

Article

# Analysis of the Effect of Uncertainty in Rainfall-Runoff Models on Simulation Results Using a Simple Uncertainty-Screening Method

Mun-Ju Shin <sup>1</sup> and Chung-Soo Kim <sup>2,\*</sup>

<sup>1</sup> Water Resources Research Team, Jeju Province Development Corporation, 1717-35, Namjo-ro, Jocheon-eup, Jeju-si, Jeju-do 63345, Korea

<sup>2</sup> Department of Land, Water and Environment Research, Korea Institute of Civil Engineering and Building Technology, 283, Goyangdae-ro, Ilsanseo-gu, Goyang-si, Gyeonggi-do 10223, Korea

\* Correspondence: alska710@kict.re.kr; Tel.: +82-(0)-31-9100-274

Received: 5 April 2019; Accepted: 28 June 2019; Published: 30 June 2019



**Abstract:** Various uncertainty analysis methods have been used in various studies to analyze the uncertainty of rainfall-runoff models; however, these methods are difficult to apply immediately as they require a long learning time. In this study, we propose a simple uncertainty-screening method that allows modelers to investigate relatively easily the uncertainty of rainfall-runoff models. The 100 best parameter values of three rainfall-runoff models were extracted using the efficient sampler Differential Evolution Adaptive Metropolis (DREAM) algorithm, and the distribution of the parameter values was investigated. Additionally, the ranges of the values of a model performance evaluation statistic and indicators of hydrologic alteration corresponding to the 100 parameter values for the calibration and validation periods was analyzed. The results showed that the Sacramento model, which has the largest number of parameters, had uncertainties in parameters, and the uncertainty of one parameter influenced all other parameters. Furthermore, the uncertainty in the prediction results of the Sacramento model was larger than those of other models. The IHACRES model had uncertainty in one parameter related to the slow flow simulation. On the other hand, the GR4J model had the lowest uncertainty compared to the other two models. The uncertainty-screening method presented in this study can be easily used when the modelers select rainfall-runoff models with lower uncertainty.

**Keywords:** uncertainty analysis; rainfall-runoff model; DREAM algorithm; indicators of hydrologic alterations; equifinality

## 1. Introduction

Conceptual rainfall-runoff models are widely used to understand hydrologic systems and to predict runoff. However, there are uncertainties in the parameters (e.g., [1]) and structure (e.g., [2]) of rainfall-runoff models, and these uncertainties can lead to uncertain results. Of course, uncertainty in the input data (e.g., [3]) can also cause uncertainty in the simulation results. Even high-quality input data have measurement errors or discretization errors (e.g., an error that occurs when daily data time-series are generated using hourly data time-series). Additionally, there is uncertainty in the simulation results even when using error-free input data (e.g., [4]). If one uses high-quality input data that cannot be further improved, the way to reduce the uncertainty of the simulation results is to reduce the uncertainty in the parameters or structure of the rainfall-runoff model.

Various studies have been conducted on the uncertainty of rainfall-runoff models. For example, Shin et al. [1] analyzed the uncertainty in the parameters of four structurally different models using the Sobol and Morris sensitivity analysis methods for five Australian catchments with spatiotemporally

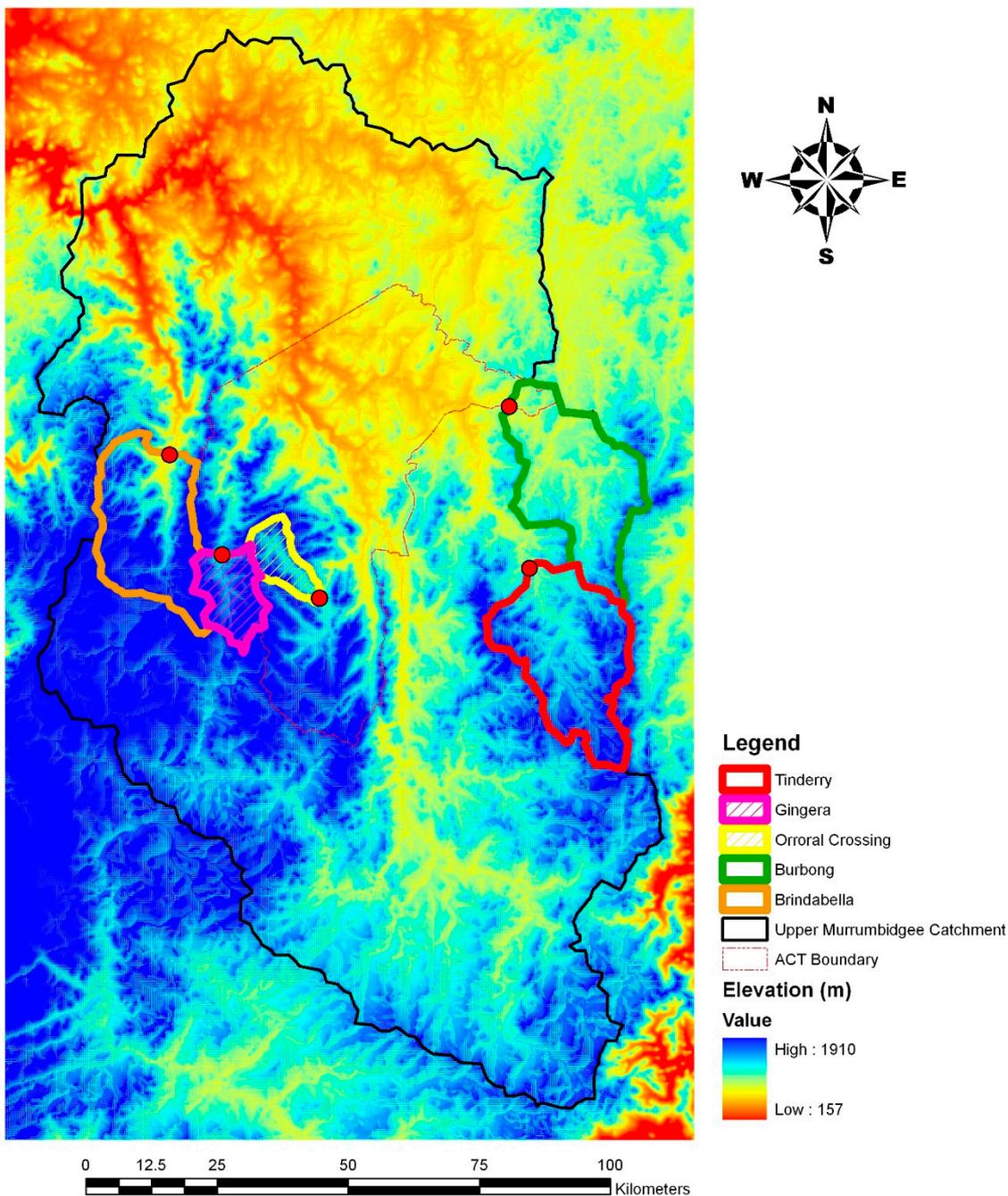
different characteristics. Van Hoey et al. [5] performed a qualitative sensitivity analysis on the model structure by analyzing the sensitivity of the simulation results by changing the model components one at a time to select the appropriate rainfall-runoff model. Furthermore, Massmann and Holzmann [6] conducted a time-varying sensitivity analysis using the Sobol method to understand the structure of rainfall-runoff models. These three studies for sensitivity analysis [1,5,6] were carried out in order to investigate the uncertainty of parameters and model structure. Beven and Freer [7] and Cho et al. [8] investigated the uncertainty of rainfall-runoff models using the Generalized Likelihood Uncertainty Estimation (GLUE) method, and Wagener et al. [9] evaluated the structure of rainfall-runoff models by identifying the model parameters using the Dynamic Identifiability Analysis (DYNIA) method. Vrugt et al. [10] investigated the improvement of the structure of rainfall-runoff models by calculating the boundary of the uncertainty of simulation results using the Shuffled Complex Evolution Metropolis Algorithm (SCEM-UA). Clark et al. [2] developed and applied the Framework for Understanding Structural Errors (FUSE) method to diagnose the differences in the structures of rainfall-runoff models and quantify the uncertainty in the structure. Shin et al. [4] used various screening methods and error-free data to investigate the uncertainty in structurally different rainfall-runoff models. However, the methods used by most of these studies on uncertainty analysis require a long learning time and are therefore difficult to apply immediately to solve the problem of model uncertainty. Therefore, it is necessary to identify an uncertainty method that modelers can immediately use to decide which rainfall-runoff model to choose and what factors to consider when choosing a model.

The aim of this study is to present a relatively easy-to-use uncertainty-screening method for modelers to select a rainfall-runoff model with less uncertainty. This study is only focused on parameter uncertainty. We show the uncertainties that may occur in rainfall-runoff models using real data and investigate the distribution of the 100 best parameter samples in order to show the uncertainty of the rainfall-runoff models. Additionally, we analyze the range of model performance evaluation statistic values and the range of hydrologic indicator values for calibration and validation periods to show how such uncertainties affect simulation results. Through the analysis of range variation, modelers can immediately determine the uncertainty of a rainfall-runoff model.

## 2. Materials

### 2.1. Catchments

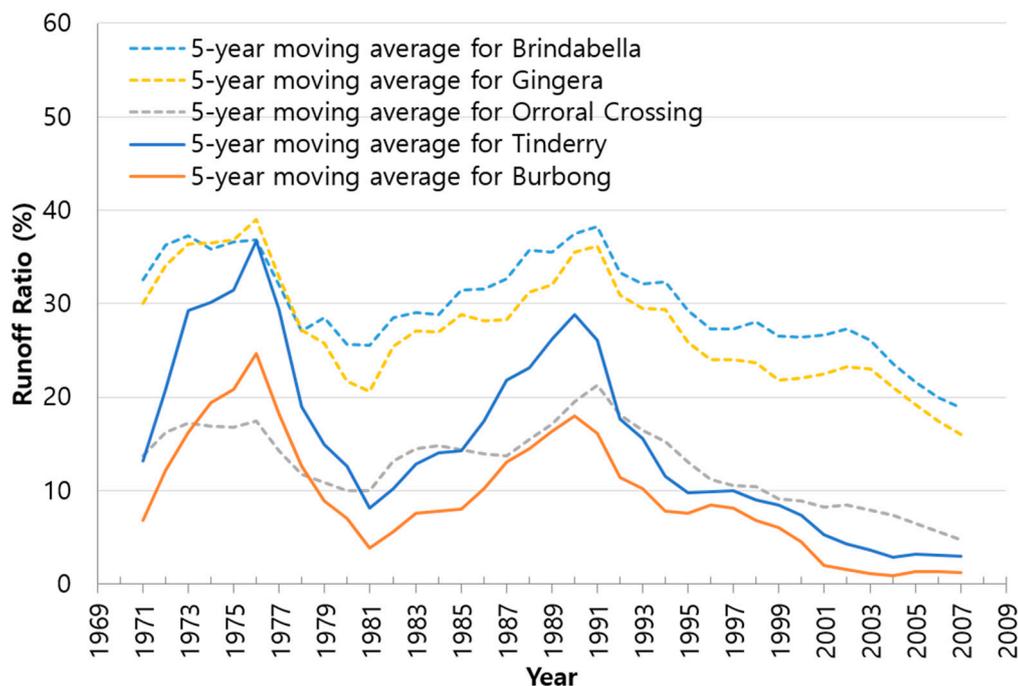
Five mountainous catchments in the Australian Capital Territory (ACT) region of Australia were selected for the uncertainty analysis (Figure 1). The chosen catchments were the following: Goodradigbee River at Brindabella (427 km<sup>2</sup>), Cotter River at Gingera (148 km<sup>2</sup>), Orroral River at Crossing (90 km<sup>2</sup>), Queanbeyan River at Tinderry (490 km<sup>2</sup>), and Molonglo River at Burbong (505 km<sup>2</sup>). The catchments have different elevation ranges and different hydroclimatological conditions ranging from wet to dry. The Brindabella and Gingera catchments are relatively wet catchments (mean annual rainfall of 1127 and 985 mm/year during 1970–2009, respectively), while the Tinderry and Burbong catchments are relatively dry catchments (mean annual rainfall of 716 and 664 mm/year during 1970–2009, respectively). The hydroclimatological characteristic of the Orroral Crossing catchment is intermediate between wet and dry (mean annual rainfall of 885 mm/year during 1970–2009). The river flows of these catchments typically have higher peaks, lower base flow, smaller runoff coefficients, and longer and more variable drier periods than those in European and North American catchments [11]. These catchment characteristics can be useful to investigate the uncertainty of rainfall-runoff models. Further information about the five catchments investigated in the present study can be found in the study of Shin et al. [1].



**Figure 1.** Locations of the five studied catchments in the Upper Murrumbidgee Catchment, Australia, which includes the Australian Capital Territory (ACT) region (red dots represent the stream-flow gauging stations).

2.2. Data

Over 40 years of data are available for these five catchments (1970–2009). These years include wet periods and very dry “Millennium drought” [12] periods. Therefore, the catchments have spatiotemporally-variable characteristics. More detail on the temporal variation in runoff response characteristics of the catchments is illustrated in Figure 2. This figure shows the changes in the annual runoff ratio for a five-year moving average for all five catchments. The annual runoff ratios of all catchments showed a similar pattern; for example, the 1970s had the largest annual runoff ratio, and the 2000s had the smallest.



**Figure 2.** Five-year moving average of the annual runoff ratio for the five studied catchments.

In this study, real daily rainfall data and potential evapotranspiration data for 40 years (1970–2009) were chosen as inputs for model simulation, while model parameters were calibrated and validated using daily observed flow data. The calibration was performed for each of four 10-year periods (the 1970s, 1980s, 1990s, and 2000s), and validation was performed for the other three 10-year periods (e.g., when the 1970s was the calibration period, the 1980s, 1990s, and 2000s were the validation periods). The year preceding each decade was used as a warm-up period. The appropriate calibration and validation periods were selected based on the studies of Anctil et al. [13], Kim et al. [14], and Yapo et al. [15]. The climate data and observed flow data were obtained from local governments (the Australian Capital Territory and New South Wales governments) and industry (ACTEW Corporation) in Australia.

### 2.3. Rainfall-Runoff Models

Three well-known and widely-used conceptual rainfall-runoff models were selected (e.g., [16–18]). These three models have different model structures. The GR4J model [19], which has four parameters, is a daily lumped rainfall-runoff model with two stores (production and routing stores) and two unit hydrographs. The process of storing rainfall, evapotranspiration, and percolation in the surface soil is controlled by the production store. In the routing store, effective rainfall is separated into 90% and 10%; the 90% of effective rainfall is routed by one unit hydrograph for slow flow generation, and the 10% of effective rainfall is routed by another unit hydrograph for quick flow generation [20]. The ranges of the parameters of the GR4J model were obtained from the work of Shin et al. [1].

The IHACRES model has various versions; this study used the Catchment Moisture Deficit (CMD) version [21], which accounts for the changes of catchment moisture at each time step. The CMD output, which is calculated by the accounting equation, and the water output, which is calculated by a nonlinear function with the raw rainfall, were combined to generate effective rainfall. Linear store uses unit hydrographs with parallel storages to convert the effective rainfall into quick and slow flows, and the quick and slow flows were summed at each time step to generate the total flows. Four parameters were calibrated by fixing two of the six parameters to a specific value. The parameter  $e$  was fixed as unity (Table 1) due to the fact that the potential evapotranspiration data were used instead of

temperature data, and the parameter  $d$  was fixed as 200 based on the study of Croke and Jakeman [21]. The definitions of the parameters  $e$  and  $d$  are described in Table 1.

The Sacramento model [22] accounts for soil moisture at various depths of interconnected soil tanks and consists of five runoff components: direct runoff from impervious area, surface runoff, interflow, supplementary base flow, and primary base flow. Excessive rainfall becomes runoff routed by a unit hydrograph, and the rest of the rainfall fills the soil moisture store. Water loss by evapotranspiration is processed in the soil moisture store, and the remaining water is separated into lateral flow and ground water. The total flow is the sum of the direct, surface, and lateral flow [23]. As in the work of Shin et al. [1], 13 of the 16 parameters of the Sacramento model were used in the present study.

Table 1 gives a description of the parameters of the three rainfall-runoff models. These models are included in the Hydrological Model Assessment and Development (Hydromad) [24], which is an R-based open-source software package (available from <http://hydromad.catchment.org>).

**Table 1.** Description of the parameters of the three rainfall-runoff models.

Parameter Name	Unit	Range	Description
<b>GR4J</b>			
$x1$	(mm)	50–5000	Maximum capacity of the production store
$x2$	(mm)	(−15)–4	Groundwater exchange coefficient
$x3$	(mm)	10–1300	One day ahead maximum capacity of the routing store
$x4$	(day)	0.5–5	Time base of Unit Hydrograph UH1
<b>IHACRES-CMD<sup>a</sup></b>			
$f$	(-)	0.5–1.3	CMD stress threshold as a proportion of $d$
$e$	(-)	1 (fixed)	Temperature to Potential Evapotranspiration (PET) conversion factor
$d$	(mm)	200 (fixed)	CMD threshold for producing flow
$\tau_s$ (tau_s)	(day)	10–1000	Time constant for slow flow store
$\tau_q$ (tau_q)	(day)	0–10	Time constant for quick flow store
$v_s$ (v_s)	(-)	0–1	Fractional volume for slow flow
<b>Sacramento</b>			
$uztwm$	(mm)	1–150	Upper zone tension water maximum capacity
$uzfwm$	(mm)	1–150	Upper zone free-water maximum capacity
$uzk$	(1/day)	0.1–0.5	Upper zone free-water lateral depletion rate
$pctim$	(-)	0.000001–0.1	Fraction of the impervious area
$adimp$	(-)	0–0.4	Fraction of the additional impervious area
$zperc$	(-)	1–250	Maximum percolation rate coefficient
$rexp$	(-)	0–5	Exponent of the percolation equation
$lztwm$	(mm)	1–500	Lower zone tension water maximum capacity
$lzfsm$	(mm)	1–1000	Lower zone supplementary free-water maximum capacity
$lzfpw$	(mm)	1–1000	Lower zone primary free-water maximum capacity
$lzsk$	(1/day)	0.01–0.25	Lower zone supplementary free-water depletion rate
$lzp$	(1/day)	0.0001–0.25	Lower zone primary free-water depletion rate
$pfree$	(-)	0–0.6	Direct percolation fraction from upper to lower zone free-water storage
$side$	(-)	0.0 (fixed)	Fraction of base flow that is draining to areas other than the observed channel
$rser$	(-)	0.3 (fixed)	Fraction of the lower zone free-water that is unavailable for transpiration purposes
$riva$	(-)	0.0 (fixed)	Fraction of the riparian vegetation area

<sup>a</sup> IHACRES-CMD is the Catchment Moisture Deficit (CMD) version of the IHACRES model.

### 3. Method

#### 3.1. Sampling of Parameters Using the Differential Evolution Adaptive Metropolis Algorithm

For the parameter calibration period of the 1980s, the parameter values of the three models were obtained using the Differential Evolution Adaptive Metropolis (DREAM) algorithm [25]. The DREAM algorithm, which is an effective Monte Carlo Markov chain sampler, was chosen to generate more samples near the optimal value. The samples generated by this algorithm approximate the posterior density function; therefore, this algorithm produces denser samples near the optimal likelihood value

and provides better resolution of the peak of dot plots compared to random or Latin hypercube sampling [4]. This algorithm simultaneously runs multiple different chains during the parameter calibration process to find the global optimum of the parameters and automatically adjusts the scale and orientation of the proposal distribution during this process. The DREAM algorithm uses a log-likelihood function as the objective function, and we used the same log-likelihood function that was applied in the work of Shin et al. [4] as follows:

$$\text{Log - likelihood} = -0.5 \times \sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2 \quad (1)$$

where  $n$  is the number of time steps,  $Q_{obs,i}$  is the observed flow at time step  $i$  (daily here),  $Q_{sim,i}$  is the simulated flow, and the sum is taken over all time steps (a 10-year period for this study).

An intensive sampling is required to produce a clear picture of the response surface. Therefore, the GR4J and IHACRES models have 100,000 function evaluations for objective function values, and the more complex Sacramento model has 1,000,000 function evaluations. Note that Cho et al. (2019) used the Isolated-Speciation-based Particle Swarm Optimization (ISPSO)-GLUE method to generate more samples in regions with high likelihood in the parameter space and investigate the variability of the model output (i.e., hydrograph) as uncertainty bounds. The purpose of this study is not to examine the uncertainty of the model structure by investigating the variability of hydrographs, but rather to evaluate the uncertainty of the parameters using equifinal samples near the optimal objective function value.

### 3.2. Investigation of Sampled Parameters Using Dotty Plots

The parameter identifiability can be investigated using dotty plots. A poorly-defined parameter has similar objective function values over the parameter range explored, and hence, the distribution of parameter values does not have a clear global maximum value. To investigate the identifiability of the parameters, the objective function values for the extracted parameter values were converted into Nash–Sutcliffe Efficiency (NSE) [26] values, which are widely used in rainfall-runoff modelling. These NSE values were then plotted in dotty plots to examine the distribution of the parameter values. The NSE is defined as follows:

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

where  $\overline{Q_{obs}}$  is the mean of the observed flow. The range of NSE is from  $-\infty$  to 1, and a value of 1 indicates that the rainfall-runoff model perfectly simulates the time series of the observed flow.

Additionally, we investigated the distribution of parameter values by adjusting the range of parameters to investigate whether the uncertainty of the model can be improved.

### 3.3. Applying the 100 Best Equifinal Calibrated Parameter Values to Validation Periods to Investigate Uncertainty in Model Structures

A total of 100 parameter sample sets with equally good NSE values for the parameter calibration periods was selected from dotty plots, and the distribution of the 100 parameter samples was subsequently investigated. The number of parameter sample sets (100) was chosen arbitrarily, and the number was sufficient for this study. The values of the 100 parameter samples can be regarded as calibrated parameter values since they have the best and equally good NSE values. Therefore, if the NSE values during the calibration period are plotted by boxplot, the range of values will be narrow.

We then simulated the rainfall-runoff models for the validation periods using the 100 parameter sample values, calculated the NSE values using the simulated flow, and examined the distribution of the NSE values for the validation periods. If the range of NSE values of the boxplot for the validation periods was significantly wider than the range for the calibration period, it means that the prediction by the rainfall-runoff model was unstable. It also means that the parameter sample sets

with equifinal [27,28] results during the calibration period do not guarantee equifinal results during the validation period; this uncertainty is due to the structural uncertainty of the rainfall-runoff model. Note that the different nature of the observed data in different periods (e.g., wet and dry) can have different NSE values (e.g., an NSE value of 0.8 in the calibration period and an NSE value of 0.7 in the validation period). However, due to the use of 100 parameter sample sets of equally high quality in the same model structure, the size of the range of NSE values in the calibration and validation periods should be similar if the parameters and model structure are well identified.

#### *3.4. Investigation of Hydrologic Indicators to Provide Specific Hydrologic Information and Investigate Uncertainty in the Rainfall-Runoff Model Structure*

Indicators of Hydrologic Alterations (IHAs) [29] were calculated using simulated hydrographs to provide modelers with specific and quantitative results. Six IHAs of Minimum (Min) and Maximum (Max) flow for one, seven, and 90 days were calculated. Here, the 1-day Max (Min), 7-day Max (Min), and 90-day Max (Min) were the annual maxima (minima) of 1-, 7-, and 90-day means, respectively. The 1-day Max represents the peak flow, and the 7-day Max represents the maximum volume of the most severe event with a seven-day duration. The value of the 90-day Max (three month flow volume) can be used for the reservoir operation. The 1-day Min, 7-day Min, and 90-day Min show the flow in drought conditions. For faster calculations and more efficient analysis of large amounts of results, the IHA package in R (<http://rpackages.ianhowson.com/rforge/IHA/>) was used, instead of the IHA software from the Nature Conservancy (<http://www.nature.org/>). IHAs have been used in many studies to investigate temporal variations in streamflow (e.g., [30–33]).

We first calculated the hydrograph for 40 years from 1970–2009 using one of the 100 best parameter sample sets. The hydrograph was then used to calculate annual IHAs, and the annual IHAs were averaged over a decade. Therefore, calibration and validation periods of 10 years had one IHA, and four IHAs from 1970–2009 were generated using one parameter set. We repeated this procedure for the 100 best parameter sample sets to calculate 100 IHAs per decade. Then, the 100 IHAs were plotted in a boxplot, and the distribution of IHAs over the calibration and validation periods of the three models was examined. Similar to the description of the distribution of NSE values in Section 3.3, a wider distribution of IHAs for the validation period implies that the model has greater predictive uncertainty. Additionally, the pros and cons of each rainfall-runoff model can be determined by examining the distribution of IHAs, and a suitable rainfall-runoff model can be selected through comparative analysis of the models. A flowchart showing the method used in this study is shown in Figure 3.

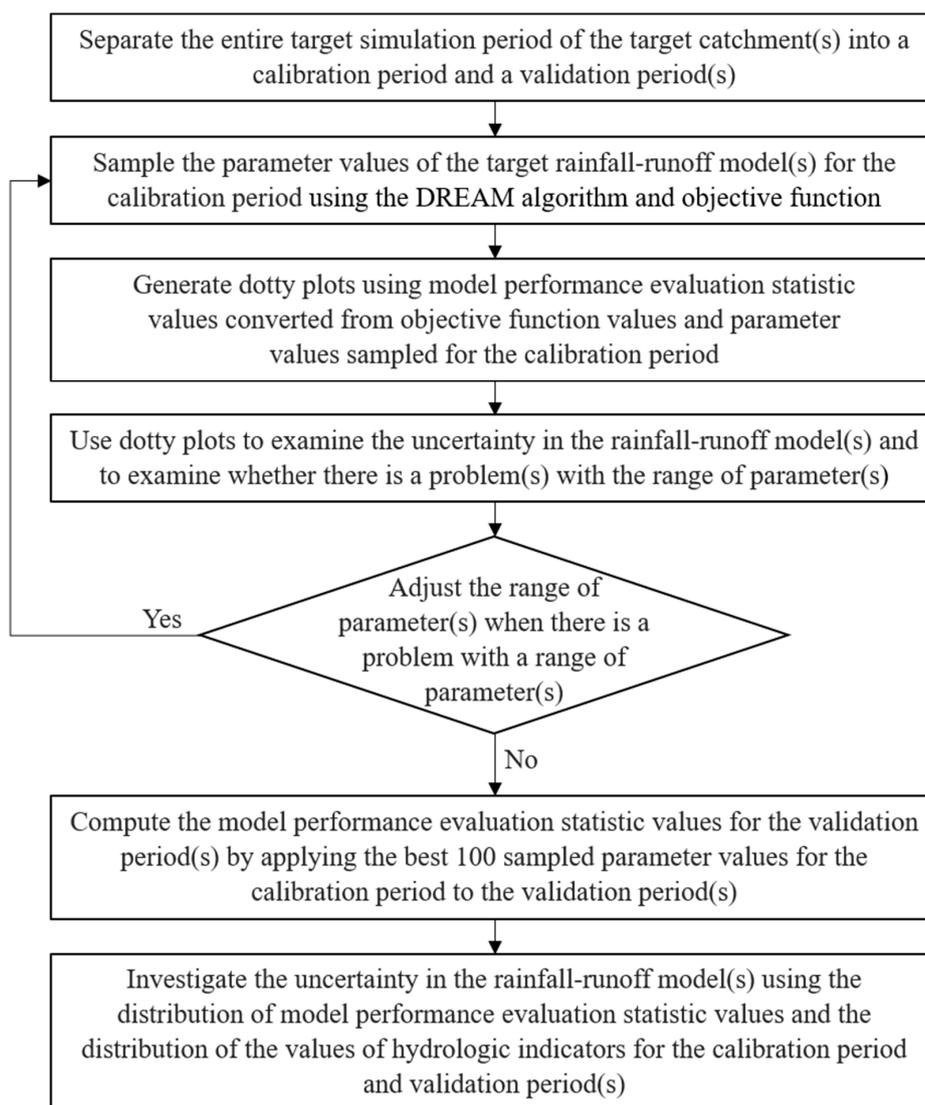


Figure 3. Flow diagram for the method.

## 4. Results and Discussion

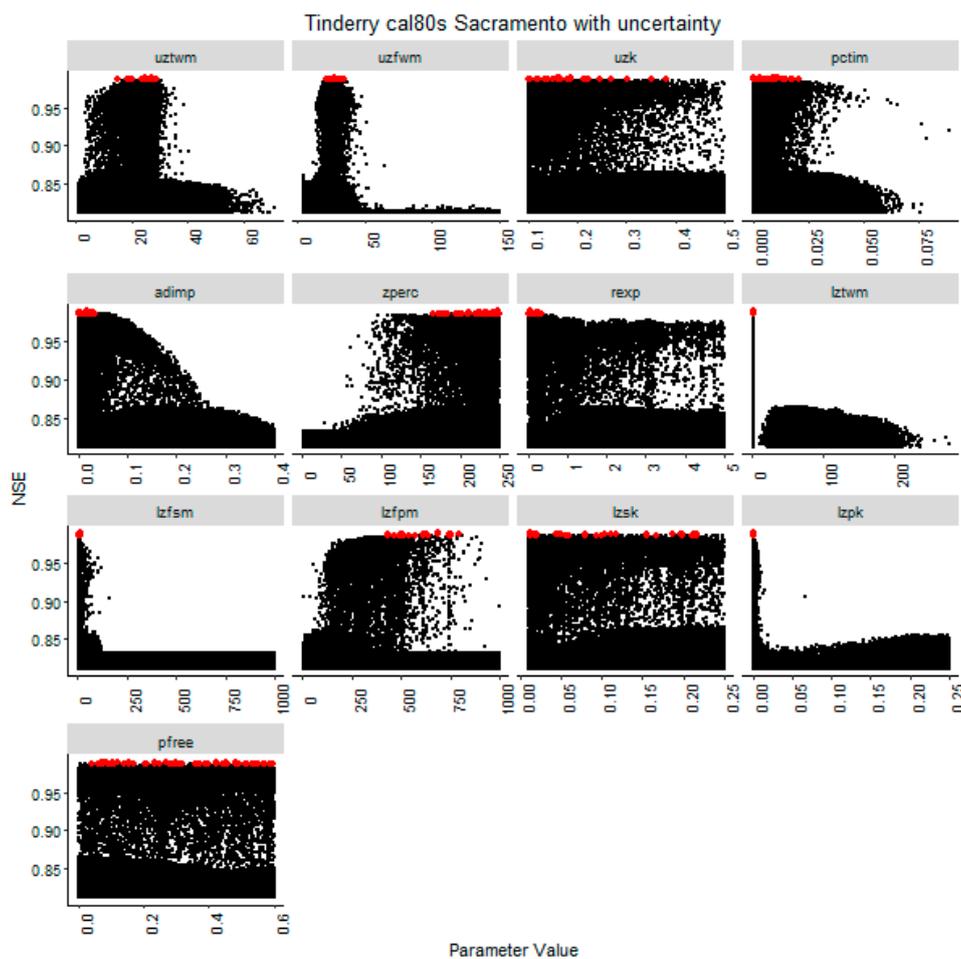
### 4.1. Investigation of Uncertainty by Investigating the Distribution of Extracted Parameter Values

Figure 4 is a dot plot that shows the parameter values of the Sacramento model calibrated using the DREAM algorithm for the actual 1980s data of the Tinderry catchment. In Figure 4, the parameter values of the Sacramento model are composed of two layers. The 100 best NSE values (red dots) in the upper layer are close to one and are not achievable when using actual data since actual data have errors, including measurement error. An investigation of 100 simulated hydrographs corresponding to these 100 NSE values showed that the Sacramento model simulated the runoff almost perfectly for a certain period (1 January 1980–27 January 1984); however, it did not simulate runoff for the rest of the period by returning N/A (Not Available). The NSE value was close to one as it was calculated only for the simulated runoff not including N/A. For the rest of the period, the Sacramento model had structural problems since it could not simulate runoff.

In Figure 4, the values of the *lztwm* parameter consisted of two distinct distributions, namely left and right. The range of the left distribution was very narrow (0–20); however, the 100 best values, whose values were close to perfect, fell within this range. Since the left distribution was abnormal with nearly perfect NSE values, there could be uncertainty in this parameter or model structure, and

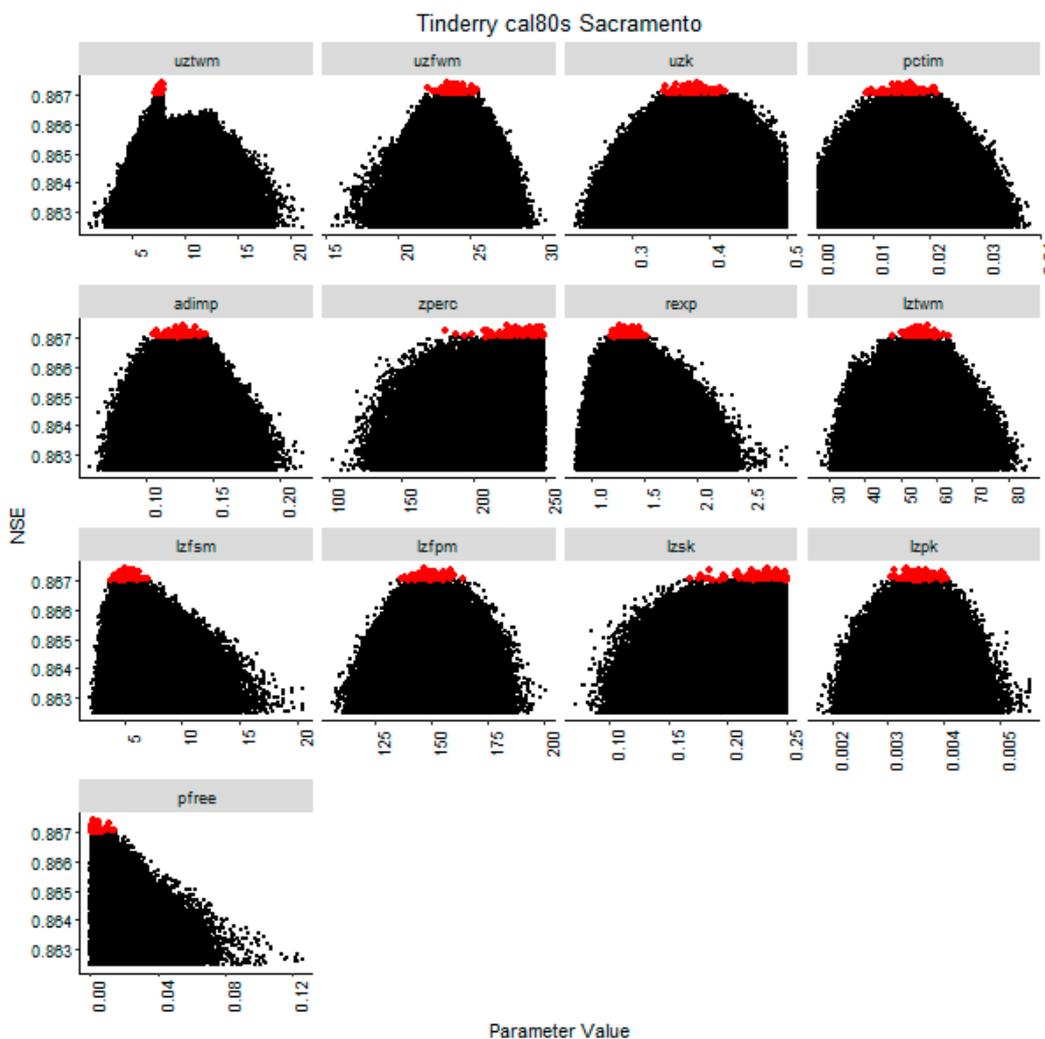
therefore, the range of the parameter needs to be investigated. We removed the range of the parameter corresponding to the left distribution and re-calibrated the values of the parameter using the modified *lztwm* parameter range and the DREAM algorithm (Figure 5). In Figure 5, the upper layer shown in Figure 4 disappears, and the distribution of the lower layer in Figure 4 becomes clearer. This means that there was uncertainty in the parameter range of the particular parameter of the Sacramento model (the *lztwm* parameter in this study). This phenomenon did not occur for the other three calibration periods for this catchment when the Sacramento model was used. Furthermore, this phenomenon did not occur for the four calibration periods for the other four catchments when the Sacramento model was applied. This method was additionally applied to the GR4J and IHACRES models; however, no such uncertainty was found. Therefore, a structural error in the Sacramento model may have caused this problem.

As shown in Figure 4, the uncertainty of the *lztwm* parameter influenced other parameters. For example, the narrow range (0–20) of the *lztwm* parameter affected the narrow range of the *lzfsm* and *lzp*k parameters. The narrow range of the *lztwm* parameter affected a slightly wider parameter range for the *uztwm* parameter (the whole range of 1–150; see Table 1) and the *uzfwm* parameter, an intermediate parameter range for the *adimp* and *zperc* parameters, and a full range for the remaining parameters. This means that each parameter of the Sacramento model interacts with the *lztwm* parameter, and the size of the interaction is different for each parameter. Through this interaction, the uncertainty of one parameter influenced all other parameters, resulting in uncertainty in the simulation results. Note that, in this study, the qualitative correlation between the parameters was analyzed. The quantitative correlation between the parameters will be analyzed in the future.



**Figure 4.** Distribution of the uncertain parameter values of the Sacramento model for the Tinderry catchment. The NSE on the vertical axis means Nash–Sutcliffe Efficiency.

As described above, the Sacramento model has problems regarding the range of parameters and the structure of the model. Therefore, as suggested by Shin et al. [4], in the Sacramento model, it is necessary to eliminate unclear parameters or merge duplicate parameters to reduce the uncertainty of the model. Note that, for the analysis in the following sections, we used the 100 best models in Figure 5 instead of the 100 best models in Figure 4.



**Figure 5.** Distribution of the parameter values of the Sacramento model for the Tinderry catchment after adjusting the range of the *lztwm* parameter.

#### 4.2. Investigation of the Distribution of Model Performance Values for Calibration and Validation Periods

Figure 6 shows the parameter values (black dots) of the three models calibrated by the DREAM algorithm using the 1980s data of the Gingera catchment. The 100 parameter values (red dots) had the best NSE values among the calibrated parameter values. In Figure 6, the 100 best parameter values of the GR4J and IHACRES models were distributed near the global maximum, while those of the Sacramento model were distributed over a wide range of parameters. The 13-parameter model (Sacramento) had a lower parameter identifiability when compared to the four-parameter models (IHACRES and GR4J), which may be due to the greater structural complexity of the Sacramento model. Note that we used the same method to investigate the distribution of the 100 best parameter values for the remaining three calibration periods, and obtained similar results (not shown here). The result of the distribution of parameter values was similar to that obtained in a previous study using error-free

data [4]. This means that the parameters of the GR4J and IHACRES models were well identified, while the parameters of the Sacramento model were not.

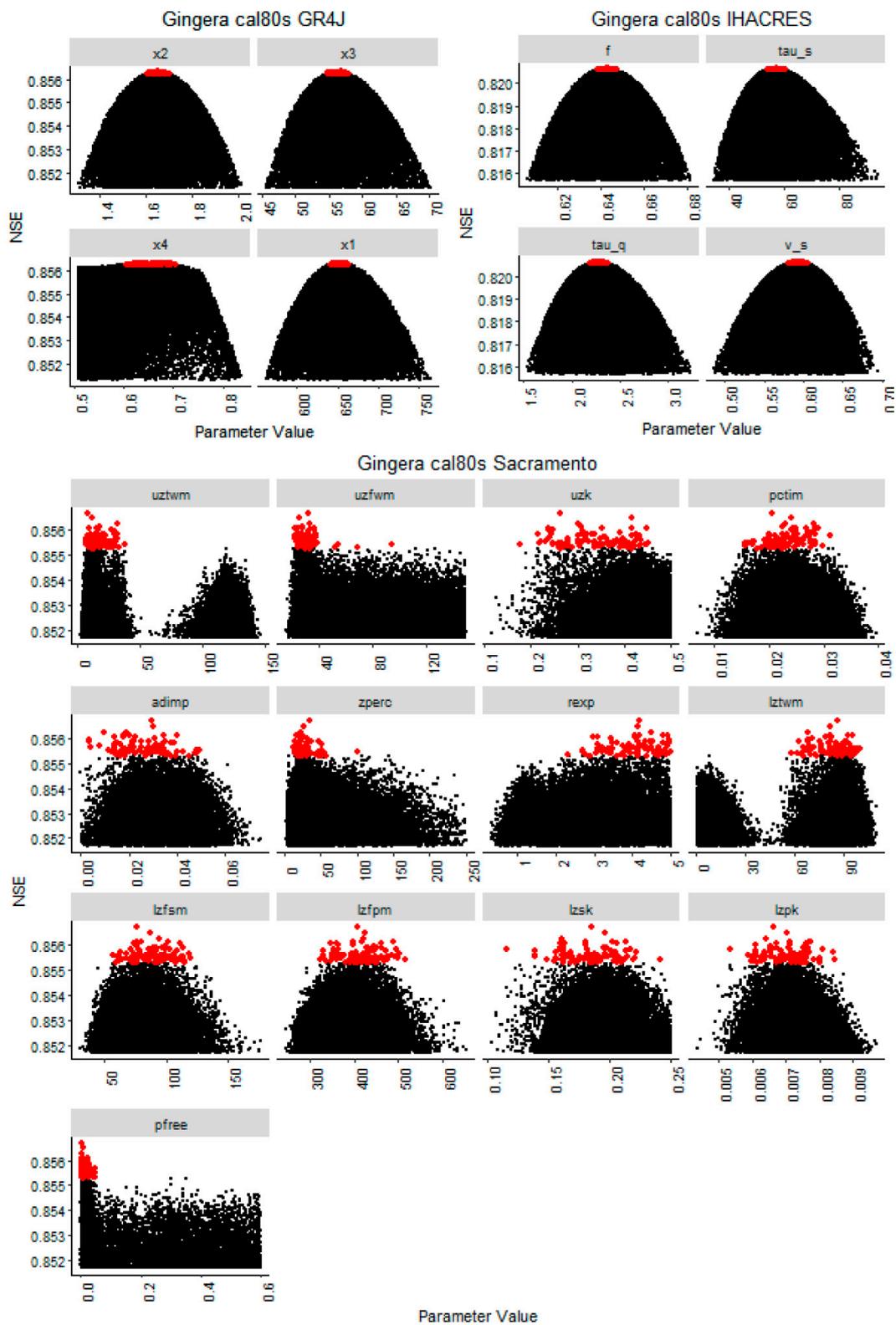


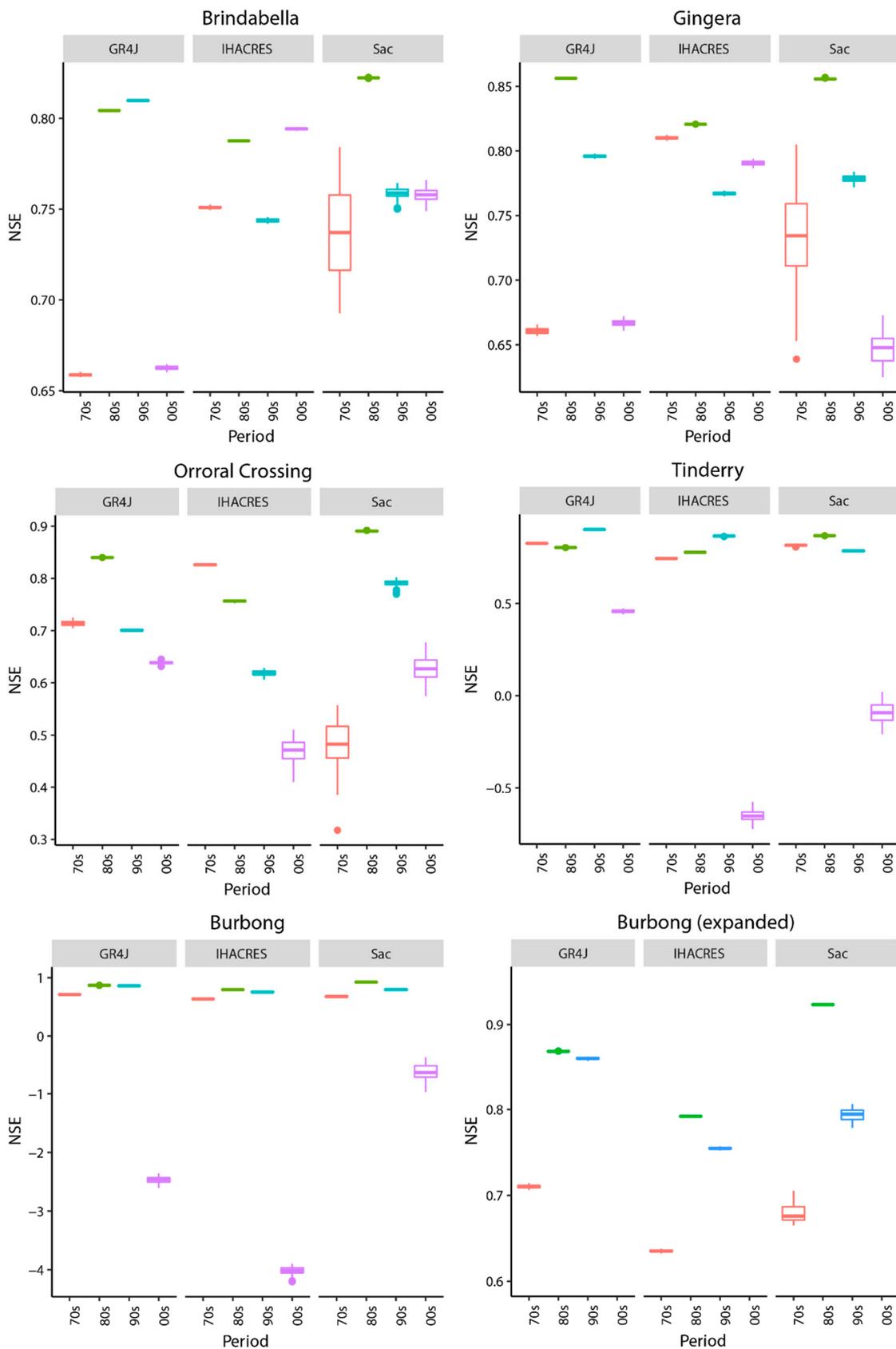
Figure 6. Distribution of the parameter values of the three models for the Gingera catchment.

Figure 7 shows boxplots of the 100 best NSE values for the calibration period (1980s) and validation periods (1970s, 1990s, 2000s) for the five catchments. From this figure, it can be seen that the range of NSE values for the GR4J and IHACRES models for the calibration and validation periods of wet catchments (Brindabella and Gingera) was very narrow. The very narrow range of the boxplots for the validation period means that these two models can predict the runoff for different periods with low uncertainty. However, the Sacramento model had a very broad range of NSE values for the 1970s, and therefore, there was great uncertainty in the predictions for the 1970s. Additionally, for the Gingera catchment, the Sacramento model had a relatively wide range of NSE values for the 2000s. This means that, for the calibration period, the 100 different parameter sets had equally good NSE values, i.e., equifinality [27,28]; however, this equifinality is not guaranteed for the validation period.

In the case of the intermediate (i.e., between wet and dry) catchment (Orroral Crossing), the GR4J model had a relatively narrow range of NSE values for all periods. However, the IHACRES model had a relatively wide range of NSE values in the 2000s. The reason for this is that the parameter of the IHACRES model for low-flow simulations ( $\tau_s$  in Figure 8) was relatively insensitive [1], and therefore, the distribution of the 100 best parameter values for the  $\tau_s$  parameter for the calibration period (1980s) in Figure 8 was somewhat broad in the range of about 600–800. As a result, the 100  $\tau_s$  parameter values with that wide distribution produced different amounts of runoff over the 2000s period, including the Millennium drought. The Sacramento model had the widest range of NSE values during the validation periods (especially for the 1970s).

For the 2000s period in the Tinderry catchment, the IHACRES and Sacramento models had a relatively wide range of NSE values, which were negative. For the 2000s period in the Burbong catchment, all three models had negative NSE values. This means that it was very difficult to predict the runoff in the dry period with the Millennium drought period using the value of the calibrated parameter for the wet period for the dry catchments. The Sacramento model, which has more parameters than the other two models, had the widest range of NSE values, and therefore had the greatest uncertainty in predicted runoff.

A comprehensive comparison of the results of the three hydrological models for the five catchments with wet and dry characteristics showed that the GR4J model (four parameters) had the lowest uncertainty in predicted runoff, and the Sacramento model (13 parameters) had the greatest uncertainty in predicted runoff. The lower parameter identifiability and greater structural uncertainty of the Sacramento model (described above) can cause greater uncertainty in runoff predictions (i.e., a broader range of NSE values).



**Figure 7.** The distribution of NSE values for calibration and validation periods for three models (GR4J, IHACRES, and Sacramento (Sac)) and five catchments.

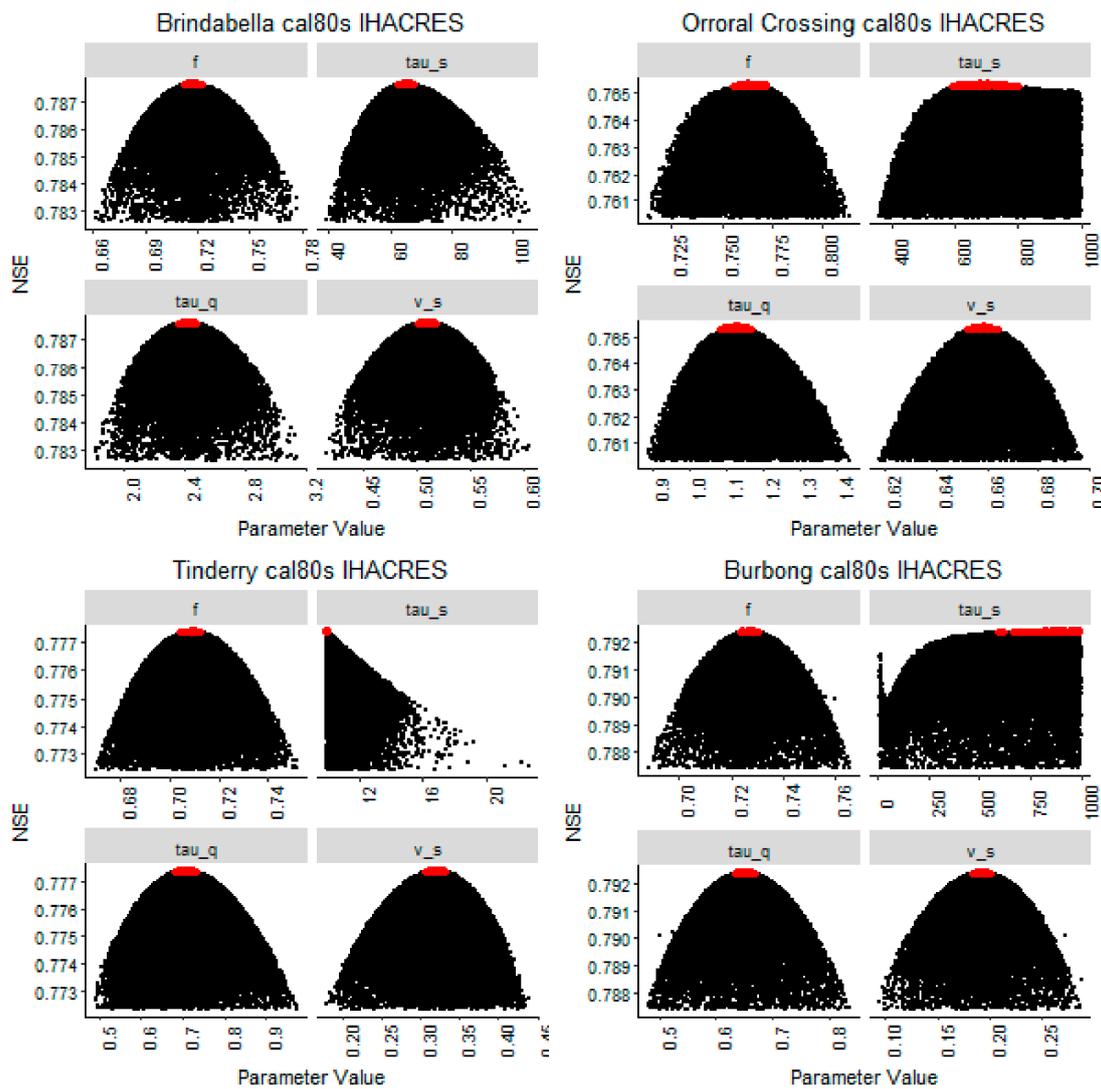


Figure 8. The distribution of the parameter values of the IHACRES model for four catchments.

The above analysis was repeatedly conducted by changing the parameter calibration period (Tables 2 and 3). Table 2 shows the difference between the maximum and minimum values of the NSE values for each period. Table 3 shows the maximum NSE value for each period. In Tables 2 and 3, the “Calibration” column represents the calibration period, and the “Period” column underneath the three models represents the simulation period. As a further explanation of the calibration and validation period, the value of 0.75 in the upper left corner of Table 3 is the maximum NSE value of the GR4J model using the 1970s period as the calibration period for the Brindabella catchment, since both the simulation period and the calibration period were the 1970s. In the same way, the value of 0.66 below the value of 0.75 is the maximum NSE value of the 1970s, which was the validation period for the 1980s calibration period, since the calibration period was the 1980s and the simulation period was the 1970s. The underlined values in Table 2 represent values that are equal to or greater than 0.06, which is an arbitrarily-selected threshold value to define a large value of the difference in NSE. The underlined values in Table 3 represent values less than zero.

**Table 2.** The differences between the maximum and minimum values of Nash–Sutcliffe Efficiency (NSE) for each period <sup>a</sup>.

Catchment	Calibration	GR4J				IHACRES				Sacramento			
		Diff_70s	Diff_80s	Diff_90s	Diff_00s	Diff_70s	Diff_80s	Diff_90s	Diff_00s	Diff_70s	Diff_80s	Diff_90s	Diff_00s
Brindabella	1970s	0.00003	0.00420	0.00627	0.00386	0.00002	0.00356	0.00707	0.00518	0.00023	0.01095	0.01716	0.01196
Brindabella	1980s	0.00286	0.00002	0.00134	0.00434	0.00299	0.00004	0.00352	0.00211	<u>0.09150</u>	0.00057	0.01414	0.01708
Brindabella	1990s	0.01189	0.00933	0.00033	0.01793	0.00481	0.00327	0.00005	0.00776	<u>0.06583</u>	0.05568	0.00272	0.03399
Brindabella	2000s	0.00723	0.00987	0.00954	0.00007	0.00660	0.00299	0.00392	0.00012	<u>0.26415</u>	<u>0.29611</u>	<u>0.09610</u>	0.01273
Gingera	1970s	0.00004	0.00560	0.00716	0.00731	0.00004	0.00493	0.01076	0.01675	0.00037	0.01295	0.02213	0.02654
Gingera	1980s	0.00885	0.00009	0.00394	0.01104	0.00449	0.00008	0.00443	0.00725	<u>0.16607</u>	0.00153	0.01199	0.04793
Gingera	1990s	0.00651	0.00257	0.00007	0.00918	0.00674	0.00373	0.00010	0.00334	<u>0.22802</u>	<u>0.07705</u>	0.00137	<u>0.06841</u>
Gingera	2000s	0.02071	0.01922	0.02102	0.00014	0.01231	0.00624	0.00287	0.00031	<u>0.84590</u>	<u>0.92870</u>	<u>0.54380</u>	0.02767
Orroral Crossing	1970s	0.00026	0.01215	0.01668	0.02329	0.00013	0.00714	0.01960	<u>0.06613</u>	0.00254	0.03060	<u>0.33796</u>	<u>0.18626</u>
Orroral Crossing	1980s	0.02002	0.00015	0.00470	0.01395	0.00529	0.00667	0.02197	<u>0.09985</u>	<u>0.23913</u>	0.00191	0.03166	<u>0.10289</u>
Orroral Crossing	1990s	0.00762	0.02289	0.00131	0.01618	0.00893	0.00672	0.00235	0.00616	<u>0.57379</u>	<u>0.08821</u>	0.00287	<u>0.11608</u>
Orroral Crossing	2000s	<u>0.11603</u>	0.05491	0.04902	0.00436	0.02505	0.01708	0.01592	0.02056	<u>2.09809</u>	<u>1.01582</u>	<u>1.64101</u>	0.05611
Tinderry	1970s	0.00001	0.00510	0.00853	<u>0.11853</u>	0.00001	0.00402	0.00564	<u>0.10515</u>	0.00012	0.01834	0.02443	<u>0.43262</u>
Tinderry	1980s	0.00222	0.00003	0.00375	0.03209	0.00523	0.00007	0.00453	<u>0.14629</u>	0.01693	0.00040	0.01646	<u>0.22845</u>
Tinderry	1990s	0.00364	0.00173	0.00008	0.03810	0.00510	0.00426	0.00008	<u>0.08202</u>	0.02236	0.01936	0.00055	<u>0.16164</u>
Tinderry	2000s	0.04085	0.05561	0.05799	0.00176	0.03814	0.03267	0.03757	<u>0.00397</u>	<u>0.21649</u>	<u>0.21524</u>	<u>0.17292</u>	<u>0.09934</u>
Burbong	1970s	0.00001	0.01606	0.01840	<u>0.51824</u>	0.00001	0.00866	0.01094	<u>0.42153</u>	0.00027	0.05746	0.04227	<u>1.08569</u>
Burbong	1980s	0.00810	0.00010	0.00495	<u>0.24704</u>	0.00557	0.00024	0.00498	<u>0.30410</u>	0.04045	0.00080	0.02779	<u>0.60965</u>
Burbong	1990s	0.00498	0.00279	0.00005	<u>0.17522</u>	0.00371	0.00124	0.00004	<u>0.16237</u>	0.03609	<u>0.15757</u>	0.02220	<u>0.60478</u>
Burbong	2000s	0.03040	0.05012	0.04277	0.00493	0.03134	0.04194	0.03552	0.03896	<u>0.32917</u>	<u>0.24963</u>	0.05623	0.01174

<sup>a</sup> Underlined values are equal to or greater than 0.06, which is an arbitrarily-selected threshold value.

Table 3. Maximum NSE values for each period <sup>a</sup>.

Catchment	Calibration	GR4J				IHACRES				Sacramento			
		Max_70s	Max_80s	Max_90s	Max_00s	Max_70s	Max_80s	Max_90s	Max_00s	Max_70s	Max_80s	Max_90s	Max_00s
Brindabella	1970s	0.75	0.70	0.69	0.58	0.78	0.75	0.67	0.75	0.81	0.77	0.72	0.75
Brindabella	1980s	0.66	0.80	0.81	0.66	0.75	0.79	0.75	0.80	0.78	0.82	0.76	0.77
Brindabella	1990s	0.68	0.80	0.82	0.64	0.72	0.77	0.77	0.77	0.26	0.68	0.82	0.68
Brindabella	2000s	0.57	0.67	0.67	0.78	0.74	0.78	0.75	0.80	0.29	0.53	0.70	0.79
Gingera	1970s	0.78	0.74	0.69	0.54	0.84	0.79	0.69	0.70	0.86	0.80	0.73	0.71
Gingera	1980s	0.67	0.86	0.80	0.67	0.81	0.82	0.77	0.79	0.80	0.86	0.78	0.67
Gingera	1990s	0.68	0.85	0.81	0.62	0.78	0.81	0.78	0.81	0.79	0.83	0.79	0.69
Gingera	2000s	0.40	0.69	0.57	0.79	0.79	0.81	0.78	0.82	0.43	0.81	0.78	0.81
Orroral Crossing	1970s	0.87	0.78	0.72	0.70	0.85	0.74	0.60	0.47	0.93	0.84	0.76	0.50
Orroral Crossing	1980s	0.72	0.84	0.70	0.65	0.83	0.76	0.63	0.51	0.56	0.89	0.80	0.68
Orroral Crossing	1990s	0.85	0.76	0.76	0.69	0.82	0.75	0.68	0.70	0.37	0.88	0.83	0.69
Orroral Crossing	2000s	0.66	0.75	0.64	0.81	0.78	0.73	0.67	0.71	0.70	0.77	0.75	0.79
Tinderry	1970s	0.86	0.73	0.77	<u>-1.01</u>	0.81	0.68	0.79	<u>-1.80</u>	0.86	0.68	0.73	<u>-1.39</u>
Tinderry	1980s	0.83	0.80	0.90	0.47	0.75	0.78	0.87	<u>-0.58</u>	0.82	0.87	0.79	0.02
Tinderry	1990s	0.81	0.80	0.93	0.61	0.77	0.76	0.89	0.03	0.79	0.80	0.94	0.58
Tinderry	2000s	0.73	0.72	0.87	0.78	0.52	0.60	0.67	0.76	0.83	0.77	0.84	0.74
Burbong	1970s	0.85	0.42	0.28	<u>-21.3</u>	0.75	0.54	0.50	<u>-17.3</u>	0.90	0.05	0.65	<u>-0.72</u>
Burbong	1980s	0.71	0.87	0.86	<u>-2.36</u>	0.64	0.79	0.76	<u>-3.90</u>	0.71	0.92	0.81	<u>-0.36</u>
Burbong	1990s	0.71	0.86	0.88	<u>-2.17</u>	0.68	0.77	0.79	<u>-3.51</u>	0.82	0.85	0.88	<u>-0.33</u>
Burbong	2000s	0.44	0.61	0.53	0.86	0.28	0.45	0.40	0.78	0.70	0.59	0.56	0.76

<sup>a</sup> Underlined values are less than zero.

As shown in Table 2, the GR4J model had relatively large differences in NSE values (underlined values) for five periods. Among these values, the value for the 1970s, which was the validation period for the 2000s calibration period of the Orroral Crossing catchment, was about 0.12, which is a large value. This means that there was great uncertainty in the simulation results when simulating the wet period using the parameters of the GR4J model calibrated for the Millennium drought period. All of the other relatively large differences in NSE were found when the 2000s period was used as the validation period, and the maximum values of NSE for these periods were all negative (see the underlined values in Table 3). Therefore, when using the simulation period including the Millennium drought period, uncertainties could be included in the results, and the result of the flow simulation should be analyzed carefully. Relatively large differences in NSE for the 2000s period also appeared in the IHACRES model. In the IHACRES model, relatively large differences were observed eight times, all of which were when the 2000s period was used as the validation period. Therefore, the simulation of the Millennium drought period was less accurate in the IHACRES model than the GR4J model. Additionally, the uncertainty of the IHACRES model for the validation period increased as the length of time between the validation period and the calibration period increased. As expected, the Sacramento model showed large differences in the NSE value (Table 2) in 35 out of a total of 80 periods (about 44%), and this model therefore had uncertainties in the simulation results regardless of calibration period and catchment characteristics. For the Sacramento model, the Millennium drought period affected 25 simulated periods, which is much higher than the five periods of the GR4J model and the eight periods of the IHACRES model. Therefore, the Sacramento model had a larger uncertainty regarding the Millennium drought period than the other two models.

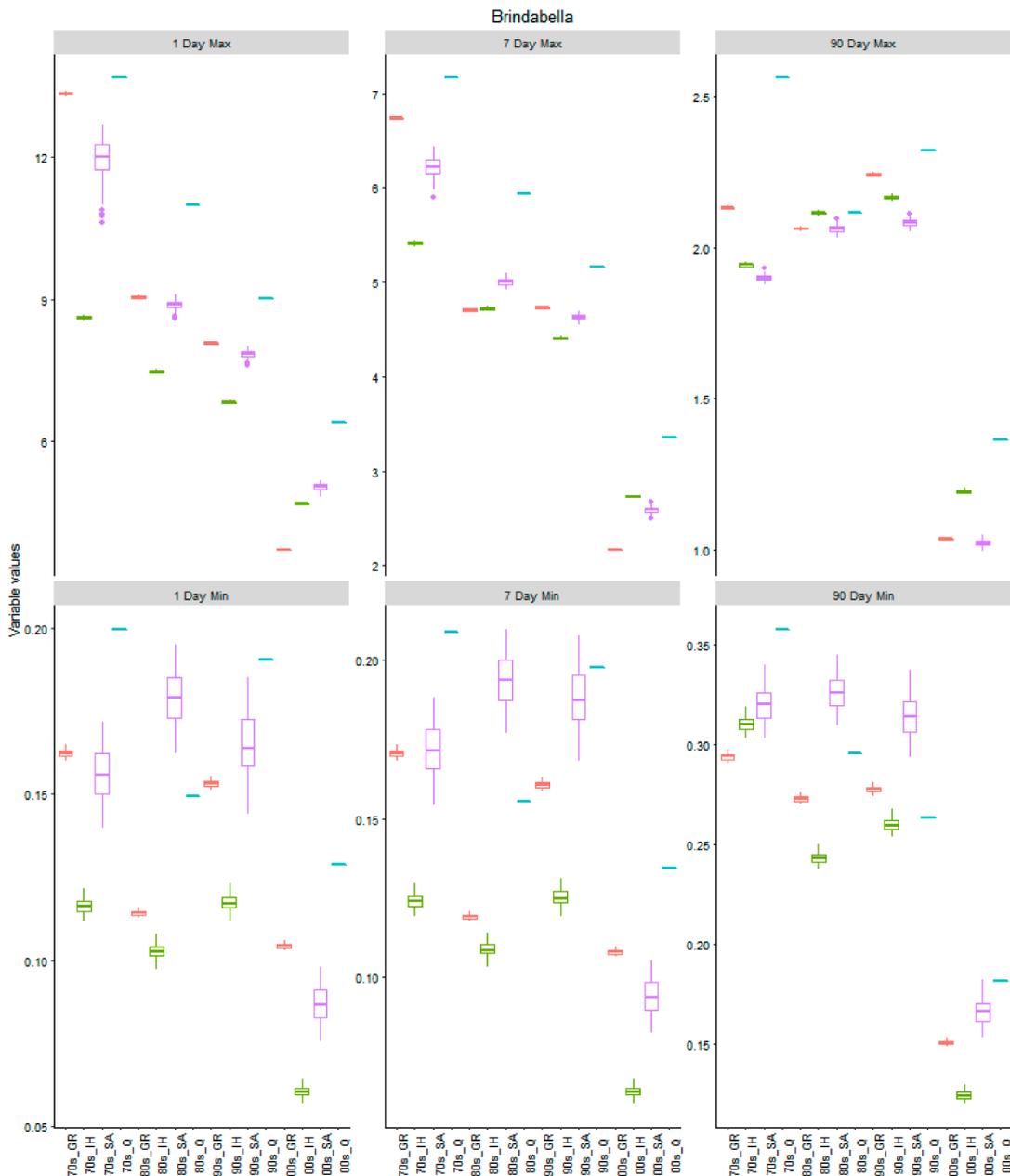
#### 4.3. Investigation of the Distribution of Hydrologic Indicators for Calibration and Validation Periods

It is not easy to determine intuitively the extent to which the difference in NSE values shown in Section 4.2 actually represents the differences in the hydrographs. Figures 9–13 are boxplots showing the difference between actual hydrographs using six IHAs. In these boxplots, the calibration period was the 1980s, and the remaining periods were the validation period.

For the wet catchments (Brindabella and Gingera), the Sacramento model (“SA” on the horizontal axis in Figures 9 and 10) had a wider range of 1-day Max, 7-day Max, and 90-day Max than the other models (Figures 9 and 10). This means that the Sacramento model had a greater uncertainty in the simulated high flow than the IHACRES (“IH” on the horizontal axis) and GR4J (“GR” on the horizontal axis) models, which have fewer parameters. For the 1-day Max, the range of values for the Sacramento model was similar for the 1990s and 2000s periods; however, the range was much wider for the 1970s period. This means that the 100 best parameter sets extracted for the 1980s period could predict the hydrographs of the 1990s and 2000s with a range of errors similar to that of the 1980s; however, the uncertainty of the peak flow prediction for the 1970s period was greater. Similar behavior was also observed for the 1-day Min and 7-day Min of the Gingera catchment. Since this behavior occurred only for the Sacramento model for wet catchments, it may be explained by the structural complexity of this model, increasing the variability of the peak flow during floods and the variability of the low flow during drought periods. For the 1-day Min, 7-day Min, and 90-day Min, the Sacramento model had a greater uncertainty than the other models.

As shown in Figures 9–13, none of the models were able to simulate accurately the observed flow values (“Q” on the horizontal axis) at all times. This means that a good performance in the calibration period does not guarantee good performance during the validation/prediction period. Usually, calibrated parameters compensate for model errors during calibration, which means that in validation/prediction periods, these model errors are no longer constrained through biased parameters. Additionally, the models that can simulate flow similar to the observed flow values were different according to the simulation period and IHA. Thus, as shown in Figure 7, even though a model has a higher NSE than the rest of the models, the model may have worse simulation results than the other models over a certain period of time.

The Sacramento model had a relatively wide range of IHAs compared to other models in intermediate (Figure 11) and dry catchments (Figures 12 and 13). The simulation results of the GR4J model had the lowest uncertainty, and the IHACRES model had a slightly broader range of 1-day Min, 7-day Min, and 90-day Min values for the Orroral Crossing and Burbong catchments. The reason why the IHACRES model had a slightly wider range is that the *tau\_s* parameter in Figure 8 has a wide range of hundreds of values, as described in Section 4.2. For the Tinderry catchment (Figure 12), the 1-day Min and 7-day Min values of the IHACRES model were close to zero. This implies that it is difficult to simulate very low flow rates of dry catchments using the IHACRES model, which supports the results of Shin et al. [4]. Note that we used the same method for the remaining three calibration periods to investigate the range of six IHAs and obtained similar results (not shown here). Therefore, this result was not sensitive to the selection of the calibration period.



**Figure 9.** Distribution of the hydrologic indicators for the Brindabella catchment. GR, GR4J; IH, IHACRES; SA, Sacramento.

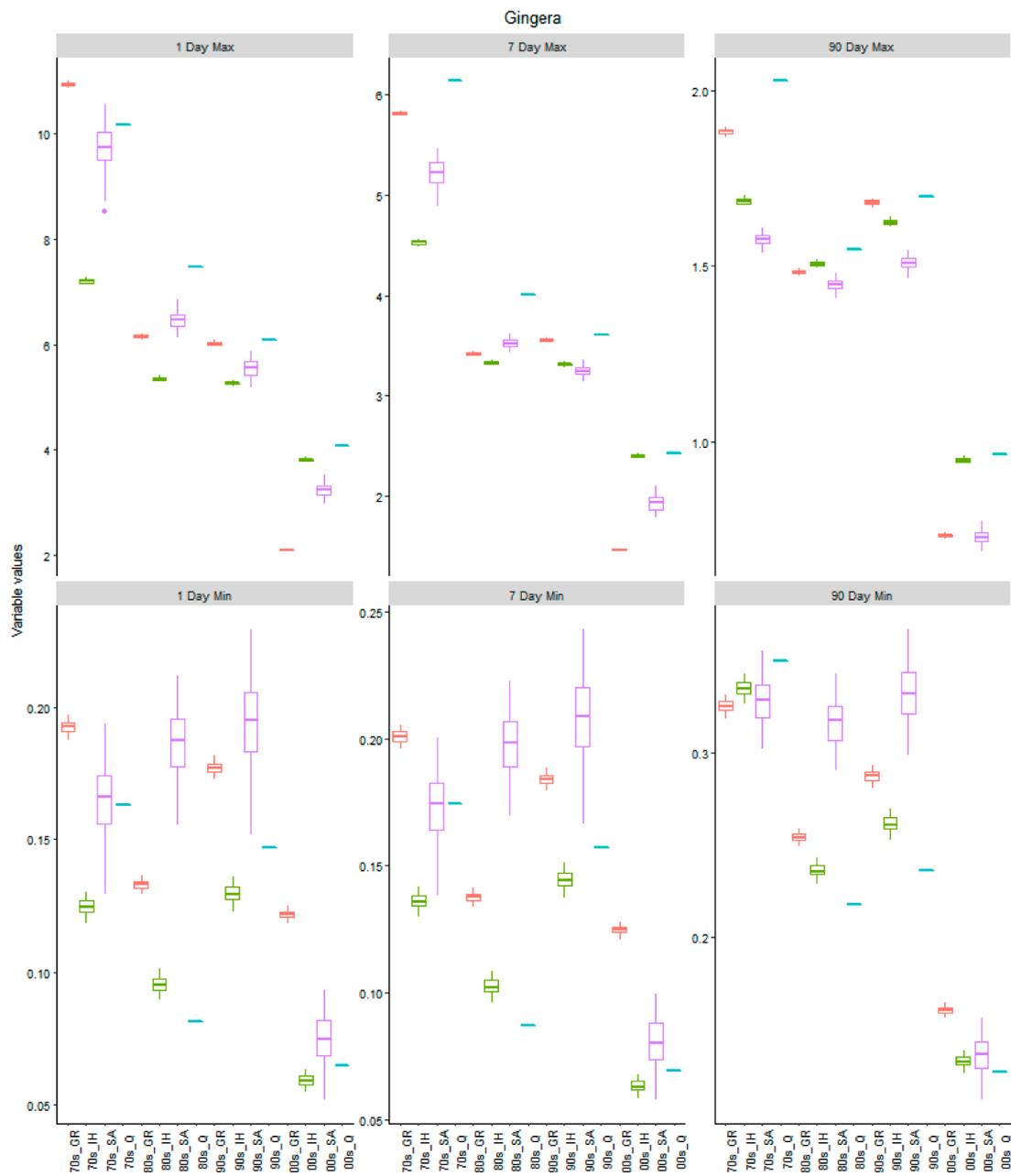


Figure 10. Distribution of the hydrologic indicators for the Gingera catchment.

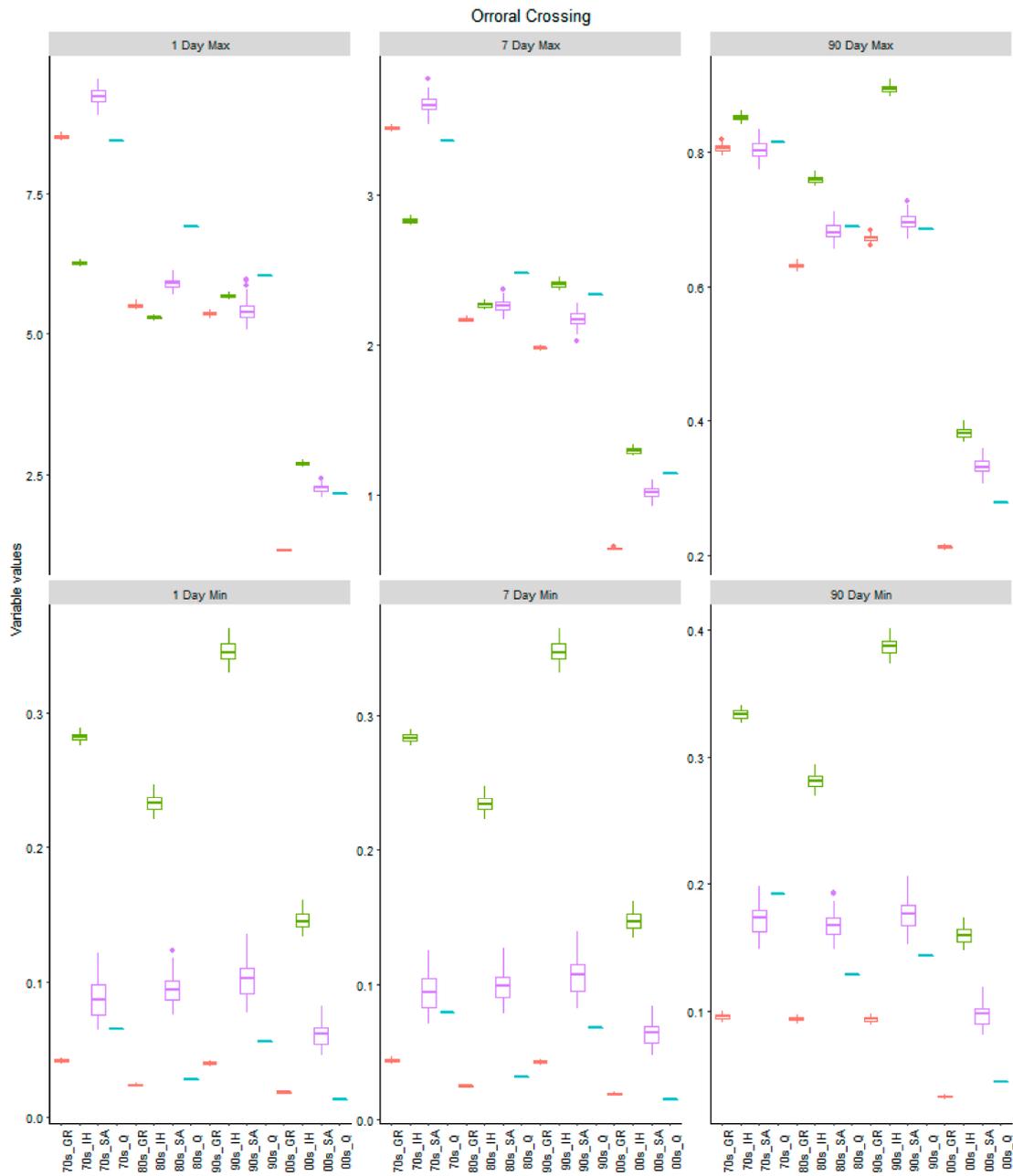


Figure 11. Distribution of the hydrologic indicators for the Orroral Crossing catchment.

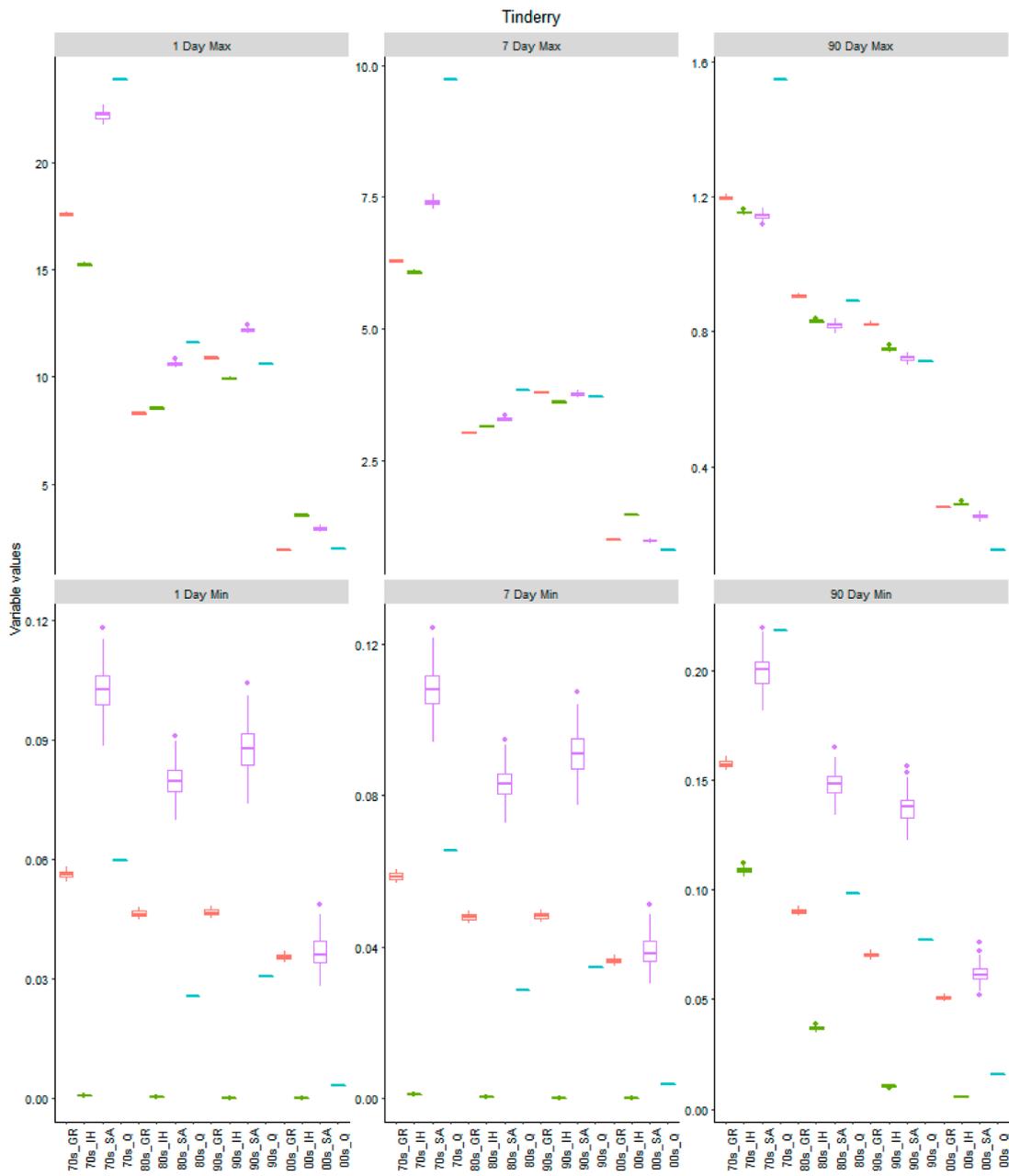


Figure 12. Distribution of the hydrologic indicators for the Tinderry catchment.

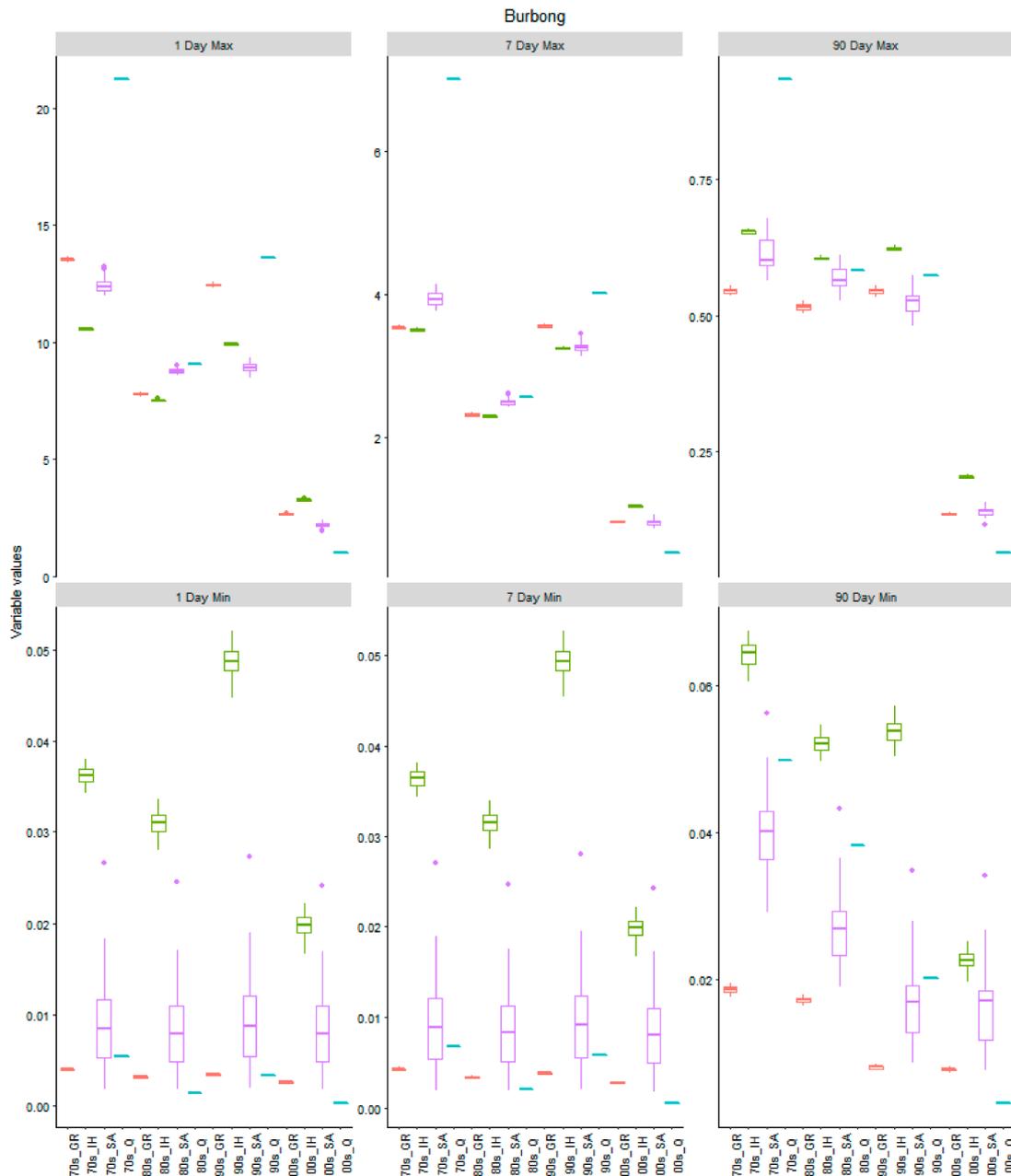


Figure 13. Distribution of the hydrologic indicators for the Burbong catchment.

### 5. Conclusions

In this study, the effects of the uncertainty of three rainfall-runoff models on simulation results were analyzed using a simple and efficient uncertainty-screening method. For the analysis of uncertainty, the distribution of the 100 best parameter values, which had equally good NSE values, was investigated, and the range of NSE and IHA values corresponding to the 100 parameter values for the calibration and validation periods was analyzed. The 100 best parameter values were extracted using the DREAM algorithm. An examination of the distribution of parameter values revealed that there was uncertainty in one parameter of the Sacramento model and that this uncertainty influenced all other parameters. Additionally, an examination of the range of NSE and IHA values showed that the simulation results of the Sacramento model, which had the largest number of parameters among the three models, had much larger uncertainty than those of the other models. The lower parameter identifiability and greater structural uncertainty of the Sacramento model can cause greater uncertainty in runoff predictions.

The IHACRES model was found to have uncertainty in the  $\tau_s$  parameter, which is related to the slow flow simulation. The simulation results of the GR4J model, which uses four parameters, were found to have the lowest uncertainty among the three rainfall-runoff models for the calibration and validation periods of the parameters. Additionally, the uncertainties in the simulation results were greatly increased when the three rainfall-runoff models used the period of the Millennium drought for the calibration or validation of the parameters. Therefore, careful attention should be paid to the analysis of the results when very different data, such as those for the Millennium drought period, are used. Furthermore, a good performance in the calibration period did not guarantee good performance in the validation period, due to the compensation of model error by potentially biased calibrated parameters. The method used in this study can be used as an uncertainty-screening method by modelers to select rainfall-runoff models with less uncertainty.

**Author Contributions:** Conceptualization, M.-J.S. and C.-S.K.; Methodology, M.-J.S.; Software, M.-J.S.; Validation, M.-J.S. and C.-S.K.; Formal Analysis, M.-J.S. and C.-S.K.; Investigation, M.-J.S. and C.-S.K.; Resources, M.-J.S.; Data Curation, M.-J.S.; Writing-Original Draft Preparation, M.-J.S.; Writing-Review and Editing, C.-S.K.; Visualization, M.-J.S.; Supervision, C.-S.K.

**Funding:** This research received no external funding.

**Acknowledgments:** We gratefully acknowledge ACTEW Corporation for supplying the rainfall data covering the five study catchments, as well as streamflow data for the Gingera and Tinderry catchments. We also gratefully acknowledge the ACT government for supplying the streamflow data for the Burbong and Ororal Crossing catchments and the NSW government for supplying the streamflow data for the Brindabella catchment.

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

## References

1. Shin, M.-J.; 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](#)]
2. Clark, M.P.; Slater, A.G.; Rupp, D.E.; Woods, R.A.; Vrugt, J.A.; Gupta, H.V.; Wagener, T.; Hay, L.E. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resour. Res.* **2008**, *44*, W00B02. [[CrossRef](#)]
3. Kavetski, D.; Kuczera, G.; Franks, S.W. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resour. Res.* **2006**, *42*, W03407. [[CrossRef](#)]
4. Shin, M.-J.; Guillaume, J.H.A.; Croke, B.F.W.; Jakeman, A.J. A review of foundational methods for checking the structural identifiability of models: Results for rainfall-runoff. *J. Hydrol.* **2015**, *520*, 1–16. [[CrossRef](#)]
5. Van Hoey, S.; Seuntjens, P.; van der Kwast, J.; Nopens, I. A qualitative model structure sensitivity analysis method to support model selection. *J. Hydrol.* **2014**, *519*, 3426–3435. [[CrossRef](#)]
6. Massmann, C.; Holzmann, H. Analysing the Sub-processes of a Conceptual Rainfall-Runoff Model Using Information About the Parameter Sensitivity and Variance. *Environ. Model. Assess.* **2015**, *20*, 41–53. [[CrossRef](#)]
7. Beven, K.; Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **2001**, *249*, 11–29. [[CrossRef](#)]
8. Cho, H.; Park, J.; Kim, D. Evaluation of Four GLUE Likelihood Measures and Behavior of Large Parameter Samples in ISPSO-GLUE for TOPMODEL. *Water* **2019**, *11*, 447. [[CrossRef](#)]
9. Wagener, T.; McIntyre, N.; Lees, M.J.; Wheater, H.S.; Gupta, H.V. Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrol. Process.* **2003**, *17*, 455–476. [[CrossRef](#)]
10. Vrugt, J.A.; Gupta, H.V.; Bouten, W.; Sorooshian, S. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.* **2003**, *39*, 1201. [[CrossRef](#)]
11. Croke, B.F.W.; Jakeman, A.J. Predictions in catchment hydrology: An Australian perspective. *Mar. Freshwater Res.* **2001**, *52*, 65–79. [[CrossRef](#)]
12. van Dijk, A.I.; Beck, H.E.; Crosbie, R.S.; Jeu, R.A.; Liu, Y.Y.; Podger, G.M.; Timbal, B.; Viney, N.R. The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resour. Res.* **2013**, *49*, 1040–1057. [[CrossRef](#)]

13. Anctil, F.; Perrin, C.; Andréassian, V. Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environ. Modell. Softw.* **2004**, *19*, 357–368. [[CrossRef](#)]
14. Kim, H.S.; Croke, B.F.W.; Jakeman, A.J.; Chiew, F.H.S. An assessment of modelling capacity to identify the impacts of climate variability on catchment hydrology. *Math. Comput. Simulat.* **2011**, *81*, 1419–1429. [[CrossRef](#)]
15. Yapo, P.O.; Gupta, H.V.; Sorooshian, S. Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data. *J. Hydrol.* **1996**, *181*, 23–48. [[CrossRef](#)]
16. Andréassian, V.; Lerat, J.; Le Moine, N.; Perrin, C. Neighbors: Nature’s own hydrological models. *J. Hydrol.* **2012**, *414*, 49–58. [[CrossRef](#)]
17. Shin, M.-J.; Eum, H.-I.; Kim, C.-S.; Jung, I.-W. Alteration of hydrologic indicators for Korean catchments under CMIP5 climate projections. *Hydrol. Process.* **2016**, *30*, 4517–4542. [[CrossRef](#)]
18. van Werkhoven, K.; Wagener, T.; Reed, P.; Tang, Y. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. *Adv. Water Resour.* **2009**, *32*, 1154–1169. [[CrossRef](#)]
19. Perrin, C.; Michel, C.; Andréassian, V. Improvement of a parsimonious model for streamflow simulation. *J. Hydrol.* **2003**, *279*, 275–289. [[CrossRef](#)]
20. Le Moine, N.; Andréassian, V.; Mathevet, T. Confronting surface- and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France). *Water Resour. Res.* **2008**, *44*, W03403. [[CrossRef](#)]
21. Croke, B.F.W.; Jakeman, A.J. A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environ. Modell. Softw.* **2004**, *19*, 1–5. [[CrossRef](#)]
22. Burnash, R.J.C.; Ferral, R.L.; McGuire, R.A. *A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers, Technical Report*; U.S. National Weather Service: Sacramento, CA, USA, 1973.
23. Podger, G (Cooperative Research Centre for Catchment Hydrology). Rainfall Runoff Library (RRL) User Guide. 2004. Available online: <https://toolkit.ewater.org.au/Tools/RRL/documentation> (accessed on 18 June 2004).
24. Andrews, F.T.; Croke, B.F.W.; Jakeman, A.J. An open software environment for hydrological model assessment and development. *Environ. Modell. Softw.* **2011**, *26*, 1171–1185. [[CrossRef](#)]
25. Vrugt, J.A.; ter Braak, C.J.F.; Diks, C.G.H.; Robinson, B.A.; Hyman, J.M.; Higdon, D. Accelerating Markov Chain Monte Carlo simulation by Differential Evolution with self-adaptive randomized subspace sampling. *Int. J. Nonlin. Sci. Num.* **2008**, *10*, 273–290. [[CrossRef](#)]
26. 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](#)]
27. Beven, K. Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Water Resour.* **1993**, *16*, 41–51. [[CrossRef](#)]
28. Beven, K. A manifesto for the equifinality thesis. *J. Hydrol.* **2006**, *320*, 18–36. [[CrossRef](#)]
29. *Indicators of Hydrologic Alteration Version 7.1: User’s Manual*; The Nature Conservancy: Arlington, VA, USA, 2009.
30. Richter, B.D.; Baumgartner, J.V.; Powell, J.; Braun, D.P. A method for assessing hydrologic alteration within ecosystems. *Conserv. Biol.* **1996**, *10*, 1163–1174. [[CrossRef](#)]
31. Sanford, S.E.; Creed, I.F.; Tague, C.L.; Beall, F.D.; Buttle, J.M. Scale-dependence of natural variability of flow regimes in a forested landscape. *Water Resour. Res.* **2007**, *43*, W08414. [[CrossRef](#)]
32. Monk, W.A.; Peters, D.L.; Baird, D.J. Assessment of ecologically relevant hydrological variables influencing a cold-region river and its delta: The Athabasca River and the Peace-Athabasca Delta, northwestern Canada. *Hydrol. Process.* **2012**, *26*, 1827–1839. [[CrossRef](#)]
33. Shrestha, R.R.; Schnorbus, M.A.; Werner, A.T.; Zwiers, F.W. Evaluating Hydroclimatic Change Signals from Statistically and Dynamically Downscaled GCMs and Hydrologic Models. *J. Hydrometeorol.* **2014**, *15*, 844–860. [[CrossRef](#)]

