

# Statistical–Dynamical Typhoon Intensity Predictions in the Western North Pacific Using Track Pattern Clustering and Ocean Coupling Predictors

SUNG-HUN KIM

*Interdisciplinary Program in Marine Meteorology, Typhoon Research Center, Jeju National University, and National Typhoon Center, Korea Meteorological Administration, Jeju, South Korea*

IL-JU MOON

*Interdisciplinary Program in Marine Meteorology, Typhoon Research Center, Jeju National University, Jeju, South Korea*

PAO-SHIN CHU

*Department of Atmospheric Sciences, School of Ocean and Earth Science and Technology, University of Hawai'i at Mānoa, Honolulu, Hawaii*

(Manuscript received 25 June 2017, in final form 12 January 2018)

## ABSTRACT

A statistical–dynamical model for predicting tropical cyclone (TC) intensity has been developed using a track-pattern clustering (TPC) method and ocean-coupled potential predictors. Based on the fuzzy *c*-means clustering method, TC tracks during 2004–12 in the western North Pacific were categorized into five clusters, and their unique characteristics were investigated. The predictive model uses multiple linear regressions, where the predictand or the dependent variable is the change in maximum wind speed relative to the initial time. To consider TC-ocean coupling effects due to TC-induced vertical mixing and resultant surface cooling, new potential predictors were also developed for maximum potential intensity (MPI) and intensification potential (POT) using depth-averaged temperature (DAT) instead of sea surface temperature (SST). Altogether, 6 static, 11 synoptic, and 3 DAT-based potential predictors were used. Results from a series of experiments for the training period of 2004–12 using TPC and DAT-based predictors showed remarkably improved TC intensity predictions. The model was tested on predictions of TC intensity for 2013 and 2014, which are not used in the training samples. Relative to the nonclustering approach, the TPC and DAT-based predictors reduced prediction errors about 12%–25% between 24- and 96-h lead time. The present model is also compared with four operational dynamical forecast models. At short leads (up to 24 h) the present model has the smallest mean absolute errors. After a 24-h lead time, the present model still shows skill that is comparable with the best operational models.

## 1. Introduction

A tropical cyclone (TC) accompanied by strong winds, storm surges, high waves, and flooding is among the most dangerous natural hazards and can pose a great threat to a global population of almost a billion people (Peduzzi et al. 2012). Accurate predictions of TC tracks and intensities are required in order to mitigate TC damage. Over the last quarter of a century, TC track forecasts have been steadily improving, while storm intensity prediction remains a major challenge (DeMaria et al. 2014), because TC

intensification involves highly complex interactions between the atmosphere, ocean, and structure of the storm (Rappaport et al. 2012), which make accurate predictions difficult. With substantial improvement in computational power, high-resolution dynamical modeling has become a useful tool for predicting TC tracks and intensities. For intensity predictions, however, traditional statistical approaches (Jarvinen and Neumann 1979; Knaff et al. 2003) are still widely used in practice, which provide consistent and basic information.

DeMaria and Kaplan (1994a) developed a Statistical Hurricane Intensity Prediction Scheme (SHIPS) combining statistical models and dynamical models. SHIPS,

---

*Corresponding author:* Il-Ju Moon, [ijmoon@jejunu.ac.kr](mailto:ijmoon@jejunu.ac.kr)

DOI: 10.1175/WAF-D-17-0082.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

which is based on a multiregression technique, uses predictors estimated from a dynamical forecast model as well as climatological and persistence predictors. The scheme has been used for hurricane intensity guidance in the North Atlantic (NA) and eastern North Pacific (ENP) at the National Hurricane Center (NHC) (DeMaria and Kaplan 1994a, 1999; DeMaria et al. 2005), and versions of SHIPS were developed for the western North Pacific (WNP), Indian Ocean (IO), and Southern Hemisphere (SH) for the operational guidance of the Joint Typhoon Warning Center (JTWC) (Schumacher et al. 2013). Knaff et al. (2005) and Knaff and Sampson (2009) developed the SHIPS-based Statistical Typhoon Intensity Prediction Schemes for WNP, IO and SH, and the models were used as the TC intensity guidance at JTWC. Until recently, most statistical–dynamical models use regression equations, where the same predictors are used at each lead time based on the assumption that TC intensity is controlled by the same environmental predictors without taking TC tracks into account. Recent studies, however, indicate that environmental predictors related to TC activity are strongly dependent on TC track patterns (Camargo et al. 2007a,b; Chu et al. 2010; Chu and Zhao 2011; Kim et al. 2011), as discussed next.

Camargo et al. (2007a,b) developed a TC track clustering technique based on a regression mixture model and showed that each cluster has a unique correlation with sea surface temperature (SST) and large-scale atmospheric circulation patterns. Chu and Zhao (2011) also developed a mixture Gaussian model to cluster TC tracks over the WNP into several types and investigated long-term changes of the TC attributes (e.g., frequency, lifespan) for each type. Based on a different clustering method (i.e., the fuzzy *c*-means method), Kim et al. (2011) showed that large-scale environment predictors influence TCs in different ways depending on each cluster. Chu et al. (2010) also used the fuzzy *c*-means method to cluster TCs in the vicinity of Taiwan and developed a statistical model for predicting the seasonal TC frequency for each cluster using a Bayesian regression scheme. Also based on track patterns, Kim et al. (2012) developed a statistical–dynamical model for predictions of seasonal TC activity, which predicts the spatial distribution of TC track density using the correlation between the seasonal TC frequency and environmental predictors calculated from the dynamical model. These studies strongly suggest that the relationship between the TC and its environmental factors depends on the track the TC takes and showed that clustering methods can be used to improve statistical–dynamical TC predictions. Most of these clustering-based studies, however, focused on predictions of TC frequency and track density, and less on TC intensity.

Independently, Chen et al. (2011) developed the Western North Pacific Tropical Cyclone Intensity Prediction Scheme (WIPS) using the stepwise regression, with the WNP divided into three subregions according to latitude and longitude: the region near the coast of east China, the South China Sea region, and the far ocean region. They suggested that using different predictors according to subgroups could improve the TC intensity skill, but they classified TCs by region, not by TC track pattern.

The primary source of energy for TCs comes from the ocean surface (Palmen 1948; Riehl 1950). It is also well known that the cooling of the upper ocean by TC-induced mixing is one of the important factors influencing TC intensity, which is mainly determined by the upper-ocean thermal structure (Mainelli-Huber 2000; Shay et al. 2000; Lin et al. 2008, 2009a,b; Moon and Kwon 2012). Leipper and Volgenau (1972) developed the concept of ocean heat content (OHC), which is an oceanic parameter representing the upper-ocean thermal structure and is expressed as an integral of calories of seawater with temperature above 26°C:

$$\text{OHC} = C_p \rho \int_0^{\text{d}26} (T - 26) dz, \quad (1)$$

where  $C_p$  is the seawater-specific heat capacity,  $\rho$  is the water density,  $\text{d}26$  is the 26°C isothermal depth, and  $T$  is the ocean temperature at a specific depth. Currently, OHC is widely used as a predictor in TC intensity prediction (Mainelli et al. 2008; DeMaria 2009; Goni et al. 2009).

On the other hand, Price (2009) suggested that depth-averaged temperature (DAT) is a better indicator for TC intensity change than the widely used OHC:

$$\text{DAT}_d = \frac{1}{d} \int_{-d}^0 T(z) dz, \quad (2)$$

where  $d$  is the depth of TC-induced vertical mixing. This is because DAT can realistically characterize the oceanic response to TCs, that is, the effect of the TC-induced sea surface cooling for shallow continental shelf and very low OHC regions where the application of OHC was limited (Price 2009). Lin et al. (2013) also argued that DAT is a better index than SSTs in estimating the maximum potential intensity (MPI), which is an upper bound of TC intensity widely used as a key predictor. Therefore, it would be interesting to examine the effect of the use of DAT on statistical predictions of TC intensity since no such attempt has been made in the literature. Particularly, it is worthwhile to explore the role of DAT in classified TC groups because the strength of the cooling effect depends on the ocean subsurface

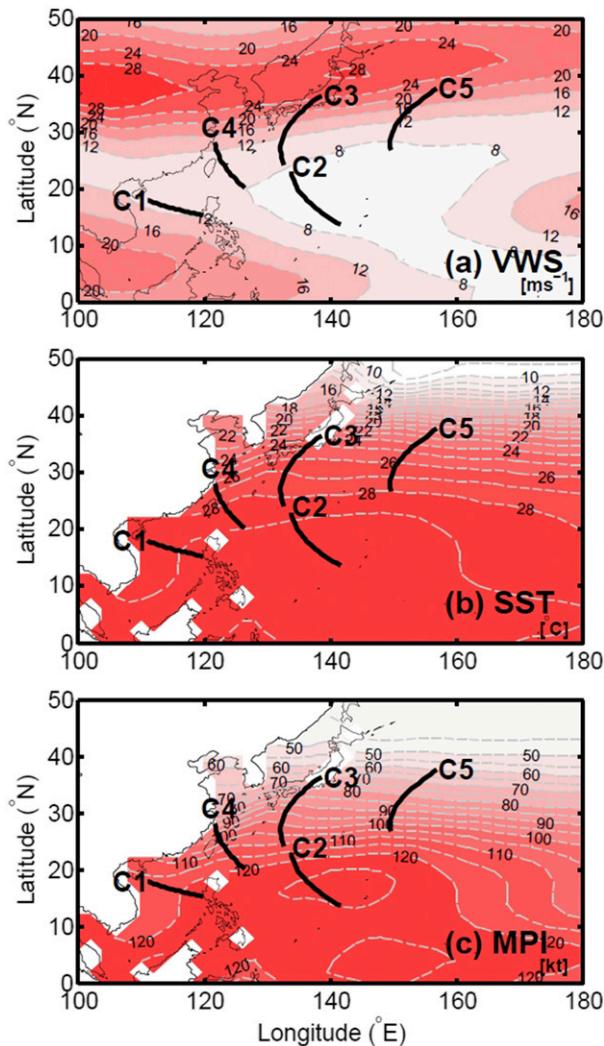


FIG. 1. Spatial distributions of (a) climatological vertical wind shear ( $\text{m s}^{-1}$ ; dashed lines), (b) SSTs ( $^{\circ}\text{C}$ ; dashed lines), and (c) MPI (kt) for typhoon season (June–November) in the WNP. Black curves represent the mean tracks of five clusters. For cluster classification, individual TC tracks were equally spaced into 21 location points, and the mean track is calculated by averaging the values at same point using all tracks belonging to a specific cluster.

thermal and haline/saline structure as well as various factors such as the storm translation speed, size, and intensity, which are related to TC tracks.

The aim of the current study, therefore, is to develop a statistical–dynamical typhoon intensity prediction scheme for the WNP using a track-pattern clustering method and DAT-based potential predictors. This new scheme is then used to investigate how much the clustering and the ocean-coupling effect may contribute to the improvement of TC intensity predictions. To test the performance of our new scheme, a series of experiments were conducted to predict the intensity of

TABLE 1. Mean maximum wind speed (kt) of TCs for five clusters at each lead time. Examples set in boldface and boldface with an asterisk represent that the differences in the mean values between each cluster and all TCs are statistically significant at the 95% and 99% confidence level based on a two-tailed  $t$  test, respectively.

Forecast time (h)	Cluster					All TCs
	C1	C2	C3	C4	C5	
24	<b>58.1*</b>	<b>69.9*</b>	<b>68.5*</b>	<b>67.0*</b>	<b>57.5*</b>	64.4
48	<b>59.4*</b>	<b>79.1*</b>	67.2	68.0	<b>57.9*</b>	67.0
72	<b>58.6*</b>	<b>83.4*</b>	<b>63.9*</b>	66.3	<b>57.4*</b>	67.7
96	<b>57.5*</b>	<b>81.6*</b>	<b>59.8*</b>	<b>62.2*</b>	<b>57.5*</b>	66.9
120	<b>54.9*</b>	<b>77.0*</b>	<b>53.5*</b>	<b>57.3*</b>	<b>59.7</b>	64.9

WNP TCs for the training period (2004–12) and the real-time prediction period (2013 and 2014).

Section 2 provides a description of the dataset and clustering method. Section 3 discusses the characteristics of classified clusters and the effect of clustering. This section also explains the static and synoptic potential predictors, including DAT-based potential predictors followed by the final predictors selected. The model performance is discussed in section 4, and our conclusions are given in section 5.

## 2. Data and clustering method

### a. Data

The present statistical–dynamical Typhoon Intensity Prediction Scheme (hereafter STIPS) is developed based on a perfect-prognosis approach (Kalnay 2003). The model is trained for the period of 2004–12 using 203 TC best-track data records and environmental reanalysis/analysis fields and is tested independently with 54 TCs in 2013 and 2014 for real-time prediction using predicted tracks and dynamical model prediction fields. The TC data for the training and test periods were obtained from the best-track data provided by the Regional Specialized Meteorological Center (RSMC) in Tokyo, Japan, and the operational 5-day track prediction results from the Korea Meteorological Administration (KMA). The intensity of the best-track data is estimated to the nearest 5 kt ( $2.514 \text{ m s}^{-1}$ ) at 6-h intervals. For this reason, model formulation as well as any discussion of intensity in this paper is conducted in terms of knots (kt;  $1 \text{ kt} = 0.51 \text{ m s}^{-1}$ ) instead of meters per second (Knaff et al. 2005).

Oceanic potential predictors were calculated using three-dimensional ocean data derived from the Hybrid Coordinate Ocean Model (HyCOM) Navy Coupled Ocean Data Assimilation (NCODA) nowcast/forecast system provided by the Naval Research Laboratory (NRL). HyCOM is a global ocean circulation model that uses a Mercator projection and a hybrid coordinate system combining isopycnic, terrain-following (sigma layers), and  $z$ -level coordinates (Bleck 2002).

TABLE 2. As in Table 1, but for the change in mean maximum wind speed (kt).

Forecast time (h)	Cluster					All TCs
	C1	C2	C3	C4	C5	
24	<b>0.1*</b>	<b>12.2*</b>	<b>-2.2*</b>	0.6	<b>-1.6*</b>	1.5
48	<b>-0.1*</b>	<b>21.1*</b>	<b>-4.8*</b>	<b>0.7</b>	<b>-2.4*</b>	3.1
72	<b>-1.8*</b>	<b>25.1*</b>	<b>-7.5*</b>	<b>-0.9*</b>	<b>-4.6*</b>	3.8
96	<b>-4.5*</b>	<b>23.3*</b>	<b>-10.3*</b>	<b>-4.3*</b>	<b>-7.3*</b>	3.3
120	<b>-6.7*</b>	<b>18.3*</b>	<b>-13.4*</b>	<b>-10.2*</b>	<b>-9.7*</b>	2.1

Atmospheric potential predictors for the training (test) period were calculated using Global Forecast System (GFS) analysis (forecast) data with  $1^\circ \times 1^\circ$  horizontal resolution and at 6-h intervals provided by the National Centers for Environmental Prediction (NCEP); this dataset has been available since 2004.

In the present STIPS model, the intensity change is predicted from the initial forecast time. For predictor selection, we used forward selection, which is the most common screening procedure in selecting a good set of predictors from a pool of potential predictors. To avoid multicollinearity, we selected the potential predictors that are not significantly correlated with each other (Aczel 1989; Fitzpatrick 1997). During the combined regression process, if the sign of each individual regression coefficient for independent predictors is changed, it is removed from the list of candidates. If the regression fitting using a combination of predictors is worse compared with an individual predictor, the single best predictor is retained. From this process, one or two predictors were finally selected.

### b. Clustering method

A fuzzy  $c$ -means clustering method (Bezdek 1981) was used to classify TC tracks (Kim et al. 2011). The fuzzy

clustering is an extension of the  $k$ -mean clustering method. This algorithm allows objects to belong to several clusters with different degrees of membership. Following the method of Kim et al. (2011), here we determined the optimal number of clusters to be five, based on four scalar validity measures (partition coefficient, partition index, separation index, and the Dunn index).

We applied the algorithm to calculate the membership coefficients for five clusters using a given track, in which the sum of the five membership coefficients equals one, and the track is finally classified into the cluster with the highest membership coefficient. For clustering, 5-day TC tracks are used rather than the entire track of the lifetime of a TC because the present version of STIPS is set to predict a total of 5 days. Therefore, the tracks of all TCs are cut into 5-day segments at 6-h intervals. TC tracks with fewer than 5 days remaining after cutting are used until the total length is 1 day. In this way, a total of 3495 track samples obtained from 203 TCs in the training period were used for cluster classification. Because of this reconstruction process of the TC track, the sample size decreases as the lead time increases. The dissimilarity between two tracks is defined as the Euclidean norm of the difference of two vectors, which contain the latitudes and longitudes for each TC track.

TABLE 3. Comparisons of standard deviations  $\sigma_{\text{WNP}}$  of the predictand (intensity change) using the mean value for the entire sample (i.e., nonclustering cases) with those  $\sigma_c$  using the mean of each cluster at each lead time. The reduction rates of  $\sigma_c$  relative to  $\sigma_{\text{WNP}}$  are indicated in parentheses.

Forecast time (h)		Cluster					All TCs
		C1	C2	C3	C4	C5	
24	$\sigma_{\text{WNP}}$	15.9	17.4	14.5	16.8	12.0	15.6
	$\sigma_c$	15.8 (0.4%)	13.7 (21.0%)	14.0 (3.3%)	16.8 (0.1%)	11.6 (3.3%)	14.8 (5.3%)
48	$\sigma_{\text{WNP}}$	22.6	29.3	22.5	25.1	17.5	24.2
	$\sigma_c$	22.4 (1.0%)	23.2 (20.8%)	21.1 (6.3%)	25.0 (0.4%)	16.7 (5.0%)	22.3 (7.9%)
72	$\sigma_{\text{WNP}}$	25.5	34.7	25.9	28.6	21.6	28.5
	$\sigma_c$	24.9 (2.3%)	27.4 (21.1%)	23.4 (9.8%)	28.3 (1.2%)	19.9 (7.6%)	25.6 (10.5%)
96	$\sigma_{\text{WNP}}$	29.0	34.6	28.3	28.0	23.7	30.2
	$\sigma_c$	27.9 (3.5%)	28.3 (18.2%)	24.8 (12.3%)	27.0 (3.6%)	21.3 (10.2%)	26.8 (11.5%)
120	$\sigma_{\text{WNP}}$	31.4	32.7	30.7	26.8	23.9	30.6
	$\sigma_c$	30.2 (3.8%)	28.3 (13.2%)	26.6 (13.4%)	23.8 (11.0%)	20.9 (12.3%)	27.1 (11.3%)

TABLE 4. List of static and synoptic potential predictors used in the present model.

Predictor	Description
iWIND	Initial max wind speed (kt)
DVMX	12-h change in intensity
LON	Longitude
LAT	Latitude
MOV	Storm translational speed ( $\text{m s}^{-1}$ )
LAND	Ratio of landmass within 200 km from center
SST	Area-averaged (0–200 km) sea surface temperature
MPI	Area-averaged (0–200 km) max potential intensity based on empirical equation
POT	MPI – iWIND
OHC	Area-averaged (0–200 km) ocean heat content
RHHI	Area-averaged (200–800 km) relative humidity at 500–300 hPa
RHLO	Area-averaged (200–800 km) relative humidity at 850–700 hPa
SH200	Area-averaged (200–800 km) 200–850-hPa vertical wind shear
SH500	Area-averaged (200–800 km) 500–850-hPa vertical wind shear
T200	Area-averaged (200–800 km) air temp at 200 hPa
U200	Area-averaged (200–800 km) zonal wind at 200 hPa
RV850	Area-averaged (0–1000 km) relative vorticity at 850 hPa

### c. Characteristics of classified clusters

In this section, we examine the characteristics of the five clusters, C1–C5, classified by the fuzzy *c*-means clustering method. The largest cluster among the five clusters is C3 (832, 23.6%). The TCs in C3 make landfall over Japan or the Korean Peninsula (Fig. 1, solid line). C1 is the second largest cluster (806, 23.2%) and includes many TCs heading westward toward Hainan Island in China, after passing through the Philippines. C4 (758, 21.3%) includes TCs that make landfall in Taiwan and along the eastern China coasts. TCs in C2 (616, 18.5%) mostly head northwestward from the eastern Philippines Sea. Type C5 has the lowest frequency of occurrence (483, 13.4%), and its TCs pass mainly over the Kuroshio Extension, farther away from the East Asian coast.

Each cluster exhibits distinctive features in mean intensity (Table 1) and mean intensity change (Table 2). TCs in C1 reside mainly in the South China Sea, where the upper-ocean heat content is relatively low compared to other tropical regions, and TCs are most likely to make landfall over Hainan Island and Guangdong. C1, along with C5, shows the lowest mean intensity (Table 1) and a distinct weakening tendency (Table 2) over the prediction lead time. TCs in C2 spend their lifetime in the tropics where the vertical wind shear (VWS) is weak (Fig. 1a) and SSTs are warmer (Fig. 1b), which are

favorable conditions for TC development. These conditions result in the strongest mean intensity and noticeable TC intensification compared to the other clusters (Tables 1 and 2). Many TCs in C2 subsequently evolve into typhoons in C3 (26%) and C4 (37%), respectively. Most TCs in C3 experience the largest weakening stage as forecast lead times increase mainly because of lower SSTs and stronger VWS along the track (Figs. 1a,b). The majority of these TCs finally make landfall over the Korean Peninsula and Japan. Statistics reveal that more than half of the TCs in C3 made landfall in the Korean Peninsula or Japan after 90-h lead time. Most TCs in C4 initially have relatively strong intensities in the tropics and then rapidly weaken after landfall across Taiwan and eastern China. C5 passes over the eastern ocean of Japan, where the VWS is strong and the SST is low (Figs. 1a,b), leading to significant weakening in intensity.

Note that in Table 1 the difference in mean values between using each cluster and all TCs (nonclustering cases) is mostly statistically significant (22 out of 25) at the 5% level based on a two-tailed Student's *t* test. These differences in mean values become even more pronounced when the change in maximum wind speed is considered (Table 2), in which 24 out of 25 have statistically significant differences. This simple comparison suggests that introducing a clustering method to STIPS can reduce the variability of TC intensity changes, which will be discussed in the next section.

### d. Benefits of cluster analysis

As seen in Tables 1 and 2, the classified clusters show distinctive characteristics in terms of the mean intensity and tendency of intensity change. To quantitatively evaluate the effect of clustering, we compared the standard deviations (SD) of the predictand for each cluster before and after clustering at each lead time (Table 3), that is, each cluster's SD from the entire sample mean  $\sigma_{\text{WNP}}$  and from the respective cluster's mean  $\sigma_c$ :

$$\sigma_{\text{WNP}} = \left[ \frac{\sum (X_c - \bar{X}_{\text{WNP}})^2}{n_c} \right]^{1/2} \quad \text{and} \quad (3)$$

$$\sigma_c = \left[ \frac{\sum (X_c - \bar{X}_c)^2}{n_c} \right]^{1/2}. \quad (4)$$

Here,  $X_c$  refers to the intensity change (i.e., predictand) of an individual storm,  $\bar{X}_c$  is the mean intensity change for each cluster,  $\bar{X}_{\text{WNP}}$  is the mean intensity change for the entire sample, and  $n_c$  is the number of samples for each cluster. Comparison of  $\sigma_c$  and  $\sigma_{\text{WNP}}$  reveals that

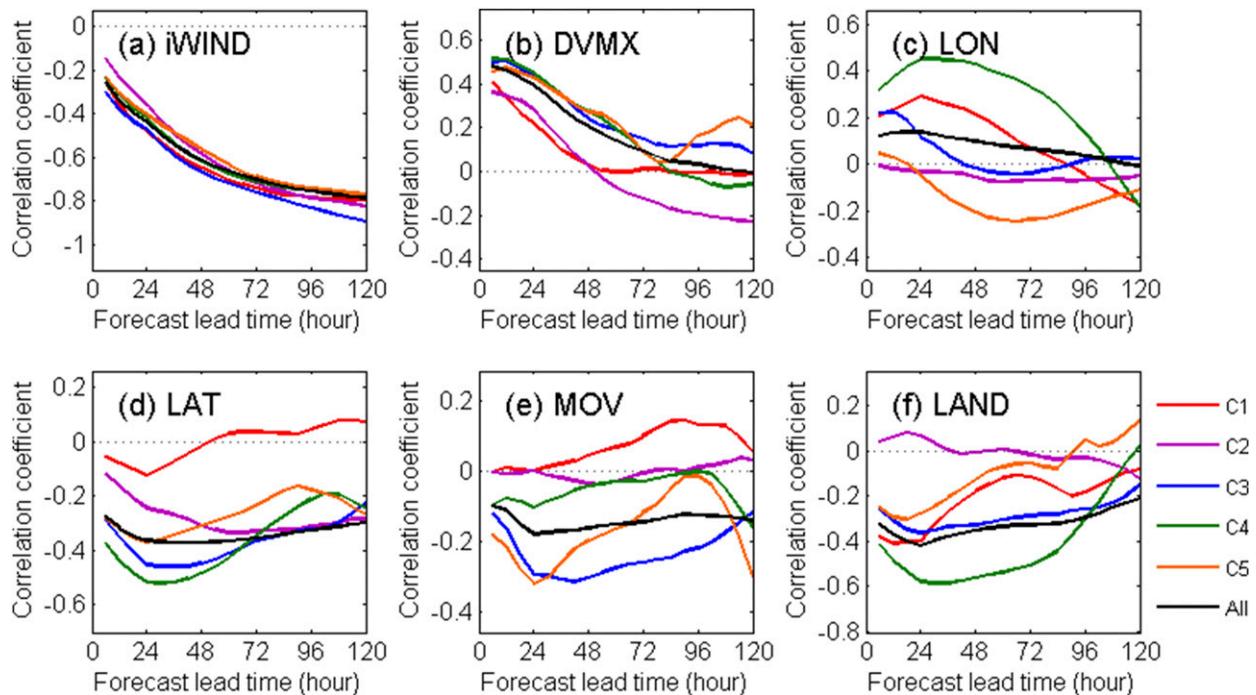


FIG. 2. The correlation coefficients between six static predictors (a) iWIND, (b) DVMX, (c) LON, (d) LAT, (e) MOV, and (f) LAND and the change in TC intensity for five clusters and the nonclustering case using all TCs at each forecast time. The correlation coefficients for individual predictors at each forecast lead time were calculated from the relation between intensity change and static information obtained using all TC track samples during the training period.

clustering reduces the variability of the predictand (i.e., intensity change) with reduction rates ranging from 0.1% to as high as 21.1% (see the values in parentheses in Table 3). The largest reduction (21.1%) was observed at 24- and 72-h lead times in C2. The significant reduction of SD suggests that the clustering-based STIPS can reduce the variability in intensity change and thus improve the TC intensity prediction skill. The clustering effects on real TC predictions will be discussed in section 3.

### 3. Statistical–dynamical typhoon intensity prediction scheme

#### a. Static and synoptic potential predictors

The present STIPS model uses various potential static and synoptic predictors, which have been widely employed in many statistical TC intensity prediction models (Elsberry et al. 1974; Jarvinen and Neumann 1979). In this study, six potential static predictors are used for the model development (Table 4): initial maximum wind speed (iWIND), intensity changes within 12h (DVMX), the longitude (LON) and latitude (LAT) of the TC center, storm translation speed (MOV), and the ratio of land

within 200 km of the TC center (LAND). They are obtained from the TC best-track information. The potential synoptic predictors, or environmental predictors, are estimated from the prediction results of the dynamical model. The present model uses 11 potential synoptic predictors (Table 4): SST, MPI, intensification potential (POT), OHC, the relative humidity of the upper and lower layers (RHHI and RHLO), the vertical wind shear of the upper and lower layers (SH200 and SH500), 200-hPa air temperature (T200), 200-hPa zonal wind (U200), and 850-hPa relative vorticity (RV850). The value averaged within 200-km radius from the storm center is used to estimate the oceanic predictors such as OHC, SST, MPI, and POT. The 200-km radius is determined based on the fact that TC-induced sea surface cooling generally occurs over strong wind regions, with its maximum at a distance greater than the maximum wind speed radius (RMW) from the TC center. In the WNP, the mean radius of 34-kt winds is about 248 km (Sampson et al. 2017), and the increase of inner SSTs within the range 2–3 times the RMW is known to contribute greatly to the increase of TC intensity (Sun et al. 2013). Atmospheric predictors such as RHHI, RHLO, SH200, SH500, T200, and U200 are estimated between 200 and 800 km from the storm center, in which doughnut-shape averaging is used to

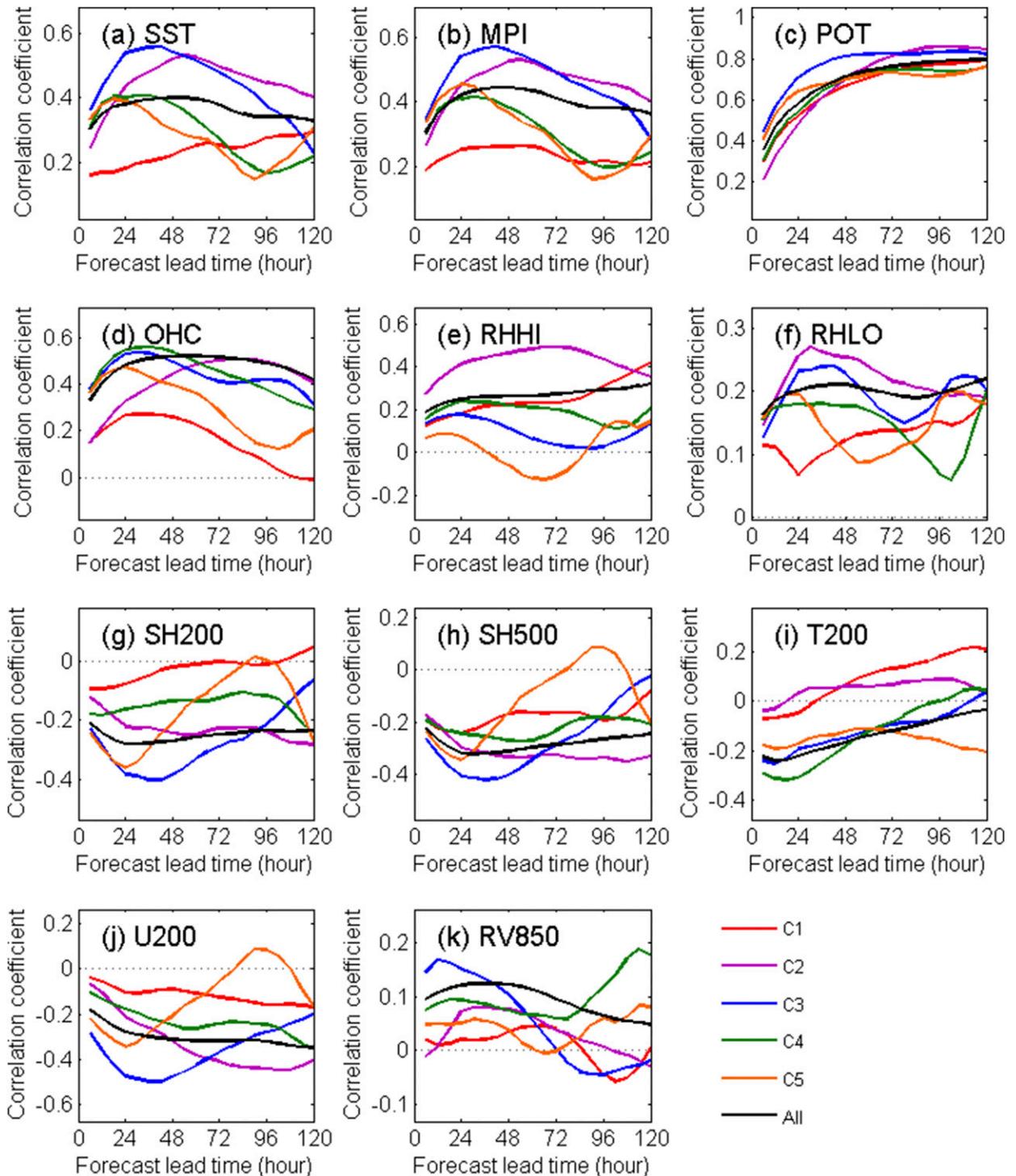


FIG. 3. As in Fig. 2, but for 11 synoptic predictors: (a) SST, (b) MPI, (c) POT, (d) OHC, (e) RHHI, (f) RHLO, (g) SH200, (h) SH500, (i) T200, (j) U200, and (k) RV850.

remove the TC effects included in the analysis dataset. For RV850, it is averaged within a 1000-km radius from the storm center.

The predictor iWIND is the most important factor in the static predictor pool and has a strong negative correlation with the intensity change for all five clusters at

TABLE 5. List of three DAT-based potential predictors used in the model.

Predictor	Description
$\text{DAT}_h$ , $h = 10, 20, 30, \dots, 120$	Ocean temp averaged from the surface down to various depths (10–120 m, 10-m interval)
$\text{MPI}_h$ , $h = 10, 20, 30, \dots, 120$	$\text{DAT}_h$ -based MPI
$\text{POT}_h$ , $h = 10, 20, 30, \dots, 120$	$\text{DAT}_h$ -based POT

longer leads (Fig. 2a). DeMaria and Kaplan (1994a) noted that strong storms have less potential for further intensification because their intensities are already closer to the MPI. DVMX represents the persistence of intensity change (Fig. 2b), which exhibits the highest correlation with intensity change in 6-h forecast time for most clusters (i.e., C3–C5), consistent with the results of Knaff et al. (2005) and Chen et al. (2011). The relationship between TC locations (longitude and latitude) and intensity change is dependent on the track types (Figs. 2c,d). TCs that move eastward after recurving (e.g., C5 in Fig. 1) tend to weaken as a result of encountering low SSTs and high westerly wind shear. This is reflected as a negative correlation between longitude and intensity change in Fig. 2c. The intensification rates of C1 and C4, which are affected by land as TCs move to the west, show a strong positive correlation with LON (Fig. 2c). For LAT, an overall negative correlation with intensity change is dominant (Fig. 2d) since TCs, on average, become weaker when they move northward because of unfavorable environmental conditions to TC development at higher latitudes.

MOV is known to be an important factor for TC–ocean interaction (Lin et al. 2008, 2009b). Slow-moving TCs tend to cause a large cold wake that could limit TC intensification. This relationship is only found in C1,

where most TCs experience a large cooling effect as a result of the shallow-ocean mixed layer of the South China Sea. For the other clusters (particularly for C3 and C5), a negative correlation between MOV and the intensity change is clearly shown. As TCs encountered the midlatitude jet stream (C3 and C5), MOV values tend to increase, but the TC intensity weakens due to strong wind shear (Fig. 2e).

LAND is defined as the ratio of landmass within a radius of 200 km from the TC center. When TCs approach land, a relatively dry and cold air mass entrains into the outer circulation of the TCs and the friction over land increases, leading to TC weakening (Kaplan and DeMaria 1995, 2001; DeMaria et al. 2006). Therefore, LAND and intensity changes are negatively correlated in most of the clusters except the nonlanding clusters (C2 and C5 in Fig. 2f). The highest negative correlation is found in C4, where most TCs with relatively strong intensity weaken rapidly during the landfall period over mountainous Taiwan and the huge landmass of eastern China (Fig. 2f).

We now turn our attention to the potential synoptic predictors (Table 4). MPI is the upper bound of TC intensity given the atmospheric vertical profile and pre-cyclone SST (Emanuel 1988). The MPI is empirically estimated according to DeMaria and Kaplan (1994a). An exponential function is derived by fitting the upper 95th percentile SST per  $0.5^\circ\text{C}$  interval in the scatterplot between SST and TC maximum intensity. Here, the SST is extracted from the HyCOM–NCODA dataset for the period of 2004–14 and averaged within a 200-km radius from the center of each TC:

$$\text{MPI} = A + B e^{C(\text{SST} - T_0)}, \quad (5)$$

where  $A = 39.91$  kt,  $B = 96.0$  kt,  $C = 0.1837^\circ\text{C}^{-1}$ , and  $T_0 = 30.0^\circ\text{C}$ . In Fig. 3b, MPI is highly correlated with TC intensity change for C3 but weakly correlated for C1

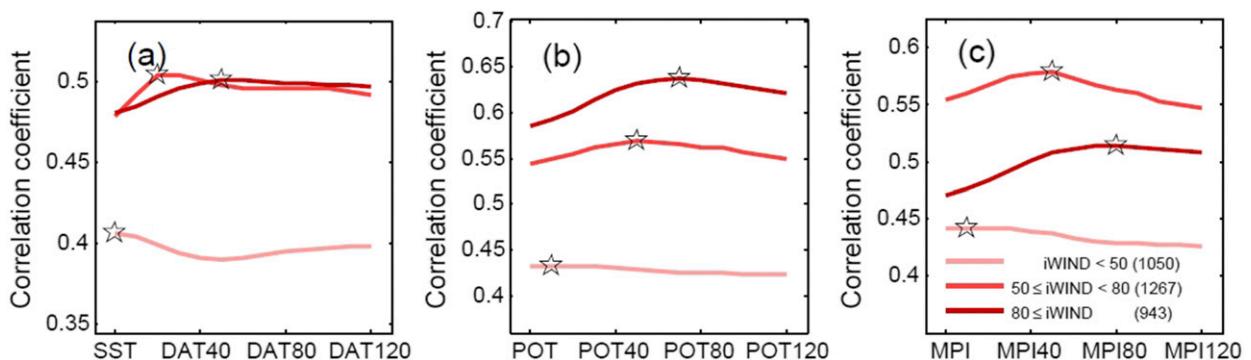


FIG. 4. The comparison of the correlation coefficients between three thermodynamical predictors (a) DAT, (b) POT, and (c) MPI and the 24-h changes in TC intensity for three groups classified by  $i\text{WIND}$ . Open pentagrams represent the location of the maximum value for the three groups. Sample sizes are indicated in parentheses in (c).

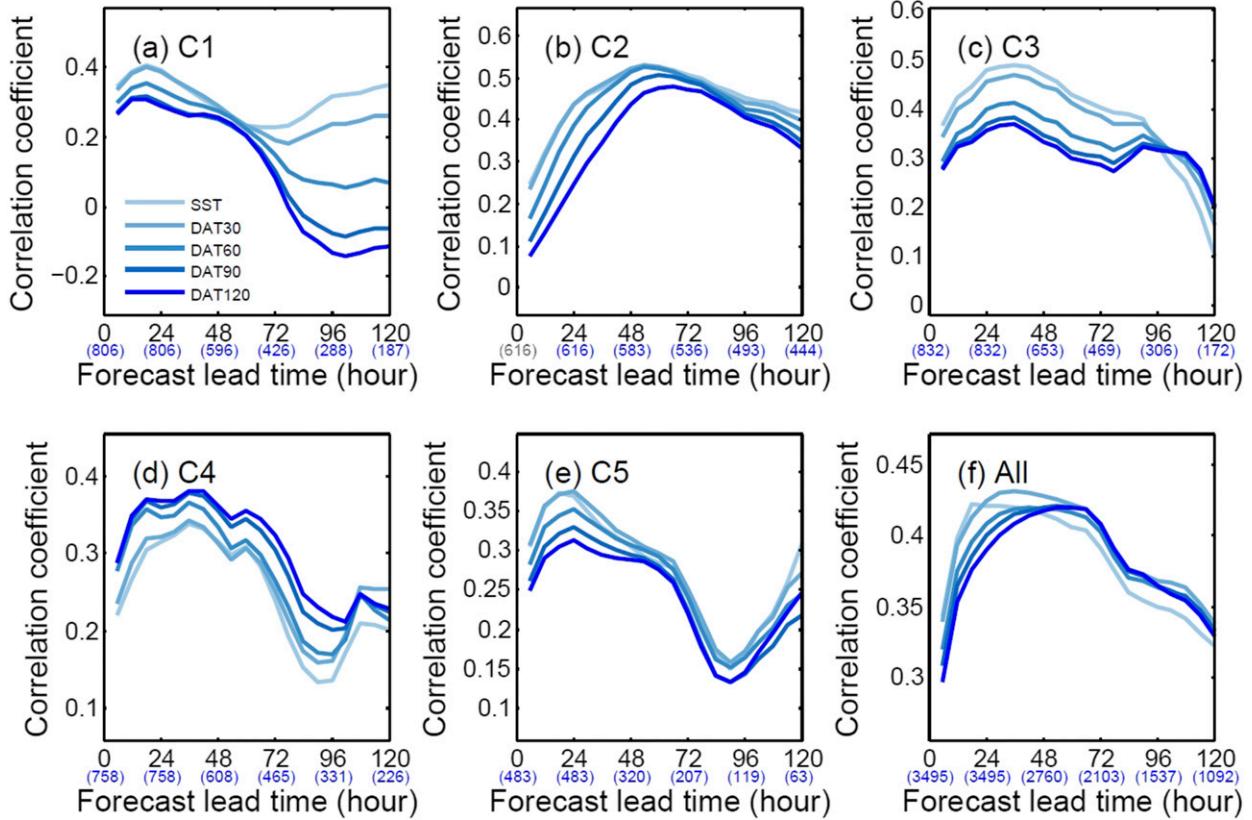


FIG. 5. Correlation coefficients between the various DATs (SST, DAT30, DAT60, DAT90, and DAT120) and TC intensity change for five clusters (a) C1, (b) C2, (c) C3, (d) C4, and (e) C5 and (f) the nonclustering case using all TCs at each lead time. Sample sizes are indicated in parentheses.

(Fig. 3b). This is partly explained by the fact that the MPI of C3 varies greatly along the TC track, but that of C1 does not change much (Fig. 1c). This pattern in MPI is very similar to that of SST (Fig. 3a).

POT is the TC potential future intensity change (DeMaria and Kaplan 1994b), defined as the difference between MPI and maximum wind at the initial time:

$$POT = MPI - iWIND. \tag{6}$$

POT is considered the most important predictor in SHIPS (DeMaria and Kaplan 1994a, 1999) and WIPS (Chen et al. 2011). Indeed, our analysis also shows that POT has the highest correlations with the predictand for lead times after 24 h, and this high correlation is maintained for all five clusters (Fig. 3c).

Lin et al. (2009a,b) suggested that warm eddies and ocean warm currents could make a critical contribution to the rapid intensification of TCs. These warm features are characterized by high OHC and can effectively mitigate the TC-induced negative feedback due to sea surface cooling. Analysis reveals that OHC is highly

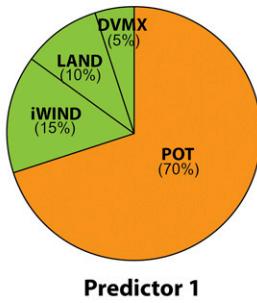
correlated with the predictand in most clusters except for C1 (Fig. 3d). For C1, three thermodynamic predictors (SST, OHC, and MPI) are not significantly correlated with the predictand. This is probably because the South China Sea has spatially homogenous thermal conditions and is surrounded by land (see Fig. 1b), and thus the TC intensity in this region is mainly controlled by factors other than thermodynamic predictors.

Relative humidity is known to affect convective buoyancy, which is a direct source of TC energy (Wu et al. 2012; Bogner et al. 2000; Knaff et al. 2005). In this

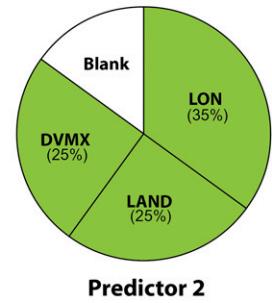
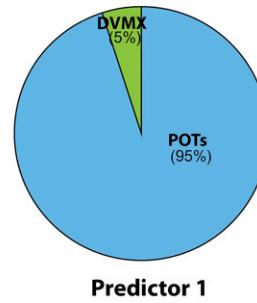
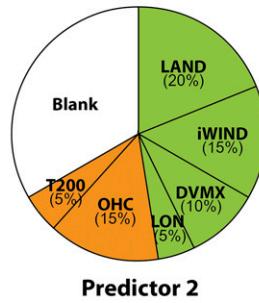
TABLE 6. Experimental designs for investigating the effects of using clustering and DAT-based predictors. The open circle and the cross indicate the methods applied and not applied, respectively.

Model	Clustering	DAT-based predictor
STIPS-SST	×	×
STIPS-DAT	×	○
CSTIPS-SST	○	×
CSTIPS-DAT	○	○

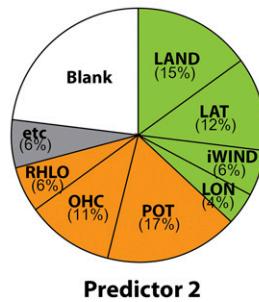
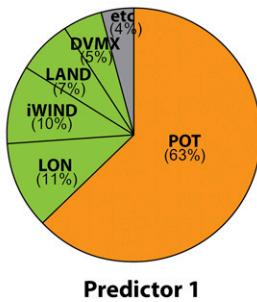
(a) STIPS-SST



(b) STIPS-DAT



(c) CSTIPS-SST



(d) CSTIPS-DAT

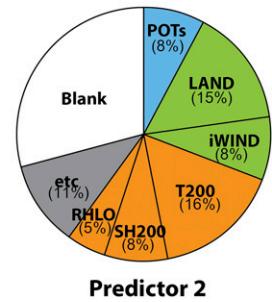
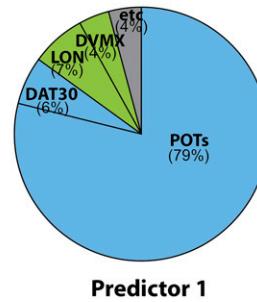


FIG. 6. Pie charts representing the ratio of the final selected predictors for (a) STIPS-SST, (b) STIPS-DAT, (c) CSTIPS-SST, and (d) CSTIPS-DAT. Predictor 1 represents the first selected predictor with the highest correlation coefficients with the predictand (intensity change). Predictor 2 represents the second selected predictor in the regression equation. Colors indicate the type of predictor (i.e., static is green, synoptic is orange, DAT based is blue, and the others are gray). If only one predictor is selected, it is marked as blank in predictor 2.

study, the relative humidity was divided into RHLO and RHHI based on Knaff et al. (2005). RHLO (RHHI) is calculated in the atmospheric layer from 850 to 700 hPa (500–300 hPa), within an annulus of 200–800-km radius (donut shape) from the center of the TC. Knaff et al. (2005) reported that RHHI is statistically important at all forecast lead times, but RHLO is not. However, our results for C2 reveal that both RHHI and RHLO have a moderate positive correlation with intensity change (Figs. 3e,f). This result is an example of how a previously unrecognized predictor could be potentially useful for a particular cluster.

Many studies have shown that VWS is an important predictor for TC intensity change (Jones 2000; DeMaria 1996; Wang and Holland 1996; Frank and Ritchie 2001; Corbosiero and Molinari 2002; Knaff et al. 2004). The VWS is defined by the magnitude of the vector difference between the two different layers, 200–850 hPa (SH200) and 500–850 hPa (SH500):

$$\text{SH200} = \sqrt{(U_{200} - U_{850})^2 + (V_{200} - V_{850})^2} \quad \text{and} \quad (7)$$

$$\text{SH500} = \sqrt{(U_{500} - U_{850})^2 + (V_{500} - V_{850})^2}, \quad (8)$$

where  $U_{200}$  ( $V_{200}$ ) is the 200-hPa zonal (meridional) wind,  $U_{500}$  ( $V_{500}$ ) is the 500-hPa zonal (meridional) wind, and  $U_{850}$  ( $V_{850}$ ) is the 850-hPa zonal (meridional) wind. Strong VWS disrupts the organized deep convection, which inhibits TC intensification. The wind-related predictors such as VWS (Figs. 3g,h), RV850 (Fig. 3k), and U200 (Fig. 3j) are highly correlated with the intensity change in C3 in which many TCs pass through the subtropical region where the VWS is strong (Fig. 1a). In contrast, a low correlation is found in C1 because most TCs in C1 move to the northwest, parallel to the vertical wind shear gradients. These results are consistent with those of Chen et al. (2011), in which VWS is an important predictor in the eastern China Sea region (corresponding to C3) and not in the South China Sea region (corresponding to C1).

#### b. DAT-based potential predictors

TC-induced vertical mixing is affected by various factors such as TC intensity, storm translation speed, the

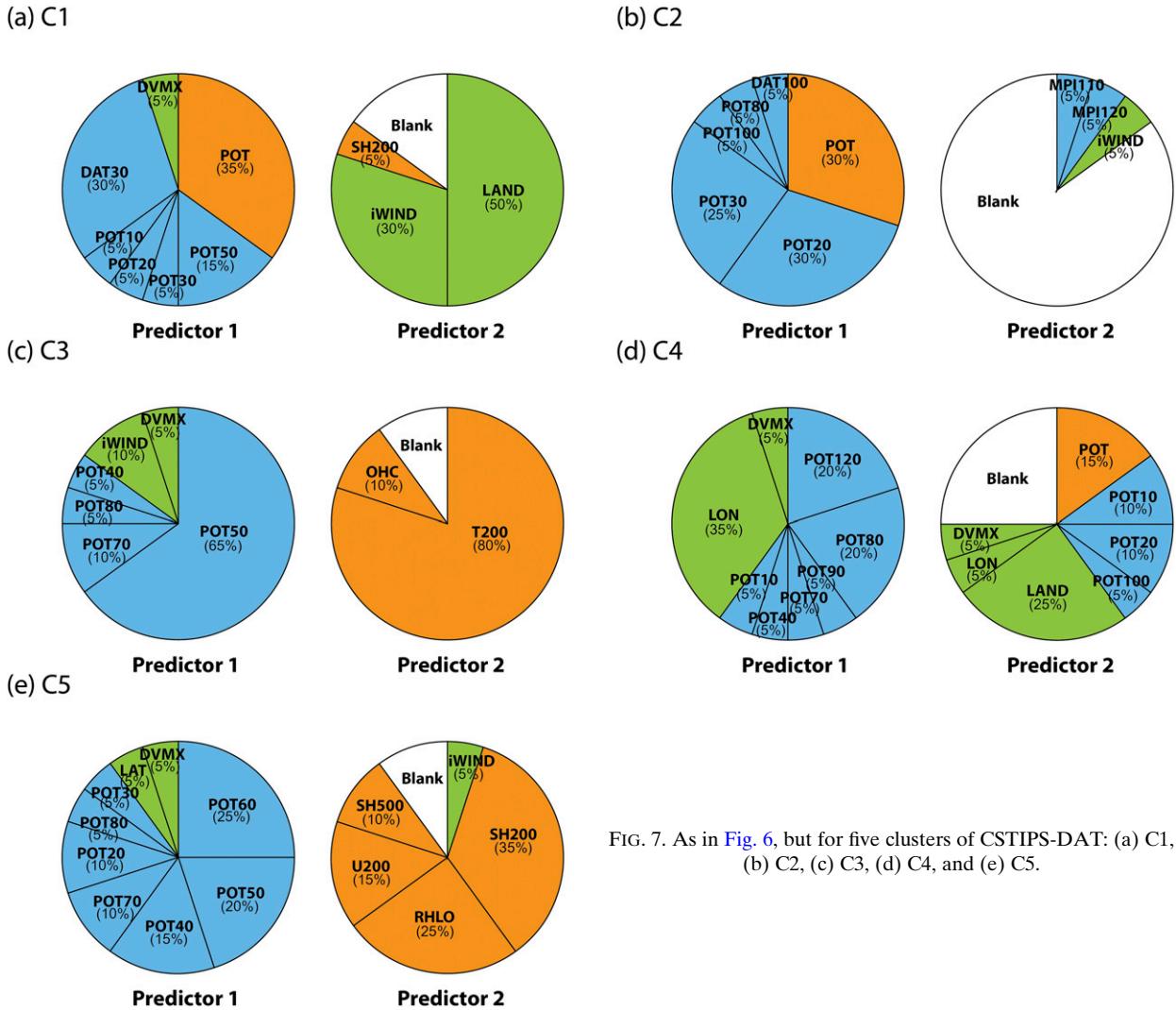


FIG. 7. As in Fig. 6, but for five clusters of CSTIPS-DAT: (a) C1, (b) C2, (c) C3, (d) C4, and (e) C5.

Coriolis parameter, and upper-ocean structure. The depth of the vertical mixing is known to be typically 60–100 m in major tropical cyclones (Price 2009). In this study, when considering the effect of the TC-induced vertical mixing on the intensity predictions, new MPIs and POTs are additionally calculated using DATs based on various mixing depths from the surface to 120 m (10-m interval) instead of using the SST (Table 5). As the storm intensity increases, DAT using deeper mixing depths is more correlated with TC intensity changes (Lin et al. 2008). To investigate the relationship between TC intensity and mixing depth, we classified TCs into three groups according to iWIND (less than 50 kt, 50–80 kt, and over 80 kt) and examined the variations of correlation coefficients between thermodynamic predictors and 24-h intensity changes as a function of different mixing depths (Fig. 4). It is shown that initially intense TCs tend to have higher correlations with the three thermodynamic predictors for deeper mixing depths.

It is interesting to ask whether this feature is also shown when TC tracks are clustered. Figure 5 presents the correlation coefficients between the DAT-based predictors and TC intensity change for five clusters and nonclustering cases at various lead times. C1 and C5, with relatively weak intensity (see Table 1), have the highest correlation coefficients up to 72-h leads when mixing depths are shallow. C4 with strong intensity has the highest correlations when mixing depths are deep (Fig. 5). The tendency of C2 and C3 with medium intensity is similar to that of C1 and C5, probably as a result of a thick mixed layer in this region leading to a small temperature change with mixing depth.

*c. Effects of using clustering and DAT-based predictors*

To investigate further the effect of using clustering and DAT-based predictors on statistical–dynamical TC

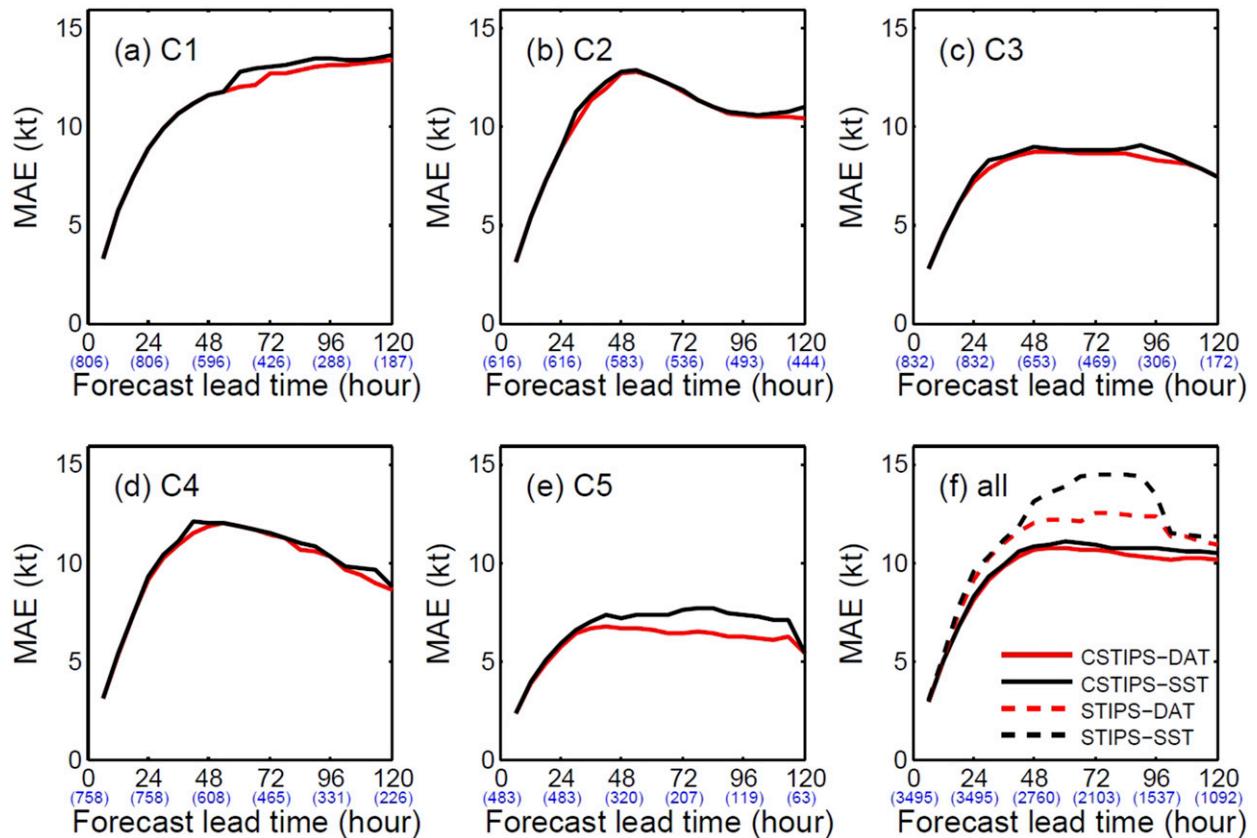


FIG. 8. Comparisons of MAEs of the maximum intensity (kt) for five clusters (a) C1, (b) C2, (c) C3, (d) C4, and (e) C5 and (f) the nonclustering cases using all TCs at each lead time during the training period (2004–12). For the five clusters in (a)–(e), two results from CSTIPS-DAT and CSTIPS-SST are compared. For the nonclustering case in (f), the results from all experiments in Table 6 are compared. Sample sizes are indicated in parentheses.

intensity predictions, four sets of experiments were conducted for the training period (Table 6): a run without the use of both clustering and DAT-based predictors (hereafter referred to as the STIPS-SST), a run using clustering only (CSTIPS-SST), a run using DAT-based predictors only (STIPS-DAT), and a run using both clustering and DAT-based predictors (CSTIPS-DAT). We first examined the predictors selected for each experiment (Fig. 6). For nonclustered models, multiple regression equations are developed for each prediction time, yielding a total of 20 equations. However, clustered models have 100 equations because five clusters are also considered. Given the large number of equations, we present the ratio of selected final predictors for each model. Here, it should be noted that the present multiple regression model allows the combination of up to five predictors to avoid overfitting, but only one or two predictors were finally selected because of multicollinearity constraints. Figure 6 shows the ratio of the final selected predictors for each set of four experiments. The first and second predictors selected from the

screening procedure for each prediction time are termed predictor 1 and predictor 2, respectively.

The most frequently selected predictor 1 was POT (including DAT-based POT) for all experiments. For STIPS-DAT, in particular, POT was selected 95% of the time as the predictor 1. The iWIND, SST, MPI, and OHC parameters, which had high correlations with intensity change (Fig. 2a), were selected less frequently. This is because, when POT was selected, the aforementioned four other predictors were excluded due to their high correlations with POT. DVMX had a particularly good performance in the short-term prediction (Fig. 2b), and this intensity-change predictor is commonly selected as the first and second predictors for two nonclustering experiments (Figs. 6a,b). LAND was also one of the main predictors selected in all four experiments. In addition, LON, OHC, RHLO, and T200 are important second predictors.

The types of selected predictors for each cluster reflect the unique characteristics of the clusters such as the tendency of mean TC intensity (or the changes), landfalls,

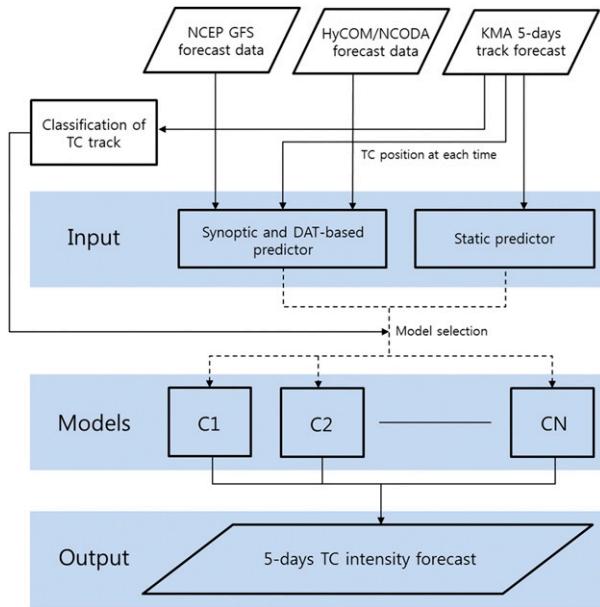


FIG. 9. Schematic of the real-time prediction procedure using CSTIPS-DAT.

upper-ocean thermal structure, and TC-induced mixing layer depth. For example, for C1 POTs using relatively shallow DAT (POT, POT10, POT20, POT30, and POT50 with depths ranging from 0 to 50 m) and LAND were chosen as major predictors (Fig. 7a). This is due to the fact that many cases of C1 experience a rapid weakening due to landfall and inadequate ocean thermal conditions with a thin mixed layer, leading to a shallow vertical mixing caused by TCs. In C2, a large percentage of the selected predictors for predictor 1 were the single POT-type predictor (95%). This arises as most TCs in C2 spend their lifetime in the tropical open ocean where oceanic thermal conditions are the most important factors. In this case, land effect and static

predictors are minor. C3, which is characterized by strong weakening tendencies with lead times (Table 2), showed that T200 was the most dominant in predictor 2. This is possible because many TCs in C3 travel to the north and are influenced by a large gradient of T200 along the track. Many TCs in C4 make landfall over the East Asian region (eastern coast of China and Taiwan), and LON and LAND were selected as major predictors. Finally, for C5, in which most TCs typically decayed over the cold open ocean, POTs with DATs of relatively shallow to medium depth (20–80 m) were important predictors. Because TCs in C5 pass through the open ocean where a large gradient of wind shear along the mean track of TC exists (Fig. 1a), vertical wind shear (SH500) was also selected. The performance of four models presented in Table 6 will be shown in the next section.

#### 4. Comparisons of model performance

##### a. Training period

In the present statistical models based on the perfect-prognosis approach, the relationship between the predictand and predictors is first established during the training period using the best-track, GFS analysis, and HyCOM–NCODA reanalysis data. This relationship is then applied to evaluating the performance of real-time intensity predictions using the predicted track and dynamical models prediction data. Figures 8 and 9 compare the prediction skills among CSTIPS-DAT, CSTIPS-SST, STIPS-DAT, and STIPS-SST for the training period. The forecast error differences between the CSTIPS-SST and CSTIPS-DAT at various leads are small for all clusters (Figs. 8a–e). The comparisons reveal that CSTIPS-DAT, which uses both clustering and DAT-based predictors, generally outperformed the other

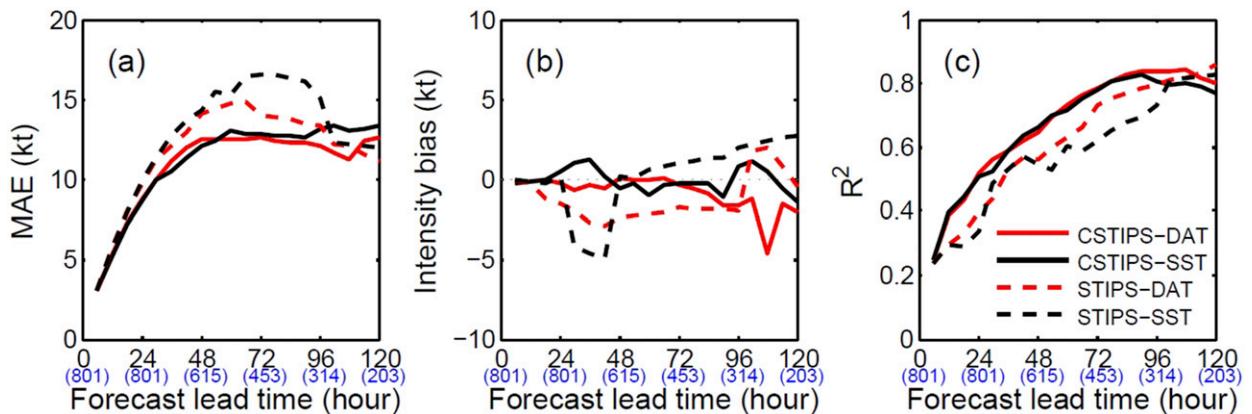


FIG. 10. Comparisons of (a) MAEs, (b) bias, and (c)  $R^2$  for the real-time TC intensity prediction. Here, the results from all experiments in Table 6 are compared. Sample sizes are indicated in parentheses.

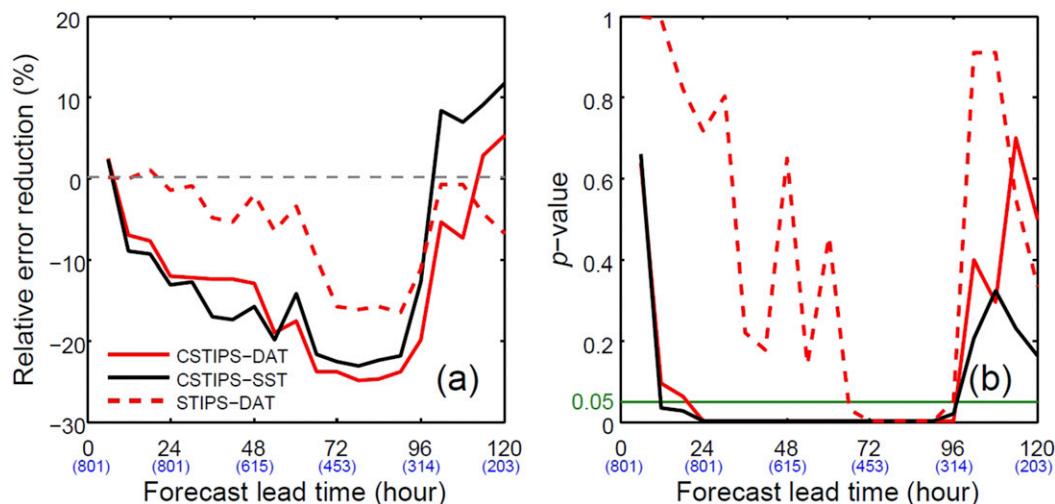


FIG. 11. The reduction rates (a) in MAEs of STIPS-DAT, CSTIPS-SST, and CSTIPS-DAT relative to the control experiment (STIPS-SST) for the test period and (b) their  $p$  values based on the two-sample  $t$  test. Sample sizes are indicated in parentheses.

three models. Applying the clustering approach between 24- and 96-h lead time (CSTIPS-SST vs STIPS-SST, and CSTIPS-DAT versus STIPS-DAT in Fig. 8f) resulted in the most significant improvement. The improvement is most remarkable at 72-h lead time. The mean absolute errors (MAEs) are reduced by 3.6 kt (25% improvement) and 1.9 kt (15%) for CSTIPS-DAT and CSTIPS-SST, respectively, compared to nonclustering settings. A similar tendency is also found in the results of  $R^2$  (not shown). The overall improvement in the CSTIP-DAT model is due to the enhanced improvement in some individual clusters. In particular, C5 showed the best prediction skills among the five clusters and contributes most to the model improvement (Fig. 8e). We found that the performance of each cluster is related to the magnitude of  $\sigma_c$  (Table 3). That is, smaller  $\sigma_c$  leads to more skillful predictions. This reaffirms that the clustering approach reduces the variability of the predictand, which ultimately results in the improvement of TC intensity prediction.

The impacts of DAT-based predictors on TC intensity prediction can be realized by comparing STIPS-DAT with STIPS-SST, and CSTIPS-DAT with CSTIPS-SST (Fig. 8f). The effects of DAT-based predictors are not as large as those of clustering, but they still further improved the performance. The largest improvement (14%) in MAE was found in the comparison with nonclustering experiments (STIPS-DAT and STIPS-SST in Fig. 8f) at lead times between 48 and 96 h.

#### b. Real-time predictions for 2013 and 2014

The model developed during the training period is tested with real-time predictions for two years: 2013 and 2014. Figure 9 illustrates a schematic of the real-time

prediction systems. The systems were implemented through the following procedure: obtain the 5-day track forecast information from KMA and atmospheric and oceanic fields at each forecast lead time from NCEP and NRL; calculate various static, synoptic, and DAT-based predictors and membership coefficients for each TC; determine the cluster to which the TC track belongs based on the highest membership coefficient; and predict the TC intensity using the regression model of the selected cluster.

The next question is, how well will the models developed in this study perform for an independent dataset? A comparison of MAE, bias, and  $R^2$  among the four models shows that the two clustered models outperformed the two nonclustered models at most lead times (Fig. 10) and that CSTIPS-DAT is slightly better than CSTIPS-SST. This feature is similar to the results of the model training. However, the overall mean MAEs for all real-time predictions (Fig. 10a) were a little larger (about 1.5 kt) than those for the training periods, likely because of the inaccuracies in real-time track predictions (Knaff et al. 2005) and the fact that the model parameters were fit for the training period, not during the independent period. To investigate the effect of the inaccurate track on the intensity errors, we compared the intensity errors for TCs having the upper 90th percentile track error with those for the remaining TCs. The comparison reveals that the errors of the former were larger than the latter by 1 kt at 24 h, 2.6 kt at 48 h, 1.6 kt at 72 h, 0.3 kt at 96 h, and 0.5 kt at 120 h. This explains that most differences in the intensity error between the training and test periods are due to the errors associated with track prediction and implies that the improvement of track prediction can lead to the reduction of intensity

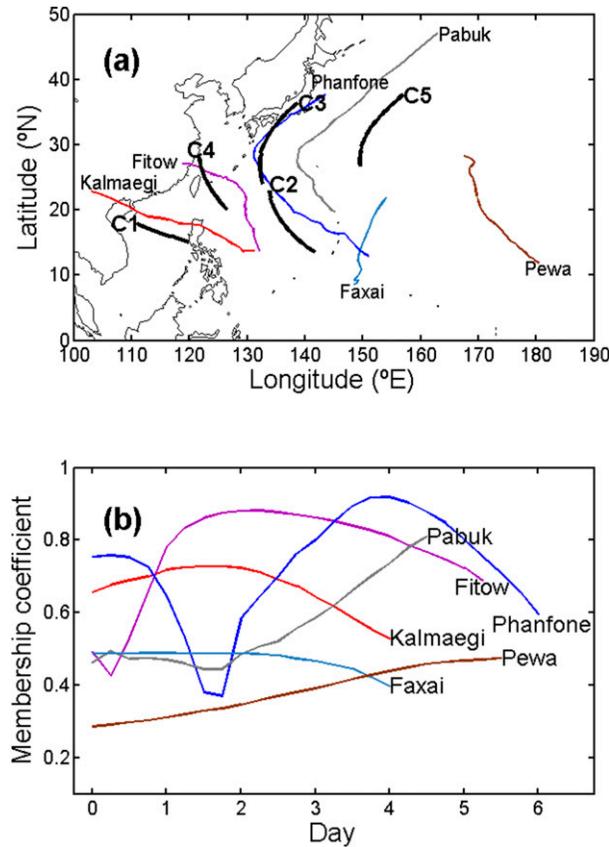


FIG. 12. (a) Tracks (colored lines) of Typhoon Pewa (2013), Fitow (2013), Pabuk (2013), Faxai (2014), Kalmaegi (2014), and Phanfone (2014) along with the mean track for five clusters (thick black lines). (b) Temporal evolution of the membership coefficient of the cluster assigned every 6 h for six TCs.

error. It also should be noted that the intensity errors in the independent experiments are also influenced by the inaccuracy in the environmental fields predicted from the numerical models (i.e., GFS and HYCOM), which is a limitation of the present perfect-prognosis-based model.

To quantify the effects of the use of clustering and DAT-based predictors, the reduction rates in MAE [i.e., relative error (RE)] were computed for STIPS-DAT, CSTIPS-SST, and CSTIPS-DAT relative to the control experiment (STIPS-SST):

$$RE = \frac{100(E - E_{STIPS-SST})}{E_{STIPS-SST}}, \quad (9)$$

where  $E$  is the MAE for each model, and  $E_{STIPS-SST}$  is the MAE of STIPS-SST. Results indicate that the use of clustering was the main contributor to the reduction in MAE (red and black solid lines in Fig. 11). For the clustered models, a large and significant reduction of

about 12%–25% between 24- and 96-h lead times was found, but not after 96-h lead time (Figs. 11a,b). This seems to be related to the lack of statistical confidence due to insufficient sample size after 96 h.

To examine the results of individual predictions in each cluster, we selected six typhoons for study: Pewa (2013), Fitow (2013), Pabuk (2013), Faxai (2014), Kalmaegi (2014), and Phanfone (2014). We analyzed their tracks (Fig. 12a), membership coefficients (Fig. 12b), and intensity prediction results using the CSTIPS-DAT model (Fig. 13). Here, the membership coefficient is an index that indicates the similarity of the individual TC track (colored lines in Fig. 12a) to the mean track of the membership cluster (black lines in Fig. 12a), that is, a larger membership coefficient implies a higher similarity of a particular TC track to its respective mean track. The cluster number assigned at each prediction time (6-h interval) is given above the x axis in Fig. 13. For example, Typhoon Phanfone started as C2 in the beginning, but its membership gradually changed to C3 with its northward movement (Fig. 12a). The comparison of predicted intensities (thin lines in Fig. 13) with the RSMC best-track data (thick lines) presents an overall good level of performance for Fitow, Pabuk, Kalmaegi, and Phanfone. For Typhoons Pewa and Faxai, however, the predicted intensities were largely overestimated. We found that the inaccurate predictions of these two typhoons are associated with their low membership coefficients (see Fig. 12b). In fact, Pewa and Faxai belonged to C5 and C2, respectively, but their tracks were far from the mean track of their respective clusters (Fig. 12a). The result suggests that an increase in the number of clusters might allow for more accurate predictions if enough samples were available. It also reconfirms the potential of cluster analysis of track patterns to improve statistical TC intensity prediction skills.

It is also interesting to compare the CSTIPS-DAT model with the latest operational dynamical models such as the Hurricane Weather Research and Forecasting Model (HWRF), the Regional Unified Model (UMR) of the KMA, the Japan Meteorological Agency’s Global Spectral Model (JGSM), and the U.S. GFS. These data are obtained from the National Typhoon Center (NTC) operating system of the KMA, and the evaluation was only made to the cases when the prediction results of all models were transferred to the NTC. As shown in Fig. 14a, for lead times up to 24 h, the CSTIPS-DAT model shows the smallest MAEs with prediction errors less than 8 kt relative to operational dynamical models for the independent test period. After 24-h leads, the CSTIPS-DAT model shows a remarkably good degree of skill and is comparable with GFS, which

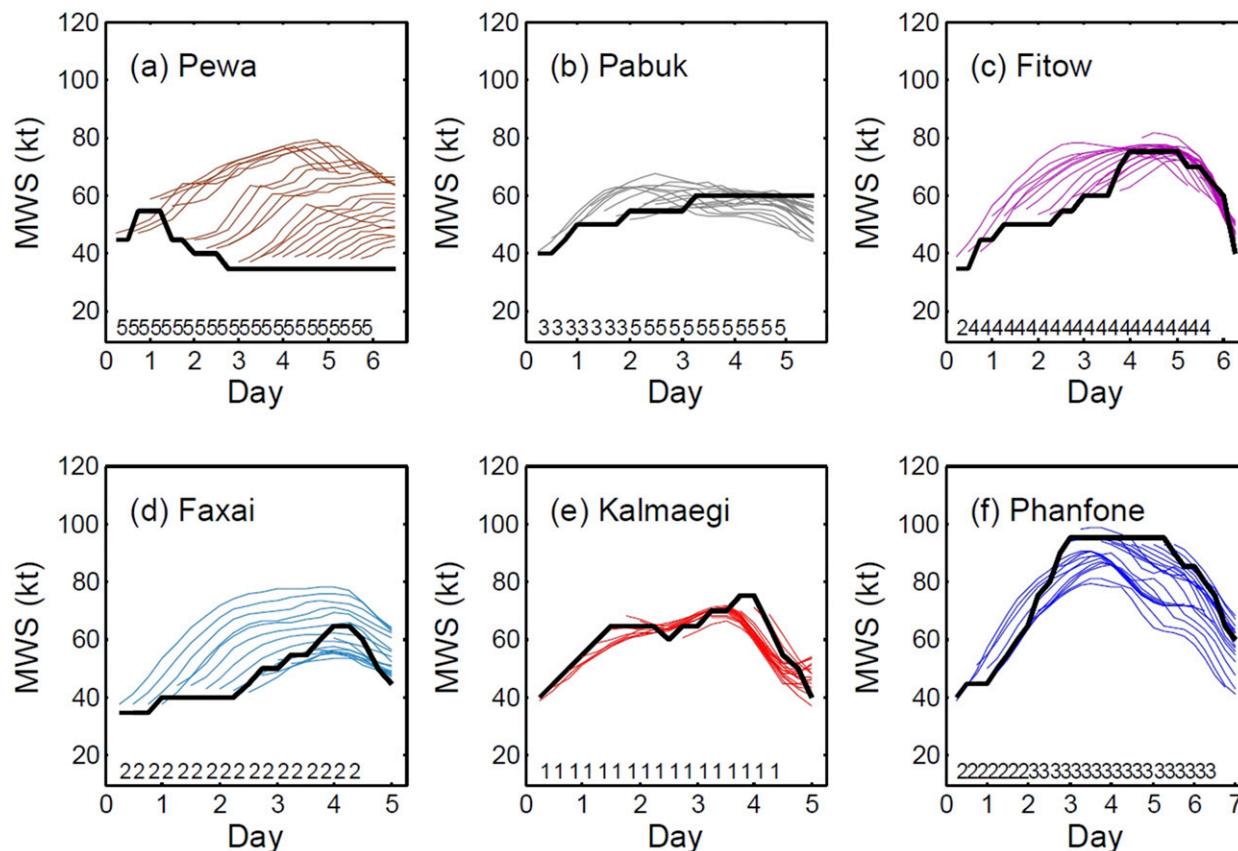


FIG. 13. Results of individual intensity predictions from CSTIPS-DAT for Typhoons (a) Pewa (2013), (b) Fitow (2013), (c) Pabuk (2013), (d) Faxai (2014), (e) Kalmaegi (2014), and (f) Phanfone (2014). The thick line denotes observations (RSMC best-track data), and the colored lines are individual CSTIPS-DAT predictions. The numbers above the  $x$  axis denote the assigned cluster number.

has the lowest MAEs among the other three operational models. The differences in MAEs between CSTIPS-DAT and GFS are not significant at most lead times except 0–18 h, but the forecast performance of CSTIPS-DAT is significantly better than that of HWRF, JGSM, and UMR at all leads (Figs. 14a,b).

## 5. Summary and conclusions

A statistical–dynamical typhoon intensity prediction model (CSTIPS-DAT) in the WNP was developed using track-pattern clustering and DAT-based potential predictors. In this model system, the tracks of TCs were classified into five clusters based on the fuzzy  $c$ -means clustering method, and the distinctive characteristics of each cluster were examined. The present model uses 6 static and 11 synoptic potential predictors as well as 3 new DAT-based ocean-coupling predictors that consider TC-induced vertical mixing effects.

To investigate the effects of using clustering and DAT-based predictors in TC intensity prediction, four

experiments were conducted for the training period of 2004–12 and real-time predictions in 2013 and 2014. Our results suggest that the use of clustering significantly reduced the MAEs relative to the nonclustered experiments in both training and real-time experiments, by about 25% (72-h lead time; CSTIPS-SST) and 12%–25% (24–96 h; clustered models), respectively. We found that the classified clusters had unique characteristics in terms of the tendency of intensity change, which reduced the variability  $\sigma_c$  of the predictand and ultimately led to the improvement of TC intensity prediction. The effect of DAT-based predictors was not as significant as that of clustering; however, performance was further improved. By using both clustering and DAT-based predictors, various sets of predictors could be skillfully selected depending on each cluster and its unique characteristics related to mean TC intensity, landfalls, upper-ocean thermal structure, and TC-induced mixing depth along the TC tracks. The test performance of the present model for 2013 and 2014 was more skillful than operational dynamical models up to 24-h lead time. After 24 h, the

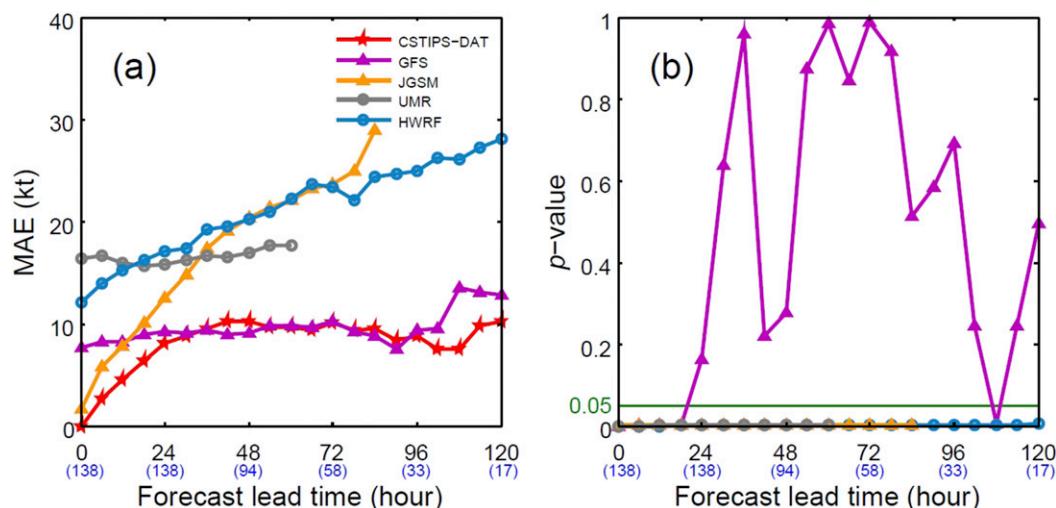


FIG. 14. Comparisons of (a) MAEs between the results of various operational dynamical models (GFS, JGSM, UMR, and HWRF) and the present model (CSTIPS-DAT) for the test period and (b) their  $p$  values for differences between individual models and CSTIPS-DAT based on the two-sample  $t$  test. Sample sizes are indicated in parentheses.

CSTIPS-DAT model still shows skill comparable with or superior to operational dynamical models.

The introduction of DAT-based predictors in STIPS has less impact on the improvement of intensity prediction than that of the clustering approach; however, it was found that DAT-based predictors, which incorporate the storm-induced mixing effect, were more correlated with intensity changes than SST-based predictors, particularly for strong TCs (more than 50 kt) (Fig. 4). It also revealed that stronger TCs tend to have a higher correlation with DAT-based thermodynamic predictors with deeper mixing depth. Among the final thermodynamic predictors selected in most clusters, the number of DAT-based predictors was much higher than the SST-based predictors (Fig. 7). These results suggest that DAT is a more physically and practically meaningful predictor related to TC intensity changes than SST, as reported in previous studies (Lin et al. 2013; Price 2009), and the inclusion of DAT-based predictors to the present model can contribute to further improvement in intensity predictions, although the effect of the DAT is not so significant when used with other predictors in the track-pattern clustering model.

The present model has limitations in the improvement it offers, particularly after 96-h lead time. This seems to be related both to the limitation of the perfect-prognosis approach and overfitting due to insufficient sample size. Also, the performance of the present model was not good for TCs with low membership coefficients, indicating the low similarity of the

individual TC tracks to the mean tracks of clusters. This implies that an increase in the number of clusters might be needed to allow for higher membership coefficients. If the sample size and the number of clusters increase in the future as more TC data are collected, it is expected that the problem of overfitting and low membership coefficients will be resolved.

*Acknowledgments.* This research was supported by the project titled “Research and Development for Numerical Weather Prediction and Earthquake Services in Korea Meteorological Administration” and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1A2B2005019). The authors wish to thank May Izumi of SOEST of the University of Hawaii for technical editing. The authors also appreciate John Knaff and two anonymous reviewers for their comments that contributed to an improved version of this manuscript. The historical western North Pacific TC best-track data are provided from the Regional Specialized Meteorological Center (RSMC), Tokyo Typhoon Center, which are available online (<http://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/besttrack.html>). The atmospheric data (GFS analysis and forecasts) were from NOAA/NCEP. The data are available online (<https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forecast-system-gfs>). The oceanic data (HyCOM/NCODA) were from the NRL, which are also available online (<https://hycom.org/dataserver/gofs-3pt0/reanalysis>).

## REFERENCES

- Aczel, A. D., 1989: *Complete Business Statistics*. Richard D. Irwin, 1056 pp.
- Bezdek, J. C., 1981: *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer Academic, 256 pp.
- Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, **4**, 55–88, [https://doi.org/10.1016/S1463-5003\(01\)00012-9](https://doi.org/10.1016/S1463-5003(01)00012-9).
- Bogner, P. B., G. M. Barnes, and J. L. Franklin, 2000: Conditional instability and shear for six hurricanes over the Atlantic Ocean. *Wea. Forecasting*, **15**, 192–207, [https://doi.org/10.1175/1520-0434\(2000\)015<0192:CIASFS>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0192:CIASFS>2.0.CO;2).
- Camargo, S. J., A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007a: Cluster analysis of typhoon tracks. Part I: General properties. *J. Climate*, **20**, 3635–3653, <https://doi.org/10.1175/JCLI4188.1>.
- , —, —, —, and —, 2007b: Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654–3676, <https://doi.org/10.1175/JCLI4203.1>.
- Chen, P., H. Yu, and J. C. L. Chan, 2011: A western North Pacific tropical cyclone intensity prediction scheme. *Acta Meteor. Sin.*, **25**, 611–624, <https://doi.org/10.1007/s13351-011-0506-9>.
- Chu, P.-S., and X. Zhao, 2011: Bayesian analysis for extreme climatic events: A review. *Atmos. Res.*, **102**, 243–262, <https://doi.org/10.1016/j.atmosres.2011.07.001>.
- , —, C.-H. Ho, H.-S. Kim, M.-M. Lu, and J.-H. Kim, 2010: Bayesian forecasting of seasonal typhoon activity: A track-pattern-oriented categorization approach. *J. Climate*, **23**, 6654–6668, <https://doi.org/10.1175/2010JCLI3710.1>.
- Corbosiero, K. L., and J. Molinari, 2002: The effects of vertical wind shear on the distribution of convection in tropical cyclones. *Mon. Wea. Rev.*, **130**, 2110–2123, [https://doi.org/10.1175/1520-0493\(2002\)130<2110:TEOVWS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2110:TEOVWS>2.0.CO;2).
- DeMaria, M., 1996: The effect of vertical shear on tropical cyclone intensity change. *J. Atmos. Sci.*, **53**, 2076–2087, [https://doi.org/10.1175/1520-0469\(1996\)053<2076:TEOVSO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<2076:TEOVSO>2.0.CO;2).
- , 2009: A simplified dynamical system for tropical cyclone intensity prediction. *Mon. Wea. Rev.*, **137**, 68–82, <https://doi.org/10.1175/2008MWR2513.1>.
- , and J. Kaplan, 1994a: A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting*, **9**, 209–220, [https://doi.org/10.1175/1520-0434\(1994\)009<0209:ASHIPS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2).
- , and —, 1994b: Sea surface temperature and the maximum intensity of Atlantic tropical cyclones. *J. Climate*, **7**, 1324–1334, [https://doi.org/10.1175/1520-0442\(1994\)007<1324:SSTATM>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1324:SSTATM>2.0.CO;2).
- , and —, 1999: An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337, [https://doi.org/10.1175/1520-0434\(1999\)014<0326:AUSHIP>2.0.CO;2](https://doi.org/10.1175/1520-0434(1999)014<0326:AUSHIP>2.0.CO;2).
- , M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543, <https://doi.org/10.1175/WAF862.1>.
- , J. A. Knaff, and J. Kaplan, 2006: On the decay of tropical cyclone winds crossing narrow landmasses. *J. Appl. Meteor.*, **45**, 491–499, <https://doi.org/10.1175/JAM2351.1>.
- , C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, <https://doi.org/10.1175/BAMS-D-12-00240.1>.
- Elsberry, R. L., G. G. Coltrane, and P. L. Krueger, 1974: Typhoon and tropical storm intensity forecasts for 24, 48 and 72 hours using statistical regression equations. Environmental Prediction Research Facility Tech. Note 15, 25 pp.
- Emanuel, K. A., 1988: Toward a general theory of hurricanes. *Amer. Sci.*, **76**, 370–379.
- Fitzpatrick, P. J., 1997: Understanding and forecasting tropical cyclone intensity change with the Typhoon Intensity Prediction Scheme (TIPS). *Wea. Forecasting*, **12**, 826–846, [https://doi.org/10.1175/1520-0434\(1997\)012<0826:UAFTCI>2.0.CO;2](https://doi.org/10.1175/1520-0434(1997)012<0826:UAFTCI>2.0.CO;2).
- Frank, W. M., and E. A. Ritchie, 2001: Effects of vertical wind shear on the intensity and structure of numerically simulated hurricanes. *Mon. Wea. Rev.*, **129**, 2249–2269, [https://doi.org/10.1175/1520-0493\(2001\)129<2249:EOVWSO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2249:EOVWSO>2.0.CO;2).
- Goni, G., and Coauthors, 2009: Applications of satellite-derived ocean measurements to tropical cyclone intensity forecasting. *Oceanography*, **22**, 190–197, <https://doi.org/10.5670/oceanogr.2009.78>.
- Jarvinen, B. R., and C. J. Neumann, 1979: Statistical forecasts of tropical cyclone intensity for North Atlantic basin. NOAA Tech. Memo. NWS NHC 10, 22 pp., <https://www.nhc.noaa.gov/pdf/NWS-NHC-1979-10.pdf>.
- Jones, S. C., 2000: The evolution of vortices in vertical shear. III: Baroclinic vortices. *Quart. J. Roy. Meteor. Soc.*, **126**, 3161–3185, <https://doi.org/10.1002/qj.49712657009>.
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.
- Kaplan, J., and M. DeMaria, 1995: A simple empirical model for predicting the decay of tropical cyclone winds after landfall. *J. Appl. Meteor.*, **34**, 2499–2512, [https://doi.org/10.1175/1520-0450\(1995\)034<2499:ASEMFP>2.0.CO;2](https://doi.org/10.1175/1520-0450(1995)034<2499:ASEMFP>2.0.CO;2).
- , and —, 2001: On the decay of tropical cyclone winds after landfall in the New England area. *J. Appl. Meteor.*, **40**, 280–286, [https://doi.org/10.1175/1520-0450\(2001\)040<0280:OTDOTC>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0280:OTDOTC>2.0.CO;2).
- Kim, H.-S., J.-H. Kim, C.-H. Ho, and P.-S. Chu, 2011: Pattern classification of typhoon tracks using the fuzzy *c*-means clustering method. *J. Climate*, **24**, 488–508, <https://doi.org/10.1175/2010JCLI3751.1>.
- , C.-H. Ho, J.-H. Kim, and P.-S. Chu, 2012: Track-pattern-based model for seasonal prediction of tropical cyclone activity in the western North Pacific. *J. Climate*, **25**, 4660–4678, <https://doi.org/10.1175/JCLI-D-11-00236.1>.
- Knaff, J. A., and C. R. Sampson, 2009: Southern Hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center, Part II: Statistical–dynamical forecasts. *Aust. Meteor. Oceanogr. J.*, **58**, 9–18.
- , M. DeMaria, C. R. Sampson, and J. M. Gross, 2003: Statistical 5-day tropical cyclone intensity forecasts derived from climatology and persistence. *Wea. Forecasting*, **18**, 80–92, [https://doi.org/10.1175/1520-0434\(2003\)018<0080:SDTCIF>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<0080:SDTCIF>2.0.CO;2).
- , S. A. Seseske, M. DeMaria, and J. L. Demuth, 2004: On the influences of vertical wind shear on symmetric tropical cyclone structure derived from AMSU. *Mon. Wea. Rev.*, **132**, 2503–2510, [https://doi.org/10.1175/1520-0493\(2004\)132<2503:OTIOVW>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<2503:OTIOVW>2.0.CO;2).
- , C. R. Sampson, and M. DeMaria, 2005: An operational Statistical Typhoon Intensity Prediction Scheme for the western North Pacific. *Wea. Forecasting*, **20**, 688–699, <https://doi.org/10.1175/WAF863.1>.
- Leipper, D., and D. Volgenau, 1972: Hurricane heat potential of the Gulf of Mexico. *J. Phys. Oceanogr.*, **2**, 218–224, [https://doi.org/10.1175/1520-0485\(1972\)002<0218:HHPOTG>2.0.CO;2](https://doi.org/10.1175/1520-0485(1972)002<0218:HHPOTG>2.0.CO;2).
- Lin, I.-I., C.-C. Wu, I.-F. Pun, and D.-S. Ko, 2008: Upper-ocean thermal structure and the western North Pacific category 5

- typhoons. Part I: Ocean features and the category 5 typhoons intensification. *Mon. Wea. Rev.*, **136**, 3288–3306, <https://doi.org/10.1175/2008MWR2277.1>.
- , C.-H. Chen, I.-F. Pun, W. T. Liu, and C.-C. Wu, 2009a: Warm ocean anomaly, air sea fluxes, and the rapid intensification of Tropical Cyclone Nargis (2008). *Geophys. Res. Lett.*, **36**, L03817, <https://doi.org/10.1029/2008GL035815>.
- , I.-F. Pun, and C.-C. Wu, 2009b: Upper-ocean thermal structure and the western North Pacific category 5 typhoons. Part II: Dependence on translation speed. *Mon. Wea. Rev.*, **137**, 3744–3757, <https://doi.org/10.1175/2009MWR2713.1>.
- , and Coauthors, 2013: An ocean coupling potential intensity index for tropical cyclones. *Geophys. Res. Lett.*, **40**, 1878–1882, <https://doi.org/10.1002/grl.50091>.
- Mainelli, M., M. DeMaria, L. K. Shay, and G. Goni, 2008: Application of oceanic heat content estimation to operational forecasting of recent Atlantic category 5 hurricanes. *Wea. Forecasting*, **23**, 3–16, <https://doi.org/10.1175/2007WAF2006111.1>.
- Mainelli-Huber, M., 2000: On the role of the upper ocean in tropical cyclone intensity change. M.S. thesis, Dept. of Meteorology and Physical Oceanography, University of Miami, Miami, FL, 73 pp.
- Moon, I.-J., and S. J. Kwon, 2012: Impact of upper-ocean thermal structure on the intensity of Korean peninsula landfall typhoons. *Prog. Oceanogr.*, **105**, 61–66, <https://doi.org/10.1016/j.pocean.2012.04.008>.
- Palmen, E., 1948: On the formation and structure of tropical hurricanes. *Geophysica*, **3**, 26–38.
- Peduzzi, P., B. Chatenoux, H. Dao, A. De Bono, C. Herold, J. Kossin, F. Mouton, and O. Nordbeck, 2012: Global trends in tropical cyclone risk. *Nat. Climate Change*, **2**, 289–294, <https://doi.org/10.1038/nclimate1410>.
- Price, J. F., 2009: Metrics of hurricane–ocean interaction: Vertically integrated or vertically averaged ocean temperature? *Ocean Sci.*, **5**, 351–368, <https://doi.org/10.5194/os-5-351-2009>.
- Rappaport, E. N., J. Jiing, C. W. Landsea, S. T. Murillo, and J. L. Franklin, 2012: The Joint Hurricane Test Bed: Its first decade of tropical cyclone research-to-operations activities reviewed. *Bull. Amer. Meteor. Soc.*, **93**, 371–380, <https://doi.org/10.1175/BAMS-D-11-00037.1>.
- Riehl, H. 1950: A model of hurricane formation. *J. Appl. Phys.*, **21**, 917–925.
- Sampson, C., E. Fukada, J. A. Knaff, B. Strahl, M. Brennan, and T. Marchok, 2017: Tropical cyclone gale wind radii estimates for the western North Pacific. *Wea. Forecasting*, **32**, 1029–1040, <https://doi.org/10.1175/WAF-D-16-0196.1>.
- Schumacher, A. S., M. DeMaria, and J. Knaff, 2013: Summary of the new statistical-dynamical intensity forecast models for the Indian Ocean and Southern Hemisphere and resulting performance. JTWC Project Final Rep., 11 pp.
- Shay, L. K., G. J. Goni, and P. G. Black, 2000: Effects of a warm oceanic feature on Hurricane Opal. *Mon. Wea. Rev.*, **128**, 1366–1383, [https://doi.org/10.1175/1520-0493\(2000\)128<1366:EOAWOF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<1366:EOAWOF>2.0.CO;2).
- Sun, Y., Z. Zhong, Y. Ha, Y. Wang, and X. D. Wang, 2013: The dynamic and thermodynamic perspectives of relative and absolute sea surface temperature on tropical cyclone intensity. *Acta Meteor. Sin.*, **27**, 40–49, <https://doi.org/10.1007/s13351-013-0105-z>.
- Wang, Y., and G. J. Holland, 1996: Tropical cyclone motion and evolution in vertical shear. *J. Atmos. Sci.*, **53**, 3313–3332, [https://doi.org/10.1175/1520-0469\(1996\)053<3313:TCMAEI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<3313:TCMAEI>2.0.CO;2).
- Wu, L., and Coauthors, 2012: Relationship of environmental relative humidity with North Atlantic tropical cyclone intensity and intensification rate. *Geophys. Res. Lett.*, **39**, L20809, <https://doi.org/10.1029/2012GL053546>.