

Article

# Improvement of Typhoon Intensity Forecasting by Using a Novel Spatio-Temporal Deep Learning Model

Shuailong Jiang<sup>1,2,†</sup>, Hanjie Fan<sup>1,†</sup> and Chunzai Wang<sup>1,3,\*</sup> 

<sup>1</sup> State Key Laboratory of Tropical Oceanography (South China Sea Institute of Oceanology, Chinese Academy of Sciences), Guangzhou 510301, China

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup> Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 510301, China

\* Correspondence: cwang@scsio.ac.cn

† These authors contributed equally to this work.

**Abstract:** Typhoons can cause massive casualties and economic damage, and accurately predicting typhoon intensity has always been a hot topic both in theory and practice. In consideration with the spatial and temporal complexity of typhoons, machine learning methods have recently been applied in typhoon forecasting. In this paper, we attempt to improve typhoon intensity forecasting by treating it as a spatio-temporal problem in the deep learning field. In particular, we propose a novel typhoon intensity forecasting model named the Typhoon Intensity Spatio-temporal Prediction Network (TITP-Net). The proposed model takes multidimensional environmental variables and physical factors of typhoons into account and fully extracts the information from the datasets by capturing spatio-temporal dependencies with a spatial attention module, which includes two-dimensional and three-dimensional convolutional operations. A series of experiments with a comprehensive framework by using TITP-Net are conducted. The MAEs of the forecasts with 18, 24, 36 and 48 h lead time obtain a significant improvement by 7.02%, 6.53%, 6.25% and 5.37% compared with some existing deep learning models and dynamical models from official agencies.

**Keywords:** deep learning; typhoon intensity; spatio-temporal model; rolling forecast; ConvLSTM



**Citation:** Jiang, S.; Fan, H.; Wang, C.

Improvement of Typhoon Intensity Forecasting by Using a Novel Spatio-Temporal Deep Learning Model. *Remote Sens.* **2022**, *14*, 5205.  
<https://doi.org/10.3390/rs14205205>

Academic Editors: Chunyan Li,  
Changming Dong, Guoqi Han and  
Jingsong Yang

Received: 14 September 2022

Accepted: 13 October 2022

Published: 18 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Typhoons are one of the most catastrophic natural disasters, bringing torrential rainfall and violent gusts that put the economy, people's lives, and property in jeopardy [1,2]. According to reports from the China Meteorological Administration (CMA), coastal regions of China are vulnerable to typhoons and suffer the effects of seven typhoons per year on average. Therefore, accurate typhoon forecasting is becoming increasingly crucial and effective in the prevention of natural disasters [3,4]. Recently, typhoon track prediction has advanced significantly due to the improvement of dynamical models, data assimilation techniques and observation technology, but predicting typhoon intensity remains a difficult task [5–8].

Changes in typhoon intensity are influenced by variables in the surrounding environment under the control of complicated physical processes, making it difficult to accurately predict typhoon intensity [9–11]. Although the definition of typhoon intensity varies depending on the oceanic zone, the maximum wind speed and minimum pressure are the most used metrics [12–14]. Both variables have little bearing on the development of typhoon prediction models [15]. In this paper, the typhoon intensity is defined as the 2-min average wind speed.

Currently, there are three types of typhoon intensity forecasting methods: numerical dynamical models, statistical regression models and deep learning models. Numerical dynamical models are based on dynamical theory, which requires a large amount of computer processing power to solve complex formulas [15–17]. Additionally, the typhoon

intensity change is attributable to the nonlinear interaction of environmental variables, which may not be fully described in a numerical dynamical model, limiting the prediction accuracy [18]. Statistical regression models are based on historical typhoon records and related variables from some agencies, including the Japan Meteorological Agency (JMA) and the Joint Typhoon Warning Center (JTWC). They are more efficient than numerical dynamical models in terms of computer processing power consumption but too difficult to accurately predict typhoon intensity [5,19,20]. Deep learning technology is a new type of scientific instrument that can be trained to conduct operations such as converting language, recognizing images, and forecasting time series. Many studies have verified that artificial intelligence has significant advantages in processing large amounts of complex data, and it has been widely used in research on typhoon forecasting. In terms of accuracy, it outperforms both numerical dynamical and statistical regression models [21–25]. For example, Yuan et al. proposed a typhoon intensity forecasting model based on long short-term memory (LSTM), which uses a rolling forecast method to make multistep forecasts [24]. By constructing a particular wide and deep framework, Xu et al. built a spatial-temporal model that combined 2D typhoon structure domain-expert knowledge and 3D typhoon structure domain-expert knowledge [26]. Zhang et al. developed a neural network framework to capture the spatio-temporal characteristics of multisource environmental variables, with notable results [27]. Wang et al. proposed a prediction algorithm for typhoon intensity changes by exploring the joint spatial features of 3D environment conditions, which improved the prediction of typhoon intensity [28]. These studies shed light on the potential of deep learning techniques for improving the accuracy and efficiency of predicting typhoon intensity. On the other hand, these studies did not pay enough attention to the complex spatio-temporal interaction of environmental factors, which is crucial for accurate typhoon intensity prediction.

From the perspective of deep learning, predicting typhoon intensity is a problem of spatio-temporal regression based on multiple dimensional variables, which is a challenging issue at present. Whether typhoon intensity can be forecast accurately depends on two aspects. First, typhoon intensity is influenced by a variety of factors, including air-sea interaction and other processes. Making full use of multiscale environmental information will improve typhoon intensity forecasting. Second, statistical approaches and some deep learning algorithms seldom consider typhoon intensity forecasting as a spatio-temporal problem, resulting in erroneous model forecasts.

In this paper, we propose a new typhoon intensity spatio-temporal prediction network (TITP-Net). Before the training procedure, we use a spatial attention module (SAM) incorporating 2D and 3D convolutional operations (2DConv and 3DConv) to better comprehend the spatial interaction of multiscale environmental factors. A unique neural network TITP-Net is then constructed by using the SAM and Convolutional LSTM (ConvLSTM), which can extract features from different dimensional data and explore the spatio-temporal relations in the process of typhoons.

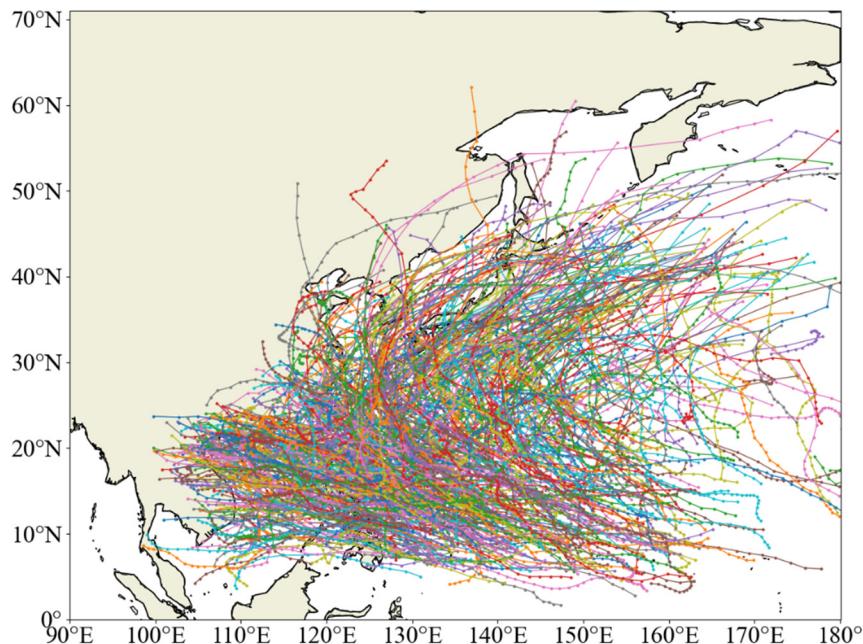
The rest of the paper is organized as follows. Section 2 discusses the datasets used in the work. Section 3 presents the methods, including the proposed network in detail, the forecasting method, and the evaluation metrics. Section 4 demonstrates the experimental results based on the detailed methodological and operational principles underlying the TITP-Net model. Section 5 presents the hybrid model performance compared with other existing methods. Finally, Section 6 provides the conclusion.

All the experiments were carried out with PyTorch 1.10.0 and TensorFlow 2.7.0 on a computer equipped with an Intel(R) Core (TM) i7-8700 CPU@3.20GHZ processor, GeForce GTX 1050Ti GPU.

## 2. Data

The Western North Pacific (WNP) is the region with the highest frequency of tropical cyclones (TCs) in the world, with an average of 27 typhoons per year [29]. Among the WNP and South China Sea coastal countries, China is the most hit by typhoons. Therefore,

typhoons in the WNP (90–180°E, 0–65°N; Figure 1) from 2001–2018 are investigated in this study. The typhoon intensity data are from the CMA. The multidimensional environmental variable datasets are from the European Centre for Medium-Range Weather Forecasts (ECMWF), reanalysis data, with a spatial resolution of 0.25° and temporal resolution of 6 h. Datasets at 0000, 0600, 1200 and 1800 UTC from 2001–2018 were obtained. The local environment is extracted from the area centered on the typhoon's latitude and longitude positions. Previous studies on tropical cyclogenesis helped determine the size of the study region. Zeng et al. chose a radius of 5° around the storm center to measure atmospheric variables [30–32]. Fu et al. and Peng et al. used a center ranging from 10 × 10° to 20 × 20° to compare developing and non-developing disturbances for tropical cyclone formation in the Northern Hemisphere summertime (June–September) from 2003 to 2008 [33,34]. In this paper, we use an area of 10 × 10° around the typhoon center as the environment area.

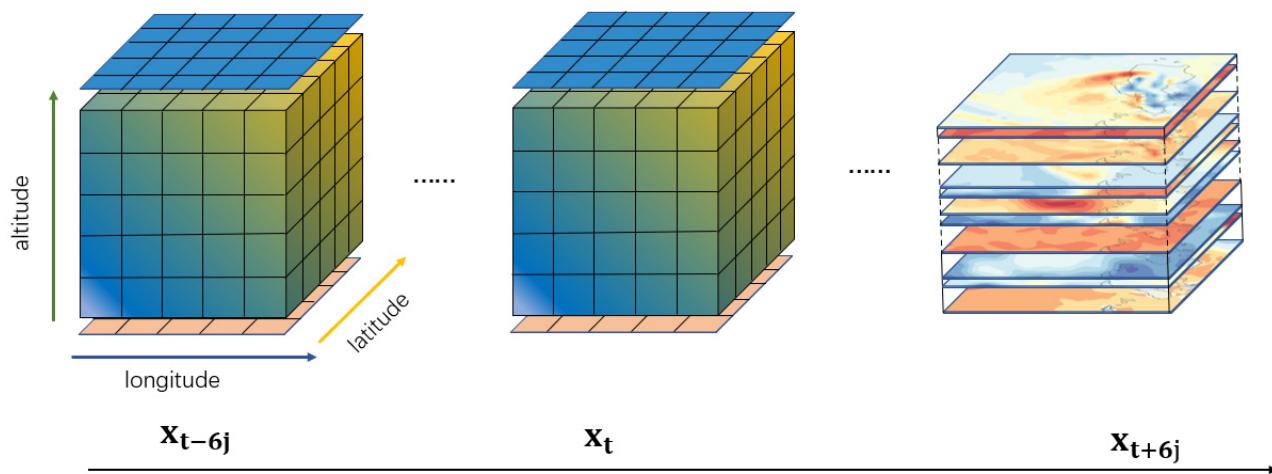


**Figure 1.** Typhoon tracks from 2001–2018 in the Western North Pacific. Different colored lines represent different typhoon tracks.

TC intensity change is affected by a combination of complicated physical processes. Previous research has shown that atmospheric and ocean characteristics are linked to typhoon intensity development [30–32,35–38]. Some studies have also proven that integrating different dimension variables into a dataset can improve intensity predictions [39]. The following environmental factors in the WNP region are used as predictors of typhoon intensity forecasting from ECMWF, as shown in Table 1, including atmospheric data, the potential vorticity (600 hPa), relative vorticity (925 hPa), and (vector) vertical wind shear (200–700 hPa), zonal wind speed (200 hPa), divergence (925 hPa), air temperature (300 hPa), relative humidity (600 hPa) and vertical velocity (200/300/400/500/600/700 hPa); oceanic data, sea surface temperature (SST) [27,28,34,40–42]. Among them, vertical velocity (200/300/400/500/600/700 hPa) is 3D data, and the rest is 2D data. Figure 2 demonstrates how we used multidimensional variables to create a comprehensive typhoon intensity environmental field to extract the temporal and spatial correlations between the various impact components. The dataset is chronologically divided into two parts: training data (dataset from 2001–2014) and test data (dataset from 2015–2018).

**Table 1.** Large-scale environmental conditions as predictors of TC formation in the WNP region. Node represents the name of abbreviations. In this case, t-1 represents the last 6 h. t-2 represent the last 12 h, etc.

Variables	Node	Time			
		t-1	t-2	t-3	t-4
Latitude of storm center ( $^{\circ}$ N)	LAT	LAT (t-1)	LAT (t-2)	LAT (t-3)	LAT (t-4)
Longitude of storm center ( $^{\circ}$ W)	LON	LON (t-1)	LON (t-2)	LON (t-3)	LON (t-4)
2 min mean maximum near center wind speed (m/s)	WIND	WIND (t-1)	WIND (t-2)	WIND (t-3)	WIND (t-4)
Sea Surface Temperate	SST	SST (t-1)	SST (t-2)	SST (t-3)	SST (t-4)
divergence (925 hPa)	DIV	DIV (t-1)	DIV (t-2)	DIV (t-3)	DIV (t-4)
relative vorticity (925 hPa)	REL_VORE	REL_VORE (t-1)	REL_VORE (t-2)	REL_VORE (t-3)	REL_VORE (t-4)
potential vorticity (600 hPa)	POT_VORE	POT_VORE (t-1)	POT_VORE (t-2)	POT_VORE (t-3)	POT_VORE (t-4)
Relative humidity (600 hPa)	REL_HUM	REL_HUM (t-1)	REL_HUM (t-2)	REL_HUM (t-3)	REL_HUM (t-4)
Vertical velocity (200/300/400/500/600/700 hPa)	VER_VEL	VER_VEL (t-1)	VER_VEL (t-2)	VER_VEL (t-3)	VER_VEL (t-4)
(vector) vertical wind shear (200–700 hPa)	SHEAR	SHEAR (t-1)	SHEAR (t-2)	SHEAR (t-3)	SHEAR (t-4)
Air temperature (300 hPa)	TEMP	TEMP (t-1)	TEMP (t-2)	TEMP (t-3)	TEMP (t-4)
Zonal wind (200 hPa)	WIND_U	WIND_U (t-1)	WIND_U (t-2)	WIND_U (t-3)	WIND_U (t-4)



**Figure 2.** Structure of the spatio-temporal environment field. From top to bottom: 2D atmospheric data, 3D atmospheric data, 2D ocean data.

To prevent network convergence failure, the dataset is normalized before training the model to unify the unit of measurement. Normalized data processing can speed up the process of finding the best gradient descent solution and enhance forecast accuracy.

Equation (1) shows the dataset normalization method used in the paper. Additionally, the data are denormalized using Equation (2) after training the model.

$$X(i) = (0.1 \times (\max - x(i)) + 0.9 \times (x(i) - \min)) / (\max - \min), \quad (1)$$

$$x(i) = (X(i) \times (\max - \min) + 0.9 \times \min - 0.1 \times \max) / 0.8, \quad (2)$$

In Equations (1) and (2),  $X(i)$  is the value after normalization,  $x(i)$  is the value before normalization, and  $\min$  and  $\max$  are the minimum and maximum values in the group.

### 3. Methods

This section focuses on the details of the TITP-Net hybrid model. Section 3.1 explains how spatial features are extracted from the environmental disturbance using the SAM. Section 3.2 introduces learning the spatio-temporal information of typhoon intensity and environmental disturbance using ConvLSTM. Section 3.3 displays the rolling forecast-

ing method. Section 3.4 summarizes the overall framework of TITP-Net hybrid model. Section 3.5 indicates the model evaluation metrics.

### 3.1. Learning Spatial Features of Multidimensional Environmental Factors by Using a Spatial Attention Module

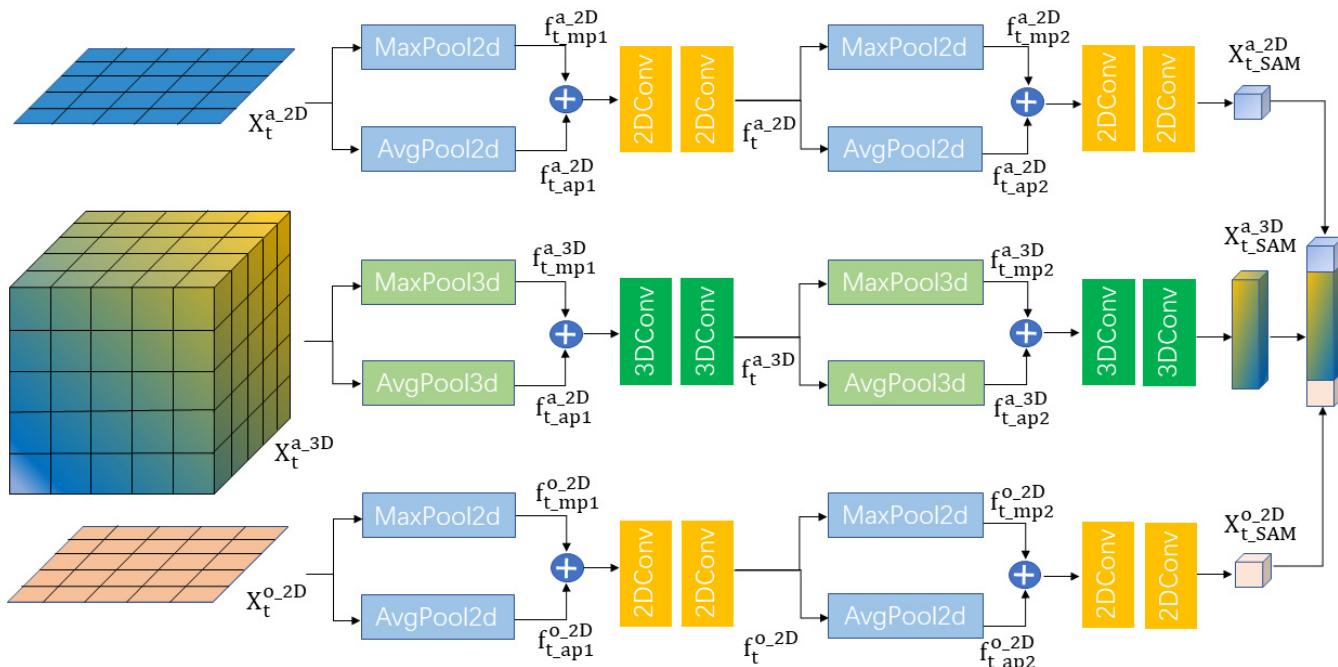
Typhoon intensity is affected by many factors and making full use of multidimensional factors will improve typhoon intensity forecasting performance. Integrating multidimensional variables into a dataset is conducive to improving intensity predictions. Previous research has shown that fully connected and convolution processes may extract features effectively [43]. In this paper, we propose a spatial attention module (SAM) to capture the spatial relationship between variables, as shown in Figure 3. Multidimensional factors determining typhoon intensity are taken into account, and several feature extraction approaches are utilized for various dimensional data. The typhoon intensity environmental field is input to the SAM, and 2DConv is used to extract spatial features from 2D data such as SST and REL\_HUM, whereas 3DConv is utilized to extract features from 3D data. At time t, the SAM receives  $X_t^{a\_2D}$ ,  $X_t^{a\_3D}$  and  $X_t^{o\_2D}$  as input for stressing crucial areas, as shown in Equations (3)–(5):

$$X_{t\_SAM}^{a\_2D} = \text{SAM}(X_t^{a\_2D}), \quad (3)$$

$$X_{t\_SAM}^{a\_3D} = \text{SAM}(X_t^{a\_3D}), \quad (4)$$

$$X_{t\_SAM}^{o\_2D} = \text{SAM}(X_t^{o\_2D}), \quad (5)$$

where  $X_t^{a\_2D}$  and  $X_t^{o\_2D}$  are 2D atmospheric factors and 2D ocean factors, respectively,  $X_t^{a\_3D}$  is the 3D atmospheric factor, and SAM is the spatial attention module.  $X_{t\_SAM}^{a\_2D}$ ,  $X_{t\_SAM}^{a\_3D}$  and  $X_{t\_SAM}^{o\_2D}$  are refined features, which are calculated using Equations (6)–(11).



**Figure 3.** Detailed structure of the spatial attention module (SAM). Including MaxPool2d, AvgPool2d, MaxPool3d, AvgPool3d, 2DConv and 3DConv. From top to bottom: the feature extraction of 2D atmospheric factors, the feature extraction of 3D atmospheric factors and the feature extraction of 2D ocean factors.

$$f_t^{a\_2D} = \text{2DConv}(\text{2DConv}([\text{AvgPool2d}(X_t^{a\_2D}); \text{MaxPool2d}(X_t^{a\_2D})])) \quad (6)$$

$$X_{t,SAM}^{a,2D} = \text{2DConv}(\text{2DConv}([\text{AvgPool2d}(f_t^{a,2D}); \text{MaxPool2d}(f_t^{a,2D})])), \quad (7)$$

$$f_t^{a,3D} = \text{3DConv}(\text{3DConv}([\text{AvgPool3d}(X_t^{a,3D}); \text{MaxPool3d}(X_t^{a,3D})])), \quad (8)$$

$$X_{t,SAM}^{a,3D} = \text{3DConv}(\text{3DConv}([\text{AvgPool3d}(f_t^{a,3D}); \text{MaxPool3d}(f_t^{a,3D})])), \quad (9)$$

$$f_t^{o,2D} = \text{2DConv}(\text{2DConv}([\text{AvgPool2d}(X_t^{o,2D}); \text{MaxPool2d}(X_t^{o,2D})])), \quad (10)$$

$$X_{t,SAM}^{o,2D} = \text{2DConv}(\text{2DConv}([\text{AvgPool2d}(f_t^{o,2D}); \text{MaxPool2d}(f_t^{o,2D})])), \quad (11)$$

For 2D atmospheric factors  $X_{t,SAM}^{a,2D}$ , a MaxPool2d operation and an AvgPool2d operation are adopted to generate two 2D maps  $f_{t_mp1}^{a,2D}$  and  $f_{t_ap1}^{a,2D}$ . Next, after two 2DConv operations, a 2D spatial map  $f_t^{a,2D}$  is fused and input to a MaxPool2d operation and an AvgPool2d operation to generate two 2D maps  $f_{t_mp2}^{a,2D}$  and  $f_{t_ap2}^{a,2D}$ . Finally, the refined feature  $X_{t,SAM}^{a,2D}$  is fused by two 2DConv operations, as shown in Equations (6) and (7). Similarly, Equations (8) and (9) are applied to 3D atmospheric factors, and the spatial attention features of 2D ocean factors are calculated using Equations (10) and (11).

### 3.2. Learn the Spatio-Temporal Feature of Large-Scale Environmental Factors by Using ConvLSTM

To learn the spatio-temporal relationships of multidimensional variables, a ConvLSTM network is used. ConvLSTM was first proposed to solve the precipitation proximity prediction problem [44]. It can not only establish temporal relations as LSTM but can also describe local spatial features as a convolutional neural network (CNN). In terms of obtaining spatio-temporal relations, ConvLSTM outperforms both LSTM and CNN-LSTM. The method replaces the input-to-state and state-to-state parts of LSTM with the form of convolution by feedforward calculation. The connection between the input and each gate is replaced by feedforward convolution, and the operation between states is also changed. With the addition of the convolution operation, not only tempo-relation can be obtained but also spatial features can be effectively extracted, similar to the convolution layer. ConvLSTM has been used to solve a variety of temporal and spatial challenges, including precipitation forecasting. The calculation of ConvLSTM is shown in Equations (12)–(17).

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i), \quad (12)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f), \quad (13)$$

$$\tilde{C}_t = \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_i), \quad (14)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t, \quad (15)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o), \quad (16)$$

$$H_t = o_t \circ \tanh(C_t), \quad (17)$$

where  $X_t$  is the input at time  $t$ ;  $H_t$  represents the network output;  $\tilde{C}_t$  represents the candidate values of the storage unit status;  $W$  is the weight matrix and  $b$  represents the deviation vector matrix;  $\circ$  is the multiplication of the corresponding elements of the matrix, also known as the Hadamard product;  $*$  denotes the convolution operation;  $i_t$  represents the value of the input gate, and  $o_t$  represents the value of the output gate;  $C_t$  stands for memory cells, which not only retain the current input features but also control whether the previous moment of information continues to pass;  $\tanh$  is the hyperbolic tangent function; and  $\sigma$  represents the sigmoid function.

### 3.3. Rolling Forecast Method

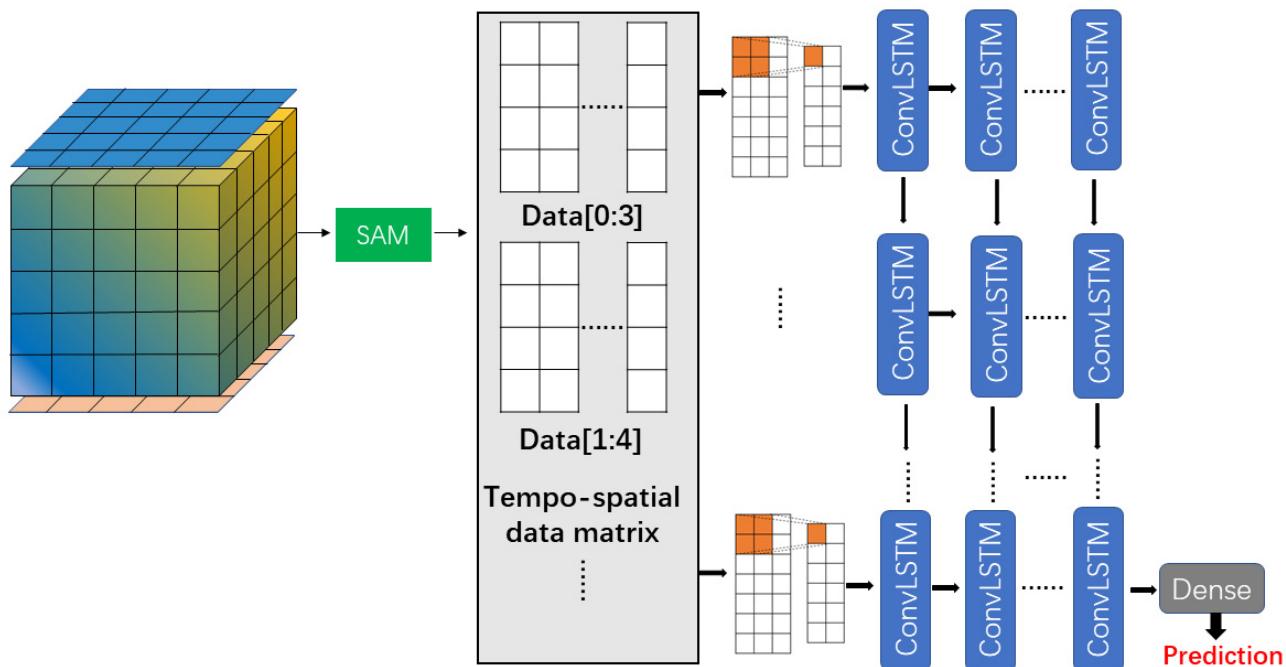
Forecasting methods based on deep learning include single-step prediction and multistep prediction. Most previous studies on TC intensity prediction using deep learning approaches produced single-step predictions [13,15]. The single-step method can only provide prediction at a particular lead time and cannot efficiently capture the detailed

evolution of TCs. Multistep prediction can predict values with long lead time, although with increasing prediction error as time step increases. In simple multistep prediction, numerous environmental variables are used for predicting TC intensity, and only output TC intensity in the next few time steps [27]. As a result, simple multistep prediction cannot catch the relevance of the two adjacent prediction values. To overcome this problem, we adopt a modified multistep approach, the rolling prediction method, which predicts not only typhoon intensity but also environmental variables in the future time steps and then uses the predicted environmental variables to further make prediction until the target time step.

The length of input time steps is defined as sequence length. Next, in the rolling prediction method, the TITP-Net model uses the environmental data and typhoon wind speed of a particular sequence length to forecast the environmental data and typhoon wind speed step by step. Assuming present time is  $t$ , we use the input spatio-temporal matrix  $X_{t-s+1:t}$  ( $s=1, 2, 3 \dots$ , sequence length) to predict  $Y_{t+1}$ , then combine the predicted  $Y_{t+1}$  and  $X_{t-s+2:t}$  as a new input spatio-temporal matrix to predict  $Y_{t+2}$ , and repeat it similarly until  $Y_{t+p}$  ( $p=1, 2, 3 \dots$ , target time step). Through the above process, the typhoon wind speed can be obtained with lead times of 6 h ( $p = 1$ ), 12 h ( $p = 2$ ), 18 h ( $p = 3$ ), 24 h ( $p = 4$ ) or even longer ( $p > 4$ ).

### 3.4. Framework of the Hybrid TITP-Net

The overall framework of TITP-Net is shown in Figure 4. The model is built with ConvLSTM and a SAM. As described in Section 3.1, the input  $X_t$  consists of 2D atmospheric data, 3D atmospheric data and 2D ocean data, which can be expressed as  $X_t = [X_t^{a\_2D}, X_t^{a\_3D}, X_t^{o\_2D}]$ . In the SAM, 2DConv is adopted to extract features from  $[X_t^{a\_2D}, X_t^{o\_2D}]$ , while 3DConv is adopted to extract features from  $[X_t^{a\_3D}]$ , learning the spatial features of atmospheric and oceanic data.



**Figure 4.** Overview of the proposed TITP-Net. The model input is the combined spatio-temporal environmental field. This diagram is a TITP-Net visualization based on the input data with timestep of 4.

The variables at time  $t$  are combined to fuse  $X_t$ . Next,  $X_t$  is input to ConvLSTM to capture the spatio-temporal data matrix for further forecasting. The detailed operation is shown in Sections 3.1–3.3 and is summarized as Equation (18):

$$Y_{t+1:t+p} = \text{TITP-Net} \left( X_{t-s+1:t}^{a\_3d}, X_{t-s+1:t}^{a\_2d}, X_{t-s+1:t}^{o\_2d} \right), \quad (18)$$

where TITP-Net represents the proposed model,  $X_{t-s+1:t}^{a\_3d}$ ,  $X_{t-s+1:t}^{a\_2d}$ ,  $X_{t-s+1:t}^{o\_2d}$  is the fused input matrix,  $Y_{t+1:t+p}$  are predicted values, and  $Y_{t+p}$  is the predicted result at the target time step  $p$ .

### 3.5. Evaluation Metrics

The model performance is evaluated based on two widely employed classic error criteria, the mean value of the absolute errors (MAE) and the root mean square error (RMSE). These two metrics are defined as follows:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (19)$$

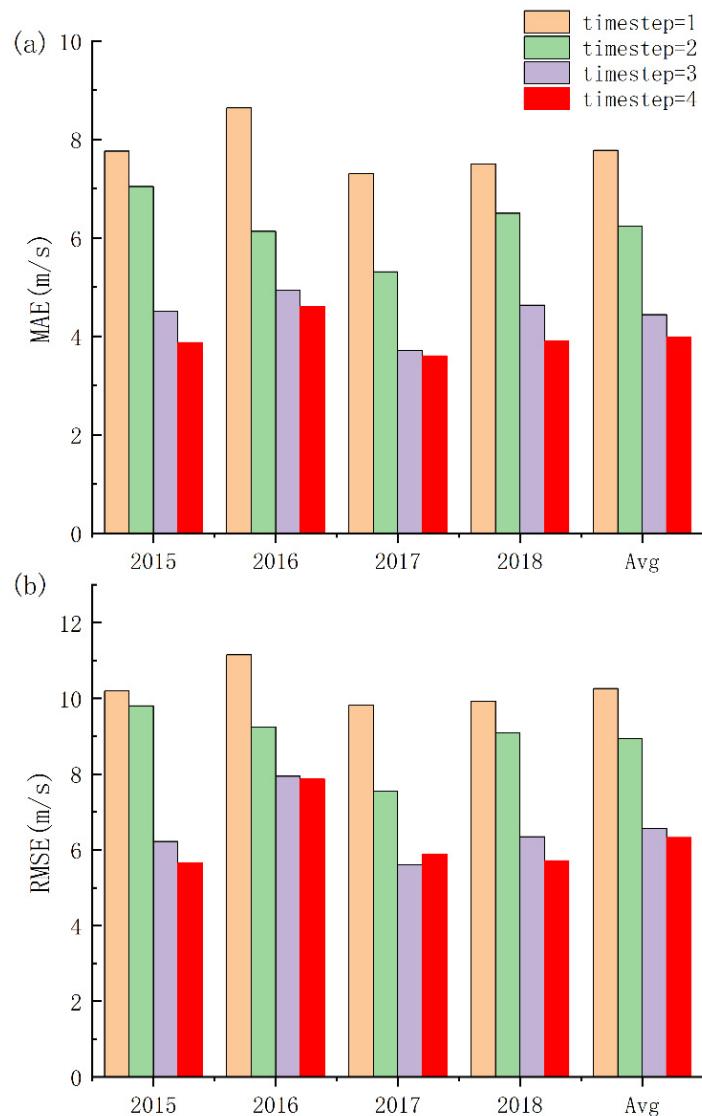
$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (20)$$

where  $y_i$  is the true value of the  $i$ th timestep;  $\hat{y}_i$  is the prediction value; and  $N$  is the number of samples. Both the MAE and RMSE are used to measure the deviation between the true value and the predicted value. The performance is better if the values of the MAE and RMSE are smaller.

## 4. Experimental Results

### 4.1. Results with Different Sequence Lengths

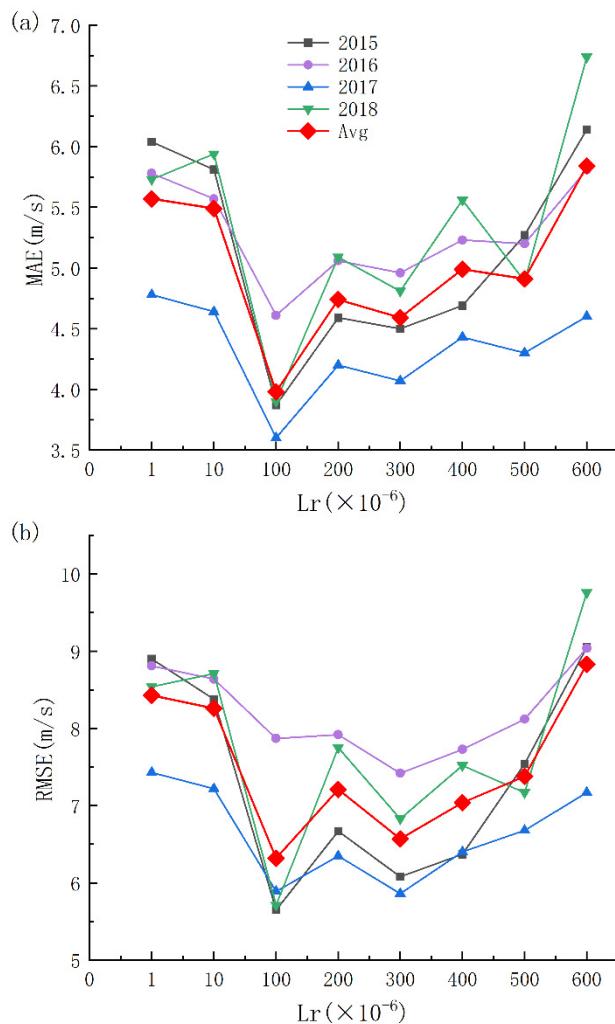
Input sequence length is an important parameter for learning spatio-temporal features, and different sequence lengths will result in various forecasting errors. Figure 5 demonstrates the 24 h ( $p = 4$ , see Section 3.3) forecasting results of different sequence lengths from 2015–2018. Considering the demand of timely predicting typhoon intensity, we use the 6 h ( $s = 1$ ), 12 h ( $s = 2$ ), 18 h ( $s = 3$ ) and 24 h ( $s = 4$ ) typhoon records as input and compare the model prediction performance at different sequence lengths. Overall, inputting 24 h ( $s = 4$ ) data achieves the best performance among the 4 sequence lengths we considered. Take 2015 as an example, the MAE results for sequence length of 6 h, 12 h, 18 h, and 24 h are 7.76 m/s, 7.04 m/s, 4.51 m/s, and 3.87 m/s, respectively, and the RMSE results are 10.20 m/s, 9.80 m/s, 6.22 m/s and 5.65 m/s, respectively. From 2015 to 2018, when inputting 6 h ( $s = 1$ ) typhoon records, the MAEs are 7.76 m/s, 8.64 m/s, 7.30 m/s and 7.50 m/s. When inputting 24 h ( $s = 4$ ) typhoon records, the model obtains better performance with MAEs of 3.87 m/s, 4.61 m/s, 3.60 m/s and 3.90 m/s. With more records input, the model can learn more data features and obtain more accurate results. The average MAE and RMSE of inputting 24 h ( $s = 4$ ) typhoon records are 3.98 m/s and 6.32 m/s, respectively, which are smaller than the others in most cases.



**Figure 5.** The experimental results of 24 h typhoon intensity forecasting with different sequence lengths based on the TITP-Net model. (a) Evaluation of the MAE from 2015–2018 and the four-year average error. Different colored histograms represent different input sequence lengths. The histograms for each sequence length in different years represent typhoon intensity forecasting results in 24h. The red bar represents the optimal sequence length with the smallest error. (b) Evaluation of the RMSE from 2015–2018 and the four-year average error.

#### 4.2. Results with Different Learning Rates

The learning rate (Lr) determines whether and when the objective function converges to the local minimum. With an appropriate learning rate, the objective function can converge to the local minimum in the shortest amount of time. Figure 6 shows the model performance based on different Lrs from 2015–2018. As shown in Figure 6a, when Lr is 0.0001, the MAE results are 3.87 m/s, 4.61 m/s, 3.6 m/s and 3.9 m/s from 2015 to 2018, and the four-year average MAE is 3.98 m/s, which is smaller than the other learning rates. From Figure 6b, it can be seen that the RMSE results of the model are minimal when the learning rate is 0.0001, and the four-year average RMSE is 6.32 m/s. Therefore, the best Lr for the model is 0.0001 based on our examination.



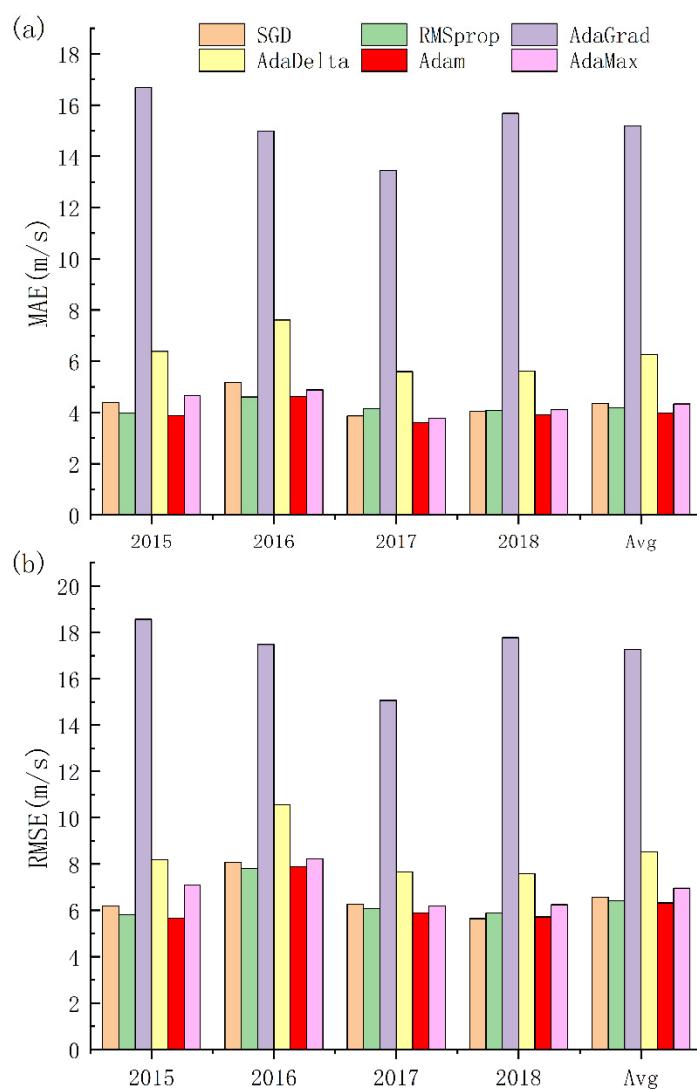
**Figure 6.** The forecasting results of 24 h ahead with various Lrs from 2015–2018 based on the TITP-Net model. (a) Evaluation of the MAE from 2015–2018 and the four-year average error. The point plots of each color represent the 24 h typhoon intensity forecast errors and the four-year average error of different Lrs in 2015–2018, respectively. The red point plot shows the four-year average error in various Lrs. (b) Evaluation of the RMSE from 2015–2018 and the four-year average error.

#### 4.3. Results with Different Optimizers

The optimizer is a crucial step for deep learning, and it is used to minimize (or maximize) the loss function by updating and calculating network parameters that affect model training and model output to approximate or reach optimal values. If a poor optimizer is chosen, it will have a great impact on the model results and dampen the learning efficiency. Some common optimizers are considered in this study, including stochastic gradient descent (SGD) [45], Root Mean Square prop (RMSprop) [46], AdaGrad [47], AdaDelta [48], adaptive moment estimation (Adam) [46,49,50] and AdaMax [50–52].

Figure 7 illustrates the typhoon intensity forecasting results with different optimizers over 24 h. From Figure 7a, the MAE of the model with the Adam optimizer has the best performance in 2015, 2017 and 2018, which are 3.87 m/s, 3.60 m/s and 3.90 m/s, respectively. The four-year forecasting average MAE with the Adam optimizer is 3.98 m/s, which is an improvement of 5.01% over the second-best performance. From Figure 7b, the RMSE of the model with the Adam optimizer has the best performance in 2015 and 2017, which are 5.65 m/s and 5.89 m/s, respectively. The four-year forecasting average RMSE is 6.32 m/s, which is an improvement of 1.56% over the second-best performance.

In plain words, Adam is suitable for typhoon intensity forecasting. The best performance of the Adam optimizer is mainly because Adam is the combination of Adaptive Gradient (AdaGrad) and RMSprop, which basically solves a series of problems of gradient descent, such as a random small sample, adaptive learning rate, and easy to become stuck in the point of a small gradient.

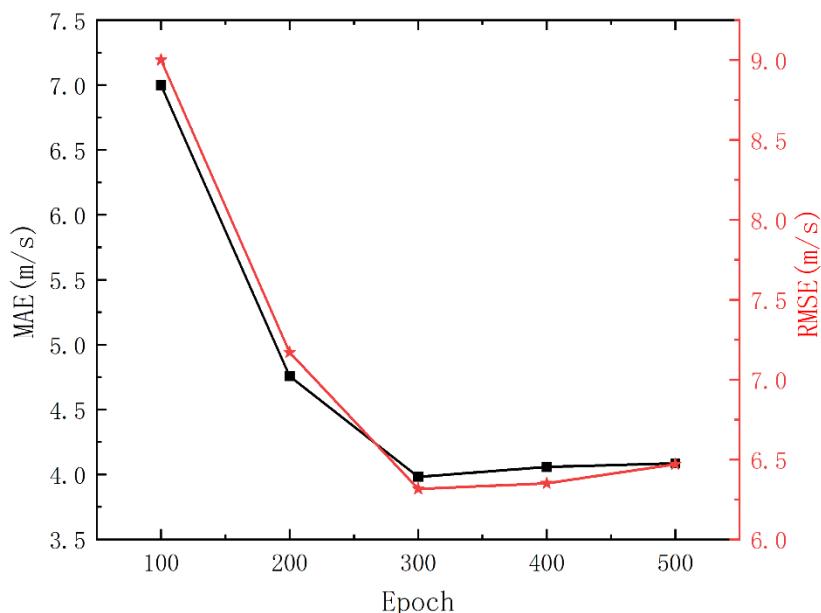


**Figure 7.** Typhoon intensity forecasting results with different optimizers over 24 h. (a) Evaluation of the MAE metric in different years and the four-year average error. Different colored histograms represent different optimizers. The histograms for each optimizer in different years represent typhoon intensity forecasting results in 24 h. The red histograms show the error of the best optimizer forecast for different years. (b) Evaluation of the RMSE metric in different years and the four-year average error.

#### 4.4. Results with Different Epochs

The epoch represents the dataset training time, which is critical for typhoon intensity forecasting. If the dataset passes too few times through the neural network, it will be underfit; otherwise, if it is passed too many times, it will be overfit. Figure 8 demonstrates the forecasting results in 24 h with the change in epochs based on the TITP-Net model. After 500 epochs, the model may be overfit. Therefore, the epochs range from 100 to 500, and the decrease rate is 100. After training the model with 300 epochs, the four-year average MAE and RMSE in 24 h are 3.98 m/s and 6.32 m/s, respectively. With 400 epochs, the four-year average MAE and RMSE in 24 h are 4.06 m/s and 6.35 m/s, respectively.

The TITP-Net model achieved the best performance based on 300 epochs compared to 400 epochs. Therefore, using 300 epochs is suitable for training the model.



**Figure 8.** TITP-Net model forecasting results with the change of epochs in 24 h. The black line represents the forecasting MAE in 24 h with various epochs. The red line represents the forecasting RMSE in 24 h with various epochs.

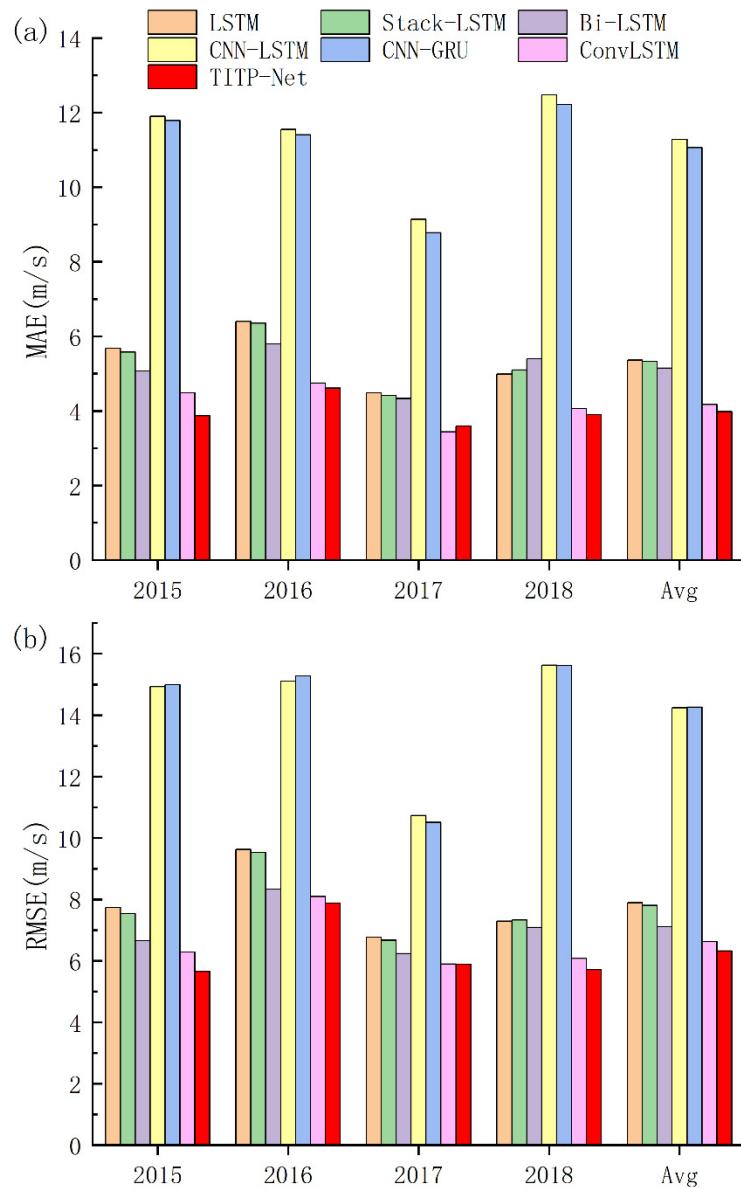
## 5. Results for Typhoon Intensity Forecasting

We used two types of baselines to validate the performance of the proposed TITP-Net model. One is other machine learning methods, and the other is up-to-date typhoon intensity forecasting reports from official agencies. Section 5.1 compares the results of some common deep learning models with the TITP-Net performance in 24 h prediction. Section 5.2 uses some prediction reports and some new forecasting results as baselines to evaluate the performance of the TITP-Net model in 24 h prediction. Section 5.3 demonstrates the typhoon intensity prediction performance in other timesteps, including 6 h, 12 h, 18 h, 24 h, 36 h and 48 h.

### 5.1. Performance Analysis of Various Deep Learning Models

To verify the proposed model effectiveness and practicability, some common deep learning models are introduced to the experiment, including LSTM, Stack-LSTM, Bi-LSTM, CNN-LSTM, CNN-GRU and ConvLSTM, as shown in Figure 9. Among these models, the CNN-LSTM and CNN-GRU models achieve the worst performance. The major cause is that the two models are originally designed for single-step prediction without considering multistep prediction using factor dependencies, which cannot balance the results among multiple prediction intervals. The performance results of the LSTM, Stack LSTM and Bi-LSTM models are better than those of the CNN-LSTM and CNN-GRU models but worse than those of the TITP-Net model. The main reason here is that LSTM cannot capture typhoon spatial features effectively. The TITP-Net model outperforms the second-best performance ConvLSTM, with MAE improvements of 13.62%, 2.95%, and 3.94% and RMSE improvements of 10.03%, 2.72%, and 6.09% in 2015, 2016 and 2018, respectively. Comparing the four-year average forecasting error with those models, the proposed TITP-Net model achieved a better performance with MAE improvements of 25.75%, 25.33%, 22.72%, 64.72%, 64.01% and 4.56% and RMSE improvements of 19.90%, 18.97%, 11.11%, 55.62%, 55.68% and 4.53% than those models. Overall, the results of TITP-Net demonstrate that it can effectively utilize the information from the sequence of environmental variables and the

relationship of multidimensional variables to improve the accuracy and performance of typhoon intensity prediction.



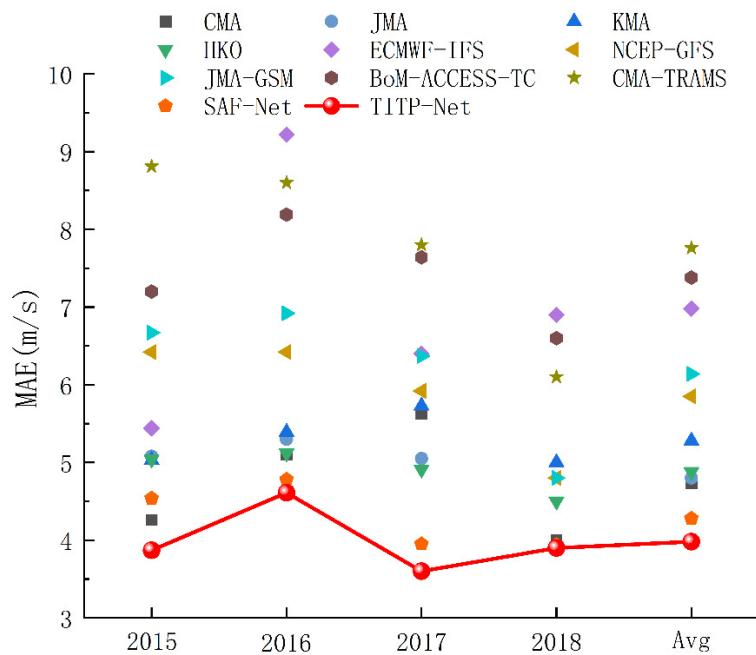
**Figure 9.** The results of typhoon intensity prediction at a lead time of 24 h based on different models. **(a)** Evaluation of the MAE metric in different years and the four-year average error. Different colored histograms represent different models, including LSTM, Stack-LSTM, Bi-LSTM, CNN-LSTM, CNN-GRU, ConvLSTM and TITP-Net. Each histograms shows the performance of typhoon intensity in 24 h prediction in different years. The red histograms show the prediction error in 24 h of the best-performance model. **(b)** Evaluation of the RMSE metric in different years and the four-year average error.

##### 5.2. The Results of Different Methods in 24 h Prediction from 2015–2018

In this Section, we utilize WMO Typhoon Committee 2015 [53], WMO Typhoon Committee 2016 [54], WMO Typhoon Committee 2017 [55], and WMO Typhoon Committee 2018 [56], released by an institution named the Typhoon Committee, which is jointly held by the Economic and Social Commission for Asia and the Pacific (ESCAP) and the World Meteorological Organization (WMO). We select the official guidance, including from the CMA, JMA, Korea Meteorological (KMA) and Hong Kong Observatory (HKO); global

models, including the Integrated Forecasting System of European Centre for Medium Range Weather Forecasting (ECMWF-IFS), Global Forecast System of National Centers for Environmental Prediction (NCEP-GFS) and Global Spectral Model of Japan Meteorological Agency (JMA-GSM); and regional models, including the Tropical cyclone model in the Australian Community Climate and Earth-System Simulator Numerical Weather Prediction systems (BoM-ACCESS-TC) and Tropical Regional Atmosphere Model for the South China Sea based on GRAPES (CMA-TRAMS) as the baselines. To further compare the accuracy of typhoon intensity prediction by the model, we quoted Xu's experimental results, and the typhoon intensity prediction errors based on the proposed SAF-Net model at a lead time of 24 h in 2015, 2016, 2017 and 2018 were 4.54 m/s, 4.78 m/s, 3.95 m/s and 3.94 m/s, respectively [26].

The MAE results of different methods are shown in Figure 10. We compare the annual forecast results of different WMO methods and select the lowest error as the WMO prediction accuracy. The minimum MAEs of typhoon intensity prediction from 2015 to 2018 are 4.26 m/s for the CMA, 5.1 m/s for the CMA, 4.91 m/s for the HKO and 3.90 m/s for the JMA, and the four-year average forecasting minimum error is 4.73 m/s for the CMA. In detail, the proposed model does not improve the typhoon prediction at a lead time of 24 h in 2018 but outperforms the typhoon prediction of the WMO from 2015 to 2017, with improvements of 9.15%, 9.61% and 26.68%. The four-year average prediction error is 15.86% lower than the WMO.

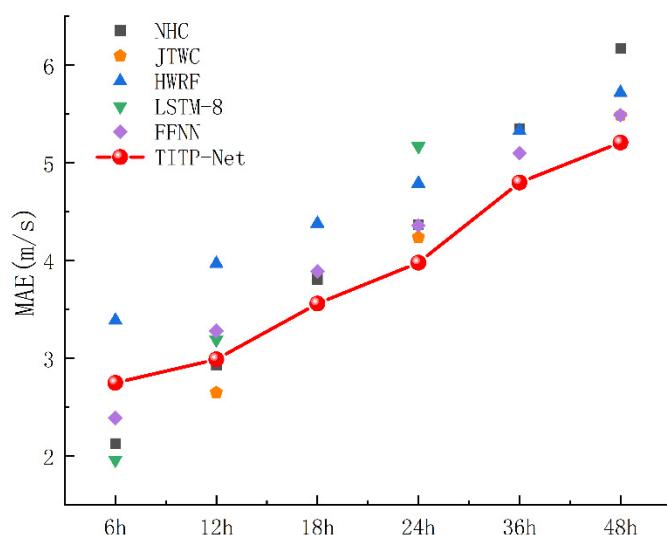


**Figure 10.** Evaluation of the MAE metrics of typhoon intensity prediction at a lead time of 24 h based on different methods from 2015–2018 and the four-year average error. Each point represents the forecasting MAE for each year and the four-year average error. The red point plot represents the performance of TITP-Net in 24 h prediction.

Next, we compare the results with a recent deep learning method, SAF-Net. The SAF-Net model obtains better results, with improvements of 6.27% and 19.55% in 2016 and 2017, respectively, and a mean improvement of 9.09% over the WMO. However, the proposed TITP-Net still outperforms SAF-Net, improving the forecasting accuracy by 14.76% in 2015, 3.56% in 2016, 8.86% in 2017, 1.02% in 2018 and 7.44% on average. In general, the TITP-Net model can outperform other methods in typhoon intensity prediction.

### 5.3. Comparison of the MAEs with Different Lead Times

To comprehensively compare the prediction performance of the model with different lead times (target time steps,  $p$ ), we further compare the proposed model with the results from official agencies and deep learning models of other scholars in 6 h ( $p = 1$ ), 12 h ( $p = 2$ ), 18 h ( $p = 3$ ), 24 h ( $p = 4$ ), 36 h ( $p = 6$ ), and 48 h ( $p = 8$ ) predictions. The objective models of official agencies, the National Hurricane Center (NHC) [57,58], JTWC [59] and Hurricane Weather Research and Forecasting (HWRF) [58], are used as the agency baselines to illustrate the practical usability of the proposed model. The results of previous studies are also used as deep learning model baselines to verify its effectiveness, including LSTM-8 [24] and FFNN [58]. The prediction errors of these methods are shown in Figure 11. Evaluation of the MAE metrics of official agencies and other scholar results in 6 h, 12 h, 18 h, 24 h, 36 h and 48 h prediction. Each point represents the forecasting MAE at 6–48 h by different methods. The red point plot represents the results in various hours prediction of TITP-Net. The result of the 6 h prediction using LSTM-8 is slightly better than the results of the official agencies. In the 12 h prediction, the JTWC has the best performance, with an improvement of 8.65% compared to the second-best result. In the 18 h, 24 h, 36 h and 48 h predictions, our model achieves better performance, with values of 7.02%, 6.53%, 6.25% and 5.37%, respectively, compared with the second-best prediction. In summary, the results show that the performance based on TITP-Net is better with longer lead times.



**Figure 11.** Evaluation of the MAE metrics of official agencies and other scholar results in 6 h, 12 h, 18 h, 24 h, 36 h and 48 h prediction. Each point represents the forecasting MAE at 6–48 h by different methods. The red point plot represents the results in various hours prediction of TITP-Net.

## 6. Conclusions

In this paper, typhoon intensity forecasting is treated as a spatio-temporal regression prediction problem in the deep learning field. Some studies have proven that integrating different dimension variables into a dataset can better understand typhoon formation and promise for improving intensity prediction. We propose a novel spatial attention module that includes 2DConv, which is used to capture features in 2D variables, and 3DConv, which is used to capture features in 3D variables. Next, a novel spatio-temporal forecasting model named TITP-Net is developed, with the ConvLSTM and spatial attention module to capture the local correlated features.

TITP-Net is trained and tested using historical typhoon information records from 2001–2018, including typhoon location and intensity from CMA, multidimensional environmental variable datasets from ECMWF. To enable model training, extensive experiments are carried out to choose the optimal model parameters: the optimal input sequence length is 4; the optimal learning rate is 0.0001; the optimal optimizer for improving model performance

is Adam; and, finally, the optimal epoch is 300 epochs. The results in the 24 h prediction are 3.87 m/s in 2015, 4.61 m/s in 2016, 3.60 m/s in 2017, 3.90 m/s in 2018 and 3.98 m/s in the four-year average.

Next, we compare our model with various methods to evaluate its performance. We introduce two types of baselines, machine learning methods and up-to-date typhoon intensity forecasting reports. Compared with machine learning methods in 24 h prediction, the TITP-Net model outperforms the other models, with an MAE improvement of 4.56% and RMSE improvement of 4.53% compared to ConvLSTM. Next, the up-to-date typhoon intensity forecasting reports are regarded as the second baselines, including the typhoon intensity forecasting reports and the recent SAF-Net deep learning method. Our model achieves the best performance, with improvements of 9.15%, 9.61%, and 26.68% in 2015–2017 and 15.86% in mean error compared to the WMO. Compared to SAF-Net, TITP-Net improves the forecasting accuracy by 14.76% in 2015, 3.56% in 2016, 8.86% in 2017, 1.02% in 2018 and 7.44% on average. In general, our proposed model outperforms the baselines.

Finally, we further compare our proposed model with the results from official agencies and the deep learning models of previous studies in 6 h, 12 h, 18 h, 24 h, 36 h and 48 h prediction. The model outperforms the baselines at 18 h, 24 h, 36 h and 48 h lead times, with improvements of 7.02%, 6.53%, 6.25% and 5.37% compared to the second-best prediction.

Our proposed model improves the prediction of typhoon intensity by fully extracting the information from the massive data with a proper method and framework. However, some further work should be carried out in the future. First, we may further make an improvement in the short-term prediction (6 h and 12 h) of typhoon intensity. Second, a more novel network based on a ConvGRU and ST-LSTM may enable a better prediction, which should be adopted in future research.

**Author Contributions:** S.J.: data acquisition; data analysis; data interpretation; study design; writing—original draft; H.F. and C.W. writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is supported by the National Natural Science Foundation of China (42192564), the National Key R&D Program of China (2019YFA0606701), the Strategic Priority Research Program of Chinese Academy of Sciences (XDB42000000 and XDA20060502), and Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0306). The author Hanjie Fan is supported by the National Postdoctoral Program of Innovative Talents (BX2021324).

**Data Availability Statement:** The typhoon intensity data are from the China Meteorological Administration (CMA, [https://tcdatatyphoon.org.cn/zjljsjj\\_zlhq.html](https://tcdatatyphoon.org.cn/zjljsjj_zlhq.html), accessed on 15 October 2022). The multidimensional variable datasets are from the European Centre for Medium-Range Weather Forecasts (ECMWF, <https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset&text=ERA5%20hourly%20data%20on%20pressure%20levels%20from%201979%20to%20present>, accessed on 15 October 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Emanuel, K. Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **2005**, *436*, 686–688. [[CrossRef](#)]
2. Uson, M.A.M. Natural disasters and land grabs: The politics of their intersection in the Philippines following super typhoon Haiyan. *Can. J. Dev. Stud. Rev. Can. Détudes Dév.* **2017**, *38*, 414–430. [[CrossRef](#)]
3. Chen, Z.; Yu, X. A Novel Tensor Network for Tropical Cyclone Intensity Estimation. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 3226–3243. [[CrossRef](#)]
4. Gray, W.M. The formation of tropical cyclones. *Meteorol. Atmos. Phys.* **1998**, *67*, 37–69. [[CrossRef](#)]
5. Demaria, M.; Mainelli, M.; Shay, L.K.; Knaff, J.A.; Kaplan, J. Further Improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Weather Forecast.* **2005**, *20*, 531–543. [[CrossRef](#)]
6. Lian, J.; Dong, P.; Zhang, Y.; Pan, J.; Liu, K. A Novel Data-Driven Tropical Cyclone Track Prediction Model Based on CNN and GRU with Multi-Dimensional Feature Selection. *IEEE Access* **2020**, *8*, 97114–97128. [[CrossRef](#)]
7. Giffard-Roisin, S.; Yang, M.; Charpiat, G.; Kumler Bonfanti, C.; Kégl, B.; Monteleoni, C. Tropical Cyclone Track Forecasting Using Fused Deep Learning from Aligned Reanalysis Data. *Front. Big Data* **2020**, *3*, 1. [[CrossRef](#)]

8. Dong, P.; Lian, J.; Yu, H.; Pan, J.; Zhang, Y.; Chen, G. Tropical Cyclone Track Prediction with an Encoding-to-Forecasting Deep Learning Model. *Weather Forecast.* **2022**, *37*, 971–987. [[CrossRef](#)]
9. Neetu, S.; Lengaigne, M.; Menon, H.B.; Vialard, J.; Mangeas, M.; Menkes, C.E.; Ali, M.M.; Suresh, I.; Knaff, J.A. Global assessment of tropical cyclone intensity statistical-dynamical hindcasts. *Q. J. R. Meteorol. Soc.* **2017**, *143*, 2143–2156. [[CrossRef](#)]
10. Gao, S.; Zhang, W.; Liu, J.; Lin, I.-I.; Chiu, L.S.; Cao, K. Improvement in typhoon intensity change classification by incorporating an ocean coupling potential intensity index into decision trees. *Weather Forecast.* **2016**, *31*, 95–106. [[CrossRef](#)]
11. Zhang, W.; Gao, S.; Chen, B.; Cao, K. The application of decision tree to intensity change classification of tropical cyclones in western North Pacific. *Geophys. Res. Lett.* **2013**, *40*, 1883–1887. [[CrossRef](#)]
12. Xin, J.; Yu, H.; Chen, P. Evaluation of Tropical Cyclone Intensity Forecasts from Five Global Ensemble Prediction Systems During 2015–2019. *J. Trop. Meteorol.* **2021**, *27*, 218–231.
13. Pan, B.; Xu, X.; Shi, Z. Tropical cyclone intensity prediction based on recurrent neural networks. *Electron. Lett.* **2019**, *55*, 413–414. [[CrossRef](#)]
14. Deo, R.V.; Chandra, R.; Sharma, A. Stacked transfer learning for tropical cyclone intensity prediction. *arXiv* **2017**, arXiv:1708.06539.
15. Chen, R.; Wang, X.; Zhang, W.; Zhu, X.; Li, A.; Yang, C. A hybrid CNN-LSTM model for typhoon formation forecasting. *Geoinformatica* **2019**, *23*, 375–396. [[CrossRef](#)]
16. Sobrevilla, K.L.M.D.; Reyes, E.O.; Hendrickx, C.A.; Yao, S.S. Typhoon Forecasting in the Philippines Using an Optimal Multilayer Feedforward Artificial Neural Network Model Trained in Resilient Propagation Algorithm. In Proceedings of the 2016 IEEE Region 10 Conference (TENCON), Singapore, 22–25 November 2016.
17. Yu, H.; Chen, G.; Zhou, C.; Wong, W.; Yang, M.; Xu, Y.; Chen, P.; Wan, R.; Hu, X. Are We Reaching the Limit of Tropical Cyclone Track Predictability in the Western North Pacific? *B Am. Meteorol. Soc.* **2021**, *103*, E410–E428. [[CrossRef](#)]
18. Pu, Z.; Kalnay, E. Numerical weather prediction basics: Models, numerical methods, and data assimilation. In *Handbook of Hydro meteorological Ensemble Forecasting*; Springer: Berlin, Germany, 2019; pp. 1–31.
19. Weber, H.C. Hurricane track prediction using a statistical ensemble of numerical models. *Mon. Weather Rev.* **2003**, *131*, 749–770. [[CrossRef](#)]
20. Goerss, J.S. Tropical cyclone track forecasts using an ensemble of dynamical models. *Mon. Weather Rev.* **2000**, *128*, 1187–1193. [[CrossRef](#)]
21. Farnoosh, A.; Azari, B.; Ostadabbas, S. Deep Switching Auto-Regressive Factorization: Application to Time Series Forecasting. *arXiv* **2020**, arXiv:2009.05135. [[CrossRef](#)]
22. Gong, J.; Qiu, X.; Wang, S.; Huang, X. Information aggregation via dynamic routing for sequence encoding. In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–25 August 2018; pp. 2742–2752.
23. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer, Vision, Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
24. Yuan, S.; Wang, C.; Mu, B.; Zhou, F.; Duan, W. Typhoon Intensity Forecasting Based on LSTM Using the Rolling Forecast Method. *Algorithms* **2021**, *14*, 83. [[CrossRef](#)]
25. Zhou, J.; Xiang, J.; Huang, S. Classification and Prediction of Typhoon Levels by Satellite Cloud Pictures through GC-LSTM Deep Learning Model. *Sensors* **2020**, *20*, 5132. [[CrossRef](#)] [[PubMed](#)]
26. Xu, G.; Lin, K.; Li, X.; Ye, Y. SAF-Net: A spatio-temporal deep learning method for typhoon intensity prediction. *Pattern Recogn. Lett.* **2022**, *155*, 121–127. [[CrossRef](#)]
27. Zhang, Z.; Yang, X.; Shi, L.; Wang, B.; Du, Z.; Zhang, F.; Liu, R. A neural network framework for fine-grained tropical cyclone intensity prediction. *Knowl.-Based Syst.* **2022**, *241*, 108195. [[CrossRef](#)]
28. Wang, X.; Wang, W.; Yan, B. Tropical Cyclone Intensity Change Prediction Based on Surrounding Environmental Conditions with Deep Learning. *Water* **2020**, *12*, 2685. [[CrossRef](#)]
29. Matsuura, T.; Yumoto, M.; Iizuka, S. A mechanism of interdecadal variability of tropical cyclone activity over the western North Pacific. *Clim. Dyn.* **2003**, *21*, 105–117. [[CrossRef](#)]
30. Zeng, Z.; Wang, Y.; Wu, C.C. Environmental dynamical control of tropical cyclone intensity—An observational study. *Mon. Weather Rev.* **2007**, *135*, 38–59. [[CrossRef](#)]
31. Zeng, Z.; Chen, L.; Wang, Y. An Observational Study of Environmental Dynamical Control of Tropical Cyclone Intensity in the Atlantic. *Mon. Weather Rev.* **2008**, *136*, 3307–3322. [[CrossRef](#)]
32. Zeng, Z.; Wang, Y.; Chen, L. A statistical analysis of vertical shear effect on tropical cyclone intensity change in the North Atlantic. *Geophys. Res. Lett.* **2010**, *37*. [[CrossRef](#)]
33. Peng, M.S.; Fu, B.; Li, T.; Stevens, D.E. Developing versus nondeveloping disturbances for tropical cyclone formation. Part I North Atlantic. *Mon. Weather Rev.* **2012**, *140*, 1047–1066. [[CrossRef](#)]
34. Fu, B.; Peng, M.S.; Li, T.; Stevens, D.E. Developing versus nondeveloping disturbances for tropical cyclone formation. Part II Western North Pacific. *Mon. Weather Rev.* **2012**, *140*, 1067–1080. [[CrossRef](#)]
35. Emanuel, K.; DesAutels, C.; Holloway, C.; Korty, R. Environmental control of tropical cyclone intensity. *J. Atmos. Sci.* **2004**, *61*, 843–858. [[CrossRef](#)]
36. Hendricks, E.A.; Peng, M.S.; Fu, B.; Li, T. Quantifying Environmental control on tropical cyclone intensity change. *Mon. Wea. Rev.* **2010**, *138*, 3243–3271. [[CrossRef](#)]
37. Merrill, R.T. Environmental influences on hurricane intensification. *J. Atmos. Sci.* **1988**, *45*, 1678–1687. [[CrossRef](#)]

38. Gray, W.M. Tropical cyclone genesis. In *Department of Atmospheric Science Paper*; Colorado State University: Fort Collins, CO, USA, 1975; p. 121.
39. Chen, R.; Zhang, W.; Wang, X. Machine learning in tropical cyclone forecast modeling: A review. *Atmosphere* **2020**, *11*, 676. [[CrossRef](#)]
40. Wijnands, J.S.; Qian, G.; Kuleshov, Y. Variable selection for tropical cyclogenesis predictive modeling. *Mon. Weather Rev.* **2016**, *144*, 4605–4619. [[CrossRef](#)]
41. Chand, S.S.; Walsh, K.J.E. Forecasting tropical cyclone formation in the Fiji region: A probit regression approach using Bayesian fitting. *Weather Forecast.* **2011**, *26*, 150–165. [[CrossRef](#)]
42. Hendricks, E.A.; Braun, S.A.; Vigh, J.L.; Courtney, J.B. A summary of research advances on tropical cyclone intensity change from 2014–2018. *Trop. Cyclone Res. Rev.* **2019**, *8*, 219–225. [[CrossRef](#)]
43. Zhang, C.; Yang, Z.; He, X.; Deng, L. Multimodal Intelligence: Representation Learning, Information Fusion, and Applications. *IEEE J. Sel. Top. Signal Process.* **2020**, *14*, 478–493. [[CrossRef](#)]
44. Xingjian, S.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems*; 2015; pp. 802–810. Available online: <https://proceedings.neurips.cc/paper/2015/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf> (accessed on 15 October 2022).
45. Bottou, L. Stochastic gradient descent tricks. In *Neural Networks: Tricks of the Trade*; Grégoire, M., Geneviève, B.O., Müller, K.R., Eds.; Springer: New York, NY, USA, 2012; Volume 7770, pp. 421–436.
46. Zou, F.; Shen, L.; Jie, Z.; Zhang, W.; Liu, W. A Sufficient Condition for Convergences of Adam and RMSProp. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 11127–11135.
47. Lydia, A.; Francis, S. Adagrad—An optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci.* **2019**, *6*, 566–568.
48. Zeiler, M.D. ADADELTA: An Adaptive Learning Rate Method. *arXiv* **2012**, arXiv:abs/1604.00133.
49. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
50. Tato, A.; Nkambou, R. Improving ADAM Optimizer. In Proceedings of the Workshop Track-ICLR 2018, Vancouver, BC, Canada, 30 April–3 May 2018.
51. Llugsí, R.; El Yacoubi, S.; Fontaine, A.; Luperá, P. Comparison between Adam, AdaMax and Adam W optimizers to implement a Weather Forecast based on Neural Networks for the Andean city of Quito. In Proceedings of the 2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM), Cuenca, Ecuador, 12–15 October 2021; pp. 1–6.
52. Xiao, B.; Liu, Y.; Xiao, B. Accurate state-of-charge estimation approach for lithium-ion batteries by gated recurrent unit with ensemble optimizer. *IEEE Access* **2019**, *7*, 54192–54202. [[CrossRef](#)]
53. Chen, G.; Lei, X.; Zhang, X.; Chen, P.; Yu, H.; Wan, R. Performance of tropical cyclone forecast in western North Pacific in 2015. *Trop. Cyclone Res. Rev.* **2016**, *5*, 47–57.
54. Chen, G.; Zhang, X.; Chen, P.; Yu, H.; Wan, R. Performance of tropical cyclone forecast in western North Pacific in 2016. *Trop. Cyclone Res. Rev.* **2017**, *6*, 13–25.
55. Chen, G.; Zhang, X.; Yang, M.; Yu, H.; Cao, Q. Performance of tropical cyclone forecast in western North Pacific in 2017. *Trop. Cyclone Res. Rev.* **2021**, *10*, 1–15. [[CrossRef](#)]
56. Chen, G.; Zhang, X.; Bai, L.; Wan, R. Performance of tropical cyclone forecast in western North Pacific in 2018. In Proceedings of the 51th Session ESCAP/WMO Typhoon Committee, Guangzhou, China, 26 February 2019.
57. Cangialosi, J.P.; Blake, E.; DeMaria, M.; Penny, A.; Latto, A.; Rappaport, E.; Tallapragada, V. Recent Progress in Tropical Cyclone Intensity Forecasting at the National Hurricane Center. *Weather Forecast.* **2020**, *35*, 1913–1922. [[CrossRef](#)]
58. Cloud, K.A.; Reich, B.J.; Rozoff, C.M.; Alessandrini, S.; Lewis, W.E.; Delle Monache, L. A feed forward neural network based on model output statistics for short-term hurricane intensity prediction. *Weather Forecast.* **2019**, *34*, 985–997. [[CrossRef](#)]
59. Huang, X.; Peng, X.; Fei, J.; Cheng, X.; Ding, J.; Yu, D. Evaluation and error analysis of official tropical cyclone intensity forecasts during 2005–2018 for the Western North Pacific. *J. Meteorol. Soc. Soc. Jpn.* **2021**, *99*, 139–163. [[CrossRef](#)]