Hydrology and
Earth System
Sciences

A statistical–dynamical approach for probabilistic prediction of sub-seasonal precipitation anomalies over 17 hydroclimatic regions in China

Yuan Li[1], **Kangning Xü**[1], **Zhiyong Wu**[1], **Zhiwei Zhu**[2], and **Quan J. Wang**[3]

[1]College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China
[2]Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD)/Joint International Research Laboratory of Climate and Environment Change (ILCEC), Nanjing University of Information Science and Technology, Nanjing 210044, China
[3]Department of Infrastructure Engineering, The University of Melbourne, Parkville 3010, Australia

Correspondence: Zhiyong Wu (wzyhhu@gmail.com)

Abstract. In this study, we develop a spatial–temporal projection-based calibration, bridging, and merging (STP-CBaM) method to improve probabilistic sub-seasonal precipitation forecast skill over 17 hydroclimatic regions in China. The calibration model is established by post-processing ECMWF raw forecasts using the Bayesian joint probability (BJP) approach. The bridging models are built using large-scale atmospheric intraseasonal predictors, including zonal wind at 200 hPa (U200) and 850 hPa (U850); an outgoing longwave radiation anomaly (OLRA); and geopotential height at 200 hPa (H200), 500 hPa (H500), and 850 hPa (H850) defined by the STP method. The calibration model and the bridging models are then merged through the Bayesian modelling averaging (BMA) method. Our results indicate that the forecast skill of the calibration model is higher compared to bridging models when the lead time is within 5–10 d. The U200- and OLRA-based bridging models outperform the calibration model in certain months and certain regions. The BMA-merged forecasts take advantage of both calibration models and bridging models. Meanwhile, the BMA-merged forecasts also show high reliability at longer lead times. However, some improvements to reliability are still needed at shorter lead times. These findings demonstrate the great potential to combine dynamical models and statistical models in improving sub-seasonal precipitation forecasts.

1 Introduction

Sub-seasonal forecasting (defined as the time range between 2 weeks and 2 months) bridges the gap between short- to medium-range weather forecasts and seasonal climate prediction (Vitart and Robertson, 2018; Liu et al., 2023). Skilful and reliable sub-seasonal precipitation forecasts are highly valuable for water resource management, flood disaster preparedness, and many other climate-sensitive sectors (White et al., 2022; Yan et al., 2023; Zhu et al., 2022). However, it is considered a difficult time range to generate skilful forecasts. The memory of atmospheric initial conditions is lost compared to short- to medium-range forecasts, while the variability of lower-boundary conditions, such as sea surface temperature, is too slow to take effect (Vitart and Robertson, 2018). Statistical models, which use observational relationships between sub-seasonal precipitation and atmospheric intraseasonal oscillations, have been developed in recent years. The spatial–temporal projection model (STPM), which extracts the coupled patterns of preceding atmospheric intraseasonal oscillations and precipitation, has shown skill in predicting sub-seasonal precipitation. Zhu and Li (2017) constructed STPMs over different climatic regions during the boreal summer monsoon season, and their results indicated that the STPMs could generate skilful forecasts for intraseasonal precipitation patterns with lead times of up to 20 d. Our previous study developed a spatial–temporal projection-based Bayesian hierarchical model (STP-BHM) to take the uncer-

tainties in the relationships between atmospheric intraseasonal oscillations and sub-seasonal precipitation into account (Li et al., 2022). However, statistical models are highly reliant on stationary relationships between predictors and predictands. Seasonal changes in climatological conditions may lead to different relationships between atmospheric intraseasonal oscillations and precipitation. Liu and Lu (2022) suggested that the impacts of boreal summer intraseasonal oscillation (BSISO) on precipitation are different between early and late summers. W. Li et al. (2023) found that the long-period BSISO event-affected region and the associated precipitation anomalies are different compared to short-period BSISO events.

With a more comprehensive understanding and better representation of potential sources of predictability, there has been much improvement in dynamical models in recent years. The Subseasonal-to-Seasonal (S2S) prediction project and the Subseasonal Experiment (SubX) project have been established to provide S2S forecasts from dynamical models. However, the sub-seasonal precipitation forecasts of global climate models (GCMs) are always of low accuracy (de Andrade et al., 2019; Li et al., 2022). The physical equations are simplified, while the small-scale processes, such as convections, cannot be well represented in most GCMs. In addition, insufficient data assimilation schemes, low capacity in capturing dynamic sources, and misrepresentation of atmosphere–ocean interactions and atmosphere–ocean interactions also contribute to the limited forecast skill (Wu et al., 2023; Zhang et al., 2021). Although post-processing methods have been proposed in recent years, the forecast skill after post-processing was still limited for lead times beyond 10–14 d (Li et al., 2021).

Despite the low forecast skill of sub-seasonal precipitation, the GCMs show much higher performance in predicting large-scale circulation patterns. Cui et al. (2021) evaluated the potential of GCMs to predict intraseasonal surface air temperature over mid- to high-latitude Eurasia. Their results indicated that the upper limit of the useful forecast skill ranged from ~ 10 to ~ 20 d. The BSISO is the predominant variability of the Asian summer monsoon, and most GCMs exhibit predictability on timescales of above 3 weeks for BSISO events (Chen and Zhai, 2017; Hsu et al., 2016; Ren et al., 2018). Lee et al. (2015) evaluated the prediction skill of BSISO indices using six coupled models in the Intraseasonal Variability Hindcast Experiment (ISVHE) project, and their results suggested that skilful BSISO prediction was about 22 d under strong initial conditions. Shibuya et al. (2021) suggested that the overall useful prediction skill of the BSISO was approximately 24 d in a global non-hydrostatic icosahedral atmospheric model (NICAM) with explicit cloud microphysics. Similar results were also found by Wu et al. (2023), i.e. that the ECMWF model showed skilful prediction of the BSISO index at a 24 d lead time.

Given the strengths and weaknesses of both statistical models and dynamical models, there has been growing interest in developing hybrid prediction models that combine forecasts from both statistical and dynamical models (Slater et al., 2023). Schepen et al. (2014) used POAMA (Predictive Ocean Atmosphere Model for Australia) forecasts of seasonal climate indices as predictors to predict seasonal precipitation over Australia. Strazzo et al. (2019b) developed a hybrid statistical–dynamical system to predict seasonal temperature and precipitation over North America. Most previous statistical–dynamical models focus on seasonal predictions. Far fewer attempts have been made on sub-seasonal timescales. Specq and Batté (2020) proposed a statistical–dynamical post-processing scheme to improve the quality of sub-seasonal forecasts of weekly precipitation using Madden–Julian Oscillation (MJO) and El Niño–Southern Oscillation (ENSO) indices as predictors. Wu et al. (2022) established a dynamical–statistical prediction model (DSPM) to improve sub-seasonal precipitation forecasts. Deep-learning models were also proposed to predict sub-seasonal extreme rainfall events, with the GCM-predicted large-scale circulation patterns used as predictors (Xie et al., 2023). Zhu et al. (2023) developed a dynamical–statistical hybrid model using the novel indices of the zonal displacements of the South Asia high and the western Pacific sub-tropical high to predict the Meiyu intraseasonal variation. Nevertheless, the relationships between large-scale circulation patterns and sub-seasonal precipitation have high uncertainty. More sophisticated hybrid models are required to further improve the probabilistic sub-seasonal precipitation forecast skill.

The calibration, bridging, and merging (CBaM) method, which employed Bayesian-theorem-based approaches to take advantage of both dynamical models and statistical models, has proven able to generate skilful and reliable seasonal precipitation and temperature forecasts over different regions (Peng et al., 2014; Schepen et al., 2014, 2016; Strazzo et al., 2019a). In calibration, the Bayesian joint probability (BJP) approach was used to post-process raw precipitation forecasts derived from GCMs. The BJP approach was also used to generate probabilistic forecasts using large-scale circulation patterns as predictors. This was also referred to as bridging. The calibrated forecasts and bridged forecasts were then merged through the Bayesian modelling averaging (BMA) method (Wang et al., 2012). Most previous studies used the CBaM method to generate seasonal forecasts. However, much less work has been done on sub-seasonal timescales for several reasons. Compared to seasonal forecasts, there are far fewer climate indices that can be used as predictors on sub-seasonal timescales. Moreover, the atmospheric intraseasonal oscillations may have different effects on precipitation anomalies in different months. As a consequence, it is much more difficult to establish bridging models for sub-seasonal precipitation forecasts. In addition, the evolution of the intraseasonal variability of precipitation varies in different stages with different periods in different regions (Liu et al., 2020; Zhu and Li, 2017). The effectiveness of the cali-

bration models will be greatly affected if seasonality is not considered.

In this study, we develop a spatial–temporal projection-based calibration, bridging, and merging (STP-CBaM) method to improve probabilistic sub-seasonal precipitation forecast skill by combining the strengths of both dynamical models and statistical models. The ECMWF sub-seasonal precipitation forecasts are calibrated using the BJP approach for each month. The bridging models are then built using large-scale atmospheric intraseasonal predictors defined by the STPM. The calibration model and bridging models are merged through the BMA method to generate skilful and reliable sub-seasonal precipitation forecasts. The STP-CBaM method will be applied to predict sub-seasonal precipitation anomalies over each hydroclimatic region during the boreal summer monsoon from May to October. The accuracy and reliability will be evaluated through a leave-one-year-out cross-validation strategy.

In the following two sections, the data and methodology are introduced. The prediction skill and reliability of the STP-CBaM method are provided in Sect. 4. Section 5 discusses the forecast skill, limitations, and future work. Key findings are summarized in Sect. 6.

2 Data

2.1 Precipitation dataset

In this study, China is divided into 17 hydroclimatic regions on the basis of both climate classifications and watershed division standards (Fig. 1). The precipitation data are derived from the latest Multi-Source Weighted-Ensemble Precipitation, version 2.8 (MSWEP V2.8) dataset. This dataset is developed by optimally merging precipitation data derived from gauge, satellite, and reanalysis datasets. It covers the period from 1979 to the near past with a spatial resolution of $0.1° \times 0.1°$. Many studies have found that the MSWEP dataset is of a high quality over China (Y. Li et al., 2023; Liu et al., 2019; Guo et al., 2023).

2.2 Reanalysis dataset and outgoing long-wave radiation (OLR) dataset

The daily mean geopotential height at 200, 500, and 850 hPa (H200, H500, H850) and zonal wind at 200 hPa (U200) and 850 hPa (U850) are derived from the ERA5 (Hersbach et al., 2020) reanalysis dataset at https://cds.climate.copernicus.eu/ (last access: 10 December 2022). The daily mean OLR data are provided by the National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory (PSL), Boulder, Colorado, USA, from their website at https://psl.noaa.gov (last access: 17 December 2022). The OLR data are developed from high-resolution infrared radiation sounder instruments and have been widely used over the globe. All daily mean data including U200, U850, OLR, H200, H500,

and H850 are bilinearly interpolated onto a horizontal resolution of $2.5° \times 2.5°$ over the period of 2001–2020.

2.3 Hindcast dataset

The ECMWF hindcast data of precipitation, U850, U200, OLR, H200, H500, and H850 are retrieved from the S2S database at http://apps.ecmwf.int/datasets/data/s2s/ (last access: 31 December 2022). Compared to other GCMs, the ECMWF model shows the highest forecast skill in various aspects (Jie et al., 2017; Wu et al., 2023; Zhang et al., 2021). In this study, we choose hindcasts when the ECMWF model version dates are in the year 2021 from May to October. Thus, the hindcasts cover the period of 2001–2020. The gridded precipitation hindcasts are area-weighted averages through 17 hydroclimatic regions as the observational data. In addition, all atmospheric hindcast fields including U200, U850, OLRA, H200, H500, and H850 are bilinearly interpolated onto a horizontal resolution of $2.5° \times 2.5°$ as the reanalysis dataset.

3 Methodology

3.1 Intraseasonal signal extraction

In this study, a non-filtering method is used to extract 10–60 d signals for both atmospheric variables (U200, U850, OLRA, H200, H500, and H850) and precipitation (Hsu et al., 2015; Zhu et al., 2015). The climatological annual cycle of observational data is first removed by subtracting a 90 d low-pass filtered climatological component. Lower-frequency signals are then removed by subtracting the last 30 d running mean. The higher-frequency signals are then removed by taking a pentad mean. The variable derived in this way represents the observational 10–60 d signals of a daily atmospheric field or precipitation. The daily intraseasonal signals are then averaged to pentad data to further reduce the noise and improve the predictability. The pentad mean 10–60 d intraseasonal precipitation is also referred to as pentad mean precipitation anomalies in the following sections.

As for the hindcast fields of the ECMWF model, the model climatology of the atmospheric variables (U200, U850, OLRA, H200, H500, and H850) and precipitation is removed as a function of the initial date and lead time. Lower-frequency signals longer than 60 d are then removed in the same way as the observations by subtracting the running mean of the last 30 d. In this process, the observed anomalies before the forecast initial date are used to make enough data for the running mean. The higher-frequency signals of the predicted variables are then removed by taking a pentad mean. The variable derived in this way represents the ECMWF model-forecasted 10–60 d signals of the daily atmospheric field or precipitation.
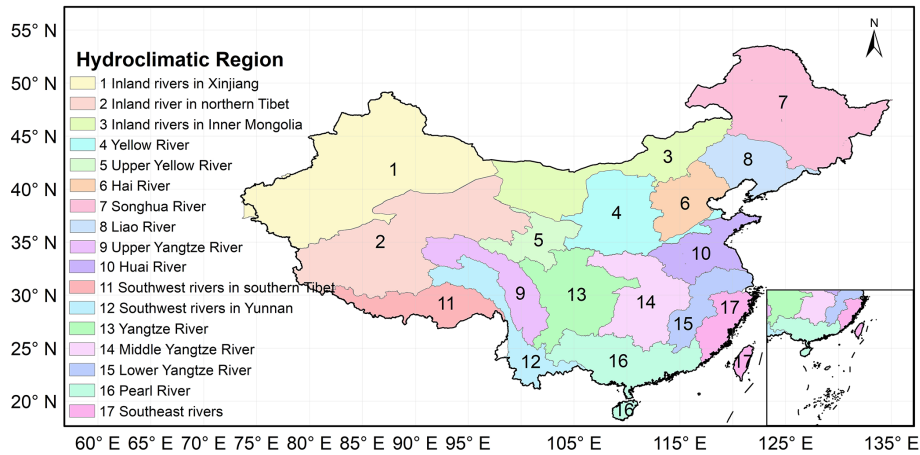
Figure 1. The 17 hydroclimatic regions over China. Publisher's remark: please note that the above figure contains disputed territories.

3.2 Model formulation

3.2.1 Predictor definition for bridging models

In this study, we establish the calibration model and bridging models for each hydroclimatic region, month, and lead time. For the calibration model, the ensemble means of ECMWF-forecasted pentad mean precipitation anomalies are used as predictors. For the bridging models, we define potential predictors using the STPM. Relevant areas of atmospheric fields that could affect 10–60 d precipitation variability are found by cell-wise correlation analysis. The effective degree of freedom is estimated following Livezey and Chen (1983).

The spatial–temporal coupled covariance patterns are constructed for grid points where the correlation is statistically significant at the 5 % level. The predictor is then defined by summing the product of the covariance field derived from the observational data and the ECMWF model-forecasted atmospheric intraseasonal signals:

$$\text{cov}\left(X_{i,p}, Y\right) = \frac{1}{T} \sum_{t=1}^{T} \left(y_t - E(y)\right)\left(x_{i,p,t} - E\left(x_{i,p}\right)\right), \quad (1)$$

$$X_p^* = \sum_{i=1}^{N} \text{cov}\left(X_{i,p}, Y\right) \cdot X_{i,p}^*, \quad (2)$$

where $X_{i,p}$ denotes the pentad mean 10–60 d signal of the pth observational atmospheric field (U200, U850, OLRA, U200, H500, H850) where the correlation is statistically significant at the 5 % level for grid i during the training period, $p = 1, 2, \ldots, 6$. Y denotes the corresponding pentad mean precipitation anomalies. $X_{i,p}^*$ denotes the pentad mean 10–60 d signal of the pth hindcast atmospheric field derived from the ECMWF model for grid i. X_p^* denotes the pth predictor defined by the STPM.

3.2.2 Calibration and bridging models

The calibration model and bridging models are established independently of each other, and each model has only one predictor and one predictand. Therefore, there is one calibration model and there are six bridging models for each hydroclimatic region, month, and lead time.

Each calibration model or bridging model is established using the BJP approach. The predictor X_k, $k = 1, \ldots, K$ and the corresponding predictand Y (pentad mean precipitation anomalies) are normalized to U_k and V using the Yeo–John transformation method.

$$u_k = \begin{cases} \frac{\left\{(x_k+1)^{\lambda_{x_k}} - 1\right\}}{\lambda_{x_k}} & \left(x_k \geq 0, \lambda_{x_k} \neq 0\right) \\ \log(x_k + 1) & \left(x_k \geq 0, \lambda_{x_k} = 0\right) \\ -\frac{\left\{(-x_k+1)^{2-\lambda_{x_k}} - 1\right\}}{2-\lambda_{x_k}} & \left(x_k < 0, \lambda_{x_k} \neq 2\right) \\ \log(-x_k + 1) & \left(x_k < 0, \lambda_{x_k} = 2\right) \end{cases} \quad (3)$$

$$v = \begin{cases} \frac{\left\{(y+1)^{\lambda_y} - 1\right\}}{\lambda_y} & \left(y \geq 0, \lambda_y \neq 0\right) \\ \log(y + 1) & \left(y \geq 0, \lambda_y = 0\right) \\ -\frac{\left\{(-y+1)^{2-\lambda_y} - 1\right\}}{2-\lambda_y} & \left(y < 0, \lambda_y \neq 2\right) \\ \log(-y + 1) & \left(y < 0, \lambda_y = 2\right) \end{cases} \quad (4)$$

λ_{x_k} and λ_y are the unknown transformation parameters for predictor X_k and predictand Y.

The matrix $\mathbf{Z}^T = [U_k V]$ is then assumed to follow a bivariate normal distribution,

$$\mathbf{Z} \sim \mathbf{N}(\boldsymbol{\mu}\boldsymbol{\Sigma}), \quad (5)$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrices to be estimated:

$$\boldsymbol{\mu}^T = \left[\mu_{U_k} \mu_V \right], \tag{6}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma} \mathbf{R} \boldsymbol{\sigma}^T, \tag{7}$$

where $\boldsymbol{\sigma}$ and $\mathbf{R}$ are the standard deviation vector and correlation coefficient matrix, respectively:

$$\boldsymbol{\sigma}^T = \left[\sigma_{U_k} \sigma_V \right], \tag{8}$$

$$\mathbf{R} = \begin{bmatrix} 1 & r_{U_k V} \\ r_{V U_k} & 1 \end{bmatrix}. \tag{9}$$

Note that the correlation coefficient matrix $\mathbf{R}$ is symmetric. Thus, there are only five unknown parameters. Here, we denote the five unknown parameters of the joint distribution as $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$.

Given a data series of $\mathbf{D}_0 = (x_{k,t} y_t), \, t = 1, \ldots, n\}$, we apply the SCE-UA (shuffled complex evolution method developed at the University of Arizona) method to estimate transformation parameters that maximize the log-likelihood function (Duan et al., 1994). The data series $\mathbf{D}_0$ is then normalized to $\mathbf{D} = \{(u_{k,t} v_t), \, t = 1, \ldots, n\}$. The posterior distribution of $\boldsymbol{\theta}$ is estimated using a Bayesian framework:

$$p(\boldsymbol{\theta}|\mathbf{D}(U_k V)) \propto p(\theta) p(\mathbf{D}(U_k V)|\boldsymbol{\theta}), \tag{10}$$

where $p(\boldsymbol{\theta})$ is the prior distribution of the parameters, and $p(\mathbf{D}(U_k V)|\boldsymbol{\theta})$ is the likelihood. As the posterior distributions of parameters $\boldsymbol{\theta}$ are not standard distributions, analytical integration is difficult. To overcome this problem, we use the new Gibbs sampling algorithm proposed by Wang et al. (2019) to draw a sample of 1000 sets of parameter values. A more detailed description of the sampling strategy can be found in Li et al. (2021).

The posterior predictive distribution of $v_k(t^*)$ is given by

$$f_{V_k}\left(v_k\left(t^*\right)\right) = \int p\left(v_k\left(t^*\right)|u_k\left(t^*\right), \boldsymbol{\theta}\right) p(\boldsymbol{\theta})$$
$$p\left(\mathbf{D}\left(U_k, V\right)|\boldsymbol{\theta}\right) \mathrm{d}\boldsymbol{\theta}, \tag{11}$$

where $x_k(t^*)$ is the new forecast value.

Again, the Gibbs sampling algorithm is used to obtain 1000 samples of $f_{V_k}(v_k(t^*))$. The samples are then back-transformed to produce the calibrated or bridged predictive density $f_k(y|x_k)$ using parameter λ_y.

3.2.3 Combining models

The merging forecasts are carried out by the BMA method proposed by Wang et al. (2012). Given all the candidate models, $f_k(y|x_k), \, k = 1, \ldots, K$, and the corresponding model weights, $w_k, k = 1, \ldots, K$, the predictive density of the BMA probabilistic forecasts can be represented as

$$f_{\mathrm{BMA}}(y|x_1, \ldots, x_K) = \sum_{i=1}^{K} w_k f_k(y|x_k), \tag{12}$$

where x_k is the predictor and y is the corresponding predictand.

To encourage even weights among the models, the prior of the model weights is assumed to follow a symmetric Dirichlet distribution, given as

$$p(w_k, k = 1, \ldots, K) \propto \prod_{k=1}^{K} (w_k)^{\alpha-1}, \tag{13}$$

where α is the concentration parameter slightly over 1 and, more specifically, $\alpha = 1 + \alpha_0/K$ and $\alpha_0 = 0.5$. The posterior distribution of model weights given $t = 1, \ldots, T$ events is as follows:

$$p\left(w_k, k = 1, \ldots, K \, \middle| \, x_k^T, y^T, \, f_k(y|x_k) k = 1, \ldots, K\right)$$
$$\propto \prod_{k=1}^{K} (w_k)^{\alpha-1} \prod_{t=1}^{T} \sum_{k=1}^{K} w_k f_k^{(t)}\left(y^t|x_k^t\right), \tag{14}$$

where $f_k^{(t)}(y^t|x_k^t)$ is the cross-validated predictive density. This indicates that the weights are assigned by the model's predictive ability rather than the fitting ability. An expectation–maximization (EM) algorithm is then used to estimate the weights by maximizing the likelihood function. Initially, all the weights are equal. The EM algorithm is then iterated until the likelihood function converges.

3.3 Evaluation

In this study, a leave-one-year-out cross-validation strategy is used to avoid any bias in skill, including predictor selection, data normalization, model building, parameter inference, and verification.

The temporal correlation coefficient (TCC) is used to evaluate the performance of the ECMWF model in predicting atmospheric intraseasonal oscillations. We should note that the ECMWF model has an initial frequency of twice a week on Tuesday and Thursday. Therefore, 160 or 180 initial dates are found for each month during the period of 2001–2020. As the atmospheric variables are auto-correlated, the effective degree of freedom is estimated following Livezey and Chen (1983).

The continuous ranked probability score (Matheson and Winkler, 1976) is used to evaluate the accuracy of probabilistic forecasts for a given lead time t:

$$\mathrm{CRPS} = \frac{1}{N} \sum_{i=1}^{N} \int \left[F_{i,t}(y) - H\left(y - o_{i,t}\right) \right]^2 \mathrm{d}y, \tag{15}$$

where $F^{i,t}()$ is the cumulative distribution function of the probabilistic forecasts for case i at lead time t and $H()$ is the Heaviside step function defined as

$$H\left(y - o_{i,t}\right) = \begin{cases} 0 & y < o_{i,t}, \\ 1 & y \geq o_{i,t}, \end{cases} \tag{16}$$
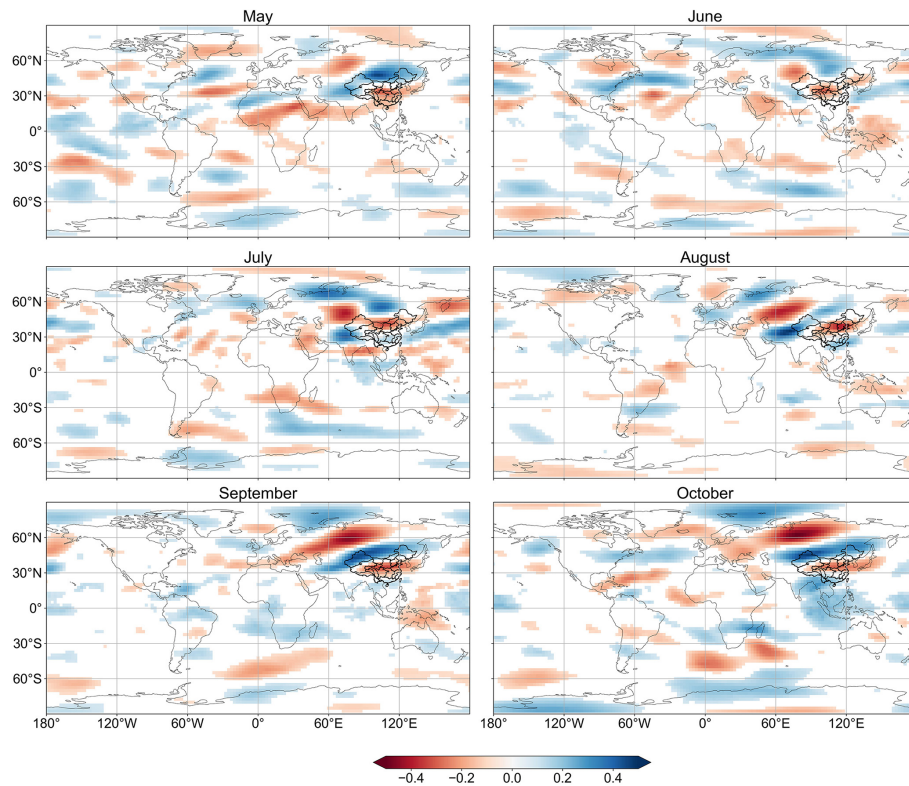
Figure 2. Correlation coefficient between pentad mean 10–60 d signals of U200 and precipitation over Region 1 (Inland rivers in Xinjiang) in different months. Correlation coefficients that are statistically significant at the 5 % level are shaded.

where $o_{i,t}$ is the corresponding observation.

A CRPS skill score is then calculated by comparing the CRPS of STP-CBaM forecasts to the CRPS of reference forecasts:

$$\text{CRPS}_{\text{SS}} = \frac{\text{CRPS}_{\text{REF}} - \text{CRPS}}{\text{CRPS}_{\text{REF}}} \times 100\,\%. \qquad (17)$$

The reference forecasts are generated using the BJP approach to fit the observations used in the training dataset. When the CRPS skill score is 100 %, the probabilistic forecasts are the same as the observations, whereas a skill score of 0 % indicates that the probabilistic forecasts show similar accuracies compared to the cross-validated climatology. A negative skill score means that the probabilistic forecasts are inferior to the cross-validated climatology.

The forecast reliability is evaluated using the α index (Renard et al., 2010). The probability integral transform (PIT) values of probabilistic forecasts for each case i at lead time t are given as

$$\pi_{i,t} = F^{i,t}\left(o_{i,t}\right), \qquad (18)$$

where $F^{i,t}()$ is the cumulative distribution function of probabilistic forecasts and $o_{i,t}$ are the corresponding observations. If the ensemble forecasts are reliable, $\pi_{i,t}$ should be uniformly distributed. The π_t values are then summarized into an α index:

$$\alpha = 1.0 - \frac{2}{N} \sum_{i=1}^{N} \left| \pi_{i,t}^* - \frac{i}{N+1} \right|, \qquad (19)$$

where $\pi_{i,t}^*$ is the sorted $\pi_{i,t}$ in increasing order. The α index ranges from 0 to 1, and a higher α index indicates higher reliability.

4   Results

4.1   Correlation analysis between atmospheric intraseasonal oscillation and precipitation anomalies

Figure 2 presents the correlation between pentad mean 10–60 d signals of U200 and precipitation over Region 1 (Inland rivers in Xinjiang) from May to October. The U200 signals near the Mongolian Plateau have a positive impact on precipitation anomalies over Region 1 in May, while the impact of U200 signals near the eastern Tibetan Plateau is negative. In June and July, the U200 signals in the West Siberian Plain and the Mongolian Plateau show positive correlations with precipitation anomalies. The spatial patterns of correlations between U200 signals and precipitation anomalies are similar in August, September, and October. The U200
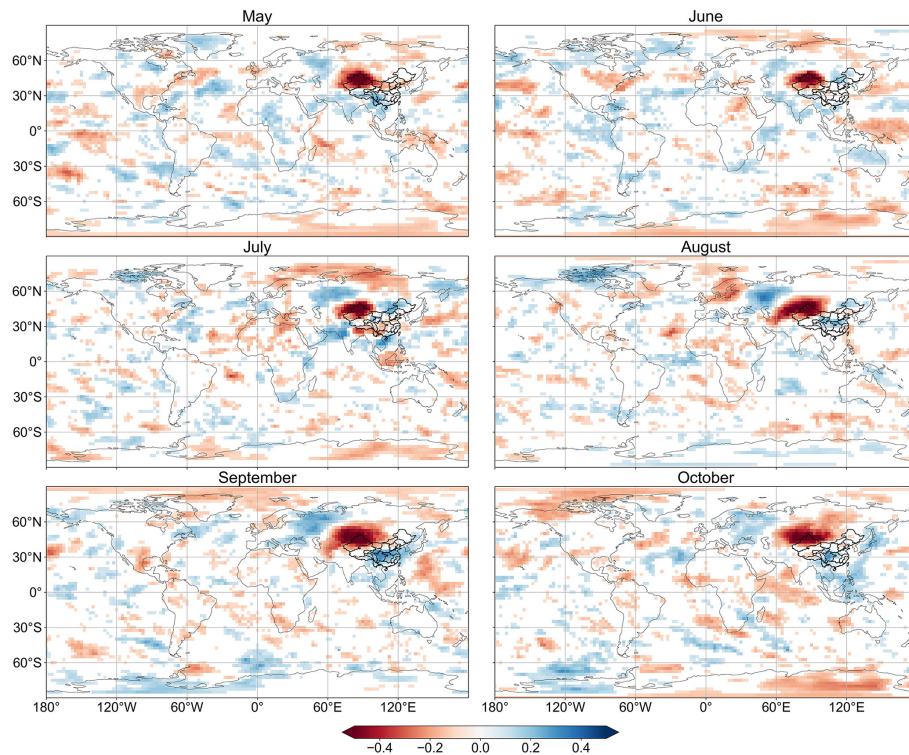
Figure 3. Same as Fig. 2 but for OLRA.

signals near the Barents Sea and the Iranian Plateau have positive impacts on precipitation anomalies over Region 1. In comparison, U200 signals over the West Siberian Plain show strong negative correlations with precipitation anomalies in these months. The OLRA signals show similar wave patterns to other atmospheric variables (Fig. 3). The spatial patterns of correlations between U850, U200, OLRA, H200, H500, H850, and precipitation anomalies are different for each month as well (Figs. S1–S4 in the Supplement).

4.2 Skill of the ECMWF model in forecasting atmospheric intraseasonal oscillations

The forecast skill of bridging models is reliant on the forecast skill of atmospheric intraseasonal oscillations derived from dynamical models. The TCC between the ensemble mean of ECMWF-forecasted U200 intraseasonal signals and the observations in May is shown in Fig. 4. The ECMWF model shows high forecast skill in predicting U200 intraseasonal signals when the lead time is within 10 d, and the correlation coefficients are mostly over 0.7 over the globe. Although the forecast skill decreases as the lead time increases, there are still regions where the forecasted U200 signals are significantly correlated with the observations. The forecast skill of the OLRA intraseasonal oscillations is lower than that of the U200 signals (Fig. 5). High forecast skill is mostly observed near the Equator from 30 to 30° E when the lead time is beyond 10 d. Similar results are also found for U850, H200,

H500, and H850, where significant correlations are found mostly near the Equator at longer lead times (Figs. S5 to S8). This suggests that the forecast skill of sub-seasonal precipitation can potentially be improved by taking advantage of both skilful prediction of atmospheric intraseasonal oscillations and stable relationships between precipitation and large-scale circulations, especially for tropical regions.

4.3 Skill of the calibration model, bridging models, and merged forecasts

Figure 6 presents the spatial distribution of the CRPS skill score of the calibration model, bridging models, and merged forecasts at different lead times in May. The calibration model shows the highest forecast skill compared to the bridging models at short lead times. The forecast skill of the calibrated forecasts decreases rapidly, and the CRPS skill scores are mostly below 10 % when the lead time is beyond 10 d. The forecast skill of the bridging models is higher than the calibration model in Region 10 (Huai River), Region 14 (Middle Yangtze River), and Region 17 (Southeast rivers) at a lead time of 15 d when OLRA is used as a predictor. The forecast skill of the bridging models is higher in Region 6 (Hai River) and Region 7 (Songhua River) when the OLRA and U200 signals are used separately as predictors at a lead time of 20 d. The merged forecasts take advantage of both the calibration model and the bridging models. Figure 7 shows the boxplots of the CRPS skill scores of the cali-
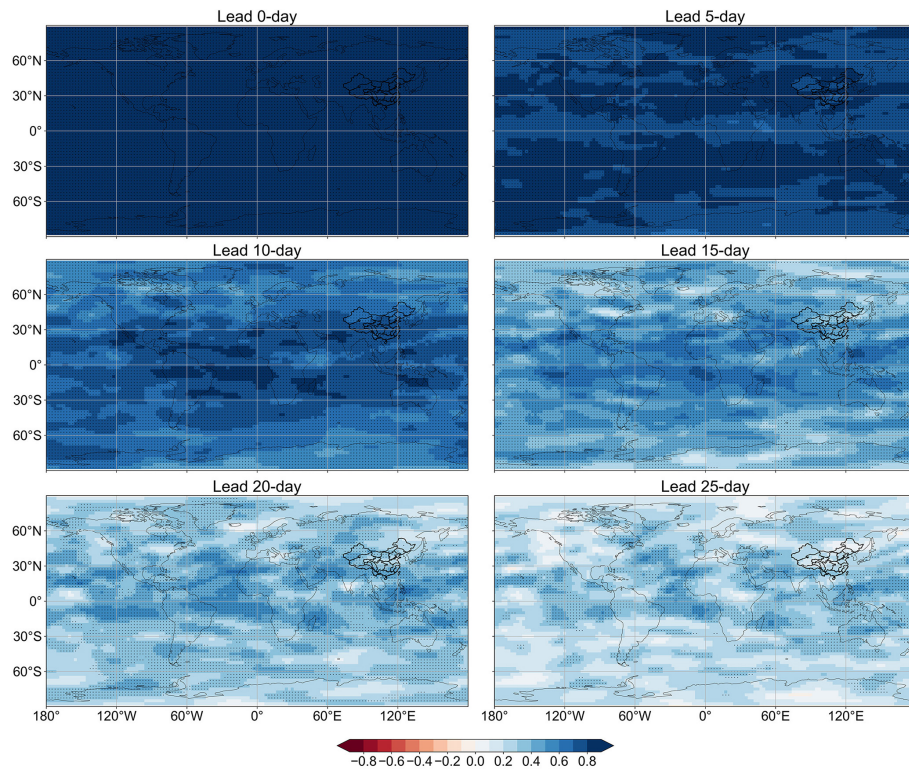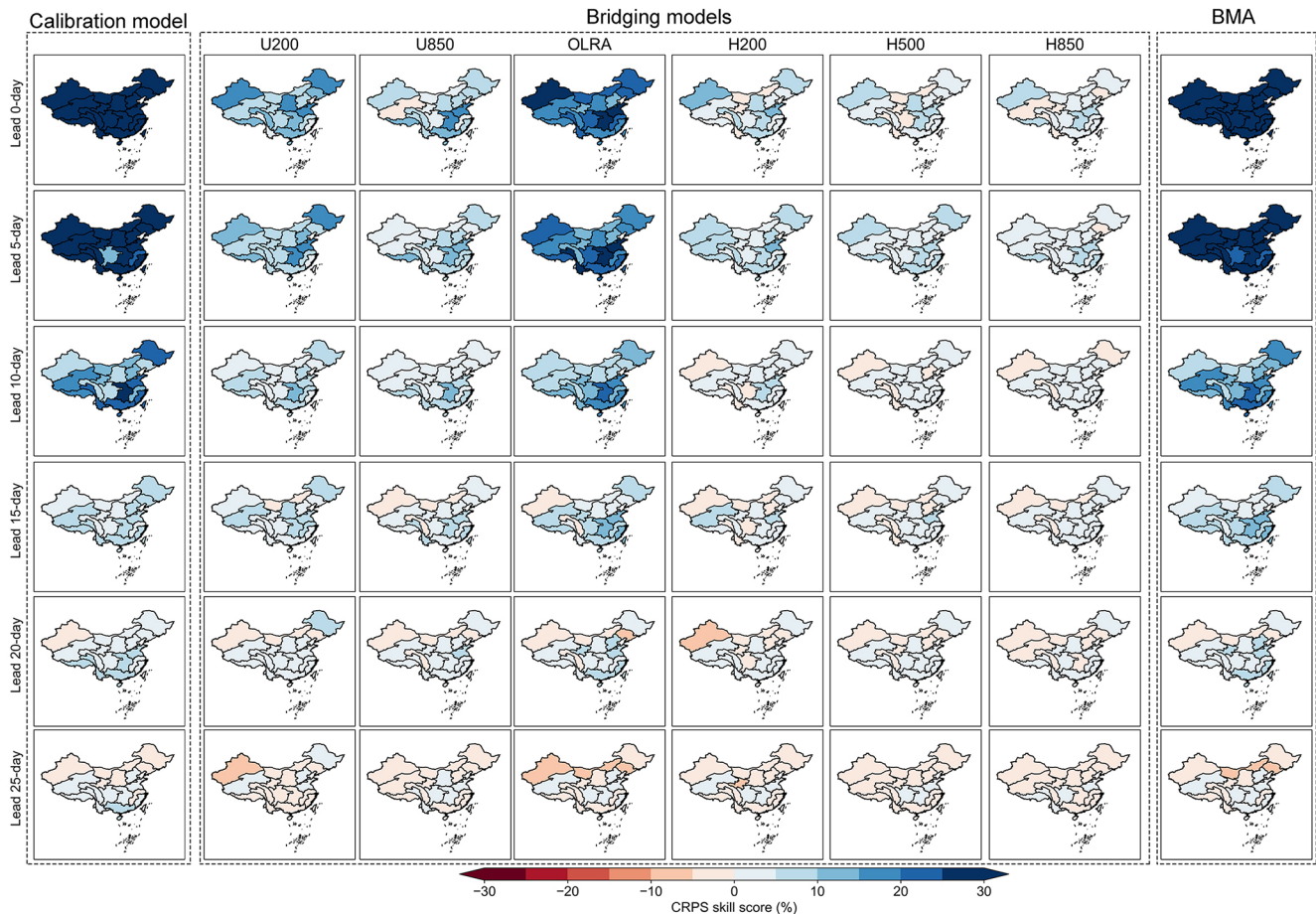
Figure 4. Temporal correlation coefficient (TCC) of the ensemble mean of U200 intraseasonal signals derived from the ECMWF model compared to the ERA5 reanalysis data in May. Correlation coefficients that are statistically significant at the 5 % level are shaded.
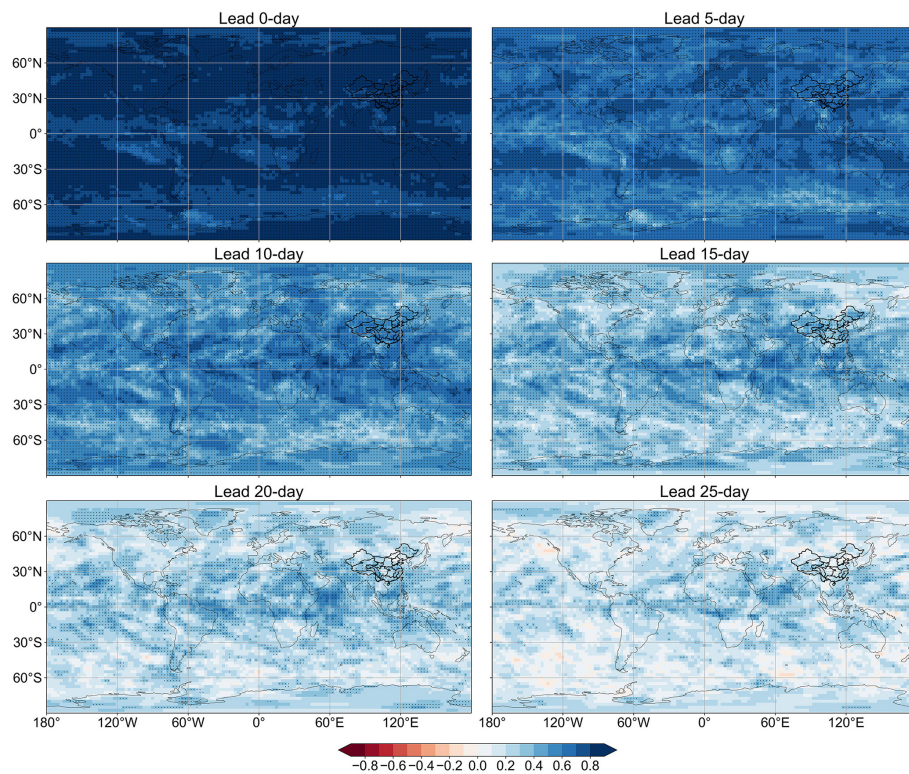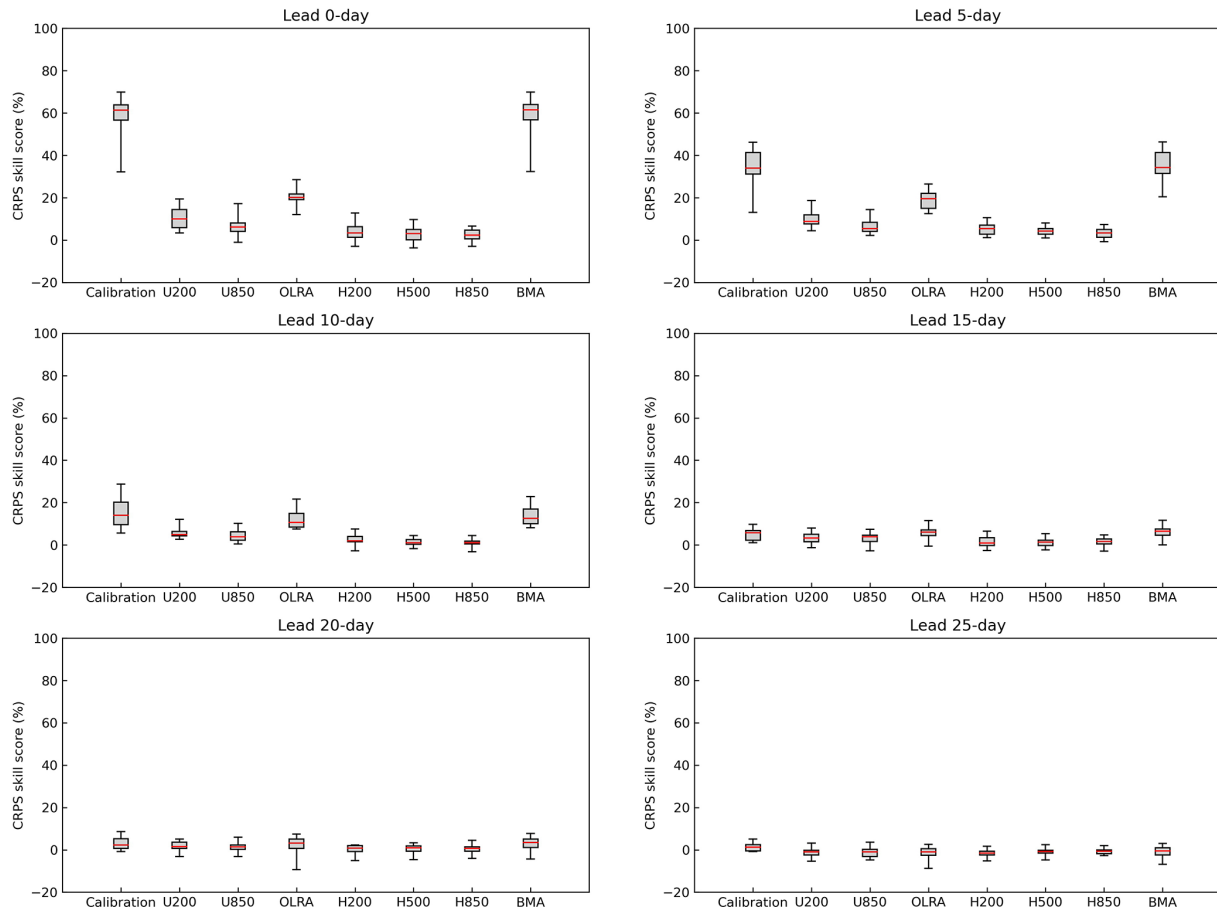

Figure 5. Same as Fig. 4 but for OLRA.

Figure 6. Spatial distribution of the CRPS skill score of the calibration model, bridging models (U200, U850, OLRA, H200, H500, H850), and merged forecasts (BMA) at different lead times in May. Publisher's remark: please note that the above figure contains disputed territories.

bration model, bridging models (U200, U850, OLRA, H200, H500, H850), and merged forecasts (BMA) at different lead times in May. The distribution of the CRPS skill score of the merged forecasts is similar to the calibration model at a lead time of 0 d. The minimum CRPS skill score of the merged forecasts is over 20 % in Region 13 (Yangtze River) at a lead time of 5 d, higher than both the calibration model and the bridging models. The bridging models, which use the U200 and OLRA as predictors, show a higher minimum CRPS skill score compared to the calibration model and other bridging models at a lead time of 10–15 d. The distributions of the CRPS skill score of the calibration model, bridging models, and BMA-merged forecasts are similar at longer lead times.

Figure 8 shows the distribution of model weights at different lead times for Region 1 (Inland rivers in Xinjiang) in May. The weights are rather stable at short lead times, when more than 90 % of the total weights are assigned to the calibration model. Similar results are also found in other regions and months (not shown). The weights of the calibration model decrease rapidly when the lead time is beyond 10 d. More weights are assigned to U200 and OLRA

at longer lead times. This is mostly consistent with the distribution of CRPS skill scores shown in Fig. 7. The CRPS skill scores of the U200- and OLRA-based bridging models are higher than the calibration model and other bridging models, especially when the lead time is between 10 and 20 d. This indicates that the U200 and OLRA signals are more useful in predicting sub-seasonal precipitation anomalies compared to other large-scale atmospheric circulation variables.

Figure 9 presents the CRPS skill score of merged forecasts at different lead times from May to October. In general, the forecast skill shows regional, monthly, and lead-time-dependent patterns. The merged forecasts show higher skill in predicting sub-seasonal precipitation anomalies over Region 16 (Pearl River) than other regions. The CRPS skill scores are positive for all months at lead times of 0–20 d. This is mainly due to the higher prediction skill of OLRA in these regions, as shown in Fig. 5. In addition, the merged forecasts show the highest skill in October, when positive skill scores are found over 14 hydroclimatic regions for all lead times except in Region 1 (Inland rivers in Xinjiang), Region 7 (Songhua River), and Region 8 (Liao River). In comparison,

Figure 7. Boxplots of the CRPS skill score of the calibration model, bridging models (U200, U850, OLRA, H200, H500, H850), and merged forecasts (BMA) at different lead times in May. The red lines are the 50th percentiles, the top and bottom of each box are the 75th and 25th percentiles, and the whiskers are the maximum and minimum skill scores.

positive skill scores are found only over three hydroclimatic regions for all lead times (Region 1, Inland rivers in Xinjiang, Region 13, Yangtze River, and Region 14, Middle Yangtze River) in July.

The α indices of merged forecasts at different lead times are shown in Fig. 10. The α index is around 0.7 at short lead times, suggesting that the forecasts have a relatively low reliability. The α index is over 0.85 when the lead time is beyond 10 d for all hydroclimatic regions and lead times. This indicates that the merged forecasts have a higher reliability at longer lead times. To figure out the relatively low reliability at short lead times, we analyse the merged forecasts over Region 3 (Inland rivers in Inner Mongolia) in May at a lead time of 0 d. The α index of the merged forecasts is around 0.6, suggesting that the merged forecasts have a low reliability. We also investigate the model weights of calibrated forecasts and bridging forecasts. The results suggest that the calibrated forecasts are more important than the bridging forecasts, while the cross-validated model weights are over 0.95. This suggests that the low reliability of merged forecasts is mostly caused by the low reliability of calibrated forecasts.

Figure S9 presents the quantile ranges of calibrated forecasts and merged forecasts against time. The quantile ranges of both the calibrated forecasts and merged forecasts are small, suggesting that the forecasts are too narrow (too confident).

5 Discussion

5.1 Discussion of forecast skill

Though the STP-CBaM model displays a good ability to generate skilful and reliable sub-seasonal precipitation forecasts over China, the forecast skill shows great diversity in different regions, months, and lead times. The calibration model shows the highest forecast skill compared to the bridging models for all regions and all months when the lead time is within 5 d (Figs. S10–S14). The U200- and OLRA-based bridging models outperform the calibration model and the other bridging models when the lead time is beyond 10 d in certain months and in certain regions. This may be explained by the strong relationship between U200, OLRA, and precip-
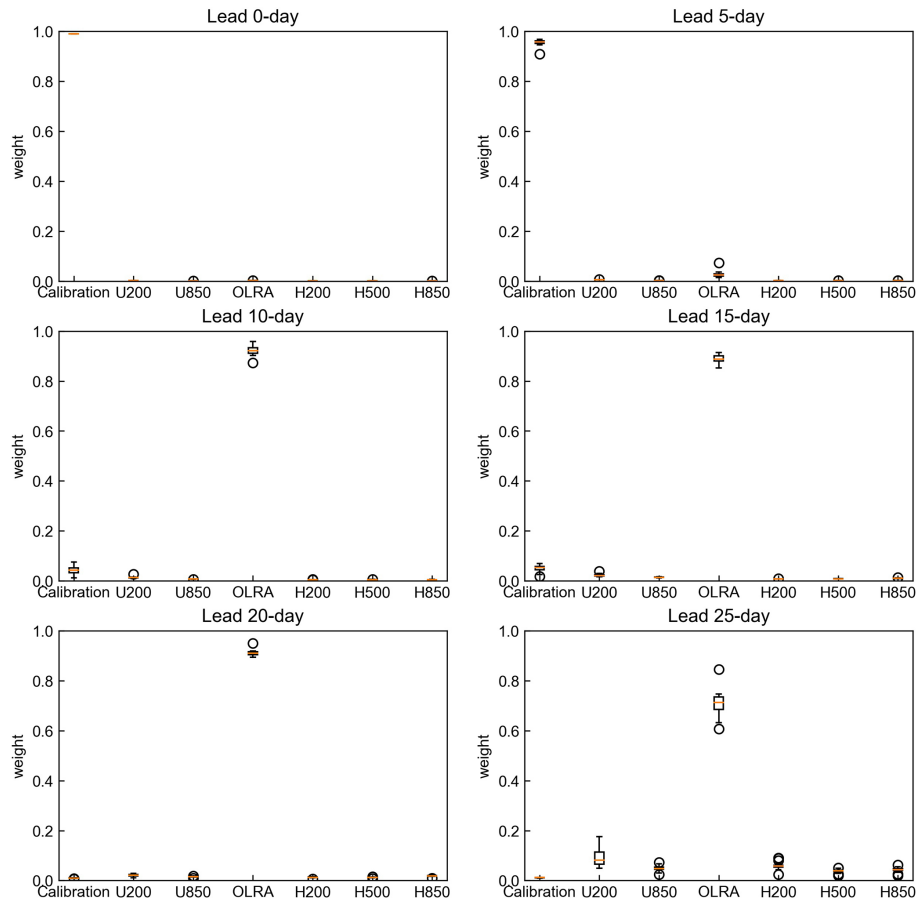
Figure 8. Boxplots showing the distribution of model weights at different lead times in cross-validation for Region 1 (Inland rivers in Xinjiang) in May.

itation anomalies and the forecast skill of U200 and OLRA in the ECMWF model in these regions (Figs. S5–S8).

However, we also note that there are several regions where the forecast skill of the calibration model is higher than the bridging models at longer lead times. This may be caused by the auto-correlations of the sub-seasonal precipitation anomalies defined in this study. In our data processing section, the observed anomalies before the forecast initial date are used to make enough data for the running mean. Thus, the predictand is not purely based on the ECMWF raw forecasts. The observational data are also introduced. The preceding observed precipitation anomalies may provide useful forecast information when the auto-correlations are high.

In addition, the limited forecast skill of large-scale circulations at mid to high latitudes in dynamical models may contribute to the limited forecast skill of bridging models as well. To figure out the potential skill of the STP-CBaM method in predicting sub-seasonal precipitation anomalies, we use the pth 10–60 d signal of the atmospheric field derived from the ERA5 reanalysis dataset as a predictor for the bridging model instead of the atmospheric field derived from the ECMWF model. Thus, the potential forecast skill is

based on ERA5 reanalysis data, while the practical forecast skill is based on the ECMWF model. Figure 11 compares the potential CRPS skill score (based on the ERA5 reanalysis) and the practical CRPS skill score (based on the ECMWF model) of merged forecasts over China. The potential CRPS skill scores are similar to the practical CRPS skill scores as the precipitation forecasts derived from the ECMWF model have a high accuracy at short lead times. The potential CRPS skill scores are much higher than the practical CRPS skill scores at longer lead times. This indicates that the forecast skill will be greatly improved when the atmospheric field is well predicted in the GCMs.

We also note that the weights do not always match the skill patterns. In this study, the posterior distributions of model weights are assigned by a model's predictive ability rather than fitting ability. Indeed, there is much literature in support of using predictive performance measures for model choice and combination based on the idea that a model is only as good as its predictions (Stock and Watson, 2006; Eklund and Karlsson, 2007). Thus, the CRPS skill score is not used when inferring model weights. This may lead to the discrepancy
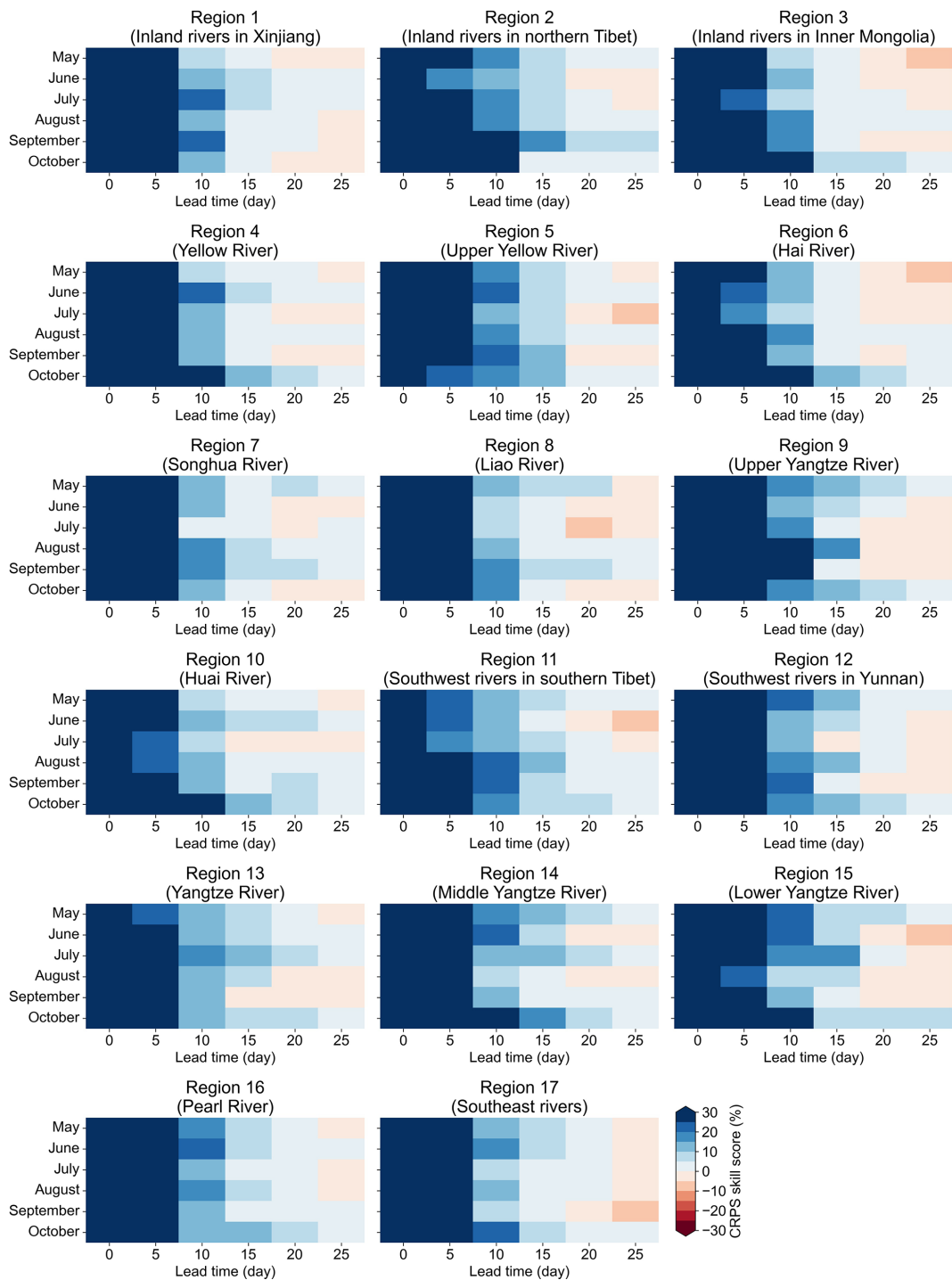
Figure 9. CRPS skill score of merged forecasts at different lead times from May to October.

between the model weights and the forecast skill score, especially when none of the models shows high predictive skill.

5.2 Limitations and future work

In this study, we aim at investigating the capability of dynamical models to improve the forecast skill of sub-seasonal precipitation anomalies using large-scale circulations as predictors. The bridging models are built based on the concurrent relationships between atmospheric intraseasonal oscillations and precipitation anomalies. Thus, the forecast skill of bridging models is highly reliant on the forecast skill of atmospheric intraseasonal oscillations derived from dynamical
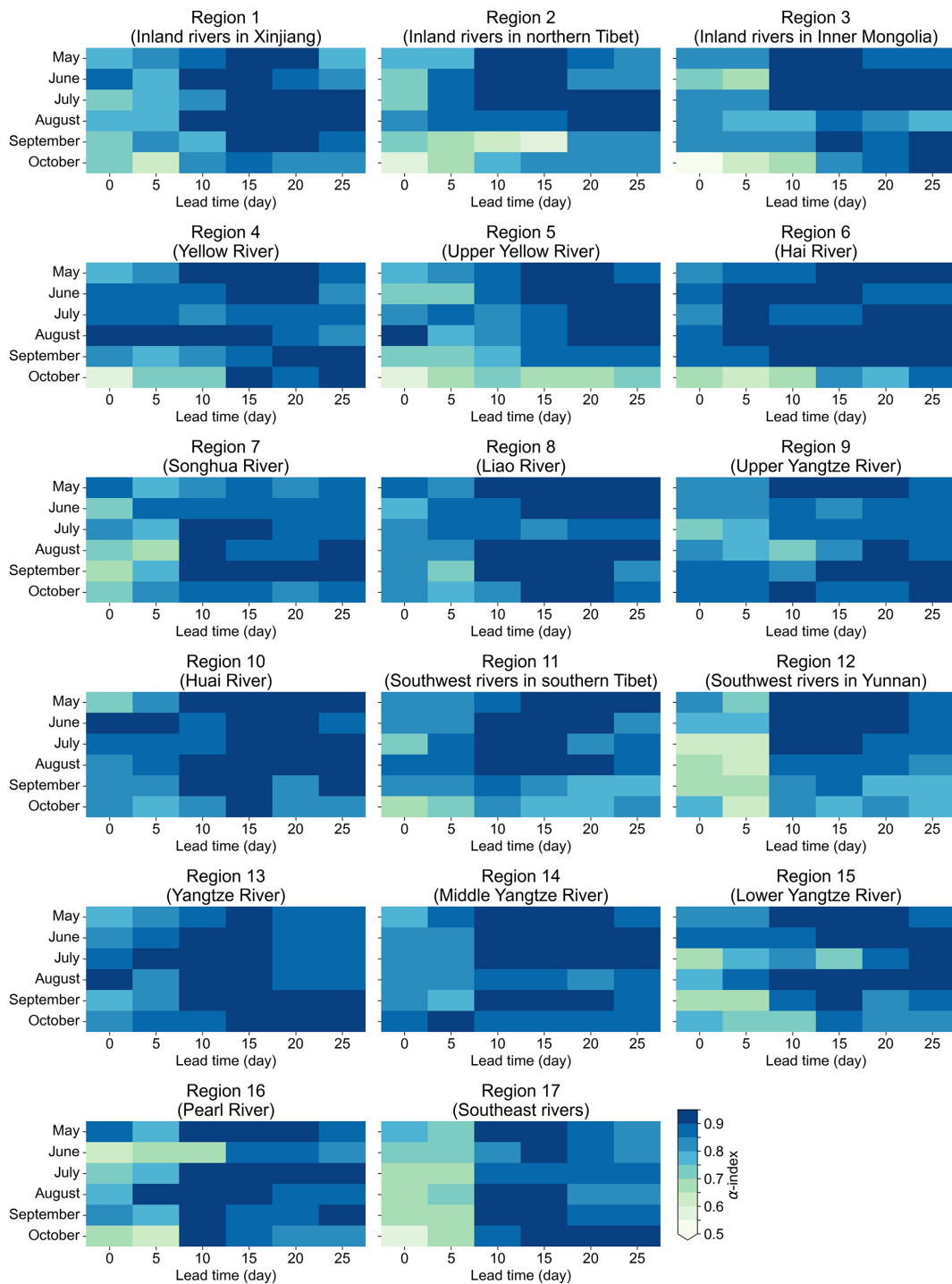
Figure 10. Same as Fig. 9 but for the α index.

models. In the future, the lagged relationships between atmospheric intraseasonal oscillations and precipitation anomalies will be considered to further improve sub-seasonal precipitation forecast skill. Although the forecast skill of the calibration model is high at short lead times, the results also suggest that the calibrated forecasts are too narrow (too confident).

We would like to focus on improving the forecast reliability, especially at short lead times, in the future.

Meanwhile, we define the predictors using the STPM for each month and each hydroclimatic region. Intraseasonal climate indices, such as the MJO index and the BSISO index, have not been considered yet. Recently, Zhu et al. (2023) pro-
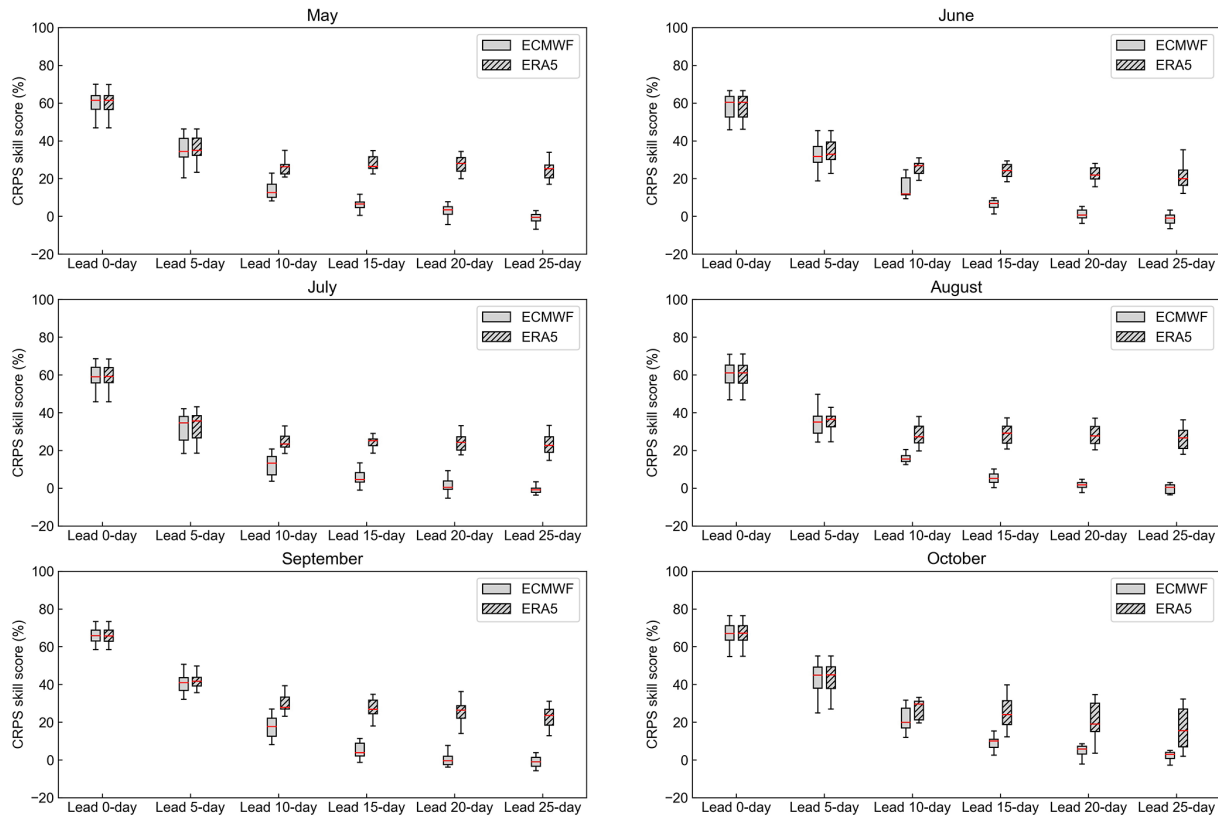
Figure 11. Practical CRPS skill score of merged forecasts based on the ECMWF model (solid) and potential CRPS skill score based on ERA5 reanalysis (hatched). The red lines are the 50th percentiles, the top and bottom of each box are the 75th and 25th percentiles, and the whiskers are the maximum and minimum skill scores.

posed two sets of novel indices based on the compound zonal displacements of the South Asia high (SAH) and the western Pacific sub-tropical high (WPH) to monitor and predict the intraseasonal variation of Meiyu. These climate indices will be introduced in the bridging models to investigate the potential improvement of the forecast skill.

In addition, we mainly focus on the prediction of intraseasonal (10–60 d) precipitation anomalies in this study. However, previous studies suggested that the intraseasonal component may only account for 7 % of the total variability in north-eastern China, while the seasonal component accounted for nearly 70 %. Thus, the relationships between seasonal precipitation anomalies and large-scale circulation patterns should also be investigated in these regions in the future.

6   Conclusions

In this study, we develop a STP-CBaM method to improve probabilistic sub-seasonal precipitation forecast skill over 17 hydroclimatic regions in China. The STP-CBaM method takes advantage of both dynamical models and statistical models. The calibration model is built by cali-

brating pentad mean precipitation anomalies derived from the ECMWF model. Bridging models are built by defining potential predictors using the spatial–temporal projection method (STPM). The calibration model and bridging models are merged through the Bayesian modelling averaging (BMA) method. Our results suggest that the forecast skill of the calibration model is higher compared to the bridging models when the lead time is within 5–10 d. The U200- and OLRA-based bridging models outperform the calibration model when the lead time is beyond 10 d in certain months and certain regions. The BMA-merged forecasts take advantage of both the calibration model and the bridging models. The BMA weights are rather stable at short lead times, where over 90 % of the total weights are assigned to the calibration model. More weights are assigned to the U200- and OLRA-based bridging models when the lead time is beyond 10 d. The results of the α index suggest that BMA-merged forecasts are reliable at longer lead times. Some improvements to reliability are still needed at shorter lead times.

gloh2o.org/mswep/ (GloH20, 2023). The OLR dataset can be sourced from https://psl.noaa.gov/thredds/catalog/Datasets/olrcdr/catalog.html (NOAA, 2023). The ECMWF hindcast data can be retrieved from the S2S database at http://apps.ecmwf.int/datasets/data/s2s/ (ECMWF, 2023).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/hess-27-4187-2023-supplement.

Author contributions. YL: conceptualization, methodology, writing (original draft preparation, review and editing). KX: methodology, writing (review and editing). ZW: conceptualization, funding acquisition, supervision, writing (review and editing). ZZ: conceptualization, review and editing. QJW: conceptualization, supervision, writing (review and editing).

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We would like to thank the two anonymous reviewers for their reviews of early versions of this paper.

Review statement. This paper was edited by Louise Slater and reviewed by two anonymous referees.

References

Chen, Y. and Zhai, P.: Simultaneous modulations of precipitation and temperature extremes in Southern parts of China by the boreal summer intraseasonal oscillation, Clim. Dynam., 49, 3363–3381, 2017.

Copernicus Climate Change Service: Fifth generation of ECMWF atmospheric reanalysis of the global climate, https://cds.climate.copernicus.eu/ (last access: 10 December 2022), 2022.

Cui, J., Yang, S., and Li, T.: How well do the S2S models predict intraseasonal wintertime surface air temperature over mid-high-latitude Eurasia?, Clim. Dynam., 57, 503–521, 2021.

de Andrade, F. M., Coelho, C. A., and Cavalcanti, I. F.: Global precipitation hindcast quality assessment of the Subseasonal to Seasonal (S2S) prediction project models, Clim. Dynam., 52, 5451–5475, 2019.

Duan, Q., Sorooshian, S., and Gupta, V. K.: Optimal use of the SCE-UA global optimization method for calibrating watershed models, J. Hydrol., 158, 265–284, https://doi.org/10.1016/0022-1694(94)90057-4, 1994.

ECMWF: S2S, ECMWF, Realtime, Daily averaged, http://apps.ecmwf.int/datasets/data/s2s/ (last access: 20 November 2023), 2023.

Eklund, J. and Karlsson, S.: Forecast Combination and Model Averaging Using Predictive Measures, Econ. Rev., 26, 329–363, https://doi.org/10.1080/07474930701220550, 2007.

GloH20: MSWEP – Multi-Source Weighted-Ensemble Precipitation, http://www.gloh2o.org/mswep/ (last access: 20 November 2023), 2023.

Guo, B., Xu, T., Yang, Q., Zhang, J., Dai, Z., Deng, Y., and Zou, J.: Multiple Spatial and Temporal Scales Evaluation of Eight Satellite Precipitation Products in a Mountainous Catchment of South China, Remote Sens., 15, 1373, https://doi.org/10.3390/rs15051373, 2023.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz©-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Q. J. Roy. Meteorol. Soc., 146, 1999–2049, 2020.

Hsu, P.-C., Li, T., You, L., Gao, J., and Ren, H.-L.: A spatial–temporal projection model for 10–30 day rainfall forecast in South China, Clim. Dynam., 44, 1227–1244, https://doi.org/10.1007/s00382-014-2215-4, 2015.

Hsu, P. C., Lee, J. Y., and Ha, K. J.: Influence of boreal summer intraseasonal oscillation on rainfall extremes in southern China, Int. J. Climatol., 36, 1403–1412, https://doi.org/10.1002/joc.4433, 2016.

Jie, W. H., Vitart, F., Wu, T. W., and Liu, X. W.: Simulations of the Asian summer monsoon in the sub-seasonal to seasonal prediction project (S2S) database, Q. J. Roy. Meteorol. Soc., 143, 2282–2295, https://doi.org/10.1002/qj.3085, 2017.

Lee, S.-S., Wang, B., Waliser, D. E., Neena, J. M., and Lee, J.-Y.: Predictability and prediction skill of the boreal summer intraseasonal oscillation in the Intraseasonal Variability Hindcast Experiment, Clim. Dynam., 45, 2123–2135, https://doi.org/10.1007/s00382-014-2461-5, 2015.

Li, W., Yang, X.-Q., Fang, J., Tao, L., and Sun, X.: Asymmetric boreal summer intraseasonal oscillation events over the western North Pacific and their impacts on East Asian precipitation, J. Climate, 36, 2645–2661, 2023.

Li, Y., Wu, Z., He, H., Wang, Q. J., Xu, H., and Lu, G.: Post-processing sub-seasonal precipitation forecasts at various spatiotemporal scales across China during boreal summer monsoon, J. Hydrol., 598, 125742, https://doi.org/10.1016/j.jhydrol.2020.125742, 2021.

Li, Y., Wu, Z., He, H., and Yin, H.: Probabilistic subseasonal precipitation forecasts using preceding atmospheric intraseasonal signals in a Bayesian perspective, Hydrol. Earth Syst. Sci., 26, 4975–4994, https://doi.org/10.5194/hess-26-4975-2022, 2022.

Li, Y., Pang, B., Zheng, Z., Chen, H., Peng, D., Zhu, Z., and Zuo, D.: Evaluation of Four Satellite Precipitation Products over Mainland China Using Spatial Correlation Analysis, Remote Sens., 15, 1823, https://doi.org/10.3390/rs15071823, 2023.

Liu, B., Zhu, C., Ma, S., Yan, Y., and Jiang, N.: Sub-seasonal processes of triple extreme heatwaves over the Yangtze River Valley in 2022, Weather Extrem., 40, 100572, https://doi.org/10.1016/j.wace.2023.100572, 2023.

Liu, F., Ouyang, Y., Wang, B., Yang, J., Ling, J., and Hsu, P.-C.: Seasonal evolution of the intraseasonal variability of China summer precipitation, Clim. Dynam., 54, 4641–4655, 2020.

Liu, J. and Lu, R.: Different Impacts of Intraseasonal Oscillations on Precipitation in Southeast China between Early and Late Summers, Adv. Atmos. Sci., 39, 1885–1896, 2022.

Liu, J., Shangguan, D., Liu, S., Ding, Y., Wang, S., and Wang, X.: Evaluation and comparison of CHIRPS and MSWEP daily-precipitation products in the Qinghai-Tibet Plateau during the period of 1981–2015, Atmos. Res., 230, 104634, https://doi.org/10.1016/j.atmosres.2019.104634, 2019.

Livezey, R. E. and Chen, W. Y.: Statistical Field Significance and its Determination by Monte Carlo Techniques, Mon. Weather Rev., 111, 46–59, https://doi.org/10.1175/1520-0493(1983)111<0046:SFSAID>2.0.CO;2, 1983.

Matheson, J. E. and Winkler, R. L.: Scoring rules for continuous probability distributions, Manage. Sci., 22, 1087–1096, https://doi.org/10.1287/mnsc.22.10.1087, 1976.

NOAA: NOAA's Outgoing Longwave Radiation – Daily Climate Data Record (OLR – Daily CDR): PSL Interpolated Version data provided by the NOAA PSL, Boulder, Colorado [data set], https://psl.noaa.gov/thredds/catalog/Datasets/olrcdr/catalog.html (last access: 20 November 2023), 2023.

Peng, Z., Wang, Q., Bennett, J. C., Schepen, A., Pappenberger, F., Pokhrel, P., and Wang, Z.: Statistical calibration and bridging of ECMWF System4 outputs for forecasting seasonal precipitation over China, J. Geophys. Res.-Atmos., 119, 7116–7135, 2014.

Ren, P., Ren, H. L., Fu, J. X., Wu, J., and Du, L.: Impact of boreal summer intraseasonal oscillation on rainfall extremes in southeastern China and its predictability in CFSv2, J. Geophys. Res.-Atmos., 123, 4423–4442, 2018.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, Water Resour. Res., 46, W05521, https://doi.org/10.1029/2009WR008328, 2010.

Schepen, A., Wang, Q. J., and Robertson, D. E.: Seasonal Forecasts of Australian Rainfall through Calibration and Bridging of Coupled GCM Outputs, Mon. Weather Rev., 142, 1758–1770, https://doi.org/10.1175/MWR-D-13-00248.1, 2014.

Schepen, A., Wang, Q., and Everingham, Y.: Calibration, bridging, and merging to improve GCM seasonal temperature forecasts in Australia, Mon. Weather Rev., 144, 2421–2441, 2016.

Shibuya, R., Nakano, M., Kodama, C., Nasuno, T., Kikuchi, K., Satoh, M., Miura, H., and Miyakawa, T.: Prediction Skill of the Boreal Summer Intra-Seasonal Oscillation in Global Non-hydrostatic Atmospheric Model Simulations with Explicit Cloud Microphysics, J. Meteorol. Soc. Jpn. Ser. II, 99, 973–992, https://doi.org/10.2151/jmsj.2021-046, 2021.

Slater, L. J., Arnal, L., Boucher, M. A., Chang, A. Y. Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R. L., Wood, A., and Zappa, M.: Hybrid forecasting: blending climate predictions with AI models, Hydrol. Earth Syst. Sci., 27, 1865–1889, https://doi.org/10.5194/hess-27-1865-2023, 2023.

Specq, D. and Batté, L.: Improving subseasonal precipitation forecasts through a statistical–dynamical approach: application to the southwest tropical Pacific, Clim. Dynam., 55, 1913–1927, https://doi.org/10.1007/s00382-020-05355-7, 2020.

Stock, J. H. and Watson, M. W.: Chapter 10 Forecasting with Many Predictors, in: Handbook of Economic Forecasting, edited by: Elliott, G., Granger, C. W. J., and Timmermann, A., Elsevier, 515–554, https://doi.org/10.1016/S1574-0706(05)01010-4, 2006.

Strazzo, S., Collins, D. C., Schepen, A., Wang, Q., Becker, E., and Jia, L.: Application of a hybrid statistical–dynamical system to seasonal prediction of North American temperature and precipitation, Mon. Weather Rev., 147, 607–625, 2019a.

Strazzo, S., Collins, D. C., Schepen, A., Wang, Q. J., Becker, E., and Jia, L.: Application of a Hybrid Statistical-Dynamical System to Seasonal Prediction of North American Temperature and Precipitation, Mon. Weather Rev., 147, 607–625, https://doi.org/10.1175/MWR-D-18-0156.1, 2019b.

Vitart, F. and Robertson, A. W.: The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events, npj Clim. Atmos. Sci., 1, 1–7, 2018.

Wang, Q., Schepen, A., and Robertson, D. E.: Merging seasonal rainfall forecasts from multiple statistical models through Bayesian model averaging, J. Climate, 25, 5524–5537, 2012.

Wang, Q., Shao, Y., Song, Y., Schepen, A., Robertson, D. E., Ryu, D., and Pappenberger, F.: An evaluation of ECMWF SEAS5 seasonal climate forecasts for Australia using a new forecast calibration algorithm, Environ. Model. Softw., 122, 104550, https://doi.org/10.1016/j.envsoft.2019.104550, 2019.

White, C. J., Domeisen, D. I., Acharya, N., Adefisan, E. A., Anderson, M. L., Aura, S., Balogun, A. A., Bertram, D., Bluhm, S., and Brayshaw, D. J.: Advances in the application and utility of subseasonal-to-seasonal predictions, B. Am. Meteorol. Soc., 103, E1448–E1472, 2022.

Wu, J., Ren, H.-L., Zhang, P., Wang, Y., Liu, Y., Zhao, C., and Li, Q.: The dynamical-statistical subseasonal prediction of precipitation over China based on the BCC new-generation coupled model, Clim. Dynam., 59, 1213–1232, 2022.

Wu, J., Li, J., Zhu, Z., and Hsu, P.-C.: Factors determining the subseasonal prediction skill of summer extreme rainfall over southern China, Clim. Dynam., 60, 443–460, https://doi.org/10.1007/s00382-022-06326-w, 2023.

Xie, J., Hsu, P.-C., Hu, Y., Ye, M., and Yu, J.: Skillful Extended-Range Forecast of Rainfall and Extreme Events in East China Based on Deep Learning, Weather Forecast., 38, 467–486, https://doi.org/10.1175/WAF-D-22-0132.1, 2023.

Yan, Y., Zhu, C., and Liu, B.: Subseasonal predictability of the July 2021 extreme rainfall event over Henan China in S2S operational models, J. Geophys. Res.-Atmos., 128, e2022JD037879, https://doi.org/10.1029/2022JD037879, 2023.

Zhang, K., Li, J., Zhu, Z., and Li, T.: Implications from Subseasonal Prediction Skills of the Prolonged Heavy Snow Event over Southern China in Early 2008, Adv. Atmos. Sci., 38, 1873–1888, https://doi.org/10.1007/s00376-021-0402-x, 2021.

Zhu, C., Liu, B., Li, L., Ma, S., Jiang, N., and Yan, Y.: Progress and Prospects of Research on Subseasonal to Seasonal Variability and Prediction of the East Asian Monsoon, J. Meteorol. Res., 36, 677–690, 2022.

Zhu, Z. and Li, T.: The statistical extended-range (10–30-day) forecast of summer rainfall anomalies over the entire China, Clim. Dynam., 48, 209–224, https://doi.org/10.1007/s00382-016-3070-2, 2017.

Zhu, Z., Li, T., Hsu, P.-C., and He, J.: A spatial-temporal projection model for extended-range forecast in the tropics, Clim. Dynam., 45, 1085–1098, https://doi.org/10.1007/s00382-014-2353-8, 2015.

Zhu, Z., Zhou, Y., Jiang, W., Fu, S., and Hsu, P.: Influence of compound zonal displacements of the South Asia high and the western Pacific subtropical high on Meiyu intraseasonal variation, Clim. Dynam., 61, 3309–3325, https://doi.org/10.1007/s00382-023-06726-6, 2023.