Assessing the performance of 33 CMIP6 models in simulating the large-scale environmental fields of tropical cyclones

Ying Han
Institute of Atmospheric Physics Chinese Academy of Sciences
Mengzhuo Zhang
Nanjing University
Zhongfeng Xu (xuzhf@tea.ac.cn)
Institute of Atmospheric Physics Chinese Academy of Sciences https://orcid.org/0000-0002-1274-6438
Weidong Guo
Nanjing University

Research Article

Keywords: tropical cyclone, multivariable integrated evaluation, CMIP6, large scale environmental field

Posted Date: May 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-339002/v1

Version of Record: A version of this preprint was published at Climate Dynamics on October 20th, 2021.
See the published version at https://doi.org/10.1007/s00382-021-05986-4.

Assessing the performance of 33 CMIP6 models in simulating the large-scale environmental fields of tropical cyclones

Ying HAN[1], Meng-Zhuo Zhang[2], Zhongfeng XU[1*] Weidong Guo[2]

[1]Key Laboratory of Regional Climate-Environment for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

[2]Institute for Climate and Global Change Research, School of Atmospheric Sciences, Nanjing University, Nanjing, 210093, China

[*] Corresponding author address:

Zhongfeng Xu

Institute of Atmospheric Physics,

Chinese Academy of Sciences,

Beijing 100029, China

E-mail: xuzhf@tea.ac.cn

Abstract：

General circulation model (GCM) biases are one of the important sources of biases and uncertainty in dynamic downscaling–based simulations. The ability of regional climate models to simulate tropical cyclones (TCs) is strongly affected by the ability of GCMs to simulate the large-scale environmental field. Thus, in this work, we employ a recently developed multivariable integrated evaluation method to assess the performance of 33 CMIP6 (phase 6 of the Coupled Model Intercomparison Project) models in simulating multiple fields. The CMIP6 models are quantitatively evaluated against two reanalysis datasets over five ocean areas. The results show that most of the CMIP6 models overestimate the mid-level humidity in almost all tropical oceans. The multi-model ensemble mean overestimates the vertical shear of the horizontal winds in the Northeast Pacific and North Atlantic. An increase in model horizontal resolution appears to be helpful in improving the model simulations. For example, there are 6–8 models with higher resolution among the top 10 models in terms of overall model performance in simulating the climatology and interannual variability of multiple variables. Similarly, there are 7–8 models with lower resolution among the bottom 10 patterns. The model skill varies depending on the region and variable being evaluated. Although no model performs best in all regions and for all variables, some models do show relatively good capability in simulating the large-scale environmental field of TCs. For example, the MPI-ESM1-2-LR, MPI-ESM1-2-HR, and FIO-ESM-2-0 models show relatively good skill in simulating the climatology and interannual variability of the large-scale environmental field in the Northern and Southern Hemispheres.

Key Words: tropical cyclone, multivariable integrated evaluation, CMIP6, large-scale environmental field

Declarations

Funding: This study was supported jointly by the National Key Research and Development Program of China (2017YFA0603803) and the National Science Foundation of China (41675105, 41775075, 42075152). The study was also supported by the Jiangsu Collaborative Innovation Centre for Climate Change.

Conflicts of interest/Competing interests (include appropriate disclosures): The authors declare no conflicts of interest.

Availability of data and material (data transparency): All data generated during this study are included in this published article.

Code availability (software application or custom code): All code used during this study have been published.

Authors' contributions: Ying Han and Zhongfeng Xu developed the idea of the study, participated in its design and coordination and draft the manuscript. Meng-Zhuo Zhang contributed to draw the figures. Weidong Guo provided review of the manuscript. All authors read and approved the final manuscript.

Ethics approval (include appropriate approvals or waivers): Not applicable

Consent to participate (include appropriate statements): All authors read and approved the final manuscript.

Consent for publication (include appropriate statements): Not applicable.

1. Introduction

Tropical cyclones (TCs) are among the most destructive natural hazards. Their strong winds and heavy rainfall pose great threats to human life and property. Therefore, the projection of TC frequency, tracks, and intensity is of great importance for human adaptation to climate change and associated decision-making.

Climate models are important tools for projecting the future changes and trends in climate events and extreme weather. However, GCMs have limited ability to simulate TCs owing to their coarse resolutions and model biases. A good approach to understanding the variations of TCs, for the purpose of forecasting both seasonal to interannual variations and long-term changes, is downscaling, which encompasses both dynamical downscaling and statistical downscaling methods (Chan et al., 2001; Fan and Wang, 2009; Wu and Yu, 2011; Knutson et al., 2013; Liu et al., 2012; Chen et al., 2020; Emanuel et al., 2008, 2020). Crucially, the performance of dynamical downscaling is strongly affected by the quality of its large-scale forcing data, e.g., the GCM output (Holland et al., 2010; Lui et al., 2020). Thus, in terms of dynamical downscaling–based simulation, it is critical to objectively evaluate the GCM's ability to simulate the large-scale fields associated with TCs genesis and development. Although some GCM bias correction methods have been developed and have proven useful in improving dynamical downscaling–based simulations over the past decade or so (e.g., Holland et al., 2010; Bruyère et al., 2014; Xu and Yang, 2012, 2015), these methods only correct part of the GCM bias. Thus, the performance of the GCM still plays a crucial role in the dynamical downscaling of future climate.

Generally, the most direct way to evaluate the ability of climate models to simulate TCs is to identify TCs from the 6-hourly outputs and compare their frequencies, tracks, and intensities with observations (Camargo, 2013; Roberts et al., 2020). However, such an approach requires downloading and processing a huge amount of data, which requires an equally huge workload. As we know, because of their coarse resolutions, GCMs cannot resolve the detailed structure of TCs well and usually underestimate the TC intensity and frequency. However, GCMs can generally resolve the TC-related large-scale fields well. Therefore, many previous studies have investigated the relationships between large-scale-environment fields with TCs using global models (Song et al., 2015). Based on the yearly genesis parameter (Gray, 1979), Emanuel and Nolan (2004) developed the TC genesis potential index (GPI), which summarizes the environmental factors influencing the genesis of TCs, such as low-level vorticity at 850 hPa, vertical wind shear between 850 and 200 hPa, relative humidity at 600 hPa, ocean temperature, and a conditionally unstable atmosphere. The parameters for an unstable atmosphere involve sea surface temperature (SST), sea level pressure (SLP), vertical atmospheric temperature, and mixing ratio. The GPI has been widely used to analyze the outputs of climate models, allowing information regarding the climate change of TCs to be obtained (Camago et al., 2007a, b; Villarini et al., 2012, 2013; Mei et al., 2019). Therefore, the GPI is an effective index for summarizing the features of the environmental fields of TCs. However, some studies have pointed out that the GPI is not always able to represent the actual TC variations on the global scale owing to certain limitations of the GPI and the complexity of TC

activities in different regions (Song et al., 2015; Zhang et al., 2010; Emanuel, 2010). For example, the GPI is not ideal for describing the frequency and location of TC genesis in the South China Sea–Northwest Pacific region because the system that affects the genesis of TCs in this region is different from that in other ocean areas (Zhao et al., 2012; Tao et al., 2020), such as the weight for thermodynamic factors and kinetic factors in GPI, and so on.

From the perspective of dynamical downscaling, GCMs provide multiple variables, e.g., temperature, wind, humidity, SST, and surface pressure, as the initial and lateral boundary conditions of regional models. Thus, the ability of GCMs to simulate these variables can directly affect the results of dynamical downscaling. Therefore, we need a comprehensive evaluation method that can quantitively evaluate model performance in simulating multiple large-scale environmental fields of TCs. The GPI has been used to evaluate model performance in simulating TCs (Camago et al., 2013; Song et al., 2015); however, this index is defined by the product of several variables (Emanuel and Nolan, 2004), which can often lead to a misleading result owing to the cancellation of biases of various variables. For example, a model may overestimate the humidity and vertical shear. The first bias favors the formation of TCs, but the second bias acts in an opposite way. Consequently, the modeled GPI could be close to the observed one and generate a good result for the wrong reasons. In terms of model evaluation, we of course should expect to avoid such a misleading result. Recently, Xu et al. (2017) devised a multivariable integrated evaluation (MVIE) method that can avoid the cancellation of biases issue in the evaluation of multiple

variables. The MVIE method is based on a vector field evaluation diagram, which is a generalization of the Taylor diagram (Taylor, 2001; Gleckler et al., 2008; Xu et al., 2016; 2017). The MVIE method can evaluate not only the performance of a model in terms of individual variables, but also its overall performance in simulating multiple variables, which can give a more accurate and comprehensive evaluation of the model performance.

This study uses the MVIE method to evaluate and compare the performances of 33 CMIP6 models in simulating multiple variables that are closely related to TC genesis and development. The evaluation is expected to provide guidance for selecting the optimal GCMs for the dynamical downscaling of TCs.

2. Data and methods

2.1 Data

The simulations used in this study are the first ensemble runs of the historical experiment from 33 CMIP6 models during the period 1979–2013 (Eyring, 2016). Table 1 provides an overview of the models used in this study. The red highlighting in Table 1 indicate the high-resolution models (resolution finer than 1°); the yellow highlighting represents the medium-resolution models (resolution of ~1°); and the rest are coarse-resolution models. Detailed model information can be obtained from http://cmip-pcmdi.llnl.gov/cmip6. In order to evaluate the performance of the CMIP6 models, two reanalysis datasets are used: the European Centre for Medium-Range Weather Forecasts Reanalysis-5 (ERA5), and the Japan Meteorological Agency and Central Research Institute of Electric Power Industry Reanalysis-55 (JRA-55). To

facilitate intercomparison, all model results and reanalysis products have been regridded to a common grid of 1.25°×1.25° by bi-linear interpolation method.

Based on previous studies, the genesis and development of TCs are closely related to the following large-scale conditions (Gray, 1968; Palmen, 1948; Riehl, 1948, 1950): (1) SST exceeding 26°C; (2) large cyclonic vertical vorticity in the lower troposphere; (3) weak vertical shear of horizontal winds between 850 and 200 hPa; (4) conditional instability through a deep tropospheric layer; (5) large values of relative humidity in the lower and middle troposphere; and (6) the disturbance should be approximately a minimum of 5° latitude from the equator, generally. Therefore, we evaluate and compare the performance of 33 CMIP6 models in terms of zonal and meridional wind (200 and 850 hPa), air temperature (200 and 850 hPa), specific humidity (600 hPa), SST, and SLP. The evaluation is carried out in five ocean basins with frequent TC activity: the Northwest Pacific, Northeast Pacific, North Atlantic, Southwest Pacific, and South Indian oceans (Table 2).

Table 1. CMIP6 models evaluated in this study.

	Models	Institution	Resolution
1	**ACCESS-CM2**	Commonwealth Scientific and Industrial Research Organisation, and Australian Research Council Centre of Excellence for Climate System Science (Australia)	~1.88° × 1.25°
2	**ACCESS-ESM1-5**	Commonwealth Scientific and Industrial Research Organisation (Australia)	~1.88° × 1.25°
3	**AWI-CM-1-1-MR**	Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (Germany)	~0.94° × 0.94°
4	**BCC-CSM2-MR**	Beijing Climate Center (China)	~1.13° × 1.13°
5	**BCC-ESM1**	Beijing Climate Center (China)	~2.81° × 2.81°
6	**CAMS-CSM1-0**	Chinese Academy of Meteorological Sciences	~1.13° × 1.12°

		(China)	
7	**CanESM5**	Canadian Centre for Climate Modelling and Analysis, Environment and Climate Change (Canada)	~2.81° × 2.81°
8	**CESM2**	National Center for Atmospheric Research, Climate and Global Dynamics Laboratory (USA)	~1.25° × 0.94°
9	**CESM2-FV2**	National Center for Atmospheric Research, Climate and Global Dynamics Laboratory (USA)	~2.5° × 1.89°
10	**CESM2-WACCM**	National Center for Atmospheric Research, Climate and Global Dynamics Laboratory (USA)	~1.25° × 0.95°
11	**CESM2-WACCM-FV2**	National Center for Atmospheric Research, Climate and Global Dynamics Laboratory (USA)	~2.5° × 1.89°
12	**CIESM**	Department of Earth System Science, Tsinghua University (China)	~1.25° × 0.94°
13	**E3SM-1-0**	Lawrence Livermore National Laboratory, Argonne National Laboratory, Brookhaven National Laboratory, Los Alamos National Laboratory, Lawrence Berkeley National Laboratory, Oak Ridge National Laboratory, Pacific Northwest National Laboratory, and Sandia National Laboratories (USA)	~1° × 1°
14	**EC-Earth3**	AEMET (Spain), BSC (Spain), CNR-ISAC (Italy), DMI (Denmark), ENEA (Italy), FMI (Germany), ICHEC, (Ireland), ICTP (Italy), IDL (Portugal), IMAU (Netherlands), IPMA (Portugal), KIT (Germany), KNMI (Netherlands), Lund University (Sweden), Met Eireann (Ireland), NLeSC (Netherlands), NTNU, (Norway), Oxford University (UK), surfSARA (Netherlands), SMHI (Sweden), Stockholm University, (Sweden), Unite ASTR (Belgium), University College Dublin, (Ireland), University of Bergen (Norway), University of Copenhagen (Denmark) University of Helsinki (Finland), University of Santiago de Compostela (Spain) Uppsala University (Sweden), Utrecht University (Netherlands), Vrije Universiteit Amsterdam (Netherlands), Wageningen University (Netherlands)	~0.70° × 0.70°
15	**FGOALS-g3**	Chinese Academy of Sciences (China)	~2° × 2.25°
16	**FIO-ESM-2-0**	First Institute of Oceanography, and Qingdao National Laboratory for Marine Science and Technology (China)	~1.25° × 0.94°
17	**GISS-E2-1-G**	Goddard Institute for Space Studies (USA)	~2.48° × 2°
18	**GISS-E2-1-H**	Goddard Institute for Space Studies (USA)	~2.48° × 2°

19	**INM-CM4-8**	Institute for Numerical Mathematics, Russian Academy of Science (Russia)	~2° × 1.5°
20	**INM-CM5-0**	Institute for Numerical Mathematics, Russian Academy of Science (Russia)	~2° × 1.5°
21	**IPSL-CM6A-LR**	Institut Pierre Simon Laplace (France)	~2.5° × 1.27°
22	**KACE-1-0-G**	National Institute of Meteorological Sciences/Korea Meteorological Administration, Climate Research Division (Korea)	~1.87° × 1.25°
23	**MCM-UA-1-0**	Department of Geosciences, University of Arizona (USA)	~3.75° × 2.25°
24	**MIROC6**	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, National Institute for Environmental Studies, and RIKEN Center for Computational Science (Japan)	~1.41° × 1.41°
25	**MPI-ESM-1-2-HAM**	ETH Zurich (Switzerland), Max Planck Institut fur Meteorologie (Germany), Forschungszentrum Julich (Germany), University of Oxford (UK), Finnish Meteorological Institute (Finland), Leibniz Institute for Tropospheric Research (Germany), and Center for Climate Systems Modeling (Switzerland)	~1.88° × 1.88°
26	**MPI-ESM1-2-HR**	Max Planck Institute for Meteorology, and Deutsches Klimarechenzentrum (Germany)	~0.94° × 0.94°
27	**MPI-ESM1-2-LR**	Max Planck Institute for Meteorology, and Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (Germany)	~1.88° × 1.88°
28	**MRI-ESM2-0**	Meteorological Research Institute (Japan)	~1.13° × 1.13°
29	**NESM3**	Nanjing University of Information Science and Technology (China)	~1.88° × 1.88°
30	**NorCPM1**	Center for International Climate and Environmental Research, Norwegian Meteorological Institute, Nansen Environmental and Remote Sensing Center, NERSC, Norwegian Institute for Air Research, University of Bergen, University of Oslo, Uni Research (Norway)	~2.5° × 1.89°
31	**NorESM2-LM**	Center for International Climate and Environmental Research, Norwegian Meteorological Institute, Nansen Environmental and Remote Sensing Center, NERSC, Norwegian Institute for Air Research, University of Bergen, University of Oslo, Uni Research (Norway)	~2.5° × 1.89°
32	**NorESM2-MM**	Center for International Climate and Environmental Research, Norwegian Meteorological Institute,	~1.25° × 0.94°

		Nansen Environmental and Remote Sensing Center, NERSC, Norwegian Institute for Air Research, University of Bergen, University of Oslo, Uni Research (Norway)	
33	**SAM0-UNICON**	Seoul National University (Korea)	~1.25° × 0.94°

Table 2. Regions and time periods evaluated in this study.

Area (only ocean)	Latitude	Longitude	Months (1979–2013)
North Atlantic	EQ–45°N	90°W–20°W	7–10
Northwest Pacific	EQ–45°N	105°E–170°W	7–10
Northeast Pacific	EQ–45°N	169°W–90°W	7–10
South Indian ocean	45°S –EQ	35°E–110°E	1–4
Southwest Pacific	45°S –EQ	135°E–145°W	1–4

2.2 Methods

Statistical metrics—namely, mean error (ME), correlation coefficient (CORR), and standard deviation (SD)—are used to measure the model ability to simulate individual variables. To measure the performance of climate models in simulating vector fields or multiple fields, we also compute the centered root-mean-square length (cRMSL) of a vector field, centered vector similarity coefficient (cVSC), and multivariable integrated evaluation index (MIEI) (Xu et al., 2016, 2017). The cRMSL and cVSC are analogous to the SD and CORR except that they measure the vector field SD and similarity, respectively. These statistics have also been applied to evaluate the vector winds in the Asian-Australian monsoon (A-AM) region simulated by CMIP5 (phase 5 of the Coupled Model Intercomparison Project) models and examine the relationship with model abilities to simulate vector wind and precipitation (Huang et al., 2019, 2020).

cRMSL and cVSC can be applied to vector fields in arbitrary dimensions. Thus, one can normalize individual scalar fields and group them into a multi-dimensional vector field. Under such a circumstance, cRMSL and cVSC can be used to evaluate model ability to simulate multiple fields. Based on cRMSL and cVSC, Xu et al., (2017) further defined the MIEI, which summarizes overall model performance in simulating multiple fields and ranks model performances across various CMIP6 models. A smaller MIEI represents a better model performance.

3. Results

3.1 Climatological mean of the multi-model ensemble mean

Figure 1 shows the difference between the multi-model ensemble (MME) mean of CMIP6 and the reanalysis data. The vertical shear of the horizontal wind between 850 and 200 hPa is defined as follows:

$$S = \frac{\sqrt{(u_{200\mathrm{hPa}} - u_{850\mathrm{hPa}})^2 + (v_{200\mathrm{hPa}} - v_{850\mathrm{hPa}})^2}}{\mathrm{abs}(H_{200\mathrm{hPa}} - H_{850\mathrm{hPa}})}, \quad (1)$$

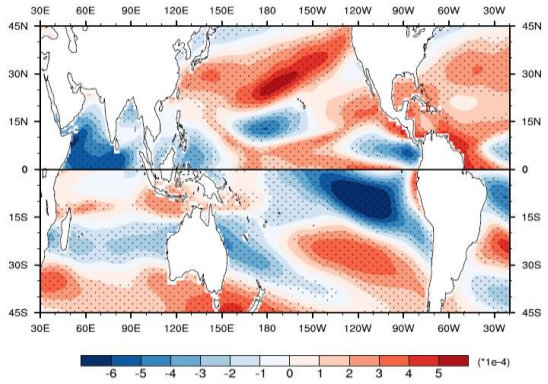
where H is geopotential height. The CMIP6 models show different errors in the vertical wind shear fields in different regions. For example, in the North Indian Ocean, the equatorial western Pacific, and southeastern Pacific, CMIP6 models generally underestimate the vertical wind shear relative to the reanalysis data, with a maximum negative deviation of 4×10^{-4} s^{-1}. In contrast, CMIP6 models significantly overestimate the wind shear in the mid-latitudes of the northern Pacific and Atlantic regions, with a maximum positive deviation of 5×10^{-4} s^{-1}. The regional mean MME wind shear in the South Indian Ocean and Southwest Pacific is close to the value of

226 the reanalysis data (Fig. 1a). In terms of specific humidity, the CMIP6 MME shows a
227 significant positive bias relative to the reanalysis data in most ocean areas, but
228 especially in the low latitudes of 5°–15° in the Northern and Southern Hemispheres,
229 where TCs usually originate (Fig. 1b). This result is generally consistent with Tian et
230 al. (2013), who found that CMIP5 models are also wetter at 600 hPa in comparison to
231 the Atmospheric Infrared Sounder (AIRS) data. Conversely, the models show a slight
232 negative bias over the central equatorial Pacific, which appears to be related to the
233 double-ITCZ problem in the spatial distribution of moisture. The double-ITCZ
234 problem has long persisted in coupled GCMs (Lin, 2007; Liu et al., 2012). Compared
235 with other ocean areas, the deviation of specific humidity in the North Atlantic is
236 relatively small. The MME SST shows a significant cold bias of approximately −1°C
237 in the northern Pacific and Atlantic oceans within 20°–40°N, and the equatorial
238 Pacific, against a warm bias of 0.5°C–2°C elsewhere (Fig. 1c). The distribution of the
239 SLP bias is generally consistent with that of the SST bias. The SST cold bias
240 corresponds to an SLP positive bias, and vice versa. The spatial maps of the
241 climatological mean air temperature between the CMIP6 MME and the reanalysis
242 data show a cold bias in 200-hPa temperature in most regions (Fig. 1f). The 850-hPa
243 temperature also shows a significant cold bias in middle and low latitudes, except for
244 weak warm biases in the Southeast and Northeast Pacific Ocean and the southern
245 oceans to the east and west of Australia (Fig. 1e).

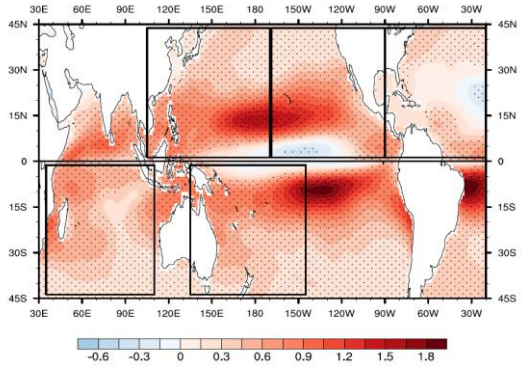
246 Clearly, the CMIP6 MME shows significant biases in large-scale environmental
247 fields, which could in turn affect the downscaled TC activities, e.g., their genesis,

track, and intensity, through the lateral boundary conditions of the regional climate model (RCM). The overestimated vertical shear of zonal wind in the Atlantic Ocean tends to hinder the genesis of TCs and weaken the intensity of TCs (Bruyère et al, 2014). In contrast, the underestimated vertical shear of zonal wind in the central Pacific, northern Indian Ocean, and western Pacific may favor the formation of TCs. A wet middle atmosphere is also conducive to the formation and maintenance of TCs. The bias of SLP could affect the track, location, and other characteristics of TCs.

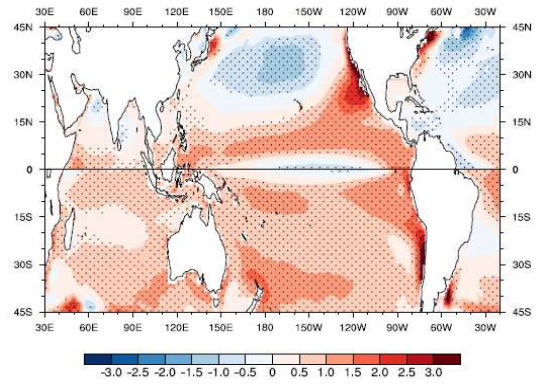
(a) Wind shear between 850 and 200 hPa (s^{-1})



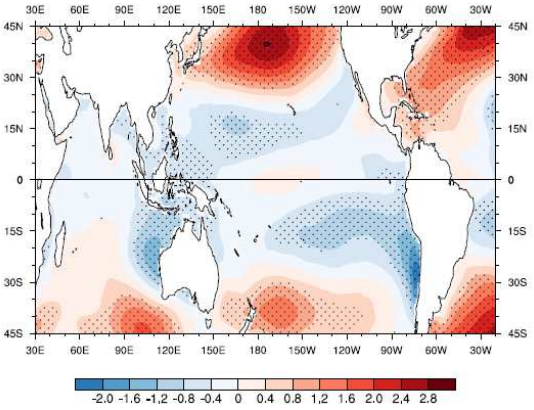
(b) 600-hPa specific humidity(g/kg)



(c) SST (°C)



(d) SLP (hPa)



(e) 850-hPa temperature (°C)

(f) 200-hPa temperature (°C)

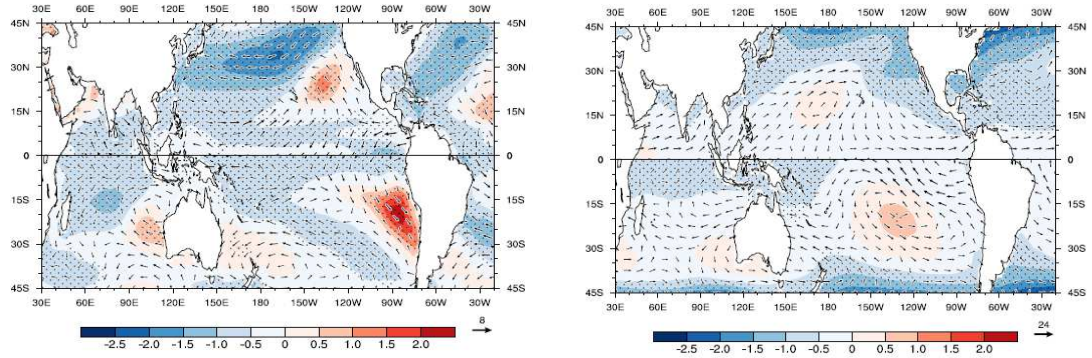


Fig. 1. Differences of climatological mean (a) vertical wind shear (s^{-1}), (b) 600-hPa specific humidity (g/kg), (c) SST (°C), (d) SLP (hPa), (e) 850-hPa air temperature (°C), and (f) 200-hPa air temperature (°C) between the multi-model ensemble mean and the reanalysis data. Black dots indicate the difference reaching the significance level of 0.05. The climatological mean is calculated over the typical TC seasons, i.e., July, August, September and October in the Northern Hemisphere, and January, February, March and April in the Southern Hemisphere. The black boxes in (a) indicate the five ocean regions with frequent TC activities, i.e., the Northwest Pacific, Northeast Pacific, North Atlantic, South Indian Ocean, and Southwest Pacific.

3.2 Inter-model spread

In order to analyze the dispersion of CMIP6 models, we calculate the SD of climatological mean variables across the 33 CMIP6 models as follows (Huang et al., 2020):

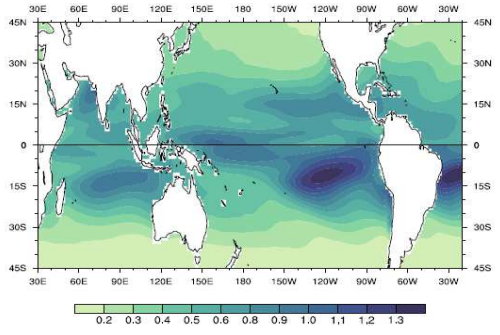
$$\sigma_p = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_{CMIP6,i} - P_{MME}\right)^2} \quad (2)$$

$$\sigma_v = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{V}_{CMIP6,i} - \overline{\boldsymbol{V}}_{MME}\right)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(u_{CMIP6,i} - \bar{u}_{MME})^2 + (v_{CMIP6,i,} - \bar{v}_{MME})^2} \quad (3)$$

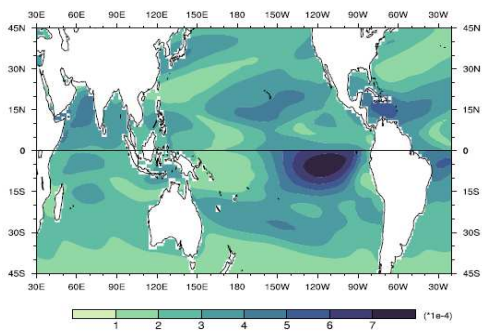
where $\sigma_p(\sigma_v)$ represents the SD of the scalar field (vector field) simulated by CMIP models relative to the MME mean, and N is equal to 33, representing the number of CMIP6 models. The SD can measure the dispersion of the CMIP6 models relative to their MME. The CMIP6 models show similar spatial patterns in terms of the inter-model spread of SLP, temperature, and SST, characterized by greater inter-model spread in the North Pacific, North Atlantic, Southeast Pacific, and southern Indian Ocean, and relatively small inter-model spread in the western Pacific (Fig. 2). In the Northwest Pacific, the inter-model spread of SLP appears to be relatively large, which

may lead to a diverse simulation of TC tracks across different models (Fig. 2d). In the Caribbean Sea and tropical eastern Pacific, the vertical wind shear shows a greater inter-model spread relative to the other regions (Fig. 2b). As we know that most TCs are formed in the tropical zone between 10° and 30°, and 87% are formed within 20° (Li, 1956; Elsberry, 1994), the greater inter-model spread of vertical wind shear in these regions indicates that the number of TCs generated in the Atlantic Ocean is likely to be significantly different among the CMIP6 models. Over the Southwest Pacific, the CMIP6 models show a relatively smaller inter-model spread in terms of the SST and 850-hPa air temperature, but a greater inter-model spread in terms of the wind shear and SLP. The inter-model spread of environmental fields is generally smaller in the South Indian Ocean than the other oceans. In terms of specific humidity, the CMIP6 models show greater inter-model spread in the tropical–subtropical regions between 30°S and 30°N (Fig. 2a). In general, CMIP6 models still show significant differences in the simulation of TC environmental fields, indicating large uncertainty in TC simulation.

(a) Specific humidity (g/kg)



(b) Wind shear between 850 and 200 hPa (s−1)



(c) SST (°C)

(d) SLP (hPa)

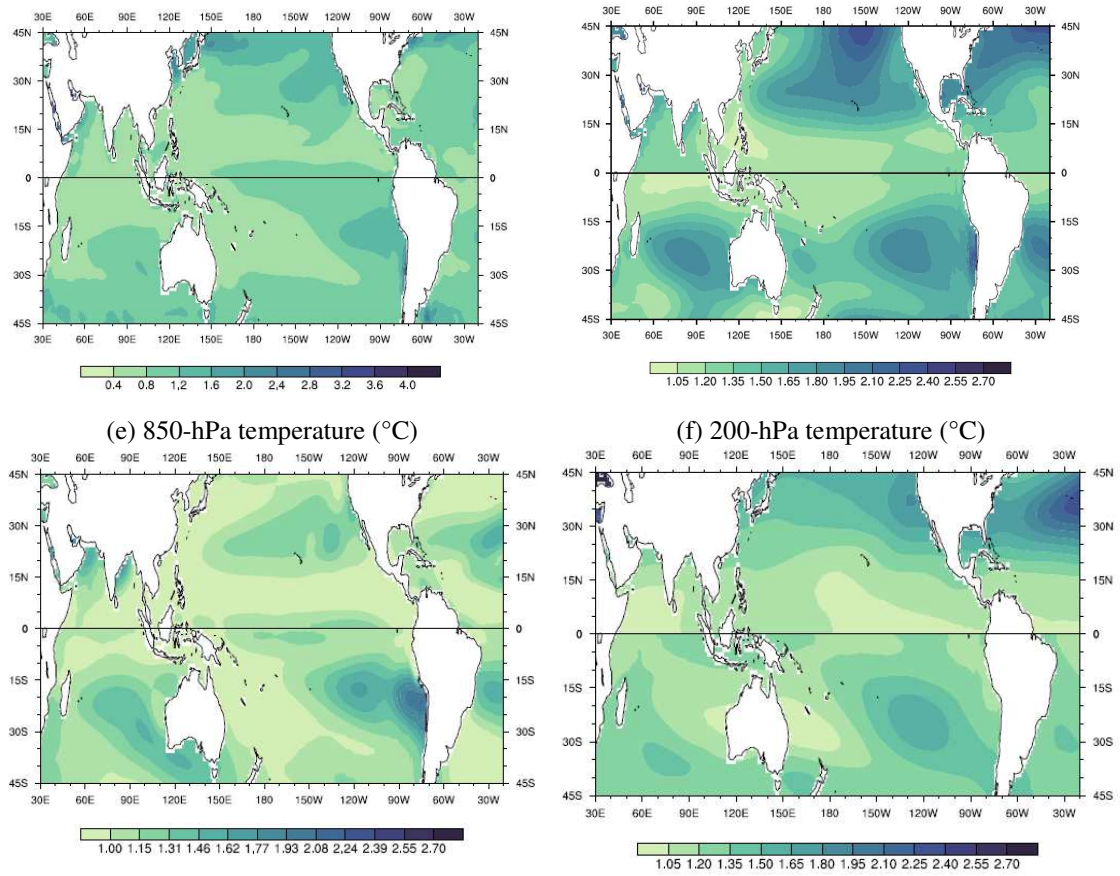


Fig. 2. Inter-model spread among 33 CMIP6 models.

3.3 Ability of individual CMIP6 models in simulating the climatology

As interpreted in section 3.2, the CMIP6 models show significant inter-model spread in terms of the simulation of various variables. In this section, we further assess the performance of individual models in simulating the large-scale environmental fields using the MVIE method developed by Xu et al. (2017). The evaluation focuses on seven variables—namely, the vector winds and air temperature at 200 and 850 hPa, 600-hPa specific humidity, SST, and SLP. These variables are on the one hand closely related to the genesis and development of TCs, but on the other hand they are also key variables driving RCMs as lateral boundary conditions during dynamical downscaling–based simulations.

3.3.1 Evaluation of multiple environmental fields in the Northwest Pacific

Each CMIP6 model is assessed against the average of two reanalysis datasets, i.e., ERA5 and JRA55. Figure 3 shows the statistical metrics of the 33 CMIP6 models in terms of the climatology of multiple variables in the Northwest Pacific region during the TC season (from July to October). The CMIP6 models can generally simulate well the spatial pattern of the climatological mean SST, SLP, 200-hPa wind field, and 850-hPa temperature field, with CORRs higher than 0.9. In contrast, the CMIP6 models show relatively poor performance in reproducing the spatial patterns of 850-hPa vector wind, 600-hPa specific humidity, and 200-hPa temperature, with CORRs ranging from 0.6 to 0.98. Model performance varies with the variable evaluated. However, the E3SM-1-0 and NorESM2-LM models show the highest cVSC among the 33 CMIP6 models, which indicates that these two models perform best in terms of simulation of the spatial pattern of these seven variables over the Northwest Pacific Ocean.

The CMIP6 models show large diversity in their simulation of the spatial SD of the different variables. For example, 25 out of 33 CMIP6 models underestimate the spatial variability of 200-hPa vector wind over the Northwest Pacific Ocean. Conversely, all models except three (MCM-UA-1-0, SAM0-UNICON, and ACCESS-ESM1-5) overestimate the SD of the 600-hPa specific humidity by 6%–49%. The CMIP6 models also significantly overestimate the 600-hPa specific humidity, characterized by an ME ranging from 15% to 165%. Clearly, most of the CMIP6 models overestimate the 600-hPa specific humidity and its spatial variability.

Such a wet bias was also found in CMIP5 models (Jiang et al., 2012; Tian et al., 2013). Su et al. (2013) and Takahashi et al. (2015) pointed out that the wet biases of CMIP5 models were likely related to underestimation of the intensity of climatological descent over ocean regions in the models.

Apart from ACCESS-ESM1-5 and BCC-CSM2-MR, all the CMIP6 models evaluated in this study underestimate the 850-hPa air temperature over the Northwest Pacific Ocean, with the ME ranging from −0.01 to −0.98, which is consistent with the results of the spatial mean fields (Figs. 1e and 3). The 200-hPa mean temperature simulated by the CMIP6 models shows greater diversity, with MEs ranging from −2.03 to 2.02. Thirty-one models (~94% of all models) show the cold bias in the lower troposphere, and 21 models (~64% of all models) show the cold bias in the upper troposphere. Tian et al. (2013) also found significant upper-tropospheric cold biases in most CMIP5 models. It seems that many CMIP6 models have inherited the cold biases from their predecessors.

Generally, there is no single model that performs best in terms of all variables and all statistics. For example, FIO-ESM-2-0 shows the best performance in terms of simulation of the spatial pattern of the climatological mean SST among the 33 CMIP6 models. However, this model moderately overestimates the domain-averaged SST by approximately 0.12°C, and the 600-hPa specific humidity in the Northwest Pacific Ocean. Based on the MIEI, some models, e.g. ACCESS-ESM1-5, BCC-CSM2-MR, CESM2, FIO-ESM-2-0, and NorESM2-MM, show relatively good performance in

terms of their simulation of multiple environmental fields of TCs in the Northwest Pacific region.

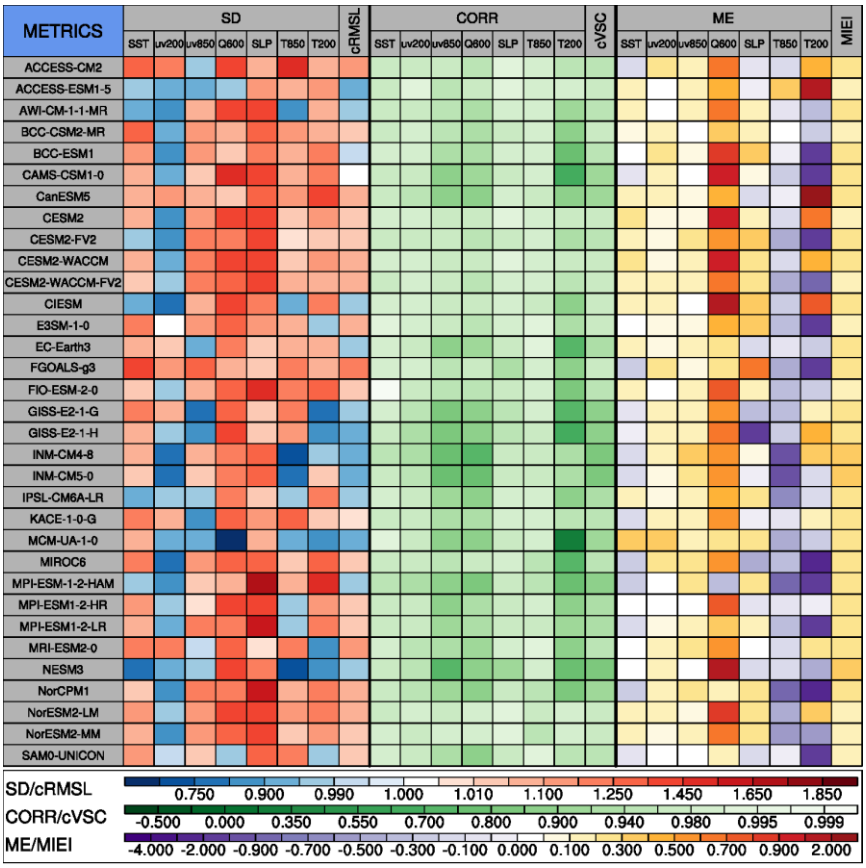


Fig. 3. Statistical metrics measuring model abilities in simulating the climatology of TC environmental fields over the Northwest Pacific Ocean during the TC season (July, August, September and October). SD: spatial standard deviation; cRMSL: centered root-mean-square length, measuring the overall SD of multiple variables; CORR: correlation coefficient; cVSC: centered vector similarity coefficient, measuring the overall pattern similarity of multiple variables; ME: mean error; MIEI: multivariable integrated evaluation index, measuring overall model performance in simulating multiple variables, which takes pattern similarity, spatial variability, and ME into account. The SD and ME are normalized by dividing the observational SD. The lighter colors represent results that are close to the observations, and vice versa. Cold and warm colors indicate that the biases are smaller and larger than the observed values, respectively.

3.3.2 Evaluation of the performance of CMIP6 models in five ocean regions

In order to assess the overall simulation capability for TC environmental fields, Fig. 4 shows the MIEIs of the 33 CMIP6 models in five ocean areas. Here, the MIEI is an uncentered statistic that measures the overall performance of the CMIP6 models

in simulating multiple environmental fields. A smaller MIEI represents better model performance. The models on the abscissa in Fig. 4 are ranked according to their average MIEI value in the five ocean areas. The models within red boxes represent the high-resolution models (resolution finer than 1°); the orange boxes indicate the medium-resolution models (resolution of ~1°); and the rest are coarse-resolution models. In general, most of the high-resolution models have smaller MIEI values, while the coarse-resolution models have larger MIEI values (Fig. 4). For example, 8 models out of the top 10 performing CMIP6 models have high or moderate horizontal resolution. In contrast, there are 8 coarse-resolution models among the bottom 10 models. These results indicate that an increase in model resolution likely favors improvement in the simulation ability. Such a relationship can also be identified by comparing models developed by the same modeling center, e.g., BCC-CSM2-MR vs. BCC-ESM1, CESM2 vs. CESM2-FV, CESM-WACCM vs. CESM2-WACCM-FV2, and NorESM2-MM vs. NorESM2-LM. Evaluation of 37 CMIP5 models also suggested that high-resolution models generally perform better in simulating the wind vector field in the Asian-Australian monsoon region (Huang et al., 2019). Of course, the above relationship is not always valid. For example, MPI-ESM1-2-LR has a coarse resolution but it ranks ahead of its high-resolution version, MPI-ESM1-2-HR. Note that a model's simulation ability is affected by other factors, such as its physical parameterization schemes, model dynamics, etc. Resolution is not the only factor that affects model performance.

Comparison of the evaluation results in different ocean areas suggests that the CMIP6 models generally show better skill in the southern Indian Ocean than the other regions (Fig. 4). According to their MIEI values, the top 10 models are shaded blue in each ocean region in Fig. 5. The CMIP6 models show obvious differences in their ability to simulate TC environmental fields in each ocean region. For example, CESM2-WACCM-FV2 only performs well in the Northwest Pacific, while BCC-CSM2-MR performs well in the Northwest Pacific and South Indian Ocean regions. In contrast, NorESM2-LM performs better over the three ocean regions in the Northern Hemisphere. Similarly, ACCESS-ESM1-5, CESM2, FIO-ESM-2-0, MPI-ESM1-2-HR, MPI-ESM1-2-LR and NorESM2-LM also perform well in these three oceans. The AWI-CM-1-1-MR and NorESM2-MM models show good performance in four ocean areas. Among all CMIP6 models evaluated in this study, SAM0-UNICON is the only model that performs relatively well in all five ocean areas in terms of simulation of the climatological mean TC environmental fields.

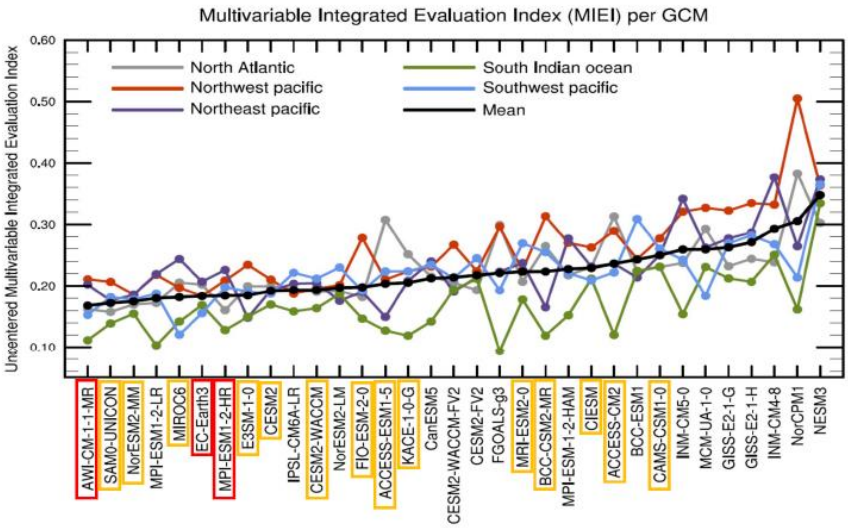


Fig. 4. MIEI values for 33 CMIP6 models in five regions. Red boxes on the abscissa indicate the

high-resolution models (resolution less than 1°); the orange boxes indicate the medium-resolution models (resolution of ~1°); and the rest are coarse-resolution models.

	NWP	NEP	NA	SIO	SWP
ACCESS-CM2	21	24	32	6	19
ACCESS-ESM1-5	2	10	31	7	20
AWI-CM-1-1-MR	10	11	3	3	2
BCC-CSM2-MR	3	26	27	4	26
BCC-ESM1	15	18	19	29	32
CAMS-CSM1-0	25	22	20	31	27
CanESM5	23	16	17	10	23
CESM2	9	9	11	21	8
CESM2-FV2	16	14	9	26	25
CESM2-WACCM	12	4	7	19	14
CESM2-WACCM-FV2	8	20	14	24	15
CIESM	20	19	21	27	13
E3SM-1-0	1	17	12	13	9
EC-Earth3	13	1	13	20	3
FGOALS-g3	18	25	29	1	11
FIO-ESM-2-0	7	23	6	12	10
GISS-E2-1-G	29	28	22	28	29
GISS-E2-1-H	30	31	25	25	31
INM-CM4-8	33	30	24	32	28
INM-CM5-0	31	27	23	15	24
IPSL-CM6A-LR	11	3	10	17	18
KACE-1-0-G	14	13	26	5	21
MCM-UA-1-0	26	29	28	30	6
MIROC6	24	5	15	11	1
MPI-ESM-1-2-HAM	28	21	18	14	17
MPI-ESM1-2-HR	19	8	2	8	12
MPI-ESM1-2-LR	17	12	5	2	7
MRI-ESM2-0	22	15	16	22	30
NESM3	32	32	30	33	33
NorCPM1	27	33	33	18	16
NorESM2-LM	4	6	8	23	22
NorESM2-MM	6	2	4	16	4
SAM0-UNICON	5	7	1	9	5

Fig. 5. Ranking of 33 CMIP6 models based on their MIEI values in five ocean regions: Northwest Pacific (NWP); Northeast Pacific (NEP); North Atlantic (NA); Southwest Pacific (SWP); South Indian Ocean (SIO). The top 10 models are marked in blue.

379

380 **3.4 Ability of individual CMIP6 models in simulating the interannual variability**

381 In this study, the interannual variability is measured by the interannual SD of
382 meteorological variables calculated over the period 1979–2014, which describes the
383 amplitude of the interannual variation and to a certain extent is related to climate
384 extremes. The interannual variability of TCs simulated by models is likely affected by
385 their ability to simulate the interannual variability of the large-scale environmental
386 fields. When selecting GCMs for dynamical downscaling–based simulation, we
387 expect them to perform well not only with respect to the climatology, but also the
388 interannual variability. Similar to Fig. 3, Fig. 6 shows the statistics of seven TC

389 environmental variables for the 33 CMIP6 models, but for the spatial field of the
390 interannual SD over the Northwest Pacific region.

391 The CMIP6 models show diverse performance in reproducing the spatial patterns
392 of the interannual variability of the different variables (Fig. 6). For example, the
393 CORRs of 200-hPa temperature variability range from −0.73 and 0.86, while the
394 CORRs of 850-hPa vector wind and 600-hPa specific humidity variability range from
395 0.01 to 0.93. The spatial patterns of SST (except in MCM-UA-1-0), SLP, 200-hPa
396 vector wind (except in MCM-UA-1-0) and 850-hPa temperature variabilities are
397 relatively better than for other variables, ranging from 0.6 and 0.98. Compared to the
398 climatological mean, the CMIP6 models show relatively poor ability in simulating the
399 spatial pattern of the interannual variability of large-scale environmental fields (Figs.
400 3, 6).

401 Most of the CMIP6 models overestimate the interannual variability of the
402 large-scale environmental fields. For example, more than 30 of the models show
403 positive biases in the amplitude of the interannual variability of 200-hPa air
404 temperature. Other variables, e.g., SST, 850-hPa and 200-hPa vector winds, and
405 600-hPa specific humidity, also show similar positive biases. NorESM2-MM
406 overestimates the amplitude of the interannual variability of SST (200-hPa air
407 temperature) by 1.2 (4.6) times. Moreover, the same series of models show similar
408 biases. For example, in terms of amplitude of the interannual variability, the
409 BCC-CSM2-MR and BCC-BCC-ESM1 models show relatively small biases for all

variables evaluated. Similarly, CESM2 and CESM2-FV2 as well as NorESM2-LM and NorESM2-MM show relatively greater positive biases.

The CMIP6 models show better skill with respect to the interannual variability in the southern Indian Ocean relative to the other ocean areas, as characterized by smaller MIEI values (Fig. 7). The models with high resolution also have certain advantages in their simulation of interannual variability, which is similar to what was found from the evaluation of climatological mean states in section 3.3.2. Among the top 10 CMIP6 models, there are 6–7 models with higher resolution. Similarly, among the bottom 10 models, there are 7–8 models with lower resolution. We do, however, find a few opposite results. For example, the low-resolution models of MPI-ESM1-2-LR and BCC-ESM1 rank higher than their corresponding high-resolution models (Fig. 7). Similar to Fig. 5, Fig. 8 ranks the 33 CMIP6 models by their MIEI values in five ocean regions, separately. The models that perform well in four ocean regions include MPI-ESM1-2-LR, MRI-ESM2-0, E3SM-1-0, and CAMS-CSM1-0. The models that perform well in three ocean regions are BCC-ESM1, CIESM, FGOALS-g3, FIO-ESM-2-0, KACE-1-0-G and MPI-ESM1-2-HR. Comparing Figs. 5 and 8, it can be seen that the NorESM-LM and NorESM-MM models perform well in terms of the climatology, but not so well in terms of the interannual variability; while BCC-ESM1 and CAMS-CSM1-0 show moderate performance in their simulation of the climatology, but good ability to simulate the interannual variability. According to the evaluation results of the 33 CMIP6 models, MPI-ESM1-2-LR, MPI-ESM1-2-HR and FIO-ESM-2-0 show

relatively good performance in simulating both the climatology and interannual variability of TC environmental fields.

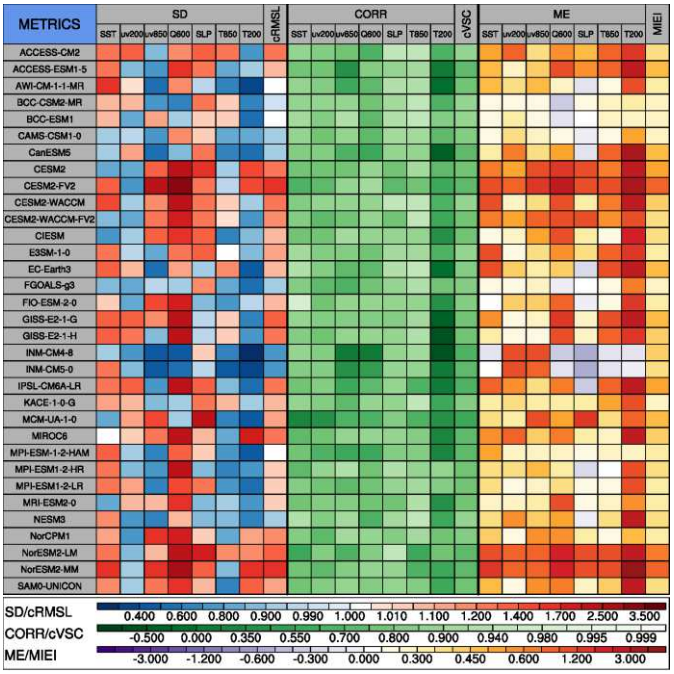


Fig. 6. Same as in Fig.3, except for the interannual variability of multiple variables. The interannual variability is measured by the temporal standard deviation

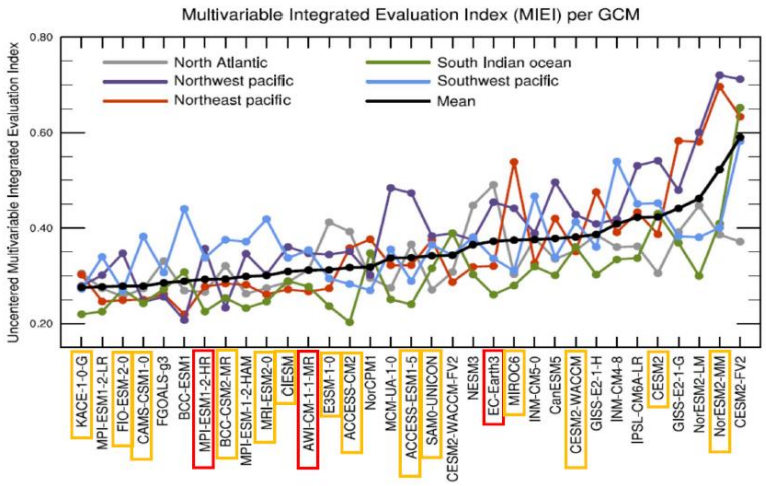


Fig. 7. MIEI values for 33 CMIP6 models in five ocean regions for the interannual variability. Red boxes on the abscissa indicate the high-resolution models (resolution less than 1°); the orange boxes indicate the

medium-resolution models (resolution of ~1°); and the rest are coarse-resolution models.

	NWP	NEP	NA	SIO	SWP
ACCESS-CM2	13	21	29	1	4
ACCESS-ESM1-5	25	18	23	7	5
AWI-CM-1-1-MR	11	7	16	15	15
BCC-CSM2-MR	2	12	17	11	20
BCC-ESM1	1	1	4	22	28
CAMS-CSM1-0	3	4	6	8	23
CanESM5	28	26	19	19	12
CESM2	30	24	14	32	30
CESM2-FV2	32	32	24	33	33
CESM2-WACCM	22	20	20	28	26
CESM2-WACCM-FV2	18	13	15	30	14
CIESM	15	8	10	17	10
E3SM-1-0	9	9	30	6	6
EC-Earth3	24	16	33	12	9
FGOALS-g3	4	6	18	14	7
FIO-ESM-2-0	12	3	1	13	1
GISS-E2-1-G	26	31	28	29	24
GISS-E2-1-H	20	28	26	20	17
INM-CM4-8	21	25	21	25	32
INM-CM5-0	19	19	25	24	31
IPSL-CM6A-LR	29	27	22	26	29
KACE-1-0-G	5	14	13	2	3
MCM-UA-1-0	27	17	9	10	16
MIROC6	23	29	12	16	8
MPI-ESM-1-2-HAM	10	11	2	5	19
MPI-ESM1-2-HR	14	10	3	3	11
MPI-ESM1-2-LR	6	2	7	4	13
MRI-ESM2-0	8	5	8	9	27
NESM3	16	15	32	21	22
NorCPM1	7	23	11	27	2
NorESM2-LM	31	30	31	18	21
NorESM2-MM	33	33	27	31	25
SAM0-UNICON	17	22	5	23	18

Fig. 8. Ranking of 33 CMIP6 models based on their MIEI values in five ocean regions for the inter-annual variability. The top 10 models are marked in blue.

4. Discussion and conclusion

In this paper, we have evaluated the abilities of 33 CMIP6 models to simulate TC environmental fields i.e., SST, SLP, high- and low-level wind speed, mid-level humidity, and high- and low-level temperature fields. These seven variables are closely related to the genesis and development of TCs; plus, they are also used as boundary conditions in RCMs to carry out dynamical downscaling–based simulations. Therefore, the performance of CMIP6 models in simulating these TC environmental variable fields can directly affect the implementation of dynamic downscaling (Holland et al., 2010; Bruyère et al., 2014).

Our results show that the MME of CMIP6 models shows a wet bias in the 600-hPa humidity relative to reanalysis data (JRA-55 and ERA5) in the tropical region from the equator to 15° in the Northern and Southern Hemisphere. Similar results were also noted by Camargo et al. (2007) and Tian et al. (2013) in CMIP5 models. However, they also pointed out that the excessive moisture of the assimilating numerical model may not indicate the real model biases for observations because the reference reanalysis data are themselves a product of the assimilating numerical model. Nevertheless, the humidity bias associated with a double-ITCZ in CMIP6 models is consistent with that in CMIP5 models (Lin, 2007; Liu et al., 2012). There is a cold bias in the upper troposphere for most CMIP5 models (Tian et al., 2013). From the results of CMIP6 models, the cold bias problem in the upper troposphere has changed in some models (36% of them) which show a warm bias in the Northwest Pacific Ocean.

The wind shear field simulated in the Southern Hemisphere is closer to the value of the reanalysis data than in the Northern Hemisphere. In the Northeast Pacific and North Atlantic, the vertical wind shear is generally larger than the value of the reanalysis data. To examine the impact of horizontal resolution on model performance, the 33 CMIP6 models are divided into three categories according to their resolution: high-resolution models (resolution less than 1°), medium-resolution models (resolution of ~1°), and coarse-resolution models (resolution of ~1.8°). In general, the high-resolution models perform relatively better than the low-resolution models, which indicates that an increase in resolution helps to improve model performance.

The model skill varies depending on the variable and region being evaluated. For example, BCC-CSM2-MR performs well in the climatology and interannual variability in the Northwest Pacific, but shows a moderate performance in other ocean regions. CMIP6 models generally show better skill in the South Indian Ocean than in other ocean regions. In terms of simulation of the climatology and interannual variability, the MPI-ESM1-2-LR, MPI-ESM1-2-HR and FIO-ESM-2-0 models show relatively high simulation skill in the global ocean areas. Furthermore, KACE-1-0-G and E3SM-1-0 perform well in all oceans.

As we know, the genesis and intensity of TCs are significantly affected by the environmental wind shear. A weak environmental wind shear is conducive to the generation and development of TCs. On the contrary, a stronger vertical wind shear tends to suppress the genesis of TCs. Previous studies (Camargo, 2013; Holland et al., 2010) have pointed out that CMIP5 models generate fewer TCs in the Northeast Pacific and North Atlantic Ocean in their simulations. According to the evaluation results of this study, the horizontal wind shear of the Northeast Pacific and North Atlantic is larger than that of the reanalysis data, which is probably an important reason why models underestimate the number of TCs in these two regions.

In addition, the mid-level humidity fields simulated by most of the 33 CMIP6 models are wetter than in the reanalysis data. A humid middle atmosphere is conducive to the genesis and development of TCs in tropical oceans. However, the overestimated vertical wind shear tends to suppress TC genesis. If one evaluates model performance simply based on a composite index, e.g., the GPI, one may obtain

misleading results owing to the cancellation between positive and negative biases. In the present study, we employ the MVIE method; and based on this method, we are able to rank GCMs based on their overall performances in simulating multiple variables such as temperature, pressure, humidity, and wind. We hope that our evaluation can provide guidance regarding the selection of GCMs for dynamical downscaling–based simulations of TCs.

Acknowledgments. We thank the climate modeling groups involved in CMIP6 for producing and making available their model outputs. The ERA5 data were provided by the European Centre for Medium-Range Weather Forecasts. The JRA-55 data were provided by the Japan Meteorological Agency. This study was supported jointly by the National Key Research and Development Program of China (2017YFA0603803) and the National Science Foundation of China (41675105, 41775075, 42075152). The study was also supported by the Jiangsu Collaborative Innovation Centre for Climate Change.

References

Bruyère, Cindy L., James M. Done, Greg J. Holland, Sherrie Fredrick (2014) Bias corrections of global models for regional climate simulations of high-impact weather. Clim dyn, 43, 1847-1856. https://doi.org/10.1007/s00382-013-2011-6

Camargo, S. J., K. A. Emanuel, and A. H. Sobel (2007a) Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. J. Climate, 20(19),4819-4834.

Camargo, S. J., A. H. Sobel, A. G. Barnston, and K. A. Emanuel (2007b) Tropical cyclone genesis potential index in climate models. Tellus(A), 59(4),428-443.

Camargo, S. J. (2013) Global and Regional Aspects of Tropical Cyclone Activity in the CMIP5 Models. J. Climate, 26: 9880–9902.

Chan J C L, J. E. Shi, K. S. Liu (2001) Improvements in the seasonal forecasting of tropical cyclone activity over the western North Pacific. Weather & Forecasting, 16: 491–498.

Chen, J., Z. Wang, C.-Y. Tam, N.-C. Lau, D.-S. D. Lau and H.Y. Mok (2020) Investigating climate change impacts on western North Pacific tropical cyclones and induced storm surges over the Pearl River Delta region based on pseudo-global warming experiments, Sci. Rep., 10: 1965. https://doi.org/10.1038/s41598-020-58824-8.

Elsberry, R. L. (1994) Global view of tropical cyclones.

Emanuel, K. A., D. S. Nolan (2004) Tropical cyclone activity and the global climate system. Preprints, 26th Conf. on Hurricanes and Tropical Meteorology, Miami, FL, Amer. Meteor. Soc., 40–241.

Emanuel, K. A., R. Sundararajan, and J. Williams (2008) Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. Bull. Amer. Meteor. Soc., 89:347–368, https://doi.org/10.1175/BAMS-89-3-347.

Emanuel K A (2010) Tropical cyclone activity downscaled from NOAA-CIRES Reanalysis, J Adv. Model Earth Syst., 1908-1958. https://doi：10.3894/JAMES.2010.2.1

Emanuel, K., & Center, L. (2020) Response of global tropical cyclone activity to increasing co2: results from downscaling cmip6 models. J. Climate, 34(1):1-54. https://doi：10.1175/JCLI-D-20-0367.1

Eyring, V., Bony, S., Meehl, G. A., Senior. C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model. Dev., 9: 1937-1958, https://doi: 10.5194/gmd-9-1937-2016.

Fan K, H. J. Wang (2009) A new approach to forecasting typhoon frequency over the western North Pacific. Weather & Forecasting, 24:974–986.

Gleckler, P. J., Taylor, K. E., and Doutriaux, C. (2008) Performance metrics for climate models, J. Geophys. Res., 113, D06104, https://doi.org/10.1029/2007JD008972.

Gray, W. M. (1979) Hurricanes: Their formation, structure and likely role in the tropical circulation. Meteorology over the Tropical Oceans, D. B. Shaw, Ed., Royal Meteor. Soc., James Glaisher House, Grenville Place, Bracknell, Berkshire, 155–218.

Gray,W.M. (1968) Global view of the origin of tropical disturbances and storms. Mon. Wea. Rev., 96:669-700.

Huang, F., Xu, Z. & Guo, W. (2020) The linkage between CMIP5 climate models' abilities to simulate precipitation and vector winds. Clim Dyn 54, 4953–4970. https://doi.org/10.1007/s00382-020-05259-6

Huang, F., Xu, Z. & Guo, W. (2019) Evaluating vector winds in the Asian-Australian monsoon region simulated by 37 CMIP5 models. Clim Dyn 53, 491–507. https://doi.org/10.1007/s00382-018-4599-z

Holland, G. J., J. Done, C. Bruyere, C. Cooper, and A. Suzuki (2010) Model investigations of the effects of climate variability and change on future Gulf of Mexico tropical cyclone activity. OTC Metocean.

Jiang, J. H., et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA 'A-Train' satellite observations, J. Geophys. Res., 117(D14), D14105, https://doi:10.1029/2011jd017237

Knutson, T. R. et al. (2013) Dynamical downscaling projections of twenty-first-century Atlantic hurricane activity: CMIP3 and CMIP5 model-based scenarios. J. Climate, 26: 6591–6617. https://doi：10.1175/JCLI-D-12-00539.1

Li, Xianzhi (1956) The comprehensive theory for typhoon genesis. Acta Meteorologica Sinica, 27:97—99.

Lin, J. L. (2007), The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis, J. Climate, 20(18):4497–4525, https://doi:10.1175/jcli4272.1.

Liu, H. L., M. H. Zhang, and W. Y. Lin (2012), An investigation of the initial development of the double-ITCZ warm SST biases in the CCSM, J. Climate, 25(1):140–155, https://doi:10.1175/2011jcli4001.1.

Liu, Z., C. S. Schwartz, C. Snyder, and S.-Y. Ha (2012) Impact of assimilating AMSU-A radiances on forecasts of 2008 Atlantic tropical cyclones initialized with a limited-area ensemble Kalman filter. Mon. Wea. Rev., 140:4017–4034, https://doi:10.1175/MWR-D-12-00083.1.

Lui, Y.S., K.-S. Tse, C.-Y. Tam, K.H. Lau and J. Chen (2020) Performance of western north Pacific tropical cyclone track and intensity evolution in MPAS-A and WRF simulations, Theor. Appl. Clim.,https://doi.org/10.1007/s00704-020-03444-5

Mei W, Kamae Y, Xie S P, et al. (2019) Variability and Predictability of North Atlantic Hurricane Frequency in a Large Ensemble of High-Resolution Atmospheric Simulations. J. Climate, 32(11).

Palmen.E. (1948) On the formation and structure of tropical hurricane，Geophysics， 4:26-38.

Riehl, H. (1948) On the formation of typhoon，J. Meteor., 5:247-264.

Riehl, H. (1950) A model of hurricane formation，J. App. Phy., 21:917-925.

Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vannière, B., et al. (2020). Projected future changes in tropical cyclones using the CMIP6 HighResMIP multimodel ensemble. Geophysical Research Letters, 47, e2020GL088662. https://doi.org/10.1029/2020GL088662

Su H, Jiang JH, Zhai C, and et al., (2013) Diagnosis of regime-dependent cloud simulation errors in CMIP5 models using “A-Train” satellite observations and reanalysis data. J Geophys Res Atmos 118(7):2762–2780. https://doi:10.1029/2012JD018575

Song, Y. J., L. Wang, X. Y. Lei, and X. D. Wang (2015) Tropical cyclone genesis potential index over the western North Pacific simulated by CMIP5 models. Adv. Atmos. Sci., 32(11), 1539–1550, https://doi: 10.1007/s00376-015-4162-3.

Takahashi H, Su H., Jiang J H. (2016) Error analysis of upper tropospheric water vapor in CMIP5 models using "A-Train" satellite observations and reanalysis data. Clim dyn, 46(9-10):2787-2803. https://doi：10.1007/s00382-015-2732-9

Tao L，Zhang Y F，Wang X B (2020) Improvement of genesis potential index for western North Pacific tropical cyclones. Trans. Atmos. Sci. ， 43(4): 603-616 ． https://doi: 10. 13878 /j．cnki．dqkxxb．20171228001

Taylor, K. E. (2001) Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res.-Atmos., 106: 7183–7192.

Tian, B., E. J. Fetzer, B. H. Kahn, J. Teixeira, E. Manning, and T. Hearty (2013), Evaluating CMIP5 Models using AIRS Tropospheric Air Temperature and Specific Humidity Climatology, J. Geophys. Res. Atmos., 118:114–134, https://doi:10.1029/2012JD018607

Villarini, G., and G. A. Vecchi (2012) Twenty-first-century projections of North Atlantic tropical storms from CMIP5 models. Nat. Climate Change, 2:604–607. https://doi：10.1038/nclimate1530

Villarini, G., and G. A. Vecchi (2013) Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. J. Climate, 26: 3231-3240. https://doi：10.1175/JCLI-D-12-00441.1

Wu, Wei, Jin-hua Yu (2011) Scenarios of 21st century tropical cyclone activity over western north pacific as projected by GFDL_RegCM. J. Tropical Meteorology, 27(6):843-852.

Xu Z. F., Z.-L. Yang (2012) An improved dynamical downscaling methods with GCM bias corrections and its validation with 30 years of climate simulations. J. Climate, 25, 6271-6286. https://doi.org/10.1175/JCLI-D-12-00005.1

Xu Z. F., Z.-L. Yang (2015) A new dynamical downscaling approach with GCM bias corrections and spectral nudging. J. Geophys. Res. Atmos, 120. https://doi:10.1002/2014JD022958.

Xu Z. F., Han, Y., Z.-L. Yang (2019) Dynamical downscaling of regional climate: A review of methods and limitations. Sci. China Earth Sci. 62:365–375. https://doi.org/10.1007/s11430-018-9261-5

Xu Z. F., Hou, Z., Han, Y., and Guo, W. (2016) A diagram for evaluating multiple aspects of model performance in simulating vector fields, Geosci. Model Dev., 9:4365–4380, https://doi.org/10.5194/gmd-9-4365-2016.

Xu Z. F., Han, Y., and Fu, C. (2017) Multivariable integrated evaluation of model performance with the vector field evaluation diagram, Geosci. Model Dev., 2017, 10:3805–3820, https://doi.org/10.5194/gmd-10-3805-2017. https://doi.org/10.5194/gmd-2017-95

Zhang, Y, H. Wang, J. Sun, and H. Drange （2010）Changes in the tropical cyclone genesis potential index over the western North Pacifc in the SRES A2 scenario. Adv. Atmos. Sci., 27. https://doi: 10.1007/x00376-010-9096-1.

614 Zhao, Junping，Wu, Liguang，Zhao, Haikun (2012) Improvement of tropical cyclone genesis potential
615 index in the western North Pacific Basin． J. Met. Sci., 32(6): 591-599
616