



Article

Applicability Assessment and Uncertainty Analysis of Multi-Precipitation Datasets for the Simulation of Hydrologic Models

Binbin Guo ^{1,2} , Jing Zhang ^{1,*}, Tingbao Xu ³, Barry Croke ^{3,4}, Anthony Jakeman ³ , Yongyu Song ¹, Qin Yang ^{1,2}, Xiaohui Lei ⁵ and Weihong Liao ⁵

¹ Beijing Key Laboratory of Resource Environment and Geographic Information System, Capital Normal University, Beijing 100048, China; guobinbin@126.com (B.G.); songdy1006@gmail.com (Y.S.); yqinss@163.com (Q.Y.)

² College of City and Tourism, Hengyang Normal University, Hengyang 421008, China

³ Fenner School of Environment and Society, The Australian National University, Canberra, ACT 2601, Australia; tingbao.xu@anu.edu.au (T.X.); barry.croke@anu.edu.au (B.C.); Tony.Jakeman@anu.edu.au (A.J.)

⁴ Mathematical Sciences Institute, The Australian National University, Canberra, ACT 2601, Australia

⁵ 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.); liaowh@iwhr.com (W.L.)

* Correspondence: 5607@cnu.edu.cn; Tel.: +86-10-6890-3139

Received: 10 October 2018; Accepted: 5 November 2018; Published: 9 November 2018



Abstract: Hydrologic models are essential tools for understanding hydrologic processes, such as precipitation, which is a fundamental component of the water cycle. For an improved understanding and the evaluation of different precipitation datasets, especially their applicability for hydrologic modelling, three kinds of precipitation products, CMADS, TMPA-3B42V7 and gauge-interpolated datasets, are compared. Two hydrologic models (IHACRES and Sacramento) are applied to study the accuracy of the three types of precipitation products on the daily streamflow of the Lijiang River, which is located in southern China. The models are calibrated separately with different precipitation products, with the results showing that the CMADS product performs best based on the Nash–Sutcliffe efficiency, including a much better accuracy and better skill in capturing the streamflow peaks than the other precipitation products. The TMPA-3B42V7 product shows a small improvement on the gauge-interpolated product. Compared to TMPA-3B42V7, CMADS shows better agreement with the ground-observation data through a pixel-to-point comparison. The comparison of the two hydrologic models shows that both the IHACRES and Sacramento models perform well. The IHACRES model however displays less uncertainty and a higher applicability than the Sacramento model in the Lijiang River basin.

Keywords: precipitation; TMPA-3B42V7; CMADS; hydrologic model; uncertainty

1. Introduction

Hydrologic models are essential tools for understanding processes of the hydrologic cycle and provide useful information for sustainable water-resource management [1]. Precipitation is the main driving factor of hydrologic processes. Accurate estimation of precipitation is crucial for reliable hydrologic predictions [2]. Traditionally, precipitation data from a ground observational network have been used as the source of areal precipitation estimates used in watershed modelling. However, ground-based precipitation observation networks are sparsely distributed and may be unable to represent the spatial variability of the precipitation completely. Moreover, precipitation measurements

are frequently missing because of malfunctioning of devices [3]. Remote sensing [4] and modelling [5] of precipitation have become viable approaches to address these problems effectively and are often used as input data to hydrologic models.

With regard to the development of remote-sensing technology, satellite-derived precipitation data are an attractive alternative in data-sparse regions because of the relatively high resolution and complete spatial coverage. A number of such remotely sensed precipitation products are currently available. These include, for example, the Climate Prediction Center morphing method (CMORPH, [6]), the Global Satellite Mapping of Precipitation (GSMaP) project [7], the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) [8] and the Global Precipitation Measurement (GPM) products [9]. Among them, the TMPA products developed by the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC), with a spatial resolution of $0.25^\circ \times 0.25^\circ$ for multiple timescales (3 hourly, daily and monthly), has received much more attention [10]; the latest research product of TMPA for post-real-time research (3B42) is version 7. Most of the applications using the TMPA-3B42V7 product indicate an excellent potential to supply reasonably high spatial and temporal resolution data for hydrometeorological applications [10–13]. However, remotely sensed precipitation data suffer from uncertainty in their retrieval algorithms and observation errors [14] due to the inference of rainfall based on observations of the conditions at the top of clouds.

While precipitation modelling is fairly accurate for coarse-scale (global-scale), organized, synoptic systems, the modelling accuracy decreases rapidly for more localized events as spatial and temporal features cannot be explicitly resolved by global models [2]. Reanalysis datasets, which are produced by assimilating multi-source data into a climate model, are a viable option of deriving reliable precipitation estimates [5]. Commonly used reanalysis datasets include the Climate Forecast System Reanalysis (CFSR) [15] from National Center for Environmental Prediction (NCEP), the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis from September 1957 to August 2002 (ERA-40 [16]) and the ECMWF Reanalysis-Interim (ERA-Interim [17]) products. While these reanalysis datasets provide important basic data for global researchers for the analysis of climate–water cycles, the spatial resolution of global reanalysis datasets is often too coarse to be used reliably in local-scale studies. Hydrologic modelling forced by reanalysis datasets has been conducted by, for example, Andreadis et al. (2017) [5], who reproduced flooding over large scales by using the Twentieth Century Reanalysis (20CRv2, [18]) dataset and downscaling techniques. Fuka et al. (2014) found the CFSR precipitation product provides a relatively reliable precipitation input for the hydrologic modelling of large-area basins.

Given the strongly underconstrained nature of precipitation inversion, data assimilation based on the large number of stations on regional scales has the potential to resolve fine-scale structures and microphysical processes with more details. The China Meteorological Assimilation Driving Datasets for the Soil and Water Assessment Tool (SWAT) model (CMADS,) developed by Dr. Xianyong Meng from the China Agricultural University (CAU), has received worldwide attention [19,20]. CMADS incorporate Space and Time Mesoscale Analysis System (STMAS) assimilation techniques [21,22] and multiple other techniques, such as loop nesting of data, projection of resampling models and bilinear interpolation. The precipitation data of the CMADS product is generated by the assimilation of multi-satellite data and precipitation from ground stations. Using CMORPH satellite products as the background field, the CMADS product assimilates hourly precipitation products of nearly 40,000 regional automatic stations and 2421 national automatic stations in China. Relative studies found the CMADS product significantly reduces the uncertainties of precipitation input for the hydrologic modelling [19]. CMADS has been verified in several basins in China and Korea [20,23–29]. However, reanalysis datasets are limited by the quality of precipitation observations and the uncertainty from the assimilation model.

A number of spatial-interpolation methods [30] are commonly used for estimating precipitation based on ground-observation data, even in data-sparse regions [31,32]. Conventional interpolation methods, such as the Thiessen polygon [33,34] and inverse-distance weighting [35], are widely used

for precipitation interpolation [32]. Ordinary kriging [36] is a geostatistical technique requiring prior calibration of a semivariogram for its parameters (range, nugget and sill). These methods are suitable for application over relatively flat areas. These methods assume that other potential drivers (particularly topography) of the spatial variation in rainfall is captured by the gauge data and information on other drivers is not needed. In data-sparse regions this is not correct and methods that explicitly include the topography are preferred. Hutchinson found that the interpolating accuracy of a precipitation surface would be enhanced significantly with an appropriate digital elevation model (DEM) [37–39]. The advantages of the ANUSPLIN package (Hutchinson and Xu, 2013) over kriging are its simplicity and there is no requirement of a separate calibration of the spatial-covariance structure. The ANUSPLIN interpolation technique has been applied in a number of studies, proving to be one of the best techniques for interpolating point precipitation data [40–42]. However, if the meteorological stations are very sparse, obtaining an accurate distribution of precipitation values through interpolation is impossible. The low density of precipitation stations is a major uncertainty source, which potentially impacts the result. Moreover, interpolation of precipitation data is unable to capture some extreme weather conditions. All interpolation techniques have difficulty in simulating sharply varying climate transitions.

Spatially distributed precipitation datasets incorporate uncertainties or errors resulting from the interpolation and retrieval algorithms, the quality of precipitation observations and the uncertainty from the assimilation model. As different precipitation datasets are limited by quantitative inaccuracies, they exhibit significant bias [43]. Smith and Kummerow [44] analyzed the water budgets of precipitation datasets from in situ, reanalysis and satellite data over the upper Colorado River basin and found the reanalysis datasets tend to overestimate in situ data, while satellite-derived precipitation data underestimate in situ data. Pfeifroth et al. [45] evaluated satellite-based and reanalysis precipitation data in the tropical Pacific and found reanalysis products overestimate small and medium precipitation amounts but underestimate high amounts. Some studies have found that runoff-generation is highly sensitive to the spatial and temporal variability of precipitation data, as a result this is found to be the main source of uncertainty in rainfall–runoff modelling [46]. Therefore, assessing the accuracy of different precipitation products and their applicability and uncertainty for hydrologic models is of great importance; the uncertainties associated with hydrologic models also play a role in the performance of hydrologic simulations [47].

Our main objective here is to assess and evaluate three general precipitation datasets in terms of their accuracy and efficacy, including the CMADS, TMPA-3B42V7 and gauge-interpolated product. The assessment is based on the simulation results from two well-known hydrologic models (IHACRES and Sacramento models). In addition, the precipitation detection capability of TRMA-3B42V7 and CMADS datasets is also evaluated through their pixel-to-point comparison to the ground-based data. The applicability of these two models is assessed. Moreover, the parameter uncertainty of each hydrologic model is also explored as this is another source of uncertainty in modelling streamflow. This research will provide more insight into precipitation analyses and hydrologic modelling.

2. Material and Methods

2.1. Study Area

The Lijiang River basin (25°12′–25°55′ N, 110°5′–110°40′ E) is located in the northeastern Guangxi Zhuang Autonomous Region of China and belongs to the upper reaches of the Guijiang River in the Pearl River system, with an area of 2591 km². The basin is an important headwater for the downstream Guilin City and has a sparse and unevenly distributed meteorological observation network.

The climate in this area is mainly sub-tropical monsoonal, with the wet season from March to August and the dry season from September to February. From 1961 to 2016, the average daily temperature was 19.10 ± 0.06 °C and the yearly average precipitation was 1900 ± 50 mm, varying between 1253.6 and 3011 mm. With elevation ranging from 32 m to 2037 m (Figure 1),

the terrain is high in the north and low in the south, with highly complex topography consisting of steep mountains and floodplains.

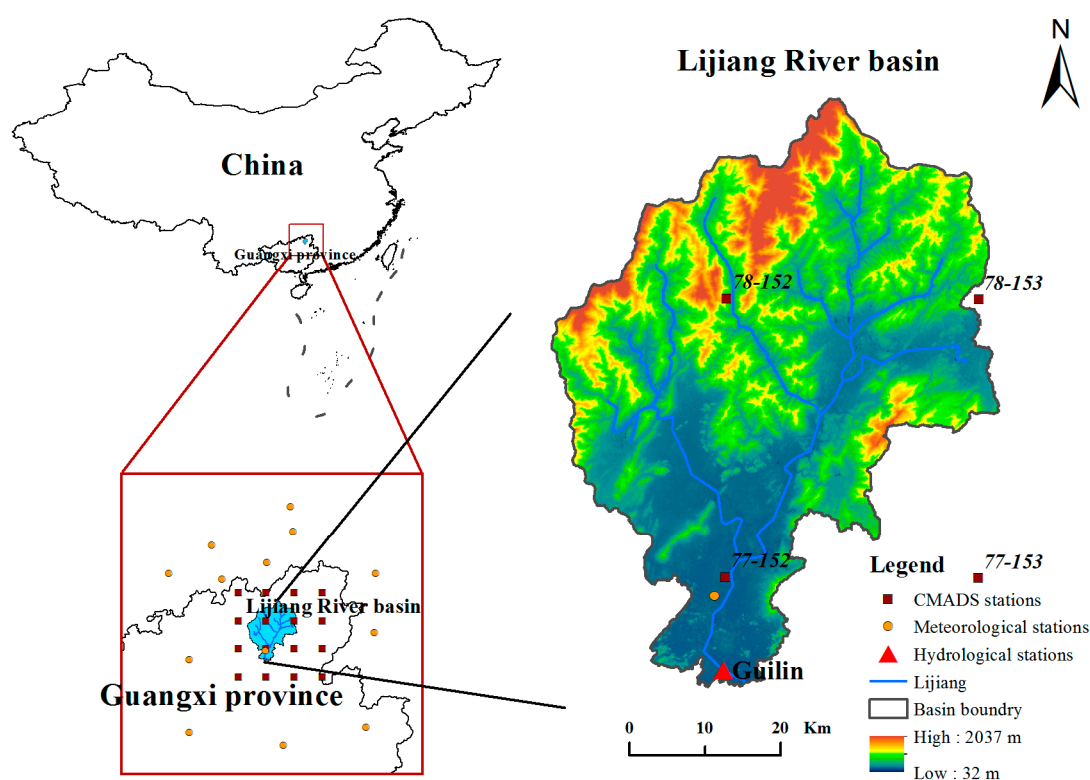


Figure 1. Location of Lijiang River basin, China and meteorological stations for the ANUSPLIN interpolation technique.

The basin is one of the most famous karst areas in the world but suffers from the fragile ecology of the karst geomorphology. The rapid development of tourism and urbanization has promoted the economy, while also causing serious environmental issues. Therefore, the development of hydrologic models in the basin is important for aiding understanding of the hydrologic processes and the formulation of scientific strategies for the management of its water resources.

2.2. Dataset Acquisition

Assessment and analysis have been conducted using three kinds of precipitation data, including in situ measurements, remote-sensing products and reanalysis data. The in situ measurements of daily precipitation from 13 meteorological stations of the national weather station network are interpolated with the ANUSPLIN technique. The mean density of meteorological stations (number of gauges per 10^4 km^2) is 0.91 (calculated using the kernel-density estimation for a search radius of 100 km; see Supplementary Materials for more details). The remote-sensing products are using the TRMM-TMPA product (also denoted TMPA-3B42V7) available from the National Aeronautics and Space Administration (NASA) official website (<https://pmm.nasa.gov/trmm>). The reanalysis data originate from the CMADS V1.0 product available from World Data System for Cold and Arid Regions (CARD) official website (<http://westdc.westgis.ac.cn>). The precipitation and maximum temperature values from 13 national meteorological stations are available from the Meteorological Data network (<http://data.cma.cn/>).

The ANUSPLIN method interpolates meteorological data from the station to the grid scale at the surface. Precipitation and maximum temperature data from meteorological stations, as well as precipitation from CMADS grid points, are interpolated to the surface grid by use of the ANUSPLIN package version 4.4, which interpolates precipitation and temperature as a function of latitude,

longitude and elevation, while accounting for the effect of topography [37–39,48]. This method interpolates precipitation and the maximum temperature to grids of size of $0.01^\circ \times 0.01^\circ$ ($\approx 1 \text{ km} \times 1 \text{ km}$), before combination with a digital elevation model (DEM) of the same resolution. The DEM for integration of the final climate grid products is derived from Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) 7.5-arc-second dataset using the Australian National University Digital Elevation Model version 5.3 (ANUDEM5.3) [49–51]. This the latest and possibly best global-terrain product to date, since it uses ground surface elevation rather than a canopy top surface as found in some satellite terrain products. Due to the smoother nature of climate surfaces relative to the underlying topography, a rebuilt 1-km resolution DEM has been used to generate the resulting climate surfaces, yielding climate grids with more realistic spatial-distribution patterns. Cross-validation statistics were calculated to evaluate the overall predictive error of the ANUSPLIN precipitation data from meteorological stations, with results demonstrating that the distributed precipitation interpolated to the surface grid by the ANUSPLIN method provides a reliable precipitation distribution for input into hydrologic models (see Supplementary Materials for more details).

The latest research product of TMPA for post-real-time research version 7 (TMPA-3B42V7), which has a spatial resolution of 0.25° , is used here. With a spatial resolution of the CMADS1.0 product of $1/3^\circ$, this study includes 16 CMADS grid points within and around the basin (Figure 1).

The daily discharge data of the Guilin hydrologic station from 2008 to 2016 (9 years) is provided by the Hydrological Bureau of the Guangxi Zhuang Autonomous Region. Time series of streamflow, maximum temperature and rainfall were used as input for hydrologic models. The spatial distributions of annual precipitation derived from CMADS, TRMM and ground-observation and maximum temperature are shown in Figure 2 (the annual average precipitation and average daily maximum temperature were calculated based on grid data from 2008 to 2016). The spatial distribution of observed precipitation and maximum temperature is interpolated from meteorological stations shown in Figure 1 using the ANUSPLIN interpolation technique. The annual precipitation estimation for CMADS, TRMM and ground-observation have similar spatial distribution patterns, with a decreasing pattern from south to north and from west to east. The average daily maximum temperature is higher in the south than in the north.

2.3. Rainfall–Runoff Models

We used two well-known conceptual rainfall–runoff models with different complexities ranging from 6–13 parameters (Table 1) to assess if the performance of the model improved with greater complexity, as well as their applicability in streamflow prediction in the Lijiang River basin.

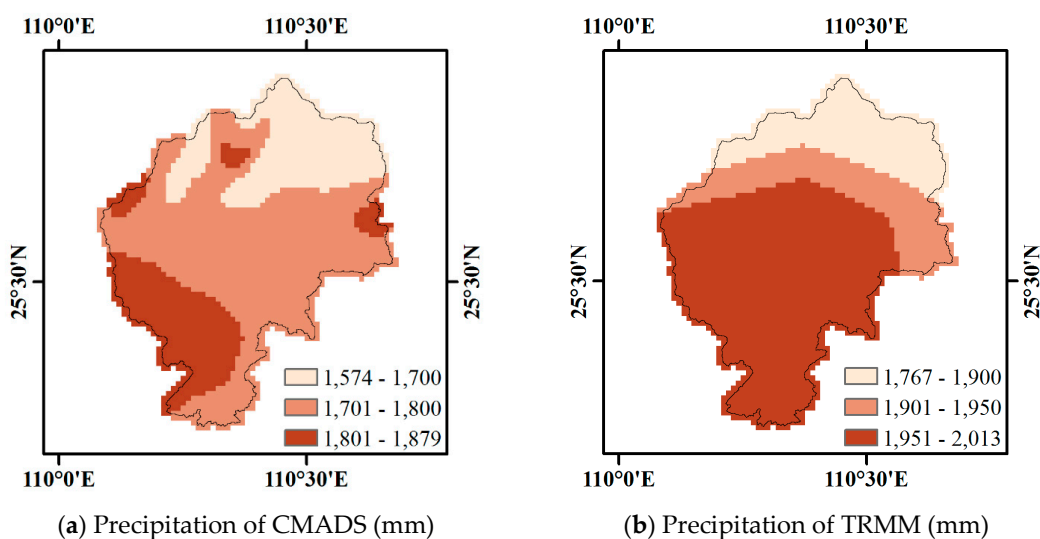


Figure 2. Cont.

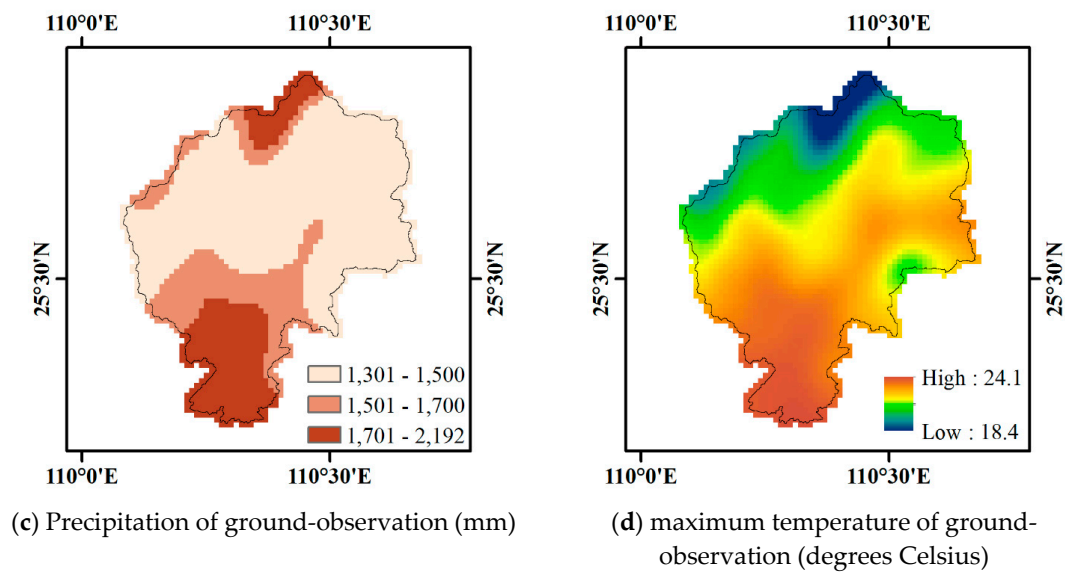


Figure 2. Spatial distribution of annual precipitation and maximum temperature estimation.

The IHACRES model (which has been used in various studies [52–54]) catchment moisture deficit (CMD) version [55] is used here as it has more consistent physical meaning for the parameters. There are two stores in the IHACRES model: the nonlinear store for the generation of effective rainfall, which uses a nonlinear function to deal with the raw rainfall, as well as using accounting equations to calculate the CMD output, and the linear store which converts the effective rainfall into quick and slow flow using unit hydrographs.

The Sacramento model, which has been used in many studies [56,57], has five runoff components: a direct runoff from an impervious area, surface runoff, interflow, supplementary base flow and primary base flow, with the 13-parameter version of the Sacramento model used here [58]. Briefly, excess rainfall becomes runoff through a unit hydrograph, with the rest of the rainfall filling various depths of interconnected soil-moisture stores. Loss through evapotranspiration occurs at the soil stores, with the remaining water becoming interflow and groundwater; the summation of the surface and lateral flow forms the streamflow.

Table 1 gives a description of the parameters of the two rainfall–runoff models. We used the Hydrological Model Assessment and Development (Hydromad) [59] modelling package to help us construct the hydrologic models. Hydromad is an open-source software package in R and is available at <http://hydromad.catchment.org>. It provides a modelling framework for environmental hydrology and supports simulation, estimation, assessment and visualization of flow response to time series of rainfall and other drivers.

2.4. Model Performance Evaluation Criteria

The Nash–Sutcliffe efficiency (*NSE*) performance measure [60] is a form of the mean squared error widely used in hydrology as a criterion for assessing hydrologic-model performance. The *NSE* objective function focuses on fitting high flowrates [61]. Here, it is also computed in terms of square-root-transformed and logarithmic-transformed flows (denoted NSE_{sq} and NSE_{log} hereafter), which makes it possible to assess the model efficiency for low flowrates [62]. The use of these three criteria (NSE , NSE_{sq} and NSE_{log}) gives a more general overview of the model efficiency and are defined as

$$NSE = 1 - \frac{\sum_{i=1}^n (obs_i - sim_i)^2}{\sum_{i=1}^n (obs_i - \overline{obs})^2}, \quad (1)$$

$$NSE_{sq} = 1 - \frac{\sum_{i=1}^n (\sqrt{obs_i} - \sqrt{sim_i})^2}{\sum_{i=1}^n (\sqrt{obs_i} - \sqrt{\overline{obs}})^2}, \tag{2}$$

$$NSE_{log} = 1 - \frac{\sum_{i=1}^n (\log(obs_i + \epsilon) - \log(sim_i + \epsilon))^2}{\sum_{i=1}^n (\log(obs_i + \epsilon) - \log(\overline{obs} + \epsilon))^2}, \tag{3}$$

respectively. The relative bias (*rel.bias* [mm]) is also used as a model-performance criterion, with the optimal value of zero and is defined as

$$rel.bias = \frac{\sum_{i=1}^n (sim_i - obs_i)}{\sum_{i=1}^n obs_i}, \tag{4}$$

where *i* is the index for individual days in the period, *n* the total number of days, *sim* denotes the simulated runoff, *obs* the observed runoff and \overline{obs} the mean observed runoff averaged over the period.

Table 1. Parameters for each model.

Parameter Name	Unit	Range	Description
IHACRES-CMD			
<i>f</i>	-	0.01–3	CMD stress threshold as a proportion of <i>d</i>
<i>e</i>	-	0.01–1.5	Temperature to potential evapotranspiration (PET) conversion factor
<i>d</i>	mm	50–550	CMD threshold for producing flow
<i>tau_s</i>	day	30–600	Time constant for slow flow store
<i>tau_q</i>	day	1–10	Time constant for quick flow store
<i>v_s</i>	-	0.1–1	Fractional volume for slow flow
Sacramento			
<i>UZTWM</i>	mm	1–150	Upper zone tension water maximum capacity
<i>UZFWM</i>	mm	1–150	Upper zone free water maximum capacity
<i>UZK</i>	1/day	0.1–0.5	Upper zone free water lateral depletion rate
<i>PCTIM</i>	-	0.000001–0.1	Fraction of the impervious area
<i>ADIMP</i>	-	0–0.4	Fraction of the additional impervious area
<i>ZPERC</i>	-	1–250	Maximum percolation rate coefficient
<i>REXP</i>	-	0–5	Exponent of the percolation equation
<i>LZTWM</i>	mm	1–500	Lower zone tension water maximum capacity
<i>LZFSM</i>	mm	1–1000	Lower zone supplementary free water maximum capacity
<i>LZFPM</i>	mm	1–1000	Lower zone primary free water maximum capacity
<i>LZSK</i>	1/day	0.01–0.25	Lower zone supplementary free water depletion rate
<i>LZPK</i>	1/day	0.0001–0.25	Lower zone primary free water depletion rate
<i>PFREE</i>	-	0–0.6	Fraction percolating from upper to lower zone free water storage

The optimization evolutionary technique used to calibrate parameter values is the Shuffled Complex Evolution (SCE) [63] algorithm, which is a popular method for parameter calibration and has proven to be both effective and relatively efficient, providing a similar performance to other evolutionary optimization algorithms [64]. The value of the objective functions for the calibration of parameters can be used as model-performance statistics.

2.5. Performance of Precipitation Detection

The expression for these statistical measures are based on a contingency table (Table 2).

Table 2. Contingency table for the ground observations and the Satellite/reanalysis estimate with a threshold of 1.0 mm.

Satellite/Reanalysis Estimate	Ground Observation	
	Observation \geq 1.0 mm	Observation $<$ 1.0 mm
Estimate \geq 1.0 mm	<i>H</i>	<i>F</i>
Estimate $<$ 1.0 mm	<i>M</i>	<i>Z</i>

Where H , F , M and Z represent the number of hits (true positives), the number of misses (false positives), the number of false alarms (false negatives) and true negatives respectively, based on a threshold of 1.0 mm. The precipitation detection capability of TRMA-3B42V7 and CMADS datasets is evaluated through their pixel-to-point comparison with the ground-based data. We use six statistical measures, including the Proportion Correct (PC), Probability of Detection (POD), Frequency Bias Index (FBI), False Alarm Ratio (FAR), Critical Success Index (CSI) and Heidke skill score (HSS) to estimate their precipitation detection capability. Those statistical measures are defined in Table 3 [14,43].

Table 3. Statistical measures for validation of precipitation detection capability.

Metric	Formula	Range	Optimal Value
Proportion Correct	$PC = \frac{H+Z}{N}$	0–1	1
Probability of Detection	$POD = \frac{H}{H+M}$	0–1	1
Frequency Bias Index	$FBI = \frac{H+F}{H+M}$	0–+∞	1
False Alarm Ratio	$FAR = \frac{F}{F+H}$	0–1	0
Critical Success Index	$CSI = \frac{H}{M+H+F}$	0–1	1
Heidke Skill Score	$HSS = \frac{2*(Z*H-F*M)}{(Z+F)*(F+H)+(M+H)*(Z+M)}$	−∞–1	1

Note: $N = H + F + M + Z$.

The PC, POD, FBI and FAR were used to measure the misdetection and false alarms from satellite/reanalysis data. CSI and HSS are more comprehensive contingency metric were used to evaluate the strength of the correlation between the ground observations and the satellite/reanalysis estimate. CSI combines the advantages of both POD and FAR. HSS can safely be compared on different datasets and also measures the overall detection skill accounting for matches due to random chance.

2.6. Generalized Likelihood Uncertainty Estimation Method

Precipitation input and the uncertainties in the parameter values of a hydrologic model are the two major factors affecting the performance of hydrologic and water-resource modelling in a basin. These are assessed here using the generalized likelihood uncertainty estimation (GLUE) method [65,66], which is a stochastic method for quantifying the uncertainty of model predictions. It can be summarized in the following steps:

- (1) A large number of models are run with randomly chosen parameter sets selected from a probability distribution; here, 100,000 group parameters are chosen obeying a uniform distribution.
- (2) Definition of the “likelihood” function (here, the performance measures NSE and NSElog) and calculation of likelihood values corresponding to each parameter set.
- (3) Definition of a cut-off threshold value for the likelihood function to distinguish between the “behavioral” parameter sets and the “non-behavioral” parameter sets.
- (4) Rescaling of the cumulative likelihood values of all behavioral models to unity.
- (5) Calculation of the percentiles of the cumulative distribution of the likelihood measure. The GLUE method integrates the outputs of all behavioral models in an ensemble prediction. For each timestep of the simulation, the output prediction is obtained as the median of the distribution of all ensemble members, with its uncertainty bounds estimated as the 5% and 95% percentiles of the distribution.

3. Results

3.1. Evaluation of Model Performance

By setting one year as the warm-up period, with 2008–2012 as the calibration period and *NSE* as the objective function, the shuffled complex evolution algorithm is used to calibrate the parameter values of the two hydrologic models. Using the resulting calibrated parameters, the overall performance (*NSE*) of observed and simulated values for the precipitation datasets and models for the period 2008–2016 is shown in Table 4. Tables 5 and 6 depict the calibrated optimal parameters sets and daily *NSE* for each precipitation dataset applied respectively to the IHACRES model and Sacramento model (2008–2012).

Table 4. Overall performance (daily *NSE* (monthly *NSE*)) of precipitation datasets for models using *NSE* as the objective function for calibration.

	IHACRES	Sacramento
Gauged	0.57 (0.83)	0.52 (0.80)
TRMM	0.56 (0.89)	0.56 (0.87)
CMADS	0.69 (0.93)	0.70 (0.92)

Table 5. Calibrated optimal parameters sets and daily *NSE* for each precipitation dataset applied to the IHACRES model (2008 to 2012).

Datasets	<i>f</i>	<i>e</i>	<i>d</i>	<i>tau_q</i>	<i>tau_s</i>	<i>v_s</i>	<i>NSE</i>
Gauged	1.132	0.05149	80.55	2.420	30.00	0.10	0.61
TRMM	1.060	0.06742	147.45	5.061	30.00	0.10	0.52
CMDAS	3.000	0.08322	50.00	3.055	30.00	0.10	0.69

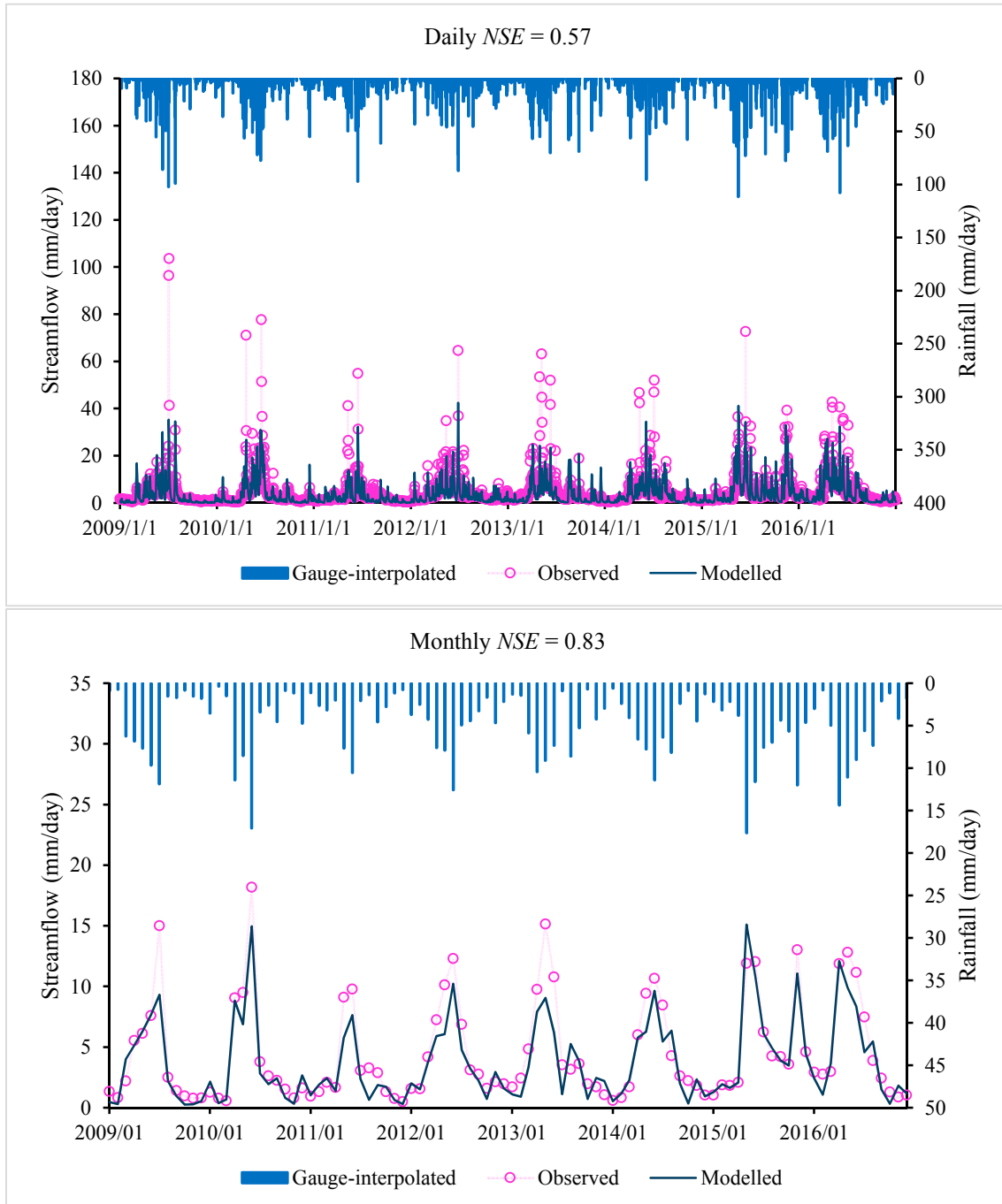
Table 6. Calibrated optimal parameter sets and daily *NSE* values for each precipitation dataset applied to the Sacramento model (2008–2012).

Datasets	<i>uztwm</i>	<i>uzfwm</i>	<i>uzk</i>	<i>pctim</i>	<i>adimp</i>	<i>zperc</i>	<i>rexp</i>
Gauged	1.000	93.5	0.322	0.0499	0.0656	149.7	3.420
TRMM	1.000	140.1	0.102	1.01×10^{-6}	1.76×10^{-8}	140.8	1.205
CMDAS	1.002	150.0	0.158	0.0509	9.48×10^{-8}	159.4	4.844
Datasets	<i>lztwm</i>	<i>lzfsm</i>	<i>lzfpn</i>	<i>lzsk</i>	<i>lzpk</i>	<i>pfree</i>	<i>NSE</i>
Gauged	1.000	998.9	944.7	0.250	0.250	0.0100	0.57
TRMM	1.320	1000.0	119.1	0.152	0.212	0.2156	0.51
CMDAS	1.963	1000.0	1.00	0.227	0.228	0.0842	0.68

Selecting the years 2008–2012 as the calibration period and the years 2012–2016 as the validation period, the daily and monthly performance of the IHACRES model for different precipitation products are shown in Table 7, with daily and monthly observed and modelled flow and rainfall shown in Figure 3. The additive merit of NSE_{sq} and NSE_{log} is also calculated. Since NSE_{sq} and NSE_{log} shift the focus from high flows to progressively lower flows, using NSE_{sq} and NSE_{log} help us judge the performance of the models in simulating over a broader range of flows. The performance of the IHACRES model shows that the CMADS dataset has the best performance among all three precipitation products during the calibration period. The TMPA-3B42V7 and gauge-interpolated product have a similar performance but perform slightly worse than the CMADS dataset. During the validation period, the *NSE*, NSE_{sq} and NSE_{log} values when using the CMADS dataset show a better performance than other precipitation datasets, which indicates the CMADS dataset performs better than other precipitation datasets in simulating both high flow and low flows. Overall, all the three precipitation datasets perform well, with the CMADS dataset performing slightly better.

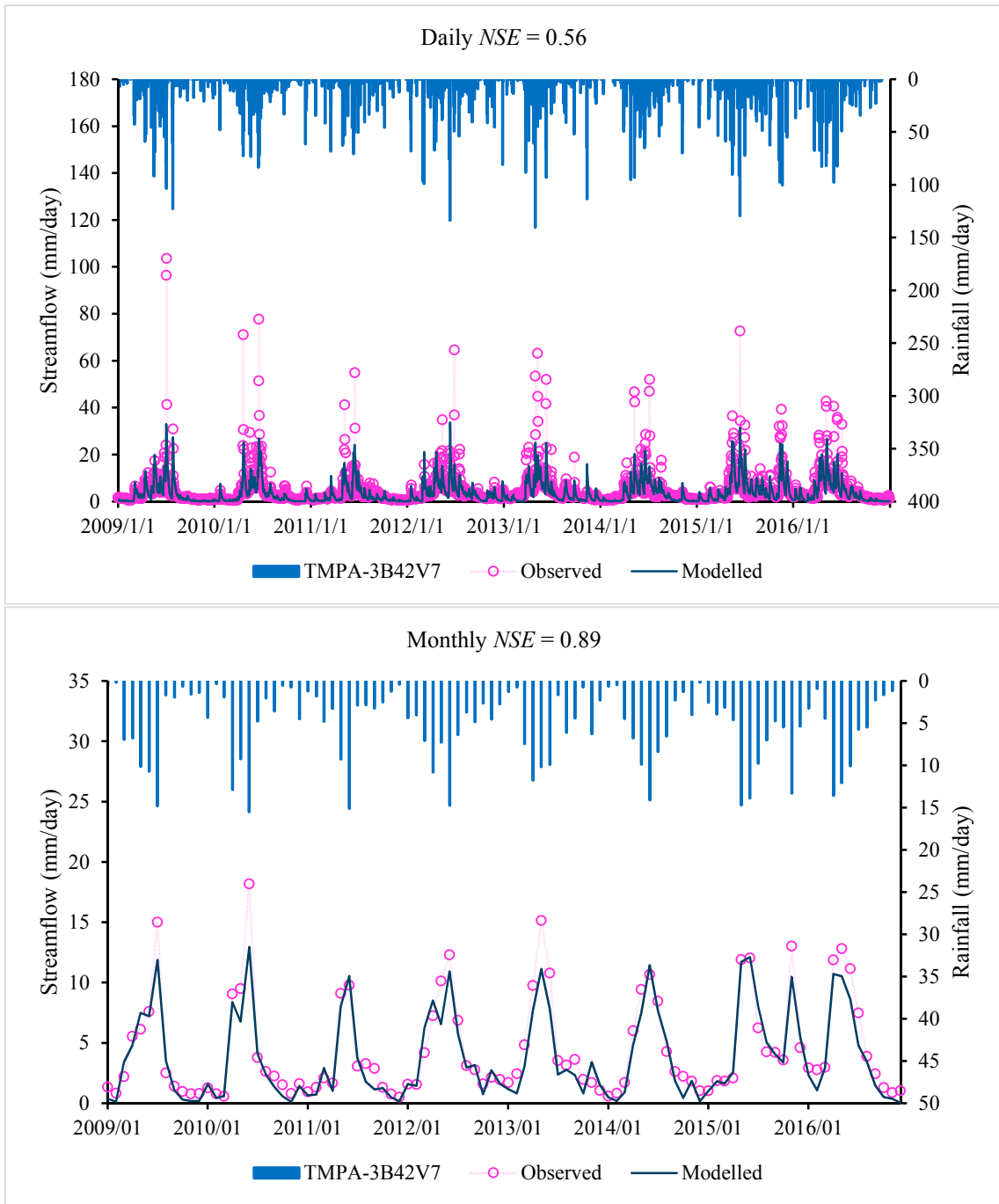
Table 7. Model performance of the IHACRES model (calibrated using *NSE*) for the calibration period and validation periods.

	Datasets	Daily rel.bias	DailyNSE	DailyNSE _{sq}	Daily NSE _{log}	Monthly NSE
Calibration period	Gauged	−0.18	0.61	0.63	0.51	0.86
Validation period	Gauged	−0.13	0.52	0.54	0.45	0.81
Calibration period	TMPA-3B42V7	−0.12	0.52	0.62	0.56	0.89
Validation period	TMPA-3B42V7	−0.15	0.61	0.59	0.49	0.89
Calibration period	CMADS	−0.21	0.69	0.57	0.32	0.93
Validation period	CMADS	−0.07	0.70	0.63	0.49	0.93



(a) Gauge-interpolated

Figure 3. Cont.



(b) Tmpa-3B42V7

Figure 3. Cont.

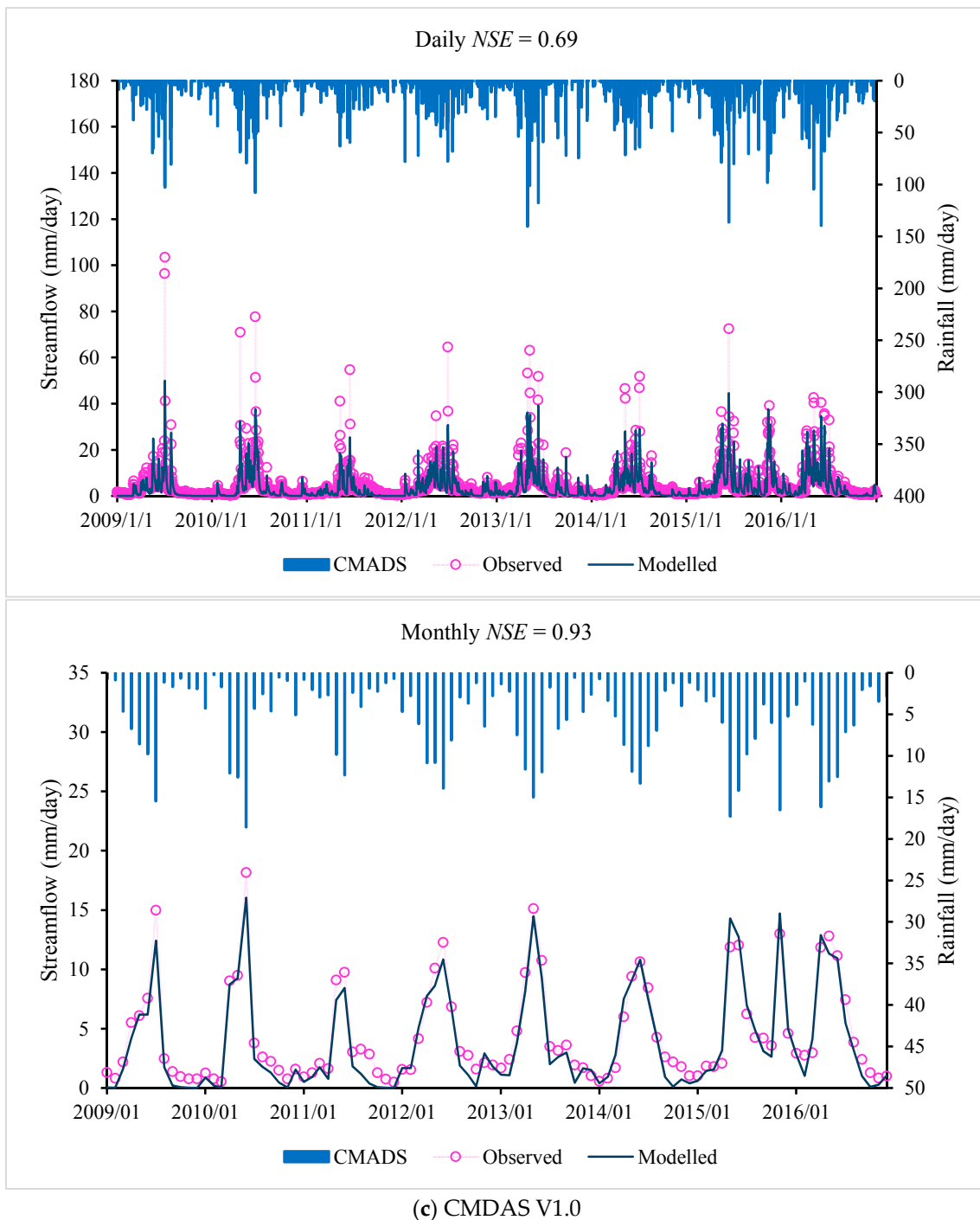
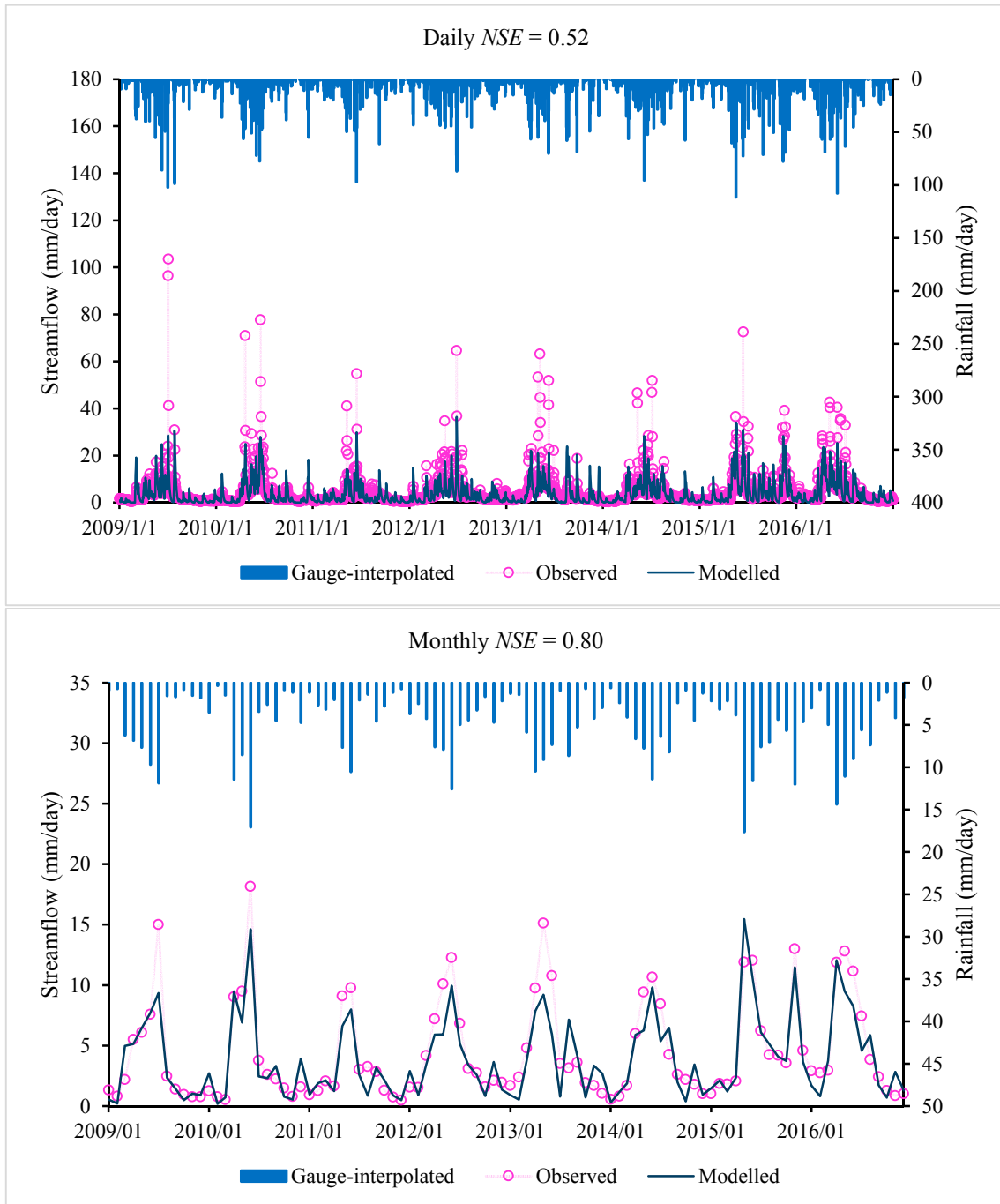


Figure 3. Observed and IHACRES-model-simulated daily and monthly runoffs for (a) Gauge-interpolated, (b) TMPA-3B42V7 and (c) CMDAS rainfall datasets (for the IHACRES model calibrated using NSE as the objective function).

The daily and monthly model performance of the Sacramento model for different precipitation products are shown in Table 8, with time series plotted in Figure 4. Again, the performance of the Sacramento model show that the precipitation product of the CMADS dataset performs best for both the calibration and validation periods, followed by the TMPA-3B42V7 datasets, with the performance of the gauge-interpolated product slightly worse.

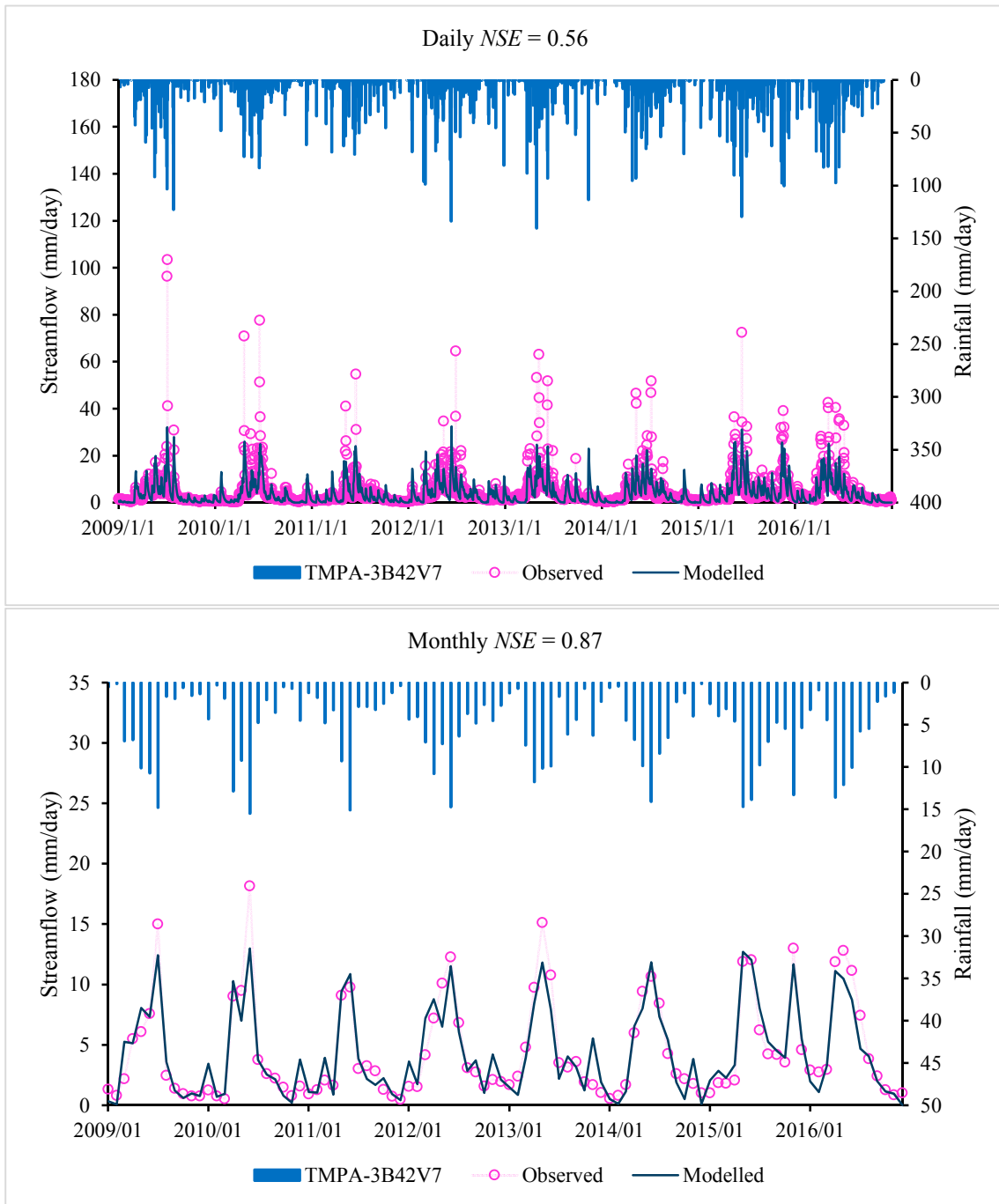
Table 8. Model performance of the Sacramento model (calibrated using *NSE*) for the calibration and verification periods.

	Datasets	Daily rel.bias	DailyNSE	DailyNSE _{sq}	Daily NSE _{log}	Monthly NSE
Calibration period	Gauged	−0.12	0.57	0.56	0.41	0.84
Validation period	Gauged	−0.11	0.47	0.41	0.29	0.77
Calibration period	TMPA-3B42V7	0.02	0.51	0.56	0.47	0.86
Validation period	TMPA-3B42V7	−0.06	0.61	0.54	0.42	0.89
Calibration period	CMADS	−0.05	0.68	0.63	0.46	0.93
Validation period	CMADS	0.01	0.71	0.57	0.40	0.91



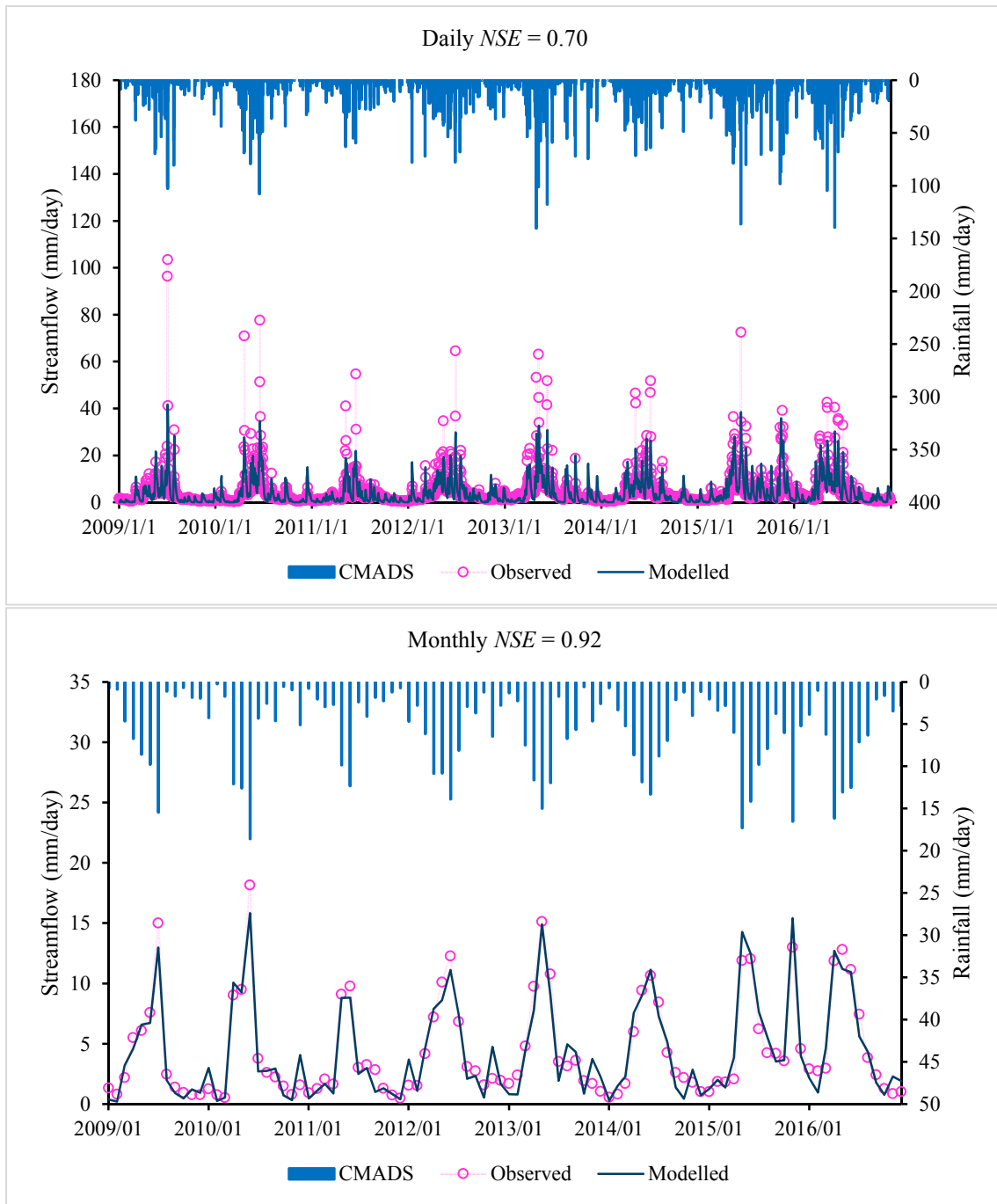
(a) Gauge-interpolated

Figure 4. Cont.



(b) TPA-3B42V7

Figure 4. Cont.



(c) CMDAS V1.0

Figure 4. Cont.

Figure 4. Observed and Sacramento-model-simulated daily and monthly runoffs for (a) Gauge-interpolated, (b) TMPA-3B42V7 and (c) CMDAS rainfall datasets (for the Sacramento model calibrated using *NSE* as the objective function).

These results show that, among the three precipitation datasets considered here for the Lijiang River basin, the CMADS precipitation datasets have a higher accuracy and better applicability in calibrating and validating the rainfall–runoff models. The reason that the gauge-interpolated rainfall

always provides the worst result in flow simulations is due to the meteorological stations available in the Lijiang river basin being too sparsely distributed to permit reliable interpolation.

3.2. Precipitation Detection

We examined the performance of TRMA-3B42V7 and CMADS using the six statistical measures (PC, POD, FBI, FAR, CSI and HSS) through the pixel-to-point comparison with the ground-based data. Following the studies of Dai [67] and Vu et al. [24], a minimum precipitation threshold of 1.0 mm per day was used for the precipitation and non-precipitation event for ground observation and satellite/reanalysis estimate. The contingency statistics of TRMA-3B42V7 and CMADS were evaluated each year through the precipitation datasets from 2008 to 2016. Figure 5 shows the contingency statistics calculated for the TRMA-3B42V7 and CMADS.

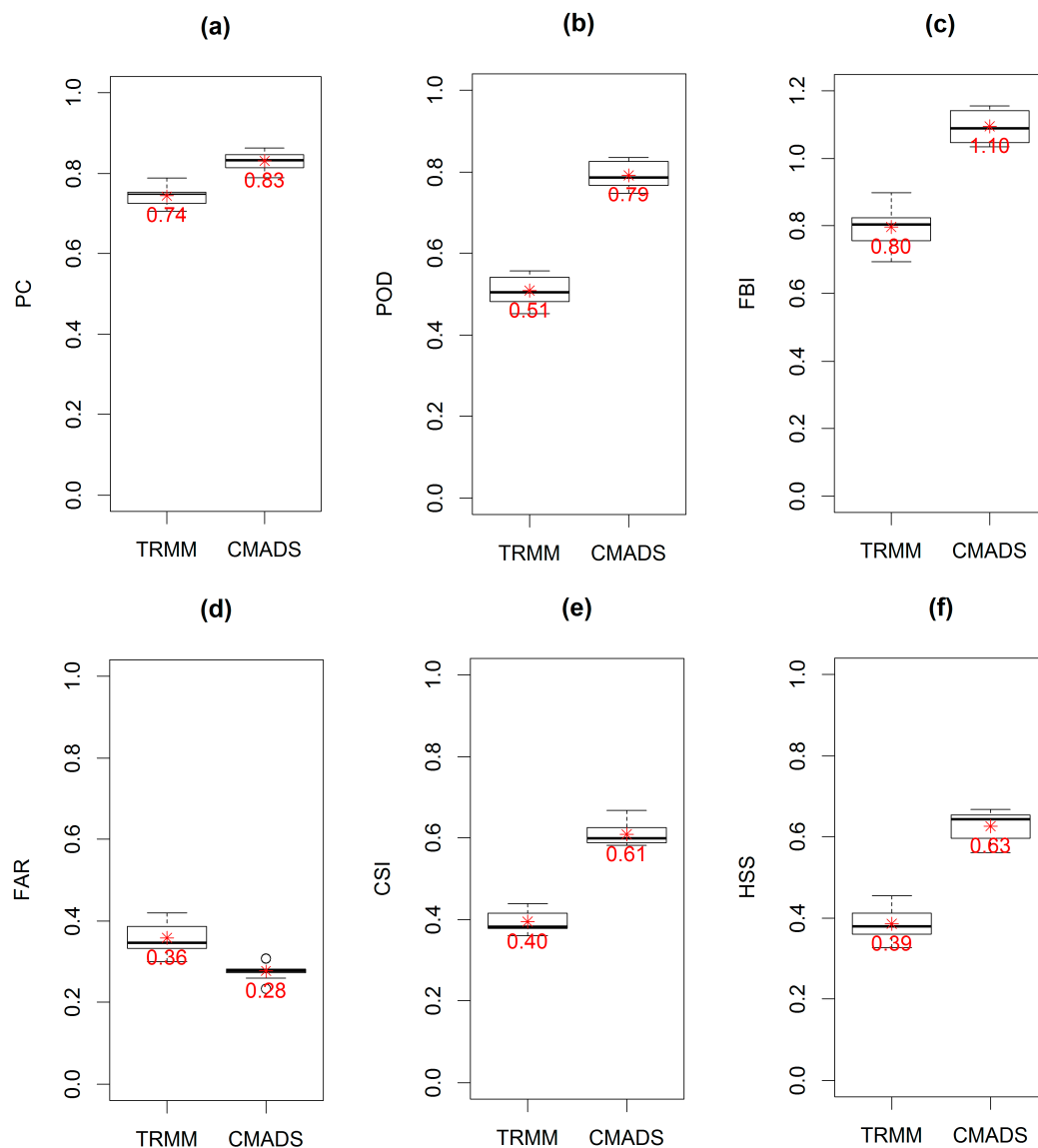


Figure 5. The box plots for the contingency statistics of (a) Proportion Correct (PC), (b) Probability of Detection (POD), (c) Frequency Bias Index (FBI), (d) False Alarm Ratio (FAR), (e) Critical Success Index (CSI), (f) Heidke Skill Score (HSS). The labelled asterisk dot represent the mean value and the middle line in the box represent the median value. Each box ranges from the lower (25th) to upper quartile (75th).

For the statistics PC, POD and CSI, the CMADS scheme (with an average of 0.83, 0.79 and 0.61) shows higher values than TRMM (with an average of 0.74, 0.51 and 0.40). For the statistic FBI, CMADS gives a mean of 1.1 (ranging from 1.03 to 1.16) and is closer to the perfect score than the FBI of TRMM with a mean of 0.8 (ranging from 0.69 to 0.9). With respect to the FAR statistic, CMADS has a smaller FAR with an average of 0.28 (ranging from 0.23 to 0.31) than TRMM with an average of 0.36 (ranging from 0.30 to 0.42). Finally, the HSS statistic for CMADS has a larger value with a mean of 0.63 (ranging 0.56 to 0.67) than TRMM with a mean of 0.39 (ranging from 0.33 to 0.45). These results indicate that the CMADS scheme shows better performance than TRMA-3B42V7 for all the six contingency statistics. Overall, compared to TRMM data, CMADS show better agreement with the ground observation data in Lijiang river basin.

3.3. Uncertainty Analysis

All precipitation products are limited by quantitative inaccuracies and they can exhibit significant bias and errors in spatial and temporal variability. As the runoff-generation is highly sensitive to the spatial and temporal variability of precipitation data, the spatial and temporal variability of precipitation is one of the main source of uncertainty in rainfall-runoff modelling.

Parameter uncertainty is another source of uncertainty in rainfall-runoff modelling. The parameter uncertainty of each hydrologic model has been explored in this study. A GLUE uncertainty analysis is applied to assess parameter uncertainty of hydrologic models here.

In the first case, for the IHACRES model, 100,000 samples are chosen from a uniform distribution for each parameter and the performance measures NSE and NSE_{log} are used as the “likelihood” functions. Using a threshold value of $NSE > 0.67$ (or $NSE_{log} > 0.78$) for the CMADS product, the GLUE algorithm finds 2000 (1416) behavioral solutions in 100,000 simulations with the IHACRES model. Using a threshold value of $NSE > 0.56$ (or $NSE_{log} > 0.69$) for the TMPA-3B42V7 product, the GLUE algorithm finds 3180 (2819) behavioral solutions in 100,000 simulations with the IHACRES model.

The red (blue) dots and lines in Figure 6 represent the distribution and boundary of behavioral parameters when using NSE_{log} (NSE) as the likelihood function, with the calibrated parameter set for each rainfall dataset indicated. From the distribution of behavioral parameter sets, we see the parameters d , f and tau_s have behavioral values distributed across the full parameter range, indicating these have the greatest uncertainty. In comparison, the distribution of v_s for behavioral parameter sets is constricted to smaller values ($< \sim 0.5$), particularly when using considering NSE . The values of tau_q are more constrained when considering NSE , due to the focus NSE gives to high flows compared to NSE_{log} . Generally, there is little interaction between most of the parameters. The main exception is the e and f parameters. The value of the e parameter is constrained to $< \sim 0.1$ providing $f > \sim 1$, increasing rapidly for smaller values of f . This indicates a highly non-linear interaction between these parameters. It should also be noted that the optimal value of the e parameter is considerably smaller than that found in Australia (0.166) found by Chapman (2001) [68], due to the influence of other factors (e.g., atmospheric transmissivity).

A similar analysis is applied for the Sacramento model, where 100,000 samples are chosen obeying a uniform distribution, and NSE and NSE_{log} are again defined as the likelihood functions. Using a threshold value of $NSE > 0.45$ (or $NSE_{log} > 0.53$), the GLUE algorithm finds 929 (359) behavioral solutions in 100,000 simulations for the CMADS product with the Sacramento model. Using a threshold value of $NSE > 0.32$ (or $NSE_{log} > 0.43$), the GLUE algorithm finds 1294 (140) behavioral solutions in 100,000 simulations for the TMPA-3B42V7 product with the Sacramento model. The pairwise correlation of behavioral parameters for the Sacramento model is shown in Figure 7.

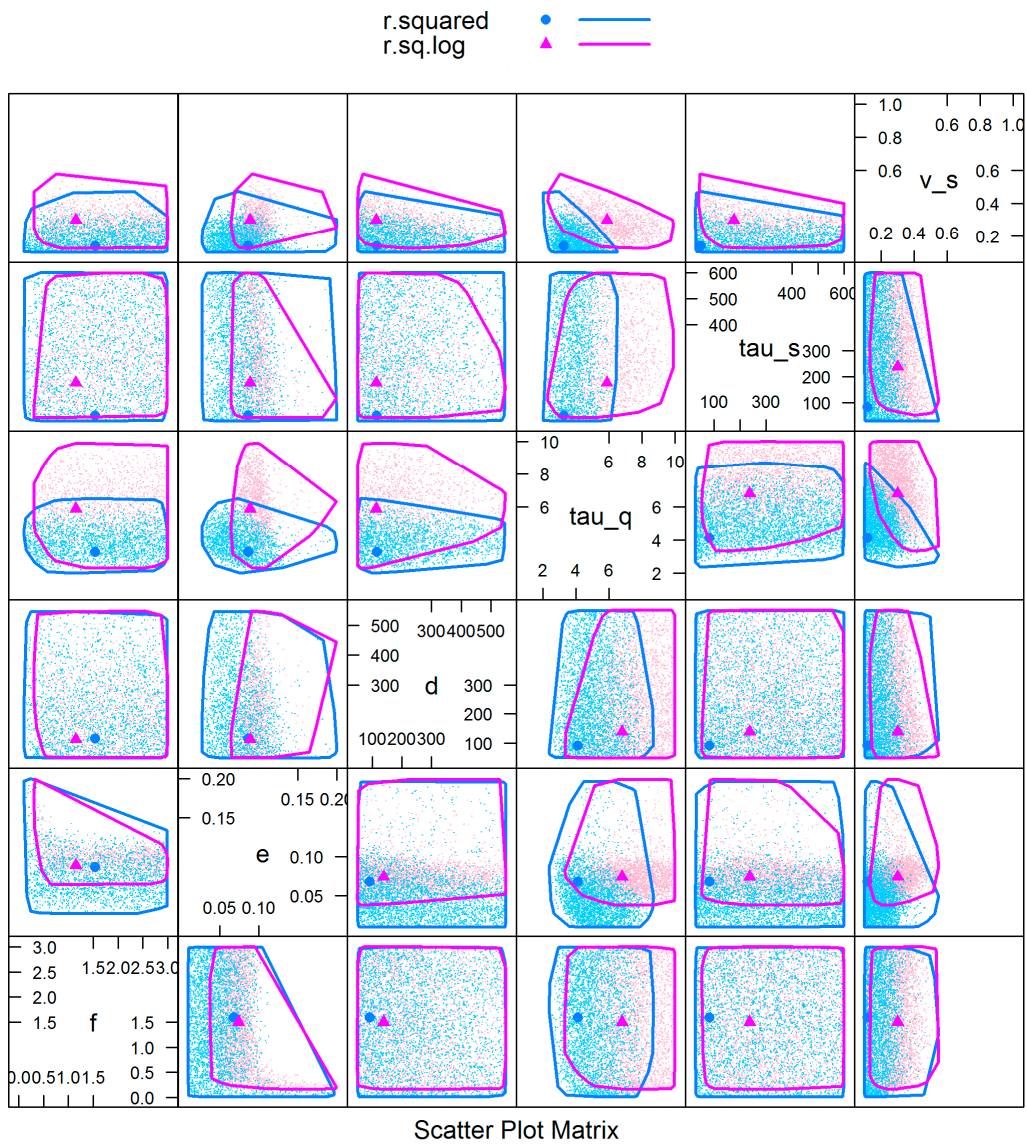


Figure 6. Two-dimensional projections of pairwise correlation of behavioral parameters for the IHACRES model using the CMADS (above diagonal) and TPA-3B42V7 (below diagonal) precipitation datasets. The heavy dots represent the location of the best objective function value obtained from the GLUE sample.

The meaning of the red (blue) dot and line in Figure 7 is similar to that in Figure 6. From the distribution of behavioral parameters, we can see the parameters *uztwm*, *adimp* and *lztwm* show less uncertainties overall. When *NSE* is selected as the likelihood function, the distribution of parameters *lztwm*, *uztwm* and *adimp* are relatively low, while the distributions of parameters *lzpk* and *lzsk* are relatively high; when *NSE_{log}* is selected as the likelihood function, the distribution of the parameters *uztwm*, *lztwm* and *adimp* is relatively low and the distribution of the parameter *pfree* is relatively high, which indicates that these parameters may be more sensitive and less uncertain.

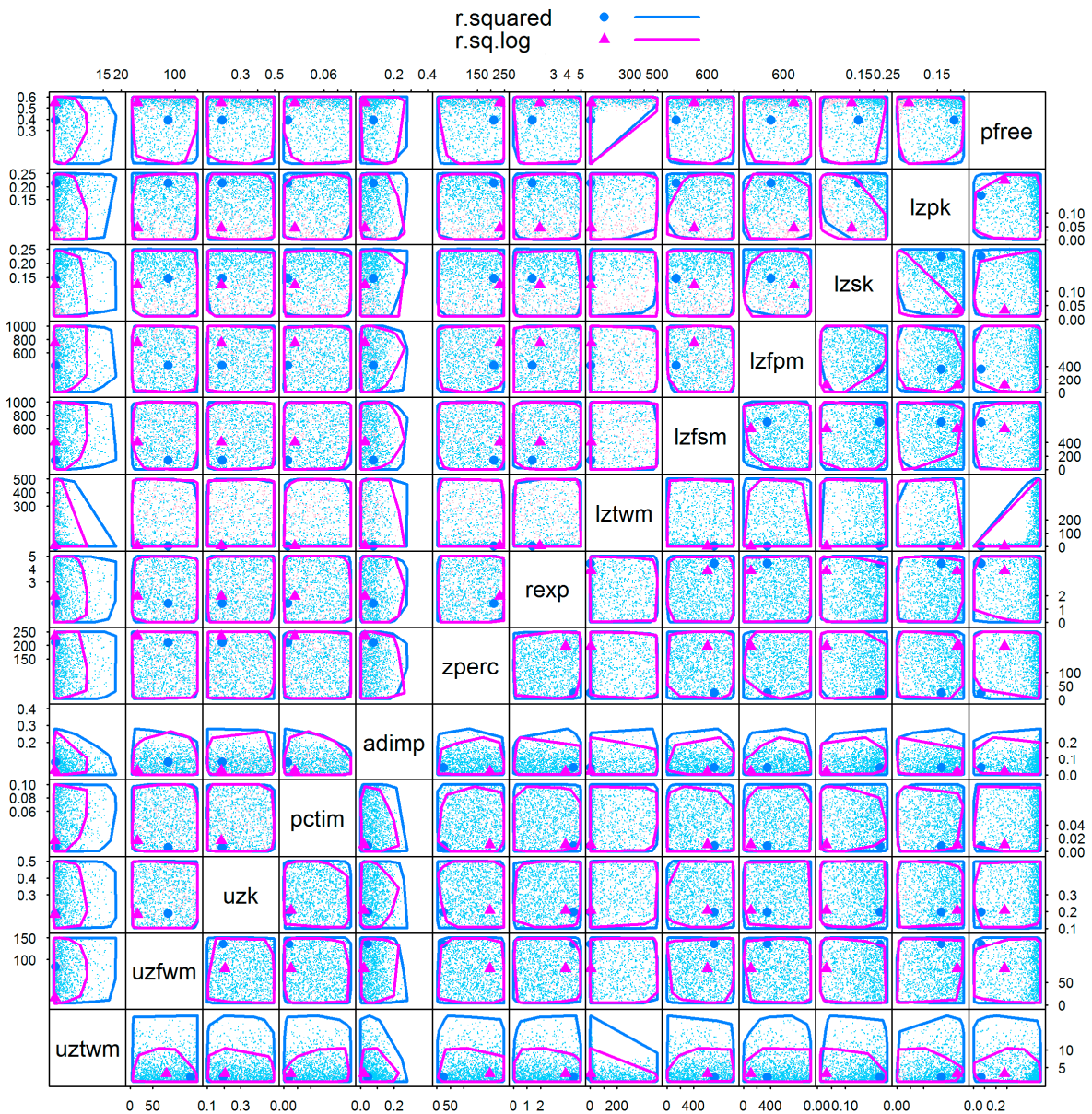


Figure 7. Two-dimensional projections of the pairwise correlation of behavioral parameters for the Sacramento model using CMADS (above diagonal) and TMPA-3B42V7 (below diagonal) precipitation datasets. The heavy dots represent the location of the best objective function value obtained from the GLUE sample.

In the second case, model performance is considered satisfactory when *NSE* is greater than 0.5 [69,70], with likelihood-function values > 0.5 defined as behavioral parameter sets. The behavioral parameter space may be used as further criteria for the evaluation of different precipitation products. The behavioral parameter space describes the number (and percentage) of behavioral solutions in 100,000 simulations, with 100,000 parameter sets generated by the same Monte Carlo random sampling method, with the same criteria of acceptability employed (i.e., the same threshold value and the objective function) for the different precipitation datasets. The statistics of the number and percentage of behavioral parameter sets for different precipitation schemes driving the IHACRES model are shown in Table 9.

Table 9. Number (percentage) of behavioral parameter sets for the IHACRES model.

	<i>NSE</i>	<i>NSE_{log}</i>
Gauged	576 (0.58%)	9144 (9.14%)
TMPA-3B42V7	1056 (1.06%)	9929 (9.93%)
CMADS	5186 (5.19%)	12,095 (12.10%)

Using a threshold value of $NSE > 0.5$ (or $NSE_{log} > 0.5$), the GLUE algorithm finds 5186 (12,095) behavioral solutions in 100,000 simulations for the CMADS precipitation product with the IHACRES model, while the GLUE algorithm achieves 1056 (9929) and 576 (9144) behavioral solutions in 100,000 simulations for the TMPA-3B42V7 and gauge-interpolated products with the IHACRES model. Therefore, the behavioral parameter space of the IHACRES model driven by the CMADS precipitation is larger than the behavioral parameter space driven by the other two precipitation inputs. The CMADS product gives a better performance than the TMPA-3B42V7 and gauge-interpolated products, because CMADS assimilated datasets are based on the large number of stations (nearly 40,000 regional automatic stations and 2421 national automatic stations in China), which gives it any priority in reflecting the actual processes of areal precipitation.

A similar analysis is applied for the Sacramento model as well. The statistics of the number and percentage of behavioral parameter sets for different precipitation schemes driving the Sacramento model are shown in Table 10. For the Sacramento model, the behavioral parameter space is very sparse when using the GLUE method for the Lijiang River basin. Using a threshold value of $NSE > 0.5$ (or $NSE_{log} > 0.5$), the GLUE algorithm finds 32 (60) behavioral solutions in 100,000 simulations for the CMADS precipitation dataset within the Sacramento model, while the GLUE algorithm achieves 0 (11) and 0 (5) behavioral solutions in 100,000 simulations for the TMPA-3B42V7 and gauge-interpolated products with the Sacramento model. Similar to their performance with the IHACRES model, the behavioral parameter space of the Sacramento model driven by the CMADS precipitation dataset is larger than the behavioral parameter space driven by the other two precipitation inputs. The CMADS product shows a better performance than the TMPA-3B42V7 and gauge-interpolated products, which, as mentioned before, is probably because the CMADS assimilated datasets are based on the strongly underconstrained large number of stations.

Table 10. Number (percentage) of behavioral parameter sets for the Sacramento model.

	<i>NSE</i>	<i>NSE_{log}</i>
Gauged	0(0.00%)	5(0.01%)
TMPA-3B42V7	0(0.00%)	11(0.01%)
CMADS	32(0.03%)	60(0.06%)

4. Discussion

Our work presents a comparative analysis for different precipitation datasets and their applicability for hydrologic modelling, including gauge-interpolated datasets, TMPA-3B42V7 and CMADS precipitation products. Two hydrologic models; IHACRES and Sacramento, are evaluated in the Lijiang River basin, as well as the accuracy of different precipitation datasets for hydrologic modelling.

The results show that the IHACRES and Sacramento models demonstrate a good and similar performance in the Lijiang River basin. Driven by the CMADS precipitation, the *NSE* values of the IHACRES (Sacramento) model are 0.69 (0.68) and 0.70 (0.71) for the calibration and validation periods, respectively. Figure 6 shows there are three sensitive parameters (f , e and v_s) for the IHACRES model. Figure 7 shows there are three sensitive parameters ($uztwm$, $adimp$ and $lztwm$) for the Sacramento model. The number of effective parameters is similar in both Sacramento model (with more parameters) and IHACRES model (with less parameters), consistent with sensitivity results in Shin et al. [64].

The uncertainty analysis carried out for the IHACRES and Sacramento models show that the uncertainty in the model predictions is greater for the Sacramento model. To sum up, IHACRES and Sacramento perform similarly in terms of simulation performance and number of effective parameters, the latter model having far more insensitive parameters. What's more, the IHACRES model has a reduced uncertainty compared with the Sacramento model. Based on these analyses, the authors conclude that IHACRES generally outperforms Sacramento in Lijiang river basin. It confirms previous findings (e.g., Orth et al. [71]) that more parameters may lead to over-fitting without an improved performance of the hydrologic model.

Of the three precipitation datasets (gauge-interpolated product, TMPA-3B42V7 and CMADS products), the CMADS product gives the best performance in simulating the rainfall–runoff process in the Lijiang River basin. The overall performance (based on DailyNSE values) of the CMADS product is 0.69 and 0.70 for the IHACRES and Sacramento models, respectively, with the overall performance of the TMPA-3B42V7 (0.56 and 0.56) and gauge-interpolated (0.57 and 0.52) products correspondingly much lower. From the analysis of Figures 3 and 4, the hydrologic model driven by the CMADS product shows a superior skill in capturing the flow peaks because the CMADS reanalysis data are based on a large number of stations. However, these datasets overestimate the simulated flood peak and underestimate low flowrates, which is probably because the models are calibrated using the performance measure NSE , being an objective function that puts more emphasis on high flowrates. The model calibrated using Nash–Sutcliffe efficiency on transformed streamflow NSE_{log} , which gives more emphasis to low flowrates, is presented in the Supplementary Materials, as well as the model performance calibrated using NSE_{log} as an objective function. Similar conclusions can be reached with the models calibrated using the performance measure NSE_{log} as the objective function. The CMADS precipitation datasets perform best in all three precipitation datasets, followed by the TMPA-3B42V7 precipitation and then the gauge-interpolated product. Comparing the model performance using NSE_{log} and NSE as the objective functions for calibration, the model calibrated using the performance measure NSE performed better in simulating peak flows, while underestimating low flowrates. In contrast, the model calibrated using the performance measure NSE_{log} performs well in simulating low flowrates but underestimates the flood peak in the simulations.

The GLUE pairwise correlation of behavioral parameters (Figures 5 and 6) give us an intuitional view of parameter uncertainties. From the distribution of behavioral parameters for the IHACRES model (Figure 5), we see the parameters d , tau_s and tau_q have the greatest uncertainties overall. Further analysis reveals that different precipitation products reshape the distribution of tau_q greatly. The parameter tau_q , which represents the time constant for quick flow store, is very sensitive to the precipitation input. For the Sacramento model (Figure 6), the parameters $uztwm$, $adimp$ and $lztwm$ show less uncertainties overall than other parameters.

The superiority of the CMADS product can also be found in the number of GLUE behavioral parameters (or their occupation percentage among all the uniformly distributed parameter sets), as well as their GLUE relative measurements coverage. From Tables 7 and 8, we find the CMADS driven hydrologic models are responsible for more behavioral parameters than the hydrologic models driven by the other two precipitation datasets.

Although the two hydrologic models introduced here are widely used, the precipitation input data are basin-averaged precipitation. The comparison and applicability of different precipitation datasets maybe be affected by this “average” precipitation, since the spatial distribution and variability of different precipitation datasets may be weakened by the effect of spatial averaging. The improved accuracy of the CMADS precipitation dataset may be more obvious with the simulation of a distributed or semi-distributed model as the rainfall input. Moreover, different uncertainty-analysis methods may affect the efficiency of the uncertainty analysis, which thus requires further research.

5. Conclusions

Precipitation is a fundamental component of the global water cycle. Precipitation datasets range from conventional ground-based datasets to remote-sensing products and reanalysis datasets, such as the gauge-interpolated product, the TMPA-3B42V7 precipitation products and the CMADS datasets invoked here. The two hydrologic models (IHACRES and Sacramento) are introduced to evaluate their applicability in the basin, the impact of model complexity and the applicability of different precipitation datasets on the hydrologic modelling.

CMADS gives best results when used in IHACRES and Sacramento to simulate flow in the Lijiang River basin. The CMADS precipitation datasets (DailyNSE = 0.69 for the IHACRES model; DailyNSE = 0.70 for the Sacramento model) give improved applicability and accuracy compared with the gauge-interpolated datasets (DailyNSE = 0.57 for the IHACRES model; DailyNSE = 0.52 for the Sacramento model) and TMPA-3B42V7 datasets (DailyNSE = 0.56 for the IHACRES model; DailyNSE = 0.56 for the Sacramento model) in the Lijiang River basin. From the analysis of Figures 3 and 4, we conclude that the CMADS precipitation-driven hydrologic models give better skill in capturing the streamflow peaks. Interpolation of gauge data performed worst, reflecting the impact of low gauge density

The precipitation detection ability of TRMA-3B42V7 and CMADS is also evaluated using six statistical measures (*PC*, *POD*, *FBI*, *FAR*, *CSI* and *HSS*) through a pixel-to-point comparison to the ground-based data. CMADS (with an average of 0.83, 0.79, 1.1, 0.28, 0.61 and 0.63) shows better performance and is closer to the perfect score than TRMM (with an average of 0.74, 0.51, 0.80, 0.36, 0.4 and 0.39).

Based on the analysis of Table 7, for the IHACRES model and using *NSE* as the likelihood function, the number and percentage of behavioral parameters for the CMADS, TMPA-3B42V7 and gauge-interpolated product are 5186 (5.19%), 1056 (1.06%) and 576 (0.58%). Using *NSE_{log}* as the likelihood function, the number and percentage of behavioral parameters for the corresponding precipitation datasets are 12,095 (12.1%), 9929 (9.93%) and 9144 (9.14%). Similar phenomena can be found with the same analysis of the Sacramento model in Table 8 but the behavioral parameter sets for that model are very sparse. We conclude that the CMADS precipitation-driven hydrologic models are more accurate, as they are responsible for more behavioral parameters than the hydrologic models driven by the other two precipitation datasets. The TMPA-3B42V7 datasets show slightly better performance than the gauge-interpolated product in this case study, indicating that global datasets are particularly useful in poorly gauged areas.

The performance of the IHACRES model (DailyNSE = 0.69 driven by the CMADS product) and Sacramento model (DailyNSE = 0.70 driven by the CMADS product) give respectable results. While both models work well, IHACRES gives lower predictive uncertainty compared to Sacramento, which implies that the general applicability of the IHACRES model is preferable to the Sacramento model in the Lijiang River basin.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4441/10/11/1611/s1>, Figure S1: Density map of precipitation stations over the region (units: gauges per km²), Table S1: Mean absolute error (MAE), root mean square error (RMSE) and relative error of cross-validation for daily interpolated precipitation grids, Table S2: Overall performance (daily *NSE_{log}* (monthly *NSE*)) of precipitation datasets for models using *NSE_{log}* as the objective function for calibration, Table S3 Model performance of IHACRES model (calibrated using *NSE_{log}*) for the calibration period and validation periods, Table S4 Model performance of Sacramento model (calibrated using *NSE_{log}*) for the calibration period and validation periods, Figure S2: Observed and IHACRES-model-simulated daily and monthly runoffs for (a) Gauge-interpolated, (b) TMPA-3B42V7 and (c) CMDAS rainfall datasets (for the IHACRES model calibrated using *NSE_{log}* as the objective function), Figure S3: Observed and Sacramento-model-simulated daily and monthly runoffs for (a) Gauge-interpolated, (b) TMPA-3B42V7 and (c) CMDAS rainfall datasets (for the Sacramento model calibrated using *NSE_{log}* as the objective function)).

Author Contributions: B.G. carried out the experiments, B.G. and J.Z. wrote the draft manuscript. B.G., T.X., B.C. and A.J. performed data analysis and designed the experiments. Y.S., Q.Y., X.L. and W.L. collected the data for the experiments. All authors contributed to the discussions and reviewed the manuscript.

Funding: This work is supported by National Key R&D Program of China (2017YFC0406004) and NSFC (41271004).

Conflicts of Interest: The authors declare no conflicts of interests.

References

- Feng, D.; Zheng, Y.; Mao, Y.; Zhang, A.; Wu, B.; Li, J.; Tian, Y.; Wu, X. An integrated hydrological modeling approach for detection and attribution of climatic and human impacts on coastal water resources. *J. Hydrol.* **2018**, *557*, 305–320. [[CrossRef](#)]
- Michaelides, S.; Levizzani, V.; Anagnostou, E.; Bauer, P.; Kasparis, T.; Lane, J.E. Precipitation: Measurement, remote sensing, climatology and modeling. *Atmos. Res.* **2009**, *94*, 512–533. [[CrossRef](#)]
- Steiner, M.; Smith, J.A.; Burges, S.J.; Alonso, C.V.; Darden, R.W. Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.* **1999**, *35*, 2487–2503. [[CrossRef](#)]
- Himanshu, S.K.; Pandey, A.; Patil, A. Hydrologic Evaluation of the TMPA-3B42V7 Precipitation Data Set over an Agricultural Watershed Using the SWAT Model. *J. Hydrol. Eng.* **2018**, *23*, 5018003. [[CrossRef](#)]
- Andreadis, K.M.; Schumann, G.J.P.; Stampoulis, D.; Bates, P.D.; Brakenridge, G.R.; Kettner, A.J. Can Atmospheric Reanalysis Data Sets Be Used to Reproduce Flooding Over Large Scales? *Geophys. Res. Lett.* **2017**, *44*, 10369–10377. [[CrossRef](#)]
- Joyce, R.J.; Janowiak, J.E.; Arkin, P.A.; Xie, P. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *J. Hydrometeorol.* **2004**, *5*, 487–503. [[CrossRef](#)]
- Kubota, T.; Shige, S.; Hashizume, H.; Aonashi, K.; Takahashi, N.; Seto, S.; Hirose, M.; Takayabu, Y.N.; Ushio, T.; Nakagawa, K.; et al. Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2259–2275. [[CrossRef](#)]
- Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.; Hong, Y.; Bowman, K.P.; Stocker, E.F. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [[CrossRef](#)]
- Kidd, C.; Huffman, G. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [[CrossRef](#)]
- Himanshu, S.K.; Pandey, A.; Yadav, B. Assessing the applicability of TMPA-3B42V7 precipitation dataset in wavelet-support vector machine approach for suspended sediment load prediction. *J. Hydrol.* **2017**, *550*, 103–117. [[CrossRef](#)]
- Erazo, B.; Bourrel, L.; Frappart, F.; Chimborazo, O.; Labat, D.; Dominguez-Granda, L.; Matamoros, D.; Mejia, R. Validation of Satellite Estimates (Tropical Rainfall Measuring Mission, TRMM) for Rainfall Variability over the Pacific Slope and Coast of Ecuador. *Water* **2018**, *10*, 213. [[CrossRef](#)]
- Jiang, S.; Liu, S.; Ren, L.; Yong, B.; Zhang, L.; Wang, M.; Lu, Y.; He, Y. Hydrologic Evaluation of Six High Resolution Satellite Precipitation Products in Capturing Extreme Precipitation and Streamflow over a Medium-Sized Basin in China. *Water* **2018**, *10*, 25. [[CrossRef](#)]
- Yang, Y.; Tang, G.; Lei, X.; Hong, Y.; Yang, N. Can Satellite Precipitation Products Estimate Probable Maximum Precipitation: A Comparative Investigation with Gauge Data in the Dadu River Basin. *Remote Sens.* **2018**, *10*, 41. [[CrossRef](#)]
- Tang, G.; Behrangi, A.; Long, D.; Li, C.; Hong, Y. Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *J. Hydrol.* **2018**, *559*, 294–306. [[CrossRef](#)]
- Saha, S.; Moorthi, S.; Pan, H.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Kistler, R.; Woollen, J.; Behringer, D.; et al. The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.* **2010**, *91*, 1015–1058. [[CrossRef](#)]
- Uppala, S.M.; Kållberg, P.W.; Simmons, A.J.; Andrae, U.; Bechtold, V.D.C.; Fiorino, M.; Gibson, J.K.; Haseler, J.; Hernandez, A.; Kelly, G.A.; et al. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 2961–3012. [[CrossRef](#)]
- Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.A.; Balsamo, G.; Bauer, P.; et al. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [[CrossRef](#)]

18. Compo, G.P.; Whitaker, J.S.; Sardeshmukh, P.D.; Matsui, N.; Allan, R.J.; Yin, X.; Gleason, B.E.; Vose, R.S.; Rutledge, G.; Bessemoulin, P.; et al. The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 1–28. [[CrossRef](#)]
19. Meng, X.; Wang, H. Significance of the China Meteorological Assimilation Driving Datasets for the SWAT Model (CMADS) of East Asia. *Water* **2017**, *9*, 765. [[CrossRef](#)]
20. Zhou, S.; Wang, Y.; Chang, J.; Guo, A.; Li, Z. Investigating the Dynamic Influence of Hydrological Model Parameters on Runoff Simulation Using Sequential Uncertainty Fitting-2-Based Multilevel-Factorial-Analysis Method. *Water* **2018**, *10*, 1177. [[CrossRef](#)]
21. Meng, X.; Wang, H.; Wu, Y.; Long, A.; Wang, J.; Shi, C.; Ji, X. Investigating spatiotemporal changes of the land-surface processes in Xinjiang using high-resolution CLM3.5 and CLDAS: Soil temperature. *Sci. Rep.* **2017**, *7*, 13286. [[CrossRef](#)] [[PubMed](#)]
22. Meng, X.; Wang, H.; Shi, C.; Wu, Y.; Ji, X. Establishment and Evaluation of the China Meteorological Assimilation Driving Datasets for the SWAT Model (CMADS). *Water* **2018**, *10*, 1555. [[CrossRef](#)]
23. Zhao, F.; Wu, Y.; Qiu, L.; Sun, Y.; Sun, L.; Li, Q.; Niu, J.; Wang, G. Parameter Uncertainty Analysis of the SWAT Model in a Mountain-Loess Transitional Watershed on the Chinese Loess Plateau. *Water* **2018**, *10*, 690. [[CrossRef](#)]
24. Vu, T.; Li, L.; Jun, K. Evaluation of Multi-Satellite Precipitation Products for Streamflow Simulations: A Case Study for the Han River Basin in the Korean Peninsula, East Asia. *Water* **2018**, *10*, 642. [[CrossRef](#)]
25. Cao, Y.; Zhang, J.; Yang, M.; Lei, X.; Guo, B.; Yang, L.; Zeng, Z.; Qu, J. Application of SWAT Model with CMADS Data to Estimate Hydrological Elements and Parameter Uncertainty Based on SUFI-2 Algorithm in the Lijiang River Basin, China. *Water* **2018**, *10*, 742. [[CrossRef](#)]
26. Shao, G.; Guan, Y.; Zhang, D.; Yu, B.; Zhu, J. The Impacts of Climate Variability and Land Use Change on Streamflow in the Hailiutu River Basin. *Water* **2018**, *10*, 814. [[CrossRef](#)]
27. Gao, X.; Zhu, Q.; Yang, Z.; Wang, H. Evaluation and Hydrological Application of CMADS against TRMM 3B42V7, PERSIANN-CDR, NCEP-CFSR, and Gauge-Based Datasets in Xiang River Basin of China. *Water* **2018**, *10*, 1225. [[CrossRef](#)]
28. Tian, Y.; Zhang, K.; Xu, Y.; Gao, X.; Wang, J. Evaluation of Potential Evapotranspiration Based on CMADS Reanalysis Dataset over China. *Water* **2018**, *10*, 1126. [[CrossRef](#)]
29. Meng, X.; Wang, H.; Lei, X.; Cai, S.; Wu, H.; Ji, X.; Wang, J. Hydrological modeling in the Manas River Basin using soil and water assessment tool driven by CMADS. *Teh. Vjesn.* **2017**, *24*, 525–534.
30. Li, J.; Heap, A.D. Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Model. Softw.* **2014**, *53*, 173–189. [[CrossRef](#)]
31. Benoit, L.; Mariethoz, G. Generating synthetic rainfall with geostatistical simulations. *Wiley Interdiscip. Rev. Water* **2017**, *4*, e1199. [[CrossRef](#)]
32. Ly, S.; Charles, C.; Degré, A. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: A review. *Biotechnol. Agron. Soc. Environ.* **2013**, *17*, 392–406.
33. Vente, C. *Handbook of Applied Hydrology: A Compendium of Water-Resources Technology*; McGraw-Hill: New York, NY, USA, 1964.
34. Thiessen, A.H. Precipitation averages for large areas. *Mon. Weather Rev.* **1911**, *39*, 1082–1089. [[CrossRef](#)]
35. Philip, G.M.; Watson, D.F. A precise method for determining contoured surfaces. *APPEA J.* **1982**, *22*, 205–212. [[CrossRef](#)]
36. Cressie, N.A. *Statistics for Spatial Data*; Wiley Online Library: New York, NY, USA, 1993.
37. Hutchinson, M.F. Interpolation of rainfall data with thin plate smoothing splines. Part II: Analysis of topographic dependence. *J. Geogr. Inf. Decis. Anal.* **1998**, *2*, 152–167.
38. Hutchinson, M.F. Interpolation of rainfall data with thin plate smoothing splines. Part I: Two dimensional smoothing of data with short range correlation. *J. Geogr. Inf. Decis. Anal.* **1998**, *2*, 139–151.
39. Hutchinson, M.F. Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Inf. Syst.* **1995**, *9*, 385–403. [[CrossRef](#)]
40. Arowolo, A.O.; Bhowmik, A.K.; Qi, W.; Deng, X. Comparison of spatial interpolation techniques to generate high-resolution climate surfaces for Nigeria. *Int. J. Climatol.* **2017**, *371*, 179–192. [[CrossRef](#)]
41. Taesombat, W.; Sriwongsitanon, N. Areal rainfall estimation using spatial interpolation techniques. *Sci. Asia* **2009**, *35*, 268–275. [[CrossRef](#)]

42. Guo, B.; Xu, T.; Zhang, J.; Barry, C.; Jakeman, A.; Seo, L.; Lei, X.; Liao, W. A comparative analysis of precipitation estimation methods for streamflow prediction. In Proceedings of the 22nd International Congress on Modelling and Simulation (MODSIM2017), Hobart, Tasmania, Australia, 3–8 December 2017; Modelling and Simulation Society of Australia and New Zealand Inc.: Canberra, Australia, 2017; pp. 43–49.
43. Sahlou, D.; Moges, S.A.; Nikolopoulos, E.I.; Anagnostou, E.N.; Hailu, D. Evaluation of High-Resolution Multisatellite and Reanalysis Rainfall Products over East Africa. *Adv. Meteorol.* **2017**, *2017*, 4957960. [[CrossRef](#)]
44. Smith, R.A.; Kummerow, C.D. A Comparison of in Situ, Reanalysis, and Satellite Water Budgets over the Upper Colorado River Basin. *J. Hydrometeorol.* **2013**, *14*, 888–905. [[CrossRef](#)]
45. Pfeifroth, U.; Mueller, R.; Ahrens, B. Evaluation of Satellite-Based and Reanalysis Precipitation Data in the Tropical Pacific. *J. Appl. Meteorol. Clim.* **2013**, *52*, 634–644. [[CrossRef](#)]
46. Smith, M.B.; Koren, V.I.; Zhang, Z.; Reed, S.M.; Pan, J.; Moreda, F. Runoff response to spatial variability in precipitation: An analysis of observed data. *J. Hydrol.* **2004**, *298*, 267–286. [[CrossRef](#)]
47. Yang, J.; Jakeman, A.; Fang, G.; Chen, X. Uncertainty analysis of a semi-distributed hydrologic model based on a Gaussian Process emulator. *Environ. Model. Softw.* **2018**, *101*, 289–300. [[CrossRef](#)]
48. Hutchinson, M.F.; Xu, T. *Anusplin Version 4.4 User Guide*; Fenner School of Environment and Society, The Australian National University: Canberra, Australia, 2013.
49. Hutchinson, M.F. *Anudem Version 5.3 User Guide*; Fenner School of Environment and Society, The Australian National University: Canberra, Australia, 2011.
50. Hutchinson, M.F. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *J. Hydrol.* **1989**, *106*, 211–232. [[CrossRef](#)]
51. Zheng, X.; Xiong, H.; Yue, L.; Gong, J. An improved ANUDEM method combining topographic correction and DEM interpolation. *Geocarto Int.* **2016**, *31*, 492–505. [[CrossRef](#)]
52. Post, D.A.; Chiew, F.H.S.; Teng, J.; Viney, N.R.; Ling, F.L.N.; Harrington, G.; Crosbie, R.S.; Graham, B.; Marvanek, S.; McLoughlin, R. A robust methodology for conducting large-scale assessments of current and future water availability and use: A case study in Tasmania, Australia. *J. Hydrol.* **2012**, *412–413*, 233–245. [[CrossRef](#)]
53. Kim, K.B.; Kwon, H.; Han, D. Exploration of warm-up period in conceptual hydrological modelling. *J. Hydrol.* **2018**, *556*, 194–210. [[CrossRef](#)]
54. Kan, G.; Zhang, M.; Liang, K.; Wang, H.; Jiang, Y.; Li, J.; Ding, L.; He, X.; Hong, Y.; Zuo, D.; et al. Improving water quantity simulation & forecasting to solve the energy-water-food nexus issue by using heterogeneous computing accelerated global optimization method. *Appl. Energy* **2018**, *210*, 420–433.
55. Croke, B.; Jakeman, A.J. A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environ. Model. Softw.* **2004**, *19*, 1–5. [[CrossRef](#)]
56. Petheram, C.; Rustomji, P.; Chiew, F.H.S.; Vleeshouwer, J. Rainfall-runoff modelling in northern Australia: A guide to modelling strategies in the tropics. *J. Hydrol.* **2012**, *462–463*, 28–41. [[CrossRef](#)]
57. Huang, C.; Newman, A.J.; Clark, M.P.; Wood, A.W.; Zheng, X. Evaluation of snow data assimilation using the ensemble Kalman filter for seasonal streamflow prediction in the western United States. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 635–650. [[CrossRef](#)]
58. Burnash, R.J.; Ferral, R.L.; McGuire, R.A. *A Generalized Streamflow Simulation System, Conceptual Modeling for Digital Computers*; U. S. Dept. of Commerce, National Weather Service: Sacramento, CA, USA, 1973.
59. Andrews, F.T.; Croke, B.F.W.; Jakeman, A.J. An open software environment for hydrological model assessment and development. *Environ. Model. Softw.* **2011**, *26*, 1171–1185. [[CrossRef](#)]
60. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
61. Krause, P.; Boyle, D.P.; Se, F.B. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* **2005**, *5*, 89–97. [[CrossRef](#)]
62. Pushpalatha, R.; Perrin, C.; Moine, N.L.; Andréassian, V. A review of efficiency criteria suitable for evaluating low-flow simulations. *J. Hydrol.* **2012**, *420–421*, 171–182. [[CrossRef](#)]
63. Duan, Q.; Sorooshian, S.; Gupta, V. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **1992**, *28*, 1015–1031. [[CrossRef](#)]
64. Shin, M.; Guillaume, J.H.A.; Croke, B.F.W.; Jakeman, A.J. Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in R. *J. Hydrol.* **2013**, *503*, 135–152. [[CrossRef](#)]

65. Beven, K.; Binley, A. The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Process.* **1992**, *6*, 279–298. [[CrossRef](#)]
66. Hornberger, G.M.; Spear, R.C. An Approach to the Preliminary Analysis of Environmental Systems. *J. Environ. Manag.* **1981**, *12*, 7–18.
67. Dai, A. Precipitation Characteristics in Eighteen Coupled Climate Models. *J. Clim.* **2006**, *19*, 4605–4630. [[CrossRef](#)]
68. Chapman, T.G. Estimation of daily potential evaporation for input to rainfall-runoff models. In *MODSIM2001: Integrating Models for Natural Resources Management across Disciplines*; Modelling and Simulation Society of Australia and New Zealand Inc.: Canberra, Australia, 2001; pp. 293–298.
69. Safeeq, M.; Mauger, G.S.; Grant, G.E.; Arismendi, I.; Hamlet, A.F.; Lee, S. Comparing Large-Scale Hydrological Model Predictions with Observed Streamflow in the Pacific Northwest: Effects of Climate and Groundwater. *J. Hydrometeorol.* **2014**, *15*, 2501–2521. [[CrossRef](#)]
70. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASABE* **2007**, *50*, 885–900. [[CrossRef](#)]
71. Orth, R.; Staudinger, M.; Seneviratne, S.I.; Seibert, J.; Zappa, M. Does model performance improve with complexity? A case study with three hydrological models. *J. Hydrol.* **2015**, *523*, 147–159. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).