
Performance of CMIP5 Models in the Simulation of Climate Characteristics of Synoptic Patterns over East Asia

WANG Yongdi 1,2 (王永弟), JIANG Zhihong 1* (江志红), and CHEN Weilin 1 (陈威霖)

1 Key Laboratory of Meteorological Disaster of Ministry of Education, Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing 210044

2 School of Geography and Remote Sensing, Nanjing University of Information Science and Technology, Nanjing 210044

(Received December 30, 2014; in final form May 25, 2015)

Supported by the National Natural Science Foundation of China (41230528, 41205162), National (Key) Basic Research and Development (973) Program of China (2012CB955204), and Priority Academic Program Development of Jiangsu Higher Education Institutions.

Corresponding author: zhjiang@nuist.edu.cn.

©The Chinese Meteorological Society and Springer-Verlag Berlin Heidelberg 2015
ABSTRACT

The evolution of daily synoptic weather patterns is the main driver of day-to-day weather change. These patterns are generally associated with changes in temperature, precipitation, etc., especially during extreme weathers. Evaluating the ability of climate models to reproduce the frequency and intensity of daily synoptic patterns is essential for increasing confidence in future projections. In this study, we investigated the ability of 34 global climate models (GCMs) included in the Coupled Model Intercomparison Project Phase 5 (CMIP5) to simulate over East Asia synoptic patterns and their evolution features in winter and summer. Daily synoptic patterns in sea level pressure and their occurrence frequencies were identified using an objective clustering algorithm; self-organizing maps (SOMs). The evaluation consists of correlating the frequencies of these patterns in the 34 CMIP5 models with the frequencies in the National Centers for Environmental Prediction (NCEP) reanalysis during the baseline period of 1980–1999. The results illustrated that most of these models were able to reproduce the synoptic patterns of the NCEP reanalysis. In addition, the frequencies of temporal SLP anomaly patterns were reproduced by most of the models over the baseline period, but the frequencies of spatial SLP anomaly patterns were only reproduced by a few GCMs. Overall, the models performed better in summer than in winter. Comprehensive evaluation shows that the four top-performing models for both winter and summer are bcc-csm1-1-m, NorESM1-M, MRI-CGCM3, and CCSM4. They show good performance in simulating the daily synoptic patterns in sea level pressure and in reproducing their occurrence frequencies. The results showed that the SOM was an effective tool for differentiating characteristics of synoptic circulation patterns and for evaluating the ability of climate models to simulate the frequency of daily synoptic patterns. The results shown in this paper can also help users to choose a better model for future climate projection and downscaling over the East Asia.

Key words: CMIP5, climate model evaluation, self-organizing maps, atmospheric circulation pattern
1. Introduction

Global climate models (GCMs) are a useful tool for simulating the present climate and for projecting the future climate. Evaluating the ability of climate models in simulating various features of large-scale circulations is essential, not only for increasing people’s confidence, both of professionals and laymen, in future projections but also for providing useful information for selecting appropriate GCMs for research such as future climate prediction and downscaling.

Using the outputs of the Coupled Model Intercomparison Project Phase 3 (CMIP3), researchers have evaluated the ability of GCMs to simulate circulations over East Asia (Li et al., 2011a, b; Zhou et al., 2009, 2010, 2011). Some research has shown that the CMIP3 models have some ability in regional climate simulation over East Asia (Zhou et al., 2006; Liu et al., 2009). Large differences, however, existed among different GCMs. The models’ ability in simulating the East Asian monsoon was weak overall (Zhao et al., 2013). For example, large model biases existed in the Qinghai-Tibet Plateau area (Zhang et al., 2011), and the simulations of the western Pacific subtropical high and the Mongolian high pressure were poor (Guo et al., 2012). Recently, several studies on the Coupled Model Intercomparison Project Phase 5 (CMIP5) model simulations over East Asia were performed (Xu et al., 2012a, b; Zhang, 2012; Chen, 2013; Sperber et al., 2013; Song et al., 2014a, b, c; He et al., 2014, 2015). Jiang et al. (2013) assessed the simulations of the climate conditions during the winter and summer monsoons over East Asia. They concluded that the overall strength of the East Asian winter monsoon did not have any trend when considered against the background of global warming. Simulation of the western Pacific subtropical high was evaluated by Liu et al. (2014). The authors concluded that most models showed a strong ability in simulating spatial patterns and amplitudes of the geopotential height field and zonal wind changes. Sperber et al. (2013) found that, in terms of the multi-model mean, the CMIP5 models outperformed the CMIP3 models in all of the diagnostics. Song et al. (2014a) reported that the CMIP5
multimodel ensemble mean showed a significant improvement over CMIP3 for the interannual East Asian Summer Monsoon (EASM) pattern.

Many studies have evaluated the climatological monthly-mean state of circulations, annual variation, and inter-annual and inter-decadal variability of circulation patterns. In many cases, however, it is also necessary to evaluate synoptic patterns and their frequency distribution on daily time scales. In recent years, studies of extreme weather events against the backdrop of global change have attracted increasing attention. A strong link exists between changes in extreme weather events and changes in synoptic patterns in terms of variation in both intensity and frequency (Paraschivescu et al., 2012; Dayan et al., 2012; Raziei et al., 2012; Cohen et al., 2013). This means that the climate characteristics of synoptic-scale cyclone activity (such as frequency, intensity, and duration) are closely linked to those of extreme weather events. However, according to Liu et al. (2005, 2006a, b) and Reusch et al. (2005, 2007), model evaluation of the climate characteristics for synoptic scale patterns (referred to as circulation patterns or CP hereafter) using traditional statistical methods (such as empirical orthogonal function or principal component analysis (PCA) and statistical indicators) is less accurate and less intuitive. Recently, Radic et al. (2011) used self-organizing maps (SOMs), an advanced nonlinear data analysis method, to evaluate Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4) models in terms of their ability in simulating the North America CP. A multitude of characteristic synoptic patterns and their occurrence frequencies within the study area were acquired. Model ability in simulating the synoptic scale CP was evaluated quantitatively by calculating the correlation coefficient between the frequencies of modeled patterns and frequencies of NCEP reanalysis patterns. Their research suggested that the SOM method was an effective tool to study a model’s ability in simulating synoptic patterns and their features. To our knowledge, there is less research on model evaluations of synoptic scale weather features over the East Asia.
The SOM was proposed by Kohonen in 1982, and was used in the study of atmospheric sciences by Hewitson et al. (2002). After the variability in multi-dimensional meteorological datasets has gone through self-learning, and been trained and clustered according to the neural networks principle, numerous characteristic synoptic patterns (e.g., winning neuron) and their evolution can be obtained (Reusch et al., 2007; Ning et al., 2012). This method has more advantages than PCA in distinguishing actual atmospheric synoptic patterns and their evolution (Reusch et al., 2005; Liu et al., 2006a, b). At present, SOM has been used in climate diagnosis, model evaluation, statistical downscaling, etc. The SOM method was used to examine the relationship between the local daily rainfall and humidity fields of large-scale circulation in northeastern and southeastern Mexico (Cavazos et al., 2000). Daily synoptic patterns were extracted from sea level pressure (SLP) fields and assessed in terms of the ability of GCM models to simulate atmospheric circulation patterns by SOMs (Finnis et al., 2009). Using SOMs, Schuenemann et al. (2009) classified the synoptic patterns in the North Atlantic region from 1961 to 1999. Detailed studies of the differences in various patterns between model output and observational data, including corresponding precipitation, were also performed (Schuenemann et al. 2009). The deviations in annual precipitation simulation by 15 IPCC AR4 models were analyzed and attributed to differences in intra-pattern variability, differences in pattern frequency, and differences in pattern frequency acting on intra-pattern variability. These results showed that SOMs could be used as an effective tool to analyze the possible causes of differences between simulation and observation.

The simulation results of more than 50 models are available from the CMIP5 archives. Compared with the CMIP3 models, most CMIP5 models have been significantly improved in terms of model structure, resolution, physical processes, etc. (Taylor et al., 2012). However, traditional statistical methods are inefficient in the assessing of daily circulation patterns climatic characteristics. Therefore, it is important to assess the simulation ability of CMIP5 models in terms of the climate characteristics of
daily CPs over the East Asia. In this study, we use the SOM-based methodology introduced by Radic et al. (2011) to obtain the main synoptic patterns in daily SLP and their occurrence frequencies in both summer and winter. By comparing the similarity of characteristic synoptic pattern frequency distribution in the GCMs and NCEP, we evaluate how well the CMIP5 models reproduce the observed synoptic scale atmospheric circulations over the East Asia.

2. Data

Our evaluation focuses on the 20-yr datasets of CMIP5 outputs during 1980–1999, which are based on the simulations (Run r1i1p1 of the “historical” simulation of CMIP5 models) by 34 GCMs (archived by the PCMDI at http://pcmdi9.llnl.gov/esgf-web-fe/). These models are listed in Table 1. To represent the synoptic patterns, we use the daily SLP field as the input data of the SOM training process.

Table 1. Basic information on the 34 models used in this study, including model identification, originating center, and atmospheric model resolution

<table>
<thead>
<tr>
<th>Models</th>
<th>Institution</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESS1-0</td>
<td>Commonwealth Scientific and Industrial Research Organisation and Bureau of Meteorology, Australia</td>
<td>1.875° × 1.25°</td>
</tr>
<tr>
<td>ACCESS1-3</td>
<td>Commonwealth Scientific and Industrial Research Organisation and Bureau of Meteorology, Australia</td>
<td>1.875° × 1.25°</td>
</tr>
<tr>
<td>bcc-csm1-1</td>
<td>Beijing Climate Center, China Meteorological Administration, China</td>
<td>2.8° × ~2.8°</td>
</tr>
<tr>
<td>bcc-csm1-1-m</td>
<td>Beijing Climate Center, China Meteorological Administration, China</td>
<td>1.125° × ~1.12°</td>
</tr>
<tr>
<td>BNU-ESM</td>
<td>Beijing Normal University, China</td>
<td>2.8° × ~2.8°</td>
</tr>
<tr>
<td>CanESM2</td>
<td>Canadian Centre for Climate Modelling and Analysis, Canada</td>
<td>2.8° × ~2.8°</td>
</tr>
<tr>
<td>Model</td>
<td>Institution</td>
<td>Resolution</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>CCSM4</td>
<td>National Center for Atmospheric Research (NCAR), USA</td>
<td>$1.25^\circ \times ~0.9^\circ$</td>
</tr>
<tr>
<td>CESM1-FASTCHEM</td>
<td>National Science Foundation/Department of Energy NCAR, USA</td>
<td>$1.25^\circ \times ~0.9^\circ$</td>
</tr>
<tr>
<td>CMCC-CM</td>
<td>Centro Euro-Mediterraneo per i Cambiamenti, Italy</td>
<td>$0.75^\circ \times ~0.75^\circ$</td>
</tr>
<tr>
<td>CMCC-CMS</td>
<td>Centro Euro-Mediterraneo per i Cambiamenti, Italy</td>
<td>$1.875^\circ \times ~1.875^\circ$</td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>Centre National de Recherches Meteorologiques, Meteo-France, France</td>
<td>$1.4^\circ \times ~1.4^\circ$</td>
</tr>
<tr>
<td>CSIRO-Mk3-6-0</td>
<td>Australian Commonwealth Scientific and Industrial Research Organization, Australia</td>
<td>$1.875^\circ \times ~1.875^\circ$</td>
</tr>
<tr>
<td>EC-EARTH</td>
<td>Royal Netherlands Meteorological Institute, The Netherlands</td>
<td>$1.125^\circ \times 1.125^\circ$</td>
</tr>
<tr>
<td>FGOALS-g2</td>
<td>Institute of Atmospheric Physics, Chinese Academy of Sciences, China</td>
<td>$2.8^\circ \times ~1.65^\circ$</td>
</tr>
<tr>
<td>FGOALS-s2</td>
<td>Institute of Atmospheric Physics, Chinese Academy of Sciences, China</td>
<td>$2.8^\circ \times ~2.8^\circ$</td>
</tr>
<tr>
<td>GFDL-CM3</td>
<td>Geophysical Fluid Dynamics Laboratory, USA</td>
<td>$2.5^\circ \times 2.0^\circ$</td>
</tr>
<tr>
<td>GFDL-ESM2G</td>
<td>Geophysical Fluid Dynamics Laboratory, USA</td>
<td>$2.5^\circ \times ~2.0^\circ$</td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>Geophysical Fluid Dynamics Laboratory, USA</td>
<td>$2.5^\circ \times ~2.0^\circ$</td>
</tr>
<tr>
<td>GISS-E2-R</td>
<td>NASA Goddard Institute for Space Studies, USA</td>
<td>$2.5^\circ \times 2.0^\circ$</td>
</tr>
<tr>
<td>HadCM3</td>
<td>Met Office Hadley Centre, UK</td>
<td>$3.75^\circ \times 2.5^\circ$</td>
</tr>
<tr>
<td>HadGEM2-CC</td>
<td>Met Office Hadley Centre, UK</td>
<td>$1.875^\circ \times 1.25^\circ$</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>Met Office Hadley Centre, UK</td>
<td>$1.875^\circ \times 1.25^\circ$</td>
</tr>
<tr>
<td>inmcm4</td>
<td>Institute for Numerical Mathematics (INM), Russia</td>
<td>$2^\circ \times ~1.5^\circ$</td>
</tr>
<tr>
<td>IPSL-CM5A-LR</td>
<td>Institut Pierre-Simon Laplace, France</td>
<td>$3.75^\circ \times ~1.895^\circ$</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>Institut Pierre-Simon Laplace, France</td>
<td>$2.5^\circ \times ~1.27^\circ$</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute</td>
<td>$2.8^\circ \times ~2.8^\circ$</td>
</tr>
</tbody>
</table>
We compare GCM outputs with the NCAR/NCEP reanalysis data (Kalnay et al., 1996), which is available from 1948 to the present. We assume that the NCAR/NCEP reanalysis is a reasonable proxy of the observation. To facilitate GCM intercomparison and validation against the NCAR/NCEP reanalysis, all model simulation outputs were interpolated to a common grid (with a resolution of 2.5°) using bilinear interpolation. The East Asia defined in this study is from 20 to 180°E, and from 0 to 80°N. Note that the SLP field over the Qinghai-Tibetan Plateau (74–104°E, 25–40°N) is not included in the analysis, because the model result there is not realistic.

3. Method

3.1 Brief description of SOM

The SOM is an unsupervised, artificial neural network based on competitive learning (Kohonen et al., 1982, 2001). Characteristic synoptic patterns represent the nonlinear phases of the synoptic scale circulation by self-learning and repeated competition. The SOM consists of two layers: input and competitive (output). Training and mapping are the two main processes in SOM training. In this study,
The daily SLP field is used as the input layer of the SOM, and certain types of synoptic circulation mode are found as winning nodes (known as nodes) in the output layer. Each category is referred to as a node, and the number of nodes is specified by the user.

The general SOM training algorithm is summarized as follows:

a) Establishment of the SOM network and its initialization. The SOM begins with a user-defined map size, and the number of circulation types is determined as needed. In order to represent the input data, the nodes of an SOM are first initialized (initialized SOM nodes are also known as reference vectors). The SOM training proceeds by each reference vector learning iteratively from the input data. Then, a random weight vector is assigned to each node.

b) Calculation of the Euclidean distance (EUD) between the weight of each node’s reference vector and an input vector. This finds the minimum EUD between the weight vector of the jth node and the input data. The EUD is defined as

\[ EUD = \sqrt{\sum_{i=1}^{n} (X_i - W_{ij})^2} \]  

where \( W_{ij} \) is the weight vector of the jth node, \( X \) is the input data, and \( X_i \) is the ith input data.

c) Finding the winning node. The EUD between the input sample and each node in the competitive (output) layer is calculated, and the node that is closest (having the smallest EUD) to the input vector is known as the winning node.

d) Updating the weight of each node. The node positions that are in the neighborhood of the winning neuron are updated. They are moved to the position that is closer to the input vector. Each node learns from the input data and shifts its position accordingly. Then, the SOM becomes topologically organized depending on the smallest EUD, resulting in similar nodes being adjacent to each other.

Updating of the weight vector can be described as
where \( t \) is the iteration variable, \( W \) is the weight vector, \( V \) is the input vector, and \( \alpha \) is the learning rate (over time). Considering the winning node to be at the center, and the radius \( R \) from the center, \( \Theta \) is the predefined neighborhood surrounding node, \( j \) (often called the neighborhood function). All input data of this repeated process are recorded as SOM training. Training is performed for a chosen iteration limit.

e) Updating the learning rate and neighborhood size. The next input vector is entered and steps b–e are repeated.

f) Determining the termination condition. Training is interrupted when the number of iterations reaches the preset condition. The final winning node is the characteristic synoptic pattern extracted by the SOM.

Details of the SOM method can be found in the literature (e.g., Kohonen 2001; Chu et al., 2012). The choice of SOM node size is actually determined by the number of major synoptic patterns.

In mathematics, the size of SOM nodes is determined in accordance with a principle: maximizing similarity within clusters and minimizing similarity between clusters. In this study, we use several different SOM sizes (the numbers of nodes being \( 3 \times 3 \), \( 4 \times 4 \), \( 5 \times 5 \), and \( 6 \times 6 \)) to compare the training results. Relevant discussion and analysis are provided in section 4.4.

### 3.2 Data preprocessing

In order to evaluate the ability of these models in terms of simulating daily synoptic patterns and their evolution, two types of anomalies are considered in our SOM analysis:

a) The daily-averaged SLP over the 20-yr baseline period (1980–1999) is subtracted from the daily SLP at each grid point (for the NCEP reanalysis and for each GCM output). The resulting anomaly fields are employed in the SOM training. These temporal SLP anomalies are herein referred to as temporal SLP.
b) For daily data, the average of each grid point within the study area is subtracted from the daily SLP at each grid point. The resulting anomaly fields are used for training. These spatial SLP anomalies are herein referred to as spatial SLP.

3.3 Evaluation of CMIP5 models

After temporal SLP training and spatial SLP training in winter and summer, respectively, we obtain two kinds of SOM patterns in winter and summer. The simulation ability of each model can be assessed by comparing the frequencies of modeled patterns with the frequencies of observations. A good model can often reproduce the actual atmospheric synoptic patterns (here represented by the NCEP reanalysis data), and each weather pattern should have the same occurrence of frequency as that of the NCEP reanalysis. Therefore, a model’s simulation ability can be measured by the correlation coefficient between the frequencies of modeled patterns and frequencies of NCEP reanalysis patterns. A ranking of model simulation ability in terms of weather pattern characteristics can be obtained using these coefficients.

4. Results and analysis

4.1 Synoptic patterns in observations

4.1.1 Characteristic synoptic patterns of SOM in temporal SLP

Following Radic et al. (2011), using temporal SLP anomalies from the NCEP reanalysis, we obtained the synoptic patterns (with a map size of $4 \times 4$) of temporal SLP and analyzed their evolution characteristics in winter and summer for the period of 1980–1999. Figure 1a and b illustrates the characteristic daily SLP anomaly patterns in winter and summer, respectively, for this period over East Asia. The temporal SLP, with the daily-averaged SLP over the 20-yr baseline period (1980–1999) subtracted, reflects the daily evolution of major weather systems over East Asia because it indicates the daily anomalies in the reference period.
For winter, the most striking feature in the SOM pattern is the evolution of the Mongolian high, which is the dominant system of the East Asian Winter Monsoon. As shown in Fig. 1a, there is initially a notable positive anomaly in the node (1, 1). From node (1, 1) to node (1, 4) (top panel), this positive anomaly becomes gradually weaker and moves eastward, accompanied by a negative anomaly appearing over the western region. From node (1, 1) to node (1, 4) (left column) the positive anomalies move westward, and, at the same time, a negative anomaly emerges in the northeastern region. The above negative anomaly becomes stronger from node (1, 4) and (4, 1) to node (4, 4). In other words, a cycle of one SOM pattern represents the evolution of the Mongolian high from a positive anomaly to a negative anomaly.

During the summer months, the situation was more complicated than that in winter. Although the Mongolian high still existed, its magnitude change was much smaller than in winter, and its range of effect was mainly at high latitudes. At middle and lower latitudes, the Indian low-pressure system existed. Meanwhile, there were some intensity changes of the subtropical high-pressure system in the western Pacific. Therefore, the SOM patterns in Fig. 1b comprehensively reflect the major systems of the summer SLP. The characteristic anomalous changes of some systems can be identified, such as the Mongolian high at high latitudes, and the Indian low and West Pacific subtropical high-pressure system in middle and lower latitudes.

4.1.2 The characteristic synoptic patterns of SOM in spatial SLP

Since spatial SLP is based on the domain-averaged daily SLP, it reflects the spatial pattern of the major weather systems in the area, which focus on spatial variability. Figure 2a depicts the evolution of winter SLP based on the spatial SOM analysis. It can be seen that the Mongolian high is still the most important high pressure system. From node (1, 1) to node (4, 4), an obvious variation process of the major pressure system position can also be seen. For example, nodes (1, 1) → (4, 1) → (4, 4) and (1, 1)
(1, 4) → (4, 4) show the process of the Mongolian high pressure moving southward from the east and west sides, respectively.
Fig. 1. The $4 \times 4$ SOM patterns (hPa) in temporal SLP based on the 1980–1999 NCEP reanalysis data: a) winter and b) summer.
Fig. 2. The $4 \times 4$ SOM patterns (hPa) in spatial SLP based on the 1980–1999 NCEP reanalysis data: a) winter and b) summer.
In the process from (1, 1) to (4, 4), two opposite changes are seen in the Aleutian low and Icelandic low pressure systems. The Icelandic low pressure system enhances and expands gradually, and is accompanied by the Aleutian Low system waning slightly. The change process of the summer major climate system is shown in Fig. 2b, and the West Pacific subtropical high and Indian low pressures are very clear. The nodes from (1, 1) to (4, 4) show that the subtropical high pressure gradually expands and strengthens from east to west. In this process, the Indian low does not show any significant changes.

4.2 Performance of CMIP5 models in terms of occurrence frequency of synoptic patterns

After creating the SOM classifications, the model performance could be evaluated. A model that can show similar synoptic patterns, and similar frequencies of these patterns, to the NCEP reanalysis will be regarded as a good model. The node frequency (in %) for a certain pattern was calculated as the total number of days of that pattern (node) divided by the total number of days in that season over the 20-yr baseline period. According to Radic et al. (2011), the success of a model depends on how well these frequencies from the model simulations correlate with the frequencies from the NCEP reanalysis. In order to determine the simulation ability of the synoptic patterns and frequencies of various CMIP5 models, an assessment was performed as follows.

a) In order to obtain the occurrence frequencies in the model, each GCM is classified according to the SOM trained by the NCEP reanalysis.

b) The frequencies of modeled patterns and those of the NCEP reanalysis patterns are compared, and the correlation coefficient between them is calculated. Thus, the simulation ability of these patterns and their frequencies in the SLP field can be assessed quantitatively.

We use the National Climate Center of China model, bcc-csm1-1-m, as an example to describe the assessment results using the SOM method.
Fig. 3. Comparison of the 4 × 4 SOM pattern frequencies of the NCEP reanalysis with those of bcc-csm1-1-m in temporal SLP: a) winter and b) summer.

Figure 3 plots the frequency of each node from the NCEP reanalysis and bcc-csm1-1-m. The nodes in Figs. 3a and b correspond to the nodes in Figs. 1a and b, respectively. In the patterns of temporal SLP anomalies, a significant positive correlation (at the 95% confidence level) is shown between the two (winter: \( r = 0.66 \); summer: \( r = 0.90 \)). A high positive correlation indicates that, for a given season over the baseline period, each pattern occurred in bcc-csm1-1-m as often as in the NCEP reanalysis.

Fig. 4. Comparison of the 4 × 4 SOM pattern frequencies of the NCEP reanalysis with those of bcc-csm1-1-m in spatial SLP: a) winter, and b) summer.
In the patterns of spatial SLP anomalies (Figs. 4a and b) the correlation is also significantly positive (winter: $r = 0.63$; summer: $r = 0.73$). This indicates that the simulation ability of bcc-csm1-1-m in spatial SLP was slightly weaker than that in temporal SLP. However, the occurrence of SLP spatial patterns in bcc-csm1-1-m was almost equal to their occurrence in the NCEP reanalysis. In other words, spatial SLP patterns that occurred frequently in the NCEP reanalysis also occurred frequently in bcc-csm1-1-m. We conclude that this model can simulate occurrence frequencies of various SOM patterns in East Asia.

![Fig. 5. Scatter plots of simulation frequency and NCEP reanalysis frequency in temporal and spatial SLP: a) winter and b) summer. The x-axis shows correlation coefficients of temporal SLP, and the y-axis shows correlation coefficients of spatial SLP.](image)
temporal and spatial SLP. The results for the temporal and spatial $4 \times 4$ SOMs over the study region are shown in Fig. 5. A correlation greater than 0.5 (at the 95% confidence level) is defined as significant.

For temporal SLP anomaly patterns (the x-axis of Fig. 5), many GCMs have significantly positive correlations ($r > 0.5$) with the NCEP reanalysis in both winter and summer; specifically, the percentage in winter is 66.7% (23 out of 34) and in summer is 88.2% (30 out of 34). According to the high percentage of significantly positive correlations, we can conclude that the frequencies of temporal SLP anomaly patterns in the NCEP reanalysis over the baseline period can be reproduced by most GCMs, and that these models’ simulation ability is better in summer than in winter.

Most of the CMIP5 GCMs can simulate fairly well the frequencies of temporal SLP anomaly patterns, but analysis of spatial SLP anomaly patterns shows less encouraging results. The y-axis of Fig. 5 indicates that the percentage of significantly positive correlations in winter is 20.6% (7 out of 34), while in summer it is only 17.6% (6 out of 34). The small percentage of significantly positive correlations means that very few CMIP5 models are able to reproduce the frequencies of spatial SLP anomaly patterns in the NCEP reanalysis in both winter and summer over the baseline period.

Based on these results of both temporal and spatial SLP, we can conclude that most of the CMIP5 models have strong simulation ability in temporal, but are poor in spatial, SLP. Our results in this study over the East Asia are consistent with those in North America (Radic et al., 2011).

4.3 Comparison of model simulation ability in temporal and spatial SLP

To further investigate these differences in simulation ability, the contrast between model frequency and NCEP reanalysis frequency is discussed below.

The dotted lines plotted in Fig. 5 show the correlation coefficient value of 0.5, which is the threshold of the 95% confidence level. The models located above the horizontal dotted line have good ability in simulating weather pattern frequencies in spatial SLP, and those on the right of the vertical
dotted line have superior performance for temporal SLP. That is to say, the models having good
simulation ability of occurrence frequency should appear in the top right quarter, and the models having
poor ability should appear in the lower left quarter. The number of models that have significant
correlations in temporal SLP is significantly higher than that in spatial SLP. Comparing winter and
summer results, the number of models in the bottom left quarter during summer is significantly lower
than that during winter. This shows that the model simulation ability of synoptic patterns in summer is
better than that in winter. This conclusion is consistent with that in section 4.2. Overall, the top five
models in winter are NorESM1-M, bcc-csm1-1-m, CCSM4, MRI-CGCM3, and EC-EARTH; and the
top five models in summer are bcc-csm1-1-m, NorESM1-M, MRI-CGCM3, CCSM4, and MPI-ESM-P.
Irrespective of the season, bcc-csm1-1-m, NorESM1-M, MRI-CGCM3, and CCSM4 all show excellent
performance.

4.4 Influence of the number of nodes

In order to single out the models that have superior performances over the East Asia, the models
are ranked according to the correlations obtained in section 4.3. Note that the choice of the SOM size
might impact our evaluation of GCM performance. Therefore, in order to find a reasonable compromise
between detail and interpretability of the SLP pattern characteristics for each season, different SOM
sizes were employed in our experiments. Our final choice for both spatial domains was to use four SOM
sizes: 3 × 3, 4 × 4, 5 × 5, and 6 × 6. We can show how much SOM size influences the model evaluation
by using more than one SOM size. Figure 6 shows the ranking curves with different nodes. The trends of
these four curves are similar, and the correlation coefficients between any two sets are in the range of
0.63 and 0.94. Therefore, the rankings are not very sensitive to the number of SOM output nodes.
Because of this, we will only consider the 4 × 4 SOM node size to further rank the models.
Fig. 6. Ranking of model simulation ability using different numbers of SOM nodes. Four different symbols represent four different SOM nodes. Names of the 34 CMIP5 models are shown on the x-axis, and the ranking is shown on the y-axis.

5. Conclusions

Using the SOM method, we identified and classified the characteristics of daily synoptic patterns of SLP in the NCEP reanalysis. We then analyzed the performance of the 34 GCMs over the East Asia. Emphasis was given to the evaluation of a model’s ability to simulate features of the characteristic synoptic patterns of daily SLP. The reference data was the NCEP reanalysis over the period of 1980–1999. The better a GCM agreed with the NCEP reanalysis, the greater the ability of the GCM. In order to find the best climate models over East Asia, the synoptic scale circulation patterns and their occurrence frequency were obtained by the SOM method. We then judged how well the GCMs compared to the NCEP reanalysis by calculating the correlation coefficient of synoptic scale circulation
pattern occurrence frequency between the two approaches. The main conclusions are summarized as follows.

a) The SOM technique could be employed as an effective tool for model assessment. Various synoptic patterns of the atmospheric circulation, including their evolution, can be identified effectively by SOM technology. The simulation ability of a model could be assessed by comparing the correlation coefficient between the modeled and NCEP reanalysis frequencies.

b) Frequencies of temporal SLP anomaly patterns can be reproduced by most of the CMIP5 models over the baseline period, and model simulation ability was better in summer than in winter. However, very few GCMs were successful in terms of spatial SLP. Only a small number of models were good at reproducing both kinds of anomalies.

c) The five top-performing models in winter are NorESM1-M, bcc-csm1-1-m, CCSM4, MRI-CGCM3, and EC-EARTH; and the top five in summer are bcc-csm1-1-m, NorESM1-M, MRI-CGCM3, CCSM4, and MPI-ESM-P. The models that perform well in both winter and summer are bcc-csm1-1-m, NorESM1-M, MRI-CGCM3, and CCSM4. Therefore, these four models should be selected preferentially for studying the synoptic pattern changes under future warming in East Asia.

Model assessment results obtained in this study can provide some guidance on mode selection for future climate projections and downscaling in further research over East Asia. However, it is important to keep in mind that model evaluation is a complicated task, and any evaluation metric has some subjectivity. Furthermore, in order to increase the reliability of the results of the model assessment, the SOM method should also be compared with other classification methods of synoptic patterns, such as K-means clustering.

Acknowledgments. We acknowledge the National Centers for Environmental Prediction (NCEP) for the NCEP reanalysis data. We acknowledge the international modeling groups, the Program for
Climate Model Diagnosis and Intercomparison (PCMDI), and the WCRP’s Working Group on Coupled Modeling (WGCM) for their roles in making available the WCRP CMIP5 multi-model data sets. We acknowledge the helpful comments of two anonymous reviewers and the editor, who helped to improve this manuscript.

References


——, L. Zou, 2010: Understanding the predictability of East Asian summer monsoon from the reproduction of land-sea thermal contrast change in AMIP-type simulation. J. Climate, 23(22): 6009-6026.