Errors in Tropical Cyclone Intensity Forecast by RSMC Tokyo and Statistical Correction Using Environmental Parameters

Kosuke Ito
University of the Ryukyus, Okinawa, Japan

Abstract

Tropical cyclone (TC) intensity forecasts issued by the Regional Specialized Meteorological Centre (RSMC) Tokyo - Typhoon Center are systematically compiled to analyze the long-term behavior of errors and to explore the potential for improvement in the forecast accuracy using a statistical correction approach. In this study, a dataset is constructed from annual statistics and every single forecast listed on annual reports on the activities of the RSMC Tokyo. This study found that (1) the accuracy of annual mean forecast has not improved over 26 years and that (2) forecast errors tend to be larger in the rapidly developing TCs. Further analysis reveals that recent forecast output (2008–2014) contains biases associated with the magnitude of the vertical shear of horizontal wind, convective available potential energy, upper ocean temperature, maximum potential intensity (MPI) and ocean coupling potential intensity (OC PI). To evaluate the adverse effect of such biases in the current forecast system, a simple statistical correction is applied. It improved TC intensity forecast by 7.8–16.9% when an OC PI is employed.

(Citation: Ito, K., 2016: Errors in tropical cyclone intensity forecast by RSMC Tokyo and statistical correction using environmental parameters. SOLA, 12, 247–252, doi:10.2151/sola.2016-049.)

1. Introduction

The forecasts of tropical cyclone (TC) intensity have been of particular importance in the atmospheric sciences because of the potential severity of resulting disasters. Since 1 July 1989, the Japan Meteorological Agency (JMA) has been serving as a Regional Specialized Meteorological Centre (RSMC) Tokyo - Typhoon Center issuing TC advisories in the Western North Pacific (WNP) within the framework of the World Weather Watch program of the World Meteorological Organization. As the WNP contains the largest number of very intense TCs of any region (D’Asaro et al. 2011), official forecasts published by the RSMC Tokyo benefit many circum-Pacific countries. It is generally believed that forecasts of TC intensity, represented by minimum sea level pressure (MSLP) and 10-min averaged maximum wind speed at 10 m height (Vmax), has been challenging comparing to track forecasts (Wang and Wu 2004). For example, DeMaria et al. (2014) demonstrated that the best intensity guidance model exhibits the statistically significant improvements in the WNP over three decades, but not official intensity forecasts conducted by the Joint Typhoon Warning Center. Yu et al. (2013) analyzed the output of the JMA global spectral model (JMA-GSM), which is currently used for RSMC Tokyo’s official TC intensity forecast, and has proven to be useful for intensity forecasts. However, errors in official intensity forecasts issued by the RSMC Tokyo have never been investigated in detail to the author’s knowledge; these forecasts comprise a synthesis of the capability of numerical models, statistical corrections, and the forecaster’s skills. Considering the social and economic importance of the RSMC Tokyo’s official forecasts, it is crucial to investigate the nature of their intensity errors and, if possible, to assess improvement possibilities by relating the errors of TC intensity forecasts to the environmental conditions, as Bhatia and Nolan (2013) did for hurricanes in the Atlantic Ocean.

The first aim of this work is to construct and analyze a dataset of intensity (MSLP and Vmax) forecast errors defined as the deviation in RSMC Tokyo official forecast relative to the RSMC Tokyo best track. The second aim is to quantify biases (defined as a composite mean error in a category) in regard to environmental parameters. Finally, we evaluate the potential for improvement using environmental parameters to diminish biases through a statistical correction. They contribute to reviewing the behavior of TC intensity forecast errors issued by the RSMC Tokyo and to the design of future TC intensity forecast systems.

2. Datasets

Figure 1 briefly summarizes the history of TC intensity forecasts, available error data, and numerical models in JMA. To date, the JMA has been primarily employing a JMA-GSM with a grid spacing of 20 km as a base model since November 2007 and applying corrections to numerical model results based on a Dvorak C1 number and MSLP relationship, sea surface temperature (SST) and MSLP relationship, and statistical forecasts in order to publish the official forecasts (Koide et al. 2014). Official six-hourly individual intensity forecast errors have been recorded in the appendices of RSMC Tokyo annual reports (hereafter referred to as ARs) since 1992 in addition to annual errors available since 1989. For Vmax errors, both knots and meters per second are used as the unit in ARs depending on the year; we converted knots to meters per second using 1 knot = 0.5144 m s⁻¹. Forecast accuracy is not verified before a TC reaches the threshold of tropical storm status or an extratropical cyclone subjected to transition from the TC.

To relate the forecast errors and environmental physical parameters, we focus on the period from 2008 to 2014, following the implementation of the JMA-GSM in November 2007 because their relationship is dependent on a base numerical model in addition to guidance, and quality of the best track. Environmental physical parameters considered here are the magnitude of the vertical shear of horizontal wind, ambient convective available potential energy (CAPE), surface temperature, upper ocean temperature (UOT), and maximum potential intensity (MPI), which are relevant to the TC intensity (e.g. Emanuel et al. 2004; Lin et al. 2013; Wang and Wu 2004). MPI was originally proposed by Emanuel (1986) as a theoretical upper limit of MSLP and Vmax given ideal environmental conditions, including fixed SST, for an axisymmetric vortex. We employ the version of this concept revised by Bister and Emanuel (1998), hereafter referred to as E-MPI. We also tested an ocean coupling potential intensity (OC PI) proposed by Lin et al. (2013) as an index in which the SST of E-MPI is replaced with the depth-averaged UOT, to account for the atmosphere-ocean coupling effect.

The magnitude of the vertical shear is defined as the difference of wind vectors between 850 hPa and 200 hPa, averaged within 300 km from the TC center, as described by Ueno and Kunii (2009). UOT is represented by the depth-averaged ocean temperature from surface to 100 m, \( T_{100_0} \) (Price 2009). E-MPI for Vmax and MSLP are expressed as follows (Bister and Emanuel 1998, 2002):
\[ V_{\text{max}}^2 = \frac{T_s - T_0}{C_D} C_p T_m \ln \left( \frac{\theta_e - \ln \theta_D}{\ln \theta_s} \right) \]  
\[ 2C_p T_m \ln \frac{P_s}{P_v} = V_{\text{max}}^2 \]  
\[ 2C_p T_m \ln \frac{P_s}{P_v} = V_{\text{max}}^2 + \text{CAPE}_s \]

where \( V_{\text{max}} \) is the maximum wind speed, \( P \) is the sea-level pressure, \( C_D \) is the drag coefficient, \( C_p \) is the surface enthalpy exchange coefficient (assuming \( C_D = C_p \)), \( T_m \) is the specific heat at constant pressure, \( T \) is temperature, \( \theta_e \) is the saturated equivalent potential temperature, and \( \theta_D \) is the equivalent potential temperature. \( \text{CAPE}_s \) is the \text{CAPE} of boundary layer air under the eyewall. Subscripts \( S, O, B, c, v \) and \( \text{env} \) represent the values at sea surface, outflow boundary, TC center, the radius of \( V_{\text{max}} \), and environment, respectively. The actual algorithm is based on the work of Bister and Emanuel (2002). The outflow layer is defined at the level of neutral buoyancy. To calculate \( \text{OC}_\text{PI} \), \( T_m \) in E-MPI is replaced with \( T_{100} \). Note that \( \text{OC}_\text{PI} \) cannot be calculated where \( \text{UOT} \) to a depth of 100 m is not available. For fair comparison between E-MPI and \( \text{OC}_\text{PI} \), E-MPI is referred to as E-MPIO when \( \text{OC}_\text{PI} \) is available. E-MPIL refers to E-MPI applied to coastal regions or land, representing about 8% and 10% of total forecast cases, respectively.

Wind fields, atmospheric temperature, humidity and ground temperature data (including SST) are obtained from the operational version of JMA-GSM and Merged Satellite and In-situ Data Global Daily SST (See Sections 3.2 and 5.2 of JMA 2013). Six-hourly data is stored with a grid spacing of 0.5° by 0.5° in the Kyoto University Research Institute for Sustainable Humanosphere database (available at http://database.rish.kyoto-u.ac.jp/arch/jmadata/data/gpv/original/). The daily UOT is retrieved from the Four-dimensional Variational Ocean ReAnalysis for the WNP over 30 years (Usui et al. 2016, manuscript submitted to J. Oceanogr.). Based on the JMA Meteorological Research Institute Multivariate Ocean Variational Estimation system (Usui et al. 2006). Because a 4D-Var-based fine-mesh dataset (15°N–60°N and 118°E–180°E) with a horizontal grid spacing of 0.1° does not cover the whole region which the RSMC Tokyo is responsible for, we used the 3D-Var-based parent domain output with the grid spacing of 0.5°.

During analysis, the vertical shear is centered at the location of MSLP, based on the JMA-GSM rather than the best track, and is temporally averaged over the integration time. Environmental \( \text{CAPE} \), ground temperature (including SST), \( \text{UOT} \), and the relative humidity are defined as the values at the TC center position in the best track, 2 days prior to the verification time (Lin et al. 2013). For simplicity, the temperature and humidity at the lowest atmospheric level of JMA-GSM are used for the values of boundary layer.

### 3. TC intensity error

Figure 2 shows annual mean forecast errors in TC position (including track forecasts before JMA started to serve as RSMC Tokyo), MSLP and \( V_{\text{max}} \). While track forecast errors today have decreased to less than half of those obtained 30 years ago, the long-term trend of intensity forecast error exhibits an increasing behavior. It might be because of the recent increase of rapid intensification\(^1\) events rather than the degeneration of the forecast system as discussed in Section 5. Mean biases of MSLP and \( V_{\text{max}} \) are closer to zero on average for forecast time (FT) of 24 and 48 h, whereas forecasts of TC intensity have typically exceeded actual intensity at FT = 72 h (~4.3 hPa and ~3.4 m s\(^{-1}\) on average) in the last seven years (figures not shown).

We classify forecast errors based on the development rate defined as the \( V_{\text{max}} \) change in the 24 h preceding a verification time (Fig. 3). Figure 3a shows that MSLP forecasts generally have a positive bias for intensifying TCs and a negative bias for decaying TCs, suggesting that forecasts fail to capture the amplitude of intensity change in the rapid intensification and decay. These biases contribute to larger RMSEs in the forecast of intensifying and decaying TCs (Fig. 3b). These features are generally consistent with the quality of MSLP forecasts (data not shown). Mean biases and standard deviations of MSLP in the forecasts of RSMC Tokyo are categorized according to environmental conditions for open ocean cases (Fig. 4). Corresponding figures for the coastal region and land are shown in supplemental material A. Figure 4a shows that a MSLP forecast tends to be much lower than the best-track MSLP, which might have caused false alarms, when the magnitude of vertical shear exceeds 8 m s\(^{-1}\). The vertical wind shear works to suppress TC intensity (Wang and Wu 2004), but weakening processes associated with vertical wind shear may not be sufficiently reproduced. The MSLP forecast expresses a negative bias when \( \text{CAPE} \) is lower than 1000 J kg\(^{-1}\) or \( T_s \) is lower than 27°C (Figs. 4b and c). Figure 4d shows that MSLP forecast tends to be too low when \( T_{100} \) is lower than 27°C, and too high when \( T_{100} \) is approximately 29°C or higher. Since TCs are more likely to intensify with a higher \( T_{100} \), it may bring about failures to capture TC intensification. Biases mentioned above are statistically significant with at least the 95% confidence level. Figures 4e and f indicate that RSMC Tokyo predicts excessively weak TCs over the ocean when E-MPIO or \( \text{OC}_\text{PI} \) expects very intense TCs and vice versa. It is notable that the coefficient of correlation between an error and the deviation for MSLP in \( \text{OC}_\text{PI} \) was −0.33 at FT = 72 h, while that between an error and the deviation in E-MPIO was −0.27. This indicates a close relationship between

---

\(^1\) Here, the rapid intensification (decay) is defined as an increase (decrease) in the 10-min maximum sustained winds of more than 15 m s\(^{-1}\) in a 24-h period. It roughly corresponds to the definition employed in National Hurricane Center (NHC), while NHC uses a 1-min averaged maximum sustained winds.
intensity error and an index that accounts for atmosphere-ocean coupling. The above-mentioned biases generally become larger with increasing FT, highlighting the large benefits of the correction with environmental parameters in improving a relatively long-term forecast.

Biases in Vmax are consistent with those in MSLP (Fig. 5)\(^2\). That is, excessively intense TCs were predicted in RSMC Tokyo official forecasts when the vertical wind shear was greater than 8 m s\(^{-1}\), when CAPE was lower than 1000 J kg\(^{-1}\), when \(T_s\) or \(T_{100}\) was lower than 27°C, or when potential intensities expected relatively weak TCs compared to RSMC Tokyo forecast. Inversely, excessively weak TCs were predicted when \(T_{100}\) was approximately 29°C or higher, or when potential intensities expected very intense TCs.

4. Statistical correction

In the previous section, we found biases in official forecasts by stratifying the data according to several environmental parameters. Hence, information on environmental parameters is useful to diminish intensity errors with almost no additional computational cost.

As a first step to evaluate room for improvement, the following simple linear regression is considered:

\[
F' = F + \alpha_1 \text{SHEAR} + \alpha_2 \text{CAPE} + \alpha_3 \text{MSLP}_{\text{PI}} + \alpha_4 \text{MSLP}_{\text{RSMC}} + \beta + \varepsilon
\]  

(4)

where \(F\) and \(F'\) represent the official TC intensity forecast issued by RSMC Tokyo and a statistically corrected forecast, respectively. The coefficients \(\alpha_1, \alpha_2, \alpha_3, \alpha_4\) and \(\beta\) are determined for each forecast time (FT = 24, 48, 72 h) and each TC intensity metric (MSLP or Vmax) using a linear regression analysis to minimize residual \(\varepsilon\). MSLP_{PI} represents potential intensity for MSLP either based on E-MPI or OC_{PI}. The dataset is divided into seven groups depending on the year (2008−2014). In order to correct the official forecasts in each group, we first computed regression coefficients from the dataset of the other six years and applied these regression coefficients to the corrections. Note that MSLP is used to correct Vmax forecasts because the relationship between biases and deviation in potential intensities for MSLP exhibits a more linear trend than that for Vmax (Figs. 4e, 4f, 5e and 5f) and fits a linear regression. For this correction, physical environmental parameters SHEAR, CAPE and MSLP_{PI} are calculated from predicted variables available at the initial time of the forecast to evaluate the benefits in realistic conditions. More specifically, (1) the center position used to calculate environmental conditions is changed from the RSMC Tokyo best track to the location of the MSLP in the JMA-GSM forecast, and (2) relative humidity,
Fig. 4. Biases in the MSLP forecasts stratified according to physical environmental parameters: (a) magnitude of the vertical shear of horizontal wind, (b) CAPE, (c) $T_s$, (d) $T_{100}$, and (e) deviation in MSLP derived from E-MPIO relative to MSLP derived from RSMC Tokyo forecast. (f) Same as (e) but for MSLP derived from OC_PI, respectively. Dots represent the mean biases in each category and the vertical bar centered at the dot represents the standard deviation of errors in forecasts published by RSMC Tokyo. Solid lines indicate the number of cases. Left vertical axis is for the mean bias and standard deviation, while the right axis corresponds to the number of cases. Black, red, and green indicate the results of FT = 24, 48, and 72 h, respectively.

Fig. 5. Vmax biases according to physical environmental parameters, as defined for Fig. 4. Note that the horizontal axes represent potential intensities for Vmax in (e) and (f) rather than for MSLP.
temperature profile, CAPE, \( T_s \) and \( T_{100} \) at the JMA-GSM forecast center position 3 days prior to the verification time for \( FT = 72 \) h are used.

Table 1 describes the optimized coefficients as well as the improvement rate (defined as percentage of reduction in RMSE) compared to the official forecasts of RSMC Tokyo. As expected, TC intensity forecasts over the open ocean are improved by 4.9–12.1% and 7.8–16.9% when E-MPIO and OC_PI are used as a potential intensity, respectively. Improvement rates become larger with increasing FTs or by employing OC_PI. The forecasts are also improved by 10.9–35.3% for a TC that made landfall or is around a coastal region. These improvements for both \( V_{\text{max}} \) and MSLP at each FT are statistically significant, with at least the 99% confidence level, when applying a paired sample two-tailed \( t \)-test to seven groups of RMSEs. The total RMSE of corrected forecasts on each year was calculated using an OC_PI (open ocean) and E-MPIO (coastal region and land) (Fig. 2). It exhibits the robust reduction of RMSE across all years. Comparing the corrected forecasts during 2008-2014 with operational forecasts until 2007, RMSEs of corrected forecasts were statistically smaller in terms of MSLP at \( FT = 72 \) h with the confidence level of 95%, while any other difference did not reach the confidence level of 95%.

### 5. Discussion

An interesting result is that the long-term trend of intensity forecast error shows an increasing behavior. A possible explanation is that a rapid change of TC intensity has occurred frequently over the last 10 years (Fig. 6). Such frequency of rapid intensification is positively correlated with RMSEs (Figs. 2b and c). The difficulty to predict rapid changes in TC intensity (Fig. 3) could yield the behavior of TC intensity error. This increase of rapid intensification events is also seen in the dataset published by the Joint Typhoon Warning Center (Wang et al. 2015). Nevertheless, the reason for the increase of rapid intensification events is still not clear. It could result from a climatological feature (Pun et al. 2013) or merely an artifact due to changes in the policy of constructing a best track dataset, Dvorak tables, and collecting observations.

### 6. Concluding remarks

Errors in operational TC intensity forecasts by the RSMC Tokyo were analyzed based on their annual reports. This analysis reveals that (1) annual mean forecasts accuracy has not improved in over 26 years and (2) forecast errors tend to be larger in developing stages. Although official TC intensity forecasts rely on many components including numerical modeling, data assimilation schemes, observations, statistical correction, and forecaster’s skill, we found that recent official forecasts contain biases with respect to the vertical wind shear, CAPE, UOT, and potential intensities. The proposed simple statistical correction shows the potential usefulness of taking these environmental parameters into account.

This statistical model can be refined by tuning the horizontal or temporal scale of represented environmental conditions, considering the TC translation speed (Lin et al. 2008) and the vertical wind shear at different elevations. Moreover, sophisticated oceanic data allow for further improvements in the forecast accuracy.
Edited by: H. Kamahori


Supplementary materials include 2 figures and 2 tables.

References


Manuscript received 2 June 2016, accepted 12 August 2016 SOLA: https://www.jstage.jst.go.jp/browse/sola/