SOLA, *2018*, *Vol. 14*, *138−143*, *doi:10.2151/sola.2018-024*

Tropical Cyclone Intensity Prediction in the Western North Pacific Basin Using SHIPS and JMA/GSM

Munehiko Yamaguchi[1], Hiromi Owada[1, 2], Udai Shimada[1], Masahiro Sawada[1], Takeshi Iriguchi[1], Kate D. Musgrave[3], and Mark DeMaria[4]

[1]*Meteorological Research Institute of the Japan Meteorological Agency*, *Tsukuba*, *Japan*
[2]*Japan Meteorological Agency*, *Tokyo*, *Japan*
[3]*Cooperative Institute for Research in the Atmosphere*, *Colorado State University*, *Fort Collins*, *Colorado*, *USA*
[4]*NOAA/National Hurricane Center*, *Miami*, *Florida*, *USA*

Abstract

This study investigates prediction of TC intensity in the western North Pacific basin using a statistical-dynamical model called the Statistical Hurricane Intensity Prediction Scheme (SHIPS), with data sources in operations at the Japan Meteorological Agency (JMA) such as the JMA/Global Spectral Model forecast fields. In addition to predicting the change in the maximum wind (Vmax) as in the original SHIPS technique, another version of SHIPS for predicting the change in the minimum sea-level pressure (Pmin) has been developed. With 13 years of training samples, a total of 26 predictors were selected from among 52 through stepwise regression. Based on three years of independent samples, the root mean square errors of both Vmax and Pmin by the 26-predictor SHIPS model were found to be much smaller than those of the JMA/GSM and a simple climatology and persistence intensity model, which JMA official intensity forecasts are currently mainly based on. The prediction accuracy was not sensitive to the number of predictors as long as the leading predictors were included. Benefits of operationalizing SHIPS include a reduction in the errors of the JMA official intensity forecasts and an extension of their forecast length beyond the current 3 days (e.g., 5 days).

(**Citation**: Yamaguchi, M., H. Owada, U. Shimada, M. Sawada, T. Iriguchi, K. D. Musgrave, and M. DeMaria, 2018: Tropical cyclone intensity prediction in the western North Pacific basin using SHIPS and JMA/GSM. *SOLA*, **14**, 138−143, doi:10.2151/sola.2018-023.)

1. Introduction

A tropical cyclone (TC) is one of the most intense and feared storms in the world and it is of great importance to elucidate mechanisms by which it generates, moves, intensifies, and changes in structure. Improving the accuracy of TC forecasts is also important as they are directly linked to disaster risk reduction, mitigation, and preparedness. The accuracy of track forecasts has significantly improved over the last few decades (e.g., Yamaguchi et al. 2017). However, forecasting TC intensity such as the maximum sustained surface wind speed (Vmax) and the minimum sea-level pressure (Pmin) still has many challenges. Ito (2016) verified the Japan Meteorological Agency (JMA) official intensity forecasts since 1992 and showed that the errors have not been decreased. As recommended at the Eighth World Meteorological Organization (WMO) International Workshop on Tropical Cyclones (ITWC-8) in 2014, improving the prediction of TC intensity is a common challenge for both the TC research and operational communities[1].

Various initiatives have been undertaken to improve intensity forecasts. The Hurricane Forecast Improvement Project (HFIP, Gall et al. 2013) is an example of a successful initiative as the errors of TC intensity prediction by the Hurricane Weather Research and Forecasting (HWRF) model have been reduced under HFIP. Thanks to powerful supercomputers, advances in numerical weather prediction (NWP) models and data assimilation schemes, and the increased number of observations including satellites and aircraft, a dynamical approach (i.e., use of direct outputs from regional NWP models) has shown promising results (e.g., Cangialosi 2018).

Another methodology to improve intensity forecasts has been the continued development of statistical-dynamical models such as the Statistical Hurricane Intensity Prediction Scheme (SHIPS, DeMaria and Kaplan 1994, DeMaria et al. 2014). SHIPS uses multiple linear regression with TC intensity change as the predictand and environmental predictors related to it (e.g., vertical wind shear and ocean heat content). SHIPS has proven to be useful for Atlantic and eastern and central North Pacific TCs in terms of not only being one of the best-performing intensity products (DeMaria et al. 2014), but also providing information about what variables are contributing to TC intensity change. Such information is of importance for the forecasters to gain insight into the processes involved in TC intensity changes, which is useful for explaining their forecasts to users, including the media and the public. Thus, SHIPS and other similar statistical-dynamical models are used worldwide at operational centers including the Regional Specialized Meteorological Center (RSMC) in Miami (Sampson and Knaff 2014).

The JMA official intensity forecasts are currently based mainly on the JMA Global Spectral Model (JMA/GSM, JMA 2013), with a horizontal resolution of about 20 km, and a climatology and persistence intensity guidance product called the Statistical Hurricane Intensity FORecast (SHIFOR, Jarvinen and Neumann 1979). However, no statistical-dynamical model is available at the RSMC Tokyo – Typhoon Center (RSMC Tokyo – Typhoon Center 2016). Thus, this study investigates the feasibility of applying a SHIPS product for TCs in the western North Pacific basin with the aim of operationalizing SHIPS at JMA in the future. Knaff et al. (2005) demonstrated the effectiveness of a statistical-dynamical intensity model in the western North Pacific basin, but that study was based on the Navy Operational Global Atmospheric Prediction System (NOGAPS) environmental predictors and the Joint Typhoon Warning Center (JTWC) best track dataset. Thus, this study will examine the accuracy of SHIPS with database available at JMA. In the original SHIPS, the predictand was the change in Vmax. Because JMA forecasters usually look first at intensity change in terms of Pmin, another version of SHIPS is developed in which the predictand is the change in Pmin.

Section 2 describes methodology and data used in this study. The verifications are given in Section 3 and a summary is provided in Section 4.

Corresponding author: Munehiko Yamaguchi, Meteorological Research Institute of the Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, Ibaraki 305-0052, Japan. E-mail: myamagu@mri-jma.go.jp

[1] Recommendations from the IWTC-8 are available at http://www.wmo.int/pages/prog/arep/wwrp/tmr/documents/ListofRecommendations.pdf

2. Methodology and data

2.1 General methodology

SHIPS is a statistical-dynamical model to predict intensity change of TCs based on multiple linear regression. The procedure to construct and evaluate the SHIPS in this study is as follows:

1) Define variables to be considered as potential predictors of intensity change;
2) Prepare database of the variables;
3) Applying stepwise regression to select the most statistically significant predictors;
4) Compute regression coefficients for the selected predictors; and
5) Evaluate SHIPS performance with independent samples.

As indicated above, Pmin will be a predictand as well as Vmax. The list of 52 potential predictors is given in Table 1. Most of these predictors are calculated along a TC track and are retrieved using a model diagnostic code from McNoldy et al. (2012). Stepwise regression with both forward selection and backward elimination are performed with training samples to find the predictors that best explain the predictands. Based on the results of

Table 1. List of predictors. Column 1 shows the names of predictors and Column 2 describes the meaning of them. The number in Column 3 means the order in which the predictor is selected in the forward selection of stepwise regression for Vmax (labeled as V) and Pmin (labeled as P), respectively (zero means that the predictor is not selected), at 24, 48, 72, 96, and 120 hours. The check marks in Column 4 indicates that the predictor is selected as one of the 26 predictors for Vmax and Pmin, respectively.

	Column 1	Column 2	Column 3										Column 4	
		Description	24		48		72		96		120		Vmax	Pmin
			V	P	V	P	V	P	V	P	V	P		
1	PERV	12-hour change of the maximum sustained wind speed to the initial time of a prediction	13	2	0	14	0	0	0	0	0	0	✓	
2	PERP	12-hour change of the minimum sea level pressure to the initial time of a prediction	2	16	0	0	0	35	0	0	0	0		✓
3	VMPE	VMAX times PERV	14	18	23	0	33	0	0	0	0	0	✓	
4	PMPE	(MSLP minus 880) times PERP	12	17	22	10	34	36	0	0	0	38		✓
5	POT	Maximum potential intensity minus VMAX	1	1	1	1	1	0	1	0	1	31	✓	✓
6	POT2	Square of POT	10	14	12	19	0	1	0	1	0	1	✓	✓
7	MSLP	Minimum sea level pressure at the initial time of a prediction	15	13	0	6	0	6	0	3	30	3	✓	✓
8	VMAX	Maximum sustained wind speed at the initial time of a prediction	0	19	0	0	5	34	5	32	5	27		
9	VMA2	Square of VMAX	6	20	5	0	13	32	0	0	0	0	✓	
10	VMSH	VMAX times SHDC	22	4	0	5	0	5	0	0	0	16	✓	
11	SHVM	SHDC divided by VMAX	16	0	26	15	0	10	0	31	29	26	✓	✓
12	PMSH	(MSLP minus 880) times SHDC	18	0	27	26	0	20	0	0	10	5		✓
13	SHRS	850−500 hPa vertical wind shear magnitude (r = 0−500 km)	0	0	18	0	20	26	0	0	9	32		
14	SHLT	SHDC times the sine of latitude	20	7	29	13	3	15	3	13	15	20	✓	✓
15	SHDC	Same as SHRD, but round to the nearest 10	0	0	0	0	28	0	0	0	0	0	✓	✓
16	SHSH	Square of SHDC	0	15	0	8	11	11	8	6	3	15	✓	✓
17	SHRD	850−200 hPa vertical wind shear magnitude (r = 0−500 km)	0	0	0	0	27	0	12	5	0	39		
18	SHGC	Generalized vertical shear parameter (DeMaria 2010)	3	9	3	9	6	16	0	0	19	0	✓	✓
19	T150	150 hPa temperature (r = 200−800 km)	23	8	7	16	0	29	0	0	0	13		
20	T200	200 hPa temperature (r = 200−800 km)	24	0	0	29	0	30	0	0	0	28	✓	✓
21	T250	250 hPa temperature (r = 200−800 km)	19	0	0	30	8	31	11	0	7	24	✓	✓
22	EPOS	Difference of equivalent potential temperature between lifted surface parcel and environment	0	0	0	17	9	12	0	9	0	35	✓	✓
23	EPSS	Same as EPOS except that saturation equivalent potential temperature is used for environment	0	24	21	0	0	8	0	7	0	6		
24	ENEG	Same as EPOS except that negative values of the difference are only considered	0	0	0	23	18	21	13	15	8	36		
25	ENSS	Same as ENEG except that saturation equivalent potential temperature is used for environment	0	0	0	0	19	24	0	28	0	17		
26	TGRD	The magnitude of the temperature gradient between 850 and 700 hPa (r = 0−500 km)	0	0	0	0	35	28	0	0	11	37	✓	✓
27	T000	1000 hPa temperature (r = 200−800 km)	0	0	19	0	0	9	0	8	0	7		
28	E000	1000 hPa equivalent potential temperature (r = 200−800 km)	0	0	20	25	0	0	0	30	0	29		
29	R000	1000 hPa relative humidity (r = 200−800 km)	9	11	9	11	14	0	7	0	0	0		
30	RHHI	500−300 hPa relative humidity (r = 200−800 km)	0	0	0	0	30	0	23	26	17	23		
31	RHMD	700−500 hPa relative humidity (r = 200−800 km)	0	0	14	0	16	33	6	27	0	30	✓	✓
32	RHLO	850−700 hPa relative humidity (r = 200−800 km)	0	0	0	24	0	23	0	29	0	11		
33	D200	200 hPa divergence (r = 0−1000 km)	0	0	25	18	24	14	15	12	16	14	✓	✓
34	COHC	Ocean heat content (OHC)	7	6	6	7	7	7	10	4	2	4	✓	✓
35	OHC2	Square of OHC2	5	12	4	4	4	4	4	14	13	8	✓	✓
36	PHCN	Climatology of OHC	21	21	0	0	0	0	9	0	6	0		
37	CSST	Climatology of sea surface temperature	0	0	0	0	0	0	0	0	0	0		
38	SDDC	Direction of SHDC	25	0	24	0	32	0	0	17	14	12		
39	SHTD	Direction of SHRD	28	22	17	21	23	19	21	23	22	21		
40	SHTS	Direction of SHRS	27	23	15	22	22	18	18	21	23	25		
41	U200	200 hPa zonal wind (r = 200−800 km)	0	0	0	0	21	0	22	24	26	34		
42	ZNAL	Zonal storm motion	0	0	0	0	26	0	17	22	25	33	✓	✓
43	V20C	200 hPa zonal wind (r = 0−500 km)	30	0	16	27	29	17	20	20	12	19		
44	V300	300 hPa tangential wind (r = 0−500 km)	0	0	0	0	17	27	0	11	0	9		
45	V500	500 hPa tangential wind (r = 0−500 km)	29	25	11	28	12	25	19	25	21	0		
46	Z850	850 hPa vorticity	11	0	10	0	0	0	0	10	27	10	✓	✓
47	TADV	Temperature advection between 850 and 700 hPa (r = 0−500 km)	0	26	0	20	25	22	14	16	24	18	✓	✓
48	PC30	Percent area of IR Tb < -30°C (r = 50−200 km)	31	0	0	0	31	0	24	0	20	22	✓	✓
49	SDIR	Standard deviation of IR Tb (r = 0−200 km)	8	10	8	12	10	13	0	18	0	0	✓	✓
50	OSLP	Absolute of (MSLP minus 970)	17	5	28	3	0	3	0	0	28	0		✓
51	TWAT	Tendency of 850 hPa tangential wind (r = 0−500 km)	4	3	2	2	2	2	2	2	4	2	✓	✓
52	DTL	Distance to land at the initial time of a prediction	26	0	13	0	15	0	16	19	18	0		

the stepwise regression, as well as an investigation on the practical number of predictors (see Section 2.3), a total of 26 predictors is selected. The regression coefficients for these predictors are calculated from the training samples. Finally, the bias and the root mean square errors (RMSEs) of the Vmax and Pmin predictions up to 5 days are evaluated with independent samples.

2.2 Data sources

The training period includes years 2000 to 2012 and data sources to create the regression coefficients are as follows:

1) Vmax (10-min average wind) and Pmin in the JMA best track[2] are used as verification of the predictands and for predictors such as *PERV*, *VMPE*, *PERP* and *PMPE* (see Table 1);
2) Japanese 55-year Reanalysis (JRA-55, Kobayashi et al. 2015) for predictors related to the atmospheric environment in Table 1;
3) TC central position in the JMA best track for predictors such as the vertical wind shear of TCs and Ocean Heat Content (OHC) along the TC track;
4) Infrared (IR) brightness temperature (Tb) of geostationary satellites (e.g., Bessho et al. 2016) for *PC30* and *SDIR*;
5) Centennial Observation-Based Estimates of Sea-Surface Temperature (COBE-SST, Ishii et al. 2005), which is used in JRA-55, to compute an empirical maximum potential intensity (MPI, DeMaria and Kaplan 1994) for *POT* and *POT2*; and
6) OHC calculated with daily oceanic reanalysis data (MOVE/MRI.com, Usui et al. 2006) for *COHC* and *OHC2*.

For the 2013-2015 independent evaluation period, the best track values are replaced by real-time analysis data from forecasters on duty at JMA. For the atmospheric environment predictors, the JMA/GSM forecast fields are used instead of JRA-55 reanalysis. The atmospheric environment predictors are time-mean values and are calculated along a TC track predicted by the JMA/GSM for the independent samples instead of along the best track TC locations that are used for the training samples. The JMA/GSM predictions are available four times a day (00, 06, 12, and 18 UTC). The forecast periods are 84 hours at 00, 06, and 18 UTC initial times, but are extended to 264 hours for the 12 UTC initial time. For the sea-surface temperature, the Merged satellite and in-situ data Global Daily Sea-Surface Temperature (MGDSST, Kurihara et al. 2006) are used as it is the JMA/GSM ocean boundary condition.

2.3 Selection of predictors

The selected or eliminated predictors with the stepwise regression differ from one prediction time to another. As an example, Table 1 shows the predictors selected in the forward selection at prediction times of 24, 48, 72, 96, and 120 hours, respectively. The numbers in Column 3 of the table indicate the order in which the predictor is selected (zero means that the predictor is not selected). The smaller the number is, the more useful the predictor is considered to explain TC intensity change from the initial time to the target prediction time. For example, the persistence of the TC intensity change over the previous 12 hours from the initial time of the prediction (labelled as PERV and PERP) is one of the leading predictors at a prediction time of 24 hours, while they are not selected at 120 hours.

From an operational forecasting point of view, not only the forecast accuracy but also the simplicity of the forecast system is of great importance. For example, understanding the factors contributing to the intensity change would be difficult and complex if the number of predictors and/or predictors themselves are different at each prediction time. DeMaria and Kaplan (1994) also mentioned that "For consistency of the prediction, the same predictors are used at all forecast intervals, even if the coefficients are not significantly different from zero at all intervals. In general, when a predictor is not significant for a particular forecast interval, the regression coefficient becomes fairly small". Therefore, 26 predictors from among the 52 predictors are selected for Vmax and Pmin, respectively (see check mark in Column 4 of Table 1), based on verifications as will be shown in Section 3. In the selection, we consider the relative importance of each predictor shown in Column 3 of Table 1. Also we select only one predictor from among a group of similar predictors in order for the forecasters to easily understand the impact of the predictor on the predicted intensity change. For example, EPOS, EPSS, ENEG and ENSS are similar with each other, so we select only EPOS.

The regression coefficients of the 26 predictors, which indicate the relative importance of each predictor to the predictands, are shown in Fig. 1. The positive (negative) values indicate contributions to intensifying (weakening) TCs. The leading predictors (predictors with large amplitudes of regression coefficients) include those related to the current TC intensity such as MSLP, ocean heat content (COHC), and 850−200 hPa vertical wind shear magnitude (SHDC). These are similar to the leading predictors for the Atlantic and eastern North Pacific versions of SHIPS. The time variations of the regression coefficients tend to be larger for predictors that are related to initial TC intensity such as MSLP and POT with larger contributions for the shorter forecast ranges. The time variations seen in COHC and SHDC might be attributed to the average TC life cycle. The time variations of SHDC are also probably due to the interaction with other measures of vertical shear such as SHGC.

2.4 Target TCs

All TCs in the RSMC Tokyo – Typhoon Center's area of responsibility (i.e., western North Pacific basin between 0°N−60°N and 100°E−180°E) that reached an intensity classification of tropical storm or stronger during their lifetime are considered in this study. When storms are classified as a tropical depression or an extratropical cyclone, however, these times are not included in the training samples except 6 hours before they are classified as tropical storm and at the time when the storms have declined in intensity to tropical depression or transitioned to an extratropical cyclone. As Vmax is not archived in the JMA best track when the storms are classified as a tropical depression or an extratropical cyclone, Vmax is set at 30 kt during such classifications. Furthermore, times after the TC centers are over land are excluded from the training samples.

The predictions for the independent samples are performed for initial times when the JMA/GSM TC tracking data are available. The TC tracking is initiated when forecasters on duty at JMA

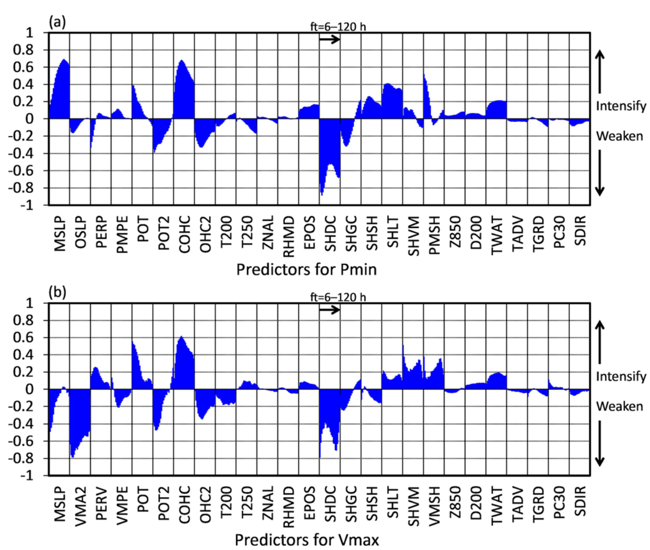

Fig. 1. Regression coefficients of the selected 26 predictors for (a) Pmin and (b) Vmax. Note that the regression coefficients for Pmin are shown with the opposite sign for easier comparison to Vmax. For each predictor, the regression coefficients are shown as a function of forecast times (ft) from 6 to 120 hours at the left margin of each abscissa box to 120 h at the right margin.

[2] JMA best tracks are available at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/trackarchives.html

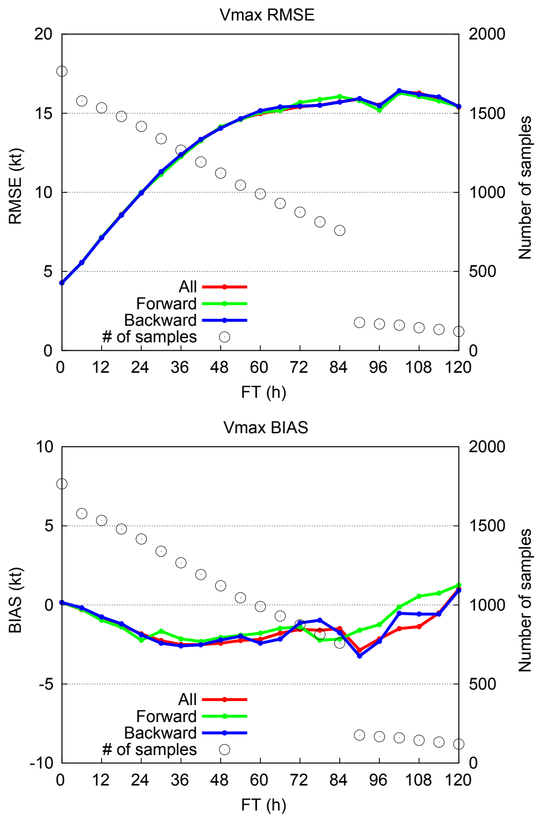

Fig. 2. Independent sample RMSE (top) and bias (bottom) of Vmax (kt) predictions, in which red, green and blue lines are for predictions with all 52 predictors and with predictors kept in the forward selection and backward elimination, respectively. Black open circles show the number of samples, corresponding to y-axis on the right.

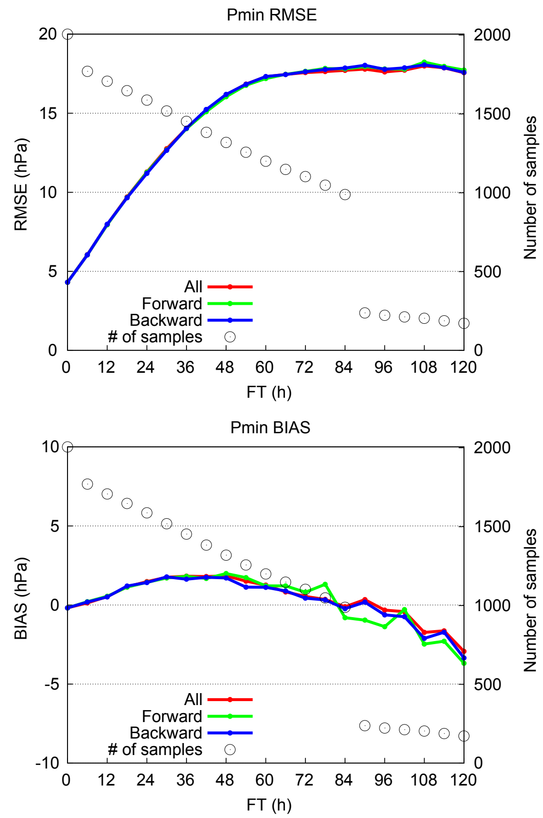

Fig. 3. As in Fig. 2, except for Pmin (hPa).

analyze TCs including those with tropical depression intensity in the western North Pacific basin. The predictions are stopped after the TCs make landfall in the TC track by JMA/GSM.

The SHIPS predictions are verified against Vmax and Pmin in the JMA best track. All prediction times when the JMA/GSM tracks the target storms are verified. As mentioned above, however, Vmax is not archived in the JMA best track when the storms are classified as a tropical depression or an extratropical cyclone. Thus such prediction times are excluded in the verification of Vmax in the independent sample. Also note that the prediction times when the verifying TC still exists in the prediction but not in the best track, or vice versa, are not included in the verification (e.g., the samples are homogenous).

3. Verification results

Verifications are shown in Figs. 2 and 3 for the independent sample predictions of Vmax and Pmin, respectively. The predictions with all 52 predictors and with predictors kept after the forward selection and backward elimination (see Table 1 for the predictors remained) result in a similar trend in both the RMSEs and bias, which indicates that the predictions with the limited number of the predictors can provide the same accuracy as those with all predictors. The bias is due to both the application of the regression equations to independent cases and to the difference in data sets used for calculating the atmospheric environmental predictors (i.e., the JMA-Reanalysis for the training samples and JMA/GSM forecast data for the independent samples). Using the JMA/GSM analyses for computation of regression coefficients instead of the JRA-55 reanalysis would be expected to reduce the bias. However, that would have required a large sample (e.g., more than 10 years) of JMA/GSM analysis fields, which would be expensive to obtain. As the model used for JRA-55 is based on the JMA/GSM with a horizontal resolution of 60 km as of December 2009 (JMA 2007, 2013), the use of JRA-55 is a reasonable choice for computation of regression coefficients.

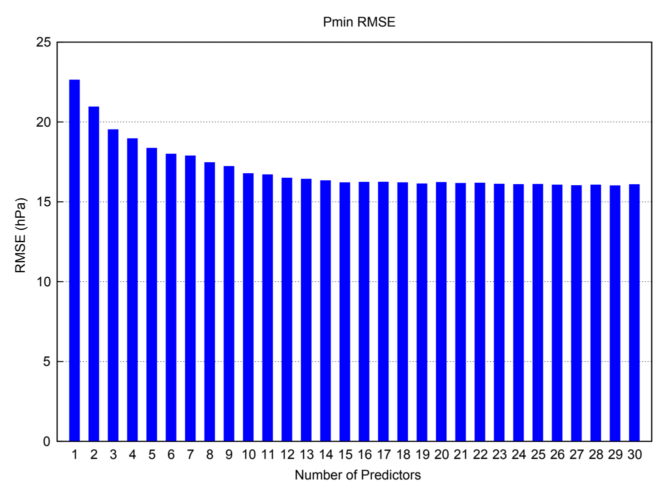

Fig. 4. Independent sample RMSE of the 48-h predictions of Pmin with addition of predictors from 1 to 30 chosen by the forward selection. See text and Table 1 for the details of the selected predictors.

As described above, for simplicity of interpretation by forecasters, it is preferable to use fewer predictors. The RMSEs of the predictions of Pmin with 1 to 30 predictors are shown in Fig. 4, which indicates the RMSE sensitivity to the number of predictors. The prediction time is 48 hours and the predictors are added one by one in the order of the forward selection shown in Table 1. For

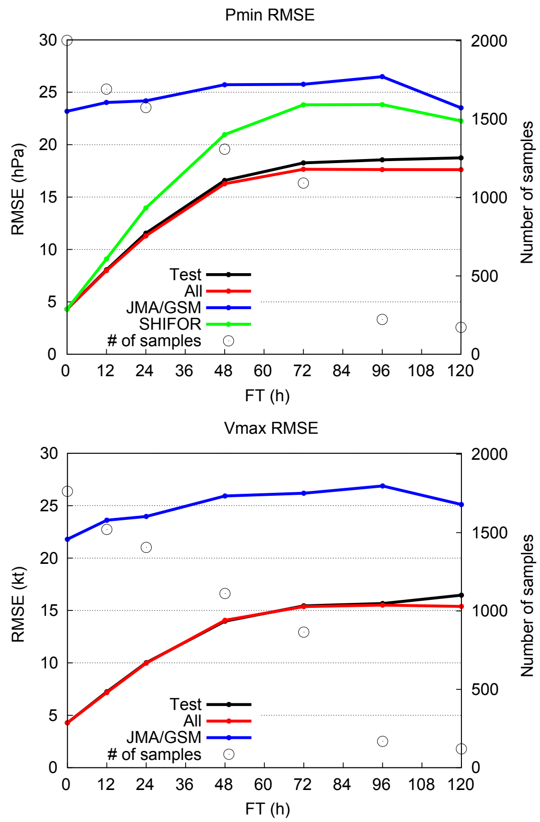

Fig. 5. Independent sample RMSE of predictions of (top) Pmin (hPa) and (bottom) Vmax (kt). SHIPS with the selected 26 and with all 52 predictors, JMA/GSM, and SHIFOR (Pmin only) are black, red, blue, and green lines, respectively. Black open circles show the number of samples corresponding to y-axis on the right.

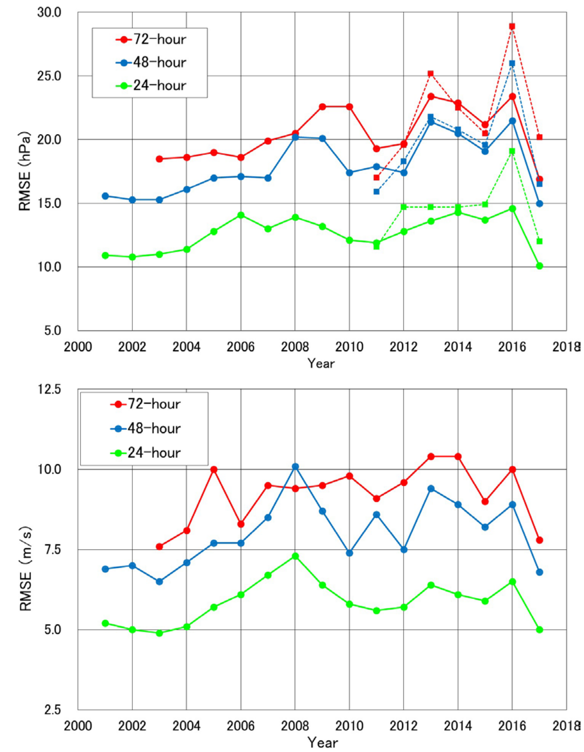

Fig. 6. Annual-average errors of JMA's operational forecasts of TC intensity in terms of (top) Pmin and (bottom) Vmax. Green, blue and red lines are for 1, 2, and 3 day forecasts, respectively. Dashed lines in the top panel shows the errors of SHIFOR.

example, POT, TWAT, OSLP, OHC2 and VMSH are used for the predictions with five predictors. Note that the RMSE decreases with the increasing number of the predictors up to about 11, but becomes saturated as more predictors are added. This result indicates that the RMSE is not sensitive to the number of the predictors as long as the leading predictors are included.

The benefit of operationalizing SHIPS at JMA is demonstrated by comparing the RMSEs of SHIPS to those of JMA/GSM and SHIFOR. Figure 5 shows the RMSEs of SHIPS predictions, Pmin and Vmax, with 52 (All) and the selected 26 (Test) predictors, the JMA/GSM, and SHIFOR (only for Pmin). The verification samples are the same as those for Fig. 3. While the RMSEs of Pmin SHIPS with the 26 predictors are slightly larger than those with the 52 predictors, they are much smaller than those of JMA/GSM and SHIFOR, which are the techniques that the JMA official intensity forecasts are currently based on. For Vmax, the RMSEs with the 26 predictors are much smaller than those of JMA/GSM and almost the same as those with the 52 predictors, except for 120 hours. The percentage improvement rates of Pmin SHIPS against SHIFOR at 24, 48, 72, 96, and 120 hours are 17, 21, 23, 22, and 16%, respectively. This result indicates great potential for significant improvements to the JMA official intensity forecasts through operationalizing SHIPS. As Sampson and Knaff (2014) show, the length of official TC intensity forecasts at major operational centers is 5 days while it is only 3 days at JMA as of August 2018. SHIPS would also allow an extension of the intensity forecast length to 5 days.

As described in Section 2.3, a total of 26 predictors from among the 52 predictors are selected for Vmax and Pmin, respectively. Although this reduction in the number of predictors results in a small increase in error, especially for Pmin (Fig. 5), the use of 26 predictors is a compromise between forecast accuracy and manageability of forecast operations.

Time series of annual-average errors of the RSMC Tokyo – Typhoon Center official forecasts of TC intensity, Pmin and Vmax, are shown in Fig. 6. The error statistics data were provided by the RSMC Tokyo – Typhoon Center. During both the difficult (large SHIFOR values) 2016 season and the relatively easier (small SHIFOR values) 2017 season, the RSMC's official intensity forecast errors were considerably below the SHIFOR values, which is an indication of skill. Since the SHIPS was first introduced for experimental use at the JMA Forecast Division during 2016, we conclude that this error reduction relative to SHIFOR indicates the benefit of using the SHIPS product. This improvement is most evident in 2016, which is significant due to the relatively large number of TCs that experienced rapid intensification during that season. Even during the easier 2017 season, the error reduction is considerable.

4. Summary

This study investigated tropical cyclone (TC) intensity prediction in the western North Pacific basin using a statistical-dynamical model called the Statistical Hurricane Intensity Prediction Scheme (SHIPS, e.g., DeMaria and Kaplan 1994), with input data available in operations at the Japan Meteorological Agency (JMA) such as the JMA Global Spectral Model (JMA/GSM) forecast fields. The predictands of SHIPS in this study are the maximum sustained surface wind speed (Vmax, 10-min average) and the minimum sea-level pressure (Pmin). Starting from 52 predictors, this study demonstrated through stepwise regression that the predictions with the limited number of the predictors can provide the same accuracy in the root mean squared errors (RMSEs) and bias as those with all predictors (Figs. 2 and 3). That is, the prediction accuracy is not so sensitive to the number of the predictors as long as the leading predictors are included (Fig. 4).

From the perspective of possible future operations of SHIPS at JMA, we selected 26 predictors for both Vmax and Pmin based on the verification results shown in Figs. 2 to 4 and Table 1, and used these 26 predictors at every forecast time. That consistency simplifies interpretation of the predictions by the forecasters. The SHIPS forecasts with the 26 predictors were then compared to those from the JMA/GSM forecasts and a climatology and persistence approach called SHIFOR, which the JMA official intensity forecasts are currently based on. The RMSEs of SHIPS with the 26 predictors were found to be much smaller than those of the JMA/GSM and SHIFOR (Fig. 5), which indicates that the JMA official TC intensity forecasts can be significantly improved through operationalizing SHIPS. Moreover, SHIPS would allow an extension of a forecast length of the JMA official intensity forecasts from the current 3 days to 5 days.

Future studies will include adding new predictors to further improve the prediction accuracy of SHIPS. For example, Shimada et al. (2018) demonstrated improved accuracy by using a microwave satellite-derived rainfall dataset from the Global Satellite Mapping of Precipitation (GSMaP, Kubota et al. 2007; Aonashi et al. 2009). Another future study will include the use of the JMA official TC track forecasts instead of JMA/GSM track predictions when calculating environmental predictors such as the ocean heat content and/or vertical wind shear. The impact is expected to be significant when the difference between the official track forecasts and the JMA/GSM track predictions is large. Such a change may require a new method to remove the GSM representation of the vortex to avoid errors in the vertical wind shear and other atmospheric predictors, as shown by DeMaria (2010).

Acknowledgments

We are deeply grateful to C. Sampson, J. Kaplan, and Dr. J. Knaff for their assistance. We acknowledge the RSMC Tokyo - Typhoon Center for providing the data used in this study, and for their generous support of this study. The opinions in this paper are those of the authors and should not be regarded as official RSMC Tokyo – Typhoon Center views. Drs. Yamaguchi and Sawada are partly supported by MEXT KAKENHI Grant 16K05556.

Edited by: K. Ito

References

Aonashi, K., J. Awaka, M. Hirose, T. Kozu, T. Kubota, G. Liu, S. Shige, S. Kida, S. Seto, N. Takahashi, and Y. N. Takayabu, 2009: GSMaP passive, microwave precipitation retrieval algorithm: Algorithm description and validation. *J. Meteor. Soc. Japan*, **87A**, 119−136.

Bessho, K., and co-authors, 2016: An introduction to Himawari-8/9— Japan's new-generation geostationary meteorological satellites. *J. Meteor. Soc. Japan*, **94**, 151−183.

Cangialosi, J. P., 2018: *National Hurricane Center Forecast Verification Report*, National Hurricane Center, 73 pp. (Available online at https://www.nhc.noaa.gov/verification/pdfs/Verification_2017.pdf, accessed 8 June 2018).

DeMaria, M., and J. Kaplan, 1994: A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting*, **9**, 209−220.

DeMaria, M., 2010: Tropical cyclone intensity change predictability estimates using a statistical-dynamical model. Extended Abstract, *29th AMS Conference on Hurricane and Tropical Meteorology*, 10−14 May 2010, Tucson, AZ.

DeMaria, M., C. Sampson, J. Knaff, and K. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387−398.

Gall, R., and co-authors, 2013: The hurricane forecast improvement project. *Bull. Amer. Meteor. Soc.*, **94**, 329−343.

Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865−879.

Ito, K., 2016: Errors in tropical cyclone intensity forecast by RSMC Tokyo and statistical correction using environmental parameters. *SOLA*, **12**, 247−252.

Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity. *NOAA Tech. Memo. NWS NHC-10*, National Technical Information Service, US, 22 pp.

JMA, 2007: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. *WMO Technical Progress Report on the Global Data-Processing and Forecasting System and Numerical Weather Prediction*, Japan Meteorological Agency, Japan, 194 pp. (Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/nwp-top.htm, accessed 8 June 2018).

JMA, 2013: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. *Appendix to WMO Numerical Weather Prediction Progress Rep.*, Japan Meteorological Agency, Japan, 177 pp. (Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2013-nwp/index.htm, accessed 8 June 2018).

Knaff, J. A., C. R. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting*, **20**, 688−699.

Kobayashi, S., and co-authors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5−48.

Kubota, T., and co-authors, 2007: Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Trans. Geosci. Remote Sens.*, **45**, 2259−2275, doi:10.1109/MICRAD.2006.1677106.

Kurihara, Y., T. Sakurai, and T. Kuragano, 2006: Global daily sea-surface temperature analysis using data from satellite microwave radiometer, satellite infrared radiometer and in-situ observations (in Japanese). *Wea. Bull.*, **73**, S1−S18.

McNoldy, B. D., K. D. Musgrave, and M. DeMaria, 2012: Diagnostics and verification of the tropical cyclone environment in regional models. *30th Conf. on Hurr. And Trop. Met.*, 15−20 April, Ponte Vedra Beach, FL. (Available online at https://ams.confex.com/ams/30Hurricane/webprogram/Paper204581.html, accessed 8 June 2018).

RSMC Tokyo – Typhoon Center, 2016: *Annual Report on the Activities of the RSMC Tokyo - Typhoon Center*, RSMC Tokyo – Typhoon Center, Japan, 95 pp. (Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/AnnualReport/2016/Text/Text2016.pdf, accessed 3 June 2018).

Sampson, C., and J. Knaff, 2014: Advances in intensity guidance. *WMO 8th International Workshop on Tropical Cyclones (IWTC-8)*, 26 pp. (Available online at http://www.wmo.int/pages/prog/arep/wwrp/new/documents/Topic2.7_AdvancesinIntensityGuidance.pdf, accessed 3 June 2018).

Shimada, U., H. Owada, M. Yamaguchi, T. Iriguchi, M. Sawada, K. Aonashi, M. DeMaria, and K. D. Musgrave, 2018: Further improvements to the statistical hurricane intensity prediction scheme using tropical cyclone rainfall and structural features. *Wea. Forecasting* (in press) (Available online at https://doi.org/10.1175/WAF-D-18-0021.1, accessed 3 June 2018).

Usui, N., S. Ishizaki, Y. Fujii, H. Tsujino, T. Yasuda, and M. Kamachi, 2006: Meteorological Research Institute multivariate ocean variational estimation (MOVE) system: Some early results. *Adv. Space Res.*, **37**, 806−822.

Yamaguchi, M., J. Ishida, H. Sato, and M. Nakagawa, 2017: WGNE intercomparison of tropical cyclone forecasts by operational NWP models: A quarter-century and beyond. *Bull. Amer. Meteor. Soc.*, **98**, 2337−2349.

Manuscript received 15 June 2018, accepted 20 August 2018
SOLA: https://www.jstage.jst.go.jp/browse/sola/