Comparison of Three Tropical Cyclone Intensity Datasets*

YU Hui1,2† (余晖), HU Chunmei3 (胡春梅), and JIANG Leyi1,2 (蒋乐贻)

1 Shanghai Typhoon Institute, CMA, Shanghai 200030
2 Laboratory of Typhoon Forecast Technique, CMA, Shanghai 200030
3 Chongqing Meteorological Observatory, Chongqing 401147

(Received January 8, 2007)

ABSTRACT

Analyzed in this paper are the 16-yr (1988-2003) tropical cyclone (TC) intensity data from three major forecast centers of the western North Pacific, i.e., China Meteorological Administration (CMA), Regional Specialized Meteorological Center Tokyo (RSMC Tokyo), and Joint Typhoon Warning Center (JTWC) of the United States. Results show that there are significant discrepancies (at 1% significance level) in the intensity of TCs among the three centers, with a maximum difference for the same TC over 30 m s$^{-1}$. The flight reconnaissance over TC can diminish the discrepancy to some extent.

A climatic and persistent prediction model is set up to study the impact of initial data from different forecast centers on the prediction of TC intensity. It is obtained that the root mean square error (RMSE) of a 4-yr independent test is the largest using data from JTWC, while the smallest using data from RSMC Tokyo. Average absolute deviation in 24-h intensity prediction is 2.5 m s$^{-1}$ between CMA and RSMC Tokyo data, and 4.0 m s$^{-1}$ between CMA and JTWC data, with a maximum deviation reaching 21 m s$^{-1}$. Such a problem in the initial value increases the difficulty in intensity prediction of TCs over the western North Pacific.

Key words: tropical cyclone (TC), intensity, data

1. Introduction

During the past decades, the accuracy of tropical cyclone (TC) track forecast has been improved constantly all over the world (Chen et al., 2004; Bryant, 2004), owing greatly to the development in remote sensing and reconnaissance technique (Velden and Hawkins, 2002; Marks, 2003), numerical model (Kurihara et al., 1998; Ma et al., 2004), and the understanding of TC motion mechanism (Chan, 2002). Taking China as an example, the 24-h average distance error of its official forecast has decreased from 240 km in 1985 to 138 km in 2004. However, little improvement has been made in TC intensity prediction (Elsberry, 2002; Yao et al., 2004; Wang and Wu, 2003). Due to the lack of observations over the ocean, the difficulty of TC intensity prediction comes from not only the lack of effective forecast techniques, but also the unknown true value of TC intensity.

With the development of atmospheric detection technology, the methods of monitoring TC have been increased constantly (Chu et al., 2002). In early times, the information of TC mainly came from navigation logs and observations over land. From the 1930s to 1960s, a network of radiosonde, weather radar, and satellite was gradually set up with some aircraft reconnaissance. The buoy network was set up in the 1970s when the satellite cloud drift wind data were also put into use. In 1975, Dvorak developed an intensity estimation technique of TC based on infrared/visible satellite images, which is still widely used at present. After the 1990s, the polar orbiting satellite data began to play more and more important role in the analyses of TC intensity, such as the AMSU temperature and precipitation data (Demuth et al., 2004), the QuikSCAT ocean surface wind data (Edson, 2002), the TRMM microwave data (Edson, 2000), and so on. However, most part of the network mentioned above focuses on land. The buoys, ship’s reports, and polar orbiting satellite data all have a problem of insufficient spacial and temporal resolution. Furthermore, the aircraft reconnaissance for TCs in the western North

---

*Supported jointly by funds from the Shanghai Typhoon Institute, the National Natural Science Foundation of China (40333025), the Ministry of Science and Technology of China (2005DB33J104), and the Forecasting system Laboratory of NMC CMA.
†Corresponding author: huiyu2001cn@yahoo.com.cn.
Pacific was ended in 1987. The Dvorak technique is the only tool that the forecast centers can rely on for most of the time while trying to define the TC intensity. Due to the subjective essence of the Dvorak technique and the fact that different centers can have access to different sources of accessory data, there generally exist some divergences among the best track and intensity datasets from different centers.

Yu and Kwon (2005) made a comparison for the intensity of Typhoons Prapiroon (2000) and Olga (1999) among the datasets from three major forecast centers of the western North Pacific, i.e., China Meteorological Administration (CMA), Regional Specialized Meteorological Center Tokyo (RSMC Tokyo), and Joint Typhoon Warning Center (JTWC) of the United States. It is found that, although the general trend of intensity change of the two TCs reported by different centers is similar to each other, there always exist some discrepancies during any specific period, or even an opposite trend. In fact, such a problem in TC intensity data has long been noticed and the idea of developing a unified TC dataset for the western North Pacific has been proposed as early as 2001. However, there have been no systematic analyses on the differences among different datasets and how the differences can affect the prediction of TC intensity.

In the following part of the paper, a general description of the datasets and analysis method will be given in Section 2. A comparison of the statistics of TC frequency, intensity, and intensity change from different datasets will be given in Sections 3, 4, and 5, respectively. In Section 6, a climatic and persistent model is established to study the possible impact of initial data from different centers on intensity prediction. The 30-yr (1970-1999) data from CMA are used to set up the model and 4-yr (2000-2003) data from the three centers are tested.

3. Statistics of TC frequency

The comparison of annual TC frequency from the three datasets (Fig.1) shows that the mean annual TC frequency from CMA is 26.2, between 26.7 from RSMC Tokyo and 25.1 from JTWC. Figure 1 further demonstrates that CMA has a dataset with the largest root mean square error (RMSE) 5.3 in annual TC frequency. They are 4.9 and 4.6 from RSMC Tokyo and JTWC, respectively. The maximum and minimum annual TC frequencies are 37 and 14, respectively, both from CMA. However, above differences in the mean value and RMSE of annual TC frequency are not significant at the 99% confidence level and it can be concluded that there is no statistically significant difference in TC annual frequency among the three
datasets.

The frequency of TC at different intensity groups is also compared by classifying the samples into tropical storm (TS, \( V_{\text{max}} \) between 17.2 and 24.4 m \( \text{s}^{-1} \)), severe tropical storm (STS, \( V_{\text{max}} \) between 24.5 and 32.6 m \( \text{s}^{-1} \)), and typhoon (TY, \( V_{\text{max}} \) larger than 32.6 m \( \text{s}^{-1} \)) according to the criteria used by CMA from 1989 to 2005. It is shown in Table 1 that annual TYs from CMA (15.8) are more than RSMC Tokyo (14.3) and JTWC (14.8). Furthermore, the RMSE of TY frequency from CMA is the largest (4.7) as compared to 4.2 and 3.9 from RSMC Tokyo and JTWC. There is no obvious difference in the annual number of STS from CMA and RSMC Tokyo with the same mean value of 5.7 and small differences in RMSE, maximum and minimum frequencies. However, the JTWC tends to report less STSs with an annual mean frequency of 4.6. The RMSE of STS frequency is also the smallest from JTWC (1.9). Relatively large discrepancy exists in TS frequency. The annual mean from CMA and RSMC Tokyo is 4.7 and 6.7, respectively, while that is 5.7 from JTWC. However, all the differences mentioned above are not significant at the 99% confidence level either.

### Table 1. Statistics of annual TY, STS, and TS frequencies according to the datasets from CMA, RSMC Tokyo, and JTWC

<table>
<thead>
<tr>
<th></th>
<th>TY</th>
<th>STS</th>
<th>TS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CMA</td>
<td>RSMC Tokyo</td>
<td>JTWC</td>
</tr>
<tr>
<td>Mean</td>
<td>15.8</td>
<td>14.3</td>
<td>14.8</td>
</tr>
<tr>
<td>RMSE</td>
<td>4.7</td>
<td>4.2</td>
<td>3.9</td>
</tr>
<tr>
<td>Max</td>
<td>21</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Min</td>
<td>6</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

In order to find out possible impact of aircraft TC reconnaissance on intensity estimation, statistical features are also analyzed for the period with flight reconnaissance (1977-1986) and compared to that without flight reconnaissance (1988-2003). It is obtained that the discrepancy between CMA, RSMC Tokyo, and JTWC in the mean annual TC frequency becomes smaller in the latter period. The average absolute deviation between CMA and RSMC Tokyo (CMA and JTWC) reduces from 1.0 to 0.9 (from 2.2 to 1.9). No remarkable discrepancy exists in the RMSE of annual TC frequency difference among the three centers for the two periods. Taking into consideration the statistical significance test mentioned above, it can be concluded that the difference in TC frequency among the three datasets tends to become smaller after the end of aircraft reconnaissance, which might be mainly due to the extensive usage of satellite data and the better international exchange of information.

### 4. Statistics of TC intensity

The statistics for TC intensity from the three datasets (Table 2) demonstrate that there generally exist differences no matter with or without aircraft reconnaissance.

For the period of 1988-2003, the mean value and RMSE of TC intensity from CMA are 30.8 and 10.3 m \( \text{s}^{-1} \), both between those from RSMC Tokyo (30.2 and 9.5 m \( \text{s}^{-1} \)) and JTWC (32.5 and 12.4 m \( \text{s}^{-1} \)). The difference in mean values between any two datasets is significant at the 99% confidence level.

As for the changes brought about by the end of aircraft reconnaissance, the discrepancy between CMA and RSMC Tokyo reduces from 0.9 to 0.6 m \( \text{s}^{-1} \) in

![Fig.1. Statistics for annual TC frequency from 1988 to 2003 according to the datasets from CMA, RSMC Tokyo, and JTWC.](image-url)
Table 2. Statistics of TC intensity according to the datasets from CMA, RSMC Tokyo, and JTWC

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (m s(^{-1}))</td>
<td>30.8</td>
<td>30.2</td>
<td>32.5</td>
<td>31.3</td>
<td>30.4</td>
<td>30.2</td>
</tr>
<tr>
<td>RMSE (m s(^{-1}))</td>
<td>10.3</td>
<td>9.5</td>
<td>12.4</td>
<td>11.3</td>
<td>10</td>
<td>10.9</td>
</tr>
<tr>
<td>Sample size</td>
<td>8999</td>
<td>9038</td>
<td>8284</td>
<td>5707</td>
<td>5591</td>
<td>4837</td>
</tr>
</tbody>
</table>

the mean intensity, and from 1.3 to 0.8 m s\(^{-1}\) in the RMSE. However, that between CMA and JTWC increases from 1.1 to 1.7 m s\(^{-1}\) in the mean value, and from 0.4 to 2.1 m s\(^{-1}\) in the RMSE. The discrepancy between RSMC Tokyo and JTWC enlarges more obviously with mean intensity difference increasing from 0.2 to 2.3 m s\(^{-1}\) and RMSE difference from 0.9 to 2.9 m s\(^{-1}\). The statistical test at 99% confidence level indicates that there is no significant difference in mean intensity between RSMC Tokyo and JTWC during the period of 1977-1986. Then it can be deduced that the absence of flight reconnaissance enlarges significantly the discrepancy between CMA or RSMC Tokyo and JTWC in TC intensity.

According to the distribution of TCs at different intensity bands (Fig.2), the ratio from CMA is close to that from JTWC at all the bands with \(V_{\text{max}} \leq 40\) m s\(^{-1}\), while remarkable discrepancy exists between RSMC Tokyo and any of them. The ratio of TS and TY with \(V_{\text{max}} \leq 40\) m s\(^{-1}\) from RSMC Tokyo is over 5% higher than that from CMA and JTWC. At STS band, RSMC Tokyo is on the remarkably lower side. RSMC Tokyo and JTWC are in good agreement at the bands of 40-50 m s\(^{-1}\) (about 16%), with CMA on the higher side (about 18%). JTWC reports most TCs stronger than 50 m s\(^{-1}\), while RSMC Tokyo the least. Another notable characteristic of RSMC Tokyo dataset is that the ratio tends to descend with the increase of \(V_{\text{max}}\).

Statistics for the difference in intensity between CMA and the other two centers by comparing the samples one by one (Table 3) demonstrate that CMA tends to be more consistent with RSMC Tokyo during the period of 1988-2003, with the mean absolute difference decreasing from 3.0 to 2.5 m s\(^{-1}\) in the period from 1977 to 1986, while that between CMA and JTWC increases from 3.8 (1977-1986) to 4.1 m s\(^{-1}\) (1988-2003). Maximum discrepancy between CMA and the other two centers is more than 30 m s\(^{-1}\). A typical example is TC 9521 (Fig.3), for which the intensity from CMA is obviously weaker than the two other centers in its decaying phase. The maximum difference between CMA and JTWC reaches 31 m s\(^{-1}\).

5. Statistics of TC intensity

Statistics of TC intensity change in 24 h are compared among different datasets (Table 4). It can be seen that JTWC reports the most radical change in

![Fig.2](image-url)  
**Fig.2.** Ratio of TCs at different intensity bands according to the datasets from CMA, RSMC Tokyo, and JTWC.
Table 3. Statistics of difference in intensity between CMA and RSMC Tokyo (CMA minus RSMC Tokyo), CMA and JTWC (CMA minus JTWC) by comparing the samples one by one

<table>
<thead>
<tr>
<th>Year</th>
<th>CMA–RSMC Tokyo</th>
<th>CMA–JTWC</th>
<th>CMA–RSMC Tokyo</th>
<th>CMA–JTWC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988-2003</td>
<td>2.5</td>
<td>4.1</td>
<td>3.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Mean absolute diff. (m s⁻¹)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max. (m s⁻¹)</td>
<td>−16</td>
<td>−31</td>
<td>−14</td>
<td>−18</td>
</tr>
<tr>
<td>Min. (m s⁻¹)</td>
<td>26</td>
<td>20</td>
<td>32</td>
<td>29</td>
</tr>
<tr>
<td>Sample size</td>
<td>8448</td>
<td>7504</td>
<td>5198</td>
<td>4638</td>
</tr>
</tbody>
</table>

TC intensity, with the average being 2 m s⁻¹ and the RMSE 9.3 m s⁻¹, both of which are the largest among the three centers. The minimum and maximum values from JTWC are −40 m s⁻¹ (weakening) and 38 m s⁻¹ (intensifying), respectively. Both are the extrema as compared to those from the other two centers. RSMC Tokyo appears to be the most conservative in the intensity change of TC and reports a mean value of 0.8 m s⁻¹, an RMSE of 6.8 m s⁻¹, and a maximum of only 21 m s⁻¹. Statistics from CMA are medium. The statistical test at 99% confidence level further demonstrates that the difference in 24-h intensity change is significant among the three datasets.

As for the distribution of 24-h intensity change (Fig.4), the CMA dataset has a unimodal distribution with the peak value (27%) in 0-5-m s⁻¹ band. There are about 19% cases with no intensity change in 24 h and a similar ratio for 5-10-m s⁻¹ band. A bimodal distribution is obtained from JTWC dataset, with the first peak value (21.7%) in 5-10-m s⁻¹ band and the second (14.1%) in −5-0-m s⁻¹ band. However, there are three peak values according to the RSMC Tokyo dataset, i.e., 5-10 m s⁻¹ (about 22%), 0 m s⁻¹ (16.7%), and −10−5 m s⁻¹ (15.7%) bands, respectively.

6. Impact of initial data uncertainty on the prediction of TC intensity as inferred from a climatic and persistent prediction model

According to the above analyses, there is remarkable discrepancy in TC intensity among different datasets. To have an idea on the possible impacts of such a kind of data uncertainty on the prediction of TC intensity, a climatic and persistent prediction model is set up and 4-yr (2000-2003) independent experiments are carried out using initial values from different forecast centers.

6.1 The climatic and persistent prediction method for TC intensity

The stepwise regression is performed on 9 climatic and persistent predictors (TC latitude, longitude, and moving velocity at initial time, 12 and 24 h before the
initial time, respectively) to set up the prediction equations for 12-, 24-, 36-, 48-, 60-, and 72-h intensity change based on the CMA dataset from 1970 to 1999. Only the samples from May to October are taken into consideration and different equations are established for different months. As shown in Table 5, the multiple correlation coefficients \( R \) vary from 0.938 for 12 h to 0.615 for 72 h and the standard deviation of residuals from 4.36 to 10.74 m s\(^{-1}\), implying a deteriorating performance as the leading time increases.

### Table 5. Sample size and regression statistics of the climatic and persistent prediction model

<table>
<thead>
<tr>
<th></th>
<th>12 h</th>
<th>24 h</th>
<th>36 h</th>
<th>48 h</th>
<th>60 h</th>
<th>72 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>17971</td>
<td>15946</td>
<td>14917</td>
<td>13418</td>
<td>11885</td>
<td>10661</td>
</tr>
<tr>
<td>( R )</td>
<td>0.938</td>
<td>0.844</td>
<td>0.759</td>
<td>0.69</td>
<td>0.646</td>
<td>0.615</td>
</tr>
<tr>
<td>Standard deviation of residuals (m s(^{-1}))</td>
<td>4.36</td>
<td>6.90</td>
<td>8.51</td>
<td>9.64</td>
<td>10.28</td>
<td>10.74</td>
</tr>
</tbody>
</table>

### 6.2 Uncertainty in prediction as induced by the initial value problem

Independent experiments are carried out using 4-yr (2000-2003) data from the three forecast centers respectively. It is found that the discrepancy in initial inputs can result in significantly different error statistics (Table 6). Taking the 24-h prediction as an example, the root mean square error (RMSE) is the largest using JTWC dataset (8.07 m s\(^{-1}\)), while the smallest using RSMC Tokyo dataset (5.48 m s\(^{-1}\)). The performance of CMA dataset is in between, 5.95 m s\(^{-1}\).

Analyses on the regression equations exhibit that the intensity of TC and its change in 12- or 24-h are major predictors for all the leading times. Generally, a linear multiple regression equation can be given as

\[
y = a + bx_1 + cx_2 + dx_3 + \ldots,
\]

where \( y \) is the predictand, \( x_1, x_2, x_3, \ldots \) are predictors, and \( a, b, c, d, \ldots \) are regression coefficients.

If only a single regression is taken into consideration, Eq.(1) can be written as

\[
y = a + bx.
\]

Substitute Eq.(2) into the RMSE formula,

\[
RMSE = \left[ \frac{1}{N} \sum_{k=1}^{N} (y - \hat{y})^2 \right]^{\frac{1}{2}}.
\]

Here, \( y \) and \( \hat{y} \) are the predicted and observed intensity changes of TC, respectively. Others are commonly used symbols. If assuming the statistical features of the predictand (~\( \hat{y} \)) remain invariable, \( RMSE \) will increase with the rise of mean value and square deviation of the predictor (\( x \)). Therefore, it is reasonable for us to have the largest \( RMSE \) if using the JTWC dataset, and the smallest \( RMSE \) if using the RSMC Tokyo dataset. The trait of linear fitting equation determines that there exists a simple linear relationship between the discrepancy in prediction and that in initial values.

Also taking the 24-h forecast as an example, the differences in the intensity prediction of any specified TC are further analyzed for the three datasets. It is found that the mean absolute deviation between the CMA and RSMC Tokyo (CMA and JTWC) datasets is 2.5 (4.0) m s\(^{-1}\), with a maximum reaching 16 (21) m s\(^{-1}\). Such a problem of the initial value increases the difficulty in the intensity prediction of TCs in the western North Pacific.

### Table 6. RMSE of the climatic and persistent prediction model with 4-yr (2000-2003) initial inputs from different forecast centers

<table>
<thead>
<tr>
<th></th>
<th>24 h</th>
<th>48 h</th>
<th>72 h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CMA</td>
<td>RSMC</td>
<td>JTWC</td>
</tr>
<tr>
<td>Sample size</td>
<td>959</td>
<td>959</td>
<td>959</td>
</tr>
<tr>
<td>RMSE (m s(^{-1}))</td>
<td>5.95</td>
<td>5.48</td>
<td>8.07</td>
</tr>
</tbody>
</table>
different intensity datasets of TCs from three major forecast centers, i.e., CMA, RSMC Tokyo, and JTWC, for the western North Pacific. The potential effect of the uncertainty in TC intensity on its prediction is also studied based on a simple statistical model. The main results are as follows.

(1) There is significant discrepancy in the intensity datasets of TC from different forecast centers, with the average values of TC intensity significantly different at 99% confidence level, and the maximum difference for a specified TC more than 30 m s$^{-1}$.

(2) The aircraft reconnaissance over TC can reduce the discrepancy to some extent in the determination of TC intensity among different centers.

(3) The intensity change of TC reported by JTWC is the most radical, with both the average and RMSE of 24-h intensity change maximum among the three centers. The next is CMA and then RSMC Tokyo.

(4) There is no remarkable discrepancy in annual frequency of TC from the three centers.

(5) According to the results from a simple statistical model, it is found that the discrepancy in initial inputs can result in significantly different error statistics, with RMSE the largest using JTWC dataset, while the smallest using RSMC Tokyo dataset. The maximum deviation between the predictions for a given TC based on CMA and JTWC datasets can reach 21 m s$^{-1}$.

Above findings indicate that the problem of basic data aggrandizes the difficulty of TC intensity prediction in the western North Pacific, which should be considered seriously in operation. Furthermore, the prediction uncertainty is only discussed based on a quite simple statistical model in this paper, which might be much more significant if any nonlinear model is used, such as a numerical prediction model. This is worthy of further study.

It should be noted that the initial value problem discussed in this study is different from that in the ensemble prediction (Zhou et al., 2003) which generally creates members by randomly perturbing the initial fields. The significant differences among TC intensity datasets is a special problem for the TCs in western North Pacific, which might be solved successfully one day by carrying out flight reconnaissance, improving the ability in analyzing remote sensing data, increasing in situ observations over ocean or strengthening the cooperation and intercommunication around the world.

Acknowledgements. The authors would like to thank Professor Duan Yihong, Professor Johnny C. L. Chan, and Dr. Xu Ming from Shanghai Typhoon Institute for their comments on this study.

REFERENCES

Bryant, A. C., 2004: Annual Tropical Cyclone Report. Joint Typhoon Warning Center/USA.


Kurihara, Y., M. A. Bender, R. E. Tuleya, and R. J. Ross, 1998: The GFDL hurricane prediction system and


