

## Article

# Forecasting Summer Rainfall and Streamflow over the Yangtze River Valley Using Western Pacific Subtropical High Feature

Ranran He <sup>1,2</sup>, Yuanfang Chen <sup>1,\*</sup>, Qin Huang <sup>1</sup>, Wenpeng Wang <sup>1</sup> and Guofang Li <sup>1</sup>

<sup>1</sup> College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; heranran2006@163.com (R.H.); hqhhu@outlook.com (Q.H.); wangwenpeng@hhu.edu.cn (W.W.); liguofang@hhu.edu.cn (G.L.)

<sup>2</sup> College of Civil Engineering, Bengbu University, Bengbu 233030, China

\* Correspondence: chenyanfang@hhu.edu.cn

**Abstract:** The western Pacific subtropical high (WPSH) is one of the key systems affecting the summer rainfall over the Yangtze River Valley in China. In this study, the forecasting capacity of the WPSH for summer rainfall and streamflow is evaluated based on the WPSH index (WPSHI) derived from the NCEP/NCAR reanalysis dataset. It has been found that WPSHI can identify extreme flood years with a higher skill than normal wet years. Specifically, exceedance probability forecasting based on WPSHI has higher skills for higher thresholds of rainfall. For streamflow, adding WPSHI as a predictor only enhances the skill for higher thresholds of streamflow relative to models based on antecedent streamflow. Under the same framework, performances of two postprocessing approaches for dynamical forecasts, i.e., the model output statistics (MOS) approach and the reanalysis-based (RAN) approach are compared. Hindcasts from Climate Forecast System version 2 from the National Center for Environmental Prediction (CFSv2) are used to calculate WPSHI, which is used as the predictor for rainfall and streamflow. The result shows that the RAN approach performs better than the MOS approach. This study emphasizes the fact that the forecasting skill of exceedance probability would largely depend on the selected threshold of the predictand, and this fact should be noticed in future studies in the long-term forecasting field.

**Keywords:** western Pacific subtropical high; the Yangtze River Valley; model output statistics (MOS); reanalysis-based (RAN) approach



**Citation:** He, R.; Chen, Y.; Huang, Q.; Wang, W.; Li, G. Forecasting Summer Rainfall and Streamflow over the Yangtze River Valley Using Western Pacific Subtropical High Feature. *Water* **2021**, *13*, 2580. <https://doi.org/10.3390/w13182580>

Academic Editor: Fernando António Leal Pacheco

Received: 8 August 2021  
Accepted: 14 September 2021  
Published: 18 September 2021

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



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

## 1. Introduction

Managing water resources and controlling risks of flood damages largely depend on the knowledge of the future rainfall and streamflow, leading to a relatively important role of seasonal hydrological forecasting. For the data-driven method, the basic step for making seasonal forecasts is to explore empirical relationships between predictors and rainfall (streamflow). A statistical model that can be directly used for operational prediction must utilize lag relationships, i.e., the relationship between antecedent ocean–atmospheric signals and rainfall (streamflow) in the following season. This method has been frequently used in the seasonal forecasting field [1,2].

At present, postprocessing outputs from dynamical forecasting systems is another frequently used approach for seasonal rainfall forecasting [3–7]. The main reason for postprocessing is that general circulation models (GCMs) often have better skills for forecasting large scale circulations than local precipitation [5,6]. Thus, forecasted circulation variables can be treated as bridges between GCM forecasts and local rainfall [5,7,8]. Streamflow can also be forecasted based on downscaling outputs of GCMs. Specifically, there are two ways for downscaling of streamflow. The first method is to use a two-step procedure, i.e., downscaling GCM outputs to local precipitation and temperature, then using them to force a hydrological model to output streamflow [9]. Another method is to downscale general circulation variables to streamflow directly and skip the hydrological model [10–12].

Considering the postprocessing methodology mentioned above, it is fundamentally important to investigate synchronous relationships between rainfall (streamflow) and circulation variables. The advantage of utilization of synchronous relationships is that their physical mechanism is relatively clearer than lag relationships. In China, it is well known that the western Pacific subtropical high (WPSH) is one of the most important circulation systems affecting summer monsoon rainfall. The spatial distribution of summer rainfall over China largely depends on the location and intensity of the WPSH. When the WPSH extends southwestward, flood often occurs in the Yangtze River Valley, and the summer rain band often locates more southern [13,14]. This mechanism can explain the extreme flood years, such as 1998 in South China. Accordingly, the WPSH is considered as an important factor affecting summer rainfall, and is treated as a key predictor in operational forecasting [15]. Wang, Xiang and Lee [8] have shown that WPSH has higher predictability, and also has a higher potential for seasonal forecasts for summer rainfall.

The Yangtze River Valley is the most important region in China, and the inter-annual variation of the summer monsoon leads to frequent floods in this region. Although the relationship between the WPSH and East Asian summer monsoon has been studied [5,8,14,16], some issues about the seasonal forecasting of summer rainfall and streamflow of this region are still needed to be further explored, which are the main themes of this study. The objectives of this study are summarized as follows.

The first goal of this study is to assess the forecasting skills for both summer rainfall and streamflow over the Yangtze River Valley based on the perfect knowledge about WPSH. For this task, we consider the effect of the definition of the positive event. Specifically, for a given threshold  $T$  of the predictand  $Y$  (rainfall or streamflow), the positive event can be defined as  $Y \geq T$ . In this setting, the forecasting procedure will be a binary classification problem. We view the forecasting skill as the function of  $T$ , and focus on forecasting skills corresponding to different  $T$  (in other words, different definitions of the positive event). To the best of the authors' knowledge, limited efforts have been made for understanding the relationship between the forecasting skill and the threshold  $T$ . If the characteristic of this relationship is well understood, one can define a positive event that can be forecasted with a much higher skill.

The second goal of this study is to compare different postprocessing approaches for dynamical forecasting systems. Two different postprocessing procedures for dynamical forecasts, i.e., the model output statistics (MOS) approach and the reanalysis-based (RAN) approach, are tested and compared for predicting summer rainfall and streamflow over the Yangtze River Valley. A review of the literature suggests that such comparison has not been tried in previous studies in the long-term forecasting field. The Climate Forecast System version 2 from the National Center for Environmental Prediction (CFSv2) is used in this study. The forecasted WPSH index (WPSHI) by CFSv2 is used as the predictor for forecasting summer rainfall and streamflow over the Yangtze River Valley.

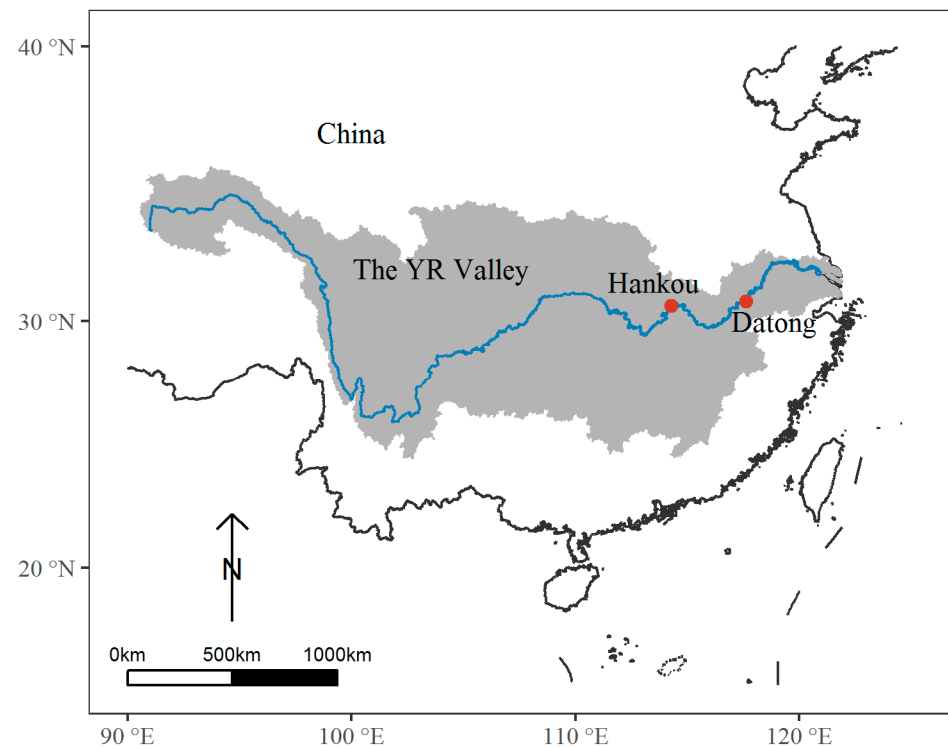
The basic technique used in this study is logistic regression, which is used for generating exceedance probability forecasts of rainfall and streamflow. Note that probability forecasting can describe uncertainty of the forecast, which is useful for decision-makers [17]. It should also be noted that summer streamflow is downscaled from WPSHI directly. This approach allows us to downscale seasonal rainfall and streamflow based under the same framework. Based on this framework, probability forecast can be applied for downscaling of streamflow. For streamflow, both antecedent streamflow and WPSHI are used as predictors, for considering both the initial state of the valley and the skill from the climate in the target season (i.e., summer).

The structure of this manuscript is organized as follows. The dataset used in this study and the definition of WPSHI are described in Section 2. In Section 3, we provide an analysis of the predictability of rainfall based on the receiver operator characteristic (ROC) analysis. Sections 4 and 5 present methods and results of a series forecasting experiments. At last, discussions and conclusions are stated in Sections 6 and 7, respectively.

## 2. Dataset Description

### 2.1. Rainfall Data and Streamflow

The NOAA's PRECipitation REConstruction over Land (PREC/L) dataset [18] is used as the observed rainfall data. Summer (June–July–August) streamflow of two stations of the main stream of the Yangtze River, i.e., Hankou station in the middle reaches and Datong station in the lower reaches, is also explored in this study. The record of Hankou and Datong used in this study covers the period of 1960–2018. The location of the Yangtze River Valley and two stations, i.e., Hankou and Datong, are shown in Figure 1.



**Figure 1.** Locations of the Yangtze River Valley (YR Valley) in China and two hydrological stations explored in this study.

### 2.2. Reanalysis Dataset and Hindcasts of CFSv2

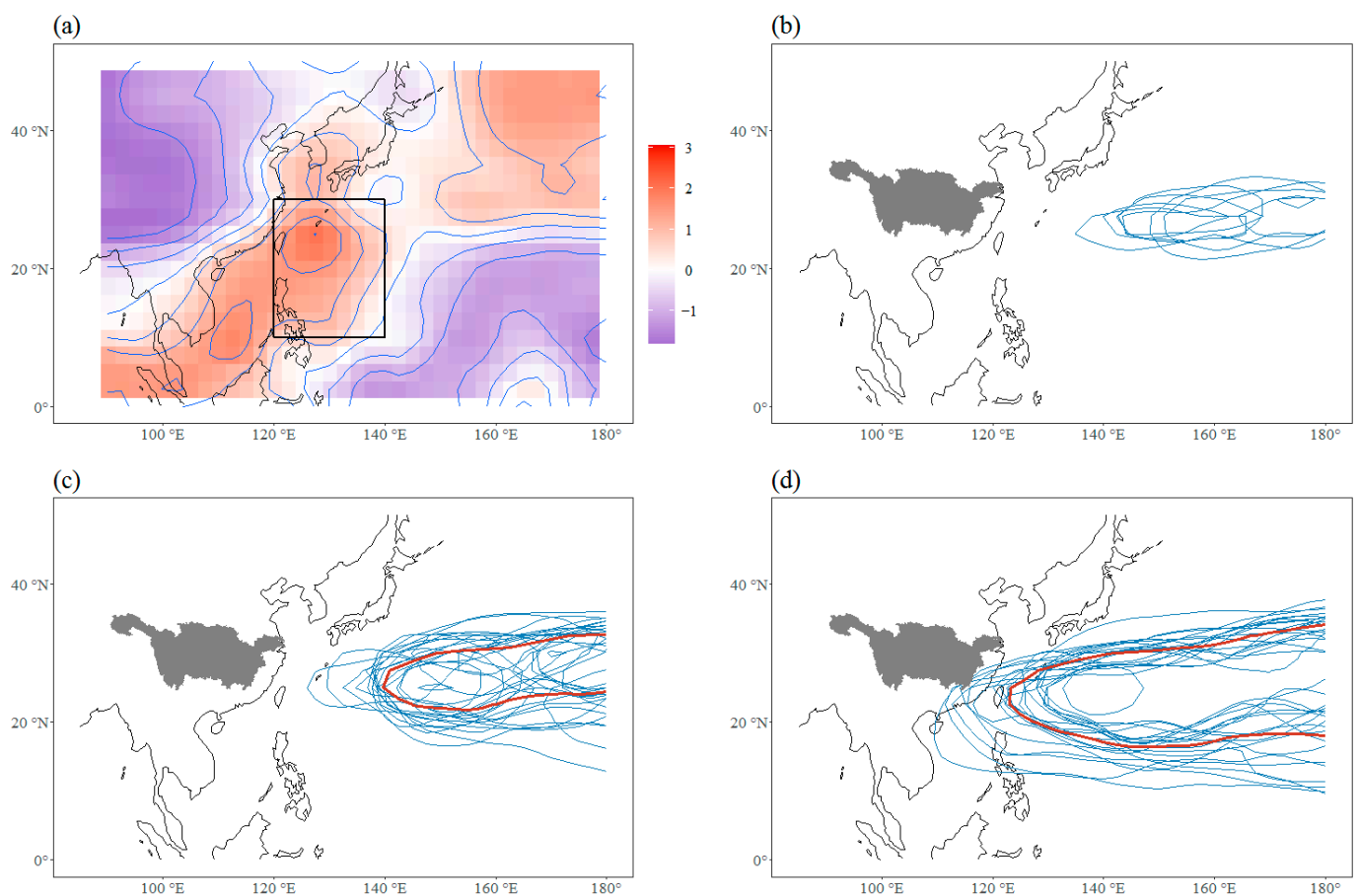
The NCEP/NCAR reanalysis dataset [19] is used in this study as the reanalysis fields of meteorological variables. The geopotential height of 500-hPa (Z500) forecasted by the CFSv2 system [20] is used as the forecasted fields. The hindcast dataset of CFSv2 from 1982 to 2010 is used in this study, and the operational forecasts from 2011 to 2018 are used to extend the hindcast dataset to cover the period of 1982–2018. The skills of CFSv2 for global and the East Asian summer monsoon have been evaluated by previous studies [21–23]. It has been found that CFSv2 can simulate many features of the East Asian monsoon system [23]. However, CFSv2 often underestimates the intensity of the monsoon system, which is true for both the Southern Asian monsoon and the East Asian monsoon [23].

For the middle time of each month, there are 24 members of forecasts released, which is initiated at successive five days from the previous month (after 7th) to the current month. The ensemble of these 24 models is used in this study. The forecasts with the released dates in February, March, April, and May are selected, and the corresponding leading times are 4 months, 3 months, 2 months, and 1 month, respectively.

### 2.3. Definition of WPSHI

The starting point of our analysis is to define a western Pacific subtropical high index (WPSHI) reflecting the characteristic of WPSH. The Z500 fields from CFSv2 and

NCEP/NCAR reanalysis dataset are used to calculate the time series of WPSHI. The definition of WPSHI used in this study is proposed by Sui et al. [24], who used the area mean value of Z500 in JJA within the region (120° E–140° E, 10° N–30° N) for constructing the WPSHI. This region corresponds to the largest variability of Z500 at the western North Pacific. Figure 2a shows the standardized deviation of the Z500 field of the reanalysis data. The calculation procedure is as follows. Firstly, the 1-order difference operator is applied on the time series data for each grid to remove the low frequency change and only retain the inter-annual component. Then, the standardized deviation is calculated for each grid. At last, the standardized deviation is normalized based on the zonal mean and zonal standardized deviation values. Although based on different procedure and different time range, the region of the largest variability shown in Figure 2 is similar with the result shown in Sui et al. [24].



**Figure 2.** (a) The standard deviation (SD) of the Z500 field, and the region for defining WPSHI. The SD has been standardized, and the details can be seen in the text; (b) the contour lines of the 5880 geopotential metre (gpm) of the 500-hPa geopotential height field, which indicates the location of the WPSH for years with WPSHI < −0.5; (c) the 5880 gpm lines for WPSHI between −0.5 and 0.5; (d) the 5880 gpm lines for WPSHI > 0.5. For (b–d), the thin lines are the 5880 lines for each year in the corresponding grade, and the thick red line is the 5880 gpm line of the corresponding mean Z500 field.

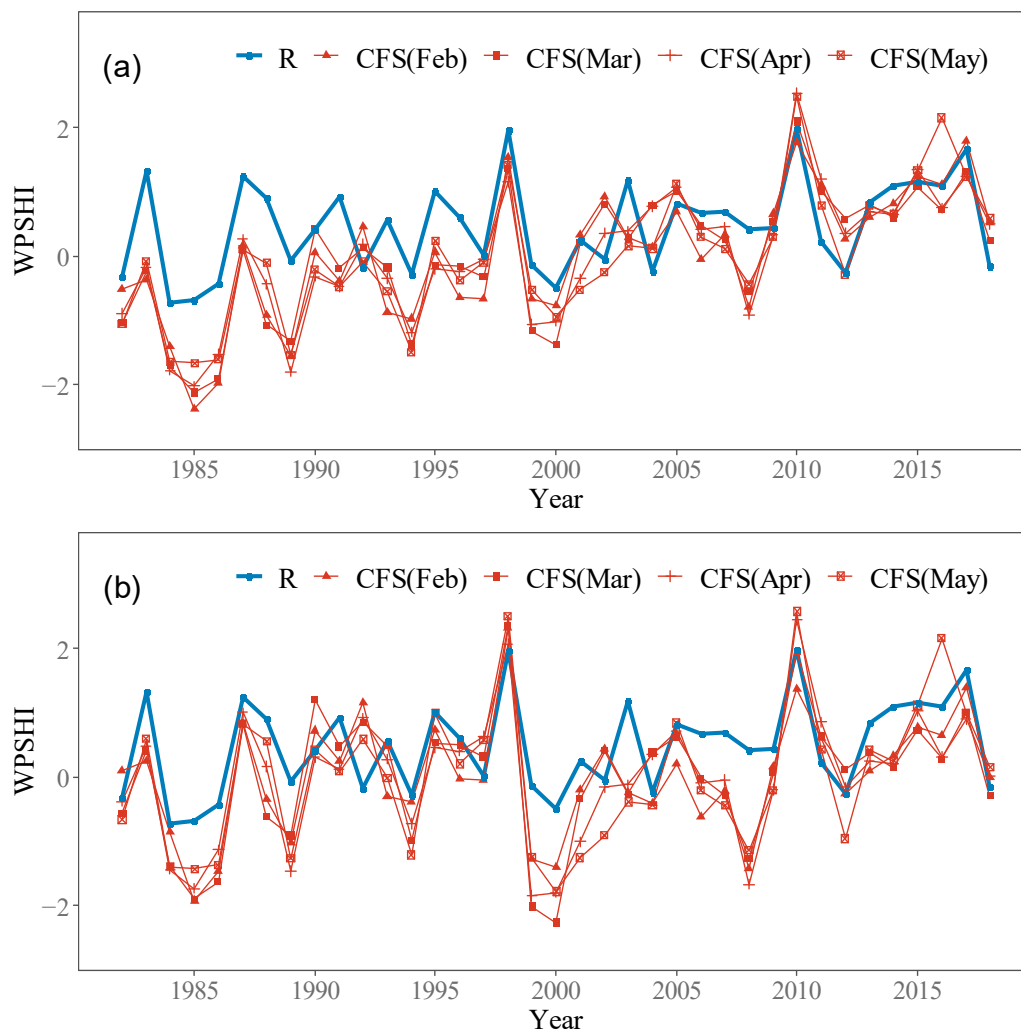
The WPSHI of the reanalysis data (denoted as WPSHI(R)) is standardized based on the mean and standard deviation of the whole period 1960–2018, based on the following equation:

$$X_s = \frac{X - \bar{X}}{Sd} \quad (1)$$

where  $X$  is the time series needed to be standardized (here  $X$  is mean Z500 within the region (120° E–140° E, 10° N–30° N)),  $\bar{X}$  and  $Sd$  are the mean and standardized deviation

of the  $X$ , and  $X_5$  is the final standardized series. Figure 2b–d show the position of the WPSH for different ranges of WPSHI, and it can be seen that WPSHI has a good indicative capacity for the position of the WPSH.

When we standardize the WPSHI of CFSv2 (denoted as WPSHI(CFS)), the discontinuity of the bias of the CFSv2 forecast is considered. One feature of the outputs of CFSv2 that should be disposed of carefully is the abrupt change in 1999 in the CFSv2 forecast [20,25]. Kumar et al. [25] have shown that this abrupt change comes from the forecast bias for SST in the equatorial Pacific and leads to changes in other variables. Figure 3a shows the WPSHI(CFS) that is standardized based on the mean value of 1982–2018. It can be seen that the forecasting bias is not stationary. Before 1999, there is an apparent larger negative bias, which is true for all leads. Note that stationary bias does not affect the postprocessing procedure, while nonstationary bias does. Thus, the final WPSHI(CFS) is calculated by the following method. First, calculate the average value of the target zone of the Z500 field from the CFSv2 forecast; then for the period of 1982–1998 and 1999–2018, the forecasting climatology of each period are subtracted from sub-series of each period, respectively; at last, the anomaly series is divided by the standard deviation calculated by the whole period. The result of this method is shown in Figure 3b.



**Figure 3.** Comparison between WPSHI from reanalysis dataset (WPSHI(R)) and from CFSv2 (WPSHI(CFS)) with different released months. WPSHI (CFS) is standardized based on two different methods. (a) Standardized based on the mean and the standard deviation of 1982–2018. (b) For 1982–1998 and 1999–2018, the mean value of each period is subtracted respectively, then standardized by the standard deviation of the whole period, i.e., 1982–2018.



### 3. ROC Analysis

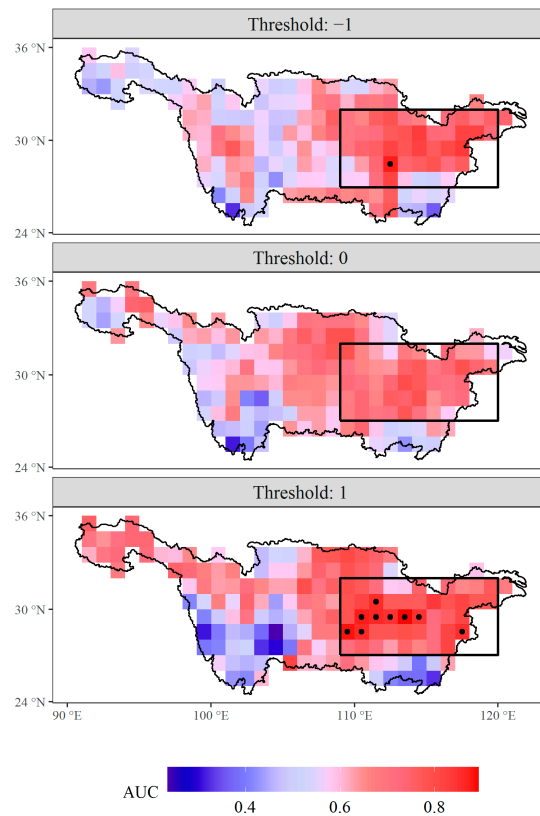
Several metrics can be used to measure the linkage between the predictor and the predictand, with the most frequently used one likely being the Pearson correlation coefficient. However, using the correlation coefficient neglects some important characteristics of the linkage between predictors and the predictand, which will be discussed here. Specifically, for rainfall of a given location, if a threshold  $T$  of rainfall is specified, we can define two classes, i.e., a positive event with the rainfall larger than  $T$ , and a negative event with the rainfall less than  $T$ . Then, we can test the capacity of the predictor to distinguish the samples from these two classes. What we want to show is that this ability can be seen as a function of the threshold  $T$ , i.e., it would change with  $T$ .

A term which is possible to be confused is the rainfall threshold. Sometimes, it is used in the field of early warning of hydro-geological disasters [26]. In this case, rainfall is the indicator for the target event. While in this paper, the rainfall threshold  $T$  is used to define the positive event, i.e., the rainfall larger than  $T$ , which is the target event identified by WPSHI.

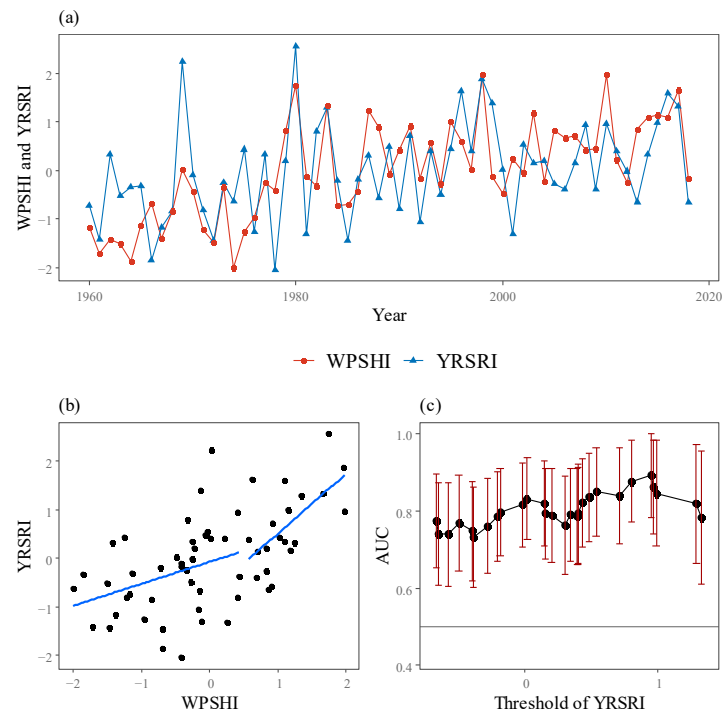
The Receiver Operator Characteristic (ROC) curve is used here for evaluating the forecasting ability. An ROC curve uses the hit rate (also known as the sensitivity) as the y coordinate, versus the false alarm rate as the x coordinate. The area under the ROC curve (AUC) can evaluate the capacity of WPSHI for discriminating between the positive and the negative event. One simple explanation of AUC could be the probability to rank a positive/negative sample pair, which is selected randomly from the sample set [27]. In this approach, building a model for generating a formal probability forecast is not needed. It is claimed that AUC should be treated as the potential skill of the predictor [28]. In this section, we use AUC to evaluate the potential skill of WPSHI for indicating the class of rainfall.

We have calculated the AUC values of WPSHI(R) for indicating class of the standardized anomaly of rainfall with three thresholds, i.e.,  $-1$ ,  $0$ , and  $1$ , respectively (Figure 4). It is clearly illustrated that AUC is higher for the threshold  $1$ . This fact enlightens us that we can find a better threshold of the predictand for which the binary classification has a higher predictability. Figure 4 also shows that the grid located at the middle and lower reaches of Yangtze River Valley have higher AUC values, indicating the higher predictability of this region. Considering this fact, in the following analysis, we define the Yangtze River Summer Rainfall Index (YRSRI) as the mean value of JJA rainfall over the box region ( $27^{\circ}$ – $32^{\circ}$  N,  $109^{\circ}$ – $120^{\circ}$  E) shown in Figure 4, which covers most of the middle and lower reaches of the Yangtze River Valley. The YRSRI is also standardized by Equation (1).

The relationship between YRSRI and WPSHI is also analyzed (Figure 5). From Figure 5a, it can be seen that WPSHI(R) and YRSRI are well correlated, and both series show the same abrupt change as the late 1970s. This upward jump of WPSHI means that the WPSH extends southwestward, leading to the wet anomaly over the Yangtze River Valley from the late 1970s [29]. Another important fact is that for extreme flood years such as 1980, 1983, and 1998, the WPSHI(R) has better indicative capacity for the YRSRI. This is also illustrated by Figure 5b, which shows the scatterplot between WPSHI(R) and YRSRI. Clearly, YRSRI only responds to WPSHI at the interval with higher WPSHI values (larger than  $0.5$ ). Specifically, for the interval of  $WPSHI < 0.5$  and  $WPSHI > 0.5$ , the Pearson correlation coefficient between WPSHI and the YRSRI is  $0.09$  and  $0.85$ , respectively. Additionally, linear regression lines are fitted for the years of  $WPSHI < 0.5$  and  $WPSHI > 0.5$ , respectively, and the slopes are  $0.45$  and  $1.23$ . AUC is calculated for different thresholds of YRSRI (Figure 5c). Still, it is important to note that the highest AUC is reached when the threshold is near  $1$ . This is consistent with the result shown in Figure 4.



**Figure 4.** AUC for each grid of rainfall over the Yangtze River Valley based on WPSHI(R) as the indicator. Grids with AUC larger than 0.85 are labelled as black points. The box region ( $27^{\circ}$ – $32^{\circ}$  N,  $109^{\circ}$ – $120^{\circ}$  E) shown in the figure is used to calculate the Yangtze River Summer Rainfall Index (YRSRI).



**Figure 5.** (a) Time series of the YRSRI and WPSHI (R), and (b) the scatter plot of these two variables. (c) The AUC values and their 95% confidence intervals corresponding to different thresholds of YRSRI.

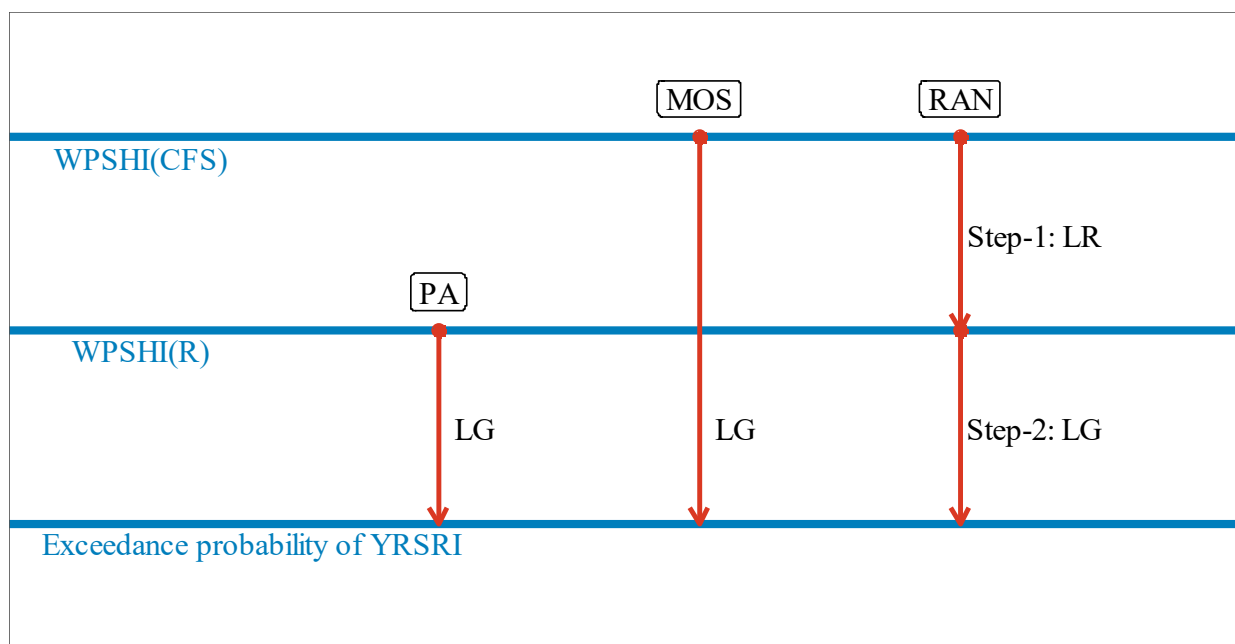
All results of this section indicate that the predictability largely depends on the threshold of YRSRI, and the response of summer rainfall to the WPSH is asymmetric and nonlinear. For larger WPSHI, i.e., when the WPSH is westward extending, the rainfall of Yangtze River Valley is more sensitive to the variation of WPSHI. This feature leads to different forecasting skills for different thresholds, and this will be investigated in the following sections.

#### 4. Modelling Methodology

Due to the limitation of the relatively small number of samples, we avoid using sophisticated models, and a simple model, i.e., logistic regression, is used as the basic tool for making probability forecasts for the binary classification. Based on logistic regression, three testing procedures based on cross validation are implemented in this study. Technical details of the three testing procedures, logistic regression, and the performance metrics are described as follows.

##### 4.1. Three Testing Procedures

In this study, three testing procedures, i.e., predictability assessment (PA), model output statistics (MOS), and the reanalysis-based (RAN) approach are explored. These approaches, except PA, have been discussed in Marzban et al. [30] in the background of weather prediction. For illustrating differences among the above three testing procedures, Figure 6 shows the corresponding schematic diagrams. The details are stated as follows.



**Figure 6.** Three test procedures explored in this study for YRSRI. In the figure, LG means logistic regression and LR means linear regression. For predicting streamflow, the procedures are similar with what have been shown in this figure, and the only difference is that the antecedent streamflow is used as another predictor for forecasting the exceedance probability of streamflow.

We first describe the procedures for predicting YRSRI. The procedures of PA and MOS are quite straightforward. For PA, the procedure builds the relationship between WPSHI(R) and YRSRI by logistic regression to forecast the exceedance probability of a given threshold. Note that PA is not the real forecast, as the reanalysis data cannot be retrieved for making operational forecasts. The result of PA reflects the predictability of the YRSRI based on the perfect knowledge of the WPSHI in the following summer. Differently from PA, MOS builds the relationship between WPSHI(CFS) and YRSRI. The advantage of MOS is that it



is unbiased when making forecasts, which is not true for the perfect prog approach. This fact makes MOS a popular method in the field of seasonal forecasts [31]. However, the disadvantage is that the number of samples for training model is lower depending on the length of hindcasts. In this study, as the hindcasts of CFSv2 are from 1982, only the years of 1982–2018 can be used for building MOS models.

The reanalysis-based (RAN) approach [30] consists of two steps for training the model and making forecasts. The first step is to build an empirical model (linear regression is used here) to map WPSHI(CFS) to WPSHI(R), and the second step is to map WPSHI(R) to the predictand. Note that the error in the first step comes from the model deficiency, while the error in the second step comes from the chaos of the climatic system. The advantage of the RAN approach is that more samples can be used to train the model of the second step, which is independent with the hindcasts of the dynamical model.

The testing procedures for streamflow are similar with that for YRSRI, and the only difference is that the antecedent streamflow is used as another predictor when forecasting JJA streamflow for PA, MOS, and the step 2 of the RAN approach.

All the tests of the above procedures are similar with the leave-one-out cross validation (LOOCV); however, the difference with the LOOCV is that the whole training years and the whole testing years are not the same in some cases. Specifically, the model is tested for each year in the test year set, i.e., 1982–2018, no matter which testing procedure is used. For testing a given year in 1982–2018, the current testing year is excluded from the training set. For PA, the training set includes the years from 1960–2018; for MOS, the training set includes the years from 1982–2018; for RAN, the training set for step 1 includes the years from 1982–2018, while for the step 2, it contains the years from 1960–2018.

#### 4.2. Logistic Regression

Logistic regression, which is a frequently used model for making probabilistic classifications, is used in this study to make class forecasts for rainfall and streamflow. An example of an application of the Logistic regression on seasonal rainfall forecast can be seen in Prasad et al. [32].

For a two class problem of a target variable  $Y$ , suppose that  $Y = 1$  means the positive class and  $Y = 0$  means the negative class, and  $p = P(Y = 1)$ , i.e., the probability of the positive class. The logistic model supposes that the logit value, i.e.,  $\log\left(\frac{p}{1-p}\right)$ , is a linear function of the predictor  $X$ :

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X \quad (2)$$

The coefficients of models can be estimated by the maximum likelihood estimation method [33]. When the coefficients have been estimated, the probability  $p$  can be calculated by:

$$p = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X)} \quad (3)$$

#### 4.3. Exceedance Probability Forecast

Exceedance probability forecasts of a given threshold are based on the logistic regression model. Here, we describe the method to generate exceedance probability forecasts of all thresholds. In the following text, the predictand  $Y$  is the YRSRI or summer streamflow:

1. The series of thresholds are selected based on the observation of the predictand  $Y$ . First, sort the  $Y$  values in the samples in 1982–2018 as the descending order  $\{Y_{[1]}, Y_{[2]}, Y_{[3]}, \dots, Y_{[n]}\}$ , where  $n$  is the number of all samples. Then, the thresholds used here are  $\{Y_{[5]}, Y_{[7]}, Y_{[8]}, \dots, Y_{[n-5]}\}$ . This setting will make at least 5 samples for the positive or negative class.
2. Choose one threshold  $T$  in step 1 and one test year in the sample set (1982–2018). All samples can be divided into two classes based on the value of the predictand  $Y$ , i.e.,

years with  $Y \geq T$  (the positive class) and years with  $Y < T$  (the negative class). Use the training set to fit a logistic regression, and then use the fitted model to forecast  $P(Y \geq T)$  for the test year.

3. Repeat step 2 for all threshold  $T$  and all test years.

Note that for the threshold  $Y_{[i]}$ , the exceedance probability of the climatology forecast is  $i/n$ .

For a larger or smaller threshold  $T$ , the sample set cut by  $T$  is not balanced, i.e., the ratio of the number of positive class and the negative class ( $n_+/n_-$ ) is not 1. In this case, how large  $p = P(Y \geq T)$  can allow us to forecast the occurrence of the positive class is a key problem. Note that using the probability output from the model, or simply choose  $p = 0.5$  as the decision threshold, will be misleading [34]. The proper selection is that when  $\frac{p}{1-p} > \frac{n_+}{n_-}$ , we make a positive class forecast. Thus, if the forecasted  $p$  is larger than the climatology forecast, the positive class will be forecasted.

It should also be noted that for a given test year, the exceedance probability  $P(Y \geq T)$  might not be monotonous decreasing as the threshold  $T$  increasing, which must be true in theory. We use the shape constrained P-splines (SCOP-splines) [35] to smooth the exceedance probability curve as monotonous decreasing. The calculation is implemented based on the R package scam.

#### 4.4. Skill Metric along the Threshold

Although neglected by other researchers, we want to show that the ability of WPSHI for discriminating the positive/negative classes of rainfall (streamflow) largely depends on the threshold. Thus, the skill for the exceedance probability forecasts is not calculated for each year (as in Piechota et al. [36]) but for each threshold of the YRSRI or streamflow.

For a given threshold  $T$ , Brier score (BS) is used to calculate skill scores. The definition of BS is:

$$BS = \frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2 \quad (4)$$

in which  $f_i$  is the  $i$ th forecast of the probability of the positive class,  $o_i$  is the observation of the  $i$ th sample (1 means positive and 0 means negative), and  $n$  is the number of the samples. The value of BS is between 0 and 1.  $BS = 0$  means perfect forecast and  $BS = 1$  corresponds to the lowest skill forecast. Note that  $f_i$  is calculated based on the model trained by the sample set excluding the  $i$  sample, as the normal leave-one-out cross validation.

Based on BS, the BS skill score (BSS) can be calculated by:

$$BSS = 1 - \frac{BS}{BS_{CLIM}} \quad (5)$$

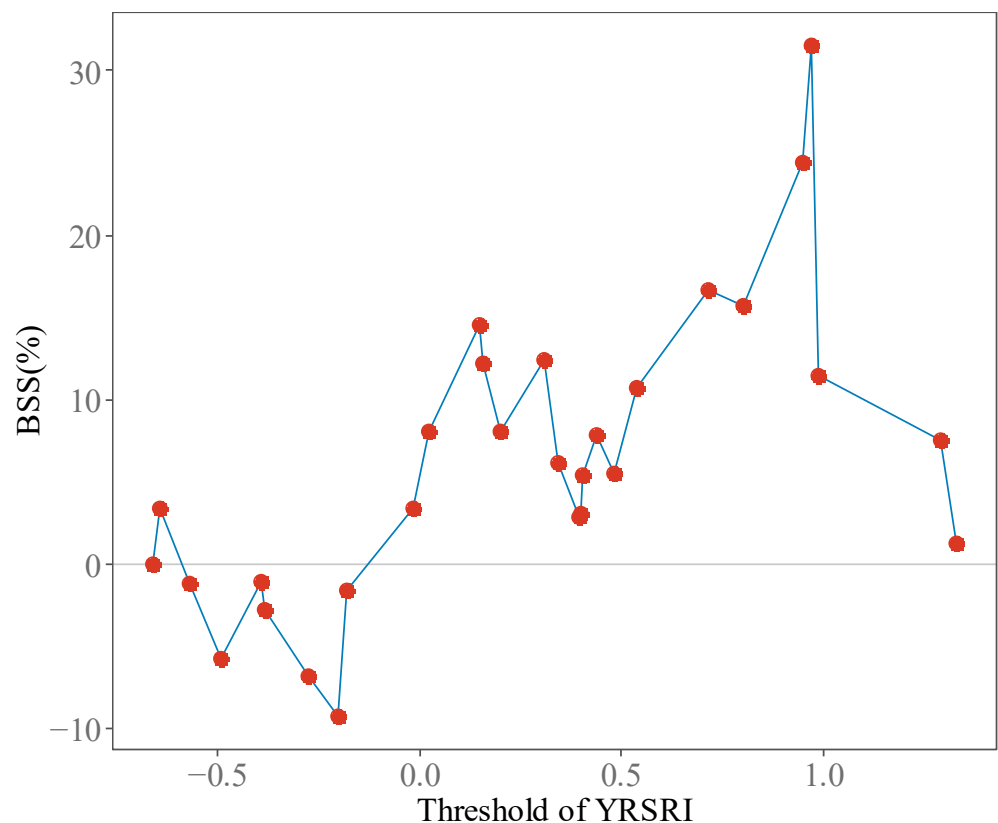
where  $BS_{CLIM}$  is the BS of the climatic forecast.

For different thresholds, BSS can be calculated respectively. Thus, we can get skill scores for different threshold  $T$ .

## 5. Results

### 5.1. Results of PA

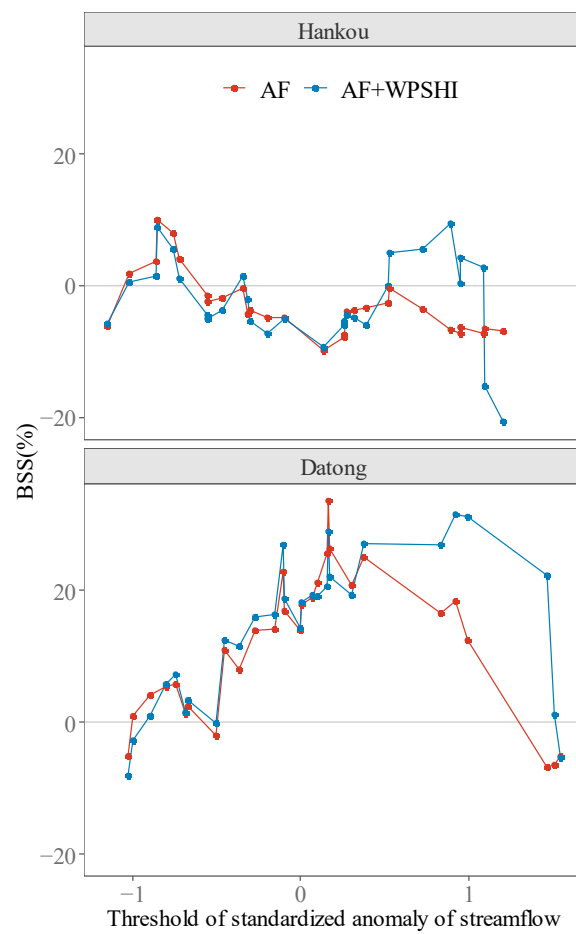
For evaluating the predictability of the YRSRI based on the WPSHI, skill is tested with WPSHI(R) as the predictor, and this is the test procedure predictability assessment (PA) that has been described in Section 4.1. Figure 7 shows the BSS of the logistic regression models corresponding to various thresholds. One important feature shown in Figure 7 is that, generally, BSS is positive only when the threshold of the YRSRI is larger than 0. Furthermore, BSS reaches the peak value ( $BSS = 31.5\%$ ) when the threshold of YRSRI is 0.97. This result is consistent with the relationship between AUC and the threshold, which has been shown in Figure 5c.



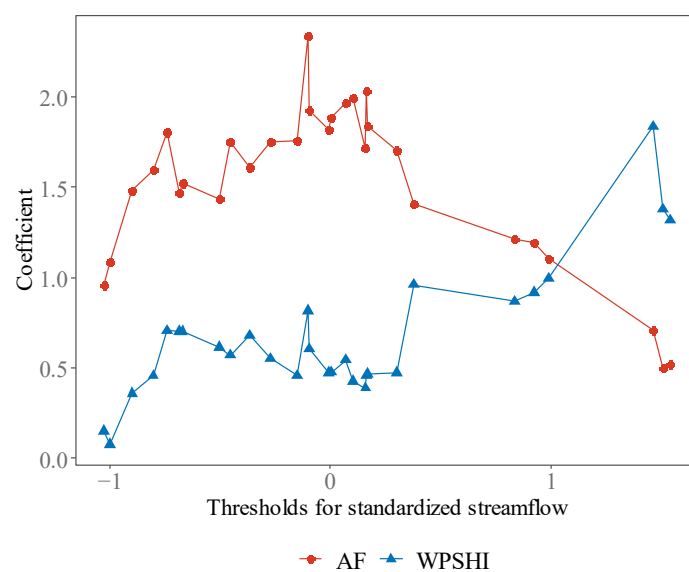
**Figure 7.** Brier skill score (BSS) of the forecast for different thresholds of the Yangtze River Valley summer rainfall index (YRSRI) based on WPSHI(R).

Not surprisingly, the above effect of threshold on rainfall will also impact the predictability of streamflow. As summer streamflow is also affected by the antecedent state of wetness of the basin, two models are built and tested for streamflow of Hankou and Datong station. The first model only uses antecedent streamflow (streamflow of May) as the predictor (denoted as AF hereafter), and the second model utilizes both antecedent streamflow and WPSHI (R) as predictors (denoted as AF + WPSHI hereafter). If the model AF + WPSHI has a higher skill than the model AF, it can thus be concluded that WPSHI provides some skill independent of the memory of the basin.

Figure 8 shows the BSS for summer streamflow of Hankou and Datong. Note that summer streamflow of Datong has relatively higher predictability than Hankou, which is reflected in the BSS of the model AF for all thresholds. The most interesting result is that, for both stations, WPSHI enhances the skill only for higher thresholds of streamflow, and this feature is clearer for Datong than Hankou. The above results are consistent with the result for the YRSRI, i.e., WPSHI shows higher skill for larger thresholds of the predictand. The skill reflected in Figure 8 can also be explained by the coefficients in the logistic regression models (Figure 9). As all predictors have been standardized, the coefficients can reflect the influence of each predictor. The pattern in Figure 9 indicates that the WPSHI plays a much dominant role in classification corresponding to higher thresholds of streamflow.



**Figure 8.** Brier skill score (BSS) of the forecast for summer streamflow of Hankou and Datong based on WPSHI(R). Two types of models are shown in the graph, i.e., the model only using antecedent streamflow (AF) and the model using both antecedent streamflow and WPSHI(R) as predictors (AF + WPSHI).

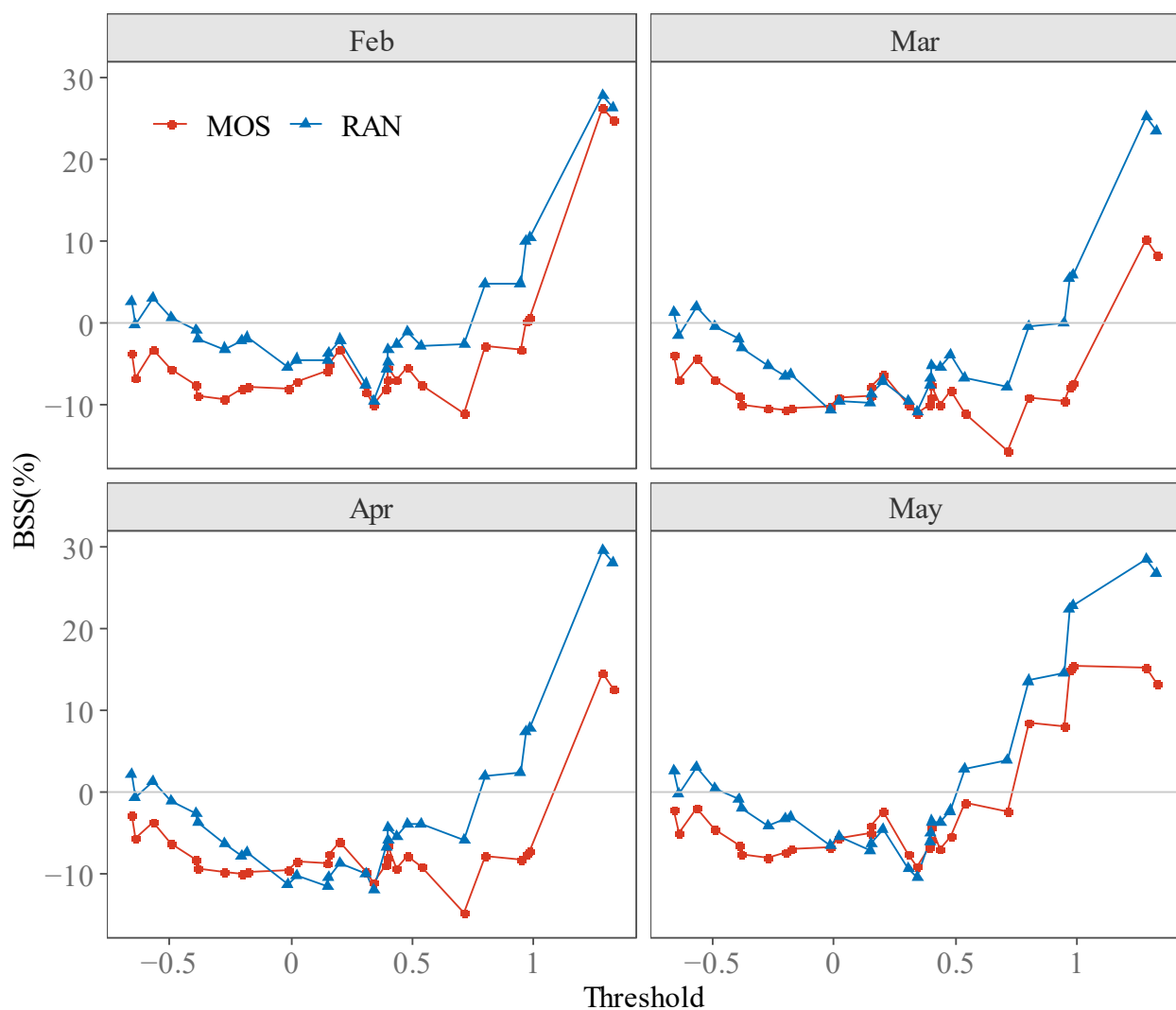


**Figure 9.** Coefficients of the logistic regression model forecasting summer streamflow of Hankou and Datong based on WPSHI(R) and antecedent streamflow (AF). The coefficients are the averaged values of models for all testing years.

### 5.2. Results of MOS and RAN

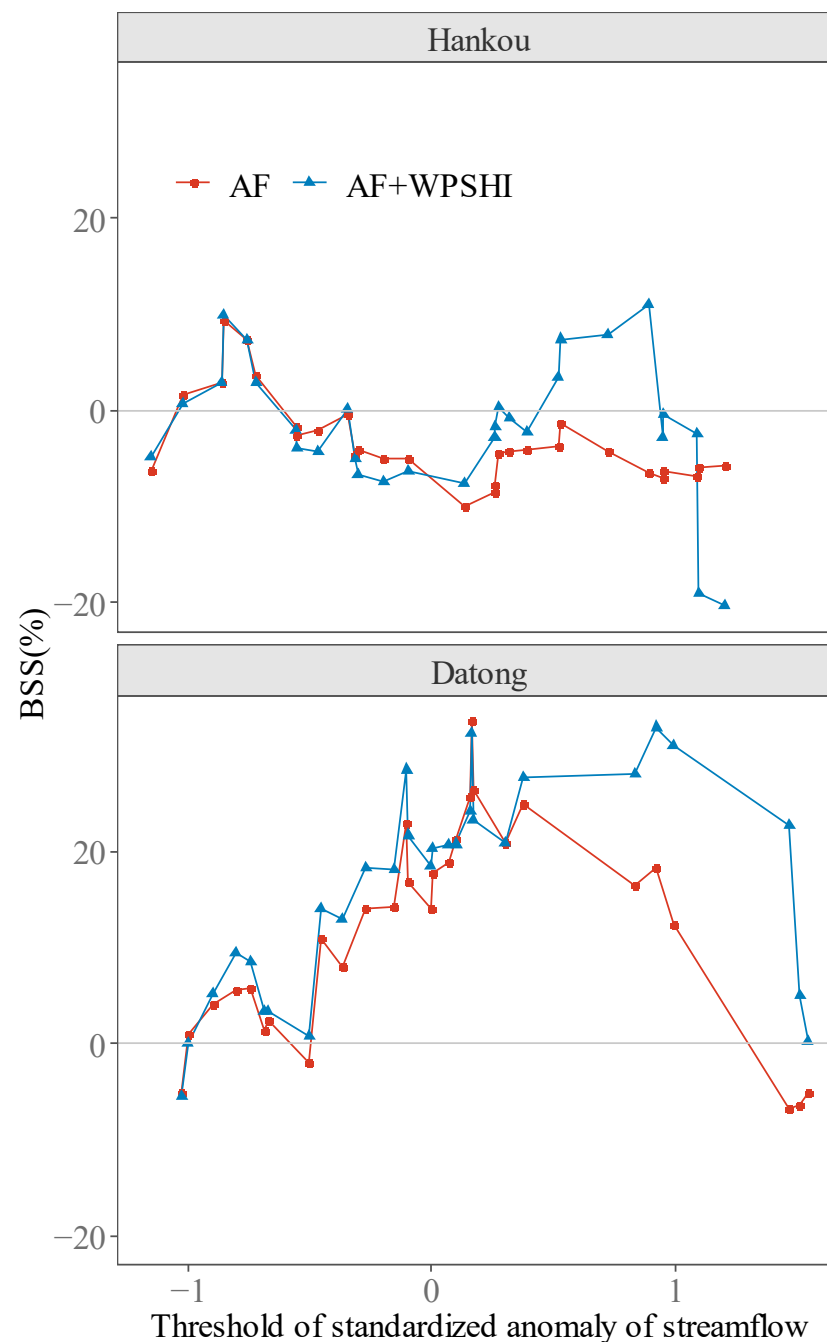
This section provides the results of two processing methods, i.e., the testing procedure MOS and RAN. As shown in Figure 6, the first step is to forecast WPSHI(R) based on WPSHI(CFS). A linear regression model is used for this task, and a leave-one-out test is used to evaluate the skill of this linear regression model. For the four releasing months (i.e., February, March, April, and May), the Nash–Sutcliffe efficiency coefficient is 0.43, 0.46, 0.48, and 0.62, respectively.

Figure 10 shows the skill scores of different leading times based on various thresholds. Note that in almost all cases, BSS of the RAN approach is larger than the MOS approach, indicating the advantage of the RAN approach. Similar to the characteristic we have shown in Figure 7, BSS is also higher for larger thresholds of YRSRI. This result indicates that WPSHI has higher skills for discriminate extreme events, especially flood summers.



**Figure 10.** BSS by the MOS approach and the RAN approach for the YRSRI corresponding to four releasing months of CFSv2.

The skill of streamflow forecasting based on the RAN approach corresponding to CFSv2 released in May is shown in Figure 11. Not surprisingly, the feature shown in Figure 8 is still obvious in Figure 11, which indicates the enhancement of the skill for larger thresholds of streamflow.



**Figure 11.** BSS by the RAN approach for summer streamflow of Hankou and Datong, corresponding to the forecast of CFSv2 released in May.

## 6. Discussion

Although sophisticated models and various predictors have been used for building seasonal forecasting models for the Yangtze River Valley, the effort to understand the roles of some key predictors based on traditional statistical methods is still quite useful, as this will lead to prediction with better interpretability. In this study, we focus on the predictive capacity of the West Pacific Subtropical High Index (WPSHI) for summer rainfall and streamflow over the Yangtze River Valley. WPSHI can be well forecasted by CFSv2, which makes WPSHI useful as a bridge for generating forecasts of rainfall and streamflow based on postprocessing of outputs of dynamical prediction systems. Thus, exploring the synchronous relationship between WPSHI and rainfall (streamflow) is beneficial to making skillful seasonal predictions.



We have demonstrated that there is a nonlinear response of summer rainfall over the Yangtze River Valley to the WPSHI, and rainfall is more sensitive to WPSHI when the value of WPSHI is higher. Because of this feature, WPSHI only shows higher skills for forecasting the exceedance probability of rainfall corresponding to larger thresholds. Similarly, for streamflow, WPSHI only enhances the skill for higher thresholds relative to the persistence forecast (i.e., the model with antecedent streamflow as the predictor). The above result means that WPSHI is only a good indicator for identifying extreme flood summers of the Yangtze River Valley. These findings allow us to select a new strategy for making long-term hydrological forecasts, i.e., to find a proper definition of the positive event based on selecting a proper threshold with higher predictability. We found that previous studies have not adequately explored exceedance probability forecasting, such as Piechota et al. [36], and this study provides a new perspective to treat probability forecasting.

With the same framework, two post-processing approaches, which have been applied in the field of weather forecasting, are also explored. We have shown that the RAN approach has a better performance than MOS. As discussed in the work of Marzban et al. [30], the main advantage of the RAN approach is that more samples can be used to training the model of step 2 as shown in Figure 6, as step 2 is independent from the dynamical model, thus the number of samples is not limited by the hindcasts available. Although many forecasting models based on downscaling technology have been explored [4,37], little effort has been applied for comparison between different postprocessing approaches. For example, for making forecasts for North China summer rainfall, Guo et al. [4] built a downscaling model based on reanalysis data, and then substituted CFSv2 forecasting values (bias-removed) to make real forecasts. In fact, this is the perfect prog (PP) approach. It is possible that the skill could be enhanced when the circulation variables are not just removed of bias but reforecasted, as what the RAN approach does. More comparisons are still needed in future studies.

This study provides a framework for generating probability forecasting for rainfall and streamflow of the Yangtze River Valley, then represents a contribution for the development of an early warning system (EWS) [38] for the study area. We have shown that the forecast can be skillful for larger thresholds of rainfall even from February. For converting probability forecasts to binary forecasts, tools such as the ROC curve are useful for making the trade-off between the benefit of hit and the cost of a false alarm, which is beyond the topic of this paper.

## 7. Conclusions

In this study, we built forecasting models for summer rainfall and streamflow over the Yangtze River Valley based on the knowledge of the western Pacific subtropical high (WPSH). Several conclusions can be listed here:

1. The rainfall over the Yangtze River Valley is more sensitive to the variability of WPSHI when WPSHI is high, while when WPSHI is less than 0.5, the rainfall shows low sensitivity. Furthermore, the middle and lower reaches of Yangtze River Valley show higher sensitivity to the variability of WPSHI than other regions. This characteristic leads to higher forecasting skill of exceedance probability forecasts corresponding to larger thresholds of rainfall.
2. The analysis of predictability of summer streamflow of the Yangtze River Valley shows that WPSHI can only enhance the forecasting skill for binary classification corresponding to larger thresholds of streamflow.
3. A comparison between two postprocessing approaches shows that the RAN approach shows a higher skill than model output statistics (MOS), as RAN can utilize more samples than MOS.
4. When building a long-term forecasting model for generating exceedance probability forecasts, one should notice the effect of the threshold, and find a proper threshold with a higher skill.

**Author Contributions:** Conceptualization, R.H., Y.C. and G.L.; formal analysis, R.H.; methodology, R.H., Y.C. and W.W.; software, R.H.; validation, Q.H., W.W. and G.L.; writing—original draft, R.H.; writing—review & editing, Y.C. and Q.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by the National Key Research and Development Program of China (2018YFC1508200, 20175018612), and the National Natural Science Foundation of China (51709075, 41701015).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

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

## References

- Shams, M.S.; Anwar, F.A.H.M.; Lamb, K.W.; Bari, M. Relating ocean-atmospheric climate indices with australian river streamflow. *J. Hydrol.* **2018**, *556*, 294–309. [[CrossRef](#)]
- Araghinejad, S.; Burn, D.; Karamouz, M. Long-lead probabilistic forecasting of streamflow using ocean-atmospheric and hydrological predictors. *Water Resour. Res.* **2006**, *42*, W03431. [[CrossRef](#)]
- Kullmann, H.; Baldauf, S.A.; Bakker, T.C.M.; Frommen, J.G. Seasonal forecast for local precipitation over northern taiwan using statistical downscaling. *J. Geophys. Res. Atmos.* **2008**, *113*, D12.
- Guo, Y.; Li, J.; Li, Y. Seasonal forecasting of north china summer rainfall using a statistical downscaling model. *J. Appl. Meteorol. Climatol.* **2014**, *53*, 1739–1749. [[CrossRef](#)]
- Lee, S.; Lee, J.; Ha, K.; Wang, B.; Schemm, J. Deficiencies and possibilities for long-lead coupled climate prediction of the western north pacific-east asian summer monsoon. *Clim. Dyn.* **2011**, *36*, 1173–1188. [[CrossRef](#)]
- Chen, H.; Sun, J.; Wang, H. A statistical downscaling model for forecasting summer rainfall in china from demeter hindcast datasets. *Weather. Forecast.* **2012**, *27*, 608–628. [[CrossRef](#)]
- Peng, Z.; Wang, Q.J.; Bennett, J.C.; Schepen, A.; Pappenberger, F.; Pokhrel, P.; Wang, Z. Statistical calibration and bridging of ecmwf system4 outputs for forecasting seasonal precipitation over china. *J. Geophys. Res. Atmos.* **2014**, *119*, 7116–7135. [[CrossRef](#)]
- Wang, B.; Xiang, B.; Lee, J.-Y. Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2718–2722. [[CrossRef](#)]
- Yuan, X.; Wood, E.F. Downscaling precipitation or bias-correcting streamflow? Some implications for coupled general circulation model (cgcm)-based ensemble seasonal hydrologic forecast. *Water Resour. Res.* **2012**, *48*, 12519. [[CrossRef](#)]
- Tisseuil, C.; Vrac, M.; Lek, S.; Wade, A.J. Statistical downscaling of river flows. *J. Hydrology* **2010**, *385*, 279–291. [[CrossRef](#)]
- Sachindra, D.; Huang, F.; Barton, A.; Perera, B. Least square support vector and multi-linear regression for statistically downscaling general circulation model outputs to catchment streamflows. *Int. J. Climatol.* **2013**, *33*, 1087–1106. [[CrossRef](#)]
- Ghosh, S.; Mujumdar, P.P. Statistical downscaling of gcm simulations to streamflow using relevance vector machine. *Adv. Water Resour.* **2008**, *31*, 132–146. [[CrossRef](#)]
- Huang, R.; Chen, J.; Huang, G. Characteristics and variations of the east asian monsoon system and its impacts on climate disasters in china. *Adv. Atmos. Sci.* **2007**, *24*, 993–1023. [[CrossRef](#)]
- Luo, S.; Jin, Z. Statistical analysis for sea surface temperature over the south china sea, behavior of subtropical high over the west pacific and monthly mean rainfall over the changjiang middle and lower reaches. *Chin. J. Atmos. Sci.* **1986**, *10*, 409–418.
- Zhang, J.; Chen, X. The operational seasonal forecasting of the summer rainfall in china. *Adv. Atmos. Sci.* **1987**, *4*, 349–362. [[CrossRef](#)]
- Huang, R.; Chen, J.; Wang, L.; Lin, Z. Characteristics, processes, and causes of the spatio-temporal variabilities of the east asian monsoon system. *Adv. Atmos. Sci.* **2012**, *29*, 910–942. [[CrossRef](#)]
- Blschl, G.; Bierkens, M.; Chambel, A.; Cudennec, C.; Zhang, Y. Twenty-three unsolved problems in hydrology (uph)-A community perspective. *Hydrol. Sci. J.* **2019**, *64*, 1141–1158. [[CrossRef](#)]
- Chen, M.; Xie, P.; Janowiak, J.E.; Arkin, P.A. Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrometeorol.* **2002**, *3*, 249–266. [[CrossRef](#)]
- Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J. The ncep/ncar 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–471. [[CrossRef](#)]
- Saha, S.; Moorthi, S.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Behringer, D.; Hou, Y.T.; Chuang, H.; Iredell, M. The ncep climate forecast system version 2. *J. Clim.* **2014**, *27*, 2185–2208. [[CrossRef](#)]
- Lang, Y.; Ye, A.; Gong, W.; Miao, C.; Di, Z.; Xu, J.; Liu, Y.; Luo, L.; Duan, Q. Evaluating skill of seasonal precipitation and temperature predictions of ncep cfsv2 forecasts over 17 hydroclimatic regions in china. *J. Hydrometeorol.* **2014**, *15*, 1546–1559. [[CrossRef](#)]

22. Silva, G.A.M.; Dutra, L.M.M.; Rocha, R.P.D.; Ambrizzi, T.; Leiva, É. Preliminary analysis on the global features of the ncep cfsv2 seasonal hindcasts. *Adv. Meteorol.* **2014**, *2014*, 117–128. [[CrossRef](#)]
23. Yang, S.; Zhang, Z.; Kousky, V.E.; Higgins, R.W.; Yoo, S.H.; Liang, J.; Fan, Y. Simulations and seasonal prediction of the asian summer monsoon in the ncep climate forecast system. *J. Clim.* **2008**, *21*, 3755–3775. [[CrossRef](#)]
24. Sui, C.H.; Chung, P.H.; Li, T. Interannual and interdecadal variability of the summertime western north pacific subtropical high. *Geophys. Res. Lett.* **2007**, *34*, 93–104. [[CrossRef](#)]
25. Kumar, A.; Chen, M.; Zhang, L.; Wang, W.; Xue, Y.; Wen, C.; Marx, L.; Huang, B. An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the ncep climate forecast system (cfs) version 2. *Mon. Weather. Rev.* **2011**, *140*, 3003–3016. [[CrossRef](#)]
26. Luca, D.L.D.; Versace, P. Diversity of rainfall thresholds for early warning of hydro-geological disasters. *Adv. Geosci.* **2017**, *44*, 53–60. [[CrossRef](#)]
27. Fawcett, T. An introduction to roc analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
28. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*; Elsevier: Amsterdam, The Netherlands, 2011.
29. Gong, D.Y.; Ho, C.H. Shift in the summer rainfall over the yangtze river valley in the late 1970s. *Geophys. Res. Lett.* **2002**, *29*. [[CrossRef](#)]
30. Marzban, C.; Sandgathe, S.; Kalnay, E. Mos, perfect prog, and reanalysis. *Mon. Weather. Rev.* **2006**, *134*, 657–663. [[CrossRef](#)]
31. Salimun, E.; Tangang, F.T.; Juneng, L.; Zwiers, F.W.; Merryfield, W.J. Skill evaluation of the cancm4 and its mos for seasonal rainfall forecast in malaysia during the early and late winter monsoon periods. *Int. J. Climatol.* **2016**, *36*, 439–454. [[CrossRef](#)]
32. Prasad, K.; Dash, S.K.; Mohanty, U.C. A logistic regression approach for monthly rainfall forecasts in meteorological subdivisions of india based on demeter retrospective forecasts. *Int. J. Climatol.* **2010**, *30*, 1577–1588. [[CrossRef](#)]
33. Hastie, T.; Tibshirani, R.; Friedman, J.; Hastie, T.; Friedman, J.; Tibshirani, R. *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2009; Volume 2.
34. Provost, F. Machine learning from imbalanced data sets 101. In Proceedings of the AAAI'2000 Workshop on Imbalanced Data sets, Austin, TX, USA, 31 July 2000; pp. 1–3.
35. Pya, N.; Wood, S.N. Shape constrained additive models. *Stat. Comput.* **2015**, *25*, 543–559. [[CrossRef](#)]
36. Piechota, T.C.; Chiew, F.H.; Dracup, J.A.; McMahon, T.A. Development of exceedance probability streamflow forecast. *J. Hydrol. Eng.* **2001**, *6*, 20–28. [[CrossRef](#)]
37. Liu, Y.; Fan, K.; Zhang, Y. A statistical downscaling model for summer rainfall over china stations based on the climate forecast system. *Chin. J. Atmos. Sci.* **2013**, *37*, 1287–1296.
38. Kelman, I.; Glantz, M.H. Early warning systems defined. In *Reducing Disaster: Early Warning Systems for Climate Change*; Singh, A., Zommers, Z., Eds.; Springer: Dordrecht, The Netherlands, 2014; pp. 89–108.