Cluster Analysis of Tropical Cyclone Tracks over the Western North Pacific Using a Self-Organizing Map

HAN-KYOUNG KIM AND KYONG-HWAN SEO

Department of Atmospheric Sciences, Division of Earth Environmental System, Pusan National University, Busan, South Korea

(Manuscript received 28 May 2015, in final form 26 October 2015)

ABSTRACT

Tropical cyclone (TC) tracks over the western North Pacific (WNP) in 1979–2013 are classified by a self-organizing map technique. A false detection rate method identifies five optimal TC clusters. Physical mechanisms of the intraseasonal and interannual variations in the TC genesis frequency are investigated for each cluster. The five clusters are separated by genesis location, from the westernmost area (east of the Philippines, C1) to the easternmost (~150°E, C5) onset area over the WNP. The intraseasonal Madden–Julian oscillation (MJO) significantly affects the genesis frequency for all clusters except for C5. In particular, MJO phases 5 and 6 (1 and 2) provide significantly favorable (unfavorable) large-scale conditions for TC genesis. Two types of El Niño–Southern Oscillation influence the interannual variation of the genesis frequency for only C2 (generated over the western Philippine Sea and East China Sea) and C4 (formed near the eastern Philippine Sea). Enhanced eastern Pacific sea surface temperature (SST) anomalies lead to a ~40% decrease in the C2 TC frequency through a reversed Walker circulation with downward motion over the WNP. Conversely, increased central Pacific SST anomalies generate a cyclonic Rossby wave northwest of the forcing, inducing a significant increase (~50%) in the C4 TC frequency. The interannual variability for the C5 TCs is strongly controlled by the variation of the western Pacific subtropical high (WPSH). A positive WPSH variation reduces the C5 TC genesis frequency by 66%, while negative WPSH anomalies enhance the frequency by 50%. A prediction scheme using information from the first four 6-h TC locations demonstrates a skillful determination of TC clusters.

1. Introduction

The western North Pacific (WNP) is one of the most active basins in association with tropical cyclone (TC) genesis. About one-third of TCs developed over the global basins occur in the WNP with an annual average of ~19. This higher TC genesis frequency results from the confluence of the Asian monsoon westerlies and the easterly trade winds over the high sea surface temperature (SST) area (Li 2012). The TC genesis is also influenced by monsoon shear line, preexisting tropical cyclone, and Rossby wave trains (Ritchie and Holland 1999; Yoshida and Ishikawa 2013; Xu et al. 2013). The genesis frequency undergoes a great variation on the intraseasonal and interannual time scales because of large-scale atmospheric and oceanic variability in this basin, including El Niño–Southern Oscillation (ENSO), the intraseasonal oscillation (ISO), and the western Pacific subtropical high (WPSH) (Wang and Chan 2002; Li 2012; Li et al. 2012; Li and Zhou 2013a,b; Wang et al. 2013). As an example of the ENSO effect, the TC genesis frequency increases in the southeast (0°–17°N, 140°–180°E) and decreases in the northwest quadrant (17°–30°N, 120°–140°E), even though the total number of TCs formed in the WNP does not show a significant ENSO influence.

As for the ISO, two different types of ISOs impact the TC genesis frequency in the WNP: the 30–60-day Madden–Julian oscillation (MJO) and the 10–20-day quasi-biweekly oscillation (QBWO). Li and Zhou (2013a) showed that all MJO phases (Wheeler and Hendon 2004) significantly modulated the TC genesis frequency for 0°–30°N and 120°–150°E (i.e., the WNP1 region of the three in their study), whereas only a few QBWO phases significantly affected the TC genesis frequency. For example, when the QBWO-enhanced (suppressed) convection center was located east of the Philippine Sea, the TC genesis frequency increased (decreased) over the

Corresponding author address: Dr. Kyong-Hwan Seo, Department of Atmospheric Sciences, Pusan National University, Busan 609-735, South Korea.
E-mail: khseo@pusan.ac.kr

DOI: 10.1175/JCLI-D-15-0380.1

WNP1 region but decreased (increased) to its east [i.e., WNP2 (0°–30°N, 150°–180°E) in their study]. Wang et al. (2013) showed that the variability of the WPSH strongly modulated the TC genesis frequency and tracks in the WNP, with a stronger WPSH decreasing the number of TCs affecting East Asia (i.e., 20°–35°N, 105°–140°E), with a correlation coefficient of −0.76.

These studies imply that the particular large-scale atmospheric and oceanic variability influences the WNP TC genesis frequency in its subregions rather than in the whole WNP. The TC trajectory is closely associated with its TC genesis location since both components are determined by the same large-scale variability. For example, Li and Zhou (2013a,b) demonstrated that the TCs generated over the oceanic area east of the suppressed MJO convection have a higher (lower) possibility to undergo a recurving (straight) trajectory. Therefore, it is useful to subdivide the WNP TCs according to their trajectories, since a clustering of trajectories provides the genesis locations.

Previous studies attempted to classify the TC tracks over a specific basin using various methods based on heuristic and statistical clustering techniques (e.g., Harr and Elsberry 1995a,b; Lander 1996; Elsner and Liu 2003; Camargo et al. 2007b,c; Kim et al. 2011). Harr and Elsberry (1995a,b) heuristically classified the WNP TC tracks into three patterns based on the large-scale variability associated with the monsoon trough and the subtropical ridge: straight-moving, northward-recurving (recurving TCs formed north of 20°N), and southward-recurving (recurving TCs formed south of 20°N) trajectories. Lander (1996) separated the WNP TC tracks into straight moving, recurving, north oriented, and staying in the South China Sea.

In contrast to the heuristic method, statistics-based clustering methods showed more WNP TC track patterns. Camargo et al. (2007b,c) identified seven TC track patterns over the WNP using the "log-likelihood" and the "within-cluster error." Similarly, Kim et al. (2011) clustered the WNP TC tracks into seven patterns using four indices: the partition coefficient (Bezdek 1981) and the partition (Bensaid et al. 1996), separation (Xie and Beni 1991), and Dunn indices (Dunn 1973). However, the selection of an optimal cluster number in those studies is arbitrary since all of the above-mentioned indices are rather smooth and continuous functions of the varying number of clusters [e.g., Figs. 1 and 2 in Camargo et al. (2007b) and Fig. 2 in Kim et al. (2011)]. A more objective selection procedure is then necessary. For this, a method selecting an objective cluster number of TC tracks over the WNP based on a field significance test is proposed: the false discovery rate (FDR). Its detailed description is given in the next section.

The most well-known statistical clustering technique is the *K*-means clustering method. Elsner and Liu (2003) applied this method using the maximum intensity position and final location of TCs. However, this method suffers difficulties when the TC groups do not have clear differences. A recent study by Kim et al. (2011) adopted a fuzzy *c*-means clustering method for the extraction of WNP TC tracks. Since it is a nonlinear neural network, the above difficulty can be avoided. However, this method has the intrinsic shortcoming that an input vector can be classified as more than one cluster. In this study, we perform a self-organizing map (SOM) method to classify the WNP TC tracks. It is one of neural network models using a neighborhood function, being widely applicable even in the presence of uncertain group boundaries. SOM clustering is done automatically without any additional external input. A detailed description is given in the next section.

After the optimal clustering, we investigate how the most important physical factors, such as the large-scale atmospheric and oceanic variability associated with the MJO, ENSO, and WPSH, affect the genesis frequency for each cluster. Furthermore, we present a potentially useful application for near-real-time prediction of TC cluster classification using the current clustering results. The remainder of the paper is organized as follows: Section 2 provides a description of the clustering technique, optimum cluster number selection, and datasets used in this study. Section 3 shows the selection of the optimum cluster number using the FDR and the SOM clustering results. The relationships between the SOM cluster patterns and the large-scale intraseasonal and interannual variability are discussed in section 4. Section 5 includes a summary and a prediction of the TC cluster classification for 2014.

2. Data and methodology

a. Data and analysis period

This study utilizes the TC activity dataset from the Regional Specialized Meteorological Centers (RSMC) Tokyo–Typhoon Center, which contains information on TC name, date, type, latitude and longitude, central pressure, and maximum sustained wind speed (V_{max}) at the 3- or 6-h intervals. We only use 6-h interval data. Normally, TCs are divided in three categories depending on the wind speed (V_{max}): tropical depression (TD; $V_{max} < 17\,m\,s^{-1}$), tropical storm (TS; $17\,m\,s^{-1} \leq V_{max} < 34\,m\,s^{-1}$), and typhoon (TY; $V_{max} \geq 34\,m\,s^{-1}$). In this study, TCs are referred to as tropical storms and typhoons, so that the genesis and decaying locations of each TC are defined as the first

and last observed positions with intensity greater than TS, respectively (Kim et al. 2011). The dataset was from 1951 to the present, but we excluded presatellite years (before 1965) to avoid reliability problems (Chu 2002). Additionally, Zhou et al. (2009) suggested that the warming trend of the Indian Ocean and western Pacific was in favor of the westward extension of the WPSH since the late 1970s. Therefore, we restricted our analysis period to 1979–2013. The TC season is defined as June–October, during which about 75% of the TCs are formed in the WNP.

To analyze the large-scale environments associated with the clustered TC tracks, the daily mean horizontal wind and geopotential height data [1979–2013 from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR)] are used (Kalnay et al. 1996). We use monthly mean SST data drived from the Extended Reconstructed Sea Surface Temperature Dataset, version 3b (ERSST v3b; Smith et al. 2008). Furthermore, to determine the effect of the MJO, a daily mean outgoing longwave radiation (OLR) dataset (1979–2013) from the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting series of satellites (Liebmann and Smith 1996) is used.

b. Track interpolation

Most clustering methods, including K-means and fuzzy c-means techniques, require the same number of input vector lengths. The SOM clustering is also performed with an equal input vector length. As done in Kim et al. (2011), prior to input to the clustering, an interpolation is made using entire TC tracks to produce TC tracks with equal length. Here 20 equal segments are prepared for all TC tracks.

c. False discovery rate

In the statistics-based clustering analysis, an optimum cluster number needs to be specified prior to the analysis. For this, we introduced the FDR method, one of the field or global significance tests. This method objectively determines whether all cluster patterns are statistically distinguishable from each other.

The current input trajectories have 20 equal segments, so an SOM cluster pattern is the *mean* position of relevant trajectories with 21 latitude and 21 longitude points. To determine if SOM cluster pattern i is statistically distinguishable from SOM cluster pattern j, we first calculate the two-tailed p values for every point (42 points in our application; this total number is denoted as M) using a two-sample Student's t test for a difference of means. Here the null hypothesis meaning that the TC tracks are statistically indistinguishable for a pair of two clusters is tested at each point.

Then, instead of assessing M local t tests sequentially, we apply a global test (FDR) to evaluate the independency between clusters i and j. This test is carried out by the following equation. Let $p(m)$ denote p values in the ascending order, with m varying from 1 to M. The sorted p values are compared with $q(m/M)$, where q is a global significance test level (usually 0.05):

$$P_{\mathrm{FDR}} = \max_{1\le m\le M}\left[p(m): p(m) \le q\frac{m}{M}\right].$$

If just q is used in the above equation instead of $q(m/M)$, the test becomes a sequential local t test. By using $q(m/M)$, so m/M varies from 0 to 1, a more stringent test is made. So here P_{FDR} is a probability that the corresponding null hypothesis, which is actually true, is wrongly rejected. If none of the points satisfy the above square-bracketed condition for a pair of cluster tracks, then the two clusters are deemed indistinguishable. Even if all local tests based on the t test reject the null hypothesis (so seemingly evaluated as a distinguishable pair), the above condition may not be satisfied for even a point; the FDR test carries much stronger criteria. More detailed information is found in Benjamini and Hochberg (1995), Wilks (2006), and Johnson (2013).

d. Self-organizing map

The SOM (Kohonen 1990, 1997) is a neural network analysis tool characterized by an unsupervised training process with its networks learning to self-classify clusters from training data without external help. The structure of the SOM contains an arbitrary dimensional continuous input layer and one- or two-dimensional discrete output layers. The SOM analysis involves two processes: training and mapping. As an initial step of the training process, for an M candidate optimal cluster number, the M numbers of first cluster centroids (i.e., weight or standard vectors) are randomly generated. The next step of the training process involves matching (i.e., competition) by which a winner cluster centroid is determined. This is done by computing the Euclidean distance between an input vector (i.e., one TC track) and the M cluster centroids and selecting the cluster centroid that has the minimum Euclidean distance. The winner cluster centroid activates the close neighbor nodes within a neighborhood radius through a specific topological neighborhood function (i.e., a Gaussian distribution). Consequently, each individual node within the radius affects each other so that the SOM using the neighborhood function supports the nonlinearity of the classification. More detailed information on the neighborhood function can be found in Liu et al. (2006). The training process includes updating of the winner cluster centroid

using the neighborhood function. The updating equation in the j node is expressed as

$$W_j(t+1) = W_j(t) + \alpha(t)\mathrm{hc}_j(t)[X(t) - W_j(t)],$$

where W_j and X are the cluster centroid and input vector, respectively; α is a learning rate usually set as a monotonically decreasing function with an iteration number t; and hc_j is the neighborhood function. As stated above, the K-means clustering does not include the neighborhood function (i.e., the second term on the right of the above equation is zero), so the cluster centroid is calculated as a mere arithmetic mean of the input vectors in each cluster. After a large number of iterations, the final cluster centroid W_j and the Euclidean distance are determined. All input vectors are examined following this training procedure to obtain the final M optimal cluster centroids.

Next, mapping is performed such that each input vector is assigned to a corresponding cluster centroid based on the Euclidean distance. The SOM toolbox is freely available on the web (http://www.cis.hut.fi/somtoolbox). For more detailed descriptions, refer to Kohonen (1990, 1997), Liu et al. (2006), and Chu et al. (2012).

The advantage of this SOM clustering is that this method generates the final weight vectors in an orderly fashion such that closer weight vectors are located more nearby (which can be seen in Fig. 2), so resulting clusters are easy to interpret and diagnose. In addition, since this method uses only one feed-forward flow instead of using several feedback steps, the execution time is minimal so that, even in the case that a large size of new input data needs to be included for clustering, SOM performs fast classification to a satisfactory level.

3. Results

Optimum cluster number and clustering results

Figure 1 shows the number of statistically indistinguishable cluster pattern pairs as a function of K. In this figure, the abscissa and ordinate indicate the numbers of cluster patterns and indistinguishable pairs, respectively. For a K increasing from 2 to 5, all pairs of cluster patterns are statistically distinguishable (i.e., the number of pairs of indistinguishable pattern is zero). However, when K is equal to or larger than six, the number of indistinguishable cluster pattern pairs rises above zero. Consequently, the selected maximum number of statistically distinguishable cluster pattern pairs K is five. The selected five optimal cluster patterns seem to be stable enough for each cluster to be associated with different large-scale atmospheric and oceanic variability, which is examined in section 4.

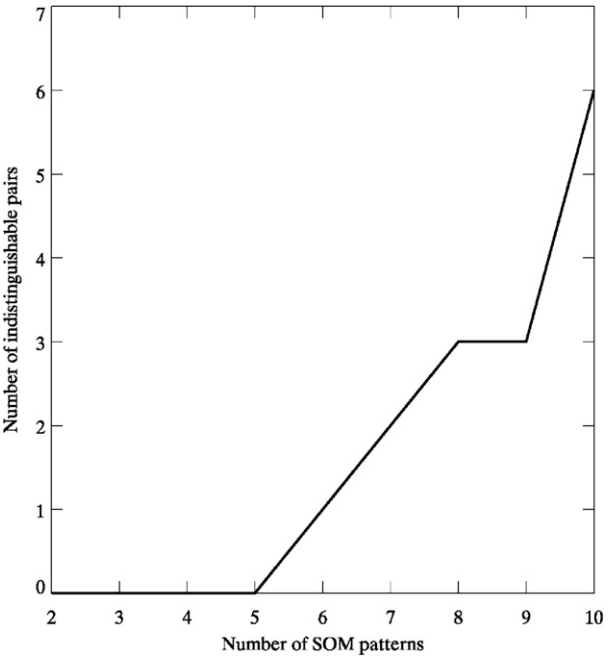

FIG. 1. Number of statistically indistinguishable TC cluster pattern pairs (at a 95% confidence level) as a function of the number of SOM cluster patterns K.

Figure 2 shows the five SOM cluster patterns for the WNP TC tracks during the TC season. The TCs in cluster 1 (C1) mainly develop around the South China Sea and northern Philippine Sea. The C1 TCs move northwestward without a curved trajectory (i.e., straight-moving track) and strike the southern China, northern Vietnam, Taiwan, and northern Philippine coasts. C1 accounts for 28.7% (189/660) of all TCs, which corresponds to the highest frequency among all clusters. On the contrary, this cluster exhibits the weakest intensity (in terms of minimum SLP of 974.4 hPa and maximum wind speed of 30.4 m s^{-1}), the shortest lifetime (4.1 days), and the shortest traveling distance (1462 km) (Table 1). These characteristics are closely related to the C1 genesis location, as C1 TCs are formed in the vicinity of the East Asian continents, resulting in a limited supply of thermal energy from the warm oceans, although genesis occurs at the lower latitude (Table 1).

The C2 TCs are formed over the Philippine Sea and East China Sea, hitting the eastern China, Taiwan, and northern Philippine coasts. Compared to C1, the C2 mean genesis position shifts east by ~10°, with TC tracks moving toward higher latitudes (Fig. 2) and, consequently, remaining for longer periods over the warm ocean. Therefore, C2 is characterized by a longer traveling distance (2166 km) and lifetime (5.5 days) and

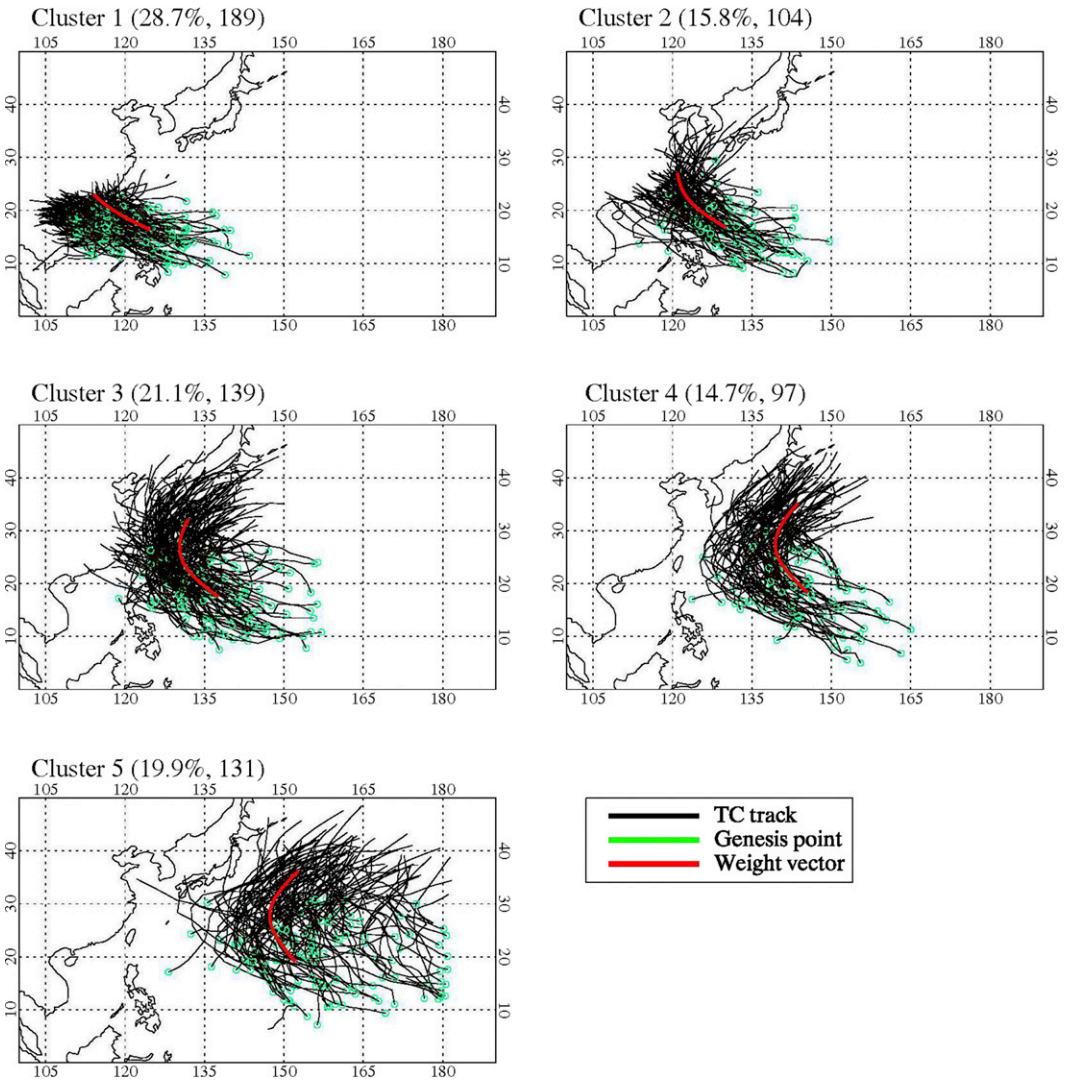

FIG. 2. Five TC cluster patterns from the SOM analysis. The black lines are the individual TC tracks, the green open circles TC genesis points, and the red line the weight vector for each cluster. The percentage and number of TCs for each cluster are shown in parentheses.

stronger intensity (955.2 hPa in SLP and wind speed of 38.8 m s^{-1}) (Table 1) than C1. C2 is the second less-observed track, with 15.8% (104/659) of the total TCs. The C3 TCs are formed around the eastern Philippine Sea and affect Taiwan, East China, Korea, and Japan with mixed straight-moving and recurving trajectories. About 20% of all TCs belong to this cluster, the second most frequent. This cluster shows the strongest intensity

TABLE 1. Mean traveling distance, lifetime, minimum sea level pressure, maximum wind speed, and genesis position for the five TC cluster patterns.

Cluster number	Traveling distance (km)	Lifetime (days)	Minimum SLP (hPa)	Maximum wind speed (m s^{-1})	Genesis position (lat, lon)
C1	1462	4.1	974.4	30.4	16.2, 120.9
C2	2166	5.5	955.2	38.8	16.4, 131.5
C3	2972	6.6	949.7	40.5	17.7, 137.9
C4	3043	6.2	953.2	38.9	18.0, 144.5
C5	2824	5.9	963.2	35.3	19.8, 157.2

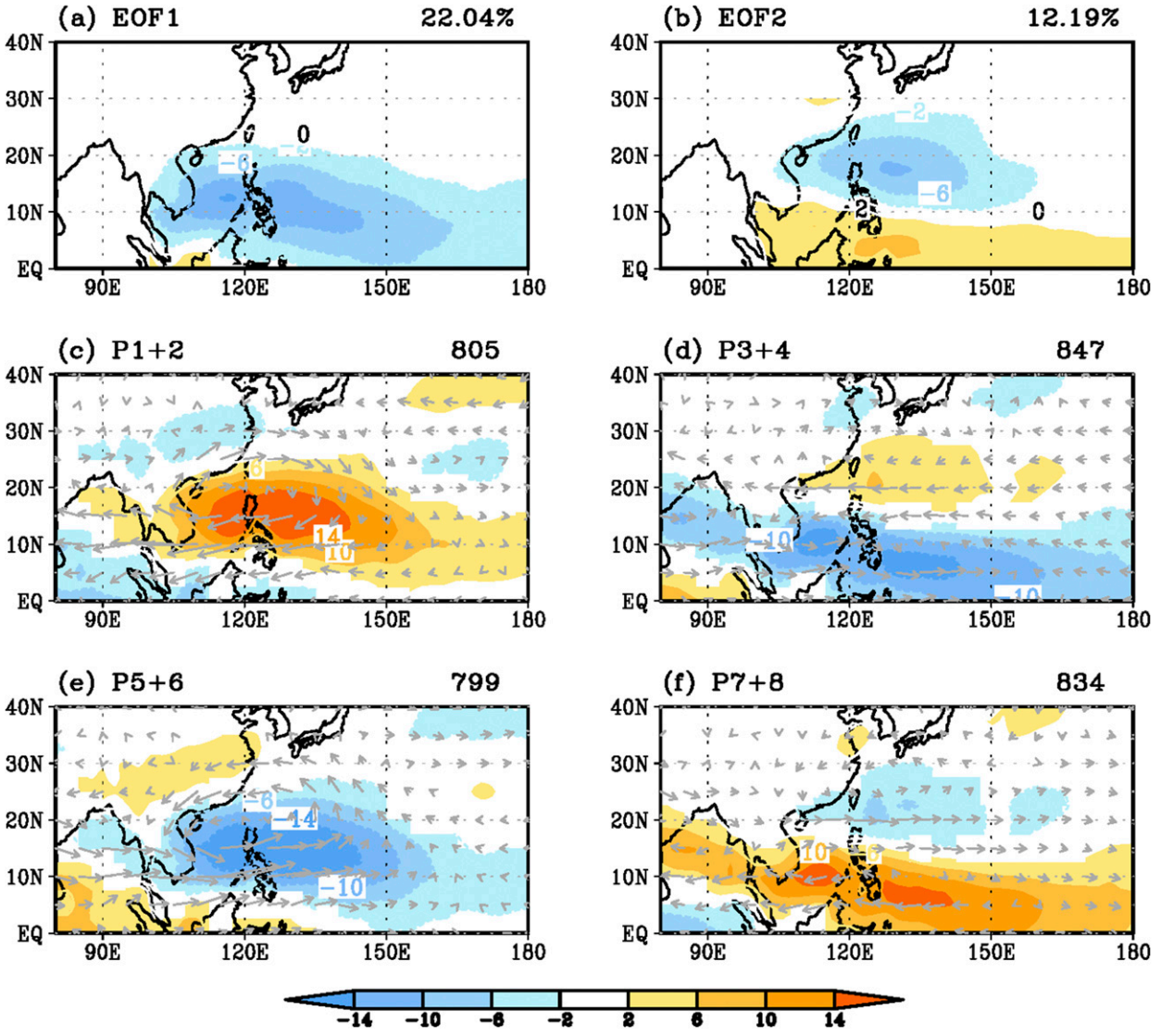

FIG. 3. (a) EOF1 and (b) EOF2 patterns for the 30–60-day filtered OLR anomalies during the TC season (June–October); (c)–(f) composites of the 30–60-day bandpass filtered 850-hPa wind anomalies (vector; $m\,s^{-1}$) and OLR anomalies (shaded; $W\,m^{-2}$) for different MJO phases. The shading and vectors are statistically significant at a 95% level. A Student's *t* test is used.

with minimum SLP of 949.7 hPa and maximum wind speed of 40.5 $m\,s^{-1}$, longest lifetime (6.6 days), and second longest traveling distance (2972 km) (Table 1). These characteristics are intimately related to the mean genesis location of this cluster (18.0°N, 137.9°E), far from the continent and at moderate latitudes, allowing for the acquisition of sufficient thermal energy from the warm oceans and ensuring their strengthening with time, reaching the Korean peninsula and Japan.

The C4 TCs develop around the eastern Philippine Sea (more east than C3), recurving toward Japan and nearby offshore. Only 14.7% (97/660) of all TCs belong to C4, the least frequent track among all clusters. This cluster shows the second longest lifetime (6.2 days) and second strongest intensity (minimum SLP of 953.2 hPa and maximum wind speed of 38.9 $m\,s^{-1}$) and the longest traveling distance (3043 km) (Table 1). The C5 contains TCs that form across the WNP with a mean genesis position in the northeastern edge (i.e., 19.8°N, 157.2°E; Table 1) of the main development region (MDR: 10°–25°N, 110°–160°E). It includes 131 TCs (i.e., ~20% of the total TCs) of straight-moving, recurving, and irregularly moving tracks, with a long traveling distance (2824 km) and lifetime (5.9 days) but weak intensity (963.2 hPa and 35.3 $m\,s^{-1}$). The TCs formed at the highest latitude among the five clusters have the largest deviation from the cluster-mean genesis position compared to the other clusters. Note that C4 and C5 TCs are

TABLE 2. DGR of the five TC cluster patterns for different MJO phases. A statistically significant increase (decrease) at the 90% and 95% level is in bold and marked by + and + + (− and − −), respectively. The statistical test is based on a Z statistic: $Z = (P - P_e)/\sqrt{P_e(1 - P_e)/N}$, where P_e and P represent the total DGR and DGR at a particular MJO phase, respectively, and N is the number of days for the particular MJO phase.

	Phase 1 + 2 (805)	Phase 3 + 4 (847)	Phase 5 + 6 (799)	Phase 7 + 8 (834)	Non-MJO (2070)
C1	**1.99 (−)**	3.54	**5.38 (+ +)**	**1.68 (−)**	4.15
C2	**0.62 (− −)**	2.95	**3.50 (+ +)**	2.04	1.40
C3	**1.86 (−)**	2.48	**4.38 (+ +)**	2.64	2.22
C4	**0.99 (−)**	1.89	**2.50 (+)**	1.56	1.93
C5	3.23	1.89	2.50	2.52	2.32

formed in the higher latitudes and easternmost longitudes among clusters (Table 1). The genesis location is more vulnerable to the lower SST and westward expansion of the subtropical anticyclone (i.e., WPSH; refer to Fig. 11). Hence, C4 or C5 may be slightly weaker than C3.

The frequency of each cluster shows a peculiar seasonal variation. For all TCs, the peak frequency appears in August and the next highest frequency occurs in September (not shown). June is the least frequent TC month. C1 develops in all TC months with a rather higher frequency for each month compared to the other clusters. C1 peak appears in September. Peaks in C2, C3, and C5 occur in August, while C4 peaks in September. C4 and C5 have a larger frequency toward the later TC period than June and July because of stronger WPSH activity during the earlier part of summer (not shown).

TCs are known to develop over the Asian summer monsoon trough region, where westerly monsoon and easterly trade winds meet (Lander 1996; Camargo et al. 2007c). All clusters are closely associated with a shear line, the zero line between zonal westerlies and easterlies. The genesis locations of C1, C2, C3, and C4 are very close to the shear line (not shown), with only C5 having a mean genesis location slightly east of the zero contour line in the easterly wind region. These results are remarkably consistent with Camargo et al. (2007c).

4. Relationship between TC cluster patterns and large-scale variability

a. Relationship between TC cluster patterns and boreal summertime MJO

The MJO is the most prominent physical mode of the intraseasonal variability. During the boreal summer, the main center of convective variability associated with the MJO shifts away from the equator to 10°–20°N, complicating the propagation patterns because of the off-equatorial thermal equator (Seo et al. 2007; Lee et al. 2012). Since the real-time multivariate (RMM) index (Wheeler and Hendon 2004) is mostly based on the boreal winter MJO, Li and Zhou (2013a,b) used a new boreal summer MJO index determined by an empirical orthogonal function (EOF) analysis for OLR data for 0°–30°N, 100°E–180°. Prior to the EOF analysis, the seasonal cycle is removed from the data, and a 30–60-day bandpass filtering is performed. Figure 3 shows the two leading EOF modes of the MJO [or boreal summer intraseasonal oscillation (BSISO)] and the composites of the filtered OLR and wind anomalies. The EOF1 mode (Fig. 3a) shows a zonally elongated negative OLR anomaly over the WNP. On the other hand, the EOF2 mode (Fig. 3b) exhibits a meridional dipole OLR anomaly structure. The maximum temporal correlation coefficient between PC1 and PC2 appears at a lag of ~10 days (not shown). The two EOF modes comprise the canonical boreal summertime MJO life cycle, with a northward-propagating component over the WNP. The results coincide well with previous summer MJO studies (Wheeler and Hendon 2004; Seo et al. 2007, 2009; Lee et al. 2012; Li et al. 2012; Li and Zhou 2013a,b). It is seen that the enhanced convection anomalies develop over the tropical Indian Ocean at phases 1 + 2 (P1 + 2) (Fig. 3c). Simultaneously, strong suppressed convective anomalies are prominent over the WNP MDR (10°–25°N, 110°–160°E) with an anticyclonic circulation anomaly northwest of the suppressed convection center (Fig. 3c). The enhanced convective center moves northeast toward the off-equatorial Maritime Continents at P3 + 4 (Fig. 3d), and the anomalies propagate to the north at P5 + 6 (Fig. 3e). For these phases, weak positive OLR anomalies are also formed over the Indian Ocean. For P7 + 8, the positive OLR anomalies gradually grow over the Philippine Sea and spread zonally (Fig. 3f).

We investigate the impact of the MJO on the TC cluster patterns and, in particular, on the TC genesis statistics. Table 2 shows the genesis frequency of TC cluster patterns for different MJO phases. Similar to previous studies (Hall et al. 2001; Li et al. 2012; Li and Zhou 2013a), a significance test has been performed by considering the fraction between a particular cluster daily genesis rate (DGR) and the total (i.e., all clusters) DGR (Table 2), where DGR is calculated when MJO

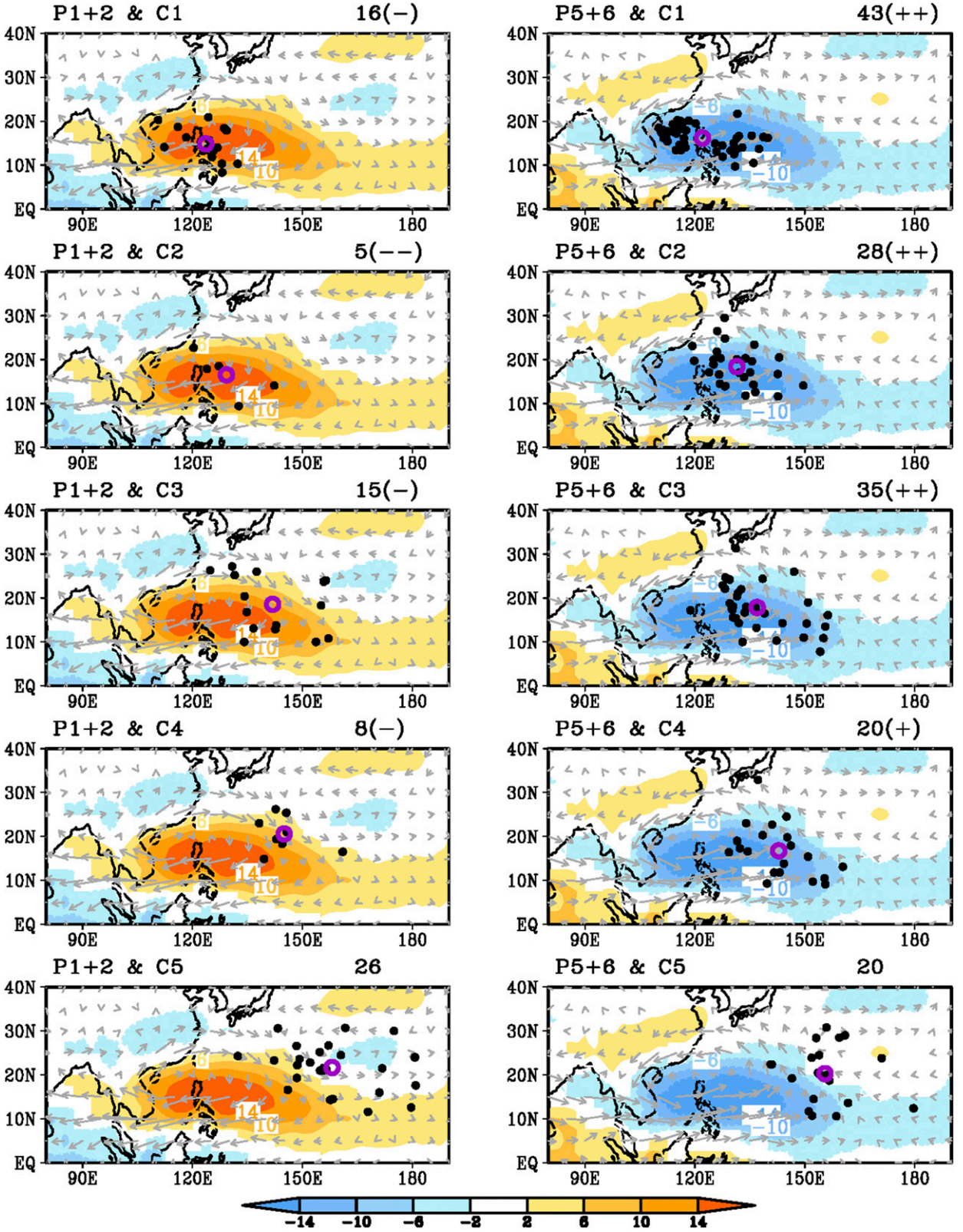

FIG. 4. Composites of the 30–60-day bandpass filtered 850-hPa wind anomalies (vector; m s^{-1}) and OLR anomalies (shaded; W m^{-2}) for phases P1 + 2 and P5 + 6 of the strong MJO events (absolute standard deviations > 1). The vectors and shading represent statistically significant variables at a 95% level. The black dots denote TC genesis points and the purple open circles their average. Frequency statistically enhanced (suppressed) at the 90% and 95% confidence level is marked by + and + + (− and − −), respectively.

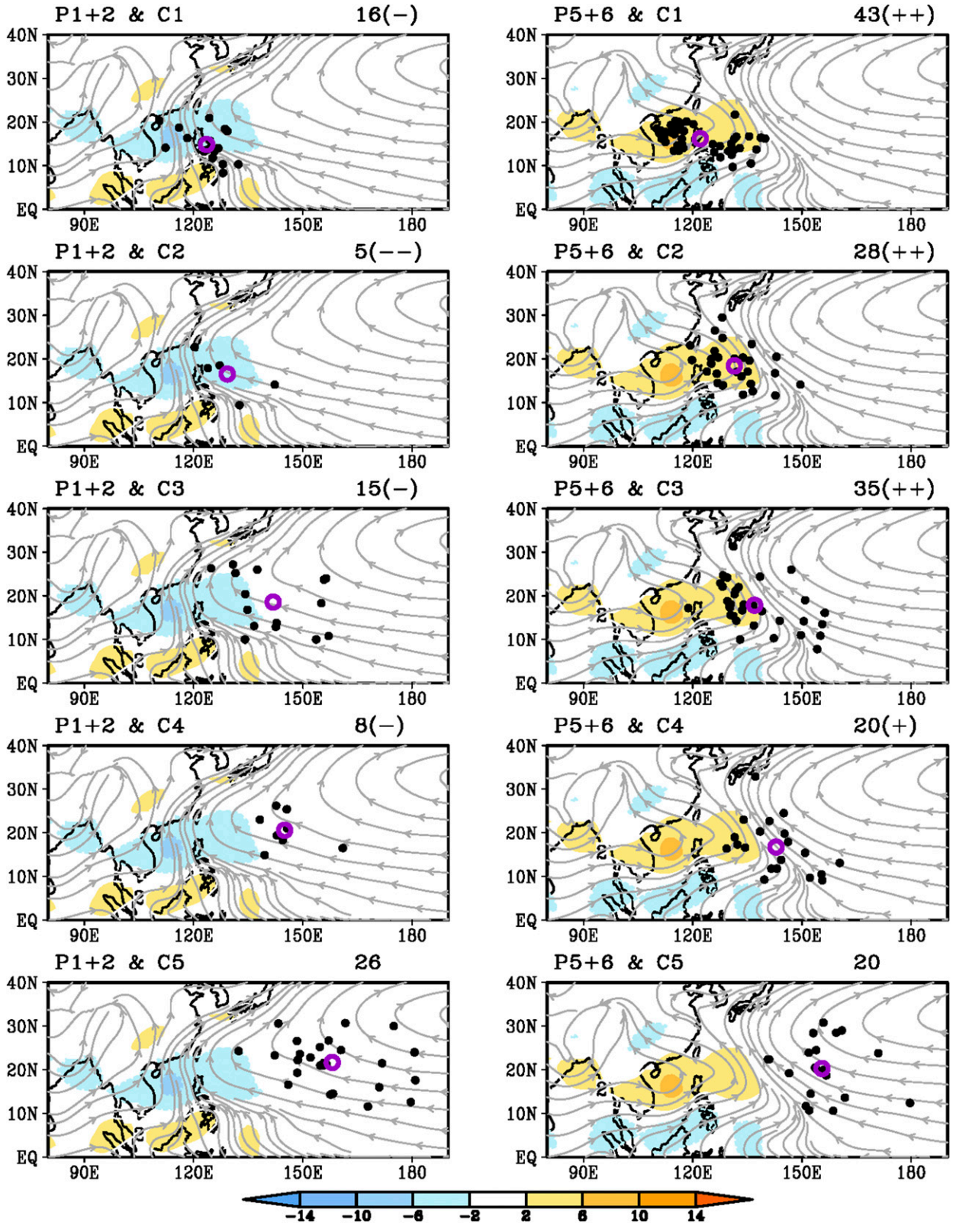

FIG. 5. Composites of the 850-hPa streamline (line; total field) and the 30–60-day bandpass filtered 850-hPa relative vorticity anomalies (shaded; 10^{-5} s^{-1}) for phases P1 + 2 and P5 + 6 of the strong MJO events (absolute standard deviations > 1). The streamlines and shading represent statistically significant variables at a 95% level. The black dots denote the TCs genesis points and purple open circles their average. Frequency statistically enhanced (suppressed) at the 90% and 95% confidence level is marked by + and + + (− and − −), respectively.

TABLE 3. Correlation coefficients between the TC frequency for each cluster and the climate indices. The bold denotes significance at a 95% confidence level. Detailed definitions of the climate indices are presented in the appendix.

	EP	CP	NAO	PDO	QBO	WPSH
C1	−0.19	−0.06	0.18	−0.22	0.14	−0.08
C2	**−0.34**	0.05	−0.02	0.22	0.04	−0.16
C3	0.16	0.17	0.05	0.08	−0.09	0.03
C4	0.05	**0.35**	−0.04	−0.08	0.14	−0.31
C5	−0.12	−0.19	−0.27	0.22	−0.28	**−0.71**

amplitude exceeds one standard deviation. For P1 + 2, the TC frequencies for C1, C2, C3, and C4 are very low, which is statistically significant at a 90% or 95% level. This results from the suppressed convection residing in the WNP MDR, providing unfavorable conditions for the development of TCs. Similarly, P5 + 6 (opposite to P1 + 2) shows an increased TC genesis frequency over the MDR for C1–C4. In general, the other phases (P3 + 4 and P7 + 8) do not show any significant changes in the genesis frequency since the maximum variance centers appear at a lower latitude, where the Coriolis force is considerably weaker. Similarly, the TC genesis frequency for C5 is not significantly affected by the large-scale MJO environment since the TC onset locations are situated east of the MJO convective anomalies. This property can be seen in the horizontal maps of the MJO convection for each phase and cluster (Fig. 4). P1 + 2 has a dry anomaly and corresponding anticyclonic circulation anomaly in the WNP MDR so that the TC onset activity is suppressed. Likewise, more TCs are found in the MDR at P5 + 6. Since C5 TCs are initiated clearly outside the convection anomalies, the MJO does not impact this cluster. Note that, as MJO amplitude becomes large, the frequency of C1, C2, and C3 also increases for phase 5 + 6 and decreases for phase 1 + 2 (not shown). The frequency change for C1 and C2 exhibits the largest sensitivity to MJO amplitude, as expected (not shown).

As supplementary information, we present the composite maps of the 850-hPa streamline (which is a total field) and the 30–60-day bandpass filtered 850-hPa relative vorticity anomalies for the different MJO phases (Fig. 5). At P5 + 6, a well-developed monsoon trough extends east along 10°–15°N, whereas P1 + 2 displays a considerably weaker monsoon trough. In addition, the WPSH for P5 + 6 is much weaker than for P1 + 2. Positive (negative) relative vorticity anomalies are also

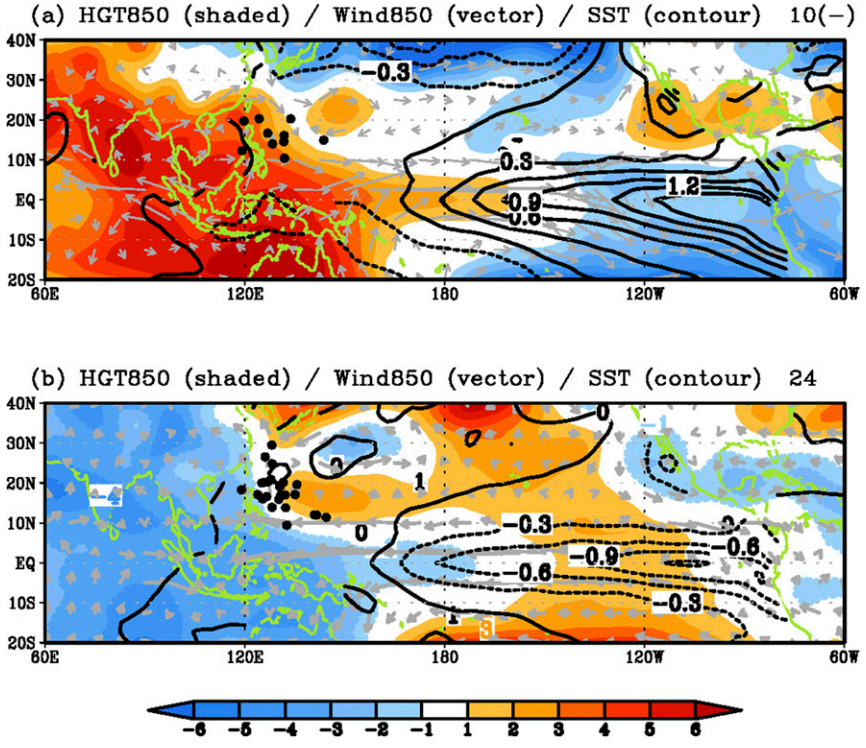

FIG. 6. Composites of the 850-hPa geopotential height (shaded; m), wind (vector; m s^{-1}), and SST (contour; °C) anomalies for (a) five warm EP years and (b) eight cold EP years. The black dots denote the TC genesis points in C2. The number of TCs is shown in the upper-right corner. Frequency statistically suppressed at the 90% confidence level is marked by −.

TABLE 4. TC mean frequency for each cluster by EP index in 1979–2013. Warm (cold) corresponds to a normalized EP index larger than 0.75 (lower than −0.75, respectively). Warm (cold) EP cases selected are 5 (8) yr. Frequency that is statistically enhanced (suppressed) at the 90% and 95% confidence level is in bold and marked by + and + + (− and − −), respectively.

	Warm	Cold	Normal	Total
C1	5.40	6.00	5.18	5.40 (189)
C2	**2.00 (−)**	3.00	3.18	2.97 (104)
C3	3.80	4.00	4.00	3.97 (139)
C4	3.20	2.25	2.86	2.77 (97)
C5	3.40	4.13	3.68	3.74 (131)

accompanied by the strong (weak) monsoon trough. These apparent circulation characteristics are consistent with the significant changes in the TC genesis frequency for C1, C2, C3, and C4. On the contrary, C5 TCs are located on the western edge and well inside the WPSH so that the effect of the MJO is weak.

b. *Relationship between TC cluster pattern and ENSO*

The relationships between the TC cluster patterns and the long-term (interannual) atmospheric and oceanic variability are investigated, and the correlations are examined for the same TC season. The potential factors influencing the TC genesis and tracks are listed in Table 3. The definition of these indices is shown in the appendix. Among these, only ENSO-related Pacific SST indices and the WPSH index are strongly related to the TC frequency for certain clusters. Hadley circulation strength, North Atlantic Oscillation (NAO), Pacific decadal oscillation (PDO), and quasi-biennial oscillation (QBO) do not significantly affect the WNP TC genesis frequency.

Previous studies have shown that the ENSO can affect the WNP TC activities (Camargo et al. 2007a,c; Wang and Chan 2002). However, most studies suggested that the ENSO influences the WNP TC formation location. For example, Wang and Chan (2002) showed that a strong El Niño (La Niña) induces TC development southeast (northwest) of the normal position through a Rossby wave cyclonic (anticyclonic) circulation response to the SST anomaly.

However, few studies showed its impact on the genesis *frequency*. Without clustering, the correlation between the total TC genesis frequency and the eastern Pacific (EP) SST is considerably low (−0.19). The total TC genesis frequency is also not significantly correlated with

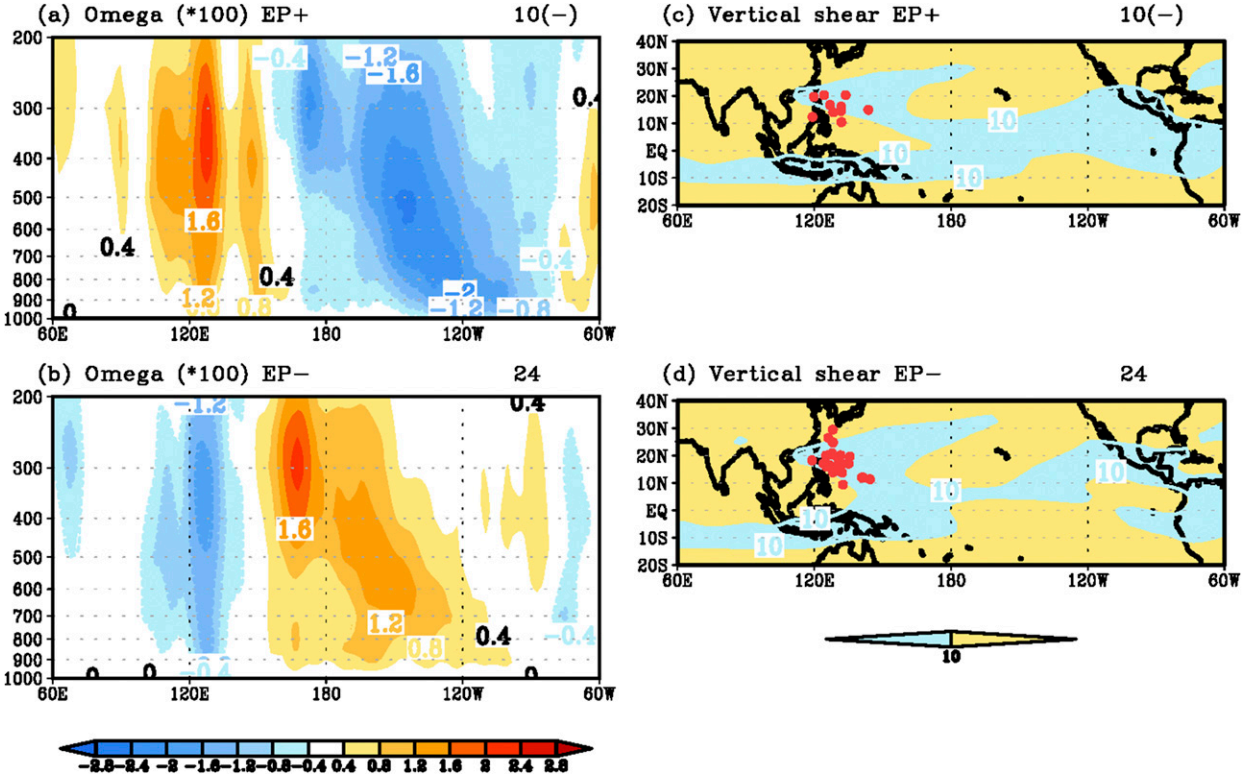

FIG. 7. Composites of the omega (intervals of 0.4 hPa s^{-1}) anomalies for (a) five warm EP years and (b) eight cold EP years. (c),(d) As in (a),(b), but for the total zonal vertical wind shear (intervals of 5 m s^{-1}). The red dots in (c) and (d) denote the C2 TC genesis points. The TC number is shown in the upper-right corner.

TABLE 5. TC frequency for each cluster by CP index in 1979–2013. Warm (cold) corresponds to a normalized EP index larger than 0.75 (lower than −0.75, respectively). Warm (cold) CP cases selected are 8 (9) yr. Frequency that is statistically enhanced (suppressed) at the 90% and 95% confidence level is in bold and marked by + and + + (− and − −), respectively.

	Warm	Cold	Normal	Total
C1	5.36	5.11	5.56	5.40 (189)
C2	2.38	3.00	3.22	2.97 (104)
C3	4.36	3.89	3.83	3.97 (139)
C4	**3.88 (+)**	2.22	2.56	2.77 (97)
C5	4.63	3.22	3.61	3.74 (131)

the central Pacific (CP) SST variation (r = 0.26). Here EP SST index is defined as the eastern Pacific SST anomaly index summing the Niño1+2 [10°S–0°, 90°–80°W] and Niño-3 [5°S–5°N, 150°–90°W] indices, and the CP SST index as the central Pacific SST anomaly averaged over [5°S–5°N, 180°E–140°W]. However, we show here that some clusters are significantly related to the ENSO type and phase. As shown in Table 3, the EP SST anomalies are negatively correlated with the C2 genesis frequency, whereas the CP SST anomalies have a significant positive relationship with the C4 frequency.

To examine in more detail the relationship with the SST anomalies over the central and eastern Pacific, we divide the EP index into positive and negative phases using a standard deviation threshold of 0.75 and −0.75, respectively (Fig. 6 and Table 4). The selected warm (cold) ENSO phase corresponds to 5 (8) yr. As seen in Table 4, only the warm EP-ENSO phase causes a statistically significant reduction in the TC genesis frequency for C2, attributed to the high-pressure anomalies extending from the Indian Ocean to the eastern Philippine Sea and main TC formation region (Fig. 6a), which arise from the downward branch of the Walker circulation forced by the positive SST anomalies over the eastern Pacific (Fig. 7a). By contrast, a cold EP-ENSO phase does not significantly increase the C2 TC genesis frequency (Fig. 6b), indicating a nonlinear relationship. This may be due to a weaker magnitude of the cold SST anomalies (EP index region mean of −0.90°C) compared to the positive SST anomalies (EP index region mean of 1.47°C) (thick lines in Fig. 6), resulting in a weaker response of the Walker circulation (Fig. 7b). A slight westward retreat of the lower geopotential height anomalies causes a favorable environment outside the C2 TC development area. Furthermore, to consider the

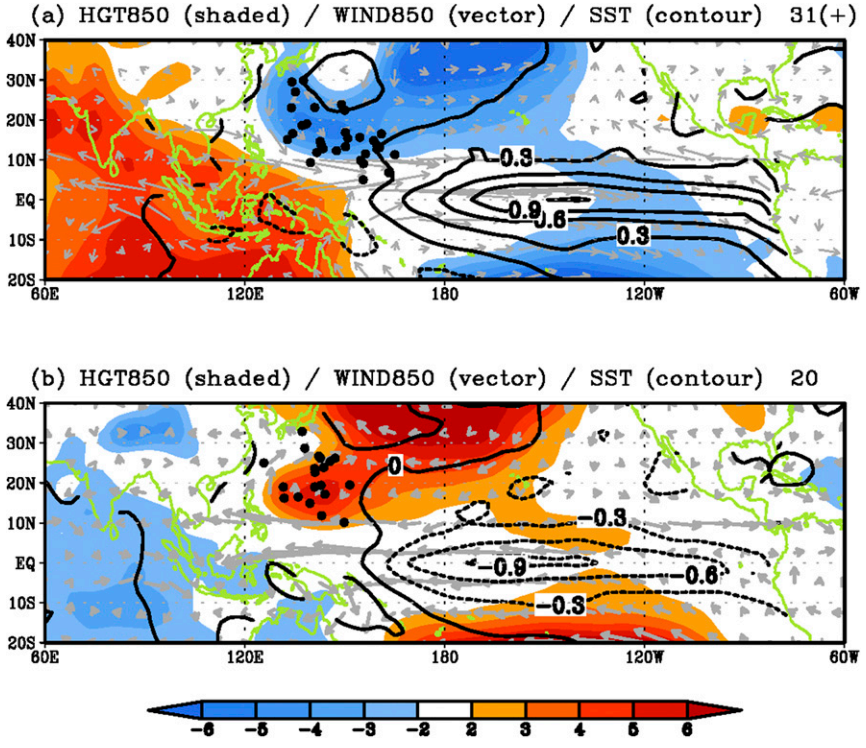

FIG. 8. Composites of the 850-hPa geopotential height (shaded; m), wind (vector; m s^{-1}), and SST (contour; °C) anomalies for (a) eight warm CP years and (b) nine cold CP years. The black dots denote the TC genesis points for C4. The number of TCs is shown in the upper-right corner. Frequency statistically suppressed at the 90% confidence level is marked by +.

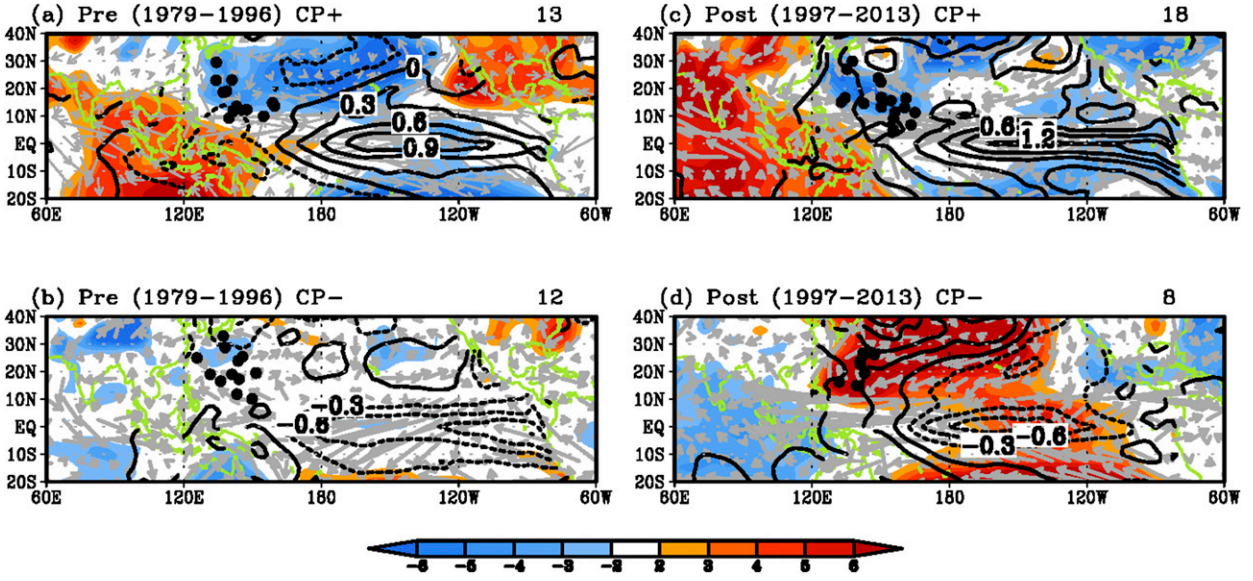

FIG. 9. Composites of the 850-hPa geopotential height (shaded; m), wind (vector; m s^{-1}), and SST (contour; °C) anomalies in 1979–96 for (a) four warm CP years and (b) four cold CP years. (c),(d) As in (a),(b), but for the four warm CP years and five cold CP years in 1997–2013. The black dots denote the TC genesis points for C4. The number of TCs is shown in the upper-right corner.

environmental wind fields the magnitude of vertical shear of the horizontal wind speed (200 − 850 hPa) is computed using the monthly data, which contain the genesis dates of TC. A smaller vertical wind shear than 10 m s^{-1} (blue area in Fig. 7d), which is favorable for the TC formation (e.g., Nolan and McGauley 2012), appears east of the negative geopotential height anomalies (Fig. 6b), leading to a near-normal TC onset frequency at this negative EP phase. The reduction in the vertical wind shear over this WNP region (120°–140°E, 10°–30°N) results from the strengthened Walker circulation (so easterly wind anomaly at the lower level and westerly wind anomaly at the upper level are enhanced) induced by the cold SST anomalies over the eastern Pacific (not shown), which, under the nominal lower-level westerly and upper-level easterly winds, reduces the vertical wind shear. In contrast, the large vertical wind shear at the positive EP phase (Fig. 7c) is collocated with the positive geopotential height anomaly, resulting in a statistically significant reduction in the onset frequency. The environmental wind shear acts to constrain the location of the TC genesis so that no significant changes appear for the negative EP SST phase.

The CP SST variation also has a significant relationship with C4 only for the positive CP ENSO phase (Table 5 and Fig. 8), which is in good agreement with

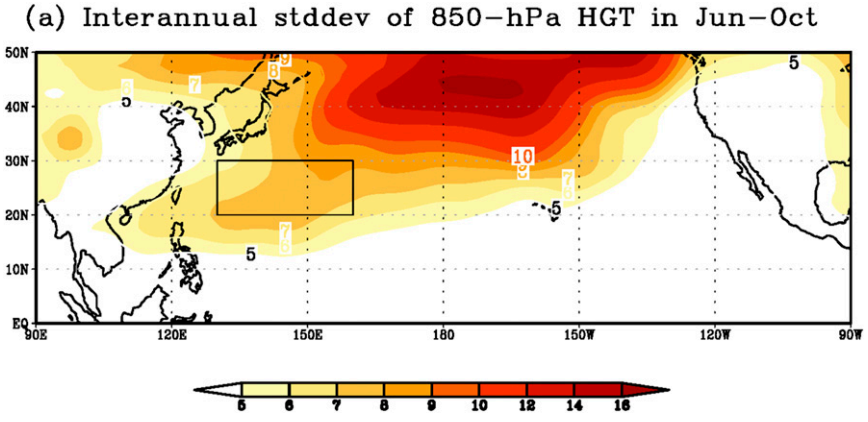

FIG. 10. Standard deviation of the interannual 850-hPa geopotential height variation in June–October. The box represents the reference area (20°–30°N, 130°–160°E) with a local peak interannual variability in the WNP.

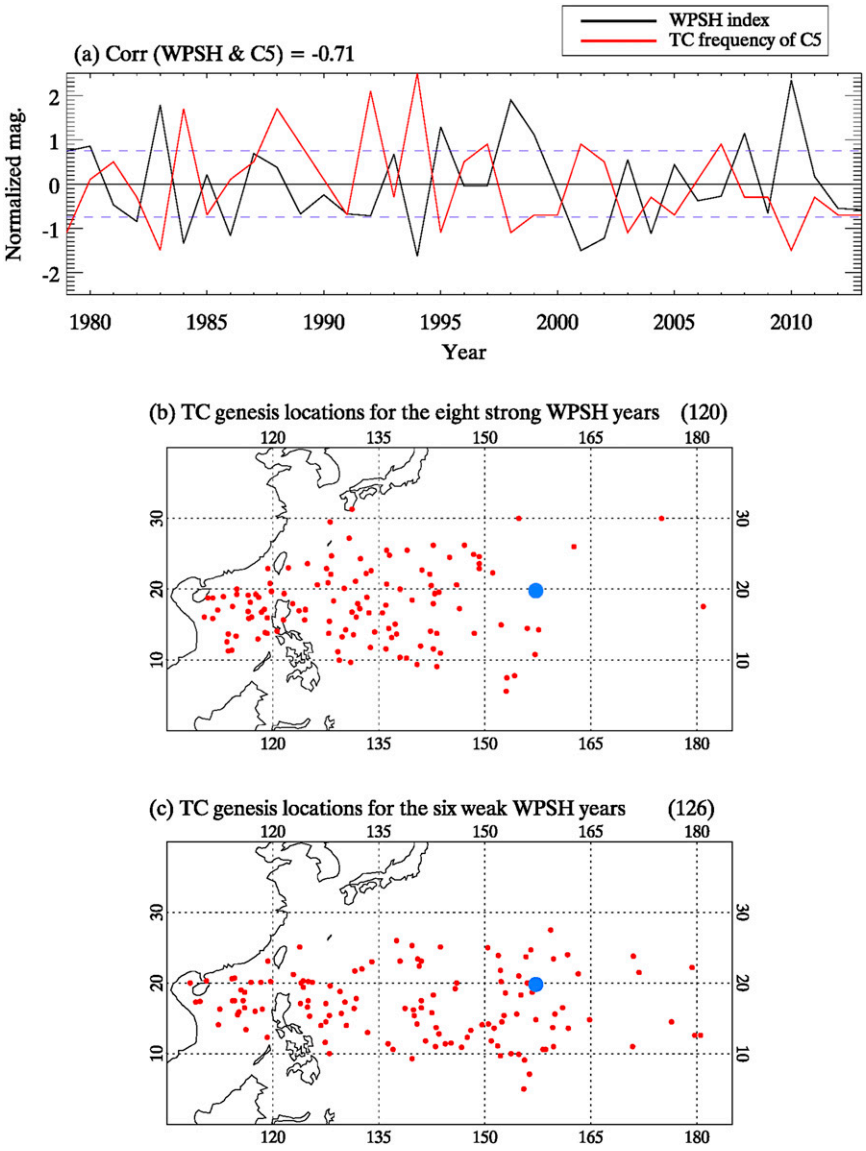

FIG. 11. (a) WPSH index (black line) and C5 TC frequency (red line). The red dots denote the genesis location of entire TCs for (b) eight strong WPSH years and (c) six weak WPSH years. The blue closed circle represents the average genesis location of C5 TCs (19.8°N, 157.2°E). The number of TCs is shown in parentheses.

Zhang et al. (2012). The enhanced C4 genesis frequency for the warm CP phase results from the descending Rossby wave response to the warm anomaly over the CP (Fig. 8a). The actual Rossby response appears off the equatorial regions in both hemispheres (Fig. 8), consistent with a Gill-type response (Gill 1980). The low-level cyclonic circulation develops northwest of the SST forcing (i.e., over the WNP). The cold CP phase seems to induce the anticyclonic circulation over the WNP (Fig. 8b), leading to a reduction in the TC genesis frequency, although not statistically significant (Table 5). This asymmetric property can be ascribed to an interdecadal variation of the cold CP phase; that is, the SST anomalies composited for four cold CP events during an earlier period (i.e., 1980s) were not strong (Fig. 9b), resulting in a very weak circulation response over the WNP (Fig. 9b) and little effect on the TC frequency change. More recently (e.g., 1997 onward; Fig. 9d), however, a strong negative SST forcing over the CP (Fig. 9d) appeared, resulting in a strong anticyclonic circulation response, which significantly reduced the TC frequency with eight TCs in five cold CP years (Fig. 9d). Oppositely, the positive CP phase behavior was more or less similar for the two

TABLE 6. TC frequency for each cluster by WPSH index in 1979–2013. Strong (weak) corresponds to a normalized CP index larger than 0.75 (lower than −0.75). Strong (weak) WPSH events selected are 8 (6) yr. Frequency that is statistically enhanced (suppressed) at the 90% and 95% confidence level is in bold and marked by + and + + (− and − −), respectively.

	Strong	Weak	Normal	Total
C1	5.63	5.33	5.33	5.40 (189)
C2	2.25	3.17	3.19	2.97 (104)
C3	3.63	3.17	4.33	3.97 (139)
C4	2.13	3.33	2.86	2.77 (97)
C5	**1.38 (− −)**	**6.00 (+)**	4.00	3.74 (131)

periods (Figs. 9a,c). The interdecadal variation of the CP SST anomalies complicated the analysis, but more recent data (Figs. 9c,d) showed a more linear statistical behavior between the warm and cold CP ENSO phases. All four CP SST cases show a TC development over small environmental wind shear regions (not shown).

To summarize, the enhanced EP SST anomalies lead to a significant decrease in the C2 TC frequency, whereas the increased CP SST anomalies significantly increase the C4 TC frequency. Note that the C4 frequency change is almost consistent with Camargo et al. (2007c) and Kim et al. (2011) for the CP El Niño phase. However, while these two previous studies demonstrated some association with the La Niña phase, our result does not show any statistically significant relationships with either the EP or CP La Niña phase. Furthermore, C2 in the current study shows a decrease in TC frequency during EP El Niño years, which has not been analyzed in their studies.

c. *Relationship between TC cluster patterns and WPSH*

It is well known that the WPSH impacts the WNP TC frequency and tracks. The interannual variation of the WPSH is due mainly to both the air–sea interaction over the western Pacific and northern Indian Ocean and the

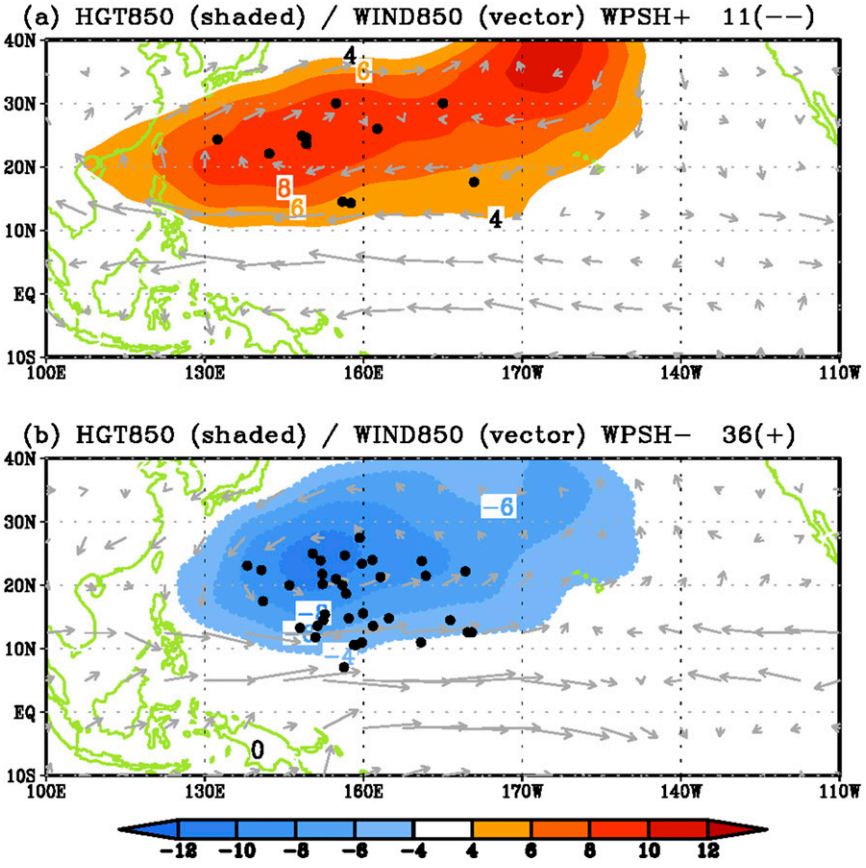

FIG. 12. Composites of the 850-hPa geopotential height (shaded; m) and wind vector anomalies (arrow; m s^{-1}) for (a) eight strong and (b) six weak WPSH years. The black dots denote the TC genesis points for C5. The number of TCs is shown in the upper-right corner. Frequency statistically enhanced (suppressed) at the 90% and 95% confidence level is marked by + and + + (− and − −), respectively.

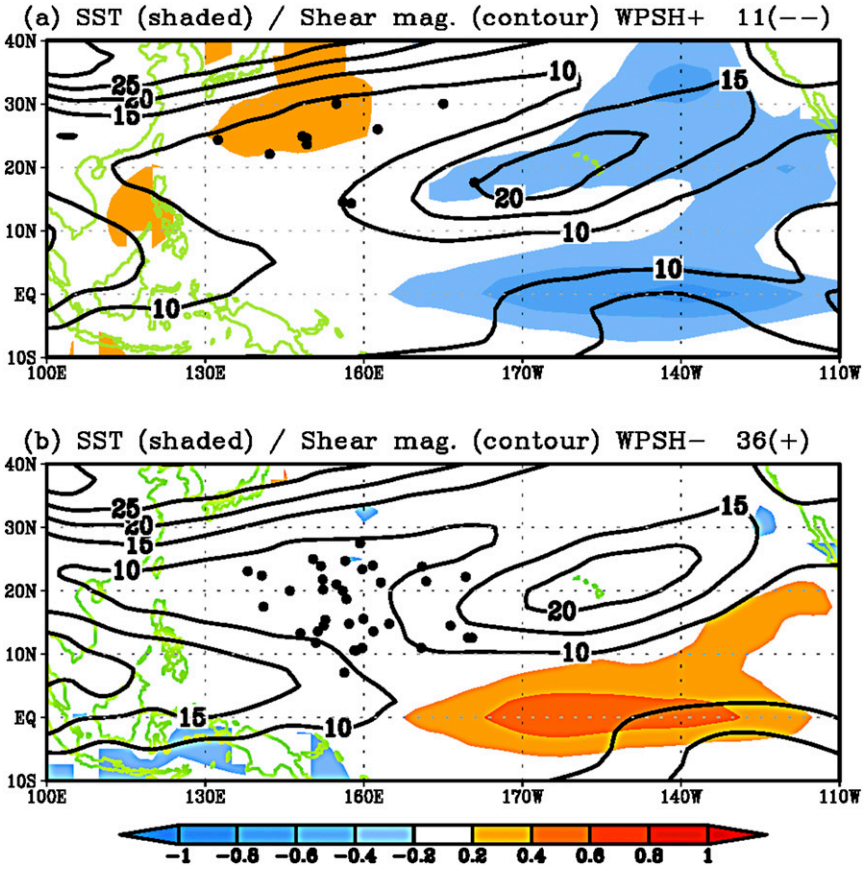

FIG. 13. Composites of the SST (shaded; °C) and total zonal vertical wind shear (contour; m s^{-1}) for (a) eight strong and (b) six weak WPSH years. The black dots denote the TC genesis points for C5. The number of TCs is shown in the upper-right corner.

descending Rossby wave response to the cold equatorial central Pacific SST anomalies (Wang et al. 2013). Wang et al. (2013) defined a WPSH index to establish a relationship between the WNP TC days and the total number of TCs impacting the East Asian coasts for June–August. However, we define the TC season from June through October. The WPSH index is then redefined by calculating the interannual standard deviation of the 850-hPa geopotential height (Fig. 10). An area defined by [20°–30°N, 130°–160°E] is selected to represent the interannual variation of the WPSH in relation to the WNP TC MDR. This reference area is slightly more northeast compared to Wang et al. (2013). Using the normalized WPSH index time series (black line in Fig. 11a) and the threshold of ±0.75, we have selected eight strong and six weak WPSH years (Table 6). As in Table 3, which demonstrates that the WPSH variation is only significantly correlated with C5 ($r = -0.71$), both phases of the WPSH variation are significantly negatively related with the C5 genesis frequency (Table 6). The interannual variations of the C5 TC frequency and the WPSH index are shown in Fig. 11a. Figures 11b and 11c show the genesis location of TCs for strong and weak WPSH years. For strong WPSH years, 15 TCs are formed per year, mostly over the western WNP (~150°E) and along the western boundary of the WPSH. In contrast, for weak years, the TC number is significantly higher (21 TCs per year) with the genesis points spread across the WNP and more TCs developing over the eastern WNP than for strong WPSH years.

Figure 12 shows the composites of the 850-hPa geopotential height and the wind vector for C5 in strong and weak WPSH years. A broad positive geopotential height and anticyclonic circulation are dominant over the WNP in strong WPSH years with only a few TCs. If one assumes that, on average, four TCs develop for this cluster, the frequency is reduced by 66% for this phase (Table 6). On the contrary, weak WPSH years show the widespread negative geopotential height and cyclonic circulation anomalies over the WNP region, creating favorable conditions for storm genesis. The negative WPSH phase

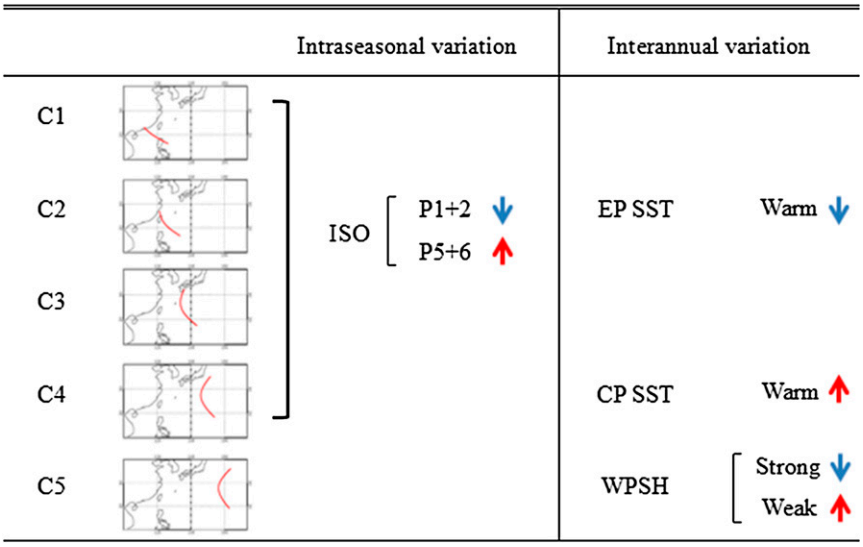

FIG. 14. Summary of the relationship between the five TC clusters and the intraseasonal and interannual atmospheric and oceanic variability. The red upward- (blue downward-) pointing arrow represents an increase (decrease) in the TC frequency.

produces six TCs per year, which corresponds to a 50% increase (Table 6).

It is interesting to note that, regardless of positive or negative WPSH years, TCs tend to develop over weak total vertical wind shear regions ($<10\,m\,s^{-1}$) in the WNP (Fig. 13). Therefore, in strong WPSH years, the weak shear region and TCs are shifted northward (Fig. 13a). For weak WPSH years, TCs develop slightly south of the negative WPSH anomaly center (Fig. 13b).

5. Summary and discussion

The moving tracks of the TCs that developed over the WNP area during 1979–2013 are classified through a self-organizing map (SOM) technique. For this, an objective determination of the optimal cluster number is performed with an FDR method, a field significance test. Five different TC clusters are identified. Previous studies demonstrated a significant influence of intraseasonal and interannual variability (e.g., the MJO, ENSO, and WPSH) on the genesis location and frequency of TCs over the WNP (Wang and Chan 2002; Li et al. 2012; Li and Zhou 2013a,b; Wang et al. 2013). In this study, we investigated these connections for each cluster. Figure 14 summarizes the relationships between the five TC clusters and the intraseasonal and interannual atmospheric and oceanic variability.

The intraseasonal MJO significantly affects the genesis frequency for all clusters except for C5 (with the easternmost genesis location). The C1–C4 TCs fall well inside of the MJO convective and circulation anomalies. For example, MJO P5 + 6 (P1 + 2), with enhanced (suppressed) convection over the WNP region, provides significantly favorable (unfavorable) large-scale conditions for TC genesis in C1, C2, C3, and C4.

Even though statistically insignificant relationships between the interannual variations of the total TC genesis frequency and the two ENSO types are present, a significant relationship appears for some clusters at a certain ENSO phase. A warm EP-ENSO event decreases the C2 TC frequency by 40% because of the formation of a downward motion and thereby high-pressure anomalies across the tropical Eastern Hemisphere from the Indian Ocean to the C2 genesis area because of a strong reversed Walker circulation forced by the EP SST anomalies. Meanwhile, the warm CP ENSO event generates a cyclonic Rossby wave circulation response northwest of the SST forcing, leading to a 50% increase in the C4 TC frequency. The negative phases of EP and CP ENSO have nonsignificant relationships with C2 and C4 because of the weaker cold SST forcing and the interdecadal variation in the amplitude of the cold CP SST anomalies, respectively. These nonlinear characteristics result in low correlations (~0.35) between the EP/CP ENSO and the C2/C4 genesis frequency.

The interannual variability in the C5 TCs is strongly controlled by the variation of the WPSH with a correlation of −0.71, the highest in this study. A positive WPSH phase reduces the C5 genesis frequency by 66%, while negative WPSH anomalies enhance it by 50%. These linear properties result in a higher correlation, in

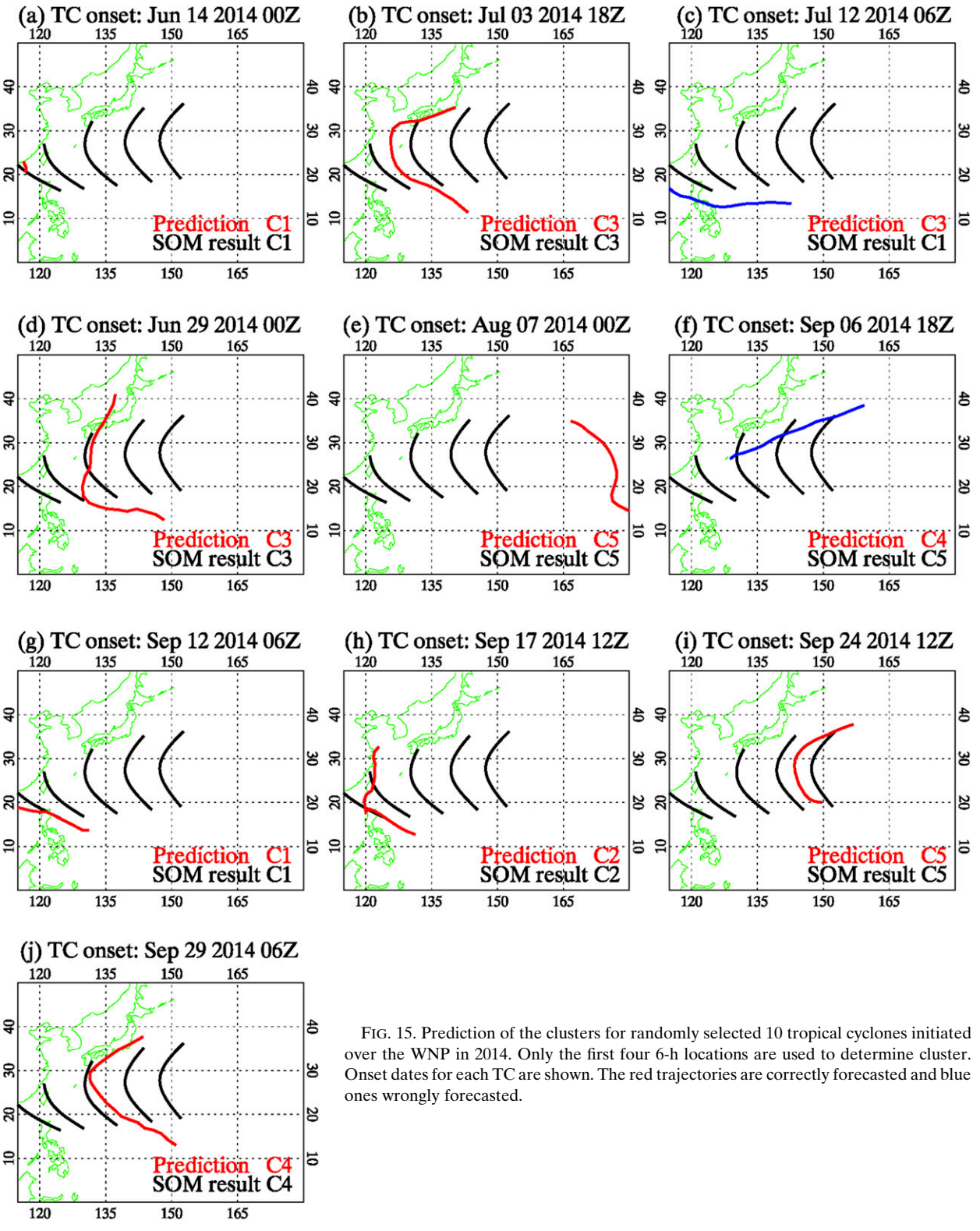

FIG. 15. Prediction of the clusters for randomly selected 10 tropical cyclones initiated over the WNP in 2014. Only the first four 6-h locations are used to determine cluster. Onset dates for each TC are shown. The red trajectories are correctly forecasted and blue ones wrongly forecasted.

contrast with the nonlinear relationships for the EP and CP ENSO. The C5 TCs tend to develop over the reduced vertical wind shear area.

The TC genesis frequency for each cluster has no significant relationship with the Hadley circulation strength, NAO, PDO, or QBO. Although the result with the variation of Indian Ocean SST is not presented here, any significant relationship is not seen with this factor. The interannual variation of the monsoon trough (not shown), the strength of which can be represented by the low-level relative vorticity (Wu et al. 2012), shows a statistically significant relationship with the C4 genesis frequency ($r = 0.39$), as expected from our results. C5 TCs appear to closely relate with the monsoon trough variation since the mean genesis position is located between 150° and 160°E (Figs. 2 and 14). However, the calculated correlation is almost zero because the mean latitude of the C5 TC genesis location is more toward the higher latitude compared to the monsoon trough extension during strong WNP monsoons (~5°–15°N).

With the current clustering analysis, the applicability for near-real-time determination of clusters is tested for the genesis dates. For independent testing, 2014 data are used, and 10 tropical cyclones are randomly selected. A TC cluster has been determined by using only the first four 6-h interval data (just one day) from the onset date and by the lowest Euclidean distance from a cluster centroid. As shown in Fig. 15, eight tropical cyclones are correctly classified, indicating high prediction skill and effective clustering in this study. Two TCs (Figs. 15c,f) are wrongly forecasted because of a rather zonally oriented pathway along 12°N and 26°–35°N, respectively. A better performance is expected when including data of TC track density and interannual large-scale atmospheric and oceanic variability, which will be the subject of future research.

The current study mainly focused on the relationships between the genesis frequency for different TC clusters and the intraseasonal and interannual large-scale variability. The characteristics of the TC tracks, activities, seasonal variation, and landfall for each cluster and their relationships with longer and large-scale atmospheric and oceanic variability will be examined in detail in the future. For example, the relationship between changes in TC track density and the stratospheric QBO (Ho et al. 2009) phase will be sought.

Acknowledgments. Thanks to Nat Johnson at the IPRC, University of Hawai'i, for the helpful discussion on FDR and SOM clustering methods. We appreciate the anonymous reviewers for their constructive suggestions and comments, which improved the manuscript. This work was funded by the Korea Meteorological Administration Research and Development Program under Grant KMIPA 2015-2113 and National Research Foundation of Korea (NRF) Grant by the Korean government (MSIP) (NRF-2014R1A2A1A11051818). The authors would like to acknowledge the support from the Korea Institute of Science and Technology Information (KISTI).

APPENDIX

Definition of Indices

EP	Eastern Pacific SST anomaly index made by summing the Niño-1+2 [10°S–0°, 90°–80°W] and Niño-3 [5°S–5°N, 150°–90°W] indices
CP	Central Pacific SST anomaly averaged over [5°S–5°N, 180°E–140°W]
NAO	The first EOF mode of the rotated principal component analysis for the 500-hPa geopotential height over 20°–90°N. This index is provided by the NOAA (http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/norm.nao.monthly.b5001.current.ascii)
PDO	The leading EOF mode of the monthly SST anomalies over the North Pacific during the tropical cyclone season. This index is provided by the University of Washington (http://jisao.washington.edu/pdo/PDO.latest)
WPSH	The maximum interannual standard deviation of the 850-hPa geopotential height averaged over [20°–30°N, 130°–160°E] during the tropical cyclone season
Hadley	The maximum value of the zonal mean mass streamfunction occurring within the latitudinal zone of 0°–30°N (Oort and Yienger 1996)
QBO	The monthly and zonal mean equatorial zonal wind at 50 hPa provided by the NOAA/Climate Prediction Center

REFERENCES

Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.*, **57B**, 289–300.

Bensaid, A. M., L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, and R. F. Murthagh, 1996: Validity-guided (re)clustering with applications to image segmentation. *IEEE Trans. Fuzzy Syst.*, **4**, 112–123, doi:10.1109/91.493905.

Bezdek, J. C., 1981: *Pattern Recognition with Fuzzy Objective Function Algorithms.* Kluwer Academic Publishers, 256 pp.

Camargo, S. J., K. A. Emanuel, and A. H. Sobel, 2007a: Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Climate*, **20**, 4819–4834, doi:10.1175/JCLI4282.1.

——, A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007b: Cluster analysis of typhoon tracks. Part I: General properties. *J. Climate*, **20**, 3635–3653, doi:10.1175/JCLI4188.1.

——, ——, ——, ——, and ——, 2007c: Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654–3676, doi:10.1175/JCLI4203.1.

Chu, J.-E., S. N. Hameed, and K. J. Ha, 2012: Nonlinear, intraseasonal phases of the East Asian summer monsoon: Extraction and analysis using self-organizing maps. *J. Climate*, **25**, 6975–6988, doi:10.1175/JCLI-D-11-00512.1.

Chu, P.-S., 2002: Large-scale circulation features associated with decadal variations of tropical cyclone activity over the central North Pacific. *J. Climate*, **15**, 2678–2689, doi:10.1175/1520-0442(2002)015<2678:LSCFAW>2.0.CO;2.

Dunn, J. C., 1973: A fuzzy relative of the ISODATA process and its use in detecting compact well-separated cluster. *J. Cybern.*, **3**, 32–57, doi:10.1080/01969727308546046.

Elsner, J. B., and K. B. Liu, 2003: Examining the ENSO-typhoon hypothesis. *Climate Res.*, **25**, 43–54, doi:10.3354/cr025043.

Gill, A. E., 1980: Some simple solutions for heat-induced tropical circulations. *Quart. J. Roy. Meteor. Soc.*, **106**, 447–462, doi:10.1002/qj.49710644905.

Hall, J. D., A. J. Matthews, and D. J. Karoly, 2001: The modulation of tropical cyclone activity in the Australian region by the Madden–Julian oscillation. *Mon. Wea. Rev.*, **129**, 2970–2982, doi:10.1175/1520-0493(2001)129<2970:TMOTCA>2.0.CO;2.

Harr, P. A., and R. L. Elsberry, 1995a: Large-scale circulation variability over the tropical western North Pacific. Part I: Spatial patterns and tropical cyclone characteristics. *Mon. Wea. Rev.*, **123**, 1225–1246, doi:10.1175/1520-0493(1995)123<1225:LSCVOT>2.0.CO;2.

——, and ——, 1995b: Large-scale circulation variability over the tropical western North Pacific. Part II: Persistence and transition characteristics. *Mon. Wea. Rev.*, **123**, 1247–1268, doi:10.1175/1520-0493(1995)123<1247:LSCVOT>2.0.CO;2.

Ho, C.-H., H.-S. Kim, J.-H. Jeong, and S.-W. Son, 2009: Influence of stratospheric quasi-biennial oscillation on tropical cyclone tracks in the western North Pacific. *Geophys. Res. Lett.*, **36**, L06702, doi:10.1029/2009GL037163.

Johnson, N. C., 2013: How many ENSO flavors can we distinguish? *J. Climate*, **26**, 4816–4827, doi:10.1175/JCLI-D-12-00649.1.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>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, doi:10.1175/2010JCLI3751.1.

Kohonen, T., 1990: The self-organizing map. *Proc. IEEE*, **78**, 1464–1480, doi:10.1109/5.58325.

——, 1997: *Self-Organizing Maps.* 2nd ed. Springer Series in Information Sciences, Vol. 30, Springer-Verlag, 426 pp.

Lander, M. A., 1996: Specific tropical cyclone track types and unusual tropical cyclone motions associated with a reverse-oriented monsoon trough in the western north Pacific. *Wea. Forecasting*, **11**, 170–186, doi:10.1175/1520-0434(1996)011<0170:STCTTA>2.0.CO;2.

Lee, J. Y., B. Wang, M. C. Wheeler, X. Fu, D. E. Waliser, and I. S. Kang, 2012: Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Climate Dyn.*, **40**, 409–509, doi:10.1007/s00382-012-1544-4.

Li, R. C. Y., and W. Zhou, 2013a: Modulation of western North Pacific tropical cyclone activity by the ISO. Part I: Genesis and intensity. *J. Climate*, **26**, 2904–2918, doi:10.1175/JCLI-D-12-00210.1.

——, and ——, 2013b: Modulation of western North Pacific tropical cyclone activity by the ISO. Part II: Tracks and landfalls. *J. Climate*, **26**, 2919–2930, doi:10.1175/JCLI-D-12-00211.1.

——, ——, J. C. L. Chan, and P. Huang, 2012: Asymmetric modulation of western North Pacific cyclogenesis by the Madden–Julian oscillation under ENSO conditions. *J. Climate*, **25**, 5374–5385, doi:10.1175/JCLI-D-11-00337.1.

Li, T., 2012: Synoptic and climatic aspects of tropical cyclogenesis in western North Pacific. *Cyclones: Formation, Triggers and Control*, K. Oouchi and H. Fudevasu, Eds., Nova Science Publishers, 61–94.

Liebmann, B., and C. A. Smith, 1996: Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull. Amer. Meteor. Soc.*, **77**, 1275–1277.

Liu, Y., R. H. Weisberg, and C. N. K. Mooers, 2006: Performance evaluation of the self-organizing map for feature extraction. *J. Geophys. Res.*, **111**, C05018, doi:10.1029/2005JC003117.

Nolan, D. S., and M. G. McGauley, 2012: Tropical cyclogenesis in windshear: Climatological relationships and physical processes. *Cyclones: Formation, Triggers and Control*, K. Oouchi and H. Fudeyasu, Eds., Nova Science Publishers, 1–34.

Oort, A. H., and J. J. Yienger, 1996: Observed interannual variability in the Hadley circulation and its connection to ENSO. *J. Climate*, **9**, 2751–2767, doi:10.1175/1520-0442(1996)009<2751:OIVITH>2.0.CO;2.

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

Seo, K.-H., J.-K. E. Schemm, W. Wang, and A. Kumar, 2007: The boreal summer intraseasonal oscillation simulated in the NCEP Climate Forecast System (CFS): The effect of sea surface temperature. *Mon. Wea. Rev.*, **135**, 1807–1827, doi:10.1175/MWR3369.1.

——, W. Wang, J. Gottschalck, Q. Zhang, J.-K. E. Schemm, W. R. Higgins, and A. Kumar, 2009: Evaluation of MJO forecast skill from several statistical and dynamical forecast models. *J. Climate*, **22**, 2372–2388, doi:10.1175/2008JCLI2421.1.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296, doi:10.1175/2007JCLI2100.1.

Wang, B., and J. C. L. Chan, 2002: How strong ENSO events affect tropical storm activity over the western North Pacific. *J. Climate*, **15**, 1643–1658, doi:10.1175/1520-0442(2002)015<1643:HSEEAT>2.0.CO;2.

——, B. Xiang, and J.-Y. Lee, 2013: Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. *Proc. Natl. Acad. Sci. USA*, **110**, 2718–2722, doi:10.1073/pnas.1214626110.

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, doi:10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Wilks, D. S., 2006: On "field significance" and the false discovery rate. *J. Appl. Meteor. Climatol.*, **45**, 1181–1189, doi:10.1175/JAM2404.1.

Wu, L., X. Wen, R. Huang, and R. Wu, 2012: Possible linkage between the monsoon trough variability and the tropical cyclone activity over the western North Pacific. *Mon. Wea. Rev.*, **140**, 140–150, doi:10.1175/MWR-D-11-00078.1.

Xie, X. L., and G. A. Beni, 1991: A validity measure for fuzzy clustering. *IEEE Trans. Pattern Anal. Mach. Intell.*, **13**, 841–845, doi:10.1109/34.85677.

Xu, Y., T. Li, and M. Peng, 2013: Tropical cyclogenesis in the western North Pacific as revealed by the 2008–09 YOTC data. *Wea. Forecasting*, **28**, 1038–1056, doi:10.1175/WAF-D-12-00104.1.

Yoshida, R., and H. Ishikawa, 2013: Environmental factors contributing to tropical cyclone genesis over the western North Pacific. *Mon. Wea. Rev.*, **141**, 451–467, doi:10.1175/MWR-D-11-00309.1.

Zhang, W., H.-F. Graf, Y. Leung, and M. Herong, 2012: Different El Niño types and tropical cyclone landfall in East Asia. *J. Climate*, **25**, 6510–6523, doi:10.1175/JCLI-D-11-00488.1.

Zhou, T., and Coauthors, 2009: Why the western Pacific subtropical high has extended westward since the late 1970s. *J. Climate*, **22**, 2199–2215, doi:10.1175/2008JCLI2527.1.