

# *A new approach to skillful seasonal prediction of Southeast Asia tropical cyclone occurrence*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Feng, X. ORCID: <https://orcid.org/0000-0003-4143-107X>,  
Hodges, K. ORCID: <https://orcid.org/0000-0003-0894-229X>,  
Hoang, L., Pura, A. G., Yang, G.-Y. ORCID:  
<https://orcid.org/0000-0001-7450-3477>, Luu, H., David, S. J.,  
Duran, G. A. M. W. and Guo, Y.-p. (2022) A new approach to  
skillful seasonal prediction of Southeast Asia tropical cyclone  
occurrence. *Journal of Geophysical Research: Atmospheres*,  
127 (12). e2022JD036439. ISSN 2169-8996 doi:  
10.1029/2022JD036439 Available at  
<https://centaur.reading.ac.uk/105539/>

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/10.1029/2022JD036439>

Publisher: American Geophysical Union

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

## A New Approach to Skillful Seasonal Prediction of Southeast Asia Tropical Cyclone Occurrence

Xiangbo Feng<sup>1</sup> , Kevin I. Hodges<sup>1</sup>, Lam Hoang<sup>2</sup> , Alvin G. Pura<sup>3</sup> , Gui-Ying Yang<sup>1</sup>, Huyen Luu<sup>2</sup>, Shirley J. David<sup>3</sup>, Ger Anne M. W. Duran<sup>3</sup>, and Yi-Peng Guo<sup>4</sup> 

<sup>1</sup>National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK, <sup>2</sup>National Centre of Hydro-Meteorology Forecasting, Vietnam Meteorological Hydrological Administration (VNMHA), Hanoi, Vietnam, <sup>3</sup>Research and Development and Training Division, Philippine Atmospheric Geophysical and Astronomical Services Administration (PAGASA), Quezon City, Philippines, <sup>4</sup>Key Laboratory of Mesoscale Severe Weather/Ministry of Education, School of Atmospheric Sciences, Nanjing University, Nanjing, China

### Key Points:

- A new approach to identify independent predictors of Southeast Asia (SEA) tropical cyclone (TC) frequency is introduced
- The preceding-season predictors represent different clusters of TC formation and propagation through a cross-season effect
- The new method outperforms existing statistical and dynamical predictions for SEA TCs

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

X. Feng,  
[xiangbo.feng@reading.ac.uk](mailto:xiangbo.feng@reading.ac.uk)

### Citation:

Feng, X., Hodges, K. I., Hoang, L., Pura, A. G., Yang, G.-Y., Luu, H., et al. (2022). A new approach to skillful seasonal prediction of Southeast Asia tropical cyclone occurrence. *Journal of Geophysical Research: Atmospheres*, 127, e2022JD036439. <https://doi.org/10.1029/2022JD036439>

Received 3 JAN 2022  
Accepted 3 JUN 2022

© 2022. The Authors.  
This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

**Abstract** Predicting the peak-season (July–September) tropical cyclones (TCs) in Southeast Asia (SEA) several months ahead remains challenging, related to limited understanding and prediction of the dynamics affecting the variability of SEA TC activity. Here, we introduce a new statistical approach to sequentially identify mutually independent predictors for the occurrence frequency of peak-season TCs in the South China Sea (SCS) and east of the Philippines (PHL). These predictors, which are identified from the pre-season (April–June) environmental fields, can capture the interannual variability of different clusters of peak-season TCs, through a cross-season effect on large-scale environment that governs TC genesis and track. The physically oriented approach provides a skillful seasonal prediction in the 41-year period (1979–2019), with  $r = 0.73$  and  $0.54$  for SCS and PHL TC frequency, respectively. The lower performance for PHL TCs is likely related to the nonstationarity of the cross-season TC–environment relationship. We further develop the statistical approach to a hybrid method using the predictors derived from dynamical seasonal forecasts. The hybrid prediction shows a significant skill for both SCS and PHL TCs, for lead times up to four or 5 months ahead, related to the good performance of models for the sea surface temperatures and low-level winds in the tropics. The statistical and hybrid predictions outperform the dynamical predictions, showing the potential for operational use.

**Plain Language Summary** Tropical cyclones (TCs) have huge impact in Southeast Asia (SEA). Seasonal forecasts of SEA peak-season (July–September) TCs remains challenging for both statistical and dynamical models. In this paper, we introduce a novel statistical approach to sequentially identify independent predictors tailored for the frequency of SEA peak-season TCs. The predictors are routinely constructed from far-field sea surface temperatures and low-level winds in the preceding-season. The main concept in this approach is that each predictor represents a unique cluster of TC formation and propagation in the peak-season. We also develop the approach to a hybrid prediction method by building the predictors from dynamical seasonal forecasts. The seasonal predictions based on the new approach perform much better than the current statistical and dynamical models. Such predictions can provide useful warning service for SEA TC activity 4–5 months ahead.

## 1. Introduction

Tropical cyclones (TCs) can have large impacts over land and marginal seas in the western North Pacific (WNP), especially in the Southeast Asia (SEA) region where countries, such as the Philippines and Vietnam, are among the most exposed in the world to TC-related hazards. Each year there are over 20 TCs on average affecting SEA, with about 50% of them occurring in the peak-season of July–September (JAS), where the TC season is usually defined as June–November. Accurate prediction of peak-season SEA TCs up to several months ahead can provide an early warning service, which may significantly reduce the damage from TCs to economic activity and social welfare.

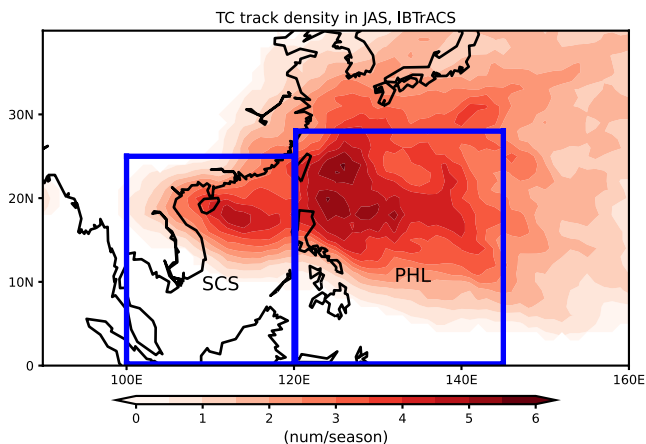
Dynamical general circulation models (GCMs) have been developed as a useful tool for seasonal TC prediction, especially in recent years with the development of high-resolution climate models related to the rapidly increased computing capability. Dynamical seasonal forecast systems have been developed in the major meteorological

centers, including the European Center for Medium-Range Weather Forecasts (ECMWF; Manganello et al., 2016; Vitart & Stockdale, 2001), Geophysical Fluid Dynamics Laboratory (Chen & Lin, 2013; Vecchi et al., 2014; Zhang et al., 2019) and the UK Met Office (Camp et al., 2015). While these systems show some degree of skill in predicting the basin-wide TC activity (i.e., genesis, track, and accumulated cyclone energy), they often perform poorly for WNP TCs at subbasin and country-level scales, in particular the SEA region (e.g., Befort et al., 2021; Camp et al., 2019; Feng, Klingaman, Hodges, & Guo, 2020; Zhang et al., 2019). The lack of regional skill is related to the model biases in the TC environmental conditions and the teleconnections from large-scale climate drivers to regional TCs (Befort et al., 2021; Feng, Klingaman, Hodges, & Guo, 2020). To improve the regional TC performance, further model evaluation and improvement such as those proposed in Feng, Klingaman, Zhang, and Guo (2020) are required.

An alternative to dynamical models is to use a statistical approach. The WNP TC activity is associated with multiple environmental factors, such as sea surface temperatures (SSTs), wind shear and the North Pacific Subtropical High, which affect genesis (e.g., Chen et al., 2006, 1998; Cheung, 2004; Chia & Ropelewski, 2002) and storm steering (e.g., Camp et al., 2019; Wang & Wang, 2019). Sources of interannual variability in these factors include the strength and position of large-scale circulations, such as the west Pacific Monsoon Trough (MT) (e.g., Chen et al., 2004; Molinari & Vollaro, 2013; Wu et al., 2012) and the Pacific Hadley and Walker circulations (e.g., Chan, 1985; Feng et al., 2021; Guo & Tan, 2018; Hsu et al., 2004; Zhang & Wang, 2015; Zhou & Cui, 2008). These factors are not mutually independent of climate variability, such as the El Niño-Southern Oscillation (ENSO) (e.g., Camargo & Sobel, 2005; Chan, 1985; Liu & Chan, 2008; Mei et al., 2015; Patricola et al., 2018; Wang & Chan, 2002; Zhang, Vecchi, Murakami, Delworth, et al., 2016; Zhao et al., 2020, 2010), the Pacific Meridional Mode (PMM) (e.g., Gao et al., 2018; Zhang, Vecchi, Murakami, Villarini, & Jia, 2016) and Atlantic Multidecadal Oscillation (e.g., Zhang et al., 2018). Statistical seasonal prediction of WNP TCs is often made through various predictors, which in principle represent the year-to-year variability of these environmental factors (e.g., Chen et al., 1998; Fan & Wang, 2009; Goh & Chan, 2010; Kim et al., 2017; Wang, Xiang, & Lee, 2013; Wang et al., 2019). Statistical methods can be further developed into hybrid statistical-dynamical methods, in which predictors are extracted from the environment variables that are dynamically predicted by GCMs (e.g., Camp et al., 2019; Kim et al., 2017; Li et al., 2017; Qian et al., 2020; Zhan & Wang, 2016).

The statistical and hybrid approaches are widely used to make basin-wide TC predictions with lead times up to several months, but they are usually less skillful for TCs at regional (subbasin and country-level) scales (e.g., Camargo & Sobel, 2005; Goh & Chan, 2010; Li et al., 2017; Wang et al., 2019). Regional TC variability is affected by compound effects of multiple large-scale drivers and local environmental conditions. The modulation of regional TCs by large-scale drivers can be nonlinear and complicated, and it may not be well represented in the conventional correlation analysis. Large-scale driver indices such as the Nino3.4 index alone cannot represent such variability well. Instead of focusing on the well-known climate variability, Wang et al. (2019) introduced a region-dependent seasonal forecast method, in which several predictors are identified in parallel for TC genesis in different sub-regions of the WNP. A similar method was also developed for the prediction of landfalling TC frequency on the Chinese coast (Tian & Fan, 2019). One caveat in these methods is that the predictors are not completely independent, and this could cause the prediction performance to be unrealistically inflated.

Another challenge to using statistical seasonal prediction methods is the strong seasonal variation in their performance. The typhoon season in the WNP is from June to November, but the interannual variability of TC activity varies greatly from month to month throughout the whole season (Feng et al., 2021). The statistical prediction systems, such as those operational in SEA (VNMHA and PAGASA contributing coauthors, personal communications), rely strongly on ENSO indices and have much lower skill in predicting the peak-season TC frequency than predicting the early-season (AMJ) or late-season (OND) frequency. This suggests that the complexity or nonlinearity of the ENSO effect (Corporal-Lodangco et al., 2016; Duy et al., 2019), including its coupling with the effects from other large-scale drivers, could not be well understood and included in these prediction systems. For example, the competing effect of the Indian Ocean SSTs in the summer of ENSO years (Xie et al., 2009, 2016) could weaken or prohibit the simultaneous or lagged relationships between the eastern Pacific SSTs (the ENSO route) and SEA TCs (e.g., Lyon & Camargo, 2009; Wang & Chan, 2002; Wang, Li, et al., 2013). The study reported in this paper is motivated by the low performance of the existing seasonal prediction (including dynamical models, which will be further described in Section 4) for the frequency of peak-season TCs in SEA.



**Figure 1.** Climatology of peak-season (JAS) tropical cyclone (TC) track density from International Best Track Archive for Climate Stewardship (IBTrACS), over 1979–2019. Blue boxes indicate the SEA subregions: the South China Sea (SCS; 0°–25°N, 100°–120°E) and east of the Philippines (PHL; 0°–28°N, 120°–140°E). TC track density is calculated at each grid point as the number of TCs over an area defined by a 5° × 5° box centered around the grid point, per season.

Recent studies suggest that the cross-season evolution of the large-scale circulation, for example, the Indo-Pacific Warm Pool Hadley Circulation (e.g., Guo et al., 2020; Guo & Tan, 2018) and the North Pacific Subtropical High (Wang & Wang, 2019), is significantly related to the interannual variability of basin-wide WNP TCs. These findings indicate that the preceding-season large-scale environment could be important sources of seasonal TC prediction. A few recent studies have attempted to establish the cross-season statistical relationships between regional TC variability (genesis and landfall frequency) and the preceding-season environmental conditions (Tian & Fan, 2019; Wang et al., 2019). However, in these studies, physical interpretations of the preceding-season predictors, including independence of the predictors, remain largely uncertain.

The study reported here aims to better understand the interannual variability and predictability of the occurrence frequency of peak-season (JAS) SEA TCs. The following research questions are central to this paper: (a) is the frequency of peak-season SEA TCs related to the pre-season environment? (b) can we derive independent predictors from the pre-season environment to build a statistical forecast? (c) can we provide physical interpretations of the predictors to elaborate the cross-season effects on SEA TCs?, and (d) what is the performance of a hybrid statistical-dynamical method in predicting the SEA TC frequency?

The paper is structured as follows. In Section 2, the data are first described.

A sequential approach to identifying independent predictors for SEA TC frequency is then introduced. Section 3 starts with the description of the statistical predictors, followed by the interpretation of the predictors by linking them to the large-scale environment. The performance of the statistical prediction is evaluated. Finally, we develop a hybrid prediction method with the predictors extracted from dynamical seasonal forecasts. Conclusions and discussion are provided in Section 4.

## 2. Data and Methodology

### 2.1. Data

#### 2.1.1. Climate Reanalysis and TC Observations

Monthly environmental fields (e.g., SSTs, winds and relative humidity) in the period 1979–2019 from the ECMWF fifth generation climate reanalysis (ERA5) (Hersbach et al., 2020) are used to construct and interpret the predictors of the SEA TC frequency. The ERA5 data are also used to verify the predicted environmental fields from the dynamical models. ERA5 is produced using a 4D-Var data assimilation system which combines the model forecasts of the ECMWF Integrated Forecast System (IFS) with the inhomogeneous observations. SSTs in ERA5 are derived from the Hadley Center Ice and Sea Surface Temperature (HadISST2) data prior to 2007 and from OSTIA thereafter. The ERA5 high-resolution deterministic reanalysis (horizontal resolution of ~31 km) is used for this study.

We focus on the occurrence frequency of peak-season (JAS) SEA TCs. The SEA is further divided into two subregions (Figure 1), the South China Sea (SCS; 100–120°E and 0–25°N) and east of the Philippines (PHL; 120°–140°E and 0°–28°N). The number of TCs developing in, and moving through, the two regions during the peak-season is counted and considered. TC tracks are obtained from the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al., 2010, 2018), which merges TC tracks from different meteorological centers. The data source of IBTrACS used in this paper is from the Japan Meteorological Agency (JMA). We evaluated the sensitivity of our method to different sources of Best Track data, for example, using the data from the China Meteorological Administration (CMA), and the track data derived from ERA5 where TC identification is independent of Best Track (Feng et al., 2021; Hodges et al., 2017). Similar results were found, although the areas for building the predictors are slightly different. Here we only present the analysis with IBTrACS. We also use the climate indices (ENSO and PMM) from the NOAA Physical Sciences Laboratory to discuss the links between the TC predictors and climate variability.

### 2.1.2. Dynamical Models and Seasonal Hindcasts

Seasonal hindcasts from GCMs are used to construct the predictors in the hybrid prediction method. These dynamical models include the ECMWF fifth generation seasonal forecast system (SEAS5), and the UK Met Office Global Seasonal Forecasting System version 5 and 6 (GloSea5 and GloSea6). SEAS5 is configured with the IFS atmosphere model (cycle 43r1) at T319 spectral horizontal resolution (about 36 km) and 91 vertical model levels, the NEMO ocean model (version 3.0) with ORCA025 grid and 75 vertical levels, and the Louvain-la-Neuve sea ice model (version 2.0) (Johnson et al., 2019). In SEAS5, the atmosphere and land surface are initialized from ERA-Interim (Dee et al., 2011), and the ocean and sea ice are initialized from the ECMWF fifth generation ocean analysis system (Zuo et al., 2019). Initial condition perturbations and a stochastic physics scheme are used to create the forecast ensemble.

GloSea5, in scientific configuration Global Coupled 2.0 (GloSea5-GC2; MacLachlan et al., 2015; Williams et al., 2015), is based on the coupled Met Office Hadley Center Global Environment Model (version 3), in which the Met Office atmospheric Unified Model (UM) (version GA6) is coupled to the NEMO ocean model (version 3.4), the Joint U.K. Land Environment Simulator (version 6.0), and the Los Alamos Sea Ice Model (version 4.1). GloSea5 uses an N216 horizontal resolution (0.83° in longitude and 0.56° in latitude) with 85 vertical levels for the atmosphere, and the tripolar ORCA025 grid with 75 vertical levels for the ocean.

GloSea6 is the successor to GloSea5, with the same model resolution and ensemble size but a higher coupling frequency (increased from 3 hourly to hourly). GloSea6, in scientific configuration Global Coupled 3.2 (GloSea5-GC3.2), has the same model components as GloSea5 but with updated versions: UM in GA7.2, NEMO in version 3.6, land model in version 8.0 and sea ice model in version 5.1. These updates, ideally, have improved model physics and parameterizations. GloSea6 is currently the Met Office operational seasonal forecast system (Williams et al., 2018). In GloSea5 and GloSea6, the forecast ensemble is generated by perturbing atmospheric physics tendencies via the Stochastic Kinetic Energy Backscatter v2 (SKEB2; Bowler et al., 2009) scheme, with the same atmospheric initialization. Uncertainties in initial conditions are partially considered by having the ensembles initialized at different dates in each month. The atmosphere and land surface are initialized from ERA-Interim, while the ocean and sea ice are initialized from the Forecast Ocean Assimilation Model (FOAM) Ocean Analysis (Blockley et al., 2014).

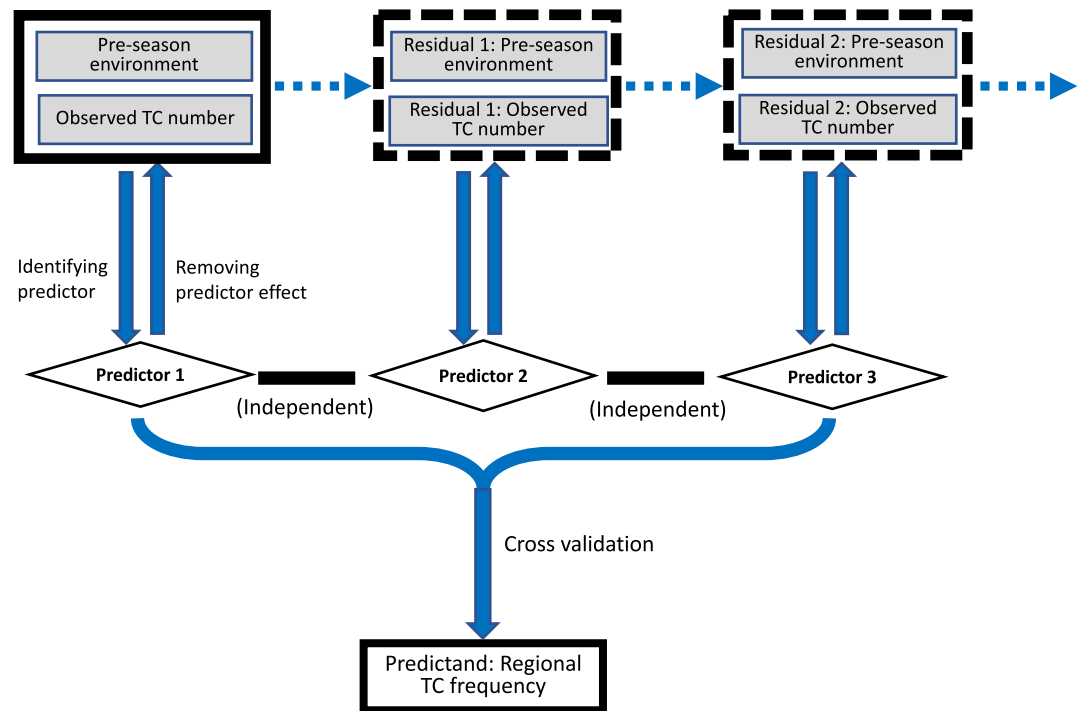
Historic seasonal forecasts (hindcasts) spanning the 23-year period 1993–2015 are used to develop the hybrid prediction system. SEAS5 hindcasts have 25 members starting on the first of the month, while GloSea hindcasts have four start dates in each month (the first, ninth, seventeenth, and 25th) with seven ensemble members per start date. For the GloSea hindcasts, the nominal month *M* includes the members initialized on the first of the month *M* and all available members with start dates back to, and including, the ninth of the month *M*-1 (28 members in total). For the SEAS5 hindcasts, the nominal month *M* includes all the members initialized on the first of the month *M* (25 members in total). Each hindcast member covers 215 days after the initialization date. For the hybrid prediction, we use hindcasts initialized in April, March and February. In the discussion section, we also evaluated the dynamical prediction of peak-season TCs in the hindcasts initialized in June and May; TCs were identified from twelve-hourly atmospheric fields, using the method described in Befort et al. (2021) and Feng, Klingaman, Hodges, and Guo (2020).

### 2.1.3. Statistical Analysis

The Pearson correlation coefficient (denoted as *r*) is used to measure the correlation between the time series of two variables. The time series are detrended before calculating the correlation. A two-tailed *t* test with *p* value of 0.05 (or 0.10) is used to test significance, with a null hypothesis of a zero correlation. In other words, the statistically significant values (*r*) are at the 95% (or 90%) confidence level, unless stated otherwise.

## 2.2. Identification of TC Predictors

In our analysis, the predictand is the frequency of TCs developing in, and moving through, the SCS and PHL in the peak-season (Figure 1). The predictors are extracted from the pre-season SSTs and low-level winds, which are critical for TC generation and development, over regions determined by correlation analysis. We also evaluated other environmental factors as potential predictors, such as relative humidity and vertical wind shear. However, these variables were found to have a less robust relationship with the TC numbers, presumably due to less



**Figure 2.** Workflow of the sequential orthogonal identification process (SOIP) to derive the predictors of the peak-season tropical cyclone (TC) frequency.

persistence of the variables across seasons. Moreover, the dynamical seasonal forecasts are found not to predict these variables as well as SSTs and low-level winds in the tropics, limiting the skill of the hybrid predictors.

We introduce a sequential orthogonal identification process (SOIP) to identify the statistical predictors for the predictand (TC frequency). Through this approach, the predictors are designed to represent different temporal components of TC variability, which are statistically independent. The schematic workflow in this approach is shown in Figure 2. The main processes are summarized in the workflow shown in Figure 2. The individual steps in the SOIP approach are described as follows:

1. The relationship (correlation and regression) between the predictand ( $F_0$ ) and pre-season environment (i.e., SST and 850 hPa zonal winds) at each grid point ( $E_0$ ) are calculated. The average environment over a continuous region where the  $p$  value of correlation is less than 0.1 is identified as the potential predictor. By doing this, several potential predictors (in the units of deg.C or m/s) can be found. Among these potential predictors, the one which has a physically meaningful relationship with the peak-season TC activity (i.e., formation, development and movement) is chosen as predictor 1 ( $P_1$ ). The predictor is then rescaled to the standard deviation of TC frequency via multiplying  $\sigma_{TC}/\sigma_p$ , where  $\sigma_{TC}$  is the interannual standard deviation of TC frequency and  $\sigma_p$  is the standard deviation of predictor. The final predictor has the same units as the predictand (times/season). Note that the shape of the region responding to the predictor can be irregular as it is essentially determined by the strength of the TC-environment relationship.
2. The effect of predictor 1 ( $P_1$ ) is removed from the predictand ( $F_0$ ) and pre-season environment, to obtain their residuals ( $F_1$  and  $E_1$ ). To do this, the linear regressions of yearly predictand ( $F_0$ ) and pre-season environment in each grid ( $E_0$ ) against the yearly predictor 1 ( $P_1$ ) are first estimated; the values of predictand and environment determined by predictor 1 via the estimated regressions are then removed from the time series of the predictand and pre-season environment at each grid point, to yield the residuals of predictand ( $F_1$ ) and environment ( $E_1$ ). This step is to ensure that in the following identification process the previously identified predictor has no statistical relation to the next predictors. In other words, the sequentially identified predictors are statistically unrelated.
3. The relationship between the predictand and environment residuals ( $F_1$  and  $E_1$ ) is further estimated. As in Step 1, the average of the environment over the regions, where the statistical relationship is significant at the 90%

- confidence level, are taken as the potential predictors. The potential predictor that has physical meaning to the peak-season TC activity is chosen and rescaled as the predictor 2 ( $P_2$ ).
- To identify another predictor ( $P_i$ ), similarly, the linear effects of the previously identified predictors ( $P_1, P_2, \dots, P_{i-1}$ ) are removed sequentially from the predictand and environment to obtain the new residuals ( $F_{i-1}$  and  $E_{i-1}$ ); the new predictor ( $P_i$ ) is then constructed from the residuals ( $F_{i-1}$  and  $E_{i-1}$ ). In principle, predictors are the regional averages of either SSTs or 850 hPa zonal winds, which have physically meaningful relationships with the predictand, and are statistically independent of each other ( $r = 0$ ) because of the sequential removal process.
  - Finally, a multivariable prediction model is built to predict the regional TC frequency. The model is formed by combining regressions between the predictand (or predictand residuals) and predictors. In each of  $N$  prediction years, the predicted total frequency of regional TCs, denoted as  $F_{total}^P$ , is expressed as:

$$F_{total}^P = \beta_0 + \sum_{i=1}^n \beta_i P_i$$

where  $\beta_0$  is the climatic average of the TC frequency (a long-term average), and  $\beta_i$  is the regression coefficient between the frequency and predictor ( $F_{i-1}$  and  $P_i$ ).  $\beta_i$  is estimated separately for individual predictors. For example,  $\beta_1$  is calculated from the time series of  $F_0$  and  $P_1$ , and  $\beta_2$  is calculated from the time series of residual  $F_1$  and  $P_2$ . For the statistical prediction, the  $\beta_i$  value is estimated through a cross-validation procedure for each year. Block elimination is employed to minimize potential skill inflation from serial correlation. For each prediction year,  $\beta_i$  is trained using  $N-5$  years of the frequency and predictor ( $F_{i-1}$  and  $P_i$ ) with a 5-year omitted block centered on the year of prediction. Then, the number of TCs in the prediction year is predicted using this estimated  $\beta_i$  and the predictor ( $P_i$ ) for the prediction year. This means that in the statistical prediction the  $\beta_i$  value varies with years. In the cross-validation procedure, we also tested the sensitivity of prediction performance to the length of omitted years, by changing the length to 1, 3, and 7 years (this will be discussed in Section 3.3). For the hybrid prediction, the  $\beta_i$  value for each predictor  $P_i$  is constant during the forecast period because the time series of  $P_i$  is directly extracted from the environment over the same region as the statistical predictor but from dynamical forecasts (i.e., no cross-validation involved).

Through the iteration of steps 1–4, a sequential set of predictors can be identified. In the rest of this paper, we only consider the first few predictors, which in total explain at least 50% of the variance of TC frequency. More predictors are needed if the percentage of the variance of TC frequency explained by predictors is set to be higher. We note that it becomes difficult to identify more than four predictors, as those areas of environment associated with TCs become smaller and sparser. We cannot rule out that the set of predictors identified by the SOIP approach could still be physically related via nonlinear and complicated processes.

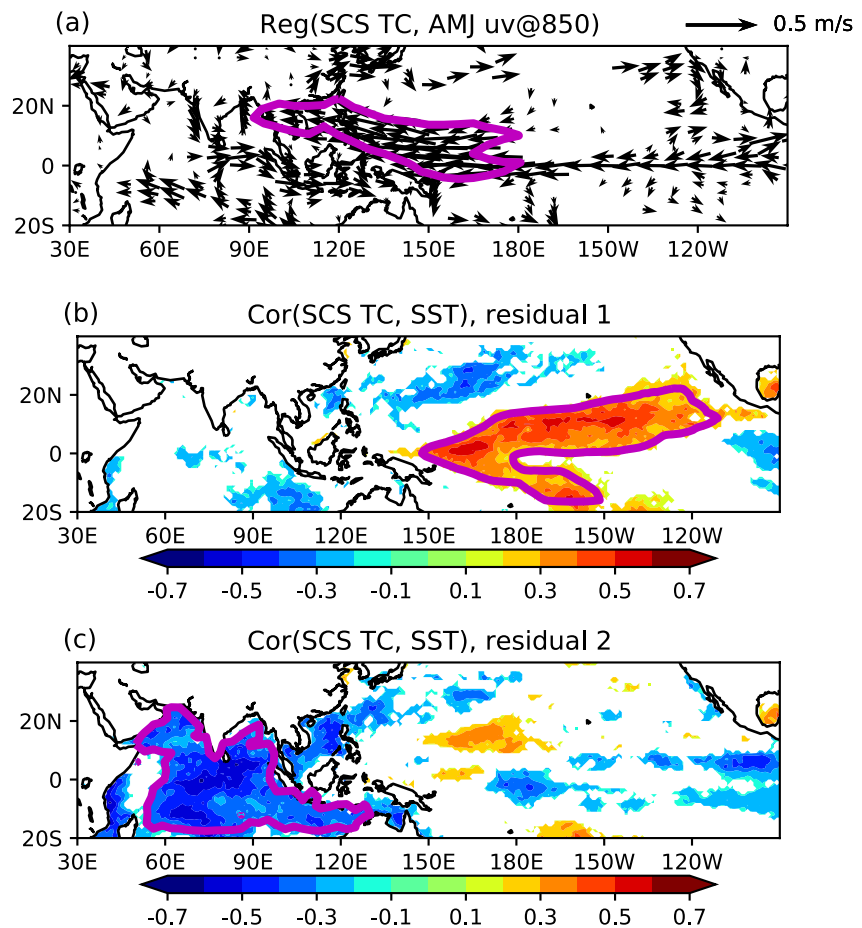
### 3. Results

#### 3.1. Constructed Predictors

For the SCS TC frequency, to capture at least 50% of the variance over the 41-year period 1979–2019, three predictors are derived using the SOIP method from the pre-season environmental fields in ERA5 (Figure 3). Predictor 1 is constructed from the average of 850-hPa easterly wind anomalies over a region with westward tilt with latitude in the western-central tropical Pacific (Figure 3a). Predictor 2 is based on the average of warm SST anomalies over a region with eastward tilt with latitude in the central-eastern tropical Pacific (Figure 3b). Predictor 3 is constructed from the average of cold SST anomalies over the Indian Ocean (Figure 3c). The time series of the three predictors and the corresponding TC frequency are shown in Figure S1 in Supporting Information S1. Predictors 1–3 are significantly correlated with the year-to-year variability of the corresponding TC frequency, with  $r = 0.48, 0.53,$  and  $0.60$  (all significant at the 95% confidence level). Combining the TC frequency associated with individual predictors yields the total frequency. The total frequency constructed from the three predictors matches well the observed total frequency, with  $r = 0.80$  (i.e., explaining ~65% of the variance). The physical meanings of the predictors will be discussed in Section 3.2.

For the PHL TC frequency, four predictors are derived from the pre-season environment (Figure 4). Predictor 1 is constructed from the average of 850-hPa easterly wind anomalies over a region in the central tropical

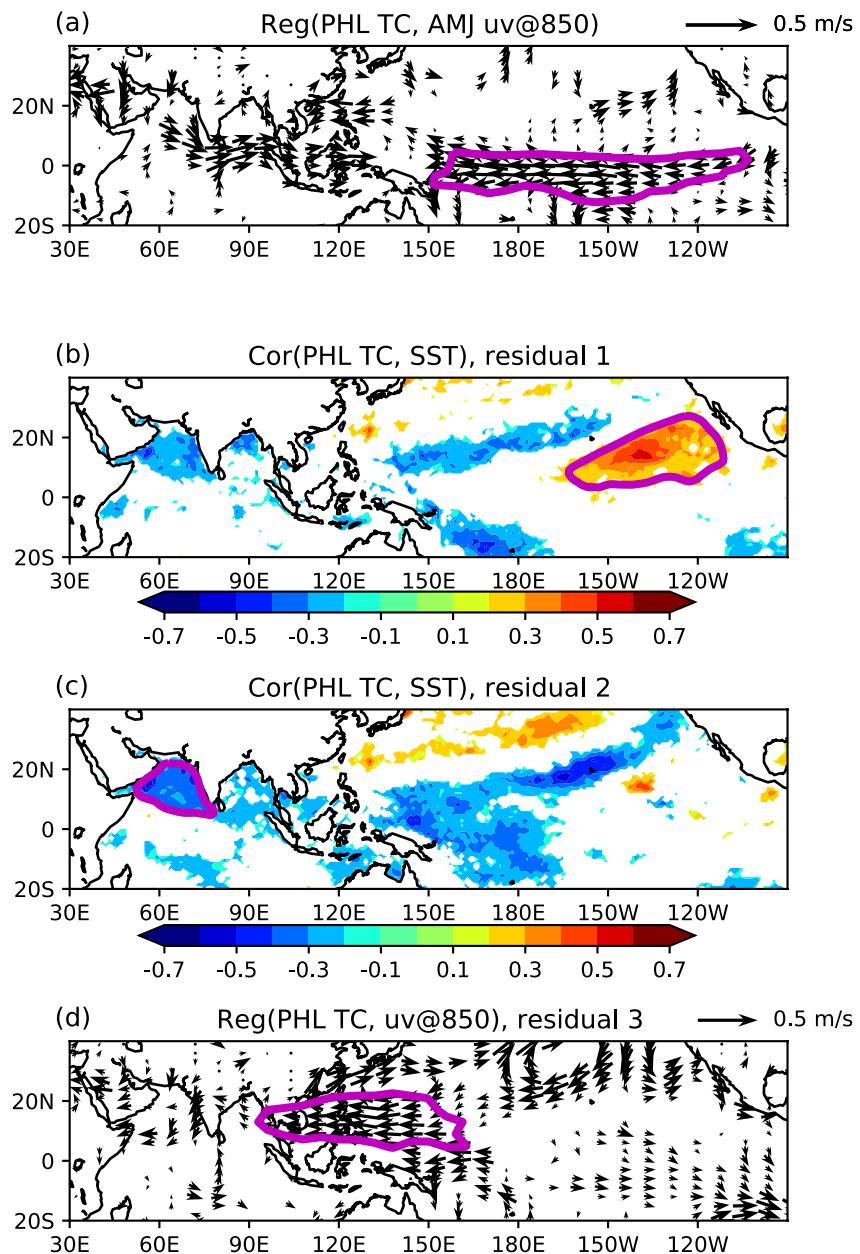




**Figure 3.** Preseason large-scale environment used to construct the South China Sea (SCS) tropical cyclone (TC) predictors over 1979–2019. (a) For predictor 1: average of 850-hPa zonal wind velocity over the area surrounded by magenta contour, where the TC-wind regressions (vectors) are significant at the 90% confidence level. (b) For predictor 2: average of sea surface temperatures (SSTs) over the area surrounded by magenta contour, where the TC-SST correlations (shading) are significant, after removing the effect of predictor 1 (i.e., residual 1). (c) For predictor 3: average of SSTs over the area surrounded by magenta contour, where the TC-SST correlations (shading) are significant, after removing the effect of predictors 1 and 2 (i.e., residual 2). *Note.* The predictors, which are the average of SSTs or winds, are rescaled to the standard deviation of observed TC frequency, in the units of times/season.

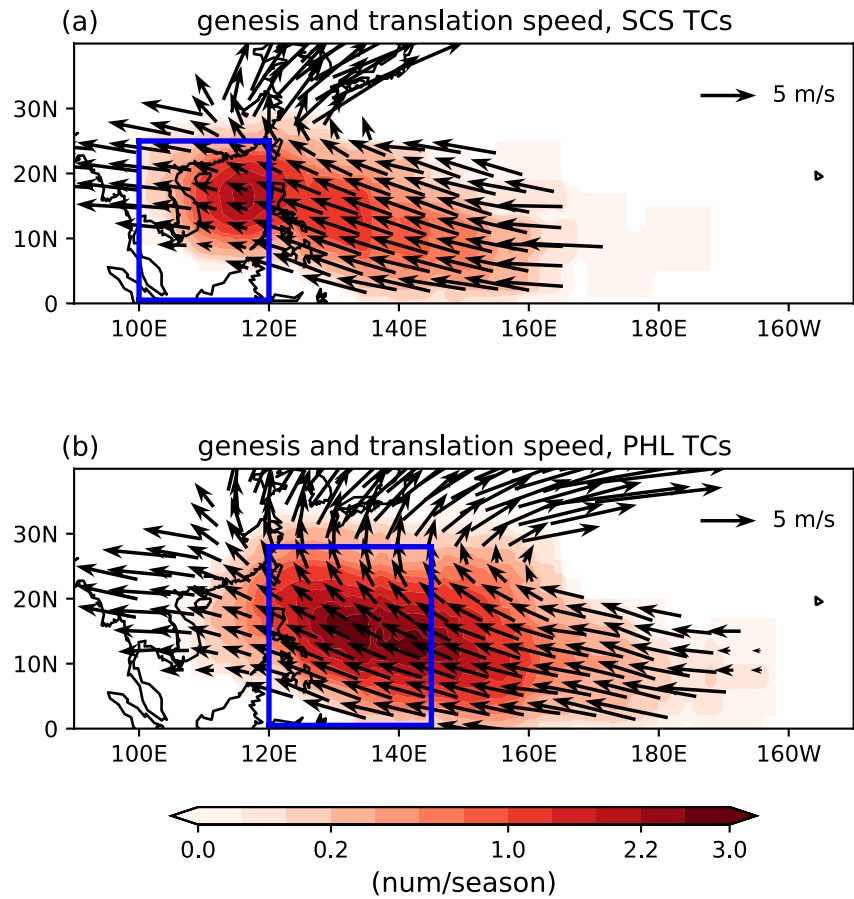
Pacific (Figure 4a). Predictor 2 is from the warm SST anomalies over a region tilted eastward with latitude in the north-eastern tropical Pacific (Figure 4b). Predictor 3 is based on the cold SST anomalies in the Arabian Sea (Figure 4c), and predictor 4 is based on the 850-hPa easterly wind anomalies over a region in the western tropical Pacific (Figure 4d). The four predictors are significantly correlated with the corresponding PHL TC frequency, with  $r = 0.37, 0.40, 0.43,$  and  $0.49$  (all significant at the 95% confidence level) (Figure S2 in Supporting Information S1), respectively. The constructed total frequency explains  $\sim 55\%$  of the observed variance, with  $r = 0.74$ .

It is worth noticing that although the predictors of the SCS and PHL TCs are identified separately, they are not independent. Table S1 in Supporting Information S1 provides the cross correlations between the predictors for the two subregions. The SCS TC predictor 1 is significantly correlated with the PHL TC predictors 1 and 4, with  $r = 0.52$  and  $0.62$ , respectively. This is related to the fact that the SCS TC predictor 1 is defined by the low-level easterly winds over a region, which extends across the two regions used to define the PHL TC predictors 1 and 4 (Figures 3 and 4). The SCS TC predictor 2 is correlated with the PHL TC predictor 2, with  $r = 0.72$ . The SCS TC predictor 3 is significantly associated with the PHL TC predictors 1, 2, and 3, with  $r = 0.48, -0.33,$  and  $0.52$ , respectively. This suggests that despite the different regions the individual components of SCS and PHL TCs share several common drivers.



**Figure 4.** As Figure 3, but for the environment used to construct the Philippines (PHL) tropical cyclone (TC) predictors.

There are a few factors that can affect the identification of the predictors. First, there is uncertainty in observing the number of TCs affecting the two regions, especially for weak storms and when the stages of the storm lifecycle are weak (Knaff et al., 2010; Torn & Snyder, 2012; Velden et al., 2006). Using different sources of TC data (either produced by different meteorological agencies or tracked from climate reanalyses) helps to understand the uncertainty in the predictor identification. We found that when the TC tracks from ERA5 (Feng, Klingaman, Hodges, & Guo, 2020; Hodges et al., 2017) are used, the regions responsible for the PHL TC predictors are slightly larger than the regions when the IBTrACS data are used (not shown). This might be related to the longer life cycle of the storm identified in climate reanalyses (Hodges et al., 2017). Excluding weak storms and the weak stages of storms may help to address the uncertainty in the predictors. However, the reduced sample size (especially for the SCS TCs) and the late-genesis stage (Feng et al., 2021) may in turn cause the TC-predictor relationship less robust. Second, the predictors, and the regions responsible for the predictors, could vary with the analysis periods. For example, during the period 1979–2019, the PHL TC frequency is significantly related to the

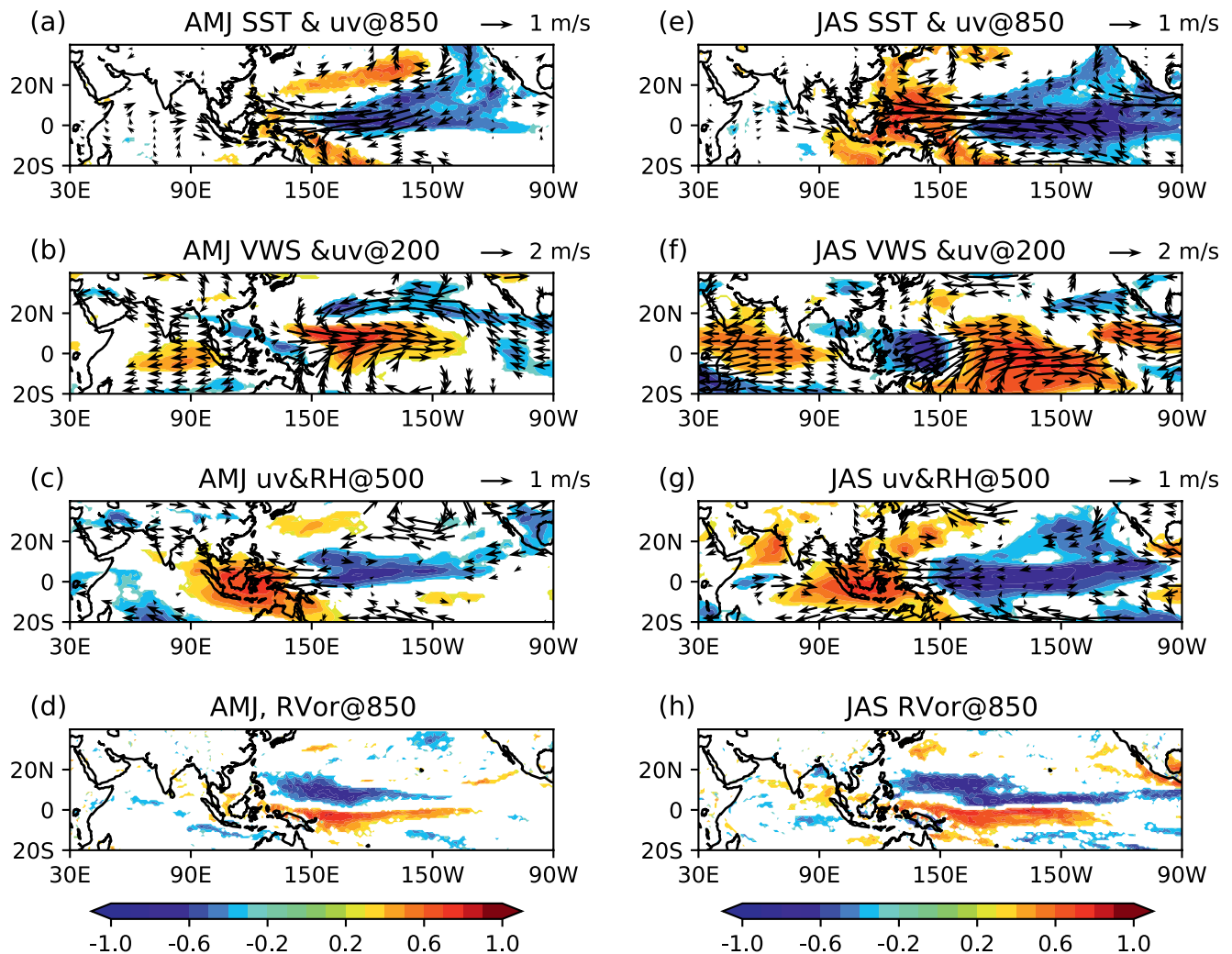


**Figure 5.** Climatology of observed genesis density (shading) and translation velocity (vectors) for South China Sea (SCS) tropical cyclones (TCs) (a) and Philippines (PHL) TCs (b) in July–September (JAS), over 1979–2019. Blue boxes indicate the areas of SCS and PHL. TC genesis density is calculated at each grid point as the number of TCs over an area defined by a  $10^\circ \times 10^\circ$  box centered around the grid point, per season.

low-level zonal winds in the equatorial central Pacific (predictor 1). But, when considering the period 1993–2015 (analyzed in Section 3.4), this relationship becomes insignificant. This uncertainty could be related to the nonstationarity of the relationships between the pre-season environment and TC activity.

### 3.2. Predictor-Environment Relationship

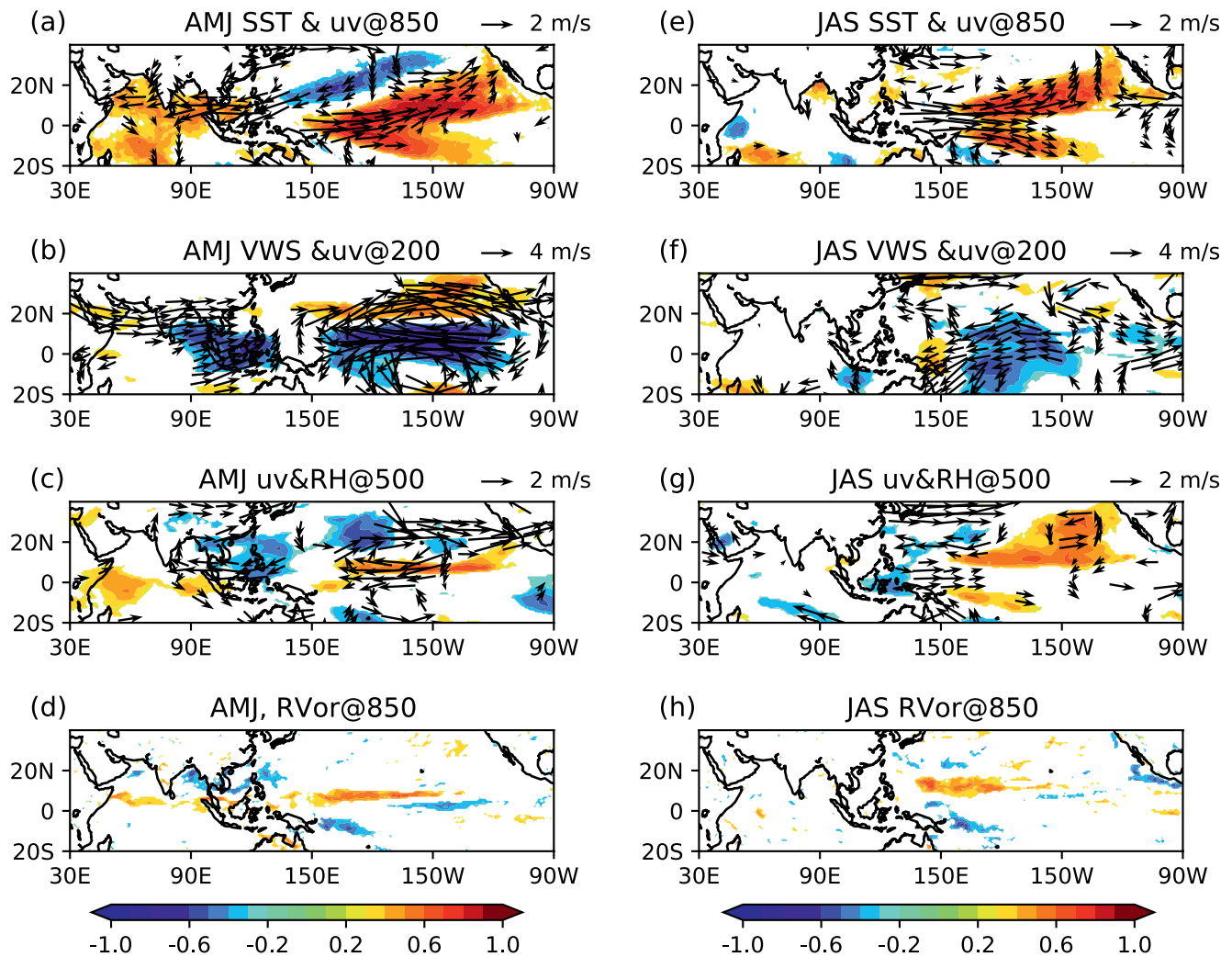
In this section, we interpret the predictors by examining the relationship between the predictors and the environmental conditions relevant to TC activity. We hypothesize that the predictors represent different clusters of TC genesis and track. TCs affecting the SCS and PHL regions include those storms that form within the region (local TCs) and others that form elsewhere but travel to this region (remote TCs). Figure 5 shows the climatology of genesis density and translation speed of the SCS and PHL TCs in the peak-season from IBTrACS. On average, there are about 7 TCs in the SCS in the peak-season, of which 61% are local TCs and 39% remote (Figure 5a). The remote TCs originate throughout a wide area in the western tropical Pacific ( $120^\circ\text{--}180^\circ\text{E}$ ), and travel north-westward to the SCS. Associated with the easterly steering flow and the North Pacific Subtropical High to the north, most SCS TCs (both local and remote) move westward or north-westward toward Vietnam and southern China (Feng et al., 2019; Wang, Xiang, & Lee, 2013). On the other hand, some local TCs travel eastward toward the Philippines, related to a south-westerly steering flow (Yang et al., 2015). The eastward-moving storms are responsible for an overall small translation speed inside the region. There are on average 13 TCs affecting the PHL in the peak-season (Figure 5b), of which 58% are local storms, 38% form to the east of the region ( $145^\circ\text{E--}160^\circ\text{W}$ ), and 4% form in the SCS and travel to the PHL.



**Figure 6.** (a–d) Relationships between South China Sea (SCS) tropical cyclone (TC) predictor 1 and the pre-season (AMJ) environment (a: SSTs and 850-hPa winds, (b) vertical wind shear (VWS) and 200-hPa winds, (c) 500-hPa relative humidity (RH) and winds, (d) 850-hPa RVor), over 1979–2019. (e–h) as (a–d), but for relationships with the peak-season (JAS) environment. Shaded areas are for correlation coefficients, and vectors for regression coefficients with winds; only relationships that are significant at the 90% confidence level are shown.

### 3.2.1. SCS TC Predictors

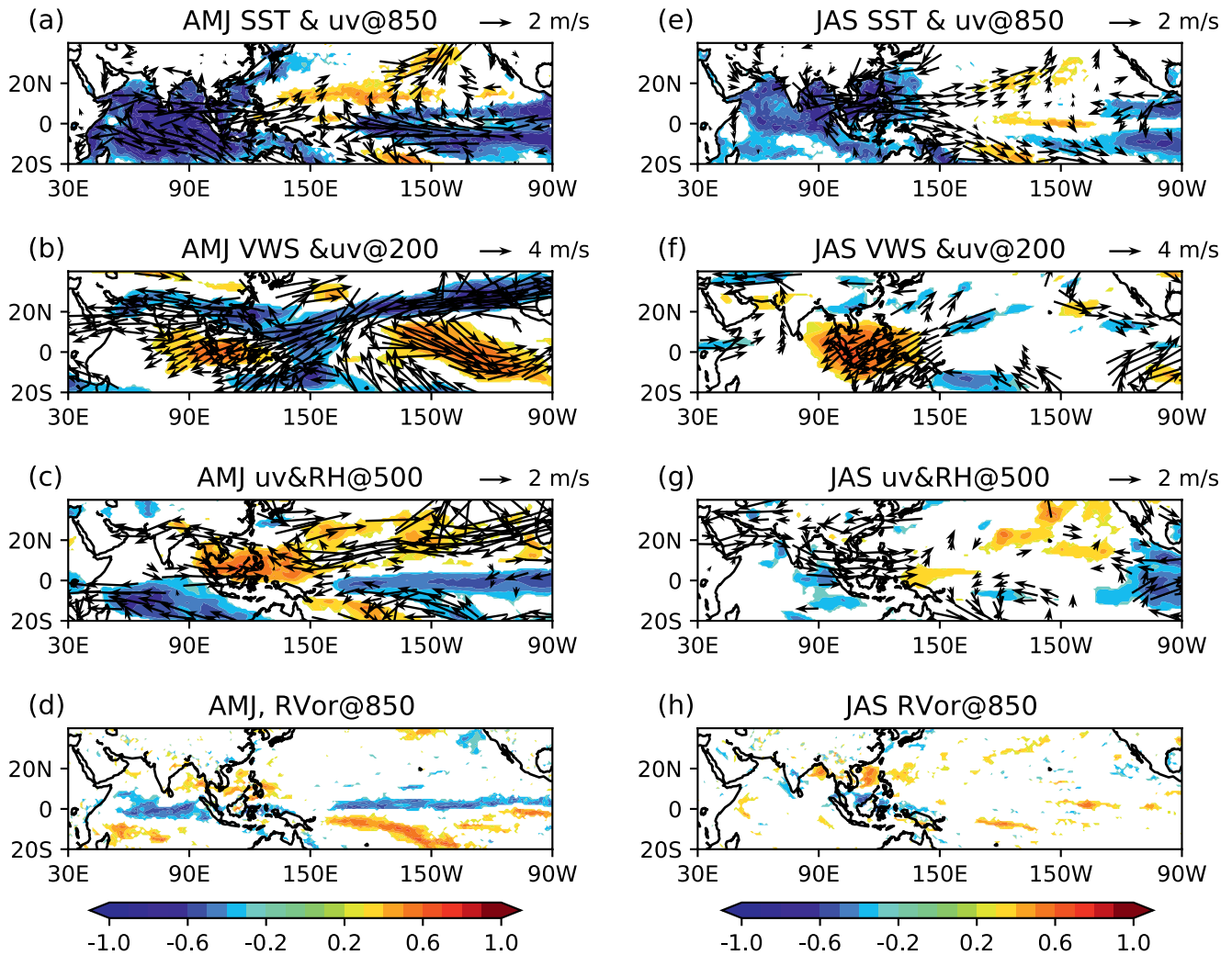
For the SCS TCs, the relationship between predictor 1 and the environmental conditions is shown in Figure 6. In the pre-season (Figures 6a–6d), positive anomalies of predictor 1 (anomalous easterly winds) indicate cold SST anomalies in the central-eastern tropical Pacific and warm SST anomalies in the subtropical Pacific and east of the Philippines. The SST changes are accompanied by an anomalous easterly wind in the lower troposphere and an anomalous cyclonic circulation in the upper troposphere. This indicates an enhanced Pacific Walker circulation related to an enhanced equatorial zonal wind-SST feedback (Cane et al., 1986). The zonal wind pattern leads to a low-level convergence and upper-level divergence in the SCS. The enhanced large-scale circulation is also associated with an increase in vertical wind shear (VWS) and a decrease in midlevel relative humidity (RH) in the central tropical Pacific. In contrast, the SCS region sees a VWS decrease and RH increase. From the pre-season to peak-season, the predictor 1-related environmental conditions strengthen (Figures 6e–6h). In the tropical Pacific, the east-to-west SST gradients and low-level easterly wind anomalies become stronger in the peak-season, indicating a strengthening or westward shift of the summer Monsoon Trough (MT) (e.g., Basconicillo et al., 2021; Chen et al., 2019; Li et al., 2017; Teng et al., 2021). In short, with positive anomalies of predictor 1, the peak-season thermodynamic and ventilating conditions are evidently more favorable for TCs to form in the SCS and PHL regions. Additionally, with positive anomalies of the predictor 1, the easterly steering



**Figure 7.** As Figure 6, but for relationships with the South China Sea (SCS) tropical cyclone (TC) predictor 2.

flow in the west Pacific significantly strengthens in the peak-season, and this steers remotely originated TCs to enter the SCS. The low-level relative vorticity (RVor) is another important factor for TC genesis by providing synoptic-scale disturbances (Molinari & Vollaro, 2013). The negative correlations between predictor 1 and RVor in the northern central Pacific (Figures 6d and 6h) suggest that the predictor 1-related TCs are less likely to form in the central-western tropical Pacific (150°–180°E).

The relationship between SCS TC predictor 2 and the environmental conditions is shown in Figure 7. The predictor is strongly associated with the pre-season environment in the central tropical Pacific (Figures 7a–7d). The predictor 2-associated environment is depicted by warm SST and low-level westerly wind anomalies in the central tropical Pacific, suggesting enhanced zonal wind-SST feedback with a deeper oceanic thermocline and less upwelling (Cane et al., 1986). The warm SST and westerly wind anomalies could correspond to an eastward shift of the MT (Basconcillo et al., 2021; Chen et al., 2019; Molinari & Vollaro, 2013; Teng et al., 2021; Wu et al., 2012), and they are closely related to the PMM pattern (Chiang & Vimont, 2004). The correlation between predictor 2 and the pre-season PMM index is  $r = 0.69$ . There is also a strong anticyclonic circulation anomaly in the upper troposphere in the north Pacific. The wind anomalies are associated with low-level convergence, upper-level divergence, VWS decrease, RH and RVor increase in the central tropical Pacific. Toward the peak-season, these environmental conditions largely persist (Figures 7e–7h). Moreover, in the peak-season, the negative effect on VWS and the positive effect on midlevel RH and low-level RVor of predictor 2 become stronger

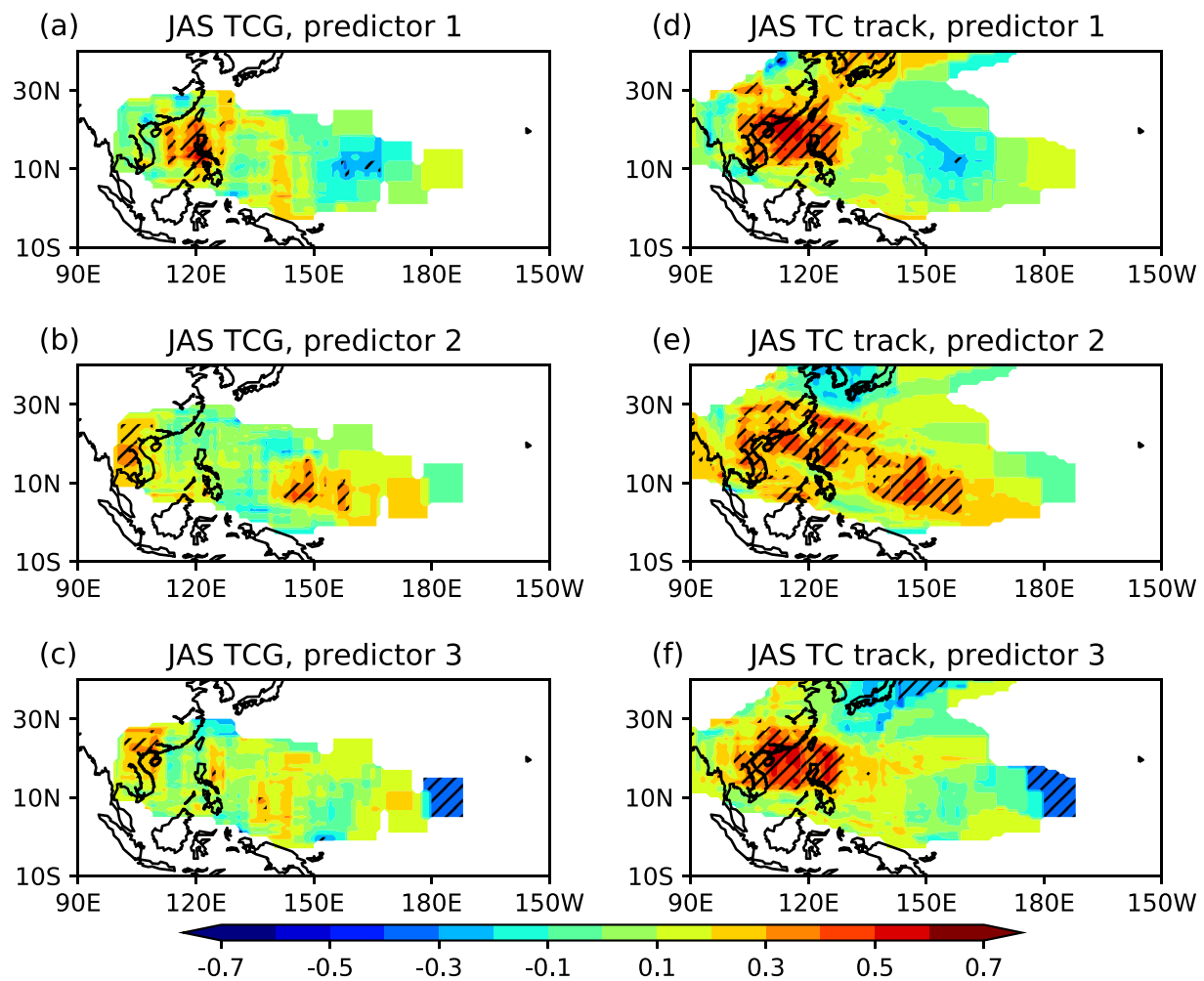


**Figure 8.** As Figure 6, but for relationships with the South China Sea (SCS) tropical cyclone (TC) predictor 3.

in the central tropical Pacific. This, together with anomalous easterly steering winds in which storms are embedded (around 20°N), encourages TCs to form in the central tropical Pacific and travel to the SCS.

For the SCS TC predictor 3, the relationship with the large-scale environment is shown in Figure 8. The most distinctive feature in the pre-season environment related to predictor 2 is the cold SST anomalies in the Indian Ocean (Figures 8a–8d). Related to this, the cross-equatorial southerly flow in the lower troposphere and the northerly flow in the upper troposphere over the Indian Ocean and west Pacific are significantly enhanced. This indicates intensified convection and a strengthened Hadley Circulation over the Indo-Pacific Warm Pool (Hoskins et al., 2020), which are strongly related to an increase in WNP TC activity (Guo et al., 2020; Guo & Tan, 2018). This large-scale circulation change can significantly reduce VWS and increase midlevel RH in the SCS and PHL regions. These environmental conditions, which are favorable for TCs to form in and travel to SEA, decay from the pre-season to peak-season (Figures 8e–8h). However, toward the peak-season, the related cross-equatorial overturning circulation in the Indo-Pacific Warm Pool shifts further east to SEA. This shift excites a strong low-level cyclonic circulation and high RVor to the north of the SCS, encouraging local TCs to form.

To support the above interpretations, we calculate the correlations between the predictors and spatial distributions of the SCS TCs. Figures 9a–9c shows the correlations with observed genesis density. (*Note.* This is only for the TCs affecting the SCS rather than all TCs in the WNP.) As expected, predictor 1 is significantly related to the TC genesis in the SCS and around the Philippines, and predictor 2 is related to the TCs that form in the open ocean (140°–180°E) and travel to the SCS. Predictor 3 has a less clear relation to any spatial clusters of genesis.



**Figure 9.** (a–c) Correlation coefficients between predictors and genesis density for South China Sea (SCS) tropical cyclones (TCs) in July–September (JAS), over 1979–2019. (d–f) as (a–c), but for correlations with TC track density. TC density is calculated at each grid point as the number of TC genesis or TC tracks over an area defined by a  $10^\circ \times 10^\circ$  box around the grid point, per season. Climatology of SCS TC genesis is shown in Figure 5a. Hatched areas show the correlation significant at the 95% confidence level.

However, the relationship between predictors and TC track density is more evident and supportive to our conclusions (Figures 9d–f). Predictors 1 and 3 represent the two clusters of TCs in the SCS, which are teleconnected to the central-eastern Pacific and Indian Ocean SSTs, respectively. Predictor two explains well the remote TCs, which is associated with the genesis-related conditions in the central tropical Pacific and the steering flow to the north. Thus, we confirm that the three derived predictors represent three different genesis and track clusters of SCS TCs (e.g., Camargo et al., 2007; Chen et al., 2017; Kim & Seo, 2016; Teng et al., 2021).

One key question is how the predictors are associated with ENSO. We found that all three predictors are significantly correlated with the preseason ENSO. Their correlations with the preseason Nino3.4 index are  $r = -0.48$ , 0.46, and  $-0.34$  (all significant at the 95% confidence level), respectively. As the predictors are statistically independent, the correlations with the Nino3.4 index indicate the diversity of the TC-ENSO relationship, possibly related to different types of ENSO (e.g., Modoki and mega-ENSO; Ashok et al., 2007; Kim et al., 2011; Wang, Li, et al., 2013; Wang, Liu, et al., 2013) and variations in the environment-ENSO teleconnections (e.g., Kim et al., 2010; Xie et al., 2009). We hypothesize that although the ENSO signal becomes weaker after the winter, the response of the remote environment in the following season could still be robust due to the lagged effect. Thus, we anticipate that the preseason ENSO (or the ENSO in the preceding winter) can affect the SCS TCs through different pathways, and a simple description of ENSO (as represented by the Nino3.4 index) cannot reveal the complexity of the ENSO impact. (*Note.* There is no significant correlation between the preseason Nino3.4 index

and SCS TC frequency.) For the peak-season Nino3.4, the correlations with the three predictors are  $r = -0.79$ ,  $0.08$ , and  $0.08$ , respectively. This suggests that the peak-season ENSO index holds a linear relationship with only a fraction of SCS TCs (the correlation between the peak-season Nino3.4 index and SCS TC frequency is  $r = -0.30$ ,  $p = 0.06$ ).

### 3.2.2. PHL TC Predictors

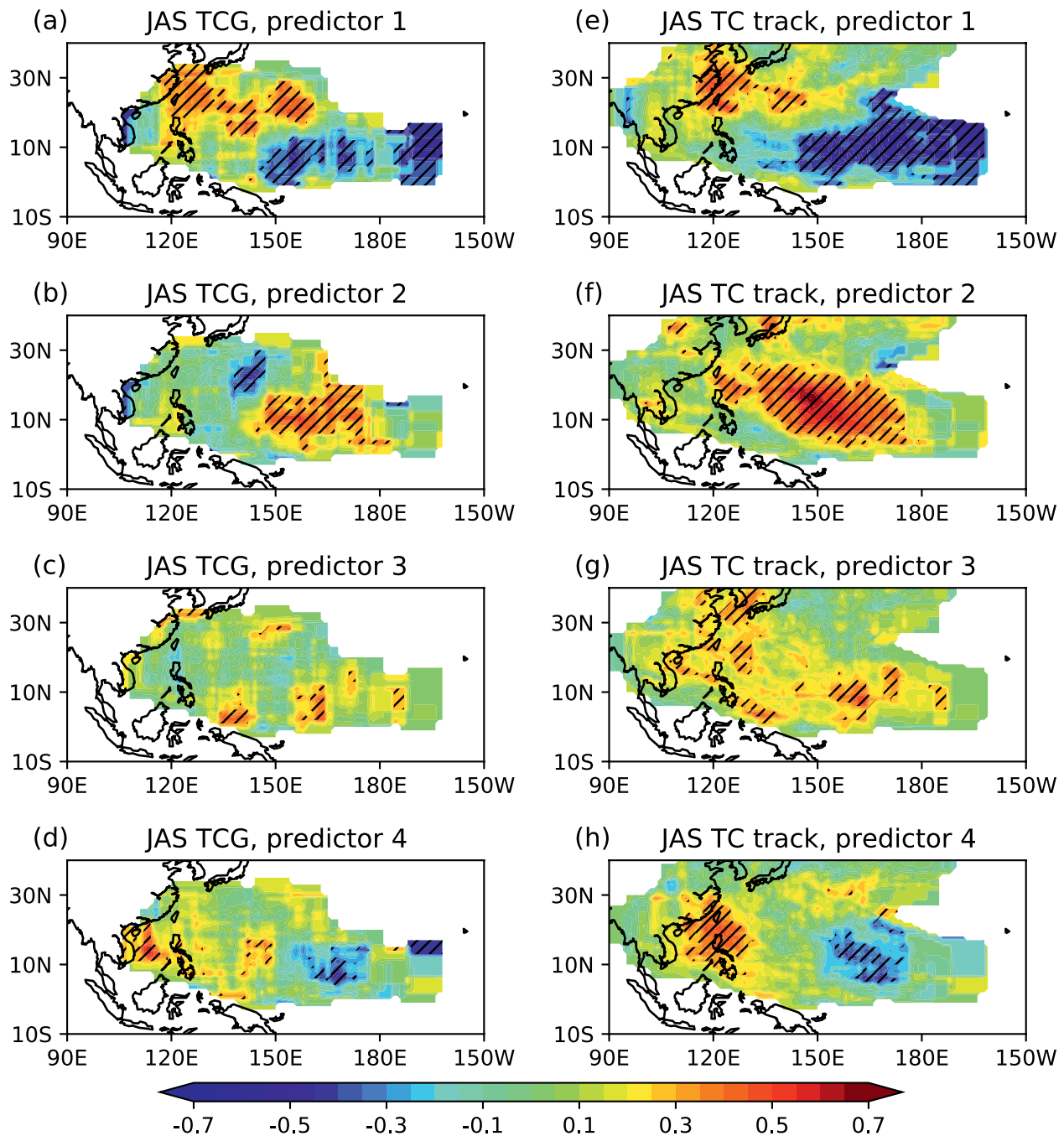
The large-scale environment related to the four predictors of the PHL TCs is shown in Figures S3–S6 in Supporting Information S1. The relationship between predictor 1 and the environment highlight the cross-season persistence of cold SST anomalies in the central-eastern tropical Pacific and the Indian Ocean (Figure S3 in Supporting Information S1). These SST changes occur with a significant enhancement of low-level convergence and upper-level divergence east of the Philippines, suggesting a strengthening regional Hadley circulation (e.g., Guo et al., 2020). The low-level westerly wind anomalies from the Indian Ocean and easterly wind anomalies from the central Pacific bring humid air to the western tropical Pacific and reduce the VWS in the SCS and PHL. These conditions persist across the two seasons. The PHL TC predictor 1 is anticorrelated with the pre-season (peak-season) ENSO index (Nino3.4), with  $r = -0.87$  ( $-0.56$ ). We also recall that the PHL TC predictor 1 is significantly associated with the SCS TC predictors 1 and 3 ( $r = 0.52$  and  $0.48$ , respectively), which represent the central-eastern Pacific and Indian Ocean SSTs separately. Thus, we anticipate that the PHL TC predictor 1 reveals two pathways of the ENSO effect on the PHL TCs, that is, the east pathway from the central Pacific and the west pathway from the Indian Ocean. Neither pre-season nor peak-season ENSO index has significant linear correlations with other predictors of the PHL TCs.

The PHL TC predictor 2, associated with the SCS TC predictor 2 ( $r = 0.72$ ), is positively correlated with the environmental conditions that are favorable for TCs to form in the central tropical Pacific, and favorable for them to travel further to the PHL (Figure S4 in Supporting Information S1). These environmental conditions well resemble the PMM spatial pattern, which was found to significantly modulate the TC activity in the south-east sector of the basin (e.g., Gao et al., 2018; Zhang, Vecchi, Murakami, Delworth, et al., 2016). The PHL TC predictor 3, associated with the SCS TC predictor 3 ( $r = 0.52$ ), reflects the favorable ventilation conditions in the southern part of the PHL, with the low-level convergence and upper-level divergence (Figure S5 in Supporting Information S1). The PHL TC predictor 4, correlated with the SCS TC predictor 1 ( $r = 0.62$ ), represents the cross-season persistence of environmental conditions that are favorable for local PHL TCs to form, that is, warm SSTs, lower VWS, cyclonic circulation anomaly and stronger low-level convergence (Figure S6 in Supporting Information S1).

As with the SCS TCs, the PHL TCs can be decomposed into different clusters (Corporal-Lodangco & Leslie, 2016). Figure 10 shows that the four predictors can capture different spatial clusters of PHL TCs. Predictor 1 is positively correlated with the PHL TCs formed in the northwest sector of the WNP, and negatively correlated with TCs formed in the southeast sector, consistent with the TC-ENSO correlations (e.g., Feng, Klingaman, Hodges, & Guo, 2020). This reemphasizes the close relation of predictor 1 to ENSO. As expected, predictor 3 is related to the TCs that form in the southern part of the PHL and move north-westward. Predictor 2 represents the westward-moving TCs that originate to the east of the PHL, while predictor 4 tends to capture both the westward-moving TCs that generate in the PHL and the eastward-moving TCs that generate in the SCS.

We note that the predictors may not be mutually independent of the variability of the MT, which is defined as the contiguous region where monthly 850-hPa RVor is positive (e.g., Molinari & Vollaro, 2013). The MT is associated with rich synoptic-scale disturbances and favorable environmental conditions (e.g., large low-level vorticity, small VWS, warm SSTs and humid air), and is crucial for the WNP TCs; over 70% of them are born within the MT. The MT is a “TC genesis bed” (e.g., Wu et al., 2012, and references therein). The identified predictors could be associated with the year-to-year variability in the position and strength of the MT. For example, for positive anomalies of the PHL TC predictor 1, the enhanced west-to-east SST gradient and low-level easterly winds along the equatorial Pacific suggest a seesaw change in the MT: strengthening east of the Philippines and weakening to the east. This favors an increase in TC activity in the PHL (Figures 10a and 10e). We anticipate plausible links between the MT and other predictors for PHL and SCS TCs (not described here). All these are consistent with the previous findings that a strengthening or an eastward shift of the MT encourages more TCs to form in the south-eastern sector of the basin, and a weakening or westward-shift of the MT favors more TCs to form in the western sector (Basconcillo et al., 2021; Chen et al., 2019; Molinari & Vollaro, 2013; Teng et al., 2021; Wu



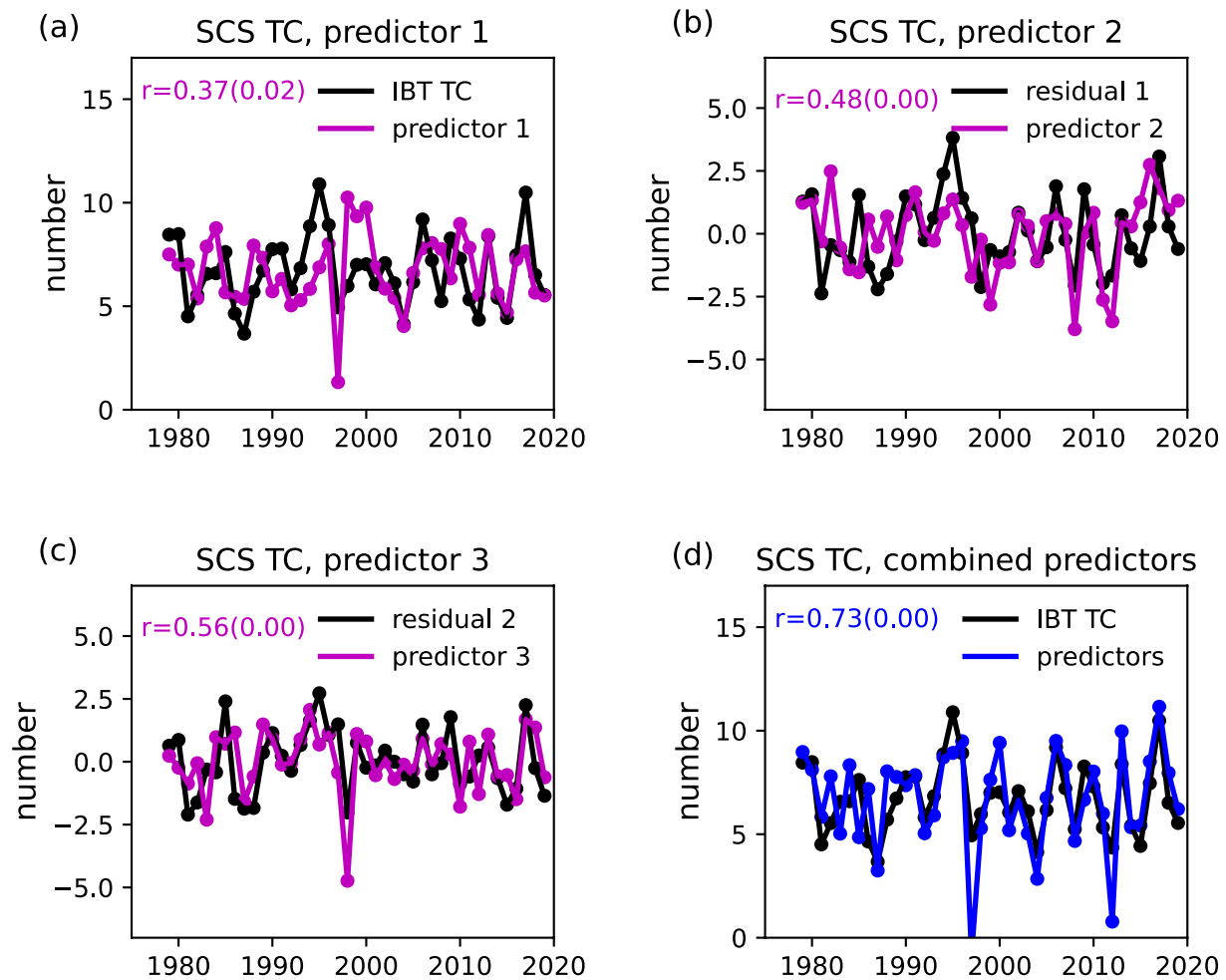


**Figure 10.** As Figure 9, but for the Philippines (PHL) tropical cyclones (TCs). Climatology of PHL TC genesis is shown in Figure 5b.

et al., 2012). The merit of the TC predictors is that they are not only associated with the genesis conditions (e.g., the MT) but also related to the background steering winds in which TCs are embedded.

### 3.3. Statistical Prediction Performance

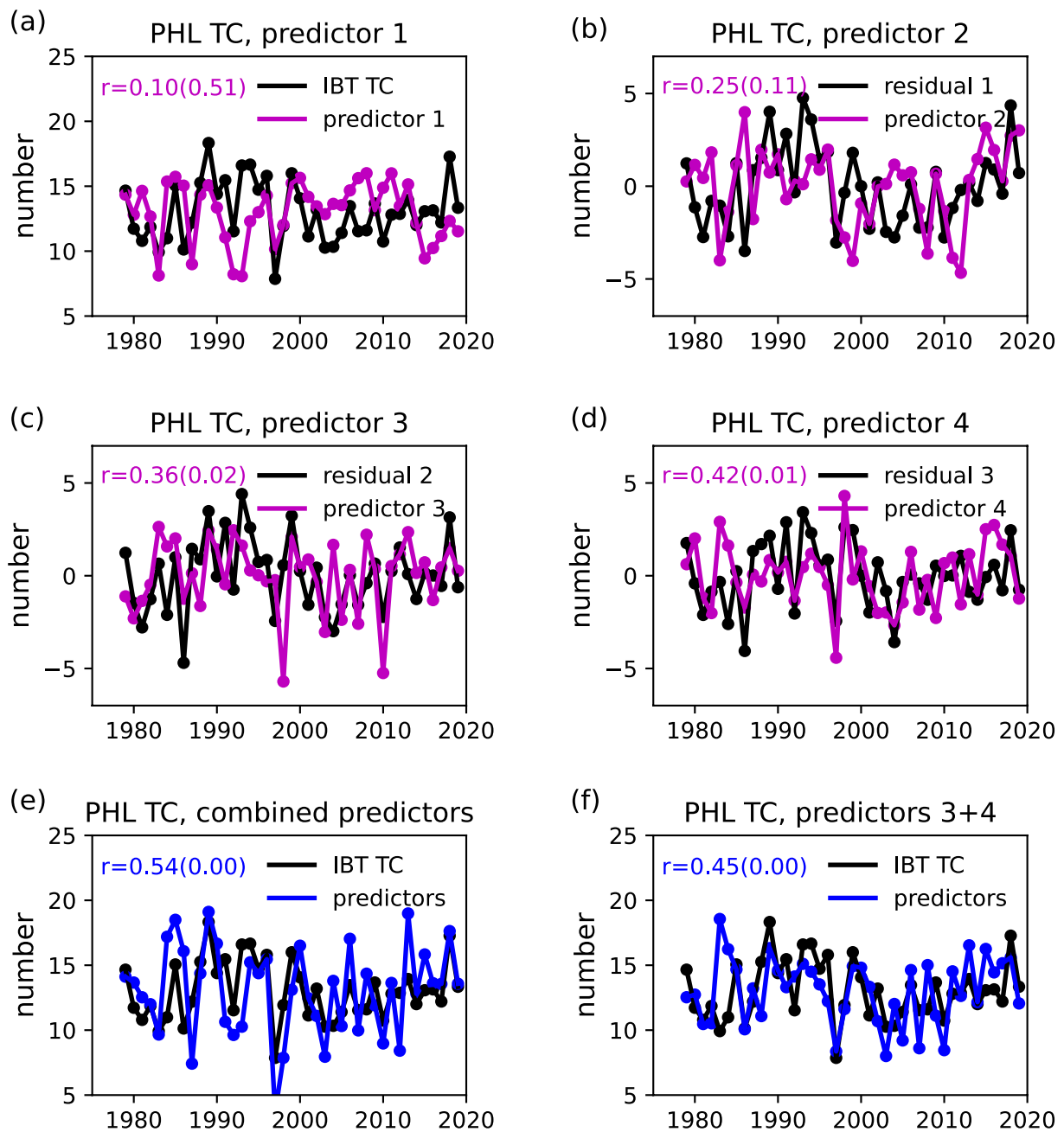
Figure 11 shows the cross-validated prediction of the SCS TC frequency using the statistically derived predictors for the period 1979–2019. For the SCS, the TC frequency predicted by predictors 1–3 is significantly correlated with the observed frequency (i.e., total frequency, frequency residual 1 and frequency residual 2), with  $r = 0.37$ ,



**Figure 11.** Cross-validated prediction of South China Sea (SCS) tropical cyclones (TC) frequency in JAS predicted by individual (a–c) and combined (d) predictors, over 1979–2019. Black lines are the observed TC frequency (or frequency residuals in b, c), magenta lines are the TC frequency (or frequency residuals in b, c) predicted by individual predictors, and blue line is the total TC frequency predicted by predictors. Correlation coefficients ( $r$  values) between the predicted and observed TC frequency (or frequency residuals in b, c) are provided in the top-left corner.

0.48, and 0.56, respectively. In the cross-validation, the predicted total frequency using three predictors agrees well the observed total frequency, with  $r = 0.73$ .

The cross-validated prediction for the PHL TCs is shown in Figure 12. The frequency predicted by predictors 1 and 2 is not significantly correlated with the observed frequency, with  $r = 0.10$  ( $p = 0.51$ ) and  $0.25$  ( $p = 0.11$ ), respectively. This contrasts with the significant TC-predictor correlations ( $r = 0.37$  and  $0.41$ ; Figure S2 in Supporting Information S1). The difference means that the two predictors, which resemble the ENSO and PMM signals, respectively, overfit the relationship with TCs. The disparity could be related to several reasons, including (but not limited to) the nonstationarity of the TC relationship with ENSO and PMM. This is especially the case for predictor 1, where ENSO could be dependent on, or modulated by, other climate modes on decadal to multidecadal timescales, leading to an inconsistent TC-ENSO teleconnection in different periods (e.g., Newman et al., 2016; Yeh et al., 2018; Wang et al., 2020; Zhao & Wang, 2016, 2019). In the cross-validation, the PHL TC predictor 1 has a better performance over 1979–1999 ( $r = 0.34$ ,  $p = 0.13$ ) than over 2000–2019 ( $r = 0.11$ ,  $p = 0.64$ ), confirming that the performance is not stationary with time. The prediction determined by predictors 3 and 4 performs well for the corresponding frequency, with  $r = 0.36$  and  $0.42$  (both significant at the 95% confidence level), respectively. The predicted total frequency of PHL TCs using the four predictors has a significant skill, with  $r = 0.54$ . The removal of predictors 1 and 2 does not have a substantial effect on the total frequency prediction (Figure 12f), with  $r = 0.45$ , suggesting the prediction source is mainly from predictors 3 and 4.



**Figure 12.** (a–e) As Figure 11, but for the Philippines (PHL) tropical cyclone (TC) frequency. (f) as (e), but for cross-validated prediction of PHL TC frequency in JAS predicted by combined predictors 3 and 4 (i.e., without predictors 1 and 2).

We also evaluated the sensitivity of the cross-validated prediction to different lengths of the omitted block (Table S2 in Supporting Information S1). When the block length increases from 1 to 7 years, the predictor performance varies marginally within 0.71–0.73 for SCS TCs and within 0.50–0.59 for PHL TCs. Thus, the performance of predictors does not change substantially when the length of the omitted block changes in the cross-validated prediction.

### 3.4. Hybrid Prediction Performance

We further develop the statistical prediction method into a hybrid statistical-dynamical approach, based on the large-scale environment produced by the dynamical seasonal hindcasts from SEAS5, GloSea5, and GloSea6. This

**Table 1**

Performance (Anomaly Correlation Coefficient) of Different Prediction Methods for SEA TC Frequency in the Peak-Season (JAS) Over 1993–2015

	Empirical prediction <sup>a</sup>	Hybrid prediction (April, March, February forecasts)			Dynamical prediction (June, May forecasts)		
		SEAS5	GloSea5	GloSea6 <sup>b</sup>	SEAS5	GloSea5	GloSea6
SCS	<b>0.80</b>	<b>0.77, 0.46, 0.38</b>	<b>0.70, 0.42, 0.32</b>	<b>0.70, 0.48</b>	−0.31, 0.04	<b>0.48, 0.04</b>	0.08, 0.03
PHL	<b>0.57</b>	<b>0.51, 0.28, 0.31</b>	<b>0.56, 0.47, 0.46</b>	<b>0.56, 0.43</b>	0.05, 0.09	0.28, 0.12	0.15, 0.26

Note. Bold numbers are significant at the 95% confidence level.

<sup>a</sup>Cross-validation. <sup>b</sup>February forecasts of GloSea6 are not available.

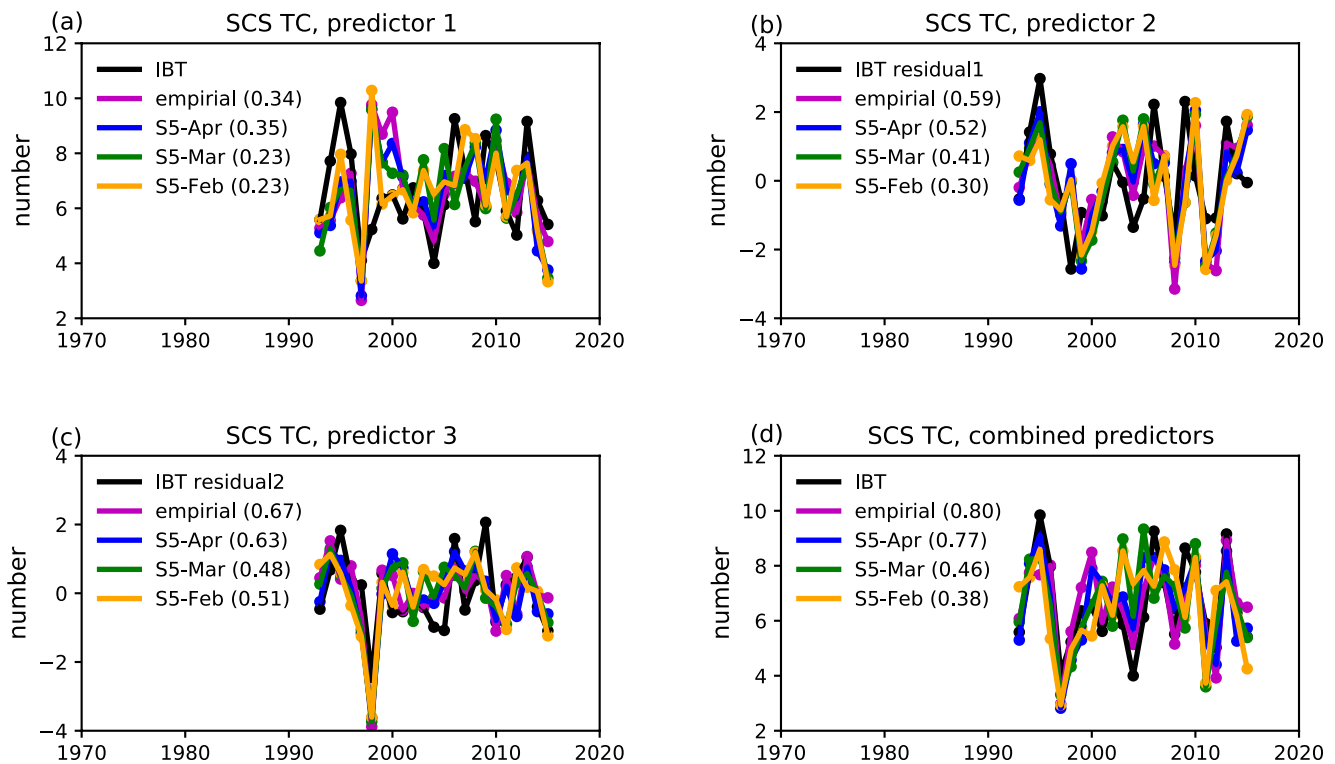
means that the hybrid predictors are “predicted” by the dynamical hindcasts. If the dynamical models perfectly predict the predictor-associated environmental conditions, the performance of the hybrid predictors will be the same as the anomaly correlation between the ERA5-based predictors and observed TC frequency, that is, the best potential performance. The hybrid prediction is directly verified against observed TC frequency, rather than through the cross-validation as in the statistical prediction. In the hybrid prediction, we use the ensemble mean of seasonal hindcasts to extract predictors. The hindcast period is 1993–2015, which is the available period of the C3S seasonal hindcasts. To evaluate the effect of lead time, we use the hindcasts initialized in April, March and February (i.e., 3–5-month lead time).

We first reidentified the statistical predictors for 1993–2015 using ERA5, as different periods may have different predictors (discussed above). The disparity between the two periods is expected to be due to the nonstationarity of the environment-TC relationship. For 1993–2015, we also identified three statistical predictors for the SCS TC frequency using the ERA5 environment (Figure S7 in Supporting Information S1). The 1993–2015 predictors are built from the same environmental variables as the 1979–2019 predictors but averaged over slightly different areas. Contrasting with the four predictors for 1979–2019, three predictors of PHL TCs were identified for the shorter period (Figure S8 in Supporting Information S1). The hybrid predictors for 1993–2015 were estimated by averaging the environmental variables in dynamical hindcasts over these “predefined” areas.

Table 1 provides the skill of the statistical prediction, and the skill of the hybrid prediction based on SEAS5, GloSea5, and GloSea6 hindcasts, for the total TC frequency. For the SCS TC frequency, the statistical prediction ( $r = 0.80$ ) performs slightly better than the hybrid prediction ( $r = 0.70$ – $0.77$ ) that uses the April hindcasts (3-month lead time). The hybrid prediction performs similarly when different dynamical models are used. When the March and February hindcasts (4–5-month lead time) are used, the skill of the hybrid prediction decreases to  $r = 0.42$ – $0.48$  (significant at the 95% confidence level) and  $r = 0.32$ – $0.38$  (not significant), respectively. For the PHL TCs, the hybrid prediction with the shortest lead time (the April hindcasts) performs best, with  $r = 0.51$ – $0.56$ . This performance is similar to the statistical prediction ( $r = 0.57$ ). As expected, when lead time increases, the hybrid prediction performance decreases. For the PHL TCs, the GloSea5-based hybrid prediction appears to perform better than that from the other two models. This is especially the case for 4–5-month lead time, for which GloSea5-based predictors have a significant skill while the SEAS5-based and GloSea6-based predictors have either no significant skill or a lower skill.

The performance of individual predictors in the hybrid prediction is also analyzed. Figures 13 and 14 demonstrate the predicted variability of TC frequency using SEAS5-based predictors. For the SCS TCs, three individual predictors with the shortest lead time (3-month lead time, the April hindcasts) have the best performance ( $r = 0.35, 0.52, \text{ and } 0.63$ ), which is close to the empirical prediction ( $r = 0.34, 0.59, \text{ and } 0.67$ ) (Figure 13). When lead time increases to 5 months, the skill of individual predictors becomes insignificant at 95% confidence level, except predictor 3 ( $r = 0.51$ ). We further find that all these predictors are predicted well in dynamical hindcasts, with  $r = 0.68$ – $0.95$  for 3–5-month lead times (Figure S9 in Supporting Information S1). It is worth noticing the hybrid predictor 1, which has the lowest skill among all predictors (Figure 13). This is due to the weak relationship between the predictor and TCs over this period, rather than the low predictability of the predictor.

For the PHL TCs, the hybrid predictor 1 has no skill ( $r < 0.20$ ), regardless of lead time (Figure 14). The lack of skill is related to the overfitting of predictor 1, which has no skill either in the cross-validated statistical prediction ( $r = -0.14$ ). Although the statistical predictor 2 performs well ( $r = 0.48$ ), the hybrid predictor 2 has no significant



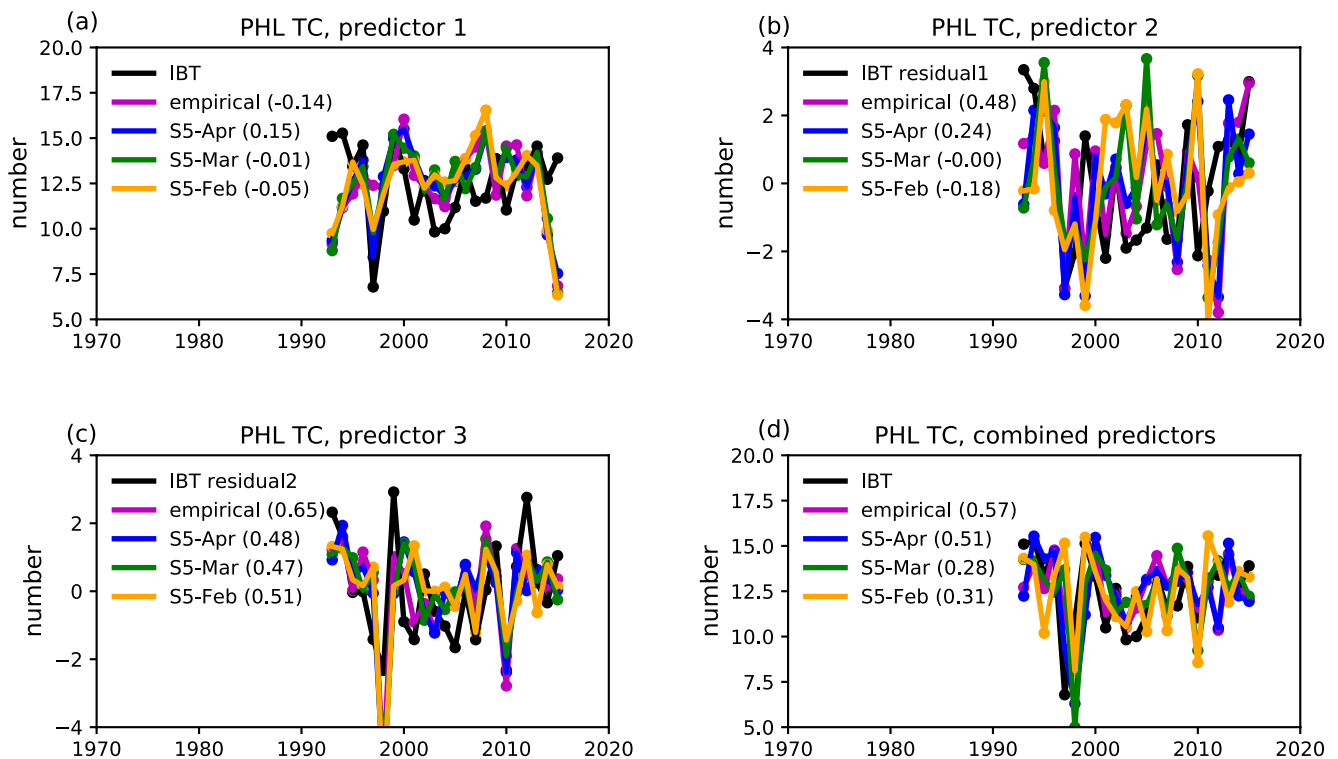
**Figure 13.** Hybrid prediction of South China Sea (SCS) tropical cyclone (TC) frequency in July–September (JAS) using individual (a–c) and combined (d) predictors derived from SEAS5 (S5) seasonal forecasts, over 1993–2015. Black lines are the observed TC frequency (or frequency residuals in b, c). Blue, green, and orange lines indicate the hybrid prediction with predictors extracted from the April, March, and February forecasts, while magenta line shows the cross-validated prediction with the statistical predictors extracted from ECMWF fifth generation climate reanalysis (ERA5). Values in the key are correlation coefficients between the predicted and observed TC frequency (or frequency residuals in b, c).

skill ( $r < 0.25$ ). This inconsistent performance is related to the low predictability of predictor 2 in dynamical hindcasts (Figure S10 in Supporting Information S1). Thus, the low performance of the hybrid predictors 1 and 2 is likely related to the nonstationarity of the TC–predictor relationship and the low predictability of the environment associated with the predictors. Predictor 3 performs well in both statistical and hybrid predictions ( $r = 0.47$ – $0.51$ ). Such good performance of hybrid prediction is related to the high predictability of predictor 3 in models. The same can be applied to GloSea5- and GloSea6-based predictors, and the results remain similar (not shown).

Figure 15 shows the skill of dynamical hindcasts in predicting the pre-season SSTs and low-level zonal winds. In the April-initialized hindcasts, all three models accurately predict the SST anomalies in the central and eastern Pacific and in the Indian Ocean ( $r > 0.7$ ), with a slightly better prediction in SEAS5. This is consistent with the generally high predictability of the predictors (Figures S9 and S10 in Supporting Information S1), which are extracted from the environment over these two basins. It is interesting to recall the relatively low predictability of the PHL TC predictor 2. Although the SST anomalies in the north-eastern Pacific are significantly predicted in hindcasts, the residuals of SSTs over this region after removing the effect of other predictors could be less accurately predicted. Additionally, the model performance for SSTs over this region degrades greatly when lead time increases (not shown). Thus, the hybrid prediction relies on the robustness/stationarity of the statistical predictor–TC relationship, and the capability of dynamical models to predict the tropical SSTs and low-level winds that form the hybrid predictors.

#### 4. Conclusions and Discussion

Accurately predicting the SEA seasonal TC activity several months ahead can provide an early warning service that may significantly reduce TC-related impacts. The ENSO index tends to have an overall weak correlation with the peak-season SEA TCs, related to the complexity (or nonlinearity) of the ENSO teleconnections and the

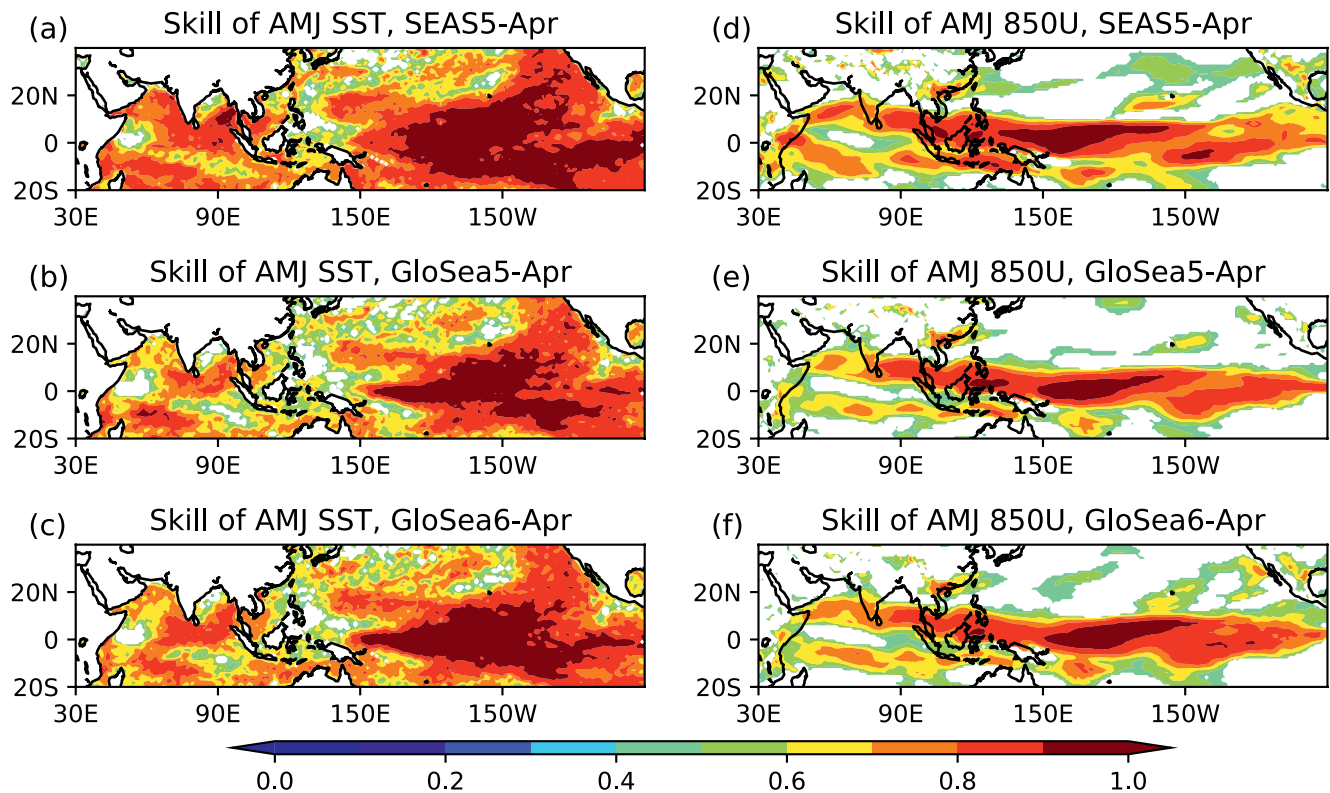


**Figure 14.** As Figure 13, but for the Philippines (PHL) tropical cyclone (TC) frequency.

possible competing or compound effects with other climate variability. It is challenging for the approaches that are solely based on the ENSO indices to predict the peak-season SEA TCs. In this paper, we introduced a novel approach (SOIP) to predict the occurrence frequency of the peak-season TCs in two subregions of SEA: SCS and PHL. Independent predictors, which have proven to have cross-season effects on different clusters of the regional TCs, are sequentially constructed from the preseason (AMJ) large-scale environment. The predictors are drawn from the SSTs and low-level zonal winds in the tropics, which represent the atmospheric and oceanic conditions that are essential for the year-to-year variability in TC generation and track. The SOIP approach is conceptually different from the traditional statistical methods, which rely on the known relationships (which are mostly linear) between climate variability (e.g., ENSO) and TC activity. Instead, the SOIP approach constructs the predictors that are tailored for the regional TCs. We showed that the predictors can well represent the interannual variability of SEA TCs that form either locally or remotely. In the 41-year period (1979–2019), the statistical approach based on ERA5 provides a skillful prediction for the observed frequency of SCS and PHL TCs, with  $r = 0.73$  and  $0.54$ , respectively. The relatively low performance for PHL TCs could be related to the nonstationarity of the predictors which resemble the ENSO and PMM signals.

We further developed the statistical approach into a hybrid prediction system. In the hybrid prediction, the predictors, which are predefined in ERA5, are predicted by three dynamical seasonal hindcasts (SEAS5, GloSea5, and GloSea6). The skill of the hybrid prediction is comparable to the statistical prediction when the April-initialized hindcasts are used ( $r = 0.7$ – $0.8$  and  $0.5$ – $0.6$  for the SCS and PHL TC frequency, respectively). When using the March-initialized hindcasts, the skill of the hybrid prediction is reduced but still mostly significant ( $r = 0.4$ – $0.5$  and  $0.3$ – $0.5$  for the SCS and PHL TCs, respectively; only not significant when using SEAS5 March hindcasts). The merit of the hybrid prediction relative to the statistical prediction is the longer lead times (4–5 months ahead). We further confirmed that the main sources of the hybrid prediction skill are from the robustness of the observed predictor-TC relationship, and the capability of dynamical models to predict the tropical SSTs and low-level winds.

The performance of seasonal hindcasts (SEAS5, GloSea5, and GloSea6) in directly predicting the SCS and PHL TCs is provided in Table 1. Over the period 1993–2015, only the GloSea5 June hindcasts show a significant skill in predicting the frequency of SCS TCs ( $r = 0.48$ ), while none of the three models predict the PHL TC frequency.



**Figure 15.** (a–c) Correlation coefficients of the preseason (AMJ) SSTs between ECMWF fifth generation climate reanalysis (ERA5) and ECMWF fifth generation seasonal forecast system (SEAS5), UK Met Office Global Seasonal Forecasting System version 5 and 6 (GloSea5 and GloSea6) April forecast ensemble mean, over 1993–2015. (d–f) as (a–c), but for correlations of 850-hPa zonal winds. Only correlations that are significant at the 95% confidence level are shown.

The new developed statistical and hybrid predictions in this paper outperform the dynamical predictions, indicating the potential for operational use. The low performance of the dynamical models for SEA TCs is consistent with previous studies (Befort et al., 2021; Feng, Klingaman, Hodges, & Guo, 2020). The biased teleconnections of climate variability, and the biased TC precursor systems in the tropics such as westward propagating tropical depressions, are among the possible reasons. Further detailed model evaluation is required to understand the model performance, which is beyond the scope of this article.

The SOIP approach in our paper shares a similar decomposing concept to the empirical orthogonal function (EOF) analysis, which is used to identify the basin-wide patterns of TC activity (e.g., Kim et al., 2005; Liu et al., 2021). Both can decompose the orthogonal variability of TCs. However, the SOIP is initially confined to the temporal variability that must be principally related to the environment, whereas the EOF analysis recognizes the geographical variability of TCs and during the analysis it does not consider the environmental conditions. We did the EOF analysis of the peak-season TC genesis and track density (both the SCS and PHL TCs). However, no robust relationships were found between the leading principal components of EOFs and the pre-season environment. Although the SOIP approach is initially developed for the frequency of peak-season SEA TCs, it can be applied to other seasons, other regions and other TC metrics (e.g., intensity). In the future, it will be worth evaluating the strength and weaknesses of the SOIP approach in those applications against other statistical and dynamical methods, to form a synthesized seasonal forecast conditional on the performance of prediction methods.

### Data Availability Statement

SEAS5, GloSea5, and GloSea6 reforecast data are archived from [Seasonal Forecast Subdaily Data on Pressure Levels](https://cds.clm.copernicus.org/) through the Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (doi:10.24381/cds.50ed0a73). ERA5 is generated by ECMWF and distributed by the C3S CDS (Hersbach et al., 2019; doi:10.24381/cds.6860a573). The climate indices (ENSO and PMM) are obtained from [https://psl.noaa.gov/gcos\\_wgsp/Timeseries/Nino34/](https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/)

(Rayner et al., 2003) and <https://psl.noaa.gov/data/timeseries/monthly/PMML/> (Chiang & Vimont, 2004) from the NOAA Physical Sciences Laboratory. IBTrACS TC data (West Pacific) are archived from <https://www.ncei.noaa.gov/products/international-best-track-archive?name=ib-v4-access> (Knapp et al., 2018; doi:10.25921/82ty-9e16). Computing and data storage facilities were provided by JASMIN (<https://jasmin.ac.uk>).

#### Acknowledgments

Xiangbo Feng, Kevin Hodges, and Gui-Ying Yang were supported by the UK Met Office Weather and Climate Science for Service Partnership for Southeast Asia, as part of the Newton Fund. Gui-Ying Yang was also supported by the National Centre for Atmospheric Science ODA national capability programme ACREW (NE/R000034/1), which is supported by NERC and the GCRF.

#### References

- Ashok, K., Behera, S. K., Rao, S. A., Weng, H., & Yamagata, T. (2007). El Niño Modoki and its possible teleconnection. *Journal of Geophysical Research*, *112*, C11007. <https://doi.org/10.1029/2006JC003798>
- Basconcillo, J., Cha, E. J., & Moon, I. J. (2021). Characterizing the highest tropical cyclone frequency in the Western North Pacific since 1984. *Scientific Reports*, *11*(1), 1–11. <https://doi.org/10.1038/s41598-021-93824-2>
- Befort, D. J., Hodges, K., & Weisheimer, A. (2021). Seasonal prediction of tropical cyclones over the north Atlantic and Western north Pacific: Dependence on model, resolution and stochastic physics. *Journal of Climate*, *35*(5), 1385–1397.
- Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., et al. (2014). Recent development of the met office operational ocean forecasting system: An overview and assessment of the new global FOAM forecasts. *Geoscientific Model Development*, *7*(6), 2613–2638. <https://doi.org/10.5194/gmd-7-2613-2014>
- Bowler, N. E., Arribas, A., Beare, S. E., Mylne, K. R., & Shutts, G. J. (2009). The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, *135*(640), 767–776. <https://doi.org/10.1002/qj.394>
- Camargo, S. J., Robertson, A. W., Gaffney, S. J., Smyth, P., & Ghil, M. (2007). Cluster analysis of typhoon tracks. Part I: General properties. *Journal of Climate*, *20*(14), 3635–3653. <https://doi.org/10.1175/JCLI4188.1>
- Camargo, S. J., & Sobel, A. H. (2005). Western North Pacific tropical cyclone intensity and ENSO. *Journal of Climate*, *18*, 2996–3006. <https://doi.org/10.1175/JCLI3457.1>
- Camp, J., Roberts, M., MacLachlan, C., Wallace, E., Hermanson, L., Brookshaw, A., et al. (2015). Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, *141*(691), 2206–2219. <https://doi.org/10.1002/qj.2516>
- Camp, J., Roberts, M. J., Comer, R. E., Wu, P., MacLachlan, C., Bett, P. E., et al. (2019). The Western Pacific subtropical high and tropical cyclone landfall: Seasonal forecasts using the Met Office GloSea5 system. *Quarterly Journal of the Royal Meteorological Society*, *145*(718), 105–116. <https://doi.org/10.1002/qj.3407>
- Cane, M. A., Zebiak, S. E., & Dolan, S. C. (1986). Experimental forecasts of EL Niño. *Nature*, *321*(6073), 827–832. <https://doi.org/10.1038/321827a0>
- Chan, J. C. (1985). Tropical cyclone activity in the northwest Pacific in relation to the El Niño/Southern Oscillation phenomenon. *Monthly Weather Review*, *113*(4), 599–606. [https://doi.org/10.1175/1520-0493\(1985\)113<0599:TCAITN>2.0.CO;2](https://doi.org/10.1175/1520-0493(1985)113<0599:TCAITN>2.0.CO;2)
- Chen, J. H., & Lin, S. J. (2013). Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. *Journal of Climate*, *26*(2), 380–398. <https://doi.org/10.1175/JCLI-D-12-00061.1>
- Chen, J. M., Pei-Hua, T., Wu, L., Jin-Shuen, L., & Hui-Shan, C. (2017). Climatological analysis of passage-type tropical cyclones from the Western North Pacific into the South China Sea. *TAO. Terrestrial, Atmospheric and Oceanic Sciences*, *28*(3), 3. <https://doi.org/10.3319/tao.2016.10.04.02>
- Chen, J. M., Wu, C. H., Gao, J., Chung, P. H., & Sui, C. H. (2019). Migratory tropical cyclones in the South China Sea modulated by intraseasonal oscillations and climatological circulations. *Journal of Climate*, *32*(19), 6445–6466. <https://doi.org/10.1175/JCLI-D-18-0824.1>
- Chen, T. C., Wang, S. Y., & Yen, M. C. (2006). Interannual variation of the tropical cyclone activity over the Western North Pacific. *Journal of Climate*, *19*(21), 5709–5720. <https://doi.org/10.1175/JCLI3934.1>
- Chen, T. C., Wang, S. Y., Yen, M. C., & Gallus, W. A., Jr. (2004). Role of the monsoon gyre in the interannual variation of tropical cyclone formation over the Western North Pacific. *Weather and Forecasting*, *19*(4), 776–785. [https://doi.org/10.1175/1520-0434\(2004\)019<0776:ROTMGI>2.0.CO;2](https://doi.org/10.1175/1520-0434(2004)019<0776:ROTMGI>2.0.CO;2)
- Chen, T. C., Weng, S. P., Yamazaki, N., & Kiehne, S. (1998). Interannual variation in the tropical cyclone formation over the Western North Pacific. *Monthly Weather Review*, *126*(4), 1080–1090. [https://doi.org/10.1175/1520-0493\(1998\)126<1080:IVITTC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<1080:IVITTC>2.0.CO;2)
- Cheung, K. K. (2004). Large-scale environmental parameters associated with tropical cyclone formations in the Western North Pacific. *Journal of Climate*, *17*(3), 466–484. [https://doi.org/10.1175/1520-0442\(2004\)017<0466:LEPAWT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0466:LEPAWT>2.0.CO;2)
- Chia, H. H., & Ropelewski, C. F. (2002). The interannual variability in the Genesis location of tropical cyclones in the northwest Pacific. *Journal of Climate*, *15*(20), 2934–2944. [https://doi.org/10.1175/1520-0442\(2002\)015<2934:TIVITG>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2934:TIVITG>2.0.CO;2)
- Chiang, J. C. H., & Vimont, D. J. (2004). Analogous meridional modes of atmosphere–ocean variability in the tropical Pacific and tropical Atlantic. [Data Set]. *Journal of Climate*, *17*(21), 4143–4158. <https://doi.org/10.1175/jcli4953.1>
- Corporal-Lodangco, I. L., & Leslie, L. M. (2016). Cluster analysis of Philippine tropical cyclone climatology: Applications to forecasting. *J Climatol Weather Forecasting*, *4*(152), 2. <https://doi.org/10.4172/2332-2594.1000152>
- Corporal-Lodangco, I. L., Leslie, L. M., & Lamb, P. J. (2016). Impacts of ENSO on Philippine tropical cyclone activity. *Journal of Climate*, *29*(5), 1877–1897. <https://doi.org/10.1175/JCLI-D-14-00723.1>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, *137*(656), 553–597. <https://doi.org/10.1002/qj.828>
- Duy, D. B., Thanh, N. D., Duc, T. Q., & VanTan, P. (2019). Seasonal predictions of the number of tropical cyclones in the Vietnam east sea using statistical models. *VNU Journal of Science: Earth and Environmental Sciences*, *35*(2), 10.25073/2588-1094/vnuees.4379
- Fan, K., & Wang, H. (2009). A new approach to forecasting typhoon frequency over the Western North Pacific. *Weather and Forecasting*, *24*(4), 974–986. <https://doi.org/10.1175/2009WAF2222194.1>
- Feng, X., Klingaman, N. P., & Hodges, K. I. (2019). The effect of atmosphere–ocean coupling on the prediction of 2016 Western North Pacific tropical cyclones. *Quarterly Journal of the Royal Meteorological Society*, *145*(723), 2425–2444. <https://doi.org/10.1002/qj.3571>
- Feng, X., Klingaman, N. P., & Hodges, K. I. (2021). Poleward migration of Western North Pacific tropical cyclones related to changes in cyclone seasonality. *Nature Communications*, *12*, 6210. <https://doi.org/10.1038/s41467-021-26369-7>
- Feng, X., Klingaman, N. P., Hodges, K. I., & 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), 10489–10504. <https://doi.org/10.1175/JCLI-D-20-0255.1>



- Feng, X., Klingaman, N. P., Zhang, S., & Guo, L. (2020). Building sustainable science partnerships between early-career researchers to better understand and predict East Asia water cycle extremes. *Bulletin of the American Meteorological Society*, *101*(6), E785–E789. <https://doi.org/10.1175/BAMS-D-20-0039.1>
- Gao, S., Zhu, L., Zhang, W., & Chen, Z. (2018). Strong modulation of the Pacific meridional mode on the occurrence of intense tropical cyclones over the Western North Pacific. *Journal of Climate*, *31*(19), 7739–7749. <https://doi.org/10.1175/JCLI-D-17-0833.1>
- Goh, A. Z. H., & Chan, J. C. (2010). An improved statistical scheme for the prediction of tropical cyclones making landfall in South China. *Weather and Forecasting*, *25*(2), 587–593.
- Guo, Y. P., Feng, X., Klingaman, N. P., & Tan, Z. M. (2020). Impact of Indo-Pacific warm pool Hadley circulation on the seasonal forecast performance for summer precipitation over the Western North Pacific. *Environmental Research Letters*, *15*(10), 104041. <https://doi.org/10.1088/1748-9326/aba97c>
- Guo, Y. P., & Tan, Z. M. (2018). Impacts of the boreal spring Indo-Pacific warm pool Hadley circulation on tropical cyclone activity over the Western North Pacific. *Journal of Climate*, *31*(4), 1361–1375. <https://doi.org/10.1175/JCLI-D-17-0422.1>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2019). ERA5 monthly averaged data on pressure levels from 1979 to present. [Dataset]. Chang. <https://doi.org/10.24381/cds.6860a573>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049.
- Hodges, K., Cobb, A., & Vidale, P. L. (2017). How well are tropical cyclones represented in reanalysis datasets? *Journal of Climate*, *30*(14), 5243–5264. <https://doi.org/10.1175/JCLI-D-16-0557.1>
- Hoskins, B. J., Yang, G. Y., & Fonseca, R. M. (2020). The detailed dynamics of the June–August Hadley cell. *Quarterly Journal of the Royal Meteorological Society*, *146*(727), 557–575. <https://doi.org/10.1002/qj.3702>
- Hsu, P. C., Chu, P. S., Murakami, H., & Zhao, X. (2014). An abrupt decrease in the late-season typhoon activity over the Western North Pacific. *Journal of Climate*, *27*(11), 4296–4312. <https://doi.org/10.1175/JCLI-D-13-00417.1>
- Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117. <https://doi.org/10.5194/gmd-12-1087-2019>
- Kim, H. K., & Seo, K. H. (2016). Cluster analysis of tropical cyclone tracks over the Western North Pacific using a self-organizing map. *Journal of Climate*, *29*(10), 3731–3751. <https://doi.org/10.1175/JCLI-D-15-0380.1>
- Kim, H. M., Webster, P. J., & Curry, J. A. (2011). Modulation of North Pacific tropical cyclone activity by three phases of ENSO. *Journal of Climate*, *24*(6), 1839–1849. <https://doi.org/10.1175/2010JCLI3939.1>
- Kim, J. H., Ho, C. H., & Chu, P. S. (2010). Dipolar redistribution of summertime tropical cyclone Genesis between the Philippine Sea and the northern South China Sea and its possible mechanisms. *Journal of Geophysical Research*, *115*, D06104. <https://doi.org/10.1029/2009JD012196>
- Kim, J. H., Ho, C. H., Sui, C. H., & Park, S. K. (2005). Dipole structure of interannual variations in summertime tropical cyclone activity over East Asia. *Journal of Climate*, *18*(24), 5344–5356. <https://doi.org/10.1175/JCLI3601.1>
- Kim, O. Y., Kim, H. M., Lee, M. I., & Min, Y. M. (2017). Dynamical-statistical seasonal prediction for Western North Pacific typhoons based on APCC multi-models. *Climate Dynamics*, *48*(1–2), 71–88. <https://doi.org/10.1007/s00382-016-3063-1>
- Knaff, J. A., Brown, D. P., Courtney, J., Gallina, G. M., & Beven, J. L. (2010). An evaluation of Dvorak technique-based tropical cyclone intensity estimates. *Weather and Forecasting*, *25*(5), 1362–1379. <https://doi.org/10.1175/2010WAF2222375.1>
- Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., & Schreck, C. J. (2018). International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4 (West Pacific). [Dataset]. NOAA National Centers for Environmental Information. <https://doi.org/10.25921/82ty-9e16>
- Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., & Neumann, C. J. (2010). The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, *91*(3), 363–376. <https://doi.org/10.1175/2009bams2755.1>
- Li, C., Lu, R., & Chen, G. (2017). Promising prediction of the monsoon trough and its implication for tropical cyclone activity over the Western North Pacific. *Environmental Research Letters*, *12*(7), 074027.
- Liu, K. S., & Chan, J. C. (2008). Interdecadal variability of Western North Pacific tropical cyclone tracks. *Journal of Climate*, *21*(17), 4464–4476. <https://doi.org/10.1175/2008JCLI2207.1>
- Liu, K. S., Chan, J. C., & Kubota, H. (2021). Meridional oscillation of tropical cyclone activity in the Western North Pacific during the past 110 years. *Climatic Change*, *164*(1), 1–22. <https://doi.org/10.1007/s10584-021-02983-8>
- Lyon, B., & Camargo, S. J. (2009). The seasonally-varying influence of ENSO on rainfall and tropical cyclone activity in the Philippines. *Climate Dynamics*, *32*(1), 125–141. <https://doi.org/10.1007/s00382-008-0380-z>
- MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., et al. (2015). Global seasonal forecast system version 5 (GloSea5): A high-resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1072–1084. <https://doi.org/10.1002/qj.2396>
- Manganello, J. V., Hodges, K. I., Cash, B. A., Kinter, J. L., III, Altshuler, E. L., Fennessy, M. J., et al. (2016). Seasonal forecasts of tropical cyclone activity in a high-atmospheric-resolution coupled prediction system. *Journal of Climate*, *29*(3), 1179–1200. <https://doi.org/10.1175/JCLI-D-15-0531.1>
- Mei, W., Xie, S. P., Zhao, M., & Wang, Y. (2015). Forced and internal variability of tropical cyclone track density in the Western North Pacific. *Journal of Climate*, *28*(1), 143–167. <https://doi.org/10.1175/JCLI-D-14-00164.1>
- Molinari, J., & Vollaro, D. (2013). What percentage of Western North Pacific tropical cyclones form within the monsoon trough? *Monthly Weather Review*, *141*(2), 499–505. <https://doi.org/10.1175/mwr-d-12-00165.1>
- Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., et al. (2016). The Pacific decadal oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427. <https://doi.org/10.1175/JCLI-D-15-0508.1>
- Patricola, C. M., Camargo, S. J., Klotzbach, P. J., Saravanan, R., & Chang, P. (2018). The influence of ENSO flavors on Western North Pacific tropical cyclone activity. *Journal of Climate*, *31*, 5395–5416. <https://doi.org/10.1175/JCLI-D-17-0678.1>
- Qian, Y., Hsu, P. C., Murakami, H., Xiang, B., & You, L. (2020). A hybrid dynamical-statistical model for advancing subseasonal tropical cyclone prediction over the Western north Pacific. *Geophysical Research Letters*, *47*, e2020GL090095. <https://doi.org/10.1029/2020GL090095>
- Rayner, N. A. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. [Data Set]. *Journal of Geophysical Research*, *108*(D14), 4407. <https://doi.org/10.1029/2002JD002670>
- Seasonal forecast subdaily data on pressure levels. Seasonal forecast subdaily data on pressure levels. [Dataset]. Chang. <https://doi.org/10.24381/cds.50ed0a73>

- Teng, H. F., Done, J. M., & Kuo, Y. H. (2021). Landfalling tropical cyclone characteristics and their multi-timescale variability connected to monsoon and easterly formation environments over the Western North Pacific. *Quarterly Journal of the Royal Meteorological Society*, *147*(738), 2953–2977. <https://doi.org/10.1002/qj.4108>
- Tian, B., & Fan, K. (2019). Seasonal climate prediction models for the number of landfalling tropical cyclones in China. *Journal of Meteorological Research*, *33*(5), 837–850. <https://doi.org/10.1007/s13351-019-8187-x>
- Torn, R. D., & Snyder, C. (2012). Uncertainty of tropical cyclone best-track information. *Weather and Forecasting*, *27*(3), 715–729. <https://doi.org/10.1175/waf-d-11-00085.1>
- Vecchi, G. A., Delworth, T., Gudgel, R., Kapnick, S., Rosati, A., Wittenberg, A. T., et al. (2014). On the seasonal forecasting of regional tropical cyclone activity. *Journal of Climate*, *27*(21), 7994–8016. <https://doi.org/10.1175/JCLI-D-14-00158.1>
- Velden, C., Harper, B., Wells, F., Beven, J. L., Zehr, R., Olander, T., et al. (2006). The Dvorak tropical cyclone intensity estimation technique: A satellite-based method that has endured for over 30 years. *Bulletin of the American Meteorological Society*, *87*(9), 1195–1210. <https://doi.org/10.1175/BAMS-87-9-1195>
- Vitart, F., & Stockdale, T. N. (2001). Seasonal forecasting of tropical storms using coupled GCM integrations. *Monthly Weather Review*, *129*(10), 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)
- Wang, B., & Chan, J. C. (2002). How strong ENSO events affect tropical storm activity over the Western North Pacific. *Journal of Climate*, *15*(13), 1643–1658. [https://doi.org/10.1175/1520-0442\(2002\)015<1643:HSEAT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1643:HSEAT>2.0.CO;2)
- Wang, B., Liu, J., Kim, H. J., Webster, P. J., Yim, S. Y., & Xiang, B. (2013). Northern Hemisphere summer monsoon intensified by mega-El Niño/southern oscillation and Atlantic multidecadal oscillation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(14), 5347–5352. <https://doi.org/10.1073/pnas.1219405110>
- Wang, B., Luo, X., & Liu, J. (2020). How robust is the Asian precipitation-ENSO relationship during the Industrial warming period (1901–2017)? *Journal of Climate*, *33*(7), 2779–2792. <https://doi.org/10.1175/JCLI-D-19-0630.1>
- Wang, B., Xiang, B., & Lee, J. Y. (2013). Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(8), 2718–2722. <https://doi.org/10.1073/pnas.1214626110>
- Wang, C., Li, C., Mu, M., & Duan, W. (2013). Seasonal modulations of different impacts of two types of ENSO events on tropical cyclone activity in the Western North Pacific. *Climate Dynamics*, *40*(11–12), 2887–2902. <https://doi.org/10.1007/s00382-012-1434-9>
- Wang, C., & Wang, B. (2019). Tropical cyclone predictability shaped by Western Pacific subtropical high: Integration of trans-basin Sea Surface temperature effects. *Climate Dynamics*, *53*(5), 2697–2714. <https://doi.org/10.1007/s00382-019-04651-1>
- Wang, C., Wang, B., & Wu, L. (2019). A region-dependent seasonal forecasting framework for tropical cyclone Genesis frequency in the Western North Pacific. *Journal of Climate*, *32*(23), 8415–8435. <https://doi.org/10.1175/JCLI-D-19-0006.1>
- Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., et al. (2018). The Met Office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357–380. <https://doi.org/10.1002/2017MS001115>
- Williams, K. D., Harris, C. M., Bodas-Salcedo, A., Camp, J., Comer, R. E., Copsey, D., et al. (2015). The met office global coupled model 2.0 (GC2) configuration. *Geoscientific Model Development*, *8*(5), 1509–1524. <https://doi.org/10.5194/gmd-8-1509-2015>
- Wu, L., Wen, Z., Huang, R., & Wu, R. (2012). Possible linkage between the monsoon trough variability and the tropical cyclone activity over the Western North Pacific. *Monthly Weather Review*, *140*(1), 140–150. <https://doi.org/10.1175/MWR-D-11-00078.1>
- Xie, S. P., Hu, K., Hafner, J., Tokinaga, H., Du, Y., Huang, G., & Sampe, T. (2009). Indian Ocean capacitor effect on Indo-Western Pacific climate during the summer following El Niño. *Journal of Climate*, *22*(3), 730–747. <https://doi.org/10.1175/2008JCLI2544.1>
- Xie, S. P., Kosaka, Y., Du, Y., Hu, K., Chowdary, J. S., & Huang, G. (2016). Indo-Western Pacific ocean capacitor and coherent climate anomalies in post-ENSO summer: A review. *Advances in Atmospheric Sciences*, *33*(4), 411–432. <https://doi.org/10.1007/s00376-015-5192-6>
- Yang, L., Du, Y., Wang, D., Wang, C., & Wang, X. (2015). Impact of intraseasonal oscillation on the tropical cyclone track in the South China Sea. *Climate Dynamics*, *44*(5–6), 1505–1519. <https://doi.org/10.1007/s00382-014-2180-y>
- Yeh, S. W., Cai, W., Min, S. K., McPhaden, M. J., Dommenges, D., Dewitte, B., et al. (2018). ENSO atmospheric teleconnections and their response to greenhouse gas forcing. *Reviews of Geophysics*, *56*, 185–206. <https://doi.org/10.1002/2017RG000568>
- Zhan, R., & Wang, Y. (2016). CFSv2-based statistical prediction for seasonal accumulated cyclone energy (ACE) over the Western North Pacific. *Journal of Climate*, *29*(2), 525–541. <https://doi.org/10.1175/JCLI-D-15-0059.1>
- Zhang, G., Murakami, H., Gudgel, R., & Yang, X. (2019). Dynamical seasonal prediction of tropical cyclone activity: Robust assessment of prediction skill and predictability. *Geophysical Research Letters*, *46*, 5506–5515. <https://doi.org/10.1029/2019GL082529>
- Zhang, G., & Wang, Z. (2015). Interannual variability of tropical cyclone activity and regional Hadley circulation over the northeastern Pacific. *Geophysical Research Letters*, *42*, 2473–2481. <https://doi.org/10.1002/2015GL063318>
- Zhang, W., Vecchi, G. A., Murakami, H., Delworth, T., Wittenberg, A. T., Rosati, A., et al. (2016). Improved simulation of tropical cyclone responses to ENSO in the Western North Pacific in the high-resolution GFDL HiFLOR coupled climate model. *Journal of Climate*, *29*(4), 1391–1415. <https://doi.org/10.1175/JCLI-D-15-0475.1>
- Zhang, W., Vecchi, G. A., Murakami, H., Villarini, G., Delworth, T. L., Yang, X., & Jia, L. (2018). Dominant role of Atlantic multidecadal oscillation in the recent decadal changes in Western North Pacific tropical cyclone activity. *Geophysical Research Letters*, *45*, 354–362. <https://doi.org/10.1002/2017GL076397>
- Zhang, W., Vecchi, G. A., Murakami, H., Villarini, G., & Jia, L. (2016). The Pacific meridional mode and the occurrence of tropical cyclones in the Western North Pacific. *Journal of Climate*, *29*, 381–398. <https://doi.org/10.1175/JCLI-D-15-0282.1>
- Zhao, H., Klotzbach, P. J., & Chen, S. (2020). Dominant influence of ENSO-like and global sea surface temperature patterns on changes in prevailing boreal summer tropical cyclone tracks over the Western North Pacific. *Journal of Climate*, *33*(22), 9551–9565. <https://doi.org/10.1175/JCLI-D-19-0774.1>
- Zhao, H., & Wang, C. (2016). Interdecadal modulation on the relationship between ENSO and typhoon activity during the late season in the Western North Pacific. *Climate Dynamics*, *47*(1), 315–328. <https://doi.org/10.1007/s00382-015-2837-1>
- Zhao, H., & Wang, C. (2019). On the relationship between ENSO and tropical cyclones in the Western North Pacific during the boreal summer. *Climate Dynamics*, *52*(1), 275–288. <https://doi.org/10.1007/s00382-018-4136-0>
- Zhao, H., Wu, L., & Zhou, W. (2010). Assessing the influence of the ENSO on tropical cyclone prevailing tracks in the Western North Pacific. *Advances in Atmospheric Sciences*, *27*(6), 1361–1371. <https://doi.org/10.1007/s00376-010-9161-9>
- Zhou, B., & Cui, X. (2008). Hadley circulation signal in the tropical cyclone frequency over the Western North Pacific. *Journal of Geophysical Research*, *113*, D16107. <https://doi.org/10.1029/2007JD009156>
- Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., & Mayer, M. (2019). The ECMWF operational ensemble reanalysis-analysis system for Ocean and Sea Ice: A description of the system and assessment. *Ocean Science*, *15*(3), 779–808. <https://doi.org/10.5194/os-15-779-2019>