# Probabilistic sub-seasonal precipitation forecasts using preceding atmospheric intraseasonal signals in a Bayesian perspective

Yuan LI[1], Zhiyong WU[1], Hai HE[1], Hao YIN[1]

[1]College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

*Correspondence to*: Zhiyong WU (wzyhhu@gmail.com)

**Abstract.** Accurate and reliable sub-seasonal precipitation forecasts are of great socioeconomic value for various aspects. The atmospheric intraseasonal oscillation (ISO), which is one of the leading sources of sub-seasonal predictability, can be potentially used as predictor for sub-seasonal precipitation forecasts. However, the relationship between atmospheric intraseasonal signals and sub-seasonal precipitation is of high uncertainty. In this study, we develop a spatial-temporal projection based Bayesian hierarchical model (STP-BHM) for sub-seasonal precipitation forecasts. The coupled co-variance patterns between preceding atmospheric intraseasonal signals and precipitation are extracted, and the corresponding projection coefficients are defined as predictors. A Bayesian hierarchical model (BHM) is then built to address the uncertainty in the relationship between atmospheric intraseasonal signals and precipitation. The STP-BHM model is applied to predict both pentad mean precipitation amount and pentad mean precipitation anomalies for each hydroclimatic region over China during the boreal summer monsoon season. The model performance is evaluated through a leave one-year-out cross-validation strategy. Our results suggest that the STP-BHM model can provide skillful and reliable probabilistic forecasts for both pentad mean precipitation amount and pentad mean precipitation anomalies at leads of 20-25 days over most hydroclimatic regions in China. The results also indicate that the STP-BHM model outperforms the National Centers for Environmental Prediction (NCEP) sub-seasonal to seasonal (S2S) model when the lead time is beyond 5 days for pentad mean precipitation amount forecasts. The Intraseasonal signals of 850-hPa and 200-hPa zonal wind (U850 and U200), 850-hPa and 500-hPa geopotential height (H850 and H500) contribute more to the overall forecast skill of pentad mean precipitation amount predictions. In comparison, the outgoing longwave radiation anomalies (OLRA) contribute most to the forecast skill of pentad mean precipitation anomalies predictions. Other sources of sub-seasonal predictability, such as soil moisture, snow cover, and stratosphere-troposphere interaction, will be included in the future to further improve sub-seasonal precipitation forecast skill.

## 1. Introduction

Accurate and reliable sub-seasonal precipitation forecasts can provide vital information for many management decisions in water resources, agriculture, and disaster mitigation (Vitart et al., 2012; Vitart and Robertson, 2018). One approach for sub-seasonal precipitation forecasts is to run dynamical models such as Global Climate Models (GCMs). Projects such as the Subseasonal-to-Seasonal Prediction Project (S2S) and the Subseasonal Experiment (SubX) have been launched to provide sub-seasonal precipitation forecasts with lead time up to 60 days from GCMs (Pegion et al., 2019; Vitart et al., 2017). However, the sub-seasonal precipitation forecasts derived directly from GCMs are of low accuracy as the physical equations are always simplified and small-scale processes cannot be well represented in the GCMs (De Andrade et al., 2019). Post-processing is always required to improve the accuracy and reliability of GCM forecasts before it can be used for other applications. Schepen et al. (2018) and our previous study (Li et al., 2020) used the Bayesian Joint Probability (BJP) method to post-process sub-seasonal precipitation forecasts over different regions, and the results suggested that the forecast skill and reliability were improved compared to raw GCM forecasts. Vigaud et al. (2020) proposed a new spatial correction method to improve sub-monthly precipitation forecasts derived from multimodel ensembles. Nevertheless, the results also indicated that the accuracy of post-processed sub-seasonal precipitation forecasts were still limited when the lead time was beyond 10-14 days.

An alternative approach for sub-seasonal precipitation forecasts is to establish statistical models based on the relationship between precipitation and preceding atmospheric or oceanic indices. Although dynamical models are more performant for short- to medium-term forecasts, statistical models are still found to be useful especially for long-term forecasts (Tuel and Eltahir, 2018; Abbot and Marohasy, 2014; Mekanik et al., 2013; Lü et al., 2011; Kirono et al., 2010). Schepen et al. (2012) suggested that the lagged climate indices were potentially useful for seasonal precipitation forecasts over Australia. Plenty of statistical algorithms, such as multiple linear regression or canonical correlation analysis, have been developed for seasonal precipitation forecasts based on the assumption that the seasonal anomalies are caused by slow-varying sea surface temperature, sea ice, snow cover, and other boundary conditions (Hwang et al., 2001; Barnston and Smith, 1996; Eden et al., 2015). A new cluster-based empirical method was proposed to predict winter precipitation anomalies over the European and Mediterranean Regions (Totz et al., 2017). This method used the sea surface temperature, geopotential height, sea level pressure, snow cover extent, and sea ice concentration as predictors. A random forest based statistical model, whose the predictors were identified from the gridded sea surface temperature, was developed to predict central and south Asia seasonal precipitation (Gerlitz et al.,

2016).

However, much fewer statistical models have been built and applied for sub-seasonal precipitation forecasts as the sources of sub-seasonal predictability are not yet fully understood. Compared to seasonal precipitation forecasts, the slow-varying boundary forcings may have limited impact on sub-seasonal precipitation as the time scale is too short. The atmospheric intraseasonal oscillation (ISO), which is the dominant mode of sub-seasonal variability, is one of the leading sources of sub-seasonal predictability (Robertson and Vitart, 2018). The boreal summer intraseasonal oscillation (BSISO) in the tropics, which is also known as Madden-Julian Oscillation (MJO) in winter, is characterized as a slow-moving system with a period of 30-90 days in the tropical atmosphere (Madden and Julian, 1971, 1972; Zhang, 2005; Woolnough, 2019; Wang and Xie, 1997). The circulation anomalies associated with the intraseasonal oscillation (ISO) are identified to have an impact on monsoon activities and heavy rainfall events (Annamalai and Slingo, 2001; Chen et al., 2004). Zhang et al. (2009) found that the rainfall patterns in Southeast China were transited from being enhanced to being suppressed when the MJO center moved from the Indian Ocean to the Western Pacific Ocean. Jia et al. (2011) suggested that the MJO influenced the rainfall patterns in China mainly by modulating the circulation in the subtropics and mid-high latitudes in winter. This suggests that the ISO signals could be potentially used for predicting sub-seasonal precipitation not only in tropical regions but in extra-tropical regions as well.

Several statistical models have been built to predict sub-seasonal precipitation based on the relationship between atmospheric intraseasonal signals and precipitation. The spatial-temporal projection (STP) model, which extracts the coupled patterns of predictors and predictand, has been developed in recent years (Hsu et al., 2020; Zhu and Li, 2017c, a, b, 2018). Hsu et al. (2015) established a set of spatial-temporal projection models (STPMs) to predict sub-seasonal precipitation at a lead time of 10-30 days over southern China. Their results suggested that the forecast skill was still promising at a 20-25-day lead time. Zhu and Li (2017c) predicted sub-seasonal precipitation by constructing STPMs over entire China, and independent forecasts of rainfall anomalies during the period of Olympic Games in 2008 and Shanghai World Expo in 2010 suggested that the STPMs were able to reproduce intraseasonal rainfall patterns at a 20-day lead time. However, we should note that the relationship between ISO signals and precipitation is highly uncertain and depend on the region and lead time. In previous studies, an optimal ensemble (OE) strategy was applied to generate probabilistic forecasts by picking up best predictors (Zhu and Li, 2017c; Zhu et al., 2015). Nevertheless, the number of best predictors was always limited. Further statistical assumptions were required

to interpret limited ensembles as probabilistic forecasts. The uncertainty in relationship between preceding ISO signals of atmospheric field and precipitation has not been fully considered yet.

There are several ways to address the above challenge. Lepore et al. (2017) established an extended logistic regression model to link the relationship between El Niño-Southern Oscillation (ENSO) and convective storm activity. Sohrabi et al. (2021) coupled the large-scale climate indices with a stochastic weather generator to provide ensemble streamflow forecasts. Compared to the above-mentioned traditional probabilistic model solutions, the Bayesian statistical models are more flexible and more efficient for assessing multiple sources of uncertainties. Wang et al. (2009) proposed a multivariate normal distribution based Bayesian joint probability (BJP) approach to predict seasonal streamflow over Australia using antecedent streamflow, ENSO indices, and other climate indicators as predictors. Peng et al. (2014) utilized the same BJP approach to predict seasonal precipitation over China using lagged oceanic-atmospheric indices. The Bayesian hierarchical model (BHM) has also been developed in recent years (Gelman and Hill, 2006). The BHMs are always constructed with several model layers. The predictand is assumed to follow distribution with unknown parameters in the first layer, and the parameters are linked with the predictors using linear regression models in the second layer. The regression coefficients are given hyperprior distributions with the BHMs. The utility of BHMs has been demonstrated in modelling spatiotemporal variability of hydrological variables in many studies (Renard, 2011; Reza Najafi and Moradkhani, 2013; Bracken et al., 2016; Lima and Lall, 2010; Lima and Lall, 2009; Devineni et al., 2013). The BHMs are also used for seasonal predictions in many fields. Chen et al. (2014) used the BHM to predict summer rainfall and streamflow over the Huai River basin, while Chu and Zhao (2007) developed a BHM model to predict seasonal tropical cyclone activity over the Central North Pacific. However, the BHMs have not been used to predict sub-seasonal precipitation before. In this study, we follow a similar BHM structure proposed by Devineni et al. (2013) to predict sub-seasonal precipitation.

China is located in east Asia, and is frequently subject to rainstorm and flood disasters during the boreal summer monsoon season. Accurate and reliable sub-seasonal precipitation forecasts can provide valuable information for mitigating the risks from rainstorm and flood disasters. However, the origin of intraseasonal precipitation variability is of high complexity owing to the mixed impact of tropical convection, forcing of Tibetan Plateau, and mid-high latitude systems (Zhu and Li, 2017c). In this study, we develop a spatial-temporal projection based Bayesian hierarchical model (STP-BHM) to predict both pentad mean precipitation amount and pentad mean precipitation anomalies over each hydroclimatic region in China during the boreal summer

monsoon season. The performance of the STP-BHM model is evaluated through a leave one-year-out cross-validation strategy.

In the following section, the datasets, main model components, including intraseasonal signal extraction, predictor definition, and Bayesian hierarchical model construction, and verification methods are introduced. The forecast skill of both pentad mean precipitation amount and pentad mean precipitation anomalies is presented in Sect. 3. Section 4 discusses the forecast skill, possible mechanism, limitations, and future work. Key findings are summarized in Sect. 5.

## 2. Data and Methodology

### 2.1 Data

In this study, China is divided into 17 hydroclimatic regions as suggested by Lang et al. (2014). The division is based on both watershed division standard and climate classifications. This will ensure that the climatic characteristics are nearly uniform in each region. The southeastern hydroclimatic regions are mostly of temperate and warm/hot summer climate without dry season (Cfb/Cfa), while the northwestern regions are mostly arid with limited precipitation (Bwk, Bsh, Bsk climate types) (Peel et al., 2007) (Figure 1). The observed precipitation is derived from the Multi-Source Weighted-Ensemble Precipitation, version 2 (MSWEP V2) dataset. The MSWEP V2 dataset is of high spatial (0.1°) and temporal (3 hourly) resolution. Compared to other gridded datasets, the MSWEP V2 exhibits more realistic spatial patterns, and higher accuracy over land (Wu et al., 2018; Beck et al., 2019). The 0.1° gridded precipitation data is area-weighted averaged through 17 hydroclimatic regions over China from May to October.

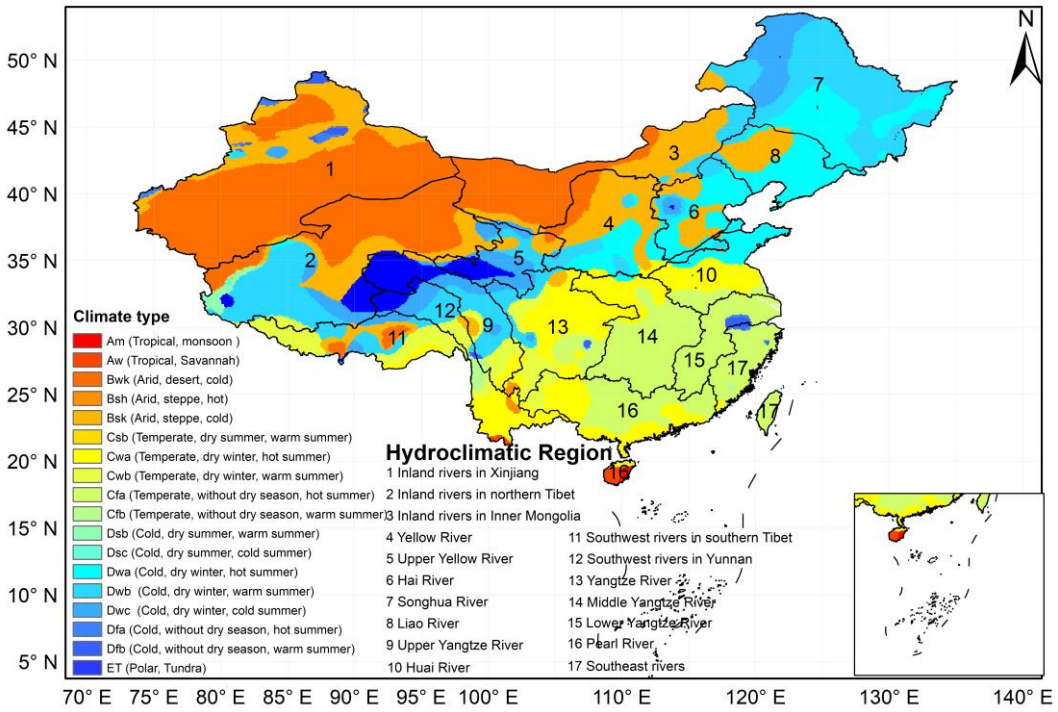


**Figure 1.** 17 hydroclimatic regions over China.

The intraseasonal oscillation is always represented by outgoing longwave radiation (OLR), zonal winds in the upper (200 hPa) and lower (850 hPa) troposphere. Although several indices, including the RMM (Realtime Multivariate MJO) index (Wheeler and Hendon, 2004) and BSISO index (Lee et al., 2013), have been proposed to monitor the propagation of oscillation, these indices may not cover patterns which might be important for sub-seasonal precipitation in certain regions. To overcome this problem, we analyze the correlation between ISO signals of preceding global OLR, zonal wind at 850 hPa (U850), zonal wind at 200 hPa (U200) and precipitation for each grid cell. In addition, the correlations with geopotential height at 850 hPa, 500 hPa, and 200 hPa (H850, H500, H200) are also analyzed. The H850, H500, and H200 have been proved to be as capable of reflecting the MJO structure as the zonal wind (Leung and Qian, 2017). The OLR data used in this study is derived from National Climate Data Center (NCDC) on a 1.0° squared resolution over the globe. The OLR data is developed from high resolution infrared radiation sounder instruments, and is valuable to a wide range of applications. A more detailed description of the OLR dataset can be found at https://www.ncdc.noaa.gov/news/new-outgoing-longwave-radiation-climate-data-record. The global gridded daily average U850, U200, H850, H500, H200 data are derived from the ERA5 reanalysis dataset at https://cds.climate.copernicus.eu/. The ERA5 reanalysis dataset is produced using advanced 4D-Var data assimilation scheme, and its horizontal resolution is approximately 30 km with 137 pressure levels in the vertical (Hersbach et al., 2020). It provides hourly record of global atmosphere, land surface and ocean waves from 1950 to present. To focus on large-scale features and increase computational efficiency, both the OLR

data and the ERA5 reanalysis data are bilinearly interpolated onto 2.5° × 2.5° latitude-longitude resolution. Moreover, we choose to focus on the period of 1979-2016 to be consistent with the temporal coverage of observed precipitation data.

The STP-BHM model we built in this study is compared to the dynamical models to provide a benchmark for sub-seasonal precipitation forecasts. In this study, we compare our results of the STP-BHM model with the NCEP model archived in the S2S Database for the same period of 1999-2010 from May to October (http://apps.ecmwf.int/datasets/data/s2s/). The NCEP hindcasts are produced by the Climate Forecast System version 2 (CFSv2), which is composed of land, ocean and atmosphere components. The system provides a 4-member ensemble run every day from 1st January 1999 to 31 December 2010. More details of the NCEP hindcasts are available at https://confluence.ecmwf.int/display/S2S/NCEP+Model+Description. The pentad mean precipitation amount forecasts of the NCEP model are generated to be consistent with the STP-BHM model.

## 2.2 Methodology

### 2.2.1 Modeling structure

The spatial-temporal projection based Bayesian hierarchical model (STP-BHM) falls into three parts as shown in Figure 2. The first part (Sect. 2.2.2) extracts intraseasonal signals of each global atmospheric field (U850, U200, OLR, H850, H500, and H200) and precipitation using a non-filtering method proposed by Hsu et al. (2015). In the second part (Sect. 2.2.3), the cell-wise correlation between ISO signals of atmospheric field and precipitation is analyzed in the six preceding pentads. The spatial-temporal coupled co-variance 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 co-variance field and atmospheric intraseasonal signals of atmospheric field at each preceding pentad. In the statistical modeling step (Sect. 2.2.4), both predictors and predictand are transformed to follow normal distributions. A Bayesian hierarchical model is then built to address the uncertainty in relationship between the predictors and predictand. The model is applied to generate probabilistic sub-seasonal precipitation forecasts after parameter inference.

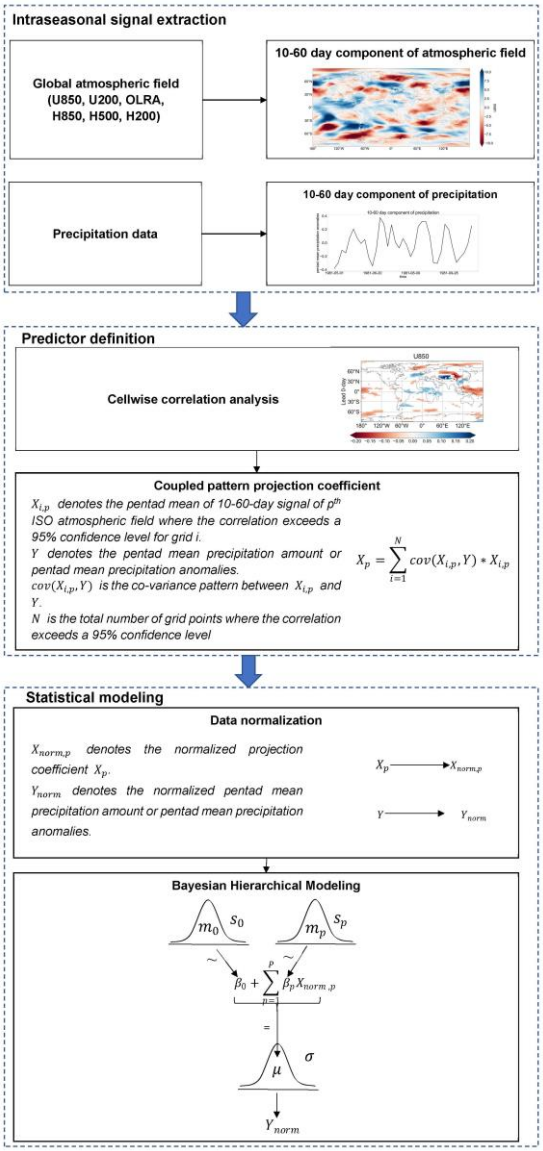


**Figure 2**. workflow of the spatial-temporal projection based Bayesian hierarchical model (STP-BHM).

### 2.2.2 Intraseasonal signal extraction

As briefly introduced in previous section, extracting meaningful intraseasonal signals is important for sub-seasonal precipitation forecasts. However, high-frequency (unpredictable) noises exist for both raw daily atmospheric variables (U850, U200, OLR, H850, H500, and H200) and raw daily precipitation. Band-pass filtering methods, such as the fast Fourier transformation, are always used to isolate intraseasonal low-frequency (10-60-day) signals from raw data (Zhang, 2005). However, traditional band-pass filtering method is impractical for real time applications as future information beyond the current date is needed. In this study, a non-filtering method proposed by Hsu et al. (2015) is used to extract 10-60-day signals of both atmospheric variables and precipitation. Compared to traditional intraseasonal signal extraction method, this approach is easy to implement and could be used for real time applications. The climatological annual cycle of raw daily data is first removed by subtracting 90-day low-pass filtered climatological component:

$$X' = X - \bar{X} \quad (1)$$

where $X$ is the raw daily data of atmospheric field or precipitation. $\bar{X}$ is the corresponding climatological 90-day low-pass filtered component derived by Lanczos filtering method during the period of 1981-2010 (Duchon, 1979). In a second step, lower-frequency signals longer than 60 days are removed by subtracting the last 30-day running mean,

$$X'' = X' - \overline{X'}^{30d} \quad (2)$$

where $\overline{X'}^{30d}$ is the last 30-day running mean of $X'$.

The higher-frequency signals are then removed by taking a pentad mean,

$$X^* = \overline{X''}^{5d} \quad (3)$$

The so-derived variable represents the 10-60-day signals of 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-day signal of precipitation is also referred as pentad mean precipitation anomalies in the following sections.

#### 2.2.3 Predictor definition

To identify relevant areas of atmospheric fields that could affect 10-60-day precipitation variability, we analyze the correlation between preceding 10-60-day signals of atmospheric fields and precipitation for each hydroclimatic region during the period of 1979-2016 from May to October. Owing to the data persistence introduced by the filtering method, the effective degree of freedom for each grid cell and each preceding pentad is estimated following Livezey and Chen (1983).

As an example, Figure 3 and Figure 4 presents the correlation between preceding pentad mean 10-60-day signals of U850, U200, OLRA, H850, H500, H200 and precipitation over Region 1 (Inland Rivers in Xinjiang) at different lead times. At leads 25 to 20 days, the significantly correlated U850 signals are mainly over western Indian Ocean. The U850 signals are then propagating eastward toward Equatorial Indian Ocean at the lead of 15 to 10 days. The U850 anomalies then gradually moved eastward and northward toward West Pacific Ocean, Mongolia plateau, Iranian plateau, and Qinghai-Tibet plateau from the lead of 10 to 0 days. The U200 signals are more pronounced compared to U850 signals. The spatial distribution of potential predictive U200 regions is rather concentrated, indicating more robust statistical relationships. The OLR anomalies appear near the Bay of Bengal at 20 to 15 day leads. At lead 5 to 0 day, the significantly correlated OLR signals are mainly over the East European Plain and West Siberian Plain.

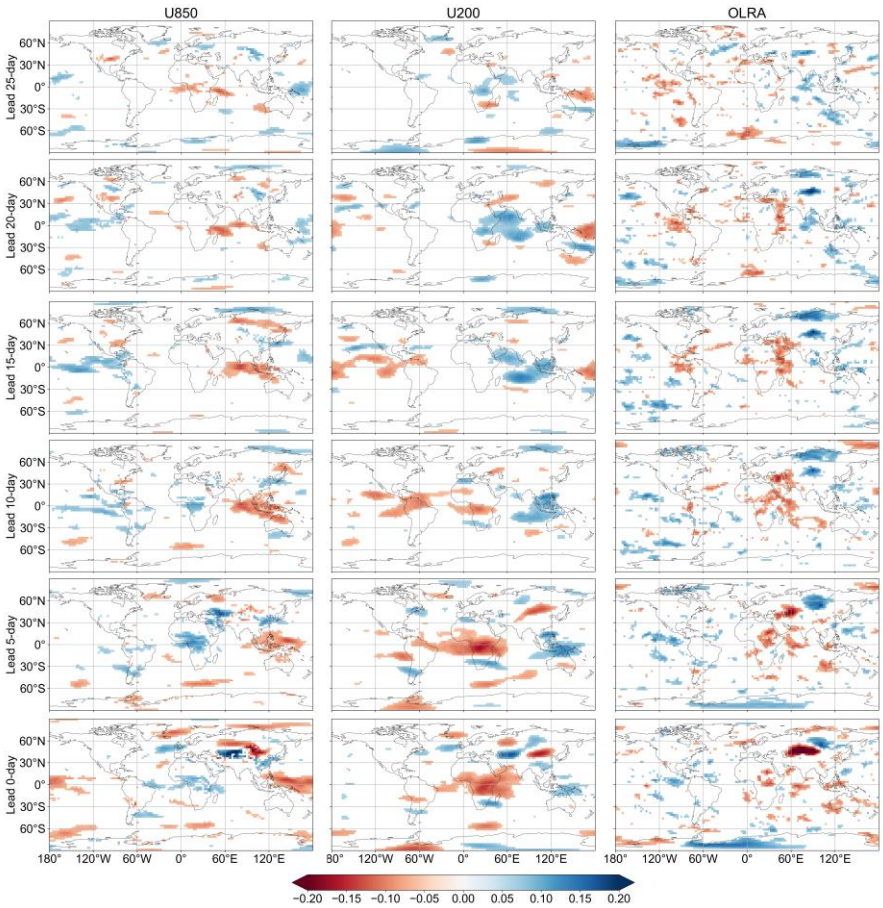


**Figure 3.** Correlation coefficient between preceding pentad mean 10-60-day signals of U850, U200, OLRA and precipitation over Region 1 (Inland Rivers in Xinjiang) at different lead times during the period of 1979~2016 from May to October. Correlation coefficients statistically significant at the 5% level are shaded.

The H850 anomalies appear near the equatorial Indian Ocean and Philippine Sea at the lead of 25 days to 20 days. At leads 15 to 20 days, the significantly correlated H850 signals are mainly over Africa. The signals gradually move eastward and northward toward Indian Ocean, Iranian plateau, and Central Asia from the lead of 10 to 0 days. Unlike the H850 fields that originated over Africa, the H200 anomaly appears to originate from Arabian Sea, southern Indian Ocean, and West Pacific Ocean from leads of 25 to 15 days. At lead 10 to 0 days, the significantly correlated H200 signals are mainly over East European Plain, West Siberian Plain, and Central Siberian plateau.

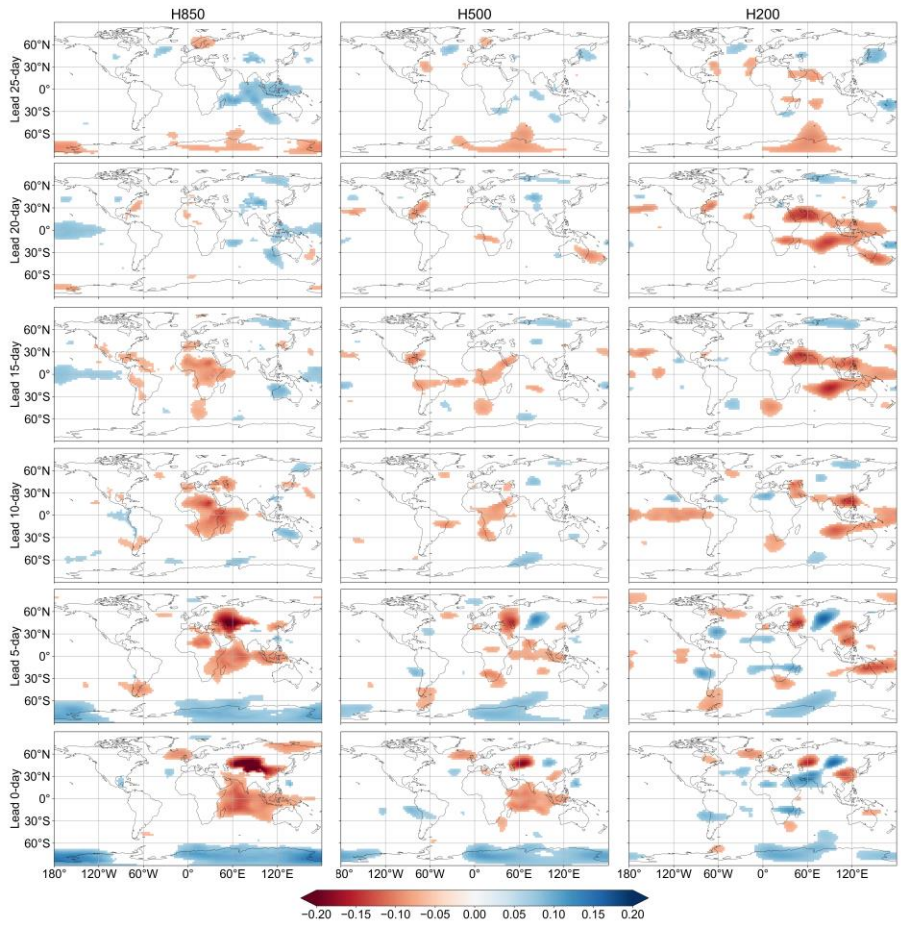


**Figure 4**. Same as Figure 3, but for H850, H500, and H200.

The correlation maps between preceding pentad mean 10-60-day signals of U850, U200, OLRA, H850, H500, H200 and precipitation over other regions are presented in Figures S1 to S32.

The spatial-temporal coupled co-variance patterns are then constructed for grid point where the correlation statistically significant at the 5% level. The predictor is then defined by summing the product of the co-variance patterns and ISO signals of atmospheric field at each preceding pentad,

$$cov(X_{i,p}, Y) = \frac{1}{T}\sum_{t=1}^{T}(y_t - E(y))(x_{i,p,t} - E(x_{i,p})) \quad (4)$$

$$X_p = \sum_{i=1}^{N} cov(X_{i,p}, Y) * X_{i,p} \quad (5)$$

where $X_{i,p}$ denotes the pentad mean 10-60-day signal of $p^{th}$ atmospheric field where the correlation statistically significant at the 5% level for grid $i$, $p = 1,2,\cdots,6$. $Y$ denotes the pentad mean precipitation amount or pentad mean precipitation anomalies. $T$ is the total number of pentads, and $N$ is the total number of grid points where the correlation statistically significant at the 5% level. Thus, there is only one predictor $X_p$ for each atmospheric field and each preceding pentad.

### 2.2.4 Statistical modelling

In previous steps, we defined predictors by analyzing the relationship between ISO signals of atmospheric field and precipitation. The so-derived predictors can be used to predict pentad mean precipitation amount as well as pentad mean precipitation anomalies. Consider, for example, predicting pentad mean precipitation amount for the period between 1st May and 5th May, 1979. In this case, pentad mean ISO signals of atmospheric field on 26th~30th April, 21th~25th April, 16th~20th April, 11th~15th April, 6th~10th April, 1st~5th April 1979 are used as predictors to generate precipitation forecasts at different lead times. A leave-one-year-out cross-validation strategy is implemented for both data normalization, model building, parameter inference, and verification to avoid any bias in skill (Michaelsen, 1987). For instance, to produce sub-seasonal precipitation forecasts in 1979, the predictors (preceding ISO signals) and predictand (pentad mean precipitation) during the period of 1980-2016 are pooled together for statistical modelling. The forecasts for the year 1979 are then issued by models trained on 1980-2016, and the performance is evaluated against the observations. This cross-validation strategy ensures that the data used for evaluation is never used for statistical modelling.

Before establishing the Bayesian hierarchical model, the predictors $\boldsymbol{X}^T = [X_1\, X_2 \cdots X_P]$ are normalized to $\boldsymbol{X}^T_{norm} = \left[X_{norm,1}\, X_{norm,2} \cdots X_{norm,P}\right]$ through the Yeo-Johnson transformation method as the input variables are allowed to be negative (Yeo and Johnson, 2000). The predictand $Y$ is normalized to $Y_{norm}$ using the Yeo-Johonson method for pentad mean precipitation anomalies. However, the pentad mean precipitation amount is highly skewed with numerous zero values. Here, we normalize the pentad mean precipitation amount $Y$ to $Y_{norm}$ using the log-sinh transformation method proposed by Wang et al. (2012). The normalization parameters are estimated using the SCE-UA (shuffled complex evolution method developed at The University of Arizona) method that maximize the log-likelihood function for both the Yeo-Johnson transformation method and log-sinh transformation method.

There are many versions and variations of BHMs. In this study ,we establish the BHM model following Devineni et al. (2013) and Chen et al. (2014). The spatial correlation of precipitation over different regions is not considered here. A traditional no-pooling BHM is built for each hydroclimatic region separately. The normalized predictand $Y_{norm}$ is assumed to follow the normal distribution,

$$Y_{norm} \sim N(\mu, \sigma^2) \tag{6}$$

We then link the parameter $\mu$ with the normalized predictors using a linear model,

$$\mu = \beta_0 + \sum_{p=1}^{P} \beta_p X_{norm,p} \tag{7}$$

where $\beta_p$ is the slope term corresponding to the normalized predictor $X_{norm,p}$, and $P$ is the total number of predictors used for prediction.

To complete the hierarchical formulation, we assume the unknown parameters, including $\sigma,\ \beta_0,\ \cdots,\ \beta_P$, follow non-informative priors:

$$\frac{1}{\sigma^2} \sim \mathrm{U}(0, 100) \tag{8}$$

$$\beta_0 \sim N(0,\ 10^4) \tag{9}$$

$$\beta_p \sim N(0,\ 10^4), \qquad p = 1, \cdots, P \tag{10}$$

This implies that the information used for posterior distribution inference is only provided by the data.

Given $\boldsymbol{\theta} = \{(\sigma,\ \beta_0,\ \beta_p),\ p = 1, \cdots, P\}$ denotes parameters in the Bayesian hierarchical model for a certain region and lead time, the full posterior of the parameters is given as:

$$p(\boldsymbol{\theta}|Y_{norm}, \boldsymbol{X}_{norm}^T) \propto p(Y_{norm}|\boldsymbol{\theta}, \boldsymbol{X}_{norm}^T)p(\boldsymbol{\theta}) \tag{11}$$

where $p(Y_{norm}|\boldsymbol{\theta}, \boldsymbol{X}_{norm}^T)$ is the likelihood, and $p(\boldsymbol{\theta})$ is the prior of parameters $\boldsymbol{\theta}$. As the posterior distributions of parameters $\boldsymbol{\theta}$ are not standard distributions, it is difficult to conduct analytical integration. In this study, we use the R package *runjags* (Denwood, 2016) to estimate the parameters of the BHM. The *runjags* offers an interface to facilitate calibrating BHMs employ a Gibbs sampling algorithm in Just Another Gibbs sampler (JAGS). The initial values of model parameters $\boldsymbol{\theta}$ are first randomly sampled from prior distributions. The parameters $\boldsymbol{\theta}$ are then updated based on the full conditional distributions. We use five independent Markov chains in each model run, with a total number of 10, 000 iterations for each chain. The convergence is ensured by the potential scale reduction factor $\hat{R}$ (Brooks and Gelman, 1998). An approximate convergence is diagnosed when the $\hat{R}$ is less than 1.1 for all parameters.

Once the parameters are sampled, the Bayesian hierarchical model can be used to predict pentad mean precipitation amount or pentad mean precipitation anomalies using preceding ISO signals. Given new preceding predictors $\boldsymbol{X}^{*T} = [X_1^*\ X_2^* \cdots X_P^*]$, the normalized predictors $\boldsymbol{X}_{norm}^{*T} = \left[X_{norm,1}^*\ X_{norm,2}^* \cdots X_{norm,P}^*\right]$ are found using the estimated transformation parameters during the training period. The posterior predictive distribution of normalized predictand is given as:

$$Y_{norm}^* \sim N(\mu^*, \sigma^2) \tag{12}$$

$$\mu^* = \beta_0 + \sum_{p=1}^{P} \beta_p X_{norm,p}^* \tag{13}$$

Again, the Gibbs sampling algorithm is used to obtain samples of $Y_{norm}^*$ by giving each of the 1000 sets of

parameter values $\boldsymbol{\theta}$. The samples of $Y^*_{norm}$ are then back-transformed to produce ensemble precipitation forecasts of $Y^*$.

### 2.2.5 Verification

In this study, we assess the performance of the STP-BHM model for both pentad mean precipitation amount and pentad mean precipitation anomalies. The Continuous Ranked Probability Score (CRPS) is used to provide an overall evaluation of the accuracy of probabilistic forecasts for both the pentad mean precipitation amount and pentad mean precipitation anomalies:

$$CRPS = \frac{1}{N}\sum_{i=1}^{N}\int[F_i(y) - H(y - o_i)]^2 dy \quad (14)$$

where $F_i()$ is the cumulative distribution function of the ensemble forecasts for pentad mean precipitation amount or pentad mean precipitation anomalies for case $i$; and $H()$ is the Heaviside step function defined as:

$$H(y - o_i) = \begin{cases} 0 & y < o_i \\ 1 & y \geq o_i \end{cases} \quad (15)$$

where $o_i$ is the corresponding observation.

The CRPS skill score is then calculated by comparing the CRPS of ensemble forecasts with the CRPS of reference forecasts:

$$CRPS_{SS} = \frac{CRPS_{REF} - CRPS}{CRPS_{REF}} \times 100\% \quad (16)$$

The reference forecasts are generated using the Bayesian hierarchical model with no predictors used for prediction. This is also referred as the cross-validated climatology. A skill score of 100% indicates that the ensemble forecasts are the same as the observations, whereas a skill score of 0% suggests that the ensemble forecasts show no improvement over the cross-validated climatology. A negative skill score means that the ensemble forecasts are inferior to the cross-validated climatology.

We also use the Brier Score (BS) to assess the capability of the STP-BHM model for predicting below-normal and above-normal events. The below-normal and above-normal events are defined using the terciles of pentad mean precipitation amount or pentad mean precipitation anomalies of cross-validated climatology.

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2 \quad (17)$$

where $p_i$ is the forecast probability of the below- or above-normal event for case $i$; and $o_i$ is the observed occurrence (0 or 1).

The Brier skill score (BSS) is then calculated as follows:

$$BSS = \frac{BS_{REF} - BS}{BS_{REF}} \times 100\% \quad (18)$$

where the $BS_{REF}$ is the $BS$ of the cross-validated climatology. The $BSS$ measures the relative skill of the forecast compared to climatology. Like the CRPS skill score, the Brier skill score takes the value 100% for perfect forecasts and 0% for the cross-validated climatology.

In this study, we use the attribute diagram to assess the reliability, resolution, and sharpness of probabilistic forecasts for both below-normal event and above-normal event. The attribute diagram shows the observed frequencies against its forecast probabilities for a given event with binary outcomes (Hsu and Murphy, 1986). The forecast probability is binned as 5 equal-width intervals, which are [0.0, 0.2), [0.2, 0.4), [0.4, 0.6), [0.6, 0.8), and [0.8, 1.0]. The corresponding observed relative frequency is plotted against the mean forecast probability in each bin. The forecasts are reliable if the scatters are along the 45-degree diagonal. The sharpness is also shown on the attribute diagram. The forecasts are sharp if the probabilities tend to be either very high (e.g. > 90%) or very low (e.g. <10%) (Peng et al., 2014). The size of each dot represents the fraction of forecasts that fall into a particular probability bin. Thus, the sharpness is indicated by the size of dots in each bin. The attribute diagram requires a large number of samples to draw robust conclusions. In this study, the probabilistic forecasts over the 17 hydroclimatic regions are pooled together to increase the sample size for each lead time.

## 3. Results

### 3.1 Forecast skill of pentad mean precipitation amount

Figure 5 presents the cross-validated CRPS skill scores for sub-seasonal forecasts of pentad mean precipitation amount at different lead times (lag times). Positive CRPS skill scores are found over all regions and all lead times, indicating that the STP-BHM model outperforms the cross-validated climatological forecasts. The CRPS skill scores are mostly over 10% in southern China even when the lead time is beyond 10 days. On the contrary, the performance of the STP-BHM model is relatively poorer in northern China with CRPS skill scores ranging from 5% to 10% at the same lead time.

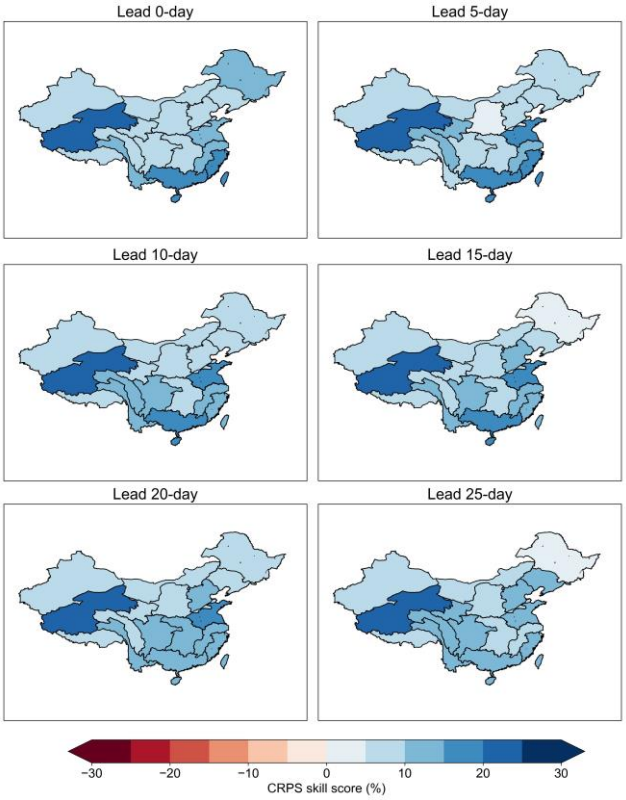


**Figure 5**. The cross-validated CRPS skill scores of the STP-BHM model for pentad mean precipitation amount forecasts at different lead times during the period of 1979-2016 from May to October.

Figure 6 illustrates the Brier skill scores of the STP-BHM model for both below-normal and above-normal events at different lead times. As can be seen in Fig. 6, the Brier skill scores are mostly above 15% for both the below-normal and above-normal events. This indicates that the STP-BHM model can provide skillful sub-seasonal forecasts for extreme events as well. Furthermore, the Brier skill scores are most ranging from 20% to 25% in southern China at leads of 20-25 days for below-normal events. The STP-BHM model shows lower forecast skills for above-normal events, which the Brier skill scores are mostly between 15% and 20%. This indicates that the below-normal events are more predictable compared to the above-normal events.

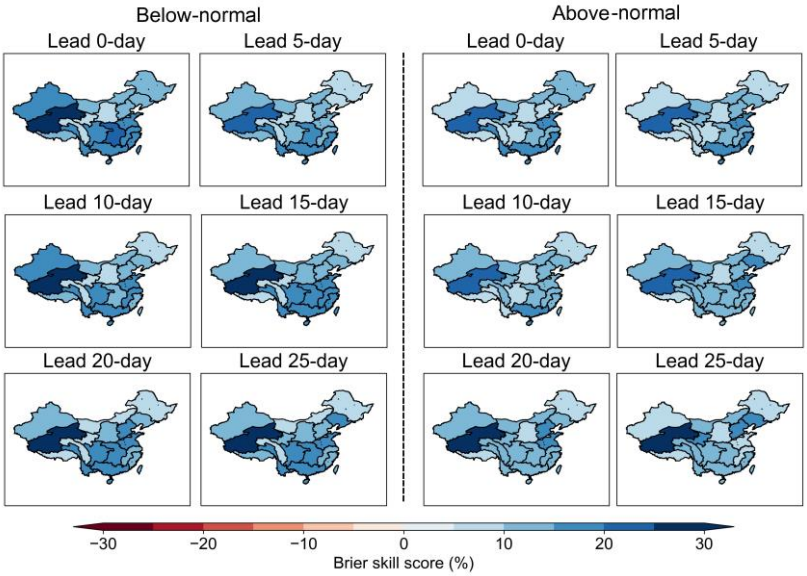


**Figure 6**. The Brier skill scores of the STP-BHM model for the prediction of below-normal and above-normal events of pentad mean precipitation amount at different lead times during the period of 1979-2016 from May to October.

To help identify the main sources of sub-seasonal precipitation predictability, we also establish the STP-BHM model for each atmospheric field separately. Figure 7 compares the CRPS skill scores of pentad mean precipitation forecasts with different predictors. In general, U850, U200, H850, and H500 show higher forecast skill compared to OLRA and H200 for almost all hydroclimatic regions and lead times. This suggests that the ISO signals of these atmospheric fields contribute more to the overall forecast skill. Compared to the STP-BHM model built with only one predictor, the forecast skill is further improved when all ISO signals of atmospheric fields are used.

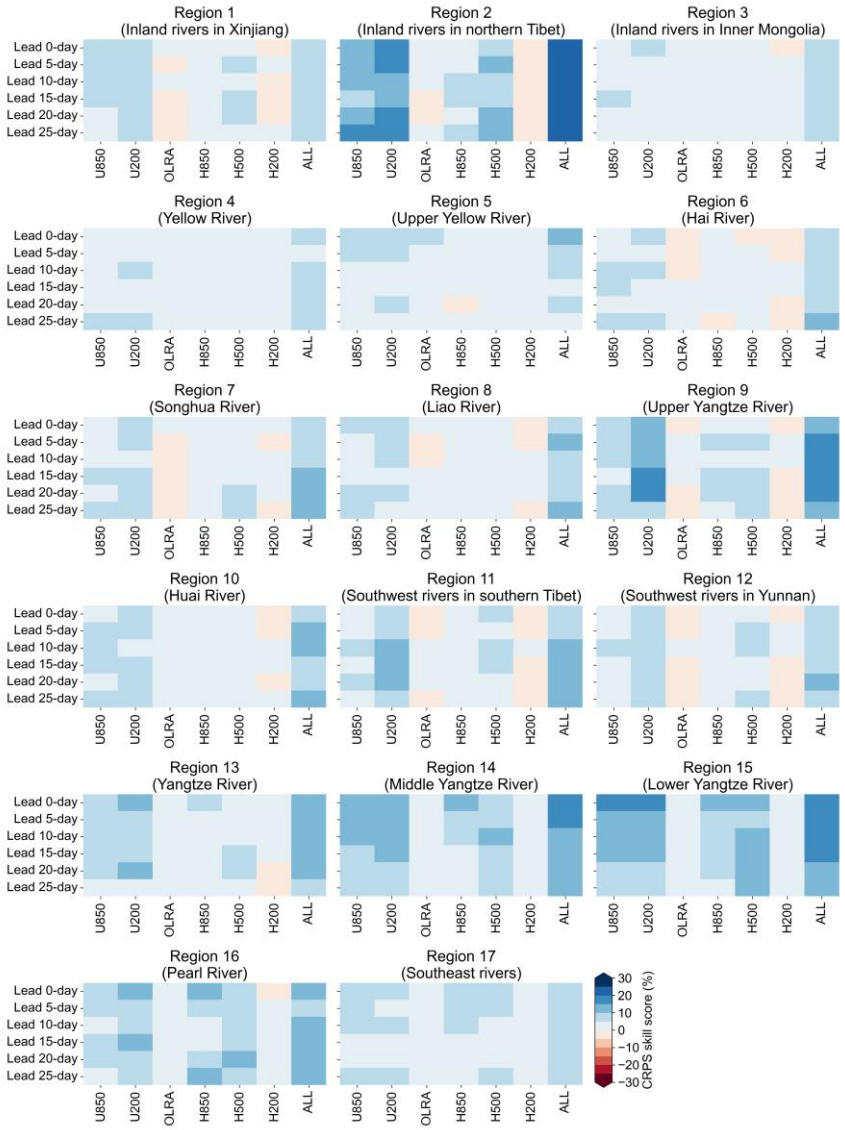


**Figure 7**. The cross-validated CRPS skill scores of the STP-BHM model for pentad mean precipitation amount forecasts with different predictors (U850, U200, OLRA, H850, H500, H200). ALL denotes that the ISO signals of all atmospheric fields are used as predictors.

The attribute diagrams of sub-seasonal forecasts of pentad mean precipitation amount for below-normal and above-normal event at different lead times are shown in Figure 8. Most points fall near the 1:1 line for both below-normal and above-normal events at all lead times. This suggests that the probabilistic forecast distributions are reliable. However, the forecast probabilities deviate slightly from the 1:1 line at higher forecast probabilities for above-normal event, indicating that the sharpness for above-normal event should be further improved.

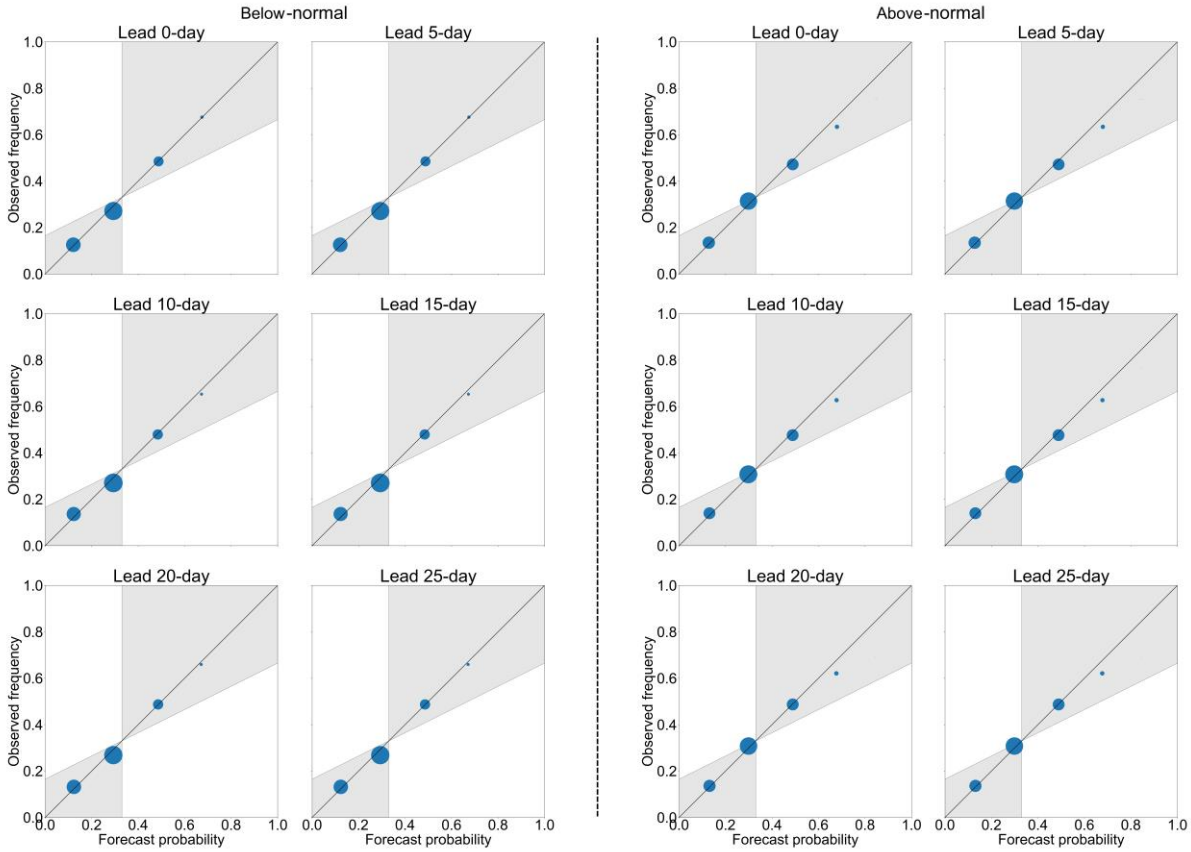


**Figure 8**. The attribute diagram of the STP-BHM model for the prediction of below-normal and above-normal events of pentad mean precipitation amount at different lead times. Forecast probability is binned with width of 0.2. The size of each dot represents the fraction of forecasts that fall into a particular probability bin.

Figure 9 compares the CRPS skill scores of the STP-BHM model and the NCEP model from May to October during the period of 1999~2010. It is not surprising that the NCEP model outperforms the STP-BHM model when the lead time is within 5 days. However, we should note that the STP-BHM model shows much higher probabilistic forecast skill compared to the NCEP model at longer lead times. Positive CRPS skill scores are observed for the STP-BHM model over most hydroclimatic regions, whereas the skill scores are mostly negative for the NCEP model.

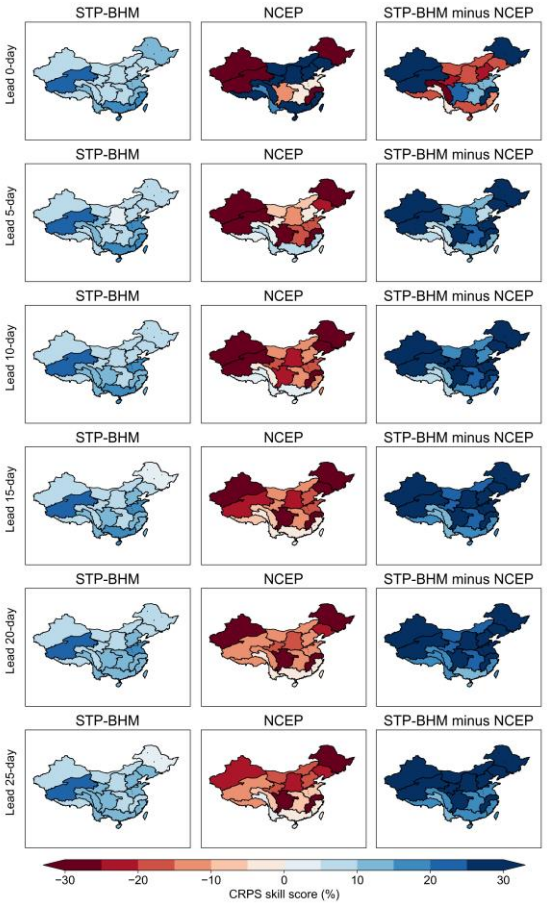


**Figure 9**. The comparison of the CRPS skill scores of the STP-BHM model and the NCEP model during the period of 1999~2010 from May to October.

### 3.2 Forecast skill of pentad mean precipitation anomalies

The cross-validated CRPS skill scores for sub-seasonal forecasts of pentad mean precipitation anomalies are shown in Figure 10. The STP-BHM model shows positive CRPS skill scores over most hydroclimatic regions, except Inland rivers in Xinjiang (Region 1) and Inland rivers in northern Tibet (Region 2). This may be explained by the relatively lower variability of pentad mean precipitation anomalies in these regions. In addition, the STP-BHM model shows higher forecast skill in eastern China with CRPS skill scores range from 10% to 15%. In comparison, the forecast skill in Inland Rivers in Inner Mongolia (Region 3), Upper Yellow River (Region 5), Upper Yangtze River (Region 9), Southwest rivers in southern Tibet (Region 11), Southwest rivers in Yunnan (Region 12), Yangtze River (Region 13), and Pearl River (Region 16) are lower. Similar results are also found by Zhu and Li (2017c), which the southwestern China shows lowest forecast skill compared to other regions.

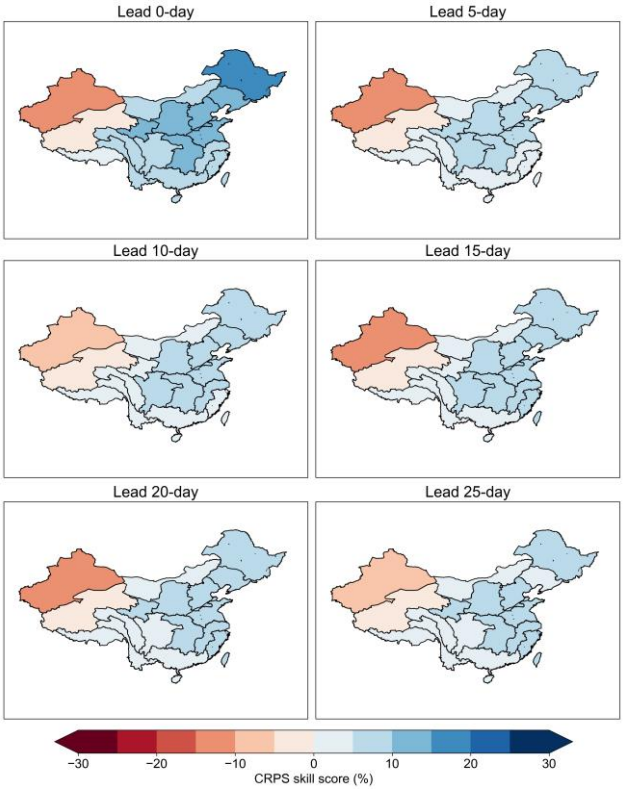


**Figure 10**. Same as Figure 5, but for pentad mean precipitation anomalies.

The Brier skill scores of pentad mean precipitation anomalies for below-normal and above-normal events are presented in Figure S33. Positive Brier skill scores are found over all regions and all lead times, indicating that the STP-BHM model outperforms the cross-validated climatological forecasts for extreme events. Meanwhile, the differences of Brier skill scores in different hydroclimatic regions are small, where the Brier skill scores are mostly ranging from 5%~15% for both below-normal and above-normal events.

Figure 11 compares the CRPS skill scores of pentad mean precipitation anomalies with different predictors. Overall, the STP-BHM model with OLRA used as predictor shows higher forecast skill compared to other predictors for almost all hydroclimatic regions and lead times. This suggests that the OLRA contributes most to the overall forecast skill of pentad mean precipitation anomalies.

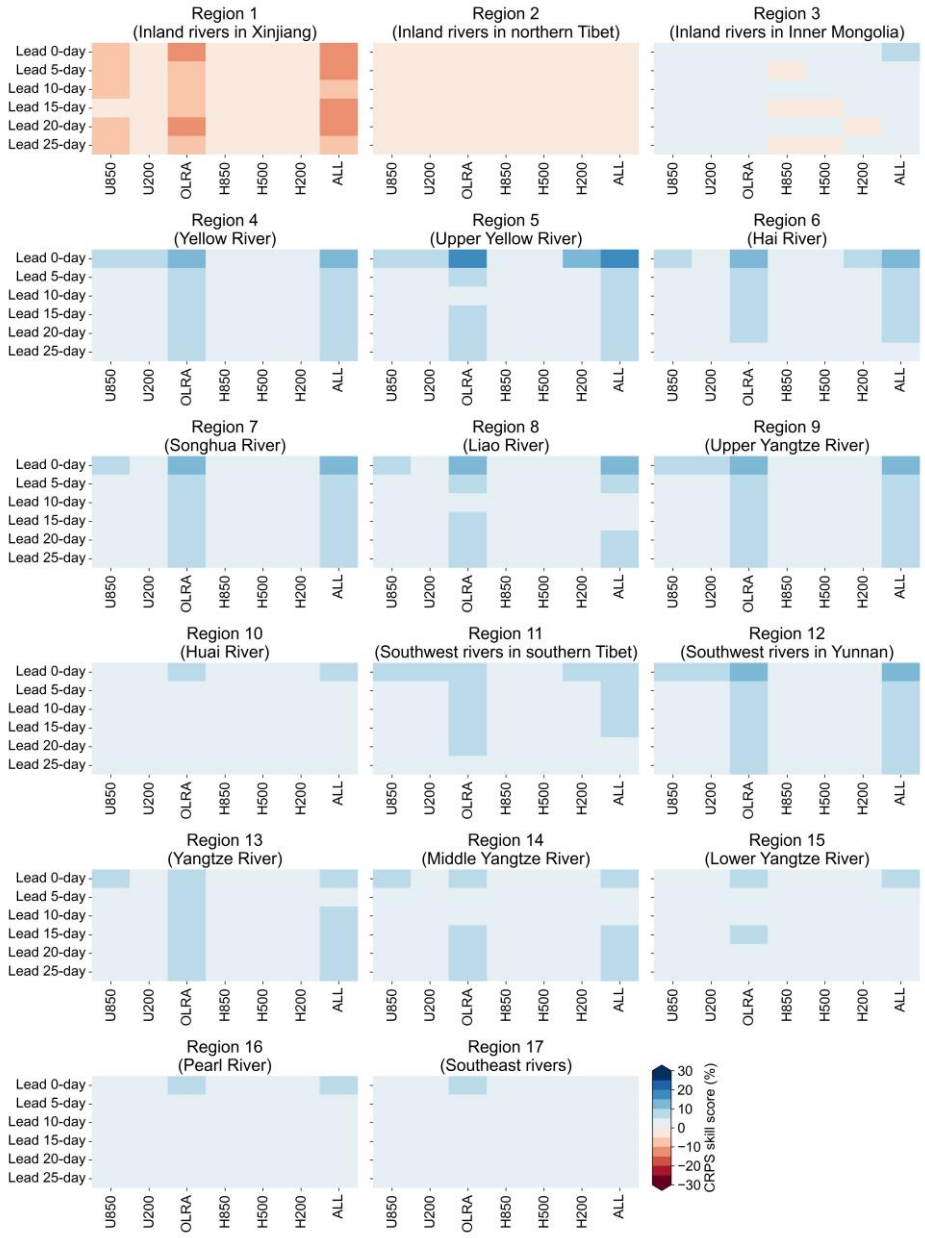


**Figure 11**. Same as Figure 7, but for pentad mean precipitation anomalies.

Shown in Figure S34 is the attribute diagram of sub-seasonal forecasts of pentad mean precipitation anomalies for below-normal and above-normal event at different lead times. Most points fall close to the 1:1 line for both below-normal and above-normal events. This suggests that the probabilistic forecast distributions are reliable for pentad mean precipitation anomalies as well. The sharpness of STP-BHM model is also observed especially for below-normal event.

## 4. Discussion

### 4.1 Forecast skill and possible mechanism

In this study, we first analyze the relationship between preceding ISO signals of atmospheric fields and precipitation. The coupled patterns are extracted and the corresponding projection coefficients are defined as predictors. A Bayesian hierarchical model is then established and applied to predict both pentad mean precipitation amount and pentad mean precipitation anomalies over China. Our results suggest that the STP-BHM model can provide skillful and reliable probabilistic forecasts for both pentad mean precipitation amount and pentad mean precipitation anomalies at a lead of 20-25 days over most hydroclimatic regions in China. However, the spatial patterns of skill scores suggest that the STP-BHM model is more skillful over southern China. This may be explained by the different characteristics of intraseasonal variability and different possible mechanism over different hydroclimatic regions. Wang (2007) analyzed the precipitation variability from April to September over China, and the results suggested that the seasonal component accounted for nearly 70% of the total variability over northeastern China. The intraseasonal (10-90 days) component only accounted for nearly 7% of the total variability, which indicate that the intraseasonal precipitation over these regions have no significant frequency peak. In comparison, the sub-seasonal component accounted for over 20% of the total variability in southeastern China. Ouyang and Liu (2020) also found that the boreal summer monsoon intraseasonal variability of precipitation over the lower Yangtze River basin was mainly dominated by the relatively low-frequency 12-20-day variability and high-frequency 8-12-day variability. Wang and Duan (2015) demonstrated that the Quasi-biweekly oscillation (QBWO, 10-20 days oscillation) was the dominant mode of intraseasonal variability of summer precipitation over the Tibetan Plateau. The relations between atmospheric intraseasonal oscillation and the low-frequency variability of precipitation vary from region to region as well. Ren and Shen (2016) suggested that the impact of tropical atmospheric intraseasonal oscillation on precipitation were more significant in regions in southern China and the Tibetan Plateau areas during the boreal summer. We also note that the forecast skill of pentad mean precipitation amounts and precipitation anomalies are different. The precipitation amounts behavior at different timescales (interannual, intraseasonal, and synoptic). The large-scale circulation anomalies (U850, U200, H850, and H500) may be dominant for the total variability of precipitation amounts. In comparison, the precipitation anomalies only represent the intraseasonal component of precipitation. The OLR plays a more important role for intraseasonal convections compared to other dynamical fields (Ventrice et al., 2013; Liu et al., 2016).

We should also note that the CRPS skill scores of the STP-BHM model are lower than NCEP dynamical models

at short lead times. The Calibration, Bridging, and Merging (CBaM) method, which makes the best use of GCM outputs, has been proved to be efficient for improving seasonal precipitation over many regions (Strazzo et al., 2019; Schepen and Wang, 2013; Peng et al., 2014). Recently, Specq and Batté (2020) proposed a similar statistical-dynamical approach to improve sub-seasonal precipitation forecasts over the southwest tropical Pacific. In the future, the statistical forecasts generated from lagged atmospheric indices should be included in the calibrated forecasts to further improve sub-seasonal precipitation forecast skill.

### 4.2 Limitations and future work

In this study, the correlation between ISO signals of atmospheric fields and precipitation is analyzed using the whole record despite the cross-validation strategy used for statistical modelling. This may introduce artificial skill into the model to some extent. However, the correlation analysis for each step of the cross-validation is difficult in practice. We analyze the spatial patterns of correlations between preceding ISO signals of U850 and precipitation over Region 1 at the lead time of 0-day for the period of 1979-2016 and 1980-2016 (Figure S35). The results show small variability between the cross-validated correlation and the whole-period correlation. In addition, the cross-validation strategy used in the statistical modelling procedure can also reduce the chance of overfitting (Vehtari and Lampinen, 2002; Delsole and Shukla, 2009).

Another limitation of this study is the treatment of zero values adopted in the statistical modelling procedure. We treat the zero values as censored data, which is also referred as “explicit” approach in Mcinerney et al. (2019). Although this treatment performed well in “low-ephemeral” and “mid-ephemeral” catchments, the performance of this “explicit” approach was poor in “high-ephemeral” (> 50% zero flows) catchments. Further development is required to overcome this problem. The copula functions are flexible in choosing marginal distributions, and have been widely used in hydrological simulations in recent years (Zhang and Singh, 2007; Vernieuwe et al., 2015; De Michele and Salvadori, 2003). Compared to the Bayesian statistics we used in this study, the copula functions are more general and the normalization may not be required when the skewed distributions are used as the marginal distribution of precipitation. This may provide a possible solution to overcome the problems caused by the large amount of zero values.

We built the Bayesian hierarchical model for each hydroclimatic region separately. However, the spatial patterns of precipitation have not been considered yet. The spatial Bayesian hierarchical model, which can

capture the spatial dependence of precipitation between different regions, could be used to provide sub-seasonal precipitation forecasts with spatial coherence (Reza Najafi and Moradkhani, 2013; Bracken et al., 2016). An alternative way to reconstruct the spatial patterns of probabilistic precipitation forecasts is to use the Schaake Shuffle method or Ensemble Copula Coupling method (Roman et al., 2013; Clark et al., 2004). Higher spatial or temporal resolution of precipitation forecasts are also needed for sub-seasonal streamflow forecasts. However, our previous studies indicated that post-processed daily precipitation forecasts from GCMs are of low accuracy when the lead time is beyond 10-14 days (Li et al., 2020). In this study, the large-scale ISO signals are only used to predict pentad mean precipitation as the noise of daily precipitation is too large. Spatial or temporal disaggregation may be required in the future to provide daily precipitation forecasts as inputs for hydrological models.

## 5. Conclusions

Accurate and reliable sub-seasonal precipitation forecasts are difficult as the predictability from atmospheric initialization is lost after two weeks, while the slowly varying boundary conditions do not have substantial impact at such a time scale. The intraseasonal oscillation (ISO) is considered as one of the leading sources of sub-seasonal predictability. However, the relationship between atmospheric intraseasonal signals and precipitation is of high uncertainty. In this study, we first analyze the correlation between preceding atmospheric intraseasonal signals (U850, U200, OLRA, H850, H500, and H200) and precipitation. The spatial-temporal coupled co-variance patterns are constructed for grid points where the correlation statistically significant at the 5% level. The predictors are then defined by summing the product of the co-variance fields and ISO signals of atmospheric field. A Bayesian hierarchical model (BHM) is then built to address the uncertainty in the relationship between the ISO signals of atmospheric fields and precipitation. The posterior distributions of the model parameters are sampled using a Gibbs sampling algorithm. The STP-BHM model is used to predict both pentad mean precipitation amount and pentad mean precipitation anomalies after parameter inference. The performance is evaluated through a leave one-year-out cross-validation strategy.

Our results suggest that the STP-BHM model we built in this study can provide skillful and reliable probabilistic forecasts for both pentad mean precipitation amount and pentad mean precipitation anomalies at a lead of 20-25 days over most hydroclimatic regions in China when all ISO signals of atmospheric fields are used as predictors. In addition, the STP-BHM model shows useful predictive skill for below-normal and above-normal events as well, and positive Brier skill scores are observed at all lead times. The spatial patterns of skill scores

suggest that the STP-BHM model is more skillful over southern China compared to other regions. The STP-BHM model outperforms the NCEP S2S dynamical model when the lead time is beyond 5 days. To help identify the main sources of sub-seasonal precipitation predictability, we also establish the STP-BHM model for U850, U200, OLRA, H850, H500, and H200, separately. The results suggest that the ISO signals of U850, U200, H850, and H500 contribute more to overall forecast skills for pentad mean precipitation amount predictions. In comparison, the OLRA contributes most to the forecast skill for predictions of pentad mean precipitation anomalies.

In this study, the spatial patterns between ISO signals of the zonal wind at 850 and 200 hPa, Outgoing Longwave Radiation, and the geopotential height at 850, 500, and 200 hPa are extracted and used to define predictors. Other sources of sub-seasonal predictability, such as soil moisture, snow cover, and stratosphere-troposphere interaction, will be included in the Bayesian hierarchical model to further improve sub-seasonal precipitation forecasts. The Calibration, Bridging, and Merging (CBaM) method can also be investigated at sub-seasonal time scale to further improve the forecast skill (Schepen and Wang, 2013; Schepen et al., 2014).

## Data availability

The precipitation dataset used in this study can be derived from http://www.gloh2o.org/mswep/. The outgoing Longwave radiation (OLR) dataset can be found at https://www.ncdc.noaa.gov/news/new-outgoing-longwave-radiation-climate-data-record, and the ERA5 dataset can be sourced from https://cds.climate.copernicus.eu/.

## Author contribution

Y.L. and Z.Y. WU designed the experiments and Y.L. carried them out. H.H preprared the data, and H.Y developed the model code and performed the simulations. Y.L. prepared the manuscript with contributions from all co-authors.

## Competing interests

The authors declare that they have no conflict of interest.

## Acknowledgements

This work was funded by the National Natural Science Foundation of China (Grant numbers: 52009027, U2240225).

## References


Abbot, J. and Marohasy, J.: Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks, Atmospheric Research, 138, 166-178, 2014.

Annamalai, H. and Slingo, J. M.: Active/break cycles: diagnosis of the intraseasonal variability of the Asian Summer Monsoon, Climate Dynamics, 18, 85-102, 10.1007/s003820100161, 2001.

Barnston, A. G. and Smith, T. M.: Specification and Prediction of Global Surface Temperature and Precipitation from Global SST Using CCA, Journal of Climate, 9, 2660-2697, 10.1175/1520-0442(1996)009<2660:SAPOGS>2.0.CO;2, 1996.

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1? Precipitation: Methodology and Quantitative Assessment, Bulletin of the American Meteorological Society, 100, 473-500, 10.1175/BAMS-D-17-0138.1, 2019.

Bracken, C., Rajagopalan, B., Cheng, L., Kleiber, W., and Gangopadhyay, S.: Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain, Water Resources Research, 52, 6643-6655, https://doi.org/10.1002/2016WR018768, 2016.

Brooks, S. P. and Gelman, A.: General methods for monitoring convergence of iterative simulations, Journal of computational and graphical statistics, 7, 434-455, 1998.

Chen, T.-C., Wang, S.-Y., Huang, W.-R., and Yen, M.-C.: Variation of the East Asian Summer Monsoon Rainfall, Journal of Climate, 17, 744-762, 10.1175/1520-0442(2004)017<0744:VOTEAS>2.0.CO;2, 2004.

Chen, X., Hao, Z., Devineni, N., and Lall, U.: Climate information based streamflow and rainfall forecasts for Huai River basin using hierarchical Bayesian modeling, Hydrology and Earth System Sciences, 18, 1539-1548, 2014.

Chu, P.-S. and Zhao, X.: A Bayesian Regression Approach for Predicting Seasonal Tropical Cyclone Activity over the Central North Pacific, Journal of Climate, 20, 4002-4013, 10.1175/JCLI4214.1, 2007.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., and Wilby, R.: The Schaake Shuffle: A Method for Reconstructing Space?Time Variability in Forecasted Precipitation and Temperature Fields, Journal of Hydrometeorology, 5, 243-262, 10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2, 2004.

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, Climate dynamics, 52, 5451-5475, 2019.

De Michele, C. and Salvadori, G.: A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-

Copulas, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2002JD002534, 2003.

DelSole, T. and Shukla, J.: Artificial Skill due to Predictor Screening, Journal of Climate, 22, 331-345, 10.1175/2008JCLI2414.1, 2009.

Denwood, M. J.: runjags: An R Package Providing Interface Utilities, Model Templates, Parallel Computing Methods and Additional Distributions for MCMC Models in JAGS., Journal of Statistical Software, 71, 1-25, 10.18637/jss.v071.i09, 2016.

Devineni, N., Lall, U., Pederson, N., and Cook, E.: A Tree-Ring-Based Reconstruction of Delaware River Basin Streamflow Using Hierarchical Bayesian Regression, Journal of Climate, 26, 4357-4374, 10.1175/JCLI-D-11-00675.1, 2013.

Duchon, C. E.: Lanczos Filtering in One and Two Dimensions, Journal of Applied Meteorology and Climatology, 18, 1016-1022, 10.1175/1520-0450(1979)018<1016:LFIOAT>2.0.CO;2, 1979.

Eden, J. M., van Oldenborgh, G. J., Hawkins, E., and Suckling, E. B.: A global empirical system for probabilistic seasonal climate prediction, Geosci. Model Dev., 8, 3947-3973, 10.5194/gmd-8-3947-2015, 2015.

Gelman, A. and Hill, J.: Data analysis using regression and multilevel/hierarchical models, Cambridge university press2006.

Gerlitz, L., Vorogushyn, S., Apel, H., Gafurov, A., Unger-Shayesteh, K., and Merz, B.: A statistically based seasonal precipitation forecast model with automatic predictor selection and its application to central and south Asia, hydrology and earth system sciences, 20, 4605-4623, 10.5194/HESS-20-4605-2016, 2016.

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, Quarterly Journal of the Royal Meteorological Society, 146, 1999-2049, 2020.

Hsu, P.-c., Zang, Y., Zhu, Z., and Li, T.: Subseasonal-to-seasonal(S2S) prediction using the spatial-temporal projection model (STPM), Transactions of Atmospheric Sciences, 43, 212-224, 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, Climate Dynamics, 44, 1227-1244, 10.1007/s00382-014-2215-4, 2015.

Hsu, W.-r. and Murphy, A. H.: The attributes diagram A geometrical framework for assessing the quality of probability forecasts, International Journal of Forecasting, 2, 285-293, 1986.

Hwang, S.-O., Schemm, J.-K. E., Barnston, A. G., and Kwon, W.-T.: Long-Lead Seasonal Forecast Skill in Far Eastern Asia Using Canonical Correlation Analysis, Journal of Climate, 14, 3005-3016, 10.1175/1520-

0442(2001)014<3005:LLSFSI>2.0.CO;2, 2001.

Jia, X., Chen, L., Ren, F., and Li, C.: Impacts of the MJO on winter rainfall and circulation in China, Advances in Atmospheric Sciences, 28, 521-533, 10.1007/s00376-010-9118-z, 2011.

Kirono, D. G., Chiew, F. H., and Kent, D. M.: Identification of best predictors for forecasting seasonal rainfall and runoff in Australia, Hydrological Processes, 24, 1237-1247, 2010.

Lang, Y., Ye, A., Gong, W., Miao, C., Di, Z., Xu, J., Liu, Y., Luo, L., and Duan, Q.: Evaluating Skill of Seasonal Precipitation and Temperature Predictions of NCEP CFSv2 Forecasts over 17 Hydroclimatic Regions in China, Journal of Hydrometeorology, 15, 1546-1559, 10.1175/JHM-D-13-0208.1, 2014.

Lee, J.-Y., Wang, B., Wheeler, M. C., Fu, X., Waliser, D. E., and Kang, I.-S.: Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region, Climate Dynamics, 40, 493-509, 10.1007/s00382-012-1544-4, 2013.

Lepore, C., Tippett, M. K., and Allen, J. T.: ENSO-based probabilistic forecasts of March–May U.S. tornado and hail activity, Geophysical Research Letters, 44, 9093-9101, https://doi.org/10.1002/2017GL074781, 2017.

Leung, J. C.-H. and Qian, W.: Monitoring the Madden–Julian oscillation with geopotential height, Climate Dynamics, 49, 1981-2006, 10.1007/s00382-016-3431-x, 2017.

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, Journal of Hydrology, 125742, 2020.

Lima, C. H. R. and Lall, U.: Hierarchical Bayesian modeling of multisite daily rainfall occurrence: Rainy season onset, peak, and end, Water resources research, 45, 2009.

Lima, C. H. R. and Lall, U.: Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow, Journal of Hydrology, 383, 307-318, 2010.

Liu, P., Zhang, Q., Zhang, C., Zhu, Y., Khairoutdinov, M., Kim, H.-M., Schumacher, C., and Zhang, M.: A Revised Real-Time Multivariate MJO Index %J Monthly Weather Review, 144, 627-642, 10.1175/mwr-d-15-0237.1, 2016.

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

Lü, A., Jia, S., Zhu, W., Yan, H., Duan, S., and Yao, Z.: El Niño-Southern Oscillation and water resources in the headwaters region of the Yellow River: links and potential for forecasting, Hydrology and Earth System

Sciences, 15, 1273-1281, 2011.

Madden, R. A. and Julian, P. R.: Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific, Journal of Atmospheric Sciences, 28, 702-708, 10.1175/1520-0469(1971)028<0702:Doadoi>2.0.Co;2, 1971.

Madden, R. A. and Julian, P. R.: Description of Global-Scale Circulation Cells in the Tropics with a 40–50 Day Period, Journal of Atmospheric Sciences, 29, 1109-1123, 10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2, 1972.

McInerney, D., Kavetski, D., Thyer, M., Lerat, J., and Kuczera, G.: Benefits of Explicit Treatment of Zero Flows in Probabilistic Hydrological Modeling of Ephemeral Catchments, Water Resources Research, 55, 11035-11060, https://doi.org/10.1029/2018WR024148, 2019.

Mekanik, F., Imteaz, M., Gato-Trinidad, S., and Elmahdi, A.: Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes, Journal of Hydrology, 503, 11-21, 2013.

Michaelsen, J.: Cross-Validation in Statistical Climate Forecast Models, Journal of Applied Meteorology and Climatology, 26, 1589-1600, 10.1175/1520-0450(1987)026<1589:CVISCF>2.0.CO;2, 1987.

Ouyang, Y. and Liu, F.: Intraseasonal variability of summer monsoon rainfall over the lower reaches of the Yangtze River basin, Atmospheric and Oceanic Science Letters, 13, 323-329, 10.1080/16742834.2020.1741322, 2020.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci., 11, 1633-1644, 10.5194/hess-11-1633-2007, 2007.

Pegion, K., Kirtman, B. P., Becker, E., Collins, D. C., LaJoie, E., Burgman, R., Bell, R., DelSole, T., Min, D., and Zhu, Y.: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment, Bulletin of the American Meteorological Society, 100, 2043-2060, 2019.

Peng, Z., Wang, Q. J., Bennett, J. C., Pokhrel, P., and Wang, Z.: Seasonal precipitation forecasts over China using monthly large-scale oceanic-atmospheric indices, journal of hydrology, 519, 792-802, 10.1016/J.JHYDROL.2014.08.012, 2014.

Ren, H. and Shen, Y.: A New Look at Impacts of MJO on Weather and Climate in China (in Chinese), Advances in Meteorological Science and Technology, 6, 97-105, 2016.

Renard, B.: A Bayesian hierarchical approach to regional frequency analysis, Water Resources Research, 47, https://doi.org/10.1029/2010WR010089, 2011.

Reza Najafi, M. and Moradkhani, H.: Analysis of runoff extremes using spatial hierarchical Bayesian modeling,

Water Resources Research, 49, 6656-6670, https://doi.org/10.1002/wrcr.20381, 2013.

Robertson, A. and Vitart, F.: Sub-seasonal to Seasonal Prediction: The Gap Between Weather and Climate Forecasting, Elsevier2018.

Roman, S., Thordis, L. T., and Tilmann, G.: Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling, Statistical Science, 28, 616-640, 10.1214/13-STS443, 2013.

Schepen, A. and Wang, Q. J.: Toward Accurate and Reliable Forecasts of Australian Seasonal Rainfall by Calibrating and Merging Multiple Coupled GCMs, Monthly Weather Review, 141, 4554-4563, 10.1175/MWR-D-12-00253.1, 2013.

Schepen, A., Wang, Q. J., and Robertson, D.: Evidence for Using Lagged Climate Indices to Forecast Australian Seasonal Rainfall, Journal of Climate, 25, 1230-1246, 10.1175/jcli-d-11-00156.1, 2012.

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

Schepen, A., Zhao, T., Wang, Q. J., and Robertson, D. E.: A Bayesian modelling method for post-processing daily sub-seasonal to seasonal rainfall forecasts from global climate models and evaluation for 12 Australian catchments, Hydrology Earth System Sciences, 22, 1615-1628, 2018.

Sohrabi, S., Brissette, F. P., and Arsenault, R.: Coupling large-scale climate indices with a stochastic weather generator to improve long-term streamflow forecasts in a Canadian watershed, Journal of Hydrology, 594, 125925, https://doi.org/10.1016/j.jhydrol.2020.125925, 2021.

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

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, Monthly Weather Review, 147, 607-625, 10.1175/MWR-D-18-0156.1, 2019.

Totz, S., Tziperman, E., Coumou, D., Pfeiffer, K., and Cohen, J.: Winter Precipitation Forecast in the European and Mediterranean Regions Using Cluster Analysis, Geophysical Research Letters, 44, 12,418-412,426, https://doi.org/10.1002/2017GL075674, 2017.

Tuel, A. and Eltahir, E. A. B.: Seasonal Precipitation Forecast Over Morocco, Water Resources Research, 54, 9118-9130, https://doi.org/10.1029/2018WR022984, 2018.

Vehtari, A. and Lampinen, J.: Bayesian Model Assessment and Comparison Using Cross-Validation Predictive Densities, Neural Computation, 14, 2439-2468, 10.1162/08997660260293292, 2002.

Ventrice, M. J., Wheeler, M. C., Hendon, H. H., Schreck, C. J., Thorncroft, C. D., and Kiladis, G. N.: A Modified Multivariate Madden–Julian Oscillation Index Using Velocity Potential %J Monthly Weather Review, 141, 4197-4210, 10.1175/mwr-d-12-00327.1, 2013.

Vernieuwe, H., Vandenberghe, S., De Baets, B., and Verhoest, N. E. C.: A continuous rainfall model based on vine copulas, Hydrol. Earth Syst. Sci., 19, 2685-2699, 10.5194/hess-19-2685-2015, 2015.

Vigaud, N., Tippett, M. K., Yuan, J., Robertson, A. W., and Acharya, N.: Spatial Correction of Multimodel Ensemble Subseasonal Precipitation Forecasts over North America Using Local Laplacian Eigenfunctions, Monthly Weather Review, 148, 523-539, 10.1175/MWR-D-19-0134.1, 2020.

Vitart, F. and Robertson, A. W.: The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events, npj Climate Atmospheric Science, 1, 1-7, 2018.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., and Fuentes, M.: The subseasonal to seasonal (S2S) prediction project database, Bulletin of the American Meteorological Society, 98, 163-173, https://doi.org/10.1175/BAMS-D-16-0017.1, 2017.

Vitart, F., Robertson, A., Kumar, A., Hendon, H., Takaya, Y., Lin, H., Arribas, A., Lee, J., Waliser, D., and Kirtman, B.: Subseasonal to seasonal prediction: Research implementation plan, WWRP/THORPEX-WCRP Report, 2012.

Wang, B. and Xie, X.: A Model for the Boreal Summer Intraseasonal Oscillation, Journal of the Atmospheric Sciences, 54, 72-86, 10.1175/1520-0469(1997)054<0072:AMFTBS>2.0.CO;2, 1997.

Wang, M. and Duan, A.: Quasi-Biweekly Oscillation over the Tibetan Plateau and Its Link with the Asian Summer Monsoon, Journal of Climate, 28, 4921-4940, 10.1175/JCLI-D-14-00658.1, 2015.

Wang, Q. J., Robertson, D. E., and Chiew, F. H. S.: A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites, Water Resources Research, 45, https://doi.org/10.1029/2008WR007355, 2009.

Wang, Q. J., Shrestha, D. L., Robertson, D. E., and Pokhrel, P.: A log-sinh transformation for data normalization and variance stabilization, Water Resources Research, 48, https://doi.org/10.1029/2011WR010973, 2012.

Wang, Z.: Climate variability of summer rainfalls in China and the possible mechanism (in Chinese), Chinese Academy of Sciences, 2007.

Wheeler, M. C. and Hendon, H. H.: An All-Season Real-Time Multivariate MJO Index: Development of an Index

for Monitoring and Prediction, Monthly Weather Review, 132, 1917-1932, 10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2, 2004.

Woolnough, S. J.: Chapter 5 - The Madden-Julian Oscillation, in: Sub-Seasonal to Seasonal Prediction, edited by: Robertson, A. W., and Vitart, F., Elsevier, 93-117, https://doi.org/10.1016/B978-0-12-811714-9.00005-X, 2019.

Wu, Z., Xu, Z., Wang, F., He, H., Zhou, J., Wu, X., and Liu, Z.: Hydrologic Evaluation of Multi-Source Satellite Precipitation Products for the Upper Huaihe River Basin, China, Remote Sensing, 10, 10.3390/rs10060840, 2018.

Yeo, I. and Johnson, R. A.: A new family of power transformations to improve normality or symmetry, Biometrika, 87, 954-959, 2000.

Zhang, C.: Madden-Julian Oscillation, Review of Geophysics, 43, https://doi.org/10.1029/2004RG000158, 2005.

Zhang, L. and Singh, V. P.: Bivariate rainfall frequency distributions using Archimedean copulas, Journal of Hydrology, 332, 93-109, https://doi.org/10.1016/j.jhydrol.2006.06.033, 2007.

Zhang, L., Wang, B., and Zeng, Q.: Impact of the Madden–Julian Oscillation on Summer Rainfall in Southeast China, Journal of Climate, 22, 201-216, 10.1175/2008JCLI1959.1, 2009.

Zhu, Z. and Li, T.: Empirical prediction of the onset dates of South China Sea summer monsoon, Climate Dynamics, 48, 1633-1645, 10.1007/s00382-016-3164-x, 2017a.

Zhu, Z. and Li, T.: Statistical extended-range forecast of winter surface air temperature and extremely cold days over China, Quarterly Journal of the Royal Meteorological Society, 143, 1528-1538, https://doi.org/10.1002/qj.3023, 2017b.

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

Zhu, Z. and Li, T.: Extended-range forecasting of Chinese summer surface air temperature and heat waves, Climate Dynamics, 50, 2007-2021, 10.1007/s00382-017-3733-7, 2018.

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