Influences of ENSO and intraseasonal oscillation on distinct tropical cyclone clusters over the western North Pacific

Yitian Qian[1], Pang-Chi Hsu[1], Hiroyuki Murakami[2, 3], Jianyun Gao[4], Huijing Wang[1], Mingkeng Duan[1]

[1] *Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD)/Joint International Research Laboratory of Climate and Environment Change (ILCEC), Nanjing University of Information Science and Technology, Nanjing, China*

[2] *Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA.*

[3] *Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA.*

[4] *Fujian Key Laboratory of Severe Weather, Fujian Institute of Meteorological Sciences, Fuzhou, 350001, China*

Submitted to *Climate Dynamics*

Corresponding author: Pang-Chi Hsu; Email: pangchi@nuist.edu.cn

Abstract

Although the influences of El Niño–Southern Oscillation (ENSO) and boreal summer intraseasonal oscillation (ISO) on basin-wide tropical cyclone (TC) activity over the western North Pacific (WNP) have been widely recognized, how the seasonal and subseasonal anomalies of sea surface temperature and atmospheric ISO variations modulate different types of WNP TCs needed further examination, as addressed in this study. Using a fuzzy *c*-means clustering method, we objectively classified the WNP TCs into seven distinct clusters with different genesis locations and trajectories. The genesis numbers of all seven TC clusters revealed significant spectral variance at the intraseasonal timescale in the bands of 10–30 and 30–90 days. Based on the diagnosis of scale-decomposed genesis potential index, we found that the increase in ISO-related mid-tropospheric moistening plays the most important role in TC genesis for all seven clusters, while anomalous circulations (low-level vorticity and mid-level vertical motion) are secondary. The trajectories associated with straight-moving and recurving TC clusters are modulated by ISO-related steering flows. These modulations of TC activities by ISO vary with the phase of ENSO. The modulations of ISO are significantly greater for TCs generated in the southeast quadrant of the WNP in El Niño years than in La Niña years, while ISO imposes a larger impact on landfalling TCs occurring in La Niña years, which are changed by the low-level winds associated with ENSO conditions. The compound effects of ENSO and ISO on TC clusters provide useful sources of subseasonal TC predictability. Our statistical model using the information of ENSO and ISO shows skillful predictions of WNP TC genesis numbers and track distributions at the lead time up to 30 days.

Keywords: tropical cyclone; intraseasonal oscillation; ENSO; cluster analysis;

subseasonal prediction; western North Pacific

1 Introduction

Tropical cyclones (TCs) are one of the costliest natural disasters affecting coastal regions over much of the world (e.g., Pielke et al. 2008; Smith and Katz 2013). Owing to its warm sea surface temperature (SST), abundant moisture, and large-scale cyclonic circulation of the monsoon trough, the western North Pacific (WNP) is the region with the highest TC occurrence, accounting for 30% of the TCs over all ocean regions globally (e.g., Gray 1968; McBride 1995). The activities of TCs over the WNP are strongly influenced by natural variabilities with different timescales. At the interannual timescale, the TC genesis location, number, track, and intensity are modulated by SST anomalies associated with the El Niño–Southern Oscillation (ENSO) (Chia and Ropelewski 2002; Wang and Chan 2002; Camargo et al. 2007a) and Pacific Meridional Mode (Zhang et al. 2016a; Murakami et al. 2017). As the dominant mode of subseasonal variability, the phase evolution and intensity of boreal summer intraseasonal oscillation (ISO) also influence the activities of WNP TCs (Li and Zhou 2013a, b; Zhao et al. 2015). Increased (decreased) TC genesis numbers are observed during the active (inactive) phase of ISO when the cyclonic anomaly, abundant moisture, small vertical wind shear, and deep convection provide a favorable environment for TC formation (Liebmann et al. 1994; Kim et al. 2008; Hsu et al. 2011; Li and Zhou 2013a; Zhao et al. 2015). By investigating the genesis potential index (GPI), Camargo et al. (2009) found the largest contribution to derive from the ISO relative humidity anomaly, followed by the low-level absolute vorticity (Zhao et al. 2015). Although the influences of ISO on basin-wide TC activities are known, the question of whether ISO imposes the same effects on different types (with different genesis locations and different trajectories) of WNP TCs remains unanswered. The modulations by ISO on TC genesis numbers in different ENSO phases have also been discussed. Li et al. (2012) revealed an asymmetric modulation of TC genesis by the Madden–Julian Oscillation (MJO) under different ENSO conditions, with an enhanced MJO–TC relationship in El Niño years compared with La Niña years. Han et al. (2019) indicated that more TCs are generated over the WNP in active phases of the 10–30-day Quasi-Biweekly Oscillation (QBWO) in ENSO neutral years than El Niño or La Niña years. Zhao et al. (2022) found that ENSO, and then either the MJO or QBWO, were the two most important predictors, after the climatology of TC genesis, over different sub-basins of the WNP, in their statistical intraseasonal TC forecast model.

The modulations of TCs by ENSO (along with other interannual SST anomalies) and ISO provide a physical basis for seasonal and subseasonal predictions of WNP TC activities (Gray et al. 1992, 1993, 1994; Klotzbach and Gray 2004, 2009; Klotzbach 2008; Leroy and Wheeler 2008; Vitart et al. 2010; Slade and Maloney 2013; Xiang et al. 2015; Murakami et al. 2015; Nakano et al. 2017; Jiang et al. 2018; Lee et al. 2018). Compared to seasonal TC prediction systems, which are relatively well developed (Gray et al. 1992, 1993, 1994; Camargo and Barnston 2009; Kim et al. 2012; Murakami et al. 2015, 2016a, b), subseasonal TC prediction models are still in their infancy (Leroy and Wheeler 2008; Vitart et al. 2010; Slade and Maloney 2013; Nakano et al. 2017; Jiang et al. 2018; Lee et al. 2018). In early work, Leroy and Wheeler (2008) developed a logistic regression model to predict the weekly TC genesis probability in the Southern Hemisphere based on the climatology of TC genesis numbers, the real-time multivariate MJO (RMM) index, and the two leading modes of the interannual SST variability associated with ENSO. Their statistical model showed an increased prediction skill, out to about the third week, by including MJO information, suggesting that ISO serves as an important source of TC genesis predictability at the subseasonal timescale.

Following the work of Leroy and Wheeler (2008), Slade and Maloney (2013) constructed a statistical model for intraseasonal TC genesis prediction over the Atlantic and eastern Pacific and indicated that a skillful prediction could be achieved at lead times of 2–3 weeks when the MJO index was used in the model. Note, however, that the statistical models of Leroy and Wheeler (2008) and Slade and Maloney (2013) only provided predictions of whether or not a TC would occur over these ocean basins; they lacked information on the genesis numbers, genesis locations, and trajectories of TCs. Wei et al. (2021) constructed a similar spatiotemporal projection method (STPM) to that proposed by Hsu et al. (2015) to predict the principal components (PCs) of the MJO and QBWO. Then, they predicted the TC genesis over the South China Sea (SCS) based on the predicted PC and historical statistical relationship between the QBWO/MJO and TC daily genesis rate.

With improvements in high-resolution dynamic models in simulating ISO and its relationship with TCs (Jiang et al. 2012; Satoh et al. 2012), it has become possible to predict the generation and movement of TCs at longer lead times of several weeks (Elsberry et al. 2010; Vitart et al. 2010; Nakano et al. 2017; Xiang et al. 2015; Jiang et al. 2018; Lee et al. 2018). For instance, Elsberry et al. (2010) and Vitart et al. (2010) showed that the European Centre for Medium-Range Weather Forecasts (ECMWF) monthly forecast model is able to provide useful guidance for the TC genesis over the WNP and the Southern Hemisphere a few weeks in advance. Based on the high-resolution (50-km) coupled model of the National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory (NOAA/GFDL), Xiang et al. (2015) reported that the genesis of Hurricane Sandy in the Atlantic in 2012, and Super Typhoon Haiyan in the western Pacific in 2013, could be correctly predicted with a lead time of about 11 days. Jiang et al. (2018) then extended the study of Xiang et al. (2015)

by assessing the prediction skill of the GFDL forecast system for all 600+ TCs that occurred during 2003–2013. They found that only about 10% of the TCs were correctly predicted at a one-week lead, indicating that subseasonal TC prediction is still a challenge for state-of-the-art high-resolution coupled models. A similar conclusion was obtained by Lee et al. (2018), who evaluated the TC genesis prediction skill of six subseasonal-to-seasonal (S2S) models and found that the ECMWF S2S model performed best, with skillful predictions for TC genesis numbers in the Atlantic and WNP in the week 2 (days 8 to 14) forecast. The other S2S models displayed limited skill at a one-week lead. The prediction skill for probabilistic TC occurrence at a regional scale (15° latitude × 20° longitude) over different basins was assessed by Gregory et al. (2020) and Lee et al. (2018), who both concluded that the ECMWF model shows the best skill, out to 3–4 weeks in advance. Qian et al. (2020) introduced a novel dynamical–statistical model for subseasonal TC forecasting over the WNP, with a 25-day lead time, which considers different TC clusters with distinct genesis locations and trajectories. Additional progress in the mechanisms of subseasonal variability and other operational TC prediction systems can be found in the review paper by Camargo et al. (2019).

In summary, the overall skill revealed by the operational systems is still limited (at around two weeks), which is because the modulation by ISO of TC subseasonal variabilities is not yet fully understood, especially regarding the compound impacts of ISO and ENSO on different kinds of TCs over the WNP. This suggests a need to further explore efficient and useful approaches to subseasonal TC prediction. Accordingly, in this study, we investigated in greater detail the modulating effects of ISO and ENSO on different types of WNP TCs, which had been previously objectively categorized by clustering analysis (Camargo et al. 2007a; Chu et al. 2010; Kim et al. 2011; Zhang et al. 2016b; Zhao et al. 2018). In this previous work, Camargo et al. (2007a) briefly demonstrated the connection between the MJO phases and ENSO and each of the TC clusters. However, the detailed processes through which ISO influences distinct TC clusters, such as the relative contributions of ISO-related dynamic and thermodynamic effects on different TC clusters, have not been elucidated. Besides, with the impact of the different phases of ENSO evolution, the different modulations of ISO on different TC clusters should also be investigated. Studies have thus far mostly focused on the subseasonal prediction of TC genesis numbers, but probabilistic information with respect to TC genesis locations and track patterns over the entire WNP would also be useful for risk management and policymakers. Based on results regarding the modulations by ISO and ENSO on individual TC clusters, we constructed statistical models by using the key ISO-related environmental factors in different ENSO phases as potential predictors of the genesis numbers of each TC cluster at the lead times of 10–40 days. Then, we obtained probability maps of TC frequency (TCF) by considering both the genesis number and climatological probability of TCF for individual clusters (Chu et al. 2010; Kim et al. 2011).

The rest of the paper is organized as follows: In section 2, we describe the methods and data used in our study. The modulating effects of ISO on distinct TC clusters, including their genesis and trajectory patterns, as well as the different impacts of ISO on distinct TC clusters under ENSO-related large-scale fields, are examined in section 3. In section 4, we present the procedures of WNP TC prediction at 10–40-day leads, along with the forecast results. Conclusions and some further discussions are provided in section 5.

2 Data and methods

2.1 Data

1) *TC dataset*

The TC genesis and track data over the WNP (0°–60°N, 100°E–180°) are from the Reginal Specialized Meteorological Centers Tokyo-Typhoon Center best-track dataset, as archived in the International Best Track Archive for Climate Stewardship (Knapp et al. 2010). We focus on tropical storms or stronger cyclones (sustained wind speed ≥ 34 knots), collectively referred to as TCs in this study.

2) *Large-scale field datasets*

The GPI was calculated using the large-scale fields obtained from the ERA-Interim dataset (Dee et al. 2011), which has a 1.5° × 1.5° resolution. The variables used include 3-D zonal and meridional wind fields, vertical velocity, relative humidity, temperature, and geopotential height. The SST data are from the NOAA High-resolution Blended Analysis of daily SST at a 0.25° × 0.25° resolution (Reynolds et al. 2007). All the large-scale fields were interpolated into a uniform resolution of 1.5° × 1.5°. The study period covers the WNP TC season (June–November) during 1982–2018.

3) *Definition of ENSO state*

El Niño (La Niña) years are defined as when the seasonal mean (July–November) Niño-3.4 (5°S–5°N, 170°–120°W) SST anomaly is greater (less) than 0.8σ, based on the NOAA High-resolution Blended Analysis of daily SST. According to this definition, six El Niño years (1982, 1987, 1997, 2002, 2009, 2015) and ten La Niña years (1984, 1985, 1988, 1989, 1998, 1999, 2000, 2007, 2010, 2011) are defined. The remaining 21 years are neutral years.

The forecast Niño-3.4 index [area-averaged SST over (5°S–5°N, 170°–120°W)] is used as the product of the North American Multi-Model Ensemble (NMME) forecasting project (Kirtman et al. 2014). The real-time Niño-3.4 index can be downloaded from https://ftp.cpc.ncep.noaa.gov/NMME/realtime_anom/ENSMEAN/. Six models provide real-time Niño-3.4 index datasets, while only four models provide early predictions from 1982. Therefore, we utilized the one-month lead real-time Niño-3.4 index products provided by these four models, including CanCM4i, CFSv2, NASA_GEOS5v2, and NCAR_CCSM4. The forecast El Niño (La Niña) phases are defined as when the real-time seasonal mean (July–November) Niño-3.4 index is greater than 0.8σ (less than -0.8σ) at a lead time of one month.

2.2 Methods

1) *Fuzzy c-means clustering analysis*

To objectively classify the WNP TC clusters (or track patterns), we used the fuzzy *c*-means clustering method (Bezdek 1981), which is based on minimizing an objective function of the *c*-means function. The mathematics of the *c*-means function can be found in Chu et al. (2010) and Kim et al. (2011), which highlight the superiority of the *c*-means clustering method in analyzing clusters with ambiguous boundaries, like TC data.

According to the results of previous studies (Camargo et al. 2007b; Chu et al. 2010; Kim et al. 2011; Zhang et al. 2016b), WNP TCs can be suitably classified into seven clusters, denoted by C1 to C7, respectively. Figure 1 and Table 1 display their basic features, including their genesis and center locations of trajectories, lifespans, trajectories lengths, maximum wind speeds, and accumulated cyclone energy (ACE; Bell 2000). “TCall” represents all TC cases over the WNP basin. Briefly, C1 is characterized by TCs that generate and develop over the SCS and might strike southern China and the Indochina Peninsula (Fig. 1a). The TCs of C1 have the shortest lifespan

(2.7 days) and trajectory length (96.8 km), as well as the smallest maximum wind speed (64.3 m s^{-1}) and ACE (9078 m^2 s^{-2}), among the seven clusters, owing to the limited basin size of the SCS and the short distance to land. The TCs of C2 have the eastmost (162.4°E) and northmost (27.0°N) genesis locations, which form and propagate over the open ocean of the WNP without making landfall (Fig. 1b). C3 is another cluster with most TCs being active in the open ocean; however, the TCs of C3 are generated more to the south and west than those of C2, and have longer lifespans (Fig. 1c, Table 1). C4 is characterized by TCs moving northwestward and turning northeastward, which affects southeastern Japan (Fig. 1d). The TCs of C5, with straightforward and west/northwest-oriented trajectories, mostly strike Taiwan and East China (Fig. 1e). C6 is a cluster of TCs that form in the southeastern part of the WNP, and has the longest lifespan and largest ACE (Table 1). Finally, the TCs of C7 are generated to the east of the Philippines and easily make landfall in South Korea and Japan.

2) *Definitions of ISO activity*

The boreal summer ISO consists of two different modes, with periods of 10–30 and 30–90 days. To separate the high-frequency (HF; period of 10–30 days) and low-frequency (LF; period of 30–90 days) ISO components, the Lanczos band-pass filter (Duchon 1979) was used. The dominant patterns of the HF/LF ISO modes over the WNP were further identified by using empirical orthogonal function (EOF) analysis of band-pass-filtered outgoing longwave radiation (OLR) over the WNP (0°–60°N, 100°E–180°) during June–November of 1982–2018. The amplitude and phase of the HF/LF ISO activities were then defined by the first two leading PCs (PC1 and PC2). An active LF/HF ISO day was selected when the amplitude of the HF/LF ISO $[(PC1^2 + PC2^2)^{1/2}]$ was greater than or equal to 1.

3) *Scale-decomposed GPI*

To quantitatively identify key factors associated with the ISO in modulating TC activity over the WNP, we examined the GPI (I_{GPI}) at the intraseasonal timescale over the TC genesis region for each TC cluster. A modified GPI that incorporates the vertical motion effect proposed by Murakami and Wang (2010) was used in this study:

$$I_{\mathrm{GPI}} = |10^5\zeta|^{3/2}\left(\frac{I_{\mathrm{RH}}}{50}\right)^3\left(\frac{V_{\mathrm{pot}}}{70}\right)^3(1+0.1V_s)^{-2}\left(\frac{-w+0.1}{0.1}\right), \qquad (1)$$

where ζ is the 850-hPa absolute vorticity (s^{-1}), I_{RH} is the relative humidity (%) at 600 hPa, V_{pot} is the TC maximum potential intensity (MPI; m s^{-1}), V_{s} is the vertical wind shear (m s^{-1}) between 850 and 200 hPa, and w is the vertical pressure velocity (Pa s^{-1}) at 500 hPa. V_{pot} was defined by Emanuel (1995) and modified by Bister and Emanuel (1998) as follows:

$$V_{\mathrm{pot}}^2 = \frac{C_k}{C_D}\frac{T_s}{T_0}(\mathrm{CAPE}^* - \mathrm{CAPE}^{\mathrm{b}}), \qquad (2)$$

where C_k is the exchange coefficient for enthalpy, C_D is the drag coefficient, T_s is SST (K), and T_0 is the mean outflow temperature (K). The quantity CAPE^* is the value of the convective available potential energy (CAPE) with reference to the surrounding environment, and $\mathrm{CAPE}^{\mathrm{b}}$ is that of the boundary layer air.

For convenience, Eq. (1) is represented by the following five terms:

$$I_{\mathrm{GPI}} = \eta\cdot\gamma\cdot\psi\cdot s\cdot\omega, \qquad (3)$$

where η denotes the GPI component associated with 850-hPa absolute vorticity; γ denotes the GPI term related to relative humidity at 600 hPa; ψ represents the GPI component of potential intensity; and s and ω represent the GPI terms of vertical wind shear and vertical motion, respectively.

To examine the relative effects of the climatological seasonal cycle and subseasonal perturbations (mainly ISO) on the five large-scale factors of the GPI, the intraseasonal components of GPI variables are defined as follows:

$$I'_{\mathrm{GPI}} = (\bar{\eta}+\eta')\cdot(\bar{\gamma}+\gamma')\cdot(\bar{\psi}+\psi')\cdot(\bar{s}+s')\cdot(\bar{\omega}+\omega') - \bar{\eta}\cdot\bar{\gamma}\cdot\bar{\psi}\cdot\bar{s}\cdot\bar{\omega} = \eta'\cdot\bar{\gamma}\cdot\bar{\psi}\cdot\bar{s}\cdot\bar{\omega} +$$

$$\bar{\eta}\cdot\gamma'\cdot\bar{\psi}\cdot\bar{s}\cdot\bar{\omega}+\bar{\eta}\cdot\bar{\gamma}\cdot\psi'\cdot\bar{s}\cdot\bar{\omega}+\bar{\eta}\cdot\bar{\gamma}\cdot\bar{\psi}\cdot s'\cdot\bar{\omega}+\bar{\eta}\cdot\bar{\gamma}\cdot\bar{\psi}\cdot\bar{s}\cdot\omega'+\mathrm{NL}. \qquad (4)$$

Here, a prime indicates the intraseasonal timescale (10–90 days) component, which was subjected to 10–90-day band-pass filtering (Duchon 1979), and an overbar represents the daily climatological mean of factors of the GPI. Thus, the changes in the GPI induced by ISO-related processes and by ISO interaction with the seasonal cycle can be represented by the sum of the five linear terms (with only one perturbation process) and nonlinear effects (with the involvement of two or more perturbation processes, such as $\eta'\cdot\gamma'\cdot\bar{\psi}\cdot\bar{s}\cdot\bar{\omega}$, $\eta'\cdot\gamma'\cdot\psi'\cdot\bar{s}\cdot\bar{\omega}$, …, $\eta'\cdot\gamma'\cdot\psi'\cdot s'\cdot\omega'$), represented by NL in Eq. (4). The scale-decomposed GPI equation has been used to discuss the effects of ISO on basin-total TC genesis over the eastern Pacific (Jiang et al. 2012) and WNP (Zhao et al. 2015). In this study, we applied this diagnostic method to each of the WNP TC clusters to reveal their formation mechanisms.

3 ISO modulations of seven TC clusters

3.1 Climatological conditions

Figure 2 presents a spectral analysis of the genesis numbers of each TC cluster, as well as all TC cases, over the WNP. In addition to the synoptic-scale (< 10 days) variability, significant spectral powers associated with the HF and LF intraseasonal variability are seen in the genesis numbers of each cluster. This suggests that TC generations in different locations (in the southeastern part of the WNP for C2–C4 and C6, or near the Philippine Sea for C5 and C7, or within the SCS for C1) are influenced by both HF and LF ISOs.

To confirm the phase relationship between ISO and TC genesis for individual clusters, Figs. 3 and 4 compare the preferred phases of HF and LF ISOs for TC formation. The evolutions of anomalous convection and circulation associated with HF

ISO are displayed in the upper panels of Fig. 3. The convective and low-pressure anomalies are initiated over the western tropical Pacific Ocean in phase 1 (Fig. 3a), and move northwestward towards the eastern Philippine Sea during phases 2–3 (Figs. 3b–c). Then, the convection and low-pressure anomalies become stronger and prevail over the Philippine Sea and SCS in phases 4–5 (Figs. 3d–e), but they weaken after moving into Southeast China during phases 6–8 (Figs. 3f–h). Meanwhile, suppressed convection accompanied by a high-pressure anomaly appears from the western tropical Pacific and moves northwestward toward the SCS (Figs. 3e–h). The statistics of all the WNP TC cases (Fig. 3p) show that about 64% (36%) of all TCs form during periods when the convective (suppressed) anomalies of HF ISO propagate from the western equatorial Pacific towards the SCS during phases 2–5 (6–8 and 1). The contrast in TC genesis numbers between active and inactive phases of HF ISO has been documented previously in Gao and Li (2011), Li and Zhou (2013a), and Zhao et al. (2015). When looking into the phase relationship between HF ISO and the TC genesis numbers of each cluster (Figs. 3i–o), we find that the most favorable (unfavorable) phase for TC genesis varies by cluster type. For example, the TCs of C1 generate the most (least) in phase 4 (phases 8 and 1) when the convective (suppressed) OLR anomalies maximize over the SCS (Fig. 3i). The active convection and low-pressure anomalies extending from the SCS to the Philippine Sea in phases 4–5 also favor the TC formations of C5 (Fig. 3m) and C7 (Fig. 3o). In contrast, the TCs of C4 and C6 tend to generate more frequently in phase 2 (Figs. 3l, and 3n) when HF ISO convection and low-pressure anomalies are located over the southeastern part of the WNP (Fig. 3b).

Figure 3 indicates that the main genesis locations for the different TC clusters vary according to the convection and low-pressure anomalies associated with HF ISO. This is also apparent in the phase relationship between LF ISO and the TC genesis numbers of individual clusters (Fig. 4). The zonally elongated OLR and geopotential height anomalies in phase 1 (Fig. 4a) and phase 8 (Fig. 4h) provide favorable conditions for the generation and propagation of the straight-moving clusters C5 and C6, while the opposite conditions appear in phase 5 (Figs. 4m and n). For the TCs of C3 and C6, i.e., TCs that generate in the southeastern part of the WNP, larger (smaller) numbers of TCs form in phase 1 (phases 4 and 5) when negative (positive) OLR anomalies show a southeastward extension. By comparing the TC genesis numbers against the MJO phase, Camargo et al. (2009) pointed out that the peak genesis numbers for most of the clusters appear in phases 5 and 6 of the MJO (based on the RMM index) as the MJO convection prevails in the tropical WNP, corresponding to phase 1 in this study.

To understand the modulating effects of ISO large-scale parameters on TC genesis for these distinct clusters, we diagnosed the scale-decomposed GPI [Eq. (4)]. Before comparing the contribution of each term to the total GPI, it was necessary to address whether the anomalous GPI can capture the TC genesis for each cluster. Figure 5 shows the composites of HF (10–30-day) and LH (30–90-day) GPI anomalies based on the TC genesis days for each cluster. Both the HF and LF GPI anomalies reasonably represent the main genesis regions of individual clusters (Figs. 5a–h, 5i–p). Notably, the HF GPI anomalies have a larger amplitude than the LF GPI for C1, C4, and C6, suggesting a larger contribution of HF ISO to these clusters as shown in the spectral analysis (Figs. 2a, d, and f). Even with a smaller amplitude, the GPI anomalies induced by LF ISO capture well the distinct genesis locations for different clusters (Figs. 5i–p). In the analysis of the phase relationship between HF/LH ISO and TC genesis (Figs. 3–4), we also find that the TCs tend to form in the locations with enhanced ISO convection, regardless of 10–30-day or 30–90-day convective signals. For the sake of simplicity, we consider the effects of both ISO modes on the 10–90-day GPI anomalies (Figs. 5q–

x). The TCs of various clusters tend to form within the maximum center of 10–90-day GPI anomalies. In other words, the 10–90-day GPI anomalies induced by the combined effect of HF and LF ISO activities could represent the potential of TC genesis for different clusters.

To identify the key processes (dynamic and/or thermodynamic effects) associated with the ISO contributing to the TC genesis of different clusters, we diagnosed and decomposed the GPI for each cluster (Fig. 6). Each term in Eq. (4) was calculated using a 6° × 6° box centered at each TC genesis location. For all the TC clusters, the relative humidity [$\gamma' = (\frac{I_{\mathrm{RH}}}{50})^3$] associated with the ISO plays the leading role in contributing to the 10–90-day GPI anomaly. The low-level absolute vorticity anomaly of the ISO [$\eta' = |10^5\zeta|^{3/2}$] is the secondary contributor, followed by the effect of ISO-related vertical motion [$\omega' = \left(\frac{-w+0.1}{0.1}\right)$]. The nonlinear effect of ISO–ISO interaction also exerts non-negligible influences on the ISO GPI anomaly, indicating that synoptic-scale activity also feeds back to the ISO (Hsu et al. 2011). In contrast, the effects of ISO-induced potential intensity and vertical wind shear anomalies are relatively small. The results are generally consistent with those of Camargo et al. (2009) and Zhao et al. (2015), who suggested that mid-level relative humidity and low-level absolute vorticity are the two most important factors affecting WNP TC genesis. Note that the results are not sensitive to the size (3° × 3° or 9° × 9°) of the box used for computing the GPI terms.

In addition to the influences on TC genesis, the ISO activities may modulate the TC trajectory by changing the background steering flow (Kim et al. 2008; Chen et al. 2009; Li and Zhou 2013b). Figure 7 shows the ISO-related anomalous steering flows (red vectors) in the four days after TC genesis for each cluster. A contrast between straight-moving TCs (C1, C5, and C6) and recurving TCs (C4 and C7) lies in the locations of ISO-related steering flow that regulates the relative intensity and locations of the monsoon trough and subtropical high. For the TCs of C1, the ISO-related cyclonic steering flow is confined to over the SCS, leading to the enhancement and westward extension of the WNP subtropical high. The enhanced anomalous southeasterly causes the TCs of C1 to move northwestward (Fig. 7a). A similar condition of a westward extension of the subtropical high together with a weakening of the monsoon trough can be observed during the TC developing stages of C5 (Fig. 7e) and C6 (Fig. 7f). On the contrary, a northeastward extension of an enhanced monsoon trough accompanied by weakening trade winds appears after the formation of C4 and C7 TCs (Figs. 7d and 7g), favoring the northward movement of TCs towards Korea and Japan.

3.2 Compound effects of ISO and ENSO on TCs

The impacts of ENSO on TC activity over the WNP have been investigated and identified as one of the most dominant modes modulating TC activities at interannual timescales (Wang and Chan 2002; Camargo and Sobel 2005). Nevertheless, the modulations by HF and LF ISOs of the different clusters of TC activities under the various backgrounds of ENSO phases still require further discussion. To assess the modulations of HF/LF ISOs on TC genesis numbers under different ENSO conditions, we calculated the daily TC genesis percentage (TCGP) in each HF/ISO phase under different ENSO conditions. The TCGP is defined as the ratio of the TC genesis number relative to the number of days for a particular HF/LF ISO phase. As demonstrated in Fig. 9, the TCGPs of the TCs that generate in the southeast quadrant of the WNP, such as those of C3, C4, and C6 (Figs. 8c, d, f), indicate greater activity in El Niño years (0.9% TCGP for C3, 2.3% TCGP for C4, and 1.8% TCGP for C6) compared to La Niña years (0.0% TCGP for C3, 1.6% TCGP for C4, and 0.4% for C6). Few TCs of C3 and C6 are generated in La Niña years. Conversely, more landfalling TCs are generated in the western WNP during La Niña years, such as the TCs of C5 (3.0% TCGP in La Niña years and 2.5% TCGP in El Niño years) and C7 (2.0% TCGP in La Niña years and 1.8% TCGP in El Niño years).

The different modulations of HF/LF ISOs on the seven TC clusters as well as all TCs over the WNP under El Niño (red bars) and La Niña (blue bars) conditions are shown in Figs. 8a–h/Figs. 9a–h, respectively. For all TCs over the WNP, the strength of the active convection (depressed convection) over the southeast quadrant of the WNP in phases 1 and 8 (phases 4–5) of the LF ISO is strengthened (offset) by the westerly wind anomalies associated with El Niño conditions [Fig. 8i (Fig. 8k)], while it is offset (strengthened) by the easterly wind anomalies with La Niña conditions [Fig. 8m (Fig. 8o)]. Therefore, a southeast (northwest) shift of positive TCGP anomalies occurs in phases 1 and 8 of El Niño (La Niña) years [Fig. 8i (Fig. 8m)] compared to climatological phases 1 and 8 (Fig. 8q). In contrast, the LF ISO-related convection over the WNP in phases 2–3 is essentially strengthened by the enhanced monsoon trough associated with La Niña conditions (Fig. 8n) compared to El Niño years (Fig. 8i), thus resulting in a larger TCGP in phases 2–3 during La Niña years (Fig. 8h). The HF ISO and ENSO also impact TC activities by changing the low-level winds associated with El Niño and La Niña years. The convections in the southeast quadrant of the WNP associated with phases 2–3 of HF ISO are enhanced (offset) by the westerly (easterly) anomalies associated with El Niño (La Niña) conditions [Fig. 9j (Fig. 9n)]. Furthermore, the modulation of ISO has different impacts on different TC clusters in El Niño and La Niña years. For the TCs of C3 and C6 generated in the southeast quadrant of the WNP, the TC genesis number anomalies of these clusters are negative in almost all phases of LF (blue bars in Figs. 8c, d and f) and HF (blue bars in Figs. 9c, d and f) ISO during La Niña years, indicating a weaker modulation of ISO in La Niña years for these two clusters compared with those in El Niño years. Strong easterly (westerly) anomalies over the tropical WNP associated with La Niña (El Niño) conditions suppress (enhance) the TC genesis numbers in the southeast quadrant of the WNP. In contrast, the TC genesis number anomalies of landfalling TCs, such as the TCs of C5 and C7, are negative in most phases (at least six out of eight phases) of LF (Figs. 8e and g) and HF (Figs. 9e and g) ISO during El Niño years, which results from the southeasterly steering flow of landfalling TCs being offset by the westerly anomalies associated with El Niño conditions.

In summary, we have examined the influences of ISO on the activities of individual TC clusters in this section. Our findings indicate that the location of TC genesis varies with the convective and low-pressure anomalies of both the HF and LF ISO modes. We have identified mid-tropospheric moistening and low-level cyclonic vorticity associated with ISO as being the leading contributors to TC genesis for all seven clusters. While El Niño and La Niña have similar impacts on TC genesis in most phases of LF and HF ISO, the modulation of LF and HF ISOs in some specific phases is still changed by the low-level winds associated with ENSO conditions. Additionally, we observed relatively weaker modulations of HF and LF ISOs on the TCs generated in the southeast quadrant of the WNP (landfalling TCs), such as the TCs of C3, C4, and C6 (C5 and C7), during La Niña (El Niño) years compared with those in El Niño (La Niña) years. These observations provide a physical basis for predicting the TC genesis numbers of each cluster using ISO-related fields as predictors. Given that each TC cluster has a distinct genesis location and trajectory, once the genesis numbers of individual clusters are predicted, the probability of TCF over the entire WNP basin can be obtained by integrating the track patterns for all clusters. Next, we provide a detailed description of the statistical prediction model and assess the WNP TC prediction skill at the subseasonal timescale.

4 Application to subseasonal TC prediction

4.1 Construction of statistical model

The modulation by ISO and ENSO of the TC genesis of each cluster (Figs. 2–9) provides a source of predictability for TC genesis at the subseasonal timescale. Then, by combining the TC track density of the seven clusters, we can obtain a probability map of TCF. This prediction method was proposed by Camargo et al. (2007a), and applied to predicting seasonal TCs by Chu et al. (2010) and Kim et al. (2012). In this study, we developed a statistical model to predict the TC genesis numbers and the probability distribution of TC tracks in every 10-day period at forecast lead times of 10–40 days.

Figure 10 illustrates the detailed steps of the prediction model's construction. The first step is to obtain distinct TC clusters by applying the clustering method to the historical TC best-track dataset (such as the result in Fig. 1). Since the TCs rarely occurred in a daily interval, our predictand here is the TC count in every 10-day period. Considering that the seasonal cycle of TC count is steady climatologically, we focus only on the prediction of TC count anomalies (deviations from a slow seasonal cycle of 10-day means). The slow seasonal cycle 10-day mean for each TC cluster retains only the mean and the first three harmonics of the daily climatology (black curves in Fig. 11) during the training period (1982–2008). Considering the different modulations of ISO under different ENSO phases, we constructed different statistical models for different ENSO phases. In other words, three statistical models were developed for each TC cluster: one for El Niño years, one for La Niña years, and another for neutral years.

Thus, in the second step, based on the lead–lag relationships between TCs and ISO anomalies with a temporal resolution of 10-day period (10-day mean) and a linear regression analysis, the potential predictors associated with ISO environments (mainly the factors of GPI and SST) are selected. Then, a linear regression model is derived for predicting anomalous TC numbers of each cluster in different ENSO phases. Using the preceding ISO-related predictors, the anomalous TC counts of individual clusters in different ENSO phases are predicted. By adding the climatological slow seasonal cycle of 10-day mean TC counts to the predicted anomalies of TC counts, the total TC counts for each cluster can be obtained. The final step (Step 3 in Fig. 10) is to construct the probability maps of TCF over the entire WNP basin. The TCF is defined as the frequency of TC occurrence in a particular 5° × 5° grid box, which indicates how often TCs passed through a specific grid box. This could be achieved by multiplying the TC counts of each cluster by the climatological probability of the track density of the corresponding ENSO phases. The results of all seven clusters are then summed. Following Kim et al. (2012), the TCF probability formula can be written as

$$P_{k,m}^{\mathrm{obs}}(lat, lon) = \frac{\mathrm{Freq}_{k,m}^{\mathrm{obs}}(lat,lon)}{N_{k,m}^{\mathrm{obs}}}, \quad (5)$$

where the variables with the superscript "obs" are observed variables in the *m*th 10-day period of the *k*th year; "lat" and "lon" are the latitude and longitude in degrees, respectively; $\mathrm{Freq}_{k,m}^{\mathrm{obs}}(\mathrm{lat}, \mathrm{lon})$ indicates the observed TCF within 5° of a grid point (lon, lat); and $N_{k,m}^{\mathrm{obs}}$ is the observed total TC count over the entire WNP.

A predicted probability map of TCF $[P_{k,m}^{\mathrm{fcst}}(\mathrm{lat}, \mathrm{lon})]$ can be approximately equal to the sum of the seven clusters of climatological probability of corresponding ENSO phases multiplied by the predicted TC counts (anomalous counts plus climatological 10-day period TC counts):

$$P^{\text{fcst}}_{k,m}(lat, lon) = \frac{1}{N^{\text{fcst}}_{k,m}}\left[\sum_{n=1}^{N} N^{\text{fcst}}_{C_n,k,m} \times \bar{P}^{\text{obs}}_{e,C_n,m}(lat, lon)\right]; \tag{6}$$

$$\bar{P}^{\text{obs}}_{e,C_n,l}(lat, lon) = \sum_{k=1}^{K} \frac{\text{Freq}^{\text{obs}}_{e,C_n,t,l}(lat, lon)}{N^{\text{obs}}_{e,C_n,t,m}}. \tag{7}$$

Here, the variables with the superscript "fcst" are the predicted variables; C_n represents the nth cluster; N is the total number of clusters (i.e., seven in this study); $\bar{P}^{\text{obs}}_{e,C_n,m}(lat, lon)$ is the mth climatological probability of C_n during K years; and e is the ENSO phase of a specific TC season (e = 1, 2, 3 refers to El Niño, La Niña, and neutral years, respectively). Here, the climatological probability is calculated using the TC data during the training period.

A "perfect reconstruction" based on "true" TC counts (instead of predicted TC counts) is used as the assessment benchmark. The perfect reconstruction probability [$P^{\text{pfrc}}_{k,m}(lat, lon)$] is written as

$$P^{\text{pfrc}}_{k,m}(lat, lon) = \frac{1}{N^{\text{obs}}_{k,m}}\left[\sum_{n=1}^{N} N^{\text{obs}}_{C_n,k,m} \times \bar{P}^{\text{obs}}_{C_n,m}(lat, lon)\right]. \tag{8}$$

We use the first 27-year period (1982–2008) as the training period and the last 10-year period (2009–2018) as the independent forecast period. During a TC season, there are 18 10-day-mean time points (1–10 June, 11–20 June, … , 18–27 November). Thus, we include 486 10-day time points (18 10-day-mean time points × 27 years) in the training period and 180 10-day-mean time points (18 10-day-mean time points × 10 years) in the forecast period. For a real-time operational application, the ISO (10–90-day) components are derived using a non-bandpass filtering method (Hsu et al. 2015). The following two steps are performed. First, the daily anomaly fields are obtained by subtracting the climatological mean and the first three harmonics of the climatological annual variability from the raw data to remove the low-frequency background state. Then, a 10-day mean is applied to match the prediction interval of a 10-day mean and to remove the synoptic-scale component (with a period shorter than 10 days).

4.2 Selection of predictors

To identify potential predictors for each TC cluster, we analyze a lead–lag correlation analysis between anomalous TC counts and preceding ISO-related dynamic and thermodynamic factors, including the GPI and its factors and SSTA, in different ENSO phases (Fig. 12). The results show that the TC genesis counts are not only correlated with local environmental conditions during the genesis period (lead 0d; top panels), but also with remote effects associated with the spatiotemporal evolution of the ISO (lead 10–40d, shown in seven lower panels). Moreover, significant correlation patterns between TCs and ISO-related dynamic and thermodynamic factors varies in different ENSO conditions (Figs. 12b–d, 12f–h). Although the predictors are selected based on the statistically significant grid points, the correlation patterns are large in scale. For instance, the anomalous TC counts of C1 are positively correlated with GPI anomalies in the main genesis regions situated over the tropical WNP (top panel in Fig. 12a). This positive GPI anomaly emerges in the tropical WNP about 20 days before TC genesis (fourth panel in Fig. 12a) and then propagates northeastward towards the western WNP region (fourth panels in Fig. 12a). Additionally, the anomalous TC counts of C6 are positively correlated with MPI anomalies over the eastern Pacific and negatively correlated with MPI anomalies over the western Pacific, as in the El Niño condition in the climatology (Fig. 12e). This finding suggests that ISO-related SST anomalies, like the El Niño-related large-scale pattern, could modulate the TC genesis frequency on the ISO timescale. The correlation maps between ISO-related dynamic and thermodynamic factors during positive and negative ENSO phases demonstrate the intraseasonal correlations between anomalous TC counts and the preceding ISO-related factors, which exclude the low-frequency influence of ENSO. For example, compared to the climatological correlations, the positive MPI anomalies over the eastern Pacific disappear and only negative MPI anomalies are found over the tropical central Pacific during El Niño years (Fig. 12e). The climatological negative MPI anomalies over the tropical WNP become positive and disorganized (Fig. 12b). In contrast to El Niño years, negative and positive MPI anomalies are found over the tropical WNP and eastern North Pacific, respectively, during La Niña years (Fig. 12c).

Considering the different correlation patterns between anomalous TC counts and ISO-related large-scale fields in different ENSO phases and the spatiotemporal evolution of the ISO, it is cumbersome and subjective to select predictors based on the highly correlated box for each ENSO phase, each forecast lead time, and each cluster. Therefore, we use a method similar to STPM in Hsu et al. (2015) to select predictors. First, in terms of the correlation coefficient (Cor) patterns for regions where anomalous TC counts (X_e) for each cluster and the ISO-related large-scale fields (Y_e) are statistically significantly correlated at the 95% confidence level (i.e., shading in Fig. 12), we focus only on those over Asia and the North Pacific (20°S– 40°N, 40°E–90°W):

$$\mathrm{Cor}(X_e, Y_e) = r(e, i, j, l), \tag{9}$$

where i and j denote the zonal and meridional grid points, l represents the lead times for the prediction time, and e indicates the ENSO phases of the TC seasons (i.e., Figs. 12b–c and Figs. 12f–h). Next, the PC predictors associated with large-scale fields are obtained by projecting anomalous large-scale fields onto Cor(X_e, Y_e) fields. The transfer function is then constructed by using a linear regression method. The parameters α and β are the regression coefficients of the prediction model during the training period. For each TC cluster, there are seven statistical models for each ENSO phase.

When we start the forecasting procedure, the ENSO phase of the TC season is first predicted based on the NMME results. The predicted ENSO phase of the TC season determines which statistical model we choose. For example, if the Niño-3.4 index exceeds 0.8σ, we choose the El Niño statistical model ($e = 1$) to predict each TC cluster. The temporal correlation coefficient between the observed and predicted Niño-3.4 index at a one-month lead during 1982–2018 is 0.97, which is statistically significant at the 99% confidence level.

4.3 Prediction skill of subseasonal TC activity

Figure 13 depicts the overall prediction skill of the statistical model based on the assessments of independent forecast results. The figure represents the temporal correlation coefficient (TCC) and root-mean-square error (RMSE) associated with each predictor at various lead times. As shown in Fig. 13, the model's prediction skill tends to diminish as the lead time increases, as evidenced by smaller TCCs and larger RMSEs. Notably, for the prediction of anomalous TC counts, the TCCs and RMSEs associated with TCall (Fig. 13a) and the sum of C1–C7 (Fig. 13b) differ among predictors at lead times of 15–30 days. Specifically, the RMSE of TCall ranges from 1.0–1.3, while that of the sum of C1–C7 increases to 1.1–1.75. Furthermore, the vertical velocity at 500 hPa exhibits the highest prediction skill for both TCC and RMSE among six predictions (purple lines in Fig. 13), whereas the GPI shows the fastest drop in prediction skill (red lines in Fig. 13). Although the SSTA predictor does not demonstrate the highest prediction skill during lead 0–10 days, its prediction skill remains stable even at lead times beyond 25 days. Additionally, the ensemble prediction's TCC is statistically significant at the 95% confidence level at a lead time of 30 days.

To demonstrate the importance of developing statistical forecast models tailored to different ENSO conditions, we compared our statistical model that incorporates ENSO effects (red lines, hereafter referred to as ENSO-fcst) with a simple statistical model based on climatological correlations that do not consider ENSO effects (blue lines, hereafter referred to as clim-fcst), as depicted in Fig. 14a. The analysis reveals that the prediction skills of ENSO-fcst and clim-fcst models are comparable at lead times of 10–15 days. However, at lead times of 20–40 days, the TCCs of the clim-fcst forecast model decline more rapidly and become statistically insignificant at the 90% confidence level, whereas the TCCs of the ENSO-fcst forecast model decrease more gradually and retain prediction skill up to a lead time of 30 days. The superior TCCs of the ENSO-fcst forecast model over the clim-fcst model indicate the advantage of developing statistical models that account for different ENSO phases. Furthermore, we observe that for the prediction of anomalous TC counts, the prediction skill of the sum of C1–C7 is higher than that of TCall at lead times of 10–30 days, indicating the need to construct statistical forecast models for individual TC clusters. We also utilized a stepwise multiple linear regression model to predict anomalous counts for each cluster at each lead time (not shown). The prediction skills were less skillful than that of the ensemble mean of linear regression models by each predictor. This may be attributable to the persistence of overfitting in the multiple linear regression.

After adding the climatological 10-day-mean TC counts to the predicted TC count anomalies, the TCCs of the sum of C1–C7 at the lead times of 10–30 days are higher than the climatological prediction, but they are lower than the climatological prediction after the 30-day forecast lead (Fig. 14b). The climatological prediction is the result of climatological 10-day-mean TC counts during the training period. The RMSEs are standard deviations of 1.00–1.35 (Fig. 14c). The reason why we utilize the intraseasonal anomaly deviations from the seasonal climatology as predictands in a linear regression model instead of total TC counts as predictands in a Poisson regression model is that the significant seasonal climatology of TC counts would contribute most of the prediction skill in a total TC counts prediction model. We also developed a Poisson regression model for the total TC counts by utilizing the same potential predictors as described in section 4.2. The prediction skill of the Poisson regression model was at a lead time of 10 days (not shown), which is lower than the result of the current linear regression model. Once the TC counts for each cluster are predicted, the probability map of TCF can be obtained by considering the climatological TCF distribution (introduced in section 4.1). Here, we show four TCF cases to provide a visual impression of TCF prediction (Fig. 15). The model can predict the regions with a higher probability of TCF in the observation (top panels in Fig. 15) at the lead time of 10–20 days. The predicted probability map of TCF of the perfect construction (second from top panels in Fig. 15) reveals errors of the model predicted biases in TC counts for a specific cluster (or some clusters).

5 Discussion and conclusions

Although ISO and ENSO modulations of basin-wide TC activities have been documented (Chia and Ropelewski 2002; Wang and Chan 2002; Camargo et al. 2007a; Li and Zhou 2013a, b; Zhao et al. 2015), how the 10–30-day and 30–90-day ISOs influence distinct TC clusters (with different genesis locations and trajectories) over the WNP and the ENSO effects on ISO–TC connection needed further examination, as addressed in this study. Specifically, we classified the WNP TCs into seven clusters using the fuzzy *c*-means clustering method (Fig. 1), consistent with previous studies (Camargo et al. 2007a, b; Kim et al. 2011; Zhang et al. 2016b; Qian et al. 2020). Based on a spectral analysis of TC genesis counts for each cluster, it was found that all seven clusters display significant variability at subseasonal timescales, with peak spectrums in the HF (10–30-day) and LF (30–90-day) ISO bands (Fig. 2). The probability of TC genesis for each cluster varies with the phase evolutions of the two ISO modes (Figs. 3 and 4). Generally, the TCs of each cluster are more likely to be generated when the ISO convective and low-pressure anomalies propagate into the main genesis regions of the specific cluster (Figs. 3 and 4).

To elucidate the modulating effects of the two ISO modes on the genesis of different clusters of TCs, the scale-decomposed GPI anomalies were diagnosed (Fig. 6), because the intraseasonal GPI distribution can capture the genesis patterns of individual clusters (Fig. 5). The diagnostic results suggest that the ISO-related relative humidity is the leading factor contributing to the intraseasonal GPI anomalies for all seven clusters. The ISO-related low-level absolute vorticity plays a secondary role. Indeed, these two dominant factors have been documented as important for WNP basin-wide TC genesis in previous studies (Zhao et al. 2015). In addition to the process of TC genesis, we also examined the influences of ISO on the movement of TCs for individual clusters. Accompanied by the intraseasonal cyclonic anomalies over the TC genesis locations, the background steering flows associated with different clusters vary geographically (Fig. 7). For the straight-moving TCs (C1, C5, and C6), the ISO-related cyclonic flows appear over the SCS and the Philippine Sea in the low latitudes, leading to enhanced easterly flow related to the westward extension of the WNP subtropical high. In contrast, the eastward shift of the WNP subtropical high induced by the strengthened ISO cyclonic flow favors the recurving TCs of C4 and C7. The results based on background steering flows of different clusters clearly illustrate the ISO effect on TC tracks, supporting previous findings using a composite of relatively limited TC cases (Chen et al. 2009).

Considering the significant influence of ENSO on TC activities at interannual timescales (Wang and Chan 2002; Camargo and Sobel 2005), the different modulation by LF (Fig. 8) and HF (Fig. 9) ISO of the intraseasonal TC genesis in different ENSO phases should also be discussed. Although the total TC counts in the WNP during El Niño and La Niña years are similar, the TCs of individual clusters are significantly influenced by ENSO (Figs. 8 and 9). There are more TCs generated in the southeast quadrant of the WNP, such as those of C3, C4, and C6, during El Niño years compared with those of the same clusters during La Niña years. The modulations of HF and LF ISO in some specific phases on the total numbers of TCs over the WNP are different during El Niño and La Niña years, which are changed by the low-level winds associated with ENSO conditions. The modulations of HF and LF ISOs on the TCs generated in the southeast quadrant of the WNP, such as the TCs of C3, C4, and C6, are relatively weaker during La Niña years compared with El Niño years. The modulations of ISOs on landfalling TCs (C5 and C7) are weaker during El Niño years.

Based on these diagnostic results, we developed linear regression models using the ISO-related large-scale fields as predictors for predicting anomalous TC genesis counts (relative to smoothed climatological 10-day-mean TC counts) every 10 days for each cluster and each ENSO phase (Fig. 10). The prediction skill for TC genesis counts, as assessed by the TCC, suggests that the statistical model has the capability of predicting anomalous TC genesis numbers at the lead time of 30 days, which is superior to the prediction skills of models that do not consider the effects of ENSO (Fig. 14). Predicted probability maps of TCF were then constructed by considering the climatological probability distribution of each cluster (Kim et al. 2012). The predicted distributions of TCF (Fig. 15) should provide useful information for disaster prevention and mitigation.

The predictability sources of these statistical TC forecast models are from the intraseasonal variabilities of large-scale fields associated with TC genesis (GPI components), while other potential sources of predictability, such as high-frequency equatorial waves (Frank and Roundy 2006; Schreck et al. 2012; Lai et al. 2020) and extratropical–tropical interaction (Camargo et al. 2018), need more exploration, which could potentially lead to improvements in TC subseasonal forecasting. In addition, not only the TC genesis and frequency distribution over the WNP, but also TC forecast products, including TC intensity and ACE, need further exploration.

Funding: This work was supported by the National Natural Science Foundation of China (42088101, 42205024), the National Postdoctoral Program for Innovative Talent (BX2021133), the China Postdoctoral Science Foundation of No. 70 General Fund (2021M701753), and the Natural Science Foundation of Jiangsu Province, China (BK20220459). We acknowledge the High-Performance Computing Center of Nanjing University of Information Science & Technology for their support of this work.

Author contributions: Y.Q. performed the research and analyzed the data; P.-C.H. designed the research; Y.Q. and P.-C.H. wrote the manuscript; H.M., J.G., H.W. and M.D. revised the manuscript.

Data Availability: The Reginal Specialized Meteorological Centers Tokyo-Typhoon Center best-track dataset is available at https://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/trackarchives.html. All reanalysis data used in this study were obtained from publicly available sources. The ECMWF ERA-Interim product is available at https://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=pl/. The NOAA High-resolution Blended Analysis of daily SST can be found at https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html. The real-time

NMME forecast Niño-3.4 index is available at https://ftp.cpc.ncep.noaa.gov/NMME/realtime_anom/ENSMEAN/.

Declarations

Ethical approval: Not applicable to this article.

Conflict of interest: The authors declare that they have no conflicts of interest.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans on Automat Contr 19:716–723. https://doi.org/10.1109/TAC.1974.1100705

Bell GD, Halpert MS, Schnell RC et al (2000) The 1999 North Atlantic and eastern North Pacific hurricane season [in "Climate Assessment for 1999"]. Bull Amer Meteor Soc 81:S19–S22.

Bezdek JC (1981) Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic, 256 pp.

Bister M, Emanuel KA (1998) Dissipative heating and hurricane intensity. Meteorol Atmos Phys 65:233–240. https://doi.org/10.1007/BF01030791

Camargo SJ, Barnston AG (2009) Experimental dynamical seasonal forecasts of tropical cyclone activity at IRI. Weather Forecast 24:472–491. https://doi.org/10.1175/2008WAF2007099.1

Camargo SJ, Camp J, Elsberry RL, et al (2018) Tropical cyclone prediction on subseasonal time-scales. Tropical Cyclone Research and Review 8:150–165. https://doi.org/10.6057/2019TCRR03.04

Camargo SJ, Robertson AW, Gaffney SJ, et al (2007a) Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. J Clim 20:3654–3676. https://doi.org/10.1175/JCLI4203.1

Camargo SJ, Robertson AW, Gaffney SJ, et al (2007b) Cluster analysis of typhoon tracks. Part I. General properties. J Clim 20:3635–3653. https://doi.org/10.1175/JCLI4188.1

Camargo SJ, Wheeler MC, Sobel AH (2009) Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. J Atmos Sci 66:3061–3074. https://doi.org/10.1175/2009JAS3101.1

Chen T-C, Wang S-Y, Yen M-C, Clark AJ (2009) Impact of the intraseasonal variability of the western North Pacific large-scale circulation on tropical cyclone tracks. Weather Forecast 24:646–666. https://doi.org/10.1175/2008WAF2222186.1

Chia H-H, Ropelewski CF (2002) The interannual variability in the genesis location of tropical cyclones in the northwest Pacific. J Clim 15:2934–2944. https://doi.org/10.1175/1520-0442(2002)015<2934:TIVITG>2.0.CO;2

Chu P-S, Zhao X, Ho C-H, et al (2010) Bayesian forecasting of seasonal typhoon activity: A track-pattern-oriented categorization approach. J Clim 23:6654–6668. https://doi.org/10.1175/2010JCLI3710.1

Dee DP, Uppala SM, Simmons AJ, et al (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597. https://doi.org/10.1002/qj.828

Duchon CE (1979) Lanczos Filtering in One and Two Dimensions. J. Appl. Meteorol. 18:1016–1022

Elsberry RL, Jordan MS, Vitart F (2010) Predictability of Tropical Cyclone Events on Intraseasonal Timescales with the ECMWF Monthly Forecast Model. Asia-Pacific J Atmos Sci 46:135–153. https://doi.org/10.1007/s13143-010-0013-4

Emanuel K (1995) Sensitivity of tropical cyclones to surface exchange coefficients and revised steady-state model incorporating eye dynamics. J. Atmos. Sci. 52:3969–3976

Frank WM, Roundy PE (2006) The role of tropical waves in tropical cyclogenesis. Mon Wea Rev, 134:2397–2417. https://doi.org/10.1175/MWR3204.1.

Gao J, Li T (2011) Factors controlling multiple tropical cyclone events in the western North Pacific. Mon Wea Rev 139: 885–894. https://doi.org/10.1175/2010MWR3340.1

Gao J, Li T (2012) Interannual variation of multiple tropical cyclone events in the western North Pacific. Adc Atmos Sci, 29:1279–1291.

https://doi.org/10.1007/s00376-012-1031-1.

Gray WM (1968) Global View of the Origin of Tropical Disturbances and Storms. Mon Wea Rev 96:669–700. https://doi.org/10.1175/1520-0493(1968)096<0669:gvotoo>2.0.co;2

Gray WM, Landsea C, Mielke P, Berry K (1994) Predicting Atlantic Basin Seasonal Tropical Cyclone Activity by 1 June. Weather Forecast 9:103–115

Gray WM, Landsea CW, Mielke PW, Berry KJ (1992) Predicting Atlantic seasonal hurricane activity 6–11 months in advance. Weather Forecast 7:440–455

Gray WM, Landsea CW, Mielke PW, Berry KJ (1993) Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. Weather Forecast 8:73–86

Han X, Zhao H, Li X, Raga GB, Wang C (2020) Modulation of boreal extended summer tropical cyclogenesis over the northwest Pacific by the quasi-biweekly oscillation under different El Niño-southern oscillation phases. International Journal of Climatology 40(2):858–873

Hsu P-C, Li T, Tsou C-H (2011) Interactions between boreal summer intraseasonal oscillations and synoptic-scale disturbances over the western North Pacific. Part I: Energetics diagnosis. J Clim 24:927–941. https://doi.org/10.1175/2010JCLI3833.1

Hsu P-C, Li T, You L, et al (2015) A spatial-temporal projection model for 10–30 day rainfall forecast in South China. Clim Dyn 44:1227–1244. https://doi.org/10.1007/s00382-014-2215-4

Jiang X, Xiang B, Zhao M, et al (2018) Intraseasonal tropical cyclogenesis prediction in a global coupled model system. J Clim 31:6209–6227. https://doi.org/10.1175/JCLI-D-17-0454.1

Jiang X, Zhao M, Waliser DE (2012) Modulation of tropical cyclones over the Eastern Pacific by the intraseasonal variability simulated in an AGCM. J Clim 25:6524–6538. https://doi.org/10.1175/JCLI-D-11-00531.1

Kim H-S, Ho C-H, Kim J-H, Chu P-S (2012) Track-pattern-based model for seasonal prediction of tropical cyclone activity in the western north Pacific. J Clim 25:4660–4678. https://doi.org/10.1175/JCLI-D-11-00236.1

Kim H-S, Kim J-H, Ho C-H, Chu P-S (2011) Pattern classification of typhoon tracks using the fuzzy *c*-means clustering method. J Clim 24:488–508. https://doi.org/10.1175/2010JCLI3751.1

Kim J-H, Ho C-H, Kim H-S, et al (2008) Systematic variation of summertime tropical cyclone activity in the western North Pacific in relation to the Madden-Julian oscillation. J Clim 21:1171–1191. https://doi.org/10.1175/2007JCLI1493.1

Kirtman BP., Min D, Infanti JM (2014) The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. *Bull. Amer. Meteor. Soc.*, 95: 585–601. https://doi.org/10.1175/BAMS-D-12-00050.1

Klotzbach PJ (2008) Refinements to Atlantic basin seasonal hurricane prediction from 1 December. J Geophys Res Atmos 113:1–11. https://doi.org/10.1029/2008JD010047

Klotzbach PJ, Gray WM (2004) Updated 6–11-month prediction of Atlantic basin seasonal hurricane activity. Weather Forecast 19:917–934. https://doi.org/10.1175/1520-0434(2004)019<0917:UMPOAB>2.0.CO;2

Klotzbach PJ, Gray WM (2009) Twenty-five years of Atlantic basin seasonal hurricane forecasts (1984–2008). Geophys Res Lett 36:1–5. https://doi.org/10.1029/2009GL037580

Knapp KR, Kruk MC, Levinson DH, et al (2010) The international best track archive for climate stewardship (IBTrACS): Unifying tropical cyclone best track data. Bull Amer Meteor Soc 91:363–376. https:// doi.org/10.1175/2009BAMS2755.1

Lai Q, Gao J, Zhang W, Guan X (2020) Influences of the equatorial waves on multiple tropical cyclone genesis over the western North Pacific. Terr Atmos Ocean Sci 31:227–238. https://doi.org/10.3319/ TAO.2020.03.20.01

Lee C-Y, Camargo SJ, Vitart F, et al (2018) Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. Weather Forecast 33:967–988. https://doi.org/10.1175/WAF-D-17-0165.1

Leroy A, Wheeler MC (2008) Statistical prediction of weekly tropical cyclone activity in the southern hemisphere. Mon Wea Rev 136:3637–3654.

https://doi.org/10.1175/2008MWR2426.1

Li RCY, Zhou W, Chan JCL (2012) Asymmetric Modulation of Western North Pacific Cyclogenesis by the Madden–Julian Oscillation under ENSO Conditions. J Clim 25:5374–5385. https://doi.org/ 10.1175/JCLI-D-11-00337.1

Li RCY, Zhou W (2013a) Modulation of western north pacific tropical cyclone activity by the ISO. Part I: Genesis and intensity. J Clim 26:2904–2918. https://doi.org/10.1175/JCLI-D-12-00210.1

Li RCY, Zhou W (2013b) Modulation of western north pacific tropical cyclone activity by the ISO. Part II: Tracks and landfalls. J Clim 26:2919–2930. https://doi.org/10.1175/JCLI-D-12-00211.1

Liebmann B, Hendon H, Glick D (1994) The relationship between tropical cyclones of the 2estern Pacific and Indian Oceans and the Madden-Julian oscillation. J Meteorol Soc Japan 72:401–411

McBride JL (1995) Tropical cyclone formation. In: Russel Elsberry (ed) Global Perspective on Tropical Cyclones, Geneva. World Meteorological Organization Rep. TCP-38, WMO/TD-No 693, pp 63–105

Murakami H, Vecchi GA, Delworth TL, et al (2017) Dominant role of subtropical pacific warming in extreme Eastern Pacific hurricane seasons: 2015 and the future. J Clim 30:243–264. https://doi.org/10.1175/JCLI-D-16-0424.1

Murakami H, Vecchi GA, Underwood SD, et al (2015) Simulation and prediction of category 4 and 5 hurricanes in the high-resolution GFDL HiFLOR coupled climate model. J Clim 28:9058–9079. https://doi.org/10.1175/JCLI-D-15-0216.1

Murakami H, Vecchi GA, Villarini G, et al (2016a) Seasonal forecasts of major hurricanes and landfalling tropical cyclones using a high-resolution GFDL coupled climate model. J Clim 29:7977–7989. https://doi.org/10.1175/JCLI-D-16-0233.1

Murakami H, Villarini G, Vecchi GA, et al (2016b) Statistical-dynamical seasonal forecast of north atlantic and U.S. landfalling tropical cyclones using the high-resolution GFDL FLOR coupled model. Mon Wea Rev 144:2101–2123. https://doi.org/10.1175/MWR-D-15-0308.1

Murakami H, Wang B (2010) Future change of North Atlantic tropical cyclone tracks: Projection by a 20-km-mesh global atmospheric model. J Clim 23:2699–2721. https://doi.org/10.1175/2010JCLI3338.1

Nakano M, Kubota H, Miyakawa T, Satoh M (2017) Genesis of super cyclone Pam (2015): modulation of low-frequency large-scale circulations and the Madden-Julian oscillation by sea surface temperature anomalies. Mon Wea Rev 145: 3143–3159. https://doi.org/10.1175/MWR-D-16-0208.1

Pielke RA, Gratz J, Landsea CW, et al (2008) Normalized hurricane damage in the United States: 1900–2005. Nat Hazards Rev 9:29–42. https://doi.org/10.1061/(ASCE)1527-6988(2008)9:1(29)

Qian Y, Hsu PC, Murakami H, Xiang B, You L, (2020) A hybrid dynamical-statistical model for advancing subseasonal tropical cyclone prediction over the western North Pacific. Geophysical Research Letters, 47, e2020GL090095. https://doi.org/10.1029/2020GL090095

Reynolds, RW, Smith TM, Liu C, et al (2007) Daily high-resolution-blended analyses for sea surface temperature. J Climate 20:5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Satoh M, Oouchi K, Nasuno T, et al (2012) The Intra-Seasonal Oscillation and its control of tropical cyclones simulated by high-resolution global atmospheric models. Clim Dyn 39:2185–2206. https://doi.org/10.1007/s00382-011-1235-6

Schreck CJ, Molinari J, Aiyyer A (2012) A global view of equatorial waves and tropical cyclogenesis. Mon Wea Rev 140: 774–788. https://doi.org/10.1175/MWR-D-11-00110.1

Slade SA, Maloney ED (2013) An intraseasonal prediction model of atlantic and east Pacific tropical cyclone genesis. Mon Wea Rev 141:1925–1942. https://doi.org/10.1175/MWR-D-12-00268.1

Smith AB, Katz RW (2013) US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. Nat Hazards 67:387–410. https://doi.org/10.1007/s11069-013-0566-5

Vitart F, Leroy A, Wheeler MC (2010) A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. Mon Wea Rev 138:3671–3682. https://doi.org/10.1175/2010MWR3343.1

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

Wei X, Yang Y, Chen L (2021) The extended-range forecast of tropical cyclogenesis over the South China Sea based on the intraseasonal oscillation. Meteorology and Atmospheric Physics 133:1577–1589. https://doi.org/10.1007/s00703-021-00830-0

Xiang B, Zhao M, Jiang X, et al (2015) The 3–4-week MJO prediction skill in a GFDL coupled model. J Clim 28:5351–5364. https://doi.org/10.1175/JCLI-D-15-0102.1

Yoshida R, Kajikawa Y, Ishikawa H (2014) Impact of boreal summer intraseasonal oscillation on environment of tropical cyclone genesis over the western North Pacific. SOLA 10: 15–18. https://doi.org/10.2151/sola.2014-004

You L, Gao J, Lin H, Chen S (2018) Impact of the intra-seasonal oscillation on tropical cyclone genesis over the western North Pacific. Int J Climatol 39: 1969–1984. https://doi.org/10.1002/joc.5927

Zhang W, Gao J, Cheung KKW, et al (2019) Intraseasonal variability of the genesis potential index and its relationship with tropical cyclogenesis over the western North Pacific. Int J Climatol 40: 22–35. https://doi.org/10.1002/joc.6190

Zhang W, Vecchi GA, Murakami H, et al (2016a) The Pacific meridional mode and the occurrence of tropical Cyclones in the western North Pacific. J Clim 29:381–398. https://doi.org/10.1175/JCLI-D-15-0282.1

Zhang W, Villarini G, Vecchi GA, et al (2016b) Statistical-dynamical seasonal forecast of western North Pacific and East Asia landfalling tropical cyclones using the high-resolution GFDL FLOR coupled model. J Adv Model Earth Syst 8:1289–1309. https://doi.org/10.1002/2013MS000282.

Zhao H, Jiang X, Wu L (2015) Modulation of northwest Pacific tropical cyclone genesis by the intraseasonal variability. J Meteorol Soc Japan 93:81–97. https://doi.org/10.2151/jmsj.2015-006

Zhao H, Lu Y, Jiang X, Klotzbach PJ, Wu L, Cao, J (2022) A Statistical Intraseasonal Prediction Model of Extended Boreal Summer Western North Pacific Tropical Cyclone Genesis. J Clim 35:2460–2478. https://doi.org/ 10.1175/JCLI-D-21-0110.1

Zhao J, Zhan R, Wang Y, et al (2018) Contribution of the Interdecadal Pacific Oscillation to the recent abrupt decrease in tropical cyclone genesis frequency over the western North Pacific since 1998. J Clim 31:8211–8224. https://doi.org/10.1175/JCLI-D-18-0202.1

Figure Captions

Figure 1 (a–g) Seven TC clusters (C1–C7) derived from the fuzzy *c*-means clustering method using (h) all WNP TC cases (TCall) during the TC season of 1982–2018. The red curve represents the central trajectory of TCs. The number and percentage of TCs for each cluster is listed in parentheses.

Figure 2 Spectral analysis of daily TC counts of (a–g) individual clusters and (h) all TC cases over the WNP during the TC seasons of 1982–2018 (unit: day^{-2}). The dashed red and blue line represents the significance test of red noise and its 90% confidence level. Red (blue) shading in the lower part of each panel represents the ratio between the spectral power and its red noise when the former is larger (smaller) than the latter, highlighting the significant (non-significant) signals.

Figure 3 (a–h) Composites of 10–30-day OLR (shading; W m^{-2}) and 500-hPa geopotential height (contours; interval: 2 hPa) anomalies in the eight phases of active HF ISO [$(PC1^2+PC2^2)^{1/2} \geq 1$]. TC genesis (red dots) of (i–o) C1–C7 and (p) TCall during each phase of HF ISO phase indices. The probability of TC genesis in each phase [excluding the days of weak ISO phase, $(PC1^2+PC2^2)^{1/2} < 1$] is shown in blue, where the dark blue number indicates the change in TC genesis counts is statistically significant at the 95% confidence level using the Monte Carlo method, in which random ISO phases were assigned to TC genesis for a large number of times (1000). If the probability of TC genesis occurrence for a certain ISO phase is larger (smaller) than the 97.5% (2.5%) percentile of the random distribution generated by 1000 simulations, it is considered statistically significant.

Figure 4 As in Fig. 3 except for the phase evolution of LF ISO and its association with TC genesis for individual clusters.

Figure 5 Composites of HF (10–30-day filtered) GPI anomalies based on TC generation days for (a–g) C1–C7 and (h) TCall during the TC seasons of 1982–2018. Panels (i–p) and (q–x) are the same as (a–h) except for the LF (30–90-day filtered) and 10–90-day filtered GPI anomalies. Blue dots mark the TC genesis locations.

Figure 6 Contributions of different terms (I_{GPI}, η, γ, ψ, s, ω) to 10–90-day GPI anomalies for (a–g) the TCs of seven clusters and (h) all TCs. These terms are composited over a 6° × 6° box centered at the genesis locations of TCs during the days of TC genesis. Black (grey) bars denote positive (negative) contributions.

Figure 7 Composites of 10–90-day filtered steering flow (red vectors; unit: m s^{-1}), total (non-filtered) steering flow (light blue streamlines) and the 5880-gpm contour (thick blue contour) based on TC genesis days and the next four days for (a–g) C1–C7. The steering flow in (h) represents the climatological-mean field during the TC seasons of 1982–2018. Steering flow is defined as the pressure-weighted vertically averaged horizontal winds between 850 and 200 hPa. Only the anomalous field exceeding the 95% significance level is shown for the ISO-related steering flow (red vectors). Black curves mark the TC trajectories of (a–g) C1–C7 and (h) TCall.

Figure 8 TC genesis percentage (TCGP) anomalies (%) of (a–g) C1–C7 and (h) TCall in each phase of LF ISO in El Niño years (red bars) and La Niña years (blue bars) during 1982–2018. The TCGP of C1–C7 and TCall in El Niño (numbers in red) and La Niña (numbers in blue) years are displayed in the right-top corner of the subplots. (i–k) Composite TCGP (shading; %) and 30–90-day band-passed OLR anomalies (contours; interval: 6 W m^{-2}) in (i, m, q) phases 1 and 8, (j, n, r) phases 2–3, (k, o, s) phases 4–5, and (l, p, t) phases 6–7 of LF ISO during (i–l) El Niño, (m–p) La Niña, and (q–t) all years. TCGP anomalies are defined as the percentage of TC genesis count divided by the numbers of days for each phase of LF ISO and El Niño (La Niña) years compared with the climatological mean for each 5° × 5° grid box.

Figure 9 As in Fig. 8 except for TC count anomalies and large-scale fields in each phase of HF ISO.

Figure 10 Steps for constructing the subseasonal TC prediction model.

Figure 11 Climatological 10-day mean TC counts of (a–g) C1–C7 and (h) TCall. The yellow bars and thick black lines are the raw climatology and the slow annual cycle (only the mean and first three harmonics of the daily climatology retained) of years during 1982–2008.

Figure 12 Temporal correlation coefficients (TCCs; shading) between TC count anomalies of (a–d) C1 and GPI, and (e–h) C6 and MPI during (a, e) all years, (b, f) El Niño years, (c, g) La Niña years, and (d, h) neutral years of 1982–2008, at 0-day, 10-day, 15-day, 20-day, 25-day, 30-day, 35-day and 40-day leads (from top to bottom). Only the TCCs that are statistically significant at the 95% confidence level are shown. TC genesis locations are represented as green dots.

Figure 13 (a) TCCs and (c) RMSEs between observed and predicted TCall count anomalies. (b, d) As in (a, c) but for the sum of C1–C7. Red, blue, orange, green, gray, purple, and pink lines in (a–d) denote the forecast skills by using GPI, Vort, RH, MPI, VWS, OMG, and SST as predictors, respectively. Thick black lines represent the skills of the ensemble results. Dashed lines in (a–b) mark the 95% confidence level of Spearman rank order correlation coefficients.

Figure 14 TCCs between the observed and predicted (a) anomalous and (b) total TC count (sum of C1–C7 in blue, TCall in red) and (c) the RMSEs of total TC counts (open dot-dashed lines indicate the results of the sum of C1–C7; solid dot lines indicate the results of TCall) at different forecast lead times from 0 to 40 days (*x*-axis). The black line in (a) is the 95% significance level. The black line in (b) is the TCC skill of climatological prediction. Blue dotted lines represent the TCC prediction skill of clim-fcst statistical models.

Figure 5 Four cases of TC frequency (TCF) probability distributions during (a) 21–30 June 2012, (b) 10–19 Aug 2018, (c) 31 July–9 August 204, and (d) 9–18 September 2010. From top to bottom: observation, perfect reconstruction, and prediction at leads of 10, 15, 20, 25 and 30 days. The TCF probability at each grid point is defined as the TCF in each 5° × 5° box divided by the total TC count over the WNP. The red typhoon symbols and curves indicate the observed TC genesis locations and trajectories for each TC case. The pattern correlation coefficients between observations and predictions are shown in the upper-right corner of each panel.

Table 1 Basic characteristics of clusters C1–C7 and all TC cases (TCall).

	Genesis location (lon, lat)	Lifespan (days)	Trajectory length (km)	Maximum wind speed (m/s)	Accumulated cyclone energy (ACE) (m^2/s^2)
C1	113.2°E, 16.9°N	2.7	96.7	64.3	9078
C2	162.4°E, 27.0°N	4.0	142.0	74.6	19958
C3	150.3°E, 17.7°N	6.4	120.4	82.3	57438
C4	145.8°E, 26.4°N	5.9	140.1	82.3	34991
C5	123.1°E, 18.3°N	4.9	107.4	87.5	27683
C6	136.8°E, 17.7°N	7.5	123.7	87.5	63761
C7	132.2°E, 24.0°N	5.7	126.3	79.7	37486
TCall	133.4°E, 21.5°N	5.1	121.8	87.5	31102

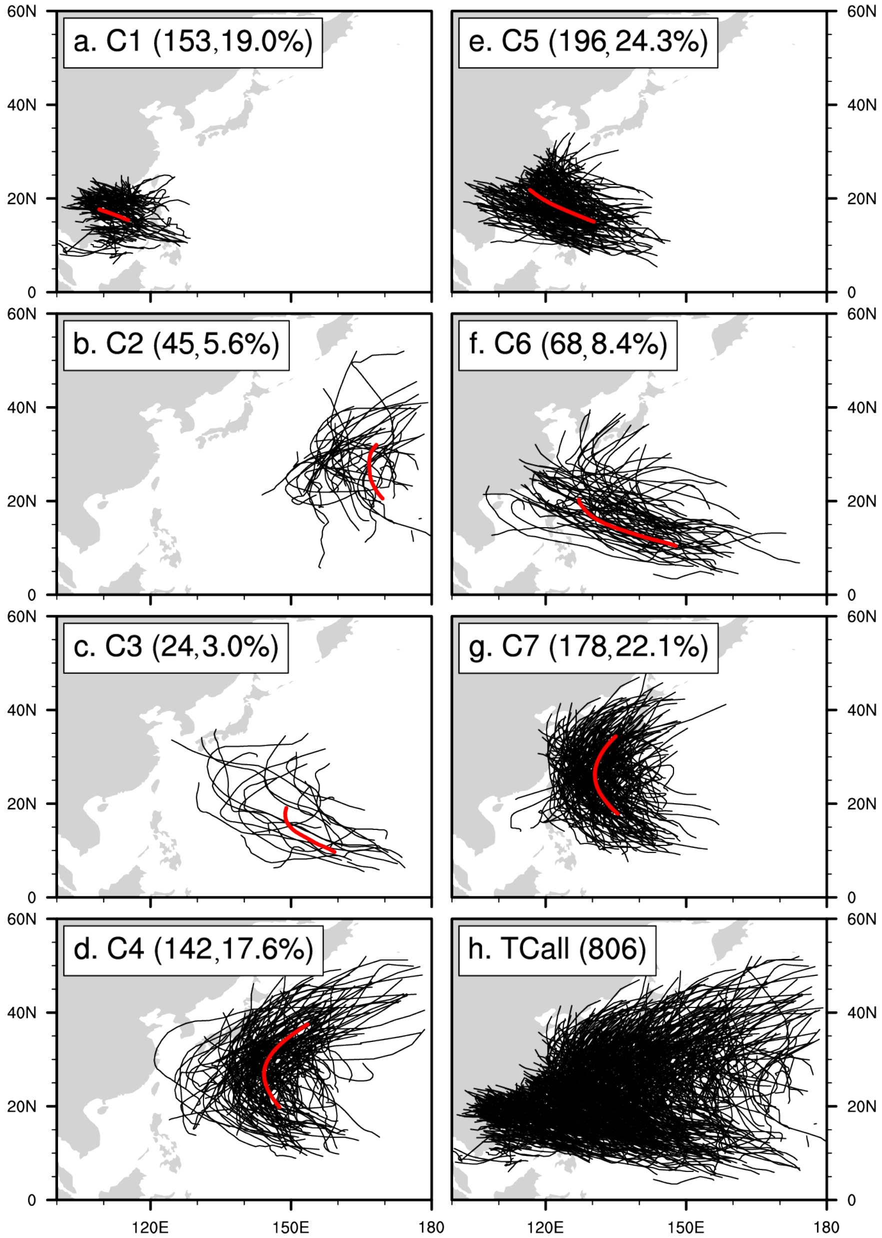


Fig. 1 (a–g) Seven TC clusters (C1–C7) derived from the fuzzy *c*-means clustering method using (h) all WNP TC cases (TCall) during the TC season of 1982–2018. The red curve represents the central trajectory of TCs. The number and percentage of TCs for each cluster is listed in parentheses.

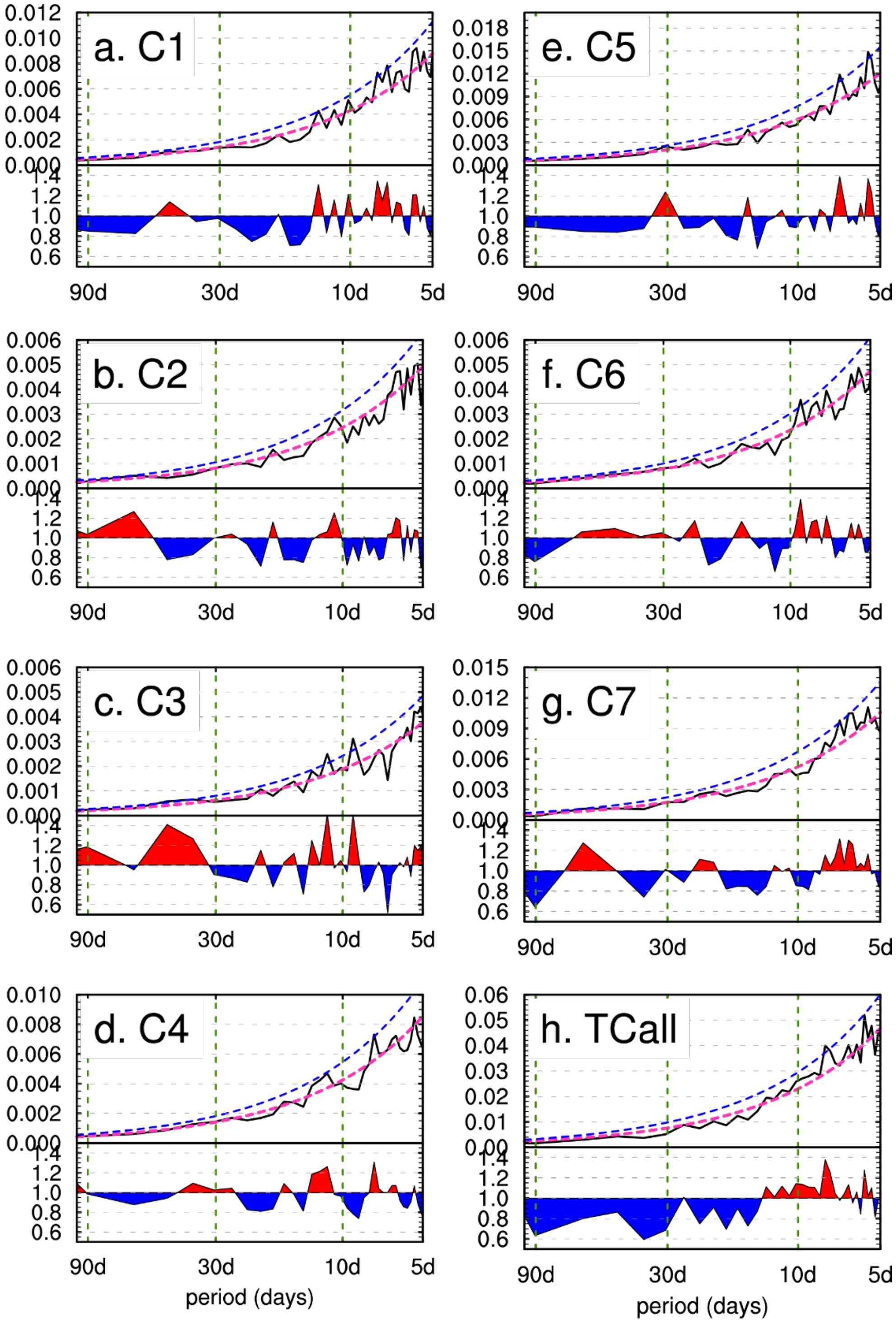


Fig. 2 Spectral analysis of daily TC counts of (a–g) individual clusters and (h) all TC cases over the WNP during the TC seasons of 1982–2018 (unit: day^{-2}). The dashed red and blue line represents the significance test of red noise and its 90% confidence level. Red (blue) shading in the lower part of each panel represents the ratio between the spectral power and its red noise when the former is larger (smaller) than the latter, highlighting the significant (non-significant) signals.

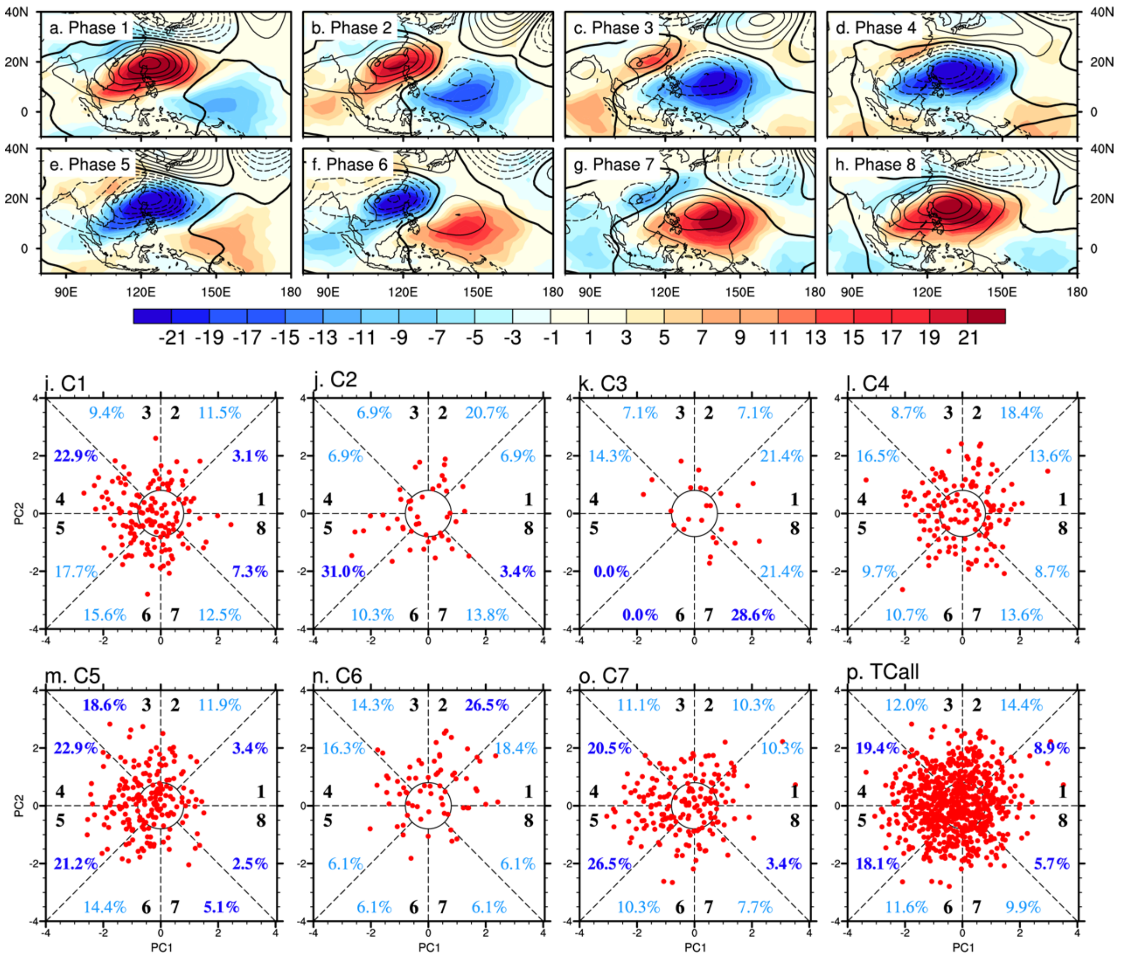


Fig. 3 (a–h) Composites of 10–30-day OLR (shading; W m^{-2}) and 500-hPa geopotential height (contours; interval: 2 hPa) anomalies in the eight phases of active HF ISO [$(PC1^2+PC2^2)^{1/2} \geq 1$]. TC genesis (red dots) of (i–o) C1–C7 and (p) TCall during each phase of HF ISO phase indices. The probability of TC genesis in each phase [excluding the days of weak ISO phase, $(PC1^2+PC2^2)^{1/2}<1$] is shown in blue, where the dark blue number indicates the change in TC genesis counts is statistically significant at the 95% confidence level using the Monte Carlo method, in which random ISO phases were assigned to TC genesis for a large number of times (1000). If the probability of TC genesis occurrence for a certain ISO phase is larger (smaller) than the 97.5% (2.5%) percentile of the random distribution generated by 1000 simulations, it is considered statistically significant.

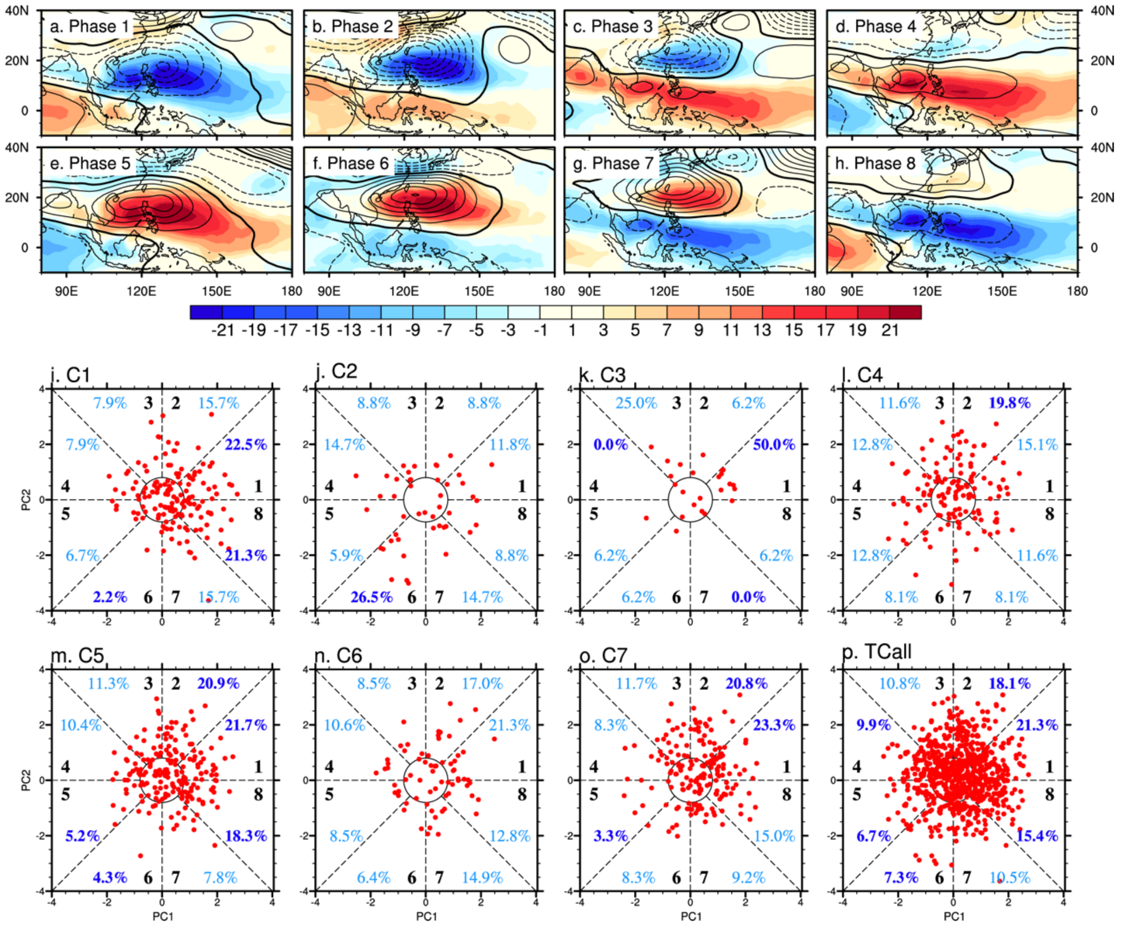
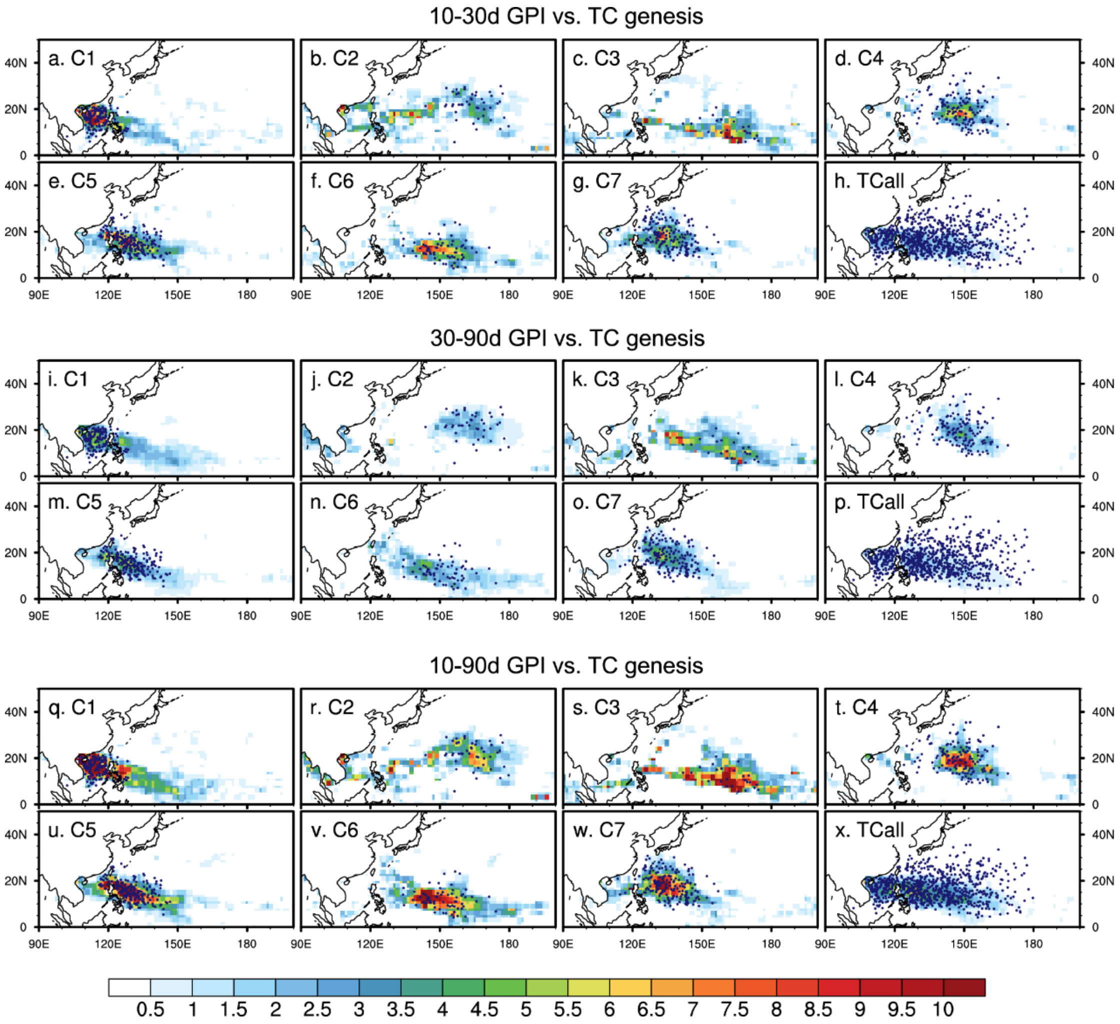


Fig. 4 As in Fig. 3 except for the phase evolution of LF ISO and its association with TC genesis for individual clusters.



1023 **Fig. 5** Composites of HF (10–30-day filtered) GPI anomalies based on TC generation
1024 days for (a–g) C1–C7 and (h) TCall during the TC seasons of 1982–2018. Panels (i–p)
1025 and (q–x) are the same as (a–h) except for the LF (30–90-day filtered) and 10–90-day
1026 filtered GPI anomalies. Blue dots mark the TC genesis locations.

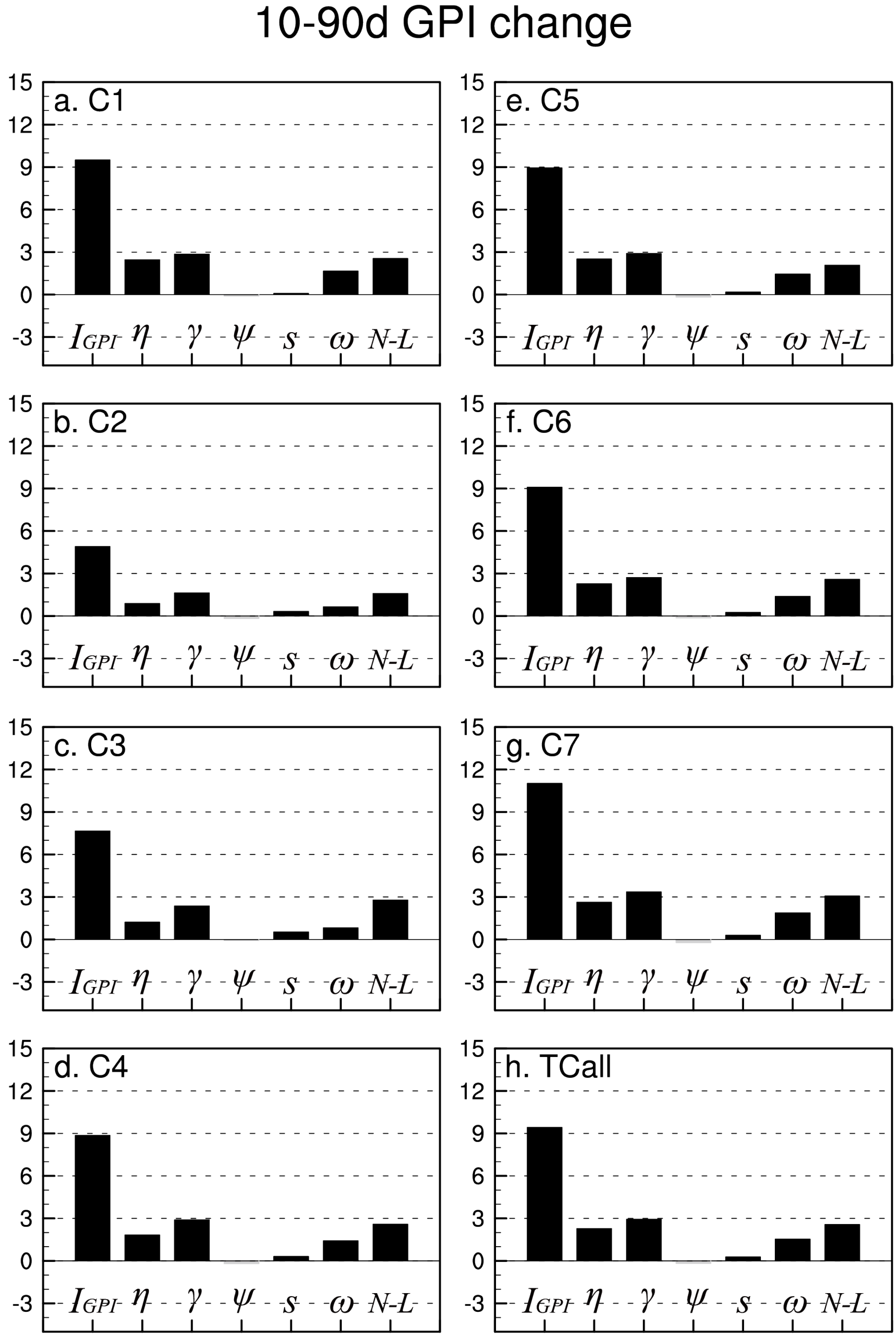


Fig. 6 Contributions of different terms (I_{GPI}, η, γ, ψ, s, ω) to 10–90-day GPI anomalies for (a–g) the TCs of seven clusters and (h) all TCs. These terms are composited over a 6° × 6° box centered at the genesis locations of TCs during the days of TC genesis. Black (grey) bars denote positive (negative) contributions.

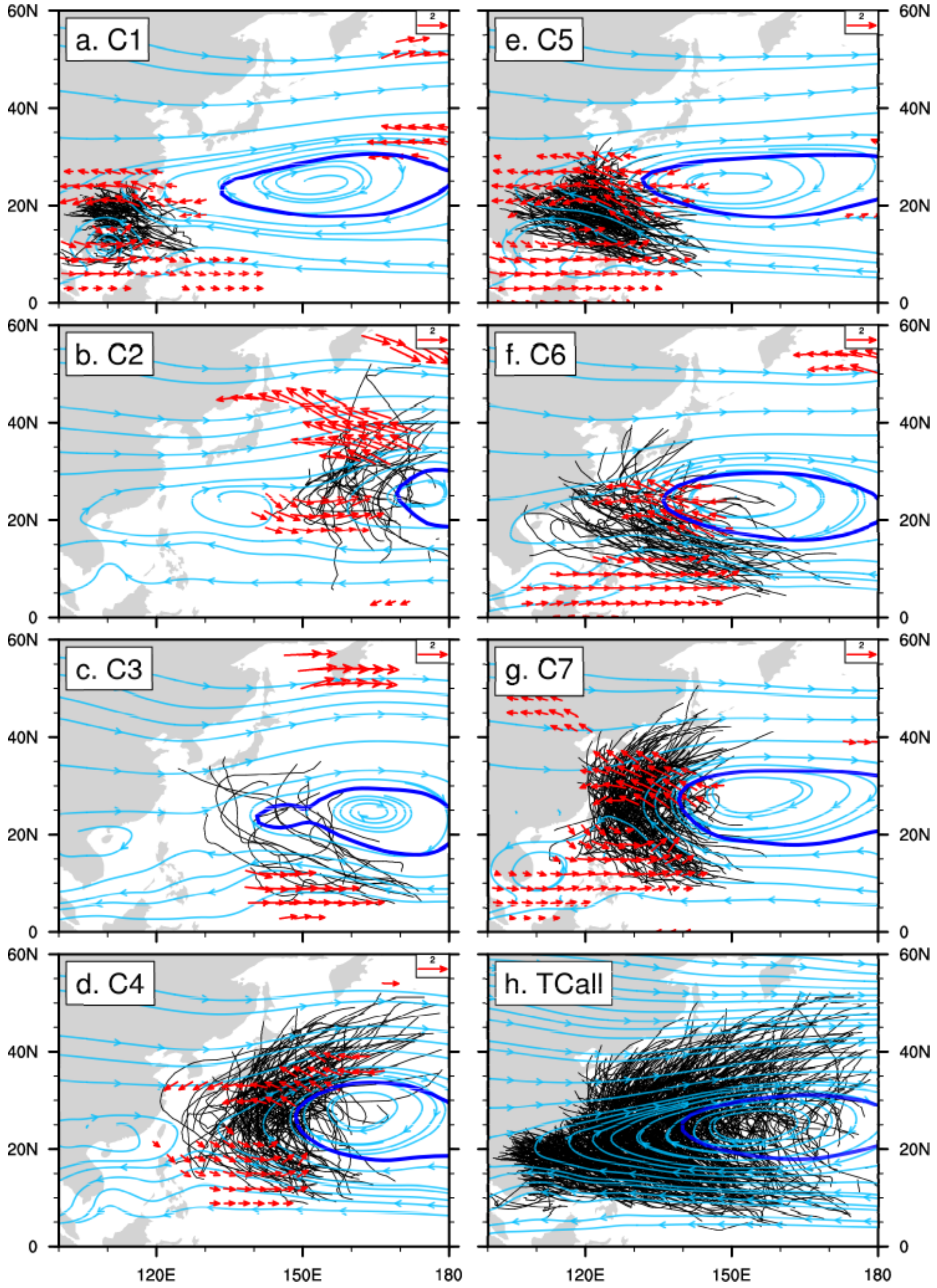


Fig. 7 Composites of 10–90-day filtered steering flow (red vectors; unit: m s^{-1}), total (non-filtered) steering flow (light blue streamlines) and the 5880-gpm contour (thick blue contour) based on TC genesis days and the next four days for (a–g) C1–C7. The steering flow in (h) represents the climatological-mean field during the TC seasons of 1982–2018. Steering flow is defined as the pressure-weighted vertically averaged horizontal winds between 850 and 200 hPa. Only the anomalous field exceeding the 95% significance level is shown for the ISO-related steering flow (red vectors). Black curves mark the TC trajectories of (a–g) C1–C7 and (h) TCall.

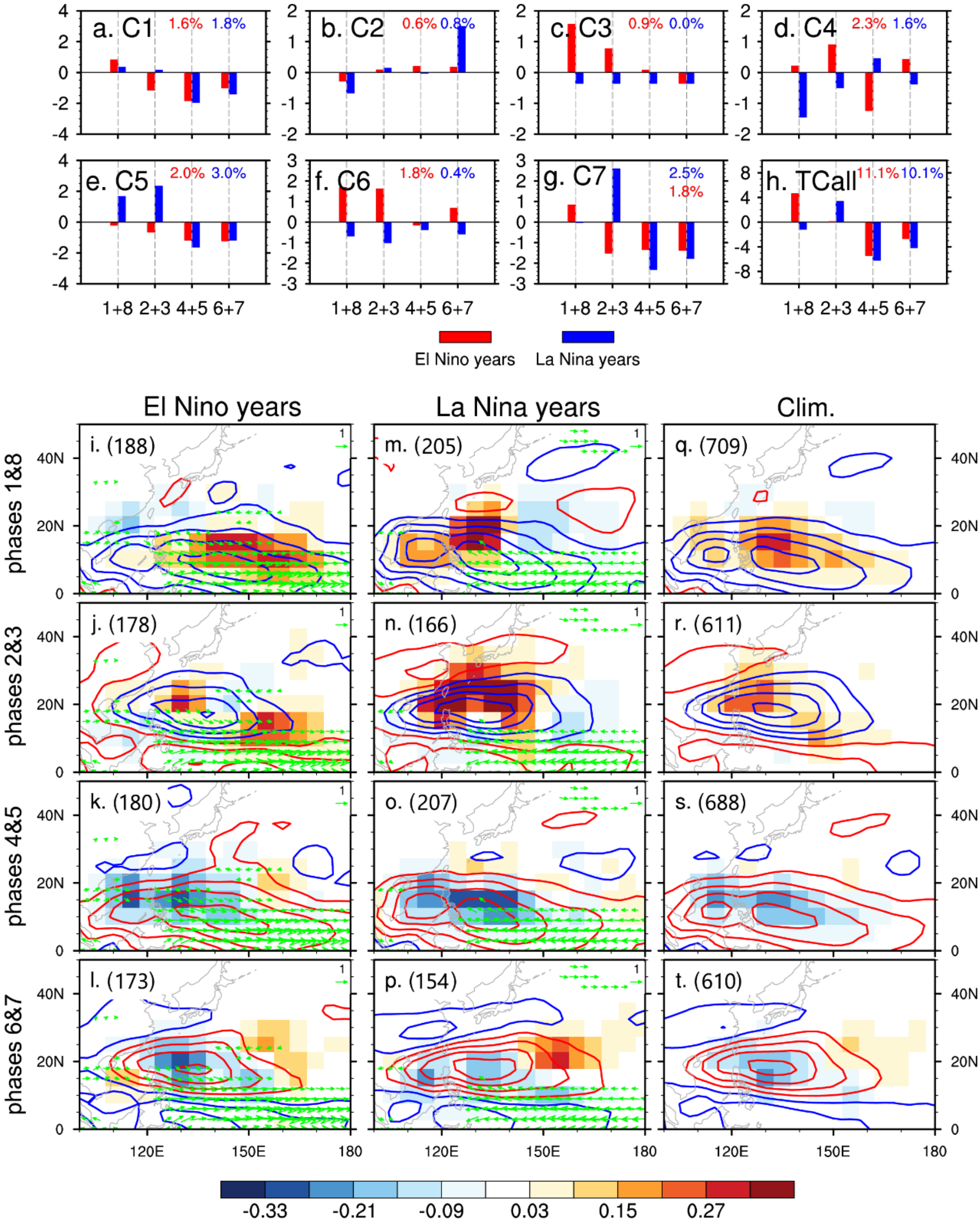


Fig. 8 TC genesis percentage (TCGP) anomalies (%) of (a–g) C1–C7 and (h) TCall in each phase of LF ISO in El Niño years (red bars) and La Niña years (blue bars) during 1982–2018. The TCGP of C1–C7 and TCall in El Niño (numbers in red) and La Niña (numbers in blue) years are displayed in the right-top corner of the subplots. (i–k) Composite TCGP (shading; %) and 30–90-day band-passed OLR anomalies (contours; interval: 6 W m^{-2}) in (i, m, q) phases 1 and 8, (j, n, r) phases 2–3, (k, o, s) phases 4–5, and (l, p, t) phases 6–7 of LF ISO during (i–l) El Niño, (m–p) La Niña, and (q–t) all

years. TCGP anomalies are defined as the percentage of TC genesis count divided by the numbers of days for each phase of LF ISO and El Niño (La Niña) years compared with the climatological mean for each 5° × 5° grid box.

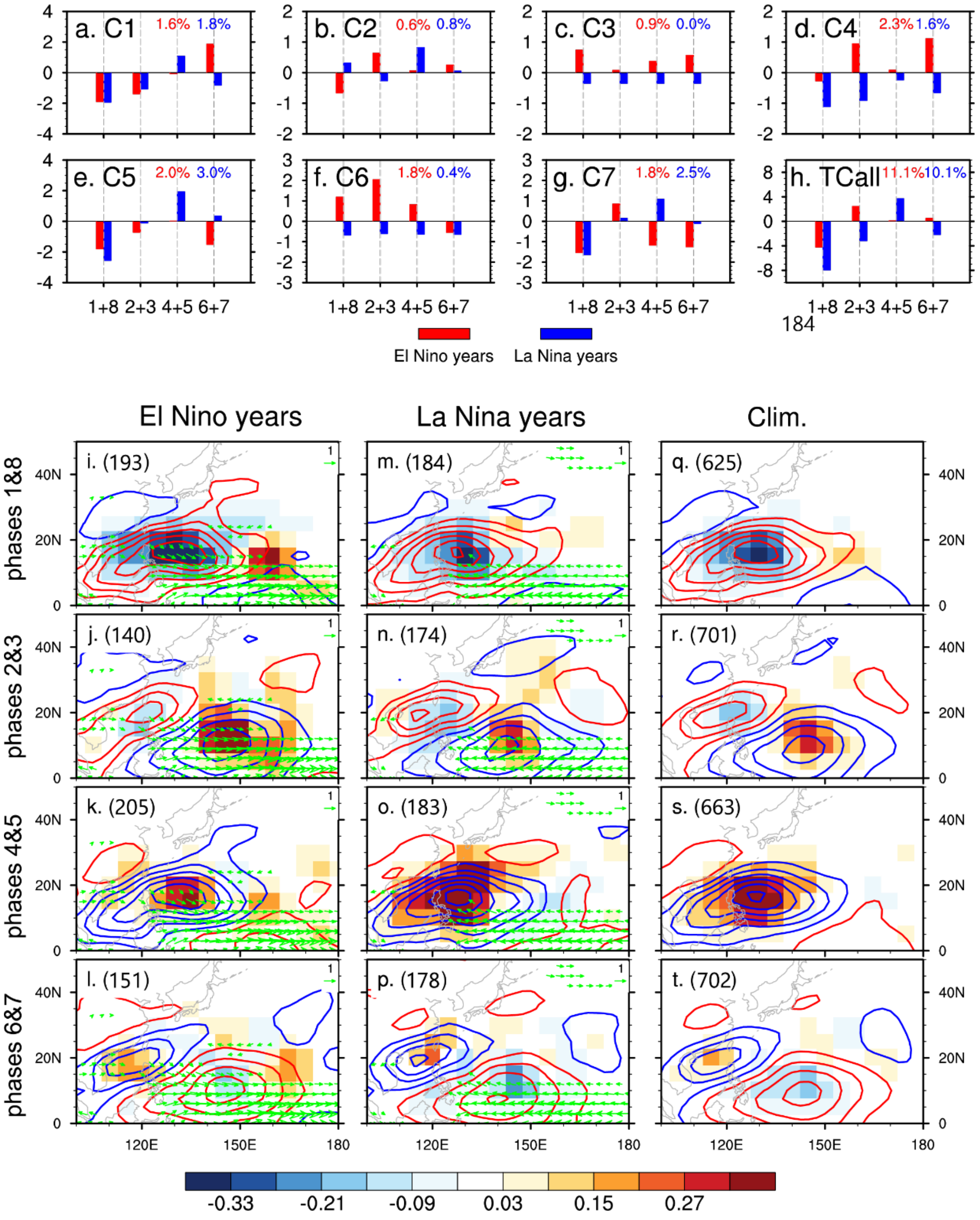
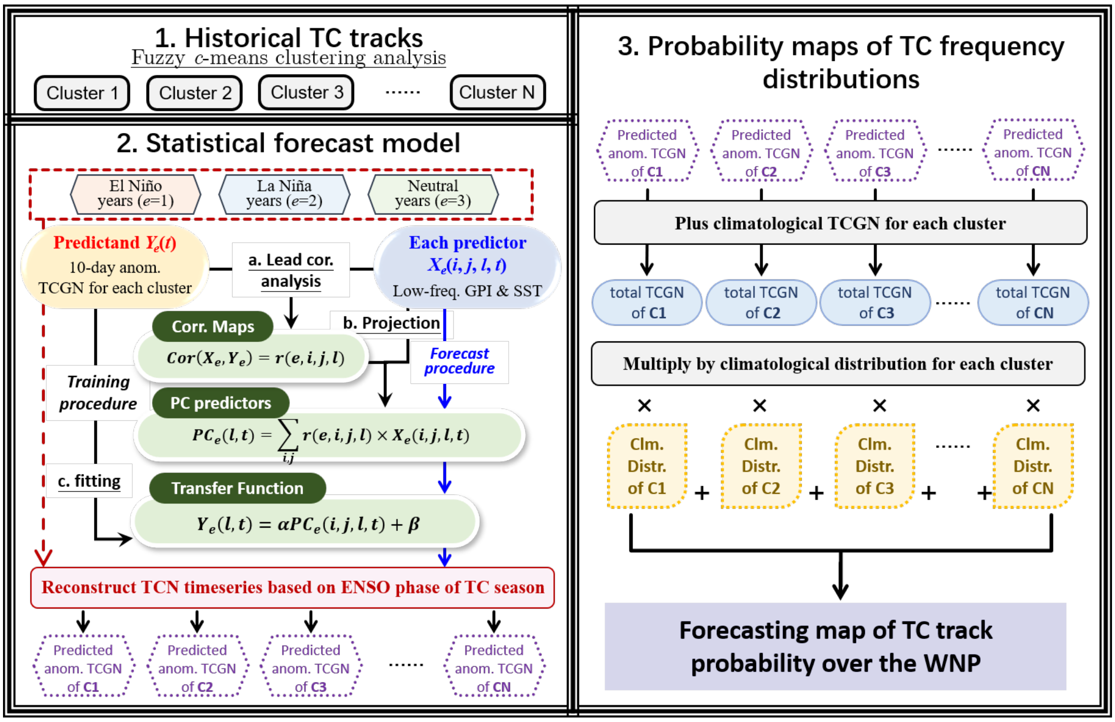


Fig. 9 As in Fig. 8 except for TC count anomalies and large-scale fields in each phase of HF ISO.



1051 **Fig. 10** Steps for constructing the subseasonal TC prediction model.

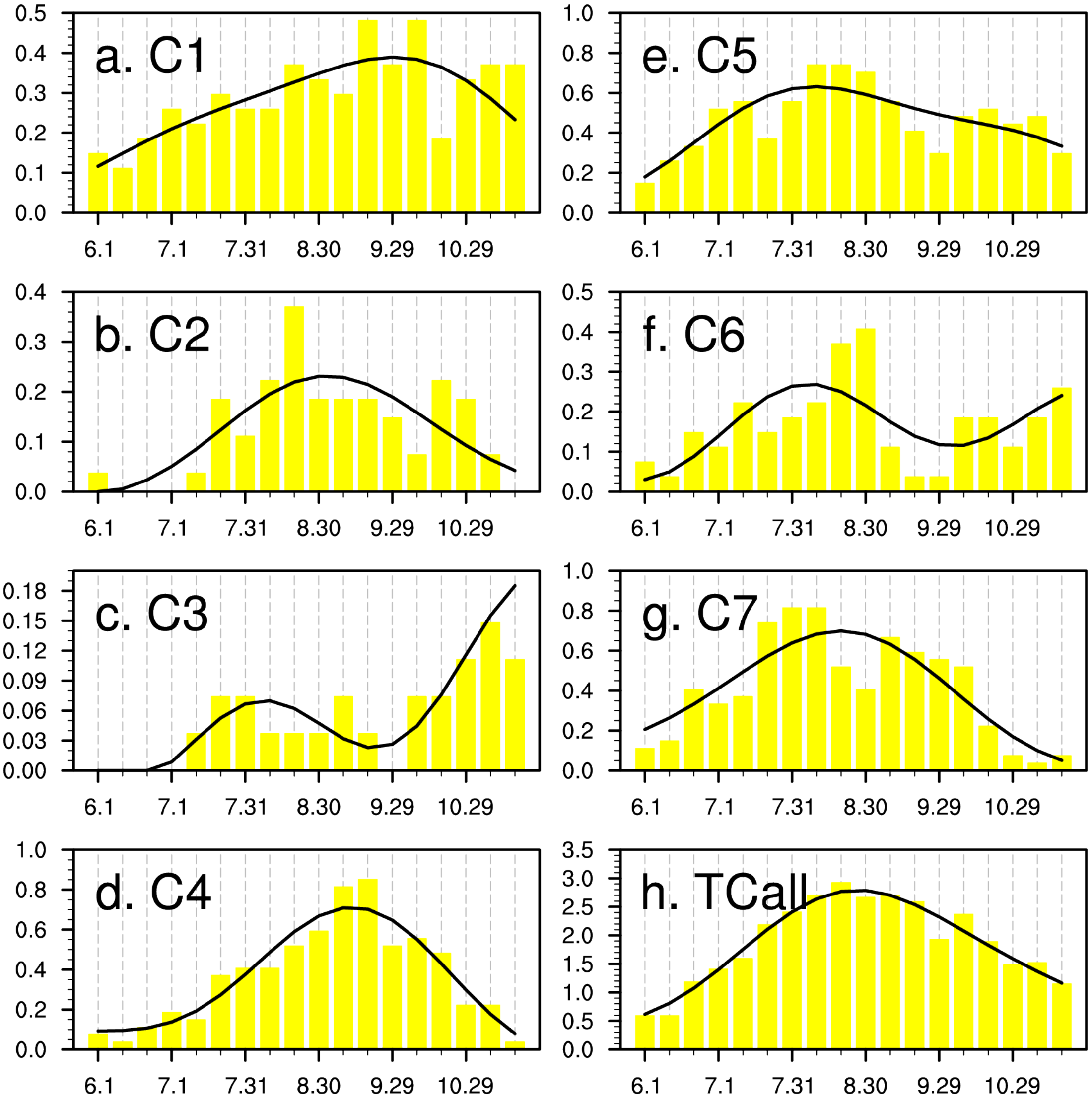


Fig. 11 Climatological 10-day mean TC counts of (a–g) C1–C7 and (h) TCall. The yellow bars and thick black lines are the raw climatology and the slow annual cycle (only the mean and first three harmonics of the daily climatology retained) of years during 1982–2008.

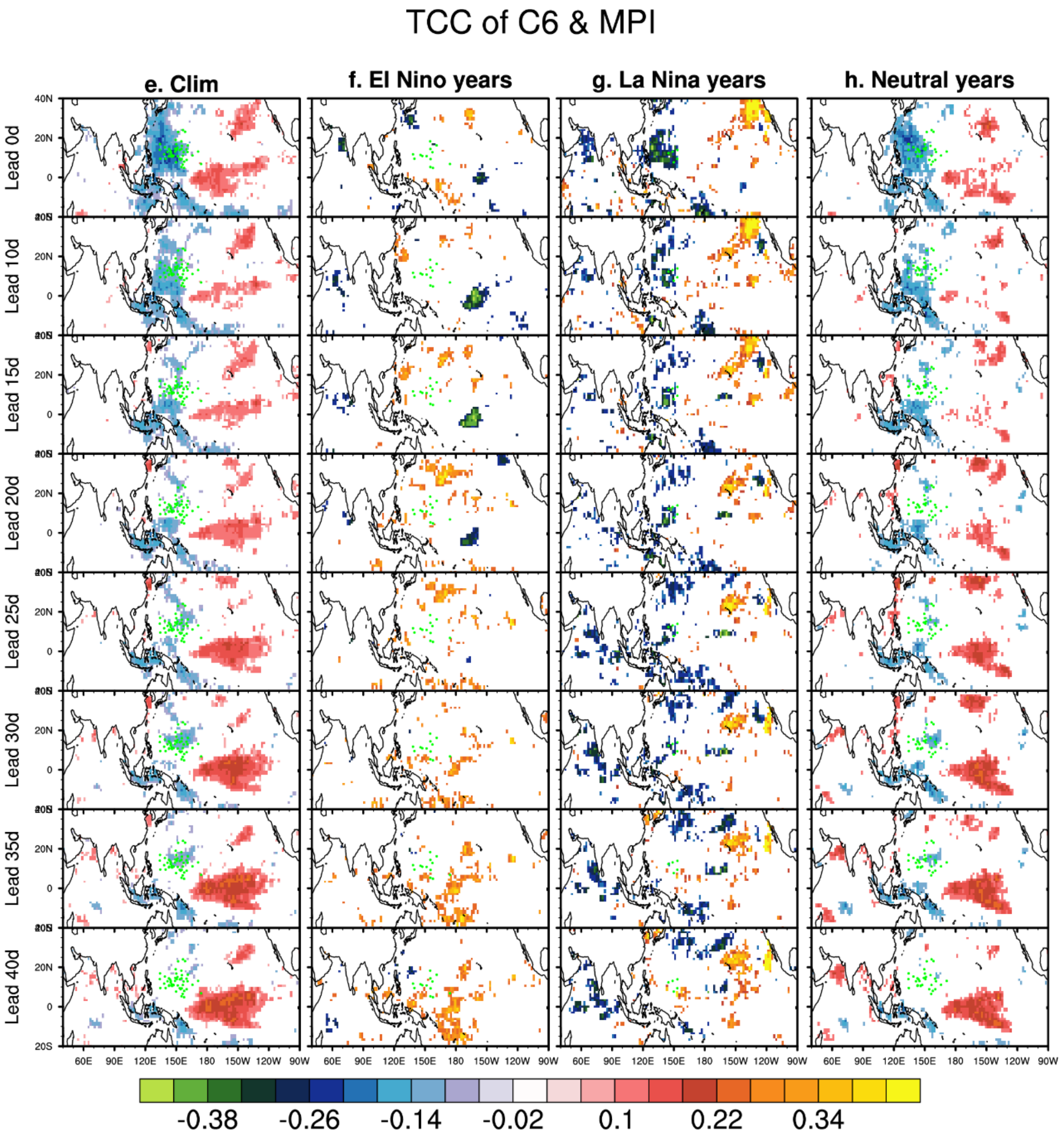
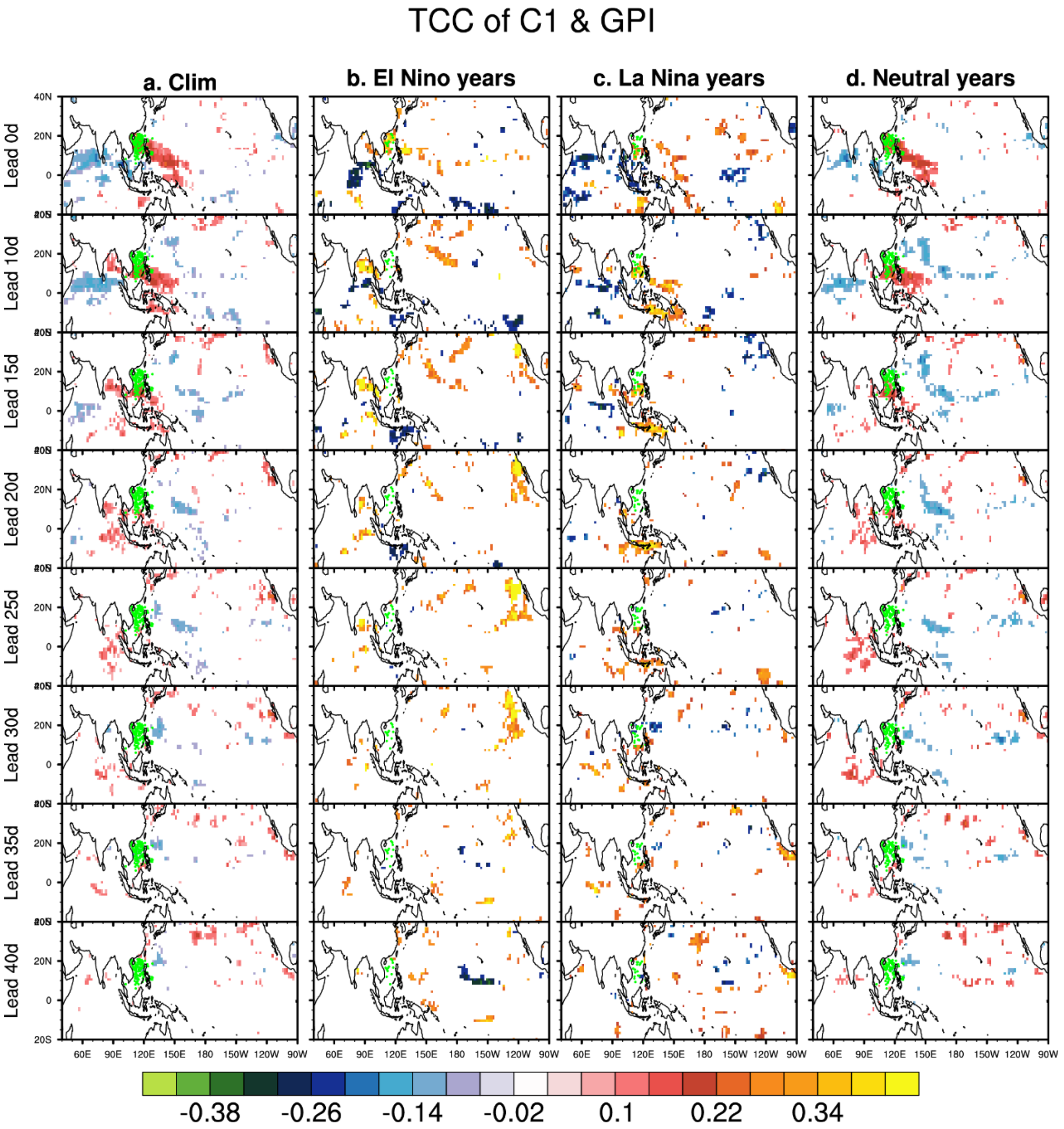


Fig. 12 Temporal correlation coefficients (TCCs; shading) between TC count anomalies of (a–d) C1 and GPI, and (e–h) C6 and MPI during (a, e) all years, (b, f) El Niño years, (c, g) La Niña years, and (d, h) neutral years of 1982–2008, at 0-day, 10-day, 15-day, 20-day, 25-day, 30-day, 35-day and 40-day leads (from top to bottom). Only the TCCs that are statistically significant at the 95% confidence level are shown. TC genesis locations are represented as green dots.



1062 Continued.

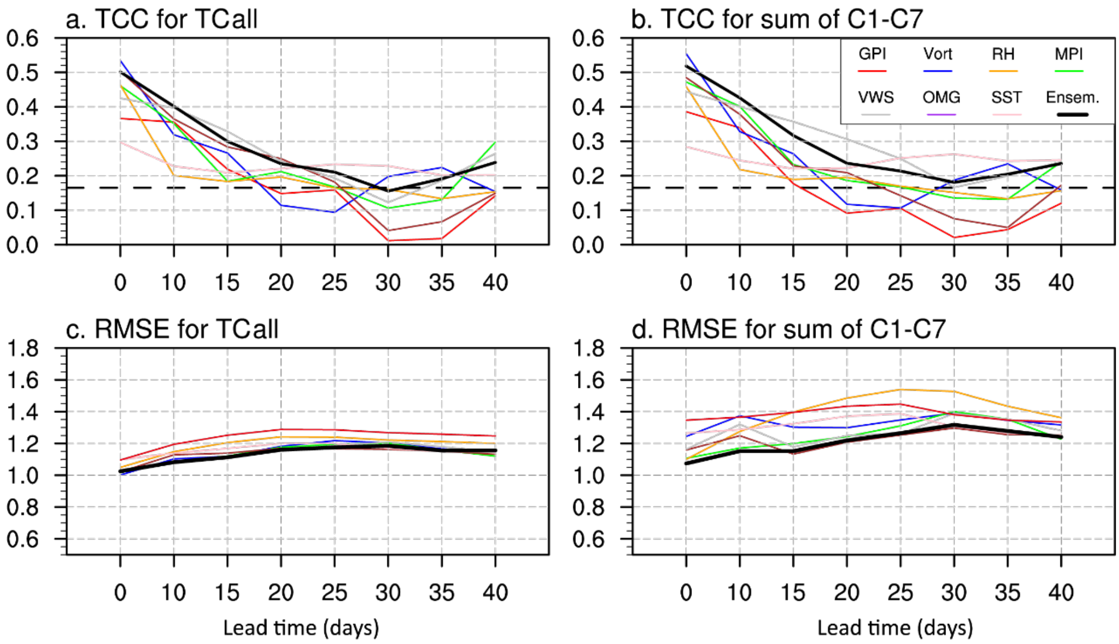


Fig. 13 (a) TCCs and (c) RMSEs between observed and predicted TCall count anomalies. (b, d) As in (a, c) but for the sum of C1–C7. Red, blue, orange, green, gray, purple, and pink lines in (a–d) denote the forecast skills by using GPI, Vort, RH, MPI, VWS, OMG, and SST as predictors, respectively. Thick black lines represent the skills of the ensemble results. Dashed lines in (a–b) mark the 95% confidence level of Spearman rank order correlation coefficients.

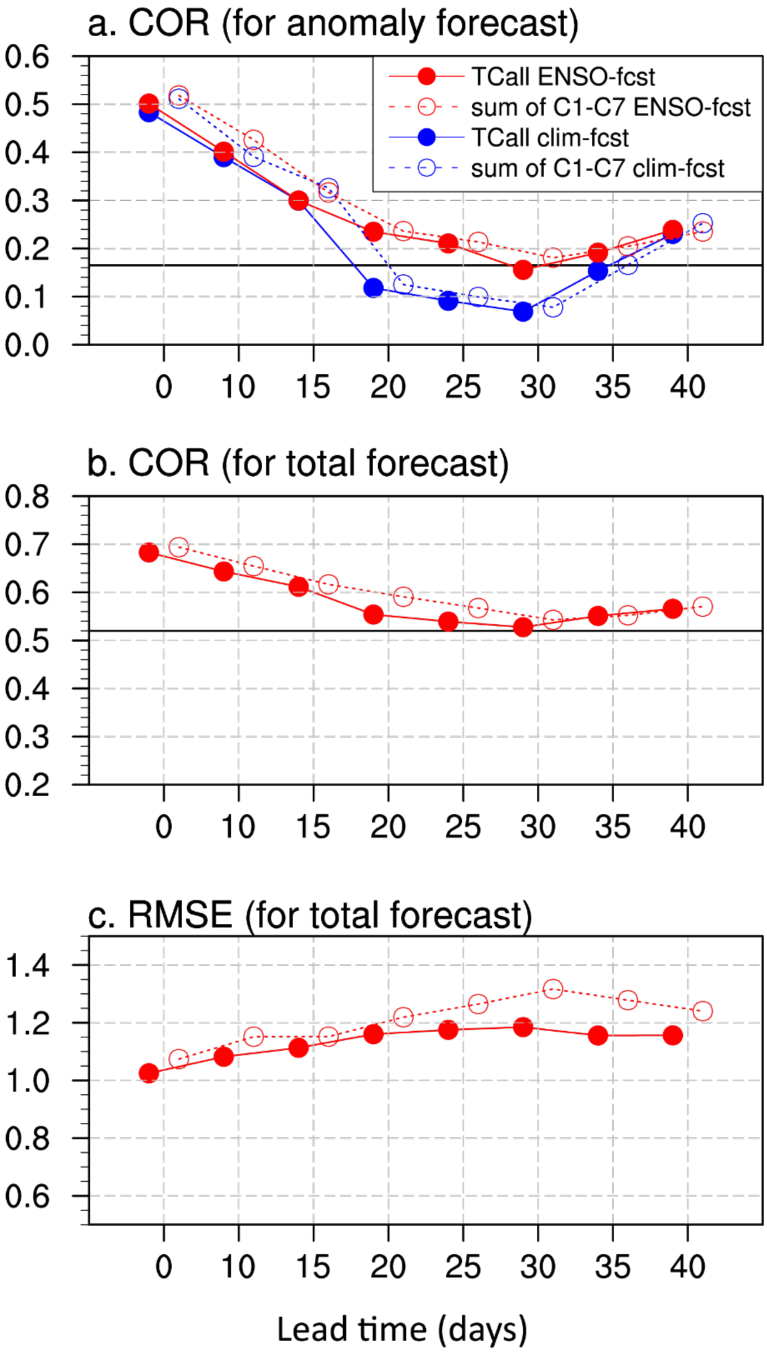


Fig. 14 TCCs between the observed and predicted (a) anomalous and (b) total TC count (sum of C1–C7 in blue, TCall in red) and (c) the RMSEs of total TC counts (open dot-dashed lines indicate the results of the sum of C1–C7; solid dot lines indicate the results of TCall) at different forecast lead times from 0 to 40 days (*x*-axis). The black line in (a) is the 95% significance level. The black line in (b) is the TCC skill of climatological prediction. Blue dotted lines represent the TCC prediction skill of clim-fcst statistical models.

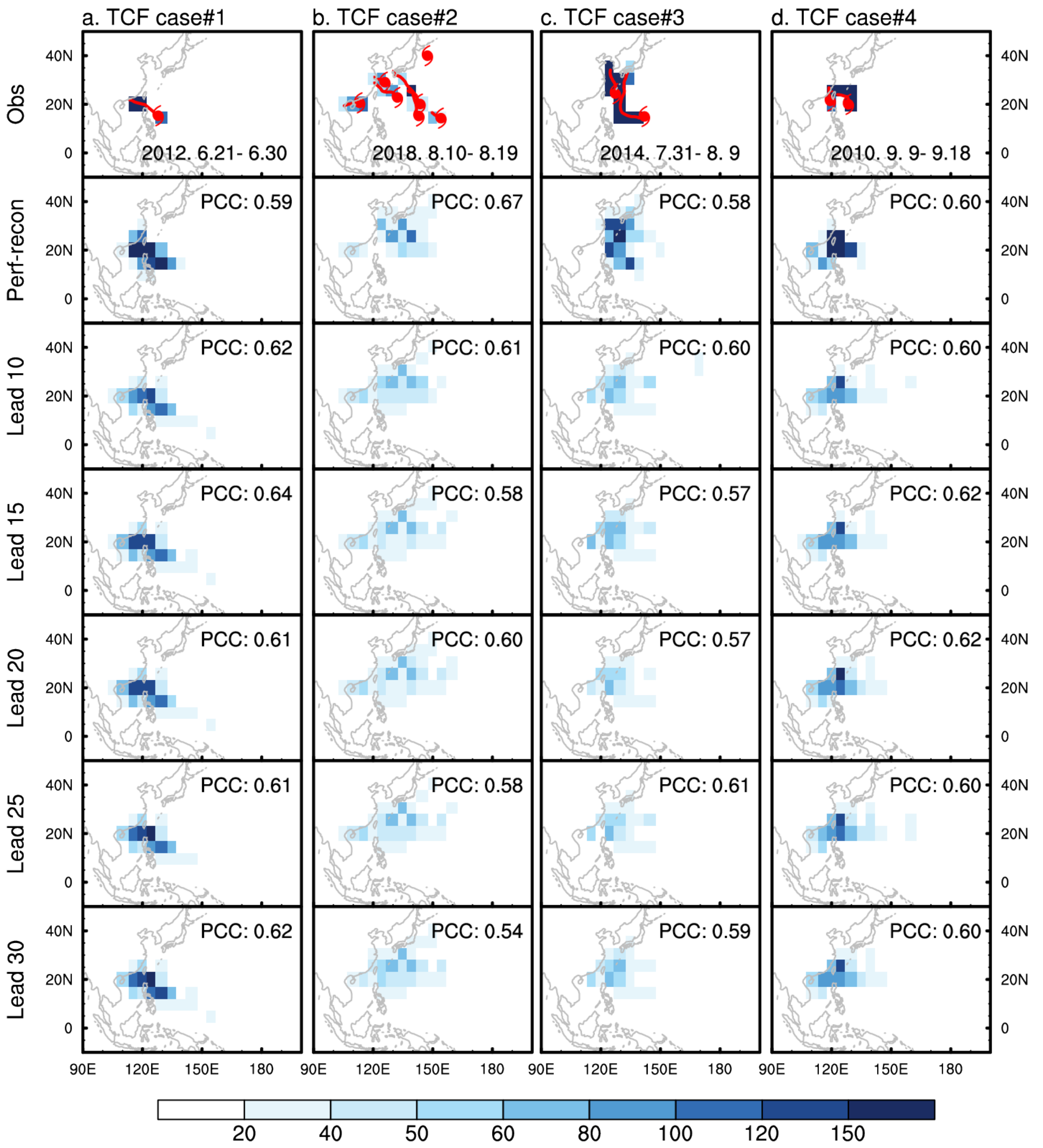


Fig. 15 Four cases of TC frequency (TCF) probability distributions during (a) 21–30 June 2012, (b) 10–19 Aug 2018, (c) 31 July–9 August 204, and (d) 9–18 September 2010. From top to bottom: observation, perfect reconstruction, and prediction at leads of 10, 15, 20, 25 and 30 days. The TCF probability at each grid point is defined as the TCF in each 5° × 5° box divided by the total TC count over the WNP. The red typhoon symbols and curves indicate the observed TC genesis locations and trajectories for each TC case. The pattern correlation coefficients between observations and predictions are shown in the upper-right corner of each panel.