

A karst runoff generation module based on the near-surface critical zone structure and threshold behaviors

Jianfei Zhao , Guofang Li, Yanan Duan , Yiming Hu, Binqun Li and Zhongmin Liang *

College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

*Corresponding author. E-mail: zmliang@hhu.edu.cn

 JZ, 0000-0002-8136-0299; YD, 0000-0003-2726-1975; ZL, 0000-0002-1079-1075

ABSTRACT

Hydrological simulation in karst areas is of great importance and challenge. It is a practical way to enhance the performance of existing hydrological models in karst areas by coupling karst modules that represent hydrological processes in these areas. The near-surface critical zone structure affects runoff generation in karst areas significantly and its complex hydrological processes could be simplified with threshold behaviors. This study proposed a three-thresholds-based karst runoff generation module (3T-KRGM), which used three reservoirs to represent water storage in the soil zone, soil–epikarst interface, and epikarst zone. The 3T-KRGM is coupled with the Xinanjiang (XAJ) model to extend the applicability of the model to karst areas. Both the improved XAJ model and the original XAJ model were used in the Shibantang watershed, which is a typical karst watershed located in southwest China. The results indicate that the performance of daily discharge simulations was obviously improved by introducing the 3T-KRGM. In addition, both the parameter sensitivity analysis and baseflow simulation demonstrate that the 3T-KRGM is rational in structure. The 3T-KRGM could also be easily coupled into other hydrological models, thus benefiting the hydrological simulation in karst areas.

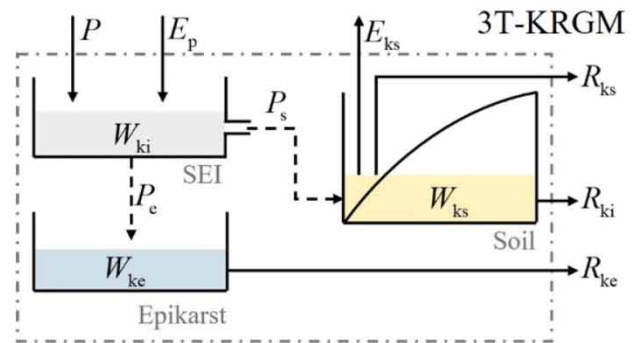
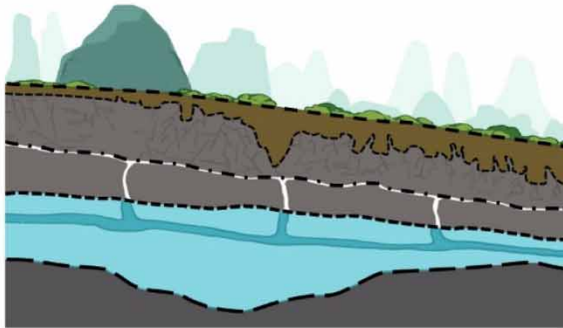
Key words: hydrological simulation, karst areas, karst runoff generation module, near-surface critical zone structure, Shibantang watershed, Xinanjiang model

HIGHLIGHTS

- We proposed a three-thresholds-based karst runoff generation module to improve model performance in karst areas.
- The theoretical basis of the module is a modification of a three-thresholds mechanism revealed by hillslope experiments.
- We coupled the module into the Xinanjiang model to extend its applicability in karst areas.
- The advantages of the module for hydrological simulation were verified from various aspects.

GRAPHICAL ABSTRACT

Module Concept

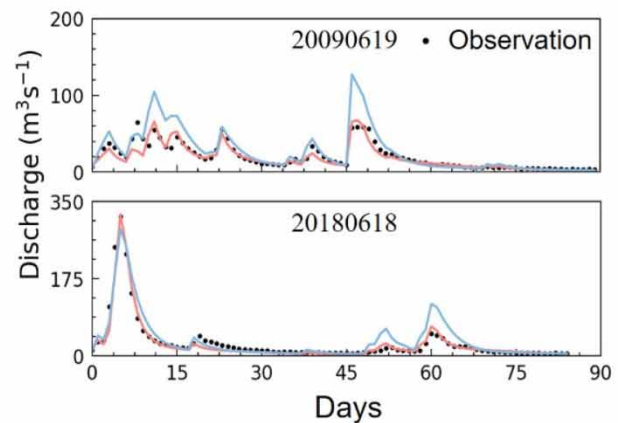
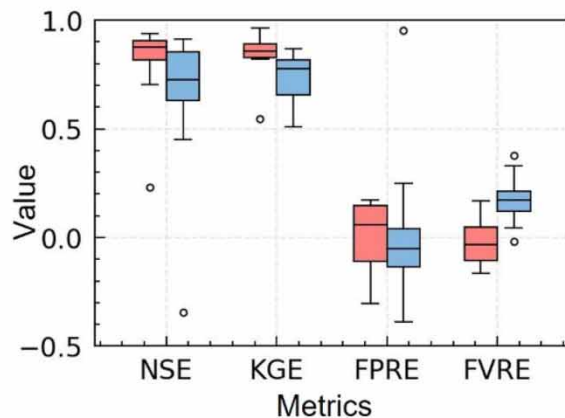


Module Coupling

Xinjiang model

Xinjiang model + 3T-KRGM

Module Validation



1. INTRODUCTION

Karst is the geological process in which water causes the dissolution of soluble rocks such as carbonate and sulfate (Hartmann *et al.* 2014; Malagò *et al.* 2016). Karst forms distinctive surface and subsurface features, including karren, sinkholes, karst conduits, and underground caves, which makes the hydrological process in karst areas highly complex (Bailly-Comte *et al.* 2009; Zhou *et al.* 2019; Chen *et al.* 2022). Globally, about 15% of the non-ice continental surface is karst areas, with 16.5% of the population living in these areas (Goldscheider *et al.* 2020). China has one of the largest continuous karst areas (located in the southwest of China) and its karst percentage is as high as 26.5% (Zhang *et al.* 2011; Goldscheider *et al.* 2020). Karst waters, which include carbonate aquifers, provide fresh water to around 25% of the population globally (Ford & Williams 2007). Hydrological simulations in karst catchments are critical not only for water resource management but also for flash flood prevention (Bonacci *et al.* 2006; Yang *et al.* 2022) and ecosystem services evaluation (Tian *et al.* 2016).

Hydrological models are essential tools for hydrological simulations. The hydrological models employed in karst areas can be split into two groups, data-driven and process-driven models, in line with other regions (Hartmann *et al.* 2014; Chen *et al.* 2022). The data-driven models, such as long short-term memory (LSTM) (Fang & Shao 2022), convert rainfall and other input into runoff directly based on transfer functions. The advantages of data-driven models are few parameters and considerable

adaptability to data situations and regions (Zhou *et al.* 2019). The lack of physical process representation is where data-driven models have long been criticized (Kuczera & Mroczkowski 1998). The process-driven models are further classified into distributed, semi-distributed, and lumped models (Paul *et al.* 2021). The distributed and semi-distributed models discretize the karst watershed into several basic units (e.g., grids or sub-basins) and assign certain governing equations, parameters, and state variables to each basic unit according to the internal karst landform (Hartmann *et al.* 2014). The advantage of distributed and semi-distributed models is that they offer a thorough insight into hydrological processes within a karst watershed while also considering temporal and spatial variability (Zhang *et al.* 2011; Xu *et al.* 2020; Li *et al.* 2021). However, the high dataset requirements and considerable computing expense restrict the application of distributed and semi-distributed models in practice (Fatichi *et al.* 2016; Simmons *et al.* 2020). The lumped models treat the whole karst watershed as a basic unit and further conceptualize its general physical processes (Hartmann *et al.* 2014; Zhou *et al.* 2019). Because of their simplicity and effectiveness, lumped models are still widely employed in practical applications (Dubois *et al.* 2020; Duran *et al.* 2020).

Lumped hydrological models use reservoirs and linkage functions between these reservoirs to provide a simplified interpretation of hydrological processes in areas of consideration (Tritz *et al.* 2011). The most widely used lumped hydrological model in China is the Xinanjiang (XAJ) model (Zhao 1992). Instead of building a lumped model for karst areas from scratch, a practical alternative is to adapt an existing lumped model by coupling modules that depict karst hydrological processes (Yang *et al.* 2022). One reason is the existence of non-karst areas in a karst-dominated watershed. The runoff generation process in karst and non-karst areas should be considered separately in a karst watershed (Zhou *et al.* 2019), which could be described by the coupled karst modules and original modules of the existing lumped model correspondingly. Another reason is that existing code and expertise in the model setup, parameter calibration, and uncertainty quantification can be reused.

The near-surface critical zone structure in karst areas denotes the soil–epikarst system and is usually characterized by high permeability and considerable water storage capacity (Perrin *et al.* 2003b). It plays an important role in runoff generation as it controls the conversion of precipitation into fast-flow and slow-flow components (Jukić & Denić-Jukić 2009). The soil and epikarst zone were treated separately in the Moisture Balance (MB) model (Jukić & Denić-Jukić 2009), and effective rainfalls were calculated based on the moisture balance of these two zones. However, the linkage between soil and epikarst zone is no more than the direct recharge to epikarst after the saturation of the soil zone in this model. Tritz *et al.* (2011) conceptualized the soil and epikarst zone as a whole and used two thresholds to represent the activation and deactivation of fast-flow paths in the soil–epikarst system. The interflow generated in the soil–epikarst system is a critical hydrological mechanism in subtropical karst areas (Fu *et al.* 2015), which could partly be attributed to the infiltration rate of the soil–epikarst interface (SEI) (Wang *et al.* 2020). The infiltration rate of the SEI functions as a threshold because it could reflect the change in the hydrological flux pattern. The complex hydrological processes in the soil–epikarst system could be simplified by identifying threshold behaviors (Wang *et al.* 2022). Thus, the near-surface critical zone structure in karst areas and related threshold behaviors should be considered in karst modules.

The objective of this study was to develop a new karst runoff generation module capable of representing hydrological processes in the soil–epikarst system explicitly. The proposed karst runoff generation module was then coupled into the XAJ model to improve model performance in karst areas. The Shibantang watershed in southwest China was selected as a study area. The parameter sensitivity of the improved XAJ model was performed to diagnose the proposed module and provide guidance for model calibration. The improved XAJ model and original XAJ model were used to simulate daily runoff in the study area, and a comparison was conducted based on the observed discharge series.

2. KARST RUNOFF GENERATION MODULE

2.1. Perceptual model

As shown in Figure 1, the typical subsurface critical zone structure of karst areas is divided into four overlaid zones: soil, epikarst, transmission, and saturation zone (Jukić & Denić-Jukić 2009; Liu *et al.* 2021). In southwest China, karst areas are usually soil-mantled despite the shallow soil depth, which is underlain by a highly weathered epikarst zone (Wang *et al.* 2020). The infiltration rate of the soil zone is usually greater than that of the SEI. Since the porosity and permeability in the epikarst zone decreased with depth, infiltration water could be detained in this zone (Williams & Andreo 2008). The

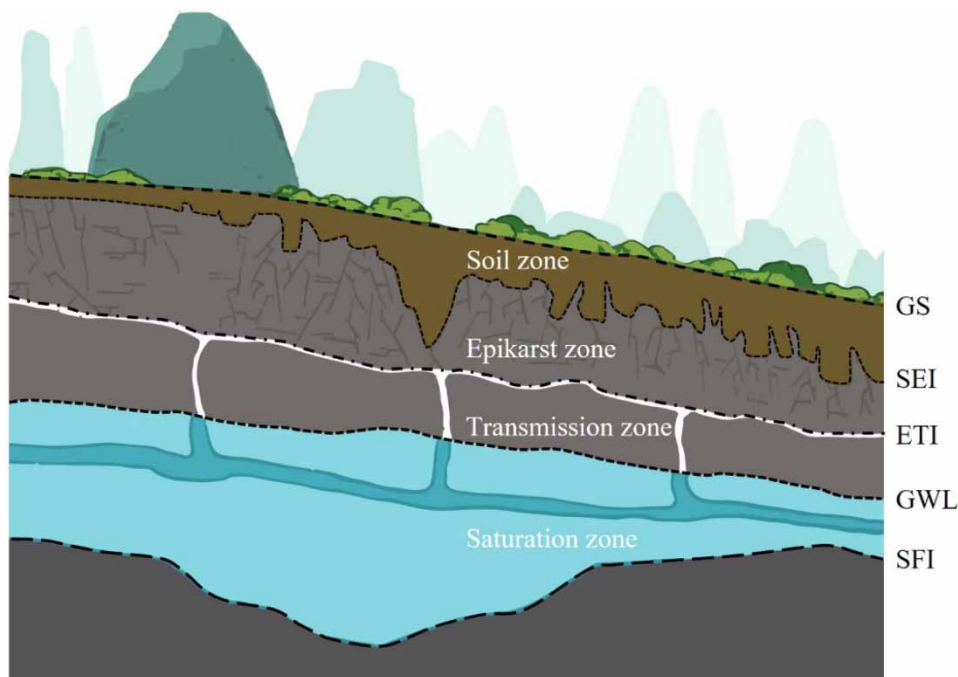


Figure 1 | The typical subsurface critical zone structure in karst areas (GS refers to the ground surface; SEI denotes soil–epikarst interface; ETI refers to the interface of epikarst and transmission zone; GWL denotes groundwater level; SFI denotes the interface of saturation zone and fresh bedrock).

transmission zone contains large fractures and karst conduits, which could transfer water from the soil–epikarst system in a fast and concentrated manner to the saturation zone (Williams & Andreo 2008; Hartmann *et al.* 2014).

The infiltrated water quickly reaches the SEI due to the shallow soil depth and the activation of the preferential pathway after the onset of precipitation (Wang *et al.* 2020). The infiltration-excess surface flow is rarely seen in karst areas as the infiltration rate exceeds the precipitation intensity in most cases. The infiltration rate of the SEI controls the amount of water that enters the epikarst. It is usually smaller than that of the soil zone, which causes water accumulation at the SEI (Fu *et al.* 2015). According to the fill-spill hypothesis (Tromp-van Meerveld & McDonnell 2006), the interflow would begin once the depressions at the SEI were filled and the saturated areas at the SEI were connected (Wang *et al.* 2020). With the rise of temporary water tables originating from the SEI, the fast and concentrated recharge to the saturation zone through karst fracture would be activated gradually (Tritz *et al.* 2011). Meanwhile, the saturated surface runoff would begin as the temporary water tables reached the ground surface (Wang *et al.* 2020, 2022). Rapid groundwater runoff would result from the fast and concentrated recharge to the saturation zone, which is characterized by rapid speed and huge quantity (Chen *et al.* 2022). The saturated surface runoff and rapid groundwater runoff could be treated together as rapid runoff. The water infiltrated into the epikarst would produce epikarst seepage runoff in the epikarst zone (Hartmann *et al.* 2014).

Three thresholds exist in the above-mentioned processes (Wang *et al.* 2020). The first threshold (T1) is the infiltration rate of the SEI, and the water would accumulate at the interface once the threshold is satisfied. The second threshold (T2) is the quantity of precipitation required to fill the depression and form connectivity at the SEI, after which the interflow runoff would be generated. The last threshold (T3) is the quantity of precipitation required to both saturate the soil zone and activate the fast and concentrated recharge path, after which it would generate rapid runoff.

2.2. Mathematical model

The proposed karst runoff generation module is based on the aforementioned perceptual model. It considers the three thresholds (T1, T2, and T3) and is thus called three-thresholds-based karst runoff generation module (3T-KRGM).

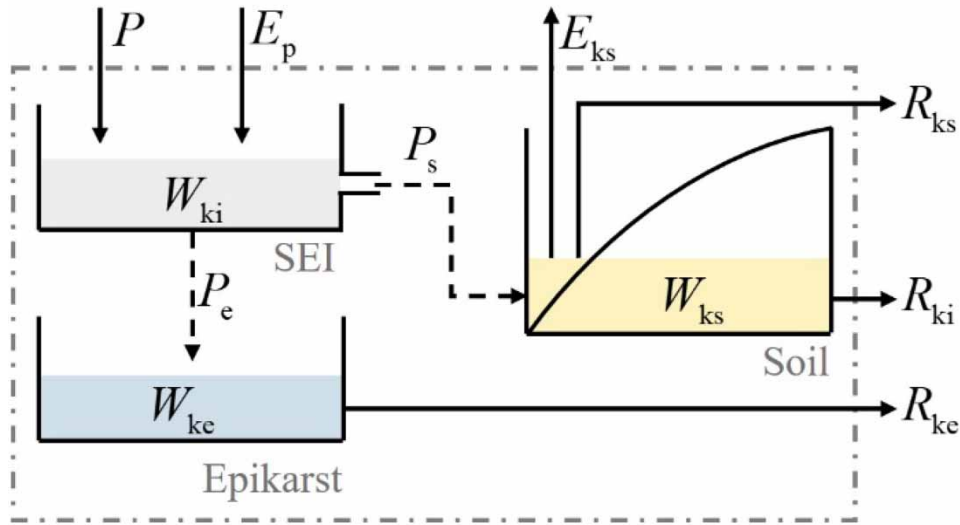


Figure 2 | The structure of the proposed three-thresholds-based karst runoff generation module (3T-KRGM). The module consists of three reservoirs: soil–epikarst interface (SEI), epikarst, and soil reservoir.

The 3T-KRGM uses three reservoirs to represent the physical storage of SEI, soil zone, and epikarst zone over the karst areas in a catchment, as shown in Figure 2.

2.2.1. SEI reservoir

The SEI reservoir is used to describe the water accumulated at the SEI. The water balance equation of the reservoir could be defined as:

$$W_{ki} = P_n + W_{ki}^0 - P_e - P_s \tag{1}$$

where W_{ki} denotes the storage of the SEI reservoir (mm), P_n is net precipitation (mm), W_{ki}^0 denotes the initial storage of the SEI reservoir (mm), P_e is the water percolated into the epikarst reservoir (mm), and P_s is the water run into the soil reservoir (mm). The portion of precipitation that exceeds evapotranspiration capacity forms net precipitation P_n . It could be expressed as:

$$P_n = \max(P - E_p, 0) \tag{2}$$

$$E_n = \max(E_p - P, 0) \tag{3}$$

where P is observed watershed-averaged precipitation (mm), E_p denotes evapotranspiration capacity (mm), and E_n denotes net evapotranspiration (mm). To minimize the number of parameters, the way to calculate E_p is set to the same as the XAJ model, which is expressed as:

$$E_p = K_e E_{obs} \tag{4}$$

where K_e is evapotranspiration coefficient (-), and E_{obs} is observed pan evaporation (mm).

Considering the spatial variation of the infiltration rate of the SEI, P_e is calculated with the GR4J type function (Perrin *et al.* 2003a) and expressed as:

$$P_e = \frac{W_{ke}^{max} [1 - (W_{ke}/W_{ke}^{max})^2] \tanh(W_{ki}/W_{ke}^{max})}{1 + W_{ke} \tanh(W_{ki}/W_{ke}^{max})/W_{ke}^{max}} \tag{5}$$

where W_{ke} denotes the storage of epikarst reservoir (mm), W_{ke}^{max} denotes the capacity of epikarst reservoir (mm), and \tanh is the hyperbolic tangent function. According to the mechanism of T1, only the part of P_n that exceeds P_e could be retained in the SEI reservoir.

According to the mechanism of T2 (fill-spill hypothesis), there is no interflow and rapid runoff generated until W_{ki} exceeds a threshold W_{ki}^t . The equation of P_s could be expressed as:

$$P_s = \max(W_{ki} - W_{ki}^t, 0) \quad (6)$$

2.2.2. Epikarst reservoir

Epikarst reservoir describes the water stored in the epikarst zone. The water balance equation of the epikarst reservoir could be expressed as:

$$W_{ke} = P_e + W_{ke}^0 - R_{ke} \quad (7)$$

where W_{ke}^0 denotes the initial value of epikarst reservoir (mm), R_{ke} is epikarst seepage runoff (mm). The equation of R_{ke} could be expressed as:

$$R_{ke} = W_{ke} \{1 - [1 + (4W_{ke}/9W_{ke}^{max})^4]^{-1/4}\} \quad (8)$$

2.2.3. Soil reservoir

The soil reservoir represents the water stored in the soil zone, which excludes the water in the SEI depression. The water balance equation of the soil reservoir could be expressed as:

$$W_{ks} = P_s + W_{ks}^0 - R_{ki} - R_{ks} - E_{ks} \quad (9)$$

where W_{ks} denotes the storage of soil reservoir (mm), W_{ks}^0 denotes the initial storage of soil reservoir (mm), R_{ki} and R_{ks} are the interflow runoff and rapid runoff (mm), and E_{ks} is the evapotranspiration from soil reservoir (mm). The equation of R_{ki} could be expressed as:

$$R_{ki} = K_{ki} \cdot W_{ks} \quad (10)$$

where K_{ki} denotes the interflow coefficient on karst pervious areas (-). The equation of R_{ks} could be expressed as:

$$R_{ks} = P_s + W_{ks} - W_{ks}^{max} + W_{ks}^{max} \left[1 - \frac{P_s + A_{ks}}{W_{ks}^{max}(1 + ek)}\right]^{ek+1} \quad (11)$$

where W_{ks}^{max} is the capacity of the soil reservoir (mm), A_{ks} denotes the watershed-averaged amount of water stored in the soil zone (mm), ek denotes the saturation water storage capacity on the karst area curve exponent (-). The equation of A_{ks} could be expressed as:

$$A_{ks} = W_{ks}^{max}(1 + ek)[1 - (1 - W_{ks}/W_{ks}^{max})^{1/(1+ek)}] \quad (12)$$

The equation of E_{ke} could be expressed as:

$$E_{ks} = \frac{[2W_{ks} - (W_{ks})^2/W_{ks}^{max}] \tanh(E_n/W_{ks}^{max})}{1 + (1 - W_{ks}/W_{ks}^{max}) \tanh(E_n/W_{ks}^{max})} \quad (13)$$

3. METHODS

3.1. The XAJ model

The XAJ model (Zhao 1992) is a conceptual hydrological model developed in 1973 and perfected in the 1980s. The critical assumption of the XAJ model is that no runoff is generated until soil moisture reaches field capacity, which is appropriate in the humid and semi-humid watersheds of China (Zang *et al.* 2021). After the total amount of runoff is evaluated, the runoff separation procedure separates generated runoff into surface, interflow, and groundwater runoff to employ different hillslope concentration methods. The XAJ model's major characteristic is interpreting the spatial distribution of tension and free water storage using two curves (Zhao 1992), which are also the basis of runoff generation and separation calculation.

The XAJ model's inputs are watershed-averaged precipitation and observed pan evaporation, while its outputs are actual evapotranspiration from the watershed and discharge at the watershed outlet. The basic structure of the XAJ model contains four modules: evapotranspiration calculation, runoff generation, runoff separation, and flow concentration. The whole watershed is subdivided into pervious and impervious areas according to the permeability of the underlying surface. The first three modules are combined to determine the generated surface, interflow, and groundwater runoff in the pervious areas. The portion of watershed-averaged precipitation that exceeds actual evapotranspiration directly forms the surface runoff in the impervious areas. The surface runoff on the whole watershed and the interflow and groundwater runoff in the pervious areas are routed from the hillslope to the watershed outlet using the flow concentration module.

3.2. The coupling of the 3T-KRGM and the XAJ model

A schematic diagram demonstrating the coupling of 3T-KRGM and the XAJ model is shown in Figure 3. To couple the 3T-KRGM (see Section 2 for details) into the XAJ model, the karst-dominated watershed is first subdivided into pervious and impervious areas. The pervious area is further separated into the karst and non-karst pervious areas, with a parameter A_{kp} denoting how much of the watershed is made up of karst pervious areas. The proposed 3T-KRGM is solely applied to the karst pervious areas, whereas the corresponding modules of the XAJ model are applied to the non-karst pervious areas and impervious areas. The 3T-KRGM shares the same inputs as the XAJ model, producing R_{ks} , R_{ki} , and R_{ke} as output. The corresponding modules of the XAJ model produce surface runoff on the non-karst pervious areas (R_s), interflow on the non-karst pervious areas (R_i), groundwater on the non-karst pervious areas (R_g), and surface runoff on the impervious areas (R_{imp}). R_{ks} , R_s , and R_{imp} formed the total surface runoff (R_s^a). R_{ki} and R_i formed the total interflow runoff (R_i^a). R_{ke} and R_g formed the total groundwater runoff (R_g^a). R_s^a , R_i^a , and R_g^a were routed to the watershed outlet and formed discharge Q by using the flow concentration module of the XAJ model. The XAJ models with and without 3T-KRGM are referred to as

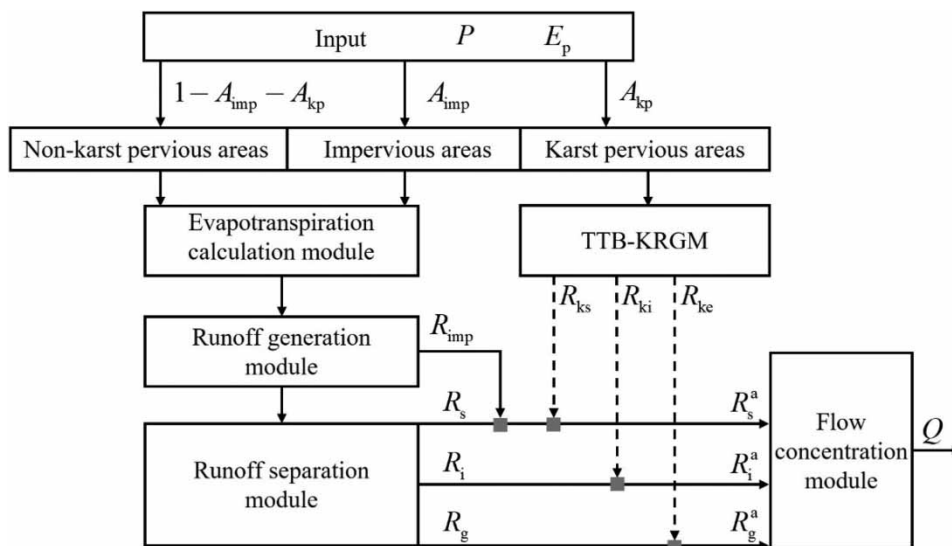


Figure 3 | The schematic diagram of the coupling of 3T-KRGM and the XAJ model. The dashed arrow denotes runoff components generated from the 3T-KRGM. The gray square denotes the summation of two runoff components.

the improved XAJ model and the original XAJ model, respectively. The parameters of the improved XAJ model and original XAJ model are listed in Table 1.

3.3. Study area and data

The Shibantang watershed (Figure 4) is located in southwest China, and the watershed outlet hydrological station is located on the Yeji River, which belongs to the Wujiang river system. The watershed has an area of 1,553 km², with an elevation range of 1,075–1,918 m a.s.l. The terrain of the watershed slopes from west to east, and its elevation varies considerably. The watershed belongs to a typical subtropical monsoon climate, and the average annual precipitation and temperature of the watershed are about 1,180 mm and 11.8 °C, respectively. In the Shibantang watershed, the surface exposed rock formations are made up of carbonate rock and sandy shale. The karst landscape is well developed and distributed, accounting for 86% of the whole watershed area (Xu *et al.* 2020). Individual karst forms are dispersed across carbonate rock areas, including stone buds, stone grooves, karst caves, karst funnel, and underground streams. In contrast, erosional landforms are primarily founded in sandy shale areas.

This study includes daily precipitation series from seven precipitation stations (Baina, Gantang, Hongjiadu, Liulong, Shachang, Shawo, and Shibantang), daily pan evaporation, and daily discharge series from Shibantang hydrological station. All data series used are from January 2006 to December 2018. The Annual Hydrological Report of China (Volume VI, Book X) is the data source for the daily precipitation and discharge, while the National Meteorological Science Data Center of China (<http://data.cma.cn/>) is the source for the daily pan evaporation. The Thiessen polygon method was used to calculate the watershed-averaged daily precipitation from all seven daily precipitation series. The elevation data of the watershed come from the SRTM 90 m DEM dataset (Geospatial Data Cloud of China, <http://www.gscloud.cn/>).

Table 1 | Parameters of the original and improved XAJ model (with signification and range)

Parameter	Signification	Range
K_e	The ratio of potential evapotranspiration to pan evaporation	[0.6,1.5]
c	Evapotranspiration coefficient of deep soil layer	[0.1,0.2]
W_{um}	Averaged tension water storage capacity of upper soil layer	[5,30]
W_{lm}	Averaged tension water storage capacity of lower soil layer	[60,90]
W_{dm}	Averaged tension water storage capacity of deep soil layer	[15,40]
A_{imp}	The ratio of impervious watershed area	[0.01,0.2]
b	The exponent of the tension water storage capacity curve	[0.1,0.4]
S_m	Averaged soil layer free water storage capacity	[10,30]
ex	The exponent of the free water storage capacity curve	[1,1.5]
K_i	Outflow coefficient of interflow from free water storage	[0.1,0.55]
K_g	Outflow coefficient of groundwater from free water storage	$0.7 - K_i$
C_i	Recession coefficient of interflow storage	[0.5,0.9]
C_g	Recession coefficient of groundwater storage	[0.95,0.998]
C_r	Recession coefficient of water storage in river network	[0.1,0.9]
l	Lag time	[1,10]
A_{kp}	The ratio of karst pervious area	[0.1,0.9]
W_{ke}^{max}	The capacity of the epikarst reservoir	[10,100]
W_{ks}^{max}	The capacity of the soil reservoir	[10,80]
ek	Saturation water storage capacity on the karst area curve exponent	[0.1,0.4]
K_{ki}	Interflow coefficient on karst pervious areas	[0.01,0.8]
W_{ki}^t	Threshold of the SEI	[0,10]

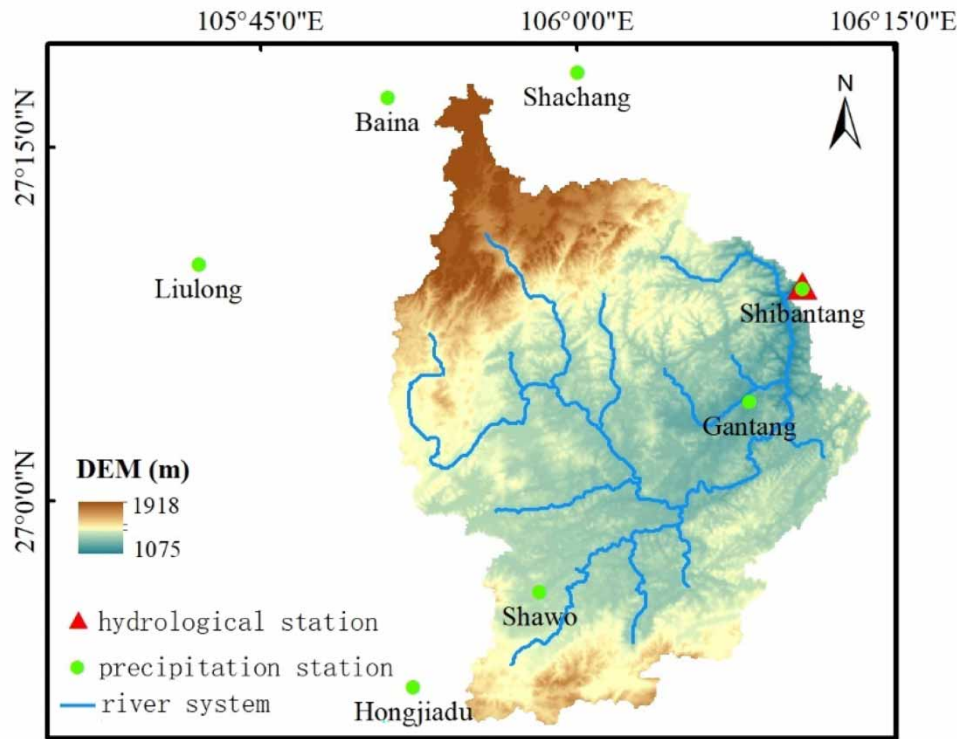


Figure 4 | The study area's location, topography, and river system with hydrological and precipitation stations.

3.4. Model setup and parameter calibration

The improved XAJ and original XAJ models were used to simulate the daily discharge series of the Shibantang watershed. According to time, the data series was divided into three periods: the model warm-up period (2006), the calibration period (2007–2013), and the validation period (2014–2018). Five performance metrics, including Nash–Sutcliffe efficiency coefficient (NSE) (Nash & Sutcliffe 1970), Kling–Gupta efficiency coefficient (KGE) (Gupta *et al.* 2009), flood peak relative error (FPRE), flood volume relative error (FVRE), and root mean square error (RMSE) were used to evaluate two models. The last three metrics were as follows:

$$\text{FPRE} = \frac{Q_s^{\max} - Q_o^{\max}}{Q_o^{\max}} \quad (14)$$

$$\text{FVRE} = \frac{\sum_{i=1}^N (Q_s^i - Q_o^i)}{\sum_{i=1}^N (Q_o^i)} \quad (15)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_s^i - \bar{Q}_o)^2} \quad (16)$$

where Q_s^{\max} is simulated flood peak (m^3/s), Q_o^{\max} is observed flood peak (m^3/s), i is the index of datum, N is the total amount of the data, Q_s^i and Q_o^i represent the i -th simulated and observed discharge, \bar{Q}_o is the average value of observed discharge. The Shuffled Complex Evolution (SCE-UA) algorithm (Duan *et al.* 1992) was used to optimize model parameters in the calibration period, and its objective function is $1 - \text{KGE}$.

3.5. Sensitivity analysis of the improved XAJ model

The sensitivity analysis of the improved XAJ model is to identify sensitive model parameters that can be used for model calibration and diagnostic evaluation (Pianosi *et al.* 2016). This study adopted the widely used PAWN method (Pianosi &

Wagener 2015; Puy *et al.* 2020) for global sensitivity analysis. The PAWN method treats the sensitivity of an input factor as the distance between the unconditional cumulative distribution function (CDF) of output obtained when all inputs vary simultaneously, and the conditional CDF obtained when varying all inputs but that input. The Kolmogorov–Smirnov (KS) statistic, i.e., the maximum absolute difference, is used to calculate the distance between the unconditional and conditional CDF. Given the complexity of the model input–output relationship, the KS statistic is calculated approximately rather than analytically. The approximation strategy detailed by Pianosi & Wagener (2018) was utilized, with the range of each input variable divided into n equally spaced intervals. The median of KS statistics across intervals was used to indicate the sensitivity of model output to model parameters. In this study, PAWN was conducted with 50,000 randomly generated parameter sets using the Latin hypercube sampling methods. The KGE performance metric was used to summarize the discharge series produced by the improved XAJ model into a scalar for sensitivity analysis. The value of n was set to 10, 20, and 40 to validate the robustness of the sensitivity analysis result according to the suggestion provided by Pianosi & Wagener (2018). The sensitivity analysis was conducted using the SALib Python library (Iwanaga *et al.* 2022).

4. RESULTS AND DISCUSSION

4.1. Sensitivity analysis of the improved model

Figure 5 shows the sensitivity of model parameters of the improved XAJ model. The greater the median of the KS statistic, the more sensitive the model parameter. The size of parameter sets for the PAWN method is fixed, while the value of n was set to a different value for the sensitivity analysis result's robustness. The result of $n = 20$ was quite comparable to the result of $n = 40$ but slightly different from the result of $n = 10$. The results of three different n values reflect a similar pattern in parameter ranking, indicating the result's reliability. Thus, the result of $n = 20$ was used for further analysis.

The most sensitive parameter is K_e , which regulates the evapotranspiration capacity and, consequently, the quantity of total runoff. The parameters L and C_r , which are used to route the runoff (originating from both karst and non-karst area) stored in the river channel to the watershed outlet, have ranks of 2 and 6, respectively. The three parameters mentioned above come from the original XAJ model and affect the improved XAJ model. K_e , L , and C_r are all sensitive model parameters according to previous studies (Zhao 1992). The fourth sensitive parameter is A_{imp} , which represents the ratio of impervious area within a watershed. In impervious areas, surface runoff is identical to net precipitation and is an important source of flood peaks. Thus, A_{imp} should be a sensitive parameter according to existing model practice. A_{kp} is the third sensitive parameter, which represents the ratio of the karst pervious areas of the watershed. It supports the importance of the 3T-KRGM because this module is only used on the karst pervious areas. K_{ki} and W_{ks}^{max} mainly controls the interflow and rapid runoff generated on the karst area (R_{ki} and R_{ks}), with ranks of 5 and 7, respectively. It indicates R_{ki} and R_{ks} play an important role in the simulation of the Shibantang watershed.

The first seven sensitive parameters (as shown in Figure 5), the remaining parameters of 3T-KRGM, and the remaining sensitive parameters of the original XAJ model (S_m , K_i , C_i , and C_g , see Table 1 for definition) were used in model parameter optimization.

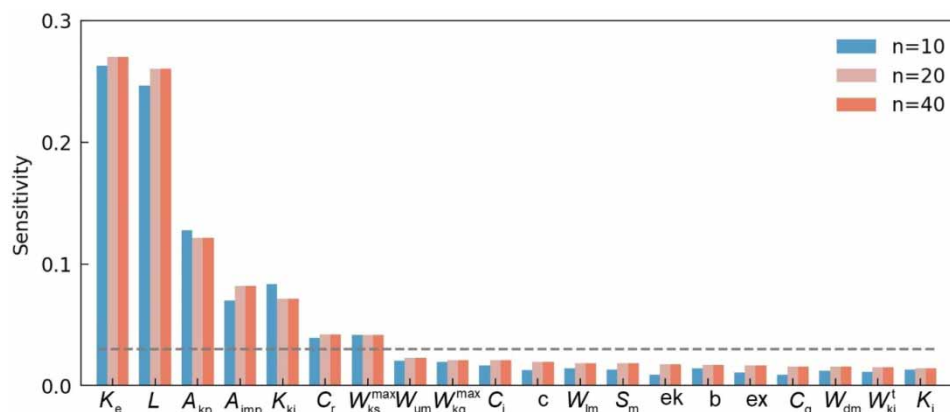


Figure 5 | The sensitivity of the improved XAJ model's parameters by the PAWN method. The parameters are in descending order using the result of $n = 20$.

4.2. Comparison of performances and parameters of the improved and original model

The original XAJ and improved XAJ models were used in the Shibantang watershed to simulate daily discharge series. The optimized parameters of the original XAJ model and improved XAJ model are listed in Table 2, and the simulation results of these two models are given in Figure 6(a), accompanied by observations. Five performance metrics were used to evaluate the simulation results for each year, including NSE, KGE, FPRE, FVRE, and RMSE (see Section 3.4 for more information). Table 3 illustrates the original and improved model's performance during the calibration and validation period, and Figure 6(b) and 6(c) displays the annual performance metrics of these two models.

The parameter K_e decides the evapotranspiration capacity according to Equation (4). The parameter W_{um} controls the water storage for direct evapotranspiration. A smaller K_e and smaller W_{um} combination means more runoff would be generated, which is revealed in Figure 6(b), as the FVRE of the original XAJ model is bigger than the improved XAJ model for most cases. The optimized values of S_m , K_i , and C_i for the original XAJ model were different from those of the improved XAJ

Table 2 | The optimized parameters of the original and improved XAJ model

Parameter	K_e	c	W_{um}	W_{lm}	W_{dm}	A_{imp}	b	S_m	ex	K_i	K_g
Original	0.86	0.14	9.9	76.8	27.5	0.10	0.23	14.9	1.14	0.4	0.3
Improved	1.10	0.14	23.1	75.2	27.0	0.11	0.21	22.1	1.28	0.3	0.4
Parameter	C_i	C_g	C_r	l	A_{kp}	W_{ke}^{max}	W_{ks}^{max}	ek	K_{kl}	W_{ki}^t	
Original	0.62	0.96	0.57	1	-	-	-	-	-	-	
Improved	0.73	0.97	0.58	1	0.81	36.1	41.6	0.26	0.03	0.85	

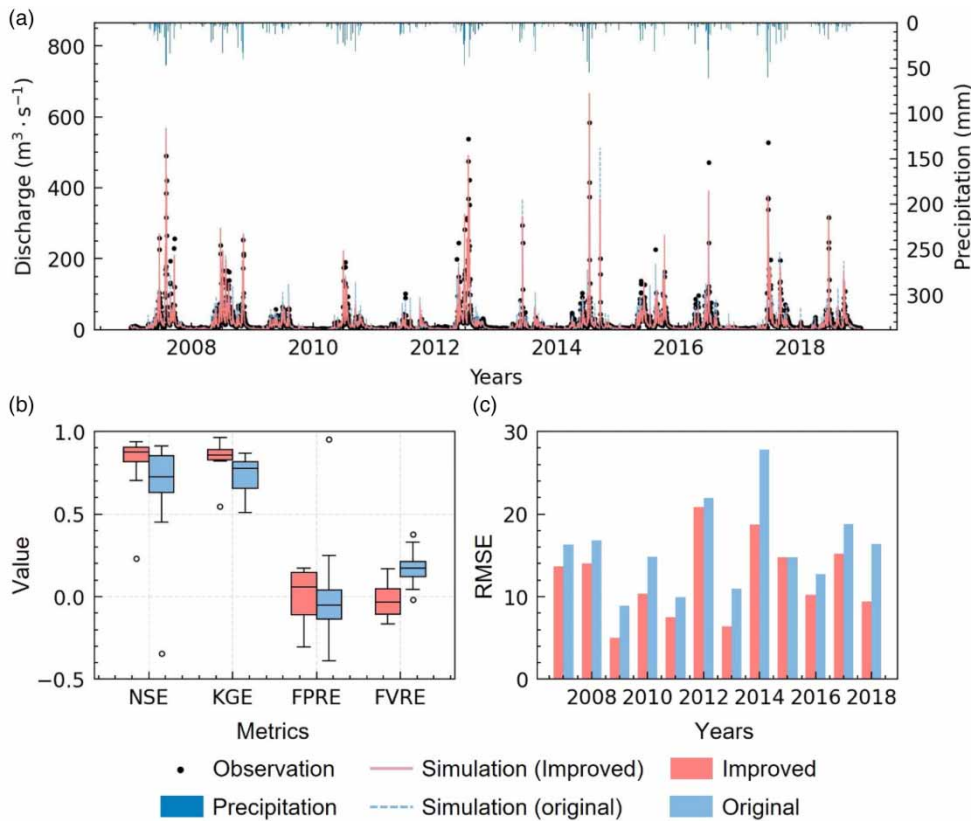


Figure 6 | The comparison of simulation results of the original and improved XAJ model. (a) Simulated daily stream discharge hydrographs accompanied by observations; (b,c) the annual performance metrics of simulation results of these two models.

Table 3 | The performance metrics of the original and improved XAJ model in the calibration and validation period

Model	Period	NSE	KGE	FPRE (%)	FVRE (%)	RMSE
Original	Calibration	0.86	0.79	-21.9	15.3	14.9
	Validation	0.76	0.83	2.4	14.0	18.8
Improved	Calibration	0.90	0.94	-8.7	-2.0	12.2
	Validation	0.86	0.88	14.0	-6.0	14.1

model, which indicates more surface runoff, more interflow runoff, and faster interflow concentration velocity. The comparison of model parameter values demonstrates that the rapid flow runoff component in karst areas was compensated with more surface runoff, more and faster interflow in the original XAJ model. The compensation is at the expense of a higher FVRE, slower hydrograph metrics (NSE and KGE), and, most importantly, the right result for the wrong reasons (Goswami *et al.* 2010). The SCE-UA optimized parameter value of the A_{kp} was 0.81, which is quite similar to the karst landscape ratio of the Shibantang watershed (Xu *et al.* 2020). It explains the rationality of model parameters of the improved XAJ model to a certain extent.

As can be seen from Table 3, the improved XAJ model was better than the original XAJ model for both calibration and validation periods in terms of the NSE, KGE, FVRE, and RMSE performance metrics. The boxplots of Figure 6(b) represent the distribution of annual performance metric values with 25%, 50% (median), and 75% quantile, as well as outliers. Compared to the original XAJ model, the NSE and KGE values of the improved XAJ model are close to 1, and the FVRE value of the improved XAJ model is close to 0, as shown in Figure 6(b). According to Figure 6(c), the RMSE values of the improved XAJ model are lower than the original XAJ model over the majority of years. Meanwhile, these two models each have advantages and disadvantages in the calibration and validation period for the FPRE metric, as shown in Table 3. The absolute mean values of the annually evaluated FPRE of the improved XAJ model for the calibration period and the validation period are 16.2 and 15.4%, respectively. In comparison, the corresponding values for the original XAJ model are 22.0 and 17.9%. It revealed that the improved XAJ model was better than the original XAJ model in terms of the FPRE metric. To conclude, the improved XAJ model performs better than the original XAJ model in the Shibantang watershed during the simulation period, indicating the model performance improvement of the proposed 3T-KRGM.

4.3. Comparison of the simulation result of the improved and original model in flood events

Three typical flood events (20090619, 20120711, and 20180618), including low, medium, and high flood peaks, covering the calibration and validation period, were selected for detailed comparison. The discharge hydrographs of these three typical flood events were extracted from the simulation results (see Figure 6(a)) of the improved XAJ model and the original XAJ model. As seen in Figure 7, the improved XAJ model fits the observation better than the original XAJ model, especially in the flood recession period. The flood peaks produced by the improved XAJ model are smaller than those produced by the original XAJ model under small intensity, as shown in event 20090619, and the second half of event 20120711 and event 20180618. In the case of big intensity, the flood peaks produced by the improved XAJ model are bigger than those produced by the original XAJ model, as shown in the first half of event 20120711 and event 20180618. The comparison of simulation results of these two models in three flood events indicates the model performance improvement of the proposed 3T-KRGM in a detailed way.

4.4. Comparison of simulation-based baseflow and UKIH-estimated baseflow

The discharge series could be divided into quick and slow-flow components from the perspective of hydrograph separation, with the slow-flow component referred to as baseflow (Le Mesnil *et al.* 2021). Along with the conventional discharge comparison, the comparison of simulation-based baseflow and estimated baseflow may offer additional information regarding model performance evaluation. In this study, the summation of R_1^a and R_g^a after slope concentration with linear reservoirs denoted as Q_1^a and Q_g^a was treated as simulation-based baseflow produced by the improved XAJ model. The simulation-based baseflow did not include the fast groundwater flow component in karst areas because it was a part of R_s^a . Based on the observed daily discharge series, the estimated baseflow was generated using the United Kingdom Institute of Hydrology (UKIH) method (Piggott *et al.* 2005). This method separated the base flow by identifying the turning points of the hydrograph

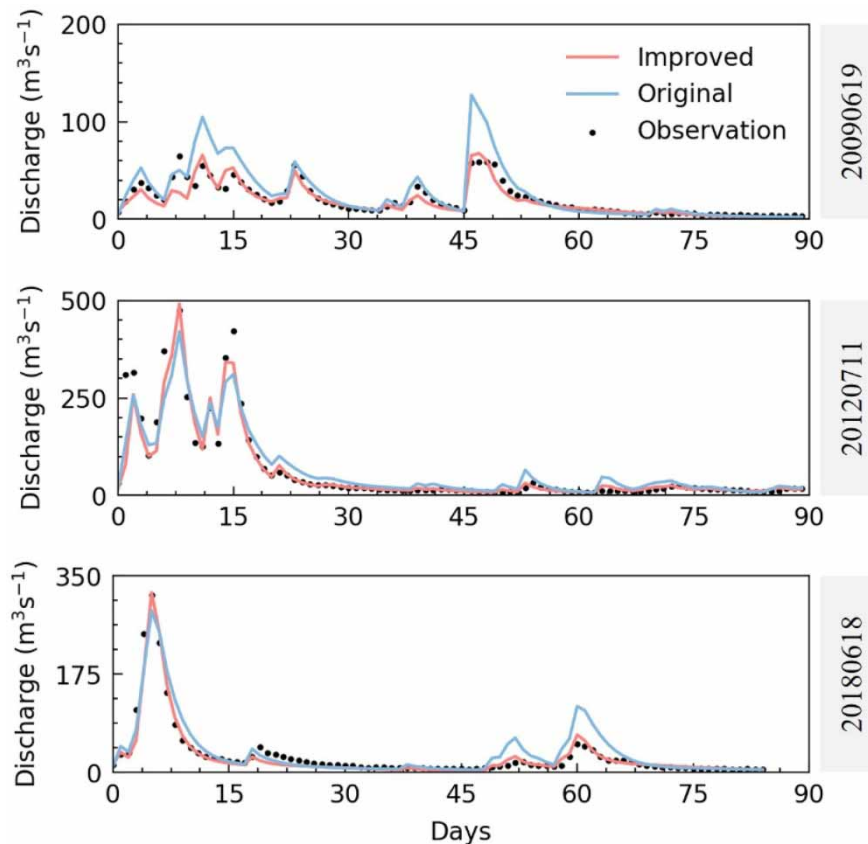


Figure 7 | Simulated daily discharge series of the original and improved XAJ model for three selected typical flood events.

and piecewise linear interpolation. [Figure 8](#) depicts comparisons between simulation-based baseflow and UKIH-estimated baseflow.

As shown in [Figure 8](#), the simulation-based baseflow agreed well with the UKIH-estimated baseflow in trends. The Pearson correlation coefficient between two baseflow series in flood events 20090619, 20120711, and 20180618 were 0.88, 0.92, and 0.72, respectively. Although there were some differences between simulation-based baseflow and UKIH-estimated baseflow, especially in the 20180618 flood event, we believed that the improved XAJ model could reproduce the baseflow well, considering the unavoidable uncertainty that existed in baseflow separation methods ([Partington et al. 2012](#)). The capacity of the improved XAJ model to reproduce the baseflow in karst areas further validates the rationality of the proposed 3T-KRGM.

5. CONCLUSIONS

A conceptual runoff generation module, called the three-thresholds-based karst runoff generation module (3T-KRGM), has been proposed to improve the performance of an existing hydrological model in karst areas. The 3T-KRGM was coupled into the XAJ model, a widely used lumped hydrological model in China. The XAJ model coupled with 3T-KRGM was referred to as the improved XAJ model, while the XAJ model not coupled with 3T-KRGM was referred to as the original XAJ model. The Shibantang watershed, with a karst ratio of 86%, was selected as the study area. The parameter sensitivity analysis of the improved XAJ model was conducted for diagnostic evaluation and model calibration. The improved XAJ model and the original XAJ model were used for the daily discharge simulation in the study area to evaluate the proposed 3T-KRGM. The main conclusions are given as follows:

- (1) The proposed 3T-KRGM considers the near-surface karst zone structure explicitly and uses three reservoirs to represent the water storage in the soil zone, soil-epikarst, and epikarst zone. The theoretical basis of 3T-KRGM is a modification of

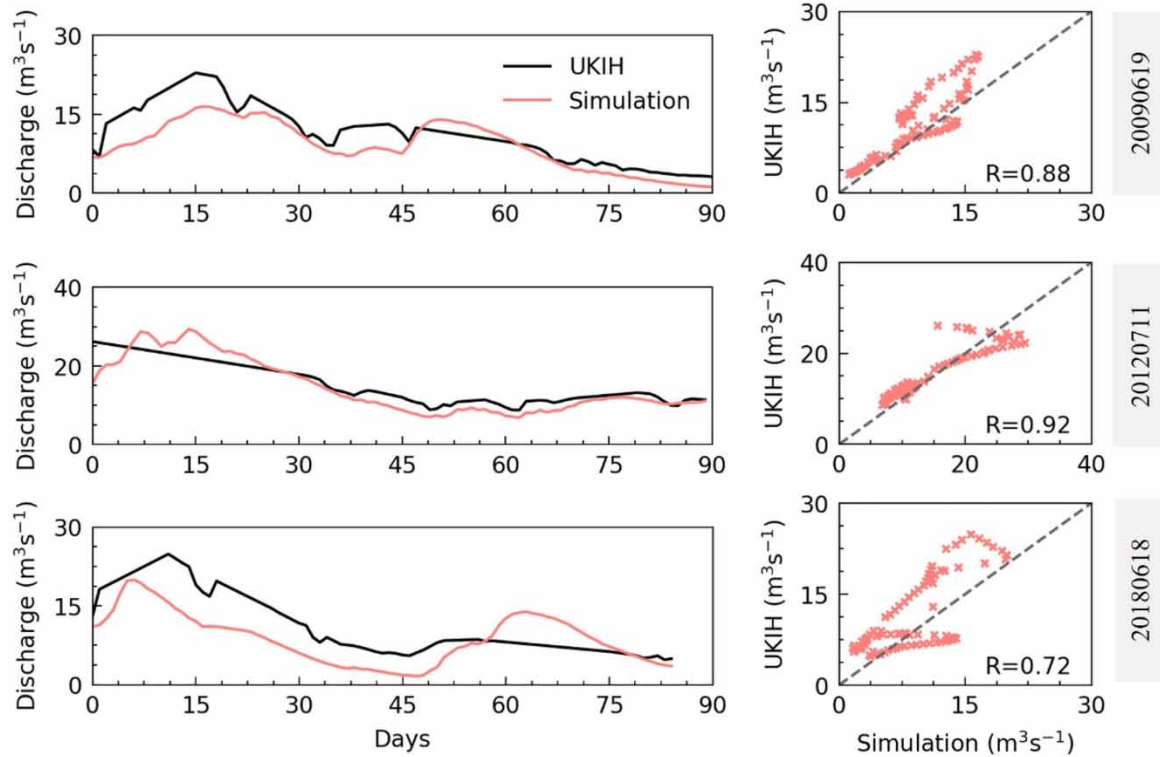


Figure 8 | The comparisons of simulation-based baseflow of the improved XAJ model and the UKIH-estimated baseflow for three selected flood events. R in the chart on the right denotes the Pearson correlation coefficient of simulation-based baseflow and UKIH-estimated baseflow.

three-threshold mechanism revealed by hillslope experiments (Wang *et al.* 2020) with consideration of rapid groundwater runoff. The 3T-KRGM has six parameters and produces rapid, interflow, and epikarst seepage runoff on karst pervious areas.

- (2) The parameter sensitivity analysis result obtained using the PAWN method reveals that three parameters from the 3T-KRGM are sensitive. Among these parameters, the ratio of karst pervious area A_{kp} is found to be the most sensitive. The calibrated value of A_{kp} is consistent with a reported value in the study area, which indicates the rationality of the 3T-KRGM's structure.
- (3) The improved XAJ model outperforms the original model with respect to performance metrics for the simulation period and the flood hydrograph for selected events in the study area. The improved XAJ model simulated slow-flow component matches UKIH-estimated baseflow during these events. The results show that the proposed 3T-KRGM improves the performance of existing hydrological model in karst areas and is reasonable.

Overall, the proposed three-thresholds-based karst runoff generation module (3T-KRGM) helps improve the XAJ model's performance in a karst-dominated watershed. The 3T-KRGM could also be coupled with other hydrological models, hence benefiting hydrological simulation in karst areas. Further research should be conducted to validate the proposed module with more watersheds.

ACKNOWLEDGEMENTS

This study was supported by the National Natural Science Foundation of China (NSFC) (grants, 41730750; 41877147). The authors appreciate the editors and anonymous reviewers for their critical comments and constructive suggestions, which helped improve the manuscript.

AUTHOR CONTRIBUTION STATEMENT

J.Z., G.L., and Z.L. conceptualized the study; J.Z., G.L., and Y.D. studied methodology and did formal analysis; J.Z. did the investigation and wrote the original draft; Y.D. did data curation; J.Z. and Y.H. visualized the study; G.L., Y.D., Y.H., B.L., and Z.L. wrote, reviewed, and edited the article; Z.L. and B.L. conducted funding acquisition; Z.L. supervised the work. All authors have read and agreed to the submitted version of the manuscript.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Bailly-Comte, V., Jourde, H. & Pistre, S. 2009 Conceptualization and classification of groundwater–surface water hydrodynamic interactions in karst watersheds: case of the karst watershed of the Coulazou River (Southern France). *Journal of Hydrology* **376** (3–4), 456–462. doi:10.1016/j.jhydrol.2009.07.053.
- Bonacci, O., Ljubenkovic, I. & Roje-Bonacci, T. 2006 Karst flash floods; an example from the Dinaric Karst (Croatia). *Natural Hazards and Earth System Sciences* **6** (2), 195–203. doi:10.5194/nhess-6-195-2006.
- Chen, S., Xiong, L., Zeng, L., Kim, J., Zhang, Q. & Jiang, C. 2022 Distributed rainfall-runoff simulation for a large-scale karst catchment by incorporating landform and topography into the DDRM model parameters. *Journal of Hydrology* **610**, 127853. doi:10.1016/j.jhydrol.2022.127853.
- Duan, Q., Sorooshian, S. & Gupta, V. 1992 Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* **28** (4), 1015–1031. doi:10.1029/91WR02985.
- Dubois, E., Doummar, J., Pistre, S. & Larocque, M. 2020 Calibration of a lumped karst system model and application to the Qachqouch karst spring (Lebanon) under climate change conditions. *Hydrology and Earth System Sciences* **24** (9), 4275–4290. doi:10.5194/hess-24-4275-2020.
- Duran, L., Massei, N., Lecoq, N., Fournier, M. & Labat, D. 2020 Analyzing multi-scale hydrodynamic processes in karst with a coupled conceptual modeling and signal decomposition approach. *Journal of Hydrology* **583**, 124625. doi:10.1016/j.jhydrol.2020.124625.
- Fang, L. & Shao, D. 2022 Application of long short-term memory (LSTM) on the prediction of rainfall-runoff in karst area. *Frontiers in Physics* **9**. doi:10.3389/fphy.2021.790687.
- Faticchi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., Downer, C. W., Camporese, M., Davison, J. H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M. & Tarboton, D. 2016 An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology* **537**, 45–60. doi:10.1016/j.jhydrol.2016.03.026.
- Ford, D. & Williams, P. D. 2007 *Karst Hydrogeology and Geomorphology*. Wiley, Chichester.
- Fu, Z. Y., Chen, H. S., Zhang, W., Xu, Q. X., Wang, S. & Wang, K. L. 2015 Subsurface flow in a soil-mantled subtropical dolomite karst slope: a field rainfall simulation study. *Geomorphology* **250**, 1–14. doi:10.1016/j.geomorph.2015.08.012.
- Goldscheider, N., Chen, Z., Auler, A. S., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Stevanovic, Z. & Veni, G. 2020 Global distribution of carbonate rocks and karst water resources. *Hydrogeology Journal* **28** (5), 1661–1677. doi:10.1007/s10040-020-02139-5.
- Goswami, M., O'Connor, K. M., Perrin, C. & Andreassian, V. 2010 A 'monster' that made the SMAR conceptual model 'right for the wrong reasons'. *Hydrological Sciences Journal* **55** (6), 913–927. doi:10.1080/02626667.2010.505170.
- Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. 2009 Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *Journal of Hydrology* **377** (1–2), 80–91. doi:10.1016/j.jhydrol.2009.08.003.
- Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. & Weiler, M. 2014 Karst water resources in a changing world: review of hydrological modeling approaches. *Reviews of Geophysics* **52** (3), 218–242. doi:10.1002/2013RG000443.
- Iwanaga, T., Usher, W. & Herman, J. 2022 Toward SALib 2.0: advancing the accessibility and interpretability of global sensitivity analyses. *Socio-Environmental Systems Modelling* **4**, 18155–18155. doi:10.18174/sesmo.18155.
- Jukić, D. & Denić-Jukić, V. 2009 Groundwater balance estimation in karst by using a conceptual rainfall–runoff model. *Journal of Hydrology* **373** (3–4), 302–315. doi:10.1016/j.jhydrol.2009.04.035.
- Kuczera, G. & Mroczkowski, M. 1998 Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research* **34** (6), 1481–1489. doi:10.1029/98WR00496.

- Le Mesnil, M., Moussa, R., Charlier, J. & Caballero, Y. 2021 Impact of karst areas on runoff generation, lateral flow and interbasin groundwater flow at the storm-event timescale. *Hydrology and Earth System Sciences* **25** (3), 1259–1282. doi:10.5194/hess-25-1259-2021.
- Li, J., Hong, A., Yuan, D., Jiang, Y., Deng, S., Cao, C. & Liu, J. 2021 A new distributed karst-tunnel hydrological model and tunnel hydrological effect simulations. *Journal of Hydrology* **593**, 125639. doi:10.1016/j.jhydrol.2020.125639.
- Liu, R., Wang, J., Zhan, H., Chen, Z., Li, W., Yang, D. & Zheng, S. 2021 Influence of thick karst vadose zone on aquifer recharge in karst formations. *Journal of Hydrology* **592**, 125791. doi:10.1016/j.jhydrol.2020.125791.
- Malagò, A., Efstathiou, D., Bouraoui, F., Nikolaidis, N. P., Franchini, M., Bidoglio, G. & Kritsotakis, M. 2016 Regional scale hydrologic modeling of a karst-dominant geomorphology: the case study of the Island of Crete. *Journal of Hydrology* **540**, 64–81. doi:10.1016/j.jhydrol.2016.05.061.
- Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290. doi:10.1016/0022-1694(70)90255-6.
- Partington, D., Brunner, P., Simmons, C. T., Werner, A. D., Therrien, R., Maier, H. R. & Dandy, G. C. 2012 Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water-groundwater flow model. *Journal of Hydrology* **458–459**, 28–39. doi:10.1016/j.jhydrol.2012.06.029.
- Paul, P. K., Zhang, Y., Ma, N., Mishra, A., Panigrahy, N. & Singh, R. 2021 Selecting hydrological models for developing countries: perspective of global, continental, and country scale models over catchment scale models. *Journal of Hydrology* **600**, 126561. doi:10.1016/j.jhydrol.2021.126561.
- Perrin, C., Michel, C. & Andréassian, V. 2003a Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology* **279** (1–4), 275–289. doi:10.1016/S0022-1694(03)00225-7.
- Perrin, J., Jeannin, P. & Zwahlen, F. 2003b Epikarst storage in a karst aquifer: a conceptual model based on isotopic data, Milandre test site, Switzerland. *Journal of Hydrology* **279** (1–4), 106–124. doi:10.1016/S0022-1694(03)00171-9.
- Pianosi, F. & Wagener, T. 2015 A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software* **67**, 1–11. doi:10.1016/j.envsoft.2015.01.004.
- Pianosi, F. & Wagener, T. 2018 Distribution-based sensitivity analysis from a generic input-output sample. *Environmental Modelling & Software* **108**, 197–207. doi:10.1016/j.envsoft.2018.07.019.
- Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. & Wagener, T. 2016 Sensitivity analysis of environmental models: a systematic review with practical workflow. *Environmental Modelling & Software* **79**, 214–232. doi:10.1016/j.envsoft.2016.02.008.
- Piggott, A. R., Moin, S. & Southam, C. 2005 A revised approach to the UKIH method for the calculation of baseflow. *Hydrological Sciences Journal* **50** (5), 911–920. doi:10.1623/hysj.2005.50.5.911.
- Puy, A., Lo Piano, S. & Saltelli, A. 2020 A sensitivity analysis of the PAWN sensitivity index. *Environmental Modelling & Software* **127**, 104679. doi:10.1016/j.envsoft.2020.104679.
- Simmons, C. T., Brunner, P., Therrien, R. & Sudicky, E. A. 2020 Commemorating the 50th anniversary of the Freeze and Harlan (1969) Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology* **584**, 124309. doi:10.1016/j.jhydrol.2019.124309.
- Tian, Y., Wang, S., Bai, X., Luo, G. & Xu, Y. 2016 Trade-offs among ecosystem services in a typical Karst watershed, SW China. *Science of the Total Environment* **566–567**, 1297–1308. doi:10.1016/j.scitotenv.2016.05.190.
- Tritz, S., Guinot, V. & Jourde, H. 2011 Modelling the behaviour of a karst system catchment using non-linear hysteretic conceptual model. *Journal of Hydrology* **397** (3–4), 250–262. doi:10.1016/j.jhydrol.2010.12.001.
- Tromp-van Meerveld, H. J. & McDonnell, J. J. 2006 Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resources Research* **42** (2). doi:10.1029/2004WR003800.
- Wang, S., Fu, Z., Chen, H., Nie, Y. & Xu, Q. 2020 Mechanisms of surface and subsurface runoff generation in subtropical soil-epikarst systems: implications of rainfall simulation experiments on karst slope. *Journal of Hydrology* **580**, 124370. doi:10.1016/j.jhydrol.2019.124370.
- Wang, S., Yan, Y., Fu, Z. & Chen, H. 2022 Rainfall-runoff characteristics and their threshold behaviors on a karst hillslope in a peak-cluster depression region. *Journal of Hydrology* **605**, 127370. doi:10.1016/j.jhydrol.2021.127370.
- Williams, P. W. & Andreo, B. 2008 The role of the epikarst in karst and cave hydrogeology; a review. *International Journal of Speleology* **37** (1), 1–10. doi:10.5038/1827-806X.37.1.1.
- Xu, C., Xu, X., Liu, M., Li, Z., Zhang, Y., Zhu, J., Wang, K., Chen, X., Zhang, Z. & Peng, T. 2020 An improved optimization scheme for representing hillslopes and depressions in karst hydrology. *Water Resources Research* **56** (5). doi:10.1029/2019WR026038.
- Yang, W., Chen, L., Chen, X. & Chen, H. 2022 Sub-daily precipitation-streamflow modelling of the karst-dominated basin using an improved grid-based distributed Xinanjiang hydrological model. *Journal of Hydrology: Regional Studies* **42**, 101125. doi:10.1016/j.ejrh.2022.101125.
- Zang, S., Li, Z., Zhang, K., Yao, C., Liu, Z., Wang, J., Huang, Y. & Wang, S. 2021 Improving the flood prediction capability of the Xin'anjiang model by formulating a new physics-based routing framework and a key routing parameter estimation method. *Journal of Hydrology* **603**, 126867. doi:10.1016/j.jhydrol.2021.126867.

- Zhang, Z., Chen, X., Ghadouani, A. & Shi, P. 2011 [Modelling hydrological processes influenced by soil, rock and vegetation in a small karst basin of southwest China](#). *Hydrological Processes* **25** (15), 2456–2470. doi:10.1002/hyp.8022.
- Zhao, R. 1992 The Xinanjiang model applied in China. *Journal of Hydrology* **135** (1), 371–381. doi:10.1016/0022-1694(92)90096-E.
- Zhou, Q., Chen, L., Singh, V. P., Zhou, J., Chen, X. & Xiong, L. 2019 [Rainfall-runoff simulation in karst dominated areas based on a coupled conceptual hydrological model](#). *Journal of Hydrology* **573**, 524–533. doi:10.1016/j.jhydrol.2019.03.099.

First received 2 December 2022; accepted in revised form 11 April 2023. Available online 24 April 2023