RESEARCH

Assessment and improvement of RegCM 4.6 coupled with CLM4.5 in simulation of land surface temperature in mainland China

Yulong Ren[1] · Xiaoqing Gao[2] · YuanPu Liu[1] · ZhenChao Li[2] · WeiGang Liu[1]

Abstract

Land surface temperature(LST) is an important indicator to study climate change and test the performance of regional climate model simulation. RegCM4.6 is the representative version of regional climate model RegCM, which is coupled with advanced third-generation land surface model NCAR CLM4.5. Currently, RegCM4.6 has become an important tool to study regional climate change in China. However, its ability to simulate land surface temperature in mainland China and the reasons for its deviation have not been systematically studied, and targeted improvement work is lacking. The present study is the first to employ LST data collected from 809 Chinese meteorological stations from the last 30 years to comprehensively assess the ability of CLM4.5 to simulate LST. Sensitivity tests of soil thermal conductivity (STC) were carried out to improve the model. Although the coupled regional climate model could accurately simulate the temporal and spatial variation of LST, a cold bias of 2~8 °C existed for all of mainland China, which was larger in seasons with more precipitation and greater soil moisture than other seasons. Deviation increased from southeast to northwest. which was caused by the incoming long-wave radiation, sensible heat, and latent heat simulated. There was a significant linear relationship between the observed and simulated LSTs, with correlation coefficients for all the stations ranged from 0.75 to 0.9 ($P < 0.001$). The observed LST increased at a rate of 0.58 °C/decade, but the simulated LST increased at a lower rate. Assessment of three different STC schemes showed that the Lu-Ren scheme was the most suitable for LST simulation in mainland China. Developing a new STC scheme that considers the role of water vapor can effectively improve the model when used in mainland China.

1 Introduction

Changes in land surface temperature (LST) can alter the balance of energy and material between the land and atmosphere, and cause major changes in precipitation, temperature, vegetation, and ecological processes (Wilson et al. 2003; Zhong et al. 2011). Thus, LST is an important indicator used for studying global climate change (Wan and Li 1997; Coll et al. 2016; Duan et al. 2017; Jones and Trewin 2015). LST is calculated by a land surface model (LSM) of a coupled climate model. The accuracy of those calculations has a direct impact on the process of simulating conditions in soil water, heat, and in simulating ecological processes. Therefore, LST is also one of the main indicators used to assess regional climate model (RCM) performance.

The RegCM is a regional climate model established by Dickinson and Giorgi in the late 1980s through expansion and modification of the radiation scheme, convection parameterization scheme, and land surface physical processes in a mesoscale model MM4 (Dickinson et al. 1989). Giorgi et al. (1993) subsequently produced RegCM2, RegCM3, and RegCM4 by improving the physical process scheme and mesoscale model. The latest mature version is RegCM4.6. In this version, a mesoscale model MM5 non-static dynamic frame option was added, which improves the model spatial resolution to 10 km and updates the radiation and convection parameterization schemes. The most widely used regional climate model in China, RegCM, is not only used for climate simulation and diagnosis but is also used as supporting tools in climate prediction.

The Community Land Model (CLM) (Zeng et al. 2002), developed by the National Center for Atmospheric Research

✉ Yulong Ren
renyl@iamcma.cn

Xiaoqing Gao
xqgao@lzb.ac.cn

[1] Institute of Arid Meteorology, CMA, Key Laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province, Key Laboratory of Arid Climate Change and Disaster Reduction of CMA, Lanzhou 730020, China

[2] Northwest Institute of Eco-Environment and Resources, CAS, Lanzhou, China

(NCAR) in the USA and based on 2nd-generation LSMs such as BATS, IAP94, and NCAR-LSM, is a typical 3rd-generation LSM. It has ten uneven soil layers, five snowfall layers, and one vegetation layer. The data of land surface cover include soil color, soil texture, percent coverage of plant functional types (PFTs) per grid, as well as leaf and stem area indices. A CLM classifies surface vegetation into 17 PFTs. Each grid point can contain 17 different PFTs, which are treated as the percentage of each PFT area to the grid area. This includes physical, chemical, hydrological, and biochemical processes such as biogeophysical processes, the hydrologic cycle, biogeochemistry, and dynamic vegetation related to climate change (Hoffman et al. 2004). This model has developed rapidly across several versions including CLM2.0, CLM3.0, CLM3.5, and CLM4.0. Version CLM4.5 revises the photosynthesis scheme, improves hydrologic process modeling, includes the distribution of wetlands in cold regions, and includes new parameterization schemes for snow cover, lake modeling, crop modeling, and the modeling of various city types. In addition, a nitrogen fixation mechanism and methane emission model in the soil vertical direction have been introduced into CLM4.5. Since the release of this LSM, CLM4.5 has been widely applied in the fields of ecology (Tang et al. 2015; Duarte et al. 2017; Bilionis et al. 2014; Chen et al. 2018; Peng et al. 2018; Brunke et al. 2016; Wu and Dickinson 2004), climate change (Umair et al. 2018; Lawrence et al. 2012), assessment of the role of greenhouse gases (Zhang et al. 2016; Zhang and Wang 1997; Akkermans et al. 2014), and hydrology (Fu et al. 2016; Liu et al. 2017; Hack et al. 2006). As a result, CLM4.5 is considered one of the most well-developed and potentially useful LSMs in the world (Lai et al. 2014). Versions of the CLM model group have also been used in studies on the simulation and assessment of LST in mainland China (Meng et al. 2017a; Wang et al. 2015; Wang et al. 2015). Sun et al. (2017) drove CLM3.5 based on the China Land Data Assimilation System with atmospheric driving data, using LST from ground observation stations to assess the quality of the model. The results show that the bias and root-mean-square error (RMSE) of simulated LST vs. observed data varied seasonally. Further, the bias and RMSE of simulated LST vs. observed data were smaller in eastern China than in its west. Meng et al. (2017b) found that the CLM3.5 model had the greatest difference between simulated and observed LSTs in Xinjiang, with a maximum difference of ~5 K in July each year. Guo et al. (2017) used NCEP atmospheric forcing data to drive CLM4.5 for simulating changes in soil temperature on the Tibetan Plateau over the past century. The simulation results were validated by observational data of soil temperature from meteorological stations and field borehole monitoring stations. The results show that CLM4.5 could reasonably simulate observed changes in soil temperature on the plateau. Chen et al. (2010) used CLM3.0 and global atmospheric near-surface forcing data from Princeton University to conduct offline simulation experiments on soil temperature in China from 1948

to 2001, further assessing the ability of CLM3.0 to simulate soil temperature at different levels. The results show that the model could simulate the spatial distribution of multiyear average soil temperature in China. The simulated soil temperature was generally lower than observed data except for in some areas. The model could well-reflect the interannual variation of soil temperature in China. Moreover, the model could basically grasp the trend of temperature changes, but the simulated trend was weaker than the observed data. Xie et al. (2017) used observation data from Nagqu Station of the Plateau Climate and Environment of the Chinese Academy of Sciences to assess the performance of model simulation for surface energy exchange in the underlying surface of an alpine meadow on the Tibetan Plateau. The results showed that CLM4.5 could effectively simulate seasonal variations and diurnal cycles of surface longwave, reflected radiation, and net radiation as well as sensible and latent heat fluxes; CLM4.5 could also simulate surface soil heat fluxes during non-freezing periods in spring, summer, and autumn on the Tibetan Plateau. However, the simulation of LST during the winter freezing period gave values smaller than that of observed data.

The vast Chinese mainland features land surface characteristics that vary substantially by region. Due to the lack of long-series national site data, most studies involving evaluations only employed a few LSM stations with short observation periods. In addition, land–atmosphere interaction feedback has a major impact on the calculation of LSTs. Most assessments considered only the forcing effect of the atmosphere on the land surface, neglecting the transport of energy and mass from that surface. Therefore, it is difficult to fully evaluate the performance of CLM4.5 for a real climate simulation. All of these factors will limit the development of CLM4.5 for mainland China. The regional climate model RegCM was fully coupled with CLM4.5. Therefore, it is of great significance to promote the development of RegCM by carrying out the LST simulation assessment and improvment of CLM45 in mainland China. So we designed a long-term (*31* years) numerical simulation test of LST for mainland China based on long-term observational LST data by using the model RegCM-CLM4.5. Finally, the experiment of improving a soil thermal conductivity model was carried out. The results will promote the development and use of the coupled regional model in mainland China.

2 Materials and methods

2.1 Data

Land surface temperature is one of the main factors of surface meteorological observation in China. Observed LST data in mainland China were collected from the daily climate dataset of 809 Chinese ground international exchange

stations compiled by the China Meteorological Administration (Fig. 1). The dataset contains daily data of meteorological conditions from China reference and basic meteorological stations recorded since January 1951. To ensure data reliability, we used daily observed LST data from the 30-year study period (January 1, 1988 to December 31, 2017).

ERA5(Hersbach et al., 2020), the latest generation of reanalysis data set from ECMWF (European Centre for Medium-Range Weather Forecasts), was used to analyze the causes of simulation biases. Adopting the advanced four-dimensional assimilation system, the data could absorb high-altitude and near-surface observations as much as possible. Compared with the fourth-generation reanalysis data set ERA-Interim, ERA5 has been greatly improved with respect to spatial-temporal resolution, dynamic framework, data utilization, number of variables, and deviation processing method (Table 1). It is currently one of the most advanced reanalysis data in the world and has become a powerful tool for weather forecast and climate change studies (Qing et al. 2021; E M et al. 2021; Li et al. 2020; Xi et al. 2021). He (2021) evaluated the surface radiation data of ERA5, and the results show that it is very close to the observation of China, which provides a long series of substitute data for studying the change of surface energy balance.

2.2 Methods

(1) Soil thermal conductivity (STC)

The calculation model of soil temperature used in CLM4.5 is as follows:

$$c\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right) \tag{1}$$

where T is the soil temperature (K), z is downward in the vertical direction (m), c is the snow/soil heat capacity (J m^{-3} K^{-1}), t is time (s), and λ is the STC (W m^{-1} K^{-1}). The results show that λ strongly influenced the calculation of soil temperature. In CLM4.5, the model proposed by Johansen (1975) was used for the calculation of λ (JH for short, same as below). This is a semi-theoretical and semi-empirical model for calculating λ. Its expression is

$$\lambda = \left(\lambda_{\text{sat}} - \lambda_{\text{dry}}\right) \bullet K_e + \lambda_{\text{dry}} \tag{2}$$

where $\lambda_{\textbf{sat}}$ and $\lambda_{\textbf{dry}}$ are the thermal conductivity for saturated and dry soil, respectively, and K_e is the Kersten function, whose expression in this model is

Fig. 1 The simulation area and the four natural regions of mainland China. The point is the location of 809 observation stations
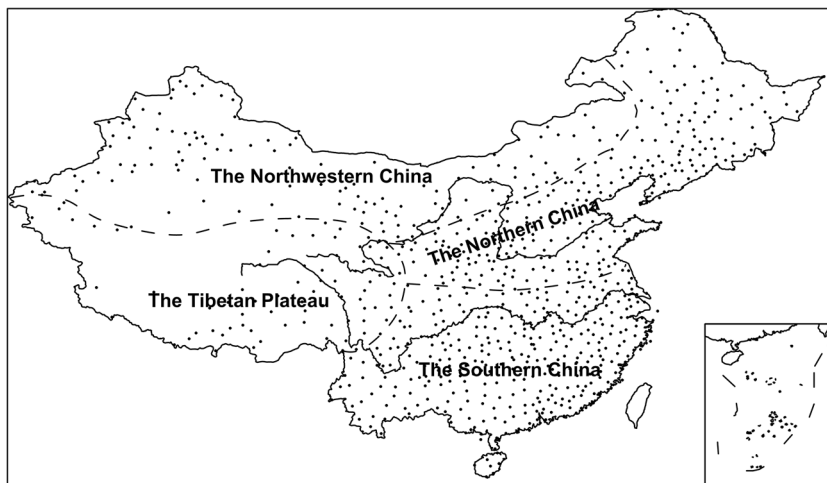
Table 1 Comparison of ERA-Interim and ERA5 data

Parameters	ERA-Interim	ERA5
Time span	1979–present	1950–present
Spatial resolution	79 km, 60 layers 0.1 hPa in top layer	31 km, 137 layers 0.01 hPa in top layer
Temporal resolution	6h	1h
Assimilation system	IFS cycle 31r2 4D-Var	IFS Cycle 41r2 4D-Var
Number of variables	About 100	More than 240
Uncertainty estimation	None	Assimilation of 10 members

$$K_e = \begin{cases} 0.7 \ \log \ S_r + \ 1.0, & 0.05 < S_r \le 0.1 \\ \log \ S_r + 1.0, & S_r > 0.1 \end{cases} \tag{3}$$

where s_r is the saturation of soil. Results have shown that K_e in the model is logarithmic, and the calculated λ was clearly smaller than the measured value (Su et al. 2016). To overcome this problem, Côté and Konrad (2005) proposed a new expression of K_e:

$$K_e = \frac{kS_r}{1 + (k-1)S_r} \tag{4}$$

where k is a parameter related to soil texture (CK for short, same as below). To make the Johansen model more suitable for low soil moisture content, Lu and Ren (2006) proposed a new scheme:

$$K_e = \exp\left(\alpha\left(1 - S_r^{\alpha - 1.33}\right)\right) \tag{5}$$

where α is a parameter related to soil texture (LR for short, same as below).

In the original CLM4.5, Ke parameterization of JH was used. In this study, two new Ke parameterizations of CK and LR were added in CLM4.5. Then, three sets of 30-year simulation were conducted, and the result was evaluated.

(2) LST change equation

To analyze the major factors influencing LST, the following LST change Eq. (6) was derived from the surface radiation balance equation (Chen and Paul, 2018):

$$\Delta T_s \cong \frac{1}{4\sigma T_s^3}\left[-SW_{in}\Delta\alpha_s + (1 - \alpha_s)\Delta SW_{in} + \Delta LW_{in} - \Delta LE - \Delta H - \Delta G\right] \tag{6}$$

In Eq. (6), T_s is the LST, SW_{in} is incoming shortwave radiation, LW_{in} is incoming longwave radiation, σ is the Stephan–Boltzmann constant (with a value of 5.67×10^{-8}W m^{-2} K^{-4}), H is the sensible heat flux, LE is the latent heat flux, and G is the ground heat flux. On the left side of the equation, ΔT_s is the LST change. $SW_{in}\Delta\alpha_s$ is the reflected solar radiation change; this equation indicates that when the surface albedo increases, the surface temperature decreases. $(1 - \alpha_s)\Delta SW_{in}$ represents the variation in incident shortwave radiation, and the equation shows that when the incident shortwave radiation increases, the surface temperature also increases. ΔLW_{in} represents the variation in incident longwave radiation related to the change of the atmospheric moisture and cloud, which is proportional to the change in LST. ΔLE is the change in latent heat, ΔH is the change in sensible heat, and ΔG is the change in surface heat flux. These three terms are inversely proportional to the change in LST. Surface emissivity is ignored in Eq. (6). In this

paper, the average data of annual, spring, summer, autumn, and winter of the model output were used to calculate the various factors of Eq. (6), and the principal factors causing the temperature changes were analyzed.

Equation (6) provides an effective way to diagnose the causes of simulation errors. However, physical variables such as albedo and surface radiation involved in Eq. (6) are all unconventional observation data, and there is no long series of site observation data in China at present. Therefore, the bilinear interpolation method was used in this paper to interpolate the LST from ERA5 reanalysis data to the stations (Fig. 1), and then evaluate its effectiveness according to real LST. Then, the data such as surface albedo and surface incoming shortwave and longwave radiations in ERA5 reanalysis data were used as the real field, and the corresponding output variables from numerical test were used to calculate the contribution of each item to the LST simulation biases in Eq. (6), so as to find out the main factors causing the LST simulation biases.

(3) Major assessment indicators

BIAS, RMSE, and Pearson correlation coefficient (r) were used as major assessment indicators. Their definitions are as follow.

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i) \tag{7}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}} \tag{8}$$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{9}$$

In Eqs. (7)–(9), i represents time. y_i is a simulated element (such as precipitation and temperature), and x_i is the corresponding observed element. The *BIAS* can be used to test whether simulated values from the model are large or small as well as the corresponding magnitude. The *RMSE* reflects the deviation of simulated from observed data. The smaller the value, the greater the simulation accuracy and the better the performance of the model. And the *Pearson correlation coefficient (r)* is a statistical quantity reflecting the linear correlation of two variables x and y. The larger the absolute value, the stronger the correlation. It was used to evaluate the temporal variation of the simulation.

2.3 Design of the numerical experiment

The simulation area covered 15.76–57.36°N and 66.25–141.13°E (Fig. 1). This was divided into a grid with 160 grid points in latitude and 145 in longitude. The horizontal grid size was 30 km, and the vertical direction was divided into 23 layers. The study used ERA-Interim reanalysis data (Dee et al., 2011) acquired from January 1987 to December 2017 for the lateral boundary. The data had a horizontal resolution of 0.75° × 0.75° (~80 km), 60 layers in the vertical, and a temporal resolution of 6 h. Sea surface temperature data were monthly average Optimum Interpolation Sea Surface Temperature (OISST) data of NOAA from the same period. Model parameters are listed in Table 2. The simulated LST was interpolated to the stations in Fig. 1 based on the bilinear interpolation method, and each evaluation indicator was calculated according to Eqs. (7)–(9).

Mainland China was divided into four regions as shown in Fig. 1 (Huang 1989). (1) The northern region is the northern part of China with a monsoon climate. (2) The southern region is south of the Qinling-Huaihe River and east of the Tibetan Plateau. It faces southeast to the East China and South China seas including the middle and lower reaches of the Yangtze River, the southern coast and southwest provinces (cities and autonomous regions) of China, and is the southern part of China with a monsoon climate. (3) The northwestern region is generally west of the Great Khingan Range and north of the Great Wall and the Kunlun-Altun Mountain Range. It embraces the non-monsoon climate portions of Inner Mongolia, Xinjiang, Ningxia, and northwestern Gansu. (4) The Tibetan Plateau forms the fourth region.

3 Results

3.1 Simulation of CLM4.5 for LST

3.1.1 Bias

The analysis of annual average LST in mainland China showed that LST decreased gradually from the southeastern coast to interior northwest. The annual average LST of the southeastern coast was > 20 °C and that between the Yangtze and Yellow rivers was about 15 °C. That of most other areas in North China was 5–10 °C. The LST of the Tibetan Plateau was the coolest, with most areas < 5 °C. The LST from northern Xinjiang to the southern Xinjiang Basin increased from 10 to 15 °C (Fig. 2a). The CLM4.5 showed favorable simulation performance in the spatial variation of annual average LST in China. The decrease in LST from the southeast coast to interior northwest was accurately simulated (Fig. 2b). However, the simulated values were clearly smaller than observed values, with most regions having a cold bias > 2 °C. Specifically, a cold bias of ~2–4 °C was observed east of 105°E, 6–8 °C west of 105°E, and 4–6 °C in other regions (Fig. 2c).

The simulations of average LST in the four seasons were very similar to the observed temperatures, with a decreasing trend from southeast to northwest. The CLM4.5 showed good simulation performance for this spatial distribution, but the bias varied greatly seasonally (data not shown). In spring, the bias was smallest in southern China. There, a cold bias of 2–4 °C was most common, with some areas having a cold bias of 0–2 °C. In spring, there was a cold bias of 6–8 °C in most of northern and northwestern China and on the Tibetan Plateau (Fig. 3a). In summer, except for some parts of southwestern and southern China where there was a cold bias of 4–6 °C; the bias in other regions was large (> 6 °C). The Tibetan Plateau had a cold bias of > 8 °C (Fig. 3b). In autumn, there was a cold bias of 2–4 °C east of 105°E, 6–8 °C west of that meridian, and 4–6 °C in North China and some of the plateau (Fig. 3c). In winter, the simulated bias of all regions in China decreased considerably. An exception was observed for the simulated bias (~6 °C) in southwestern China, in that in most regions it was < 4 °C (Fig. 3d).

Analyzing the climatic background, summer and autumn are the principal rainy seasons in China because they are strongly affected by the East Asian summer monsoon. Precipitation in most of mainland China was heavy during this season, which increased soil moisture. Winter and spring were controlled by a single westerly circulation system. Precipitation was weak ,and soil moisture decreased in most of mainland China. This indicates that the bias was closely related to soil moisture (Zhao et al.2021).

Table 2 Main simulation parameters	Dynamic structure	Test scheme
	Dynamic frame	Mesoscale model (MM5) non-static frame
	Large-scale precipitation scheme	SUBEX (Subgrid explicit water vapor scheme)
	Radiation transmission scheme	National Center for Atmospheric Research CCM3
	Sea surface flux scheme	Zeng
	Pressure gradient scheme	Hydrostatic recursion

Fig. 2 The observed and simu-
lated land surface temperatures
and the bias: (**a**) observations;
(**b**) simulation and bias. (**c**) bias
unit: °C

(a)

(b)

(c)
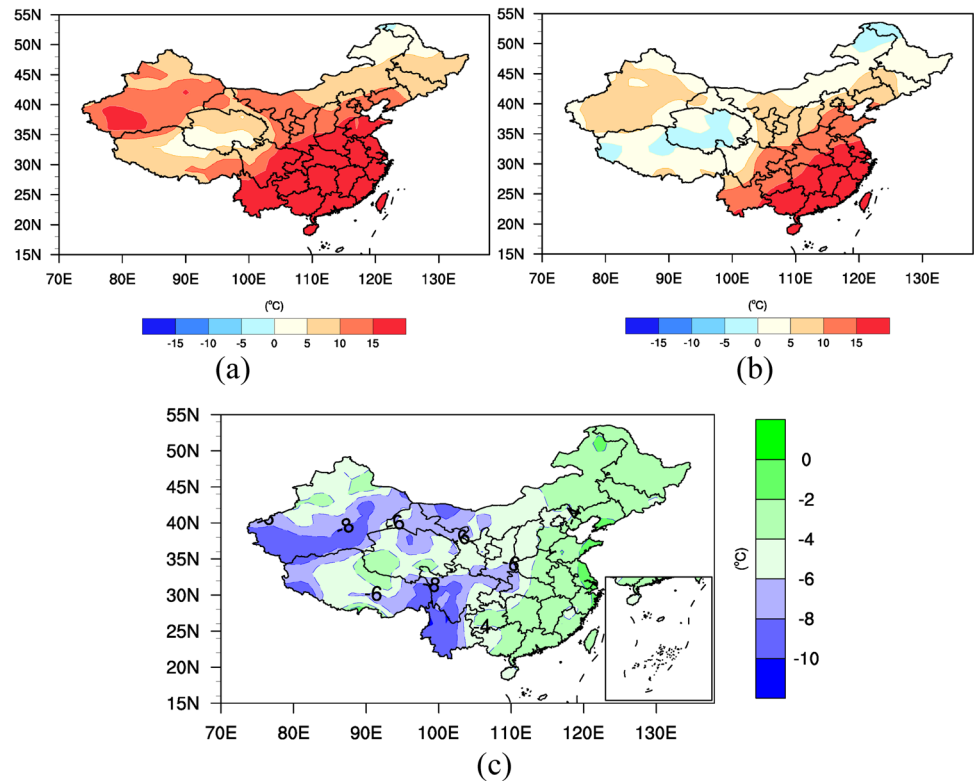
Fig. 3 LSTs bias for (**a**) spring,
(**b**) summer, (**c**) autumn, and (**d**)
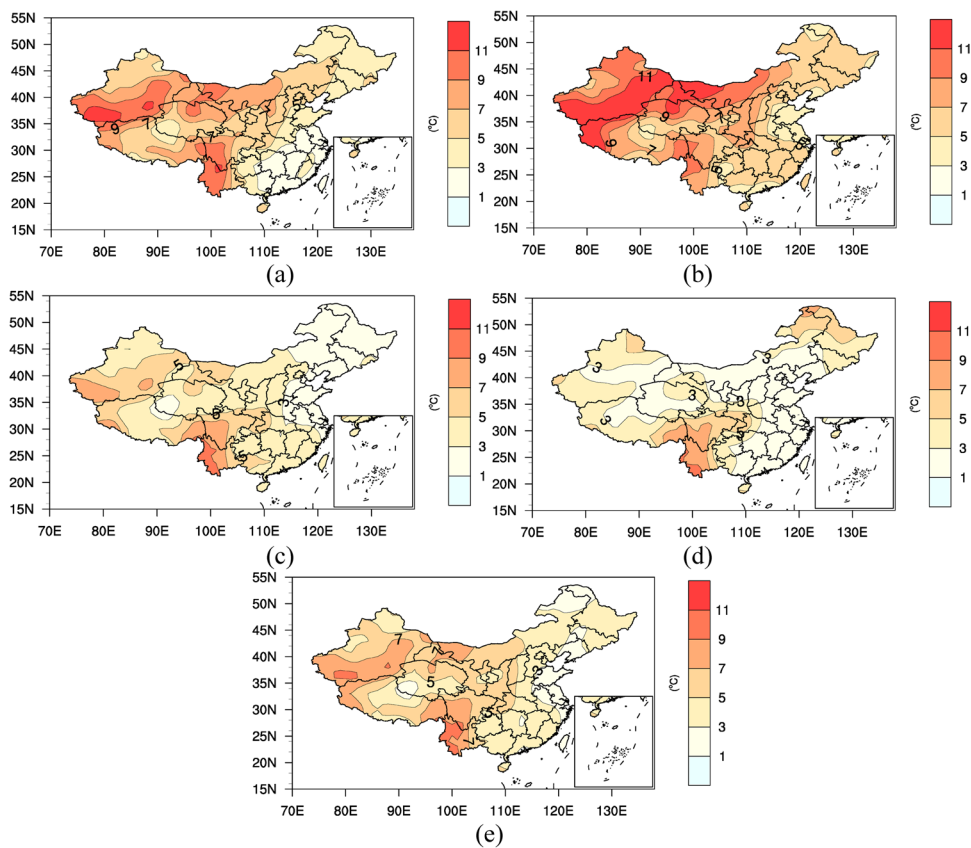winter; unit: °C

(a)

(b)

(c)

(d)
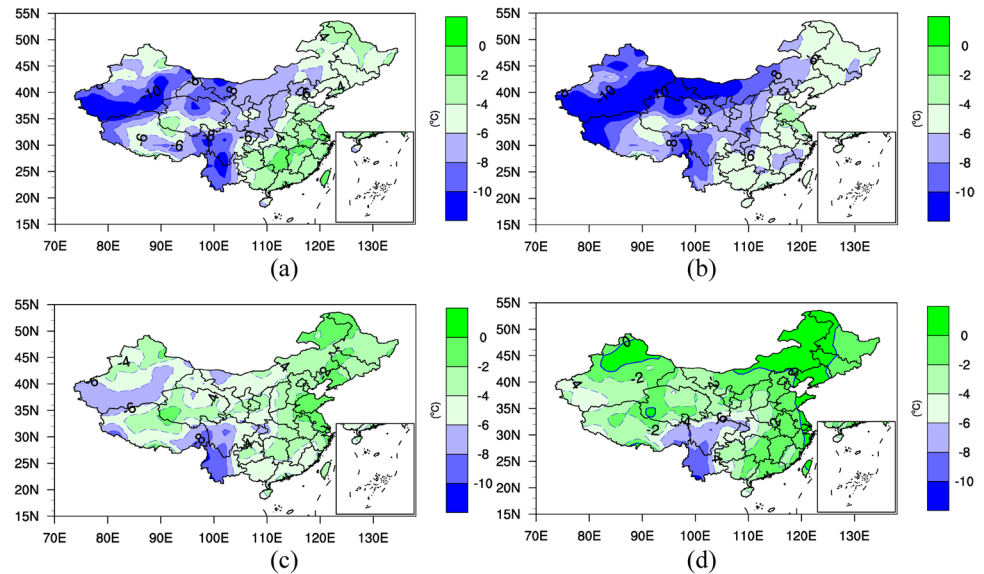
3.1.2 RMSE

The annual RMSE analysis showed that the RMSE in main-
land China increased gradually from southeast to northwest.
The RMSE of most of the southern regions was 1–3 °C,
with that of most of the northern regions at 5–7 °C. That
of the eastern part of the northwestern region was 7–9 °C,
and that of the southern Xinjiang Basin was 9–11 °C. That
of the Tibetan Plateau was 3–11 °C (Fig. 4a). In spring,

the RMSE in China was relatively large. There was again a
decrease from southeast to northwest. The RMSE in most
of the southern regions was < 5 °C, whereas that of other
regions was 9–11 °C (Fig. 4b). Summer and autumn had the
minimum RMSE among the four seasons. Except for the
plateau region where the RMSE was 9–11 °C, other regions
of China had an RMSE of 3–5 °C (Fig. 4c, d). The RMSE
distribution in winter was similar to that of the entire year
(Fig. 4e).

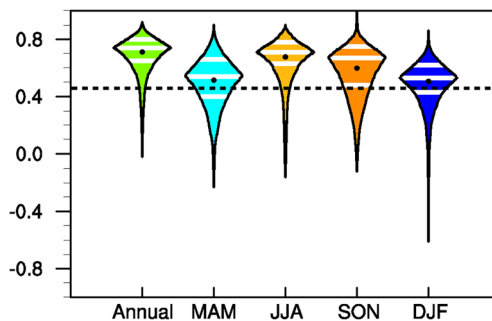**Fig. 4** Root mean square error in (**a**) all year, (**b**) spring, (**c**) summer, (**d**) autumn, and (**e**) winter. Unit: °C

4 Mechanism diagnosis of deviation

Compared with Fig. 2c, the annual mean LST bias of ERA5 in northern China is significantly reduced, and the value is generally less than 4 °C, and most areas are below 2 °C. The regional bias of eastern China and the plateau also decreased significantly (Fig. 6a). Compared with Fig. 3a–d, the LST bias of the four seasons in ERA5 is similar to the annual average, and the overall bias is small. The regional deviation in northern China is the most significant, and most of the regional deviations in the Qinghai-Tibet Plateau are

The sparsity of stations on the Qinghai-Tibet Plateau may lead to large false errors in the region, but it cannot change the fact that stations are sparse in the short term. Earlier evaluation results based on observations from over 2000 stations show that the surface temperature simulation error of the Qinghai-Tibet Plateau is larger than that of other regions in China. Therefore, the large deviation of surface temperature simulation on Qinghai-Tibet Plateau may be caused by the model itself, such as the scheme of soil thermal conductivity.

3.1.3 Correlation

Correlation coefficients between simulated and observed LST at all stations in China were between 0.75 and 0.9 (reaching significance level of $a=0.001$). The largest correlation coefficient was observed in summer, between 0.70 and 0.85 (reaching significance level of $a=0.001$). The next largest correlation coefficient was in autumn (0.5–0.75, reaching significance level of $a=0.001$). The results in spring were similar to those in autumn, and the maximum correlation coefficient was 0.80 (reaching significance level of $a=0.001$). The coefficient was between 0.35 and 0.65 in winter (reaching significance level of $a=0.05$) (Fig. 5).

Fig. 5 Correlation between observed and simulated LST. The dash line shows when the correlation became significant at $\alpha = 0.01$

also significantly reduced. Figure 7 showed the correlation between ERA5 and the real mean LST in each region. ERA5 was quite close to the real LST in northern China, northwestern China, southern China, and the Qinghai-Tibet Plateau, and the biases were within 3 °C (Fig. 7), which was much lower than the biases of CLM4.5 (which was about 5 °C). Most of the correlation coefficients in each region were above 0.84, which was much higher than the critical value (r=0.56) at 0.001 significance level.

Based on the above evaluation, this paper study the causes of CLM4.5 simulation biases by calculating the percentage of biases of surface incoming short-wave and long-wave radiations, sensible heat and latent heat between ERA5 and simulated value from CLM (Fig. 8). The results showed that compared with ERA5, the annual mean surface albedo simulated by CLM was about 20% higher, and the value was 30% higher in autumn and winter. Surface incoming shortwave radiation was 20% higher, and it was close to 80% in winter. However, the surface incoming long-wave radiation was about 20% lower, and the largest bias was showed in autumn. The sensible heat was 20% lower in the north of mainland China and 40% higher in the north of mainland China. The value was higher over the whole region, and the largest bias was approximate 80% in winter. Latent heat showed overestimation, which was 40% higher in the north of Mainland China, and was overestimated by 2 times in winter.

Biases of LST compared with ERA5 caused by various factors in Eq. (6) simulated by CLM were calculated (Fig. 9). The results show that the surface incoming short-wave radiation simulated by CLM4.5 was overestimated, which made the average LST of the regions 5 °C higher than that of ERA5. This was more significant in the southern region, which was about 7 °C higher on average, and showed the greatest effect in all regions in spring and winter. Surface albedo, long-wave radiation, sensible, and latent heat all contributed negatively to the increase of LST. Due to the small incoming long-wave radiation term, the LST in each region was −5 °C lower than ERA5, and the largest underestimation was in the southern region. The difference of LST caused by sensible heat and latent heat was close to each other, which was within 5 °C. The above analysis showed that the incoming long-wave radiation, sensible heat, and latent heat simulated by CLM4.5 were smaller than ERA5, which was the reason for the low simulation of LST.

5 Improvement of LST simulation

The above research revealed the factors and their contribution values of the cold deviation caused by the atmospheric model coupled with CLM4.5, which provides a meaningful reference for the improvement of the regional land-atmosphere coupled climate model. On the other hand, the

Fig. 6 ERA5 LST bias for (**a**) annual, (**b**) spring, (**c**) summer, (**d**) autumn, and (**e**) winter; unit: °C
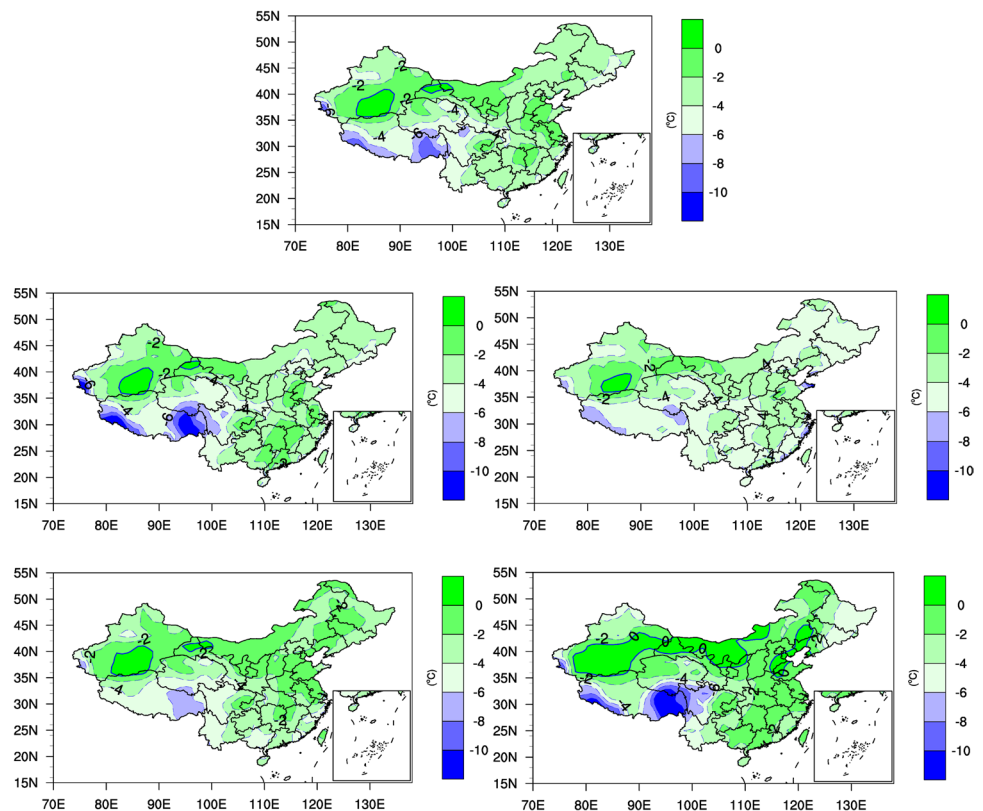
Fig. 7 Comparison between ERA5 and real average surface temperature in each region, **a**, north, **b**, northwest, **c**, South, **d**, Qinghai Tibet Plateau
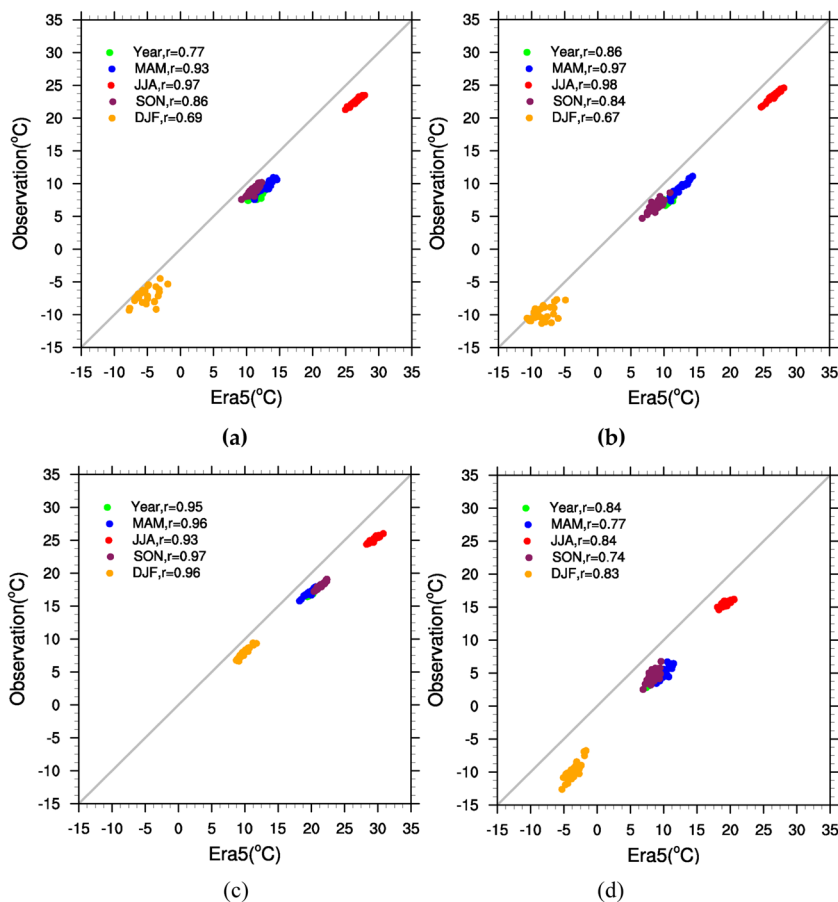

physical model for calculating the surface temperature was the root cause of the deviation. Perhaps the model of the thermal conductivity in CLM4.5 is the main reason for this bias. To study the effects of STC (λ) on the simulation of LST, two long-term (31-year) simulation tests were conducted for two schemes. The results shows that the LST in most of northern China simulated by the Côté-Konrad scheme was higher by 0.5–1.0 °C when compared with that of the Johansen scheme, whereas that temperature in most other regions was reduced by 0–0.5 °C (Fig. 10a, b). In the Lu-Ren scheme, the LST increased by 0.5–1.0 °C over that of Johansen scheme in most regions except for on the Tibetan Plateau and in the regions of Guangdong and Guangxi, where that temperature was reduced by 0–0.5 °C. The LST in most of northern China increased by 1–1.5 °C, with the increase in some areas reaching 3 °C (Fig. 10c, d). According to the bias of the three schemes, little difference was observed between the Johansen and the Côté-Konrad schemes. Meanwhile, the Lu-Ren scheme can significantly reduce the cold bias in most regions. Therefore, it is more suitable for the simulation of LSTs on the Chinese mainland.

In order to further evaluate the improvement effect of LR scheme on model performance, four LST series JH, CK, LR, and ERA5 were compared with the observational LST in each region. The results (Fig. 11) showed that ERA5 and LR had the smallest bias. The biases were −2.5 °C (ERA5) and −3.0 °C (LR) in the northern region. JH showed the largest bias, with a value of −3.5 °C. The bias of each product was larger in summer and autumn, but the bias of ERA5 and LR was still the smallest. The bias characteristics were similar between the northwestern region and the southern region; the bias of ERA5 and LR was the smallest among the four series. LST change showed (Fig. 12) that LST in each region increased significantly in the past 30 years. The four series can reflect this rising trend, but LR was the closest to the observational, with correlation coefficient $r=0.89$, reaching significance level of $a=0.001$. Compared with other regions, the simulation errors of annual average LST in the Qinghai-Tibet Plateau were larger by all series, and the LR showed the smallest bias and the most significant correlation (correlation coefficient $r=0.72$, reaching significance level of $a=0.001$). The Tibetan Plateau is a special region compared with other regions. The glacier has a large area, and much areas are covered with snow in winter and spring. CLM4.5 calculated snow temperature using the formula convection model, but the calculation scheme of snow

Fig. 8 Deviation rate of surface energy balance factors of CLM relative to ERA5, **A**. albedo, **B**. surface net incident short wave radiation, **C**. surface net incident long wave radiation, **D**. sensible heat, **E**. latent heat
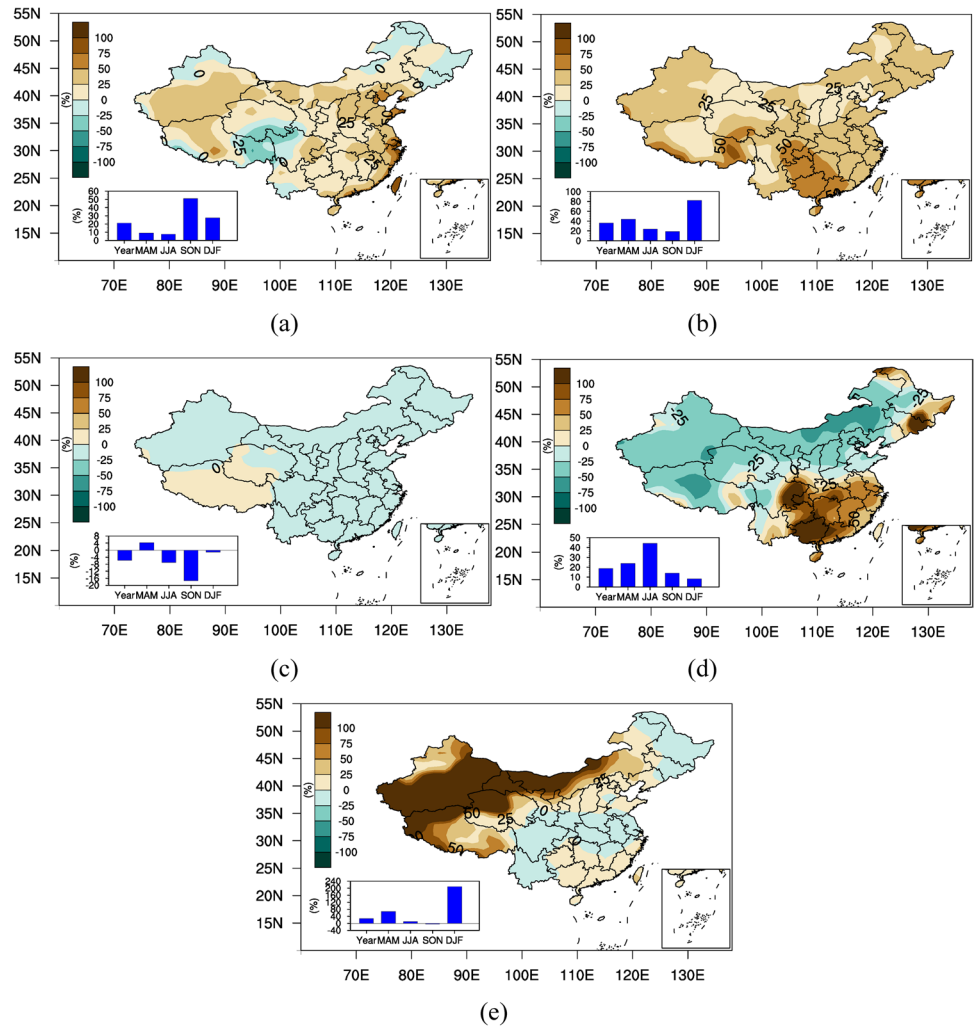

Fig. 9 LST changes caused by biases from albedo (a), surface shortwave radiation (sr), surface longwave radiation (lr), sensible heat (sh), and latent heat (le) in different regions of mainland China (**a** Northern China, **b** northwestern China, **c** Southern China, **d** Qinghai-Tibet Plateau)
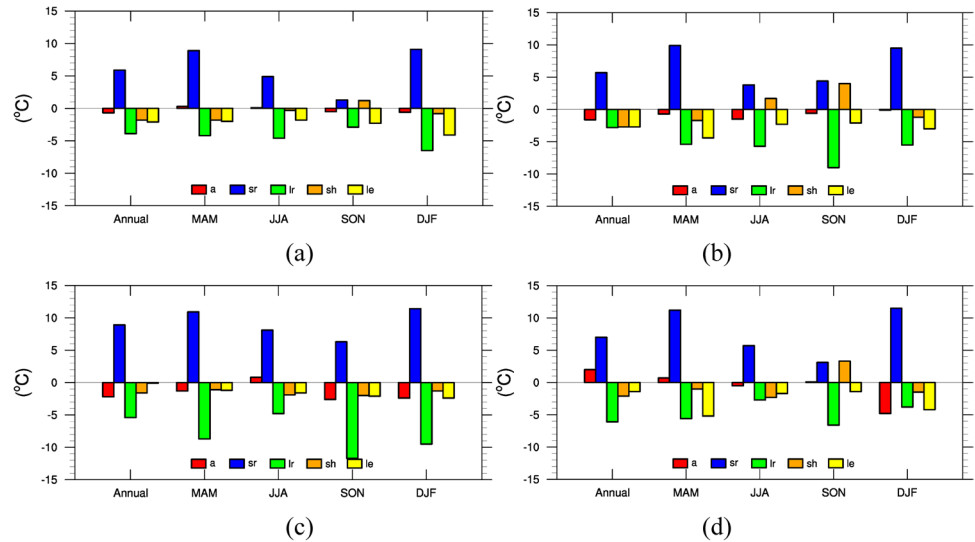
Fig. 10 Biases between observations and parameterized land surface temperatures. (**a**) Côté-Konrad scheme versus observations; (**b**) Côté-Konrad scheme versus Johansen scheme; (**c**) the Lu-Ren scheme versus observations, and (**d**) Lu-Ren scheme versus Johansen scheme (Unit:°C)
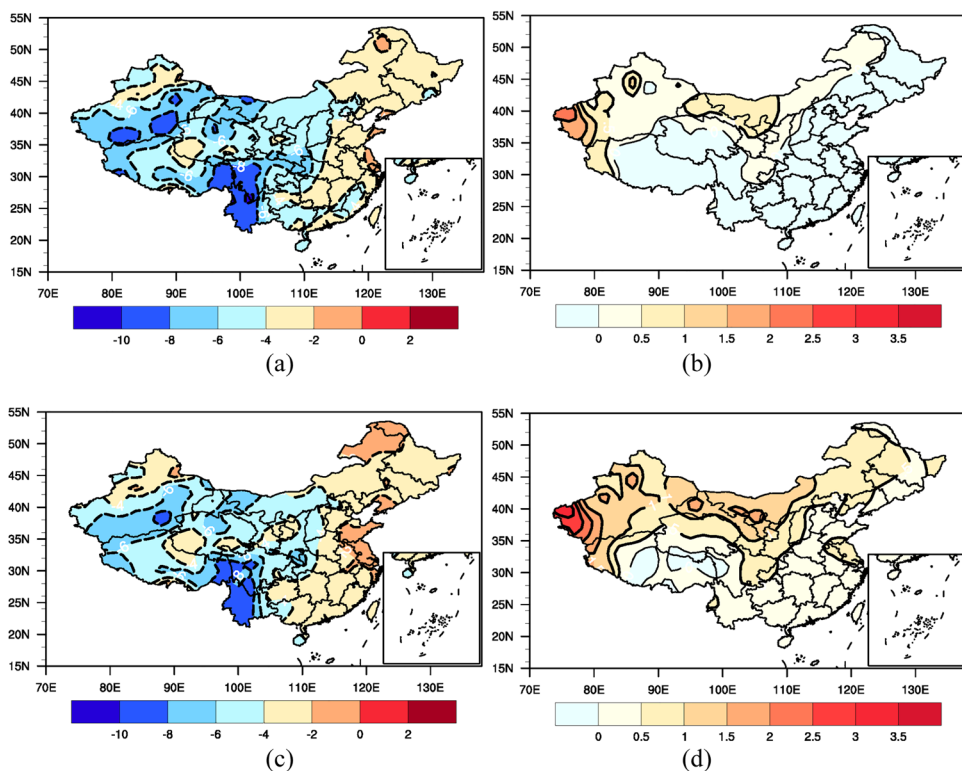
Fig. 11 Bias in (**a**) all areas, (**b**) northwest China, (**c**) north China, (**d**) south China, and (**e**) on the Tibetan Plateau; see Fig. 2 for a map of these regions. Unit: °C
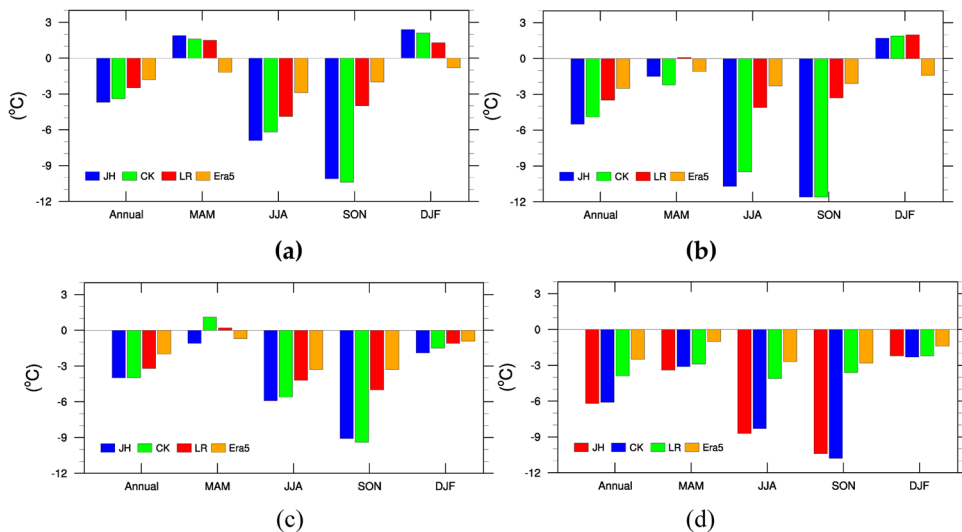

thermal conductivity was different from John's scheme, which should be in the soil, so Jodan (1991) model was adopted. Xie (2017) evaluated the simulation performance of CLM4.5 on the underlying surface of alpine meadow on the Tibetan Plateau (hereinafter referred to as the Plateau) by using a whole year's observation data from Nagqu Alpine Climate and Environment Observation and Research Station of the Chinese Academy of Sciences. The results showed that CLM4.5 can well simulate the seasonal variation and diurnal cycle characteristics of surface long

wave, reflected radiation, net radiation, sensible and latent heat flux, and surface soil heat flux in spring, summer, and non-freezing period in autumn. However, the simulation of surface temperature and sensible heat flux is lower due to the high surface albedo caused by snow cover. Therefore, parameterization scheme of plateau snow cover and albedo parameterization schemes related to snow cover need further improvement.

The Theil-Sen (Lavagnini et al. 2011) method was used to analyze the trend of annual average LST in China over

Fig. 12 Bias in (**a**) all areas,
(**b**) northwest China, (**c**) north
China, (**d**) south China, and
(**e**) on the Tibetan Plateau;
see Fig. 2 for a map of these
regions. Unit: °C

(a)

(b)

(c)

(d)

(e)
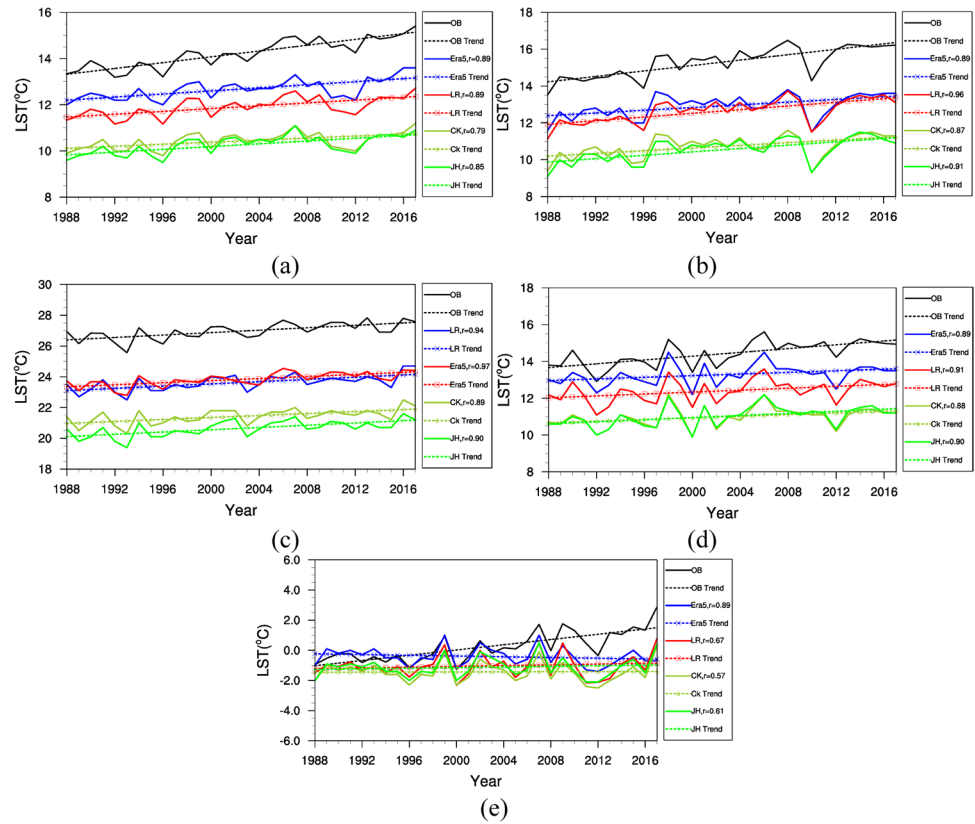
Table 3 Trend in five LSTs of
mainland China in the last 30
years (unit: °C/decade)

Period	Product	All areas	Northwestern	Northern	Southern	Plateau
Annual	OB	0.58	0.71	0.33	0.50	0.77
	LR	0.38	0.40	0.36	0.19	0.01
	ERA5	0.31	0.53	0.30	0.25	0.06
	JH	0.30	0.50	0.37	0.27	0.00
	CK	0.23	0.40	0.36	0.19	-0.01
MAM	OB	0.77	0.8	0.50	0.50	1.33
	LR	0.18	0.29	0.38	0.20	-0.02
	ERA5	0.31	0.54	0.42	0.13	0.07
	JH	0.38	0.52	0.43	0.30	0.10
	CK	0.24	0.38	0.38	0.20	-0.02
JJA	OB	0.94	1.00	0.74	0.71	1.33
	LR	0.24	0.53	0.45	0.17	-0.16
	ERA5	0.38	0.74	0.60	0.19	-0.10
	JH	0.22	0.47	0.50	0.17	0.15
	CK	0.27	0.53	0.45	0.18	-0.11
SON	OB	0.37	0.62	0.16	0.42	0.29
	LR	0.17	0.37	0.25	0.25	-0.02
	ERA5	0.25	0.44	0.20	0.33	0.14
	JH	0.30	0.50	0.28	0.35	0.06
	CK	0.22	0.33	0.27	0.25	-0.01
DJF	OB	0.67	0.68	0.45	0.67	0.88
	LR	0.28	0.34	0.38	0.29	0.24
	ERA5	0.47	0.54	0.23	0.45	0.62
	JH	0.26	0.23	0.43	0.29	0.21
	CK	0.32	0.33	0.38	0.31	0.30

the period of 1988–2017 (Table 3). The results show that this temperature increased at a rate of 0.58 °C/decade over that period. The Tibetan Plateau and northwestern regions of China had the maximum increases (0.77 and 0.71 °C/decade respectively), while the northern region had less change (an increase of 0.33 °C/decade). The LST simulated by CLM4.5 also increased, but the simulated increase was less than that of the observational data. In spring, the rate of increase of LST in China was 0.77 °C/decade, with that in the Tibetan Plateau region having the maximum increase (1.33 °C/decade), and other regions showing small increases. The CLM successfully simulated the increasing trend of LST in all regions during spring, but the simulated

increase was smaller than that of the observational data. Notably, the actual rapid increase temperature in the Plateau region was not reproduced by the simulation. In summer, the LST in all of mainland China increased the fastest at 0.94 °C/decade. However, the value simulated by CLM was 0.72 °C/decade smaller than observed. The temperature on the Plateau also increased sharply, at 1.33 °C/decade, with the simulated value 1.48 °C/decade smaller than what was observed. Similar to spring and summer, the trend of LST increase in autumn and winter was accurately simulated, but the simulated increase was smaller than what was observed. A comparison of the simulated trend of the four LST series (Table 3) shows that their variations were all smaller than

Fig. 13 Taylor diagrams of four LST series for all areas of China, northwest China, north China, south China, and on the Tibetan Plateau. (**a**) all year, (**b**) spring, (**c**) summer, (**d**) autumn, and (**e**) winter
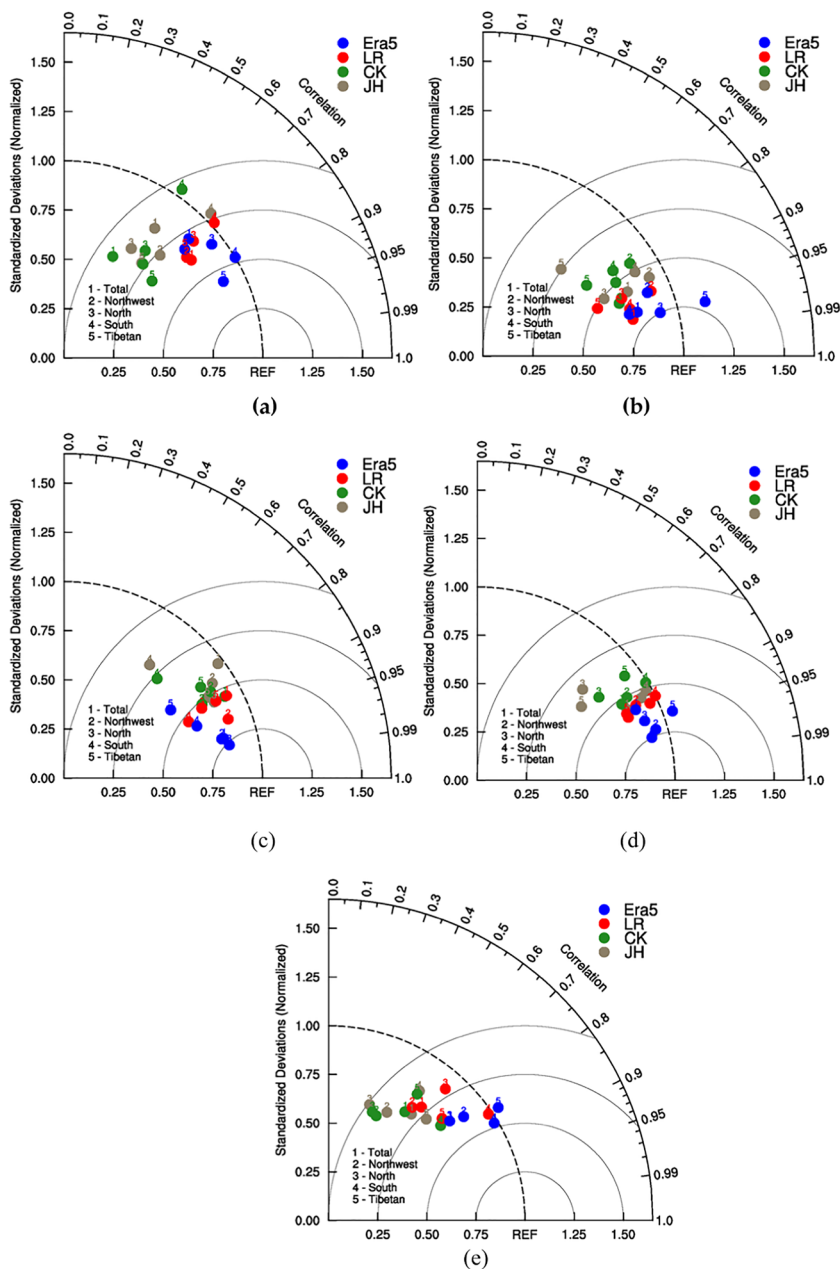
that based on observed data. The simulated LST trend of ERA5 and LR was much closer to the actual trend.

A Taylor (2001) diagram provides a visual framework for comparing a set of variables from one or more test datasets to one or more reference datasets. In the present work, the diagram was used to comprehensively assess the performance of the three types of simulations of LST in mainland China (Fig. 13). The results show that ERA5 and LR were better than the other two types of simulations (blue solid circles closer to the REF).

From all we can see that the bias in summer and autumn was greater than that in other seasons, because the heavy rainfall occurred in these two seasons, the soil moisture was higher than that in other periods. Meanwhile, the change in soil moisture had an important effect on the simulation of LST, especially in arid and semiarid areas of northern China, where evaporation at the soil surface is intense. However, the isothermal model used in CLM4.5 did not consider the effects of a change in soil moisture on soil temperature, perhaps resulting in a large simulated bias in northern and northwestern regions of China that experience low soil moisture and large variations in soil moisture. Therefore, developing a new λ calculation scheme and considering the role of water vapor in the calculation of soil temperature can provide effective means for improving the performance of the model in simulating the LSTs of mainland China.

6 Conclusions

We ran the CLM4.5 in a land–atmosphere coupling approach. The longest and latest observed LST dataset for mainland China was used for the first time to comprehensively assess LSTs in China as simulated by CLM4.5. The results show that CLM4.5 produced systematic cold deviations in the simulation of LSTs in mainland China. The RMSE in mainland China increased gradually from southeast to northwest, with the smallest value in the south (1–3 °C) and largest in the southern Xinjiang Basin (9–11 °C). Summer and autumn had the smallest RMSEs for each region in a year. Correlation coefficients between simulated and observational data for all stations in China were between 0.75 and 0.9 ($P < 0.001$). The strongest correlation was observed in summer, with the correlation coefficient from 0.70 to 0.85 ($P < 0.001$). In winter, that coefficient was the smallest, 0.35 to 0.65 ($P < 0.05$). As a result, the observed annual and seasonal average LSTs in mainland China had strong linear relationships with those simulated by CLM4.5. In the past 30 years, the LST of mainland China increased at a rate of 0.058 °C/decade, with that of the Tibetan Plateau and northwestern regions increasing fastest (0.077 and 0.071 °C/decade, respectively) and that in the northern region changing the least (0.033°C/decade). The LST simulated

by CLM4.5 also increased, but that increase was smaller than that of observational data. Diagnose by using ERA5 reanalysis data and LST variation equation showed that the negative bias caused by the incoming long-wave radiation, sensible heat, and latent heat simulated were the main factors causing the low LST simulation.

The STC sensitivity numerical tests show that STC had a major influence on the reduction of simulated LST. However, the simulated cold bias from the Lu-Ren scheme remained large, and an increasing trend of the bias was observed for the Tibetan Plateau. In addition, the bias in summer and autumn was greater than that in other seasons, which shows that a change in soil moisture had an important effect on the simulation of LST, especially in arid and semiarid areas of northern China where evaporation at the soil surface is intense. However, the isothermal model used in CLM4.5 did not consider the effect of a change in soil moisture on soil temperature, resulting in a large simulated bias in northern and northwestern regions with low soil moisture and a large variation in soil moisture. Therefore, developing a new λ calculation scheme and considering the role of water vapor in the calculation model of soil temperature are effective means for improving the performance of the model in simulating LST in mainland China.

Declarations

Ethics approval Not applicable

Consent to participate Not applicable

Consent for publication Not applicable

Conflict of interest The authors declare no competing interests.

References

Akkermans T, Thiery W, van Lipzig NPM (2014) The regional climate impact of a realistic future deforestation scenario in the Congo basin. J Climate 27(7):2714–2734. https://doi.org/10.1175/JCLI-D-13-00361.1

Bilionis I, Drewniak BA, Constantinescu EM (2014) Crop physiology calibration in clm. Geosci Model Dev 7(5):1071–1083

Brunke MA, Broxton P, Pelletier J (2016) Implementing and evaluating variable soil thickness in the community land model version 4.5 (clm4.5). J Climate 29(9):3441–3461

Chen HS, Xiong MM, Sha WY (2010) Simulation of land surface processes over China and its validation Part I:Soil temperature. Scientia Meteorol Scinica 30(5):621–630

Chen M, Griffis TJ, Baker JM (2018) Comparing crop growth and carbon budgets simulated across ameriflux agricultural sites using the community land model (clm). Agric For Meteorol 256-257:315–333

Coll C, Garciasantos V, Niclos R (2016) Test of the MODIS land surface temperature and emissivity separation algorithm with ground measurements over a rice paddy. IEEE Trans Geosci Remote Sens 54(5):3061–3069

Côté J, Konrad JM (2005) A generalized thermal conductivity model for soils and construction materials. Can Geotech J 42(3):443–458

Dee DP, Uppala SM, Simmons AJ (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quart J Roy Meteor Soc 137(656):553–597

Dickinson RE, Errico RM, Giorgi F (1989) A regional climate model for the western United States. Clim Change 15:383–422

Duan SB, Li ZL, Leng P (2017) A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data. Remote Sens Environ 195:107–117

Duarte HF, Raczka BM, Ricciuto DM (2017) Evaluating the community land model (clm4.5) at a coniferous forest site in northwestern united states using flux and carbon-isotope measurements. Biogeosciences 14(18):4315–4340

Fu C, Wang G, Goulden ML (2016) Combined measurement and modeling of the hydrological ipact of hydraulic, redistribution using clm4.5 at eight ameriflux sites. Hydrol Earth Syst Sci 20(5):2001–2018

Giorgi F, Marinucci MR, Bates GT (1993) Development of a 2nd generation regional climate model(Regcm2): convective processes and assimilation of lateral boundary-conditions. Mon Weather Rev 121(10):2814–2832

Guo DL, Li D, Liu GY (2017) Simulated change in soil temperature on the Tibetan Plateau from 1901 to 2010. Quaternary Sci 37(5):1102–1110

Hack JJ, Caron JM, Yeager SG (2006) Simulation of the global hydrological cycle in the ccsm community atmosphere model version 3 (cam3): mean features. J Climate 19(11):2199–2221

He Y, Wang K, Feng F (2021) Improvement of ERA5 over ERA-Interim in simulating surface incident solar radiation throughout China. J Climate 34(10):3853–3867

Hersbach H, Bell B, Berrisford P (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049

Hoffman F, Vertenstein M, Thornton P (2004) Community land model version 3.0 (clm3.0) developer's guide. Department of Energy, United States

Huang BW (1989) Outline of China's Comprehensive Natural Zoning. Geography 21:10–20

Johansne O (1975) Thermal properties of soils. Ph.D. thesis,. University of Trondheim

Jones DA, Trewin BC (2015) On the relationships between the El Niño-Southern Oscillation and Australian land surface temperature. Int J Climatol 20(7):697–719

Lai X, Wen J, Ceng SX (2014) Numerical simulation and evaluation study of soil moisture over China by using clm4.0 model [in Chinese]. J Atmos Sci 38(3):499–512

Lavagnini I, Badocco D, Pastore P (2011) Theil–sen nonparametric regression technique on univariate calibration, inverse regression and detection limits. Talanta 87:188

Lawrence DM, Slater AG, Swenson SC (2012) Simulation of present-day and future permafrost and seasonally frozen ground conditions in ccsm4. J Climate 25(7):2207–2225

Li F, Chavas DR, Reed KA (2020) Climatology of severe local storm environments and synoptic-scale features over North America in ERA5 reanalysis and CAM6 simulation. J Climate 33(19):8339–8365

Liu D, Mishra AK (2017) Performance of amsr_e soil moisture data assimilation in clm4.5 model for monitoring hydrologic fluxes at global scale. J Hydrol 547:67–79

Lu S, Ren T, Gong YS (2006) An improved model for predicting soil thermal conductivity from water content at room temperature. Soil Sci Soc Am J 71(1):8–14

Meng X, Wang H, Wu Y (2017a) Investigating spatiotemporal changes of the land-surface processes in Xinjiang using high-resolution clm3.5 and cldas, soil temperature. Sci Rep 7(1):13286

Meng XY, Wang H, Liu ZH (2017b) Simulation and verification of land surface soil temperatures in the Xinjiang Region by the CLM3.5 model forced by CLDAS. Acta Ecologica Sinica 37(3):979–995

Peng B, Guan K, Chen M (2018) Improving maize growth processes in the community land model, implementation and evaluation. Agric For Meteorol 250-251:64–89

Qin S, Wang K, Wu G, Ma Z (2021) Variability of hourly precipitation during the warm season over eastern China using gauge observations and ERA5. Atmospheric Res 264:105872

Su LJ, Wang QJ, Wang S (2016) Soil thermal conductivity model based on soil physical basic parameters. Transactions of the CSAE 32(2):127–133

Sun S, Shi CS, Liang X (2017) Assessment of ground temperature simulation in China by different land surface models based on station observations. J Appl Meteor 28(6):737–749

Tang J, Riley WJ, Niu J (2015) Incorporating root hydraulic redistribution in CLM4.5, Effects on predicted site and global evapotranspiration, soil moisture, and water storage. J Adv Model Earth Syst 7(4):1828–1848

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res 106(D7):7183

Umair M, Kim D, Ray RL (2018) Estimating land surface variables and sensitivity analysis for clm and vic simulations using remote sensing series. Sci Total Environ 633:470–483

Wan Z, Li ZL (1997) A physics based algorithm for retrieving land‐surface emissivity and temperature from EOS/MODIS data. IEEE Trans Geosci. Remote Sens 35:980–996

Wang XF, Yang MX, Guo JP (2015) Simulation and improvement of land surface processes in Nameqie, Central Tibetan Plateau, using the Community Land Model (CLM3.5). Environ Earth Sci 73(11):7343–7357

Wilson JS, Clay M, Martin E (2003) Evaluating environmental influences of zoning in urban ecosystems with remote sensing. Remote Sens Environ 86(3):303–321

Wu W, Dickinson RE (2004) Time scales of layered soil moisture memory in the context of land-atmosphere interaction. J Climate 17(14):2752–2764

Xi D, Chu K, Tan ZM, Gu JF (2021) Characteristics of warm cores of tropical cyclones in a 25-km-mesh regional climate simulation over CORDEX East Asia domain. Clim Dyn 57:2375–2389

Xie ZP, Hu ZY, Liu HL (2017) Evaluation of the surface energy exchange simulations of land surface model CLM4.5 in alpine meadow over the Qinghai-Xizang Plateau. Plateau Meteor 36(1):1–12

Zeng X, Shaikh M, Dai Y (2002) Coupling of the common land model to the ncar community climate model. J Climate 15(14):1832–1854

Zhang Y, Wang WC (1997) Model simulated northern winter cyclone and anticyclone activity under a greenhouse warming scenario. J Climate 10(7):1616–1634

Zhang L, Mao J, Shi X (2016) Evaluation of the community land model simulated carbon and water fluxes against observations over China flux sites. Agric For Meteorol 226-227:174–185

Zhao JZ, Wang AH, Wang HJ (2021) Soil moisture memory and its relationship with precipitation characteristics in China region. Chinese J Atmospheric Sci (in Chinese) 45(4):799–818

Zhong L, Su Z, Ma Y (2011) Accelerated changes of environmental conditions on the Tibetan plateau caused by climate change. J Climate 24(24):6540–6550

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