

# *Western North Pacific tropical cyclones in the Met Office Global Seasonal Forecast System: performance and ENSO teleconnections*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Feng, X. ORCID: <https://orcid.org/0000-0003-4143-107X>,  
Klingaman, N. ORCID: <https://orcid.org/0000-0002-2927-9303>,  
Hodges, K. ORCID: <https://orcid.org/0000-0003-0894-229X>  
and Guo, Y.-P. (2020) Western North Pacific tropical cyclones  
in the Met Office Global Seasonal Forecast System:  
performance and ENSO teleconnections. *Journal of Climate*,  
33 (24). pp. 10489-10504. ISSN 1520-0442 doi:  
<https://doi.org/10.1175/JCLI-D-20-0255.1> Available at  
<https://centaur.reading.ac.uk/92964/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1175/JCLI-D-20-0255.1>

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

## **CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

# Western North Pacific Tropical Cyclones in the Met Office Global Seasonal Forecast System: Performance and ENSO Teleconnections

XIANGBO FENG,<sup>a</sup> NICHOLAS P. KLINGAMAN,<sup>a</sup> KEVIN I. HODGES,<sup>a</sup> AND YI-PENG GUO<sup>b</sup>

<sup>a</sup> National Centre for Atmospheric Science and Department of Meteorology, University of Reading, Reading, United Kingdom

<sup>b</sup> Key Laboratory of Mesoscale Severe Weather/Ministry of Education, and School of Atmospheric Sciences, Nanjing University, Nanjing, China

(Manuscript received 9 April 2020, in final form 15 September 2020)

**ABSTRACT:** The performance of the Met Office Global Seasonal Forecast System (GloSea5-GC2) for tropical cyclone (TC) frequency for the western North Pacific (WNP) in July–October is evaluated, using 23 years of ensemble forecasts (1993–2015). Compared to observations, GloSea5 overpredicts the climatological TC frequency in the eastern WNP and underpredicts it in the western and northern WNP. These biases are associated with an El Niño–type bias in TC-related environmental conditions (e.g., low-level convergence and steering flow), which encourages too many TCs to form throughout the tropical Pacific and slows TC propagation speed. For interannual TC frequency variability, GloSea5 overestimates the observed negative TC–ENSO teleconnection in the western and northern WNP, associated with an eastward shift in the ENSO teleconnection to environmental conditions. Consequently, GloSea5 fails to predict interannual TC variability in the northeast WNP (south of Japan); performance is higher in the southwest WNP (e.g., the South China Sea) where the sign of the TC–ENSO teleconnection is correct. This study suggests the need to reduce biases in environmental conditions and associated ENSO teleconnections in GloSea5 to improve the TC prediction performance in the WNP.

**KEYWORDS:** ENSO; Tropical cyclones; Seasonal forecasting; General circulation models; Model evaluation/performance; Interannual variability


## 1. Introduction


As one of most destructive weather phenomena, tropical cyclones (TCs) can have tremendous impacts over land, associated with strong winds, heavy rainfall, and storm surges. The western North Pacific (WNP) is the most active area for TC activity, accounting for more than one-third of all TCs and the globally accumulated cyclone energy (Gray 1968; Maue 2011). WNP TCs usually form in a favorable region for convection (e.g., over the tropical oceans, where it is warm and moist). Their movement is primarily modulated by the large-scale background steering flow, along the southern side of the North Pacific subtropical high (NPSH) in the tropics and on the western side of the NPSH in the subtropics. Accurate prediction of WNP TC activity up to several months ahead can provide an early warning service, which may significantly reduce TC damage in coastal WNP regions.

Seasonal predictions of TC activity are typically made with statistical models or dynamical atmosphere–ocean general circulation models (AOGCMs), or hybrid statistical–dynamical

approaches. The statistical and hybrid approaches use empirical relationships between TC activity and slowly evolving large-scale oceanic and atmospheric conditions, by assuming that current or predicted large-scale conditions persist over the forecast period, or that current or predicted conditions have a lagged effect on TCs. We focus our discussion on AOGCMs, as these are most relevant to our study. Until recently, AOGCMs could not simulate TC dynamical features very well, due in part to low model resolution (Henderson-Sellers et al. 1998; Strachan et al. 2013, and references therein). In recent years, increases in computing capability and the development of high-resolution AOGCMs have raised the potential to use AOGCMs to directly predict seasonal TC activity. Recent studies have evaluated seasonal prediction systems for WNP TC activity, including systems from the European Centre for Medium-Range Weather Forecasts (ECMWF; Vitart and Stockdale 2001; Manganello et al. 2016), the Geophysical Fluid Dynamics Laboratory (GFDL; Chen and Lin 2013; Vecchi et al. 2014; Zhang et al. 2019), and the Met Office (Camp et al. 2015). These systems perform reasonably well for seasonal WNP TC activity; however, they struggle to skillfully predict regional-scale TC activity (e.g., Zhang et al. 2019; Camp et al. 2019), especially in the WNP marginal seas.

For the interannual variability of regional WNP TC activity, sources of seasonal climate predictability include El Niño–Southern Oscillation (ENSO) (e.g., Chan 1985; Camargo and

 Denotes content that is immediately available upon publication as open access.

 Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-20-0255.s1>.

Corresponding author: Xiangbo Feng, [xiangbo.feng@reading.ac.uk](mailto:xiangbo.feng@reading.ac.uk)

DOI: 10.1175/JCLI-D-20-0255.1

© 2020 American Meteorological Society



This article is licensed under a Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

Sobel 2005; Kim et al. 2011; Zhan et al. 2011; Li and Zhou 2012; Bell et al. 2014; Patricola et al. 2018), the Pacific meridional mode (Zhang et al. 2016), the North Pacific gyre oscillation (e.g., Zhang et al. 2013), Pacific sea ice cover (e.g., Fan 2007), and the quasi-biennial oscillation (e.g., Ho et al. 2009). Some studies also suggest important transbasin teleconnections to WNP TC activity, including those from the Indian Ocean climate state (e.g., Zhan et al. 2011; Takaya et al. 2017), the tropical North Atlantic SST (e.g., Yu et al. 2016; Gao et al. 2018), and the Atlantic multi-decadal oscillation (e.g., Zhang et al. 2018). These interbasin and transbasin large-scale phenomena control environmental conditions (e.g., vertical wind shear, relative humidity, and steering flow) relevant to TC genesis, intensity, and track. It is important to understand the ability of AOGCMs to predict the teleconnections from large-scale phenomena to regional TC activity, and the effects of these teleconnections on prediction performance.

An older version of the Met Office Global Seasonal Forecast System, version 5, using the Global Atmosphere 3.0 configuration (GloSea5-GA3), showed some skill for basinwide WNP TC frequency (Camp et al. 2015). Significant biases remained over the basin, even for one-month lead time forecasts. A more recent study reported that the latest version of GloSea5, using the Global Coupled model 2.0 configuration (GloSea5-GC2), has high skill for seasonal predictions of the NPSH index and associated TC landfall in East Asia (Camp et al. 2019). Compared to GloSea5-GA3, GloSea5-GC2 has an updated atmospheric dynamical core, as well as revisions to its convective parameterization (Williams et al. 2015). However, GloSea5-GC2 fails to predict the interannual variability of total TCs influencing East Asia, performing worse than GloSea5-GA3 (Camp et al. 2019). Successful predictions of WNP TC activity require high performance for predicting the regional-scale atmospheric circulation (e.g., the NPSH), as well as for predicting environmental factors relevant for TC genesis and development (e.g., SSTs, tropical convection, and vertical wind shear). However, the representation of these environmental conditions in GloSea5 and their possible impacts on WNP TC prediction are not well understood and need to be addressed.

In this paper, we evaluate the ability of GloSea5-GC2 to predict the climatology and variability of regional WNP TC frequency, including the effects of forecast lead time and ensemble size. We link model performance for TC activity to performance for both the local environmental factors and regional-scale circulations relevant to WNP TCs. To understand performance for interannual WNP TC variability (in track and genesis), we also investigate the GloSea5 TC–ENSO teleconnection.

The paper is structured as follows. In section 2, GloSea5 and the reforecast dataset are described, along with the verification data and analysis methods. Section 3 first evaluates the performance of GloSea5 in simulating the climatological WNP TC frequency. The TC–ENSO teleconnection in GloSea5 is then analyzed and compared with the teleconnection in observations and reanalysis. The diagnosed biases in GloSea5 TC activity are related to errors in the local environment and large-scale circulation. Finally, in this section, we evaluate the performance of GloSea5 for the interannual variability of regional TC frequency in the WNP, including the effects of ensemble size and forecast lead time. Conclusions are provided in section 4.

## 2. Model and methodology

### a. Model

We evaluate the Met Office Global Seasonal Forecast System, version 5 (GloSea5), in scientific configuration Global Coupled 2.0 (GloSea5-GC2; MacLachlan et al. 2015; Williams et al. 2015). GloSea5 is based on the coupled Met Office Hadley Centre Global Environment Model, version 3 (HadGEM3), in which the Met Office Unified Model Global Atmosphere model (version 6.0) is coupled to the Nucleus for European Modeling of the Oceans (NEMO) dynamical ocean model (version 3.0), the Joint U.K. Land Environment Simulator (JULES), and the Los Alamos Sea Ice Model (CICE). GloSea5 uses an N216 atmospheric horizontal resolution ( $0.83^\circ$  in longitude and  $0.56^\circ$  in latitude) with 85 vertical levels, and the tripolar ORCA025 grid (approximately  $0.25^\circ$ ) with 75 vertical levels for the ocean. In GloSea5, the Stochastic Kinetic Energy Backscatter v2 (SKEB2; Bowler et al. 2009) scheme is used to generate the forecast ensemble by perturbing atmospheric physics tendencies; all ensemble members for a given start date are initialized from the same atmospheric, oceanic, and land surface conditions. The atmosphere and land surface are initialized from ERA-Interim (Dee et al. 2011), while the ocean and sea ice are initialized from the Forecast Ocean Assimilation Model (FOAM) Ocean Analysis (Blockley et al. 2014). GloSea5 is described in detail in MacLachlan et al. (2015) and Williams et al. (2015).

We analyze the full set of available GloSea5 seasonal reforecasts, which span a 23-yr period from 1993 to 2015. We analyze forecasts initialized in June, May, and April. These reforecasts are used operationally for model calibration and to evaluate model performance. There are four start dates in each month: the 1st, 9th, 17th, and 25th. There are seven ensemble members per start date; each member is a 7-month forecast. We combine all ensemble members initialized each month into a 28-member ensemble. We analyze forecasts for the 4-month season of July–October (JASO) in which TCs are most active in the WNP basin ( $0^\circ$ – $60^\circ$ N,  $100^\circ$ – $180^\circ$ E). The effect of lead time is analyzed by comparing the performance of forecasts initialized in June, May, and April, which have a 1-, 2-, and 3-month lead time for the JASO season, respectively. TC activity and environmental conditions (e.g., SSTs, wind, geopotential height, and precipitation) are extracted from GloSea5 without bias correction.

### b. Observation and reanalysis data

To verify GloSea5 TC forecasts, we use observed TC tracks from the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al. 2010) and TC tracks obtained from the ECMWF interim reanalysis (ERA-Interim), during 1993–2015. To ensure a consistent comparison, we apply the same TC tracking scheme [see section 2c(1)] to the atmospheric fields of GloSea5 forecasts and ERA-Interim. We also verified GloSea5 against TC tracks from the fifth major global reanalysis produced by ECMWF (ERA5; Hersbach and Dee 2016), with similar results to those for verification against ERA-Interim.

ERA-Interim was produced using a spectral atmospheric model at horizontal resolution of  $\sim 79$  km and 4DVar data assimilation. We also use ERA-Interim to verify various GloSea5 atmospheric fields, such as winds and geopotential height, as well as SSTs. Observed precipitation is obtained

from the Global Precipitation Climatology Project (GPCP; Adler et al. 2003). The ENSO conditions are defined by the Niño-3.4 index, by which SSTs are averaged over the Niño-3.4 region (5°N–5°S, 170°–120°W).

### c. Methods

#### 1) TC TRACKING

TCs are identified and tracked in GloSea5 and ERA-Interim from 6-hourly data, using the same method as Feng et al. (2019) and Hodges and Klingaman (2019) (see references therein for more details). The basic tracking uses the TRACK scheme (Hodges 1994, 1995, 1999). In the first instance, all positive vorticity centers that exceed  $0.5 \times 10^{-5} \text{ s}^{-1}$  in the range 0°–60°N are tracked in the spectrally filtered 850-hPa vorticity on the T42 horizontal grid for GloSea5 and in the average vorticity (between 850 and 600 hPa) on the T63 horizontal grid for ERA-Interim as in Hodges et al. (2017) where different reanalyses are compared. The T42 grid is used in GloSea5 due to limited availability of levels. Following the tracking the T63 vorticity maxima at additional levels up to 200 hPa (850, 500, 300, and 200 hPa) are added to the tracks using a search within a 5° radius (geodesic) of the tracked center to facilitate a warm core search. The TC tracks are identified by applying the same criteria as used in Hodges et al. (2017), namely

- 1) the T63 relative vorticity at 850 hPa must attain a threshold of at least  $6 \times 10^{-5} \text{ s}^{-1}$ ;
- 2) the difference in vorticity between 850 and 200 hPa (at T63 spectral truncation) must be greater than  $6 \times 10^{-5} \text{ s}^{-1}$ , to provide evidence of a warm core (via thermal wind balance);
- 3) the T63 vorticity center must exist at each level between 850 and 200 hPa for a coherent vertical structure;
- 4) criteria 1 to 3 must be jointly attained for a minimum of four consecutive time steps (1 day) and only apply over the oceans; and
- 5) tracks must start within a latitudinal band (0°–30°N) and last for at least 2 days.

Additionally, a further step is made to keep only the parts of tracks from the first to the last points that satisfy the TC criteria, to focus on the main TC life cycle stages and exclude the precursor and post-extratropical transition stages. Tropical cyclogenesis is defined as the first point of each identified TC track. The tracking is performed over the full length of the forecasts and reanalysis.

#### 2) ENSEMBLE FORECAST RESAMPLING

The 28 GloSea5 ensemble members from each month of start dates are independent, because for each of the four start dates per month the seven ensemble members are perturbed randomly and continuously during the forecast by stochastic physics (see section 2a). Because the order of ensemble members is arbitrary, a new 28-member ensemble forecast dataset for 1993–2015 can be generated by randomly reordering ensemble members without replacement from each year in 1993–2015. Repeated many times, this creates a large set of resampled forecasts that can be used to generate statistical distributions of teleconnections and evaluate performance (Johnson et al. 2017; Guo et al. 2020). We reorder the 28-member GloSea5 ensemble 2000 times to create 56 000 ( $28 \times 2000$ ) 23-yr time series; larger sizes do not substantially change the results. We refer to these forecast time series as “resampled forecast members.”

We also create sets of subsampled  $M$ -member ensembles, in which  $M$  is 1 to 27, for the 23-yr forecast time series. For each year of forecasts in 1993–2015,  $M$  members are randomly selected from the original 28 members without replacement. This produces  $28!/(28 - M)!M!$  possible unique selections for each year. The  $M$ -member selection is made independently for each of the 23 years. These  $M$  members are then concatenated across the 23 years to create ensembles of synthetic 23-yr time series. This can produce  $[28!/(28 - M)!M!]^{23}$  possible unique ensembles. To reduce computational cost, we randomly choose 2000  $M$ -member ensembles; a sensitivity test shows that more ensembles do not substantially change the results. We refer to these  $M$ -member ensembles as “sub-sampled ensembles,” which are used in section 3c to evaluate the effect of ensemble size on the potential predictability and prediction performance for TC activity.

The resampling process assumes that the four start dates share the same features of uncertainty in their forecasts and that the 28 ensemble members of forecasts are randomly distributed around the mean. These resampled datasets, and the relationships derived from them, are valid only for the 28-member GloSea5 ensemble. The relationships may be sensitive to the size of the GloSea5 ensemble, because a larger or smaller ensemble may have a different ensemble spread. The results from this statistical resampling may not agree with, and are not a replacement for, results from model sensitivity experiments that test variations in ensemble size.

#### 3) SIGNAL-TO-NOISE RATIO

The interannual variability of the ensemble mean is the predictable signal in the model. In individual ensemble members, this signal is embedded within the uncertainty of the modeling system, due to, for example, imperfectly resolved physics and the chaotic nature of the climate system. We use the signal-to-noise ratio (SNR) to measure the potential predictability of external variability (predictable signal) relative to internal variability (unpredictable signal), as simulated by GloSea5. The calculation of SNR is adopted from Mei et al. (2015):

$$\text{SNR} = \sigma_f / \sigma_i, \quad (1)$$

where  $\sigma_f$  is the interannual standard deviation of the ensemble mean (i.e., the forced external variability) and  $\sigma_i$  is the standard deviation of the anomaly departures from the ensemble mean in all individual members (i.e., the random internal variability).

A high SNR indicates high potential predictability of an external climate signal in GloSea5, while a low SNR value means that GloSea5 is unable to consistently predict an external signal (i.e., due to a divergence among model ensemble members). As the SNR is only a model estimate, it may not represent the behavior of the true physical system. Further, a high SNR does not necessarily mean high prediction performance.

## 3. Results

### a. Climatology of TC track

#### 1) BIAS OF TC TRACK DENSITY

TC track density is calculated at each grid point as the number of TCs passing through an area defined by a  $10^\circ \times 10^\circ$

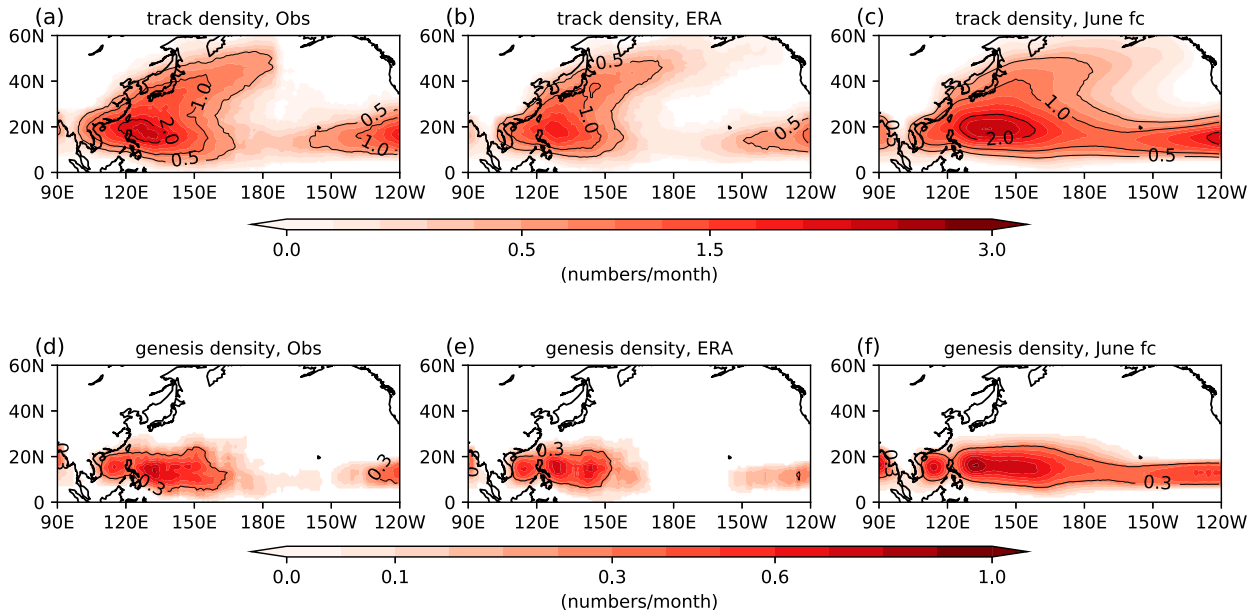


FIG. 1. (a)–(c) Climatology of TC track density in the IBTrACS, ERA-Interim, and GloSea5 June forecast ensemble mean, with contours showing the values of 0.5, 1.0, and 2.0 TCs month<sup>-1</sup>. (d)–(f) Climatology of TC genesis density in IBTrACS, ERA-Interim, and GloSea5 June forecast ensemble mean, with contours showing the values of 0.3 TCs month<sup>-1</sup>. TC density is calculated at each grid point as the number of TCs or TC geneses over an area defined by a 10° × 10° box around the grid point, per month over JASO. The analysis period is 1993–2015.

box around the grid point per month in JASO. The 28-member ensemble mean of GloSea5 TC track density averaged over 1993–2015 is verified against the track density from IBTrACS and ERA-Interim for the same period. There are discrepancies in TC track density between these two verification datasets (Figs. 1a,b). In the WNP, overall, there are fewer TCs (0.2–0.4 TCs per month; hereafter, TCs month<sup>-1</sup>) identified in ERA-Interim than in IBTrACS. This could be related to the model resolution in ERA-Interim, which is too coarse to produce enough storms that satisfy the TC criteria in the TRACK scheme [section 2c(1)]. In the central North Pacific (40°–60°N), there are more TCs in ERA-Interim, due to the extended life cycle identified by TRACK. Hodges and Emerton (2015) and Hodges et al. (2017) discuss in more detail the agreements and discrepancies between the observed and reanalysis TCs. Because the same TC identification scheme is applied to ERA-Interim and GloSea5, and because ERA-Interim and GloSea5 have similar horizontal resolutions, validating GloSea5 against ERA-Interim is fairer than validating GloSea5 against IBTrACS from the perspective of model development. We are aware that IBTrACS is widely used in verification of operational TC forecasts. Therefore, in our analysis, we use both datasets to evaluate the prediction performance of GloSea5 for WNP TCs.

In IBTrACS and ERA-Interim, WNP TC track density varies between 0 to 2.0 TCs month<sup>-1</sup>, with the highest density east of the Philippines and a local minimum in the central Pacific (Figs. 1a,b). In the ensemble mean of GloSea5 June forecasts, the climatological TC track density is more zonally oriented than in the verification datasets, without a local minimum in the central Pacific (Fig. 1c). Verifying against

IBTrACS yields a positive bias of 0.5–1.0 TCs month<sup>-1</sup> in the central-eastern WNP (10°–30°N, 130°E–180°), whereas in other regions (e.g., Japan, east of China, and southeast of the Philippines) GloSea5 has a negative bias of –0.5 to –0.2 TCs month<sup>-1</sup> (Fig. 2a). When verifying against ERA-Interim, the positive bias of track density increases to up to 1.5 TCs month<sup>-1</sup> in the central-eastern WNP, while the negative bias reduces (Fig. 2d). The ensemble spread (standard deviation across the 28-member ensemble) of TC track density between GloSea5 members is less than 0.15 TCs month<sup>-1</sup> over most of the basin (not shown). Thus, the positive biases in TC track density in the June forecasts against the different verification datasets are consistent across the ensemble members.

The GloSea5 WNP TC density biases are associated with the misrepresentation of cyclogenesis and TC paths. In IBTrACS and ERA-Interim, WNP TCs mostly originate in the western tropical Pacific (0°–20°N, 100°–160°E), while in GloSea5 TCs form across most of the tropical Pacific (Figs. 1d–f). In the western and central tropical Pacific (100°E–150°W), GloSea5 June forecasts generate more TCs than IBTrACS, with the largest bias in the central (>0.3 TCs month<sup>-1</sup>) (Fig. 2g). Cyclogenesis biases against ERA-Interim are slightly larger than those against IBTrACS (Fig. 2j), as there is less cyclogenesis in the central tropical Pacific in ERA-Interim (Fig. 1e). Figures 3a–c show TC translation velocity. In IBTrACS and ERA-Interim, most low-latitude TCs travel northwest at 2–4 m s<sup>-1</sup> along the southern edge of the NPSH, and recurve north at 120°–150°E around the western edge of the NPSH. However, in GloSea5 TCs travel too slowly in the tropics; the recurvature region is shifted east (120°E–180°) relative to observations (120°–150°E). The GloSea5 TC translation velocity has a cyclonic bias at low to middle latitudes of 1–5 m s<sup>-1</sup> (Figs. 3d–i), related to too many TCs

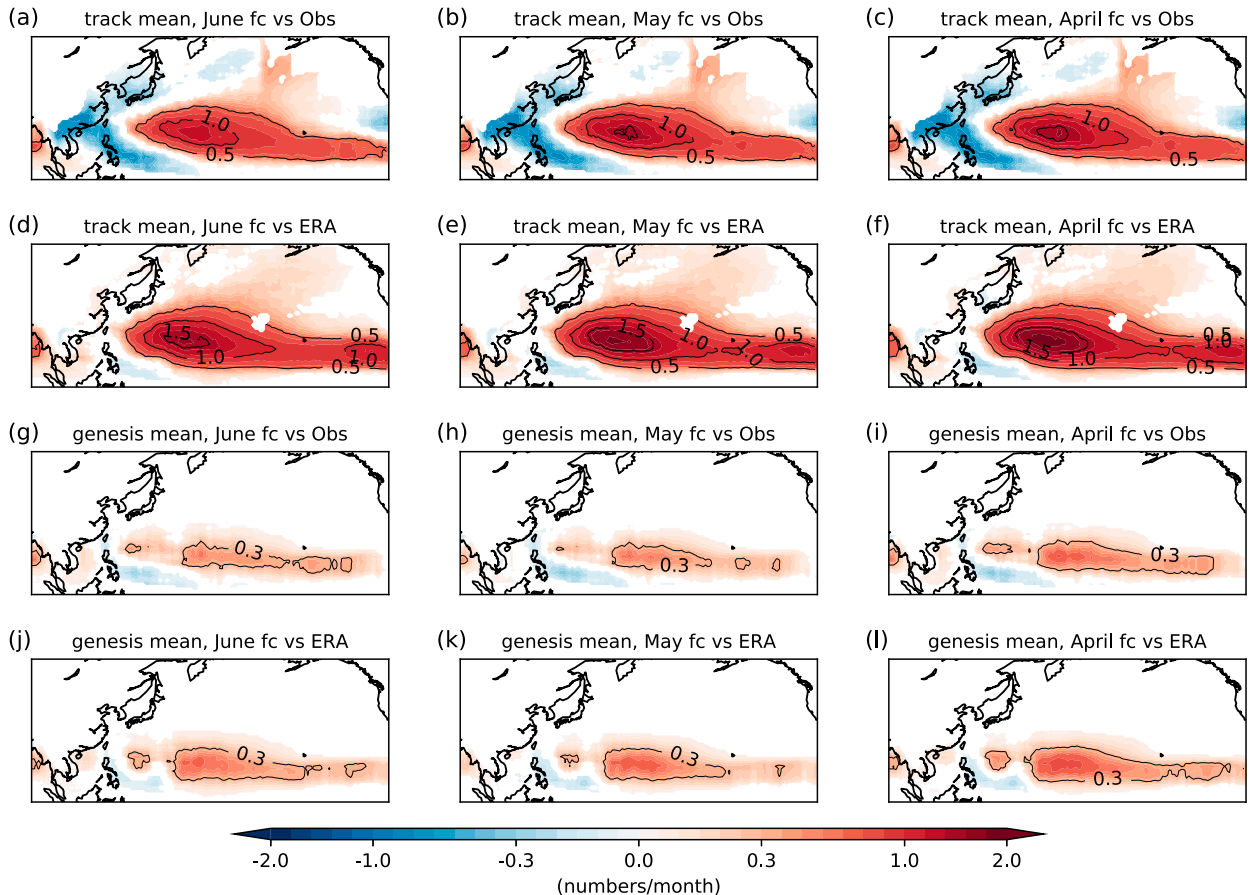


FIG. 2. (a)–(c) Bias of TC track density in GloSea5 June, May, and April forecast ensemble means, against IBTrACS, with contours showing the bias values of 0.5, 1.0, and 1.5 TCs month<sup>-1</sup>. (d)–(f) As in (a)–(c), but for bias against ERA-Interim TCs. (g)–(i) Bias of TC genesis density in GloSea5 June, May, and April forecast ensemble means, against IBTrACS, with contours showing the bias values of 0.3 TCs month<sup>-1</sup>. (j)–(l) As in (g)–(i), but for bias against ERA-Interim TCs. Only biases larger than the ensemble spread are plotted. TC density is calculated at each grid point as the number of TCs or TC geneses over an area defined by a 10° × 10° box around the grid point, per month over JASO. The analysis period is 1993–2015.

traveling into the open Pacific and too few reaching the south and west. The cyclonic bias is larger when verifying against IBTrACS than against ERA-Interim.

We analyze the effect of forecast lead time on mean TC frequency prediction by analyzing GloSea5 forecasts initialized in May and April. With lead time, TC track density biases remain consistent in geographic pattern, but grow in amplitude in the central-eastern WNP (Figs. 2a–f). Cyclogenesis biases also grow with lead time (Figs. 2g–l). The track density bias growth in the central-eastern WNP is also related to a slower translation speed east of the Philippines (Figs. 3d–i).

## 2) BIASES OF TC ENVIRONMENTAL CONDITIONS

Environmental conditions for TC genesis and movement in GloSea5 are analyzed to understand the TC frequency biases. At low levels (1000 hPa), the climatological wind in GloSea5 June forecasts has a large-scale cyclonic bias in the North Pacific with respect to ERA-Interim (Fig. 4a). The anomalous convergence associated with these biases is collocated with a positive precipitation bias in the western and central tropical Pacific (0°–20°N, 100°E–

150°W), suggesting excessive convection and moisture convergence. Midlevel (e.g., 500 hPa) relative humidity biases share the same spatial pattern as precipitation (not shown), with the largest positive bias (>10%) in the central tropical Pacific. This wetter condition favors TC genesis in the central tropical Pacific, likely contributing to the overestimated TC genesis (Figs. 2g,j). The precipitation and surface wind biases are related to the misrepresentation of the intertropical convergence zone (ITCZ), which is intensified and shifted east in the tropical Pacific (130°E–160°W) relative to GPCP (see the contour in Fig. 4a for observed climatology of precipitation).

The vertical wind shear (defined as the difference of wind flow at 200 and 850 hPa), another key factor for TC formation and development, is also biased in GloSea5 (Fig. 4b). Compared to ERA-Interim, GloSea5 has stronger wind shear in the western tropical Pacific (0°–20°N, 100°–160°E), with a bias of up to 12 m s<sup>-1</sup> in the South China Sea (SCS) and east of the Maritime Continent (MC). In contrast, in the eastern tropical Pacific (10°–20°N, 160°E–150°W), the GloSea5 vertical wind shear is underestimated by 2–10 m s<sup>-1</sup>, which may be responsible for the overestimated TC activity. These errors cause the region of

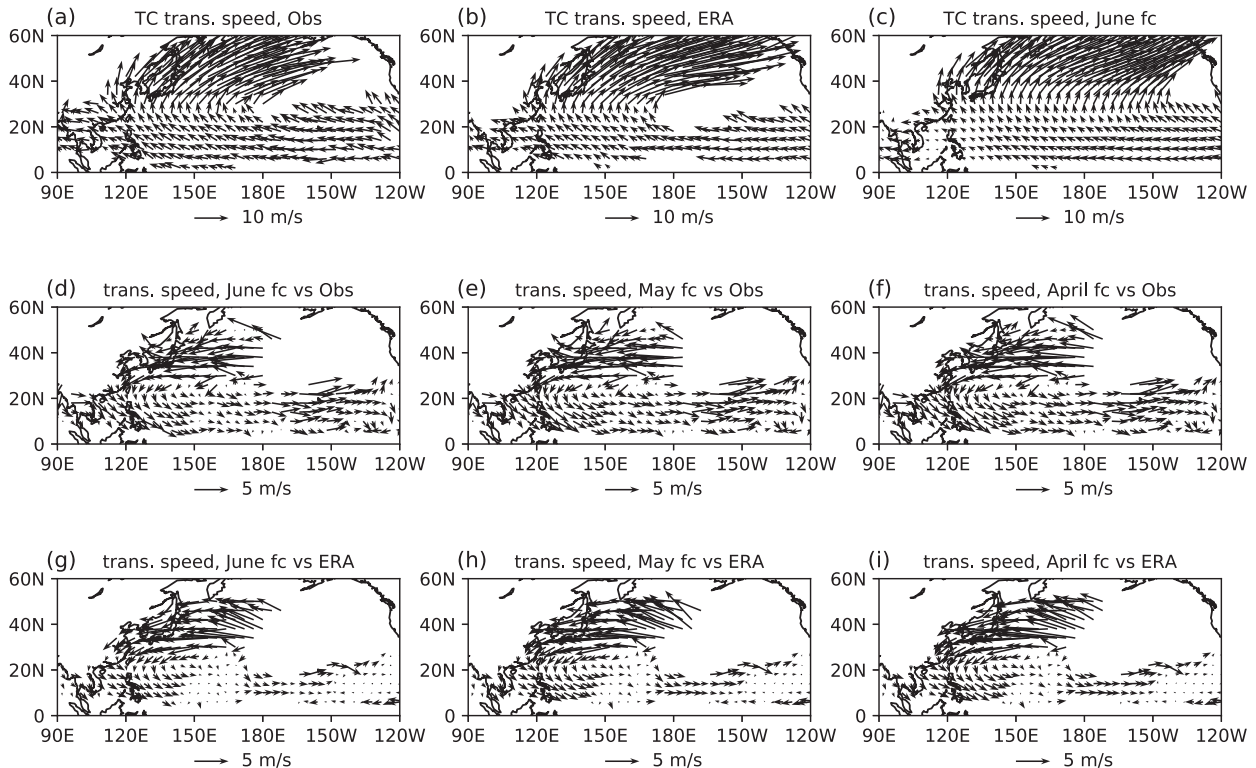


FIG. 3. (a)–(c) Climatology of TC translation velocity in IBTrACS, ERA-Interim, and GloSea5 June forecast ensemble mean. (d)–(f) Bias of TC translation velocity in GloSea5 June, May, and April forecast ensemble means, against IBTrACS. (g)–(i) As in (d)–(f), but for bias against ERA-Interim TCs. TC translation velocity is calculated at each grid point as TC traveling speed averaged by all TCs that are passing over an area defined by a  $10^{\circ} \times 10^{\circ}$  box around the grid point over JASO. The analysis period is 1993–2015.

climatologically weak shear (contour in Fig. 4b) to shift east in GloSea5. The errors are associated with an intensified, and eastward shift of, the tropical overturning circulation (i.e., a lower-tropospheric westerly bias and an upper-tropospheric north-easterly cross-equatorial bias).

The GloSea5 SSTs in the June forecasts show a warm bias in the eastern Pacific ( $0^{\circ}$ – $30^{\circ}$ N,  $180^{\circ}$ – $120^{\circ}$ W) and a cold bias in the Indo-Pacific warm pool ( $0^{\circ}$ – $20^{\circ}$ N,  $90^{\circ}$ – $150^{\circ}$ E) (Fig. 4c); the cold bias reaches over  $1^{\circ}$ C in the SCS and east of the Philippines. East of the Philippines ( $5^{\circ}$ – $10^{\circ}$ N,  $120^{\circ}$ – $150^{\circ}$ E), the cold SST and positive wind shear biases may be responsible for the slightly underestimated TC genesis (Figs. 2g,j). In the overall western tropical Pacific ( $0^{\circ}$ – $20^{\circ}$ N,  $100^{\circ}$ – $160^{\circ}$ E), although the SST and wind shear biases are both unfavorable for TC formation (Figs. 4b,c), GloSea5 still produces excessive TC genesis (Figs. 2g,j), perhaps because the SSTs remain warm enough ( $>28^{\circ}$ C) for TC genesis. The east Pacific warm SST bias extends the area of the Indo-Pacific warm pool in GloSea5, which may weaken wind shear and increase convection in the east. The cyclogenesis overestimation in the western tropical Pacific may also be related to remote biases of excessive deep convection and warm SSTs in the central tropical Pacific ( $0^{\circ}$ – $20^{\circ}$ N,  $150^{\circ}$ E– $150^{\circ}$ W), which favor development of TC precursor systems such as westward propagating tropical depressions and equatorial waves.

The biases in the TC track density and translation velocity in GloSea5 June forecasts (Figs. 2a,d and 3d,g) are also associated

with the misrepresentation of the large-scale steering flow. Compared to ERA-Interim, the GloSea5 WNP steering flow (500 hPa) has a strong cyclonic bias (Fig. 4d), which matches the TC propagation velocity bias (Figs. 3d,g) but has a larger magnitude. The wind bias is accompanied by a weakened NPSH by 20–50 m. Associated with the NPSH and steering flow biases, GloSea5 TCs tend to move westward too slowly in the tropics and recurve north farther east. This is associated with too many TCs in the central Pacific and too few in the western WNP (Figs. 2a,d). The weak NPSH bias is consistently found in a hierarchy of Met Office Unified Model (MetUM) coupled and atmosphere-only simulations (e.g., Feng et al. 2019), suggesting a systematic error. The above diagnosed biases in TC environmental factors are much larger than the ensemble spreads of those factors, indicating the biases are robust.

Changes in TC environmental conditions at longer lead times are shown in the online supplemental material (Figs. S1 and S2). Biases in environmental factors grow from the June to the May forecasts: humidity in the central tropical Pacific increases, wind shear and SST east of the MC weakens and cools, respectively, and the NPSH weakens. The growth slows from May to April forecasts, indicating possible saturation of the biases. In short, GloSea5 exhibits a significant El Niño-type bias in WNP environmental conditions that grows quickly and may saturate after 2 months' lead time. However, it is difficult to diagnose which conditions are most responsible for the TC



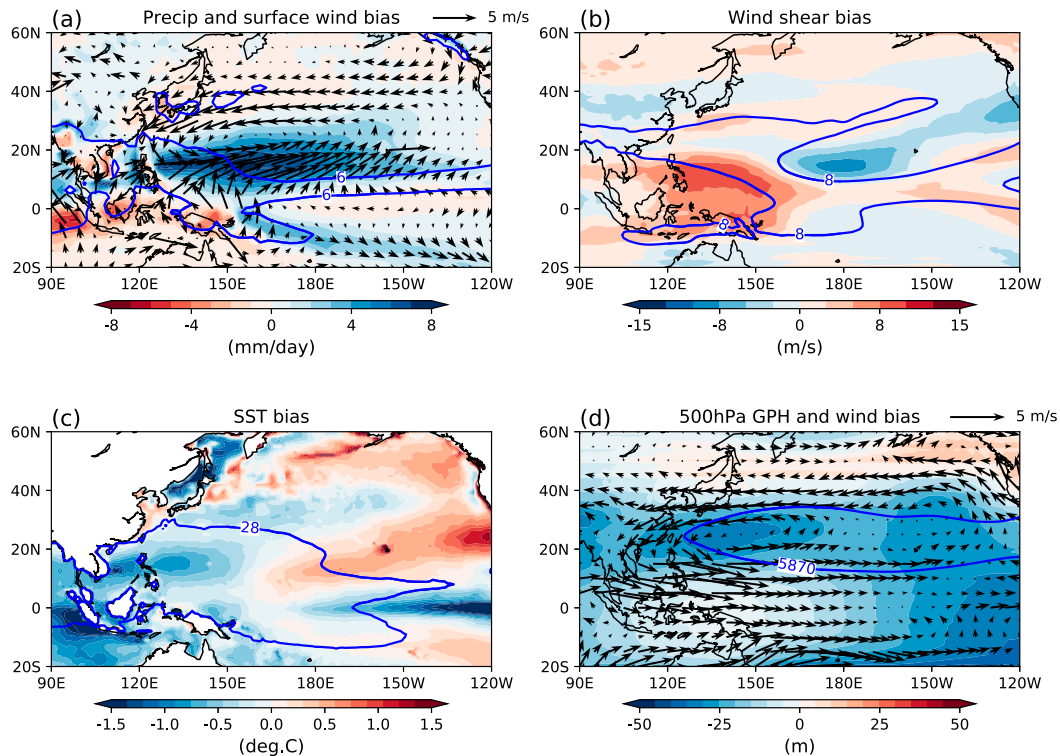


FIG. 4. (a) GloSea5 1000-hPa wind (vectors) and precipitation (shaded) biases, (b) vertical wind shear bias, (c) SST bias, and (d) 500-hPa wind (vectors) and geopotential height (GPH; shaded) biases, in June forecast ensemble mean. Blue lines are the climatology of environmental conditions in observations or ERA-Interim, i.e.,  $6 \text{ mm day}^{-1}$  of GPCP precipitation,  $8 \text{ m s}^{-1}$  of vertical wind shear,  $28^\circ\text{C}$  of SST, and  $5870 \text{ m}$  of GPH from ERA-Interim, respectively. Only biases larger than the ensemble spread are plotted. Precipitation is verified against GPCP, while other fields are verified against ERA-Interim. The analysis period is 1993–2015.

biases, because in a coupled system these environmental conditions biases are closely associated. Further detailed sensitivity experiments are required to address this challenge.

### b. ENSO teleconnection

#### 1) TELECONNECTION TO TC TRACK DENSITY

ENSO conditions are defined by the Niño-3.4 index. The GloSea5 ensemble mean predicts well the observed Niño-3.4 index in JASO (Fig. 5), both for the variability and magnitude, with  $r = 0.94, 0.90,$  and  $0.85$  for the June, May, and April forecast ensemble mean, respectively. The SNR of Niño-3.4 in the ensemble forecasts is  $>3$ , indicating a small internal variability and a large external variability.

Figures 6a, 6b, 7a, and 7b show the regional TC frequency–ENSO relationships in the two verification datasets. Both IBTrACS and ERA-Interim show strong and positive TC frequency–ENSO relationships (correlation coefficient  $r > 0.6$ ; regression coefficient  $\beta > 0.4 \text{ TCs month}^{-1}\text{ }^\circ\text{C}^{-1}$ ) in the northeast WNP (e.g., southeast of Japan) and the central and eastern tropical Pacific ( $0^\circ$ – $20^\circ\text{N}$ ,  $150^\circ\text{E}$ – $150^\circ\text{W}$ ); there are weak and negative correlations ( $-0.35 < r < 0$ ;  $-0.2 \text{ TCs month}^{-1}\text{ }^\circ\text{C}^{-1} < \beta < 0$ ) in the southwest WNP (e.g., the SCS, the southern China, and the Philippines).

For June forecasts, GloSea5 ensemble mean correctly reproduces the strong positive correlations in the central and

eastern tropical Pacific (Fig. 6c). The GloSea5 correlations are strongly negative in both the western and northern WNP ( $r < -0.6$ ), whereas in IBTrACS and ERA-Interim the correlations are strongly positive in the northeast (e.g., south of Japan) and weakly negative in the southwest (e.g., the SCS). The regression coefficient  $\beta$  between ENSO and regional TC frequency is over 50% smaller in GloSea5 ensemble mean than in either IBTrACS or ERA-Interim in most regions, except for the southwest WNP (e.g., the SCS) (Fig. 7c). We also analyzed  $\beta$  by combining the 28 individual members, but the results remain similar. This suggests that in GloSea5, the sensitivity of TC frequency to ENSO is too weak in most of the WNP, even though the TC–ENSO correlations are enhanced. At longer lead times, the negative correlation in the north and west of the basin intensifies and expands; the size of negative  $\beta$  values in the west increases (Figs. 6c–e and 7c–e).

Ensemble averaging reduces model uncertainty, increasing the contribution of external variability, for example, related to ENSO. However, observations (and ERA-Interim) have only one single “ensemble member”. Thus, it is unfair to compare the observed or ERA-Interim TC–ENSO teleconnection with that in the GloSea5 ensemble mean. We analyzed the distribution of the regional TC–ENSO teleconnection in the 56 000 “resampled forecast member” time series for 1993–2015 [section 2c(2)].

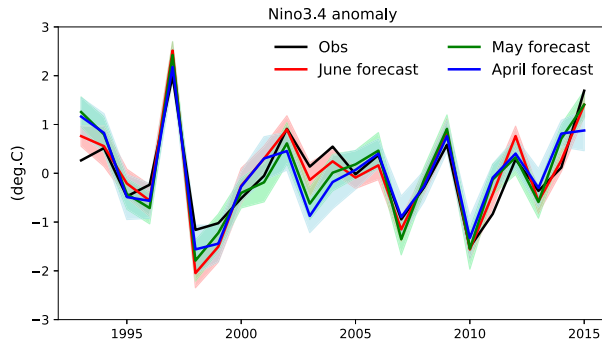


FIG. 5. Niño-3.4 index from ERA-Interim SST (black), and GloSea5 June (red), May (green), and April (blue) forecasts for JASO. Solid line represents the ensemble mean, and shaded area is for the ensemble spread.

The distributions of the ENSO–TC teleconnection ( $r$  and  $\beta$ ) in the resampled forecast members with different lead times are shown in Fig. 8. The spatial pattern of the median  $r$  (Figs. 8a–c) is similar to the ensemble-mean  $r$  (Figs. 6c–e), but the median  $r$  is much smaller ( $-0.4 < r < 0.4$ ) because ensemble averaging strengthens the correlation. The dispersion of  $r$  is nearly homogeneous over the North Pacific, indicating the consistent variability in the simulated TC–ENSO teleconnection. The median of  $\beta$  has the same spatial pattern and magnitude as the ensemble-mean teleconnection, while the dispersion of  $\beta$  is larger in the WNP and smaller in the eastern Pacific (Figs. 8d–f). By considering the dispersions of  $r$  and  $\beta$  in the resampled forecast members, in the northern and eastern WNP, the observed TC–ENSO teleconnection is outside the range of median values plus or minus spread, confirming the

robustness of the negative bias in the teleconnection in GloSea5.

## 2) TELECONNECTION TO TC ENVIRONMENTAL CONDITIONS

The observed regional TC–ENSO relationships can be interpreted using the ENSO teleconnections to the large-scale atmospheric conditions in observations and ERA-Interim (correlation coefficients for interannual variability in Figs. 9a–d, regression coefficients in Fig. S3). In El Niño years, a substantial cyclonic near-surface circulation develops in the North Pacific, with strong westerly anomalies in the central tropics ( $150^{\circ}\text{E}$ – $150^{\circ}\text{W}$ ) (Fig. 9a). These westerly anomalies bring moist and warm air from the warm pool to the central tropics, where the upper ocean is also warm in El Niño years (Fig. 9c). Moister air and an increased vertical temperature gradient enhance convection in this region. Vertical wind shear is significantly reduced in the eastern tropical WNP ( $0^{\circ}$ – $20^{\circ}\text{N}$ ,  $120^{\circ}\text{E}$ – $180^{\circ}$ ) during El Niño (Fig. 9b), due to the eastward shifted zonal circulation. These changes favor TC formation in the central tropical Pacific ( $150^{\circ}\text{E}$ – $150^{\circ}\text{W}$ ) and suppress it in the west ( $100^{\circ}$ – $150^{\circ}\text{E}$ ) (Figs. 6a,b).

In ERA-Interim, the El Niño–associated westerly steering flow anomaly in the western tropical Pacific ( $0^{\circ}$ – $10^{\circ}\text{N}$ ,  $100^{\circ}\text{E}$ – $180^{\circ}$ ) slows TCs traveling west (Fig. 9d). Consequently, the anomalous TC motion (translation velocity) is eastward in the tropics (Fig. S4a). In the midlatitudes, because the westerly steering flow anomaly is collocated with the reduced NPSH in El Niño years, TCs begin to recurve farther east. The opposite-signed patterns are found in La Niña years. These anomalies lead to positive TC density–ENSO correlations in the northern and eastern WNP (Figs. 6a,b). (Correlation coefficients for

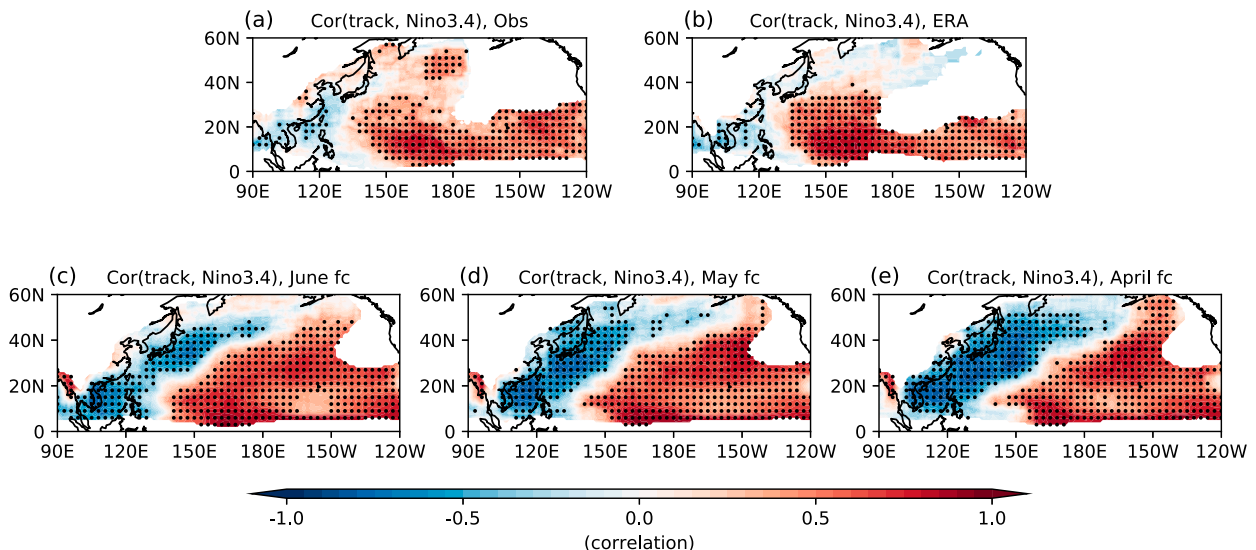


FIG. 6. (a),(b) Correlation coefficient between TC track density and ENSO in IBTrACS and ERA-Interim. (c)–(e) Correlation coefficient between TC track density and ENSO in GloSea5 June, May, and April forecast ensemble means. Dotted areas are where correlations pass significance test at the 90% confidence level. TC track density is calculated at each grid point as the number of TCs over an area defined by a  $10^{\circ} \times 10^{\circ}$  box around the grid point, per month over JASO. The analysis period is 1993–2015.

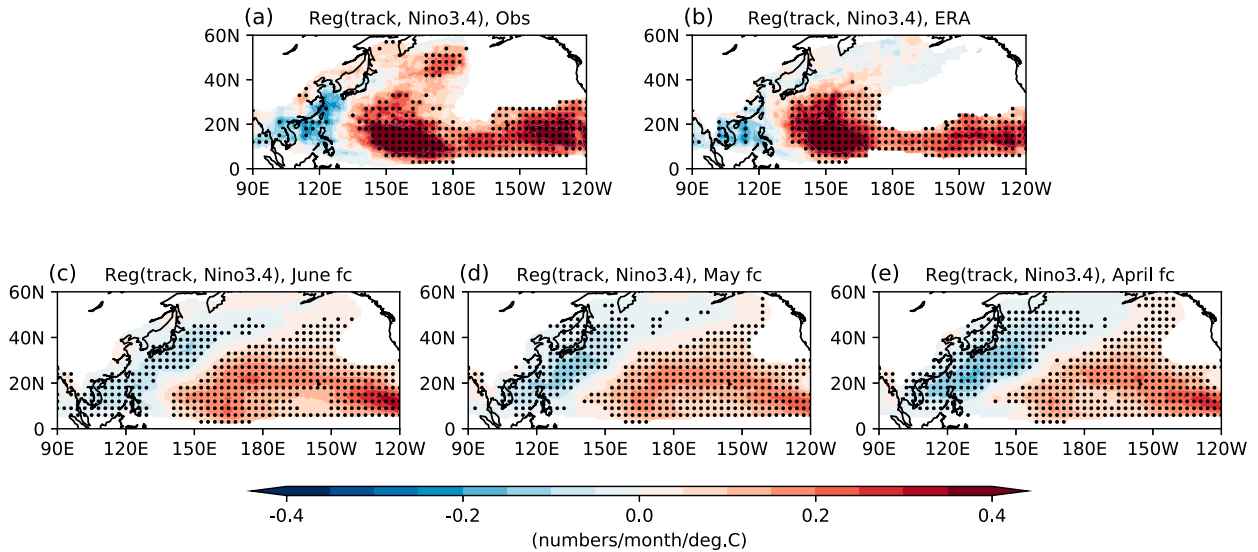


FIG. 7. (a),(b) Regression coefficient between TC track density and ENSO in IBTrACS and ERA-Interim. (c)–(e) Regression coefficient between TC track density and ENSO in GloSea5 June, May, and April forecast ensemble means. Dotted areas are where correlations pass significance test at the 90% confidence level. TC track density is calculated at each grid point as the number of TCs over an area defined by a  $10^{\circ} \times 10^{\circ}$  box around the grid point, per month over JASO. The analysis period is 1993–2015.

interannual variability are given in Figs. 9a–d, and regression coefficients in Fig. S3.)

The GloSea5 errors in the TC–ENSO teleconnection in the northern and western WNP are associated with errors in the simulated environment–ENSO teleconnections. Figures 9e–h show the correlations between ENSO and the environmental

conditions in the GloSea5 ensemble mean, for the June forecasts, with regression coefficients shown in Figs. S3e–h. In El Niño years, GloSea5 produces significantly less rainfall in the subtropical WNP ( $20^{\circ}$ – $50^{\circ}$ N,  $100^{\circ}$ E– $180^{\circ}$ ), whereas no such signal is seen in GPCP (Figs. 9a,e). Equally, increased 500-hPa relative humidity in the subtropical WNP is strongly associated

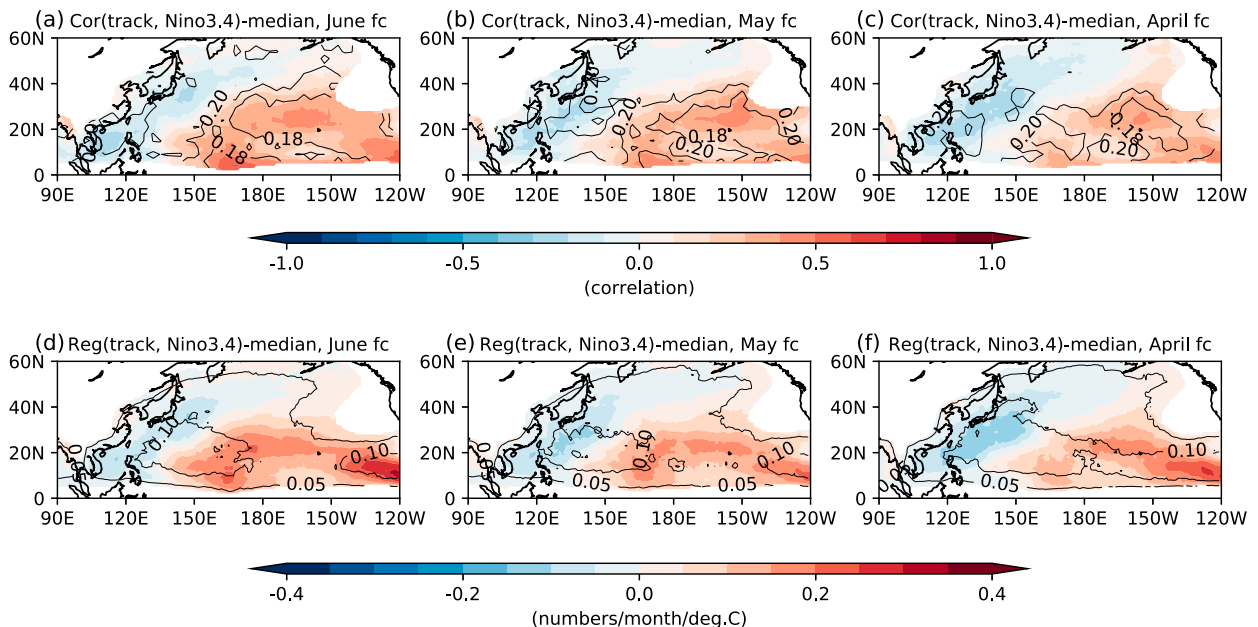


FIG. 8. (a)–(c) Median of correlation coefficients between TC track density and ENSO in GloSea5 June, May, and April resampled forecast members, with contours showing the standard deviation of correlation samples. (d)–(f) As in (a)–(c), but for regression coefficients between TC track density and ENSO. TC track density is calculated at each grid point as the number of TCs over an area defined by a  $10^{\circ} \times 10^{\circ}$  box around the grid point, per month over JASO. The analysis period is 1993–2015.

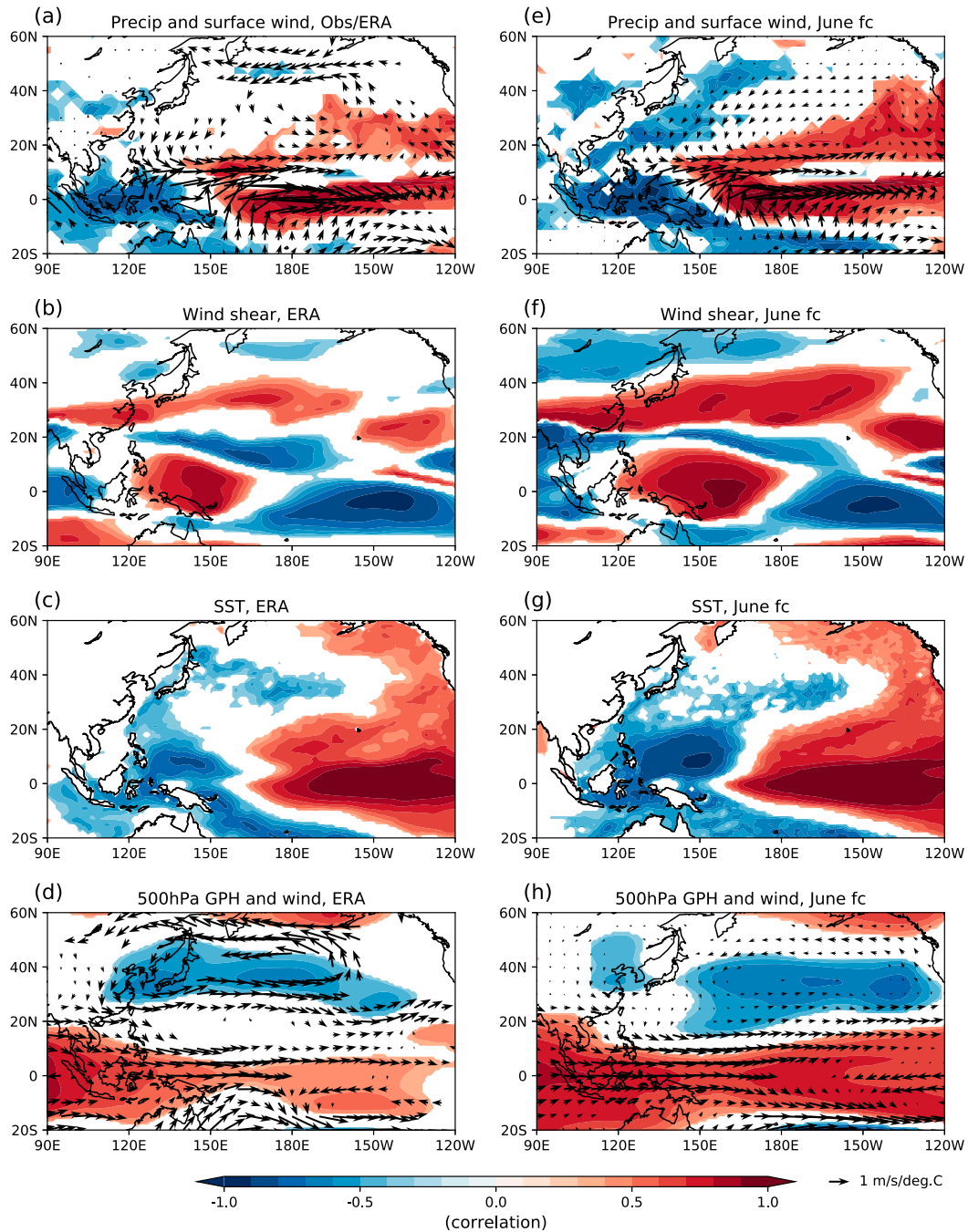


FIG. 9. (a) GPCP precipitation–ENSO correlation (shaded) and 1000-hPa wind–ENSO regression (vectors) in ERA-Interim. (b) Vertical wind shear–ENSO correlation in ERA-Interim. (c) SST–ENSO correlation in ERA-Interim. (d) 500-hPa GPH–ENSO correlation (shaded) and 500-hPa wind–ENSO regression (vectors) in ERA-Interim. (e)–(h) As in (a)–(d), but for GloSea5 June forecast ensemble mean. Only correlations or regressions that pass significance test at the 90% confidence level are plotted. The analysis period is 1993–2015.

with El Niño events in GloSea5, which is again less clear in ERA-Interim (not shown). The discrepancy could be related to the ITCZ, which is shifted farther east in El Niño years in GloSea5. The El Niño–related increase in vertical wind shear in the western tropics (east of MC) and subtropics ( $20^{\circ}$ – $40^{\circ}$ N)

of the Pacific is more robust and extensive in GloSea5 than in ERA-Interim (Figs. 9b,f). In addition, the observed western Pacific cold SST anomaly associated with El Niño is extended farther west to the SCS and farther east to the central Pacific (Figs. 9c,g). These conditions suppress TC activity in the

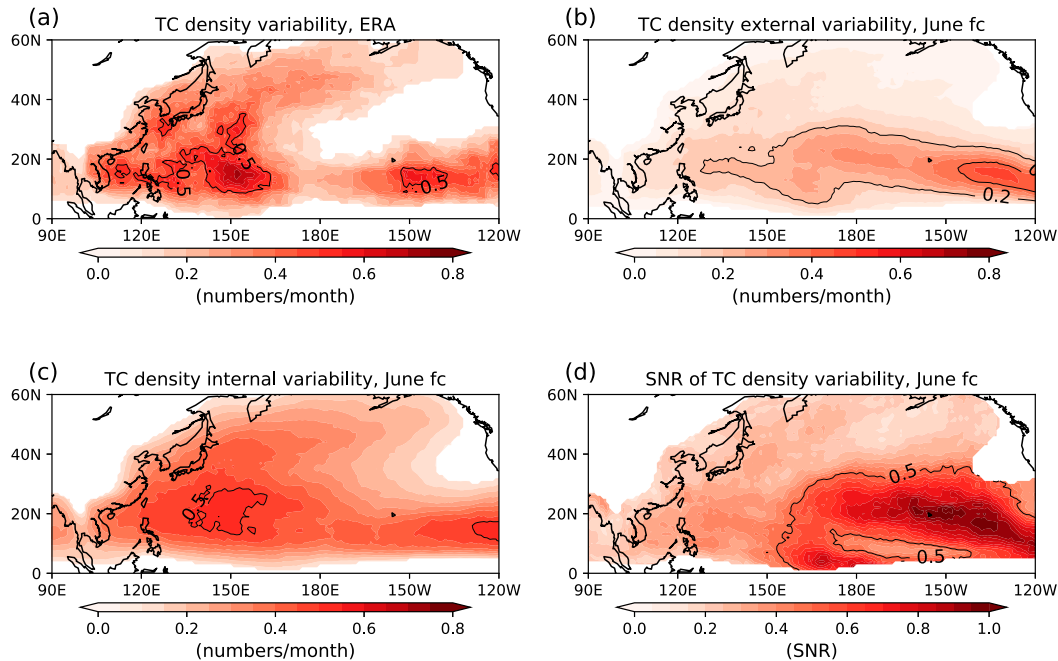


FIG. 10. (a) Interannual standard deviation of TC track density in ERA-Interim, with contours showing the values of 0.5 TCs month<sup>-1</sup>. (b) Interannual standard deviation of external variability of TC track density in GloSea5 June forecasts, with contours showing the values of 0.2 and 0.5 TCs month<sup>-1</sup>. (c) Interannual standard deviation of internal variability of TC track density in GloSea5 June forecasts, with contours showing the values of 0.5 TCs month<sup>-1</sup>. (d) SNR of interannual variability of TC track density in GloSea5 June forecasts, i.e., ratio of (b) to (c), with contours showing the values of 0.5. TC track density is calculated at each grid point with an area defined by a 10° × 10° box around the grid point, per month over JASO. The analysis period is 1993–2015.

western WNP (100°–150°E) and encourage TC activity in the eastern WNP (150°E–180°) during El Niño years, enhancing the effect of ENSO on TC track density (Fig. 6c).

The ENSO teleconnection to the steering flow in GloSea5 also deviates from that in ERA-Interim (Figs. 9d,h; see also Figs. S3d,h). In the tropics the westerly steering flow due to El Niño is stronger than expected, while in the subtropics the steering flow due to El Niño is weaker, related to the smaller changes in the NPSH. Combined with the systematically weaker steering flow in GloSea5 (Fig. 4d), the errors in the steering flow–ENSO teleconnection slow TCs traveling from east to the west in El Niño years, and cause TCs to begin to recurve farther east than observed (Fig. S4b). In the end, too many TCs travel to the eastern WNP and too few TCs travel to the north and west. The environment–ENSO teleconnections ( $r$  and  $\beta$ ) in the May and April forecasts are provided in the supplemental material (Figs. S5 and S6). At longer lead time, the teleconnections to the NPSH and steering flow in the subtropical Pacific become even weaker and shifted farther east, increasing the teleconnection bias. This is consistent with the effect of lead time on the TC–ENSO teleconnection for the WNP (Figs. 6 and 7).

### c. Interannual variability of TC frequency

#### 1) POTENTIAL PREDICTABILITY

The regional interannual variability of TC track density is shown in Fig. 10. In ERA-Interim, the TC frequency variability

is largest east of the Philippines at >0.5 TCs month<sup>-1</sup> (Fig. 10a) or 50% of climatology. In GloSea5 for the June forecasts, the ensemble mean exhibits substantially weaker variability over these regions (i.e., <0.2 TCs month<sup>-1</sup>; Fig. 10b). The internal variability due to model uncertainty is large in the WNP (Fig. 10c). The SNR of TC external variability is <0.5 in regions of the WNP, and >0.5 in the eastern Pacific (Fig. 10d); the potential predictability is stronger in the eastern Pacific, but weaker in the WNP. The weak potential predictability in the WNP TC frequency is related to the weak potential predictability in the environmental conditions (Fig. S7).

In the ensemble forecasts, the external variability is isolated through averaging the ensemble members. Efficient use of computational resources requires understanding of the minimum ensemble size to reliably represent external variability. Here, we take the 28-member ensemble mean as the “true” external variability. For each subsampled ensemble size [1 to 27 members; described in section 2c(2)], the ensemble means of the 2000 “sub-sampled ensembles” are then correlated with the “true” external variability. The average results for TC frequency in the WNP (0°–60°N, 100°E–180°) and the eastern Pacific (0°–60°N, 180°–120°W) basins are shown in Fig. 11. In the WNP, clearly, the external variability of TC frequency is captured better in GloSea5 for larger ensemble sizes (Fig. 11a). For small ensembles, potential prediction performance is higher for shorter lead time forecasts, suggesting the potential predictability increases faster with ensemble size at shorter

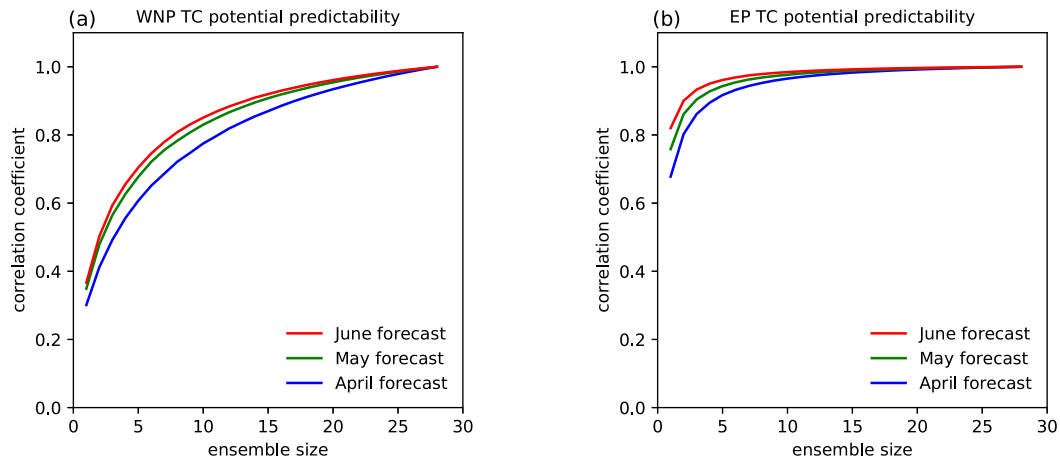


FIG. 11. (a) Average correlation coefficient for WNP TC frequency anomaly ( $0^{\circ}$ – $60^{\circ}$ N,  $100^{\circ}$ E– $180^{\circ}$ ) in JASO between the 28-member ensemble mean and GloSea5 subsampled ensemble mean, as a function of subset ensemble size, in June, May, and April forecasts. (b) As in (a), but for the eastern Pacific (EP;  $0^{\circ}$ – $60^{\circ}$ N,  $180^{\circ}$ – $120^{\circ}$ W). The analysis period is 1993–2015.

lead times. In the eastern Pacific, the potential predictability for external variability increases faster with small ensemble size than in the WNP (Fig. 11b).

## 2) ACTUAL PREDICTABILITY

The interannual correlation for regional TC frequency anomalies between GloSea5 and IBTrACS, measuring the actual predictability for TC frequency in GloSea5, is shown in Figs. 12a–c. The value of  $r$  is 0.5–0.8 in the central and eastern Pacific, where the TC–ENSO teleconnections are well represented (Figs. 6 and 7). In the SCS and northern China, where the weak TC–ENSO teleconnection is observed,  $r$  is 0.4–0.6. In the northeast WNP, where GloSea5 TC–ENSO teleconnections have the largest errors, the TC variability is negatively correlated with IBTrACS ( $r = -0.4$  to  $-0.2$ ). For the May and April forecasts, the performance is degraded in both the southwest and the northeast WNP. Performance is nearly identical when verified against ERA-Interim (Figs. 12d–f), except in northern China and the northeast WNP (south of Japan) where the performance degrades.

Prediction performance for basinwide TC frequency interannual variability as a function of ensemble size and lead time is shown in Figs. 13a and 13b. In the WNP, verified against IBTrACS, prediction performance increases with ensemble size, but is never significant at the 90% confidence level, with the highest  $r = 0.23$ , 0.16, and 0.08 for the June, May, and April forecasts, respectively. Verified against ERA-Interim TCs, the performance of the June forecasts increases (highest  $r = 0.45$ ), but it remains poor for May and April forecasts. In contrast, in the eastern Pacific, ensemble size has a very limited effect on TC variability prediction performance; the overall performance is  $r > 0.65$  for all lead times, with a slightly better performance against ERA-Interim than against IBTrACS. This is related to a good representation of the strong TC–ENSO teleconnections over this basin. It is worth noting the large discrepancies between the potential predictability (Fig. 11a) and actual prediction performance (Fig. 13a) for WNP TC frequency variability. This suggests further increasing actual performance

requires improving the underlying physical model in GloSea5, not increasing ensemble size.

It is important to emphasize the SCS ( $0^{\circ}$ – $25^{\circ}$ N,  $100^{\circ}$ – $120^{\circ}$ E), as it is the region where GloSea5 performs best for TC frequency variability in the marginal seas of the WNP (Fig. 12). The good performance in the SCS is associated with the accurate representation of the sign of the TC–ENSO teleconnection in GloSea5 (Figs. 6, 7). Figure 13c shows the prediction performance for SCS TC frequency anomalies as a function of lead time and ensemble size. The performance increases for shorter lead times and larger ensemble sizes; performance tends to saturate for ensemble sizes larger than 10. Verified against ERA-Interim, the performance of the 28-member ensemble mean for SCS TCs reaches  $r = 0.56$ , 0.47 (significant at the 95% confidence), and 0.19 for the June, May, and April forecasts, respectively. Verifying against IBTrACS increases performance, with the largest  $r = 0.63$ , 0.49 (significant at the 95% confidence), and 0.30 (significant at the 85% confidence) for the June, May, and April forecasts, respectively. The good performance of longer-lead-time forecasts in the SCS may be useful for early warnings.

## 4. Conclusions and discussion

We evaluated the performance of the U.K. Met Office GloSea5-GC2 global seasonal forecast system for predicting WNP TC frequency in JASO (July–October), from ensemble forecasts initialized in June, May, and April. Although GloSea5 captures the main features of the observed climatological TC frequency in the WNP, considerable biases remain, with TCs underpredicted in the western and northern WNP and substantially overpredicted in the east. WNP TC genesis and paths are both biased. The TC frequency bias may be related to El Niño–type biases in environmental conditions, such as mid-tropospheric moisture, wind shear, and SSTs. The WNP anticyclonic steering flow is also biased, collocated with a weakened NPSH, causing TCs to move slower in the tropics and recurve

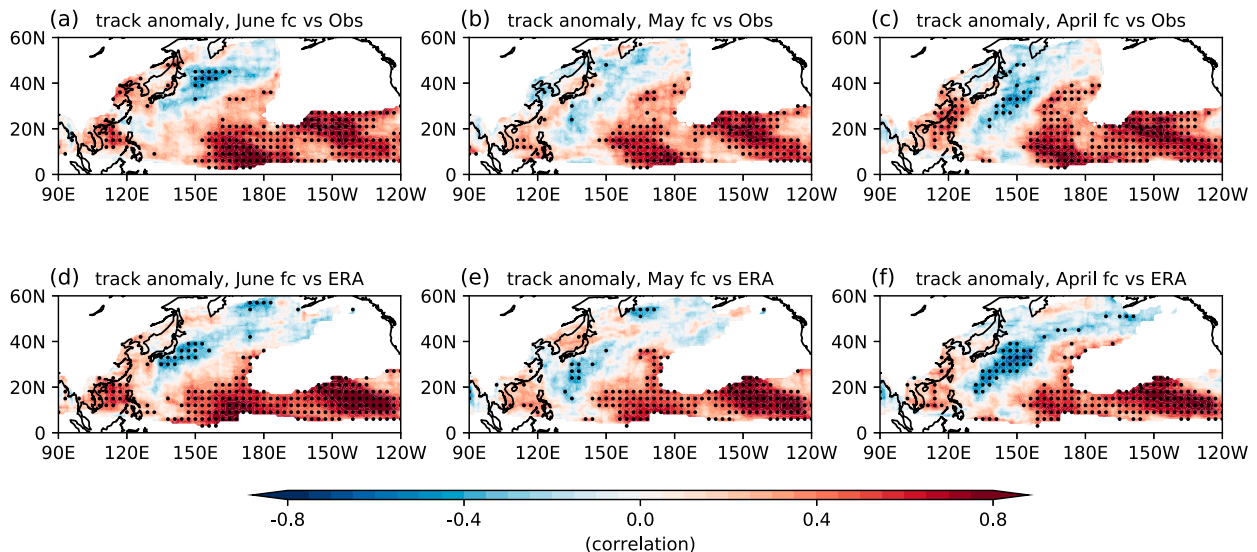


FIG. 12. (a)–(c) Correlation coefficient for TC track density between IBTrACS and GloSea5 June, May, and April forecast ensemble means. (d)–(f) As in (a)–(c), but for correlation between ERA-Interim and GloSea5 ensemble means. Dotted areas are where correlations pass significance test at the 90% confidence level. TC track density is calculated at each grid point as the number of TCs over an area defined by a  $10^\circ \times 10^\circ$  box around the grid point, per month over JASO. The analysis period is 1993–2015.

north farther east than observed. At longer lead times, the GloSea5 TC frequency bias increases in the central-eastern WNP, due to the bias growth both in cyclogenesis and in TC paths.

The WNP TC–ENSO teleconnection is misrepresented in GloSea5. A strong negative TC–ENSO correlation is simulated by GloSea5 in the north and the west, whereas in observations or reanalysis the correlation is strongly positive in the north and weakly negative in the southwest (e.g., the SCS). Although ENSO is well predicted, the impact on the TC environmental conditions (e.g., midlevel relative humidity, wind shear, and SST) expands too far zonally, which strengthens the effect of El Niño to suppress tropical cyclogenesis in the western WNP and to encourage it in the east. In El Niño years, too many TCs travel to the eastern WNP and too few TCs travel to the north and west, likely due to the strong westerly steering flow related to El Niño at the equatorial edge of the NPSH.

In GloSea5, the regional performance for the interannual TC frequency variability is strongly related to the representation of the regional TC–ENSO teleconnection. The best performance is in the regions where the TC–ENSO teleconnection is well represented; performance is moderate in the SCS, where the sign of the TC–ENSO teleconnection is correct but the magnitude is overestimated; performance is worst in the northeast WNP, where GloSea5 misrepresents both the sign and magnitude of the TC–ENSO teleconnection. In the WNP, GloSea5 performance for TC frequency variability is strongly a function of ensemble size, since the simulated TC–ENSO teleconnection has high intraensemble variability. In contrast, ensemble size has a very limited effect on performance in the eastern Pacific, where the strong TC–ENSO teleconnection is consistently represented across the ensemble members.

In GloSea5, the western tropical Pacific cyclogenesis bias may not be associated solely with local environmental conditions, but also with remote environmental conditions such as in the central tropical Pacific, where TC precursors (e.g., equatorial waves and tropical depression disturbances) are probably too active in GloSea5. SST biases in other ocean basins, such as the Atlantic and Indian Oceans, may also be relevant. Future work should investigate the representation of these precursors and remote conditions in GloSea5 and evaluate their effects on WNP TC prediction performance (e.g., Vecchi et al. 2019). Other modes of variability, such as the Indian Ocean dipole, tropical North Atlantic SST, and quasi-biennial oscillation, may also affect interannual WNP TC variability and prediction performance. Evaluating representability of teleconnections of these modes in GloSea5 is equally important to understand WNP TC prediction performance.

Reasonably representing the TC environmental conditions and their teleconnections is important for dynamical models to accurately predict the interannual WNP TC variability. The causes of GloSea5 biases in these fields are not immediately clear; biases may be related to the convection scheme and its sensitivity to moisture. Evaluating the genesis potential index (GPI; Camargo et al. 2007a; Zuo et al. 2018) was initially thought to be useful for understanding GloSea5 genesis biases. However, GPI is of limited use for model validation, because the globally fitted index does not well represent the mean or interannual variability of TC genesis at regional scales (including the WNP; e.g., Camargo et al. 2007a; Menkes et al. 2012), and also because the index was originally calibrated for the observed relationships between TC genesis and environmental factors, which may differ considerably in models (e.g., Camargo et al. 2007b). Hence, diagnosing the environmental factors in terms of GPI that are most responsible for the biases

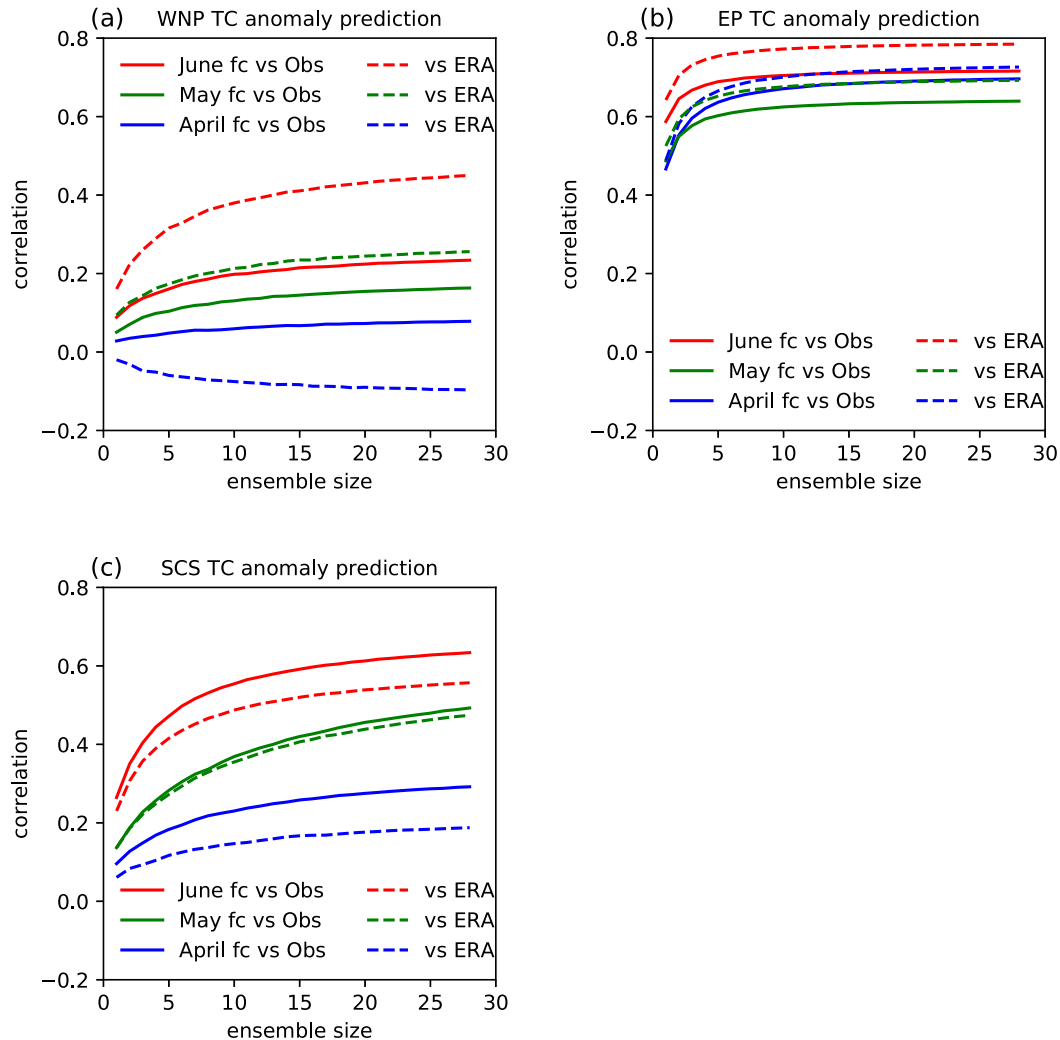


FIG. 13. (a) Average correlation coefficient for basinwide WNP TC frequency anomaly ( $0^{\circ}$ – $60^{\circ}$ N,  $100^{\circ}$ E– $180^{\circ}$ ) in JASO between IBTrACS (solid lines) and GloSea5 subsampled ensemble mean, as a function of subset ensemble size, in June, May, and April forecasts. Dashed lines are for correlation between ERA-Interim and GloSea5 forecasts. (b),(c) As in (a), but for the eastern Pacific (EP;  $0^{\circ}$ – $60^{\circ}$ N,  $180^{\circ}$ – $120^{\circ}$ W) and SCS ( $0^{\circ}$ – $25^{\circ}$ N,  $100^{\circ}$ – $120^{\circ}$ E) TC frequency anomalies, respectively. The analysis period is 1993–2015.

is not considered here in our paper. Understanding the sources of these errors remains challenging and an active area of research that requires detailed sensitivity experiments. Additionally, in the future, it will be worth comparing the regional performance of GloSea5 for WNP TCs with that of other forecast systems, including statistical and hybrid statistical–dynamical approaches.

**Acknowledgments.** Xiangbo Feng and Kevin Hodges were supported by the Met Office Climate Science for Service Partnership for China and the Weather and Climate Science for Service Partnership for Southeast Asia, as part of the Newton Fund. Nicholas Klingaman was supported by an Independent Research Fellowship from the Natural Environment Research Council (NE/L010976/1) and by the Global Challenges Research Fund, via Atmospheric hazards in developing Countries: Risk assessment and Early Warning (ACREW; NE/R000034/1). The authors are grateful

to the Met Office for making GloSea5 reforecast data available. ERA-Interim is generated by ECMWF and distributed by the Climate Data Store (CDS), under the Copernicus Climate Change Service (C3S). Computing and data storage facilities were provided by JASMIN (<https://jasmin.ac.uk>).

## REFERENCES

- Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167, [https://doi.org/10.1175/1525-7541\(2003\)004<1147:TVGP2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGP2.0.CO;2).
- Bell, R., K. Hodges, P. L. Vidale, J. Strachan, and M. Roberts, 2014: Simulation of the global ENSO–tropical cyclone teleconnection by a high-resolution coupled general circulation model. *J. Climate*, **27**, 6404–6422, <https://doi.org/10.1175/JCLI-D-13-00559.1>.



- Blockley, E. W., and Coauthors, 2014: Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new global FOAM forecasts. *Geosci. Model Dev.*, **7**, 2613–2638, <https://doi.org/10.5194/gmd-7-2613-2014>.
- Bowler, N. E., A. Arribas, S. E. Beare, K. R. Mylne, and G. J. Shutts, 2009: The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **135**, 767–776, <https://doi.org/10.1002/qj.394>.
- 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>.
- , K. A. Emanuel, and A. H. Sobel, 2007a: Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Climate*, **20**, 4819–4834, <https://doi.org/10.1175/JCLI4282.1>.
- , A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007b: Tropical cyclone genesis potential index in climate models. *Tellus*, **59A**, 428–443, <https://doi.org/10.1111/j.1600-0870.2007.00238.x>.
- Camp, J., and Coauthors, 2015: Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 2206–2219, <https://doi.org/10.1002/qj.2516>.
- , and Coauthors, 2019: The western Pacific subtropical high and tropical cyclone landfall: Seasonal forecasts using the Met Office GloSea5 system. *Quart. J. Roy. Meteor. Soc.*, **145**, 105–116, <https://doi.org/10.1002/qj.3407>.
- Chan, J. C., 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:TCATIN>2.0.CO;2](https://doi.org/10.1175/1520-0493(1985)113<0599:TCATIN>2.0.CO;2).
- Chen, J. H., and S. J. Lin, 2013: Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. *J. Climate*, **26**, 380–398, <https://doi.org/10.1175/JCLI-D-12-00061.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Fan, K., 2007: North Pacific sea ice cover, a predictor for the western North Pacific typhoon frequency? *Sci. China*, **50D**, 1251–1257, <https://doi.org/10.1007/s11430-007-0076-y>.
- Feng, X., N. P. Klingaman, and K. I. Hodges, 2019: The effect of atmosphere–ocean coupling on the prediction of 2016 western North Pacific tropical cyclones. *Quart. J. Roy. Meteor. Soc.*, **145**, 2425–2444, <https://doi.org/10.1002/qj.3571>.
- 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>.
- Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700, [https://doi.org/10.1175/1520-0493\(1968\)096<0669:GVOTOO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1968)096<0669:GVOTOO>2.0.CO;2).
- Guo, Y. P., X. Feng, N. P. Klingaman, and Z. Tian, 2020: Impact of Indo-Pacific warm pool Hadley circulation on the seasonal forecast performance for summer precipitation over the western North Pacific. *Environ. Res. Lett.*, <https://doi.org/10.1088/1748-9326/aba97c>, in press.
- Henderson-Sellers, A., and Coauthors, 1998: Tropical cyclones and global climate change: A post-IPCC assessment. *Bull. Amer. Meteor. Soc.*, **79**, 19–38, [https://doi.org/10.1175/1520-0477\(1998\)079<0019:TCAGCC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<0019:TCAGCC>2.0.CO;2).
- Hersbach, H., and D. P. Dee, 2016: ERA5 reanalysis is in production. *ECMWF Newsletter*, No. 147, ECMWF, Reading, United Kingdom, 7, <http://www.ecmwf.int/sites/default/files/elibrary/2016/16299-newsletter-no147-spring-2016.pdf>.
- Ho, C. H., H. S. Kim, J. H. Jeong, and S. W. Son, 2009: Influence of stratospheric quasi-biennial oscillation on tropical cyclone tracks in the western North Pacific. *Geophys. Res. Lett.*, **36**, L06702, <https://doi.org/10.1029/2009GL037163>.
- Hodges, K. I., 1994: A general method for tracking analysis and its application to meteorological data. *Mon. Wea. Rev.*, **122**, 2573–2586, [https://doi.org/10.1175/1520-0493\(1994\)122<2573:AGMFTA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<2573:AGMFTA>2.0.CO;2).
- , 1995: Feature tracking on the unit sphere. *Mon. Wea. Rev.*, **123**, 3458–3465, [https://doi.org/10.1175/1520-0493\(1995\)123<3458:FTOTUS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<3458:FTOTUS>2.0.CO;2).
- , 1999: Adaptive constraints for feature tracking. *Mon. Wea. Rev.*, **127**, 1362–1373, [https://doi.org/10.1175/1520-0493\(1999\)127<1362:ACFFT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<1362:ACFFT>2.0.CO;2).
- , and R. Emerton, 2015: The prediction of Northern Hemisphere tropical cyclone extended life cycles by the ECMWF ensemble and deterministic prediction systems. Part I: Tropical cyclone stage. *Mon. Wea. Rev.*, **143**, 5091–5114, <https://doi.org/10.1175/MWR-D-13-00385.1>.
- , and N. P. Klingaman, 2019: Prediction errors of tropical cyclones in the western North Pacific in the Met Office global forecast model. *Wea. Forecasting*, **34**, 1189–1209, <https://doi.org/10.1175/WAF-D-19-0005.1>.
- , A. Cobb, and P. L. Vidale, 2017: How well are tropical cyclones represented in reanalysis datasets? *J. Climate*, **30**, 5243–5264, <https://doi.org/10.1175/JCLI-D-16-0557.1>.
- Johnson, S. J., A. Turner, S. Woolnough, G. Martin, and C. MacLachlan, 2017: An assessment of Indian monsoon seasonal forecasts and mechanisms underlying monsoon interannual variability in the Met Office GloSea5-GC2 system. *Climate Dyn.*, **48**, 1447–1465, <https://doi.org/10.1007/s00382-016-3151-2>.
- 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>.
- Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS) unifying tropical cyclone data. *Bull. Amer. Meteor. Soc.*, **91**, 363–376, <https://doi.org/10.1175/2009BAMS2755.1>.
- Li, R. C., and W. Zhou, 2012: Changes in western Pacific tropical cyclones associated with the El Niño–Southern Oscillation cycle. *J. Climate*, **25**, 5864–5878, <https://doi.org/10.1175/JCLI-D-11-00430.1>.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, <https://doi.org/10.1002/qj.2396>.
- Manganello, J. V., and Coauthors, 2016: Seasonal forecasts of tropical cyclone activity in a high-atmospheric-resolution coupled prediction system. *J. Climate*, **29**, 1179–1200, <https://doi.org/10.1175/JCLI-D-15-0531.1>.
- Maue, R. N., 2011: Recent historically low global tropical cyclone activity. *Geophys. Res. Lett.*, **38**, L14803, <https://doi.org/10.1029/2011GL047711>.
- 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>.

- Menkes, C. E., and Coauthors, 2012: Comparison of tropical cyclogenesis indices on seasonal to interannual timescales. *Climate Dyn.*, **38**, 301–321, <https://doi.org/10.1007/s00382-011-1126-x>.
- 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>.
- Strachan, J., P. L. Vidale, K. Hodges, M. Roberts, and M. E. Demory, 2013: Investigating global tropical cyclone activity with a hierarchy of AGCMs: The role of model resolution. *J. Climate*, **26**, 133–152, <https://doi.org/10.1175/JCLI-D-12-00012.1>.
- Takaya, Y., Y. Kubo, S. Maeda, and S. Hirahara, 2017: Prediction and attribution of quiescent tropical cyclone activity in the early summer of 2016: Case study of lingering effects by preceding strong El Niño events. *Atmos. Sci. Lett.*, **18**, 330–335, <https://doi.org/10.1002/asl.760>.
- 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>.
- , and Coauthors, 2019: Tropical cyclone sensitivities to CO<sub>2</sub> 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>.
- Vitart, F., and T. N. Stockdale, 2001: Seasonal forecasting of tropical storms using coupled GCM integrations. *Mon. Wea. Rev.*, **129**, 2521–2537, [https://doi.org/10.1175/1520-0493\(2001\)129<2521:SFOTSU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2521:SFOTSU>2.0.CO;2).
- Williams, K. D., and Coauthors, 2015: The Met Office Global Coupled Model 2.0 (GC2) configuration. *Geosci. Model Dev.*, **8**, 1509–1524, <https://doi.org/10.5194/gmd-8-1509-2015>.
- 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, <https://doi.org/10.1007/s00382-015-2618-x>.
- 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>.
- Zhang, G., H. Murakami, R. Gudgel, and X. Yang, 2019: Dynamical seasonal prediction of tropical cyclone activity: Robust assessment of prediction skill and predictability. *Geophys. Res. Lett.*, **46**, 5506–5515, <https://doi.org/10.1029/2019GL082529>.
- Zhang, W., Y. Leung, and J. Min, 2013: North Pacific gyre oscillation and the occurrence of western North Pacific tropical cyclones. *Geophys. Res. Lett.*, **40**, 5205–5211, <https://doi.org/10.1002/grl.50955>.
- , 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>.
- , —, —, —, T. L. Delworth, X. Yang, and L. Jia, 2018: Dominant role of Atlantic multidecadal oscillation in the recent decadal changes in western North Pacific tropical cyclone activity. *Geophys. Res. Lett.*, **45**, 354–362, <https://doi.org/10.1002/2017GL076397>.
- Zuo, H., T. Li, J. Liu, and M. Peng, 2018: Physical processes controlling earlier and later onset of a typhoon season in the western North Pacific. *Climate Dyn.*, **51**, 2807–2823, <https://doi.org/10.1007/s00382-017-4046-6>.