Multi-Factor Intensity Estimation for Tropical Cyclones in the Western North Pacific Based on the Deviation Angle Variance Technique

Wei ZHONG1, Meng YUAN2*, Hexin YE1, and Xia LUO1

1 College of Meteorology and Oceanology, National University of Defense Technology, Nanjing 211101
2 Submarine College of Navy, Qingdao 266199

(Received January 13, 2020; in final form June 7, 2020)

ABSTRACT

In this paper, the infrared cloud images from Fengyun series geostationary satellites and the best track data from the China Meteorological Administration (CMA-BST) in 2015–2017 are used to investigate the effects of two multi-factor models, generalized linear model (GLM) and long short-term memory (LSTM) model, for tropical cyclone (TC) intensity estimation based on the deviation angle variance (DAV) technique. For comparison, the typical single-factor Sigmoid function model (SFM) with the map minimum value of DAV is also used to produce TC intensity estimation. Sensitivity experiments regarding the DAV calculation radius and different training data groups are conducted, and the estimation precision and optimum calculation radius for DAV in the western North Pacific (WNP) are analyzed. The results show that the root-mean-square-error (RMSE) of the single-factor SFM is 8.79–13.91 m s\(^{-1}\) by using the individual years as test sets and the remaining two years as training sets with the optimum calculation radius of 550 km. However, after selecting and using the high-correlation multiple factors from the same test and training data, the RMSEs of GLM and LSTM models decrease to 5.93–8.68 and 4.99–7.00 m s\(^{-1}\) respectively, with their own optimum calculation radii of 350 and 400 km. All the sensitivity experiments indicate that the SFM results are significantly influenced by the DAV calculation radius and characteristics of the training set data, while the results of multi-factor models appear more stable. Furthermore, the multi-factor models reduce the optimum radius within the process of DAV calculation and improve the precision of TC intensity estimation in the WNP, which can be chosen as an effective approach for TC intensity estimation in marine areas.

Key words: deviation angle variance (DAV) technique, tropical cyclone (TC), intensity estimation

Citation: Zhong, W., M. Yuan, H. X. Ye, et al., 2020: Multi-factor intensity estimation for tropical cyclones in the western North Pacific based on the deviation angle variance technique. J. Meteor. Res., 34(5), 1038–1051, doi: 10.1007/s13351-020-9216-5.

1. Introduction

Tropical cyclone (TC) is one of the most catastrophic weather systems in the world (Zhang and Guo, 2008). As the majority of meteorological observations are on the land, whereas the genesis and development of TCs mostly occur over the ocean, it is difficult for us to directly acquire the information on the TC structure, intensity, trajectory, and variation, due to the shortage of observations and limits in data extraction techniques. Therefore, the detection and forecast of TCs have always been the most challenging issue in predicting and warning of hazardous weather. To acquire as much as possible accurate information about TCs on the basis of current methods so as to strengthen the ability in TC monitoring, is the most essential way to improve TC forecasts.

Over the vast ocean, the most effective method to track and analyze TC life cycles relies on the visible and infrared cloud images (Piñeros et al., 2010). At present, the most widely utilized way to analyze TCs with satellite cloud images is the Dvorak technique proposed by Dvorak (1975) in the 1970s. In comparison with the TC eye and its surrounded clouds with a series of standard models of TCs, the Dvorak technique evaluates the TC intensity by a comprehensive analysis of cloud circulations and distribution of brightness temperature. On this
basis, the Advanced Objective Dvorak Technique (AODT) was introduced by Olander and Velden (2007), which integrated subjective analysis statistics and regression equations to predict the TC intensity and also achieved self-located TC centers. The Dvorak technique has been the prevalent analysis method in TC studies because of its good performance in practical application. Although AODT lays more emphasis on the objectivity in the analysis process, its cloud classification still relies on the subjective analysis so that users are required to consecutively analyze cloud images, and thus, the prediction results seriously depend on users’ experience (Piñeros et al., 2010).

Based on the infrared radiation (IR) cloud images for real-time TC monitoring, Piñeros et al. (2008) proposed an objective analysis method—Deviation Angle Variance Technique (DAV-T)—that can differentiate the TC intensity through the morphology and dynamic characters of clouds. As TCs with different intensities present significant differences in their organization processes, DAV-T first quantifies the TC axisymmetry and then establishes the relationship of the obtained characteristic values and TC intensity via the Sigmoid function model (SFM) to realize the real-time monitoring of TC structure and intensity. Note that the calculation radius in DAV-T needs to acclimatize TC scales, and it has obvious impacts on the distributions of DAV values in clouds.

Considering that the large-scale environment and TC formation vary in different TC genesis regions, it is necessary to conduct sensitivity experiments on the calculation radius in corresponding regions (Briegel and Frank, 1997; Ritchie and Holland, 1999; Hill and Lackmann, 2009). Piñeros et al. (2011) applied DAV-T to monitor TCs over the North Atlantic areas during 2004–2009 and employed the map minimum value (MMV) to represent the TC axisymmetric degree. Their results indicated that: (1) the root-mean-square error (RMSE) of TC intensity from predictions and national hurricane center (NHC) records was 24.8 kt, with the 2009 best tracks as test sets; (2) when the radius was calculated every 50 km within the range of 150–500 km to select the optimum parameter for the North Atlantic, the error of DAV-T was relatively small, with the calculation radius between 300 and 400 km. Ritchie et al. (2014) took the DAV of the TC circulation center (E-DAV) as the proper value to represent the TC axisymmetric degree, and found that: (1) for the northwestern Pacific region from 2007 to 2011, RMSEs of the TC intensity from predictions and Joint Typhoon Warning Center (JTWC) records were between 12.9 and 15.1 kt with any one year as the test set and the rest as the training set; (2) similarly, for the northeastern Pacific region, RMSEs of the predicted TC intensity and JTWC records were between 9.4 and 16.9 kt; (3) the error of DAV-T was the minimum with a calculation radius of 300 km and the error reduced further if the radii of 250 and 500 km were taken.

Using the above DAV-T to determine the TC intensity has actually presented obvious flaws since only a single parameter of cloud DAV is considered as predictive factor. When E-DAV was adopted as the predictor for the TC intensity, the performance was much improved, but the location of the TC eye exerted a great influence on the model performance. E-DAV alone failed to reveal the whole TC structure because this value only represented the axisymmetric degree around the TC eye. However, MMV can indicate the axisymmetric degree of deep convective clouds within the whole TC system and will not be affected by the TC location (Yuan and Zhong, 2019).

Research on objective methods of TC intensity estimation has progressed considerably in recent years. The factors were expanded to describe the TC intensity characteristics more completely. Zhang et al. (2016, 2019) extracted the deviation angle gradient co-occurrence matrix (DAGCOM) with the IR image. The relevance vector machine (RVM) was used to build the estimation model that performs much better than the traditional linear regression model. In addition, the optimal calculation radius of factors and optimal numbers of factors to the model are also investigated. On the other hand, intelligent estimation models such as machine learning has significantly improved the accuracy of results. Combinido et al. (2018) used 11.0 μm-channel infrared satellite data as the input and constructed a 2-D Convolutional Neural Network (CNN) to estimate the TC intensity in the western North Pacific (WNP). Moreover, Lee et al. (2020) proved that expanding the input data to the multi-channel infrared is more effective than the single-channel method. Until now, most of the TC intensity estimation methods using deep learning are based on CNN model, which can capture the spatial features from an image and is helpful to identify the object accurately, as well as its relation to other objects in an image. Although the value of TC intensity is closely related to the spatial feature of the TC structure at the same time, each TC case still has its own time evolution influenced by the surrounding atmosphere, external forces, internal dynamic processes, and so on. Thus, the long short-term memory (LSTM) model is introduced to trace the time evolution of the TC intensity. The LSTM model is a modified version of Recurrent Neural Network (RNN), which solves the problems of
gradient disappearance and gradient explosion of RNN, insufficient long-term memory capacity, and so on, and makes the RNN be used effectively in the long-range timing information (Graves, 2012).

Based on the DAV-T and FY series satellite data, this paper discusses the TC intensity estimation techniques over the WNP Ocean. Section 2 presents the data and methodology. The multi-factor models: generalized linear model (GLM), and LSTM model, are introduced and compared with the single-factor SFM in Section 3, which also includes the sensitivity experiments on the calculation radius. Section 4 compares the performance of the single-factor and multi-factor models with the optimum calculation radius. The final section is the conclusion and prospect for future work.

2. Data and methods

2.1 Data

The satellite data we utilized in this paper originates from the longwave (10.7 μm) IR image that are generated by Fengyun (FY) series geostationary satellites of China from 2015 to 2017. Influenced by satellite debugging or regional encryption observation, the single geostationary satellite may not supply the information in some periods. In order to cover the most area of the WNP Ocean, the observations from satellite FY-2F are mainly used here, while images from satellite FY-2G are supplemented to expand the available satellite data. Considering the different coverages of FY-2F and FY-2G, the spatial area of available satellite data is defined to be 0°–40°N, 100°–160°E. The spatial resolution is 0.05° per pixel (approximately 5 km per pixel), and the time resolution is 1 h. Piñeros et al. (2011) find out that reducing the spatial resolution of the research data will not significantly impact the calculation results. Therefore, all available satellite data used here are reshaped with the spatial resolution of 0.1° (approximately 10 km) per pixel to improve the computation efficiency.

The best track data from the China Meteorological Administration (CMA-BST: http://tcdata.typhoon.org.cn/zjiljsij_zlhsq.html) with a temporal resolution of 6 h is used in this paper to analyze the position, intensity, and division of different life history stages of typhoons (TYs). According to the National Standard for TC Class (GB/T 19201-2006), TC can be divided into eight classes: Weaker than Tropical Depression (TD), TD, Tropical Storm (TS), Severe Tropical Storm (STS), TY, Severe TY (STY), Super TY (SuperTY), and Extratropical Cyclone (ET). In order to make full use of the satellite data on high temporal resolution, the cubic spline function is used to interpolate CMA-BST data to the time intervals of the satellite data. Furthermore, it should be noted that, affected by the noise and short-timescale disturbances, physical variables calculated from the satellite data exhibit high-frequency fluctuation, especially during the complex formation process of TCs (Piñeros et al., 2008). Therefore, the Chebyshev Filter with the 0.1π-cutoff frequency (low pass) is used to smooth the calculated variables from satellite data as in Piñeros et al. (2011) in order to obtain a better match with the CMA-BST.

2.2 Methods

The organization degree of TC cloud structures varies significantly at different TC development stages. Observational studies based on satellite data have shown that, TC genesis occurs with the concentration and closure of scattered deep convective cloud clusters; TC develops with the axisymmetric process of organized annular deep convection belts around the eye; and TC weakens with the dissipation of organized convective cloud belts over land or cold ocean surface (Hubert and Timchalk, 1969; Dvorak, 1975). As a result, the symmetry of convective clouds relative to the center of the TC circulation can be regarded as a significant parameter to estimate the TC intensity. An ideal axisymmetric TC system can be supposed as an annulus deep convection belt encircling a cloud-free eye, implying that the corresponding brightness temperature in the satellite image drops off with the radius from the eyewall to periphery. In other words, the direction of all brightness temperature gradient vectors is pointing toward or away from the eye. Hence, the symmetry degree of any vortex system can be quantified through its deviation to the ideal TC system, which is the basic theory of DAV-T. The specific process of DAV-T can be found in Piñeros et al. (2008).

Specifically, the process of DAV-T is as follows. Firstly, calculating the gradient direction of brightness temperature at each pixel on the IR-image. Secondly, choosing any pixel at the IR image as the reference point \( O \), and obtaining the deviation angle \( \theta_i \) between the gradient and radial directions of all pixels within the calculating radius \( R_v \). Then, the variance of all the above deviation angles is defined to be the DAV value of this reference point (as shown in Fig. 1a). Finally, after taking all pixels in the satellite image as the reference point in turn, the DAV values at every pixel constitute the DAV map (Fig. 1b).

Following the above procedure, the IR image at 0600 UTC 13 September 2016 and the corresponding DAV map can be seen in Figs. 1b, d. There are three organized
deep convective cloud clusters with distinct vortex properties respectively in the northern Indochina Peninsula, southeast of Taiwan Island, and west of Marshall Islands, respectively. These cloud clusters correspond to the three TCs existing in the WNP Ocean at that time: the landed TY Rai (1615) at the stage of TS, SuperTY Meranti (1614) moving to the southeast coast of China, and TY Malakas (1616) that had just reached TS. In addition, there are also large areas of vigorous deep convection in the South China Sea, Philippine Peninsula, and central Pacific. Convection in the central Pacific (0°–16°N, 150°–165°E) appears to concentrate as cloud clusters, while the structure of deep convection in other places is relatively loose.

According to the relationship between the DAV and axisymmetric deep convective cloud cluster, the higher the organization degree of the cloud cluster, the better the symmetry, which means the relatively lower DAV value;
on the contrary, the more scattered the cloud, the higher the DAV value (Piñeros et al., 2008). Wood et al. (2015) used 2000 deg² as the maximum value of the organized cloud clusters based on the Geostationary Operational Environmental Satellite (GOES) observation. Considering the characteristics of observation information from FY series satellites, we set the threshold value to be 2200 deg², which can capture all TC records in the CMA-BST dataset even when TCs are at the stage of weaker than TD (Yuan and Zhong, 2019).

The DAV map (Fig. 1d) shows that there are obvious DAV lows within the blue circles, which are centered by the three recorded TCs from CMA-BST with a radius of 350 km. The DAV low related to the SuperTY stage of Meranti (1614) appears to be a circular shape, and its MMV is located at the TC center with the minimum value of 1231 deg². However, when it comes to the DAV distribution of the TS stage of TY Rai (1615) and Malakas (1616) at the same time, although the shape of the two DAV lows still presents a relatively complete circular structure, the positions of MMV deviate to the TC center, which depend more on the location of the axisymmetric deep convection in the TC system. Furthermore, the MMV value of Malakas (1616) and Rai (1615) at the TS stage is 1598 and 1783 deg², respectively.

Meanwhile, the DAV map also indicates that the deep convection over the South China Sea and Philippine Peninsula belong to the unorganized convection. However, there are sporadic DAV lows within the large aggregated cloud clusters in the central Pacific, which suggests that the clusters here may have the possibility to spin up. Hence, through the DAV-T method, we can detect TC by the IR image and quantify its axisymmetric degree, which can provide a valid basis for TC identification and intensity estimation over the ocean.

In this paper, 83 TC cases that occurred in the WNP from 2015 to 2017 are studied to obtain the TC intensity estimation model, including 9014 hourly satellite IR image samples. Four groups (A to D; Table 1) are designed to check the model efficiency. Groups A, B, and C are designed to test the intensity estimation with CMA-BST records in an individual year from 2015 to 2017 and to train the model with the remaining 2-yr samples. Group D is used to obtain the fitting results with all the 3-yr samples.

### 3. Models and results

#### 3.1 Single-factor SFM

Earlier studies have shown that the corresponding cloud cluster becomes more symmetric when the TC intensity is stronger. Therefore, both the DAV value at the TC center and MMV in the regional system correspond significantly well to the maximum wind speed \( V_{\text{max}} \) of the TC (Piñeros et al., 2011). Considering the uncertainties in TC positioning in the operational detection, in this paper, the MMV is used to construct the statistical estimation model between the DAV and \( V_{\text{max}} \) of the TC.

As the Sigmoid function not only coincides with the variation characteristics of the MMV sequence, but also converges at the upper and lower boundaries of \( V_{\text{max}} \), the estimation deviating much from the actual intensity can be avoided, which makes it a general model for TC intensity estimation (Piñeros et al., 2011; Ritchie et al., 2014). The SFM is defined as below,

\[
V_{\text{max}} = I_{\text{min}} + \frac{I_{\text{max}} - I_{\text{min}}}{1 + \exp(\alpha + \beta x)} ,
\]

where \( V_{\text{max}} \) is the fitted value of \( V_{\text{max}} \) obtained from the model; \( I_{\text{max}} \) and \( I_{\text{min}} \) correspond to the upper and lower limits of \( V_{\text{max}} \) in the best track dataset, respectively; \( x \) is the value of MMV; and \( \alpha \) and \( \beta \) are two model parameters obtained by fitting samples. To make the fitting results of the model cover all TC intensities, \( I_{\text{max}} \) and \( I_{\text{min}} \) are taken as 80 and 5 m s\(^{-1}\), respectively. Since the time resolution of satellite IR image data is higher than that of CMA-BST, the high-frequency filtered MMV (FMMV) sequence, obtained by filtering the MMV sequence with Chebyshev Filter referred to in Section 2.1, is used here, and \( V_{\text{max}} \) of the TC system in the CMA-BST dataset is interpolated. Yet the \( V_{\text{max}} \) and DAV values are not corresponding one-to-one, so it is difficult to build a stable model for intensity estimation. To solve this problem, the median of MMV values corresponding to each \( V_{\text{max}} \) in the CMA-BST dataset is taken as the fitting data, thus avoiding the problem of low intensity during the fitting process caused by different sizes of samples with various intensity levels (Piñeros et al., 2011).

Studies show that, the calculation results of DAV and application of DAV-T in the objective estimation of the TC intensity are greatly affected by the radius of calcula-

| Group | Test set | Size (h) | Training set | Size (h) |
|-------|----------|----------|--------------|----------|
| A: G-2015 | 2015 | 4022 | 2016, 2017 | 4992 |
| B: G-2016 | 2016 | 2941 | 2015, 2017 | 6073 |
| C: G-2017 | 2017 | 2051 | 2015, 2016 | 6963 |
| D: G-all | 2015, 2016, 2017 | 9014 | 2015, 2016, 2017 | 9014 |
tion. For instance, during most of the years in the North Atlantic, the optimum calculation radius for DAV-T is within 200–250 km, while it is within 350–400 km in the WNP (Ritchie et al., 2014). This is attributed to the significant differences in TC scales caused by different oceanic environmental fields and development characteristics of TCs at different stages. Thus, if the selected radius differs greatly from the actual TC scale, it would lead to deviations in the magnitude and value of the DAV map, further affecting the accuracy of intensity estimation models.

To compare the influence of different calculation radii on the model construction, the DAV is calculated with the each-time IR image data corresponding to all cases in three years. The calculation is conducted for seven different radii varying from 250 to 550 km at the interval of 50 km. The Sigmoid fitting curve of the TC $V_{max}$ obtained from the FMMV based on different radii for calculation is provided in Fig. 2a. Overall, the Sigmoid function can well reflect the correspondence of the TC $V_{max}$ in the effective intensity range to the FMMV. When the TC intensity is weaker than TD or stronger than STY, the smaller (absolute) slope of the Sigmoid function curve means that within the same TC intensity variation interval, the corresponding DAV difference is more obvious. In other words, this model is significantly better than others in the case of very weak (weaker than TD) or very strong (stronger than STY) TCs.

It can be seen from Fig. 2a that the FMMV corresponding to the equivalent $V_{max}$ tends to increase as the radius increases, indicating that the axisymmetry of the whole system reduces with the increasing calculation radius. Meanwhile, as the radius increases, the gradient of the fitting curve increases obviously when $V_{max}$ is below 60 m s$^{-1}$, while it weakens significantly when the function becomes saturated. Such a distribution characteristic of the function can lead to the following results. When the TC intensity is relatively weak, the larger the calculation radius is, the larger the gradient of the fitting function will be, thereby the corresponding relation between the $V_{max}$ and FMMV becomes more accurate. But for the high-intensity TC, the larger the calculation radius is, the weaker the gradient of the function will be, and it could easily result in the obvious fitting deviation during the model training.

To further examine the application effect of the single-factor SFM on the TC intensity estimation, the test results of the model calculated at different radii are obtained according to the settings of four experiments as mentioned above (Fig. 3). G-2015, G-2016, and G-2017, represent the test data composing of the TC cases in 2015, 2016, and 2017 respectively, while TCs in the other two years are used as the training data, and the results are represented by the RMSE between $\tilde{V}_{max}$ from the results of the model and $V_{max}$ from the CMA-BST dataset in the test year. The result of G-all test is represented by the RMSE between the hindcast $\tilde{V}_{max}$ from the model based on all samples during the period of 2015–2017 and the $V_{max}$ from the CMA-BST dataset. From the test results, it can be seen that the RMSE of the single-factor SFM is within the range of 8.79–13.77 m s$^{-1}$, the average effect of G-2017 performs slightly better than that of

Fig. 2. (a) The Sigmoid-based fitting curve of the FMMV (deg$^2$) obtained from different calculation radii in the G-all test with the maximum wind speed ($V_{max}$; m s$^{-1}$) in the CMA-BST dataset. The scattered points refer to the median of FMMV values corresponding to different $V_{max}$; the solid line is the fitting curve based on the Sigmoid function; and the seven colors represent the results with seven different radii for DAV calculation varying from 250 to 550 km at an interval of 50 km, respectively. (b) The distribution of the RMSE (labeled with the straight line; m s$^{-1}$) of the test results calculated at different radii in the four groups of experiments. Different colors refer to different experiments; G-2015, G-2016, and G-2017 represent the RMSE of the test groups, respectively; and G-all represents the fitting RMSE of the training set.
G-2016, and that of G-2015 performs the worst, and its RMSE at each radius all exceeds 13 m s\(^{-1}\). According to the aforementioned case information, 2015 is the high-occurrence year of SuperTYs, and 15 out of 29 cases reach the level of SuperTY throughout the year; whereas there are only 12 cases of SuperTY in total during 2016 and 2017. Thus, it can be seen that, the accuracy of the high-intensity TC estimation based on the Sigmoid function is affected by both the disappearance of the gradient at boundaries of the function and the limitation in the number of cases, which can lead to considerable errors in estimation.

Considering the test results obtained under different radii, the optimum radius of 450–500 km is found in G-all, G-2015, and G-2016. This optimum radius is larger than that from the tests of similar models in the Atlantic Ocean and mid-East Pacific Ocean, which is accordant with the conclusion given by Knaff et al. (2010) that the TC size in the WNP Ocean is larger than those in other sea areas, statistically. Moreover, it is interesting that, unlike the other tests, the RMSE in the G-2015 test does not decrease monotonically under different radii varying from 250 to 350 km. Its RMSE under the radius of 250 km is smaller than those under the radii of 300 and 350 km. However, the RMSE in the G-2017 test decreases with the increasing calculation radius, while the variation amplitude reduces significantly when the calculation radius is greater than 500 km. Statistical analysis of TC samples shows that in 2015, there are more SuperTYs but fewer low-intensity TC cases (only 6 cases) whose intensities are below the TS. But in 2017 there are only 4 high-intensity cases selected for testing and the number of TC cases with the intensity below the TS reaches 10. In fact, statistical studies show that there is a certain correspondence between the TC size and intensity. Especially, the size of SuperTY is the minimum com-
pared with its size at the other stages, and the size of low-intensity TCs (below the TS intensity) is relatively large. Therefore, in the G-2015 test, the calculation radius of 250 km can better depict the scale of SuperTY; while in the G-2017 test, the larger the calculation radius is, the closer it will be to the size of the low-intensity TC. This also coincides with the fact that the gradient of the Sigmoid function increases with the increasing calculation radius; thus, it is favorable to improve the resolving accuracy of the corresponding relations between the $V_{\text{max}}$ and FMMV. In a word, selection of a calculation radius suitable for the size of TC system facilitates the improvement in the accuracy of TC intensity estimation. Besides, the test results reveal that the errors among different calculation radii exhibit small differences while those among different test groups are relatively large, indicating that the stability of the single-factor SFM is poor and that the differences among the models are significant as the sample data change.

### 3.2 Multi-factor models

Studies have shown that besides the value of MMV, the relative distance (RD) between the location of MMV and the circulation center of the TC, as well as the distribution characteristics of the DAV value, are good indicators for the TC intensity (Yuan and Zhong, 2019). Based on the DAV-T and with consideration of the construction process of the statistical regression model for the TC intensity (Demaria et al., 2005), the accuracy of the TC intensity estimation could be improved by screening and introducing information on the infrared cloud images, TC location, and integration of the DAV and cloud image. Therefore, eight preselected factors (see Table 2 for description) are introduced here. These factors can be classified into three categories. (1) FMMV, MDAV, and RD are symmetrization factors from DAV map, which include the structure of organized deep clouds and their relationship with the circulation center. (2) Minimum black-body brightness temperature (TBBm), S-20, and TBBstd are convection factors from IR images, which reflect the influence of deep convection as in Demaria et al. (2005). (3) Lat and Lon are location factors (Table 2), which indicate the effect of location on the TC intensity.

#### 3.2.1 Factor effects

GLM is an extended form of the general linear model. The feature of GLM is that it is able to establish the relationship between the mathematical expectation of response variables and predictive variables in the form of linear combination through the link function. This method is capable of processing the data without changing their natural metrics, whether they are with nonlinear or non-constant variance. At the same time, while taking into account the single-factor effect, the regression model can also examine the nonlinear interactions between any two predictors (McCullagh, 1984; Sun et al., 2013). Therefore, the GLM regression model can be used to test the factor effects and can be adopted to establish the relationship between the TC intensity and various predictors (Guo et al., 2014), which is defined as below,

$$
\hat{V}_{\text{max}} = b_0 + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} b_{ij} x_i x_j + \varepsilon, \tag{2}
$$

where $\hat{V}_{\text{max}}$ is the fitted value of $V_{\text{max}}$ in the TC system obtained from the model; $x = (x_1, x_2, \cdots, x_n)$ represents each predictor for determining the TC intensity; and $b_i$ ($i = 1, 2, \cdots, n$) is the coefficient of each factor term. Accordingly, $\sum_{i=1}^{n} b_i x_i$ in the formula represents the single-factor effect of each predictor; $\sum_{i=1}^{n} \sum_{j=i+1}^{n} b_{ij} x_i x_j$ represents the interactions between every two predictors; and $\varepsilon$ is the residual.

Before construction of the multi-factor estimation model, contributions of the preselected factors to the intensity estimation should be examined to improve the estimation accuracy. In the multi-factor model, for the $i$-th predictor, the variance contribution is defined as $V_i = b_i^2 / C_{ii}$, where $C_{ii}$ is the element at line $i$ and column $i$ of the matrix $C = (X'X)^{-1}$, and $X$ is the matrix of the preselected factors. Thus, the relative variance contribution rate for the $i$-th factor can be defined as the ratio of the variance contribution of the $i$-th factor to the sum of the variance contribution of all factors; that is,

$$
P_i = V_i / \sum_{i=1}^{n} V_i.$$

First, all the pre-selected factors are introduced into the GLM. The distribution of the factors’ average variance contribution rates to the TC intensity under differ-

### Table 2. Preselected factors for the TC intensity estimation

| Name     | Description                                                                 |
|----------|-----------------------------------------------------------------------------|
| FMMV     | Filtered MMV                                                                |
| MDAV     | Averaged DAV around the MMV location within the radius of 200 km            |
| RD       | The distance between the MMV location and circulation left                 |
| TBBm     | Minimum black-body brightness temperature within the radius of 200 km        |
| TBBstd   | The standard deviation of TBB within the radius of 100–300 km               |
| Lat      | Latitude of circulation left*                                               |
| Lon      | Longitude of circulation left*                                              |

*The circulation left is defined by the local vorticity maximum at 850 hPa calculated with ECMWF reanalysis (ERA)-Interim data.

OCTOBER 2020
Zhong, W., M. Yuan, H. X. Ye, et al. 1045
ent radii is analyzed (Fig. 3a). The dark color represents the variance contribution rate of single factor, and the light color represents the total variance contribution rate of all pairwise interactions involving this factor. As can be seen from the figure, in general, the factors with a variance contribution rate of over 10% include the FM-MV, S-20, RD, Lon, and TBBstd, while the corresponding rates of the Lat, TBBm, and MDAV are relatively low. Among all the factors, the FMMV rank the highest in the variance contribution, reaching over 70%, which demonstrates the high correlation between the axisymmetry and intensity of the TC system. Meanwhile, the distributions of S-20 and TBBstd represent the region size and uniform distribution of the deep convection near the circulation center of the system, and both of the two factors exhibit high correlations with the system intensity, especially S-20. The RD factor can reflect the deviation of the large-area organized convection from the circulation center of the system. It makes up for the deficiency of the MMV’s deviation from the circulation center, which is caused by the influence of large-area organized convection under the severe asymmetry of the system and the increasing baroclinicity at the late development stage of TCs. Hence, it facilitates the improvement of the model accuracy (Yuan and Zhong, 2019). In addition, it should be noted that contribution rates of the average TBB and MMV are quite low, indicating that the TC intensity has a poor correlation to the regional-mean distribution of the convection. This may be attributed to the fact that both of the two factors are highly correlated with the key factor FMMV, thus the introduction of the two factors has little effect on the model. Considering that the interactions between factors occupy the absolute predominance in the variance contribution rate of the model, in order to further investigate the effects of the pre-selected factors and ensure the stability of the model, each predictor is removed from the model in turn, and the RMSE of the model without a certain predictor is compared with that of the model containing all the predictors. The increment is defined as the percentage increase in the RMSE before and after the predictor is removed. The larger the increment is, the more important the predictor is for the TC intensity determination; otherwise, the less important the predictor will be. As shown in Fig. 3b, the effect is the most significant when the factor FMMV is removed, with the increase in the RMSE all exceeding 25% under different radii. TBBstd and S-20 are the next. Particularly, the model is more sensitive to the two factors under smaller radii. The influence of RD is equivalent to those of TBBstd and S-20, but it increases with the increasing calculation radius and tends to be stable when the radius is > 400 km. When the two predictors, MDAV and TBBm, are removed respectively, deviations of the RMSE from that of the model with all factors are so small that they are almost negligible, which also correspond with their variance contribution rates. Besides, the position of the circulation center has a certain influence on the system intensity, and it is interesting that the effect of the longitude is stronger than that of the latitude. For the latitude, it has a greater impact on the modeling under smaller calculation radius, and the impact is quite weak when the calculation radius is greater than 400 km.

According to the contribution rate of each factor introduced above, we add them to the GLM in the sequence of FMMV, S-20, RD, Lon, TBBstd, Lat, TBBm, and MDAV, one by one, to compare the estimation results. With the data of all the three years as the validation set (G-all), it can be seen from Fig. 3c that the RMSE is gradually decreasing with the increase of the introduced factor number, especially for the first six factors. However, when the number of factors further increases, the accuracy of the GLM no longer improves significantly. Under the condition that the DAV calculation radius is 350 km, the RMSE is reduced by only 0.21 m s\(^{-1}\) compared with the model with the introduction of six factors. On the other hand, too many factors may weaken the robustness of the model due to collinearity. Therefore, it is necessary to minimize the number of factors to be introduced on the basis of ensuring the estimation accuracy. With full consideration of both the variance contribution and influence of the factor removal on the model, the six parameters FMMV, RD, S-20, TBBstd, Lat, and Lon are introduced to the GLM.

### 3.2.2 Intensity detection with GLM and LSTM models

In recent years, the models of machine learning have been gradually applied to the research and prediction of time series data. A deep learning model is a deep neural network model with multiple nonlinear mapping levels, which can extract features layer by layer from the input data, and dig deeper connections (LeCun and Bengio, 2015). In this section, we use the LSTM model, as well as GLM, to establish the multi-factor model of TC intensity estimation, and compare the results with the single-factor SFM.

The structure of the LSTM model is shown in Fig. 4. Each LSTM cell includes an input, output, and forget gates (Fig. 4a). The input gate determines the new information that can be added to the cell state. The output gate determines what will be regarded as the output value in the current state. The forget gate determines the information that will be discarded from the cell state. The
The schematic diagram of the structure of (a) LSTM cell and (b) 2-D LSTM chains, in which $f_t$, $i_t$, and $o_t$ denote the forget, input, and output gates, respectively; $σ$ and $tanh$ are activate functions $σ$ represents the Sigmoid function while $tanh$ represents the hyperbolic tangent function; and $x_t$ is the input value at the current time, while $C_t$ and $h_t$ are the state of the cell and hidden layer at the current moment, respectively.

LSTM model avoids the problem of long-term dependence by setting these three gates. The hidden layer of LSTM model consists of chain repeating cells. In this study, we use a two-dimensional LSTM (2-D LSTM) model to estimate the TC intensity. We set the number of nodes in the first layer network to 20 and the number of nodes in the second layer network to 5 (Fig. 4b).

The test results of GLM and LSTM intensity estimation models with optimum factors (Table 3) reveal that the relative effect of the test results is similar to that of the single-factor SFM among different groups, and the test result of G-2017 with relatively fewer high-intensity TC cases is significantly better than those of the other groups, while the group of G-2015 with the most high-intensity TC cases exhibits relatively poor performance. However, the RMSE range of the multi-factor models with GLM and LSTM models are 5.93–8.68 and 4.69–7.93 m s$^{-1}$, indicating that there is a significant improvement on the overall performance compared with the single-factor SFM. Meanwhile, an optimum calculation radius exists in all the four groups, which mainly ranges from 300 to 450 km. Between the two multi-factor models, LSTM model has lower RMSE in every test group even with different radii, and the optimum radius also narrows to 350–400 km, which means that the estimation results can be further improved with the introduction of the machine learning technology.

Through the comprehensive comparison, it can be seen that, physically, the multi-factor models not only consider the symmetrical characteristic of TCs but also can effectively reduce the error of intensity estimation by considering the deep convection, environment information, and their interactions. As the calculation efficiency of DAV is inversely proportional to the square of the radius, it is obvious that the determination of an optimum radius by the multi-factor models can effectively reduce the computational burden.

### 4. Comparison of the fitting and estimation results to the TC intensity

Although the optimum radii corresponding to differ-
ent validation sets are different, all of them basically fluctuate around the optimum radii corresponding to the fitting dataset. Thus, the optimum calculation radius in the fitting test is adopted for comparison between the results of the above two models. Specifically, the optimum calculation radii are 550 km for the single-factor SFM, 350 km for the multi-factor GLM, and 400 km for the multi-factor LSTM model. Biases of the results with the optimum calculation radii from the CMA-BST dataset at different stages of TCs are compared in Fig. 5. Generally, for the single-factor SFM, the median of intensity biases below the STS intensity is close to the zero line; at the stages of TY and STY, the obvious positive bias exists; and the significant negative bias is observed at the stage of STY except for the G-2017 test. However, large differences still exist among different test groups. For instance, the median of the intensity bias in the G-2016 test is slightly negative at the stages of STS and TY, while that in the G-2017 test is significantly positive at the stage of STY. Moreover, for intensities above TS, the distance between the upper and lower quartiles is larger and the tail is longer. Meanwhile, there are also many outliers and a large number of anomalies exceeding ± 20 m s\(^{-1}\) with a maximum value of over 50 m s\(^{-1}\). Generally, the estimation results from the single-factor SFM overestimate the TC intensity at the low-intensity stage and underestimate at the high-intensity stage, meanwhile the test results are very sensitive to the distribution of TC cases, revealing the poor stability of the model.

Comparatively, for the multi-factor models, distributions of the bias medians in different test groups are consistent, except for the positive bias at the STY stage in the G-2017 test, which is different from the others. Overall, it is shown that the medians in all test groups exhibit positive biases at the stages with the intensity below TD and get close to the zero line at the stages of TS, STS, and TY, yet negative biases occur at the stages above the STY intensity. In addition, it can be seen that the distance between the upper and lower quartiles as well as the tail length all decrease significantly, the number and

![Fig. 5. Boxplots for biases of the estimation results in the (a) single-factor SFM, (b) multi-factor GLM, and (c) LSTM model, from the CMA-BST dataset at different stages of the TC. “A”, “B”, “C”, and “D” in the four boxplots of different colors at each stage of the TC correspond to the G-2015, G-2016, G-2107, and G-all tests, respectively.](image)
magnitude of the outliers are within a reasonable interval, and also no bias of over 40 m s\(^{-1}\) appears (Figs. 5b, c). It indicates that the estimation results of the multi-factor models are more accordant with records in the CMA-BST dataset. By investigating differences between the two multi-factor models, it reveals that the number and magnitude of the outliers can be further reduced in the LSTM model.

Taking three SuperTYs [Dujuan (1521), Meranti (1614), and Talim (1718)] as examples, comparison of the TC intensities among the CMA-BST records, models’ estimations, and trends of four high-correlation factors after normalization, is shown in Fig. 6. It can be found more clearly that the multi-factor models reduce the estimation deviation at the peak-intensity stage and dissipation stage. Furthermore, both of the multi-factor models eliminate the large false increase at the early stage of these three cases. The reason for the greatly improved estimation accuracy of the multi-factor models is more due to the complementary effects among different types of factors. On the contrary, since the estimation results of the single-factor SFM only depend on the inverse correlation between the \(V_{\text{max}}\) and FMMV, the sharp drop of FMMV in the early stage of the TC led to a large deviation in the intensity estimation. As discussed in Yuan and Zhong (2019), FMMV can indicate the axisymmetry of deep cloud clusters around its location. The closer between the FMMV location and TC center at the same time, the better the estimation result is. However, when there exist the large-scale non-closed deep convective cloud clusters at the TC early stage, the FMMV decreases while the TC axisymmetry is poor. Fortunately, the other three factors (RD, S-20, and TBBstd), exert the positive effects on reducing this false increase by respectively showing the increasing distance between the FMMV location and TC center, limited area of deep convective cloud clusters, and large standard deviation around the circulation center. By comparing the GLM and LSTM models, it is found that the consideration of the nonlinear interaction among different factors with machine learning has avoided the high frequency oscillation of the estimation results and reduced the deviation.

Owing to different satellite data, study area, and testing methods used in recent studies of TC intensity estimation, it is not possible to directly compare the results among these studies in a comprehensive manner. But it is still meaningful to understand the effect of the fitting test from the advanced models in literature. As shown in Table 4, compared with the SFM (Ritchie et al., 2014) with seven years TC cases, GLM still slightly reduced the RMSE with no more than half samples. LSTM also achieved better results among complicated models, with the single-channel data and a much smaller sample number. The results also show that the introduction of multi-channel information could significantly reduce the fitting RMSE (Lee et al., 2020). We will further test the effects of multi-factors with longer durations of samples and more channel data in the future.

5. Conclusions

In this paper, intensity estimations for the TCs in the WNP during 2015–2017 are carried out by employing both a single-factor model SFM and two multi-factor models GLM and LSTM, based on the DAV technique. Meanwhile, sensitivity experiments with regard to the calculation radius and different training data groups are also conducted, and the estimation precision and optimum calculation radius for DAV in the WNP are ana-
lyzed. It is found that the SFM has relatively poor stability and lower accuracy in the TC intensity estimation; in addition, the calculation radius for the DAV technique is large, resulting in low computing efficiency, which cannot reach the ideal effect of the TC intensity determination. On the other hand, the two multi-factor models (GLM and LSTM) significantly decrease the RMSE of TC intensity estimation in all the test groups, mainly through introducing the information of deep convection structures. Moreover, the optimum radius for the multi-factor models ranges from 350 to 400 km, which effectially decreases the load of computation. The results demonstrate that the multi-factor models have better fitting degree with the CMA-BST data, as well as intensive deviation distribution and better stability.

Based on the recent studies on the TC intensity estimation, we gather that future work may need to pay attention to the following perspectives. (1) Since the number size of cases and multi-channel information intensity significantly affect the TC intensity estimation results, besides the effect of the multi-factors, expanding the input information and high-correlation factors to improve the accuracy of TC intensity estimation is a main task. (2) The application of machine learning is helpful to reflect the complicated nonlinear relation between the physical factors and TC intensity, and adjusting and improving LSTM model with multi-factors will be also important in the future work.

REFERENCES
Briegel, L. M., and W. M. Frank, 1997: Large-scale influences on tropical cyclogenesis in the western North Pacific. Mon. Wea. Rev., 125, 1397–1413, doi: 10.1175/1520-0493(1997)125<1397:LSIOTC>2.0.CO;2.

Combinido, J. S., J. R. Mendoza, and J. Aborot, 2018: A convolutional neural network approach for estimating tropical cyclone intensity using satellite-based infrared images. Pro. 2018 24th Int. Conference on Pattern Recognition, IEEE, Beijing, China, 1474–1480, doi: 10.1109/ICPR.2018.8545593.

Demaria, M., M. Mainelli, L. K. Shay, et al., 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). Wea. Forecasting, 20, 531–543, doi: 10.1175/WAF862.1.

Dvorak, V. F., 1975: Tropical cyclone intensity analysis and forecasting from satellite imagery. Mon. Wea. Rev., 103, 420–430, doi: 10.1175/1520-0493(1975)103<0420:TCIAAF>2.0.CO;2.

Graves, A., 2012: Long short-term memory. Supervised Sequence Labelling with Recurrent Neural Networks, A. Graves, Ed., Springer, Berlin, Heidelberg, 37–45, doi: 10.1007/978-3-642-24797-2_4.

Guo, Z., M. H. Wang, Y. Qian, et al., 2014: A sensitivity analysis of cloud properties to CLUBB parameters in the single-column Community Atmosphere Model (SCAM5). J. Adv. Model. Earth Syst., 6, 829–858, doi: 10.1002/2014MS000315.

Hill, K. A., and G. M. Lackmann, 2009: Influence of environmental humidity on tropical cyclone size. Mon. Wea. Rev., 137, 3294–3315, doi: 10.1175/2009MWR2679.1.

Hubert, L. F., and A. Timchalk, 1969: Estimating hurricane wind speeds from satellite pictures. Mon. Wea. Rev., 97, 382–383, doi: 10.1175/1520-0493(1969)97<0382:EHWSFS>2.3.CO;2.

Knaff, J. A., D. P. Brown, J. Courtney, et al., 2010: An evaluation of Dvorak technique-based tropical cyclone intensity estimates. Wea. Forecasting, 25, 1362–1379, doi: 10.1175/2010WAF2222375.1.

LeCun, Y., Y. Bengio, and G. Hinton, 2015: Deep learning. Nature, 521, 436–444, doi: 10.1038/nature14539.

Lee, J., J. G. Im, D. H. Cha, et al., 2020: Tropical cyclone intensi-
ity estimation using multi-dimensional convolutional neural networks from geostationary satellite data. Remote Sens., 12, 108, doi: 10.3390/rs12010108.

McCullagh, P., 1984: Generalized linear models. Eur. J. Oper. Res., 16, 285–292, doi: 10.1016/0377-2217(84)90282-0.

Olander, T. L., and C. S. Velden, 2007: The advanced Dvorak technique: Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery. Wea. Forecasting, 22, 287–298, doi: 10.1175/WAF975.1.

Piñeros, M. F., E. A. Ritchie, and J. S. Tyo, 2008: Objective measures of tropical cyclone structure and intensity change from remotely sensed infrared image data. IEEE Trans. Geosci. Remote Sens., 46, 3574–3580, doi: 10.1109/TGRS.2008.2000819.

Piñeros, M. F., E. A. Ritchie, and J. S. Tyo, 2010: Detecting tropical cyclone genesis from remotely sensed infrared image data. IEEE Geosci. Remote Sens. Lett., 7, 826–830, doi: 10.1109/LGRS.2010.2048694.

Piñeros, M. F., E. A. Ritchie, and J. S. Tyo, 2011: Estimating tropical cyclone intensity from infrared image data. Wea. Forecasting, 26, 690–698, doi: 10.1175/WAF-D-10-05062.1.

Ritchie, E. A., and G. J. Holland, 1999: Large-scale patterns associated with tropical cyclogenesis in the western Pacific. Mon. Wea. Rev., 127, 2027–2043, doi: 10.1175/1520-0493(1999)127<2027:LSPAWT>2.0.CO;2.

Ritchie, E. A., K. M. Wood, O. G. Rodríguez-Herrera, et al., 2014: Satellite-derived tropical cyclone intensity in the North Pacific Ocean using the deviation-angle variance technique. Wea. Forecasting, 29, 505–516, doi: 10.1175/WAF-D-13-00133.1.

Sun, Z. C., Y. B. Tao, S. Li, et al., 2013: Statistical strategies for constructing health risk models with multiple pollutants and their interactions: Possible choices and comparisons. Environ. Health, 12, 85, doi: 10.1186/1476-069X-12-85.

Wood, K. M., O. G. Rodriguez-Herrera, E. A. Ritchie, et al., 2015: Tropical cyclogenesis detection in the North Pacific using the deviation angle variance technique. Wea. Forecasting, 30, 1663–1672, doi: 10.1175/WAF-D-14-00113.1.

Yuan, M., and W. Zhong, 2019: Detecting intensity evolution of the western North Pacific super typhoons in 2016 using the deviation angle variance technique with FY data. J. Meteor. Res., 33, 104–114, doi: 10.1007/s13351-019-8064-7.

Zhang, C. J., J. F. Qian, L. M. Ma, et al., 2016: Tropical cyclone intensity estimation using RVM and DADI based on infrared brightness temperature. Wea. Forecasting, 31, 1643–1654, doi: 10.1175/WAF-D-15-01001.1.

Zhang, C. J., Q. Luo, L. J. Dai, et al., 2019: Intensity estimation of tropical cyclones using the relevance vector machine from infrared satellite image data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 12, 763–773, doi: 10.1109/JSTARS.2019.2894654.

Zhang, Q. H., and C. Y. Guo, 2008: Overview of the studies on tropical cyclone genesis. Acta Oceanol. Sinica, 30, 1–11. (in Chinese)

Tech & Copy Editor: Meili ZHANG