

Article

Improving the Spring Air Temperature Forecast Skills of BCC_CSM1.1 (m) by Spatial Disaggregation and Bias Correction: Importance of Trend Correction

Chunfeng Duan ^{1,2,*}, Pengling Wang ^{3,*}, Wen Cao ² , Xujia Wang ¹ , Rong Wu ¹ and Zhi Cheng ¹

¹ Anhui Climate Center, Hefei 230031, China; wangxj_lzu@foxmail.com (X.W.); wr_ahqx@163.com (R.W.); alexclimate@163.com (Z.C.)

² Atmospheric Science and Satellite Remote Sensing Key Laboratory of Anhui Province, Anhui Meteorological Institute, Hefei 230031, China; sgfxxy_0@163.com

³ National Climate Center, Beijing 100081, China

* Correspondence: dcf118@126.com (C.D.); wangpl@cma.gov.cn (P.W.)

Abstract: In this study, an improved method named spatial disaggregation and detrended bias correction (SDDBC) based on spatial disaggregation and bias correction (SDBC) combined with trend correction was proposed. Using data from meteorological stations over China from 1991 to 2020 and the seasonal hindcast data from the Beijing Climate Center Climate System Model (BCC_CSM1.1 (m)), the performances of the model, SDBC, and SDDBC in spring temperature forecasts were evaluated. The results showed that the observed spring temperature exhibits a significant increasing trend in most of China, but the warming trend simulated by the model was obviously smaller. SDBC performed poorly in temperature trend correction. With SDDBC, the model's deviation in temperature trend was corrected, and consequently, the temporal correlation between the model's simulation and the observation as well as the forecasting skill on the phase of temperature were improved, thus improving the MSSS and the ACC. From the perspective of probabilistic prediction, the relative operating characteristic skill score (ROCSS) and the Brier skill score (BSS) of the SDDBC for three categorical forecasts were higher than those of the model and SDBC. The SDDBC's BSS increased as the effect of the increasing resolution component was greater than that of the decreasing reliability component. Therefore, it is necessary to correct the predicted temperature trend in post-processing for the output of numerical prediction models.

Keywords: BCC_CSM1.1 (m); air temperature; spatial disaggregation and bias correction; seasonal forecasting; China



Citation: Duan, C.; Wang, P.; Cao, W.; Wang, X.; Wu, R.; Cheng, Z. Improving the Spring Air Temperature Forecast Skills of BCC_CSM1.1 (m) by Spatial Disaggregation and Bias Correction: Importance of Trend Correction. *Atmosphere* **2021**, *12*, 1143. <https://doi.org/10.3390/atmos12091143>

Academic Editor: Sean F. Milton

Received: 16 August 2021

Accepted: 2 September 2021

Published: 5 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Climate system models (CSMs) have become the main tool for climate prediction around the world [1–3]. Recently, the Beijing Climate Center (BCC) of the China Meteorological Administration (CMA) has improved the physics and resolution of its operational CSM and updated the forecast system to a second-generation climate system model (BCC_CSM) [2]. Recent studies have used the archived BCC_CSM reforecasts for different applications, such as evaluating the forecast skill of Asian–Western Pacific summer monsoon [4], Asian summer monsoon [5], Madden–Julian oscillation [6], summer precipitation [7], synoptic eddy and low-frequency flow [8], Indian Ocean basin mode and dipole mode [9], stratospheric sudden warming [10], primary East Asian summer circulation patterns [11], and winter temperature [12]. The model has shown a considerable ability to predict important climate phenomena, tropical large-scale atmospheric circulation anomalies and primary climate variability modes. However, the prediction skill of weak anomaly signals and atmospheric circulation in middle and high latitudes still needs to be improved.

Seasonal predictions are often needed at the local scale for the assessment and correction of forecasts to meet local forecasting needs. The downscaling method is an effective way to transform model prediction to the local scale and improve the forecast accuracy, which has two categories: dynamic downscaling and statistical downscaling [13–18]. Statistical downscaling uses the statistical relationship between the output of a CSM and local observations to obtain local-scale or station forecasts and is simple to apply, with high computational efficiency [13–15]. It can also be employed to correct systematic bias between forecasts and observations. The limitation of statistical downscaling methods is that they require long-term continuous forecast archives and observations to establish stable and reliable statistical relationships [13–15].

Many studies on statistical downscaling have been carried out, and new methods are emerging. The bias correction and spatial disaggregation (BCSD) method is an interpolation-based downscaling method first proposed by Wood et al. in 2002 [19,20]. This method has been widely used in hydrologic and climate prediction studies [21–30]. The BCSD method is composed of bias correction by quantile mapping and spatial disaggregation, which effectively corrects both the mean and variance of forecasts based on observations. By reversing the order of the BCSD process, the spatial disaggregation with bias correction (SDBC) method was developed [31]. This modification improved the downscaling skill, and the SDBC method was widely used for developing local-scale statistics of precipitation, surface temperature, reference evapotranspiration, and climate extremes [32–35]. In addition to downscaling methods based on interpolation, parametric approaches [36,37] and Bayesian merging techniques [38,39] are also usually used to carry out seasonal downscaling of climate forecasts. However, according to some researchers, downscaling methods with stochastic analog and constructed analog have shown good performance [22,40–45], and seasonal reforecast datasets are generally not long enough to use those analog-based downscaling methods since there is a limited number of potential historical analogs [33].

The seasonal mean air temperature is one of the fundamental products of seasonal prediction [46]. Global and regional temperature variations have shown a significant increasing trend under the background of global warming [47]. Since the 1990s, the spring air temperature over China has increased significantly, and the warming rate is the largest among the four seasons of the year (Figure 1), which makes the prediction of spring air temperatures in China more difficult. Studies have shown that IPCC CMIP5 models are able to reproduce the warming trend of observed global and regional averaged surface air temperatures. However, there are still obvious differences between the temperature trends simulated by the diverse variety of models and the observations [48–50]. Therefore, it is worth studying how well the operational seasonal forecasting model BCC_CSM could forecast the spring temperature while considering the rapid warming during spring in China since the 1990s. Bias in the model-simulated temperature trend leads to forecast error, which has a strong impact on the model forecast performance. Assuming that the operational climate model BCC_CSM could not accurately simulate this trend, correcting the model would be another problem worth studying.

Existing correction methods such as SDBC can improve both the mean and the variance of forecasts' probability distribution. However, the variation trends are ignored, and the forecast error caused by the trend deviation between model simulations and observations cannot be reduced. Thus, a method to correct model simulation trends to reduce model errors and improve forecast skills is urgently needed.

This paper intended to extend SDBC into the spatial disaggregation and detrended bias correction (SDDBC) method by removing the trends of model simulations and observations and then adding the observed trend to both. The performance of seasonal prediction of spring air temperature for the BBC model, SDBC, and SDDBC was evaluated by employing deterministic and probabilistic forecast verification methods. The influence of temperature trend correction on the seasonal predictability of the model was analyzed. The purpose was to improve the forecast skills of the model for spring air temperature. This work

addressed three main questions: (1) How well could the BCC_CSM forecast the trend of spring air temperature? (2) How well did the BCC_CSM and SDBC forecast the spring air temperature? (3) How much could the forecast skill of the BCC_CSM for air temperature be improved by modifying the SDBC through correcting the simulated trend, and what are the advantages of SDDBC over SDBC?

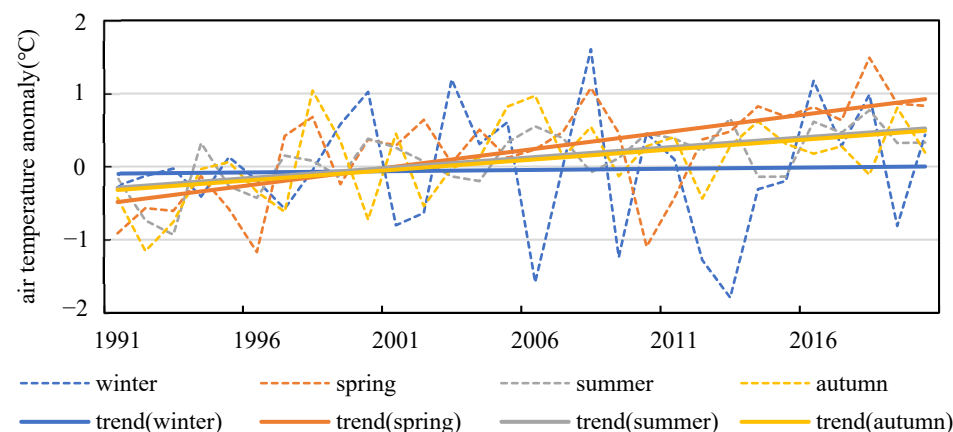


Figure 1. Time series and linear trends of seasonal mean air temperature anomalies from 1991–2020 in China.

2. Data

2.1. Observed Data

The data used in this study were the boreal spring (March to May) mean air temperature data for the period 1991 to 2020 at 160 meteorological stations over China from the China Meteorological Administration. The spatial distribution of the 160 meteorological stations is shown in Figure 2. The stations are divided into seven regions based on the geography and administration of China, including Northeast China (sub-region 1, 17 stations), North China (sub-region 2, 22 stations), East China (sub-region 3, 32 stations), South China (sub-region 4, 14 stations), Central China (sub-region 5, 16 stations), Northwest China (sub-region 6, 31 stations), and Southwest China (sub-region 7, 28 stations).

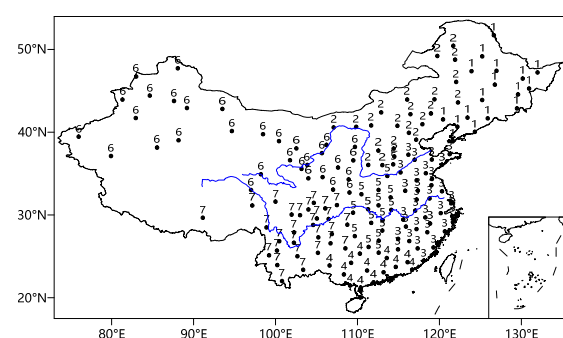


Figure 2. Location of observation stations and sub-regional divisions in China (1. Northeast China; 2. North China; 3. East China; 4. South China; 5. Central China; 6. Northwest China; 7. Southwest China).

2.2. Model Data

The BCC-CSM1.1 (m) model used in this study was developed by the Beijing Climate Center (BCC) of the China Meteorological Administration. This model consists of fully coupled components of the atmosphere, ocean, ice, and land and has been applied in research on climate change projection and climate prediction at the BCC [3]. The BCC_CSM shows a reliable performance in short-term climate prediction [4,5]. The hindcasts and forecasts of the model were initiated from the first day of each month from 1991 to 2020.

In total, 24 ensemble members were used to predict the monthly average atmospheric circulation and surface climatic factors in the next 13 months, with a resolution of $1^\circ \times 1^\circ$. In this study, the spring air temperature forecasts by the model from March 1st were used. The deterministic forecasts were determined by the ensemble mean of the 24 members. The climate state of the observation and the model is the average value from 1991 to 2010.

3. Methods

3.1. Downscaling Methods

(a) SDBC

The spatial disaggregation and bias correction (SDBC) method [31] has two steps. In the first step, the model forecasts are interpolated to a station using inverse distance weighting (IDW). The control points' number of neighbors of IDW is 4, and the weighting function is the inverse power of the distance (a power of 2 was used in this study). In the second step, the interpolated data of the model are bias-corrected based on the station's observation data using the quantile mapping technique [33,34]. The bias-corrected data at time i at station j are calculated as follows:

$$x_{i,j,corr} = F_{o,c}^{-1} \left[F_{f,c} \left(x_{i,j,f} \right) \right] \quad (1)$$

where $F(x)$ and $F^{-1}(x)$ denote the cumulative density function (CDF) of the data and its inverse, respectively; the subscripts f and o indicate model forecasts and observation data, respectively; and the subscript c indicates the calibration period. The cross-validation procedure is conducted by leaving the target year out when creating the CDFs of the observation data.

(b) SDDBC

Since the model cannot accurately simulate temperature trends, in order to reduce the effect of trend simulation errors on forecasts, the SDBC was improved by removing the trend of forecasts and observations ahead of bias correction and then adding the observation trend. The modified method based on SDBC is called spatial disaggregation and detrended bias correction (SDDBC).

This method has four steps. In the first step, the model forecasts are interpolated to a station using IDW. In the second step, the interpolated model data and the observations are detrended (Equations (2) and (3)). In the third step, the detrended data of the model are bias-corrected using the quantile mapping technique based on the detrended data of the observation. In the final step, the observed trend is added to the bias-corrected, downscaled, and detrended model data. The SDDBC method not only corrects the mean and variance of the prediction in the probability space but also further corrects the trend. Thus, bias-corrected data at time i at station j are calculated as follows:

$$\Delta x_{f,c} = x_{f,c} - f(x_{f,c}) \quad (2)$$

$$\Delta x_{o,c} = x_{o,c} - f(x_{o,c}) \quad (3)$$

$$\Delta x_{i,j,m} = x_{i,j,f} - f(x_{f,c}) \quad (4)$$

$$x_{i,j,corr} = F_{o,c}^{-1} \left[F_{f,c} \left(x_{i,j,f} \right) \right] + f(x_{o,c}) \quad (5)$$

where $f(x)$ is the optimal trend fitting of the data, and here is the linear trend fitting based on least square; Δx represents the data after removing the linear trend. In steps 2–4, the cross-validation procedure is conducted by leaving the target year out.

3.2. Evaluation Statistics

(a) RMSE

The root mean square error (RMSE) reflects the difference between forecasts and observations, with smaller values indicating better accuracy [51]. RMSE was calculated as follows [52]:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (f_i - o_i)^2}{n}} \quad (6)$$

where o_i represents observation data, f_i represents the model forecasts or model-corrected forecasts, and n is the amount of data.

(b) ACC and TCC

The anomaly correlation coefficient (ACC) reflects the similarity of anomalous spatial patterns between forecasts and observations [8]. The ACC for the year j was calculated as follows [53]:

$$\text{ACC}_j = \frac{\sum_{i=1}^m (\Delta o_{i,j} - \overline{\Delta o_j})^2 \times \sum_{i=1}^m (\Delta f_{i,j} - \overline{\Delta f_j})^2}{\sqrt{\sum_{i=1}^m (\Delta o_{i,j} - \overline{\Delta o_j})^2 \times \sum_{i=1}^m (\Delta f_{i,j} - \overline{\Delta f_j})^2}} \quad (7)$$

where $\Delta o_{i,j}$ and $\Delta f_{i,j}$ represent the observation and forecast anomalies for year j at station i , respectively; $\overline{\Delta o_j}$ and $\overline{\Delta f_j}$ are the spatial averages of the observation and forecast anomalies, respectively; and m is the number of stations.

The temporal correlation coefficient (TCC) is used to measure the forecast skill for each station. The TCC at station i was calculated as follows [53]:

$$\text{TCC}_i = \frac{\sum_{j=1}^n (o_{i,j} - \overline{o_i})^2 \times \sum_{j=1}^n (f_{i,j} - \overline{f_i})^2}{\sqrt{\sum_{j=1}^n (o_{i,j} - \overline{o_i})^2 \times \sum_{j=1}^n (f_{i,j} - \overline{f_i})^2}} \quad (8)$$

where $o_{i,j}$ and $f_{i,j}$ represent the observations and forecasts for year j at station i , respectively; $\overline{o_i}$ and $\overline{f_i}$ are the time averages of the observations and forecasts, respectively; and n is the number of years. ACC and TCC range between -1 and 1 . The closer they are to 1 , the higher the forecast skill is.

(c) MSSS

The mean squared skill score (MSSS) is a relative skill measure that compares model forecasts with the climatology forecast. MSSS is calculated as follows [54]:

$$\text{MSSS} = 1 - \frac{\sum_j w_j \text{MSE}_j}{\sum_j w_j \text{MSE}_{cj}} \quad (9)$$

where MSE_j is the mean squared error of the model forecasts, MSE_{cj} is the mean squared error of climatology forecasts, and w_j is equal to $\cos(\theta_j)$, where θ_j is the latitude of station j . MSSS ranges from $-\infty$ to 1.0 , with a value of 0 indicating that the forecast has equivalent skill to climatology, negative values indicating that the forecast has less skill than climatology, and a value of 1.0 indicating a perfect forecast.

MSSS _{j} for fully cross-validated forecasts can be expanded as follows [55]:

$$\text{MSSS}_j = \left\{ 2 \frac{S_{ffj}}{S_{oj}} r_{foj} - \left(\frac{S_{ffj}}{S_{oj}} \right)^2 - \left(\frac{[\overline{f_j} - \overline{o_j}]}{S_{oj}} \right)^2 + \frac{2n-1}{(n-1)^2} \right\} / \left\{ 1 + \frac{2n-1}{(n-1)^2} \right\} \quad (10)$$

where $r_{f_{xj}}$ is the product-moment correlation of the forecasts and observations at station j ; $\bar{x}_j$ and $\bar{f}_j$ and s_{xj} and s_{fj} are the average value and root mean square error of observations and forecasts, respectively; and n is the number of years. The first three terms of the decomposition of MSSS_j are related to phase skills (through the correlation), amplitude errors (through the ratio of the forecast to observed variances), and overall bias error of the forecasts [54].

(d) ROCSS

The relative operating characteristic (ROC) is a curve that indicates the relationship between hit rate (HR) and false alarm rate (FAR), and different sorted ensemble members are used as decision thresholds [54]. This prototypical ROC is a plot of HR (ordinate) vs. FAR (abscissa). The area under the ROC curve (AUC) can be used in the calculation of a probabilistic skill score. The approximate integral AUC is calculated as follows [56]:

$$AUC = \sum_{i=1}^{n+1} \frac{(FAR_i - FAR_{i-1}) * (HR_i - HR_{i-1})}{2} \quad (11)$$

where HR_i and FAR_i are the hit rate and the false alarm rate, respectively, and n is the amount of data of probability bins.

The ROC skill score (ROCSS) is calculated from the AUC [53]:

$$ROCSS = 2 \times AUC - 1 \quad (12)$$

(e) BSS

The Brier skill score (BSS) was employed to evaluate the skill of probabilistic forecasts in terciles (above normal, near normal, and below normal) for each station. The BSS is written as follows [53]:

$$BSS = 1 - \frac{BS_f}{BS_c} = \frac{BS_{res} - BS_{rel}}{BS_c} = BSS_{res} - BSS_{rel} \quad (13)$$

where BS_f and BS_c represent the Brier score (BS) of the forecast and climatology, respectively; BSS_{res} and BSS_{rel} are the resolution component and the reliability component of the BSS, respectively; and BS_{res} and BS_{rel} are the resolution component and the reliability component of BS, respectively. The BSS ranges between $-\infty$ and 1.0; values of 1 indicate perfect skill and values of 0 indicate that the skill of the forecast is equivalent to climatology.

4. Results

4.1. Air Temperature Trend

Figure 3 shows the spatial distribution of spring air temperature trends from observations and the BCC_CSM. Both observed and simulated temperatures increased, but at different rates. Overall, the average rate of increase in the observed air temperature over China was 0.49 °C/decade, significantly higher than that of the simulated rate of 0.3 °C/decade. Excluding parts of South China and Southwest China, the simulated air temperature warming rates were lower than the observed rates in most regions. Larger differences were found in North China, Central China, Northwest China, and East China, ranging between 0.22 and 0.25 °C/decade. North China showed the largest difference, where the simulated trend of temperature was 0.31 °C/decade, significantly lower than the observed rate of 0.55 °C/decade. There were fewer differences in South China, Southwest China, and Northeast China, ranging between 0.09 and 0.14 °C/decade. It is suggested that the trend of increasing spring air temperature was underestimated by the BCC_CSM in most parts of China.

Moreover, differences in air temperature trends between the observations and the BCC_CSM results led to an annual variation in the model error. The model error of spring air temperature over China increased significantly at an average rate of 0.18 °C/decade.

Warming rates higher than $0.2\text{ }^{\circ}\text{C}/\text{decade}$ were found over North China, Central China, Northwest China, and East China. The rates over Southwest China, Northeast China, and South China were lower. Therefore, it is necessary to correct the warming trend during the post-processing of the model forecast results to reduce the model forecast error and improve the forecast efficiency.

The spatial distribution of the temperature trend of the SDBC method was consistent with that of the BCC_CSM. The trends of most stations were also below the observed rates, since SDBC did not modify the underestimated warming rate from the BCC_CSM. Meanwhile, the temperature trends from the SDDBC method were close to the observed trends. Thus, the SDDBC method effectively solved the problem of underestimation of the spring air temperature trend by the BCC_CSM.

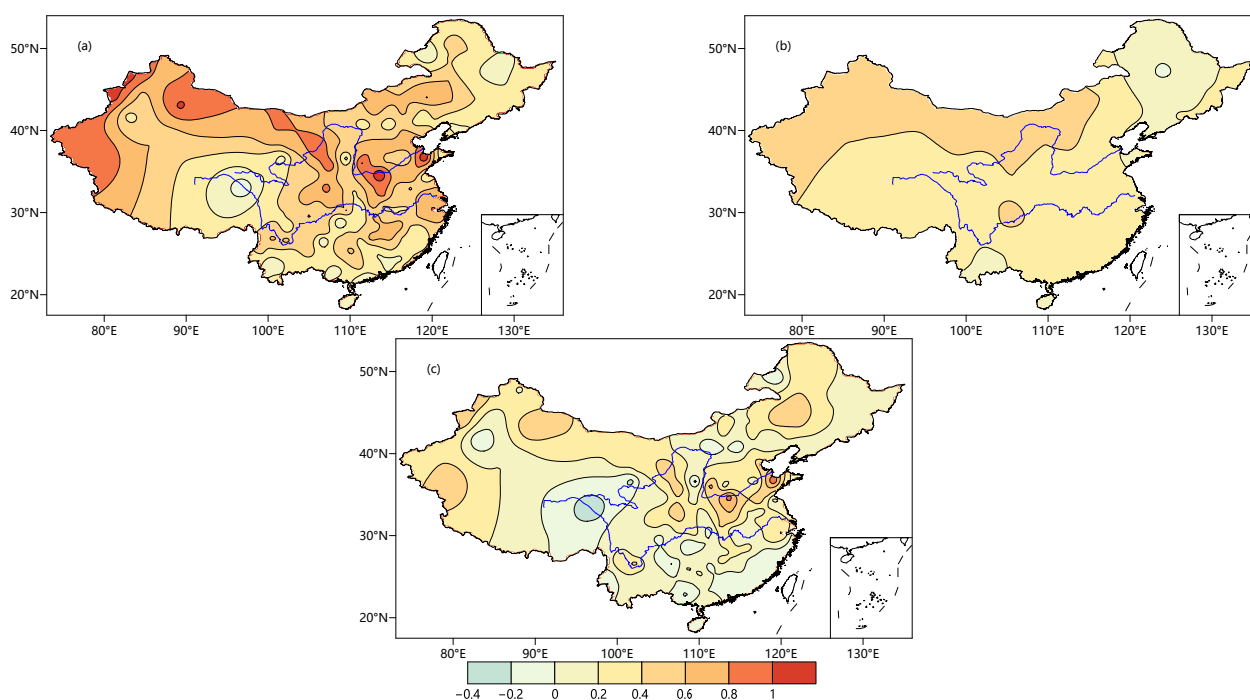


Figure 3. Spatial distribution of spring air temperature trends from (a) observation and (b) the BCC_CSM and (c) the difference between the observed trend and the BCC_CSM trend from 1991 to 2020. Units: $^{\circ}\text{C}/\text{decade}$.

4.2. Deterministic Evaluation of Forecast Skill

(a) RMSE

The BCC_CSM forecasts systematically underestimated the spring temperature over China and the seven sub-regions, and the bias was $3.88\text{ }^{\circ}\text{C}$ on average over China. A larger systematic bias of over $4\text{ }^{\circ}\text{C}$ was found over Northwest China, Northeast China, North China, and Southwest China. The greatest systematic bias was $7.26\text{ }^{\circ}\text{C}$ in Northwest China, while the smallest was in South China with a value of $0.87\text{ }^{\circ}\text{C}$. SDBC and SDDBC effectively eliminated the systematic temperature biases and presented almost no bias in all seven sub-regions.

Figure 4 shows the spatial distribution of the root mean square error (RMSE) values for three methods. The RMSE values of spring air temperature for the BCC_CSM ranged from 0.61 to $14.4\text{ }^{\circ}\text{C}$, averaging at $4.98\text{ }^{\circ}\text{C}$ over the whole country. RMSE was larger in West and North China than in East and South China. The largest RMSE was detected in Northwest China ($7.79\text{ }^{\circ}\text{C}$), followed by Southwest China ($5.59\text{ }^{\circ}\text{C}$), and South China had the smallest error ($1.78\text{ }^{\circ}\text{C}$). RMSE values for the SDBC and SDDBC methods ranged between 0.49 and $1.66\text{ }^{\circ}\text{C}$ and 0.48 and $1.71\text{ }^{\circ}\text{C}$, respectively, averaging at 0.89 and $0.87\text{ }^{\circ}\text{C}$ over China, respectively. The spatial distribution of the RMSE values for these two methods was very

similar. Values higher than 1 °C were found in Northeast China and the western part of Northwest China, while the values ranged from 0.50 to 1.00 °C in most other areas. The RMSE values were greatly reduced by the SDBC and SDDBC methods compared with the BCC_CSM. The RMSE for the SDDBC method was smaller than that for the SDBC method. Lower RMSE values suggested a useful correction of the SDBC and SDDBC methods to the BCC_CSM. Moreover, the SDDBC method performed better than the SDBC method.

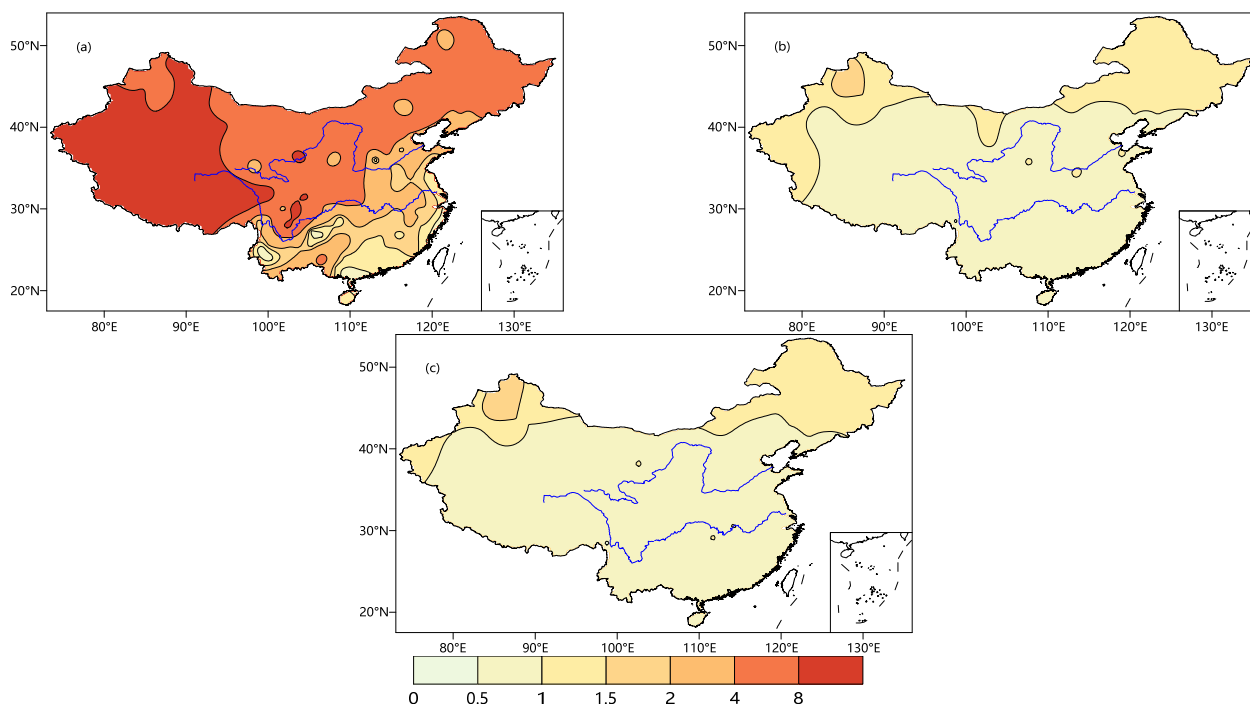


Figure 4. Spatial distribution of RMSE of spring air temperature between observations and forecasts by (a) BCC_CSM, (b) SDBC, and (c) SDDBC from 1991 to 2020 over China. Units: °C.

(b) TCC and ACC

Figure 5 shows the spatial distribution of the temporal correlation coefficient (TCC) of spring air temperature forecast for the BCC_CSM and the SDBC and SDDBC methods. The BCC_CSM was skillful, with TCC values ranging between 0.09 and 0.72 over China. The TCC in North China was significantly higher than that in South China. TCCs above 0.4 (significant at the 5% level) suggested the higher skill of the BCC_CSM in Northeast China, North China, Northwest China, North Central China, and Northeast China. In particular, the TCCs for the northern part of Northwest China, the northern part of East China, the southern part of North China, and the northern part of Central China were greater than 0.5.

The TCC for the SDBC and SDDBC methods was also obviously higher in North China than that in South China, which presented a similar spatial distribution to that for the BCC_CSM. However, the area with a TCC for SDBC above 0.4 was narrowed. Instead, the area with a TCC for SDDBC above 0.4 was expanded. Moreover, SDDBC was more skillful in the southern part of North China, the northern part of East China and the northern part of Central China, with TCCs above 0.7, while the TCC in Northeast China was smaller. In general, the narrowed area with higher TCCs for SDBC led to lower forecast skills and the expanded area with higher TCCs for SDDBC resulted in an improvement in forecast skills.

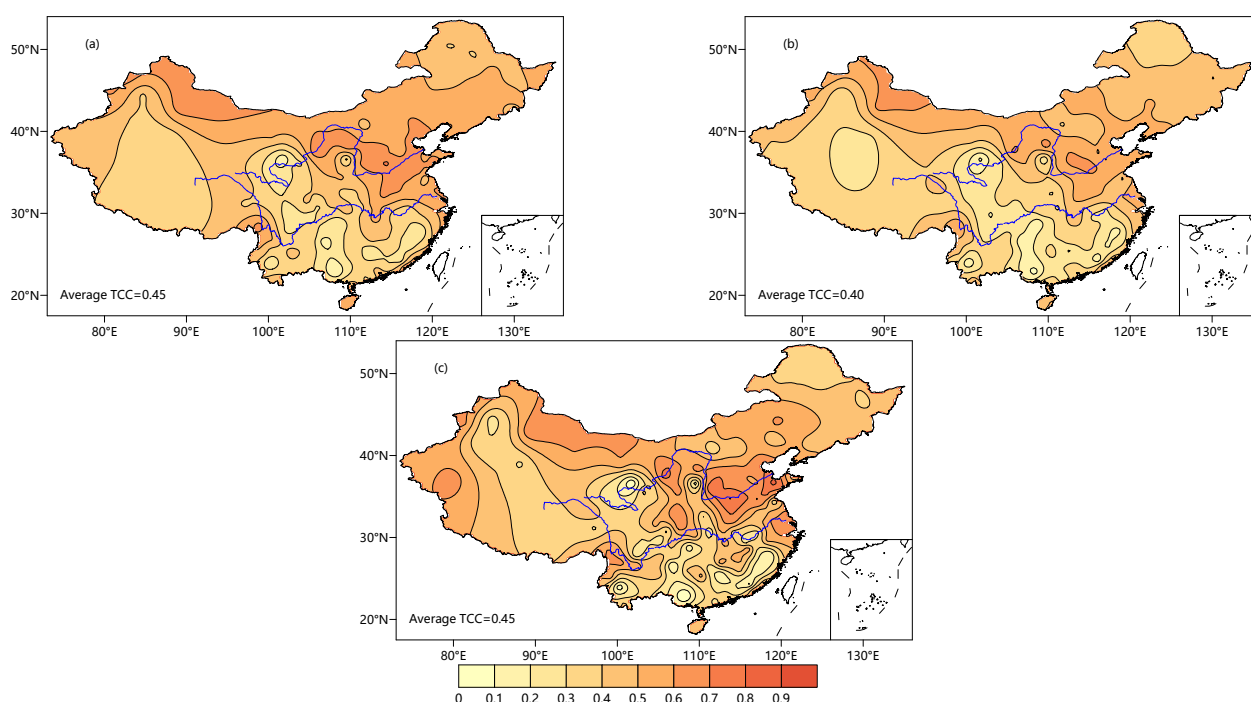


Figure 5. Spatial distribution of TCCs for spring air temperature anomalies from 1991 to 2020 based on observations and forecasts by (a) BCC_CSM, (b) SDBC, and (c) SDDBC over China.

Figure 6 shows the anomaly correlation coefficients (ACCs) of spring air temperature forecasts by the BCC_CSM and the SDBC and SDDBC methods. The BCC_CSM showed statistically significant skill at the 5% level ($ACC = 0.31$) for the whole of China. The ACCs of SDBC and SDDBC were 0.03 lower and 0.04 higher than that of the BCC_CSM, respectively.

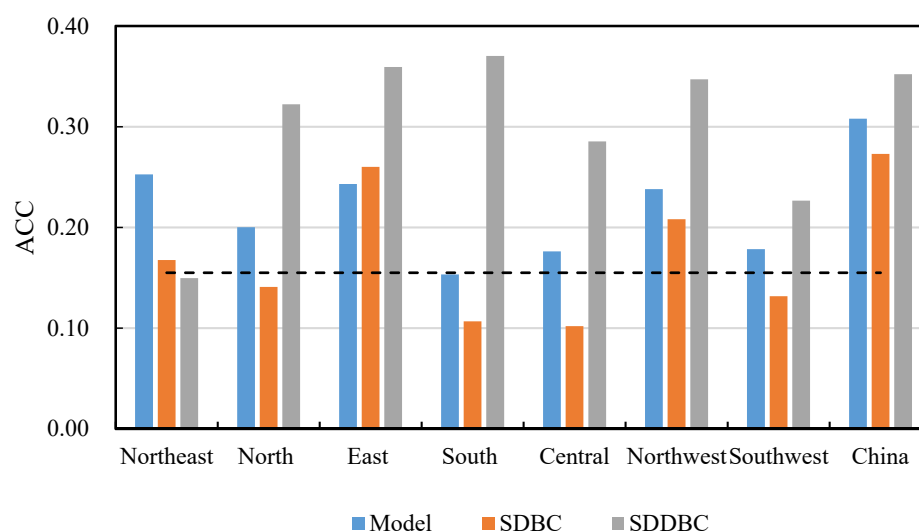


Figure 6. ACC for spring air temperature anomalies from 1991 to 2020 based on observations and forecasts by BCC_CSM, SDBC, and SDDBC over China and seven sub-regions. The dashed black line denotes statistical significance at 95% confidence level based on Student's *t*-test.

The greatest ACC of the BCC_CSM for temperature forecasting was found in Northeast China (0.25), followed by East China and Northwest China (0.24), and the smallest was in South China (0.15). The ACCs were significant at the 5% level in most regions over the whole country, except South China, which suggests a high forecast skill of the BCC_CSM.

SDBC had a lower ACC than the BCC_CSM did in most regions except East China, and the difference was the greatest in Northeast China, with a value of 0.09. SDDBC showed a higher ACC in most areas except Northeast China. The greatest improvement in ACC was found in South China (0.22), followed by North China, East China, Central China, Northwest China (0.11–0.12), and Southwest China (0.05). Thus, in terms of ACC, the SDBC method showed lower skill, while the SDDBC method was obviously more skillful than the BCC_CSM.

(c) MSSS

The MSSS of the BCC_CSM and SDBC for spring air temperature anomaly forecasts in China was 0.18. The MSSS of SDDBC was 0.22, 22% higher than the above two methods. Figure 7 shows the MSSS distribution of the three methods used to forecast spring air temperature. The BCC_CSM was skillful, with a positive MSSS in most parts of China except Southwest and South China. The distribution of MSSS was similar to that of TCC, which showed higher skill in North China than in South China. High MSSS values were found in North China, Northwest China, Northeast China, East China, and the northern part of Central China, ranging between 0.2 and 0.51. The forecasts were most skillful in North China, with an average MSSS of 0.29. The MSSS was lower in South China with an average of 0.05. The BCC_CSM showed no skill in most of Southwest China due to the negative MSSS.

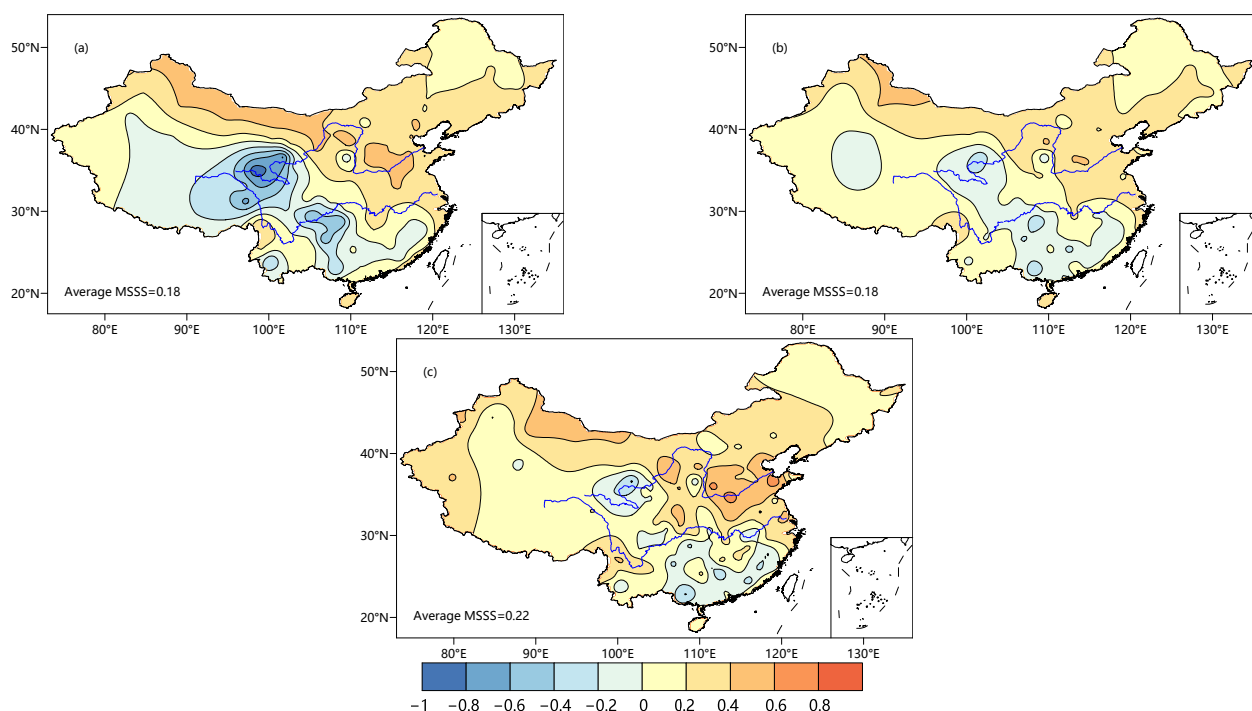


Figure 7. Spatial distribution of MSSS for spring air temperature anomalies from 1991 to 2020 based on observations and forecasts by (a) BCC_CSM, (b) SDBC, and (c) SDDBC over China.

The MSSS spatial distribution of SDBC and SDDBC was close to that of the BCC_CSM, ranging between -0.37 and 0.44 and -0.44 and 0.73 , respectively. The area of SDBC with a positive MSSS was wider than that of the BCC_CSM. The forecast skill was improved in Southwest China and the southwestern part of Northwest China because of the increased MSSS. However, the MSSS of SDBC was lower in almost all of the remaining areas compared with the BCC_CSM, especially for the MSSS turning negative from positive in South China.

The area of SDDBC with a positive MSSS was much larger than that of the BCC_CSM, while an area having no skill ($\text{MSSS} < 0$) was detected in South China. Except Northeast

China and South China with a decreased MSSS, the SDDBC forecast was most skillful among the three methods. A 23% to 34% improvement in MSSS occurred in North China, East China, Central China, Northwest China, and Southwest China.

MSSS is decomposed into phase skill, amplitude error, and systematic error. The systematic error term is close to 0 without considering the drift of climate states. Thus, the MSSS is mainly determined by the phase skill term and the amplitude error term.

Table 1 shows the overall average MSSS for spring temperature anomaly forecasts from 2001 to 2020 based on observations and forecasts by the BCC_CSM, SDBC, and SDDBC methods over China. The phase skill and amplitude error of SDBC both dropped by 0.1 compared with the BCC_CSM. Thus, because of the same influence of these two errors on the MSSS, the forecast skill of SDBC remained. However, the amplitude error of SDDBC increased by 0.01 and the phase skill increased by 0.05. The increase in phase skill being greater than the increase in amplitude error led to the improvement of the forecast skill of SDDBC.

Table 1. MSSS for spring air temperature anomalies from 1991 to 2020 based on observations and forecasts by BCC_CSM, SDBC, and SDDBC over China and seven sub-regions.

Region	Model	MSSS		Phase Skills			Amplitude Errors		
		SDBC	SDDBC	Model	SDBC	SDDBC	Model	SDBC	SDDBC
Northeast China	0.23	0.20	0.18	0.77	0.61	0.61	0.59	0.47	0.49
North China	0.29	0.25	0.31	0.74	0.58	0.74	0.49	0.38	0.49
East China	0.23	0.20	0.26	0.51	0.48	0.69	0.33	0.37	0.53
South China	0.05	−0.01	−0.03	0.43	0.40	0.42	0.44	0.48	0.52
Central China	0.21	0.18	0.24	0.51	0.45	0.64	0.38	0.37	0.52
Northwest China	0.22	0.22	0.27	0.58	0.50	0.68	0.51	0.41	0.53
Southwest China	−0.13	0.04	0.11	0.55	0.33	0.51	0.74	0.35	0.48
China	0.18	0.18	0.22	0.58	0.48	0.63	0.50	0.40	0.51

A lower phase skill of SDBC was found in all seven sub-regions, while a lower amplitude error was detected in most areas, except East and South China. The decrease in phase skill being greater than the increase in amplitude error led to the reduced forecast capability of SDBC in most sub-regions. More than half of the areas showed greater phase skill and larger amplitude error for the SDDBC method than the BCC_CSM, except Northeast, South, and Southwest China. The increase in phase skill being greater than the decrease in amplitude error led to the improvement in forecast skill of SDDBC in most sub-regions.

4.3. Probabilistic Evaluation of Forecast Skill

(a) ROC

ROC and BSS were employed to evaluate the probabilistic forecast skills for the spring air temperature anomalies in China. The skills of the BCC_CSM, SDBC, and SDDBC for above-normal (AN), near-normal (NN), and below-normal (BN) forecasts were compared. Figure 8 shows the ROC diagrams for the three methods. A ROC curve that lies along the 1:1 line indicates no skill, and a curve that is far toward the upper-left corner indicates high skill. The ROC diagrams showed that, of the three methods, the skill was the best for BN forecasts, followed by AN and NN forecasts. Thus, the weak anomaly signal, which is difficult to forecast, led to the lower predictability of the model. SDDBC showed a better performance than the BCC_CSM and SDBC did for all three (AN, NN, and BN) forecasts.

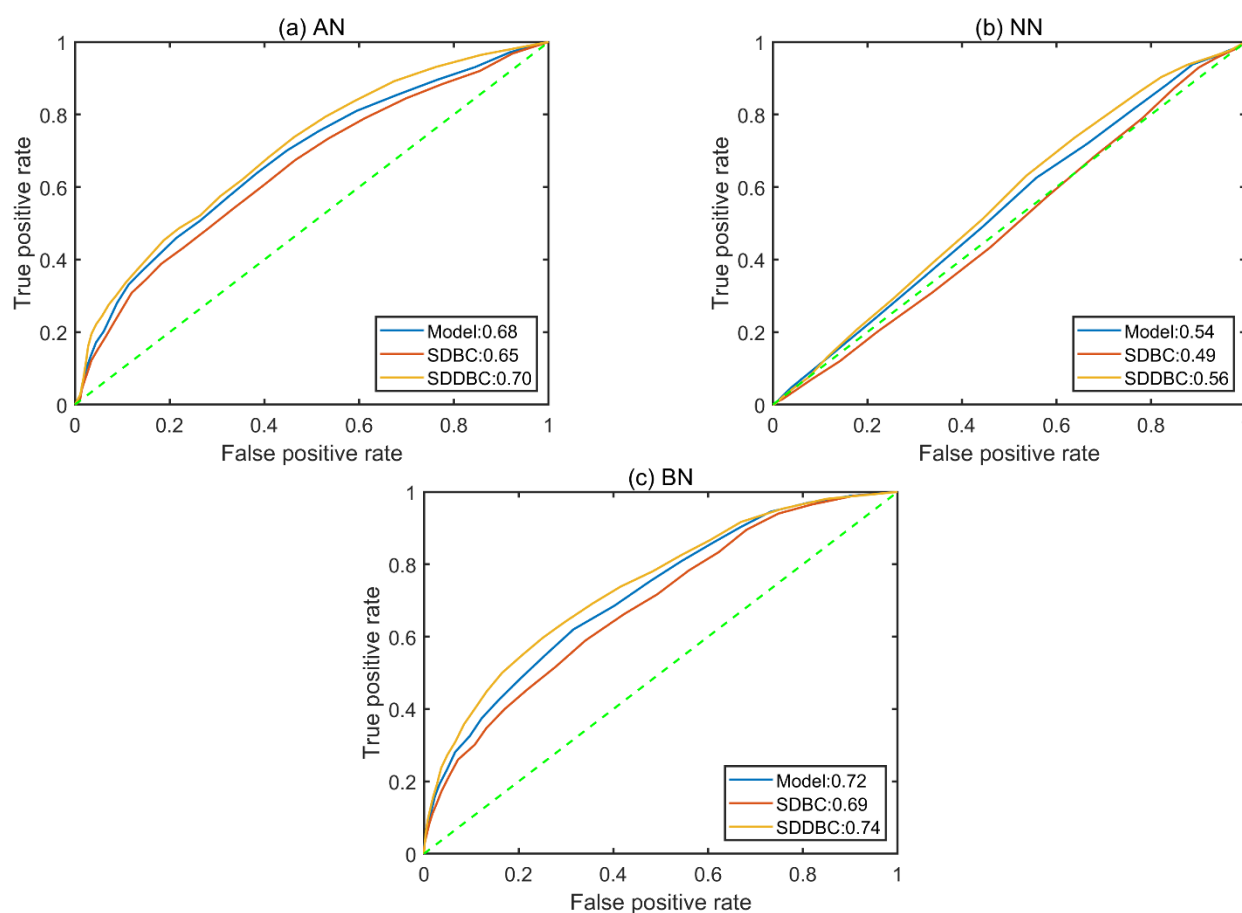


Figure 8. ROC curve for (a) AN, (b) NN, and (c) BN forecasts of spring temperature by BCC_CSM, SDBC, and SDDBC over China.

Figure 9 shows the spatial distribution of the relative operating characteristic skill score (ROCSS) of the spring air temperature for three categorical probabilistic forecasts by the BCC_CSM, SDBC, and SDDBC over China. An ROCSS above 0 denotes a skillful forecast from the method for the stations. Overall, the BCC_CSM showed better skill for AN and BN forecasts than for NN forecasts and was more skillful for BN than AN forecasts. The ROCSS in most areas was above 0.2 for AN forecasts in China. The largest ROCSS of 0.43 was found in Central China, followed by North China, East China, and Northwest China, where the ROCSS was larger than 0.35. The ROCSS for BN forecasts was higher than 0.4 over most areas of China, except South China and parts of Southwest China. The ROCSS for NN forecasts was mostly less than 0.2 and showed no forecast skill in some areas of Northwest and Northeast China. The ROCSS spatial distributions of SDBC and SDDBC were similar to that of the BCC_CSM. SDBC was less skillful than the model for all three categorical forecasts. The site-level average ROCSS for AN, NN, and BN forecasts decreased by 0.06, 0.09, and 0.06, respectively. The areas of NN forecast having no skill (ROCSS less than 0) were wider and became non-skilled from skillful in Central China, South China, and East China. The skill of SDDBC for the three categorical forecasts was better in most areas except Northeast China. In terms of AN forecasts, the ROCSS of North China, East China, Northwest China, and Southwest China enhanced significantly by 0.05 to 0.06. For BN forecasts, the ROCSS increased by more than 0.05 in most of China, particularly increasing by 0.08 to 0.09 in Northwest and Southwest China. For NN forecasts, the non-skilled area significantly narrowed. Compared with that of SDBC, the ROCSS of SDDBC for AN, NN, and BN forecasts was improved by 0.11, 0.13, and 0.11, respectively.

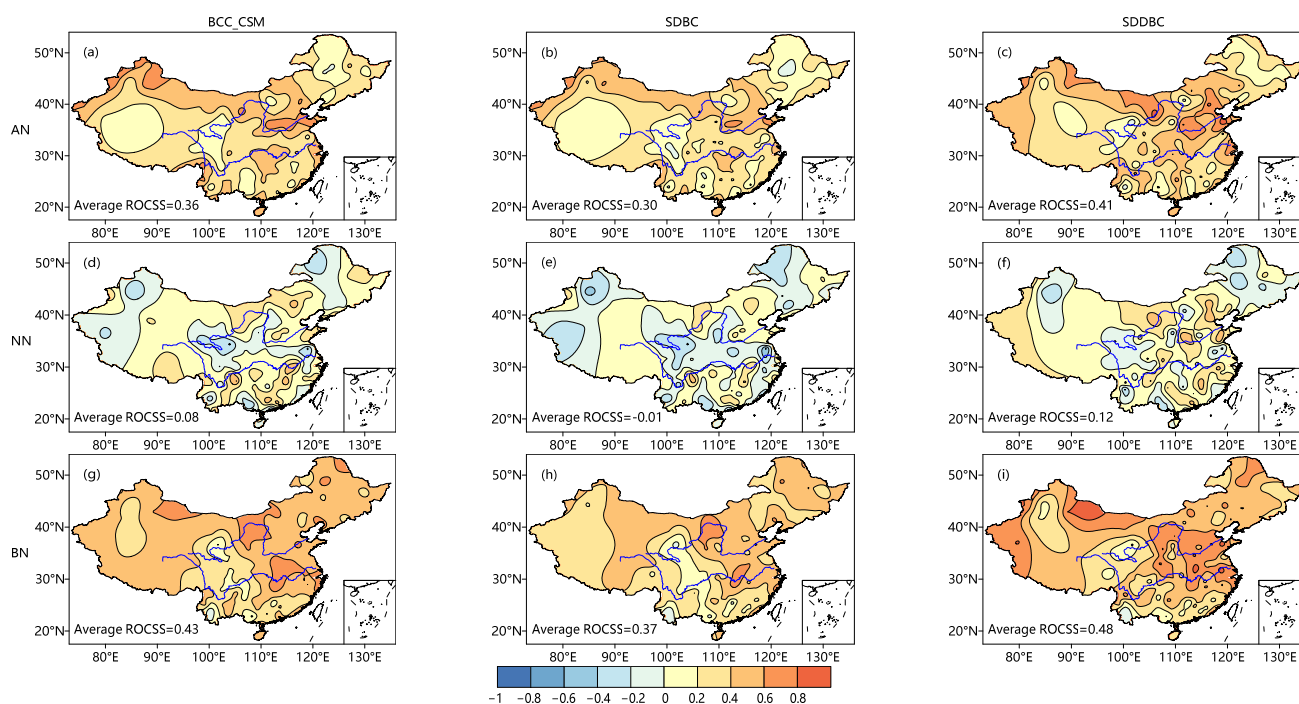


Figure 9. Spatial distribution of ROCSS for (a–c) AN, (d–f) NN, and (g–i) BN forecasts of spring air temperature by (a,d,g) BCC_CSM, (b,e,h) SDBC, and (c,f,i) SDDBC over China.

(b) BSS

Figure 10 shows the Brier skill score (BSS), the resolution component of the BSS (BSSres), and the reliability component of the BSS (BSSrel) for above-normal (AN), near-normal (NN), and below-normal (BN) forecasts of spring air temperature anomaly by the BCC_CSM, SDBC, and SDDBC over China. The negative BSS values indicate that the BCC_CSM, SDBC, and SDDBC had no skill for the NN forecasts. The three methods were found to be more skillful with greater BSS values for BN forecasts than AN forecasts. SDBC showed no skill for AN forecasts and was less skillful than the BCC_CSM for BN forecasts. Based on the BSS in Figure 10, SDDBC was obviously the most skillful for both AN and BN forecasts. This result is consistent with the findings from the ROC diagram. The BSS equation was employed to evaluate the resolution and reliability of the spring air temperature anomaly forecasts. Larger BSSres values suggest higher resolution, and smaller BSSrel values indicate stronger reliability. The BSSres values of the three methods for NN forecasts were close to zero, while the BSSrel values were larger than 0.05. The low resolution combined with the low reliability led to the unskillful NN forecasts by the three methods. The BSSres was larger for BN forecasts than AN forecasts by the three methods, suggesting the higher resolution for BN forecasts than for AN forecasts. The BSSrel for BN forecasts was smaller than for AN forecasts, indicating stronger reliability for BN forecasts. Therefore, the BSS for BN forecasts is greater than AN forecasts, and the three methods showed more skill for BN forecasts.

Compared with the BCC_CSM, the BSSres of SDBC for the three categorical forecasts decreased, and the BSSrel increased, resulting in lower resolution, reliability, and BSS. The BSSres and BSSrel of SDDBC for the three categorical forecasts both increased, resulting in higher resolution and lower reliability. However, the increase in BSSres was greater than that in BSSrel, so the improvement in resolution exceeded the decrease in reliability, leading to the increase in BSS.

Table 2 shows the BSS for three categorical probabilistic forecasts of spring temperature by BCC_CSM, SDBC, and SDDBC over China and seven sub-regions. The BCC_CSM, SDBC, and SDDBC had no skill for NN forecasts in the seven sub-regions of China. The three methods also had no skill for the AN forecasts in Northeast China. The BCC_CSM showed

the largest BSS of 0.16 for BN forecasts among the three methods in Northeast China, indicating that the skill was not improved by the correction of SDBC and SDDBC. For the other six regions, the BSS of the BCC_CSM for BN forecasts was positive. The largest BSS of 0.2 was detected in Central China, followed by North China and East China with a BSS of 0.18, suggesting better forecast skill in these areas. The BSS for AN forecasts was less than that for BN forecasts in all six regions. Northwest China had the highest BSS of 0.09 for AN forecasts, followed by North and Central China with a BSS of 0.07. There was no skill in South China because of the negative BSS. Compared with the BCC_CSM, the skill of SDBC for AN and BN forecasts decreased by between 0.03 and 0.05. SDDBC presented a larger BSS for both AN and BN forecasts and was more skillful than the BCC_CSM. The BSS for AN forecasts was improved significantly by between 0.03 and 0.04 in North China and East China, while the BSS for BN forecasts was improved significantly by between 0.04 and 0.06 in Northwest China, North China, and Central China. The lower resolution and reliability of SDBC for AN and BN forecasts in each sub-region led to the lower BSS, while the same resolution and lower reliability for NN forecasts resulted in the decreased BSS. SDDBC improved the resolution for the three categorical forecasts in all sub-regions, except Northeast China, but deteriorated the reliability in most regions. The improvement in resolution being greater than the decrease in reliability resulted in the increases in BSS. Compared with SDBC, the resolution was improved for the three categorical forecasts in each region by SDDBC, while the variation in reliability differed from area to area. This suggests that the BSS improvement of SDDBC was mainly due to the increase in resolution.

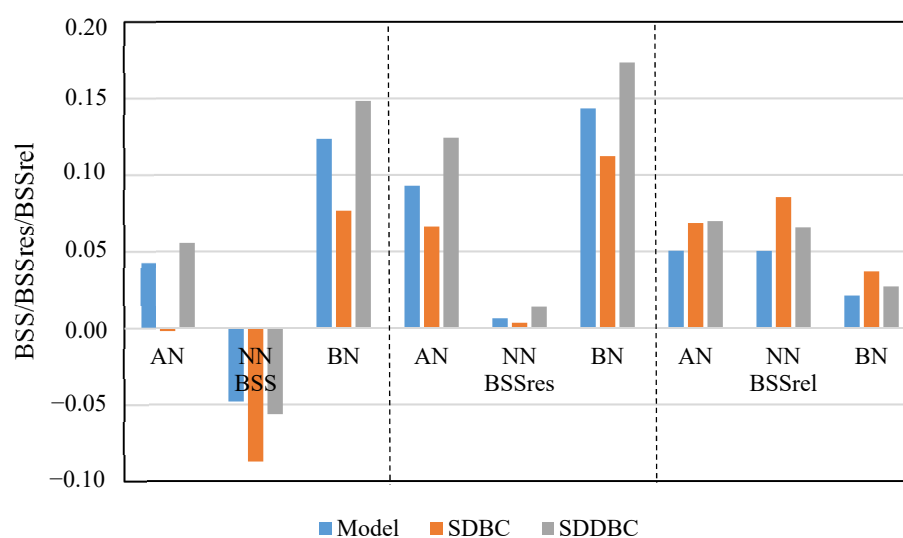


Figure 10. BSS, BSSres, and BSSrel for AN, NN, and BN forecasts of spring air temperature by BCC_CSM, SDBC, and SDDBC over China.

Table 2. BSS for AN, NN, and BN forecasts of spring air temperature by BCC_CSM, SDBC, and SDDBC over China and seven sub-regions.

Region	Above Normal			Near Normal			Below Normal		
	Model	SDBC	SDDBC	Model	SDBC	SDDBC	Model	SDBC	SDDBC
Northeast China	−0.07	−0.13	−0.11	−0.02	−0.07	−0.09	0.16	0.12	0.14
North China	0.07	0.03	0.11	−0.04	−0.07	−0.03	0.18	0.14	0.22
East China	0.06	0.01	0.08	−0.04	−0.08	−0.05	0.18	0.13	0.21
South China	−0.02	−0.07	0.00	−0.08	−0.12	−0.07	0.02	−0.04	−0.02
Central China	0.07	0.03	0.08	−0.03	−0.08	−0.05	0.20	0.15	0.23
Northwest China	0.09	0.05	0.12	−0.09	−0.12	−0.05	0.11	0.07	0.17
Southwest China	0.04	−0.01	0.03	−0.03	−0.07	−0.07	0.02	−0.03	0.05
China	0.04	0.00	0.06	−0.05	−0.09	−0.06	0.12	0.08	0.15

5. Discussion

SDBC can reduce model error by eliminating systematic bias, and this agrees with existing studies [32–35]. However, it did not improve the forecast skills. SDDBC not only retains the advantages of SDBC, but also effectively improves probabilistic and deterministic forecast skills by correcting the temperature trend bias. All the results showed that correcting temperature trend bias is urgent in the post-processing of model outputs. SDDBC performs well in forecasting climatic factors which have obvious variation trends and whose trends cannot be appropriately simulated by the model. It can be used for the forecasting of climatic factors with obvious trends, such as temperature, solar radiation, extreme temperature events, wind speed, atmospheric circulation indexes, subtropical height field, sea surface temperature, and sea ice.

Moreover, the results also showed that improving the forecasting ability of the model for the temperature trend would significantly improve the forecasting ability for seasonal temperature. Future efforts should thus focus on improving the accuracy of the model in simulating trends in climate.

In terms of the lower ACC, MSSS, ROCSS, and BSS, SDDBC was less skillful than the BCC_CSM and SDBC were in Northeast and South China. This is mainly due to the insignificant trend of spring air temperature for these two areas. Trend correction cannot improve the TCC between the BCC_CSM and the observation, and so the forecast skill drops. Therefore, SDDBC may not improve the forecast skill when the trend of climatic factors is not significant, the nonlinear characteristics of the trend are obvious, or the trend is already forecasted well by the model.

6. Conclusions

This study assessed the prediction skill of the BCC_CSM, SDBC, and SDDBC for spring air temperature over China by employing deterministic and probabilistic forecast verification methods from 1991 to 2020. The influence of temperature trend correction on the seasonal prediction capability of the model was analyzed. The main conclusions with discussions can be summarized as follows.

Although the BCC_CSM simulated a significant trend of increasing spring air temperature, the warming rate was obviously underestimated. SDDBC was more skillful than SDBC as it corrected the underestimated air temperature trend.

The BCC_CSM showed a severe cold bias for the spring temperature forecast in China. The RMSE was larger in the west and north than in the east and south. The results of TCC, ACC, and MSSS indicated that the BCC_CSM was skillful for spring air temperature forecast in China, and the skill was higher in the north than in the south. In terms of the probabilistic forecast, the BCC_CSM showed considerable skill in forecasting temperature and was found to be more skillful with a greater BSS for BN forecasts than for AN forecasts, while having minor skill for NN forecasts.

SDBC and SDDBC can effectively eliminate the systematic error of the model and obviously reduce the root mean square error (RMSE). Compared with the model, SDBC cannot improve the anomaly correlation coefficient (ACC) and the mean squared skill score (MSSS) of air temperature forecasts, while SDDBC performed better than the model in terms of ACC and MSSS and was also more skillful than SDBC due to the correction of the temperature trend bias, the increase in temporal correlation between the model, and the observation and the improvement in skill of the model for phase forecast, resulting in the better MSSS and ACC. The relative operating characteristic skill score (ROCSS) and the Brier skill score (BSS) of SDDBC for the three categorical forecasts were higher than those of the model, while the scores given by SDBC were lower. The improvement in resolution exceeded the decrease in reliability, leading to the increase in BSS.

Author Contributions: Conceptualization, C.D. and P.W.; methodology, C.D. and X.W.; software, C.D.; investigation, C.D. and P.W.; data curation, C.D., X.W., Z.C. and R.W.; writing—original draft preparation, C.D. and W.C.; writing—review and editing, C.D., W.C. and R.W.; visualization, C.D.;

project administration, Z.C. and P.W.; funding acquisition, C.D. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant numbers 41605068 and 41805060) and the Innovative Development Special Project of China Meteorological Administration (grant number CXFZ2021Z011).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: The authors are grateful to the reviewers for many valuable suggestions, which helped to improve the quality of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Saha, S.; Moorthi, S.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Behringer, D.; Hou, Y.-T.; Chuang, H.-Y.; Iredell, M.; et al. The NCEP climate forecast system version 2. *J. Clim.* **2014**, *27*, 2185–2208. [\[CrossRef\]](#)
2. Johnson, S.J.; Stockdale, T.N.; Ferranti, L.; Balmaseda, M.A.; Molteni, F.; Magnusson, L.; Tietsche, S.; Decremier, D.; Weisheimer, A.; Balsamo, G.; et al. SEAS5, the new ECMWF seasonal forecast system. *Geosci. Model. Dev.* **2019**, *12*, 1087–1117. [\[CrossRef\]](#)
3. Wu, T.; Song, L.; Li, W.; Wang, Z.; Zhang, H.; Xin, X.; Zhang, Y.; Zhang, L.; Li, J.; Wu, F.; et al. An overview of BCC climate system model development and application for climate change studies. *J. Meteorol. Res.* **2014**, *28*, 34–56. [\[CrossRef\]](#)
4. Liu, X.; Wu, T.; Yang, S.; Li, Q.; Cheng, Y.; Liang, X.; Fang, Y.; Jie, W.; Nie, S. Relationships between interannual and intraseasonal variations of the Asian–western Pacific summer monsoon hindcasted by BCC_CSM1.1 (m). *Adv. Atmos. Sci.* **2014**, *31*, 1051–1064. [\[CrossRef\]](#)
5. Liu, X.; Wu, T.; Yang, S.; Jie, W.; Nie, S.; Li, Q.; Cheng, Y.; Liang, X. Performance of the seasonal forecasting of the Asian Summer Monsoon by BCC_CSM1.1 (m). *Adv. Atmos. Sci.* **2015**, *32*, 1156–1172. [\[CrossRef\]](#)
6. Ren, H.-L.; Wu, J.; Zhao, C.-B.; Cheng, Y.-J.; Liu, X.-W. MJO ensemble prediction in BCC_CSM1.1 (m) using different initialization schemes. *Atmos. Ocean. Sci. Lett.* **2016**, *9*, 60–65. [\[CrossRef\]](#)
7. Gong, Z.; Dogar, M.M.; Qiao, S.; Hu, P.; Feng, G. Assessment and correction of BCC_CSM's performance in capturing leading modes of summer precipitation over North Asia. *Int. J. Climatol.* **2017**, *38*. [\[CrossRef\]](#)
8. Zhou, F.; Ren, H.L. Dynamical feedback between synoptic eddy and low-frequency flow as simulated by BCC_CSM1.1 (m). *Adv. Atmos. Sci.* **2017**, *34*, 1316–1332. [\[CrossRef\]](#)
9. Lu, B.; Ren, H.-L.; Eade, R.; Andrews, M. Indian Ocean SST modes and their impacts as simulated in BCC_CSM1.1 (m) and HadGEM3. *Adv. Atmos. Sci.* **2018**, *35*, 1035–1048. [\[CrossRef\]](#)
10. Rao, J.; Ren, R.; Chen, H.; Liu, X.; Yu, Y.; Yang, Y. Sub-seasonal to Seasonal Hindcasts of Stratospheric Sudden Warming by BCC_CSM1.1 (m), A Comparison with ECMWF. *Adv. Atmos. Sci.* **2019**, *36*, 17–32. [\[CrossRef\]](#)
11. Zhou, F.; Ren, H.; Hu, Z.; Liu, M.; Wu, J.; Liu, C. Seasonal Predictability of Primary East-Asian Summer Circulation Patterns by Three Operational Climate Prediction Models. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 629–646. [\[CrossRef\]](#)
12. Liu, Y.; Fan, K.; Chen, L.; Ren, H.-L.; Wu, Y.; Liu, C. An operational statistical downscaling prediction model of the winter monthly temperature over China based on a multi-model ensemble. *Atmos. Res.* **2021**, *249*, 105262. [\[CrossRef\]](#)
13. Wilby, R.; Dawson, C.; Barrow, E. SDSM—A decision support tool for the assessment of regional climate change impacts. *Environ. Model. Softw.* **2002**, *17*, 147–159. [\[CrossRef\]](#)
14. Murphy, J. An evaluation of statistical and dynamical techniques for downscaling local climate. *J. Clim.* **1999**, *12*, 2256–2284. [\[CrossRef\]](#)
15. Xu, C.Y. From GCMs to river flow. A review of downscaling methods and hydrologic modelling approaches. *Prog. Phys. Geogr.* **1999**, *23*, 229–249. [\[CrossRef\]](#)
16. Giorgi, F.; Mearns, L.O. Approaches to the simulation of regional climate change, A review. *Rev. Geophys.* **1991**, *29*, 191–216. [\[CrossRef\]](#)
17. Zhongfeng, X.U.; Han, Y.; Yang, Z. Dynamical downscaling of regional climate, A review of methods and limitations. *Sci. China Earth Sci.* **2019**, *62*, 21–31. [\[CrossRef\]](#)
18. Yoon, J.-H.; Mo, K.; Wood, E. Dynamic-model-based seasonal prediction of meteorological drought over the contiguous United States. *J. Hydrometeor.* **2012**, *13*, 463–482. [\[CrossRef\]](#)
19. Wood, A.; Maurer, E.; Kumar, A.; Lettenmaier, D.P. Long range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.* **2002**, *107*, 4429. [\[CrossRef\]](#)
20. Wood, A.; Leung, L.R.; Sridhar, V.; Lettenmaier, D.P. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Clim. Chang.* **2004**, *62*, 189–216. [\[CrossRef\]](#)

21. Salathe, E.; Mote, P.W.; Wiley, M.W. Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States Pacific Northwest. *Int. J. Climatol.* **2007**, *27*, 1611–1621. [\[CrossRef\]](#)
22. Maurer, E.P.; Hidalgo, H.G. Utility of daily vs. monthly large-scale climate data, An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 551–563. [\[CrossRef\]](#)
23. Tryhorn, L.; Degaetano, A. A comparison of techniques for downscaling extreme precipitation over the Northeastern United States. *Int. J. Climatol.* **2011**, *31*, 1975–1989. [\[CrossRef\]](#)
24. Bürger, G.; Murdock, T.Q.; Werner, A.T.; Sobie, S.R.; Cannon, A. Downscaling Extremes—An Intercomparison of Multiple Statistical Methods for Present Climate. *J. Clim.* **2012**, *25*, 4366–4388. [\[CrossRef\]](#)
25. Sharma, D.; Babel, M.S. Application of downscaled precipitation for hydrological climate-change impact assessment in the upper Ping River Basin of Thailand. *Clim. Dyn.* **2013**, *41*, 2589–2602. [\[CrossRef\]](#)
26. Lin, W.; Wen, C. A CMIP5 multimodel projection of future temperature, precipitation, and climatological drought in China. *Int. J. Climatol.* **2014**, *34*, 2059–2078. [\[CrossRef\]](#)
27. Sharma, D.; Babel, M.S. Assessing hydrological impacts of climate change using bias-corrected downscaled precipitation in Mae Klong basin of Thailand. *Meteorol. Appl.* **2017**, *25*, 384–393. [\[CrossRef\]](#)
28. Shrestha, R.R.; Schnorbus, M.A.; Cannon, A.J. A Dynamical Climate Model-Driven Hydrologic Prediction System for the Fraser River, Canada. *J. Hydrometeorol.* **2015**, *16*, 150206095227003. [\[CrossRef\]](#)
29. Touseef, M.; Chen, L.; Yang, K.; Chen, Y. Long-Term Rainfall Trends and Future Projections over Xijiang River Basin, China. *Adv. Meteorol.* **2020**, *2020*, 6852148. [\[CrossRef\]](#)
30. Lorenz, C.; Portele, T.C.; Laux, P.; Kunstmann, H. Bias-corrected and spatially disaggregated seasonal forecasts, a long-term reference forecast product for the water sector in semi-arid regions. *Earth Syst. Sci. Data* **2021**, *13*, 2701–2722. [\[CrossRef\]](#)
31. Abatzoglou, J.T.; Brown, T.J. A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.* **2012**, *32*, 772–780. [\[CrossRef\]](#)
32. Hwang, S.; Graham, W.D. Development and comparative evaluation of a stochastic analog method to downscale daily GCM precipitation. *Hydrol. Earth Syst. Sci. Discuss.* **2013**, *10*, 2141–2181. [\[CrossRef\]](#)
33. Tian, D.; Martinez, C.J.; Graham, W.D. Seasonal Prediction of Regional Reference Evapotranspiration Based on Climate Forecast System Version 2. *J. Hydrometeorol.* **2014**, *15*, 1166–1188. [\[CrossRef\]](#)
34. Tian, D.; Martinez, C.J.; Graham, W.D.; Hwang, S. Statistical Downscaling Multimodel Forecasts for Seasonal Precipitation and Surface Temperature over the Southeastern United States. *J. Clim.* **2014**, *27*, 8384–8411. [\[CrossRef\]](#)
35. Ali, S.; Eum, H.-I.; Cho, J.; Dan, L.; Khan, F.; Dairaku, K.; Shrestha, M.L.; Hwang, S.; Nasim, W.; Khan, I.A.; et al. Assessment of climate extremes in future projections downscaled by multiple statistical downscaling methods over Pakistan. *Atmos. Res.* **2019**, *222*, 114–133. [\[CrossRef\]](#)
36. Schaake, J.; Demargne, J.; Hartman, R.; Mullusky, M.; Welles, E.; Wu, L.; Herr, H.; Fan, X.; Seo, D.J. Precipitation and temperature ensemble forecasts from single-value forecasts. *Hydrol. Earth Syst. Sci. Discuss.* **2007**, *4*, 655–717. [\[CrossRef\]](#)
37. Wood, A.W.; Schaake, J.C. Correcting errors in streamflow forecast ensemble mean and spread. *J. Hydrometeorol.* **2008**, *9*, 132–148. [\[CrossRef\]](#)
38. Luo, L.; Wood, E.F. Use of Bayesian merging techniques in a multimodel seasonal hydrologic ensemble prediction system for the eastern United States. *J. Hydrometeorol.* **2008**, *9*, 866–884. [\[CrossRef\]](#)
39. Luo, L.; Wood, E.F.; Pan, M. Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. *J. Geophys. Res.* **2007**, *112*, D10102. [\[CrossRef\]](#)
40. Pierce, D.W.; Cayan, D.R.; Thrasher, B.L. Statistical Downscaling Using Localized Constructed Analogs (LOCA). *J. Hydrometeorol.* **2014**, *15*, 2558–2585. [\[CrossRef\]](#)
41. Chang, S.; Graham, W.; Geurink, J.; Wanakule, N.; Asefa, T. Evaluation of impacts of future climate change and water use scenarios on regional hydrology. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 4793–4813. [\[CrossRef\]](#)
42. Chang, S.; Graham, W.; Geurink, J.; Asefa, T. Evaluation of impact of climate change and anthropogenic change on regional hydrology. *Hydrol. Earth Syst. Sci. Discuss.* **2018**, *3*, 1–43. [\[CrossRef\]](#)
43. Maurer, E.; Hidalgo, H.; Das, T.; Dettinger, M.D.; Cayan, D.R. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1125–1138. [\[CrossRef\]](#)
44. Tian, D.; Martinez, C.J. Forecasting reference evapotranspiration using retrospective forecast analogs in the southeastern United States. *J. Hydrometeorol.* **2012**, *13*, 1874–1892. [\[CrossRef\]](#)
45. Tian, D.; Martinez, C.J. Comparison of two analog-based downscaling methods for regional reference evapotranspiration forecasts. *J. Hydrol.* **2012**, *475*, 350–364. [\[CrossRef\]](#)
46. Wang, S.; Zhu, J. A review on seasonal climate prediction. *Adv. Atmos. Sci.* **2001**, *18*, 197–208. [\[CrossRef\]](#)
47. Li, Q.; Sun, W.; Yun, X.; Huang, B.; Dong, W.; Wang, X.L.; Zhai, P.; Jones, P. An updated evaluation of the global mean Land Surface Air Temperature and Surface Temperature trends based on CLSAT and CMST. *Clim. Dyn.* **2021**, *56*, 635–650. [\[CrossRef\]](#)
48. Xin, X.-G.; Wu, T.-W.; Li, J.-L.; Wang, Z.-Z.; Li, W.; Wu, F.-H. How Well does BCC_CSM1.1 Reproduce the 20th Century Climate Change over China? *Atmos. Ocean. Sci. Lett.* **2013**, *6*, 21–26. [\[CrossRef\]](#)
49. Papalexiou, S.M.; Rajulapati, C.R.; Clark, M.P.; Lehner, F. Robustness of CMIP6 Historical Global Mean Temperature Simulations, Trends, Long-Term Persistence, Autocorrelation, and Distributional Shape. *Earth Future* **2020**, *8*. [\[CrossRef\]](#)

-
50. Kumar, S.; Kinter, J.L.; Pan, Z.; Sheffield, J. Twentieth century temperature trends in CMIP3, CMIP5, and CESM-LE climate simulations spatial-temporal uncertainties, differences and their potential sources. *J. Geophys. Res. Atmos.* **2016**, *121*, 9561–9575. [[CrossRef](#)]
 51. Wen, C.; Duan, C.; Shen, S.; Yao, Y. Evaluation and Parameter Optimization of Monthly Net Long-Wave Radiation Climatology Methods in China. *Atmosphere* **2017**, *8*, 94. [[CrossRef](#)]
 52. Stanski, H.R.; Wilson, L.J.; Burrows, W.R. *Survey of Common Verification Methods in Meteorology*; World Weather Watch Technical Report No. 8, WMO/TD No. 358; World Meteorological Organization: Geneva, Switzerland, 1998; p. 17.
 53. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*; Academic Press: Cambridge, MA, USA; Elsevier: Amsterdam, The Netherlands, 2011; pp. 328–340.
 54. World Meteorological Organization. *Standardised Verification System (SVS) for Long-Range Forecasts (LRF)*; New Attachment II-8 to the Manual on the GDPFS (WMO-No. 485); World Meteorological Organization: Geneva, Switzerland, 2006; Volume I, pp. 68–80.
 55. Murphy, A.H. Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Weather Rev.* **1988**, *16*, 2417–2424. [[CrossRef](#)]
 56. Hamill, T.M.; Juras, J. Measuring forecast skill, is it real skill or is it the varying climatology? *Q. J. R. Meteorol. Soc.* **2006**, *132*, 2905–2923. [[CrossRef](#)]