Water Resources Research

RESEARCH ARTICLE
10.1029/2022WR032484

Special Section:
Advancing process representation in hydrologic models: Integrating new concepts, knowledge, and data

Hydrological Model Adaptability to Rainfall Inputs of Varied Quality

Jiao Wang[1], Lu Zhuo[1,2], Dawei Han[1], Ying Liu[1], and Miguel Angel Rico-Ramirez[1]

[1]Department of Civil Engineering, University of Bristol, Bristol, UK, [2]School of Earth and Environmental Sciences, Cardiff University, Cardiff, UK

Abstract Numerous studies have evaluated the reliability and hydrologic utility of various rainfall data sets through hydrological modeling. However, the calibration of hydrological models compensates for errors in rainfall inputs. The drivers, conditions, and factors affecting the calibration of hydrological models given the accuracy of rainfall inputs are not well understood. Here, we explore hydrological model adaptability to rainfall inputs of varied quality and its potential mechanisms. Twenty-eight rainfall products from multiple sources are collected for a headwater catchment in the Southern United States. These rainfall data sets include measurements from rain gauges, weather radars, satellites, reanalysis products, and Weather Research and Forecasting model simulations. Such rainfall data sets with varied errors are used to independently calibrate a widely used conceptual Xin'anjiang (XAJ) hydrological model. Results suggest that the hydrological model can often adapt well to two scenarios of inaccurate rainfall inputs producing high-performance streamflow simulations. This adaptive ability is controlled by an adaptable threshold of the overall bias of the rainfall inputs. Moreover, hydrological model adaptability to rainfall inputs is further influenced by how event-based rainfall bias shapes the overall rainfall bias, especially from those of heavy rainstorms. The hydrological model can adapt to those rainfall inputs that contain important information content for model calibration. Notably, the adaptability to rainfall inputs of the XAJ model is mainly controlled by a bias reduction through adjustment of evapotranspiration and soil moisture storage, yielding satisfactory effective rainfall. The study quantitatively sheds new light on hydrological model adaptability to rainfall input quality.

Plain Language Summary Accurate rainfall is an ideal input for the hydrological model to simulate streamflow reliably. However, some inaccurate rainfall data sets can be compensated by hydrological model calibration to generate good streamflow. It is still unclear what kind of rainfall inputs the model calibration can and cannot adapt to. Here, we explore a large number of rainfall data sets from different sources and their corresponding streamflow simulation performance after independent model calibrations. It is found that the hydrological model can adapt well to two scenarios of inaccurate rainfall data. The adaptive ability of the hydrological model calibration to rainfall inputs is not only affected by the accuracy over the entire period but also by the accuracy of individual rainfall events, especially those of the most severe rainstorms. Therefore, a good hydrological simulation does not always mean its rainfall input data are reliable. This study enhances a quantitative understanding of how the model calibration adapts to the errors in rainfall input data. The findings are expected to provide quantitative guidance for bias correction and data fusion of rainfall products.

Key Points:

- Two scenarios of inaccurate rainfall inputs and a threshold of overall rainfall bias are identified for hydrological model adaptability
- Hydrological model adaptability is further influenced by how event-based bias of individual rainstorms shapes the overall rainfall bias
- Hydrological model adaptability is mainly controlled by a bias reduction through adjustment of evapotranspiration and soil moisture storage

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Wang,
jiao.wang@bristol.ac.uk;
jiao.wang_hydro@outlook.com

Citation:

Wang, J., Zhuo, L., Han, D., Liu, Y., & Rico-Ramirez, M. A. (2023). Hydrological model adaptability to rainfall inputs of varied quality. *Water Resources Research*, *59*, e2022WR032484. https://doi.org/10.1029/2022WR032484

Received 6 APR 2022
Accepted 13 JAN 2023

Author Contributions:

Conceptualization: Jiao Wang, Lu Zhuo, Dawei Han
Formal analysis: Jiao Wang
Funding acquisition: Jiao Wang
Investigation: Jiao Wang

1. Introduction

Hydrological models serve as useful tools for flood forecasting and water resource management (Bárdossy & Singh, 2008). Rainfall data are the primary input for hydrological models. Usually, rain gauge observations are considered the most reliable rainfall source (so-called "ground truth"). However, due to the insufficient rain gauge density, rainfall estimations from weather radars, satellites, reanalysis, and Numerical Weather Prediction (NWP) model products have also been widely applied for hydrological applications.

A plethora of studies employed hydrological modeling to evaluate the reliability of various state-of-the-art rainfall data sets in different regions. Many of them suggested that despite the varied accuracy of rainfall data compared with observations, some rainfall data sets could reproduce satisfactory streamflow simulations after model recalibration, and could even perform as well as using the observed rainfall (Beck et al., 2017; Essou et al., 2016; Jiang & Bauer Gottwein, 2019; Tarek et al., 2020; Zhuo, Han, et al., 2015). Also, some rainfall data sets behaved equally well in simulating streamflow (Essou et al., 2016). However, some rainfall data sets

Methodology: Jiao Wang, Lu Zhuo, Dawei Han
Software: Jiao Wang, Ying Liu
Supervision: Lu Zhuo, Dawei Han, Miguel Angel Rico-Ramirez
Validation: Jiao Wang
Visualization: Jiao Wang, Ying Liu
Writing – original draft: Jiao Wang
Writing – review & editing: Jiao Wang, Lu Zhuo, Dawei Han, Miguel Angel Rico-Ramirez

still exhibited poor performance no matter how many model calibrations were carried out (Jiang et al., 2012; Srivastava et al., 2014). Notably, these studies used the dynamic calibration approach of the model parameters, namely, recalibrating the hydrological model when using different rainfall inputs (Oudin et al., 2006; Raimonet et al., 2017). Consequently, each rainfall forcing corresponds to an optimal parameter set in each calibrated model. This is understandable because the model requires calibration to estimate parameter values, but no rainfall data set can be regarded a priori as a reference for calibration (Beven, 2012; Raimonet et al., 2017; Singh & Bárdossy, 2012). Certainly, such rainfall data sets could have opposite hydrological performance in other regions due to remarkable temporal-spatial discrepancies. No rainfall product can always exhibit excellent performance everywhere (Barrett et al., 1994; Beck et al., 2017). This study aims to explore hydrological model adaptability, which is defined as the calibration ability of a hydrological model to adapt to different scenarios, with a focus on the errors in rainfall inputs. Hence, the inter-comparison of different data sets in different regions is of minimal concern and out of the scope of this study.

It is unclear what kind of information is transmitted from rainfall inputs to the streamflow outputs via the hydrological model (Beven & Smith, 2015; Müller et al., 2021). Previous studies on the error sources in hydrological modeling suggested that rainfall quality might be the most important trigger, and its systematic error has a larger effect on the model output than the random error (Oudin et al., 2006; Paturel et al., 1995; Renard et al., 2010; Xu et al., 2006). This is because the systematic error in rainfall data will lead to a bias in the water balance, thereby resulting in a systematic error in the model calibration. The hydrologic model calibration is capable of compensating for errors in model inputs, outputs and deficiencies in model structure (Clark & Vrugt, 2006). This allows some biased rainfall data sets to perform well in hydrological simulations through model calibration (Jiang et al., 2012). As reported by previous studies (Clark & Vrugt, 2006; Kabir et al., 2022; McMillan et al., 2011; Xu et al., 2006), the model calibration is highly dependent on the rainfall data and can compensate for inaccurate rainfall to some extent. Additionally, different hydrological models adapt to the quality of rainfall inputs variously due to their different parametrizations (Oudin et al., 2004, 2006). However, what kind of inaccurate rainfall inputs would make the model calibration fail to adapt is still unclear. Even some of the fundamental questions regarding the adaptability mechanisms of hydrological models to rainfall input errors remain unanswered.

Here, we aim to explore three key scientific questions: (a) What rainfall input errors can and cannot be adapted by a hydrological model? (b) What are the specific requirements of hydrological model adaptability to rainfall input errors? (c) Why can the hydrological model adapt to some inaccurate rainfall inputs? Our objectives are threefold: (a) to trace the potential patterns of hydrological model adaptability to the quality of rainfall inputs; (b) to identify the restrictive conditions (i.e., range or thresholds, scenarios) of hydrological model adaptability to rainfall inputs; and (c) to discover the underlying mechanisms of hydrological model adaptability. Several hourly rainfall data sets from multiple sources are compiled for a headwater catchment in the Southern United States. The rainfall data sets include measurements from rain gauges, weather radars, satellites, reanalysis products, and NWP model simulations. This is the first time to employ such a large number of rainfall data sets to study hydrological model adaptability. We use these data sets to independently calibrate a widely used conceptual Xin'anjiang (XAJ) hydrological model. The impact of different rainfall input scenarios on model performance is analyzed to identify the patterns and conditions for hydrological model adaptability. The mechanism of this adaptability is further discussed.

2. Data and Methods

2.1. Study Area

To minimize human interference, a near-natural headwater-type catchment is selected as a case study. The upper Lobutcha catchment is one of the headwater-type catchments of the Upper Pearl River Basin in the central part of Mississippi State, United States (Figures 1a and 1b), with its elevation decreasing from north to south (Figure 1c). This catchment has a drainage area of 635 km^2, with a relatively sparse population and a small amount of development. The major waterway is the Lobutcha creek, one of the upper tributaries of the Pearl River, which enters the Pearl River from the northwest side (Mississippi Department of Environmental Quality, 2008). The flow from Lobutcha creek is diverted into Ross Barnett Reservoir (130 km^2), which is the state's largest source of drinking water (Wersal et al., 2006). The catchment is dominated by a humid subtropical climate, making it suitable for the application of the XAJ hydrological model. The precipitation is unevenly distributed throughout the year and mainly occurs from January to May.

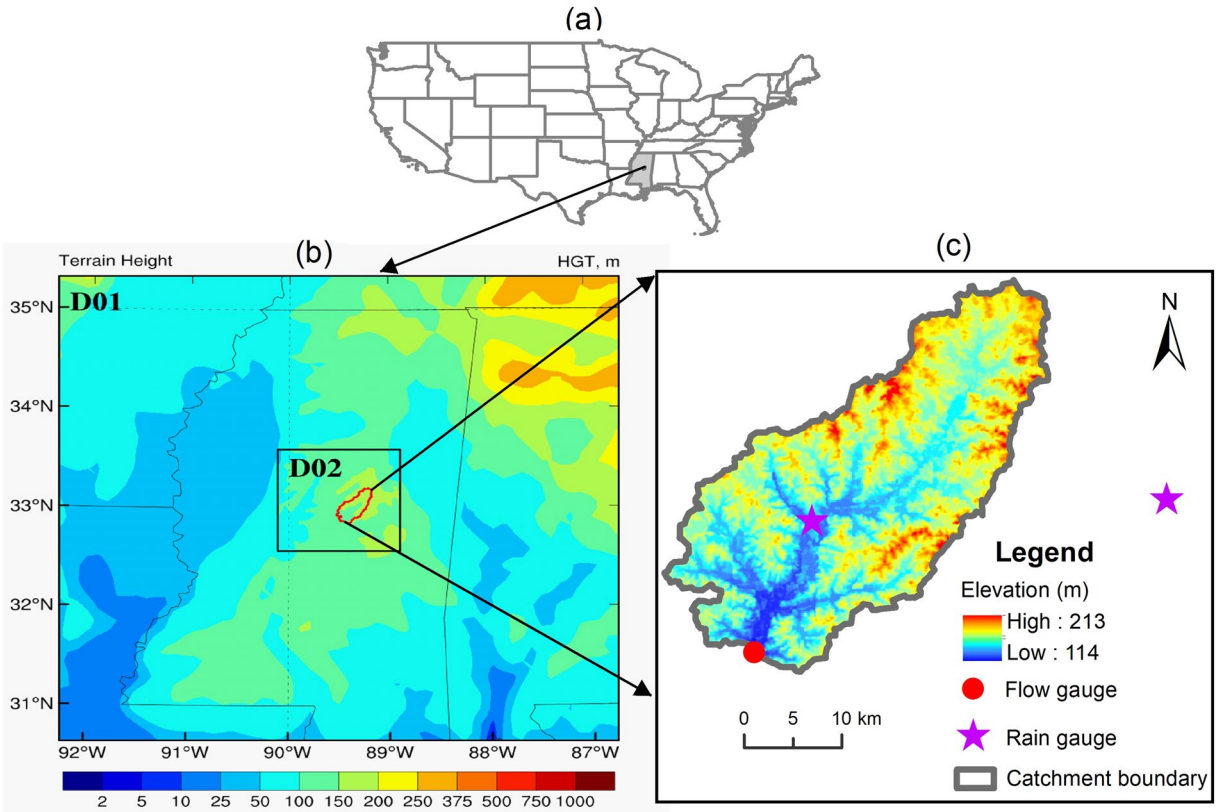

Figure 1. The location map (a and b), two nested domains (D01 and D02) employed in the Weather Research and Forecasting model experiments (b), and elevation distribution (c) of the study area.

2.2. Hydrological Data Sets

The hydrological data sets cover the period from 1 January to 30 March 2016. During this period, a heavy storm event occurred on 10 March 2016, and the National Weather Service declared a state of emergency for the state of Mississippi. About 8–10 inches of rain fell in 24 hr in some areas during this event (https://floodlist.com/america/usa/floods-texas-louisiana-mississippi-oklahoma-march-2016). The storm event caused the flow at the outlet of the study catchment to peak in that year.

To explore the underlying mechanism of hydrological model adaptability to rainfall inputs of varied quality, this study employs twenty-eight up-to-date rainfall data sets. Such data sets are generated in different ways and have different spatial-temporal resolutions. Their brief information is summarized in Table 1.

The rain gauge data are obtained from the Hourly Precipitation Data set (HPD). It is collected from the NOAA/NWS Fischer-Porter in situ network of precipitation gauges (Lawrimore, 2018). HPD data are used as observed rainfall in this study. The weather radar data from the NEXRAD Mosaics are developed and archived by Iowa Environmental Mesonet at the Iowa State University. They used level-III base radar reflectivity from approximately 130 WSR-88D sites across the continental United States to generate high-resolution products (Schultz et al., 2016). The rainfall rate with the WSR-88D radar network uses the default *Z-R* relationship (Woodley et al., 1975).

The Global Precipitation Measurement (GPM) consists of one Core Observatory and approximately 10 constellation satellites. Its level-3 products are produced with the IMERG algorithm, which inter-calibrates and merges precipitation estimates from all constellation microwave sensors and microwave-calibrated infrared satellite estimates (Huffman et al., 2019). The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN-CCS, simply CCS) uses infrared images as the only input, employing the cloud top temperature and rainfall (T_b-R) relationships to produce rainfall estimates indirectly (Hong et al., 2004).

Table 1
Overview of Various Rainfall Data Sets Used in This Study

Type	Name	Coverage	Spatial/temporal resolutions
Gauge	HPD V2	Mainly in CONUS, 1948–present	Point/1 hr
Radar	NEXRAD Mosaic	23°–50°N, 126°W–65°E, 1995–present	0.5 km/5 min
Satellite	GPM IMERG	Global, 2000–Present	0.1°/0.5 hr
	PERSIANN-CCS	60°S–60°N, 180°W–180°E, 2003–present	0.04°/1 hr
Reanalysis	NLDAS-2	25°–53°N, 125°W–67°E, 1979–present	0.125°/1 hr
	GLDAS-2.1	60°S–90°N, 180°W–180°E, 2000–present	0.25°/3 hr
	ERA5	Global, 1959–present	0.25°/1 hr
	ERA5-Land	Global, 1950–present	0.1°/1 hr
NWP simulations	WRF-downscaled (20 scenarios)	Study area, study period	3 km/1 hr

Note. CONUS means the Contiguous United States.

Phase 2 of the North American Land Data Assimilation System (NLDAS-2) is an offline data assimilation system with uncoupled land surface models driven by observation-based atmospheric forcing (Rui & Mocko, 2019; Xia et al., 2012). The GLDAS-2.1 is an offline terrestrial modeling system, which makes use of ground-based observations and satellite products (Rui & Beaudoing, 2019; Rodell et al., 2004). The generation of ERA5 engages with four-dimensional variational data assimilation of satellite and in-situ observations (Copernicus Climate Change Service, 2018). The ERA5-Land is a land-only reanalysis product involved with precipitation readjustment. It is forced by the low atmospheric meteorological fields of ERA5 and additional lapse-rate correction (Copernicus Climate Change Service, 2019).

In addition to rain gauges, radar, satellites and reanalysis products, the NWP models are also widely used for simulating rainfall estimations. The Weather Research and Forecasting (WRF) model is the state-of-the-art mesoscale NWP model and can be used to produce various meteorological data (e.g., rainfall, evapotranspiration, soil moisture, wind, etc.). It enables the downscaling of these variables at much finer spatial and temporal resolutions (J. Liu et al., 2012; Rogelis & Werner, 2018; Srivastava et al., 2014), which would meet the growing requirement for studying floods in small and medium catchments. However, the WRF rainfall outputs are highly sensitive to physical parameterization, such as cloud microphysics, cumulus and planetary boundary layer (C. Liu et al., 2011; Y. Liu et al., 2021). In this study, 20 parameterizations are applied to the WRF model separately to produce different rainfall data sets, aiming to extend the range of the rainfall data set's quality. An added benefit is that this can also provide new insights into the hydrological performance of the WRF rainfall estimations under different WRF parameterizations, which have rarely been investigated in previous studies. All scenarios are computed on two nested domains, D01 and D02 (Figure 1b). The rainfall estimates from the inner domain D02 are used in this study. The schemes used in each parameterization scenario and the related websites can be found in Table S1 in Supporting Information S1.

We use the potential evapotranspiration (PET) from the NLDAS-2 product due to the lack of evapotranspiration observations to drive the XAJ hydrological model (Zhuo, Han, et al., 2015). The PET data is kept the same in all XAJ simulations. The observed hourly streamflow data at the catchment outlet are provided by the United States Geological Survey.

2.3. Hydrological Modeling

The XAJ model is selected for hydrological simulation in this study. It is computationally efficient and easy to use with the minimal preparation needed for input data. It has shown good performance in humid and semi-humid catchments worldwide (Yang et al., 2020; Zhuo & Han, 2016b; Zhuo, Han, et al., 2015). The XAJ model consists of four modules: evapotranspiration module, runoff production module, runoff separation module and runoff concentration module (Zhao et al., 1980). First, the evapotranspiration module computes the actual evapotranspiration in the three soil layers (upper, lower and deepest layers). Then a tension water capacity curve is introduced in the runoff production module, and the runoff is produced when the soil water content of its aeration zone exceeds the field capacity. The total runoff is separated into the surface runoff, interflow, and groundwater

Table 2
The Range of the Xin'anjiang Model Parameters Used in the Study Area

Symbol	Definition (unit)	Hydrological process	Range
K	Ratio of evapotranspiration	Evapotranspiration, Runoff production	0.20–1.50
WUM	Areal mean tension water storage capacity of the upper soil layer (mm)	Evapotranspiration	10–50
WLM	Areal mean tension water storage capacity of the lower soil layer (mm)	Evapotranspiration	20–150
WDM	Areal mean tension water storage capacity of the deepest soil layer (mm)	Evapotranspiration	50–400
C	Coefficient of potential evaporation in the deepest soil layer	Evapotranspiration	0.10–0.70
WM	Areal mean tension water storage capacity (WM = WUM + WLM + WDM)	Runoff production	/
IMP	Percentage of impervious areas in the catchment (%)	Runoff production	0.001–0.10
B	Exponent of the tension water-capacity distribution curve	Runoff production	0.10–0.90
SM	Areal mean free water-storage capacity (mm)	Runoff separation	20–50
EX	Exponent of the free water-capacity distribution curve	Runoff separation	0.001–2.00
KI	Outflow coefficient of the free water storage to interflow relationships	Runoff separation	0.10–0.70
KG	Outflow coefficient of the free water storage to groundwater relationships	Runoff separation	0.10–0.70
CI	Recession constant of the interflow storage	Runoff concentration	0.01–0.996
CG	Recession constant of the groundwater storage	Runoff concentration	0.01–0.996
CS	Recession constant for routing through the channel system with each sub-basin	Runoff concentration	0.10–0.70
LAG	Lag parameter of the river networks concentration	Runoff concentration	0.001–6.00
V	Parameter of the Muskingum method (m/s)	Runoff concentration	0.20–1.0
dX	Parameter of the Muskingum method	Runoff concentration	0.001–0.40

runoff in the runoff separation module. Last, the surface runoff at the outlet of the catchment is computed using the Muskingum method in the runoff concentration module, while the inter runoff and groundwater are routed to river channels through linear reservoirs. A detailed description of the XAJ model can be found in previous studies (Zhao, 1992; Zhao et al., 1980). Based on some primary experiments and the aforementioned studies, the range values of each XAJ parameter used for the study catchment is shown in Table 2.

For better hydrological simulations, longer time series data would be ideally required. However, due to the huge computing resources requested by the WRF simulations and the issue of incomplete rainfall observations, this study uses a 3-month period covering the flash flooding event from March 2016. Each rainfall data set is converted into the catchment scale data at an hourly time step and is used as an input for the XAJ hydrological modeling individually. The areal mean rainfall is calculated by averaging the observations of the two rain gauges (Figure 1c) using the Thiessen polygon method (Han & Bray, 2006). For grid-type data sets, it is through calculating all the grids (fully and partially) located inside the catchment boundary by the weighted average method (Zhuo & Han, 2016a). The period from 1 to 6 January is used as the spin-up of the WRF model. The time required for a hydrological model to warm up varies with the model structure and may be affected by rainfall amount and initial wetness (Kim et al., 2018). Ideally, a longer time would help determine the appropriate initial condition for each of the 28 rainfall input data sets. However, only 3 months are used in this study due to the vast computational resources required to run 20 WRF scenarios. The initial conditions (three-layer soil moisture) are optimized as a part of the calibration process. This ensures that the entire study period is used. The first 12-day data are then neglected to minimize potential unstable errors in the early stages of calibration. Ultimately, the following analysis is based on hourly data from 19 January to 30 March 2016.

Normally in hydrology, we use part of the data to calibrate the hydrological model and the rest for validation. This is helpful to judge if the model is over-parameterized and overfitted. Such a method is of great importance to guarantee the feasibility of the constructed hydrological model and the reliability of its hydrological prediction. However, this study focuses on exploring how the hydrological model adapts to varied-quality rainfall inputs through model calibration. It is therefore meaningless to detect over-parameterization and overfitting. Thus, we calibrate the XAJ parameters by using the whole data period without dividing them into calibration and validation periods. We use a dynamic calibration (C1) approach for each of the rainfall input data sets. Twenty-eight XAJ

models are calibrated and independent of each other. In particular, the XAJ parameters are calibrated by using a standard gradient-based automatic optimization (GAO) method (Lagarias et al., 1998). The streamflow simulation performance is then evaluated by the Nash-Sutcliffe Efficiency (NSE) criteria (Nash & Sutcliffe, 1970), the most common indicator in hydrology (e.g., Beck et al., 2017; J. Liu & Han, 2010; Kim et al., 2018; Zhuo, Dai, & Han, 2015).

Meanwhile, we conduct 100,000 experiments for each rainfall input data set to understand their general hydrological simulation performance. The independent random parameter set in each experiment is generated by the Latin Hypercube Sampling (LHS) Strategy (Iman & Conover, 1980; McKay et al., 2000). The feasible ranges of parameters (see Table 2) are used to dominate the generation of random parameter sets (Helton & Davis, 2003; Beven & Freer, 2001). Correspondingly, the optimal XAJ parameter set for each rainfall data set is obtained based on the best of the LHS and GAO results. To intuitively compare the optimal parameter values across all rainfall data sets, each parameter value is normalized to [0,1], using:

$$\mathrm{Pa}_{\mathrm{norm}} = (\mathrm{Pa} - \mathrm{Pa}_{\mathrm{min}})/(\mathrm{Pa}_{\mathrm{max}} - \mathrm{Pa}_{\mathrm{min}}) \tag{1}$$

where Pa refers to each XAJ model parameter, $\mathrm{Pa}_{\mathrm{min}}$ and $\mathrm{Pa}_{\mathrm{max}}$ are the minimum and maximum parameter values across all rainfall data sets, and $\mathrm{Pa}_{\mathrm{norm}}$ represents the normalized parameter value.

Due to the equifinality effect in hydrological modeling (Beven & Freer, 2001), there may be other parameter sets that allow the model to achieve equally good simulation results. The model parameter sensitivity can complementarily help us to understand the contributions of different parameter sets to hydrological model adaptability. A widely used global sensitivity analysis technique, called Elementary Effect test (Campolongo et al., 2007; Morris, 1991; Pianosi et al., 2015), is used to analyze the parameter sensitivity to model output. We also apply a static calibration (C2) method to investigate the effect of C1 method and better understand the hydrological model adaptability. In the C2 method, only rain gauge data are used for the hydrological model calibration and the optimal parameter set is kept constant when using other rainfall data sets.

2.4. Rainfall Quality Indicators

The quality of rainfall data sets needs to be quantified to explore the connections between hydrological model performance and its rainfall input accuracy. Based on previous studies on evaluating rainfall quality (e.g., Beck et al., 2017; Jiang & BauerGottwein, 2019; Jiang et al., 2012; Kabir et al., 2022), four commonly used indicators are selected to characterize the reliability of each data set against rain gauge data, encompassing NSE, Pearson's Correlation Coefficient (PCC), Root Mean Square Error (RMSE, mm), and Bias (BIAS, mm). NSE measure the goodness of fit between the rainfall observation and estimations. It varies from negative infinity $-\infty$ to 1, and a value of 1 means the simulations perfectly match the observations. PCC measure the degree of linear correlation between estimation and observation, and its optimal value is 1. RMSE reflects the average magnitude of the error and the perfect value of RMSE is 0. BIAS indicates the total systematic bias in the estimation compared to the observation, and its optimal value is 0. Their equations are summarized below,

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{T}(S_i - O_i)^2}{\sum_{i=1}^{T}\left(O_i - \overline{O}\right)^2} \tag{2}$$

$$\mathrm{PCC} = \frac{\sum_{i=1}^{T}\left(O_i - \overline{O}\right)\sum_{i=1}^{T}\left(S_i - \overline{S}\right)}{\sqrt{\sum_{i=1}^{T}\left(O_i - \overline{O}\right)^2\sum_{i=1}^{T}\left(S_i - \overline{S}\right)^2}} \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(S_i - O_i)^2} \tag{4}$$

$$\mathrm{BIAS} = \sum_{i=1}^{T}(S_i - O_i) \tag{5}$$

where O_i and S_i are rain gauge data and other rainfall data sets, respectively; $\overline{O}$ and $\overline{S}$ are the mean values of O_i and S_i, respectively; and T is the total number of hours.

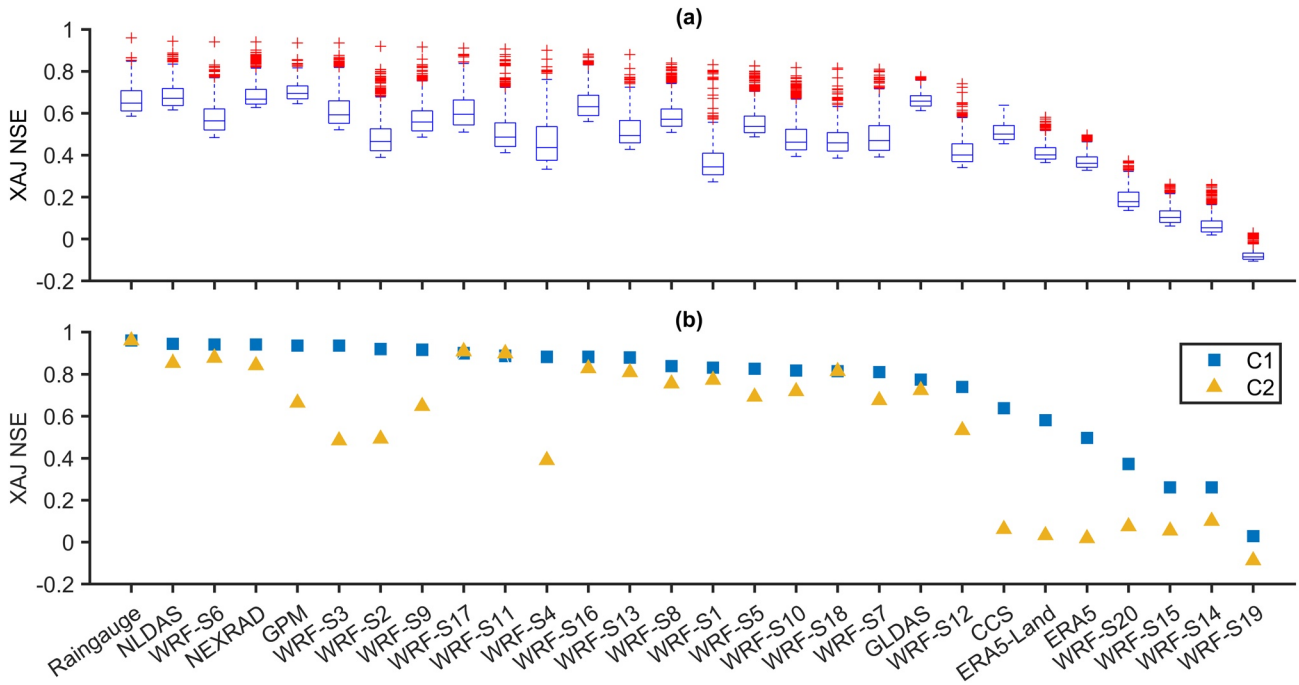

Figure 2. Overall performance of the best 1,000 Xin'anjiang (XAJ) simulations forced by the 28 hourly rainfall data set (a), and the XAJ Nash-Sutcliffe Efficiency (NSE) results by using dynamic (C1) and static (C2) calibration methods (b).

3. Results and Discussion

3.1. What Rainfall Input Errors Can and Cannot Be Adapted by a Hydrological Model?

To provide insight into the overall hydrological performance of each rainfall data set, we first compare the best 1,000 XAJ NSE results among all LHS experiments and GAO calibration results (Figure 2a). It is readily discernible that the streamflow simulation performance driven by different rainfall data sets varies substantially. Variously predictive skills can be seen in terms of NSE across different simulations. Figure 2b provides the best simulation performance of each rainfall forcing using the C1 method. Most rainfall data sets can generate satisfactory streamflow simulations (XAJ NSE > 0.8) with certain XAJ parameter settings. The corresponding XAJ modeled streamflow are shown in Figure S1 in Supporting Information S1. The streamflow simulated by the rain gauge data is in good agreement with the flow gauge observations (XAJ NSE = 0.96), indicating that the XAJ model can be applied to the study area.

Figure 2b also includes the XAJ NSE results calibrated by the C2 method. It can be seen that the hydrological performance has been greatly improved by using the dynamic (input-specific) calibration method (C1). For example, the XAJ NSE values of GPM, WRF-S3, WRF-S2, WRF-S9, and WRF-S4 simulations using the C1 method have reached more than 0.9, while those using the C2 method are about 0.45–0.65. Similarly, CCS, ERA5-Land and ERA5 simulations show significant improvements with XAJ NSE increasing from 0.02–0.06 to 0.5–0.65. This indicates that the hydrological model can adapt to rainfall inputs to improve the simulation performance. However, it is worth pointing out that the hydrological model is incapable of adapting to six rainfall data sets. ERA5, ERA5-Land, WRF-S14, WRF-S15, WRF-S19, and WRF-S20 rainfall data sets fail to produce acceptable hydrological performance (XAJ NSE < 0.6) no matter what parameter settings are selected. This suggests that the ability to adapt to rainfall data might have some limits. The hydrological model may lose its adaptability if the quality of rainfall data falls below a threshold.

Figure 3 presents the relationship between the rainfall quality indicators (PCC, NSE, RMSE, and BIAS) and the corresponding hydrological performance (XAJ NSE). It can be seen that many rainfall inputs with a PCC (NSE) inferior to 0.6 (0.2) can still yield satisfactory hydrological performance with a XAJ NSE greater than 0.8 (Figures 3a and 3b). Similar findings can also be observed when the rainfall RMSE is even greater than 1.5 mm (Figure 3c). It appears that some inaccurate rainfall estimates can still reproduce streamflow well. There

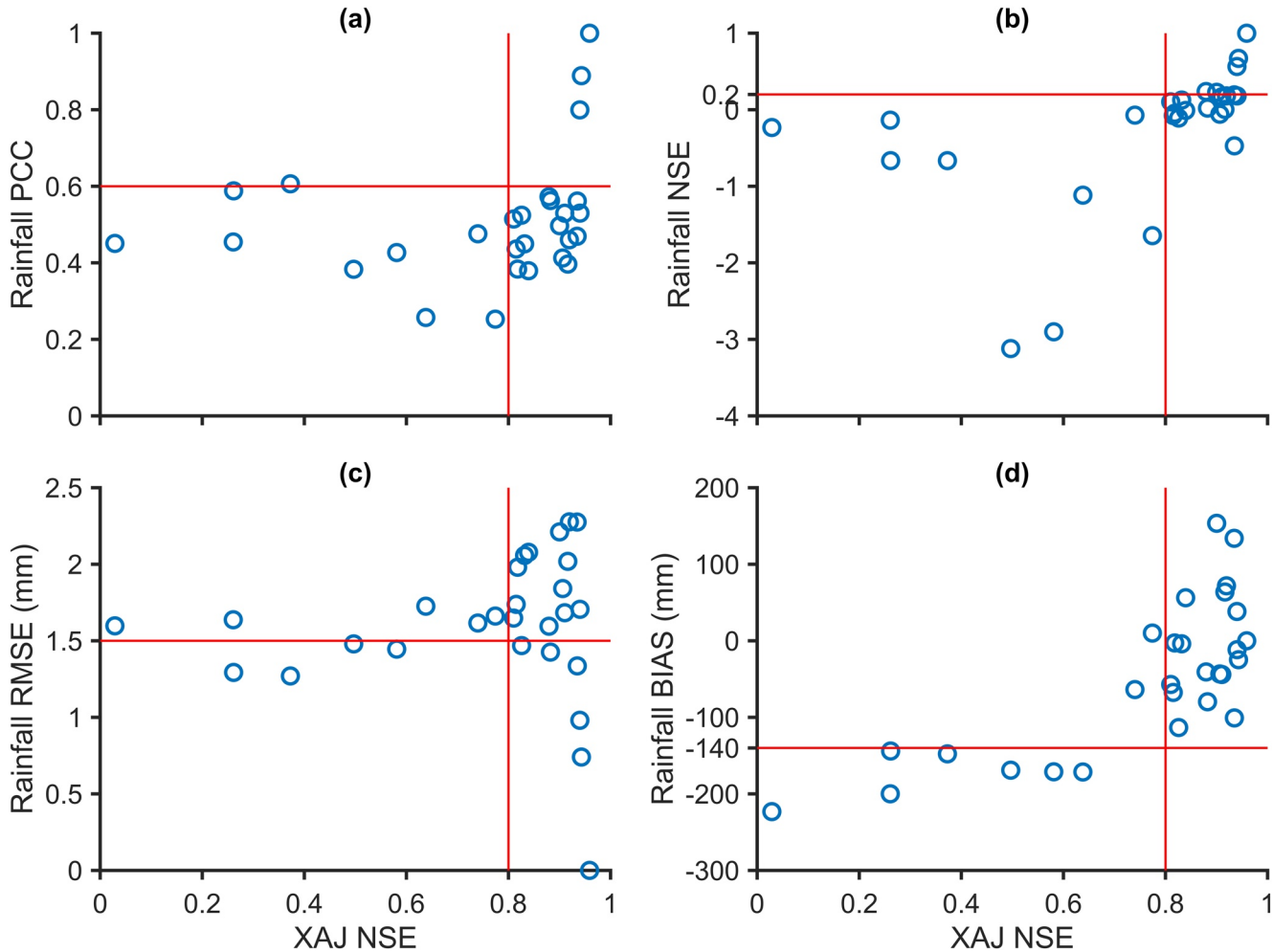

Figure 3. Each scatterplot shows the relationship between the rainfall quality indicators, Pearson's Correlation Coefficient (PCC) (a), Nash-Sutcliffe Efficiency (NSE) (b), Root Mean Square Error (RMSE) (c), Bias (BIAS) (d), and the corresponding streamflow simulation performance (Xin'anjiang NSE).

are weak relations between these three rainfall quality indicators (PCC, NSE, and RMSE) and the corresponding streamflow simulation performance (XAJ NSE). However, it is evident from Figure 3d that all good XAJ simulations are driven by the rainfall input data sets with BIAS that are above −140 mm. In other words, when the rainfall data set shows an underestimation greater than 140 mm (i.e., below −140 mm), the model fails to exhibit excellent predictive skill in flow simulations (XAJ NSE < 0.8). These results suggest that rainfall BIAS is likely to be a better indicator to assess the hydrological applicability of rainfall data sets.

In general, a large rainfall bias can affect the water balance in hydrological modeling (Paturel et al., 1995; Xu et al., 2006). A severe underestimation usually indicates a failure to detect some rainfall events, making it difficult for simulated streamflow to achieve the same peak flow as the flow gauge observations. The hydrological model would not be able to adapt to rainfall inputs below the underestimation threshold, which is diagnosed as −140 mm in this particular study. Interestingly, even within the threshold of −140 mm, it is found from Figure 3d that some rainfall data sets with a BIAS greater than 100 mm generate better hydrological performance than those with low BIAS values. This implies that the overall rainfall bias is not the only factor determining the corresponding hydrological performance and further investigation is needed.

Figure 4 depicts a comparison of the cumulative rainfall from 28 rainfall products against rain gauge observations, which have 518 mm of rainfall in total over the entire period. It can be found from Figures 4a and 4b that the rainfall data sets with a low overall BIAS (|BIAS| < 52 mm) can reproduce streamflow with good accuracy (XAJ NSE > 0.9). On the contrary, the data sets shown in Figure 4f underestimate the rainfall by more than 140 mm and

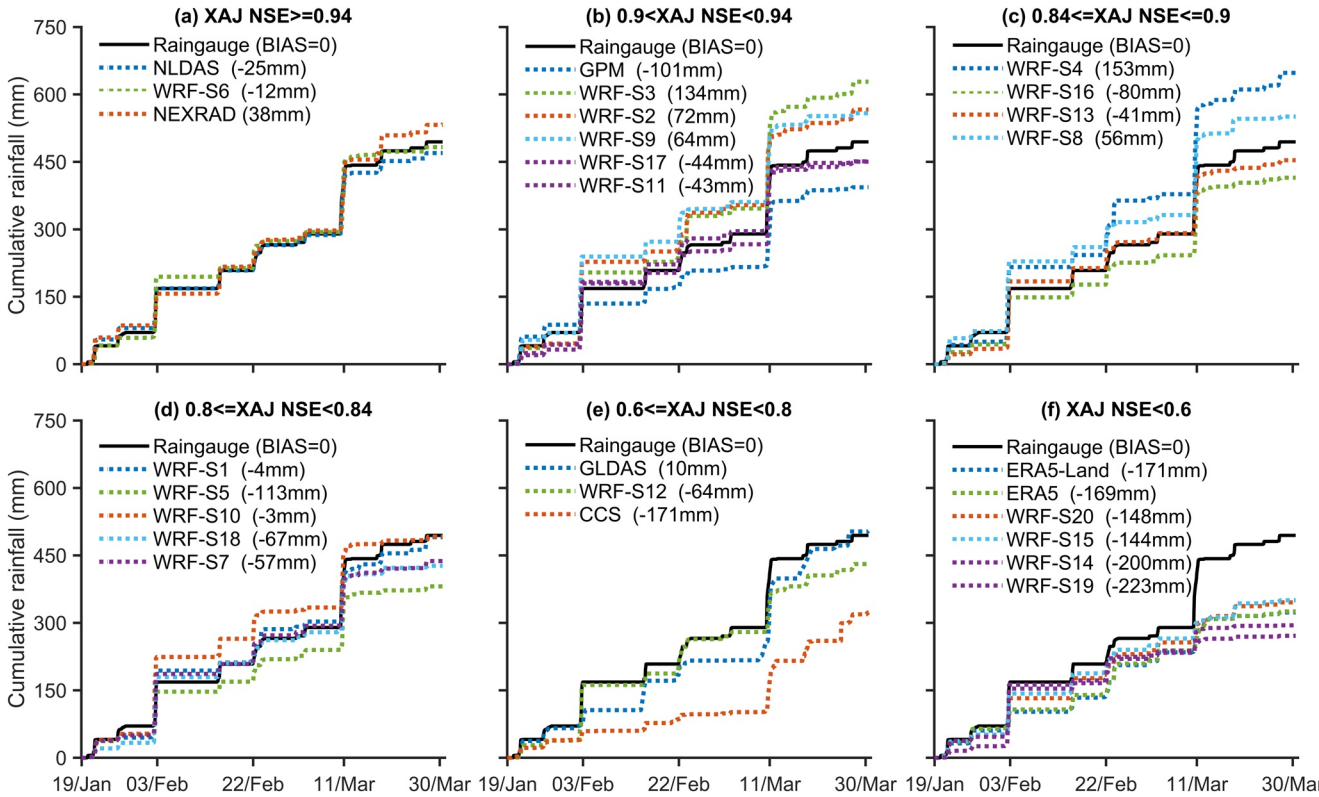

Figure 4. Cumulative rainfall comparison of 28 rainfall input data sets during the entire study period. The number in the bracket of the legend shows an overall Bias (BIAS) value of each rainfall data set against the rain gauge observations. The rainfall products are presented in descending order of their Xin'anjiang modeling's Nash-Sutcliffe Efficiency (XAJ NSE) values, including 6 categories, XAJ NSE >= 0.94 (a), 0.9 < XAJ NSE < 0.94 (b), 0.84 < XAJ NSE <= 0.9 (c), 0.8 <= XAJ NSE < 0.84 (d), 0.6 <= XAJ NSE < 0.8 (e), and XAJ NSE < 0.6 (e). The *x*-axes highlight some heavy rainfall events (e.g., 3 February, 22 February, and 11 March).

fail to yield good hydrological performance (XAJ NSE < 0.6). It is also interesting that the rainfall data sets from GPM, WRF-S3, and WRF-S4 with a BIAS of −101, 134, and 153 mm, respectively, yield satisfactory streamflow simulations (XAJ NSE > 0.9). They outperform WRF-S1, WRF-S10, and GLDAS rainfall data sets that have smaller BIAS (about −4 to 10 mm) and XAJ NSE values (about 0.8). This suggests that the rainfall input with a lower overall BIAS does not always lead to a better streamflow simulation. These findings emphasize the limitation of using the overall rainfall BIAS as a measure of performance in representing the relation between rainfall input quality and hydrological model performance.

Notably, large discrepancies in estimation accuracy can be seen across different rainfall events (as shown by the sudden increases in rainfall in Figure 4). The data sets which are shown in Figure 4f exhibit substantial underestimation of rainfall during the 11 March event and an overall poor hydrological performance. Similarly, GLDAS significantly underestimates the rainfall event on 3 February. Other rainfall data sets also present different over- and underestimations in different rainfall events. These findings highlight the significance of event-based bias in rainfall inputs. Our next step is to further analyze the specific rainfall BIAS of each rainfall event.

3.2. What Are the Specific Requirements of Hydrological Model Adaptability to Rainfall Input Errors?

The whole study period can be divided into eight rainfall events based on the observed rainfall and streamflow. The detailed information of each rainfall event is presented in Figure S2 in Supporting Information S1. The total rainfall amount during the entire study period is 518 mm. Event 6 is the severest event which happened around 11 March, resulting in about 153 mm rainfall in total. Event 2 (98 mm) is the second severest rainstorm which mainly happened on 3 February. These are the two major rainstorms that are the focus in the following analysis.

It is important to understand how the bias of each rainfall event shapes the overall rainfall bias. Figure 5 provides information on the overall pattern and fluctuations of the bias across all events. It shows the event-based rainfall

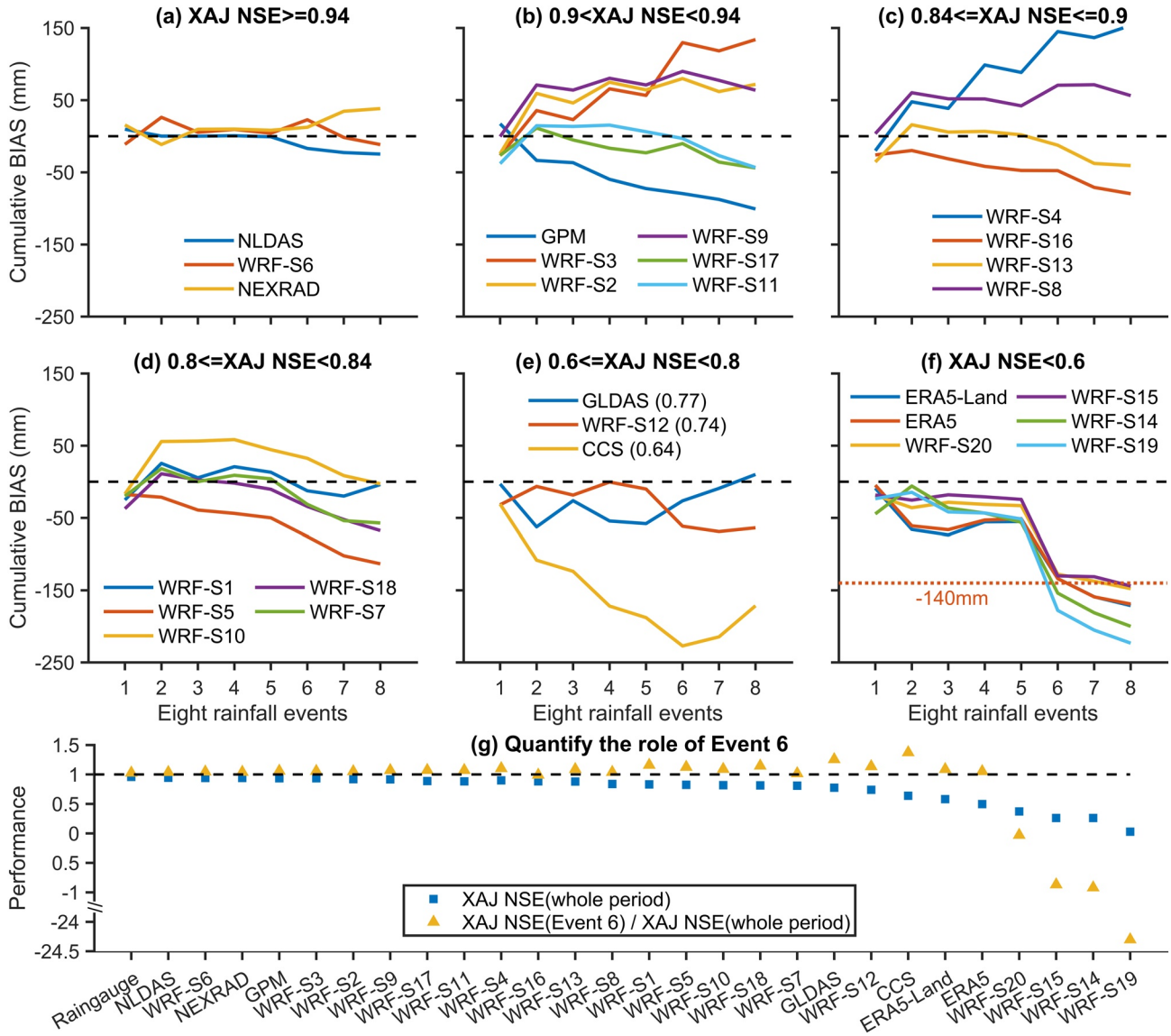

Figure 5. Cumulative event-based Bias (BIAS) from Event 1 to Event 8 in each rainfall data set (a–f). The red dashed line shows an overall rainfall bias of −140 mm, which is the threshold for the rainfall input data to reproduce streamflow well in this study. Panel (g) shows the role of the heaviest Event 6 in the overall streamflow simulations.

BIAS of each rainfall data set against the rain gauge data. The cumulative bias of the two major rainstorms (Event 6 and Event 2) can be found in Figures S3 and S4 in Supporting Information S1. The following key observations can be made. First, the accuracy of rainfall estimates during the heaviest Event 6 plays a key role in the overall streamflow simulations. With an overall rainfall BIAS higher than −140 mm and acceptable BIAS in Event 6, a rainfall data set can reproduce streamflow with a XAJ NSE greater than 0.7. The input rainfall products that have good rain estimates for the major rainstorms and low overall BIAS during the whole period, can reproduce satisfactory streamflow simulations (Figures 5a–5d), such as NLDAS, WRF-S6, NEXRAD, WRF-S17, and WRF-S11. On the contrary, even with low BIAS during other events, the rainfall input data set that substantially underestimates the most severe Event 6 could still perform poorly in terms of streamflow simulations (Figure 5f). Moreover, GLDAS and CCS (Figure 5e) fail to detect the second most severe Event 2, making them hard to reach the second peak flow and only yield acceptable streamflow simulations (0.6 < XAJ NSE < 0.8). Similarly, due to large overestimations of Event 2, WRF-S1 and WRF-10 (Figure 5d) are incapable of generating excellent hydrological performance. Nevertheless, their XAJ NSE values are about 0.8 as they show good estimates for Event 6 well and have low rainfall BIAS throughout the period.

These findings suggest that the hydrological simulation performance (XAJ NSE) is not only affected by the overall rainfall BIAS but also influenced by the event-based BIAS, especially the BIAS during major rainfall events. That is reasonable because the two major rainstorms receive a substantial portion of the total rainfall (~50%) during the study period, and the rainfall during Event 6 leads to peak flow during the study period. Besides, NSE is highly sensitive to peak flows (Beck et al., 2017; Gupta et al., 2009; Krause et al., 2005). Hence, the overall streamflow simulation is highly influenced by these heavy rainfall events, particularly Event 6, but less affected by other light or medium rainfall events. To quantify the role of Event 6 in the overall streamflow simulation, we calculate the XAJ NSE for Event 6 alone and then divide it by the overall XAJ NSE for the entire period. The ratio results for all rainfall data sets are shown in Figure 5g. Clearly, in most cases, the ratio is close to 1 when the overall XAJ NSE is greater than 0.8. On the other hand, when the overall XAJ NSE is less than 0.5, the ratio becomes negative, which is caused by poor streamflow simulation performance during Event 6 (XAJ NSE < 0). The finding quantitatively indicates the remarkable contribution of the heaviest rainstorm Event 6 to the overall streamflow simulations.

Second, with good estimations during smaller events and an overall rainfall BIAS of greater than −140 mm, the rainfall data sets that consistently under- or overestimate during most rainstorms can still perform well in reproducing streamflow simulations. In other words, the hydrological model can adapt to this error well. For example, although the WRF-S3 and WRF-S4 overestimate Events 2, 4, and 6 by about 60 mm, they can still produce good streamflow simulations with high XAJ NSE values (0.93 and 0.9) (Figures 5b and 5c). Despite an overall underestimation of 101 mm, GPM underestimates these three events by 51 mm, 23 and 7 mm, respectively, it still outperforms most rainfall data sets (XAJ NSE = 0.93). Similar results can also be found on WRF-S9, WRF-S2, and WRF-S8. Note that CCS has consistent underestimations for Events 1–6 and just exhibits acceptable XAJ simulations (XAJ NSE = 0.64). This is likely because it fails to detect Event 2 and underestimates the entire study period by more than 140 mm. Remarkably, the overall rainfall BIAS of these rainfall data sets is caused by consistent under- or overestimations during most rainstorms, rather than only Event 6 or Event 2. In contrast, for the data sets shown in Figure 5f, it is the large underestimations of Event 6 that mainly contribute to the overall negative rainfall BIAS.

Third, the XAJ hydrological model appears to be less likely to accommodate the rainfall inputs that severely underestimate the rainstorm events. Instead, the model can adapt well to the heavy overestimations of most rainstorms. It can be seen that most data sets in Figures 5b–5d present overestimations during rainstorms, some even have substantial overestimations, and they are still capable of yielding satisfactory streamflow simulation performance. For example, WRF-S3 overestimates Event 2 and Event 6 by 62 and 73 mm, respectively, and reproduces streamflow very well with a XAJ NSE of 0.93. However, the rainfall data sets with large underestimations during most rainstorms are found to fail to generate good streamflow simulations (Figure 5f). For example, ERA5-Land presents an underestimation for Event 2 (Event 6) of 57 mm (79 mm), and it underperforms WRF-S3 in terms of streamflow simulation (XAJ NSE = 0.58).

Fourth, the results suggest that the use of higher resolution rainfall data in this study area would not necessarily lead to improved streamflow simulations. This is in agreement with the study of Kabir et al. (2022). It is evident from Figure 5a and Table 1 that NLDAS (0.125°) has comparable performance with NEXRAD (0.5 km) in terms of rainfall estimations, and both reproduce streamflow very well. On the other hand, with a higher spatial resolution (0.04°), CCS does not yield more accurate rainfall estimates and more skilful streamflow simulations. Similarly, ERA5-Land (0.1°) and ERA5 (0.25°) data produce equivalent rainfall estimates and hydrological simulations. These results indicate that the spatial resolution of rainfall inputs is not the primary trigger of bias in streamflow simulations in this study area. Instead, it is the accuracy of the total amount of rainfall in each rainstorm (especially the heaviest one) that is the most important. However, it cannot be ignored that the situation would be quite different in complex topographic regions. For example, as shown by Rasmussen et al. (2011), the use of lower spatial resolution for the WRF model would result in an underestimation of rainfall estimations. In that case, the higher spatial resolution data are more likely to more effectively capture the spatial distribution of rainfall and thus provide more information about rainfall characteristics.

The above results demonstrate that the streamflow simulation is not only influenced by the overall bias of the rainfall estimates but also strongly influenced by its event-based bias, especially the storm event, which substantially leads to peak flow. The ideal scenario is that each rainfall event has a small bias, thereby resulting in a small overall rainfall bias. However, the complexity of rainfall characteristics makes it challenging to capture different

rainfall events. Our results suggest that the rainfall data set can reproduce streamflow well if its overall bias is mainly caused by persistent over- or underestimation of most rainstorms. Particularly, there is a threshold of the overall rainfall bias for the hydrological model calibration to adapt to rainfall inputs. In other words, the hydrological model can adapt to inaccurate rainfall inputs as long as the rainfall data contains important information content for model calibration (J. Liu & Han, 2010; Beven & Smith, 2015).

3.3. Why Can the Hydrological Model Adapt to Some Inaccurate Rainfall Inputs?

Although the accuracy of rainfall estimates varies, some rainfall products have comparable predictive skills in streamflow simulations. The question is, why can the XAJ hydrological model adapt to rainfall inputs of different quality (bias)? Since the XAJ model setups and PET data remain the same, the optimal XAJ parameters generated by each model calibration play an important role.

The XAJ hydrological model is one of the saturation-excess-based models. Rainfall first satisfies the demand for evapotranspiration. The actual evapotranspiration (ET_a) is computed based on rainfall (P) inputs, PET inputs, parameter K, soil moisture (SM) content, and soil water storage capacity in the three soil layers. The runoff production module assumes that runoff is not generated until the SM content in the aeration zone reaches the field capacity (Zhao, 1992). The streamflow at the outlet is then computed by the runoff separation and concentration modules. Hence, the total runoff in the runoff production module is the predecessor of streamflow in the runoff concentration module. Here, we calculate the effective rainfall (or net rainfall, P_E) after actual evapotranspiration and SM replenishment using the following equation:

$$P_E = P - \mathrm{ET}_a - \Delta_{\mathrm{SM}} \tag{6}$$

where Δ_{SM} represents the change in SM content. The latter two components are computed using the total amount of water available in the three soil layers.

To examine P_E during the simulation process and to get an intuitive understanding of how event-based rainfall BIAS influences the streamflow simulations, we select three representative rainfall data sets to illustrate the comparisons. Figure 6 shows the cumulative rainfall, streamflow, cumulative P_E and SM changes. GPM has an underestimation of 101 mm in total and consistently underestimates most rainstorms (Figure 6a). On the contrary, WRF-S3 (Figure 6b) consistently overestimates rainstorm events, leading to an overall overestimation of 134 mm for the whole study period. WRF-S15 (Figure 6c) estimates most events well, and its overall rainfall BIAS (−144 mm) is mainly caused by a significant underestimation of Event 6 (106 mm). It is found that GPM and WRF-S3 behave equally well in reproducing streamflow with a XAJ NSE of 0.93 (Figures 6d and 6e). Also, their cumulative P_E curves (red lines) are very close to those using rain gauge data (Figures 6g and 6h). However, WRF-S15 rainfall only yields a streamflow simulation with a XAJ NSE of 0.26 (Figure 6f). Although its total P_E is slightly higher than those using rain gauge data (Figure 6i), large differences can be found during individual rainfall events. In particular, its P_E is overestimated during Events 1–5 and 7–8, but is severely underestimated during the most severe Event 6. This highlights again the importance of the estimation accuracy of individual rainfall events, especially the heaviest event that leads to peak flow. If a significant overall bias of rainfall input data is predominantly ascribed to the underestimation of the most severe rainstorm, then the hydrological model is unable to adapt to the rainfall error.

Specifically, as the parameter primarily responsible for transforming PET into ET_a, K plays a critical role in the water balance of the hydrological simulations. It can adapt to the overall bias of rainfall inputs in the evapotranspiration module. For WRF-S3 rainfall data, K is optimized as 1.5 (Figure 6e), which reaches the maximum and means scaling up the PET data to adapt to its consistent underestimation of rainstorms and large overall underestimation. When the XAJ model is driven by GPM rainfall which consistently underestimates rainstorms and holds an overall underestimation, K is optimized as 0.48 (Figure 6a) to reduce the PET inputs. However, with a large underestimation of Event 6 mainly resulting in an overall rainfall bias below the adaptable threshold (−140 mm), WRF-S15 fails to reproduce streamflow well (Figures 6c and 6f). In this case, the model calibration has optimized K with a minimum value (0.2) in an attempt to reduce the PET inputs to compensate for the considerably underestimated rainfall inputs. Nonetheless, the model still fails to adapt to the WRF-S15 rainfall data set.

Meanwhile, SM storage also affects the water balance on hydrological model adaptability. If the rainfall input data exhibit underestimations, such as GPM and WRF-S15 (Figures 6j and 6l), less rainwater is absorbed by the

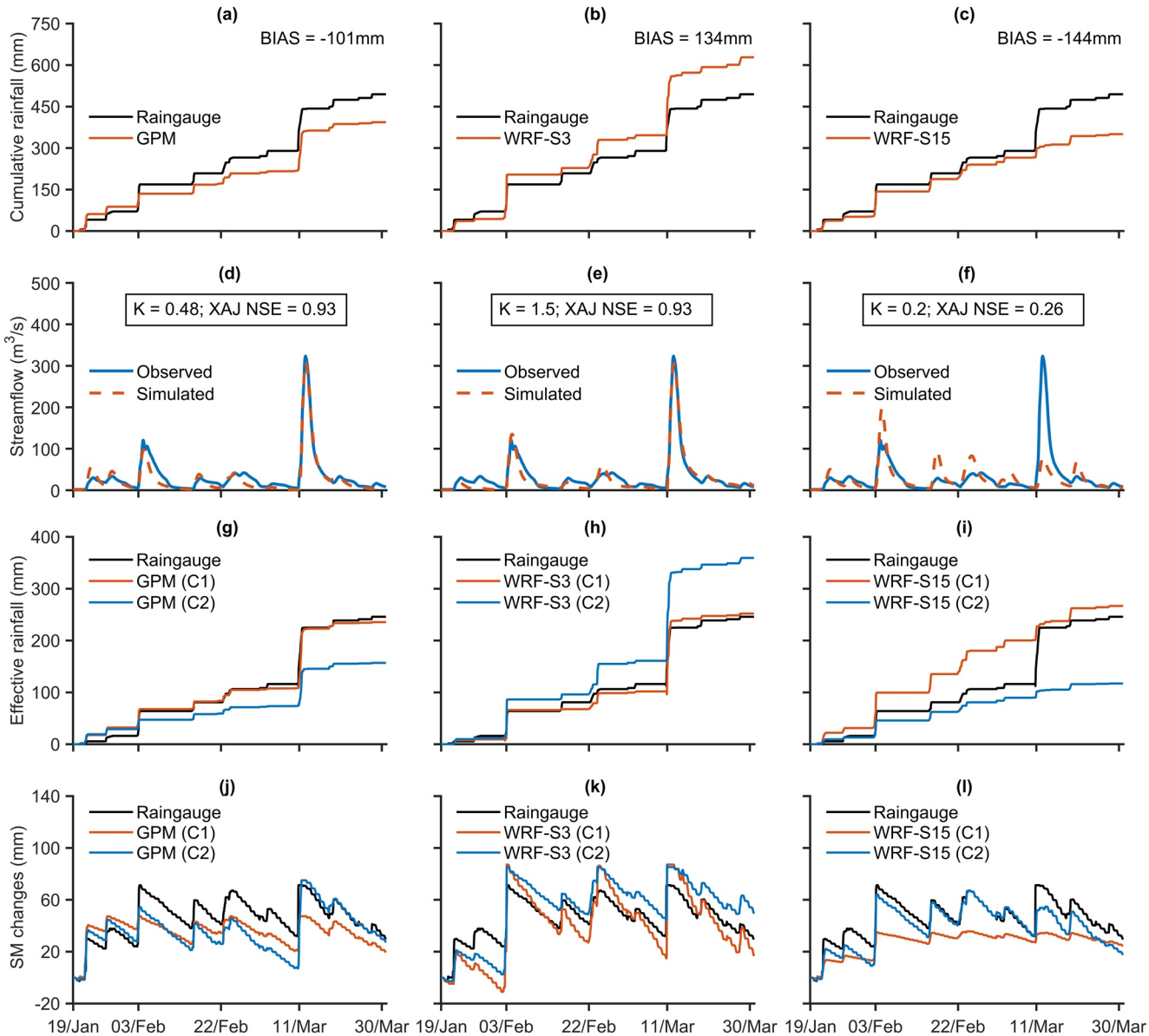

Figure 6. Cumulative rainfall (a–c), streamflow (d–f), cumulative effective rainfall (g–i), and soil moisture change (j–l) of three representative rainfall data sets. Effective rainfall is calculated by subtracting actual evapotranspiration and soil moisture changes.

soil layers. In contrast, more water would be infiltrated into the soil layers when the model is driven by overestimated rainfall data sets (Figure 6k). Particularly, soil water can follow the regular pattern of rainfall input data that consistently over- or underestimate most rainstorms (Figures 6a, 6b, 6j and 6k). However, it is difficult for SM storage to act in accordance with changeable rainfall data sets, such as those that estimate many rainfall events well but fail to detect severe rainstorms (Figures 6c and 6l).

The specific values of all XAJ model parameters for each rainfall data set are shown in Table S2. Figure 7 displays the normalized parameter values calculated by Equation 1 and the parameter sensitivity for each rainfall data set. Especially, when the rainfall input shows an overestimation (underestimation) against the rain gauge observations, K appears to be greater (smaller) than K_raingauge. This is expected since K is a vital parameter that acts directly on the overall bias of the rainfall inputs. The parameter sensitivity results can also help us to understand the contributions of different parameters to hydrological model adaptability. Although the values of most parameters vary across rainfall input data sets, K, WLM, and WDM are the most sensitive parameters for the majority of rainfall input data sets. Thus, by combining both normalized parameter values and parameter sensitivity results,

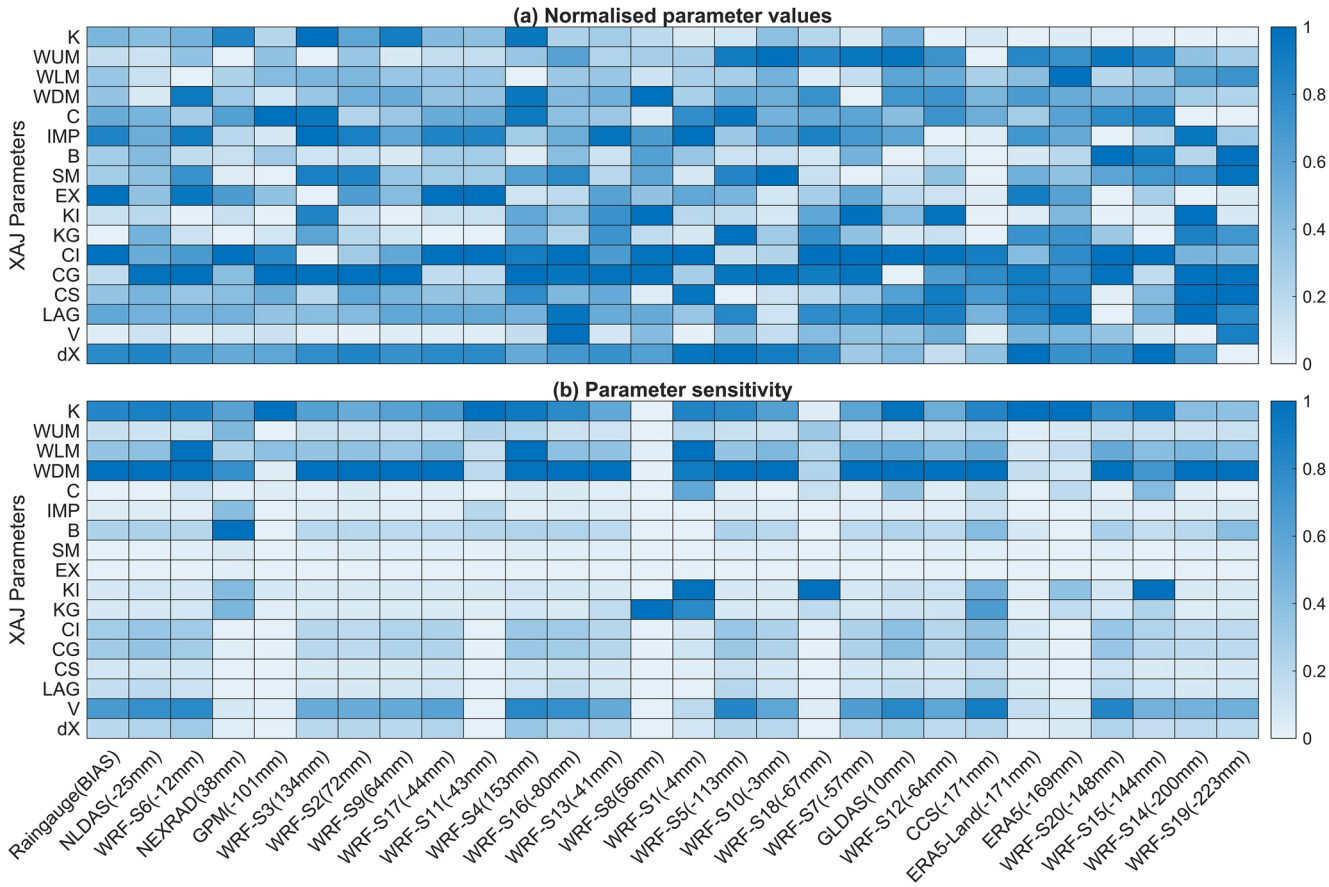

Figure 7. Normalized values of the Xin'anjiang (XAJ) parameters (a) and the parameter sensitivity for each rainfall input data set (b).

it is found that the parameters dealing with water balance significantly affect the adaptability of the hydrological model to rainfall input data.

Overall, it is the joint action of the adjustments to actual evapotranspiration and soil moisture storage in the hydrological model calibration that allow the model to adapt to the bias in rainfall data and to yield satisfactory effective rainfall, thereby reproducing streamflow well. The XAJ hydrological model is a conceptual model that assumes three soil layers in evapotranspiration and soil moisture. Their values are calibrated and adjusted for reproducing optimal performance in each XAJ model. However, there is no observed data for these two variables and we can only assume that the dynamics of these varibales are captured by the model.

3.4. Are the Findings Similar in Another Study Period?

The results of the above study are based on a single season. Are the findings from other periods the same or similar? The data from 1 January to 30 December 2014 have been used for verification. Owing to the considerable computing resources required to run the WRF model, we just select the rainfall data from five sources, rain gauges, CCS, NLDAS, ERA5, and ERA5-Land products. As stated in our previous results (Figures 1, 4 and 5 in Sections 3.1 and 3.2), NLDAS represents the good rainfall data whose bias is adapted by the hydrological model well, and CCS shows moderate performance, while ERA5 and ERA5-Land represent those data sets with significant bias. Thus, although the WRF rainfall data sets are excluded, the selected products are sufficient to demonstrate the validity of our overall conclusions. Figure 8 demonstrates some of the verification results. The corresponding modeled streamflow driven by CCS, ERA5 and ERA5-Land are shown in Figure S5 in Supporting Information S1, and the observed rainfall and streamflow of the four major rainstorms that occurred in 2014 are presented in Figure S6 in Supporting Information S1. The specific values of all XAJ model parameters calibrated for the five rainfall input data sets are shown in Table S3 in Supporting Information S1.

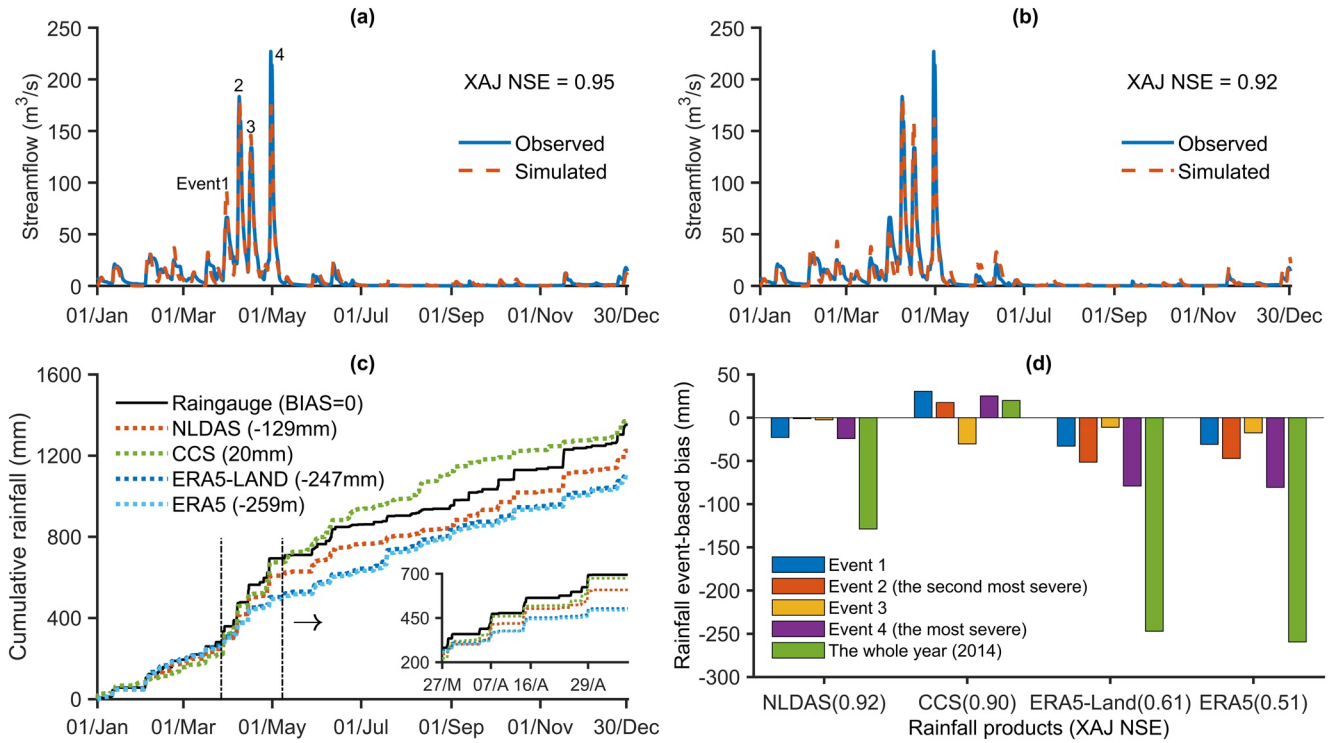

Figure 8. Results in the verification period (from 1 January 2014 to 30 December 2014), including the modeled streamflow driven by rain gauge data compared with the observed streamflow (a), the modeled streamflow driven by North American Land Data Assimilation System (NLDAS) rainfall (b), the overall cumulative rainfall of each rainfall data set (c), and the bias of the four major rainstorms in each rainfall product (d).

As shown in Figures 8b and 8c, the hydrological model adapts well to NLDAS whose large negative bias is mainly caused by consistent underestimation of rainstorms. However, the hydrological model fails to adapt to the rainfall input data whose overall bias severely exceeds the threshold. This can be seen by ERA5 and ERA5-Land, which heavily underestimate the whole study period and the most severe Event 4 (Figures 8c and 8d), and thereby fail to reproduce good streamflow, especially the peak flow (Figures S5b and S5c in Supporting Information S1). These findings suggest again that there is a threshold for the overall rainfall bias that makes the hydrological model adapt to the inaccurate rainfall inputs, and the hydrological model adaptability is further affected by the event-based rainfall bias, especially from those of the most severe rainstorms.

3.5. Implications From Hydrological Model Adaptability to Rainfall Input Quality

Based on our aforementioned analysis, the hydrological model calibration can adapt to some inaccurate rainfall data sets to a certain degree. However, such a model calibration could produce reasonable results by hiding the errors of rainfall inputs (Clark & Vrugt, 2006; Kabir et al., 2022). With the compensational ability of model recalibrations, it is challenging to evaluate or compare the reliability of rainfall data sets based on their hydrological simulation performance. Besides, even for the rain gauge data, it is difficult to be inferred from the observed streamflow data considering the catchment has an impact on amplifying (as a high pass filter) or dampening (as a low pass filter) the rainfall variability (Müller et al., 2021; Oudin et al., 2004; Xu et al., 2006).

However, does it mean there is no hint on the rainfall quality from the corresponding hydrological simulations? When we compare the hydrological simulation performance with event-based rainfall BIAS and overall rainfall BIAS (Figures 4–6 and Figure S1 in Supporting Information S1, see Section 3.2), the rainfall data set whose overall BIAS is mainly caused by significant underestimations of the severest rainstorm, is found to produce poor streamflow simulations (XAJ NSE < 0.6). While good streamflow simulations (XAJ NSE > 0.8) are generated by the rainfall data sets (a) holding a small overall bias and small bias during most rainstorm events, or (b) holding a large overall bias which is caused by consistent under- or overestimations of most rainstorms, and this can be mostly compensated by adjusting the evapotranspiration and soil moisture storage.

The dynamic model calibration can provide some implications for the accuracy of rainfall input to some extent. Some limitations of the rainfall inputs can be identified by checking the streamflow simulations. We can also get a pre-overview of the hydrological applicability of a rainfall data set in terms of overall rainfall bias and event-based rainfall bias. The cumulative rainfall curve against observations is informative to judge if a rainfall data set can reproduce streamflow well. In addition, the bias in the rainfall input data set is retained using the static calibration (C2) method (Figures 6g–6i). How the model adapts to rainfall input data can be speculated by comparing the effective rainfall results in the two calibration methods (C1 and C2).

This study provides a quantitative understanding of the lumped hydrological model's compensational capability on rainfall input data sets with varied quality. This is important since a large number of studies concluded the hydrological utility of different rainfall products based on the dynamic model calibration. It needs to be stated that the rainfall input data should be bias-corrected well before hydrological application, rather than using the biased rainfall data directly. It is not recommended to use the model calibration to compensate for the biased rainfall data. Otherwise, the good simulation outputs based on model calibration adjustment would give us wrong messages about the hydrological applicability of rainfall data sets. The results in this study are helpful for quantitative guidance on rainfall bias correction and rainfall selection before the hydrological application. Furthermore, the hydrological model would be expected to perform better with the data fusion approach that quantitatively combines various rainfall data sets. Such an approach will allow us to make the best use of different rainfall products that contain different information.

3.6. Limitation and Recommendations

Numerous studies have examined the reliability of various rainfall data sets and their corresponding hydrological utility. However, few studies have been conducted on exploring the adaptability mechanism of the hydrological model calibration on rainfall input accuracy. Our study aims to fill in the knowledge gap, but certainly, there are some limitations. Owing to the huge computing resources required to perform the hourly WRF simulation, we just use 3-month time series (covering eight rainfall events) of 28 rainfall products over a catchment as a case study. The verification is based on 1-year time series of five representative rainfall products. Further studies are necessary to explore a wider range of catchments over longer periods which covers different characteristics of rainfall events. As only the XAJ model is considered in this study, there is a question on whether the results could be generalized to other hydrological models. The concept and structure of the XAJ model are representatives of many typical conceptual hydrological models. Thus, the findings in our study can provide useful insight into other similar hydrological models (e.g., the PDM model) albeit it is expected that the threshold and patterns of hydrological model adaptability might be different depending on the specific hydrological model used. Besides, it makes little difference how PET is derived (e.g., using the energy balance method, the aerodynamic method, or the combined method) as long as PET is used driving in the hydrological model. In addition, various land surface models are gaining increasing attention in hydrologic applications. Despite increased complexities, their hydrological modeling components are still based on conventional hydrological models. For example, the latest JULES model from the UK is a cutting-edge land surface model, and its hydrological model component is based on two well-known hydrological models (Chou et al., 2022), namely, PDM and TOPMODEL. The Noah-MP land surface model has been coupled into the WRF-Hydro model in modeling subsurface flow routing, overland flow routing and channel flow routing. The widely used option in Noah-MP is the simple TOPMODEL-based runoff with SIMGM. Moreover, some studies have investigated the integrated framework of the NWP model (e.g., WRF) and hydrological models (e.g., PDM and TOPMODEL) (Rogelis & Werner, 2018; Srivastava et al., 2014). Therefore, the results are applicable to the hydrological component in various land surface models as well. Further investigation is certainly warranted to provide a detailed and mechanistic understanding of the hydrological model adaptability of rainfall quality worldwide using the proposed method in our study.

The degree of consistency in the event-based bias is found to be of great importance in this study. It might be related to the correlation metric PCC. To gain insights into how rainfall PCC during different events affects the overall streamflow simulation performance, we also re-calculate PCC values for all events, using only data extracted for that event. This allows us to see if there are any event-based patterns in model performance and if PCC can measure the importance of the estimation accuracy of individual rainfall events. As shown in Figure 9, there is no clear relationship between the XAJ NSE and the overall rainfall PCC, as well as the event-based rainfall PCC (especially the heaviest Event 6). For example, the WRF-S6 rainfall has an overall the PCC of

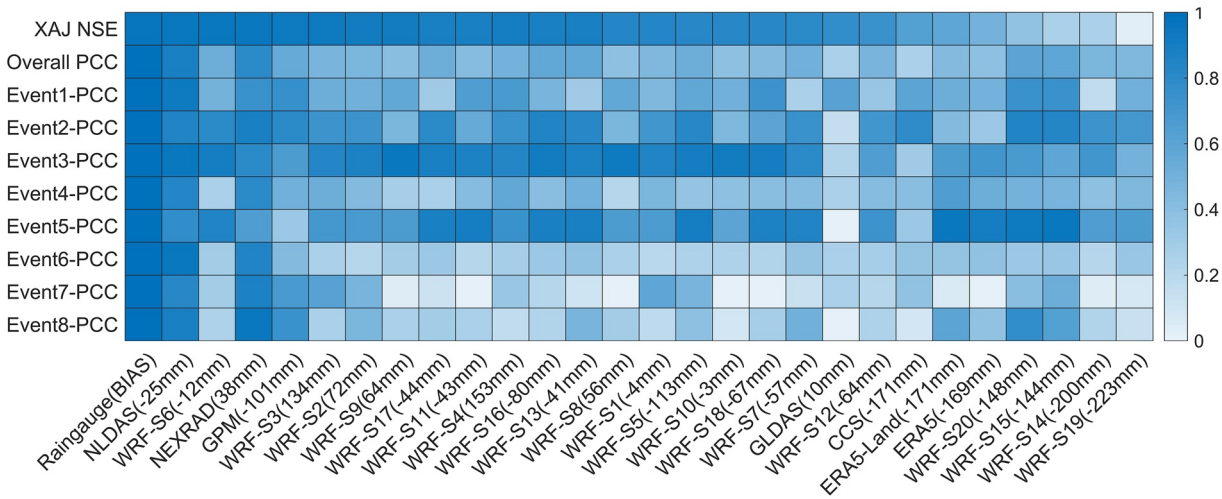

Figure 9. Relationship among the Xin'anjiang (XAJ) Nash-Sutcliffe Efficiency (NSE), the overall rainfall Pearson's Correlation Coefficient (PCC), and event-based PCC. The rainfall products are presented from left to right in descending order of their XAJ NSE values.

0.53 and PCC values of less than 0.3 in Events 4, 6, 7, and 8, it can still reproduce good streamflow simulations with a high XAJ NSE (0.94). On the other hand, the WRF S14, S15, S19, and S20 rainfall products with overall PCC scores ranging from 0.45 to 0.61, only generate streamflow simulations with a XAJ NSE inferior to 0.4. This could be due to the fact that PCC only measures the linear correlation between two sets of data and it has limitations in quantifying the bias. Thus, the degree of consistency in the event-based bias found in our study needs to be further quantified in the future. In addition, different rainfall events may have different importance degrees. An indicator combined with PCC, BIAS and the importance weighting of individual events, could be one of the potential solutions. It needs to be determined by a large number of experiments and would be a useful hydrological performance indicator for rainfall products, as well as to provide insights into rainfall bias correction.

Moreover, different metrics have different emphases and give us different insights into hydrological simulation performance. Correspondingly, the hydrological performance metric has an impact on judging the quality of rainfall input data. Since our study pays more attention to floods, NSE is used with more focus on volume errors and timing errors of high flows. It is also the most widely used in hydrological modeling. The hydrological model adaptability patterns found in our study are more related to high rainfall events. However, for other hydrological applications, such as water resources management and hydraulic structures design, the volume errors of high flows (floods) and low flows (droughts) would be given more consideration. Thus, other metrics might be more suitable in those cases. Additionally, further studies could also be conducted to explore the multi-functional hydrological models that serve both floods and droughts simultaneously. In this case, calibrating the model for high flows and low/moderate flows separately would improve hydrological simulation performance (e.g., Text S1 and Figure S7 in Supporting Information S1). There is a need to further explore the effect of different evaluation metrics on hydrological model adaptability in various hydrological applications.

Also, except for NSE, PCC, RMSE, and BIAS, some indicators related to conditional rainfall errors would add merits, particularly in analyzing how errors in large rainstorm events quantitatively affect hydrological model adaptability. This needs more research to identify the specific thresholds in different rainstorms. The threshold is also affected by the contribution (weight) of each rainfall event to the overall hydrological simulation. More indicators about overall rainfall quality, individual events and hydrological response performance, would help us get a better understanding of the hydrological model adaptability. Furthermore, different indicators may give us different facets of rainfall quality. The multi-variate framework analysis is useful for exploring the single and joint effects of different rainfall indicators on the hydrological response variable. Tree-based models can identify the non-linear and non-monotonic dependencies, particularly when we know little about the patterns and relationships between predictor variables and the response variable (Merz et al., 2013). A multivariate analysis of various rainfall indicators is expected to reveal the different characteristics of rainfall data and their interactions. However, such approaches require a large number of data sets and experiments to produce reliable results.

4. Conclusions and Summary

This study explores the adaptability mechanism of hydrological model calibration on rainfall inputs with varied quality. Twenty-eight rainfall data sets from rain gauges, weather radars, satellites, reanalysis products and WRF simulations are used as forcings to drive the lumped XAJ hydrological model. All XAJ model simulations are optimized independently using a dynamic calibration approach. Potential connections between the hydrological simulation performance and the accuracy of its rainfall inputs are investigated over the whole study period and across different rainfall events. The adaptability mechanism of the hydrological model is identified by analyzing the characteristics of rainfall errors and the compensational pattern of model calibrations. The results are verified using data from another period and similar patterns are found.

Despite some limitations, this study provides major advances in uncovering the hydrological model adaptability on rainfall input quality, providing a quantitative-based understanding of the adaptability mechanism (i.e., drivers, conditions and factors) of the model calibration to rainfall errors. The results are also helpful for quantitative guidance on bias correction, data fusion and input selection of rainfall products before the hydrological application.

We synthesize our key findings as follows. First, the hydrological model can often adapt to some inaccurate rainfall inputs showing good streamflow simulation to some extent. The overall rainfall BIAS helps in predicting hydrological model adaptability to rainfall data sets. Furthermore, the hydrological model adaptability has an adaptable threshold of overall rainfall BIAS, which is −140 mm for our study (total rainfall is about 518 mm). Second, hydrological model adaptability is not only affected by the overall rainfall BIAS but also significantly influenced by the event-based rainfall BIAS. How the event-based rainfall bias of rainstorms shapes the overall rainfall bias affects hydrological model adaptability. In particular, the rainfall accuracy during the heaviest rainstorm which results in peak flow has a large impact. The hydrological model would be unable to adapt well to the large overall rainfall bias that is mainly caused by a significant underestimation of the severest rainstorm. Third, the hydrological model is expected to adapt to a rainfall product that contains important information content for model calibration. It can adapt well to two scenarios of inaccurate rainfall data sets: (a) have a small overall rainfall bias and small rainfall bias during most rainstorms, or (b) have a large overall bias which is caused by consistent under- or overestimations of most rainstorms. Fourth, the adaptability to rainfall inputs of the lumped XAJ hydrological model is mainly controlled by a bias reduction through adjustment of evapotranspiration and soil moisture storage. Fifth, the cumulative rainfall curve against observations is beneficial to get a pre-overview of whether the hydrological model calibration can adapt to the errors in rainfall input.

More effort is needed to further explore the degree of consistency in the event-based bias of rainfall data sets. The multi-variate analysis framework involving more indicators about overall rainfall quality, individual rainstorm events and hydrological response variables, would be promising to provide a bigger picture and detailed understanding of hydrological model adaptability to rainfall quality in different hydrological scenarios.

Data Availability Statement

The Shuttle Radar Topography Mission (SRTM) 1 Arc Second data are available at https://earthexplorer.usgs.gov/; the HPD rain gauge data are available at https://www.ncei.noaa.gov/data/coop-hourly-precipitation/v2/access/; the NEXRAD Mosaic data are available at https://mesonet.agron.iastate.edu/; the GPM IMERG, NLDAS-2, and GLDAS-2.1 data are available at https://disc.gsfc.nasa.gov/; the PERSIANN-CCS data are available at https://chrsdata.eng.uci.edu/; the ERA5 and ERA5-Land data are available at https://cds.climate.copernicus.eu; and the streamflow data are available from the United States Geological Survey (USGS) stream gaging station at https://www.usgs.gov/. The detailed introduction of the WRF model is available at https://www2.mmm.ucar.edu/wrf/users/. The sensitivity analysis codes are available at https://www.safetoolbox.info/. Additional codes are available from the corresponding author upon request.

Acknowledgments
J. Wang is grateful to the China Scholarship Council and the University of Bristol for providing the joint scholarship (No 201906380142) for her Ph.D. study at the University of Bristol. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/. Special acknowledgment goes to the data set developers from the NCEI/NOAA, Iowa Environmental Mesonet, NASA, CHRS, ECMWF, and USGS for producing and sharing the data sets employed in this study, and the WRF model developers. The authors are also grateful to Editor Peter Troch, Associate Editor, and three anonymous reviewers for their insightful and constructive comments that led to substantial improvements in the paper.

References

Bárdossy, A., & Singh, S. K. (2008). Robust estimation of hydrological model parameters. *Hydrology and Earth System Sciences*, *12*(6), 1273–1283. https://doi.org/10.5194/hess-12-1273-2008

Barrett, E. C., Adler, R. F., Arpe, K., Bauer, P., Berg, W., Chang, A., et al. (1994). The first WetNet precipitation intercomparison project (PIP-1): Interpretation of results. *Remote Sensing Reviews*, *11*(1–4), 303–373. https://doi.org/10.1080/02757259409532268

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., et al. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, *21*(12), 6201–6217. https://doi.org/10.5194/hess-21-6201-2017

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, *249*(1–4), 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8

Beven, K., & Smith, P. (2015). Concepts of information content and likelihood in parameter calibration for hydrological simulation models. *Journal of Hydrologic Engineering*, *20*(1), A4014010. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000991

Beven, K. (2012). *Rainfall-runoff modelling: The primer* (2nd ed.). Wiley-Blackwell, Chichester. https://doi.org/10.1002/9781119951001

Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, *22*(10), 1509–1518. https://doi.org/10.1016/j.envsoft.2006.10.004

Chou, H. K., Heuminski de Avila, A. M., & Bray, M. (2022). Evaluating the Atibaia River hydrology using JULES6.1. *Geoscientific Model Development*, *15*(13), 5233–5240. https://doi.org/10.5194/gmd-15-5233-2022

Clark, M. P., & Vrugt, J. A. (2006). Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters. *Geophysical Research Letters*, *33*(6), L06406. https://doi.org/10.1029/2005gl025604

Copernicus Climate Change Service (C3S). (2018). *ERA5 hourly data on single levels from 1959 to present*. C3S Climate Data Store (CDS). Accessed 20 June 2021. https://doi.org/10.24381/cds.adbb2d47

Copernicus Climate Change Service (C3S). (2019). *ERA5-land hourly data from 1950 to present*. C3S Climate Data Store (CDS). Accessed 16 August 2021. https://doi.org/10.24381/cds.e2161bac

Essou, G. R. C., Arsenault, R., & Brissette, F. P. (2016). Comparison of climate datasets for lumped hydrological modeling over the continental United States. *Journal of Hydrology*, *537*, 334–345. https://doi.org/10.1016/j.jhydrol.2016.03.063

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Han, D., & Bray, M. (2006). Automated Thiessen polygon generation. *Water Resources Research*, *42*(11), W11502. https://doi.org/10.1029/2005wr004365

Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety*, *81*(1), 23–69. https://doi.org/10.1016/s0951-8320(03)00058-9

Hong, Y., Hsu, K. L., Sorooshian, S., & Gao, X. G. (2004). Precipitation estimation from remotely sensed imagery using an Artificial neural network cloud classification system. *Journal of Applied Meteorology*, *43*(12), 1834–1853. https://doi.org/10.1175/jam2173.1

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., et al. (2019). *Algorithm Theoretical Basis Document (ATBD) Version 06 NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG)*. NASA/GSFC. Retrieved from https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf

Iman, R. L., & Conover, W. J. (1980). Small sample sensitivity analysis techniques for computer models with an application to risk assessment. *Communications in Statistics – Theory and Methods*, *9*(17), 1749–1842. https://doi.org/10.1080/03610928008827996

Jiang, L. G., & BauerGottwein, P. (2019). How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China. *Journal of Hydrology*, *572*, 486–500. https://doi.org/10.1016/j.jhydrol.2019.03.042

Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M. (2012). Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. *Journal of Hydrology*, *452–453*, 213–225. https://doi.org/10.1016/j.jhydrol.2012.05.055

Kabir, T., Pokhrel, Y., & Felfelani, F. (2022). On the precipitation-induced uncertainties in process-based hydrological modeling in the Mekong River Basin. *Water Resources Research*, *58*(2), e2021WR030828. https://doi.org/10.1029/2021wr030828

Kim, K. B., Kwon, H. H., & Han, D. (2018). Exploration of warm-up period in conceptual hydrological modelling. *Journal of Hydrology*, *556*, 194–210. https://doi.org/10.1016/j.jhydrol.2017.11.015

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, *5*, 89–97. https://doi.org/10.5194/adgeo-5-89-2005

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM Journal on Optimization*, *9*(1), 112–147. https://doi.org/10.1137/S1052623496303470

Lawrimore, J. (2018). *Algorithm theoretical basis document COOP-hourly precipitation data (HPD) network version 2*. National Centers for Environmental Information (NCEI), National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce. Retrieved from https://www.ncei.noaa.gov/data/coop-hourly-precipitation/v2/doc/CHPD-v2-ATBD-20181023.pdf

Liu, C., Ikeda, K., Thompson, G., Rasmussen, R., & Dudhia, J. (2011). High-resolution simulations of wintertime precipitation in the Colorado Headwaters region: Sensitivity to physics parameterizations. *Monthly Weather Review*, *139*(11), 3533–3553. https://doi.org/10.1175/mwr-d-11-00009.1

Liu, J., Bray, M., & Han, D. (2012). Sensitivity of the Weather Research and Forecasting (WRF) model to downscaling ratios and storm types in rainfall simulation. *Hydrological Processes*, *26*(20), 3012–3031. https://doi.org/10.1002/hyp.8247

Liu, J., & Han, D. (2010). Indices for calibration data selection of the rainfall-runoff model. *Water Resources Researc*h, *46*(4), W04512. https://doi.org/10.1029/2009wr008668

Liu, Y., Chen, Y., Chen, O., Wang, J., Zhuo, L., Rico-Ramirez, M. A., & Han, D. (2021). To develop a progressive multimetric configuration optimisation method for WRF simulations of extreme rainfall events over Egypt. *Journal of Hydrology*, *598*, 126237. https://doi.org/10.1016/j.jhydrol.2021.126237

McKay, M. D., Beckman, R. J., & Conover, W. J. (2000). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *42*(1), 55–61. https://doi.org/10.1080/00401706.2000.10485979

McMillan, H., Jackson, B., Clark, M., Kavetski, D., & Woods, R. (2011). Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *Journal of Hydrology*, *400*(1–2), 83–94. https://doi.org/10.1016/j.jhydrol.2011.01.026

Merz, B., Kreibich, H., & Lall, U. (2013). Multi-variate flood damage assessment: A tree-based data-mining approach. *Natural Hazards and Earth System Sciences*, *13*(1), 53–64. https://doi.org/10.5194/nhess-13-53-2013

Mississippi Department of Environmental Quality (MDEQ). (2008). *Pearl River Basin*. MDEQ, Mississippi, United States. Retrieved from https://www.mdeq.ms.gov/wp-content/uploads/2008/10/PearlRiverBasin.pdf

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. https://doi.org/10.2307/1269043

Müller, M. F., Roche, K. R., & Dralle, D. N. (2021). Catchment processes can amplify the effect of increasing rainfall variability. *Environmental Research Letters*, *16*(8), 084032. https://doi.org/10.1088/1748-9326/ac153e

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Oudin, L., Andréassian, V., Perrin, C., & Anctil, F. (2004). Locating the sources of low-pass behavior within rainfall-runoff models. *Water Resources Research*, *40*(11), W11101. https://doi.org/10.1029/2004wr003291

Oudin, L., Perrin, C., Mathevet, T., Andréassian, V., & Michel, C. (2006). Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. *Journal of Hydrology*, *320*(1–2), 62–83. https://doi.org/10.1016/j.jhydrol.2005.07.016

Paturel, J. E., Servat, E., & Vassiliadis, A. (1995). Sensitivity of conceptual rainfall-runoff algorithms to errors in input data — Case of the GR2M model. *Journal of Hydrology*, *168*(1–4), 111–125. https://doi.org/10.1016/0022-1694(94)02654-T

Pianosi, F., Sarrazin, F., & Wagener, T. (2015). A MATLAB toolbox for global sensitivity analysis. *Environmental Modelling & Software*, *70*, 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009

Raimonet, M., Oudin, L., Thieu, V., Silvestre, M., Vautard, R., Rabouille, C., & Le Moigne, P. (2017). Evaluation of gridded meteorological datasets for hydrological modeling. *Journal of Hydrometeorology*, *18*(11), 3027–3041. https://doi.org/10.1175/jhm-d-17-0018.1

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., et al. (2011). High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate*, *24*(12), 3015–3048. https://doi.org/10.1175/2010jcli3985.1

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, *46*(5), W05521. https://doi.org/10.1029/2009wr008328

Rodell, M., Houser, P. R., Jambor, U. E. A., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rogelis, M. C., & Werner, M. (2018). Streamflow forecasts from WRF precipitation for flood early warning in mountain tropical areas. *Hydrology and Earth System Sciences*, *22*(1), 853–870. https://doi.org/10.5194/hess-22-853-2018

Rui, H., & Mocko, D. (2019). *Readme document for North America Land Data Assimilation system phase 2 (NLDAS-2) products*. NASA Goddard Earth Science Data Information and Services Center (GES DISC). Retrieved from https://hydro1.gesdisc.eosdis.nasa.gov/data/NLDAS/README.NLDAS2.pdf

Rui, H., & Beaudoing, H. (2019). *Readme document for NASA GLDAS Version 2 Data products*. NASA Goddard Earth Science Data Information and Services Center (GES DISC). Retrieved from https://data.mint.isi.edu/files/raw-data/GLDAS_NOAH025_M.2.0/doc/README_GLDAS2.pdf

Schultz, D. M., Fairman, J. G., Kirshbaum, D. J., Gray, S. L., & Barrett, A. I. (2016). Climatology of banded precipitation over the contiguous United States. *Monthly Weather Review*, *144*(12), 4553–4568. https://doi.org/10.1175/mwr-d-16-0015.1

Singh, S. K., & Bárdossy, A. (2012). Calibration of hydrological models on hydrologically unusual events. *Advances in Water Resources*, *38*, 81–91. https://doi.org/10.1016/j.advwatres.2011.12.006

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., & Islam, T. (2014). Sensitivity and uncertainty analysis of mesoscale model downscaled hydro-meteorological variables for discharge prediction. *Hydrological Processes*, *28*(15), 4419–4432. https://doi.org/10.1002/hyp.9946

Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, *24*(5), 2527–2544. https://doi.org/10.5194/hess-24-2527-2020

Wersal, R. M., Madsen, J. D., & Tagert, M. L. (2006). *Aquatic plant survey of Ross Barnett Reservoir for 2005*. GeoResources Institute, Mississippi State University.

Woodley, W. L., Olsen, A. R., Herndon, A., & Wiggert, V. (1975). Comparison of gage and radar methods of convective rain measurement. *Journal of Applied Meteorology and Climatology*, *14*(5), 909–928. https://doi.org/10.1175/1520-0450(1975)014<0909:COGARM>2.0.CO;2

Xia, Y., Mitchell, K., Ek, M., Cosgrove, B., Sheffield, J., Luo, L., et al. (2012). Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *Journal of Geophysical Research*, *117*(D03110), 1–23. https://doi.org/10.1029/2011jd016051

Xu, C. Y., Tunemar, L., Chen, Y. D., & Singh, V. P. (2006). Evaluation of seasonal and spatial variations of lumped water balance model sensitivity to precipitation data errors. *Journal of Hydrology*, *324*(1–4), 80–93. https://doi.org/10.1016/j.jhydrol.2005.09.019

Yang, X., Magnusson, J., Huang, S., Beldring, S., & Xu, C. Y. (2020). Dependence of regionalization methods on the complexity of hydrological models in multiple climatic regions. *Journal of Hydrology*, *582*, 124357. https://doi.org/10.1016/j.jhydrol.2019.124357

Zhao, R. J. (1992). The Xinanjiang model applied in China. *Journal of Hydrology*, *135*(1–4), 371–381. https://doi.org/10.1016/0022-1694(92)90096-E

Zhao, R. J., Zhang, Y. L., & Fang, L. R. (1980). The Xinanjiang model. In *Hydrological Forecasting Proceedings Oxford Symposium, IAHS, Oxford* (Vol. 129, pp. 351–356).

Zhuo, L., Dai, Q., & Han, D. (2015). Meta-analysis of flow modeling performances—To build a matching system between catchment complexity and model types. *Hydrological Processes*, *29*(11), 2463–2477. https://doi.org/10.1002/hyp.10371

Zhuo, L., & Han, D. (2016a). Could operational hydrological models be made compatible with satellite soil moisture observations? *Hydrological Processes*, *30*(10), 1637–1648. https://doi.org/10.1002/hyp.10804

Zhuo, L., & Han, D. (2016b). Misrepresentation and amendment of soil moisture in conceptual hydrological modelling. *Journal of Hydrology*, *535*, 637–651. https://doi.org/10.1016/j.jhydrol.2016.02.033

Zhuo, L., Han, D., Dai, Q., Islam, T., & Srivastava, P. K. (2015). Appraisal of NLDAS-2 multi-model simulated soil moistures for hydrological modelling. *Water Resources Management*, *29*(10), 3503–3517. https://doi.org/10.1007/s11269-015-1011-1