Evaluation of events of extreme temperature change between neighboring days in CMIP6 models over China

Shuaifeng Song · Xiaodong Yan

Received: 7 June 2021 / Accepted: 8 July 2022 / Published online: 19 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract
In the context of global warming, the frequency and intensity of extreme weather and climate events are increasing. However, the impact of these changes that is directly felt by people is the day-to-day temperature change. Extreme temperature changes between neighboring days (ETCNs) carry substantial disease risks and socioeconomic impacts. Evaluation studies of ETCN events with global climate models (GCMs) remain unknown in China. This study quantitatively evaluates the performances of 35 GCMs and the multi-model ensemble (MME) of the Coupled Model Intercomparison Project 6 (CMIP6) in simulating the extreme cooling (EC) and extreme warming (EW) events of two consecutive days as defined by relative thresholds. The results showed that from 1981 to 2013, the annual average EW frequencies showed an increasing trend over China, but a decreasing trend for EC events, and the frequency of EW events was higher than that of EC events. EW events mostly occurred in spring, while EC events occurred in autumn. Additionally, the performances of the CMIP6 models were quite different between EC and EW events. The models could capture the annual cycle of EC and EW events well, and the simulations of EW events were generally more reliable than those of EC events. Furthermore, most CMIP6 models overestimated the frequency of EW events but underestimated the frequency of EC events in China. The CMIP6 models could capture the trends in EC events in China but fail to simulate them in EW events. The interannual variability of EW events exhibited relatively better performance than that of EC events. The CMIP6 MME effectively improved the capabilities of the models to simulate the climatology of ETCN events. Individual CMIP6 models exhibited better performances than the CMIP6 MME in terms of the trend and interannual variability. Finally, according to the overall ranking of the CMIP6 models, MPI-ESM-1–2-HAM and FGOALS-f3-L achieved the best performance in simulating EW and EC events, respectively. This study selected the optimal models in different regions at the seasonal and annual scales, providing theoretical support for the frequency projection and modeling improvement of ETCN events.

Keywords ETCN · CMIP6 models · Multi-model ensembles · China

1 Introduction
Global climate change is a great challenge for humankind. According to the fifth assessment report of the IPCC, from 1880 to 2012, the global average surface temperature showed an upward linear trend, increasing by 0.85 °C (IPCC, 2013). Compared with the mean climate state, extreme weather and climate events are more sensitive to global warming (Aguilar et al. 2009; Ha and Yun 2012; Schoetter et al. 2015; Lewis et al. 2017), and their frequency and intensity are also increasing (Sui et al. 2018), exerting profound impacts on ecosystems and human society (Easterling et al. 2000). The day-to-day temperature changes and fluctuations induced by climate change are directly felt by people (Li and Yan 2010). An extreme temperature change between neighboring days (ETCN) refers to a dramatic change between the daily temperature of a given day and the temperature of the previous day. Recent studies found that ETCN changes not only have huge effects on ecosystems but also increase the risk of disease and death. These impacts indicate that in addition to the commonly used temperature indicators (such as the average temperature, maximum temperature, and minimum temperature), ETCN is an important meteorological index that reflects the impacts of meteorological environmental...
changes and human health-related changes (Goldberg et al. 2011; Zhan et al. 2017).

In recent years, the frequency, intensity, and duration of ETCNs have increased (Xu et al. 2015), increasing the risk of disease and death (Lin et al. 2013; Cheng et al. 2014). For example, Guo et al. (2011) showed that when the 24-h temperature decreased by 3 °C in Brisbane, the nonaccidental death rate increased by 15.7%, the cardiovascular death rate increased by 35.3%, and people under the age of 65 were most easily affected. Kang et al. (2021) also found that for a 1 °C increase in temperature variability, a 6% increase was observed in cardiovascular disease in China. More studies have found that changes in ETCNs are related to infectious diseases, such as hand, foot, and mouth disease (Xu et al. 2015; Cheng et al. 2016), pneumonia (Xu et al. 2014), and respiratory tract infections (Liu et al. 2015). In addition, ETCNs also affect the growth and reproduction of animals and plants. In Australia, studies on lizards have shown that sharp temperature fluctuations can change the shape of lizards and have a significant effect on embryo formation (Shine and Elphick 2001). Overall, China has a large population and a complex and diverse climate and is especially vulnerable to extreme weather events, resulting in serious economic losses and casualties. Given these characteristics, the simulation of ETCNs is essential for enhancing our confidence in reducing model uncertainties and developing adaptation strategies to reduce the risk of diseases, especially in different regions and seasons in China (Ren et al. 2011).

Global climate models are important tools used for climate simulations and projections of future climate change (Yang et al. 2018). The Coupled Model Intercomparison Project (CMIP) launched by the World Climate Research Program (WCRP) has evolved through CMIP1 (1995), CMIP2 (1997), CMIP3 (2004), and CMIP5 (2013) since its inception. Today, the CMIP has entered the CMIP6 phase, with CMIP6 providing data on the current simulations and future projections of approximately 112 climate models from nearly 33 institutions around the world (Zhou et al. 2019). Compared with the previous-generation CMIP5 model, CMIP6 is significantly improved in its spatial resolution, physical process, coupled carbon cycle, and so on (Eyring et al. 2016). The ability of a model to simulate the current climate has a direct impact on the accuracy of its future projections. Therefore, it is necessary to test the simulation ability of models and to objectively evaluate the accuracy and uncertainty of the simulation results (Grose et al. 2020).

Many scholars have evaluated the ability of CMIP6 models to simulate air temperature (Luo et al. 2020), extreme precipitation (Akinsanola et al. 2020), terrestrial evapotranspiration (Wang et al. 2020), sea ice (Notz and Community 2020), the Hadley circulation (Grise and Davis 2020), soil moisture (Yuan et al. 2021), wind speed (Krishnan and Bhaskaran 2020), snow depth (Chen et al. 2021), the Indian Ocean dipole (McKenna et al. 2020), the East Asian summer monsoon index (Xin et al. 2020), and so on. Regarding ETCNs, Peng et al. (2019) used daily temperature data from 804 meteorological stations across China to calculate the frequency of consecutive days with temperature differences greater than 1 °C. However, the temperature change threshold set by the author was too small to reflect extreme situations of temperature changes between neighboring days. Zhou et al. (2020) studied the spatiotemporal pattern of global day-to-day temperature changes using station-collected data and 27 CMIP6 models. The results showed that events in which the daily average temperature decreased by more than 10 °C decreased. However, the author only studied the trend changes of EC events and did not consider changes in EW events or seasonal changes. At the beginning of the twenty-first century, the World Meteorological Organization (WMO) and WCRP jointly established the Expert Team on Climate Change Detection and Indexes (ETCCDI). According to daily temperature and precipitation data, the ETCCDI defined 27 representative extreme climate indexes for the study of global and regional extreme climate change (Kim et al. 2020a, b; Akinsanola et al. 2020; Ge et al. 2021; Fan et al. 2020); the ETCN index was not defined.

Therefore, research on the evaluation of different meteorological elements or variables in China using CMIP6 models has been extensive, but to our knowledge, few relevant evaluative studies of ETCN events using CMIP6 models have been conducted in different seasons over China. To what extent do CMIP6 models capture the annual cycle, spatial pattern, trend change, and interannual variability of ETCN events? Which models should be selected as the optimal models based on a set of metrics in different zones in China? Is the performance of multi-model ensemble better than that of individual models? Motivated by these questions, the purpose of this study is to quantitatively evaluate the performance of individual CMIP6 models and the multi-model ensemble in simulating ETCN events, which are defined by relative thresholds and divided into EC events and EW events in eight subregions in China. This study provides a reliable scientific theory for the improvement of climate models and the projection of future EW and EC events in China.

2 Data and methods

2.1 Data

As a reference from observations, we used the daily maximum temperature (Tmax) or daily minimum temperature (Tmin) data provided by the National Climate Center of the China Meteorological Administration. This dataset has been widely used in research on climate change in China. The research period of this study is from January 1, 1981, to February 28, 2014. To better evaluate the ability of the
To simulate the ETCN event frequencies in different regions of China, China was divided into eight regions in this paper. These subregions over land were adapted from Wu et al. (2020) as shown in Fig. 1.

GCM daily Tmax and Tmin data were downloaded from the CMIP6 dataset (https://esgf-node.llnl.gov/search/cmip6/). The dataset covers the same period as the observations. We chose the “Experiment ID” as the “historical” term and the “Variant Label” as the “r1i1p1f1” term, for a total of 35 models. Table 1 provides an overview of the institutions, model names, and resolution information.

2.2 ETCN event definitions

A day-to-day temperature change is estimated as the difference between the daily temperature on two neighboring days (Gough, 2008; Zhan et al. 2017; Lei et al. 2020):

\[ TCN = T_i - T_{i-1} \quad (i = 1, 2, 3 \ldots n) \]  

where \( TCN \) denotes a change in the daily Tmax or Tmin for the day \( i \), and \( T_i, T_{i-1} \) denotes the Tmax or Tmin for day \( i \) (the previous day is denoted as \( i - 1 \)). The term \( n \) is the total number of days in the four seasons, which are spring (MAM), summer (JJA), autumn (SON), and winter (DJF). \( TCN < 0 \) indicates a cooling event, while \( TCN > 0 \) indicates a warming event.

Slight warming or cooling events have little impact on human health. Thus, this study only considers the occurrence of sudden strong warming or cooling events. First, the value of TCN is sorted from highest to lowest. And then, using the top and bottom tenth percentiles as the criteria (Shi et al. 2018; Tan et al. 2021), the 10th percentile and 90th percentile are thus selected as the thresholds for ETCN events in this study. Note that ETCN events are divided into extreme cooling (EC) and extreme warming (EW) events. Here, an EC event refers to a decrease in the Tmin (\( TCN < 0 \)) under the 10th percentile, while an EW event refers to an increase in the Tmax (\( TCN > 0 \)) greater than the 90th percentile (Song and Yan 2021; Cai et al. 2020; Shi et al. 2018).

2.3 Bilinear interpolation

Due to the different resolutions of CMIP6 models, the bilinear interpolation method is used in this paper to uniformly interpolate the data of each model to the meteorological stations (Jiang et al. 2020). Then, the model simulation results are compared with the observation results to evaluate the ability of each climate model to simulate ETCNs in China.

2.4 Linear trend analysis

The long-term trend of the frequencies of ETCN events was analyzed using the linear tendency estimation method (Tokarska et al. 2020). A simple linear regression was performed between the temperature variable \( y \) and the corresponding time \( x \):

\[ y = ax + b \]  

where \( a \) is a linear regression coefficient representing the rate of change in ETCN events. A positive or negative value indicates an increasing or decreasing trend of ETCN events, respectively.

The trend results were tested for significance using the \( t \) test at the 95% confidence level.

![Fig. 1 Spatial distribution of the 2298 stations and elevation of the topography over mainland China.](image)
To assess the performances of the simulations against the observations, two statistical indexes were considered: the root mean square error (\(RMSE\)) and bias (\(BIAS\)), which measure the overall accuracy of the simulations. These three indexes were calculated as follows:

\[
BIAS = \frac{1}{n} \sum_{i=1}^{n} (O_i - M_i) 
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - M_i)^2}
\]

where \(O_i\) and \(M_i\) describe the observed and simulated ETCN event frequencies and \(n\) is the number of samples; a good model will have \(BIAS\) and \(RMSE\) values of almost 0.

We selected the Taylor skill score (\(TS\)) to assess the spatial pattern of the models. The \(TS\) is defined according to Wang et al. (2018) as follows:

\[
TS = 4(1 + R)^2 / \left[ \left( \frac{\sigma_M}{\sigma_O} + \frac{\sigma_O}{\sigma_M} \right) (1 + R)^2 \right]
\]

where \(R\) indicates the spatial correlation coefficient between the observed and simulated frequencies of ETCN events and \(R\) is the maximum attainable correlation coefficient. The parameters \(\sigma_M\) and \(\sigma_O\) are the standard deviations of the simulations and observations, respectively. A score closer to 1 indicates a better consistency between the observations and simulations, while a score of 0 indicates no match at all. \(TS\) focuses on the performances of the correlation coefficient and variance, according to Guo et al. (2021).

The annual cycle is also a very crucial evaluation index for model performance, which can capture the monthly evolution of ETCN events. The annual cycle index (\(ACI\)) reflects the amplitude of seasonal variation, which is calculated as.

\[
ACI = \frac{M_{max} - M_{min}}{O_{max} - O_{min}}
\]

where \(M_{max}\) and \(M_{min}\) are the maximum and minimum values of four seasons (spring, summer, autumn, and winter) of the frequencies based on the CMIP6 model, respectively. Likewise, \(O_{max}\) and \(O_{min}\) are the observed data. Better performance is achieved when \(ACI\) is closer to 1.

To analyze the abilities of the models with respect to temporal variations, we asked the models to reproduce the temporal standard deviation. Therefore, we used the interannual variability skill (\(IVS\)) to evaluate the models (Zhu et al. 2020; Kim et al. 2020a, b). The formula is as follows:

\[
IVS = \left( \frac{STD_m}{STD_O} - \frac{STD_O}{STD_m} \right)^2
\]
where $STD_m$ and $STD_r$ are the interannual standard deviations of the simulated and observed data, respectively. Small IVS values indicate better performances between the CMIP6 models and the reference datasets (Fan et al. 2020).

### 3 Multi-model ensemble mean

To eliminate systematic errors within the model, we use the multi-model ensemble average (MME) of 35 GCMs (Mudryk et al. 2020; Bai et al. 2021), calculated as follows:

$$MME(t) = \frac{1}{N} \sum_{i=1}^{N} P_n(t)$$

(8)

where $MME(t)$ is the ensemble mean at time $t$, $N$ is the total number of GCMs, and $P_n(t)$ is the frequency of the nth GCM at time $t$.

#### 3.1 Comprehensive rating index

According to these multiple evaluation indexes, the rank of each model can be obtained. The comprehensive rating index (CRI) was introduced to calculate the comprehensive rank of each model (Rivera and Arnould, 2020; Zhang et al. 2018).

$$CRI = 1 - \frac{1}{1 \times n \times m} \sum_{i=1}^{n} rank_i$$

(9)

In the above formula, $m$ is the number of models (here, we used 35); $n$ is the number of indicators used for the evaluation; and $rank_i$ is the rank of the model according to its simulation ability based on evaluation index $i$. The closer $rank_i$ is to 1, the closer CRI is to 1, indicating a better simulation effect of the model.

### 4 Results

#### 4.1 Annual cycle

The observation results show that from 1981 to 2013, the annual average frequency of EW events in China was higher than that of EC events; these frequencies were 19.042 days and 17.980 days, respectively. In terms of different seasons, the frequencies of EW events in spring, summer, and winter are higher than those of EC events, and the frequency of EW events in autumn is slightly lower than that of EC events. EW events occur most often in spring (5.0 days) and least often in winter (4.584 days). In contrast, EC events occur most often in autumn (4.751 days) and least often in spring (4.323 days). Therefore, ETCN events are most likely to occur in the spring and autumn seasons. Additionally, obvious regional differences exist in the spatial distributions of EW and EC events. Most EW events occur in the SC region, with an average annual frequency of about 19.344 days (Figs. 2 and 3). EW events mainly occur in the WNW and SC regions in spring and autumn, respectively, and increasingly occur in the NEC regions in summer and winter. Notably, the TP region has the fewest EW events in all seasons. For the EC events, the frequency of EC events is the lowest in the WNW region, especially in spring. Generally, EC events are more frequent in the NEC region in spring, autumn, and winter, while in the TP region in summer. The annual cycle of the frequencies of EW and EC events in China is simulated by 35 CMIP6 models. The results show that the CMIP6 models can capture annual cycle features well (except for the SC region). Compared to EC events, the ACI values of EW events are generally closed to 1 in China; thus, CMIP6 models reproduce the annual cycle of EW events more robustly than that of EC events. However, CMIP6 models generally overestimate the magnitude of the annual cycle, especially in the TP region. The ACI value is in a range of 1.221 (GISS-E2-1-G) to 3.783 (INM-CM4-8) for EW events and 1.306 (CMCC-ESM) to 3.484 (MPI-ESM-1–2-HAM) for EC events. CMIP6 MME exhibits the limited ability to improve the performance of the annual cycle in China, and the improvement varies from region to region. For example, the CMIP6 MME could obviously improve the simulation ability of EW events in the NEC, WNW, and NC regions, with the ACI values of 1.020, 1.020, and 1.016, respectively. For the EC events, the CMIP6 MME shows better reproducibility of the annual cycle of EC events in the NEC region, and individual CMIP6 models are better for simulating events in other regions.

#### 4.2 Climatology

Figures 4 and 5 show the difference in the spatial distributions between the EW and EC events simulated by CMIP6 models and those observed in China. Generally, CMIP6 models can adequately capture the spatial patterns of EW and EC events in China, but some bias still exists, as shown in Figs. 6 and 7. Most CMIP6 models overestimate the frequency of EW events in China; the annual average bias ranges from approximately 0.279 days (GISS-E2-1-G) to 1.452 days (EC-Earth3-CC), and the RMSE values are 2.173 days (TaiESM1) to 3.853 days (NESM3). However, the frequency of EW events simulated by most models in the TP region in autumn is underestimated, with biases of approximately −0.589 days (INM-CM5-0) to 0.075 days (KIOST-ESM), and the RMSEs are 0.970 days (TaiESM1) to 1.572 days (NorCPM1). Generally, the SWC region simulations are the closest to the observations, especially
Fig. 2 Annual cycles of EW events frequency anomalies (relative to annual average) in China (i) and subregions (a–h). The gray and red lines represent the individual CMIP6 models and CMIP6 MME, respectively. Accordingly, the gray and red tests indicate the value of the annual cycle index (ACI). The observation is shown with black lines.

Fig. 3 Same as Fig. 2 but for EC events
Fig. 4 The spatial distribution and differences of the EW events between the observations and the selected GCMs as well as between the observations and the CMIP6 MME from 1981 to 2013. The solid black lines indicate the subregions.
in summer. The biases between the simulations and observations are approximately \(-0.179\) days (INM-CM4-8) to 0.242 days (BCC-CSM2-MR) in this region, and the RMSEs are 1.309 days (CMCC-ESM2) to 1.688 days (E3SM-1–0).

In addition, the annual average frequency of EC events in China is underestimated as a whole, which is contrary to the conclusion obtained for EW events. The biases range from approximately \(-2.517\) days (NorESM2-LM) to

Fig. 5  Same as Fig. 4 but for EC events
Fig. 6  Box-and-whisker plots for relative bias between the observations and the selected GCMs in the different subregions. The boxes indicate the interquartile model spread (range between the 25th and 75th percentiles). The red box indicates EW events, and the blue box indicates EC events. The solid green marks within the boxes show the CMIP6 MME results.

Fig. 7  Same as Fig. 6, but for RMSE.
0.476 days (INM-CM5-0), and the RMSEs are 1.915 days (FGOALS-f3-L) to 3.213 days (NorESM2-MM). However, the CMIP6 model simulations vastly overestimate autumn and winter EC events in China, especially in the TP and WNW regions, with more than 70% of the models displaying positive biases. Interestingly, unlike most CMIP6 models, INM-CM4-8 and INM-CM5-0 models overestimate the frequency of EC events in China and subregions, especially in autumn and winter.

It is indicated that the TS results of the EW events are generally higher than that of EC events over China (Fig. 8). The TS scores of most CMIP6 models simulating EW events are above 0.4. Furthermore, models from the same institution generally exhibited consistent performances. For example, the simulation abilities of EC-Earth-Consortium’s three models (EC-Earth3, EC-Earth3-Veg, EC-Earth3-Veg-LR) are relatively better, especially that of EC-Earth3-AerChem, and the TS score is 0.621 over China. Thus, models with high resolutions are revealed to output better simulations of EW events than models with low resolutions, which is consistent with the previous study (Sillmann et al., 2013). From the perspective of different seasons, the TS scores for simulations of EW events in spring are high, at 0.305 (NESM3) to 0.643 (TaiESM1), followed by those of EW events in autumn, at 0.187 (GISS-E2-1-G) to 0.542 (EC-Earth3-AerChem). However, the TS scores for the simulations of EW events in summer are generally low, ranging from 0.173 (IPSL-CM6A-LR) to 0.477 (GFDL-ESM4). Regarding the different regions, the TS scores obtained for the simulations of EW events in the WNW region are generally higher, with an annual average ranging from 0.369 (AWI-CM-1–1-MR) to 0.688 (BCC-CSM2-MR). In contrast, the TS values in summer in the SC region and in winter in the NC, WNW, and NEC regions are lower.

Unlike the EW events, the TS scores of most models simulating the annual average frequency of EC events in China exhibit a lower TS value. Among them, the TS score of NESM3 is the highest, at 0.525. In terms of different seasons, the TS scores for the simulations of spring EC events in China range from 0.173 (INM-CM4-8) to 0.588 (E3SM-1–0), followed by those for the simulations of autumn EC events, which range from 0.181 (MPI-ESM-1–2-HAM) to 0.485 (KIOST-ESM); the TS scores of the winter EC event are generally low, ranging from 0.073 (EC-Earth3-CC) to 0.418 (NESM3). Regarding the different regions, the TS scores for the simulations of EW events in the WNW region are generally higher, with an annual average ranging from 0.369 (AWI-CM-1–1-MR) to 0.688 (BCC-CSM2-MR). The TS scores obtained in the WNW region are lower in spring, summer and winter, but the TS values are higher in autumn ranging from 0.225 (CanESM5) to 0.673 (NESM3).

The improvement of CMIP6 MME varies from evaluation indexes and regions. According to the result of BIAS, CMIP6 MME failed to improve the simulation ability. Remarkably, in terms of RMSE, CMIP6 MME can effectively outperform the individual CMIP6 models in all

![Fig. 8](image-url) Same as Fig. 6, but for TS

© Springer
regions. CMIP6 MME exhibits higher performance for the TS score in simulating EC events in autumn in the WNW and winter ENW regions. The TS value is 0.709 and 0.415, respectively, which is better than all individual CMIP6 models.

5 Trend analysis

The observation results reveal that the annual average EW frequencies show a significant increasing trend (1.092 days/10a) over China. The increasing rate of EW events is fastest in summer (0.311 days/10a) and slowest in autumn (0.1 days/10a). As determined from the spatial distributions, except for the NEC, WNW, and TP regions, the annual average EW frequencies in most regions are significant upward trends. The increasing rate of EC events is fastest over SWC regions, especially in summer (0.838 days/10a). As a comparison, the annual average EC event frequencies show a decreasing trend (−0.317 days/10a) with no statistical significance from 1981 to 2013 in China. The decreasing rate of EC events is fastest in NC and TP regions, with a rate of −0.748 days/10a ($P < 0.05$) and −1.056 days/10a ($P < 0.05$), respectively. However, in ENW and NEC regions, the frequencies of EC events show an increasing trend, with a rate of 0.221 days/10a ($P > 0.05$) and 0.325 days/10a ($P > 0.05$), respectively. In terms of the four seasons, the trend of EC event frequencies in summer, autumn, and winter exhibits consistency for the annual average frequencies in China, as previously noted. However, in spring, the frequencies of EC events show a slight upward trend (0.065 days/10a), notably in ENW and JH regions.

Whether a model can simulate the trend changes in EW and EC events is an important index used to examine the simulation ability of the model. As shown in Figs. 9 and 10, large differences exist among the simulation results of the models, especially for EW events. Most CMIP6 models could not reflect the trend changes in EW or EC events in China, and a few models have optimal simulation results. For the simulation of EW events, 11 models display a better performance in simulating the increasing trend of the annual average frequency of EW events in China (Fig. 11). The NorESM2-MM model is the closest to the observations, with a rate of 0.734 days/10a. Most models could simulate the observed increasing trend in winter but poor performance in summer. Generally, the spring NorESM2-MM, summer IITM-ESM, autumn GISS-E2-1-G, and winter AWI-ESM-1–1-LR simulate trends of EW events in China are good agreement with the observations. From the perspective of different regions, the models could exhibit the decreasing trend of EW events in the NEC region in summer.

Fig. 9 CMIP6 models simulating the time series of the frequency of EW events in China
and autumn, and could also capture the increasing trend of winter EW events in the ENW and JH regions. However, CMIP6 models have poor performance for reproducing the trend in the JH, ENW, and SWC regions in summer, and the NC regions in autumn.

Compared to the EW events, most models show a relatively better performance for the simulation in the trend of the frequencies of EC events (Fig. 12). Except for the IITM-ESM model, CMIP6 models could reasonably simulate the decreasing trend of the annual average frequencies of EC events in China. The outputs of the FGOALS-f3-L model are closest to the observations, with a rate of $-0.319$ days/10a. Most models could simulate the decreasing trend in summer and winter but poor performance for spring and autumn. The spring CanESM5, summer EC-Earth3-Veg, autumn IPSL-CM6A-LR, and winter ACCESS-CM2 simulated trends of EC events in China are in good agreement with the observations in China, and the optimal models in four seasons are different from the EW events. Regarding the subregions, more than 50% of the models could simulate the trends of EC events in the NC in all seasons. Most models show a better performance for the trends of EC events in the NC, JH, SC, SWC, and ENW regions in summer, while poor performance in the NEC region.

CMIP6 MME could reasonably capture the decreasing trends in EC events in China; however, it fails to simulate the increasing trends in EW events (Fig. 13). In addition, the EC and EW events simulated by CMIP6 MME are quite different among different regions. The CMIP6 MME could obviously improve the simulation ability of winter EC events in China, especially in the JH region. The trend obtained is $-0.107$ days/10a, which is second to the FGOALS-f3-L model and is better than the outputs of another individual model. In terms of EW events, the CMIP6 MME does not outperform the individual model in China. Regionally, a significant improvement in CMIP6 MME has been found in the NEC, TP, and WNW regions in autumn, especially in the NEC region in autumn. The trend of EC events simulated by CMIP6 MME is $-0.133$ days/10a in the NEC region, which is close to the observed trend ($-0.153$ days/10a). However, for the other regions, the effect of the CMIP6 MME simulation is not as good as that of the simulations of most individual models. Although some models could successfully simulate the trends of the EW and EC events frequencies, the simulations of the change rates are much weaker and differ significantly from the observed results.

### 5.1 Interannual variability

The assessment of temporal variability is another index used to measure the simulation ability of a model. In this
Fig. 11 Spatial distribution of annual EW events trends during 1981–2013 from the observations, MME, and the selected GCMs. The black solid lines indicate the subregions.
Fig. 12  Same as Fig. 11 but for EC events
In this study, different evaluation indexes are used to select the optimal individual CMIP6 models in simulating the annual cycles, spatial patterns, climatology, interannual variability, and trends of EW and EC events in China. Because of the discrepancy inherent in model ranking, a comprehensive evaluation index (CRI) is adopted to select the most
skillful models in different regions and seasons based on the models’ overall rankings. As shown in Tables 2 and 3, the optimal models for simulating EW and EC events differed among different seasons and regions. For the annual average EW events in China, MPI-ESM-1–2-HAM has the best simulation performance. TaiESM1 in summer, AWI-ESM-1–1-LR in winter, and NorESM2-MM in spring and autumn are the preferred models for simulating EW events. For the annual average EC events in China, the FGOALS-f3-L model has the best simulation performance, and the spring TaiESM1, summer FGOALS-g3, autumn IPSL-CM6A-LR, and winter MPI-ESM-1–2-HAM are the best models for EC events. Notably, the high-resolution models perform better for EW events in the NC, JH, and SC regions and EC events in the NEC and JH regions, whereas this expectation is not extendable to other regions.

It is worth noting that the CMIP6 MME significantly improves the climatology simulations of EW and EC events in winter in China. This is mainly because the CMIP6 MME reduces the bias and uncertainty among climate models and outputs information that is closer to the observed results (Wang et al. 2018; Almazroui et al. 2017). However, the

| Table 2 | The optimal CMIP6 models for EW events by different evaluation indicators based on the CRI |
|-----------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
|               | Spring                          | Summer                         | Autumn                          | Winter                          | Annual                          |
| NEC             | MPI-ESM-1–2-HAM                 | GISS-E2-1-G                    | ACCESS-M1-5                     | CMCC-ESM2                       | CMCC-ESM2                       |
| NC              | FGOALS-f3-L*                    | TaiESM1*                       | ACCESS-M1-1-G                   | GFDL-ESM4                       | TaiESM1*                        |
| JH              | TaiESM1*                        | TaiESM1*                       | ACCESS-M1-5                     | AWI-ESM-1–1-LR                  | CMCC-ESM2                       |
| SC              | TaiESM1*                        | EC-Earth3-CC*                  | TaiESM1*                        | AWI-ESM-1–1-LR                  | NorESM2-MM                      |
| SWC             | TaiESM1*                        | IITM-ESM                       | NorESM2-MM                      | ACCESS-CM2                      | GISS-E2-1-G                     |
| TP              | ACCESS-CM2                      | EC-Earth3-Veg-LR               | MPI-ESM1-2-HR*                  | MPI-ESM1-2-LR                   | EC-Earth3-Veg-LR                |
| WNW             | NorCPM1                         | MIROC6                         | EC-Earth3-CC*                   | CMCC-ESM2                       | GISS-E2-1-G                     |
| ENW             | NorESM2-MM                      | TaiESM1*                       | NorESM2-MM                      | AWI-ESM-1–1-LR                  | MPI-ESM-1–2-HAM                |
| China           | NorESM2-MM                      | TaiESM1*                       | NorESM2-MM                      | AWI-ESM-1–1-LR                  | MPI-ESM-1–2-HAM                |

*“* indicates a high-resolution model (resolution over 1°).
CMIP6 MME does not significantly improve the overall performance of EW and EC events in China, and the CMIP6 MME is not the best scheme. A possible reason for this result is that the CRI score includes the IVS index value. Previous studies have shown that the CMIP6 MME has a large bias when simulating interannual variability, which may be due to the models of the CMIP6 MME canceling each other out and smoothing the natural variability within the climate system. Thus, the long-term trends of EW and EC events are underestimated, which is consistent with the conclusions of temperature studies conducted by Zhao et al. (2014) and Li et al. (2019). Therefore, the CMIP6 MME method should be used cautiously when simulating the trends and interannual variability rates of EW and EC events.

6 Discussion and conclusion

In this paper, we quantitatively assessed the performances of 35 CMIP6 GCMs in simulating the annual cycle, spatial pattern, climatology, interannual variability, and trend in ETCN events in China and different regions within China using multiple assessment methods and CRI metrics combined with daily Tmin and Tmax. These results provided a reference for model improvements and projections in the future. The following conclusions were obtained.

The observation results show that from 1981 to 2013, the annual average frequency of EW events in China is higher than that of EC events. EW events mostly occur in spring, while EC events mainly occur in autumn. From the trend, the annual average EW frequencies show an increasing trend over China but a decreasing trend for EC events.

The 35 CMIP6 models could robustly reproduce the annual cycles of EW and EC events well. The simulation of the annual cycle of EW events is better than that of EC events. The CMIP6 models could capture the spatial patterns of EW and EC events in China. However, most CMIP6 models overestimate the frequency of EW events but underestimate the frequency of EC events in China. The TS scores of the EW events are generally higher than that of EC events over China.

Most CMIP6 models could capture the trends in EC events in China but fail to simulate the trend in EW events. The interannual variability of EW events exhibits a relatively better performance than that of EC events. Combining the evaluation indicators of the annual cycle, spatial pattern, climatology, trend and the interannual variability in EW events and EC events, it can be concluded that the optimal models for simulating EW and EC events differ among different seasons and regions in China. MPI-ESM-1–2-HAM best simulates EW events, while FGOALS-f3-L best simulates EC events in China.

Although the CMIP6 climate models could reproduce the spatial pattern of the frequency of ETCN events in China, some biases still exist in the simulation of ETCN events. Previous studies have shown that uncertainties in climate models mainly result from the subgrid parameterization process (Bony and Dufresne 2005; Luan et al. 2016). Furthermore, dynamic downscaling was a useful way to reduce model uncertainty based on the best model for simulating ETCN events, as was selected here; this method could improve simulations of regional climate changes (Seo and Ok 2013). To eliminate model uncertainty, we could improve the resolution of the climate model to reduce its dependence on the physical parameterization process (Randall et al. 2003; Eyring et al. 2019; Paik et al. 2020). The high-resolution model effectively improved the model performance in simulating the EW and EC events, especially in the JH region. However, this improvement was difficult to apply to the TP regions. This indicated that physical processes play an important role in areas with complex topography, which is still a challenging topic for the improvement of global climate models (Jiang et al. 2015).

In addition, the uncertainty of observed data was another main cause of biases, such as the data uncertainty in the TP region related to the sparse station density (Yin et al. 2019). In the future, multisource observation data would be used

| Table 3 Same as Table 2 but for EC events |
|----------------------------------------|
| Spring | Summer | Autumn | Winter | Annual |
| NEC    | EC-Earth3-AerChem* | EC-Earth3* | EC-Earth3-CC* | INM-CM4-8 | KIOST-ESM |
| NC     | EC-Earth3* | FGOALS-g3 | KIOST-ESM | ACCESS-CM2 | GFDL-ESM4 |
| JH     | FGOALS-f3-L* | FGOALS-g3 | IPSL-CM6a-LR | FGOALS-f3-L* | EC-Earth3-Veg* |
| SC     | GFDL-ESM4 | TaiESM1* | GFDL-ESM4 | GFDL-CM4 | NESM3 |
| SWC    | GFDL-ESM4 | CMCC-ESM2 | MPI-ESM1–1–2-HAM | AWI-ESM1–1–1-LR | CMCC-ESM2 |
| TP     | GISS-E2-1-G | NEMS3 | GFDL-ESM4 | ACCESS-EASM1–5 | ACCESS-EASM1–5 |
| WNW    | KIOST-ESM | GFDL-ESM4 | EC-Earth3-CC* | GFDL-CM4 | GFDL-ESM4 |
| ENW    | GISS-E2-1-G | GFDL-ESM4 | ACCESS-EASM1–5 | GFDL-ESM4 | GISS-E2-1-G |
| China  | TaiESM1* | FGOALS-g3 | IPSL-CM6a-LR | MPI-ESM1–1–2-HAM | FGOALS-f3-L* |
for comparative analyses to improve the accuracy of model evaluations (Li and Yan 2010). The MME method used in this study was an equal-weight ensemble. Each model could be given different weights according to the ranking of the models’ performance to obtain more accurate simulation results (Knutti 2010); this would be the content of our next work.

Acknowledgements We would like to thank the China Meteorological Administration for temperature data and information. We also acknowledge the Coupled Model Intercomparison Project (CMIP) launched by the World Climate Research Program (WCRP) for CMIP6 data.

Author contribution All authors have read and agreed to the published version of the manuscript. SS analyzed data and wrote original draft. XY conceived and designed the study.

Funding This work was supported by the National Key Research and Development Program of China (grant number 2019YFA0606904).

Code availability MATLAB code developed for this study is available from the authors upon request.

Data availability All data sets used in this study are publicly available.

Declarations

Ethics approval The manuscript is conducted within the ethical manner advised by the Theoretical and Applied Climatology. Permissions or licenses were obtained.

Consent to participate Not applicable.

Consent for publication Authors agree to publish this paper.

Conflict of interest The authors declare no competing interests.

References

Aguilar E, Aziz Barry A, Brunet M, Ekang L, Fernandes A, Massoukina M, Mbah J, Mhandado Nascimento Peterson Thamba Umba Tomou Zhang ADJTCOMX (2009) Changes in temperature and precipitation extremes in western central Africa, Guinea Conakry, and Zimbabwe, 1955-2006. J Geophys Res Atmos 114(D2):356–360. https://doi.org/10.1029/2008JD011010

Akinsanola AA, Kooperman GJ, Pendergrass AG, Hannah WM, Reed KA (2020) Seasonal representation of extreme precipitation indices over the United States in CMIP6 present-day simulations. Environ Res Lett 15(9):094003. https://doi.org/10.1088/1748-9326/ab92c1

Akinsanola AA, Ongoma V, Kooperman GJ (2021) Evaluation of CMIP6 models in simulating the statistics of extreme precipitation over Eastern Africa. Atmos Res 254:105509. https://doi.org/10.1016/j.atmosres.2021.105509

Almazroui M, Islam MN, Saeed S, Alkalhaf AK, Dambul R (2017) Assessment of uncertainties in projected temperature and precipitation over the Arabian Peninsula using three categories of Cmip5 multimodel ensembles. Earth Syst Environ 1(2):1–20. https://doi.org/10.1007/s41748-017-0027-5

Bai H, Xiao D, Wang B, Liu DL, Feng P, Tang J (2021) Multi-model ensemble of CMIP6 projections for future extreme climate stress on wheat in the North China plain. Int J Climatol 41:E171–E186. https://doi.org/10.1002/joc.6674

Cai Q, Guan ZY, Xu M (2020) Temporal and spatial characteristics of extreme cooling events in eastern China in winter monsoon period during 1960 and 2012. Trans Atmos Sci 43(3):458–468. https://doi.org/10.13878/j.cnki.dqkkxb.20190719001

Chen S, Liu J, Ding Y, Zhang Y, Cheng X, Hu Y (2021) Assessment of snow depth over Arctic sea ice in CMIP6 models using satellite data. Adv Atmos Sci 38(2):168–186. https://doi.org/10.1007/s00376-020-0213-5

Cheng J, Zhu R, Xu Z, Xu X, Wang X, Li K, Su H (2014) Temperature variation between neighboring days and mortality: a distributed lag non-linear analysis. Int J Public Health 59(6):923–931. https://doi.org/10.1007/s00385-014-0611-5

Cheng J, Zhu R, Xu Z, Wu J, Wang X, Li K, Wen L, Yang H, Su H (2016) Impact of temperature variation between adjacent days on childhood hand, foot and mouth disease during April and July in urban and rural Hefei, China. Int J Biometeorol 60(6):883–890. https://doi.org/10.1007/s00484-015-1082-y

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO (2000) Climate extremes: observations, modeling, and impacts. Science 289(5487):2068–2074. https://doi.org/10.1126/science.289.5487.2068

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(5):1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Eyring V, Cox PM, Flato GM, Gleckler PJ, Abramowitz G, Caldwell P, Collins WD, Gier BK, Hall AD, Hoffman FM, Hurtt GC, Jahn A, Jones CD, Klein SA, Krasting JP, Kwiatkowski L, Lorenz R, Maloney E, Meehl GA, Pendergrass AG, Pincus R, Ruane AC, Russell JL, Sanderson BM, Santer BD, Sherwood SC, Simpson IR, Stouffer RJ, Williamson MS (2019)Taking climate model evaluation to the next level. Nat Clim Change 9:102–110. https://doi.org/10.1038/s41558-018-0355-y

Ge F, Zhu S, Luo H, Zhi X, Wang H (2021) Future changes in precipitation extremes over Southeast Asia: insights from CMIP6 multi-model ensemble. Environ Res Lett 16(2):024013. https://doi.org/10.1088/1748-9326/abd7ad

Goldberg MS, Gasparuzzi A, Armstrong B, Valois MF (2011) The large change in temperature between neighbouring days increases mortality: a distributed lag non-linear analysis. Int J Public Health 59(6):923–931. https://doi.org/10.1007/s00385-014-0611-5

Gough WA (2008) Theoretical considerations of day-to-day temperature variability applied to Toronto and Calgary Canada data. Theor Appl Climatol 94(1):97–105. https://doi.org/10.1007/s00704-007-0346-9

Grise KM, Davis SM (2020) Hadley cell expansion in CMIP6 models. Atmos Chem Phys 20(9):5249–5268. https://doi.org/10.5194/acp-20-5249-2020

Guo Y, Barnett AG, Yu W, Pan X, Ye X, Huang C, Tong S (2011) A short-term influence of temperature on neurodevelopmental outcomes in children with non-syndromic autism spectrum disorder. Autism Res 4(4):245–253. https://doi.org/10.1002/aur.129

Guo G, Bao A, Chen T, Zheng G, Wang Y, Jiang L, De Maeyer P, Assouline S, Grams S, Han J (2015) Assessment of CMIP6 in simulating precipitation over arid Central Asia. Atmos Res 252:105451. https://doi.org/10.1016/j.atmosres.2021.105451

Ha KJ, Yun KS (2012) Climate change effects on tropical night days in Seoul, Korea. Theor Appl Climatol 109(1–2):191–203. https://doi.org/10.1007/s00704-011-0573-y
Jiang Z, Li W, Xu J, Li L (2015) Extreme precipitation indices over China in CMIP5 models Part i: Model Evaluation. J Climate 28(21):8603–8619. https://doi.org/10.1175/JCLI-D-15-0009.1

Jiang J, Zhou T, Chen X, Zhang L (2020) Future changes in precipitation over Central Asia based on CMIP6 projections. Environ Res Lett 15(5):054009. https://doi.org/10.1088/1748-9326/ab7d03

Kim YH, Min SK, Zhang X, Sillmann J, Sandstad M (2020) Evaluation of the CMIP6 multi-model ensemble for climate extreme indices. Weather Clim Extreme 29:100269. https://doi.org/10.1016/j.wace.2020.100269

Knutti R (2010) The end of model democracy? Climatic Change 102:395–404. https://doi.org/10.1007/s10584-010-9800-2

Krishnam A, Bhaskaran PK (2020) Skill assessment of global climate model wind speed from CMIP5 and CMIP6 and evaluation of projections for the Bay of Bengal. Clim Dynam 55(9):2667–2687. https://doi.org/10.1007/s00382-020-05406-z

Lewis SC, King AD, Perkins-Kirkpatrick SE (2017) Defining a new normal for extremes in a warming world. B Am Meteorol Soc 98(6):1139–1151. https://doi.org/10.1175/BAMS-D-16-0183.1

Li Z, Yan Z (2010) Application of multiple analysis of series for homogenization to Beijing daily temperature series (1960–2006). Adv Atmos Sci 27(4):777–787. https://doi.org/10.1007/s00376-010-9052-0

Li X, Xu C, Li L, Luo Y, Yang Q, Yang Y (2019) Evaluation of air temperature of the typical river basin in desert area of Northwest China by the CMIP5 models: a case of the Kaidu-Kongqi River Basin. Resources Science 41(6):1141–1153. https://doi.org/10.1007/s00477-019-8512-1

Lin H, Zhang Y, Xu Y, Xu X, Liu T, Luo Y, Xiao J, Wu W, Ma W (2013) Temperature changes between neighboring days and mortality in summer: a distributed lag non-linear time series analysis. PLoS one 8(6):e66403. https://doi.org/10.1371/journal.pone.0066403

Liu Y, Guo Y, Wang C, Li W, Lu J, Shen S, Xia H, He J, Qiu X (2015) Association between temperature change and outpatient visits for respiratory tract infections among children in Guangzhou, China. Int J Env Res Pub He 12(1):439–454. https://doi.org/10.3390/ijerph120100439

Luan Y, Yu Y, Zheng W (2016) Review of development and application of high resolution global climate system model. Advance in Earth Sci 31(3):258–268. https://doi.org/10.11867/j.issn.1001-8186.2016.03.0258

Luo N, Guo Y, Guo Z, Chen K, Chou J (2020) Assessment of CMIP6 and CMIP5 model performance for extreme temperature in China. Atmos Ocean Sci Lett 13(6):589–597. https://doi.org/10.1008/16742834.2020.1808430

McKenna S, Santosso A, Gupta AS, Taschetto AS, Cai W (2020) Indian Ocean Dipole in CMIP5 and CMIP6: characteristics, biases, and links to ENSO. Sci Rep-UK 10(1):1–13. https://doi.org/10.1038/s41598-020-68268-9

Mudryk L, Santolaria-Otín M, Krinner G, Méndez M, Derksen C, Brutel-Vuilmet C, Brady M, Essery R (2020) Historical Northern Hemisphere snow cover trends and projected changes in the CMIP6 multi-model ensemble. Cryosphere 14(7):2495–2514. https://doi.org/10.5194/tc-14-2495-2020

Peng K, Deng J, Zhang Y, Gong Z, Qin B (2019) Short-term temperature fluctuation in the spring in China during 1957–2015. Climatic and Environmental Research 24(1):125–134. https://doi.org/10.3878/j.issn.1006-9585.2018.17060

Randall D, Khairoutdinov M, Arakawa A, Grabowski W (2003) Breaking the cloud parameterization deadlock. B Am Meteorol Soc 84(11):1547–1564. https://doi.org/10.1175/BAMS-84-11-1547

Ren G, Chen Y, Zou XK, Zhou YQ, Zhang L (2011) Change in climatic extremes over mainland China based on an integrated extreme climate index. Clim Res 50(1–2):113–124. https://doi.org/10.3354/cr01023

Schoetter R, Cattiaux J, Douville H (2015) Changes of western European heat wave characteristics projected by the CMIP5 ensemble. Clim Dynam 45(5–6):1601–1616. https://doi.org/10.1007/s00382-014-2434-8

Seo KH, Ok J (2013) Assessing future changes in the East Asian summer monsoon using CMIP3 models: results from the best model ensemble. J Climate 26:1807–1817. https://doi.org/10.1175/JCLI-D-12-00109.1

Shi J, Cui L, Wen K, Tian Z, Wei P, Zhang B (2018) Trends in the consecutive days of temperature and precipitation extremes in China during 1961–2015. Environ Res 161:381–391. https://doi.org/10.1016/j.envres.2017.11.037

Shine R, Elphick MJ (2001) The effect of short-term weather fluctuations on temperatures inside lizard nests, and on the phenotypic traits of hatchling lizard. Biol J Linn Soc 72(4):555–565. https://doi.org/10.1046/j.1099.1

Sillmann J, Kharin VV, Zhang X, Zwiers FW, Bronaugh D (2013) Climate extremes indices in the CMIP5 multimodel ensemble Part 1 Model evaluation in the present climate. J Geophys Res-Atmos 118(4):1716–1733. https://doi.org/10.1002/jgrd.50203

Song S, Yan X (2021) Changes in the frequency of extreme cooling events in winter over China and their relationship with Arctic oscillation. Sustainability 13(20):11491. https://doi.org/10.3390/su132011491

Sui Y, Lang X, Jiang D (2018) Projected signals in climate extremes over China associated with a 2 C global warming under two RCP scenarios. Int J Climatol 38:e678–e697. https://doi.org/10.1002/joc.5399

Tan ML, Juneng L, Tangang FT, Chung JX, RadinFirdaus RB (2021) Changes in temperature extremes and their relationship with ENSO in Malaysia from 1985 to 2018. Int J Climatol. https://doi.org/10.1002/joc.6864

Tianjun Zhou, Liwei Zou, Xiaolong Chen (2019) Commentary on the coupled model intercomparison project phase 6 (CMIP6). Clim Chang Res 15(5):445–456

Wang B, Zheng L, Liu DL, Ji F, Clark A, Yu Q (2018) Using multi-model ensembles of CMIP5 global climate models to reproduce observed monthly rainfall and temperature with machine learning methods in Australia. Int J Climatol 38(13):4891–4902. https://doi.org/10.1002/joc.5705

Wang Z, Zhan C, Ning L, Guo H (2020) Evaluation of global terrestrial evaportranspiration in CMIP6 models. Theor Appl Climatol 143:521–531. https://doi.org/10.1007/s00704-020-03437-4

Wu Y, Miao C, Duan Q, Shen C, Fan X (2020) Evaluation and projection of daily maximum and minimum temperatures over China using the high-resolution NEX-GDDP dataset. Clim Dynam 55(9):2615–2629. https://doi.org/10.1007/s00382-020-05404-1

Xin X, Wu T, Zhang J, Yao J, Fang Y (2020) Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon. Int J Climatol 40(15):6423–6440. https://doi.org/10.1002/joc.6590

Xu ZW, Hu WB, Tong SL (2014) Temperature variability and childhood pneumonia: an ecological study. Environ Health-Glob 13:51. https://doi.org/10.1186/1476-069X-13-51

Xu J, Zhao D, Su H, Xie M, Cheng J, Wang X, Li K, Yang H, Wen L, Wang B (2015) Impact of temperature variability on childhood hand, foot and mouth disease in Huainan, China. Public Health 134:86–94. https://doi.org/10.1016/j.puhe.2015.10.029

Yang T, Tao Y, Li J, Zhu Q, Su Lu, He X, Zhang X (2018) Multicriterion model ensemble of CMIP5 surface air temperature over China. Theor Appl Climatol 132(3):1057–1072. https://doi.org/10.1007/s00704-017-2143-4
Yin H, Sun Y, Donat MG (2019) Changes in temperature extremes on the Tibetan Plateau and their attribution. Environ Res Lett 14(12):124015. https://doi.org/10.1088/1748-9326/ab503c

Zhan Z, Zhao Y, Pang S, Zhong X, Wu C, Ding Z (2017) Temperature change between neighboring days and mortality in United States: a nationwide study. Sci Total Environ 584:1152–1161. https://doi.org/10.1016/j.scitotenv.2017.01.177

Zhang Y, You Q, Chen C, Ge J, Adnan M (2018) Evaluation of down-scaled CMIP5 coupled with VIC model for flash drought simulation in a humid subtropical basin, China. J Clim 31:1075–1090. https://doi.org/10.1175/JCLI-D-17-0378.1

Zhao T, Chen L, Ma Z (2014) Simulation of historical and projected climate change in arid and semi-arid areas by CMIP5 models. Sci Bull 59:412–429. https://doi.org/10.1111/j.1432-0003-2013-

Zhu H, Jiang Z, Li J, Li W, Sun C, Li L (2020) Does CMIP6 inspire more confidence in simulating climate extremes over China? Adv Atmos Sci 37(10):1119–1132. https://doi.org/10.1007/s00376-020-9289-1

Bony S, Dufresne JL (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. Geophys Res Lett 32(20). https://doi.org/10.1029/2005GL023851

Fan X, Miao C, Duan Q, Shen C, Wu Y (2020). The performance of CMIP6 versus CMIP5 in simulating temperature extremes over the global land surface. J Geophys Res-Atmos 125(18):e2020JD033031. https://doi.org/10.1029/2020JD033031

Gao Z, Zhu J, Guo Y, Luo N, Fu Y, Wang T (2021). Impact of land surface processes on a record-breaking rainfall event on May 06 07 2017 in Guangzhou China. J Geophys Res-Atmos 126(5):e2020JD032997

Grose MR, Narsey S, Delage FP, Dowdy AJ, Bador M, Boschat G, Boschat G, Chung C, Kajtar JB, Rauniyar S, Freund MB, Lyu K, Rashid H, Zhang X, Wales S, Trenham C, Holbrook NJ, Cowan T, Alexander L, Arblaster JM, Power S (2020). Insights from CMIP6 for Australia’s future climate. Earths Future 8(5): e2019EF001469 https://doi.org/10.1029/2019EF001469

IPCC ARS WGI Climate change: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York: Cambridge University Press, 2013.

Kang Y, Tang H, Zhang L., Wang S, Wang X, Chen Z, Zheng C, Yang Y, Wang Z, Huang G, Gao R (2021). Long-term temperature variability and the incidence of cardiovascular diseases: A large, representative cohort study in China. Environ Pollut. 116831. https://doi.org/10.1016/j.envpol.2021.116831

Kim MK, Yu DG, Oh JS, Byun YH, Boo KO, Chung IU, Park JS, Min SK, Sung HM (2020). Performance evaluation of CMIP5 and CMIP6 models on heatwaves in Korea and associated teleconnection patterns. J Geophys Res-Atmos. 125(23): e2020JD032583. https://doi.org/10.1029/2020JD032583

Lei X, Liu L, Chen R, Liu C, Hong J, Cao L, Lu Y, Dong X, Chen X, Qiu X, Xia M, Ding B, Qian L, Wang L, Zhou W, Gui Y, Kan H, Zhou Y, Zhang X (2020). Temperature changes between neighboring days and childhood asthma: a seasonal analysis in Shanghai, China. Int J Biometeorol, 1–10 https://doi.org/10.1007/s00484-020-02057-x

Notz D, Community S (2020). Arctic sea ice in CMIP6. Geophys Res Lett 47(10):e2019GL086749. https://doi.org/10.1029/2019GL086749

Paik S, Min SK, Zhang X, Donat MG, King AD, Sun Q (2020). Determining the anthropogenic greenhouse gas contribution to the observed intensification of extreme precipitation. Geophys Res Lett 47 (12). 10.1029/2019GL086875 e2019GL086875

Rivera JA, Arnould G (2020). Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: climatic features and long-term trends (1901–2014). Atmos Res 241. https://doi.org/10.1016/j.atmosres.2020.104953

Tokarska KB, Stolpe MB, Sippel S, Fischer EM, Smith CJ, Lehner F, Knutti R (2020). Past warming trend constrains future warming in CMIP6 models. Sci Adv 6(12): eaaz9549. https://doi.org/10.1126/sciadv.aaz9549

Yuan S, Quiring SM, Leaor ZT (2021). Historical changes in surface soil moisture over the contiguous United States: an assessment of CMIP6. Geophys Res Lett 48(1): e2020GL089991. https://doi.org/10.1029/2020GL089991

You Q, Wu F, Wang H, Jiang Z, Pepin N, Kang S (2020). Projected changes in snow water equivalent over the Tibetan Plateau under global warming of 1.5° and 2° C. J Climate 33(12):5141-5154. https://doi.org/10.1175/JCLI-D-19-0719.1

Zhou X, Wang Q, Yang T (2020). Decreases in days with sudden day-to-day temperature change in the warming world. Global Planet Change 192:103239. https://doi.org/10.1016/j.gloplacha.2020.103239

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.