# Variability and predictability of basin-wide and sub-basin tropical cyclone genesis frequency in the Northwest Pacific

Wei Mei[1*] and Shuo Li[1]

Department of Earth, Marine and Environmental Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC

* Corresponding author: Wei Mei, wmei@email.unc.edu

## Abstract

The variability and predictability of tropical cyclone genesis frequency (TCGF) during 1973−2010 at both basin-wide and sub-basin scales in the Northwest Pacific are investigated using a 100-member ensemble of 60-km-resolution atmospheric simulations that are forced with observed sea surface temperatures (SSTs). The sub-basin regions include the South China Sea (SCS) and the four quadrants of the open ocean. The ensemble-mean results well reproduce the observed interannual-to-decadal variability of TCGF in the southeast (SE), northeast (NE), and northwest (NW) quadrants, but show limited skill in the SCS and southwest (SW) quadrant. The skill in the SE and NE quadrants is responsible for the model's ability to replicate the observed variability in basin-wide TCGF. Above-normal TCGF is tied to enhanced relative SST (i.e., local SST minus tropical-mean SST) either locally or to the southeast of the corresponding regions in both the observations and ensemble mean for the SE, NE and NW quadrants, but only in the ensemble mean for the SCS and SW quadrant. These results demonstrate the strong SST control of TCGF in the SE, NE and NW quadrants; both empirical and theoretical analyses suggest that ensembles of ~10, 20, 35 and 15 members can capture the SST-forced TCGF variability in these three sub-basin regions and the entire basin, respectively. In the SW quadrant and SCS, TCGF contains excessive

noise, particularly in the observations, and thus shows low predictability. The variability and predictability of the large-scale atmospheric environment and synoptic-scale disturbances and their contributions to those of TCGF are also discussed.

## 1. Introduction

The Northwest Pacific (NWP) is the basin where tropical cyclones (TCs) are the most active in terms of both genesis frequency and lifetime peak intensity (e.g., Chan and Shi 1996; Chia and Ropelewski 2002; Camargo and Sobel 2005). These violent storms can bring about major societal and economic impacts to the countries and regions in East and Southeast Asia, where large and dense population resides (e.g., Zhang et al. 2009; Woodruff et al. 2013). Thus, it is of great importance to have a good understanding and accurate prediction of the variability in NWP TC activity (e.g., Knutson et al. 2010; Kossin et al. 2016; Lee et al. 2020). In this study, we focus on TC genesis frequency (TCGF) at both basin-wide and sub-basin scales in the NWP.

The variations in basin-wide TCGF over the NWP have been extensively studied during the past two decades, with a focus on the role of sea surface temperatures (SSTs). The SST factors that have been identified include the central-Pacific El Niño–Southern Oscillation (ENSO), the Pacific Meridional Mode (PMM), and SST anomalies in the tropical Indian and Atlantic Oceans (e.g., Wang et al. 2013; Wang and Wang 2019; Zhan et al. 2019). A positive phase of central-Pacific ENSO (also known as El Niño Modoki and Date Line El Niño; Larkin and Harrison 2005; Ashok et al. 2007; Kao and Yu 2009) tends to encourage basin-wide TC genesis by inducing favorable atmospheric conditions, such as above-normal low-level vorticity, over the majority of the NWP (e.g., Chen and Tam 2010; Kim et al. 2011; Mei et al. 2015; Liu and Chen 2018; Patricola et al. 2018; Wu et al. 2018; Zhao and Wang 2019). More recently, the PMM, which is characterized by a meridional dipole pattern of SST anomalies and has strong associations with the central-Pacific ENSO (e.g., Larson and Kirtman 2014; Capotondi and Sardeshmukh 2015; Amaya 2019), has also been proposed as a mechanism driving the variability in NWP basin-wide TCGF. Specifically, a positive phase of the PMM promotes TC formation in the NWP, mainly via its effect on dynamic factors (e.g., reduced vertical wind shear; Zhang et al. 2016; Liu et al. 2019). In addition, tropical SST anomalies outside the Pacific may also influence TC genesis in the NWP. Positive SST anomalies in the tropical Indian Ocean, which often emerge during the summers following El Niños, can substantially suppress TC genesis in the NWP by generating an anticyclonic circulation anomaly in the lower troposphere over the NWP (e.g., Xie et al. 2009,2016; Du et al. 2011; Zhan et al. 2011; Li 2012; Tao et al. 2012; Ha et al. 2015). Anomalous SST warming in the tropical North Atlantic has also been linked to below-normal TC activity in the NWP and the mechanisms may involve the Walker circulation and SSTs in the Indian Ocean and subtropical North Pacific (e.g., Huo et al. 2015; Yu et al. 2016; Zhang et al. 2017; Gao et al. 2018).

TC genesis in the NWP also exhibits strong spatial variations, and TCs forming in different parts of the NWP have considerable differences in their characteristics, including track orientation, landfalling location, and lifetime peak intensity (e.g., Camargo 2007a,b; Mei and Xie 2016; Kim and Seo 2016; Nakamura et al. 2017). Accordingly, a good understanding of the variability and changes in regional TCGF is more important than that in basin-wide TCGF (e.g., Liu and Chan 2003; Vecchi et al. 2014). A well-known factor responsible for the spatial inhomogeneity in NWP TCGF variability is the traditional or eastern-Pacific ENSO. This type of ENSO has opposite effects on the large-scale environment in the southeast and northwest portions of the NWP, and thereby leads to a shift in TC genesis without significantly altering basin-wide TCGF (e.g., Chan 1985; Lander 1994; Wang and Chan 2002; Camargo and Sobel 2005; Chen et al. 2006; Choi et al. 2015). Using primarily observations, a recent study by Wu et al. (2019) shows that TCGF in the southeast and northwest quadrants can also be affected by SSTs in the tropical Indian Ocean, and TCGF variability in the northeast quadrant may be related to SSTs in the tropical North Atlantic.

Despite the past efforts on understanding regional variations in TCGF, a comprehensive examination of TCGF variability over all sub-basin regions of the NWP (e.g., the South China Sea) using both observations and the simulations that explicitly produce TCs is still lacking. In addition, both the observations and simulations exhibit substantial internal variability, and the internal variability has been suggested to have strong spatial dependence. It is accordingly expected that the predictability of TCGF may vary considerably across different sub-basin regions of the NWP.

In this study, we attempt to fill these gaps and explore the variability and predictability[1] of both basin-wide and sub-basin TCGF in the NWP using a set of 60-km-resolution atmospheric simulations with an unprecedented ensemble size. After describing the datasets and methods in use (section 2) and comparing the observed and simulated TCGF climatology (section 3), we investigate the SST-forced interannual-to-decadal variability in the simulated TCGF, compare it with the observations, and study the underlying physical mechanisms (section 4). We then in section 5 explore the internal variability and predictability of TCGF as well as those of the large-scale atmospheric environment and synoptic-scale disturbance activity, the two modulators of TCGF. Concluding remarks are given in section 6. For the convenience of reading, the acronyms and abbreviations used in this paper are listed in Table 1.

## 2. Data and Methods

### *2.1 Observational and reanalysis data*

Owing to the discrepancies among available TC best track datasets (e.g., Barcikowska et al. 2017), we use three best track datasets produced respectively by the Joint Typhoon Warning Center (Chu et al. 2002), Shanghai Typhoon Institute of the China Meteorological Administration (Ying et al. 2014), and Japan Meteorological Agency, all of which provide the location and intensity of TCs at 6-h intervals. We use the mean of the three best track data to represent the

[1] Note that the predictability discussed in this study is "potential predictability", since the simulations in use are forced with perfect boundary (observed SST) conditions (Kang and Shukla 2006).

observations. Given that relatively large uncertainties exist in TC data during the early time period and the simulations (see section 2.2) end in 2010, here we focus on TCs reaching at least tropical storm intensity during 1973–2010. As an example of the uncertainties in TC best track data, Fig. S1 in the online supplemental material shows that the annual mean initial intensity (measured by sustained surface wind speed) of TCs is larger prior to 1973 than after, and thus the genesis locations may not be accurate before 1973. TCGF is calculated as the total number of TCs forming in the entire basin or each sub-basin region (defined in section 2.3) on a yearly basis.

In addition, monthly SSTs and atmospheric variables (including sea level pressure, specific and relative humidity, temperature, 850- and 250-hPa horizontal winds, and 500-hPa vertical pressure velocity) from the Japanese 55-year Reanalysis (JRA-55; Kobayashi et al. 2015) are used to study the physical mechanisms underlying the year-to-year variability in observed TCGF.

*2.2 Simulations*

The simulations in use are the historical simulations from the Database for Policy Decision Making for Future Climate Change (d4PDF; Mizuta et al. 2017). They were run with the Meteorological Research Institute Atmospheric General Circulation Model (AGCM), version 3.2, of 60-km resolution, and forced with observed monthly SSTs and sea ice concentration (COBE-SST2; Hirahara et al. 2014) as well as climatological monthly sea ice thickness. The simulations cover the period from 1951 to 2010, and consist of 100 member simulations that differ in initial conditions and slightly in the prescribed SSTs.[2] The simulations replicate the year-to-year variations in large-scale atmospheric circulation associated with global tropical SST variability (Kamae et al. 2017a,b; Ueda et al. 2018).

[2] The initial conditions in different members were the snapshots on different dates in former simulations with the same AGCM. The small perturbations added to the observed SSTs represent SST sampling and analysis errors, and were generated using empirical orthogonal functions (EOFs) that represent the interannual variability in SSTs. Details on how the initial conditions and SSTs were perturbed can be found in the appendix of Mizuta et al. (2017).

The simulations generate TC-like vortices. These vortices are detected and tracked using sea level pressure, 850-hPa vorticity, 850-hPa, 300-hPa and surface winds, warm-core temperature, and duration of the tracks following a methodology that combines Murakami et al. (2012) and Mei et al. (2014). We have examined some randomly selected vortices and found that they share many similarities with TCs in the observations. As in Yoshida et al. (2017), the simulations reproduce many statistics of TCs in the observations, such as the geographical distribution of climatological TC occurrence [e.g., Fig. 2 in Yoshida et al. (2017)]. We have shown in Mei et al. (2019) that the simulations capture 70% of the variability in observed North Atlantic hurricane frequency. We, however, recognize a notable caveat of using the AGCM simulations is that these simulations may not generate correct surface wind speeds and fluxes, which in turn may affect TC-related thermodynamic parameters (e.g., potential intensity) and TC activity (Emanuel and Sobel 2013).

In the simulations, we consider storms with a lifetime peak intensity of at least 11 m $s^{-1}$ as TCs. The choice of 11 m $s^{-1}$ is made to match the annual averaged TCGF in the simulations to that in the observations during 1973–2010 (i.e., ~25.74 per year). This threshold value is smaller than what is used in the observations (i.e., 17.5 m $s^{-1}$), primarily owing to the relatively low resolution of the model and the differences in the average time period over which storm intensity is estimated (e.g., Bacmeister et al. 2018; Li and Sriver 2018). Using different threshold values produces similar results and does not alter the main conclusions of this study (e.g., Table S1 in the online supplemental material).

In each member simulation, a variable (e.g., basin-wide TCGF) can be divided into two components: forced component tied to the imposed SST, and internal component due to the randomness in atmospheric processes. Accordingly, the SST-forced component is approximated by the ensemble mean of the 100-member simulations, and its year-to-year variations are referred to as forced variability. The internal variability is determined by the deviations of individual member simulations from the ensemble mean. It is worth noting that the perturbations added to the SSTs have negligible effects on the forced variability and no significant effects on the internal variability in the atmospheric variables and simulated TC frequency (Mizuta et al. 2017; Mei et al. 2019).

*2.3 Definition of the five sub-basin regions*

In studying the ENSO effects on TC genesis over different portions of the NWP, Wang and Chan (2002) use 140°E/17°N as the boundaries for the four quadrants of the open ocean during the early and peak TC season and 150°E/17°N during the late season. In a recent study by Wu et al. (2019), 150°E/15°N are adopted as the boundaries to obtain nearly equal areas for the four quadrants of the open ocean. Here we employ 144°E/16°N as the dividing longitude/latitude, which are located around the middle of those used in previous studies, for the four quadrants of the open ocean (Fig. 1); changing the boundaries by a couple of degrees reaches very similar conclusions. The South China Sea (SCS) comprises the fifth sub-basin region of the NWP.

*2.4 Correlation skill of the ensemble mean in reproducing the observed variability*

Here we extend the results shown in section 4b (page 3161) of Mei et al. (2019) to make them applicable to generic cases. Let $z(t)$ be the observed time series of a given variable (e.g., basin-wide TCGF in this study) $z(t) = x_o(t) + \varepsilon(t)$, where $x_o(t)$ represents the signal in the observations and $\varepsilon(t)$ represents the noise with mean 0 and standard deviation $\sigma_\varepsilon$. Let $y_i(t)$ be the simulated time series of the same variable of interest in the *i*th member simulation: $y_i(t) = x_m(t) + e_i(t), i = 1, 2, \ldots, N$, where $x_m(t)$ represents the signal in the simulations, $e_i(t)$ represents the model noise with mean 0 and standard deviation $\sigma_e$ (i.e., the internal variability defined in section 2.2), and $N$ is the total number of member simulations.

Assume that $x_o(t)$, $\varepsilon(t)$, $x_m(t)$, and $e_i(t)$ are weakly stationary time series, and that $e_i(t), i = 1, 2, \ldots, N$, are uncorrelated with each other and are uncorrelated with $\varepsilon(t)$ at all leads and lags. Define the ensemble mean of the $N$ simulations as $\bar{y}(t) = N^{-1}\sum_{i=1}^{N} y_i(t)$. With some algebra, we can write the population correlation between $\bar{y}(t)$ and the observations (i.e., $z(t)$) as

$$\rho_{\mathrm{en}} = \mathrm{cor}\big(\bar{y}(t), z(t)\big) = \frac{\mathrm{cov}(\bar{y}(t), z(t))}{\sqrt{\mathrm{var}(\bar{y}(t))\mathrm{var}(z(t))}} = \frac{\rho_{m,o}\sigma_m\sigma_o}{\sqrt{(\sigma_m^2 + N^{-1}\sigma_e^2)(\sigma_o^2 + \sigma_\varepsilon^2)}}, \tag{1}$$

where $\rho_{m,o}$ is the population correlation between $x_o(t)$ and $x_m(t)$, $\sigma_m$ is the standard deviation of $x_m(t)$ and is also known as the forced variability defined in section 2.2, and $\sigma_o$ is the standard deviation of $x_o(t)$. It follows that the population correlation between the *i*th member simulation and the observations is

$$\mathrm{cor}\big(y_i(t), z(t)\big) = \frac{\rho_{m,o}\sigma_m\sigma_o}{\sqrt{(\sigma_m^2 + \sigma_e^2)(\sigma_o^2 + \sigma_\varepsilon^2)}}, i = 1, 2, \ldots, N,$$

which is denoted by $\rho$. Then we have the following identity that links $\rho_{\mathrm{en}}$ and $\rho$:

$$\rho_{\mathrm{en}} = \rho\sqrt{\frac{\sigma_m^2 + \sigma_e^2}{\sigma_m^2 + N^{-1}\sigma_e^2}} = \rho\sqrt{\frac{1 + \mathrm{SNR}^{-2}}{1 + N^{-1}\mathrm{SNR}^{-2}}},$$

where $\mathrm{SNR} = \sigma_m/\sigma_e$ is known as the signal-to-noise ratio in the model simulations.

In practice, $\rho$ is estimated by $\bar{r} = N^{-1}\sum_{i=1}^{N} r_i$, the mean of the sample correlation $r_i$ between individual member simulations and the observations. Accordingly, the sample correlation between the ensemble mean and the observations, $r_{\mathrm{en}}$, can be estimated as

$$r_{\mathrm{en}} = \bar{r}\sqrt{\frac{1 + \widehat{\mathrm{SNR}}^{-2}}{1 + N^{-1}\widehat{\mathrm{SNR}}^{-2}}}, \tag{2}$$

where $\widehat{\mathrm{SNR}}$, the estimator of SNR, is defined in the following subsection.

*2.5 Computing the signal-to-noise ratio (SNR) in practice*

Using the same notation as in section 2.4, we obtain the following identities:

$$\sigma_y^2 = \mathrm{var}\big(y_i(t)\big) = \sigma_m^2 + \sigma_e^2\,,$$

$$\sigma_{\bar{y}}^2 = \mathrm{var}\big(\bar{y}(t)\big) = \sigma_m^2 + N^{-1}\sigma_e^2\,.$$

It follows that

$$\sigma_m^2 = \frac{1}{N-1}\left(N\sigma_{\bar{y}}^2 - \sigma_y^2\right),$$

$$\sigma_e^2 = \frac{N}{N-1}\left(\sigma_y^2 - \sigma_{\bar{y}}^2\right).$$

We calculate the SNR as

$$\mathrm{SNR} = \frac{\sigma_m}{\sigma_e} = \sqrt{\frac{1}{N}\frac{N\sigma_{\bar{y}}^2 - \sigma_y^2}{\sigma_y^2 - \sigma_{\bar{y}}^2}}\,. \tag{3}$$

In practice, we estimate $\sigma_{\bar{y}}^2$ as $\hat{\sigma}_{\bar{y}}^2$, the sample variance of the ensemble mean $\bar{y}(t)$, and estimate $\sigma_y^2$ by averaging the sample variances of $y_i(t)$ in individual member simulations, i.e., $\hat{\sigma}_y^2 = N^{-1}\sum_{i=1}^{N} s_i^2$, where $s_i^2$ is the sample variance of $y_i(t)$. We then obtain the estimated SNR, denoted as $\widehat{\mathrm{SNR}}$. The SNR calculated using this method is more accurate and stable than that computed using the method described in Mei et al. (2014, 2015, 2019), as illustrated in Fig. S2 of the online supplemental material and numerically confirmed using the toy model described in section 4b (page 3160) of Mei et al. (2019).

*2.6 Calculations of a genesis potential index and synoptic-scale disturbance activity*

A genesis potential index (GPI), which integrates four thermodynamic and dynamic factors and represents the favorability of the large-scale atmospheric environment in which TCs develop, is calculated following Emanuel (2010) as:

$$\mathrm{GPI} = \frac{a|\eta|^3\left[\max(V_{\mathrm{PI}} - 35, 0)^2\right]}{\chi^{4/3}(25 + V_{\mathrm{sh}})^4}\,, \tag{4}$$

where $a$ is a constant and in this study set to be $10^{16}$, $\eta$ is the 850-hPa absolute vorticity, $V_{\mathrm{PI}}$ is the TC potential intensity, $\chi$ is the 600-hPa entropy deficit, and $V_{\mathrm{sh}}$ is the magnitude of the 250–850-hPa wind shear vector [see also Korty et al. (2012) and Tang and Emanuel (2012)].

To quantify the contribution of each component (e.g., 850-hPa vorticity) to the changes in the GPI, we recompute the GPI using the original, year-to-year varying values for that component but the climatology of 1973–2010 for the remaining three components, following Camargo et al. (2007c). This procedure is carried out for all the four components of the GPI. Note that the sum of the contributions of the four individual components is close but not exactly equal to the changes in the GPI, owing to the nonlinearity of the GPI formula, as discussed in Camargo et al. (2007c).

The synoptic-scale disturbance activity is assessed following Li et al. (2010) and Vecchi et al. (2019). Specifically, for each year at each grid it is defined as the variance of 2–8-day bandpass filtered 850-hPa relative vorticity during the peak TC season (e.g., June–November for the NWP). To minimize the effects of TCs, the vorticity within 500 km of each TC location is removed before computing the variance.

## 3. Climatology of TCGF

Figure 1 displays the geographical distribution of climatological TC genesis over the NWP at 2°×2° grids in the observations and ensemble mean of the simulations. Generally, the model reproduces the large-scale pattern and magnitude of TC genesis in the observations. For instance, in both the observations and simulations, TCs form primarily in the SCS and south of 24°N over the open ocean; and over the open ocean, TC genesis exhibits a southeast-northwest orientation. It is worth noting that the spatial distribution is smoother in the simulations than in the observations, simply because of the averaging effect of the ensemble mean. In addition, in the SCS TC genesis on average is located slightly more south in the simulations than in the observations.

The model is also skillful at replicating the climatological seasonal cycle of the observed TCGF in the NWP and its five sub-basin regions (Fig. 2). The skill is higher in the northwest (NW), northeast (NE) and southeast (SE) quadrants than in the SCS and the southwest (SW) quadrant in terms of the magnitude and phase of the seasonal cycle. Specifically, the simulations well capture the peak month of TCGF in the former three regions: August for the NW and NE quadrants and October for the SE quadrant; and in both the observations and simulations, no TCs occur between December and April in the NW and NE quadrants. In the SCS, the simulated seasonal cycle has a magnitude similar to that in the observations, but lags in phase by approximately one month, with fewer TCs forming during May−August and more TCs during October−March in the simulations. In the SW quadrant, the magnitude of the seasonal cycle in the simulations is only around half of that in the observations, with fewer TCs during June−October and more TCs during December−March. Interestingly, in this region both the observed and simulated seasonal cycles exhibit a local dip in August, which is a robust feature in nearly all the 100 member simulations and warrants a further investigation.

Based on the seasonal cycle of TCGF, the following months are defined as active TC seasons for the NWP and its sub-basin regions: June−November for the entire basin, the SCS, and the SW quadrant; July−October for the NW and NE quadrants; and July−December for the SE quadrant. In the following sections, SSTs, the large-scale atmospheric condition, and synoptic-scale disturbance activity averaged or defined in these months will be used to explore the physical mechanisms underlying the year-to-year variability of TCGF in the corresponding regions.

## 4. Forced variability in TCGF and its connections to the large-scale environment and synoptic-scale disturbances

### *4.1 Interannual-to-decadal variations in TCGF*

Figure 3a shows the year-to-year variations of basin-wide TCGF in the ensemble mean of the simulations (thick black curve) and in the observations (thick red curve). The ensemble mean captures nearly half of the variance in observed TCGF, with a correlation coefficient of around 0.7. In both the observations and ensemble mean, basin-wide TCGF shows an upward trend from the mid-1970s to late 1990s, experiences a substantial drop in 1998 when a strong La Niña event takes place, and stays at a relatively low level afterwards.

The skill of the model at reproducing the interannual-to-decadal variability in basin-wide TCGF is attributable to two factors: (i) the model's skill at replicating TCGF variability in the SE and NE quadrants of the basin, and (ii) the dominance of TCGF of these two quadrants in the variability of basin-wide TCGF. For these two quadrants, the correlation coefficients between the simulated and observed TCGF are 0.88 and 0.45, respectively (Figs. 3c,f). Meanwhile, TCGF in these two quadrants evolve in a manner similar to basin-wide TCGF. They respectively account for around 31% and 29% of the total variance in basin-wide TCGF (the correlation coefficients $r$ = 0.56 and 0.54, respectively) in the observations and around 60% and 27% ($r$ = 0.78 and 0.52, respectively) in the ensemble mean (Table 2), while they are nearly uncorrelated with each other in both the observations and simulations ($r$ = -0.10 and 0.01, respectively; Table 2).

The model also shows skills at capturing a considerable portion of the observed TCGF variability in the NW quadrant ($r$ = 0.48; Fig. 3b). This, however, does not contribute to the model's good performance in reproducing the observed variability in basin-wide TCGF, because TCGF in this quadrant explains less than 1% of the variations in basin-wide TCGF (the correlation coefficients between TCGF in the NW quadrant and basin-wide TCGF are -0.03 and -0.09 in the observations and ensemble mean, respectively; Table 2). It is worth noting that TCGF in this quadrant negatively covaries with TCGF in the SE quadrant ($r$ = -0.54 and -0.48 in the observations and ensemble mean, respectively), largely owing to the ENSO effect; this will be discussed in section 4.2.

On the contrary, the model has limited skill at reproducing the observed TCGF variability in the SW quadrant of the basin and the SCS: the correlation coefficients between the simulated and observed TCGF are 0.14 and 0.32 respectively for these two sub-basin regions, both insignificant at the 0.05 level (Figs. 3d,e). This result echoes the underperformance of the model in simulating the observed climatological seasonal cycle over the two regions (section 3 and Figs. 2d,e). These discrepancies between the simulations and observations are primarily due to the intrinsically low predictability of TC genesis in the observations over these two regions, which will be discussed in section 5.

### *4.2 Linkages to SSTs and the large-scale atmospheric environment*

The high skill of the model at reproducing the observed TCGF variability in the SE, NW and NE quadrants demonstrates the strong SST control of TC genesis in these sub-basin regions given the fact that the simulations are forced with observed SSTs. Next, we shall attempt to identify the regions where SSTs are important for TCGF of the entire NWP and its sub-basin regions. We have calculated the correlation coefficients between TCGF in individual regions and global original SST and those between TCGF and global relative SST (i.e., SST minus tropical-mean SST); the calculations were carried out separately for the observations and ensemble mean. The results for relative SST are broadly consistent with those for original SST, and here we focus on relative SST (Fig. 4) because (1) relative SST better represents dynamic and thermodynamic processes (e.g., convection), and (2) using relative SST produces more consistent results between the simulations and observations, facilitating the interpretation of the linkages between TCGF and

SSTs. For the sake of completeness, we also show the results for original SST in Fig. S3 of the online supplemental material.

In both the observations and simulations, an active NWP TC season is characterized by above-normal relative SSTs over the off-equatorial tropical central North Pacific and below-normal relative SSTs in the Indo-West Pacific and tropical North Atlantic (Figs. 4a,d).[3] This anomalous SST pattern is consistent with the findings in previous studies, which show that NWP TCs may be modulated by SSTs in the tropical Pacific, Indian and Atlantic Oceans (e.g., Clark and Chu 2002; Du et al. 2011; Zhan et al. 2011; Wang et al. 2013; Mei et al. 2015; Yu et al. 2016; Zhang et al. 2016; Patricola et al. 2018; Zhao and Wang 2019; Wu et al. 2020). Next, we shall identify the SST pattern and large-scale atmospheric conditions that are responsible for the year-to-year variations of TCGF in individual sub-basin regions.

A La Niña-like state favors TC genesis in the NW quadrant of the basin in both the observations and simulations (Figs. 4b,e). When SSTs in the central-to-eastern equatorial Pacific are colder than usual, the NW quadrant experiences above-normal relative SSTs. The increased relative SSTs in this quadrant tend to enhance relative humidity in the middle troposphere of the region (Fig. 5a and Fig. S5a in the online supplemental material) via intensified convection (Figs. S6b,e in the online supplemental material), and thereby promote TC genesis. This result is in line with Camargo et al. (2007c) and Li et al. (2022), both of which emphasize the importance of relative humidity in modulating TC genesis over this region.

High TCGF in the NE quadrant of the basin is associated with above-normal relative SSTs over 150°E–160°W, 10°–25°N in both the observations and simulations (Figs. 4c,f). High relative

[3] We note that the effect of the relative SST anomalies over the tropical central Pacific is more prominent in the model ensemble mean than in the observations. This discrepancy can be primarily attributed to the fact that the observations represent one realization of all possibilities that could occur (cf. Fig. S4 in the online supplemental material and Fig. 4a), whereas the ensemble mean approximates the average of all possibilities.

SSTs in this region produce excessive latent heat release via strengthened convective activity (Figs. S6c,f in the online supplemental material), which generate an anomalous cyclonic circulation to its northwest as a Rossby wave response in the lower troposphere (Figs. 4c,f). The enhanced low-level vorticity, along with increased relative humidity, provides favorable environment nurturing TC genesis in the NE quadrant of the NWP (Fig. 5b and Fig. S5b in the online supplemental material).

In the simulations, the year-to-year variations of TCGF over the SCS are also linked to a La Niña-like SST pattern (Fig. 4j). This pattern resembles the SST pattern for TCGF of the NW quadrant despite a slightly westward shift (cf. Figs. 4e,j), corresponding to a significantly positive correlation between TCGF in these two regions (Table 2). The enhanced relative SSTs in the SCS and east of the Philippines increase moisture in the middle troposphere (Fig. 5c) and generate a cyclonic circulation anomaly to the northwest (Fig. 4j), strengthening low-level vorticity (Fig. 5c); both increased mid-level relative humidity and low-level vorticity facilitate TC genesis in the SCS. In the observations, a similar but insignificant SST pattern is detected for TCGF of the SCS (Fig. 4g). This implies that the internal variability in the observed TCGF is much stronger than the forced variability in this region, consistent with the relatively low skill of the simulations at replicating observed TCGF in this region (Fig. 3d).

In the simulations, the SST pattern responsible for TCGF variations in the SW quadrant is more or less similar to that for the NE quadrant but with stronger correlations in the lower latitudes (Fig. 4k). Above-normal relative SSTs over 130°–160°E, 10°S–10°N increase local relative humidity and generate an anomalous cyclonic circulation to their northwest in the lower troposphere, promoting low-level vorticity and TC genesis in the SW quadrant (Fig. 5d). In the observations, a similar but insignificant SST pattern emerges (Fig. 4h).[4] This is analogous to what occurs in the SCS, but the correlation in the observations is even weaker (cf. Figs. 4g,h), indicating the even stronger internal variability in the observed TCGF and poorer skill of the model in the SW quadrant (Fig. 3e).

In both the observations and simulations, high TCGF in the SE quadrant is tied to an El Niño-like condition (Figs. 4i,l). Above-normal relative SSTs in the central-to-eastern tropical Pacific and below-normal relative SSTs in the Indo-West Pacific produce a large anomalous low-level cyclonic circulation over the majority of the North Pacific and an anomalous low-level anticyclonic circulation covering the SCS and tropical North Indian Ocean (Figs. 4i,l). This dipole pattern of low-level circulation encourages TC genesis in the SE quadrant via enhanced low-level vorticity and reduced vertical wind shear, and discourages TC genesis in the SCS via reduced low-level vorticity (Fig. 5e and Fig. S5e in the online supplemental material), explaining a negative correlation between TCGF in these two sub-basin regions (Table 2). The accompanied below-normal relative SSTs in the NW quadrant suppress TC genesis in the region via reduced relative humidity in the mid-troposphere (Fig. 5e and Fig. S5e in the online supplemental material), as discussed above, accounting for a negative correlation between TCGF in this region and that in the SE quadrant (Table 2).

In short, in the simulations above-normal TCGF in all five sub-basin regions can be linked to enhanced relative SSTs either locally or to the southeast of the region (Figs. 4e,f,j,k,l), which themselves are associated with changes in both local and remote SSTs (Figs. S3e,f,j,k,l in the online supplemental material). The promoted TC genesis is attributable to increased mid-level

[4] The lack of significant correlations between TCGF and SSTs over the SW quadrant and the SCS also exist in individual member simulations (Fig. S7 in the online supplemental material), because of the extremely strong internal variability in atmospheric processes and TCGF. The 100-member ensemble mean can effectively remove most of the internal variability, producing significant relationships between TCGF and SSTs shown in Figs. 4j,k.

relative humidity in the NW quadrant, to increased low-level vorticity and mid-level relative humidity in the NE and SW quadrants and the SCS, and to increased low-level vorticity and reduced vertical wind shear in the SE quadrant (Fig. 5). The observational results are consistent with those from the simulations in the NW, NE and SE quadrants in terms of the associated SST pattern and large-scale atmospheric conditions (Figs. 4b,c,i and Figs. S5a,b,e in the online supplemental material). They are, however, mostly insignificant for the SCS and the SW quadrant (Figs. 4g,h and Figs. S5c,d in the online supplemental material), echoing the low predictability of observed TCGF in these two sub-basin regions (Figs. 3d,e).

*4.3 Role of synoptic-scale disturbances*

Synoptic-scale disturbances provide seeds for TC genesis (Fu et al. 2007, 2012; Zong and Wu 2015), and their effect on TC genesis under global warming has been emphasized in previous studies (e.g., Yoshimura and Sugi 2005; Yoshimura et al. 2006; Li et al. 2010; Vecchi et al. 2019). However, it is unclear whether they significantly affect NWP TCGF on interannual-to-decadal time scales, and whether their effect on TCGF is similar across different sub-basin regions.

Figures 6a–e show the observed correlation between synoptic disturbance activity (defined in section 2.6) and TCGF in the five sub-basin regions. In both the NW and SE quadrants, high TCGF tends to be associated with above-normal synoptic disturbance activity (Figs. 6a,e). As will be discussed in section 5.4, synoptic disturbance activity in these two regions is also more predictable than that in other sub-basin regions, contributing to the higher predictability of TCGF in the two regions. We also note that a season with active synoptic disturbances in the SE quadrant is likely a season with inactive disturbances in the NW quadrant; this might also contribute to the significantly negative correlation between TCGF in these two sub-basin regions (Table 2). In addition, in the NE quadrant the year-to-year variability of disturbance activity may play a marginal role in modulating TC genesis (Fig. 6b). On the contrary, the correlation between synoptic disturbance activity and TCGF is weak and statistically insignificant in both the SCS and the SW quadrant (Figs. 6c,d).

In the simulations, synoptic disturbances significantly modulate TC genesis in all sub-basin regions, with high synoptic disturbance activity favoring TC genesis (Figs. 6f–j). The similarities and differences between the simulated and observed synoptic disturbance activity will be discussed in section 5.4.

## 5. Internal variability

The spread of gray curves in Fig. 3 suggests that TCGF exhibits strong internal variability, in addition to the forced variability induced by SSTs. The magnitude of the internal variability may differ considerably among the five sub-basin regions, given the differences in the level of SST control. In this section, we shall take advantage of the 100-member ensemble and address the following four aspects pertaining to the internal variability in TCGF of the entire NWP and its sub-basin regions: (1) difference in the noise level between the observed and simulated TCGF; (2) relationship between the noise level in the simulated TCGF and the skill of the model; (3) number of member simulations needed to skillfully capture the observed variations in TCGF; and (4) internal variability in the large-scale atmospheric environment and synoptic-scale disturbance activity and their possible contributions to the internal variability in TCGF.

### *5.1 A comparison between the observed and simulated TCGF in terms of the noise level*

The level of noise in TCGF between the observations and simulations can be compared using the ratio of predictable component (RPC; Eade et al. 2014)[5], a quantity that for a sufficiently large ensemble size (e.g., 100 member simulations in this study) is expressed as

[5] It is worth noting that the predictable component defined here is fundamentally different from that reviewed in DelSole and Tippett (2007): the former measures the fraction of variance that is predictable in either observations or

$$\mathrm{RPC} = \frac{\mathrm{PC_{obs}}}{\mathrm{PC_{model}}} \approx \frac{r_{\mathrm{en}}}{\sqrt{\hat{\sigma}_{\bar{y}}^2/\hat{\sigma}_y^2}}, \tag{5}$$

where $\mathrm{PC_{obs}}$, the predictable component of the observations, is approximated as the sample correlation between the observations and the ensemble mean of the simulations ($r_{\mathrm{en}}$); and $\mathrm{PC_{model}}$, the predictable component of the model, is expressed in terms of the ratio of the sample variance of the ensemble mean, $\hat{\sigma}_{\bar{y}}^2$, and the averaged sample variance of individual member simulations, $\hat{\sigma}_y^2$ (see section 2.5 for the notation). The RPC, thus, reflects the difference in the level of predictability between the observations and simulations[6]: an RPC value greater than (smaller than) 1 suggests that the observations are more predictable than the simulations, and that the model is overdispersive (underdispersive) and thus underconfident (overconfident).

The RPC values for TCGF of the entire NWP and its five sub-basin regions are shown in the last column of Fig. 7a. It is greater than 1 in the SE and NW quadrants, suggesting that TCGF in the real world has a lower noise level and thus is more predictable than TCGF in the simulations over these two regions. In contrast, the RPC value is smaller than 1 in the other three sub-basin regions, particularly the SW quadrant. This indicates that the model is overconfident in simulating and predicting the year-to-year variability of TCGF in these regions, and that TCGF in the observations actually has a higher level of noise than that in the simulations. For basin-wide TCGF,

---

model simulations, while the latter is defined as the projection vector that minimizes the ratio of the forecast distribution variance and the climatological distribution variance.

[6] Plugging the sample version of Eq. (1): $r_{\mathrm{en}} = \frac{r_{m,o}\hat{\sigma}_m\hat{\sigma}_o}{\sqrt{(\hat{\sigma}_m^2+N^{-1}\hat{\sigma}_\varepsilon^2)(\hat{\sigma}_o^2+\hat{\sigma}_\varepsilon^2)}}$, where $r_{m,o}$ is the sample correlation between the signal in the simulations and that in the observations, into Eq. (5), we obtain

$$\mathrm{RPC} \approx \frac{r_{m,o}}{1+N^{-1}\widehat{\mathrm{SNR}}^{-2}}\sqrt{\frac{1+\widehat{\mathrm{SNR}}^{-2}}{1+\widehat{\mathrm{SNR}}_o^{-2}}} \approx r_{m,o}\sqrt{\frac{1+\widehat{\mathrm{SNR}}^{-2}}{1+\widehat{\mathrm{SNR}}_o^{-2}}},$$

where $\widehat{\mathrm{SNR}}_o = \hat{\sigma}_o/\hat{\sigma}_\varepsilon$ is the estimated SNR in the observations. Thus, strictly speaking the RPC is determined by two factors: the difference between the signal in the simulations and that in the observations (i.e., $r_{m,o}$), and the ratio of the level of predictability in the observations to that in the simulations. In this study, we assume $r_{m,o} = 1$, and the RPC only reflects the difference in the level of predictability between the observations and simulations.

the RPC value is slightly greater than 1, implying that the model simulations are slightly noisier and thus slightly less predictable than the observations.

To further illustrate the discrepancies in the level of noise (and thus predictability) between the observations and simulations, we perform the following calculations for TCGF of the entire basin as well as individual sub-basin regions. First, we compute the correlation coefficient between the observations and each of the 100 member simulations, and plot the histogram of the obtained 100 correlation coefficients in the form of a probability density as a red curve in Fig. 8. Second, for each individual member simulation, we compute its correlation coefficient with the other 99 member simulations, and plot the probability density of the obtained 99 correlation coefficients as a gray curve in Fig. 8; the average of the 100 gray curves is plotted as a black curve. Third, we compute the correlation coefficient between each member simulation and the ensemble mean of the 100 member simulations, and plot the probability density of the obtained 100 correlation coefficients as a blue curve in Fig. 8. Lastly, we mark the correlation coefficient between the observations and the ensemble mean as a vertical dotted magenta line in Fig. 8.

For TCGF of the NW and SE quadrants and of the entire basin, the center of the red curve is located to the right of the center of the black curve and the value marked by the magenta line is higher than the mean value implied by the blue curve (Figs. 8a,b,f; Table S2 in the online supplemental material). These results suggest that in these two sub-basin regions or when considering the TCGF over the entire NWP, individual member simulations on average are more similar to the observations than to each other and the observations have a higher predictability than the simulations. On the contrary, for TCGF of the NE and SW quadrants and the SCS, the mean value implied by the red curve is smaller than that by the black curve and the value denoted by the magenta line is smaller than the mean value implied by the blue curve (Figs. 8c–e; Table S2 in the online supplemental material). These indicate that in these three sub-basin regions, the observations on average have a higher level of noise than individual member simulations and accordingly individual simulations are more similar to each other than to the observations.

These results from the comparisons of the probability distributions of the correlation coefficients are consistent with the RPC values discussed above. Both show that in the NW and SE quadrants and when viewing the entire NWP as a whole, TCGF in the observations has a lower level of noise and thus is more predictable than that in the simulations; and that the opposite holds true for TCGF of the NE and SW quadrants and the SCS.

### *5.2 Noise level in the simulated TCGF and its relationship with the skill of the ensemble mean at reproducing the observations*

In both the NW and SE quadrants, TCGF is more predictable in the observations than in the simulations, with an RPC value of 1.37 and 1.04, respectively. The larger RPC value in the NW quadrant suggests that the difference in the noise level between the observations and simulations is greater in the NW quadrant than in the SE quadrant. This implies that the model would have a greater skill at reproducing the observed variability of TCGF in the NW quadrant than in the SE quadrant, others being equal. However, the correlation between the observed and simulated TCGF is weaker for the NW quadrant (0.48 for the NW quadrant vs. 0.88 for the SE quadrant; Figs. 3b,f and the last column of Fig. 7c). Such a contradiction can be reconciled by taking into account the SNR of TCGF in the model simulations, as it is one of the two factors determining the correlation skill of the model, according to Eq. (5):

$$r_{\mathrm{en}} \approx \mathrm{RPC} \cdot \sqrt{\hat{\sigma}_{\bar{y}}^2 / \hat{\sigma}_y^2} = \mathrm{RPC} \cdot \sqrt{\frac{1+N^{-1}\widehat{\mathrm{SNR}}^{-2}}{1+\widehat{\mathrm{SNR}}^{-2}}} \approx \mathrm{RPC} \cdot \sqrt{\frac{1}{1+\widehat{\mathrm{SNR}}^{-2}}}, \tag{6}$$

where the last approximation holds for a large $N$ (e.g., 100 in this study).

The last column of Fig. 7b shows the SNR of simulated TCGF in individual sub-basin regions and the entire NWP. TCGF in the SE quadrant has the highest SNR (i.e., 1.58), with the internal variability accounting for only 29% of the total variability. In contrast, the SNR of TCGF in the NW quadrant (i.e., 0.36) is the lowest among the five sub-basin regions, with as much as 89% of the total variability due to the internal variability. It is apparent that the simulated TCGF in the NW quadrant exhibits a much higher level of noise (or a larger disagreement among individual member simulations) than that in the SE quadrant. Accordingly, despite of a higher RPC value, observed TCGF contains more noise in the NW quadrant, leading to a lower skill of the model at replicating the observed TCGF variations there (i.e., a smaller $r_{\mathrm{en}}$; the last column of Fig. 7c).

On the other hand, the SNR of TCGF in the SW quadrant is around 50% higher than that in the NW quadrant (i.e., 0.54 vs. 0.36). However, the much higher noise level in the observations in the SW quadrant (RPC = 0.30 vs. 1.37 in the NW quadrant) ruins the model's ability to capture the observed TCGF variability in this region ($r_{\mathrm{en}}$ = 0.14 vs. 0.48 in the NW quadrant). This result implicates that the SNR, representing the noise level in a model, by itself cannot be used to quantify the skill of the model at reproducing and predicting the observations.

The SNRs of TCGF in the SCS and the NE quadrant are 0.45 and 0.62, respectively, and their respective RPC values are 0.76 and 0.85. Both the relatively low SNR and RPC contribute to the poor performance of the model in capturing the observed TCGF variability in the SCS. The SNR of basin-wide TCGF is 0.92, primarily owing to the low noise level in the SE quadrant.

### *5.3 Number of member simulations needed to skillfully capture the observed TCGF variability*

As discussed earlier, averaging across member simulations can reduce the noise level in the ensemble mean and thereby improve the model's skill at capturing the observed variability. In this subsection, we shall examine the dependence of the model's skill on the number of member simulations. To achieve this, we adopt a random sampling method to independently draw $N$ ($N$ = 1, 2, 3, …, 98, 99, 100) member simulations from the entire 100 member simulations to form an ensemble, and then compute the correlation coefficient between the obtained ensemble mean and the observations. For each choice of $N$, we repeat the procedure 2000 times, yielding a collection of 2000 correlation coefficients. We then visualize the distribution of the collection of the correlation coefficients using a box-and-whisker plot.

Figure 9a shows the results for basin-wide TCGF. As expected, increasing the ensemble size tends to reduce the random variations retained in the ensemble mean, and as a result, narrow down the range of the correlation coefficient and increase its mean value. The mean value increases dramatically when the ensemble size increases from 1 to 10, and converges toward 0.7 (i.e., the correlation coefficient between the observations and the ensemble mean of all 100 member simulations) with a further increase in ensemble size. Overall, an ensemble of 15 simulations is needed to maximize the skill of the model at capturing the observed variability in basin-wide TCGF over the NWP.

A similar pattern can be found in the distribution of the correlation coefficient between the ensemble mean and the observations for individual sub-basin regions, with the range of the correlation coefficient narrowing and the mean value increasing as the ensemble size grows (Figs. 9b–f). A comparison of the six subplots in Fig. 9 reveals two distinct aspects as follows. (i) For a specific ensemble size, the spread of the correlation coefficient is negatively associated with the SNR. (ii) As ensemble size increases, the mean value converges faster when the SNR is higher and/or $r_{\mathrm{en}}$ in Eq. (2) is smaller, with the effect of the SNR dominating over the effect of $r_{\mathrm{en}}$.

Next, we derive a mathematical formula to quantify (ii), as it provides particularly helpful guidance on the designs of numerical experiments in terms of the number of needed ensemble members. Taking the derivative of $r_{\mathrm{en}}$ with respect to $N$ in Eq. (2), we obtain

$$\frac{\mathrm{d}r_{\mathrm{en}}}{\mathrm{d}N} = \frac{\bar{r}\cdot\widehat{\mathrm{SNR}}^{-2}}{2N^2}\sqrt{\frac{1+\widehat{\mathrm{SNR}}^{-2}}{(1+N^{-1}\widehat{\mathrm{SNR}}^{-2})^3}}\ . \tag{7}$$

When $N$ gets bigger, $r_{\mathrm{en}}$ levels off. By continuity, we can always find an integer $N_{\mathrm{min}}$ numerically such that when $N > N_{\mathrm{min}}$ the rate of change in $r_{\mathrm{en}}$ is smaller than $p \cdot r_{\mathrm{en\_max}}$, where $p$ is a predetermined tolerance level (e.g., $p = 2.5 \times 10^{-3}$) and $r_{\mathrm{en\_max}} = \bar{r}\sqrt{1+\widehat{\mathrm{SNR}}^{-2}}$.

When $N^{-1}\widehat{\mathrm{SNR}}^{-2}$ is sufficiently small (a condition often fulfilled in cases with $\widehat{\mathrm{SNR}} > 0.2$), we apply the Taylor series expansion to Eq. (7) and obtain

$$\frac{\mathrm{d}r_{\mathrm{en}}}{\mathrm{d}N} \approx \frac{\bar{r}\cdot\widehat{\mathrm{SNR}}^{-2}\sqrt{1+\widehat{\mathrm{SNR}}^{-2}}}{2N^2+3N\cdot\widehat{\mathrm{SNR}}^{-2}}\,. \tag{8}$$

By setting the derivative in Eq. (8) to $p \cdot \bar{r}\sqrt{1+\widehat{\mathrm{SNR}}^{-2}}$, we obtain

$$N_{\mathrm{min}} = \frac{2}{3p+\sqrt{8p\cdot\widehat{\mathrm{SNR}}^{2}+9p^2}}\,. \tag{9}$$

In practice, we take $N_{\mathrm{min}}$ as the ceiling of the right-hand side of Eq. (9). Figure 10 displays $N_{\mathrm{min}}$ as a function of the SNR for various values of $p$, with solid curves showing numerical solutions based on Eq. (7) and dashed curves corresponding to Eq. (9). As expected, a smaller ensemble size is needed for simulating variables with a higher SNR.

We then proceed to estimate the number of member simulations required to capture the observed TCGF variability in individual sub-basin regions based on Fig. 9 with $p = 2.5 \times 10^{-3}$. It is evident that the required ensemble size differs considerably among the sub-basin regions. Specifically, for the SE quadrant, where the SNR is very large, 10 member simulations are sufficient to replicate the observed variations in TCGF (Fig. 9f). For the NW and NE quadrants, 35 and 20 members are needed, respectively (Figs. 9b,c). These AGCM-based estimations are shown as black symbols in Fig. 10 and in line with our theoretical results (blue curves in Fig. 10). For the SCS, more than 100 simulations will probably yield a correlation skill of the model significant at the 0.05 level (Fig. 9d). For the SW quadrant, increasing the ensemble size does not help improve the skill of the model (Fig. 9e), because of the very high noise level in the observations.

### *5.4 Internal variability and predictability of the large-scale environment and synoptic-scale disturbance activity*

As discussed in sections 4.2 and 4.3, both the large-scale atmospheric environment and synoptic-scale disturbances can modulate TCGF. It is natural to expect that their internal variability contributes to the internal variability in TCGF. In this subsection, we examine the variability and predictability of both the large-scale environment and synoptic-scale disturbances over individual sub-basin regions of the NWP.

The first five columns in Fig. 7b show the SNRs of the GPI and its four components over the five sub-basin regions. The SNR of the GPI is greater than 1 in all sub-basin regions, except the NW quadrant, suggesting the relatively low noise level in the simulated large-scale atmospheric environment. Among the four components of the GPI, thermodynamic factors (i.e., potential intensity and mid-level saturation deficit) have higher SNRs than dynamic factors (i.e., vertical wind shear and low-level vorticity), and vertical wind shear has the highest level of noise among the four components; the exceptions are low-level vorticity and vertical wind shear in the SE quadrant. These results indicate that thermodynamic variables generally have higher similarities across member simulations than dynamic variables, except in the SE quadrant where ENSO exerts strong influences on dynamic fields.

In all individual sub-basin regions, the GPI and the components dominating the forced variability in TCGF (section 4.2; e.g., saturation deficit for the NW quadrant and low-level vorticity for the SE quadrant) have higher SNRs than TCGF (Fig. 7b). This suggests that the large-scale environment has a lower level of noise than TCGF in the simulations and thus contributes relatively little to the large noise in the simulated TCGF, similar to what occurs in the North Atlantic basin (Mei et al. 2019). When considering the variations in the SNR across sub-basin regions, a good correspondence exists between the GPI and TCGF. Specifically, the SNR is the highest in the SE quadrant for both the GPI and TCGF, and the lowest in the NW quadrant.

The RPC values of the GPI are smaller than 1 in all sub-basin regions, except in the NE quadrant, indicating that the large-scale environment has a higher noise level in the observations than in the simulations and that the model is overconfident in predicting it (Fig. 7a). Despite this, the model is still skillful at reproducing the observed variability in the large-scale environment (Fig. 7c) because of the relatively large SNRs in the simulations (Fig. 7b). Among the four components of the GPI, the RPC values of thermodynamic factors are generally larger than those of dynamic factors (Fig. 7a). This, along with the higher SNRs, results in a higher skill of the model at simulating and predicting the thermodynamic factors in the observations (Fig. 7c).

In contrast to the SNR, the RPC of the GPI is not unanimously higher than that of TCGF in individual sub-basin regions (Fig. 7a). Instead, it is higher in the SCS and the SW and NE quadrants but lower in the SE and NW quadrants. In the former three sub-basin regions, the higher RPCs, together with higher SNRs, lead to a higher skill of the model at simulating the observed variability in the large-scale environment than that in TCGF (Fig. 7c). This indicates that factors other than the large-scale environment play a more important role in limiting the model's skill at reproducing the observed TCGF variability in these regions. The lower RPCs in the latter two regions (i.e., the SE and NW quadrants) reduce the model's ability to capture the observed variability in the large-scale environment, making it comparable to the model's skill at replicating that of TCGF in these two regions (Fig. 7c).

Figures 11a,b show the spatial distribution of the RPC and SNR of synoptic-scale disturbance activity over the NWP, respectively. The SNR is relatively large in the deep tropics, particularly the SE quadrant (Fig. 11b). The area with a lower noise level in the observations than in the simulations (i.e., RPC > 1) is located only sporadically over the SE and NW quadrants of the basin (Fig. 11a). As a result, the model shows skills at replicating the observed year-to-year variations in synoptic disturbance activity over the SE quadrant and a small portion of the NW and NE quadrants, but not in the other two sub-basin regions (Fig. 11c).

Based on the results in this subsection and in subsections 4.2 and 4.3, we can reach the following conclusions. (1) In the SE, NW and NE quadrants, the model's skill at replicating the observed large-scale atmospheric environment and synoptic-scale disturbance activity contributes to the model's skill at reproducing the observed TCGF variability (particularly in the SE quadrant). (2) In the SCS and the SW quadrant, the very high noise level in the observed TCGF and synoptic-scale disturbance activity contributes to the weak associations between them and between TCGF and the large-scale atmospheric environment in the observations. The high noise level and these weak associations in turn are largely responsible for the model's poor performance in replicating the observed TCGF variations.

## 6. Summary and Conclusions

Using best track data and a large ensemble of 60-km-resolution atmospheric simulations forced with observed sea surface temperatures (SSTs), this study has examined the variability and predictability of both basin-wide and sub-basin tropical cyclone (TC) genesis frequency (TCGF) in the Northwest Pacific (NWP). The sub-basin regions include the South China Sea (SCS) and the four quadrants of the open ocean that are divided by 144°E and 16°N. The simulations well reproduce the geographical distribution of climatological TC genesis in the observations in terms of both the large-scale pattern and amplitude (Fig. 1). The model is also able to simulate the climatological seasonal cycle of the observed TCGF in the entire NWP and individual sub-basin regions, particularly in the northwest (NW), northeast (NE) and southeast (SE) quadrants (Fig. 2).

The ensemble mean of the simulations is skillful at replicating the year-to-year variability of the observed TCGF in the NW, NE and SE quadrants of the basin (Figs. 3b,c,f), indicating the strong SST control of TC genesis in these sub-basin regions. The model's skill in the SE and NE quadrants is responsible for the model's ability to capture the observed interannual-to-decadal variability in basin-wide TCGF (Fig. 3a), since TCGF of these two sub-basin regions dominates the variability of basin-wide TCGF in both the observations and simulations (Table 2). On the contrary, the ensemble mean shows limited skill at reproducing the observed TCGF variations in the SCS and the southwest (SW) quadrant (Figs. 3d,e), primarily owing to the high noise level and low predictability of TC genesis in the observations over these regions.

We then proceeded to explore the physical mechanisms behind TCGF variability in individual sub-basin regions. In the ensemble mean of the simulations, above-normal TCGF is attributable to increased mid-level relative humidity in the NW quadrant, to increased low-level vorticity and mid-level relative humidity in the NE and SW quadrants and the SCS, and to increased low-level vorticity and reduced vertical wind shear in the SE quadrant (Fig. 5). These favorable large-scale atmospheric conditions, in turn, can be linked to enhanced relative SSTs (i.e., local SSTs minus tropical-mean SST) either locally or to the southeast of the corresponding regions (Fig. 4), which themselves are associated with changes in both local and remote SSTs (e.g., SSTs in the tropical Indian and Atlantic Oceans; Fig. S3 in the online supplemental material). The observations (Fig. 4 and Fig. S5 in the online supplemental material) show results that are statistically significant and consistent with the simulations in the NW, NE and SE quadrants, but insignificant results in the SCS and the SW quadrant, echoing the low predictability of TCGF in the observations over the latter two regions (Figs. 3d,e).

In the ensemble mean, enhanced synoptic-scale disturbance activity also tends to promote TC genesis in all sub-basin regions (Figs. 6f–j). In the observations, however, the effect of synoptic disturbance activity is prominent in the SE and NW quadrants, marginally significant in the NE quadrant, and insignificant in the SCS and the SW quadrant (Figs. 6a–e). The stronger correlations in the ensemble mean are due in part to the fact that a considerable portion of random variations are averaged out in the ensemble mean. The connections between synoptic disturbance activity in individual sub-basin regions and SSTs remain unclear, and are currently being explored using both observations and simulations and will be presented in a follow-up manuscript.

We have also investigated the internal variability and predictability of TCGF in the NWP, taking advantage of the unprecedentedly large ensemble of simulations. We started by comparing the level of noise between the observations and simulations. In the NW and SE quadrants and the entire NWP, TCGF in the simulations has a higher level of noise and thus is less predictable than that in the observations (Figs. 8a,b,f); in other words, the model is overdispersive and underconfident (i.e., RPC > 1; Fig. 7a). In contrast, in the NE and SW quadrants and the SCS, the model is underdispersive and overconfident, and the observed TCGF is less predictable than the simulated TCGF (i.e., RPC < 1; Fig. 7a and Figs. 8c–e).

We then quantified the noise level in the simulations by means of the signal-to-noise ratio (SNR; Fig. 7b). The SNR of TCGF is highest in the SE quadrant, and is smaller than 1 in other regions with the smallest value in the NW quadrant. It is slightly less than 1 for TCGF of the entire basin, primarily owing to the low noise level in the SE quadrant. We further showed that neither the SNR nor RPC alone can be used to quantify the skill of the model at replicating and predicting the observations, as the noise levels in both the simulations and observations are important.

We also assessed the impact of ensemble size on the skill of the model at reproducing the observations using the simulations (Fig. 9) and a theoretical analysis [Eqs. (7),(9)]. The results show that 15 members are sufficient to capture the observed year-to-year variability in basin-wide TCGF over the NWP (Figs. 9a,10). For individual sub-basin regions, 10, 20 and 35 members are needed to replicate the observed TCGF variability in the SE, NE and NW quadrants, respectively (Figs. 9b,c,f,10). For the SCS, more than 100 members would produce a correlation skill marginally significant at the 0.05 level (Fig. 9d). For the SW quadrant, where a very high level of noise is present in the observations (Figs. 7a,b), increasing ensemble size is futile (Fig. 9e). These results provide helpful information on the number of ensemble members needed to capture the observed variability and to obtain reliable predictions. This can be instructive for future designs of numerical experiments that target studying and predicting TCGF in the NWP.

Lastly, we evaluated the internal variability and predictability of the seasonal-mean large-scale atmospheric environment and synoptic-scale disturbance activity and their potential contributions to the internal variability in TCGF. In the simulations, the large-scale environment generally has a SNR greater than 1 and exhibits a lower level of noise than TCGF in all sub-basin regions (Fig. 7b), suggesting relatively less noise in the simulated large-scale environment. Despite the fact that for the large-scale environment the noise level in the observations is higher than that in the simulations (i.e., RPC $< 1$; Fig. 7a), the large SNR leads to good performance of the model in reproducing the observed variability in the large-scale environment (Fig. 7c). A comparison of the model's skill at reproducing the large-scale environment with the model's skill at replicating TCGF suggests that (i) the former substantially contributes to the latter in the SE, NW and NE quadrants; and that (ii) in all sub-basin regions, factors other than the large-scale environment are more important contributors to the internal variability in TCGF, particularly in the SCS and the SW quadrant.

The model also shows skill at reproducing the observed variability in synoptic-scale disturbance activity in the SE quadrant and a small portion of the NW and NE quadrants but not in the other two sub-basin regions (Fig. 11c), contributing to the model's skill at reproducing the observed TCGF variability in the SE, NW and NE quadrants. The relatively high skill of the model at replicating synoptic-scale disturbance activity in the NW quadrant is primarily due to the noise level in the observations being lower than that in the simulations (i.e., RPC > 1; Fig. 11a), whereas in the NE quadrant it is primarily due to the low noise level in the simulations (i.e., relatively high SNR; Fig. 11b). In the SE quadrant, both factors contribute (Figs. 11a,b).

In short, the ensemble mean of the simulations is skillful at reproducing the observed interannual-to-decadal variability of TCGF in the SE, NE and NW quadrants, but shows limited skill in the SCS and the SW quadrant. The remarkably good performance over the SE quadrant (Fig. 3f) is due to (i) the high skill of the model at replicating the observed variability of the large-scale environment and synoptic-scale disturbance activity and (ii) the strong modulation of TCGF by the large-scale environment and synoptic-scale disturbance activity in both the observations and simulations. On the contrary, the particularly poor performance over the SW quadrant (Fig. 3e) is attributed to (i) the weak connections between TCGF and the large-scale environment/synoptic-scale disturbance activity in the observations, which are due to the presence of extremely strong internal variability, and (ii) the inability of the model to reproduce the observed variability in the large-scale environment and synoptic-scale disturbance activity. The low predictability of TCGF in the SCS and the SW quadrant may reduce the predictability of seasonal TC landfalling activity over the Indochina, South China, and the Philippines (Fig. S8 in the online supplemental material). More research is needed to fully understand the variability of TCGF in the SW quadrant as well as in the SCS. We are especially interested in the effects of surrounding landmasses, which could inhibit seeds in these two sub-basin regions from reaching TC state.

As noted in section 2.2, one caveat of the present study is that the employed simulations are Atmospheric Model Intercomparison Project (AMIP)-type simulations and the missing air-sea interaction can lead to biases in surface energy fluxes, which in turn may affect the simulated TC activity (particularly the intensity). Simulations with coupled models would mitigate this issue, though they prevent an accurate quantification of the internal variability induced by atmospheric processes. In addition, the results presented here are based on simulations produced by one model. It would be desirable to test our results with other AGCMs of similar and/or higher resolutions.

*Acknowledgements*

This work was supported by a startup fund from the University of North Carolina at Chapel Hill and National Science Foundation Grant AGS-2047721. This study used d4PDF produced with the Earth Simulator jointly by science programs (SOUSEI, TOUGOU, SI-CAT, DIAS) of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. We thank the editor and three anonymous reviewers for their comments, which greatly helped improve the manuscript.

**References**

Amaya, D. J., 2019: The Pacific Meridional Mode and ENSO: a Review. *Curr. Clim. Change Rep.*, **5**, 296–307, https://doi.org/10.1007/s40641-019-00142-x.

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007.

Bacmeister, J. T., K. A. Reed, C. Hannay, P. Lawrence, S. Bates, J. E. Truesdale, N. Rosenbloom, and M. Levy, 2018: Projected changes in tropical cyclone activity under future warming scenarios using a high-resolution climate model. *Climatic Change*, **146**, 547–560, https://doi.org/10.1007/s10584-016-1750-x.

Barcikowska, M., F. Feser, W. Zhang, and W. Mei, 2017: Changes in intense tropical cyclone activity for the western North Pacific during the last decades derived from a regional climate model simulation. *Clim. Dyn.,* **49,** 2931–2949.

Camargo, S. J., and A. H. Sobel, 2005: Western North Pacific tropical cyclone intensity and ENSO. *J. Climate*, **18**, 2996–3006, https://doi.org/10.1175/JCLI3457.1.

Camargo, S. J., A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007a: Cluster analysis of typhoon tracks. Part I: General properties. *J. Climate*, **20**, 3635–3653.

Camargo, S. J., A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007b: Cluster analysis of typhoon tracks. Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654–3676.

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

Capotondi, A., and P. D. Sardeshmukh, 2015: Optimal precursors of different types of ENSO events. *Geophys. Res. Lett.*, **42**, 9952–9960, https://doi.org/10.1002/2015GL066171.

Chan, J. C. L., 1985: Tropical cyclone activity in the northwest Pacific in relation to the El Niño/Southern Oscillation phenomenon. *Mon. Wea. Rev.*, **113**, 599–606, https://doi.org/10.1175/1520-0493(1985)113,0599:TCAITN.2.0.CO;2.

Chan, J. C. L., and J. E. Shi, 1996: Long-term trends and interannual variability in tropical cyclone activity over the western North Pacific. *Geophys. Res. Lett.*, **23**, 2765–2767.

Chen, G., and C. Y. Tam, 2010: Different impacts of two kinds of Pacific Ocean warming on tropical cyclone frequency over the western North Pacific. *Geophys. Res. Lett.*, **37**, L01803, https://doi.org/10.1029/2009GL041708.

Chen, T. C., S. Y. Wang, and M. C. Yen, 2006: Interannual variation of tropical cyclone activity over the western North Pacific. *J. Climate*, **19**, 5709–5720, https://doi.org/10.1175/JCLI3934.1.

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

Choi, Y., K. J. Ha, C. H. Ho, and C. E. Chung, 2015: Interdecadal change in typhoon genesis condition over the western North Pacific. *Climate Dyn.*, **45**, 3243–3255, https://doi.org/10.1007/s00382-015-2536-y.

Chu, J.-H., C. R. Sampson, A. S. Levine, and E. Fukada, 2002: The joint typhoon warning center tropical cyclone best-tracks, 1945–2000. Naval Research Laboratory Rep. NRL/MR/7540-02-16, 22 pp.

Clark, J. D., and P.-S. Chu, 2002: Interannual variation of tropical cyclone activity over the central North Pacific. *J. Meteor. Soc. Japan*, **80**, 403–418, https://doi.org/10.2151/jmsj.80.403.

DelSole, T., and M. K. Tippett, 2007: Predictability: Recent insights from information theory. *Rev. Geophys.*, **45**, RG4002.

Du, Y., L. Yang, and S.-P. Xie, 2011: Tropical Indian Ocean influence on northwest Pacific tropical cyclones in summer following strong El Niño. *J. Climate*, **24**, 315–322, https://doi.org/10.1175/2010JCLI3890.1.

Eade, R., D. Smith, A. Scaife, E. Wallace, N. Dunstone, L. Hermanson, and N. Robinson, 2014: Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.*, **41**, 5620–5628, https://doi.org/10.1002/2014GL061146.

Emanuel, K. A., 2010: Tropical cyclone activity downscaled from NOAA-CIRES reanalysis, 1908–1958. *J. Adv. Model. Earth Syst.*, **2**, https://doi.org/10.3894/JAMES.2010.2.1.

Emanuel, K. A., and A. Sobel, 2013: Response of tropical sea surface temperature, precipitation, and tropical cyclone-related variables to changes in global and local forcing. *J. Adv. Model. Earth Syst.*, **5**, 447–458, https://doi.org/10.1002/jame.20032.

Fu, B., T. Li, M. S. Peng, and F. Weng, 2007: Analysis of tropical cyclogenesis in the western North Pacific for 2000 and 2001. *Wea. Forecasting*, **22**, 763–780.

Fu, B., M. S. Peng, T. Li, and D. E. Stevens, 2012: Developing versus nondeveloping disturbances for tropical cyclone formation. Part II: Western North Pacific. *Mon. Wea. Rev.*, **140**, 1067–1080, https://doi.org/10.1175/2011MWR3618.1.

Gao, S., Z. Chen, and W. Zhang, 2018: Impacts of tropical North Atlantic SST on western North Pacific landfalling tropical cyclones. *J. Climate*, **31**, 853–862, https://doi.org/10.1175/JCLI-D-17-0325.1.

Ha, Y., Z. Zhong, X. Yang, and Y. Sun, 2015: Contribution of East Indian Ocean SSTA to western North Pacific tropical cyclone activity under El Niño/La Niña conditions. *Int. J. Climatol.*, **35**, 506–519, https://doi.org/10.1002/joc.3997.

Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1.

Huo, L., P. Guo, S. N. Hameed, and D. Jin, 2015: The role of tropical Atlantic SST anomalies in modulating western North Pacific tropical cyclone genesis. *Geophys. Res. Lett.*, **42**, 2378–2384, https://doi.org/10.1002/2015GL063184.

Kamae, Y., W. Mei, S.-P. Xie, M. Naoi, and H. Ueda, 2017a: Atmospheric rivers over the northwestern Pacific: Climatology and interannual variability. *J. Climate*, **30**, 5605–5619.

Kamae, Y., and Coauthors, 2017b: Forced response and internal variability of summer climate over western North America. *Climate Dyn.*, **49**, 403–417.

Kang, I.-S., and J. Shukla, 2006: Dynamic seasonal prediction and predictability of the monsoon. *The Asian Monsoon*, B. Wang, Ed., Springer Praxis, 585–612.

Kao, H.-Y, and J.-Y. Yu, 2009: Contrasting eastern-Pacific and central-Pacific types of ENSO. *J. Climate*, **22**, 615–632, https:// doi.org/10.1175/2008JCLI2309.1.

Kim, H.-K., and K.-H. Seo, 2016: Cluster analysis of tropical cyclone tracks over the western North Pacific using a self-organizing map. *J. Climate*, **29**, 3731–3751, https://doi.org/10.1175/JCLI-D-15-0380.1.

Kim, H.-M., P. J. Webster, and J. A. Curry, 2011: Modulation of North Pacific tropical cyclone activity by three phases of ENSO. *J. Climate*, **24**, 1839–1849, https://doi.org/10.1175/ 2010JCLI3939.1.

Knutson, T. R., and Coauthors, 2010: Tropical cyclones and climate change. *Nat. Geosci.*, **3**, 157–163.

Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, https://doi.org/10.2151/jmsj.2015-001.

Kossin, J. P., K. A. Emanuel, and S. J. Camargo, 2016: Past and projected changes in western North Pacific tropical cyclone exposure. *J. Climate*, **29**, 5725–5739, https://doi.org/10.1175/JCLI-D-16-0076.1.

Korty, R. L., S. J. Camargo, and J. Galewsky, 2012: Tropical cyclone genesis factors in simulations of the Last Glacial Maximum. *J. Climate*, **25**, 4348–4365, https://doi.org/10.1175/JCLI-D-11-00517.1.

Lander, M., 1994: An exploratory analysis of the relationship between tropical storm formation in the western North Pacific and ENSO. *Mon. Wea. Rev.*, **122**, 636–651, https://doi.org/10.1175/1520-0493(1994)122,0636:AEAOTR.2.0.CO;2.

Larkin, N. K., and D. E. Harrison, 2005: Global seasonal temperature and precipitation anomalies during El Niño autumn and winter. *Geophys. Res. Lett.*, **32**, L16705.

Larson, S. M., and B. P. Kirtman, 2014: The Pacific meridional mode as an ENSO precursor and predictor in the North American multimodel ensemble. *J. Climate*, **27**, 7018–7032.

Lee, T.-C., T. R. Knutson, T. Nakaegawa, M. Ying, and E. J. Cha, 2020: Third assessment on impacts of climate change on tropical cyclones in the Typhoon Committee Region – Part I: Observed changes, detection and attribution. *Tropical Cyclone Research and Review*, **9**, 1–22, https://doi.org/10.1016/j.tcrr.2020.03.001.

Li, H., and R. L. Sriver, 2018: Tropical cyclone activity in the high resolution Community Earth System Model and the impact of ocean coupling. *J. Adv. Model. Earth Syst.*, **10**, 165–186.

Li, S., W. Mei, and S.-P. Xie, 2022: Effects of tropical sea surface temperature variability on Northern Hemisphere tropical cyclone genesis. *J. Climate*, in press.

Li, T., M. Kwon, M. Zhao, J.-S. Kug, J.-J. Luo, and W. Yu, 2010: Global warming shifts Pacific tropical cyclone location. *Geophys. Res. Lett.*, **37**, L21804.

Li, T., 2012: *Synoptic and Climatic Aspects of Tropical Cyclogenesis in Western North Pacific in Cyclones: Formation, Triggers, and Control*. K. Oouchi and H. Fudevasu, Eds., Noval Science Publishers, 61–94.

Liu, C., W. Zhang, M. Stuecker, and F.-F. Jin, 2019: Pacific Meridional Mode-Western North Pacific Tropical Cyclone Linkage Explained by Tropical Pacific Quasi-Decadal Variability. *Geophys. Res. Lett.*, **46**, 13346–13354.

Liu, K. S., and J. C. L. Chan, 2003: Climatological characteristics and seasonal forecasting of tropical cyclones making landfall along the south China coast. *Mon. Wea. Rev.*, **131**, 1650–1662, https://doi.org/10.1175//2554.1.

Liu, K. S., and J. C. L. Chan, 2013: Inactive period of western North Pacific tropical cyclone activity in 1998–2011. *J. Climate*, **26**, 2614–2630.

Liu, Y., and G. H. Chen, 2018: Intensified influence of the ENSO Modoki on boreal summer tropical cyclone genesis over the western North Pacific since the early 1990s. *Int. J. Climatol.*, **38**, e1258–e1265. https://doi.org/10.1002/joc.5347.

Mei, W., S.-P. Xie, and M. Zhao, 2014: Variability of tropical cyclone track density in the North Atlantic: Observations and high-resolution simulations. *J. Climate*, **27**, 4797–4814, https://doi.org/10.1175/JCLI-D-13-00587.1.

Mei, W., S.-P. Xie, M. Zhao, and Y. Wang, 2015: Forced and internal variability of tropical cyclone track density in the western North Pacific. *J. Climate*, **28**, 143–167, https://doi.org/10.1175/JCLI-D-14-00164.1.

Mei, W., and S.-P. Xie, 2016: Intensification of landfalling typhoons over the northwest Pacific since the late 1970s. *Nat. Geosci.*, **9**, 753–757, https://doi.org/10.1038/ngeo2792.

Mei, W., Y. Kamae, S.-P. Xie, and K. Yoshida, 2019: Variability and Predictability of North Atlantic Hurricane Frequency in a Large Ensemble of High-Resolution Atmospheric Simulations. *J. Climate,* **32**, 3153–3167, https://doi.org/10.1175/JCLI-D-18-0554.1.

Mizuta, R., and Coauthors, 2017: Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bull. Amer. Meteor. Soc.*, **98**, 1383–1398, https://doi.org/10.1175/BAMS-D-16-0099.1.

Murakami, H., R. Mizuta, and E. Shindo, 2012: Future changes in tropical cyclone activity projected by multi-physics and multi-SST ensemble experiments using the 60-km-mesh MRIAGCM. *Climate Dyn.*, **39**, 2569–2584, https://doi.org/10.1007/s00382-011-1223-x.

Nakamura, J., S. J. Camargo, A. H. Sobel, N. Henderson, K. A. Emanuel, A. Kumar, T. E. LaRow, H. Murakami, M. J. Roberts, E. Scoccimarro, P. L. Vidale, H. Wang, M. F. Wehner, and M. Zhao, 2017: Western North Pacific tropical cyclone model tracks in present and future climates. *J. Geophys. Res. Atmos.*, **122**, 9721–9744.

Patricola, C. M., S. J. Camargo, P. J. Klotzbach, R. Saravanan, and P. Chang, 2018: The influence of ENSO flavors on western North Pacific tropical cyclone activity. *J. Climate*, **31**, 5395–5416, https://doi.org/10.1175/JCLI-D-17-0678.1.

Sugi, M., Y. Yamada, K. Yoshida, R. Mizuta, M. Nakano, C. Kodama, and M. Satoh, 2020: Future changes in the global frequency of tropical cyclone seeds. *SOLA*, **16**, 70–74.

Tang, B. H., and K. Emanuel, 2012: A ventilation index for tropical cyclones. *Bull. Amer. Meteor. Soc.*, **93**, 1901–1912, https://doi.org/10.1175/BAMS-D-11-00165.1.

Tao, L., L. Wu, Y. Wang, and J. Yang, 2012: Influence of tropical Indian Ocean warming and ENSO on tropical cyclone activity over the western North Pacific. *J. Meteor. Soc. Japan*, **90**, 127–144, https://doi.org/10.2151/jmsj.2012-107.

Ueda, H., K. Miwa, and Y. Kamae, 2018: Seasonal modulation of tropical cyclone occurrence associated with coherent Indo-Pacific variability during decaying phase of El Niño. *J. Meteor. Soc. Japan*, **96**, 381–390, https://doi.org/10.2151/jmsj.2018-044.

Vecchi, G. A., and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Climate*, **27**, 7994–8016, https://doi.org/10.1175/JCLI-D-14-00158.1.

Vecchi, G. A., and Coauthors, 2019: Tropical cyclone sensitivities to $CO_2$ doubling: Roles of atmospheric resolution, synoptic variability and background climate changes. *Climate Dyn.*, **53**, 5999–6033, https://doi.org/10.1007/s00382-019-04913-y.

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.

Wang, B., 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, https://doi.org/10.1073/pnas.1214626110.

Wang, C., and B. Wang, 2019: Tropical cyclone predictability shaped by western Pacific subtropical high: Integration of trans-basin sea surface temperature effects. *Climate Dyn.*, **53**, 2697–2714, https://doi.org/10.1007/s00382-019-04651-1.

Woodruff, J. D., J. L. Irish, and S. J. Camargo, 2013: Coastal flooding by tropical cyclones and sea-level rise. *Nature*, **504**, 44–52, https://doi.org/10.1038/nature12855.

Wu, L., H. Zhang, J.-M. Chen, and T. Feng, 2018: Impact of two types of El Niño on tropical cyclones over the western North Pacific: Sensitivity to location and intensity of Pacific warming. *J. Climate*, **31**, 1725–1742, https://doi.org/10.1175/JCLI-D-17-0298.1.

Wu, R., Y. Yang, and X. Cao, 2019: Respective and combined impacts of regional SST anomalies on tropical cyclogenesis in different sectors of the western North Pacific. *J. Geophys. Res. Atmos.*, **124**, 8917–8934, https://doi.org/10.1029/2019JD030736.

Wu, R., X. Cao, and Y. Yang, 2020: Interdecadal change in the relationship of the western North Pacific tropical cyclogenesis frequency to tropical Indian and North Atlantic Ocean SST in early 1990s. *J. Geophys. Res. Atmos.*, **125**, e2019JD031493.

Xie, S.-P., K. Hu, J. Hafner, Y. Du, G. Huang, and H. Tokinaga, 2009: Indian Ocean capacitor effect on Indo-western Pacific climate during the summer following El Niño. *J. Climate*, **22**, 730–747.

Xie, S.-P., Y. Kosaka, Y. Du, K. Hu, J. S. Chowdary, and G. Huang, 2016: Indo-western Pacific ocean capacitor and coherent climate anomalies in post-ENSO summer: A review. *Adv. Atmos. Sci.*, **33**, 411–432, https://doi.org/10.1007/s00376-015-5192-6.

Yamada, Y., and Coauthors, 2021: Evaluation of the contribution of tropical cyclone seeds to changes in tropical cyclone frequency due to global warming in high-resolution multi-model ensemble simulations. *Prog. Earth Planet. Sci.*, **8**, 11.

Ying, M., W. Zhang, H. Yu, X. Lu, J. Feng, Y. Fan, Y. Zhu, and D. Chen, 2014: An overview of the China Meteorological Administration tropical cyclone database. *J. Atmos. Oceanic Technol.*, **31**, 287–301, doi:10.1175/JTECH-D-12-00119.1.

Yoshida, K., M. Sugi, R. Mizuta, H. Murakami, and M. Ishii, 2017: Future changes in tropical cyclone activity in high-resolution large-ensemble simulations. *Geophys. Res. Lett.*, **44**, 9910–9917, https://doi.org/10.1002/2017GL075058.

Yoshimura, J., and M. Sugi, 2005: Tropical cyclone climatology in a high-resolution AGCM—Impacts of SST warming and $CO_2$ increase. *SOLA*, **1**, 133–136, doi:10.2151/sola.2005-035.

Yoshimura, J., M. Sugi, and A. Noda, 2006: Influence of greenhouse warming on tropical cyclone frequency. *J. Meteor. Soc. Japan*, **84**, 405–428.

Yu, J., T. Li, Z. Tan, and Z. Zhu, 2016: Effects of tropical North Atlantic SST on tropical cyclone genesis in the western North Pacific. *Climate Dyn.*, **46**, 865–877.

Zhan, R., Y. Wang, and X. Lei, 2011: Contributions of ENSO and east Indian Ocean SSTA to the interannual variability of northwest Pacific tropical cyclone frequency. *J. Climate*, **24**, 509–521, https://doi.org/10.1175/2010JCLI3808.1.

Zhan, R., Y. Wang, and J. Zhao, 2019: Contributions of SST anomalies in the Indo-Pacific Ocean to the interannual variability of tropical cyclone genesis frequency over the Western North Pacific. *J. Climate*, **32**, 3357–3372, https://doi.org/10.1175/JCLI-D-18-0439.1.

Zhang, Q., Q. Liu, and L. Wu, 2009: Tropical cyclone damages in China 1983–2006. *Bull. Amer. Meteor. Soc.*, **90**, 489–495, https://doi.org/10.1175/2008BAMS2631.1.

Zhang, W., G. A. Vecchi, H. Murakami, G. Villarini, and L. Jia, 2016: The Pacific meridional mode and the occurrence of tropical cyclones in the western North Pacific. *J. Climate*, **29**, 381–398, https://doi.org/10.1175/JCLI-D-15-0282.1.

Zhang, W., G. A. Vecchi, G. Villarini, H. Murakami, A. Rosati, X. Yang, L. Jia, and F. Zeng, 2017: Modulation of western North Pacific tropical cyclone activity by the Atlantic meridional mode. *Climate Dyn.*, **48**, 631–647, https://doi.org/10.1007/s00382-016-3099-2.

Zhao, H., and C. Wang, 2019: On the relationship between ENSO and tropical cyclones in the western North Pacific during the boreal summer. *Climate Dyn.*, **52**, 275–288.

Zong, H., and L. Wu, 2015: Synoptic-scale influences on tropical cyclone formation within the western North Pacific monsoon trough. *Mon. Wea. Rev.*, **143**, 3421–3433.

**List of Tables**

Table 1: List of acronyms and abbreviations used in this paper.

Table 2: Pairwise correlation coefficients of TCGF among different regions: the entire basin (NWP), the NW quadrant (NW), the NE quadrant (NE), the SCS (SCS), the SW quadrant (SW), and the SE quadrant (SE). The numbers outside and inside the parentheses show the results for the observations and the ensemble mean of the simulations, respectively. The numbers in bold are the correlation coefficients significant at the 0.01 level.

**List of Figures**

Figure 1: Geographical distribution of the climatological TC genesis (per year in 2°×2° grids) in the NWP in (a) the observations and (b) the ensemble mean of the simulations. Dashed blue lines delineate the divisions of the five sub-basin regions: the northwest quadrant (NW), the northeast quadrant (NE), the South China Sea (SCS), the southwest quadrant (SW), and the southeast quadrant (SE).

Figure 2: Simulated and observed seasonal evolutions of the climatological TCGF in (a) the entire NWP and (b–f) its five sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, thin gray curves show the results for individual member simulations and thick black curve shows the results for the ensemble mean of the simulations; thin red curves show the results for three different best track data sets (section 2.1) and the thick red curve shows their average. Note that the scale of $y$ axis in (a) is different from that in (b)–(f).

Figure 3: Temporal evolution of the simulated (ensemble mean; thick black) and observed (mean of the best track data; thick red) annual TCGF in (a) the entire NWP and (b–f) its five sub-basin

regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, thin gray curves show the results for individual member simulations, thin red curves show the results for individual best track data sets, and the number in the upper-right corner shows the correlation coefficient between the thick red and black curves. Note that the scale of $y$ axis in (a) is different from that in (b)–(f).

Figure 4: Correlation coefficients between relative SSTs and TCGF of the entire NWP or one of its sub-basin regions (color shading) and coefficients of 850-hPa winds regressed on the normalized TCGF (vectors; m $s^{-1}$; the regressions are performed separately for the zonal and meridional components and then the regression coefficients from the two regressions form the vectors). (a)–(c),(g)–(i) show the results for the observations, and (d)–(f),(j)–(l) are for the ensemble mean of the simulations. (a),(d) show the results for the entire NWP; (b),(e) for the NW quadrant; (c),(f) for the NE quadrant; (g),(j) for the SCS; (h),(k) for the SW quadrant; and (i),(l) for the SE quadrant. SSTs and winds are averaged over the peak season of individual regions (e.g., June–November for the entire NWP; section 3). Green curves delineate the areas where relative SSTs are significantly correlated with TCGF at the 0.05 level. Wind vectors are shown only in the areas where the regression coefficients are significant at the 0.1 level and greater than 0.05 m $s^{-1}$ in magnitude.

Figure 5: Coefficients of the GPI and of the respective contributions of its four components (i.e., potential intensity $V_{\mathrm{PI}}$, mid-level saturation deficit $\chi$, vertical wind shear $V_{\mathrm{sh}}$ and low-level vorticity $\eta$) regressed on the normalized TCGF of (a) the NW quadrant, (b) the NE quadrant, (c) the SCS, (d) the SW quadrant, and (e) the SE quadrant in the ensemble mean of the simulations. Black dots denote the areas where the regression coefficients are significant at the 0.05 level.

Figure 6: Correlation coefficients between synoptic-scale disturbance activity and TCGF of (a),(f) the NW quadrant, (b),(g) the NE quadrant, (c),(h) the SCS, (d),(i) the SW quadrant and (e),(j) the SE quadrant. (a)–(e) show the results for the observations, and (f)–(j) are for the ensemble mean of the simulations. Black dots denote the areas where the correlation coefficients are significant at the 0.05 level.

Figure 7: (a) Bar plots of the RPC of the GPI and its four components (i.e., potential intensity $V_{\mathrm{PI}}$, mid-level saturation deficit $\chi$, vertical wind shear $V_{\mathrm{sh}}$ and low-level vorticity $\eta$) as well as the TCGF over individual sub-basin regions (shown in different colors). For the TCGF, the result for the entire basin (cyan) is also displayed. (b) As in (a), but for the SNR. (c) As in (a), but for the correlation coefficient between the observations and the ensemble mean of the 100 simulations. Black dots show the mean value of the correlation coefficients in individual member simulations, with vertical black lines denoting $\pm$one standard deviation.

Figure 8: Histograms of the correlation coefficients between TCGF in one member simulation and that in the other 99 member simulations (gray curves), between TCGF in the observations and that in the 100 member simulations (red curve), and between TCGF in one member simulation and the ensemble mean of the 100 member simulations (blue curve) for (a) the entire NWP and (b–f) individual sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, black curve shows the average of the gray curves, and vertical dotted magenta line shows the correlation coefficient between TCGF in the observations and that in the ensemble mean of the 100 member simulations.

Figure 9: Box-and-whisker plots of the correlation coefficients between TCGF in the simulations and that in the observations as a function of ensemble size for (a) the entire NWP and (b–f) individual sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and

the SE quadrant. In each panel, green dots show the theoretical results based on Eq. (2), and the horizontal dashed purple line denotes the critical value of the correlation coefficient significant at the 0.05 level.

Figure 10: Number of ensemble members needed to reach the level-off of the model skill at reproducing observed/SST-forced TCGF variability versus the SNR of simulated TCGF for various predetermined tolerance levels ($p$): $p = 0.001$ (red), $p = 0.0025$ (blue), $p = 0.005$ (green), and $p = 0.01$ (magenta). Solid curves show numerical solutions based on Eq. (7) and dashed curves corresponding to Eq. (9). Black symbols show the empirical results based on the AGCM simulations with $p = 0.0025$: the circle for the SE quadrant, the square for the entire NWP, the diamond for the NE quadrant, the star for the SCS, and the triangle for the NW quadrant; the result for the SW quadrant is not shown because of the very low model skill in this sub-basin region.

Figure 11: (a) Spatial distribution of the RPC of synoptic-scale disturbance activity during the NWP TC season (i.e., June–November). Green contours delineate the areas with the RPC equal to 1. (b) As in (a), but for the SNR. Blue and green contours delineate the areas with the SNR equal to 0.5 and 1, respectively. (c) As in (a), but for the correlation coefficient between the observations and the ensemble mean of the 100 member simulations. Black dots denote the areas where the correlation coefficients are significant at the 0.05 level.

**Tables**

Table 1: List of acronyms and abbreviations used in this paper.

| | | | |
|---|---|---|---|
| **AGCM** | Atmospheric General Circulation Model | **RPC** | Ratio of predictable component |
| **AMIP** | Atmospheric Model Intercomparison Project | **SCS** | South China Sea |
| **d4PDF** | Database for Policy Decision Making for Future Climate Change | **SE** | Southeast |
| **ENSO** | El Niño–Southern Oscillation | **SNR** | Signal-to-noise ratio |
| **GPI** | Genesis potential index | **SSTs** | Sea surface temperatures |
| **NE** | Northeast | **SW** | Southwest |
| **NW** | Northwest | **TCGF** | Tropical cyclone genesis frequency |
| **NWP** | Northwest Pacific | **TCs** | Tropical cyclones |
| **PMM** | Pacific Meridional Mode | | |

Table 2: Pairwise correlation coefficients of TCGF among different regions: the entire basin (NWP), the NW quadrant (NW), the NE quadrant (NE), the SCS (SCS), the SW quadrant (SW), and the SE quadrant (SE). The numbers outside and inside the parentheses show the results for the observations and the ensemble mean of the simulations, respectively. The numbers in bold are the correlation coefficients significant at the 0.01 level.

| Regions | NW | NE | SCS | SW | SE |
|---|---|---|---|---|---|
| NWP | -0.03 (-0.09) | **0.54** (**0.52**) | 0.11 (-0.23) | 0.29 (0.27) | **0.56** (**0.78**) |
| NW | | 0.24 (0.35) | -0.05 (**0.53**) | -0.10 (0.17) | **-0.54** (**-0.48**) |
| NE | | | 0.06 (0.31) | 0.08 (0.33) | -0.10 (0.01) |
| SCS | | | | -0.02 (0.37) | -0.25 (**-0.69**) |
| SW | | | | | -0.16 (-0.21) |

# Figures

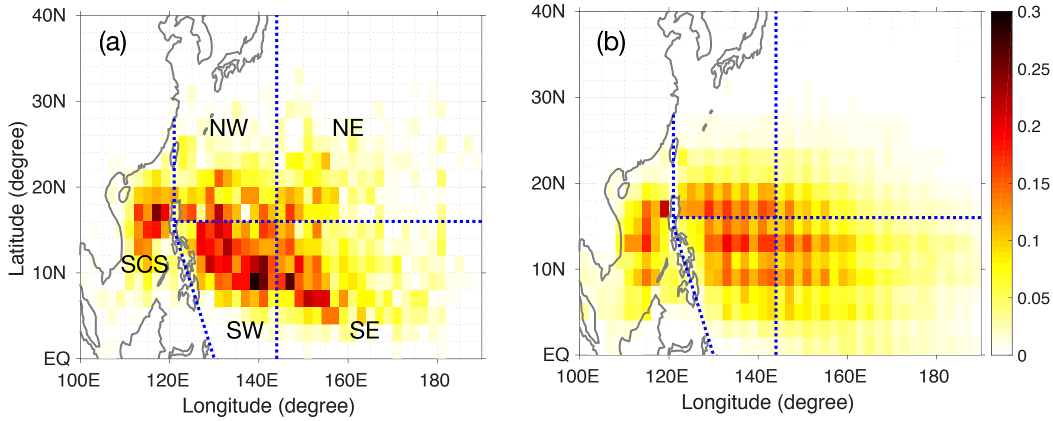


Figure 1: Geographical distribution of the climatological TC genesis (per year in 2°×2° grids) in the NWP in (a) the observations and (b) the ensemble mean of the simulations. Dashed blue lines delineate the divisions of the five sub-basin regions: the northwest quadrant (NW), the northeast quadrant (NE), the South China Sea (SCS), the southwest quadrant (SW), and the southeast quadrant (SE).

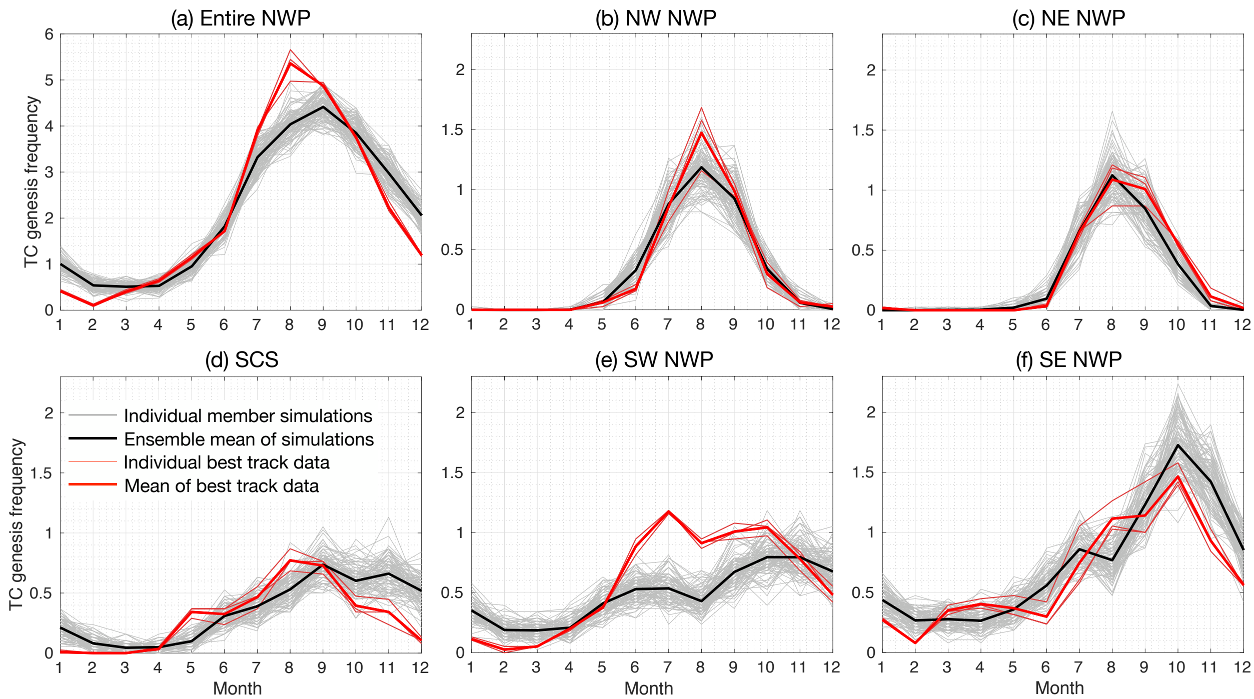


Figure 2: Simulated and observed seasonal evolutions of the climatological TCGF in (a) the entire NWP and (b–f) its five sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, thin gray curves show the results for individual member simulations and thick black curve shows the results for the ensemble mean of the simulations; thin red curves show the results for three different best track data sets (section 2.1) and the thick red curve shows their average. Note that the scale of $y$ axis in (a) is different from that in (b)–(f).

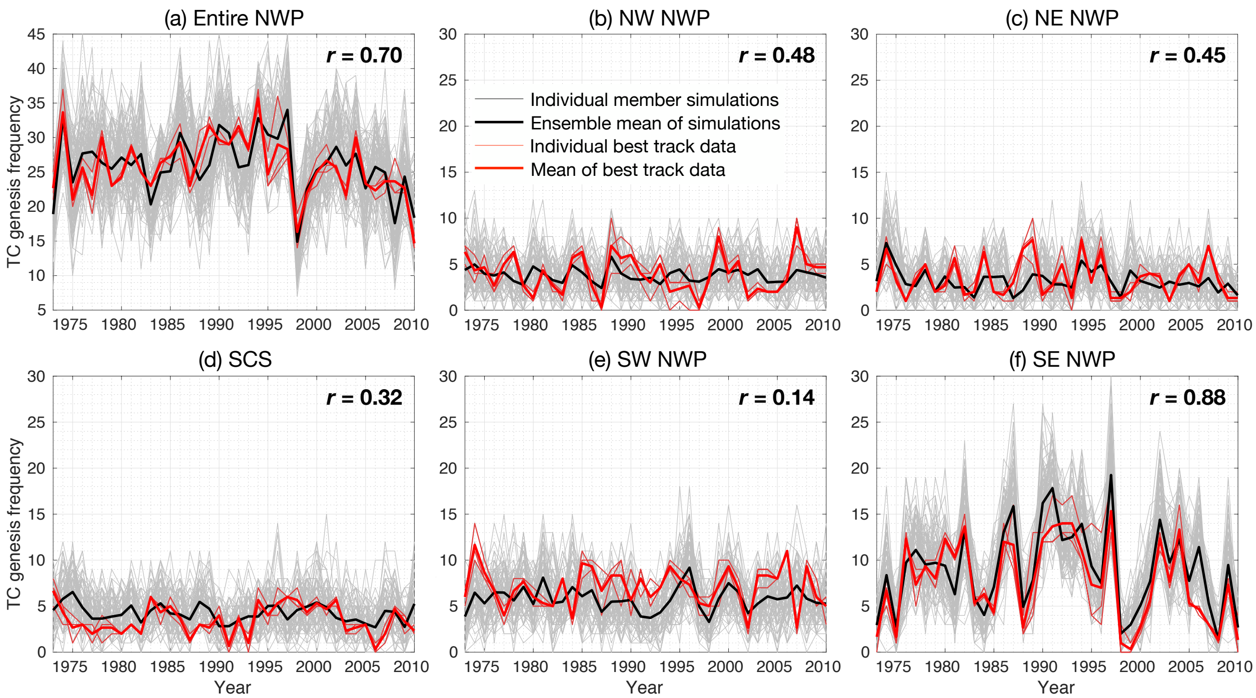


Figure 3: Temporal evolution of the simulated (ensemble mean; thick black) and observed (mean of the best track data; thick red) annual TCGF in (a) the entire NWP and (b–f) its five sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, thin gray curves show the results for individual member simulations, thin red curves show the results for individual best track data sets, and the number in the upper-right corner shows the correlation coefficient between the thick red and black curves. Note that the scale of $y$ axis in (a) is different from that in (b)–(f).

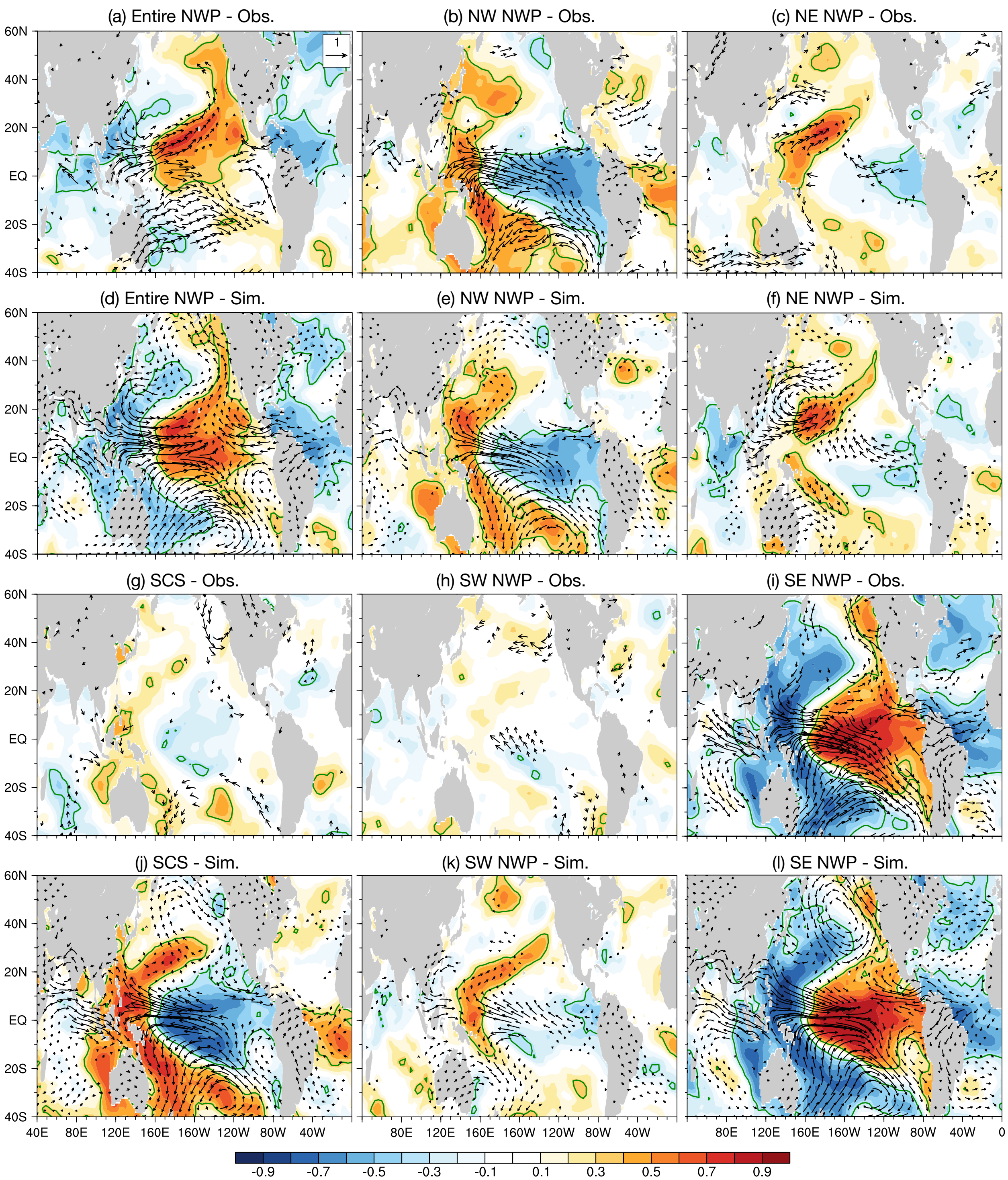


Figure 4: Correlation coefficients between relative SSTs and TCGF of the entire NWP or one of its sub-basin regions (color shading) and coefficients of 850-hPa winds regressed on the normalized TCGF (vectors; m $s^{-1}$; the regressions are performed separately for the zonal and

meridional components and then the regression coefficients from the two regressions form the vectors). (a)–(c),(g)–(i) show the results for the observations, and (d)–(f),(j)–(l) are for the ensemble mean of the simulations. (a),(d) show the results for the entire NWP; (b),(e) for the NW quadrant; (c),(f) for the NE quadrant; (g),(j) for the SCS; (h),(k) for the SW quadrant; and (i),(l) for the SE quadrant. SSTs and winds are averaged over the peak season of individual regions (e.g., June–November for the entire NWP; section 3). Green curves delineate the areas where relative SSTs are significantly correlated with TCGF at the 0.05 level. Wind vectors are shown only in the areas where the regression coefficients are significant at the 0.1 level and greater than 0.05 m $s^{-1}$ in magnitude.

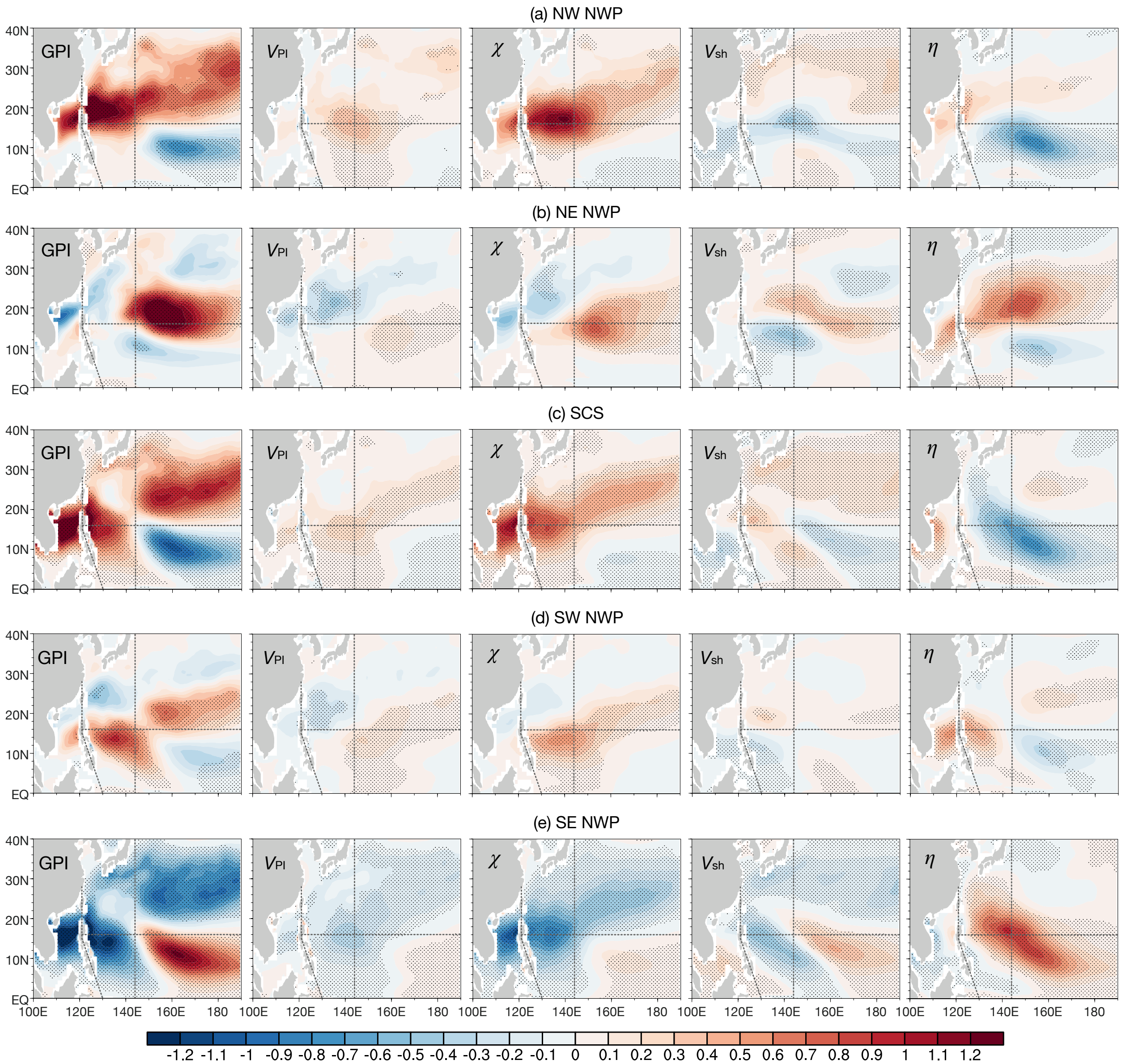


Figure 5: Coefficients of the GPI and of the respective contributions of its four components (i.e., potential intensity $V_{\mathrm{PI}}$, mid-level saturation deficit $\chi$, vertical wind shear $V_{\mathrm{sh}}$ and low-level vorticity $\eta$) regressed on the normalized TCGF of (a) the NW quadrant, (b) the NE quadrant, (c) the SCS, (d) the SW quadrant, and (e) the SE quadrant in the ensemble mean of the simulations. Black dots denote the areas where the regression coefficients are significant at the 0.05 level.

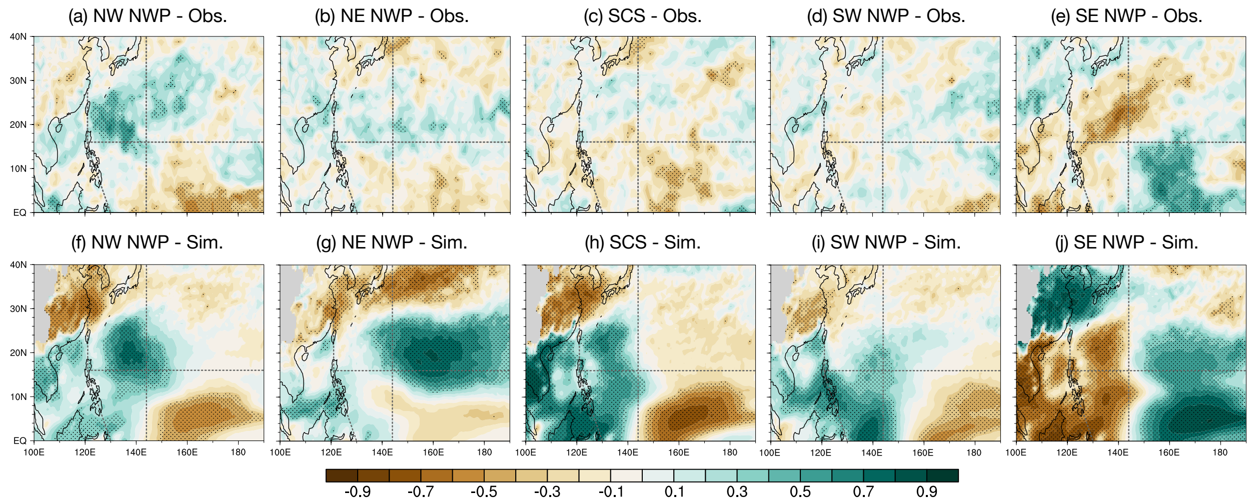


Figure 6: Correlation coefficients between synoptic-scale disturbance activity and TCGF of (a),(f) the NW quadrant, (b),(g) the NE quadrant, (c),(h) the SCS, (d),(i) the SW quadrant and (e),(j) the SE quadrant. (a)–(e) show the results for the observations, and (f)–(j) are for the ensemble mean of the simulations. Black dots denote the areas where the correlation coefficients are significant at the 0.05 level.

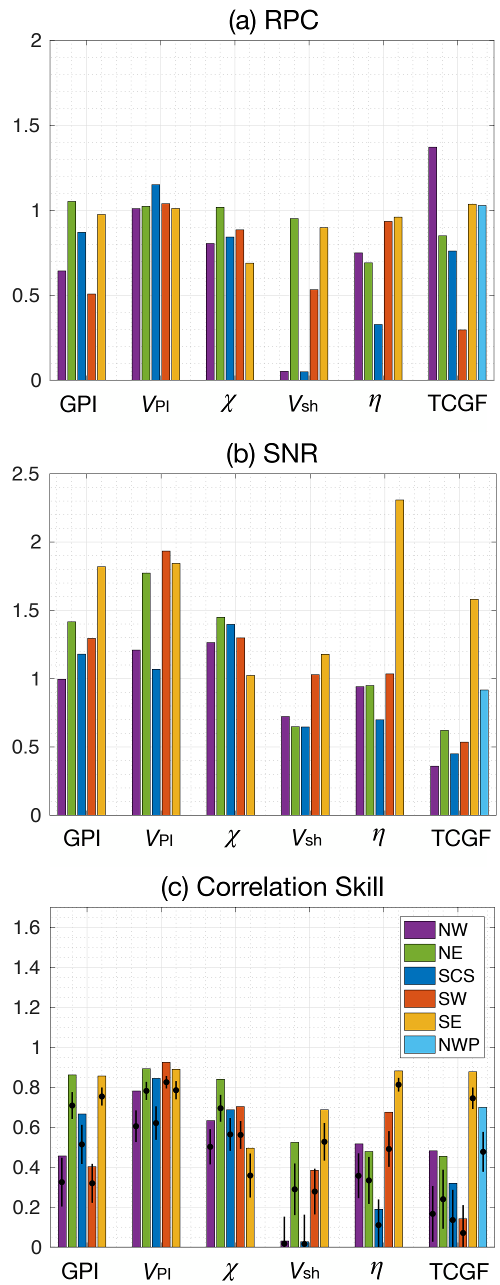


Figure 7: (a) Bar plots of the RPC of the GPI and its four components (i.e., potential intensity $V_{\mathrm{PI}}$, mid-level saturation deficit $\chi$, vertical wind shear $V_{\mathrm{sh}}$ and low-level vorticity $\eta$) as well as the TCGF over individual sub-basin regions (shown in different colors). For the TCGF, the result for the entire basin (cyan) is also displayed. (b) As in (a), but for the SNR. (c) As in (a), but for the correlation coefficient between the observations and the ensemble mean of the 100 simulations.

Black dots show the mean value of the correlation coefficients in individual member simulations, with vertical black lines denoting $\pm$one standard deviation.

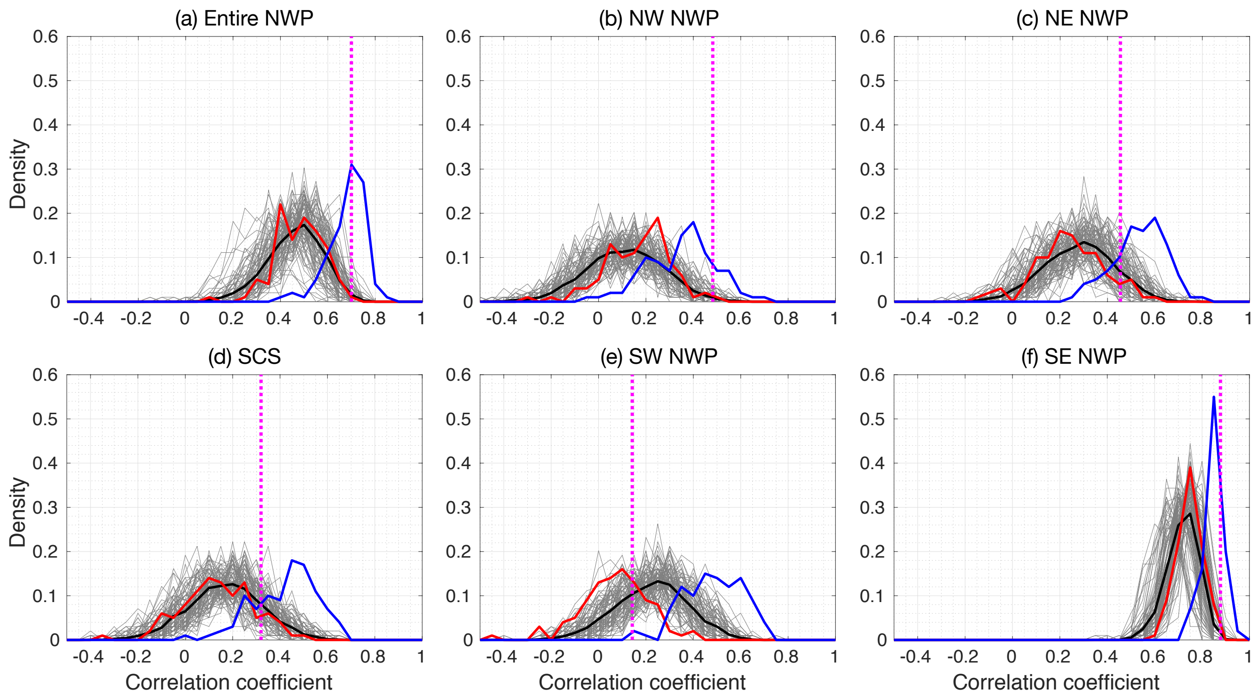


Figure 8: Histograms of the correlation coefficients between TCGF in one member simulation and that in the other 99 member simulations (gray curves), between TCGF in the observations and that in the 100 member simulations (red curve), and between TCGF in one member simulation and the ensemble mean of the 100 member simulations (blue curve) for (a) the entire NWP and (b–f) individual sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, black curve shows the average of the gray curves, and vertical dotted magenta line shows the correlation coefficient between TCGF in the observations and that in the ensemble mean of the 100 member simulations.

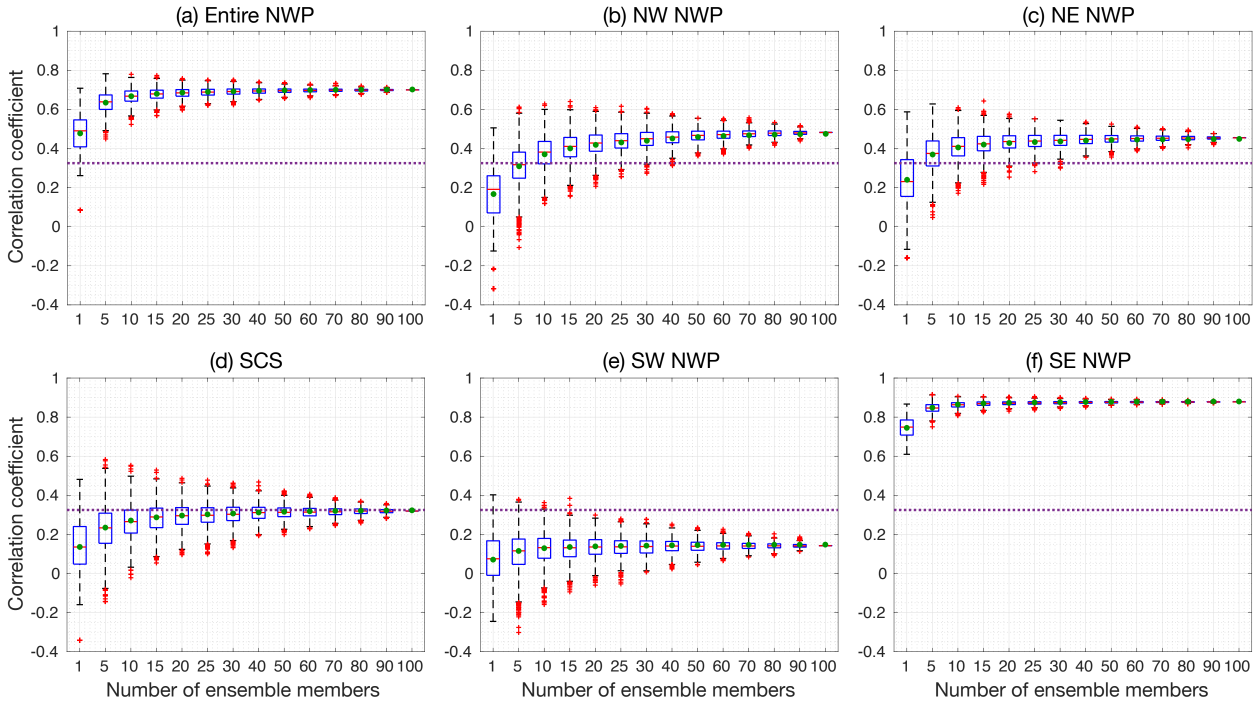


Figure 9: Box-and-whisker plots of the correlation coefficients between TCGF in the simulations and that in the observations as a function of ensemble size for (a) the entire NWP and (b–f) individual sub-basin regions: the NW quadrant, the NE quadrant, the SCS, the SW quadrant, and the SE quadrant. In each panel, green dots show the theoretical results based on Eq. (2), and the horizontal dashed purple line denotes the critical value of the correlation coefficient significant at the 0.05 level.

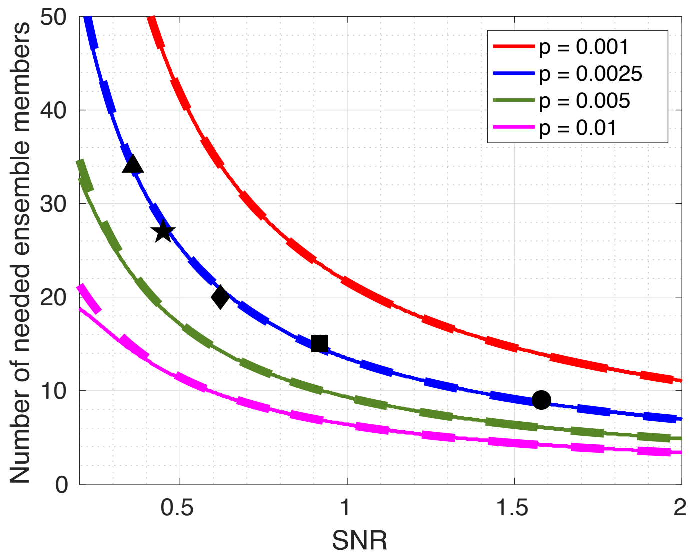


Figure 10: Number of ensemble members needed to reach the level-off of the model skill at reproducing observed/SST-forced TCGF variability versus the SNR of simulated TCGF for various predetermined tolerance levels ($p$): $p = 0.001$ (red), $p = 0.0025$ (blue), $p = 0.005$ (green), and $p = 0.01$ (magenta). Solid curves show numerical solutions based on Eq. (7) and dashed curves corresponding to Eq. (9). Black symbols show the empirical results based on the AGCM simulations with $p = 0.0025$: the circle for the SE quadrant, the square for the entire NWP, the diamond for the NE quadrant, the star for the SCS, and the triangle for the NW quadrant; the result for the SW quadrant is not shown because of the very low model skill in this sub-basin region.

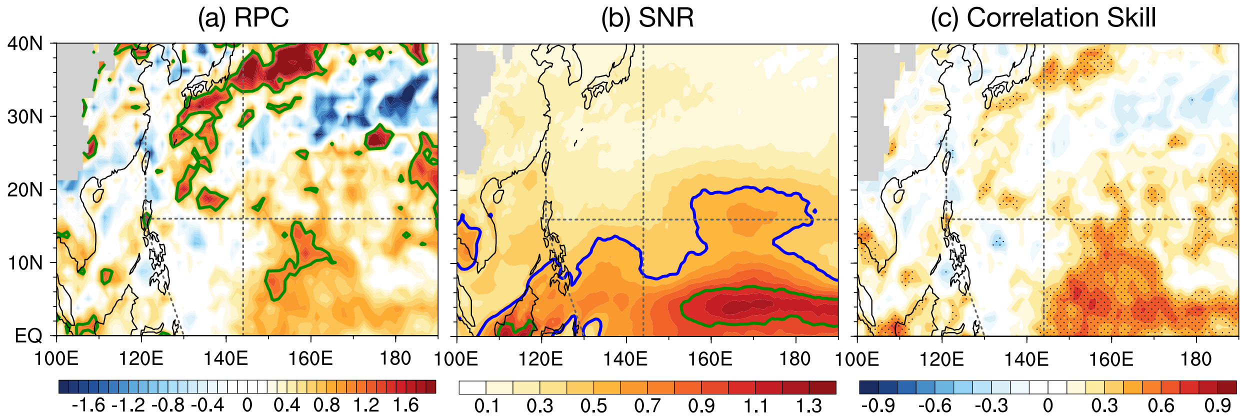


Figure 11: (a) Spatial distribution of the RPC of synoptic-scale disturbance activity during the NWP TC season (i.e., June–November). Green contours delineate the areas with the RPC equal to 1. (b) As in (a), but for the SNR. Blue and green contours delineate the areas with the SNR equal to 0.5 and 1, respectively. (c) As in (a), but for the correlation coefficient between the observations and the ensemble mean of the 100 member simulations. Black dots denote the areas where the correlation coefficients are significant at the 0.05 level.