CMIP5, CMIP6 and reanalysis data

The TC data in the pre-training process came from the CMIP5 and CMIP6 models. The TC tracking scheme of the Camargo-Zebiak algorithm [55] is adopted. The TC data in the fine-tuning and verification periods were from the IBTrACS Version 4, which contains necessary variables, such as TC intensity, and TC latitude and longitude, with a temporal resolution of 6 hr and a spatial resolution of 0.1°. For quantitative description of TC spatial distribution, TC track density is defined as the frequency of TC occurring in 5°×5° bin at 6-hr interval from June to November.

Numerical Models and TC track method

To compare with SeaUnet simulation results, we select the seasonal prediction results of four numerical models (i.e., CFSv2, SEAS5, GloSea5-GC2, and Glosea6). The overviews of these four numerical models are as follow and the detail of them are shown in Table S1.

1. CFSv2: The CFS version 2 was developed at the Environmental Modeling Center at NCEP. It is a fully coupled model representing the interaction between the Earth's atmosphere, oceans, land and seaice.2. SEAS5: this is the fifth system run at ECMWF to produce real-time seasonal forecasts. Seasonal forecasts provide predictions of how the average atmospheric, ocean and land surface conditions over particular areas and periods of time are likely to be different from the long-term average.

3. GloSea5-GC2: GloSea5-GC2 is an ensemble prediction system built around the high resolution version of the Met Office climate prediction model: HadGEM3 family atmosphere-ocean coupled climate model.

4. Glosea6: GloSea6 is an ensemble prediction system built around the high resolution version of the Met Office climate prediction model: HadGEM3 family atmosphere—land—ocean—sea-ice coupled climate model.

Meridional and zonal winds at 10 m, 300 hPa, 500 hPa, and 850 hPa; temperatures of 200, 500 and 850 hPa; vorticity at 850 hPa; and sea-level pressure were used for identifying TC. Because the temporal resolution of some numerical models (i.e., SEAS5, GloSea5-GC2 and Glosea6) is 12 hr, we interpolated all TC results to 6 hr by linear interpolation. Similar to Sun et al. [56] and Kim et al. [57], the TC identification criteria are as follows:

1. The maximum relative vorticity at 850 hPa exceeds $4.0\times10^{-5}\ s^{-1}$.

2. The grid point with the minimum sea-level pressure within a radius of 500 km from the position of the maximum relative vorticity is defined as the center of the TC.

3. The maximum wind speed at 10 m within a radius of 800 km from the TC center is greater $10.8\ m\cdot s^{-1}$, and the lifetime maximum wind speed at 10 m is greater than $17.2\ m\ s^{-1}$.

4. TC-center temperature at 300 hPa is 1°C higher than the area-averaged temperature within a radius of 200-600 km from the TC center.

5. TC lifetime must be at least 48 hr.

6. The genesis location is located south of 30°N over the ocean.

Architecture of the SeaUnet model.

As shown in Figure. 1, SeaUnet is a fully convolutional neural network architecture, where the network backbone is Se_ResUnet [58]. Note that, different from traditional ResUnet, Se_ResUnet introduces SE attention module (Figure S1), which can assign weights to feature maps of different channels. All network models are written with TensorFlow2.3 library. Referring to the study of Song et al. [58], SeaUnet is composed of feature extraction network and feature fusion network (Figure 1). The main output layer of the model is composed of a convolution layer with kernel size of 1×1 and stride of 1×1, whose number of filter is 1; this is used to match the TC track density label. The secondary output layer of the model is composed of two fully connected neural networks with 128 neurons and one neuron, respectively. The secondary output layer corresponds to TCTF.

A residual layer in feature extraction network is composed of two Attention blocks. The Attention block structure is illustrated in Figure S1. It consists of a convolutional layer, a batch normalization (BN) layer, a Sqeeze-Excitation (SE) layer (Figure S1a), and a Spatial Attention (SA) layer (Figure S1b). ReLU is introduced into the Attention block as the activation function after the BN layer to realize nonlinear transformation. The first three residual layers are followed by a maximum pooling layer with kernel size of 2×2 and stride of 2×2 for down-sampling, and the fourth residual layer is followed by a deconvolution layer for up-sampling. The kernel sizes of the convolution layer in the four residual layers are 3×3, 3×3, 2×2, and 1×1, respectively; and the stride is 1×1.

In feature fusion network, each of the first two residual layers is followed by a deconvolution layer with kernel size of 2×2 and stride of 2×2 for up-sampling. The last residual layer is followed by a bottleneck layer. The kernel sizes of the convolution layer in the three residual layers are 2×2, 3×3 and 3×3, respectively; and the stride is 1×1. The bottleneck layer consists of two convolution layers with kernel size of 3×3, stride of 1×1 and a BN layer. The main output layer of the model is composed of a convolution layer with kernel size of 1×1 and stride of 1×1, whose number of filter is 1, which is used to match the TC track density label.

Attached to the last residual layer is a zero-padding layer, whose main purpose is to match the size of the heat map. The zero-padding layer is then connected with an average pooling layer with kernel size of 1×2 and stride of 1×2. The secondary output layer of the model is composed of two fully connected neural networks with 128 neurons and one neuron, respectively. The secondary output layer corresponds to the TC track frequency.

Evaluation indicators.

$$RMSE^t = \sqrt{\frac{1}{MN}\sum_{i}^{M}\sum_{j}^{N}\left(obs_{i,j}^t - pre_{i,j}^t\right)^2} \tag{S1}$$

where t represents time, $obs_{i,j}^t$ represents the observed TC track density at grid (i, j) at time step t, and $pre_{i,j}^t$ represents the predicted value of the model. M and N represent the meshes in meridional and zonal directions, respectively.

Relation between the anomalies of three variables (SubSST, H_{500} and U_{shear}) and Heat map

The H_{500}, u_{shear}, and SubSST are selected as variables in the SeaUnet according to priori knowledge in the previous studies. Figure S6 demonstrates the relation between the anomalies of these variables and Heat map in 2010 and 2018. For SubSST, the positive distance level coincides with the high contributing regions in the heat map (e.g., the Pacific Ocean in 2010, the Indian Ocean in 2010, the central tropical Pacific in 2018 and the sea east of Japan in 2018). For H_{500}, the positive distance level coincides with the high contributing regions in the heat map (e.g., the Arctic region in 2010 and the sea east of Japan in 2018). For U_{shear}, the positive distance level also coincides with the high contributing regions in the heat map (e.g., the Pacific Ocean in 2010, the Indian Ocean in 2010, and the central tropical Pacific in 2018).

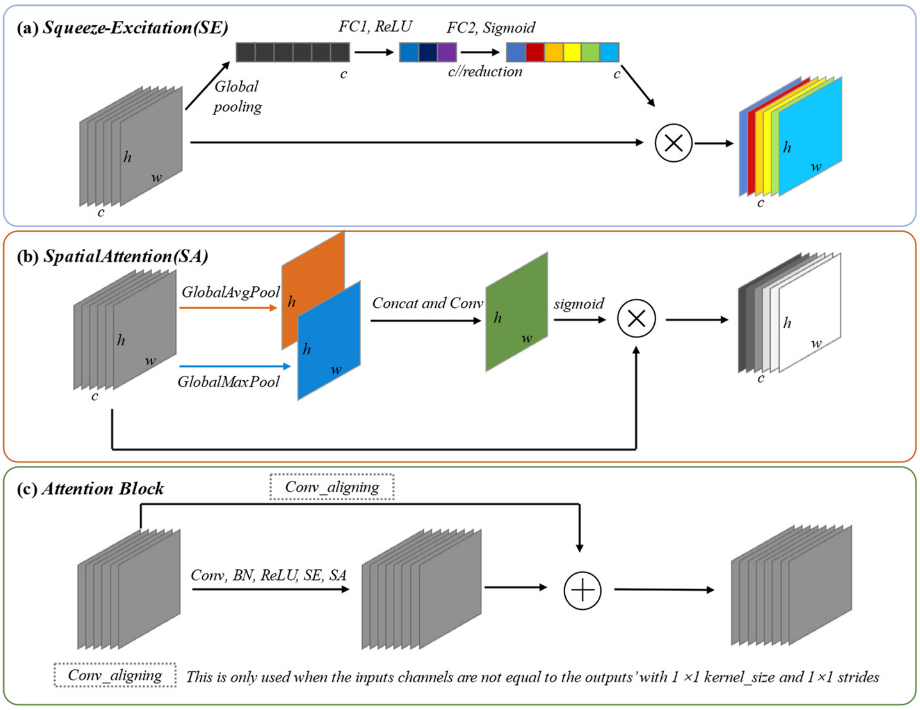

Figure S1. Diagram of attention block structure.

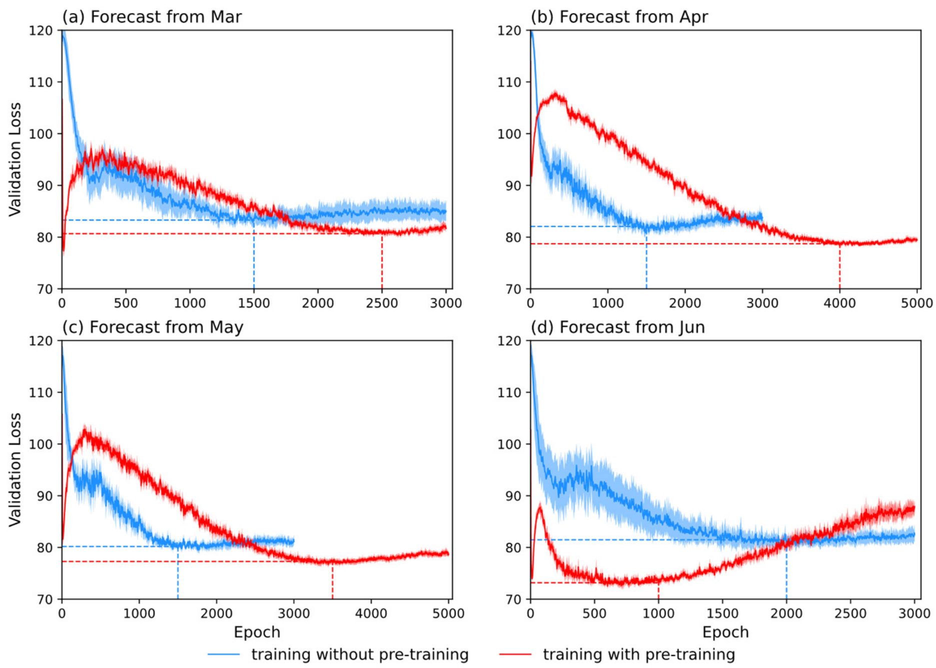

Figure S2. Prediction results during the fine-tuning process of the model from 2001 to 2005, the loss on the validation set changes with the number of training iterations. The realization represents the error loss value, and the shaded area represents 50% of the variance of 10 runs.

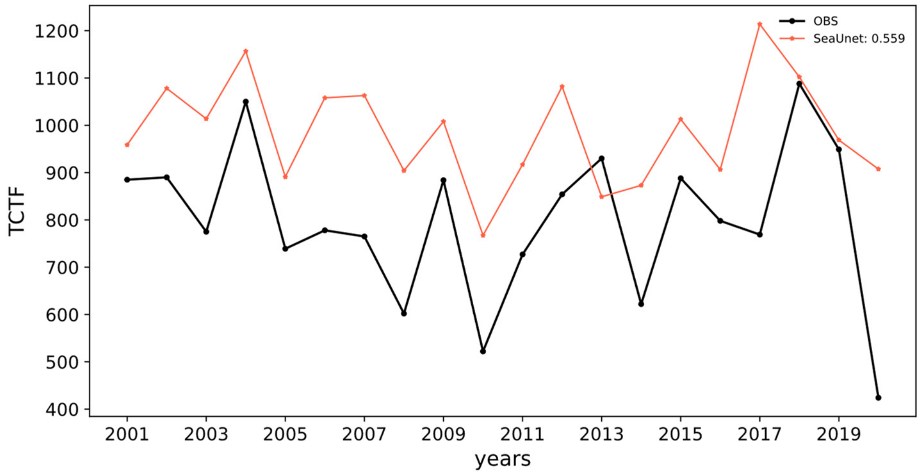

Figure S3. WNP TC track frequency (TCTF) predicted by the ensemble mean of SeaUnet and observed from 2001 to 2020. The number after the colon in the legend represents the correlation coefficient between the predicted and observed values.

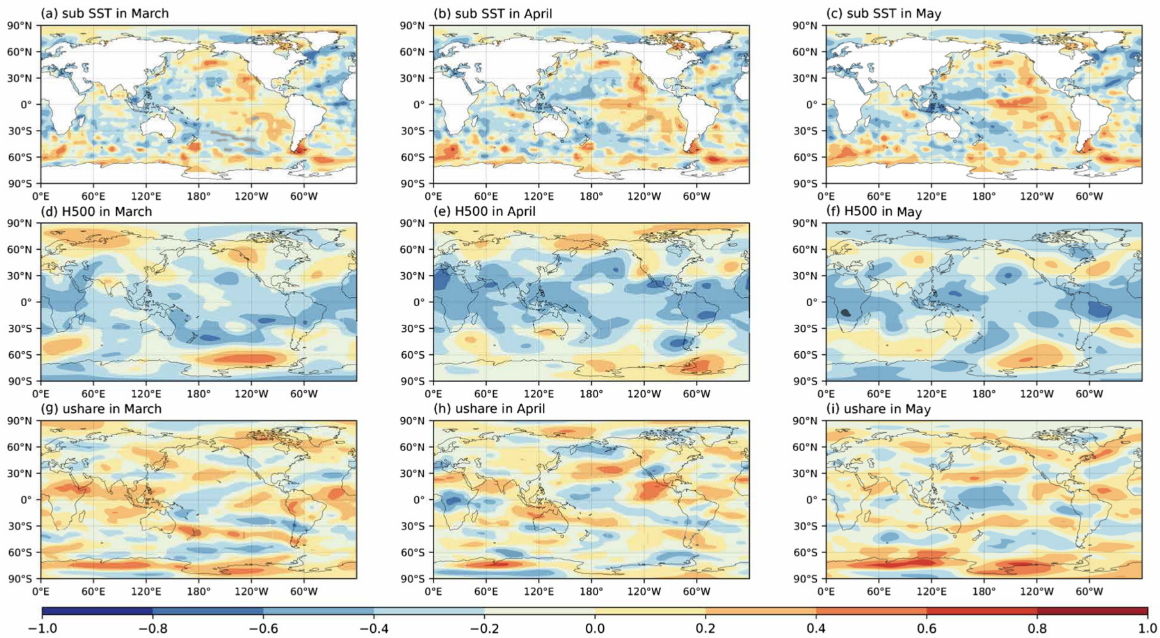

Figure S4. Correlation between precursors and TCTF in the WNP from June to November. The slash indicating the region that passes the 90% confidence test. The calculation time range is from 1991 to 2020. a, the precursor is subsurface SST in March. b, the precursor is subsurface SST in April. c, the precursor is subsurface SST in May. d, the precursor is H500 in March. e, the precursor is H500 in April. f, the precursor is H500 in May. g, the precursor is ushear in March. h, the precursor is ushear in April. i, the precursor is ushear in May.

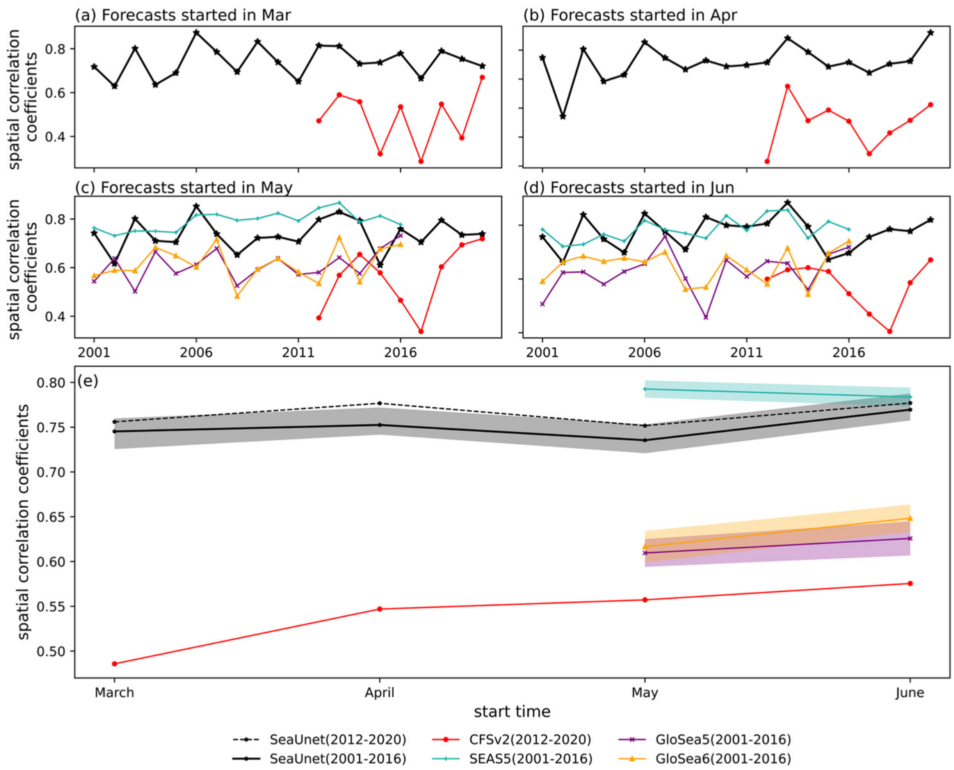

Figure S5. Spatial correlation coefficient between predicted TC track density results. Annual spatial correlation coefficient of the predicted TC track density of SeaUnet, SEAS5, Glosea5, and Glosea6 ensemble mean with respect to the observation in different starting months (i.e., (a) start in March; (b) start in April; (c) start in May; (d) start in June), with SeaUnet for the time range from 2001 to 2020, CFSv2 for the time range from 2012 to 2020, SEAS5, Glosea5 and Glosea6 for the time range from 2001 to 2016. (e) spatial correlation coefficient of each model changing with starting time. The solid line represents the average error of the ensemble mean of each model in the selected year, and the shading represents the 25% variance of the ensemble members. To ensure the rationality of comparison, the time range of the solid line for SeaUnet is from 2001 to 2016, and the time range of the dotted line is from 2012 to 2020; the time range for CFSv2 is from 2012 to 2020; and the time range for SEAS5, Glosea5 and Glosea6 is from 2001 to 2016.

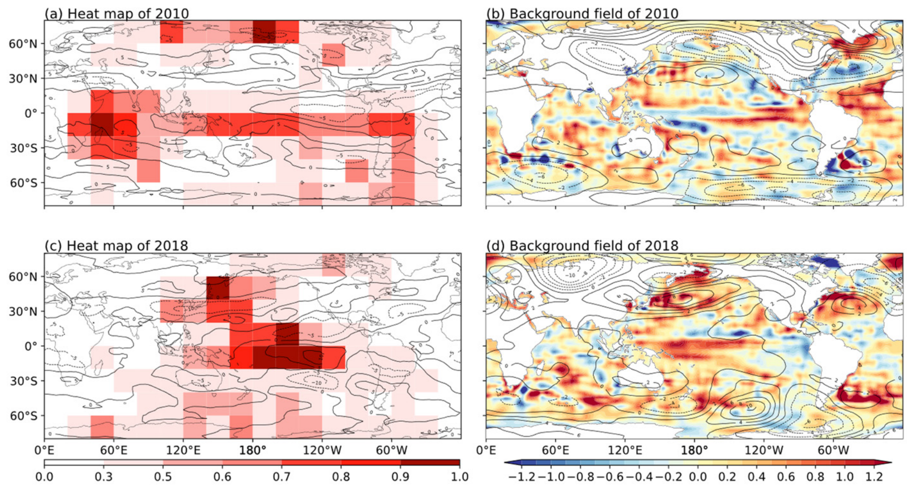

Figure S6. (a) Heat map for 2010, shading indicate heat map, contours indicate u_{shear}; (b) environmental field distance level averaged in 2010 (shading: SubSST distance level; contours: H_{500} distance level); (c) same as (a) but for 2018; (d) same as (b) but for 2018. The climate mean state is calculated from 1981 to 2010.

Table S1 Selected numerical models for seasonal prediction.

Model	Spatial resolution	Temporal resolution	Initial time	Forecast timeliness	Time range	Ensemble size	Research institution
SEAS5	1°×1°	12 hr	1 May and 1 June	7 months	2001-2016	25	European Centre for Medium-Range Weather Forecasts
Glosea5-GC2	1°×1°	12 hr	1 May and 1 June	7 months	2001-2016	8	Met office
Glosea6	1°×1°	12 hr	1 May and 1 June	7 months	2001-2016	8	Met office
CFSv2	0.5°×0.5°	6 hr	1 March, 1 April, 1 May, and 1 June	9 months	2012-2020	—	National Centers for Environmental Prediction

Table S2 Climate models of CMIP5 and CMIP6 used in this study.

Category	Model name	Atmospheric model resolution （lon×lat）	Ocean model resolution （lon×lat）	Research institution
CMIP5 (Historical, RCP4.5, RCP8.5)	CanESM2	128×64	256×192	Centre for Climate Modeling and Analysis
	CSIROMk3.6.0	192×96	192×189	Commonwealth Scientific and Industrial Research Organization in collaboration with the Queensland Climate Change Centre of Excellence
	FGOALS-g2	128×60	360×196	Institute of Atmospheric Physics, Chinese Academy of Sciences
	GFDL-CM3	144×90	360×200	Geophysical Fluid Dynamics Laboratory
	GFDL-ESM2M	144×90	360×200	Geophysical Fluid Dynamics Laboratory
	HadGEM2-ES	192×144	360×216	Hadley Center for Climate Prediction and Research, Met Office
	MIROC5	256×224	360×368	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
	MPI-ESM-LR	256×220	192×96	Max Planck Institute for Meteorology (MPI)
	MRI-CGCM3	320×160	360×368	Meteorological Research Institute
CMIP6 (Historical)	BCC-CSM2-MR	320×360	360×181	National Center for Climate Research, Beijing
	EC-Earth3	512×256	360×181	European Union Earth System Model Alliance
	MIROC6	256×128	360×181	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology
	MPI-ESM1-2-HR	384×192	360×181	Max Planck Institute for Meteorology (MPI)
	MRI-ESM2-0	320×160	360×181	Meteorological Research Institute
	NorESM2-MM	288×192	360×181	Norwegian Climate Centre