Comparing the Hydrological Responses of Conceptual and Process-Based Models with Varying Rain Gauge Density and Distribution

Zhaokai Yin 1,*, Weihong Liao 2,*, Xiaohui Lei 2,3, Hao Wang 2,3 and Ruojia Wang 4,5

1 State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, Tianjin 300072, China; yinzhaokai@tju.edu.cn
2 State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China; lxh@iwhr.com (X.L.);
wanghao@iwhr.com (H.W.)
3 Science and Technology Innovation Center for Smart Water, Northeastern University, Shenyang 110819, China
4 Department of Information Management, Peking University, Beijing 100871, China; Wangruojia@pku.edu.cn
5 Institute of Ocean Research, Peking University, Beijing 100871, China

* Correspondence: liaowh@iwhr.com; Tel.: +86-10-6878-5503

Received: 23 August 2018; Accepted: 5 September 2018; Published: 7 September 2018

Abstract: Precipitation provides the most crucial input for hydrological modeling. However, rain gauge networks, the most common precipitation measurement mechanisms, are sometimes sparse and inadequately distributed in practice, resulting in an imperfect representation of rainfall spatial variability. The objective of this study is to analyze the sensitivity of different model structures to the different density and distribution of rain gauges and evaluate their reliability and robustness. Based on a rain gauge network of 20 gauges in the Jinjiang River Basin, south-eastern China, this study compared the performance of two conceptual models (the hydrologic model (HYMOD) and Xinanjiang) and one process-based distributed model (the water and energy transfer between soil, plants and atmosphere model (WetSpa)) with different rain gauge distributions. The results show that the average accuracy for the three models is generally stable as the number of rain gauges decreases but is sensitive to changes in the network distribution. HYMOD has the highest calibration uncertainty, followed by Xinanjiang and WetSpa. Differing model responses are consistent with changes in network distribution, while calibration uncertainties are more related to model structures.

Keywords: rainfall-runoff modeling; rainfall; rain gauge network; sensitivity analysis

1. Introduction

Precipitation is one of the most crucial inputs in catchment runoff modeling and measuring rainfall is essential for determining hydrological catchment response [1,2]. However, because precipitation is generated by extremely complicated, non-linear, and sensitive atmospheric physical process [3], it shows highly spatial and temporal variability at the basin scale [4–7].

Due to the importance of precipitation in modeling, prior work has evaluated the effect of uncertainty in precipitation on the response of simulated streams in hydrological models. Wilson et al. [8] found that peak runoff, total runoff volume, and peak timing are considerably influenced by the spatial distribution and precision of rainfall measurements. Singh [9] provided a detailed literature review on the influence of spatial–temporal variability in hydrological factors on rainfall runoff modeling. Sun et al. [10] found that runoff prediction errors at the catchment scale were significantly related to the representation of rainfall data spatial variability. Shen et al. [11] showed
that noticeable uncertainty in stream flow and non-point source pollution modeling was caused by spatial rainfall uncertainty obtained from different precipitation interpolation methods.

Traditionally, there are three mechanisms for measuring or observing precipitation: rain gauges, weather radar, and satellite-based sensors [12–14]. Rain gauge networks remain the most common method for measuring precipitation [15–17] due to their higher accuracy in representing rainfall at their respective locations [18,19] and longer recording period for investigating long-term rainfall runoff processes [20]. Because a rain gauge is a point measurement for precipitation, representing rainfall spatial variability must be affected by the density and distribution of rain gauge networks. To address this relationship, Pardo-Igúzquiza [21] tried to optimize the rain gauge distribution and quantity in a given catchment. Chen et al. [22] proved that more pluviometers distributed in a catchment resulted in higher precision in areal precipitation calculation. Xu et al. [23] and Girons Lopez et al. [24] found there is a threshold value of rain gauge density, before which the presentation of rainfall improves significantly with increasing rain gauge numbers.

Significant early research focused on the effect of density and distribution of rain gauge networks on catchment modeling. However, most were only concerned with the impact of density. Michaud and Sorooshian [25] showed that poor rain gauge densities caused an inadequate simulation of flood peak in a midsized semi-arid catchment. Chaplot et al. [26] selected rain gauges with a certain empirical distribution. Andreassian et al. [27], Anctil et al. [28], and Xu et al. [23] randomly selected several subsets of different numbers of gauges to achieve the same level of rain gauge density and evaluated the relationship between density and modeling results. Similarly, Drogue and Khediri [29] employed approximately 100 subsets with the same number of rain gauges and used the mean average of rainfall data for their evaluated.

In the study of Bardossy and Das [30], distributions of rain gauges with different densities were obtained from an optimization algorithm. Some studies have considered pluviometer density and distribution at the same time [31], while other researchers have investigated the effect of udometer distribution on conceptual models [27,29,32]. These studies have evaluated the individual responses of different process-based models [26,31], compared a couple of models [33], and even applied a neural network to rainfall-runoff modeling [28]. However, a comprehensive comparison of the performances of different models impacted by different rain gauge network distributions, especially networks with varying numbers of gauges, has yet to be presented.

Identifying the effect of different rain gauges' distribution on the hydrological response of conceptual and process-based models will be helpful in (1) optimizing rain gauge networks and (2) selecting hydrological models for a given basin. Therefore, this study aims to advance the discipline by comparing the impact of rain gauge distribution on typical conceptual and process-based distributed models. Furthermore, the potential influence of uncertainty in model calibrations is investigated. The paper is organized as follows: Section 2 introduces the study area, datasets, applied models and design scenarios. Section 3 provides the simulation results in two parts: for the whole validation period and a typical month within the validation period. Section 4 discusses the results and compares them with previous research. Section 5 provides the study conclusions.

2. Materials and Methods

2.1. Study Area

The Jinjiang River Basin is located in the north-west Ganjiang River Basin, China, with a drainage area of 6215 km² above the Gaoan hydrological station (Figure 1). The elevation of the catchment is higher in the north-west, where most tributaries of the Jinjiang River originate, ranging from 18 to 1096 m above sea level. The Jinjiang River Basin is located in the subtropical region with a warm and humid climate and receives approximately 1300–2100 mm of precipitation per year, with average runoff reaching about 184 m³/s at Gaoan station based on measured data. As the second largest tributary of the Ganjiang River, the Jinjiang River flows into the Ganjiang River about 30 km above
Nanchang City, the capital city of Jiangxi Province, China. No controlled reservoirs have been built on the main stream of the Jinjiang River, resulting in a direct threat to Nanchang City during flood periods.

Figure 1. Map of the study area showing the locations of streams, stream gauges, and rainfall gauges overlain on a digital elevation model (DEM) of the Jinjiang River Basin.

2.2. Datasets

Daily precipitation data from 20 rain gauges and observed streamflow data from the Gaoan hydrological station for 2008–2013 were provided by the Jiangxi Hydrological Bureau. The distribution of the pluviometers with their names and location of the Gaoan hydrological station are shown in Figure 1. The Thiessen polygon method [34] was used to interpolate the recorded rain data in the sub-basin areas.

Other meteorological data, including temperature, wind speed, relative humidity, and solar radiation, was obtained from the China Meteorological Assimilation Driving Datasets for the SWAT (Soil and Water Assessment Tool) model (CMADS) version 1.1 [35,36]. The CMADS datasets were developed based on the China Meteorological Assimilation Land Data Assimilation System (CLADS) with a temporal resolution of 1 day and spatial resolution of 0.25° × 0.25° [37]. The datasets were tested in the Heihe River Basin, Manas River Basin, and Hunhe River Basin [37–39], demonstrating a reliable performance in reflecting the observed meteorological data for China.

The digital elevation map (DEM) was developed by the Data Center for Resources and Environment Sciences, Chinese Academy of Sciences (RESDC), at a resolution of 1 km × 1 km. Land-use data with the same resolution as the DEM for the basin in 2000 were developed by the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences. The 10 km × 10 km cell soil data and related soil physical data were gathered from the Institute of Soil Science, Chinese Academy of Sciences.

2.3. Models

In this study, three hydrological models were applied to quantify hydrologic responses. Each model was a typical representation of a different model type, ranging from conceptual to physically based: the hydrologic model (HYMOD) [40], Xinanjiang model (XAJ) [41], and water and energy transfer between soil, plants and atmosphere model (WetSpa) [42]. All three models were used as semi-distributed models with the same watershed subdivisions shown in Figure 1.
Each model operated on every sub-watershed separately with the same parameters while the rainfall and meteorological data were unique. The outputs from upper sub-basins were used as the inputs to the following sub-basins. Total streamflow for the entire basin was generated through all sub-basins according to hydraulic connections using the Muskingum method. Observed daily rainfall, potential evapotranspiration and runoff data at the Gaoan station were used as input data for all three models. The WetSpa model required the DEM, land use, and soil type for the study area.

2.3.1. Hydrologic Model (HYMOD)

HYMOD is a watershed-scale conceptual rainfall-runoff model based on the probability-distributed theory introduced by Moore [40]. It has been used in several rainfall runoff modeling studies [43–46]. This model describes runoff generation using a rainfall excess model and two sets of linear tanks (one slow flow tank and three identical tanks for quick flow) to reflect the routing process. The five parameters used in HYMOD are provided in Table 1; schematics for its application can be found in prior work [44,46].

2.3.2. Xinanjiang Model (XAJ)

The XAJ model, developed by Zhao et al. [41,47], has been widely applied in rainfall runoff simulation, hydrologic forecasting, and other applications [48,49]. This model is more parameterized in describing the details of hydrological responses than HYMOD because it has three modules: three-layer evapotranspiration, runoff generation, and runoff routing [47,50]. XAJ describes the underlying surface of the study area as three vertical layers, upper, lower, and deep. Runoff generation, routing, and evaporation are separately performed in the three layers. Descriptions of the 15 parameters, and their ranges, are provided in Table 1; the calculation process is summarized in Li et al. [50].

2.3.3. Water and Energy Transfer Between Soil, Plants and Atmosphere Model (WetSpa)

WetSpa is a typical physically based distributed model initially proposed by Wang et al. [42] and further perfected and applied to simulating hydrological processes by many researchers [51–54]. In this model, hydrological processes include precipitation, evapotranspiration, interception, infiltration, surface runoff, depression storage, percolation, ground water drainage, and interflow. Water and energy balances are considered within horizontal grid cells and in four vertical layers, the canopy, root layer, transmission layer, and saturated zone. Therefore, the model accounts for the highly spatial variance in meteorological data, terrain, land cover, and soil across a catchment. Values of some parameters, such as permeability and evapotranspiration rate, in a single grid cell can be determined from a priori knowledge from land use and soil types. Global parameters still need to be calibrated; some are presented in Table 1.

2.3.4. Model Calibration and Validation

Calibration is a necessary procedure for all hydrological models before simulating catchment rainfall runoff. The three described models are calibrated against 2009–2011 observed daily streamflow data with a 1-year period (2008) for spin-up. The dynamically dimensioned search algorithm (DDS) [55] is used to optimize parameters for each model. The DDS algorithm is a heuristic global random search algorithm, in which the search is performed in all dimensions of the decision space and the number of dimensions is reduced as the search continues. This process makes the DDS a very efficient optimization algorithm, which has been successfully applied to simulating rainfall runoff and predicting hydrology [56,57]. With a total 1000 search iterations, the model parameters are optimized by maximizing the objective function, the Nash–Sutcliffe efficiency (NSE) [58], given by:

\[
NSE = 1 - \frac{\sum_{i=1}^{N} (obs_i - sim_i)^2}{\sum_{i=1}^{N} (obs_i - \bar{obs})^2},
\]

(1)
in which \( obs \) and \( sim \) are the observed and simulated daily runoff, respectively, \( \bar{obs} \) is the mean value of observed runoff, \( i \) is the \( i \)th day, and \( N \) is the number of days. A value of 1 for NSE represents a perfect fit. Considering the equifinality in model calibration \([59]\), each model is calibrated for 30 trials to achieve a better statistical representativeness of the results in the validation period.

### Table 1. Hydrologic model (HYMOD), Xinanjiang model (XAJ), and water and energy transfer between soil, plants and atmosphere model (WetSpa) parameters.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>Description</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYMOD</td>
<td>Alpha</td>
<td>Partitioning factor between quick and slow flow</td>
<td>0.60–0.99</td>
</tr>
<tr>
<td></td>
<td>( b_{exp} )</td>
<td>Degree of spatial variability of soil moisture capacity in the catchment</td>
<td>0.00–2.00</td>
</tr>
<tr>
<td></td>
<td>( C_{max} ) (mm)</td>
<td>Maximum soil moisture capacity in the catchment</td>
<td>0.00–1000</td>
</tr>
<tr>
<td></td>
<td>( R_q ) (day)</td>
<td>Residence time of the linear quick flow tanks</td>
<td>0.00–0.99</td>
</tr>
<tr>
<td></td>
<td>( R_s ) (day)</td>
<td>Residence time of the linear slow flow tank</td>
<td>0.001–0.10</td>
</tr>
<tr>
<td></td>
<td>( U_M ) (mm)</td>
<td>Averaged soil moisture storage capacity of the upper layer</td>
<td>5–100</td>
</tr>
<tr>
<td></td>
<td>( L_M ) (mm)</td>
<td>Averaged soil moisture storage capacity of the lower layer</td>
<td>50–300</td>
</tr>
<tr>
<td></td>
<td>( D_M ) (mm)</td>
<td>Averaged soil moisture storage capacity of the deep layer</td>
<td>5–100</td>
</tr>
<tr>
<td></td>
<td>( B )</td>
<td>Exponential of the distribution to tension water capacity</td>
<td>0–1</td>
</tr>
<tr>
<td>XAJ</td>
<td>( I_m ) (%)</td>
<td>Percentage of impervious and saturated areas in the catchment</td>
<td>0–0.1</td>
</tr>
<tr>
<td></td>
<td>( C )</td>
<td>Coefficient of the deep layer</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>( S_w ) (mm)</td>
<td>Exponent of the free water capacity curve influencing the development of the saturated area</td>
<td>1.0–1.5</td>
</tr>
<tr>
<td></td>
<td>( K_f )</td>
<td>Outflow coefficients of the free water storage to groundwater relationships</td>
<td>0.05–0.70</td>
</tr>
<tr>
<td></td>
<td>( K_i )</td>
<td>Outflow coefficients of the free water storage to interflow relationships</td>
<td>( K_f + K_i = 0.7–0.8 )</td>
</tr>
<tr>
<td></td>
<td>( C_x )</td>
<td>Recession constants for groundwater storage</td>
<td>0.9–0.999</td>
</tr>
<tr>
<td></td>
<td>( C_l )</td>
<td>Recession constants for lower interflow storage</td>
<td>0.05–0.95</td>
</tr>
<tr>
<td></td>
<td>( C_s )</td>
<td>Recession constant in the lag and route method for routing through the channel system within each sub-basin</td>
<td>0.001–0.8</td>
</tr>
<tr>
<td></td>
<td>( L ) (day)</td>
<td>Lag in time empirical value</td>
<td>0–3</td>
</tr>
<tr>
<td>WetSpa</td>
<td>( K_i )</td>
<td>Interflow scaling factor</td>
<td>0–10</td>
</tr>
<tr>
<td></td>
<td>( K_g(\text{day}^{-1}) )</td>
<td>Groundwater recession coefficient</td>
<td>0–0.05</td>
</tr>
<tr>
<td></td>
<td>( K_r )</td>
<td>Initial soil moisture factor</td>
<td>0–2</td>
</tr>
<tr>
<td></td>
<td>( K_d(\text{mm} \cdot \text{day}^{-1}) )</td>
<td>Correction factor for potential evapotranspiration</td>
<td>0–2</td>
</tr>
<tr>
<td></td>
<td>( K_{gw}(\text{mm}) )</td>
<td>Initial groundwater storage</td>
<td>0–500</td>
</tr>
<tr>
<td></td>
<td>( K_{gw}(\text{mm}) )</td>
<td>Groundwater storage scaling factor</td>
<td>0–2000</td>
</tr>
<tr>
<td></td>
<td>( K_t(\circ \text{C}) )</td>
<td>Base temperature for snowmelt</td>
<td>−1–1</td>
</tr>
<tr>
<td></td>
<td>( K_{d}(\text{mm} \cdot \text{C}^{-1} \cdot \text{day}^{-1}) )</td>
<td>Temperature degree-day coefficient</td>
<td>0–10</td>
</tr>
<tr>
<td></td>
<td>( K_{rd}(\text{C}^{-1} \cdot \text{day}^{-1}) )</td>
<td>Rainfall degree-day coefficient</td>
<td>0–0.05</td>
</tr>
<tr>
<td></td>
<td>( K_{so}(\text{mm}) )</td>
<td>Surface runoff coefficient</td>
<td>0–5</td>
</tr>
<tr>
<td></td>
<td>( K_{p}(\text{mm}) )</td>
<td>Rainfall scaling factor</td>
<td>0–500</td>
</tr>
</tbody>
</table>

* Parameter ranges for HYMOD from Moradkhani et al. [45]. Parameter ranges for XAJ from Li et al. [50] and Lü et al. [60]. The selection and range in parameters for WetSpa from Shafii et al. [61].

The optimized parameter values from the calibration are used to model runoff for 2012–2013. The statistical criteria used to quantitatively evaluate model performance are relative bias (BIAS), reflecting the ability of recreating the water balance; the Pearson correlation coefficient (CC), which manifests the agreement between the simulated and the observed runoff; and NSE, as described above, which evaluates the goodness of the model results. The first two indexes are given by the following expressions:

\[
\text{BIAS} = \frac{\sum_{i=1}^{N} (\text{sim}_i - \text{obs}_i)}{\sum_{i=1}^{N} \text{obs}_i} \times 100\%,
\]

\[
\text{CC} = \frac{1}{N} \frac{\sum_{i=1}^{N} (\text{sim}_i - \overline{\text{sim}})(\text{obs}_i - \overline{\text{obs}})}{\sigma_{\text{sim}}\sigma_{\text{obs}}},
\]
where, $\sigma_{\text{sim}}$ and $\sigma_{\text{obs}}$ are standard deviations of simulated and observed runoff series, $\overline{x_{\text{sim}}}$ is the average value of assumed runoff, and the other symbols are the same as defined for Equation (1). The perfect value for BIAS is 0; for CC it is 1. The standard deviation $\sigma$ is calculated as follows:

$$
\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \overline{x})^2},
$$

where, $x_i$ is the $i$th value of observed or simulated daily runoff, $\overline{x}$ is the arithmetic average value of observed or simulated daily runoff, and the remaining symbols are the same as defined previously.

2.3.5. Scenario Settings

The correlations between rain data for a single gauge and streamflow data for Gaoan station are provided in Table 2. Notably, all rain gauges have higher relevance with runoff data with 1 day or 2 days delay, even for LZH, which is the nearest udometer to Gaoan station. However, not a single gauge is significantly related to the streamflow data, and the reverse is the same. To estimate the model uncertainties resulting from the accuracy of precipitation measurements, a total of 16 scenarios with varying rain gauge numbers and distributions are evaluated in this study. As presented in Table 2, rain gauge partition is described in terms of the direction that the barycenter selected udometer sets. Most rain gauges are distributed in one direction, while at least one rain gauge is set to the opposite direction to provide necessary rainfall spatial variety. According to the shape of Jinjiang River Basin, the directions are classified as central, north-east, north-west, south-east, and south-west. Given the five rain gauge distributions and four selected rain gauge quantities, i.e., 5, 10, 15, and 20, different pluviometer densities are defined in the study area. For each direction, rain gauges in scenarios with lower quantity are selected from those with higher quantity, i.e., 10 rain gauges in scenario 10_C are selected from those in scenario 15_C, while udometers in 10_SE are selected from those from 15_SE. Therefore, higher quantity scenarios contain all rain gauges in lower quantity scenarios in the same direction.
Table 2. Correlation between rain and streamflow data and gauge selection for different scenarios.

<table>
<thead>
<tr>
<th>Gauge Name</th>
<th>CC_0 *</th>
<th>CC_1</th>
<th>CC_2</th>
<th>CC_3</th>
<th>20_C **</th>
<th>15_C</th>
<th>10_C</th>
<th>5_C</th>
<th>15_SW</th>
<th>15_SE</th>
<th>15_NW</th>
<th>15_NE</th>
<th>10_SW</th>
<th>10_SE</th>
<th>10_NW</th>
<th>10_NE</th>
<th>5_SW</th>
<th>5_SE</th>
<th>5_NW</th>
<th>5_NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSH</td>
<td>0.142</td>
<td>0.492</td>
<td>0.581</td>
<td>0.360</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CSH</td>
<td>0.131</td>
<td>0.496</td>
<td>0.532</td>
<td>0.328</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CHP</td>
<td>0.139</td>
<td>0.548</td>
<td>0.575</td>
<td>0.363</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CH</td>
<td>0.145</td>
<td>0.459</td>
<td>0.538</td>
<td>0.356</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>FCH</td>
<td>0.158</td>
<td>0.564</td>
<td>0.589</td>
<td>0.378</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>FP</td>
<td>0.140</td>
<td>0.538</td>
<td>0.574</td>
<td>0.403</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>HSH</td>
<td>0.143</td>
<td>0.478</td>
<td>0.568</td>
<td>0.365</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>LSH</td>
<td>0.151</td>
<td>0.512</td>
<td>0.389</td>
<td>0.376</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>LX</td>
<td>0.147</td>
<td>0.512</td>
<td>0.569</td>
<td>0.361</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>LZH</td>
<td>0.143</td>
<td>0.534</td>
<td>0.537</td>
<td>0.337</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>LC</td>
<td>0.143</td>
<td>0.562</td>
<td>0.584</td>
<td>0.355</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NG</td>
<td>0.128</td>
<td>0.513</td>
<td>0.547</td>
<td>0.337</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SNHW</td>
<td>0.136</td>
<td>0.523</td>
<td>0.547</td>
<td>0.346</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XG</td>
<td>0.148</td>
<td>0.510</td>
<td>0.548</td>
<td>0.357</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XY</td>
<td>0.151</td>
<td>0.497</td>
<td>0.592</td>
<td>0.386</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XF</td>
<td>0.155</td>
<td>0.568</td>
<td>0.569</td>
<td>0.361</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>XJ</td>
<td>0.140</td>
<td>0.539</td>
<td>0.593</td>
<td>0.373</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>YGL</td>
<td>0.161</td>
<td>0.548</td>
<td>0.552</td>
<td>0.355</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>YQ</td>
<td>0.145</td>
<td>0.478</td>
<td>0.569</td>
<td>0.383</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ZHSH</td>
<td>0.142</td>
<td>0.492</td>
<td>0.581</td>
<td>0.360</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

* CC_0, CC_1, CC_2, and CC_3 refer to the correlation between rain and streamflow data on the same day and after 1 day, 2 days, and 3 days delay; ** The number in front of the underline is the number of rain gauges used in the scenario and the letters followed by the underline refer to the direction of most rain gauges: C is central and SW, SE, NW, NE indicate south-west, south-east, north-west, and north-east, respectively; *** The ticks refer to the rain gauges selected in the named scenario.
3. Results

Daily streamflow reproduction performance of the three models for 16 rain gauge scenarios during the validation period is investigated in two parts. In part I, general model performance during the entire validation period (1 January 2012 to 31 December 2013) is assessed using the statistical results from the evaluating indicators described in Section 2.3.4. In part II, model performance is discussed for a single month (June 2012) with a detailed daily flow hydrograph comparison.

3.1. Model Performance in the Calibration and Validation Period

To compare model performance between HYMOD, XAJ, and WetSpa for the 16 rain gauge scenarios, the statistical criteria in the calibration and validation period are provided in Tables 3–5. Boxplots for the validation period are shown in Figure 2.

Table 3 indicates stable NSE and CC values between the calibration and validation periods for HYMOD in all 16 scenarios. However, there is a clear improvement in the BIAS value from the calibration to validation period. XAJ shows similar behavior (Table 4). WetSpa shows a slight improvement in NSE and CC values in most scenarios (Table 5).

Figure 2a–c illustrates the accuracy of the streamflow simulations, as indicated by NSE. Generally, metrics for the validation period are acceptable for all three models despite different pluviometer selection scenarios; average NSE values are greater than 0.6. In more detail, results for XAJ are better than those for HYMOD and WetSpa, and the latter two are comparable in most scenarios. Notably, there are discrepancies between the NSE values for the three models. Interquartile ranges for HYMOD are significantly larger than those for XAJ, followed by those for WetSpa, which has the narrowest interquartile NSE range. WetSpa shows the best stability in model performance in every calibration trial, while HYMOD has the largest uncertainty.

When rain gauge selections are taken into consideration, results for all three models perform consistently. With the reduction in rain gauge numbers (scenario 20_C, 15_C, 10_C and 5_C), the mean NSE values for all models show lower variability. The NSE amplitudes for HYMOD and XAJ are rather small, showing robustness for a simple decline in the quantity of udometers. In contrast, the values for WetSpa show greater sensitivity toward changes in rain gauge quantity.

The impacts of the spatial distribution of rain gauges show similar trends in terms of NSE metrics for all models. With the same number of rain gauges (15), all models in scenario 15_SE and 15_NW perform comparably with those in 15_C. However, the accuracy of the results in scenario 15_SW is generally lower. Unexpectedly, model performances for 10_SE and 5_SE are poorest among scenarios with the same number of rain gauges. Similarly, model performance worsens with a decrease in rain gauges from 15 to 5 when the rain gauges are concentrated in the north-west part of the basin (scenarios 15_NW, 10_NW, and 5_NW). Furthermore, variations in the number of rain gauges concentrated in the south-west and north-east has few impacts on streamflow simulations, except for WetSpa in scenario 5_NE, which shows a dramatic reduction in both average and variation in NSE values. Similar results are found for the CC values, as illustrated in Figure 2d–f.

Figure 2g–h shows the ability of the three models to recreate the water balance for the 16 rain gauge scenarios. Generally, HYMOD and XAJ exhibit better performance in water balance simulations; HYMOD overestimates and XAJ underestimates the average within about 5%. WetSpa poorly estimates water balance with a BIAS of approximately 20%. Interquartile ranges for BIAS indicate that WetSpa has most stable model performance, while HYMOD has the lowest under multiple calibrations, as indicated previously.
Table 3. HYMOD model performance in the calibration and validation periods.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Nash–Sutcliffe Efficiency (NSE)</th>
<th>Relative bias (BIAS) (%)</th>
<th>Nash–Sutcliffe Efficiency (NSE)</th>
<th>Relative bias (BIAS) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration</td>
<td>Validation</td>
<td>Calibration</td>
<td>Validation</td>
</tr>
<tr>
<td></td>
<td>Mean *</td>
<td>Q1 **</td>
<td>Q3 ***</td>
<td>Mean</td>
</tr>
<tr>
<td>20_C **</td>
<td>0.773</td>
<td>0.739</td>
<td>0.835</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td>0.768</td>
<td>0.732</td>
<td>0.823</td>
<td>0.883</td>
</tr>
<tr>
<td>15_C</td>
<td>0.735</td>
<td>0.674</td>
<td>0.815</td>
<td>0.870</td>
</tr>
<tr>
<td></td>
<td>0.735</td>
<td>0.676</td>
<td>0.815</td>
<td>0.871</td>
</tr>
<tr>
<td>10_C</td>
<td>0.748</td>
<td>0.681</td>
<td>0.818</td>
<td>0.876</td>
</tr>
<tr>
<td></td>
<td>0.748</td>
<td>0.691</td>
<td>0.818</td>
<td>0.871</td>
</tr>
<tr>
<td>5_C</td>
<td>0.751</td>
<td>0.695</td>
<td>0.823</td>
<td>0.876</td>
</tr>
<tr>
<td></td>
<td>0.752</td>
<td>0.677</td>
<td>0.799</td>
<td>0.864</td>
</tr>
<tr>
<td>15_SW</td>
<td>0.734</td>
<td>0.666</td>
<td>0.797</td>
<td>0.869</td>
</tr>
<tr>
<td></td>
<td>0.734</td>
<td>0.670</td>
<td>0.775</td>
<td>0.861</td>
</tr>
<tr>
<td>15_SE</td>
<td>0.739</td>
<td>0.685</td>
<td>0.828</td>
<td>0.870</td>
</tr>
<tr>
<td></td>
<td>0.738</td>
<td>0.675</td>
<td>0.829</td>
<td>0.871</td>
</tr>
<tr>
<td>15_NW</td>
<td>0.763</td>
<td>0.721</td>
<td>0.830</td>
<td>0.881</td>
</tr>
<tr>
<td></td>
<td>0.758</td>
<td>0.719</td>
<td>0.817</td>
<td>0.877</td>
</tr>
<tr>
<td>15_NE</td>
<td>0.745</td>
<td>0.721</td>
<td>0.810</td>
<td>0.868</td>
</tr>
<tr>
<td></td>
<td>0.730</td>
<td>0.703</td>
<td>0.796</td>
<td>0.857</td>
</tr>
<tr>
<td>10_SW</td>
<td>0.770</td>
<td>0.743</td>
<td>0.829</td>
<td>0.887</td>
</tr>
<tr>
<td></td>
<td>0.743</td>
<td>0.715</td>
<td>0.801</td>
<td>0.871</td>
</tr>
<tr>
<td>10_SE</td>
<td>0.742</td>
<td>0.689</td>
<td>0.814</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>0.743</td>
<td>0.634</td>
<td>0.751</td>
<td>0.858</td>
</tr>
<tr>
<td>10_NW</td>
<td>0.725</td>
<td>0.657</td>
<td>0.797</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>0.737</td>
<td>0.672</td>
<td>0.792</td>
<td>0.863</td>
</tr>
<tr>
<td>10_NE</td>
<td>0.759</td>
<td>0.702</td>
<td>0.824</td>
<td>0.881</td>
</tr>
<tr>
<td></td>
<td>0.742</td>
<td>0.684</td>
<td>0.810</td>
<td>0.871</td>
</tr>
<tr>
<td>5_SW</td>
<td>0.730</td>
<td>0.684</td>
<td>0.789</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>0.734</td>
<td>0.688</td>
<td>0.789</td>
<td>0.864</td>
</tr>
<tr>
<td>5_SE</td>
<td>0.718</td>
<td>0.690</td>
<td>0.795</td>
<td>0.857</td>
</tr>
<tr>
<td></td>
<td>0.647</td>
<td>0.603</td>
<td>0.707</td>
<td>0.829</td>
</tr>
<tr>
<td>5_NW</td>
<td>0.685</td>
<td>0.635</td>
<td>0.764</td>
<td>0.843</td>
</tr>
<tr>
<td></td>
<td>0.693</td>
<td>0.627</td>
<td>0.760</td>
<td>0.842</td>
</tr>
<tr>
<td>5_NE</td>
<td>0.747</td>
<td>0.716</td>
<td>0.817</td>
<td>0.872</td>
</tr>
<tr>
<td></td>
<td>0.729</td>
<td>0.683</td>
<td>0.802</td>
<td>0.866</td>
</tr>
</tbody>
</table>

* Mean refers to the arithmetic average value; ** Q1 refers to the 1st quartile value; *** Q3 refers to the 3rd quartile value.
Table 4. XAJ model performance in the calibration and validation periods.

<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Calibration</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NSE</td>
<td>CC</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Q1</td>
</tr>
<tr>
<td>20_C</td>
<td>0.805</td>
<td>0.796</td>
</tr>
<tr>
<td>15_C</td>
<td>0.811</td>
<td>0.806</td>
</tr>
<tr>
<td>10_C</td>
<td>0.812</td>
<td>0.811</td>
</tr>
<tr>
<td>5_C</td>
<td>0.782</td>
<td>0.772</td>
</tr>
<tr>
<td>15_SW</td>
<td>0.791</td>
<td>0.787</td>
</tr>
<tr>
<td>15_SE</td>
<td>0.827</td>
<td>0.821</td>
</tr>
<tr>
<td>15_NW</td>
<td>0.801</td>
<td>0.791</td>
</tr>
<tr>
<td>15_NE</td>
<td>0.804</td>
<td>0.800</td>
</tr>
<tr>
<td>10_SW</td>
<td>0.783</td>
<td>0.768</td>
</tr>
<tr>
<td>10_SE</td>
<td>0.804</td>
<td>0.787</td>
</tr>
<tr>
<td>10_NW</td>
<td>0.795</td>
<td>0.789</td>
</tr>
<tr>
<td>10_NE</td>
<td>0.809</td>
<td>0.806</td>
</tr>
<tr>
<td>5_SW</td>
<td>0.700</td>
<td>0.681</td>
</tr>
<tr>
<td>5_SE</td>
<td>0.818</td>
<td>0.809</td>
</tr>
<tr>
<td>5_NW</td>
<td>0.765</td>
<td>0.765</td>
</tr>
<tr>
<td>5_NE</td>
<td>0.767</td>
<td>0.752</td>
</tr>
<tr>
<td>Scenarios</td>
<td>Calibration</td>
<td>Validation</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>NSE  Q1 Q3</td>
<td>BIAS (%)</td>
</tr>
<tr>
<td>Mean</td>
<td>Mean Q1 Q3</td>
<td>Mean Q1 Q3</td>
</tr>
<tr>
<td>20_C</td>
<td>0.736 0.726 0.738 0.873 0.867 0.877</td>
<td>37.173 37.888 41.076 0.773 0.781 0.831</td>
</tr>
<tr>
<td>15_C</td>
<td>0.739 0.723 0.737 0.874 0.867 0.873</td>
<td>32.699 35.497 37.932 0.787 0.773 0.789</td>
</tr>
<tr>
<td>10_C</td>
<td>0.722 0.707 0.724 0.870 0.862 0.870</td>
<td>38.330 39.602 42.736 0.751 0.733 0.756</td>
</tr>
<tr>
<td>5_C</td>
<td>0.661 0.645 0.664 0.839 0.831 0.843</td>
<td>38.849 38.624 41.266 0.716 0.709 0.722</td>
</tr>
<tr>
<td>15_SW</td>
<td>0.702 0.678 0.709 0.860 0.854 0.863</td>
<td>35.191 35.874 38.798 0.726 0.713 0.734</td>
</tr>
<tr>
<td>15_SE</td>
<td>0.726 0.722 0.733 0.869 0.866 0.874</td>
<td>36.688 36.733 39.689 0.772 0.765 0.778</td>
</tr>
<tr>
<td>15_NW</td>
<td>0.703 0.701 0.714 0.860 0.860 0.866</td>
<td>39.389 38.772 41.964 0.767 0.767 0.775</td>
</tr>
<tr>
<td>15_NE</td>
<td>0.734 0.726 0.737 0.874 0.869 0.877</td>
<td>39.874 40.299 42.661 0.756 0.747 0.761</td>
</tr>
<tr>
<td>10_SW</td>
<td>0.727 0.713 0.731 0.865 0.860 0.869</td>
<td>34.026 34.220 37.596 0.723 0.710 0.726</td>
</tr>
<tr>
<td>10_SE</td>
<td>0.706 0.696 0.715 0.864 0.859 0.867</td>
<td>38.548 37.182 40.586 0.687 0.676 0.702</td>
</tr>
<tr>
<td>10_NW</td>
<td>0.692 0.675 0.696 0.855 0.848 0.859</td>
<td>43.935 45.957 47.672 0.763 0.757 0.770</td>
</tr>
<tr>
<td>10_NE</td>
<td>0.726 0.710 0.726 0.869 0.862 0.870</td>
<td>36.227 38.853 41.945 0.754 0.744 0.755</td>
</tr>
<tr>
<td>5_SW</td>
<td>0.647 0.650 0.655 0.814 0.817 0.819</td>
<td>31.600 31.541 33.406 0.733 0.730 0.737</td>
</tr>
<tr>
<td>5_SE</td>
<td>0.718 0.704 0.722 0.872 0.866 0.875</td>
<td>40.607 40.459 42.972 0.673 0.663 0.679</td>
</tr>
<tr>
<td>5_NW</td>
<td>0.651 0.635 0.654 0.837 0.829 0.840</td>
<td>46.650 47.065 50.304 0.720 0.714 0.729</td>
</tr>
<tr>
<td>5_NE</td>
<td>0.615 0.617 0.666 0.799 0.811 0.831</td>
<td>36.948 36.992 43.946 0.624 0.564 0.746</td>
</tr>
</tbody>
</table>
Figure 2. NSE (a–c), CC (d–f), and BIAS (g–i) distributions for HYMOD (a,d,g), XAJ (b,e,h), and WetSpa (e,f,i) for 16 scenarios illustrating performance during the entire validation period. Colors of the boxes indicate the locations of concentrated rain gauges: central (black), south-west (red), south-east (green), north-west (blue), and north-east (gray). Boxplots show the 0.10, 0.25, 0.50, 0.75, and 0.90 percentiles.
3.2. Monthly Analysis for 2012

Monthly performance is another important mechanism for comparing models’ responses due to
different rain gauge distributions. Figure 3 illustrates the monthly average runoff and precipitation
in the Jinjiang River Basin. Clearly, the wet season in 2012 ranges from March to July, resulting in
a flood season with average flow above 200 m$^3$/s. However, dry seasons are not particularly dry;
most months, except for August and October, have average precipitation over 100 mm per month.

![Figure 3. Monthly average runoff and precipitation in the Jinjiang River Basin in 2012.](image)

Monthly comparisons for HYMOD, XAJ, and WetSpa are shown in Figure 4. As shown in
Figure 4a–c, all three models tend to simulate streamflow in the flood season better, but poorly
reproduce runoff in the dry season, especially in August and October. Specifically, HYMOD
shows better adaptability than XAJ and WetSpa, achieving acceptable NSE values in more months,
especially in January and September. However, WetSpa performs worse than the other models in
March, May, July, and November. In addition, HYMOD monthly performances show no distinct trends
indicating an impact of rain gauge distributions. XAJ demonstrates higher usability in the flood season.
However, it is more likely to be affected by the distribution of rain gauges. June, for example, has the
highest NSE value (0.50) in scenario 15_SE, but the lowest NSE value (0) in 5_NW. Finally, WetSpa only
achieves acceptable results in the flood season. NSE values in the dry season are no higher than 0.1,
extcept for December, in which NSE values are about 0.3.

The overall deviations for all three models vary significantly from month to month (Figure 4d–f).
Of the three models, HYMOD generates relatively accurate overall estimates in most months. However,
it overestimates total flood volume in August and underestimates total flood volume in January,
November, and December. XAJ produces results that vary between overestimates and underestimates.
The WetSpa model achieves the best performance in March, but produces more than 30% overestimates
from July to October, which confirms the results in Figure 4c.
Figure 4. Heatmaps of monthly performance for HYMOD (a,d), XAJ (b,e), and WetSpa (c,f) based on NSE (a–c) and BIAS (d–f). For each heatmap, X axes refer to the months in 2012 and Y axes refer to the different rain gauge distribution scenarios. The colors represent the average NSE or BIAS values for all trials in the current scenario and current month.

3.3. Model Performance in June 2012

To investigate the performance of three models based on data precision for different precipitation values, daily time series of the simulated and observed hydrographs in a typical wet season month in the Jinjiang River basin (June 2012) are compared. June 2012 has two complete flood peaks, and the two peak flows are neither too high nor too low. Therefore, analyzing model performance in June 2012 is appropriate.

As presented in Figure 5, the accumulative precipitation in June 2012 ranges from 187.1 mm to 355.3 mm. The rain center is located in the central area of the basin captured by rain gauge LC. Therefore, three scenarios, 20_C as a benchmark, 10_SE absent LC, and 5_NE with LC (see Table 2), are investigated as LC presence could be a critical factor affecting runoff modeling.
Figure 5. Interpolated mean rainfall distribution in June 2012 over the Jinjiang River Basin using 20 rain gauges.

Figure 6a indicates that in scenario 20_C, HYMOD results generally agree with daily observed streamflow in June 2012. However, hydrographs in every trial range widely compared with XAJ (Figure 6b) and WetSpa (Figure 6c). XAJ and WetSpa misrepresent the first peak flow in June, while capturing the second. XAJ underestimates the lower streamflow, while WetSpa slightly overestimates them.

Absent the key rain gauge LC in scenario 10_SE (Figure 6d–f), the standard deviations in measured rain data on 25 June 2012 are clearly smaller than in 20_C. The remaining rain gauges do not correctly reflect the spatial variability in actual precipitation for the entire basin. All three models underestimate the second peak flow that they precisely captured in 20_C. For 5_NE, all models overestimate the second peak flow as the rainfall data from LC are assumed to represent the areal rainfall input from most sub-watersheds (Figure 6g–i).

Meanwhile, variations in simulated second peak flow for WetSpa increase notably and the simulated runoff time series can be divided into three groups. Most hydrographs in one group tend to appreciably overestimate the second peak flow. While those in another group tend to significantly overestimate the second peak flow, e.g., XAJ. The remainder in the third group tend to miss the second peak flow.
Figure 6. Comparison of observed and simulated runoff from scenarios 20_C (a–c), 10_SE (d–f), and 5_NE (g–i) in HYMOD (a,d,g), XAJ (b,e,h), and WetSpa (c,f,i) in June 2012. Error bars represent the standard deviation of precipitation data measured by rain gauges in the indicated scenario.
4. Discussion

The sensitivity of HYMOD, XAJ, and WetSpa models is compared to characterize the spatial distribution of rainfall in a medium scale semi-humid watershed in southeast China. The NSE, CC, and BIAS values for 16 rain gauge distribution scenarios are used to assess the performance of hydrological models with varying rain gauge deployment locations. The HYMOD and XAJ models can be regarded as representative conceptual models with different complexity, while the WetSpa model is considered a typical process-based distributed model.

4.1. Influence of Rain Gauge Distribution

For all three models, runoff varied slightly with a decreasing number of rain gauges (from 20 to 5). This result differs from the previous studies of Andreassian et al. [27] and Xu et al. [23], which showed that model efficiency declines with loss of rainfall spatial knowledge. However, it agrees with Chaplot et al. [26], who found that model performance remains stable despite decreasing gauge concentration. One explanation is that former studies used lumped hydrological models that applied average surface rainfall as input while the SWAT used in Chaplot et al. [26] is a distributed model that takes advantage of rainfall data interpolated into each sub-catchment. In this study, all three models are applied as semi-distributed models, and thus rainfall spatial variety is captured with a limited number of rain gauges. While all models are able to cope with a restricted number of rain gauges, their distribution is more likely to result in uncertainty in simulated streamflow. These results suggest that appropriate rain gauge configuration is of higher importance than the number or density of rain gauges in a certain catchment.

4.2. Comparing Different Model Structures

Because all three models are regarded as semi-distributed models that operate on sub-basins with distributed rainfall data, differences in simulated runoff between them is likely derived from model complexity and the rationality of parameterization. The general model precision and uncertainty performance in scenario 20_C is consistent with multiple studies [52,61–64].

The factor leading to larger uncertainty in model calibration for HYMOD and smaller uncertainty for WetSpa can be attributed to model structure complexity. In the model calibration process, optimization algorithms are supposed to reach optimal values. However, reaching the global optimal value requires significant computing resources and time. Therefore, most trials in the calibration period just reach local optimal values. It is possible that a higher complexity model structure could result in local optimal values closer to global optimal values according to the concept of equifinality introduced by Beven [59]. However, HYMOD has a simpler structure than XAJ and WetSpa with fewer parameters, leading to a parameter space consisting of fewer dimensions and fewer parameter combinations that can reach the ‘best’ performance, i.e., the global optimal value, during calibration. Most local optimal values are worse than the global optimal values, but the corresponding parameter sets are possibly reached during calibration. Therefore, HYMOD performance varies widely for each calibration trial.

Different performances in the water balance simulations are a more complex problem. As illustrated in Figure 2h, XAJ tends to underestimate total water volume in general. Sun et al. [65] reported similar results and concluded that underestimates of free water storage lead to underestimates of total water volume. In addition, WetSpa systematically overestimates total runoff volume. This behavior may be due to a consistent overestimate of low flows and an underestimate of high flows, where the total volume of low flows is higher than that of high flows (see Figure 4c,f,g).

4.3. Detailed Discussion on Model Performance

Finally, it is important to examine the inconsistent simulation of the two floods in June 2012. As illustrated in Figure 6, it is clear that observed streamflow is sensitive to rainfall; for example,
small flood peaks on 8 and 18 June were obtained after relatively small rainfall events, showing that basin storage (including soil and groundwater) is rather limited and thus surface flow can be generated directly. When verifying the simulated flow from XAJ and WetSpa, however, responses to small rainfall events are systematically underestimated. While in the two major floods, the simulated streamflow increases and reaches the flood peaks with a one-day delay compared with the observed data, but decreases back to normal at the same day, or even earlier. Similar behavior appears in other flood simulations, as presented in Table 6. These deviations are likely due to overestimating the water storage capacity of the basin. Hence, in the first flood period with rainfall concentrated in a single day, more precipitation turns into storage and generates streamflows with a longer delay, resulting in an average underestimation of $-32.1\%$, $-37.4\%$, and $-36.3\%$ in peak flow and $-21.1\%$, $-33.3\%$, and $-21.3\%$ in total volume for HYMOD, XAJ, and WetSpa, respectively (see Table 6). In the second flood period, with rainfall scattered over two days, precipitation in the first day fulfills the water storage and thus precipitation on the following day generates surface flow directly, resulting in a better reproduction of the second peak flow. The underestimation drops remarkably to $-12.2\%$, $-2.7\%$, and $-7.0\%$ in peak flow and $-4.1\%$, $-14.7\%$, and $-4.9\%$ in total volume (shown in Table 6). Such uncertainty derived from the model structure or model parameters needs to be addressed in further studies.

5. Conclusions

This study compared the performance and uncertainties of the HYMOD, XAJ, and WetSpa conceptual and process-based models with varying rain gauge numbers and distributions in the Jinjiang River Basin, China. Long time series, monthly, and daily performance was analyzed using several statistical indicators.

The results for all three models showed that a reduction in the number of rain gauges only resulted in worse performance when the rain gauge distribution was inhomogeneous. This observation demonstrates that appropriate rain gauge configuration is of greater importance than their deployment density in a certain catchment. Furthermore, the stable performance between different model structures indicates that rainfall spatial variability does not impact model performance due to the model mechanism.

The analyses also show that the simpler conceptual model HYMOD suffers from greater variations during multiple calibrations, while the complex process-based model WetSpa is more stable across different calibration trials. This behavior indicates that the uncertainty in the model calibration is more related to model structure than rainfall spatial variability.

Notably, this study was conducted for a middle-scale semi-humid river basin, which is suitable for rainfall runoff simulation. Future work should build on these results by extending the types of study areas encompassing a range of climates, surface areas, and environmental conditions.
Table 6. Detailed model performances for sample floods in scenario 20_C.

<table>
<thead>
<tr>
<th>Start Date</th>
<th>End Date</th>
<th>Actual Peak Flow (m³/s)</th>
<th>Total Average Rainfall (mm)</th>
<th>HYMOD</th>
<th>XAJ</th>
<th>WetSpa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Delay 1 (day)</td>
<td>BIASp 2 (%)</td>
<td>BIASv 3 (%)</td>
</tr>
<tr>
<td>29 February 2012</td>
<td>13 March 2012</td>
<td>842</td>
<td>148.3</td>
<td>0</td>
<td>-4.9</td>
<td>-9.8</td>
</tr>
<tr>
<td>9 June 2012</td>
<td>14 June 2012</td>
<td>1330</td>
<td>74.3</td>
<td>1</td>
<td>-32.1</td>
<td>-21.1</td>
</tr>
<tr>
<td>25 June 2012</td>
<td>1 July 2012</td>
<td>1080</td>
<td>83.0</td>
<td>1</td>
<td>-12.2</td>
<td>-4.1</td>
</tr>
<tr>
<td>4 April 2013</td>
<td>9 April 2013</td>
<td>726</td>
<td>62.9</td>
<td>0</td>
<td>-9.8</td>
<td>-10.6</td>
</tr>
<tr>
<td>14 May 2013</td>
<td>21 May 2013</td>
<td>772</td>
<td>96.3</td>
<td>0</td>
<td>-11.2</td>
<td>1.8</td>
</tr>
</tbody>
</table>

1 Denotes the delay days that the simulated flood peak occurs in 30 trials minus the date of the observed flood peak; 2 refers to the mean BIAS calculated using the simulated flood peak volume versus observed flood peak volume; 3 denotes the mean BIAS calculated using the simulated total flood volume versus observed total flood volume.
Author Contributions: Data curation, R.W.; Investigation, Z.Y.; Project administration, X.L. and H.W.; Resources, W.L.; Supervision, X.L. and H.W.; Visualization, Z.Y.; Writing—original draft, Z.Y.; Writing—review and editing, Z.Y., W.L. and R.W.

Funding: This research was funded by the [National Key R&D Program of China] grant number [2017YFB0203104] and [National Natural Science Fund] grant number [51709273].

Acknowledgments: This paper was jointly supported by the National Key R&D Program of China (2017YFB0203104) and National Natural Science Fund (51709273). We are grateful to the Jiangxi Hydrological Bureau, Data Center for Resources and Environment Sciences, Chinese Academy of Sciences, The Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences and Institute of Soil Science, Chinese Academy of Sciences for providing data.

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

References
13. Li, D.; Christakos, G.; Ding, X.; Wu, J. Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of Xiaoxi catchment (Taihu lake basin, China). *J. Hydrol.* 2018, 556, 1139–1152. [CrossRef]


31. Chang, C.L.; Chao, Y.C. Effects of spatial data resolution on runoff predictions by the BASINS model. *Int. J. Environ. Sci. Technol.* 2013, 11, 1563–1570. [CrossRef]


35. Meng, X. *China Meteorological Assimilation Driving Datasets for the SWAT Model*, 1st ed.; version 1; Cold and Arid Regions Science Data Center at Lanzhou: Lanzhou, China, 2016. [CrossRef]


50. Li, H.; Zhang, Y.; Chiew, F.H.S.; Xu, S. Predicting runoff in ungaged catchments by using Xinanjiang model with MODIS leaf area index. *J. Hydrol.* **2009**, *370*, 155–162. [CrossRef]


52. Safari, A.; De Smedt, F.; Moreira, F. WetSpa model application in the Distributed Model Intercomparison Project (DMIP2). *J. Hydrol.* **2012**, *418–419*, 78–89. [CrossRef]


54. Safari, A.; De Smedt, F.; Moreda, F. WetSpa model application in the Distributed Model Intercomparison Project (DMIP2). *J. Hydrol.* **2012**, *418–419*, 78–89. [CrossRef]

55. Safari, A.; De Smedt, F. Improving WetSpa model to predict streamflows for gaged and ungaged catchments. *J. Hydroinform.* **2014**, *16*, 758–771. [CrossRef]


