Responses to Comments on “Sub-seasonal precipitation forecasts using preceding atmospheric intraseasonal oscillation signals in a Bayesian perspective” (Referee #1)

Anonymous Referee #1 Received and published on 28 March 2022.

Our responses are in blue and proposed manuscript revisions underlined.

General comment

The authors established a Bayesian hierarchical model (BHM) to predict the 10-60d precipitation for 17 hydroclimatic regions over China during the boreal summer monsoon season (May to October) by using the previous atmospheric intraseasonal signals. Both deterministic and probabilistic evaluations showed that the BHM provides skillful subseasonal forecasts over southeastern and southwestern hydroclimatic regions at a lead time of 20-25 days while the skills are poor over northeastern China, owing to the underestimation of intraseasonal variability.

The authors have conducted numerous calculations and employed many different statistical analysis methods. However, the explanation for their choice of the calculation and methods are deficient. Moreover, I cannot tell whether the BHM proposed in this paper show any superior skills than other statistical models or even dynamical S2S models. From this point of view, I incline to reject the manuscript, but I give an opportunity to the authors to improve the manuscript.

The authors thank the referee’s valuable comments. As introduced in the introduction section, several statistical models have been developed to generate sub-seasonal precipitation forecasts. The Spatial-Temporal Projection Model (STPM), which extracts the coupled patterns of preditors and predictand, has been widely used in recent years (Hsu et al., 2020; Zhu and Li, 2017a, b, c, 2018). The STPM1 is based on the singular value decomposition (SVD) analysis, while the STPM2 is constructed by analyzing the spatial-temporal coupled co-variance patterns between predictors and predictand (Fig. 1). A more detailed description of STPM1 and STPM2 can be found in Hsu et al. (2015).

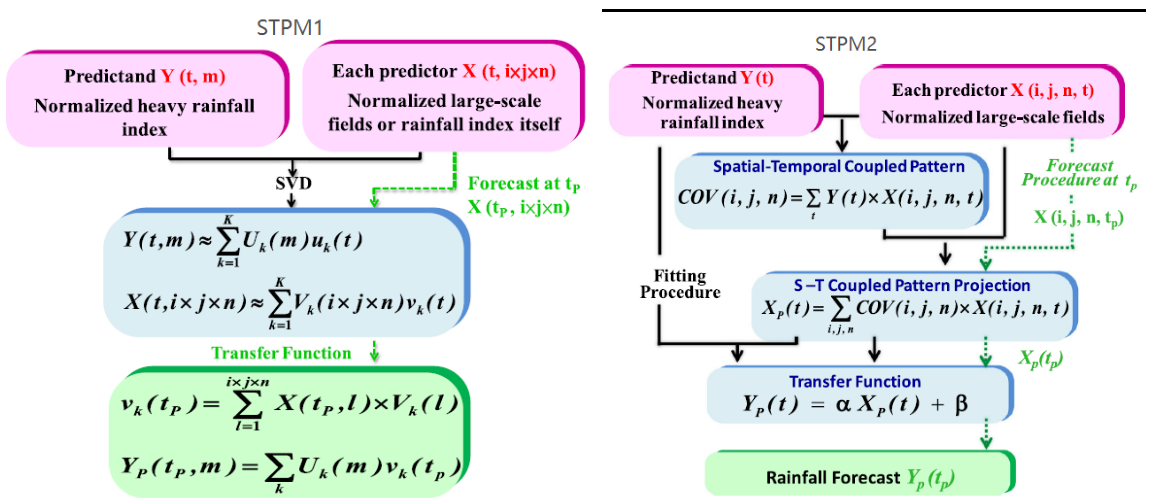

Fig. 1. Major steps of STPM1 and STPM2 prediction model (Hsu et al., 2015).

However, we should note that the uncertainty of sub-seasonal precipitation forecasts may be underestimated in STPM models. In previous studies, an optimal ensemble (OE) strategy was applied to pick up best predictors and generate probabilistic forecasts (Zhu and Li, 2017a; Zhu et al., 2015). Nevertheless, the ensemble number (number of best predictors) was always limited. Further statistical assumptions were required to interprete limited ensembles as probabilistic forecasts. Compared to OE-based probabilistic forecasts, the Bayes-theorem based statistical models are more flexible and more efficient for assessing multiple sources of uncertainties. The Bayes-theorem based models have been widely used for various aspects, and the predictive probability distributions could be generated through Markov chain Monte Carlo sampling algorithms. Thus, we will develop a **STPM2-BHM** probabilistic forecast model by taking full advantages of both STPM and Bayesian statistical modelling. We will no longer define potential predictors by averaging ISO signals in the areas of significant correlations. Instead, the predictors will be defined by extracting the coupled patterns between pentad precipitation anomalies and atmospheric intraseasonal oscillation signals, which is also known as **STPM2** in Fig. 1. The **BHM** model is then built to address the **parameter uncertainty** in the **transfer function** shown in Fig. 1.

the following changes in Sect. 2.2.2 and Sect. 2.2.3 will be made:

In section 2.2.2, we defined potential predictors by averaging U850, U200, OLR, H850, H500, H200 signals in the areas of significant correlations at different lead times. The number of potential predictors was then narrowed down using the Least Absolute Shrinkage and Selection Operator (LASSO) regression and stepwise regression approaches.

In the revised manuscript, the LASSO and stepwive regression approaches will not be used to define potential predictors. Instead, we will follow the steps of STPM2 (Fig. 1) to define coupled pattern projection coefficients as predictors. In addition, the predictand Y will be the pentad mean precipitation anomalies as suggested by the referee. We first construct spatial-temporal coupled covariance patterns (COV) where the predictand Y (10-60-day component of precipitation) and predictors X (10-60-day component of ISO signals) are significantly correlated,

$$cov(Y,X) = \frac{1}{T}\sum_{t=1}^{T}(y_t - E(y))(x_t - E(x)) \quad (1)$$

where t is the number of pentads during the training period, y_t is the ISO of precipitation, and x_t is the ISO of atmospheric fields.

The coupled pattern projection coefficient X_p , which can also be regarded as the predictor, is then obtained by multiplying the covariance and predictors X, and summing the product for each grid point where a 95 % significant level is reached,

$$X_p = cov(Y,X) \times X \quad (2)$$

Comapred to previous studies, we will build a bayesian hierarchical model to address the parameter uncertainty of the transfer function $Y_p = \alpha X_p + \beta$ shown in Fig. 1.

Here, we assume the normalized predictand Y_p follows the normal distribution,

$$Y_p \sim N(\mu_p, \sigma_p^2) \tag{3}$$

We then link the parameter μ_p with the normalized coupled pattern projection coefficient X_p using a linear model,

$$\mu_p = \alpha X_p + \beta \tag{4}$$

To complete the hierarchical formulation, we assume the unknown parameters, including σ_p, α, and β follow non-informative priors:

$$\frac{1}{\sigma_p^2} \sim \mathrm{U}(0, 100) \tag{5}$$

$$\alpha \sim N(0,\ 10^4) \tag{6}$$

$$\beta \sim N(0,\ 10^4) \tag{7}$$

The posterior distributions of these parameters will be obtained using the Markov chain Monte Carlo algorithm as well.

Fig. 2 presents the leave-one-year-out cross-validated sub-seasonal forecasts in 1981 as an example of the **STPM2-BHM** model. Here, the spatial-temporal coupled covariance patterns are derived from 10-60-day component of precipitation and U850. The predictand is the pentad mean anomalies of precipitation over Region 1 (Inland Rivers in Xinjiang). The **STPM2-BHM** model shows high forecast skills at different lead times, and prediction skills are mainly come from intraseasonal component of U850 filed.

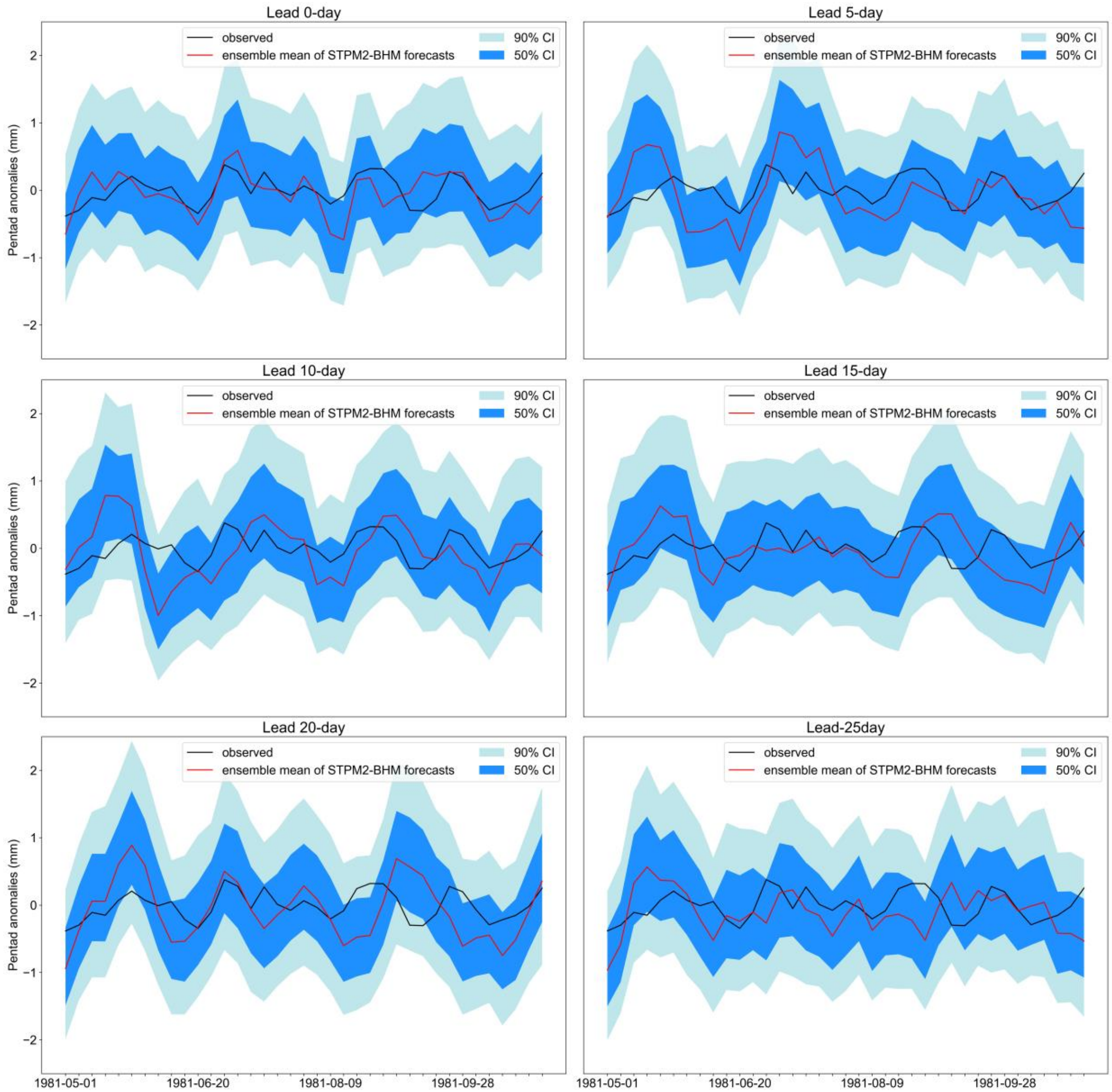

Fig. 2. Sub-seasonal forecasts of pentad mean precipitation anomalies over Region 1 (Inland Rivers in Xinjiang) during the boreal summer monsoon in 1981. The ensemble mean of STPM2-BHM forecasts are shown by the red line, observations by the black line, alongside 50% (shaded in blue) and 95% (shaded in powderblue) confidence intervals. CI = confidence interval.

We would like to incorporate the above proposals in the revised manuscript. A more detailed analysis of the results will be given to ensure that the **STPM2-BHM** model is reliable and robust for generating probabilistic forecasts of pentad mean precipitation anomalies over China.

Major comments:

1. The intraseasonal variability and the intraseasonal oscillation are different terms. The authors focus on the prediction of intraseasonal precipitation (10-60d) over China during summer (May to October). Although the selected predictors are atmospheric intraseasonal signals, no specific BSISO or MJO pattern can be found in the previous correlation maps. The title may be more consistent with the content after removing "oscillation".

Thanks for this comment. <u>We will incorporate this suggestion in the revised manuscript.</u>

2. The selected intraseasonal signals and the physical processes of their influencing on precipitation over China should be provided.

Thanks for this comment. <u>We will provide the intraseasonal signals and physical processes of their influencing on precipitation over China as supplementary file in the revised manuscript.</u>

3. For each region and each pentad from May to October, a BHM is built to forecast precipitation at different lead time. The detail information should be shown in caption of Fig.2, Fig.3, Fig.5. Are the results in these figures for a specific pentad or the average mean from May to October? If the latter is the case, will the skill for each pentad be similar throughout the whole summer?

Thanks for this comment. The results shown in Fig. 2, Fig. 3, and Fig. 5 in the manusctipt are the overall forecast skills by pooling all forecasts and observations from 1979~2016 together. We also agree that the forecast skills should be different for each pentad as the impacts of physical processes on precipitation vary at different time. Fig. 3 gives an example of the correlation coefficients between the ensemble mean of **<u>STPM2-BHM model</u>** forecasts and obervations for each pentad during the boreal summer monsoon over Region 1. Overall, the correlations show great diversity at different pentads from May to October. <u>A more comprehensive analysis would be given in the revised manuscript to address this comment.</u>

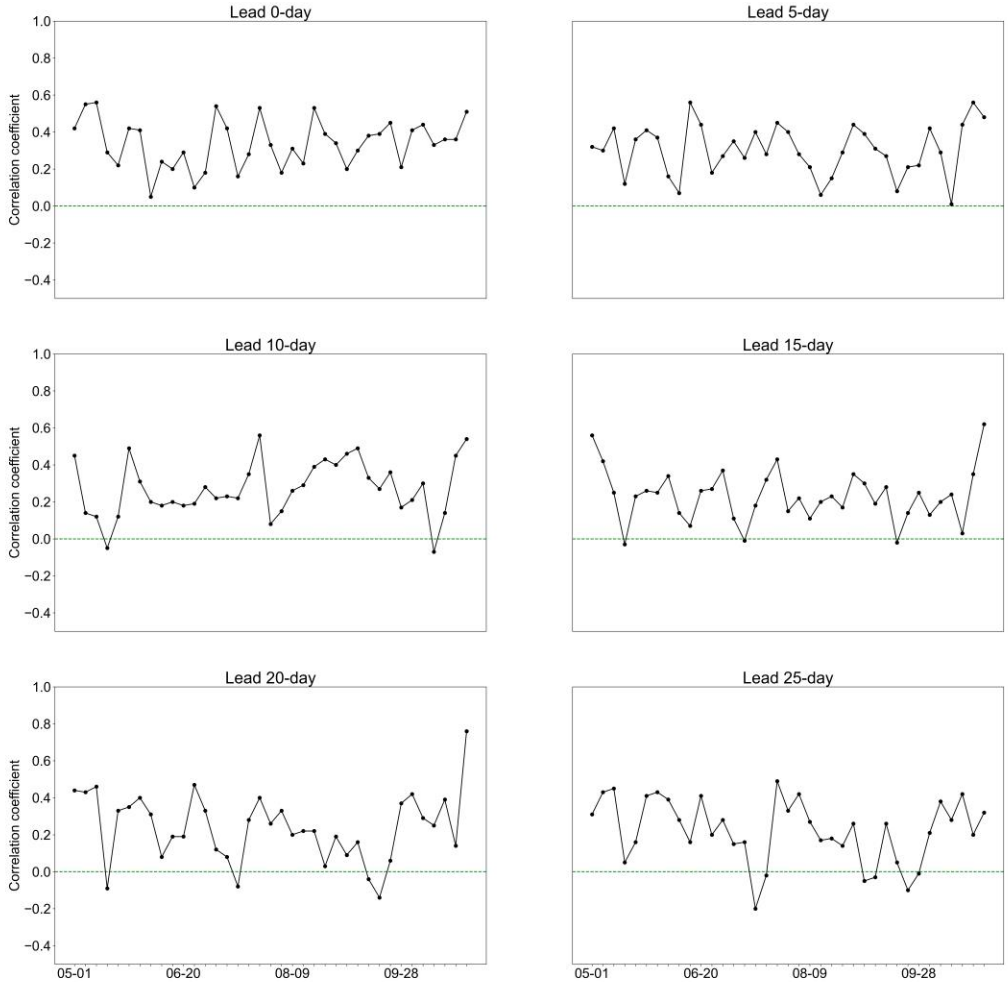

Fig. 3. Correlation coefficients between ensemble mean of STPM2-BHM model forecasts and the observations over Region 1 (Inland Rivers in Xinjiang). The predictors are obtained by analyzing the spatial-temporal coupled covariance patterns between ISO of precipitation and U850.

4. Figure 1 shows the division of the hydroclimatic regions. However, this is not a scientific way to divide China with respect to rainfall variation. Does the precipitation in each region have the coherent intraseasonal variation? If not, the correlation map is meaningless because they are calculated based on the areal-mean precipitation. Moreover, do we really need 17 regions?

The authors appreciate this suggestion. We agree that the intraseasonal variation of rainfall vary in different parts of China. Zhu and Li (2017a) used the rotated empirical orthogonal function (REOF) metod to divide the entire China into 10 sub regions as shown in Fig. 4. However, we would like to keep the division of 17 hydroclimatic regions in the revised manuscript for several reasons. We admit that the division proposed by Zhu and Li (2017a) could ensure that the precipitation in each region have coherent intraseasonal variation. However, this division may be difficult for other applications, especially for hydrological modelling purpose. In this study, the sub-seasonal precipitation forecasts for each hydroclimatic region could be potentially used as inputs of conceptual hydrologic models to generate sub-seasonal streamflow forecasts. Meanwhile, the division of 17 hydroclimatic regions is based on both watershed division standard and climate classifications. This will ensure that the climatic characteristics are nearly uniform in each region. A more detailed description of the division could be found in Lang et al. (2014). To resolve this comment in the revised manuscript, we will highlight the reason why we choose to divide China into 17 hydroclimatic regions in Sect. 2 Data and Methodology.

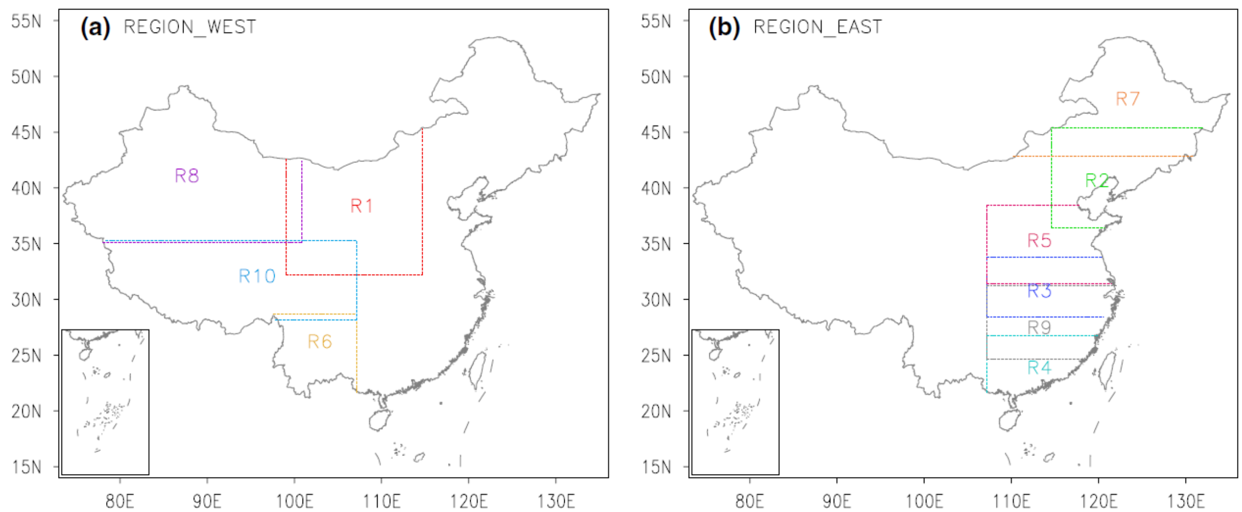

Fig. 4. The division of China based on REOF of the 10–80-day summer rainfall (a) regions in west of China (b) regions in east of China (Zhu and Li, 2017a).

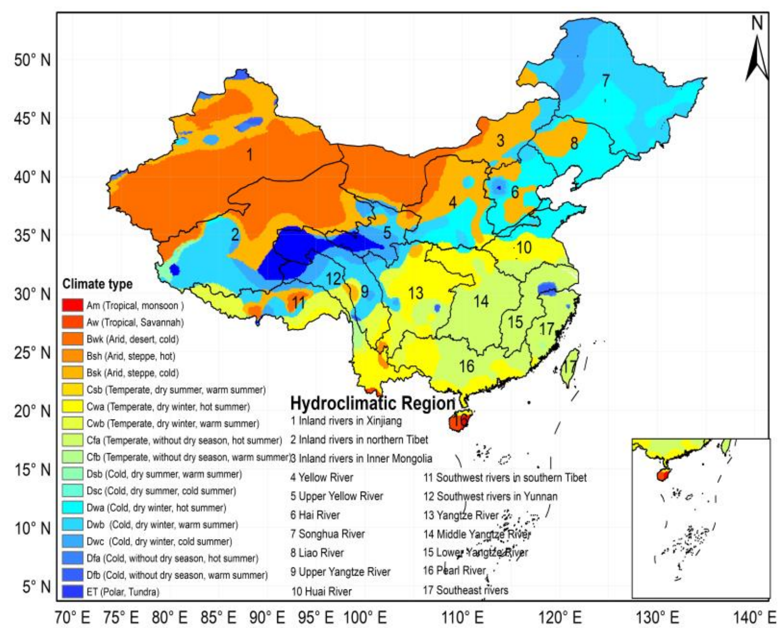

Fig. 5. 17 hydroclimatic regions over China.

1) No evidences are provided to justify the advantage of this prediction model. Does this model have better performance than the ECMWF S2S model? Or the spatial-temporal projection models (STPM)? The authors need to make some comparison.

Thanks for this comment. As mentioned above, we will build a **STPM2-BHM model** by taking full advantages of both STPM and Bayesian statistical modelling. Compared to STPM2 model introduced in previous studies, the **STPM2-BHM model** could provide predictive density functions by addressing parameter uncertainties. Thus, we would like to focus on probabilistic forecast skills. We will provide a more detailed description of the **STPM2-BHM model** in Sect. 2. Meanwhile, our previous study used the Bayesian joint probability (BJP) approach to calibrate ECMWF S2S model forecasts at different spatiotemporal scales (Li et al., 2021). The results suggested that the probabilistic forecast skills were almost zero when the lead time was beyond 10 days. To resolve this comment, we will provide more discussion on the results of **STPM2-BHM model** and our previous work.

2) From Figure 4, I can see the prediction skills mainly came from the annual cycle (which is quite stable), rather than the pentad variation (the intraseasonal component). How about the skills if only anomaly precipitation is verified? I think the skill is very limited.

Thanks for this comment. We will revise the method section as mentioned above, and the pentad mean precipitation anomalies is defined as the predictand Y in the **STPM2-BHM model**. The results shown in Fig. 2 and Fig. 3 indicate that the **STPM2-BHM model** could provide skilful sub-seasonal precipitation forecasts. A more comprehensive evaluation will be given to ensure that the newly built **STPM2-BHM model** is skilful for forecasting sub-seasonal precipitation anomalies.

Minor comments:

1. Page 1 Line 9, ".." is ".".

Thanks for this comment. We will incorporate this suggestion in the revised manuscript.

2. Page 9 Line 208, the order of Fig. 2 is confusing for the reader to discern the evolution of intraseasonal atmospheric signals from Lead 25d to Lead 0d. The figure can be sliced to two figures with the first one showing the correlation between preceding U850, U200, OLR and 10-60d precipitation from Lead 25d to Lead 0d, and the second one showing the remaining H850, H500, and H200.

Thanks for this comment. We will incorporate this suggestion in the revised manuscript (Fig. 6 and Fig. 7).

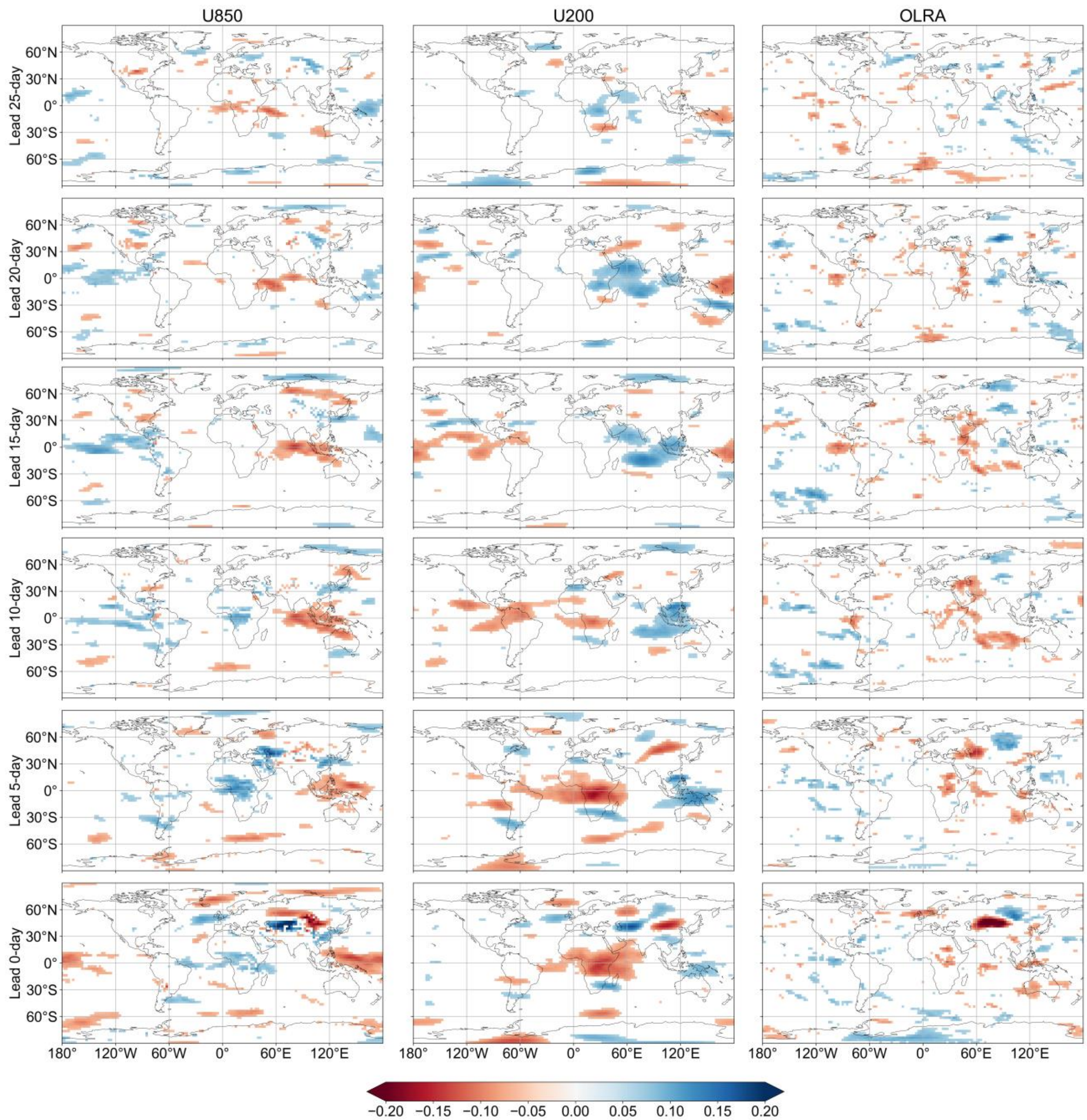

Fig. 6. Correlation between preceding ISO signals of U850, U200, OLRA and 10-60-day component of precipitation over Region 1 (Inland Rivers in Xinjiang) at different lead times. Correlation coefficients statistically significant at the 5% level are shaded.

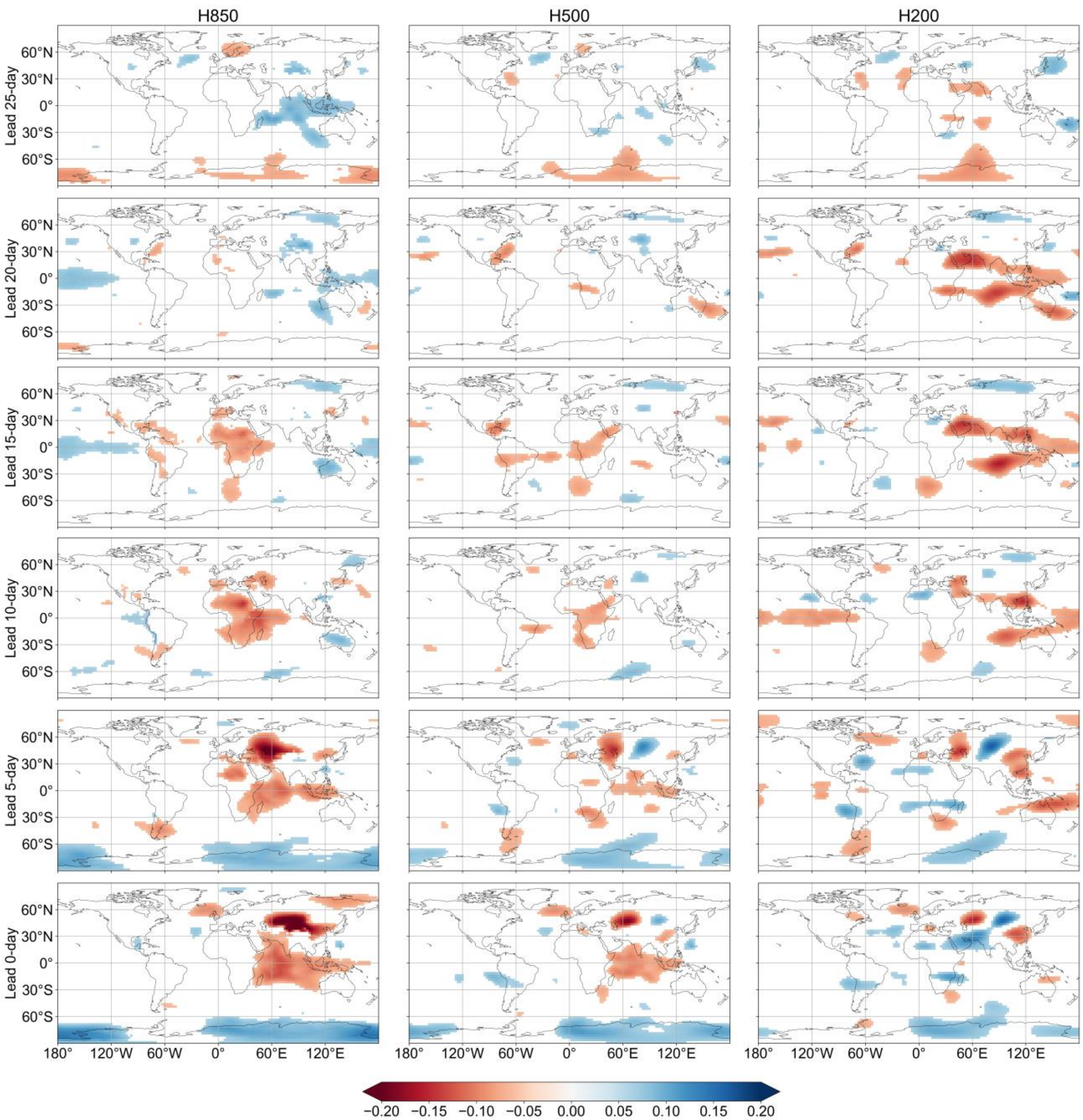

Fig. 7. Same as Fig. 6, but for H850, H500, and H200.

3. Page 15 Line 355, Fig. 3. The skill of Kling-Gupta Efficiency (KGE) in region 2, region 9 and region 12 increases with time, why? Could you please show r, β and γ before you show KGE? Because correlation coefficient and bias are the basic metric for forecast verification.

Thanks for this comment. The KGE values increase with lead time over Region 2, Region 9, and Region 12 is mostly owing to the potential predictors we defined in Sect. 2.2.2. Averaging U850, U200, OLR, H850, H500, H200 signals in the areas of significant correlations may lose some useful signals. <u>In the revised manuscript, we will define the predictors by multiplying the co-variance field and each</u>

predictor and sum the product for each grid point (at each lag) where a 95% significant level is reached as the STPM2. Meanwhile, the pentad mean precipitation anomalies will be treated as the predictand Y. The KGE may not be suitable for evaluating the forecast accuracy as the cross-validated observation mean of anomalies is nearly zero. We will provide correlation coefficient and bias of the ensemble mean of STP2-BHM model forecasts instead of the KGE in the revised manuscript.

4. Page 15 Line 355, Fig. 3. The prediction skill (KGE) of region 1 is the best in 17 regions, but in Fig. 4, the BHM model shows no skills for extreme events. Please explain the reason.
Thanks for this comment. The low forecast skills of the BHM model for extreme events is mainly owing to the definition of potential predictors. Instead, we will built a **STPM2-BHM model** in the revised manuscript. The predictors are defined by the coupled pattern projection coefficient. Fig. 2 suggests that the newly built STPM2-BHM model is capable of predicting extreme events. To resolve this comment, ranked probabilistic forecast skills at different percentiles will also be given to evaluate the forecast skills of **STPM2-BHM model** for extreme events.

5.Page 16 Line 365. What is the standard of efficient prediction in KGE and Continuous Ranked Probability Score (CRPS)? In the paper, the authors use “0.2” and “positive” as the standards, what is the reason?
Thanks for this comment. Positive KGE values are always used as indicative 'good' simulations in hydrological simulations (Knoben et al., 2019). A CRPS 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. In the revised manuscript, we will use the correlation coefficient and bias to evaluate deterministic forecast skills of the **STPM2-BHM model**, the CRPS skill score will be used to provide an overall evaluation of probabilistic forecasts, and the RPS skill scores at different percentiles will be used to evaluate the probabilistic forecast skills of extreme events.

6.Page 18 Line 385. The prediction skill over northeast China is relatively lower than that over southeastern and southwestern China. Although the number of samples will be induced, the results of southeastern and southwestern China can better demonstrate the skill of BHM.
Thanks for this comment. We will incorporate this suggestion in the revised manuscript.

7.Page 18 Line 385. There is no caption of a detail description of the size of dots.
Thanks for this comment. The size of the dots indicates the fraction of forecasts in that probability bin. We will incorporate this suggestion in the revised manuscript.

8. Line 355, during the boreal summer monsoon season.

Thanks for this comment. We will incorporate this suggestion in the revised manuscript.

9. Line 55-70, So far, there are many statistical models for subseasonal prediction (some of them were already used in operational subseasonal prediction). The authors may want to read or cite the following publications, and make comparisons with their model.

Zhu Z., T. Li, P.-C. Hsu, J. He, 2015: A spatial-temporal projection model for extended-range forecast in the tropics. Clim. Dyn., 45(3), 1085-1098. doi: 10.1007/s00382-014-2353-8.

Zhu Z., T. Li, 2018: Extended-range forecasting of Chinese summer surface air temperature and heat waves. Clim. Dyn., 50(5-6), 2007-2021. doi: 10.1007/s00382-017-3733-7.

Zhu Z., T. Li, 2017: The statistical extended-range (10–30-day) forecast of summer rainfall anomalies over the entire China. Clim. Dyn., 48(1), 209-224. doi: 10.1007/s00382-016-3070-2.

Zhu Z., T. Li, 2017: Empirical prediction of the onset dates of South China Sea summer monsoon. Clim. Dyn., 48(5), 1633-1645. doi: 10.1007/s00382-016-3164-x.

Zhu Z., T. Li, 2017: Statistical extended-range forecast of winter surface air temperature and extremely cold days over China. Q. J. R. Meteor. Soc., 704(143), 1528-1538. doi: 10.1002/qj.3023.

Zhu Z., S. Chen, K. Yuan, Y. Chen, S. Gao, Z. Hua, 2017: Empirical subseasonal predicting summer rainfall anomalies over the middle and lower reaches of Yangtze River basin based on the atmospheric intraseasonal oscillation. Atmos., 8(10), 185. doi:10.3390/atmos8100185.

Zhu Z., T. Li, L. Bai, J. Gao, 2017: Extended-range forecast for the temporal distribution of clustering tropical cyclogenesis over the western North Pacific. Theor. Appl. Climatol., 130(3), 865-877. doi: 10.1007/s00704-016-1925-4.

Li W., P. Hsu, J. He, Z. Zhu, W. Zhang, 2016: Extended-range forecast of spring rainfall in southern China based on the Madden–Julian Oscillation. Meteorol. Atmos. Phys., 128(3), 331-345. doi: 10.1007/s00703-015-0418-9.

Thanks for this comment. We will revise the introduction section to have a more detailed description of recent progresses on subseasonal predictions.

10. Line 75-80, “However, we should note that the relationships between ISO signals and precipitation are of high uncertainty for different regions at different lead times”

Yes, that is why in Zhu and Li (2017), they used REOF to divided the mainland China into 10 subregions based on the coherent nature of the 10-90 variation in each subregion. They predicted 10-30day predictand at once because considering the whole process of intraseasonal variability with the time-varying and spatial varying information. The authors may want to read the paper via the following link:

http://dqkxxb.cnjournals.org/dqkxxb/article/abstract/20200120

Thanks for this comment. We have read this article, and a **STPM2-BHM model** will be built to take advantages of both STPM and Bayesian modelling.

Reference:

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.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol. Earth Syst. Sci., 23, 4323-4331, 10.5194/hess-23-4323-2019, 2019.

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.

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, 598, 125742, https://doi.org/10.1016/j.jhydrol.2020.125742, 2021.

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, 2017a.

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, 2017b.

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, 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.