Evaluating Boreal Summer Circulation Patterns of CMIP6 Climate Models Over The Asian Region

Lulei Bu

Fudan University

Zhiyan Zuo (✉ zuozhy@fudan.edu.cn)

Fudan University   https://orcid.org/0000-0001-5708-723X

Ning An

China Academy of Meteorological Sciences

Abstract

Our confidence in future climate projection depends on the ability of climate models to simulate the current climate, and model performance in simulating atmospheric circulation affects the ability to simulate extreme events. This study uses the self-organizing map (SOM) method to evaluate the frequency, persistence, and transition characteristics of models in the Coupled Model Intercomparison Project Phase 6 (CMIP6) for different ensembles of the 500 hPa daily geopotential height (Z500) in Asia, and then ranks all ensembles according to a comprehensive ranking metric (MR). Our results show that the SOM method is a powerful tool for assessing the daily-scale circulation simulation skills in Asia, and the results are not significantly affected by different map sizes. Positive associations between the performance of ensembles for any two of frequency, persistence, and transition were found, indicating that an ensemble that performs well for one metric is good for the others. The results of the MR ranking show that the r10i1p1f1 ensemble of CanESM5 gives the best overall simulation of 500 hPa circulation in Asia, and this is also the ensemble that best simulates frequency characteristics. The MR simulation skills of the 10 best ensembles for the position of the Western North Pacific Subtropical High (WNPSH) are far better than those of the 10 worst. Such differences may lead to errors in the simulation of extreme events. This study will help future studies in the choice of ensembles with higher circulation simulation skills to improve the credibility of their conclusions.

1. Introduction

The general circulation of the atmosphere is a key factor affecting climate variation, whether on global or regional scales, because it drives the circulation of energy and water vapor (Maidens et al. 2021; Zhao C. et al. 2019). With the intensification of global warming, various extreme events have occurred frequently and have had a great impact on society and the environment (Gu et al. 2016; Kong et al. 2020; Sales et al. 2018). Some extreme events, such as extreme high temperature and extreme precipitation, are particularly affected by the atmospheric general circulation (Boschat et al. 2014; Fischer et al. 2010; Gibson et al. 2017; Liu et al. 2015; Loikith et al. 2019; Lu et al. 2020; Ohba et al. 2020). The atmospheric circulation at 500 hPa is of great importance because it presents a strong relationship between higher level circulation and surface variables (Gao et al. 2019; Horton et al. 2015; Mioduszewski et al. 2016).

Global climate models (GCMs) are powerful tools for simulating the current climate and predicting future climate change (Wang et al. 2015). The Sixth Coupled Model Intercomparison Project Phase 6 (CMIP6) was initiated by the World Climate Research Programme (WCRP), with the purpose of answering new scientific questions facing the field of climate change and providing data support to achieve the scientific goals established by the WCRP 'Grand Challenge' plan. The CMIP6 includes about 112 climate models from 33 institutions around the world participating in 23 sub-programs. These data will support the next 5 to 10 years of global climate research (Eyring et al. 2016). The evaluation of CMIP6 climate models is an important requirement for further research on downscaling and projection. However, most of the current assessments of CMIP6 are for surface variables, such as temperature and precipitation (Almazroui et al. 2020). Assessments of the circulation in Asia are rare.

There are many methods for evaluating model circulation, but their main purpose is always downscaling. The mainstream downscaling methods include principal component analysis (PCA), empirical orthogonal function (EOF) analysis, K-means clustering, and the self-organizing map (SOM) method. Previous studies have shown that the PCA and EOF methods are not accurate enough and not very intuitive when evaluating model circulation patterns (Wang et al. 2015). K-means clustering is an effective circulation classification method (Agel et al. 2017), but its biggest drawback is that it can only represent some discrete atmospheric systems and cannot organize them into a continuum (Gao et al. 2019). However, the actual atmospheric circulation develops continuously.

The SOM method solves these issues. The SOM is an unsupervised neural network algorithm that maps high-dimensional input data to a two-dimensional space. In the process of iteration, it not only updates the winning node, but also updates its neighboring nodes according to the weight. The SOM method was first proposed by Kohonen (1998), and first applied by Hewiston and Crane (2006) in the field of climate downscaling, and has since been widely used in the field of climate research. It can organize long time sequences of atmospheric circulation into a continuum, so that not only can the characteristics of a specific circulation type be seen, but also how this circulation might develop, because one particular node tends to change from its neighboring nodes. This method is effective in connecting abnormal atmospheric circulation patterns and extreme high temperature and precipitation events (Agel et al. 2017; Loikith et al. 2017).

Some previous studies have used the SOM method to evaluate the ability of CMIP5 models to simulate weather-scale circulation patterns (Cassano et al. 2007; Wang et al. 2015). The circulation simulation capabilities of models can be ranked by comparing the correlation coefficient or root mean square error (Mioduszewski et al. 2016) between models and reanalysis products for the frequency, persistence, and transition metrics (Gibson et al. 2016). Models with better simulation capabilities for one characteristic (such as frequency) of circulation tend to be better for other characteristics (persistence and transition; (Gibson et al. 2016). These research results provide a basis for studying the causes of extreme events and for scenario projections.

The Western North Pacific Subtropical High (WNPSH) is the most important circulation system that affects the summer in Asia at 500 hPa. Its position, shape, and strength dominate the climate of Asia (Zhang et al. 2020; Zhao C. et al. 2019). Monsoon and typhoon activities over the western Pacific are closely related to the WNPSH (Chen et al. 2020). For example, water vapor from the tropical ocean is transported to China around the western ridge of the WNPSH. The convergence of warm humidity and cold air from high latitudes means that rainfall often occurs on the northwestern edge of the WNPSH (Liu et al. 2014; Preethi et al. 2017). Therefore, after the overall assessment of model simulation of the circulation in Asia, it is necessary to test the skill in simulating the WNPSH.

This study has two main objectives. First, to use the SOM method to evaluate the climate models of CMIP6 based on three metrics for frequency, persistence, and transition and to obtain rankings from the

performance of these different aspects and a comprehensive ranking metric (MR). Second, to check whether the top-ranked model also gives a better simulation of the WNPSH.

This paper is organized as follows: Sect. 2 introduces the data and methodology. Section 3 presents the rankings of CMIP6 models from the performance of different metrics. Section 4 checks the relationship between SOM-based model rankings and WNPSH performance. A discussion is presented in Sect. 5 and the overall conclusions in Sect. 6.

2. Data And Methods

2.1 Reanalysis data

In this study, the geopotential height at the 500 hPa isobaric level (Z500) of the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v5 (ERA5) was mainly used, and data from the Japanese Meteorological Agency (JMA) Reanalysis (JRA-55) were also used to test whether different reanalysis products gave different results (Ebita et al. 2011). ERA5 is based on the Integrated Forecasting System (IFS) Cy41r2 and has benefited from long-term development in model physics, core dynamics and data assimilation (Hersbach et al. 2020). Compared with the previous generation JMA reanalysis, JRA-55 uses a more advanced data assimilation scheme, increased model resolution, a new variational bias correction for satellite data, and several additional observational data sources (Ebita et al. 2011). we use the absolute Z500 instead of the Z500 anomaly because the absolute Z500 has a clearer physical meaning and is more intuitive, whereas using the Z500 anomaly to train the SOM may mix different circulation regimes (An et al. 2021). When we started the study, ERA5 was available from 1979, and considering that the historical experiment of most of the CMIP6 models ended in 2014, we chose 1979 to 2014 as the research period. We consider only boreal summer (June to August) and calculate the daily average for every 6 hours (0000, 0600, 1200, and 1800 coordinated Universal Time (UTC)), so the time range is 3312 days in total. The analysis domain extends from 40 to 180 °E and from 0 to 60°N, following the suggestion from Gao et al. (2019) that the climate of China is affected by the circulation in this region.

2.2 Climate model data

We analyze the absolute Z500 of 140 ensembles from 23 selected climate models participating in CMIP6 (Table 1). Compared with CMIP5, CMIP6 has more institutions and models participating, and also has more sub-experiments. CMIP6 aims to provide a scientific basis for the Intergovernmental Panel on Climate Change (IPCC) 6th Assessment Report (AR6). The Historical experiment is started from the model state of the piControl experiment at a certain time but is driven by various external forcings based on observations that have changed over time since 1850 (Zhou et al. 2019), and the piControl experiment is an experiment that maintains the external forcing (e.g., greenhouse gas, solar radiation, aerosol, land use) at the level of 1850 to drive the global coupled model for long-term integration of more than 500 years.

Table 1

Basic information on the models used in this study, including model identification, originating center, atmospheric model resolution, and ensemble numbers.

Model	Institution	Resolution	Ensemble numbers
ACCESS-CM2	Commonwealth Scientific and Industrial Research Organisation, Australia	250 km	1
AWI-ESM-1-1-LR	Commonwealth Scientific and Industrial Research Organisation, Germany	250 km	1
BCC-CSM2-MR	Beijing Climate Center, China Meteorological Administration, China	100 km	1
BCC-ESM1	Beijing Climate Center, China Meteorological Administration, China	250 km	1
CanESM5	Canadian Centre for Climate Modeling and Analysis, Canada	500 km	25
CESM2	National Center for Atmospheric Research, United States	100 km	10
CESM2-WACCM	National Center for Atmospheric Research, United States	100 km	3
CESM2-WACCM-FV2	National Center for Atmospheric Research, United States	100 km	3
CNRM-CM6-1-HR	Centre National de Recherches Météorologiques, France	50 km	1
CNRM-ESM2-1	Centre National de Recherches Météorologiques, France	250 km	3
FGOALS-f3-L	Institute of Atmospheric Physics, Chinese Academy of Sciences, China	100 km	3
GFDL-CM4	Geophysical Fluid Dynamics Laboratory, USA	250 km	1
INM-CM4-8	Institute for Numerical Mathematics, Russia	100 km	1
INM-CM5-0	Institute for Numerical Mathematics, Russia	100 km	10
IPSL-CM6A-LR	Institut Pierre-Simon Laplace, France	250 km	32
MIROC6	National Institute for Environmental Studies, The University of Tokyo, Japan	250 km	10
MPI-ESM-1-2-HAM	Max Planck Institute for Meteorology, Germany	250 km	2
MPI-ESM1-2-HR	Max Planck Institute for Meteorology, Germany	100 km	10

Model	Institution	Resolution	Ensemble numbers
MPI-ESM1-2-LR	Max Planck Institute for Meteorology, Germany	250 km	10
MRI-ESM2-0	Meteorological Research Institute, Japan	100 km	5
NorESM2-LM	Norwegian Climate Centre, Norway	250 km	3
NorESM2-MM	Norwegian Climate Centre, Norway	100 km	3
TaiESM1	Research Center for Environmental Changes, Taiwan, China	100 km	1

2.3 Methodology

In this study, the SOM method is used to classify the 500 hPa atmospheric circulation characteristics in Asia. Each node classified represents a type of circulation pattern. The input data are the daily absolute Z500 interpolated onto an Equal-Area Scalable Earth-type (EASE) grid at a spatial resolution of 2.5° × 2.5° to ensure that the correct distance is calculated by the SOM procedure (Gibson et al. 2017). Each node first initializes its parameters (i.e., weight coefficient), and the number of parameters of each node is the same as the dimension of the input data. Then, for each input data field, the closest node is found according to the distance function (such as Euclidean distance), and the node with the smallest distance is the 'winning node'. After finding the 'winning node', its neighboring nodes are updated, and after updating the weight coefficient of the node, the weight vector of each node is updated according to the gradient descent method. The algorithm iterates until it converges or meets the termination condition set by the user.

This study uses the second version of the SOM toolbox on MATLAB developed by Helsinki University of Technology (http://www.cis.hut.fi/projects/somtoolbox/). The 'hexagonal' lattice and 'sheet' map shape are set up on the topological structure, the initialization scheme is the random initialization scheme, and other settings are the default settings (Jiang et al. 2015; Liu et al. 2006). Gibson et al. (2017) has shown that allowing the neighborhood radius of the first 50% of training iterations to decrease from 5 to 1, and keeping the neighborhood radius at 1 for the last 50% of training iterations gives a SOM suitable for classifying a greater number of different weather types. Therefore, in this study, a total of 1000 trainings iterations were performed during the training of the SOM. The neighborhood radius of the first 500 trainings decreased from 5 to 1, and the neighborhood radius of the last 500 trainings was kept at 1.

One of the most important steps in the SOM approach is determining the number of nodes. Too few SOM nodes may cause the classification results to confuse different circulation features, and too many SOM nodes will cause the classification results to share similar features (Ford et al. 2017; Loikith et al. 2019; Mioduszewski et al. 2016). To determine the best map size for this study, we experimented with 11 different map sizes, including 2 × 3, 3 × 4, 4 × 4, 4 × 5, 5 × 5, 5 × 6, 5 × 7, 6 × 6, 6 × 7, 7 × 7, and 7 × 8 (Fig. 1). For each configuration, we calculated the spatial correlation coefficient between the actual value

of each day and the value of the assigned node to quantitatively evaluate the difference between different map sizes.

We call the SOM result obtained from ERA5 the 'Master SOM' (Fig. 2), and then map each day onto the 'Master SOM' for each ensemble. Mapping is accomplished by finding the node in the 'Master SOM' that is closest to (i.e., has the smallest squared distance from) the daily state of the ensemble (Cassano et al. 2007). Three metrics were used to evaluate the CMIP6 models from the SOM derived nodes: frequency, persistence and transition metrics (Gibson et al. 2016). Node frequency refers to the probability of occurrence of this node, and the frequency metric is defined by dividing the number of days the node appears by the total number of days. The 95% confidence level for the node frequency of occurrence is given by

$$p \pm 1.96\sqrt{p(1-p)/l} \tag{1}$$

where l represents the number of days, and p represents the node frequency for a randomly derived data set. In other words, $p = 1/N$ for an SOM with N nodes. Node frequency is considered significant if it exceeds these limits. For the master SOM with 20 nodes the expected frequency of occurrence for each node is 0.05 with a 95% confidence interval of ± 0.74%. Node persistence represents the continuity of a node. In this study, the persistence metric is expressed by dividing the number of events that lasted more than two days at this node by the number of events that lasted only one or two days at this node. The transition metric refers to the transition probability from one node to another. Each node can transit to nodes other than itself, so the SOM with N nodes can have $N(N-1)$ kinds of transitions in total.

The performance of different ensembles on different metrics is described by calculating the Pearson correlation coefficient between different ensembles and ERA5. For example, the frequency performance of a particular ensemble is represented by the correlation coefficient of the frequency metrics between this ensemble and ERA5 for all nodes. The higher the correlation coefficient, the better the frequency performance of the ensemble, and the higher the frequency ranking. After ranking according to different metrics, the comprehensive rankings can be obtained according to the ranking metric (MR), which is defined as

$$MR = 1 - \frac{1}{nm}\sum_{i=1}^{n} rank_i \tag{2}$$

where m is the number of ensembles, and n is the number of metrics. The closer the MR value is to 1, the higher the comprehensive performance of the ensemble (Li et al. 2015).

In order to describe the WNPSH, we calculated the western ridge point index and the northern boundary position index for ERA5 and each ensemble. The western ridge point index refers to the longitude of the westernmost position of the 5880-gpm isobar at the height of 500 hPa in the domain 10 to 60 °N, and 90 to 180 °E. The northern boundary position index refers to the average latitude of the 5880-gpm isobar on

the north side of the subtropical high along each meridian at the height of 500 hPa and in the same domain (Liu et al. 2019).

3. Results

3.1 Determination of SOM map size

In order to determine the optimal number of nodes, we tested 11 sets of SOM configurations with different map sizes, and calculated the spatial correlation coefficient between the actual Z500 and the assigned nodes, as shown in Fig. 1. As the number of nodes increases, the spatial correlation coefficient also increases. However, a larger number of nodes may produce nodes that share similar circulation characteristics and a smaller number of nodes may confuse different circulation characteristics. According to most research on Asia, the number of SOM nodes ranges from 12 to 20 (Gao et al. 2019; Li M. et al. 2020; Liu et al. 2015; Wang et al. 2015). If we search to minimize the quantization error within groups and to maximize the topographic error between groups, 20 (4 × 5) nodes are optimal (Li M. et al. 2020; Yin et al. 2010).

3.2 Master SOM

Figure 2 shows the SOM map obtained using the daily absolute Z500 from ERA5 with the aforementioned SOM settings. In this Master SOM, node 'a1' refers to the node in the top-left corner of the SOM plane and node 'e4' refers to the node in the bottom-right corner. The orange shading in nodes a3, a4, b3, b4, c3, c4, d3, d4, e2, e3, and e4 to the east highlights the area with geopotential height exceeding 5880 gpm, where the 5880-gpm isobar is considered the boundary of the WNPSH (Liu et al. 2019). The orange shading of nodes d2, e1, a3, a4, b3, b4, c3, c4, d3, d4, e2, e3, and e4 to the west indicates the North Africa High (NAH), which often changes with the movement of the WNPSH (La et al. 2002).

According to the Master SOM, neighboring nodes share similar circulation characteristics, and the difference along the diagonal between node a1 and node e4 is the largest. The circulation pattern shown at node a1 is characterized by the East Asian Trough (EAT) near the Kamchatka Peninsula in middle and high latitudes and the subtropical high represented by the 5840-gpm isobar at middle and low latitudes. Cold advection from the west of the EAT transports cool air from the north to Asia, alleviating the high temperature in summer. The circulation pattern shown at node e4 is more conducive to generating high temperature. The WNPSH extends westward to southern China, and its intensity reaches 5880-gpm, whereas the westerly jets in the middle to high latitudes are stable and block the cold air from northern latitudes from moving south, leading to the development of high temperatures in Asia. If this type of circulation pattern lasts for a long time, coupled with unsupported regional soil moisture conditions, persistent high temperatures occur (Boschat et al. 2014; Ding et al. 2011). From node a1 to e4, the EAT gradually weakens, but the WNPSH continues to expand westward and southward, which gradually has a greater impact on the Asian weather. The NAH also gradually intensifies from node a1 to e4, but its position does not change much. In addition, the existence of the strong WNPSH of node e4 is conducive

to the transport inland of water vapor from the Pacific Ocean, which favors heavy rainfall in southern China, Japan and South Korea.

The node frequency of the Master SOM shows the frequency of occurrence of each circulation pattern of ERA5 in all 3312 days (Fig. 3 (a)); node frequencies that are statistically (at the 95% level) above or below those expected by chance are shown in red. The frequencies of nodes a1 and e4 are the highest, both exceeding 10%. The frequencies of nodes a2, b2, b3, d2, and e3 are the lowest, only ~ 2%, suggesting these are transition circulation patterns that appear less frequently. It shows that the SOM method is effective in identifying weather patterns that occur less frequently. The frequencies of nodes a3, a4, b1, b4, c1, d4, and e1 are close to the random probability of 5% and are not statistically significant at the 95% level.

The node persistence represents the duration of each node, and the solid black line in Fig. 4 shows the continuity of each node in the Master SOM. The persistence metric of most nodes is greater than 1 except for nodes a2, b2, b3, d2, and e3, suggesting that, for most nodes, the number of events lasting more than two days is greater than the number of events lasting two days or less. Nodes a2, b2, b3, d2, and e3 have more events with a duration of two days or less. These nodes also have the lowest frequencies (~ 2%; Fig. 3 (a)), and on average appear once or twice in summer each year. Nodes a1 and e4 are the most persistent, and the frequencies of these two nodes are the highest (Fig. 3 (a)), indicating that the circulation patterns corresponding to these two nodes are the two dominant circulation patterns in summer from 1979 to 2014.

Node transition is represented by the probability of a node transitioning to another. Figure 5 (a) shows the transition probability of each node in the Master SOM. Generally speaking, a node tends to transfer to its neighboring nodes. This is why the high probability values are concentrated near the diagonal from a1 to e4. The transition probability of different nodes and the circulation pattern corresponding to each node in the Master SOM (Fig. 2) can be combined to analyze the physical explanation of the change of different circulation patterns. Node e4 has the highest probability of transferring to nodes d4, e2, and e3, with a total probability of 0.86, which shows that the position of the WNPSH gradually retreats eastward and northward at low and middle latitudes. However, the probability of node e4 transferring to other nodes is almost zero. The transition probability of node d4 to nodes c4 and e3 is higher, at 0.31 and 0.25, respectively. This means that if the WNPSH is the same shape as in node d4, it is easier for it to shrink (to nodes c4 and e3), but almost never to expand (to node e4), Interestingly, the transition probability of node e4 to d4 is much higher than that of d4 to e4. This shows that although a system is not final and may change in intensity, the probability of changing opposite ways is different, at least for the most important system (the WNPSH) in this study region. Among the transition metrics of all nodes, that from node b4 to a4 is the largest, with a value of 0.47, followed by the transition metric from node a4 to a3, at 0.43. The transition metrics of node b3 to other nodes (nodes a2, a3, a4, b2, b4) are similar but not high (~ 0.1 in each case).

3.3 Ensemble ranking of different metrics

We calculated the correlation coefficients between all ensembles shown in Table 1 and ERA5 for the frequency, persistence, and transition metrics to evaluate the performance skills of each ensemble on these three different metrics. Based on this, we obtain three sets of rankings for all ensembles (Table 2). Our research found that there is a fairly strong link between the performance for different metrics, as shown in Fig. 6. Positive associations between any two of frequency performance, persistence performance and transition performance were found, indicating that good simulation skill for the frequency characteristic is related to skills in simulating persistence and transition characteristics of circulation. There are, however, a few exceptions. For example, the persistence performance of the r23i1p1f1 ensemble of IPSL-CM6A-LR is good, with a correlation coefficient of 0.97 with ERA5 and a ranking of 2 for persistence metric (Table 2). However, the transition performance of this ensemble is relatively weaker, with a correlation coefficient of 0.40 with ERA5 and a ranking of 121 for transition metric (Table 2). Figure 6 shows that overall the relationship between frequency performance and transition performance is the strongest, with a correlation coefficient of 0.57, while the correlation coefficient between persistence and transition performance is 0.50, and that between frequency and persistence performance is 0.45. These have all passed the 99% significance test. As a consequence, we calculate the MR metric considering all three different metrics, and the final ensemble rankings are based on the MR metric.

Table 2

Ensemble rankings of CMIP6 based on frequency performance, persistence performance, transition performance, and MR.

Model	Ensemble	Frequency ranking	Persistence ranking	Transition ranking	MR ranking
CanESM5	r10i1p1f1	1	5	21	1
NorESM2-MM	r1i1p1f1	19	1	13	2
CanESM5	r19i1p1f1	13	4	22	3
CESM2-WACCM-FV2	r2i1p1f1	4	14	25	4
NorESM2-LM	r2i1p1f1	11	31	6	5
CanESM5	r23i1p1f1	3	49	2	6
CanESM5	r1i1p1f1	17	36	3	7
CanESM5	r18i1p1f1	2	20	35	8
MPI-ESM1-2-LR	r5i1p1f1	23	7	29	9
CanESM5	r16i1p1f1	9	21	36	10
CanESM5	r12i1p1f1	29	35	4	11
CanESM5	r24i1p1f1	20	9	40	12
CanESM5	r17i1p1f1	28	33	9	13
CanESM5	r4i1p1f1	16	23	38	14
MPI-ESM1-2-LR	r4i1p1f1	26	10	41	15
MPI-ESM1-2-LR	r10i1p1f1	45	22	11	16
NorESM2-LM	r1i1p1f1	10	53	23	17
NorESM2-LM	r3i1p1f1	7	30	49	18
CanESM5	r22i1p1f1	33	6	48	19
MPI-ESM1-2-LR	r9i1p1f1	15	59	15	20
BCC-ESM1	r1i1p1f1	12	13	65	21
CanESM5	r21i1p1f1	50	3	39	22
CESM2-WACCM-FV2	r1i1p1f1	14	46	34	23
CanESM5	r9i1p1f1	21	42	32	24
MPI-ESM1-2-LR	r8i1p1f1	67	32	1	25

Model	Ensemble	Frequency ranking	Persistence ranking	Transition ranking	MR ranking
MPI-ESM1-2-LR	r2i1p1f1	68	8	28	26
CanESM5	r15i1p1f1	8	61	42	27
CanESM5	r7i1p1f1	69	38	5	28
CanESM5	r25i1p1f1	27	41	50	29
CESM2-WACCM-FV2	r3i1p1f1	18	72	30	30
MPI-ESM1-2-LR	r1i1p1f1	62	15	43	31
CanESM5	r13i1p1f1	34	28	58	32
NorESM2-MM	r2i1p1f1	24	43	57	33
BCC-CSM2-MR	r1i1p1f1	6	58	61	34
CanESM5	r5i1p1f1	61	12	53	35
CanESM5	r6i1p1f1	54	17	56	36
MPI-ESM1-2-LR	r3i1p1f1	44	70	17	37
CESM2-WACCM	r2i1p1f1	38	18	75	38
MIROC6	r10i1p1f1	113	11	8	39
CanESM5	r8i1p1f1	5	75	55	40
CESM2	r10i1p1f1	51	44	45	41
MPI-ESM1-2-HR	r8i1p1f1	31	25	87	42
GFDL-CM4	r1i1p1f1	76	48	20	43
MRI-ESM2-0	r5i1p1f1	74	29	44	44
CanESM5	r20i1p1f1	71	24	52	45
CanESM5	r11i1p1f1	32	89	27	46
MPI-ESM1-2-LR	r6i1p1f1	43	60	46	47
MRI-ESM2-0	r1i1p1f1	66	51	33	48
ACCESS-CM2	r1i1p1f1	77	19	54	49
CanESM5	r14i1p1f1	22	83	47	50
MPI-ESM1-2-LR	r7i1p1f1	65	57	31	51
CESM2	r2i1p1f1	49	40	68	52

Model	Ensemble	Frequency ranking	Persistence ranking	Transition ranking	MR ranking
CESM2-WACCM	r1i1p1f1	39	50	73	53
MRI-ESM2-0	r4i1p1f1	73	54	37	54
CanESM5	r2i1p1f1	72	34	59	55
NorESM2-MM	r3i1p1f1	25	80	62	56
MPI-ESM1-2-HR	r1i1p1f1	30	52	88	57
MIROC6	r9i1p1f1	127	37	14	58
CanESM5	r3i1p1f1	64	63	51	59
MIROC6	r6i1p1f1	119	55	7	60
MRI-ESM2-0	r3i1p1f1	70	56	60	61
MPI-ESM1-2-HR	r3i1p1f1	37	81	74	62
MPI-ESM1-2-HR	r6i1p1f1	36	99	67	63
MIROC6	r3i1p1f1	114	77	16	64
CESM2	r4i1p1f1	63	67	78	65
IPSL-CM6A-LR	r23i1p1f1	87	2	121	66
MIROC6	r2i1p1f1	125	68	24	67
CESM2	r7i1p1f1	52	76	89	68
MIROC6	r1i1p1f1	120	73	26	69
CESM2	r6i1p1f1	57	85	80	70
CESM2	r1i1p1f1	47	108	69	71
MPI-ESM1-2-HR	r5i1p1f1	41	114	70	72
MPI-ESM1-2-HR	r4i1p1f1	35	107	83	73
AWI-ESM-1-1-LR	r1i1p1f1	75	86	66	74
CESM2	r9i1p1f1	55	100	72	75
IPSL-CM6A-LR	r26i1p1f1	94	27	106	76
TaiESM1	r1i1p1f1	56	111	63	77
MPI-ESM1-2-HR	r9i1p1f1	40	106	84	78
MIROC6	r7i1p1f1	136	90	10	79

Model	Ensemble	Frequency ranking	Persistence ranking	Transition ranking	MR ranking
MPI-ESM1-2-HR	r10i1p1f1	42	113	82	80
IPSL-CM6A-LR	r3i1p1f1	84	39	114	81
MPI-ESM1-2-HR	r7i1p1f1	60	102	76	82
MIROC6	r5i1p1f1	126	96	19	83
CESM2-WACCM	r3i1p1f1	48	124	71	84
CESM2	r3i1p1f1	46	123	77	85
CESM2	r8i1p1f1	53	115	79	86
MRI-ESM2-0	r2i1p1f1	110	74	64	87
MIROC6	r8i1p1f1	123	116	12	88
IPSL-CM6A-LR	r7i1p1f1	79	71	101	89
MIROC6	r4i1p1f1	128	109	18	90
IPSL-CM6A-LR	r13i1p1f1	85	64	109	91
IPSL-CM6A-LR	r28i1p1f1	91	62	110	92
IPSL-CM6A-LR	r10i1p1f1	88	65	112	93
MPI-ESM1-2-HR	r2i1p1f1	58	125	85	94
CESM2	r5i1p1f1	59	119	91	95
INM-CM5-0	r9i1p1f1	133	16	125	96
IPSL-CM6A-LR	r6i1p1f1	108	26	140	97
IPSL-CM6A-LR	r4i1p1f1	80	92	104	98
IPSL-CM6A-LR	r19i1p1f1	96	66	118	99
IPSL-CM6A-LR	r11i1p1f1	78	79	128	100
IPSL-CM6A-LR	r5i1p1f1	109	47	129	101
CNRM-CM6-1-HR	r1i1p1f2	82	78	131	102
CNRM-ESM2-1	r2i1p1f2	124	87	81	103
IPSL-CM6A-LR	r2i1p1f1	93	69	135	104
IPSL-CM6A-LR	r1i1p1f1	98	95	108	105
IPSL-CM6A-LR	r27i1p1f1	81	118	103	106

Model	Ensemble	Frequency ranking	Persistence ranking	Transition ranking	MR ranking
IPSL-CM6A-LR	r24i1p1f1	104	84	115	107
IPSL-CM6A-LR	r8i1p1f1	83	101	120	108
IPSL-CM6A-LR	r12i1p1f1	107	103	97	109
INM-CM5-0	r10i1p1f1	139	45	124	110
IPSL-CM6A-LR	r14i1p1f1	99	82	130	111
IPSL-CM6A-LR	r25i1p1f1	111	88	113	112
FGOALS-f3-L	r1i1p1f1	118	104	95	113
IPSL-CM6A-LR	r31i1p1f1	92	126	102	114
INM-CM5-0	r6i1p1f1	135	94	98	115
INM-CM5-0	r7i1p1f1	129	91	107	116
IPSL-CM6A-LR	r16i1p1f1	86	122	119	117
MPI-ESM-1-2-HAM	r1i1p1f1	116	121	93	118
CNRM-ESM2-1	r1i1p1f2	122	120	90	119
IPSL-CM6A-LR	r20i1p1f1	102	93	137	120
IPSL-CM6A-LR	r21i1p1f1	89	105	139	121
FGOALS-f3-L	r2i1p1f1	115	128	94	122
CNRM-ESM2-1	r3i1p1f2	112	135	92	123
MPI-ESM-1-2-HAM	r2i1p1f1	121	133	86	124
INM-CM5-0	r8i1p1f1	130	110	100	125
INM-CM5-0	r3i1p1f1	131	97	116	126
IPSL-CM6A-LR	r17i1p1f1	90	129	127	127
INM-CM5-0	r4i1p1f1	132	98	117	128
FGOALS-f3-L	r3i1p1f1	117	137	96	129
IPSL-CM6A-LR	r18i1p1f1	101	117	132	130
IPSL-CM6A-LR	r9i1p1f1	95	139	123	131
IPSL-CM6A-LR	r32i1p1f1	105	134	126	132

Model	Ensemble	Frequency ranking	Persistence ranking	Transition ranking	MR ranking
IPSL-CM6A-LR	r15i1p1f1	97	132	136	133
INM-CM5-0	r5i1p1f1	134	112	122	134
INM-CM5-0	r2i1p1f1	137	127	105	135
IPSL-CM6A-LR	r22i1p1f1	100	138	133	136
IPSL-CM6A-LR	r30i1p1f1	103	131	138	137
IPSL-CM6A-LR	r29i1p1f1	106	136	134	138
INM-CM4-8	r1i1p1f1	140	140	99	139
INM-CM5-0	r1i1p1f1	138	130	111	140

Frequency performances for the highest and lowest ranked ensembles are given in Fig. 3 (b) and (c). The r10i1p1f1 ensemble of CanESM5 is the best ensemble for frequency performance. The frequencies of diagonal nodes a1 and e4 are very high, while some transitional nodes in the middle, such as nodes b3, c3, d3, and e3 have very low frequencies. The correlation coefficient between this ensemble and ERA5 reaches 0.73. However, this ensemble also obviously overestimates the frequency of some nodes, such as a3, a4, e1, and e2, and underestimates others, such as b1 and b2. What is interesting is that nodes with a more westward and southward WNPSH appear more frequently in the ensemble that has the highest frequency performance. The r1i1p1f1 ensemble of INM-CM4-8 has the lowest frequency performance, and the correlation coefficient between this ensemble and ERA5 is only 0.21. From Fig. 3 (c), almost all of the Z500s of this ensemble fall in nodes a1, b1, c1, d1, and e1, indicating that this ensemble has a systematic negative bias in the simulation of Z500, which causes the Z500 in this ensemble to be allocated to those nodes with weaker Z500 circulation.

Figure 4 shows the highest and lowest ranked ensembles for persistence performance. The r1i1p1f1 ensemble of NorESM2-MM is the highest performing ensemble for persistence simulation, with a correlation coefficient of 0.97 with ERA5. The persistence metrics change of each node in this ensemble is almost exactly the same as that of ERA5, but the persistence metrics of all nodes except d4 are lower than those of ERA5. This shows that in the r1i1p1f1 ensemble of NorESM2-MM, there are fewer events with a duration of more than two days, and more events with a duration of less than two days for each node. The r1i1p1f1 ensemble of INM-CM4-8 has the lowest persistence performance, with correlation coefficient of 0.39 with ERA5. It is also the lowest ranked ensemble for node frequency performance. The persistence metrics of this ensemble in some nodes are very high (nodes a1, b1, c1, d1, and e1), but in other nodes are zero (nodes a2, a3, b4, d1, e1, c3, c4, d4, and e4), which leads to the strange oscillation of the blue curve in Fig. 4. Compared with Fig. 3, it is easy to understand that some nodes do not have even one day allocated to them, so the persistence metric of these nodes is of course 0, whereas the persistence metric of other nodes is very high.

Transition performance represents the ability of CMIP6 to simulate the intensity and location changes of the weather in the region. Figure 5 (b) and (c) show the highest and lowest ranked ensembles for transition performance. The r8i1p1f1 ensemble of MPI-ESM1-2-LR is the best ensemble for transition performance, and the correlation coefficient with ERA5 is 0.86. The node transition distribution of this ensemble (Fig. 5(b)) is similar to that of ERA5 (Fig. 5(a)), but the transition probability is lower overall. This may be because this ensemble's persistence metrics are larger overall. That is, the persistence of each node is longer. For example, node a1 may be more inclined to stay at node a1 for several days, and other nodes have similar characteristics. The r6i1p1f1 ensemble of IPSL-CM6A-LR is the lowest ranked ensemble for transition performance, and the correlation coefficient with ERA5 is only 0.37. The transition distribution of this ensemble is very different from that of ERA5. First, node a1 almost always transfers to nodes a2 and b1, and the probability of transferring to other nodes is almost zero. However, in ERA5, node a1 not only transfers to nodes a2 and b1, but sometimes transfers to other nodes such as a3 and c1. Secondly, some nodes, such as a2, a3, b3, and e3, do not even have a probability of transferring to other nodes. This is because these nodes do not have any days assigned to them, so there will be no transition probability for this node. However, the transition probability of some nodes in this ensemble is more consistent with ERA5. For example, this ensemble and ERA5 all have a high probability of transitioning from node d1 to e1.

The above results show that the highest performing ensembles for node frequency, persistence and transition in CMIP6 are the r10i1p1f1 ensemble of CanESM5, the r1i1p1f1 ensemble of NorESM2-MM and the r8i1p1f1 ensemble of MPI-ESM1-2-LR, respectively. The lowest performing ensemble for frequency and persistence is the r1i1p1f1 ensemble of INM-CM4-8, and the lowest performing ensemble for transition is the r6i1p1f1 ensemble of IPSL-CM6A-LR. To better describe the ability of CMIP6 models to simulate the overall circulation pattern in the Asian region, the MR metric is used here. According to the MR metric (Table 2), the top (1) ranking is the r10i1p1f1 ensemble of CanESM5, which is also the top ranking ensemble for frequency performance, and the second ranking ensemble is the r1i1p1f1 ensemble of NorESM2-MM, which is the top ranking for persistence performance. The lowest ranked ensemble is the r1i1p1f1 ensemble of INM-CM5-0. The r1i1p1f1 ensemble of INM-CM4-8 is ranked last for frequency performance and persistence performance, but due to its relatively high ranking for transition performance, it ranked second last rather than last by the MR metric.

3.4 Model ranking and the WNPSH simulation

The WNPSH is a key circulation system that controls the western Pacific summer monsoon and typhoon activities, and is an important indicator of summer weather in countries along the Northwest Pacific (Chen et al. 2020). In order to test whether the top-ranked ensembles can better represent the actual climate in Asia, we calculated the daily WNPSH metrics for ERA5 and all ensembles shown in Table 1. The probability density functions (pdf) of the 10 top and 10 lowest ensembles are then calculated, and compared with ERA5.

Figures 7 and 8 show the pdf distributions after normal distribution fitting of the western ridge point index and the northern boundary position index of ERA5 and the top and lowest 10 ensembles, respectively. In

Fig. 7, the average western ridge point index of ERA5 is around 130°E, and that of the top 10 ensembles is around 125°E, which is about 5°E west of ERA5. The pdf distribution of the lowest 10 ensembles shows that the western ridge point index is around 160°E on average, which is about 30°E east of that of ERA5 and even 10°E east of the average of the CMIP5 simulations (Zhao C. et al. 2019). This may be because only the CMIP5 r1i1p1f1 ensembles were evaluated, and here we evaluated all ensembles. In Fig. 8, the average northern boundary position index of ERA5 is around 31°N, and the average of the top 10 ensembles is around 29°N, about 2° to the south. The average northern boundary position index of the last 10 ensembles is around 40°N, which shows that the last 10 ensembles shift the northern boundary of the WNPSH northward by about 10°. As extreme events like heatwaves, droughts, and typhoons are greatly affected by the location of the WNPSH (Choi et al. 2019; Zhao Y. et al. 2019), such differences will inevitably lead to errors in the simulation of extreme events.

4. Discussion

In order to test whether different reanalysis data give different results, we generate another 'Master SOM' using JRA-55 reanalysis data (Ebita et al. 2011; Kobayashi et al. 2016), and then apply the same procedure to rank all ensembles listed in Table 1. The rankings (Supplementary Table 1) are similar to those shown in Table 2, which is obtained using ERA5 reanalysis data. For example, the 25 ensembles of CanESM5 are still ranked high, and the 32 ensembles of IPS-CM6A-LR are still ranked low, although there are slight differences in the specific ranking numbers. Therefore, we consider that the SOM results obtained from these two different reanalysis datasets are consistent.

The simulation capability of a model is affected by many factors, such as its dynamic core, parametric scheme, resolution and so on. Our research results show that the model's ability to simulate Z500 does not seem to have much to do with resolution. The top ranked ensemble for the MR metric, the r10i1p1f1 ensemble of CanESM5, has a relatively coarse spatial resolution of 500 km, while the second ranked r1i1p1f1 ensemble of NorESM2-MM has a relatively fine horizontal resolution of 100 km. The lower-ranking ensembles are mainly from the IPSL-CM6A-LR and INM-CM5-0 models, with resolutions of 250 km and 100 km, respectively. However, we have only four different resolutions here, 500 km, 250 km, 100 km, and 50 km, and there is only one ensemble with 50-km resolution, so we cannot draw a significant conclusion. Subsequent work needs to add more samples with different resolutions to analyze the relationship between circulation simulation skills and model resolution. In addition, the influence of different dynamic cores and parameterization schemes on the circulation simulation is worth further analysis.

Extreme events have seriously affected the world in recent years (Kong et al. 2020; Lee et al. 2020; Li D. et al. 2020; Lu et al. 2020; Meehl et al. 2004). Many previous studies have pointed out the close connections between circulation and extreme weather events (Boschat et al. 2014; Ding et al. 2011; Horton et al. 2015; Li et al. 2018; Liu et al. 2015; Pezza et al. 2011; Raymond et al. 2017). As mentioned above, the circulation pattern of node e4 favors the occurrence of extreme high temperature and heavy rainfall events in summer. Also, node e4 has a long duration (Figure 4), which will be more likely to lead to long

lasting extreme high temperature events and even compound extreme events (Faranda et al. 2020). Our next work will focus on the extent to which extreme events in Asia, such as extreme high temperatures and heavy precipitation, are affected by the circulation, and how the circulation affects these extreme events. If these issues can be studied thoroughly, it will help improve the accuracy of projections of extreme events and reduce the economic losses and casualties in Asia caused by extreme events.

5. Conclusions

The Asian summer climate is significantly affected by the large-scale atmospheric circulation. The ability of models to simulate circulation characteristics is one of the most important factors affecting the future progress of Asian regional climate research. This paper uses the SOM method to evaluate the ability of CMIP6 models in simulating Z500 in the Asian region. Our results show that the r10i1p1f1 ensemble of CanESM5 is the best ensemble for frequency performance, and it is also the top ensemble for comprehensive performance, measured as the MR metric. The r1i1p1f1 ensemble of NorESM2-MM is the best ensemble for persistence performance, and the r8i1p1f1 ensemble of MPI-ESM1-2-LR is the top ensemble for transition performance. The r1i1p1f1 ensemble of INM-CM4-8 is the lowest ranked ensemble for frequency performance and persistence performance. The r6i1p1f1 ensemble of IPSL-CM6A-LR has the lowest transition ranking. The r1i1p1f1 ensemble of INM-CM5-0 is the ensemble with the lowest ranking according to the MR metric. Generally speaking, the rankings of different ensembles from the same model are relatively close. For example, the MR rankings of the 25 ensembles of CanESM5 are relatively high, whereas the MR rankings of the 32 ensembles of IPSL-CM6A-LR are relatively low. In addition, pairwise correlation coefficients between the frequency performance, persistence performance and transition performance are all around 0.5 and significant, which means that good simulation skill for one type of circulation characteristic (like frequency) is related to skills in simulating other types of circulation characteristics (persistence and transition).

Judging from the pdf distributions of the western ridge point index and the northern boundary position index of the WNPSH, the top-ranked ensembles indeed simulate the position of the WNPSH better, although it is a little farther west and south than in reality, indicating that the rankings based on the SOM method are credible. The lower-ranked ensembles mainly simulate the WNPSH farther eastward and northward. The simulation of the western ridge point index is 30° farther east, and the simulation of the northern boundary position index is 10° farther north. Overall, the evaluation of the WNPSH's location supports the rankings based on the SOM method.

Declarations

Acknowledgments

Conflicts of interest/Competing interests The authors declare no competing interests.

Availability of data and material (data transparency) All original data can be downloaded from the URLs shown in the Acknowledgments. Processed data is available upon request from the corresponding authors (zuozhy@fudan.edu.cn, ann2012@mail.bnu.edu.cn).

Code availability (software application or custom code) All analysis code is available upon request from the corresponding authors (zuozhy@fudan.edu.cn, ann2012@mail.bnu.edu.cn).

References

1. Agel L, Barlow M, Feldstein SB, Gutowski WJ (2017) Identification of large-scale meteorological patterns associated with extreme precipitation in the US northeast. Clim Dyn 50:1819–1839. https://doi.org/10.1007/s00382-017-3724-8

2. Almazroui M, Saeed S, Saeed F, Islam MN, Ismail M (2020) Projections of Precipitation and Temperature over the South Asian Countries in CMIP6. Earth Syst Environ 4:297–320. https://doi.org/10.1007/s41748-020-00157-7

3. An N, Zuo Z (2021) Investigating the influence of large-scale circulation patterns on regional dry and wet heat waves in North China. Clim Dyn (in Press)

4. Boschat G, Pezza A, Simmonds I, Perkins S, Cowan T, Purich A (2014) Large scale and sub-regional connections in the lead up to summer heat wave and extreme rainfall events in eastern Australia. Clim Dyn 44:1823–1840. https://doi.org/10.1007/s00382-014-2214-5

5. Cassano JJ, Uotila P, Lynch AH, Cassano EN (2007) Predicted changes in synoptic forcing of net precipitation in large Arctic river basins during the 21st century. J Geophys Res Biogeosci 112 https://doi.org/10.1029/2006jg000332

6. Chen X, Zhou T, Wu P, Guo Z, Wang M (2020) Emergent constraints on future projections of the western North Pacific Subtropical High. Nat Commun 11:2802. https://doi.org/10.1038/s41467-020-16631-9

7. Choi W, Kim KY (2019) Summertime variability of the western North Pacific subtropical high and its synoptic influences on the East Asian weather. Sci Rep 9:7865. https://doi.org/10.1038/s41598-019-44414-w

8. Ding T, Qian W (2011) Geographical patterns and temporal variations of regional dry and wet heatwave events in China during 1960–2008. Adv Atmos Sci 28:322–337. https://doi.org/10.1007/s00376-010-9236-7

9. Ebita A et al (2011) The Japanese 55-year Reanalysis "JRA-55". An Interim Report Sola 7:149–152. https://doi.org/10.2151/sola.2011-038

10. Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

11. Faranda D, Vrac M, Yiou P, Jézéquel A, Thao S (2020) Changes in Future Synoptic Circulation Patterns: Consequences for Extreme Event Attribution. Geophys Res Lett 47 https://doi.org/10.1029/2020gl088002

12. Fischer EM, Schär C (2010) Consistent geographical patterns of changes in high-impact European heatwaves. Nat Geosci 3:398–403. https://doi.org/10.1038/ngeo866

13. Ford TW, Schoof JT (2017) Characterizing extreme and oppressive heat waves in Illinois. J Geophys Res Atmos 122:682–698. https://doi.org/10.1002/2016jd025721

14. Gao M, Yang Y, Shi H, Gao Z (2019) SOM-based synoptic analysis of atmospheric circulation patterns and temperature anomalies in China. Atmos Res 220:46–56. https://doi.org/10.1016/j.atmosres.2019.01.005

15. Gibson PB, Perkins-Kirkpatrick SE, Uotila P, Pepler AS, Alexander LV (2017) On the use of self-organizing maps for studying climate extremes. J Geophys Res Atmos 122:3891–3903. https://doi.org/10.1002/2016jd026256

16. Gibson PB, Uotila P, Perkins-Kirkpatrick SE, Alexander LV, Pitman AJ (2016) Evaluating synoptic systems in the CMIP5 climate models over the Australian region. Clim Dyn 47:2235–2251. https://doi.org/10.1007/s00382-015-2961-y

17. Gu S, Huang C, Bai L, Chu C, Liu Q (2016) Heat-related illness in China, summer of 2013. Int J Biometeorol 60:131–137. https://doi.org/10.1007/s00484-015-1011-0

18. Hersbach H et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049. https://doi.org/10.1002/qj.3803

19. Horton DE, Johnson NC, Singh D, Swain DL, Rajaratnam B, Diffenbaugh NS (2015) Contribution of changes in atmospheric circulation patterns to extreme temperature trends. Nature 522:465–469. https://doi.org/10.1038/nature14550

20. Jiang N, Luo K, Beggs PJ, Cheung K, Scorgie Y (2015) Insights into the implementation of synoptic weather-type classification using self-organizing maps: an Australian case study. Int J Climatol 35:3471–3485. https://doi.org/10.1002/joc.4221

21. Kobayashi C, Iwasaki T (2016) Brewer-Dobson circulation diagnosed from JRA-55. J Geophys Res Atmos 121:1493–1510. https://doi.org/10.1002/2015jd023476

22. Kong D, Gu X, Li J, Ren G, Liu J (2020) Contributions of Global Warming and Urbanization to the Intensification of Human-Perceived Heatwaves Over China. J Geophys Res Atmos 125

https://doi.org/10.1029/2019jd032175

23. La J, Shun-Wu Z, Feng D (2002) The Climatic Character of North Africa Subtropical High and Its Effect on Summer Precipitation in China. J Nanjing Inst Meteorol (in Chinese) 25:816–822. https://doi.org/10.13878/j.cnki.dqkxxb.2002.06.013

24. Lee S-H, Kang JE, Park CS, Yoon DK, Yoon S (2020) Multi-risk assessment of heat waves under intensifying climate change using Bayesian Networks. Int J Disast Risk Re 50 https://doi.org/10.1016/j.ijdrr.2020.101704

25. Li D, Yuan J, Kopp RE (2020) Escalating global exposure to compound heat-humidity extremes with warming. Environ Res Lett 15 https://doi.org/10.1088/1748-9326/ab7d04

26. Li J, Ruan C (2018) The North Atlantic–Eurasian teleconnection in summer and its effects on Eurasian climates. Environ Res Lett 13 https://doi.org/10.1088/1748-9326/aa9d33

27. Li M, Jiang Z, Zhou P, Le Treut H, Li L (2020) Projection and possible causes of summer precipitation in eastern China using self-organizing map. Clim Dyn 54:2815–2830. https://doi.org/10.1007/s00382-020-05150-4

28. Li W, Jiang Z, Xu J, Li L (2015) Extreme Precipitation Indices over China in CMIP5 Models. Part I: Model Evaluation. J Clim 28:8603–8619. https://doi.org/10.1175/jcli-d-15-0099.1

29. Liu W, Wang L, Chen D, Tu K, Ruan C, Hu Z (2015) Large-scale circulation classification and its links to observed precipitation in the eastern and central Tibetan Plateau. Clim Dyn 46:3481–3497. https://doi.org/10.1007/s00382-015-2782-z

30. Liu Y, Li W, Zuo J, Hu Z-Z (2014) Simulation and projection of the western pacific subtropical high in CMIP5 models. J Meteorol Res 28:327–340. https://doi.org/10.1007/s13351-014-3151-2

31. Liu Y, Liang P, Sun Y (2019) The Asian Summer Monsoon Characteristics, Variability, Teleconnections and Projection. Candice Janco, Cambridge

32. Liu Y, Weisberg RH, Mooers CNK (2006) Performance evaluation of the self-organizing map for feature extraction. J Geophys Res 111 https://doi.org/10.1029/2005jc003117

33. Loikith PC, Lintner BR, Sweeney A (2017) Characterizing Large-Scale Meteorological Patterns and Associated Temperature and Precipitation Extremes over the Northwestern United States Using Self-Organizing Maps. J Clim 30:2829–2847. https://doi.org/10.1175/jcli-d-16-0670.1

34. Loikith PC, Pampuch LA, Slinskey E, Detzer J, Mechoso CR, Barkhordarian A (2019) A climatology of daily synoptic circulation patterns and associated surface meteorology over southern South America. Clim Dyn 53:4019–4035. https://doi.org/10.1007/s00382-019-04768-3

35. Lu C et al (2020) An Unusual Heat Wave in North China During Midsummer, 2018. Front Earth Sci 8:238. https://doi.org/10.3389/feart.2020.00238

36. Maidens A, Knight JR, Scaife AA (2021) Tropical and stratospheric influences on winter atmospheric circulation patterns in the North Atlantic sector. Environ Res Lett 16 https://doi.org/10.1088/1748-9326/abd8aa

37. Meehl GA, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305:994–997. https://doi.org/10.1126/science.1098704

38. Mioduszewski JR, Rennermalm AK, Hammann A, Tedesco M, Noble EU, Stroeve JC, Mote TL (2016) Atmospheric drivers of Greenland surface melt revealed by self-organizing maps. J Geophys Res Atmos 121:5095–5114. https://doi.org/10.1002/2015jd024550

39. Ohba M, Sugimoto S (2020) Impacts of climate change on heavy wet snowfall in Japan. Clim Dyn 54:3151–3164. https://doi.org/10.1007/s00382-020-05163-z

40. Pezza AB, Van Rensch P, Cai W (2011) Severe heat waves in Southern Australia: synoptic climatology and large scale connections. Clim Dyn 38:209–224. https://doi.org/10.1007/s00382-011-1016-2

41. Preethi B, Mujumdar M, Prabhu A, Kripalani R (2017) Variability and teleconnections of South and East Asian summer monsoons in present and future projections of CMIP5 climate models. Asia-Pac J Atmos Sci 53:305–325. https://doi.org/10.1007/s13143-017-0034-3

42. Raymond C, Singh D, Horton RM (2017) Spatiotemporal Patterns and Synoptics of Extreme Wet-Bulb Temperature in the Contiguous United States. J Geophys Res Atmos 122:13,108–113,124. https://doi.org/10.1002/2017jd027140

43. Sales K et al (2018) Experimental heatwaves compromise sperm function and cause transgenerational damage in a model insect. Nat Commun 9:4771. https://doi.org/10.1038/s41467-018-07273-z

44. Wang Y, Jiang Z, Chen W (2015) Performance of CMIP5 models in the simulation of climate characteristics of synoptic patterns over East Asia. J Meteorol Res 29:594–607. https://doi.org/10.1007/s13351-015-4129-4

45. Yin C, Li Y, Ye W, Bornman JF, Yan X (2010) Statistical downscaling of regional daily precipitation over southeast Australia based on self-organizing maps. Theor Appl Climatol 105:11–26. https://doi.org/10.1007/s00704-010-0371-y

46. Zhang D, Martin GM, Rodríguez JM, Ke Z, Chen L (2020) Predictability of the Western North Pacific Subtropical High Associated with Different ENSO Phases in GloSea5. J Meteorol Res 34:926–940. https://doi.org/10.1007/s13351-020-0055-1

47. Zhao C, Jiang Z, Sun X, Li W, Li L (2019) How well do climate models simulate regional atmospheric circulation over East Asia? Int J Climatol 40:220–234. https://doi.org/10.1002/joc.6205

48. Zhao Y et al (2019) The Large-Scale Circulation Patterns Responsible for Extreme Precipitation Over the North China Plain in Midsummer. J Geophys Res Atmos 124:12794–12809. https://doi.org/10.1029/2019jd030583

49. Zhou TJ, Zou LW, Chen XL (2019) Commentary on the Coupled Model Intercomparison Project Phase 6 (CMIP6). Clim Change Res 15:445–456. https://doi.org/10.12006/j.issn.1673-1719.2019.193
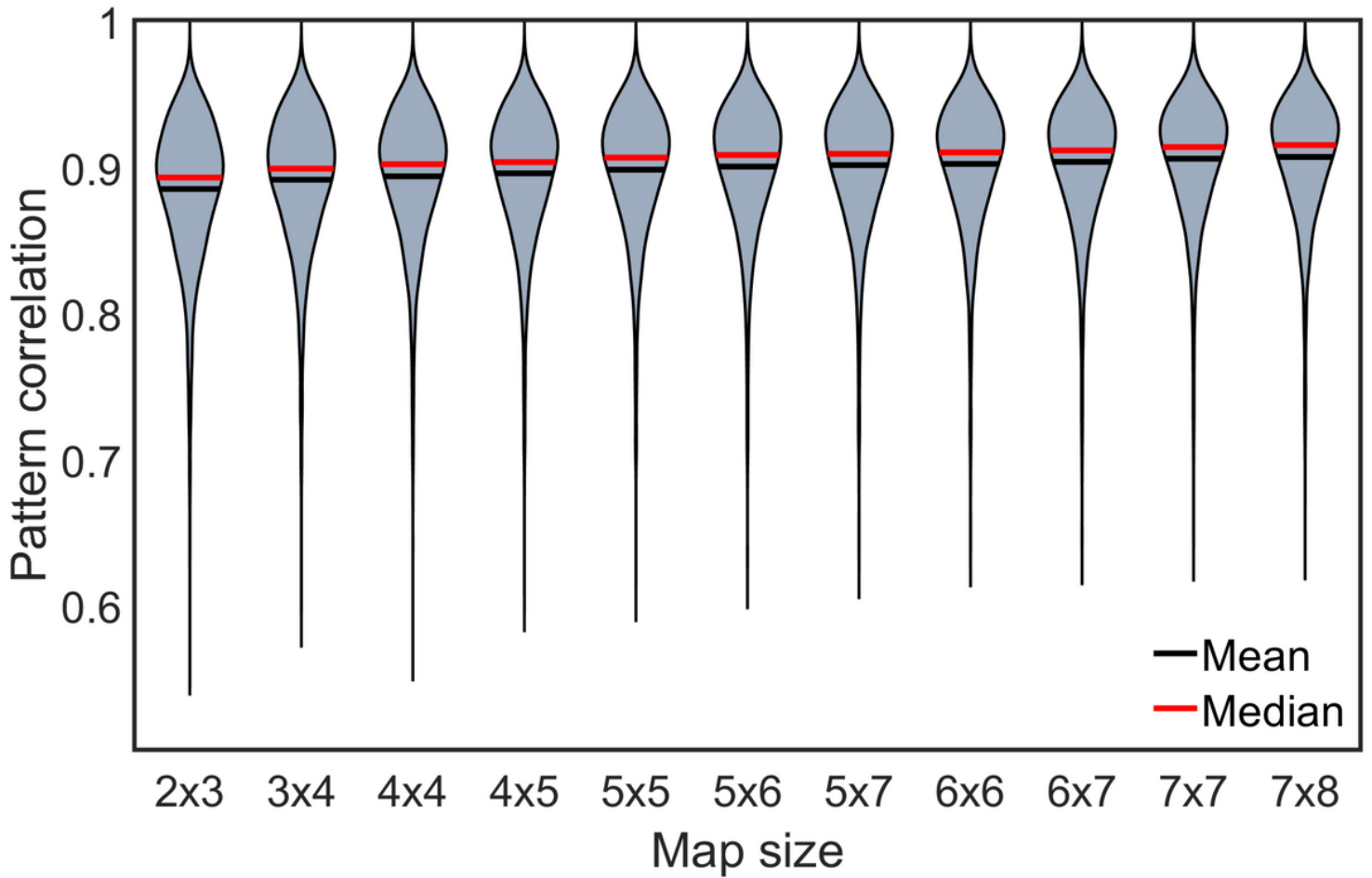
Figures

Figure 1

The "violin" plot showing pattern correlations under different node configurations between each daily absolute 500 hPa geopotential height and the SOM node it was assigned to, displaying the mean (black line), median (red line), and probability distribution (gray violin).
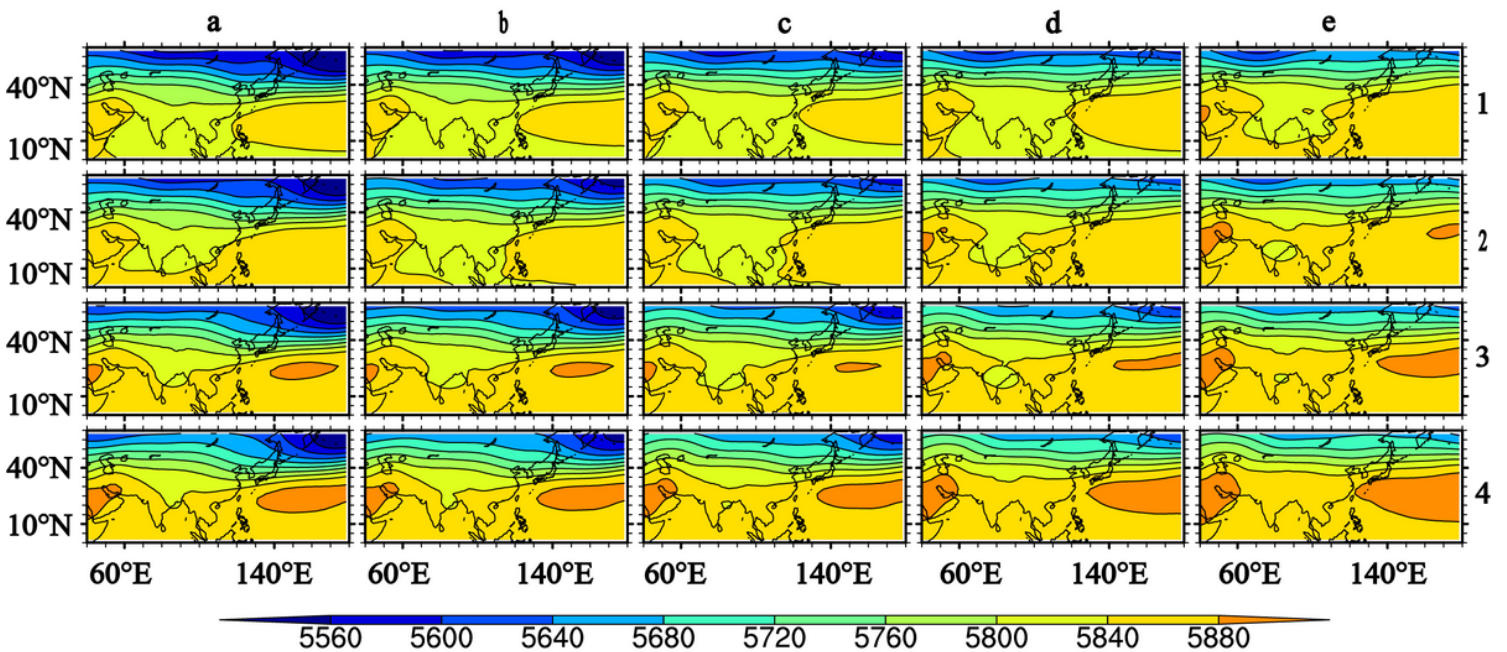
Figure 2

Master 20 node (4 × 5) SOM derived from ERA5 daily absolute 500 hPa geopotential height (shading and black contours every 40 gpm) in summer (June–August) over 1979–2014. In terms of the node referencing, node 'a1' refers to the node in the top-left corner of the SOM plane and node 'e4' refers to the node in the bottom-right corner. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.
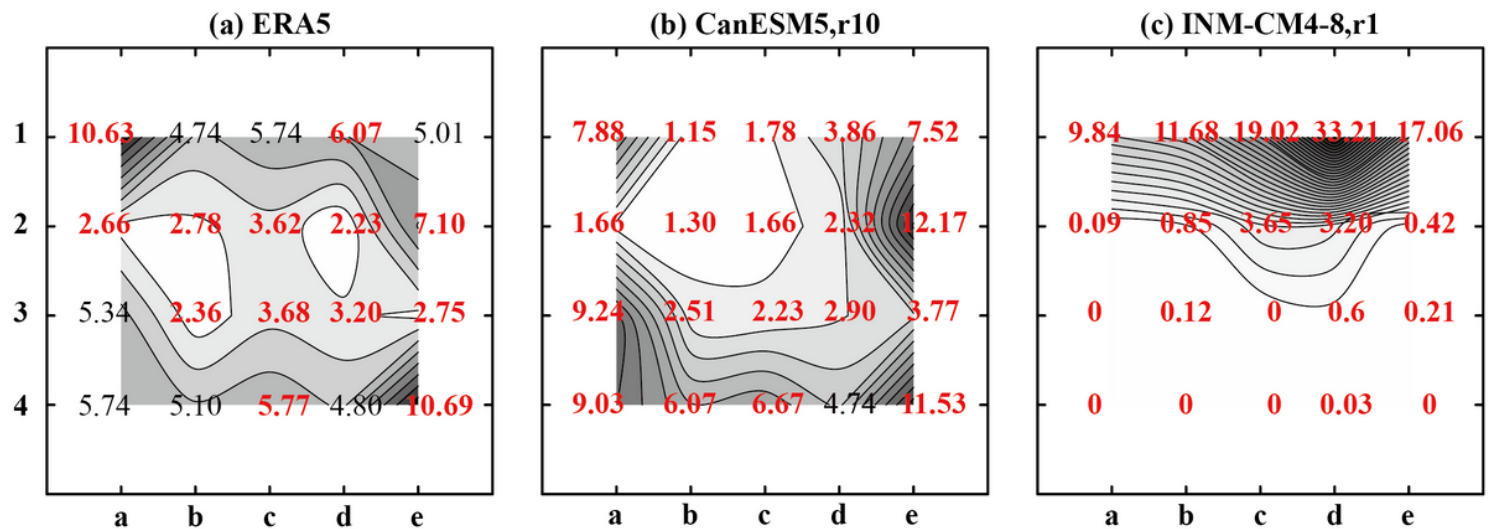
Figure 3

Contour plots of node frequencies corresponding to (a) the Master SOM from ERA5, (b) the highest ranked ensemble for frequency performance, and (c) the lowest ranked ensemble for frequency performance. The highest/lowest ranked ensembles for frequency performance are defined here based on the correlation coefficients of frequency metrics between ensembles and ERA5. Particular node frequencies statistically (at $p < 0.05$) above or below those expected values by chance are shown in red. Node locations correspond to those in Figure 2. The correlation coefficient between ERA5 and (b) the highest ranked ensemble for frequency performance is 0.73, while the correlation coefficient between ERA5 and (c) the lowest ranked ensemble for frequency performance is 0.21.
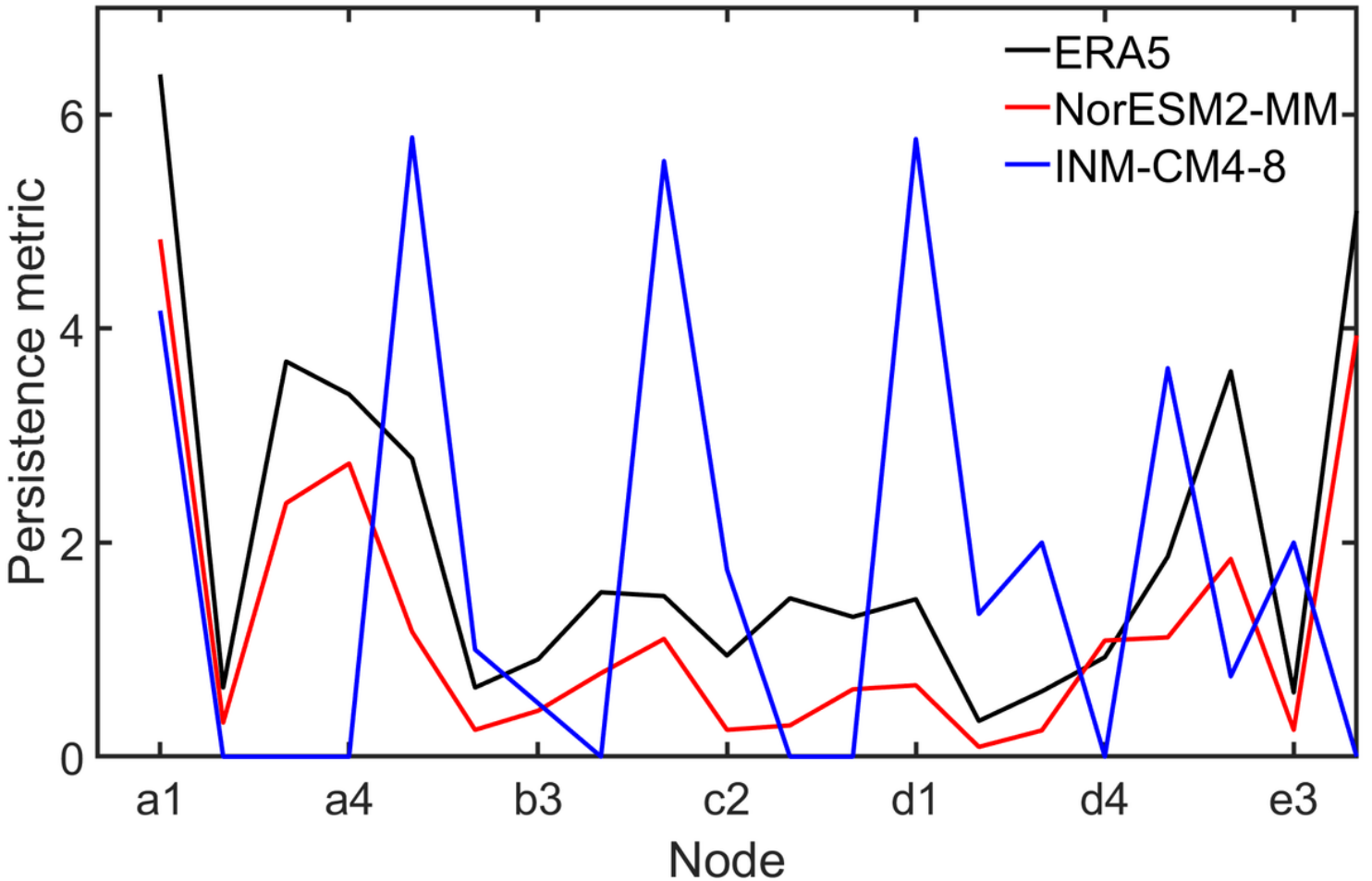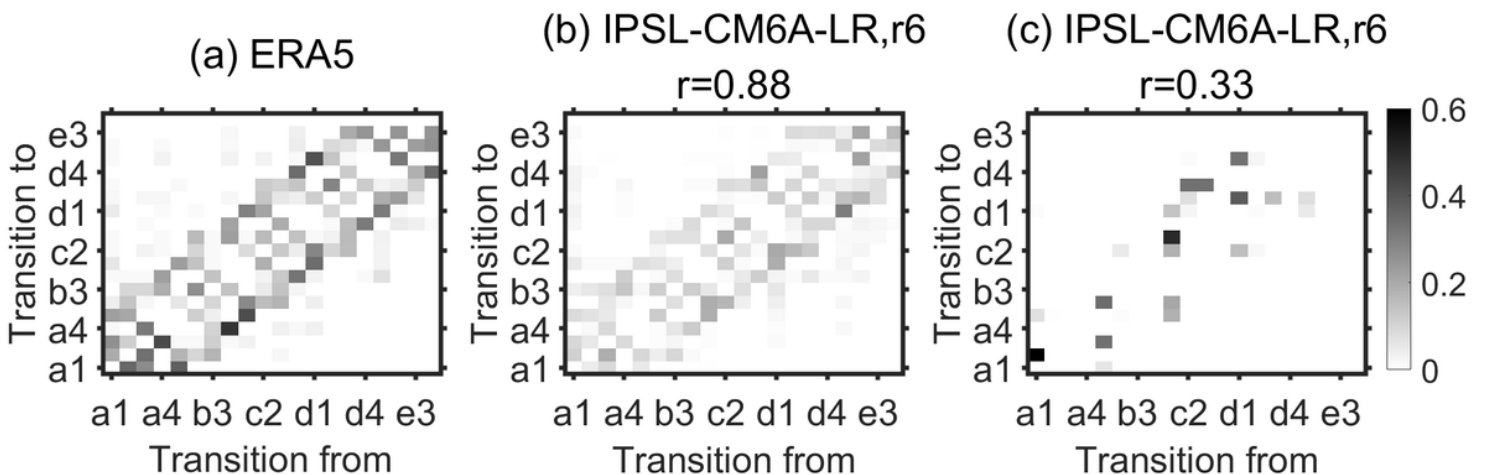
Figure 4

Line graphs of persistence metric corresponding to ERA5 (black line), the highest ranked ensemble for persistence performance (red line), and the lowest ranked ensemble for persistence performance (blue line). The highest/lowest ranked ensembles of persistence performance are defined here based on the correlation coefficients of persistence metrics between ensembles and ERA5. Node locations correspond to those in Figure 2. The correlation coefficient between ERA5 and the highest ranked ensemble for persistence performance is 0.97, while the correlation coefficient between ERA5 and the lowest ranked ensemble for persistence performance is 0.39.

Figure 5

Transition probability plots (grayscale indicates probability) for (a) ERA5, (b) the highest ranked ensemble for transition performance, and (c) the lowest ranked ensemble for transition performance. The highest/lowest ranked ensembles for transition performance are defined here based on the correlation coefficients of transition metrics between ensembles and ERA5. Node locations correspond to those in Figure 2. The correlation coefficient between ERA5 and (b) the highest ranked ensemble for transition performance is 0.86, while the correlation coefficient between ERA5 and (c) the lowest ranked ensemble for transition performance is 0.37.
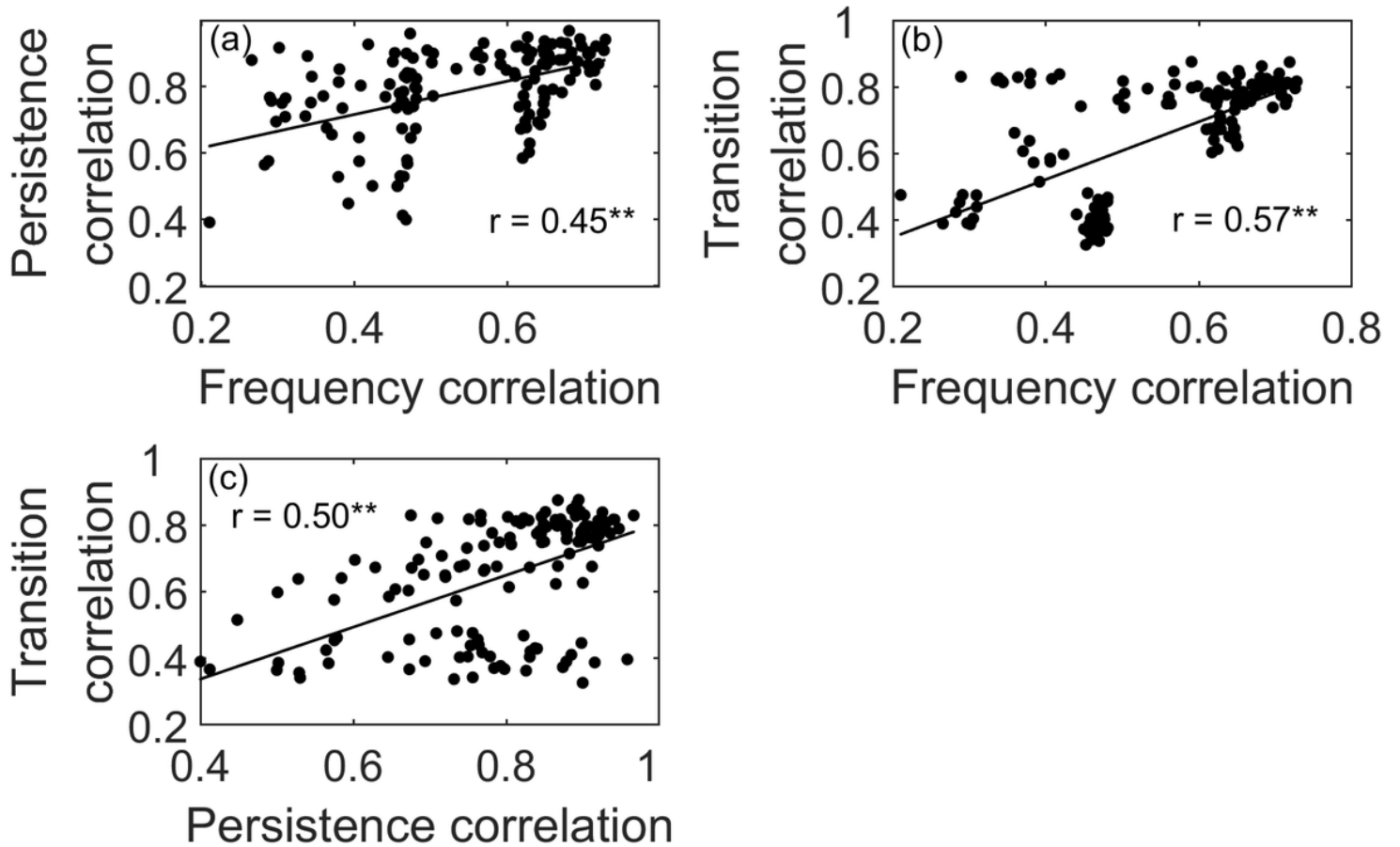

Figure 6

Scatterplots showing relationship between the performance of different metrics in the various CMIP6 ensembles and ERA5. The correlation coefficients between (a) frequency performance and persistence performance, (b) frequency performance and transition performance and (c) persistence performance and transition performance are 0.45, 0.57 and 0.50, respectively.
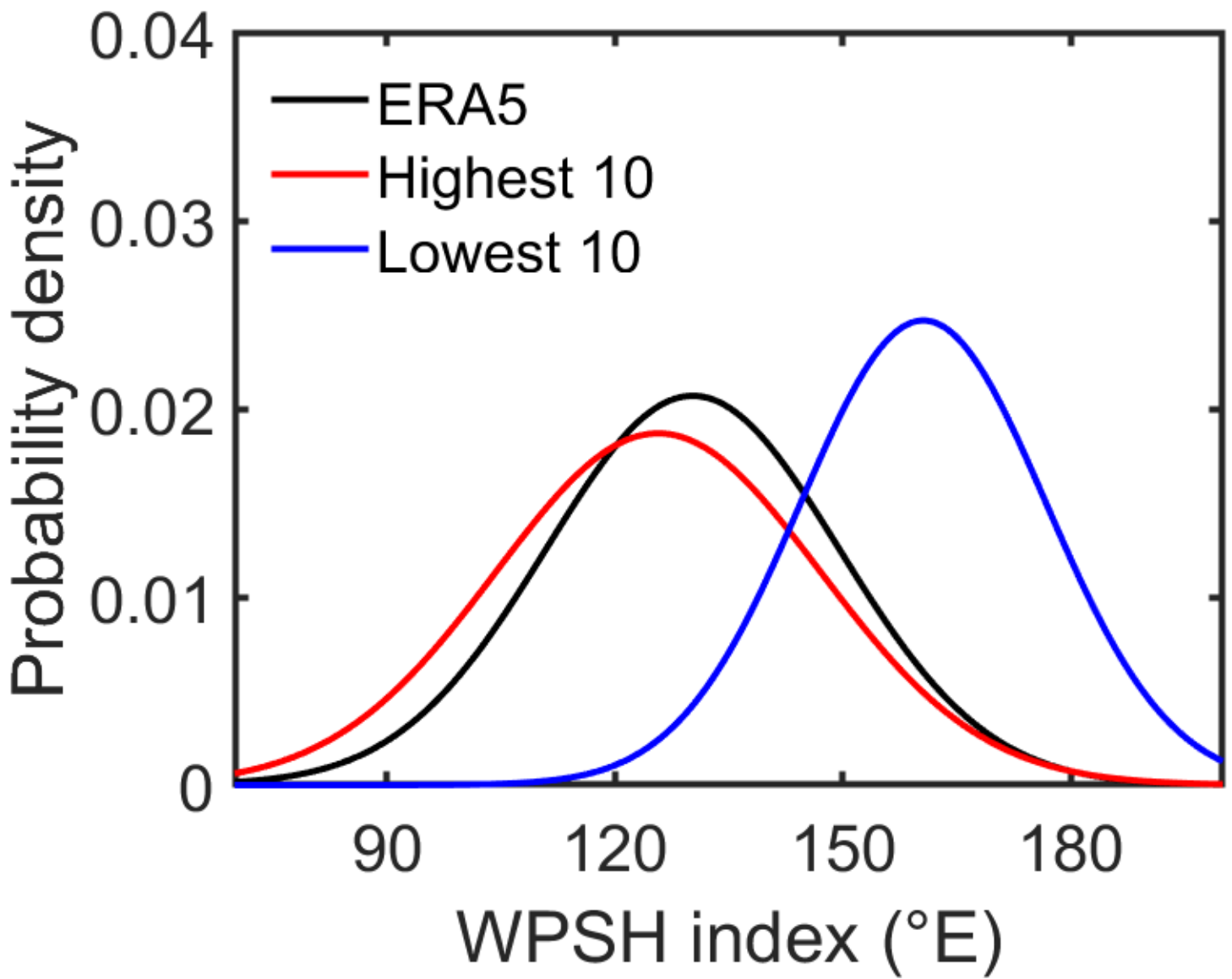
Figure 7

Probability density function distribution of the WNPSH western ridge point index from ERA5 (black curve) and the 10 highest ranked ensembles (red curve) and 10 lowest ranked ensembles (blue curve) after fitting to normal distributions.
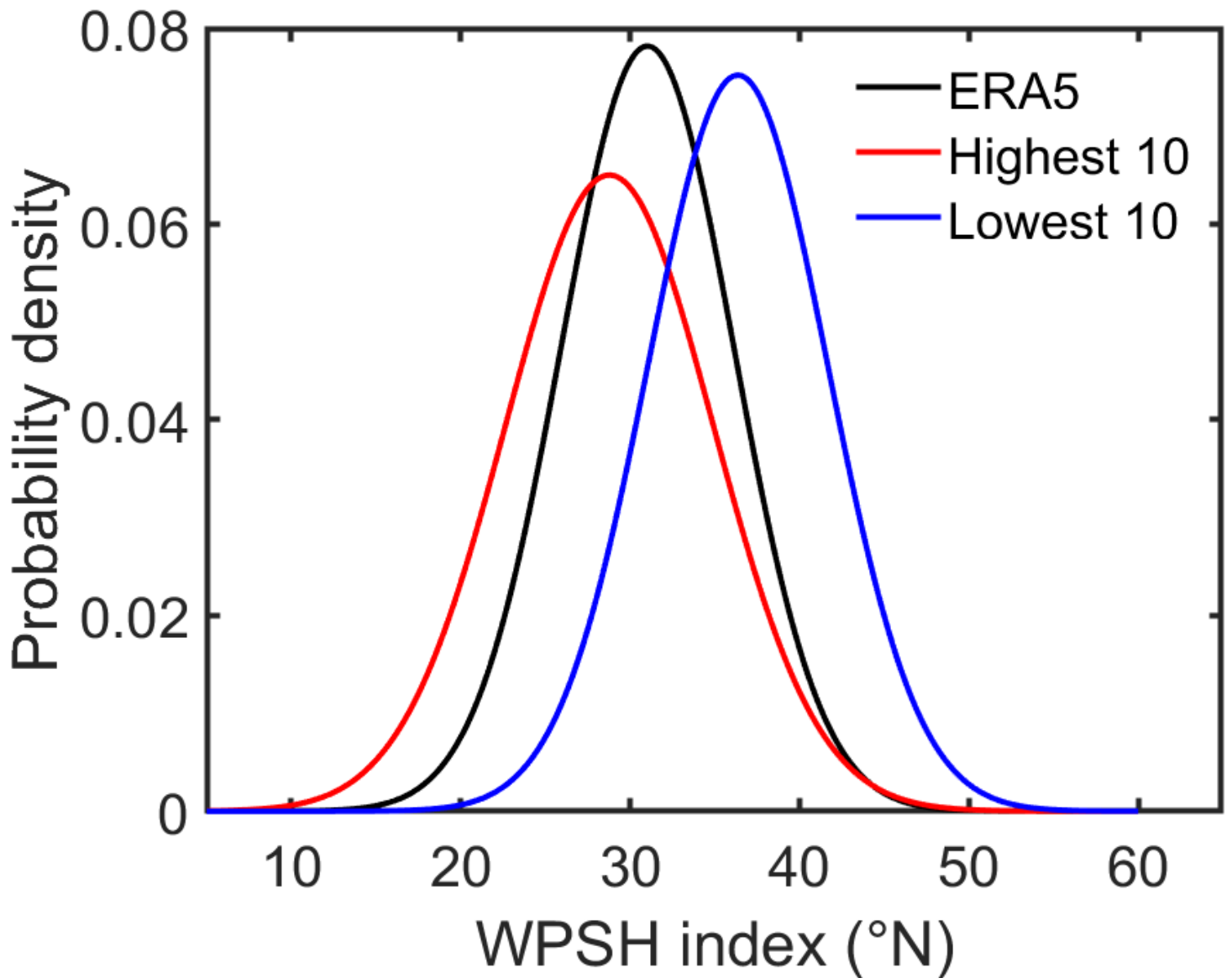
Figure 8

Probability density function distribution of the WNPSH northern boundary position index of ERA5 (black curve) and the 10 ranked ensembles (red curve) and 10 lowest ranked ensembles (blue curve) after fitting to normal distributions.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial0420.docx