attention mechanism in specific prediction tasks.

Our study utilizes the aforementioned models to conduct experiments and evaluate their effectiveness in predicting tropical cyclone tracks. The model evaluation metrics include prediction accuracy, model stability, and performance over different time ranges. Our goal is to determine which model is most suitable for typhoon track prediction and to explore how to further optimize the models to enhance prediction accuracy.

2.4. Performance metrics

To comprehensively assess the performance of deep learning models in predicting tropical cyclone tracks in this study, we employed a range of detailed metrics. These metrics specifically measure the accuracy and reliability of the models in predicting the future positions (latitude and longitude) of tropical cyclones, with the error margins quantified in degrees of geographic coordinates.

Mean Absolute Error (*MAE*): This key indicator measures the average absolute difference between the model's predicted track (expressed as latitude and longitude coordinates) and the actual observed track, crucial for assessing the accuracy of the predicted track.

Mean Squared Error (*MSE*): This metric reflects the average of the squared differences between the predicted and actual tracks across all data points, indicating the overall magnitude of the prediction error.

Root Mean Squared Error (*RMSE*): The square root of the MSE, RMSE provides an assessment in the same units as the track error, offering a direct view of the deviation in the model's predicted track.

Coefficient of Determination (R^2): This statistic describes the model's ability to explain the variation in tropical cyclone tracks, with values closer to 1 indicating higher prediction accuracy.

Mean Absolute Percentage Error (*MAPE*): This metric displays the average deviation between the predicted and actual tracks in percentage terms, providing an intuitive representation of the model's prediction accuracy.

Together, these metrics offer a comprehensive evaluation of the model's predictive performance, ensuring an effective assessment of the accuracy and reliability of future tropical cyclone track predictions.

2.5. Experimental design

This study is dedicated to comprehensively evaluating the impact of various deep-learning parameter configurations on tropical cyclone track prediction, aiming to identify the most effective model settings for enhancing prediction accuracy. To this end, we have undertaken a series of detailed experimental steps to ensure a thorough and accurate assessment of model performance. Detailed information about the parameters used in the experiments is presented in Table 1.

All experiments in this study were conducted using the Pytorch deep learning framework (Paszke et al., 2019) and performed on an NVIDIA 3070 GPU. The specific experimental procedures are as follows:

(1) Data Loading and Preprocessing:
We acquired tropical cyclone data from the China Meteorological Administration's Tropical Cyclone Data Center, standardizing time intervals and eliminating incomplete tracks to ensure data consistency and completeness. The dataset, covering the years 1949 to 2022, was divided into training and validation sets in an 8:2 ratio, with the first 80% of the data used for training and the remaining 20% for validation.

(2) Experimental Setup:
To ensure reproducibility, we set a fixed random seed and adjusted various model parameters to evaluate their impact on tropical cyclone track prediction. These parameters included model types (RNN, LSTM, GRU), the use of attention mechanisms, lengths of input and output sequences, training set proportions, batch sizes, hidden layer sizes, number of network layers, and learning rates.

(3) Training Process:
We used the Adam optimizer and trained the models based on the set hyperparameters (such as batch size, learning rate, and number of training epochs), paying special attention to the specific effects of different parameter configurations on model performance.

(4) Performance Evaluation:
The predictive performance of the models was assessed on the validation set using multiple metrics (*MAE, MSE, RMSE,* R^2, and *MAPE*) to quantify the impact of different configurations on prediction accuracy.

(5) Results Visualization:
We plotted performance metrics to visually demonstrate the model's performance during training, as well as the impact of different hyperparameter configurations on the outcomes. Notably, these visualizations were based on the data from 2023.

Table 1
| Parameters setting.

Parameter	Setting
Input length (hours)	6 / 12 /18 / 24 / 48 / 72
Output length (hours)	6 / 12 /18 / 24 / 48 / 72
Layers	1 / 2 / 3 / 4 / 5 / 6
Hidden size	4 / 8 / 16 / 32 / 64 / 128
Batch size	128
Learning rate	0.001
Epochs	301
Train ratio	0.8
Optimizer	Adam

3. Experimental results and analysis

In this chapter, we systematically present and analyze the experimental results. This includes the specific impacts of input sequence length, prediction period, and model complexity on the prediction outcomes. Furthermore, through the visualization of specific case studies, we demonstrate the practical application and effectiveness of the models in predicting tropical cyclone tracks.

3.1. Impact of input length on prediction results

In the context of applying deep learning to tropical cyclone track prediction, the choice of input length has an important impact on the efficiency of model learning and prediction accuracy. Understanding how input length affects model performance involves not only the amount of data but also whether the model can capture key dynamic features of cyclone track changes. In addition, appropriate input length helps models learn and internalize complex climate patterns that are critical to predictions. Therefore, choosing the optimal input length can ensure the robustness of the model in dealing with natural variability and imbalanced data, especially when distinguishing between frequent tropical depressions and rare super typhoons. By analyzing the effects of different input lengths in experiments, we can finely tune the model to improve the overall performance of tropical cyclone track predictions. This subsection focuses on analyzing the performance of three models (RNN, LSTM, and GRU) at varying input lengths. The aim is to understand how input length affects the models' accuracy and efficiency. Our experiments employed a uniform configuration: the output length was fixed at 6 h, the batch size at 128, the hidden size at 64, and a three-layer network structure was used. Such uniformity in settings allowed us to precisely assess the specific impact of different input lengths on model performance. The experimental results are shown in Table 2 and Fig. 4.

The experimental results indicate that for the RNN model, an increase in input length from 6 to 24 h leads to a reduction in MAE and MSE, suggesting that prediction accuracy improves with longer input lengths. However, as the input length continues to increase, this trend of improvement becomes less pronounced and even shows a decrease in performance in some cases. This could be attributed to the traditional RNN's susceptibility to the vanishing gradient problem when dealing with longer input sequences, hindering its ability to effectively learn and remember long-term historical information.

Conversely, the LSTM and GRU models overall outperform the RNN at different input lengths. Particularly, the LSTM shows lower MAE and MSE at input lengths of 18 and 24 h, indicating its significant advantage in handling long-term dependencies. The LSTM's unique gating mechanism effectively overcomes the vanishing gradient problem faced by

Table 2
| Influence of input sequence lengths on performance of models predicting tropical cyclone tracks.

Input Lengths		6h	12h	18h	24h	48h	72h
RNN	*MAE*	0.890	1.004	0.630	0.610	**0.543**	0.562
	MSE	6.130	7.187	**4.492**	4.932	5.221	4.705
	RMSE	2.476	2.681	**2.119**	2.221	2.285	2.169
	R^2	0.977	0.972	0.979	0.981	0.981	**0.982**
	MAPE	1.905	1.942	1.716	1.498	**1.100**	1.368
LSTM	*MAE*	0.717	0.452	**0.438**	**0.438**	0.443	0.468
	MSE	4.602	3.564	3.492	**2.778**	4.102	3.276
	RMSE	2.145	1.888	1.869	**1.667**	2.025	1.810
	R^2	0.982	0.986	0.987	**0.988**	0.985	0.987
	MAPE	1.538	1.006	0.969	1.101	**0.882**	1.111
GRU	*MAE*	0.706	0.933	0.487	**0.471**	0.491	0.486
	MSE	4.088	5.378	4.434	**3.071**	3.315	4.412
	RMSE	2.022	2.319	2.106	**1.753**	1.821	2.100
	R^2	0.984	0.977	0.984	**0.985**	0.984	**0.985**
	MAPE	1.507	1.896	1.071	1.296	1.361	**0.954**

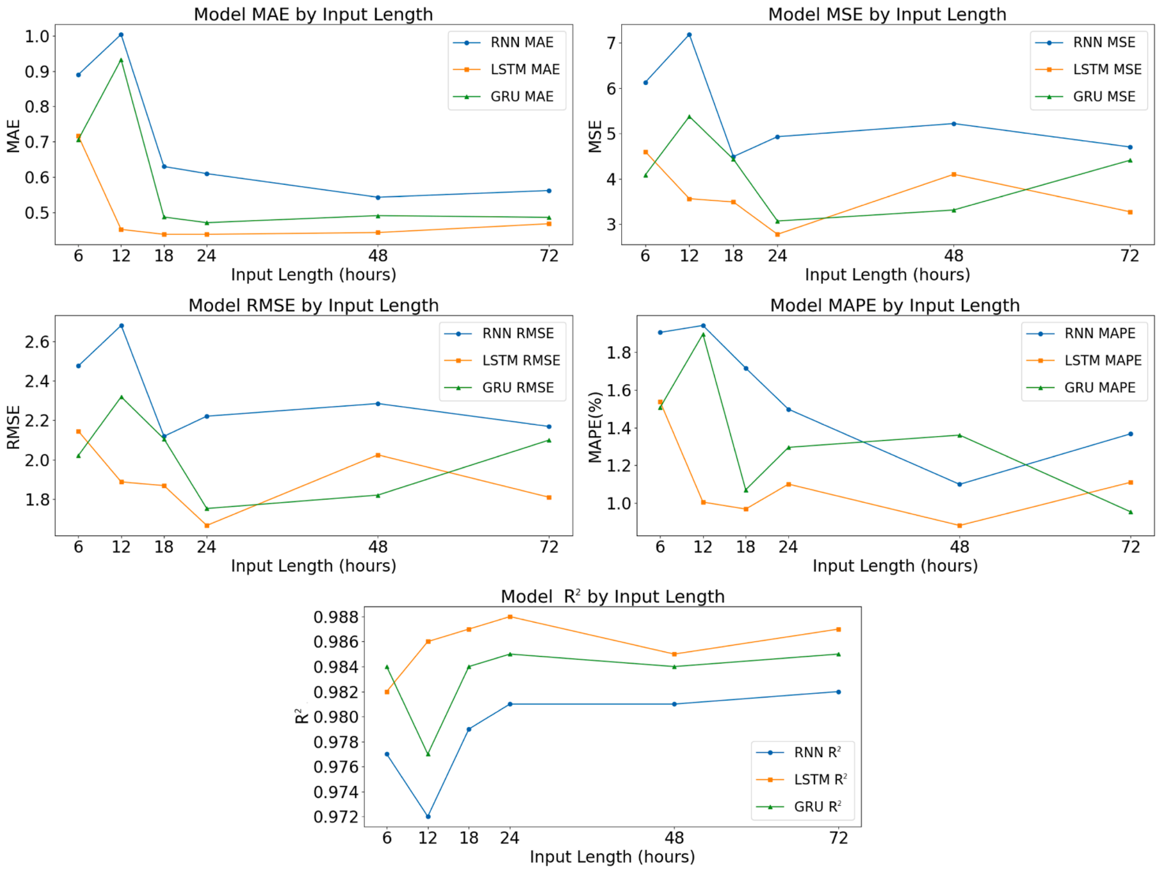

Fig. 4. Comprehensive performance evaluation of deep learning models across varying input sequence lengths for tropical cyclone track prediction.

traditional RNN, thus performing better in processing long-term dependencies. Similarly, the GRU model also demonstrates good performance across various input lengths, especially with a 48-hour input length, where it has relatively lower MAE and MSE. The GRU's streamlined structure, which reduces the number of parameters compared to LSTM, likely achieves more efficient learning at specific input lengths.

Overall, our study reveals the significance of input length in tropical cyclone track prediction models. An appropriate input length can provide the model with sufficient historical information to more accurately capture and predict the behavior patterns of cyclones. Our findings emphasize the importance of considering these details when selecting models and configuring input parameters, which is crucial for enhancing the accuracy of tropical cyclone track predictions.

3.2. *Impact of prediction length on prediction results*

We further investigated the impact of prediction length on model performance. To ensure consistency in the experiments, we adopted a uniform experimental configuration: the input length was fixed at 24 h, the batch size was set to 128, the hidden size was 64, and a three-layer network structure was utilized. The experimental results are presented in Table 3.

As depicted in Fig. 5, for the RNN model, an increasing trend in the prediction length (from 6 to 72 h) exhibited a decline in performance across all metrics. Notably, at longer prediction lengths (48 and 72 h), there was a significant increase in MAE, MSE, and RMSE, and a notable decrease in the R^2 value, indicating a substantial reduction in prediction accuracy. The LSTM model generally outperformed the RNN across all prediction lengths. Despite a decrease in performance with increased prediction length, the extent of this decline was relatively minor for LSTM, particularly maintaining higher accuracy at medium prediction lengths (such as 24 h). The GRU model showed similar performance to LSTM across different prediction lengths, demonstrating higher stability. Compared to LSTM, GRU experienced a slightly more pronounced decline in performance at longer prediction lengths, but still overall performed better than the RNN.

The experimental results indicate that with increasing prediction lengths, the performance of all models decreases, but LSTM and GRU models are relatively more robust in handling long-term prediction tasks. This may be due to the specialized structural design of LSTM and GRU, which enables them to more effectively handle long-term dependencies and memory information. The RNN model, due to its inherent structural limitations, performs poorly in long-term prediction

Table 3
| Influence of output sequence lengths on performance of models predicting tropical cyclone tracks.

Output Lengths		6h	12h	18h	24h	48h	72h
RNN	*MAE*	**0.610**	0.885	1.021	1.218	2.210	2.748
	MSE	**4.932**	5.639	5.832	7.707	15.727	23.301
	RMSE	**2.221**	2.375	2.415	2.776	3.966	4.827
	R^2	**0.981**	0.964	0.972	0.962	0.915	0.872
	MAPE	**1.498**	2.604	2.844	2.870	4.882	6.292
LSTM	*MAE*	**0.438**	0.565	0.752	0.918	1.690	2.465
	MSE	**2.778**	3.672	4.340	5.688	9.900	18.202
	RMSE	**1.667**	1.916	2.083	2.385	3.146	4.266
	R^2	**0.988**	0.986	0.982	0.976	0.947	0.900
	MAPE	**1.101**	1.278	1.781	2.018	3.946	5.706
GRU	*MAE*	**0.471**	0.773	0.781	1.119	1.825	2.573
	MSE	**3.071**	4.993	5.330	6.763	11.528	19.657
	RMSE	**1.753**	2.235	2.309	2.601	3.395	4.434
	R^2	**0.985**	0.979	0.978	0.971	0.941	0.891
	MAPE	**1.296**	1.844	1.772	2.307	4.175	5.921

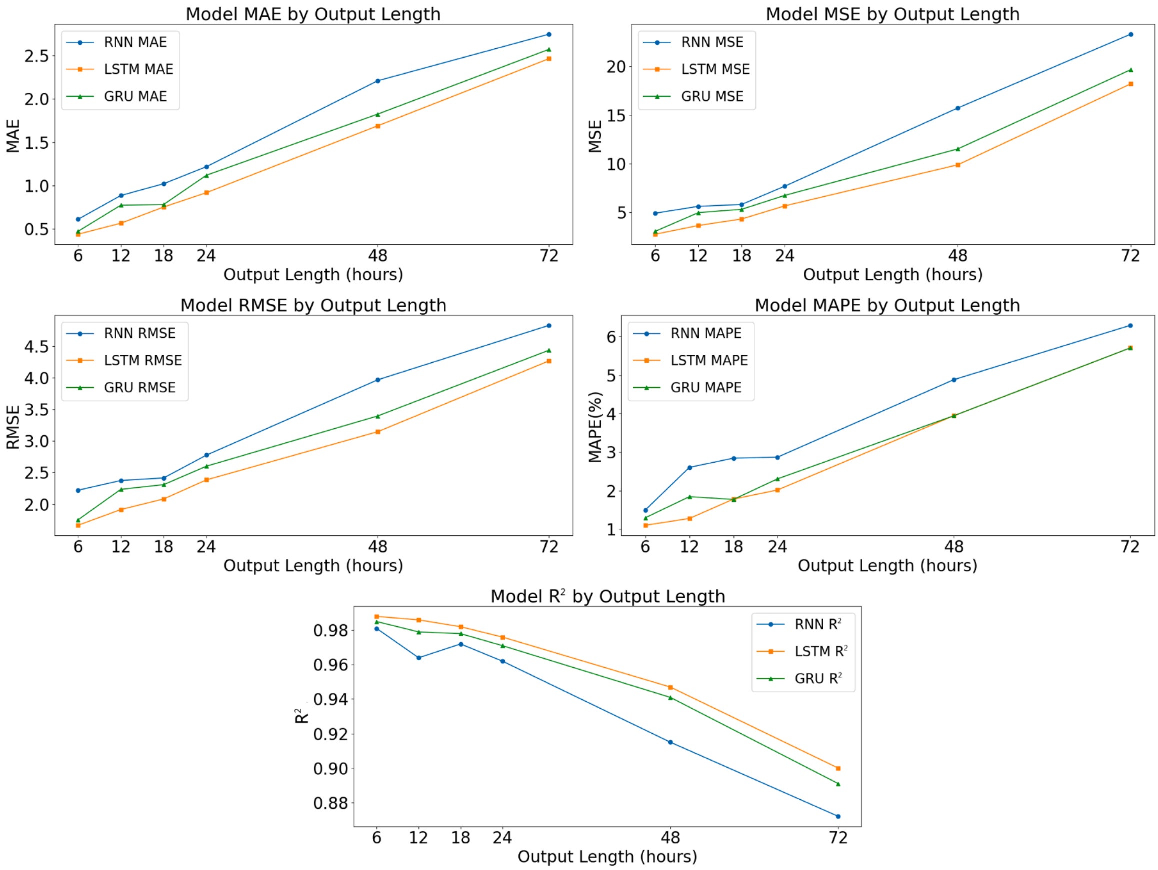

Fig. 5. Comprehensive performance evaluation of deep learning models across varying output sequence lengths for tropical cyclone track prediction.

tasks, particularly at longer prediction lengths. These findings are significant for improving tropical cyclone track prediction models. Choosing the appropriate model and adjusting prediction length settings can significantly enhance the accuracy and reliability of predictions. Especially for long-term predictions, LSTM and GRU models may be preferable choices. These insights offer valuable guidance for the future application of deep learning models in disaster warning and emergency response strategies.

3.3. Impact of model complexity on prediction results

In this subsection, we explore the impact of model complexity on tropical cyclone track prediction results. By systematically analyzing the performance of three different models (RNN, LSTM, and GRU) at varying levels of network structural complexity, we aim to uncover how model complexity influences the accuracy and efficiency of predictions.

(1) Number of Network Layers

The number of network layers is also a key parameter in deep learning models, which has an important impact on model performance. The formation and evolution of tropical cyclones is a multi-scale and multi-complex process. Choosing different numbers of network layers can be seen as introducing different levels of feature representation and abstraction into the model, which helps the model better capture the multi-scale features in the dynamic mechanism of tropical cyclones. However, deeper networks may also increase model complexity and computational resource requirements, so a balance between model performance and computational efficiency needs to be considered when selecting the number of network layers. This subsection is dedicated to analyzing the performance of the three deep learning models across different numbers of network layers, to gain a deeper understanding of how network layers affect the models' accuracy and efficiency. To ensure consistency in the experiments, we adopted a uniform experimental configuration: an input length of 24 h, an output length of 6 h, a batch size of 128, and a hidden size of 64. The experimental results are presented in Table 4 and Fig. 6.

The experimental data indicates that the RNN model exhibits a nonlinear change in performance with increasing network layers. Specifically, as the network layers increase from one to three, there is a slight upward trend in performance metrics such as MAE and MSE, suggesting a minor decline in performance. However, as the network layers continue to increase to four or more, the performance metrics begin to deteriorate significantly, especially in the six-layer network structure, where the degradation is particularly pronounced. This suggests that in RNN models, an excessive number of layers leads to

Table 4
| Influence of layers on performance of models predicting tropical cyclone tracks.

Layers		1	2	3	4	5	6
RNN	*MAE*	**0.431**	0.585	0.610	0.695	1.089	10.456
	MSE	**3.091**	3.900	4.932	3.661	8.136	204.138
	RMSE	**1.758**	1.975	2.221	1.913	2.852	14.288
	R^2	**0.983**	0.979	0.981	0.981	0.951	−0.010
	MAPE	**1.219**	1.601	1.498	1.850	3.183	24.681
LSTM	*MAE*	**0.376**	0.428	0.438	0.440	0.481	0.517
	MSE	**2.261**	2.594	2.778	3.851	3.243	3.303
	RMSE	**1.504**	1.611	1.667	1.962	1.801	1.817
	R^2	**0.991**	0.990	0.988	0.986	0.986	0.985
	MAPE	**0.868**	0.917	1.101	0.969	1.247	1.381
GRU	*MAE*	**0.455**	0.462	0.471	0.561	0.712	0.877
	MSE	**2.867**	3.518	3.071	4.282	6.294	5.229
	RMSE	**1.693**	1.876	1.753	2.069	2.509	2.287
	R^2	**0.987**	**0.987**	0.985	0.984	0.977	0.965
	MAPE	1.209	**1.099**	1.296	1.422	1.546	2.954

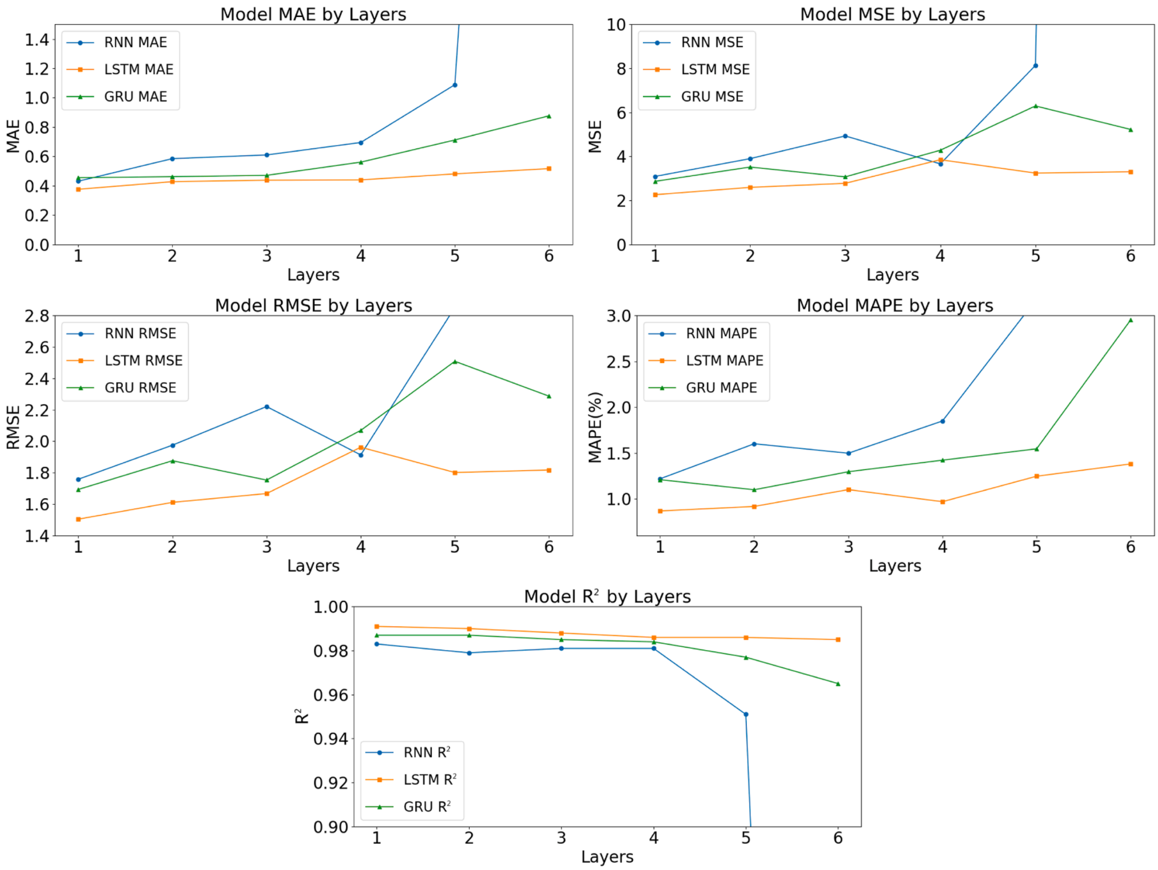

Fig. 6. Comprehensive performance evaluation of deep learning models across varying layers for tropical cyclone track prediction.

performance degradation, possibly due to the deep network structure causing gradient vanishing issues, impacting the model's ability to learn long-term dependencies. Owing to the design features of LSTM and GRU that counteract the gradient vanishing problem (such as LSTM's gating mechanism and GRU's simplified structure), they generally exhibit higher robustness with an increase in network layers. Nevertheless, this does not mean that LSTM and GRU will always maintain improved or stable performance with an increase in network layers. Too many layers may lead to overfitting or increase the computational burden during model training and inference.

In conclusion, our study highlights the importance of network layers in tropical cyclone track prediction models, especially for RNN models, where an excessive number of layers may be counterproductive, leading to significant performance degradation. This finding provides important guidance for the design of deep learning models, particularly when dealing with meteorological data characterized by complex time series. It suggests that the design of deep learning models requires meticulous parameter adjustment and testing to find the network structure most suited for specific prediction tasks.

(2) Hidden Sizes

The dynamic mechanism of tropical cyclones is complex, and its dynamic mechanism may involve multi-scale and nonlinear changes. In the model, a larger hidden layer size can provide more capacity, allowing the model to better capture these complex dynamic characteristics. However, the choice of hidden layer size also needs to consider computational resources and overfitting issues. Excessively large hidden layer sizes may cause the model to overfit the training data, thus requiring a trade-off between model performance and computational efficiency. Therefore, when selecting the size of the hidden layer, it is necessary to comprehensively consider the dynamic mechanism of tropical cyclones, data complexity, and computing resources to ensure that the model can fully utilize this information to make accurate predictions. In this subsection, we analyze the impact of hidden sizes on the performance of tropical cyclone track prediction models. For experimental consistency, we maintained a uniform configuration: an input length of 24 h, an output length of 6 h, a batch size of 128, and a three-layer network structure. The experimental results are depicted in Table 5 and Fig. 7.

The experimental findings indicate that for the RNN, increasing the hidden size from very low (such as 4, 8) to moderate (like 32, 64) leads to a significant decrease in performance metrics such as MAE, MSE, and RMSE, while R^2 increases. This implies that the model's predictive accuracy is notably enhanced with a larger hidden size. However, with too small a hidden size (4, 8), the model performs poorly, likely due to

Table 5
| Influence of hidden sizes on performance of models predicting tropical cyclone tracks.

Hidden Sizes		4	8	16	32	64	128
RNN	*MAE*	10.427	10.444	1.150	**0.589**	0.610	0.628
	MSE	203.700	203.737	6.646	5.751	4.932	**4.533**
	RMSE	14.272	14.274	2.578	2.398	2.221	**2.129**
	R^2	−0.006	−0.006	0.963	0.978	**0.981**	0.979
	MAPE	24.927	24.914	3.879	**1.280**	1.498	1.707
LSTM	*MAE*	10.434	0.565	0.572	0.457	**0.438**	0.463
	MSE	203.679	4.067	3.255	3.828	**2.778**	2.838
	RMSE	14.272	2.017	1.804	1.957	**1.667**	1.685
	R^2	−0.006	0.980	0.987	0.986	0.988	**0.989**
	MAPE	24.943	1.544	1.363	**0.942**	1.101	1.052
GRU	*MAE*	5.129	1.076	0.494	0.533	**0.471**	0.540
	MSE	65.383	7.940	3.691	4.036	**3.071**	3.783
	RMSE	8.086	2.818	1.921	2.009	**1.753**	1.945
	R^2	0.489	0.964	0.983	0.984	**0.985**	0.983
	MAPE	19.727	3.062	1.271	**1.238**	1.296	1.405

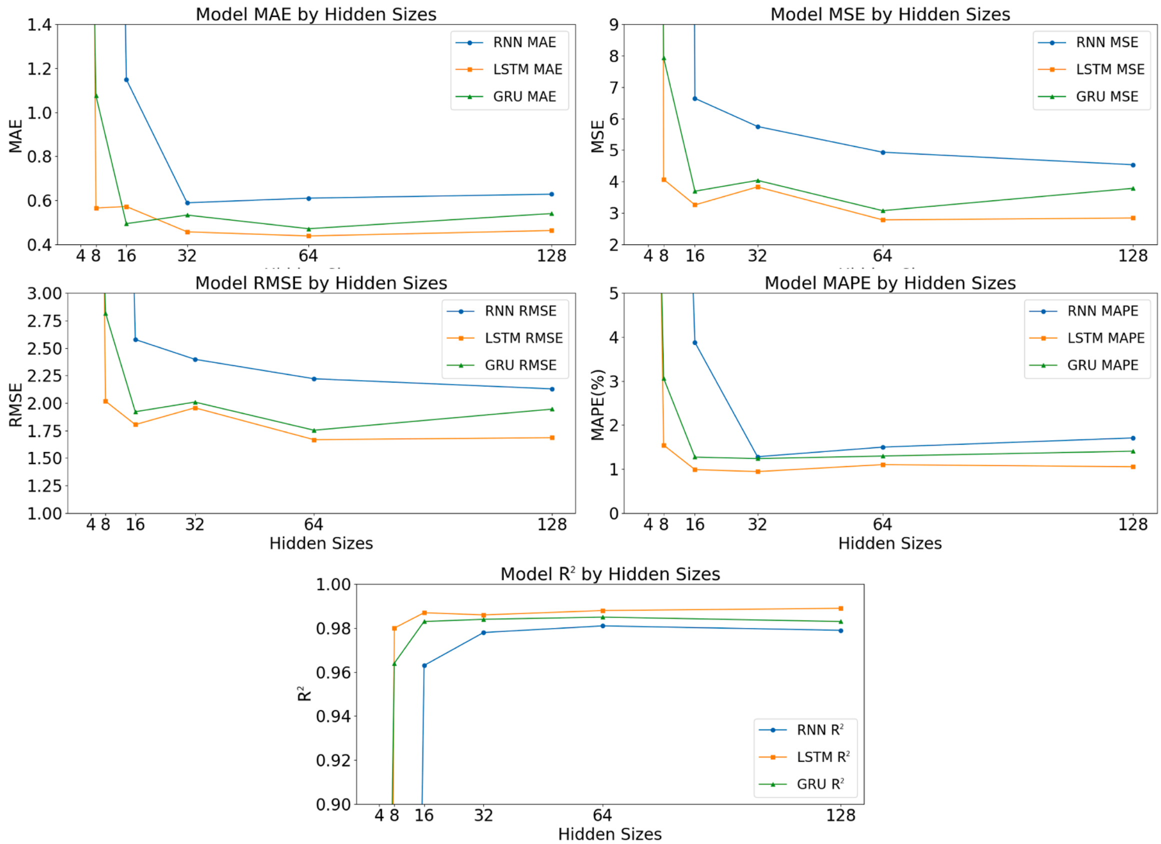

Fig. 7. Comprehensive performance evaluation of deep learning models across varying hidden sizes for tropical cyclone track prediction.

severe underfitting. The improvement is attributable to the larger hidden size providing a richer representation of data, enabling the RNN to more effectively learn and predict complex tropical cyclone track patterns.

Similarly, the LSTM model underperforms at very low hidden sizes (4, 8), indicating limited predictive capacity. As the hidden size increases, its performance gradually improves, particularly at medium to high counts (such as 64, 128), where the model shows better predictive accuracy and stability. The GRU model also exhibits inferior performance at very low hidden sizes, aligning with the trends seen in RNN and LSTM. With an increase in hidden size, GRU's performance progressively improves, especially at medium and higher configurations, demonstrating enhanced predictive accuracy and stability.

In conclusion, the hidden size significantly impacts the performance of all three models. At very low hidden sizes, RNN, LSTM, and GRU all demonstrate poor performance, possibly due to an inability to capture sufficient data features and complexities. With increasing hidden sizes, all models show marked improvements, particularly at moderate configurations. This finding underscores the importance of appropriately choosing hidden sizes to achieve optimal predictive performance.

(3) Use of Bi-Directional Networks

We conducted an in-depth analysis of the impact of employing Bi-Directional Networks on tropical cyclone track prediction. To ensure consistency in the experiments, we used a uniform configuration: an input length of 24 h, an output length of 6 h, a batch size of 128, a hidden size of 64, and a three-layer network structure. The experimental results are detailed in Table 6.

Bi-directional RNN outperformed uni-directional RNN across all performance metrics. Specifically, Bi-Directional RNN showed reductions in MAE, MSE, and RMSE, and an increase in R^2, indicating greater accuracy and reliability in predicting tropical cyclone tracks. This improvement is likely due to the Bi-Directional RNN's ability to utilize both past and future information, gaining a more comprehensive contextual understanding and thereby providing a more holistic data interpretation.

Table 6
| Influence of bidirectionality on performance of models predicting tropical cyclone tracks.

Bidirectionality	RNN	Bi-RNN	LSTM	Bi-LSTM	GRU	Bi-GRU
MAE	0.610	0.595	**0.438**	0.454	0.471	0.510
MSE	4.932	4.081	2.778	**2.720**	3.071	3.932
RMSE	2.221	2.020	1.667	**1.650**	1.753	1.983
R^2	0.981	0.983	**0.988**	**0.988**	0.985	0.986
MAPE	1.498	1.486	1.101	1.133	1.296	**1.007**

However, for LSTM and GRU models, transitioning to a bi-directional structure did not yield the expected performance enhancement. In fact, Bi-LSTM and Bi-GRU showed slight decreases in performance on some metrics. This could be due to the increased complexity of the bi-directional network, which in this case did not provide additional valuable information to the model, and might have introduced noise or led to overfitting during the learning process.

These results are crucial for understanding the relationship between model complexity and predictive performance. They indicate that increasing the complexity of the model (such as adopting a bi-directional network) does not always lead to improved performance. Particularly in the case of LSTM and GRU, the more complex bi-directional structure did not enhance predictive performance. This offers important insights for our scientific question: in tropical cyclone track prediction, the choice and configuration of the model must be carefully considered to ensure the model's complexity matches the demands of the prediction task. Excessively increasing model complexity may not bring the expected performance improvements and might instead lead to reduced efficiency or decreased prediction accuracy. Therefore, when designing prediction models, there should be a balance between model complexity and the specific requirements of the prediction task to achieve optimal predictive performance.

(4) Attention Mechanism

In this section, we investigate the impact of the attention mechanism on tropical cyclone track prediction models. The attention mechanism, a crucial deep learning technique, enhances model predictive accuracy and efficiency by enabling the model to differentially weigh various parts of the input data. To ensure experimental consistency, we adopted a uniform configuration: an input length of 24 h, an output length of 6 h, a batch size of 128, a hidden size of 64, and a three-layer network structure. The experimental results are shown in Table 7.

The experimental data indicate that the integration of the attention mechanism in the RNN results in a reduction in MAE, MSE, and RMSE, with a slight decrease in R^2. This suggests that the addition of the attention mechanism positively impacts the predictive accuracy of the RNN, albeit with limited improvement in the model's interpretability.

For LSTM, the introduction of the attention mechanism led to a slight increase in MAE, MSE, and RMSE, and a slight decrease in R^2. This may suggest that in the context of LSTM, the attention mechanism does not significantly enhance model performance. This could be due to the LSTM's inherent gating mechanism, which already focuses on important information to some extent, making the additional attention mechanism less advantageous. Similarly, the integration of the attention mechanism in GRU resulted in a slight decrease in performance metrics, indicating that for the GRU model, the attention mechanism may not effectively enhance the model's ability to process time series data.

In summary, our research results emphasize the need for careful consideration of integrating the attention mechanism when selecting and designing deep learning models for tropical cyclone track prediction. For some models (like RNN), the attention mechanism may be a beneficial addition, but for others (such as LSTM and GRU), it may not bring the expected performance improvement. This finding provides important insights for optimizing the application of deep learning models in complex prediction tasks.

3.4. Case visualization

We conducted a comprehensive evaluation of all tropical cyclones in 2023. The model was configured with a 24-hour input length, a batch size of 128, 64 hidden size, and a single network layer, achieving a balance between computational efficiency and prediction accuracy. The summarized prediction results are shown in Table 8.

We selected Mawar, Doksuri, and Saola for individual visualization due to their unique formation processes and significant impacts. Mawar was chosen for its rapid intensification and wide impact area, Doksuri for its abrupt and significant track changes, and Saola for its complex path and large intensity fluctuations. These characteristics provided diverse data to validate the model's adaptability and accuracy.

Typhoon Mawar formed over the Northwest Pacific on May 19, 2023, and reached the status of a super typhoon as classified by the China Meteorological Administration on May 23, becoming the first super typhoon of 2023. Mawar's rapid intensification and wide impact area posed significant challenges for track prediction. Our LSTM model demonstrated high accuracy in predicting Mawar's path, with an average error of 21.84 km (Fig. 8).

Table 7
| Influence of attention mechanism on performance of models predicting tropical cyclone tracks.

Attention	RNN	RNN+Att	LSTM	LSTM+Att	GRU	GRU+Att
MAE	0.610	0.600	**0.438**	0.460	0.471	0.488
MSE	4.932	4.240	**2.778**	3.852	3.071	3.250
RMSE	2.221	2.060	**1.667**	1.963	1.753	1.801
R^2	0.981	0.980	**0.988**	0.984	0.985	0.984
MAPE	1.498	1.629	**1.101**	1.131	1.296	1.322

Table 8
| 2023 tropical cyclones test cases.

Typhoon Name	Date Formed	Date Dissipated	Track Error (km)
2023-MAWAR	2023–05–19	2023–06–02	21.84
2023- GUCHOL	2023–06–06	2023–06–12	34.94
2023-TALIM	2023–07–13	2023–07–18	38.95
2023-DOKSURI	2023–07–20	2023–07–31	37.56
2023-KHANUN	2023–07–26	2023–08–11	27.25
2023-LAN	2023–08–07	2023–08–17	22.62
2023-DORA	2023–08–01	2023–08–22	72.03
2023-SAOLA	2023–08–23	2023–09–04	26.12
2023-DAMREY	2023–08–23	2023–08–29	53.01
2023-HAIKUI	2023–08–27	2023–09–10	36.30
2023-KIROGI	2023–08–30	2023–09–06	58.27
2023-KOINU	2023–09–29	2023–10–09	21.57
2023-BOLAVEN	2023–10–07	2023–10–14	30.30

Typhoon Doksuri formed on July 20, 2023, and quickly intensified into a super typhoon. It reached peak intensity on July 25 and made landfall in the Philippines on July 26, followed by a second landfall in China's Fujian province on July 28. Doksuri's abrupt and significant track changes tested the model's ability to adapt to dynamic trajectory shifts. The LSTM model predicted Doksuri's path with an average error of 37.56 km (Fig. 9).

Typhoon Saola formed east of the Philippines on August 23, 2023, and rapidly intensified into a super typhoon. It made landfall in Guangdong province, China, on September 2, causing strong winds and heavy rainfall in southern China. Saola's complex path and large intensity fluctuations provided a challenging test case for the model's adaptability. The LSTM model predicted Saola's path with an average error of 26.12 km (Fig. 10).

In the experiments, the model rapidly completed the track prediction tasks, taking only 8.23 s, 7.76 s, and 8.16 s for the predictions of Mawar, Doksuri, and Saola, respectively. This rapid prediction capability is crucial for meteorological monitoring and early warning systems, enabling near real-time updates of typhoon track information and providing a valuable time window for emergency response and disaster preparedness. In addition to response speed, rapid computation also implies optimized use of computing resources. In resource-limited environments, such as mobile meteorological stations or field emergency response centers, the ability to provide predictions within a short time using fewer computing resources is an important consideration in model design.

In summary, the LSTM model demonstrated high accuracy and adaptability in predicting the tracks of typhoons in 2023. Detailed analyses of Mawar, Doksuri, and Saola validated the model's effectiveness and robustness in handling various types of typhoon track predictions, showcasing its potential for real-time applications.

4. Conclusion and discussions

This research successfully demonstrates the immense potential of deep learning technologies in predicting typhoon tracks. We conducted a comprehensive comparison and evaluation of the performance of various deep learning models, including RNN, LSTM, and GRU, in predicting typhoon tracks in the Northwest Pacific region. Additionally, we explored the impact of different model parameters (such as input data length and prediction window size) and model complexity on prediction results. Through specific typhoon case analyses, our models not only exhibited outstanding prediction accuracy but also maintained high efficiency in processing large amounts of data. These achievements not only prove the practicality of deep learning in typhoon track prediction but also provide new effective tools for disaster prevention, mitigation, and emergency response. Based on our research, we have drawn the following conclusions:

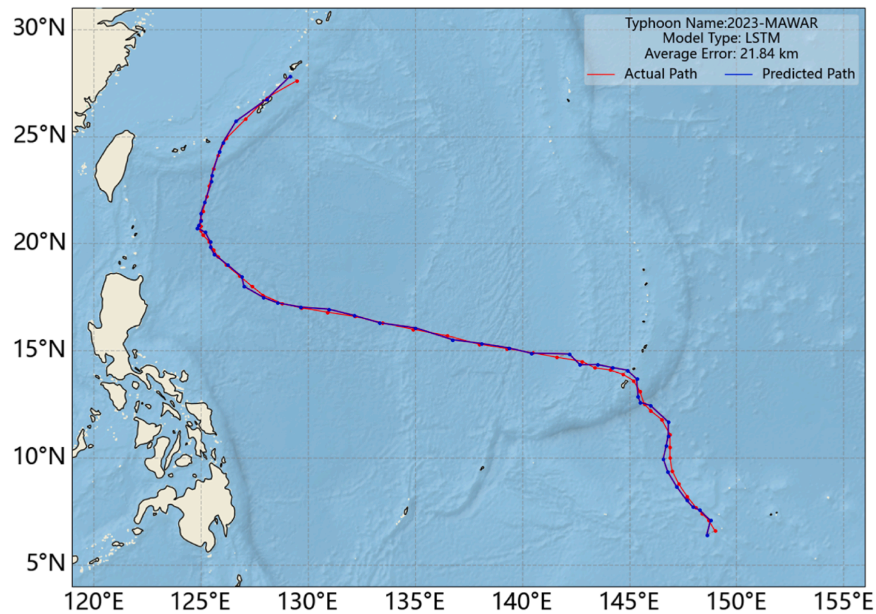

Fig. 8. Comparative visualization of the predicted and actual tracks of typhoon Mawar with quantified error analysis.

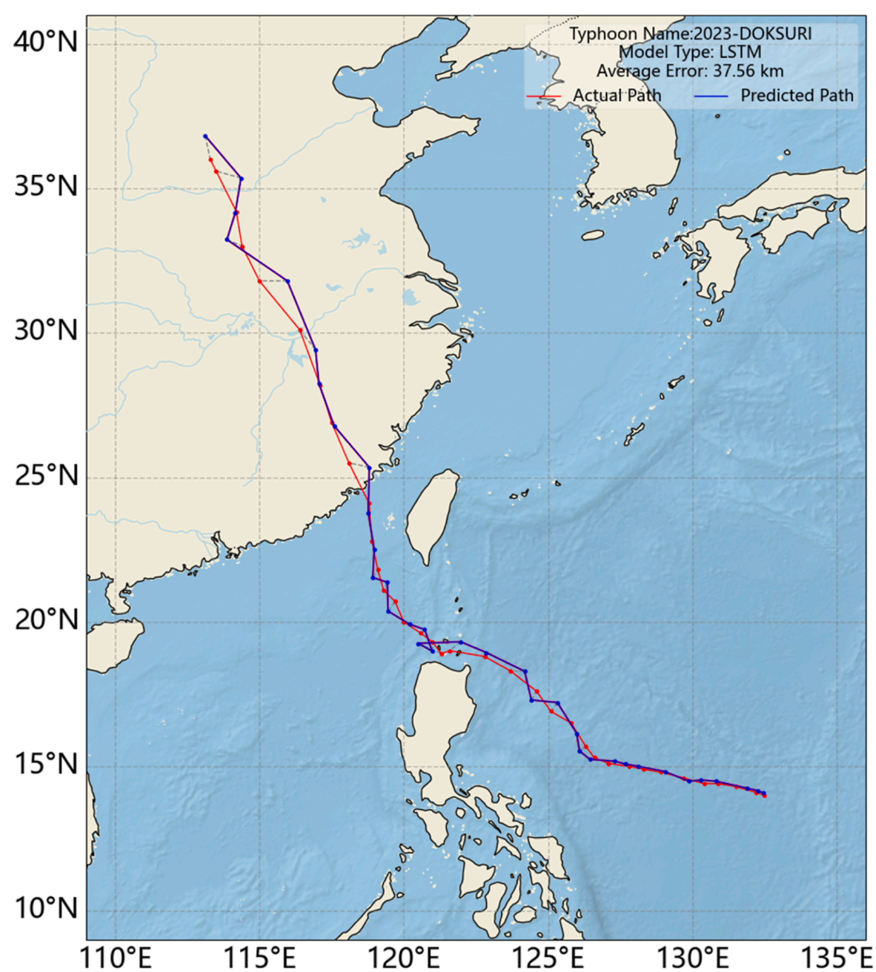

Fig. 9. Comparative visualization of the predicted and actual tracks of typhoon Doksuri with quantified error analysis.

(1) RNN: Suitable for handling time series data with strong short-term dependencies, but it suffers from the vanishing gradient problem when dealing with long-term dependencies in tropical cyclone track prediction. This leads to a decrease in accuracy, particularly when long time spans are required. RNN models may perform well in short-term path predictions, but their

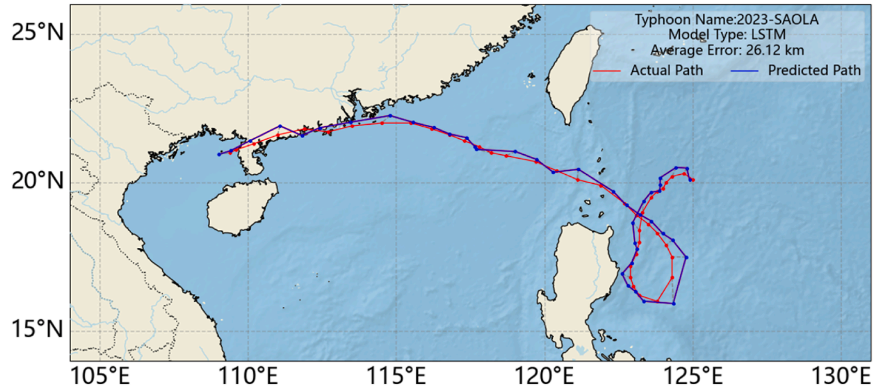

Fig. 10. Comparative visualization of the predicted and actual tracks of typhoon saola with quantified error analysis.

effectiveness significantly declines in long-term predictions. Due to overall performance limitations compared to LSTM and GRU, RNN is not the preferred choice for tropical cyclone track prediction.

(2) LSTM: By introducing forget gates, input gates, and output gates, LSTM effectively addresses the issue of long-term dependencies. LSTM is suitable for tropical cyclone track prediction when long-term memory is required, especially in scenarios where path prediction needs to account for complex weather changes and long-time span information. Our research results indicate that LSTM significantly outperforms RNN models in predicting tropical cyclone tracks during the Northwest Pacific tropical cyclone season.

(3) GRU: While simpler in structure than LSTM, GRU achieves similar effectiveness in handling long-term dependencies and boasts higher computational efficiency. GRU models are particularly advantageous in real-time tropical cyclone track prediction applications that require rapid response. Due to its simplified structure, GRU consumes fewer computational resources while maintaining high prediction accuracy. Our research shows that GRU performs on par with LSTM in predicting tropical cyclone tracks during the Northwest Pacific tropical cyclone season, making it a reasonable choice for real-time prediction and resource-constrained environments.

Despite the progress made, our research has some non-negligible limitations. Firstly, the performance of our models relies heavily on the quality and quantity of data. Incompleteness or inaccuracies in the dataset can significantly impact the effectiveness of model training and the accuracy of predictions. Secondly, while deep learning models excel in analyzing complex meteorological data, their interpretability remains a challenge, especially when decision-makers need to understand the reasons behind model predictions. Moreover, the adaptability and generalization capability of the models to new or anomalous weather patterns not seen in historical data requires further research and improvement. To address this, future work should focus on:

(1) Integrating Physical Dynamics: Our future research will aim to incorporate more of the physical mechanisms underlying typhoon formation and evolution. By embedding principles from meteorological sciences, particularly those related to typhoon dynamics such as thermal structures, wind patterns, and moisture variations, we can enrich the data inputs and enhance the models' contextual understanding.

(2) Expanding Data Diversity: In line with integrating physical dynamics, expanding our dataset to include a wider range of meteorological variables will be crucial. This approach will allow for a more comprehensive analysis of how various factors contribute to typhoon track development.

Through these efforts, we aspire to advance our understanding of deep learning applications in tropical cyclone track prediction and move toward more accurate and reliable models, providing strong support for reducing the impacts of natural disasters.

CRediT authorship contribution statement

Peng Hao: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Yaqi Zhao:** Formal analysis, Data curation, Conceptualization. **Shuang Li:** Writing – review & editing, Methodology. **Jinbao Song:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation. **Yu Gao:** Software.

Declaration of competing interest

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.

Data availability

The link is given in the article.

Funding information

This work was supported by the National Key Research and Development Program of China [grant number 2023YFC3008202] and the National Natural Science Foundation of China [grant number 41830533].

References

Aberson, S.D., 1998. Five-day tropical cyclone track forecasts in the North Atlantic basin. Weather Forecast. 13 (4), 1005–1015.
Alemany, S., Beltran, J., Perez, A., Ganzfried, S., 2019. Predicting hurricane trajectories using a recurrent neural network. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 468–475.
Bauer, P., Thorpe, A., Brunet, G., 2015. The quiet revolution of numerical weather prediction. Nature 525 (7567), 47–55.
Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., Tian, Q., 2023. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619 (7970), 533–538.
Bishop, C.H., Toth, Z., 1999. Ensemble transformation and adaptive observations. J. Atmospheric Sci. 56 (11), 1748–1765.
Buizza, R., Leutbecher, M., Isaksen, L., 2008. Potential use of an ensemble of analyses in the ECMWF ensemble prediction system. Q. J. R. Meteorol. Soc. 134 (637), 2051–2066.
Chen, J.-H., Lin, S.-J., Zhou, L., Chen, X., Rees, S., Bender, M., et al., 2019a. Evaluation of tropical cyclone forecasts in the next generation global prediction system. Mon. Weather Rev. 147 (9), 3409–3428.

Chen, R., Wang, X., Zhang, W., Zhu, X., Li, A., Yang, C, 2019b. A hybrid CNN-LSTM model for typhoon formation forecasting. Geoinformatica 23, 375–396.

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint.

Dey, R., Salem, F.M., 2017. Gate-variants of gated recurrent unit (GRU) neural networks. In: Proceedings of the 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS). IEEE, pp. 1597–1600.

Easterling, D.R., Meehl, G.A., Parmesan, C., Changnon, S.A., Karl, T.R., Mearns, L.O., 2000. Climate extremes: observations, modeling, and impacts. Science 289 (5487), 2068–2074.

Gao, S., Zhao, P., Pan, B., Li, Y., Zhou, M., Xu, J., et al., 2018. A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network. Acta Oceanolog. Sin. 37, 8–12.

Giffard-Roisin, S., Yang, M., Charpiat, G., Kumler Bonfanti, C., Kégl, B., Monteleoni, C, 2020. Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data. Front. Big Data 1.

Graves, A. (2012). Long short-term memory. Supervised sequence labelling with recurrent neural networks, 37–45.

Gray, W.M., 1968. Global view of the origin of tropical disturbances and storms. Mon. Weather Rev. 96 (10), 669–700.

Haiden, T., Janousek, M., Vitart, F., Bouallègue, Z.B., Ferranti, L., Prates, F., et al., 2018. Evaluation of ECMWF forecasts, Including the 2018 Upgrade. European Centre for Medium Range Weather Forecasts Reading, UK.

Hao, P., Li, S., Song, J., Gao, Y., 2023. Prediction of sea surface temperature in the South China Sea based on deep learning. Remote Sens. 15 (6), 1656.

Higa, M., Tanahara, S., Adachi, Y., Ishiki, N., Nakama, S., Yamada, H., et al., 2021. Domain knowledge integration into deep learning for typhoon intensity classification. Sci. Rep. 11 (1), 12972.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Hodges, K., Klingaman, N., 2019. Prediction errors of tropical cyclones in the western North Pacific in the Met Office global forecast model. Weather Forecast. 34 (5), 1189–1209.

Jiang, G.Q., Xu, J., Wei, J., 2018. A deep learning algorithm of neural network for the parameterization of typhoon-ocean feedback in typhoon forecast models. Geophys. Res. Lett. 45 (8), 3706–3716.

LeCun, Y., Bengio, Y.J.T.H.O.B.T., and Networks, N. (1995). Convolutional networks for images, speech, and time series. 3361(10), 1995.

Lu, X., Yu, H., Ying, M., Zhao, B., Zhang, S., Lin, L., et al., 2021. Western North Pacific tropical cyclone database created by the China Meteorological Administration. Adv. Atmos. Sci. 38, 690–699.

Matsuoka, D., Nakano, M., Sugiyama, D., Uchida, S., 2018. Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model. Prog. Earth Planetary Sci. 5 (1), 1–16.

Neumann, C.J. (1972). An alternate to the HURRAN (hurricane analog) tropical cyclone forecast system. NOAA Technical Memorandum NWS SR-62, 24.

Peduzzi, P., Chatenoux, B., Dao, H., De Bono, A., Herold, C., Kossin, J., et al., 2012. Global trends in tropical cyclone risk. Nat. Clim. Change 2 (4), 289–294.

Pradhan, R., Aygun, R.S., Maskey, M., Ramachandran, R., Cecil, D.J., 2017. Tropical cyclone intensity estimation using a deep convolutional neural network. IEEE Trans. Image Process. 27 (2), 692–702.

Song, T., Li, Y., Meng, F., Xie, P., Xu, D., 2022. A novel deep learning model by Bigru with attention mechanism for tropical cyclone track prediction in the Northwest Pacific. J. Appl. Meteorol. Climatol. 61 (1), 3–12.

Wang, L., Wan, B., Zhou, S., Sun, H., Gao, Z., 2023. Forecasting tropical cyclone tracks in the northwestern Pacific based on a deep-learning model. Geosci. Model. Dev. 16 (8), 2167–2179.

Wang, Y., Zhang, W., Fu, W., 2011. Back Propogation (BP)-neural network for tropical cyclone track forecast. In: Proceedings of the 2011 19th International Conference on Geoinformatics. IEEE, pp. 1–4.

Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., et al., 2014. An overview of the China Meteorological Administration tropical cyclone database. J. Atmos. Oceanic Technol. 31 (2), 287–301.

Yu, Y., Si, X., Hu, C., Zhang, J., 2019. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. 31 (7), 1235–1270.

Zaremba, W., Sutskever, I., and Vinyals, O. (2014). Recurrent neural network regularization. arXiv preprint (2014).

Zhang, H., Jing, M., Zhang, H., Li, L., Zheng, Y., Tang, J., et al., 2023. Deep learning approach for forecasting sea surface temperature response to tropical cyclones in the Western North Pacific. Deep Sea Res. Part I 197, 104042.