



Investigating the appropriate model structure for simulation of a karst catchment from the aspect of spatial complexity

Yong Chang¹ · Jichun Wu¹ · Guanghui Jiang^{2,3} · Xiaoe Zhao¹ · Qiang Zhang^{2,3}

Received: 30 June 2018 / Accepted: 19 December 2018 / Published online: 2 January 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Multi-model frameworks are widely used to identify the appropriate model structure for the study catchment. However, most frameworks mainly consider the process complexity of the model, and few of them consider the spatial complexity. In this paper, we investigated the appropriate model structure for a karst catchment from the aspect of spatial complexity. The purpose is twofold: (1) to investigate whether the spatial complexity is needed to simulate the spring discharge of this karst catchment and (2) to investigate whether the increase of model's spatial complexity can make up its deficiency on the process complexity. Three simple lumped models with different process complexities were chosen to gradually increase the spatial heterogeneity of their parameters to investigate the appropriate model structure for simulating the discharge of a karst spring. The results show that the performances of three lumped models highly improve when adding the routing function to them. However, further considering the spatial parameter heterogeneity, only one model shows obvious performance improvement and other two models show limited improvement. Moreover, this model with relatively complex spatial parameter heterogeneity still shows worse performance than another lumped model. This indicates an increase of models' spatial complexity cannot always make up their process deficiencies. The final comparison results indicated that the lumped model or their semi-lumped version with flexible process complexity is enough to simulate the discharge of this karst spring and no extra spatial complexity is needed. Our studies also indicated that the increase in spatial complexity of the model cannot always fully compensate its deficiency in process complexity.

Keywords Lumped model · Semi-distributed model · Process complexity · Spatial complexity · Karst catchment

This article is a part of Topical Collection in Environmental Earth Sciences on Characterization, Modeling, and Remediation of Karst in a Changing Environment, guest edited by Zexuan Xu, Nicolas Massei, Ingrid Padilla, Andrew Hartmann, and Bill Hu.

✉ Yong Chang
www.kr@163.com

¹ Key Laboratory of Surficial Geochemistry, Department of Hydrosociences, School of Earth Sciences and Engineering, Ministry of Education, Nanjing University, Nanjing 210093, China

² Key Laboratory of Karst Dynamics, Ministry of Land and Resources & Guangxi, Institute of Karst Geology, Guilin 541004, China

³ International Karst Research Center Auspices of UNESCO, Guilin 541004, China

Introduction

One important task in the hydrology is to quantify the catchment behavior under different climatic conditions or land uses, which is helpful to develop strategies for the water resources planning and management. The hydrologic model is a powerful tool to accomplish this task. From the last century to now, various hydrologic models were proposed and had been applied to many different catchments (Beven 2012; Hartmann et al. 2014; Singh and Woolhiser 2002). However, many models still suffer from the potential weakness of their model structures, besides the observational uncertainties, which may provide good calibration results but poor predictions (Hrachowitz and Clark 2017). How to find the appropriate model structure for the study catchment is still a big challenge in hydrology.

The multi-model framework is widely used recently to understand the dominant hydrological process and find appropriate model structure for the study catchment (Chang

et al. 2017; Clark et al. 2008; Coxon et al. 2014; Euser et al. 2013; Fenicia et al. 2008, 2014). This framework always contains many potential model structures with different complexity and the performances of these different models are compared to identify the appropriate model structure. Although this approach is widely used, most multi-model frameworks mainly consider different process complexities of the models to find the appropriate lumped model. As stated by Hrachowitz and Clark (2017), “the fundamental question that needs to be addressed in any model application is which model complexity is supported by the available data, including both process complexity, . . . , and spatial complexity, . . .”. Besides the process complexity, the spatial complexity of the model may also be an important factor that should be considered in the model. Natural systems always exhibit strong spatial heterogeneity introduced by geology, soil types, vegetation, or topography (Das et al. 2008; Kumar et al. 2010; Smith et al. 2004). Therefore, the hydrologic models should consider the spatial heterogeneity of the catchment to represent a more realistic model structure. Whether the consideration of spatial heterogeneity in the hydrologic model could improve model performance has been studied for many years and different results are presented. Several studies indicated that the lumped model is versatile enough to represent the spatial heterogeneity of the catchment and it could provide very similar or even better performance than the distributed model (Ghavidelfar et al. 2011; Reed et al. 2004; Refsgaard and Knudsen 1996; Vansteenkiste et al. 2014). However, many other studies indicated that the model performance on the streamflow could be improved by the consideration of spatial complexity (Yan and Zhang 2014; Atkinson et al. 2003; Boyle et al. 2001b; Euser et al. 2015; Han et al. 2014; Kumar et al. 2010; Nijzink et al. 2016). Therefore, whether the spatial heterogeneity is needed in the hydrological model and what is the level of spatial complexity required in the model is still unclear and will depend on the system being studied.

Karst aquifers are famous for their strong spatial heterogeneity of water flow and storage, such as diffuse recharge vs. point recharge, rapid, and often turbulent flow in the conduit network vs. slow laminar flow in the fissured matrix (Ford and Williams 2007; Goldscheider and Drew 2007). Although it has more complex hydrological processes than other aquifers, many simple models are still applied in these aquifers (Bailly-Comte et al. 2012; Chang et al. 2015; Fleury et al. 2007; Hartmann et al. 2012b, 2013b; Tritz et al. 2011). Several studies are also conducted to identify the appropriate model structure through the comparison of multi-models (Chang et al. 2017; Hartmann et al. 2013a). However, these works are also limited to the process complexity of the lumped model. Theoretically, the karst catchment may have higher spatial heterogeneity than other aquifers, because the karst development is more susceptible to the external factors,

such as lithology, vegetation, topography, hydrodynamic condition, and so on (Ford and Williams 2007). Therefore, it is necessary to investigate whether the spatial heterogeneity is needed for the simulation of these karst aquifers.

Since the model performance on the streamflow may be improved by the increase in its process complexity or spatial complexity, it brings another question: whether the model with simple process complexity and high spatial complexity could provide similar or even better performance than the lumped model of high process complexity? In other words, would an increase in spatial complexity of the model effectively compensate for its process deficiency? If it does, we may find another appropriate model structure for simulation of the study catchment except for the lumped model.

In this paper, we investigated the possible appropriate model structure of a karst aquifer from the aspect of spatial complexity. Chang et al. (2017) had discussed the appropriate model structure of this karst aquifer through a multi-model framework. However, that framework only explores the process complexity of the lumped model supported by the available data and the spatial complexity of the model is not considered. Therefore, the work in this paper can be regarded as an important supplement to the previous work. In general, this paper has two purposes. One of them is to understand whether extra spatial complexity is needed for the lumped model to simulate this karst aquifer? Another purpose is to find whether the increase of model's spatial complexity can effectively make up its process deficiency and whether another appropriate model structure with the relatively simple process and complex spatial complexity exists. Three lumped models with different process complexities were chosen from the multi-model framework proposed by Chang et al. (2017), and then, these models were gradually changed into the complex semi-distributed model by adding the routing function and gradually considering the spatial parameter heterogeneity. The appropriate spatial complexity of each model is determined through the comparison of model performance and sensitivity analysis as presented in Chang et al. (2017). Subsequently, the performances of different model structures are compared to find whether the increase in model's spatial complexity could effectively compensate its process deficiency. Finally, the most appropriate model structure for this karst aquifer is screened out.

Site descriptions

The Yaji experimental site is located in the southeast of Guilin city, China. This karst aquifer is developed in a very thick formation (several 100 m) of upper Devonian pure limestone and the geomorphology belongs to a typical peak-cluster-depression landform. The climate of this

place belongs to typical subtropical monsoon with annual precipitation of about 1915 mm (Yuan et al. 1996). The main rainy season is from April to August and the proportion is above 70% of annual precipitation. The storm is frequent with the highest rainfall intensity of 285.9 mm/day according to the previous records. The average annual temperature is about 18.8 °C with the high temperature in summer and low temperature in winter. Under climatic conditions, the karstification degree of this aquifer is very high and various karst features, such as cave, shaft, sinkholes, and karren, can all be found in this site. In general, the karst development is mainly controlled by NEE-oriented structures and this direction is also the main flow direction of the groundwater. Due to the massive deforestation in the last century, the vegetation in this area is mainly secondary shrub with the vegetation coverage of 60–80%.

There are three springs (S31, S291, and S29) located in the west of this experimental site and their resurgences are mainly controlled by the regional NNE-oriented structure. Among these three springs, S31 is the biggest spring and many studies, including spring hydrochemistry and hydrologic modeling, had been done in this spring catchment (Chang et al. 2015, 2017; Liu et al. 2004; Jiang et al. 2011; Yuan et al. 1996). The catchment of this spring mainly contains three depressions (Nos. 1, 3, and 4) and the rainfall is the only recharge resource of the spring (Fig. 1). For each depression, there is at least one sinkhole in its bottom and the previous tracer tests indicate that these sinkholes all connect to the spring S31 directly through the conduit. The epikarst zone is well developed in this experimental site and several epikarst springs distribute in each depression. These sinkholes drain the water in the depression directly to the

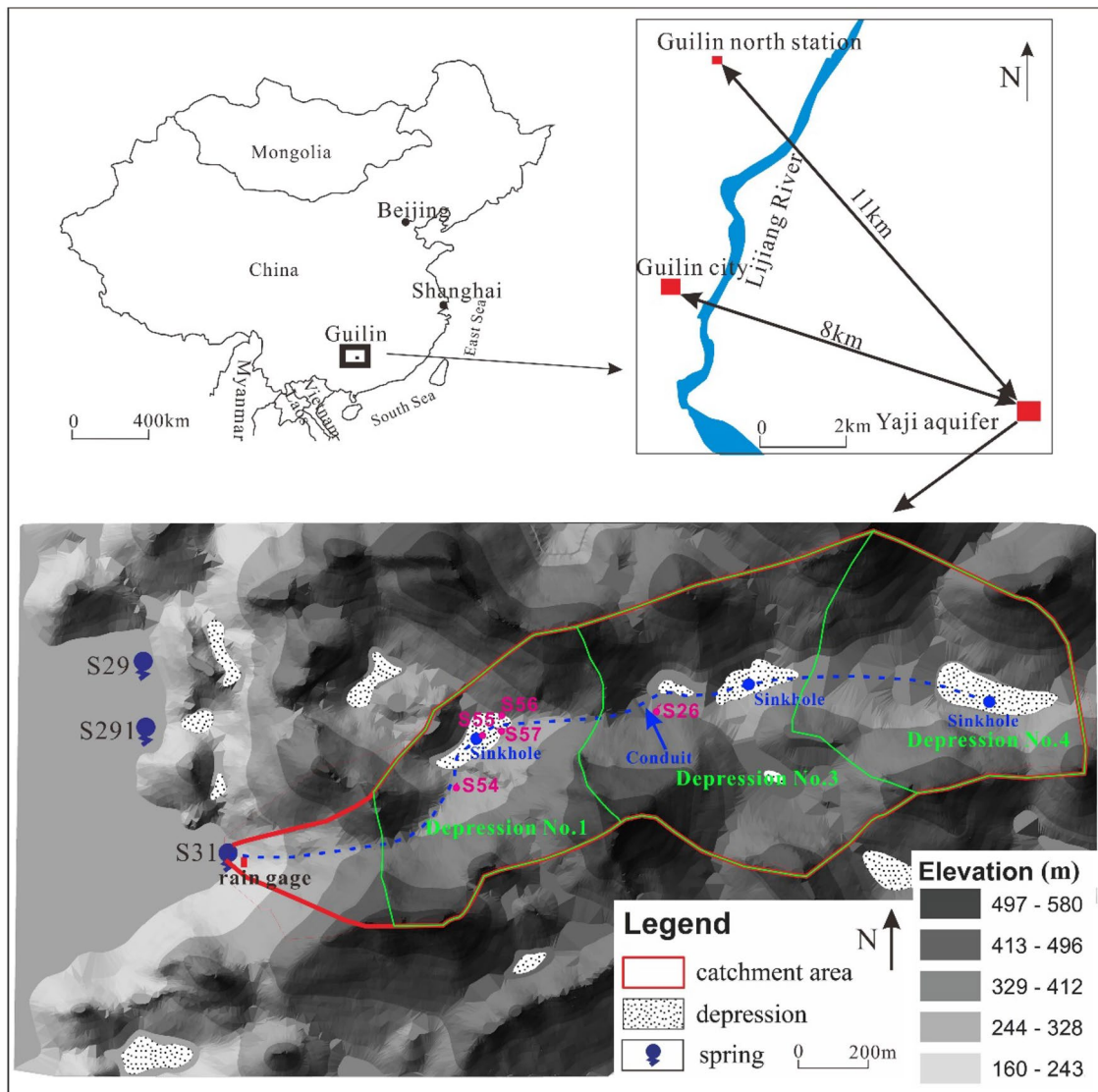


Fig. 1 Location and catchment area of spring S31

spring. Unfortunately, the conduit in this spring catchment is not accessible and its accurate position is not available. However, its rough location could be speculated according to the locations of sinkholes and surface topography (the blue-dashed line in Fig. 1).

Although the surface catchment area of S31 is very small (about 1 km²), we also found obvious difference characteristics among different depressions. The bottom elevation of depression No. 1 is much lower (about 50 m) than depression Nos. 3 and 4 which cause the much steeper slope of the depression No. 1. Meanwhile, much more epikarst springs are distributed in depression No. 1 than other depressions. There are about four different epikarst springs (S54, S55, S56, and S57) in depression No. 1, whereas only one epikarst spring (S26) distributes in depression No. 3 and no epikarst spring is found in depression No. 4 (Fig. 1). The different distributions of the epikarst spring in different depressions indicate their different karstification degree. According to landform and karstification degree, the catchment should at least be divided into two different hydrological units: (1) depression No. 1 and (2) depression Nos. 3 and 4. For example, Chang (2015) used these two spatial parameter partitions in a distributed model to simulate the discharge of spring S31. However, whether the spatial parameter partition is truly needed to simulate the discharge of this spring is not well understood.

There is one rain gauge near the spring S31 to record the rainfall. The nearest meteorological station (Guilin north station) is located 11 km northwest of the spring S31 which can provide the daily average meteorological data for the model. Since 2016, a small weather station is set up near spring S31 and can provide the daily average meteorological data for the model.

Methods

In general, the spatial complexity of the model can be considered from two aspects: (1) spatial parameter heterogeneity (Atkinson et al. 2003; Boyle et al. 2001; Das et al. 2008) and (2) spatial heterogeneity of hydrological process (Euser et al. 2015; Gao et al. 2014; Savenije 2010). Due to a lack of knowledge of the spatial heterogeneity of hydrological process in the study site, we only consider the spatial parameter heterogeneity of the model. In this paper, three lumped models were chosen to evaluate their performances in the spring hydrograph with the increase of their spatial complexities. The study karst catchment belongs to a typical peak-cluster-depression landform which mainly contains three depressions. Each depression is relatively independent which behaves like sub-catchment in the non-karst catchment. They are mainly connected to the spring by the conduit network. Therefore, we consider the spatial parameter heterogeneity of the model based on the depression units.

Given that the conduit network is the main passage to connect each depression to the spring, we first add a routing function for considering the flow process in karst conduit in the lumped model to establish the lumped-routing model. Subsequently, the whole catchment is divided into three units and the semi-lumped model is set up according to the spatial distribution of the depressions and conduit network. Based on this semi-lumped model, the parameters for each unit are considered as different values gradually to establish the semi-distributed models. The appropriate spatial parameter heterogeneity of each model structure can be considered as the one having relatively good model performance and all identifiable parameters, which can be determined according to the performance comparison and parameter sensitivity analysis as presented in Chang et al. (2017).

Three lumped models

Chang et al. (2017) had used a multi-model framework including 12 different models to identify the appropriate model structure for the spring S31. In this paper, we chose three different lumped structures from that multi-model framework to establish the corresponding semi-distributed models. These three models all consist of the linear storage reservoir and an evapotranspiration reservoir. The first lumped model S1 is the simplest one which just uses a linear storage reservoir to simulate the behavior of the whole catchment (Fig. 2). The second lumped model M1 uses two parallel linear storage reservoirs to consider the quick and slow flow separately in the catchment and this model structure has been widely used to simulate the spring discharge of karst catchment (Fleury et al. 2007; Hartmann et al. 2012a, b). The third lumped model C1 uses a linear storage reservoir with two outlets of different height to simulate the behavior of the whole catchment. This model has been proved to be the most appropriate lumped model to simulate the discharge of spring S31 in the previous study (Chang et al. 2017). For each model, the same evapotranspiration reservoir is used to calculate the effective precipitation on the catchment. The potential evapotranspiration is calculated by Eagleman's method (Eagleman 1967) which just need air temperature and relative humidity. The actual evapotranspiration is assumed to be linear with the saturation degree of evapotranspiration reservoir. The parameter needed for this reservoir the maximal capacity S_{\max} . More information about this evapotranspiration can be found in these studies (Chang et al. 2015, 2017; Jukić and Denić-Jukić 2009). The main difference between the three lumped models and the corresponding models in the multi-model framework proposed by Chang et al. (2017) is that the triangular transfer function is not considered in these three models. The triangular transfer function is mainly used to consider the time lag between the rainfall and spring discharge, and this lag time would be considered by adding the routing function in these

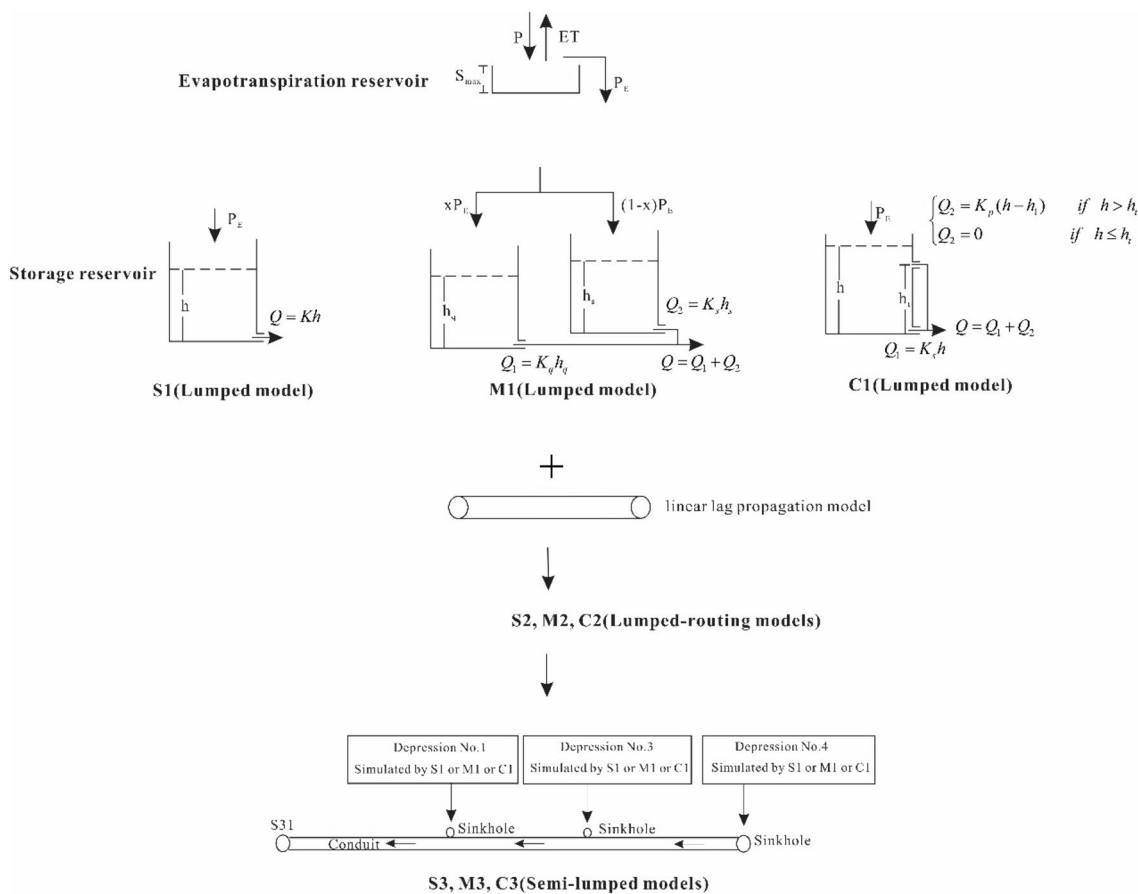


Fig. 2 Structures of lumped models (S1, M1, and C1), lumped-routing models (S2, M2, and C2) and semi-lumped models (S3, M3, and C3) to simulate the discharge of the karst spring

three lumped models in Sect. 3.2. In general, S1 is the simplest model with two parameters (S_{max}, K) which only considers one linear hydrological process, whereas M1 and C1 are relatively complex models with four parameters (S_{max}, x or h_i, K_s, K_q) which both consider two different linear hydrological processes (Tables 1, 2).

Lumped-routing and semi-lumped models

Based on each lumped model, a routine function is added first (Fig. 2). This routine function is mainly used to consider the flow routine in the conduit. Since the karstification degree in Yaji experimental site is very high, the flow in the conduit is very close to the flow in the river channel. Therefore, the simple linear lag propagation model for the river was used to simulate the flow process in the conduit (Bentura and Michel 1997; Lerat et al. 2012a):

$$g(\lambda, q_1, q_m, \dots, q_n)(t) = \sum_{i=1}^n [\alpha_i q_i(t - \delta_i) + (1 - \alpha_i) q_i(t - \delta_i - 1)] \tag{1}$$

With

$$\delta_i = [L_i/\lambda]$$

$$\alpha_i = 1 + [L_i/\lambda] - L_i/\lambda,$$

where $[]$ is the integer part, L_i is the distance between the point i and the spring (L), and λ is the routing parameter which is homogeneous to a celerity (L/T). By adding the lag propagation function on the lumped models, three different lumped-routing models (S2, M2, and C2) are presented with one extra parameter (L/λ) needed to be calibrated (Table 1). However, these models are still lumped, because any spatial information is not considered. In addition to this simple lag propagation model, we also had used the Hayami kernel function (Lerat et al. 2012b; Moussa 1996), which is the analytical solution of the linearized Saint-Venant equation by neglecting the inertia terms, to simulate the flow routing in the conduit. However, it gives very similar results with much more parameters. Therefore, only the results based on the simple lag propagation model are presented in this paper.

To further discretize the lumped model, we assume that the water in the depression all recharges the conduit through the sinkhole directly and the diffuse recharge is

Table 1 Parameter descriptions and their ranges

| Models | Parameters | Ranges | Units | Descriptions |
|--------|-------------|-------------|-------------------|--|
| S1–S8 | L/λ | 0–300 | min | Lag time, only used in S1 |
| | λ | 3–33 | m/min | Routing parameter |
| | K | 1e-04–1e-02 | min ⁻¹ | Coefficient of the linear storage reservoir |
| | S_{\max} | 0–0.2 | m | Maximal capacity of the evapotranspiration reservoir |
| M1–M8 | L/λ | 0–300 | min | Lag time, only used in M1 |
| | λ | 3–33 | m/min | Routing parameter |
| | x | 0–1 | – | Partition coefficient |
| | K_s | 1e-05–5e-03 | min ⁻¹ | Coefficient of the slow storage reservoir |
| | K_q | 5e-04–5e-02 | min ⁻¹ | Coefficient of the quick storage reservoir |
| | S_{\max} | 0–0.2 | m | Maximal capacity of the evapotranspiration reservoir |
| C1–C8 | L/λ | 0–300 | min | Lag time, only used in C1 |
| | λ | 3–33 | m/min | Routing parameter |
| | h_t | 0–0.04 | m | Threshold for the occurrence of upper outlet |
| | K_s | 5e-05–5e-03 | min ⁻¹ | Coefficient of the lower outlet |
| | K_q | 5e-04–5e-02 | min ⁻¹ | Coefficient of the upper outlet |
| | S_{\max} | 0–0.2 | m | Maximal capacity of the evapotranspiration reservoir |

Table 2 Brief descriptions of different models

| | Models | Parameter number | Lumped parameters | Parameters considering spatial heterogeneity | Descriptions | |
|------------------|--------|---------------------|--|---|--|--|
| Lumped | S1 | 2 | K, S_{\max} | – | The initial lumped model | |
| | M1 | 4 | x, K_s, K_q, S_{\max} | – | | |
| | C1 | 4 | h_t, K_s, K_q, S_{\max} | – | | |
| Lumped-routing | S2 | 3 | $L/\lambda, K, S_{\max}$ | – | Adding the linear lag propagation function | |
| | M2 | 5 | $L/\lambda, x, K_s, K_q, S_{\max}$ | – | | |
| Semi-lumped | C2 | 5 | $L/\lambda, h_t, K_s, K_q, S_{\max}$ | – | Spatial discretization | |
| | S3 | 3 | λ, K, S_{\max} | – | | |
| Semi-distributed | M3 | 5 | $\lambda, x, K_s, K_q, S_{\max}$ | – | Considering two spatial partitions of the evapotranspiration reservoir | |
| | C3 | 5 | $\lambda, h_t, K_s, K_q, S_{\max}$ | – | | |
| | S4 | 4 | λ, K | S_{\max} (Two partitions) | | Considering three spatial partitions of the evapotranspiration reservoir |
| | M4 | 6 | λ, x, K_s, K_q | S_{\max} (Two partitions) | | |
| | C4 | 6 | λ, h_t, K_s, K_q | S_{\max} (Two partitions) | | Considering two spatial partitions of the linear storage reservoirs |
| | S5 | 5 | λ, K | S_{\max} (Three partitions) | | |
| | M5 | 7 | λ, x, K_s, K_q | S_{\max} (Three partitions) | | Considering three spatial partitions of the linear storage reservoirs |
| | C5 | 7 | λ, h_t, K_s, K_q | S_{\max} (Three partitions) | | |
| | S6 | 4 | λ, S_{\max} | K (Two partitions) | | Considering two spatial partitions of the linear storage reservoirs |
| | M6 | 8 | λ, S_{\max} | x, K_s, K_q (Two partitions) | | |
| C6 | 8 | λ, S_{\max} | h_t, K_s, K_q (Two partitions) | Considering three spatial partitions of the linear storage reservoirs | | |
| S7 | 5 | λ, S_{\max} | K (three partitions) | | | |
| M7 | 11 | λ, S_{\max} | x, K_s, K_q (Three partitions) | Considering three spatial partitions of the evapotranspiration reservoir and the linear storage reservoirs simultaneously | | |
| C7 | 11 | λ, S_{\max} | h_t, K_s, K_q (Three partitions) | | | |
| S8 | 7 | λ | K, S_{\max} (Three partitions) | Considering three spatial partitions of the evapotranspiration reservoir and the linear storage reservoirs simultaneously | | |
| M8 | 13 | λ | x, K_s, K_q, S_{\max} (Three partitions) | | | |
| C8 | 13 | λ | h_t, K_s, K_q, S_{\max} (Three partitions) | | | |

neglected. This assumption is reasonable, since the karstification degree in Yaji experimental site is very high and most water should translate through the main conduit. Under this assumption, the lumped models (S1, M1, and C1) are used to simulate the discharge of each depression, and then, the discharge is routed by the conduit to the spring using the simple linear lag propagation model [Eq. (1)]. Therefore, we can set up the semi-lumped model without consideration of spatial parameter heterogeneity (Fig. 2). The area of each depression is determined according to the surface landform (Fig. 1). The conduit length between each depression and spring could be roughly estimated through the speculated conduit location. Although each depression may have several sinkholes at the bottom to drain the point recharge, we simplify these different sinkholes into one which locates at the bottom center of each depression (Fig. 1). Finally, we convert lumped-routing models (S2, M2, and C2) into three semi-distributed models (S3, M3, and C3) by considering the spatial distribution of sinkholes and conduit network (Fig. 2). As the conduit length between each depression and spring can be roughly obtained according to the speculated conduit location (Fig. 1), the parameter L is known and only one extra parameter (λ) is added in contrast to the lumped models (S1, M1, and C1).

It should be noted that the semi-distributed models described above only consider the three depressions. However, the catchment of spring S31 also contains a small hillslope above the spring in addition to the three depressions (Fig. 1). Actually, it is debatable to include the hillslope into the catchment of S31, since there is no sinkhole in this hillslope and only very limited diffuse infiltrations may recharge the spring S31. Most water should flow out of the catchment through the lateral flow in the epikarst zone. Meanwhile, the area of this hillslope is very small (about 0.07 km^2) in contrast to the whole catchment (1.09 km^2). Therefore, neglecting this hillslope may not cause serious water imbalance problem. To maintain consistency between the lumped and semi-distributed model, the catchment area used in the lumped models is set to 1.02 km^2 .

Semi-distributed models

Based on semi-lumped models (S3, M3, and C3), the spatial parameter heterogeneity is gradually considered to establish semi-distributed models to explore the appropriate spatial parameter partition for each model structure for simulation of the spring discharge (Table 2).

1. To consider the spatial heterogeneity of effective precipitation for each depression which may be caused by different surface slope, soil depth, or vegetation coverage, the parameter S_{\max} for each depression is gradually assumed to be different values. The values of other

parameters for different depressions are set to the same. First, we only consider S_{\max} of depression No. 1 is different from depression Nos. 3 and 4 (two parameter partitions) and we can get model S4, M4, and C4. Then, the values of S_{\max} in three depressions are all considered to be different (three parameter partitions) to set up more complex semi-distributed models S5, M5, and C5.

2. We only consider the spatial parameter heterogeneity of the linear storage reservoirs (K in S3, K_q, K_s and x in M3, and K_q, K_s , and h_t in C3) and keep the parameter (S_{\max}) to be the same in different depressions. Similar to the situation (1), two and three parameter partitions are considered separately (S6, M6, C6 and S7, M7, C7).
3. S_{\max} and the parameters of the linear storage reservoirs for each depression are all considered to be different to establish the fully semi-distributed model (S8, M8, C8).

In general, there are eight different models for each model structure and more details about these different lumped, semi-lumped, and semi-distributed models are shown in Table 2. With the consideration of spatial parameter heterogeneity of evapotranspiration and linear storage reservoirs, the parameter number increases gradually.

Optimization

The multi-objective optimization method is used in this paper to calibrate these different models and compare their performances. Two different objectives which put different emphasis on the high flows and low flows are chosen to evaluate the model performance:

$$F_1 = \frac{\sum (Q_m - Q_s)^2}{\sum (Q_m - \bar{Q}_m)^2} \tag{2}$$

$$F_2 = \frac{\sum (\ln(Q_m + \epsilon) - \ln(Q_s + \epsilon))^2}{\sum (\ln(Q_m + \epsilon) - \ln(\bar{Q}_m + \epsilon))^2}, \tag{3}$$

where Q_m and Q_s represent the measured and simulated discharge, and \bar{Q}_m is the average measured discharge in the simulated period. ϵ is a small constant to avoid a calculation problem when measured or simulated discharge gets to zero. In this paper, this small value is set to $0.002 \text{ m}^3/\text{s}$ subjectively (proximately 3% of the average discharge of this spring in the simulation period) and it would not strongly affect the evaluation of the model performance. The multi-objective optimization algorithm, AMALGAM, proposed by Vrugt and Robinson (2007) is used to search the optimal pareto front of these two criteria. The pareto front shows the trade-off between the two objectives. Each point on the front may not be considered better than other points on one

objective without causing a simultaneous deterioration of the other objective. The AMALGAM algorithm merges the strengths of several multi-objective algorithms to create the offspring of high-quality self-adaptively according to the performance of each algorithm and shows a significant evolution speed to the Pareto front. More information about this algorithm can be found in these two studies (Vrugt et al. 2009; Wöhling et al. 2008). The parameter range is determined according to the previous research (Chang et al. 2017) and is shown in Table 1. The population size in AMALGAM is set to 100 and the maximum number of iteration is set to 1000. The multi-objective optimization method is widely used to compare the performance of different models (Boyle et al. 2001; Chang et al. 2017; Fenicia et al. 2008; Lee et al. 2011). The improvement of model performance can be identified as the optimal Pareto front progressively moves toward the origin of the axes.

Sensitivity analysis

The regional sensitivity analysis (RSA) method (Freer et al. 1996; Hornberger et al. 1985) is chosen to evaluate the parameter identifiability indirectly. The parameter with low sensitivity is considered to be poorly identified. This analysis method is based on a random sampling of parameter space and could be easily used with multi-objectives. For each objective, the parameter population is ranked into 10 groups of equal size directly according to the objective function values and the cumulative distribution of the parameters in each group is plotted to evaluate the sensitivity of each parameter to this objective (Freer et al. 1996; Wagener et al. 2001). The Kolmogorov–Smirnov statistic (KS) is used to quantitatively assess the dispersion degree of these cumulative curves as proposed by Chang et al. (2017). The high value of KS indicates the high sensitivity of the parameter. For each parameter, two KS values can be got from two different objectives and the large value is chosen to represent the parameter sensitivity and identifiability.

The RSA method is implemented within the SAFE toolbox (Pianosi et al. 2015; Wagener et al. 2001). For each model in the multi-model framework, 10,000 random parameter groups are sampled in their defined spaces through the Latin hypercube method. The parameter range for the model calibration (Table 1) is also used for the sensitivity analysis.

Simulation periods

The rainfall–discharge data in two periods, from Jan to June in 2013 and 2017, respectively, are used for model calibration and validation, respectively. For each period, the first two months are set to the warm-up period to eliminate the influence of the initial condition on the simulation result. To fully explore the data information in two periods, two

calibration–validation procedures for each model were carried out. In the first procedure, the data in 2013 were used for calibration and data in 2017 for validation; in the second procedure, the data for model calibration and validation were switched. The sensitivity analysis was conducted in both periods. As the most available data except for the rainfall data in 2017 all have a time resolution of 15 min, the step time in the model is set to 15 min. For the rainfall data in 2017, the available time interval is half an hour. When using these data into the models, we simply divided the accumulation rainfall in half an hour into two equal values in each interval of 15 min.

Results

Model performances

Models S1–S8

Figure 3 shows the calibration and validation results of S1–S8 and the optimal combination of two objectives having the smallest distant to the origin of the axes on the optimal Pareto front for each model is shown in Table 3. The comparison results are consistent in two separate calibration–validation procedures. When the simple routing function is added to the model S1, S2 shows a little better performance than S1. The semi-lumped model S3 shows complete same Pareto front with the lumped model S2, indicating their same performances. S4 and S5 show very limited performance improvement in contrast to S3 when considering the spatial parameter heterogeneity of evapotranspiration reservoir. The model performance is highly improved when two spatial partitions of routing reservoir are considered in S6. However, when further considering three spatial partitions of routing reservoir in S7, it shows marginal performance improvement in contrast to S6. The final fully semi-distributed model S8 also shows limited performance improvement in contrast to S7.

Models M1–M8

Figure 4 shows the calibration and validation results of M1–M8 and the optimal combination of two objectives for each model is shown in Table 3. The comparison results among different models are also very similar in two calibration–validation procedures. The performance of M2 is much better than M1 when the simple routing function is added. However, the semi-lumped model M3 shows very similar performance with the lumped model M2. When further considering the spatial parameter heterogeneity of evapotranspiration and routing reservoir, the model performance only shows marginal improvement. The final semi-distributed

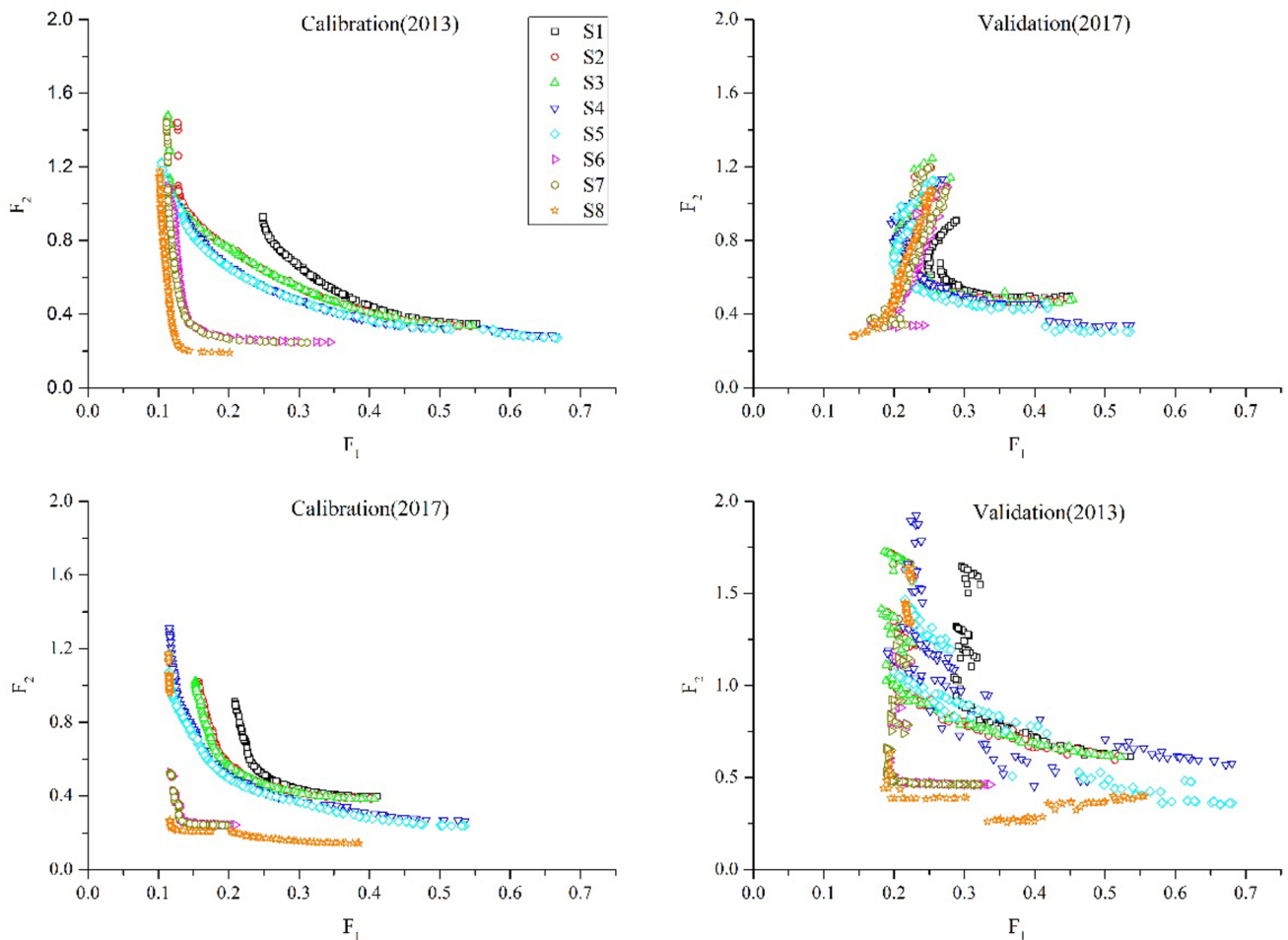


Fig. 3 Calibration and validation results of S1–S8 in two calibration–validation procedures. Calibration (2013) means that models are calibrated using the data from Jan to June in 2013; validation (2017)

means that models are validated using the data from Jan to Jun in 2017 and optimal parameter values in Calibration (2013)

model M8 also shows very small performance improvement in contrast to M3.

Models C1–C8

Figure 5 shows the calibration and validation results of C1–C8 and the optimal combination of two objectives for each model is shown in Table 3. The performance of C2 is always much better than C1 in both calibration and validation periods. The performance comparisons among other models show a little difference in different calibration–validation procedure. When using the data in 2013 for the calibration, the semi-distributed model C3 shows a litter better performance than C2. While considering the spatial parameter heterogeneity of evapotranspiration reservoir, C4 and C5 provide complete similar calibration results to C3. When considering the two spatial partitions of the routing reservoir, C6 shows a little better calibration performance than

C3. While further increasing the spatial partition from two to three, the model performance shows no obvious improvement. The final fully semi-distributed model C8 also shows no obvious performance improvement in contrast to C6 or C7 in the calibration period. However, When using the data in 2017 for the calibration, the model performance improves gradually with the increase of spatial partition of evapotranspiration and routing reservoir, the model performance (C3–C7) always be improved in the calibration period. The final fully semi-distributed model C8 shows the best calibration result. Although these different models show obvious different performance in the calibration period, the performance differences among these models are limited in the validation period. Only some pareto solutions of C4, C6, and C7 show a little performance improvement. It is worth to be mentioned that some pareto solutions of C8 show a little worse performance than C6 or C7 in the validation period which indicates that C8 is in the high risk of over-fitting.

Table 3 Relative best combination of two objectives for each model in calibration and validation period

| Models | First calibration–validation procedure | | | | Second calibration–validation procedure | | | |
|--------|--|----------------|-------------------|----------------|---|----------------|-------------------|----------------|
| | Calibration (2013) | | Validation (2017) | | Calibration (2017) | | Validation (2013) | |
| | F ₁ | F ₂ | F ₁ | F ₂ | F ₁ | F ₂ | F ₁ | F ₂ |
| S1 | 0.44 | 0.40 | 0.31 | 0.50 | 0.32 | 0.42 | 0.47 | 0.62 |
| S2 | 0.40 | 0.41 | 0.30 | 0.48 | 0.30 | 0.42 | 0.45 | 0.62 |
| S3 | 0.40 | 0.41 | 0.34 | 0.46 | 0.30 | 0.42 | 0.44 | 0.65 |
| S4 | 0.38 | 0.38 | 0.42 | 0.36 | 0.28 | 0.39 | 0.40 | 0.45 |
| S5 | 0.36 | 0.40 | 0.43 | 0.31 | 0.32 | 0.34 | 0.37 | 0.51 |
| S6 | 0.19 | 0.28 | 0.17 | 0.33 | 0.15 | 0.26 | 0.22 | 0.47 |
| S7 | 0.20 | 0.26 | 0.17 | 0.33 | 0.14 | 0.25 | 0.20 | 0.48 |
| S8 | 0.14 | 0.20 | 0.14 | 0.28 | 0.13 | 0.21 | 0.33 | 0.26 |
| M1 | 0.43 | 0.39 | 0.33 | 0.48 | 0.31 | 0.42 | 0.46 | 0.63 |
| M2 | 0.19 | 0.28 | 0.17 | 0.35 | 0.15 | 0.25 | 0.24 | 0.50 |
| M3 | 0.19 | 0.27 | 0.16 | 0.34 | 0.14 | 0.25 | 0.21 | 0.48 |
| M4 | 0.18 | 0.27 | 0.26 | 0.25 | 0.14 | 0.23 | 0.20 | 0.53 |
| M5 | 0.18 | 0.27 | 0.26 | 0.25 | 0.14 | 0.22 | 0.40 | 0.36 |
| M6 | 0.19 | 0.27 | 0.17 | 0.33 | 0.15 | 0.22 | 0.23 | 0.50 |
| M7 | 0.17 | 0.27 | 0.16 | 0.33 | 0.15 | 0.22 | 0.21 | 0.52 |
| M8 | 0.14 | 0.20 | 0.14 | 0.29 | 0.13 | 0.20 | 0.22 | 0.36 |
| C1 | 0.22 | 0.14 | 0.21 | 0.29 | 0.17 | 0.22 | 0.23 | 0.45 |
| C2 | 0.08 | 0.11 | 0.18 | 0.25 | 0.11 | 0.18 | 0.11 | 0.42 |
| C3 | 0.07 | 0.10 | 0.18 | 0.25 | 0.10 | 0.17 | 0.10 | 0.41 |
| C4 | 0.07 | 0.10 | 0.17 | 0.20 | 0.09 | 0.12 | 0.14 | 0.33 |
| C5 | 0.07 | 0.10 | 0.17 | 0.25 | 0.09 | 0.10 | 0.19 | 0.21 |
| C6 | 0.06 | 0.09 | 0.18 | 0.23 | 0.09 | 0.15 | 0.10 | 0.40 |
| C7 | 0.06 | 0.09 | 0.18 | 0.21 | 0.09 | 0.14 | 0.08 | 0.39 |
| C8 | 0.05 | 0.08 | 0.20 | 0.21 | 0.09 | 0.08 | 0.21 | 0.28 |

The best combination of the two objectives on the optimal pareto front for each model is determined by point with the smallest distance to the origin of the axes of

Parameter sensitivity

The RAS method is used to analyze the parameter sensitivity and identifiability. The analysis results are shown in Figs. 6, 7 and 8. In general, the results of parameter sensitivity are similar when using data in two different periods. For each model set, the KS value of parameter λ is relatively consistent in different models. KS values of other parameters decrease with the increase of spatial parameter partition. When the spatial partition of the parameter is considered, the parameter for the spatial zone with large area has relatively large KS value and high sensitivity, since the spatial partition with a large area has more contribution to the spring discharge.

As pointed in Chang et al. (2017), when KS is less than 0.1, the ten cumulative curves are very close to the straight line and the parameter is insensitive and may be hard to be identified. In this paper, we also use this value to determine whether the parameter is sensitive or insensitive (red dash lines in Figs. 6, 7, 8). For the models S1–S8, λ and K are all sensitive even though three spatial partitions of K are

considered. For the models M1–M5 and C1–C5, the parameters (x or h_r , K_s , and K_q) of the routing reservoir are all sensitive except x and K_q in M1. It is quite surprising that even the simple lumped model M1 also has two insensitive parameters, x and K_q . This indicates the obvious structural defects of M1. The sensitivities of these two parameters are highly improved in M2 when the routing function is added. When considering two spatial parameter partitions of routing reservoir in M6 or C6, x (or h_r) and K_q in the first partition are insensitive. When further considering three spatial partitions, K_q in the third partition also get close to be insensitive. For the parameter S_{\max} , it is only sensitive in the lumped models or semi-lumped model in the first calibration period (data in 2013). When the spatial partition is considered, S_{\max} in each partition is always insensitive. The parameter S_{\max} is always insensitive in all models in the second calibration period (data in 2017) regardless of the consideration of spatial partition or not. This indicates that the spring discharge may not contain sufficient information to identify the spatial variability of the effective precipitation in this study site.

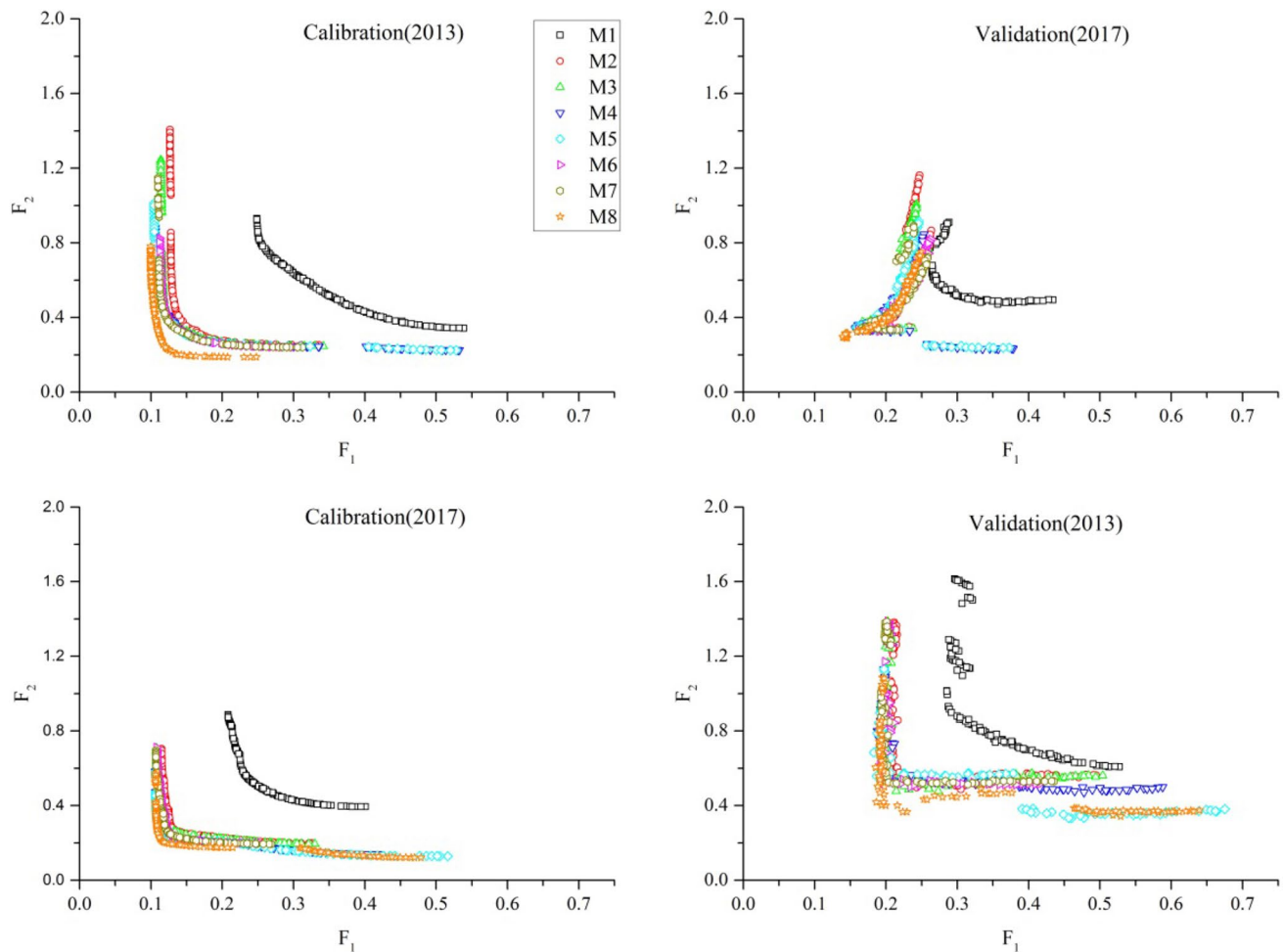


Fig. 4 Calibration and validation results of M1–M8 in two calibration–validation procedures

Discussion

Comparisons between lumped models and semi-lumped models

For the three different model structures, the analysis results all show the model performance improves when the simple routing function is added in the original lumped models. Especially, for the model M1, the parameter identifiabilities of x and K_q are also highly improved in M2 when the routing function is added. Therefore, the routing function is truly needed in the lumped model. This is mainly because there is an obvious lag time (about several hours) between the rainfall and discharge of spring S31, and the time step of the model (15 min) is much less than this long lag time. The lumped model with simple routing reservoir could hardly reflect this lag process. This is also the main reason that the simple time translation or triangular transfer function is often used in the simple lumped model to offset this lag time

in the previous studies to simulate the discharge of spring S31 (Chang et al. 2015, 2017; Yuan et al. 1996).

When the spatial distribution of each depression and conduit network is considered to discretize the lumped model, the model performances of semi-distributed models (S3, M3, and C3) show marginal improvement in contrast to lumped-routing models (S2, M2, and C2). In this paper, the main difference between the lumped and semi-lumped models is that the semi-lumped model could consider the different arrival times of water in three depressions in spring S31 due to the different distances between the sinkhole of each depression and spring. Their similar performances indicate this difference is not very important for the simulation of spring S31. This is mainly due to the relatively small catchment area and short conduit length. However, it should be noted that the location of the conduit network in the catchment of spring S31 is mainly speculated through the surface landform and distribution of the sinkholes which may still have a large difference from the distribution of the actual

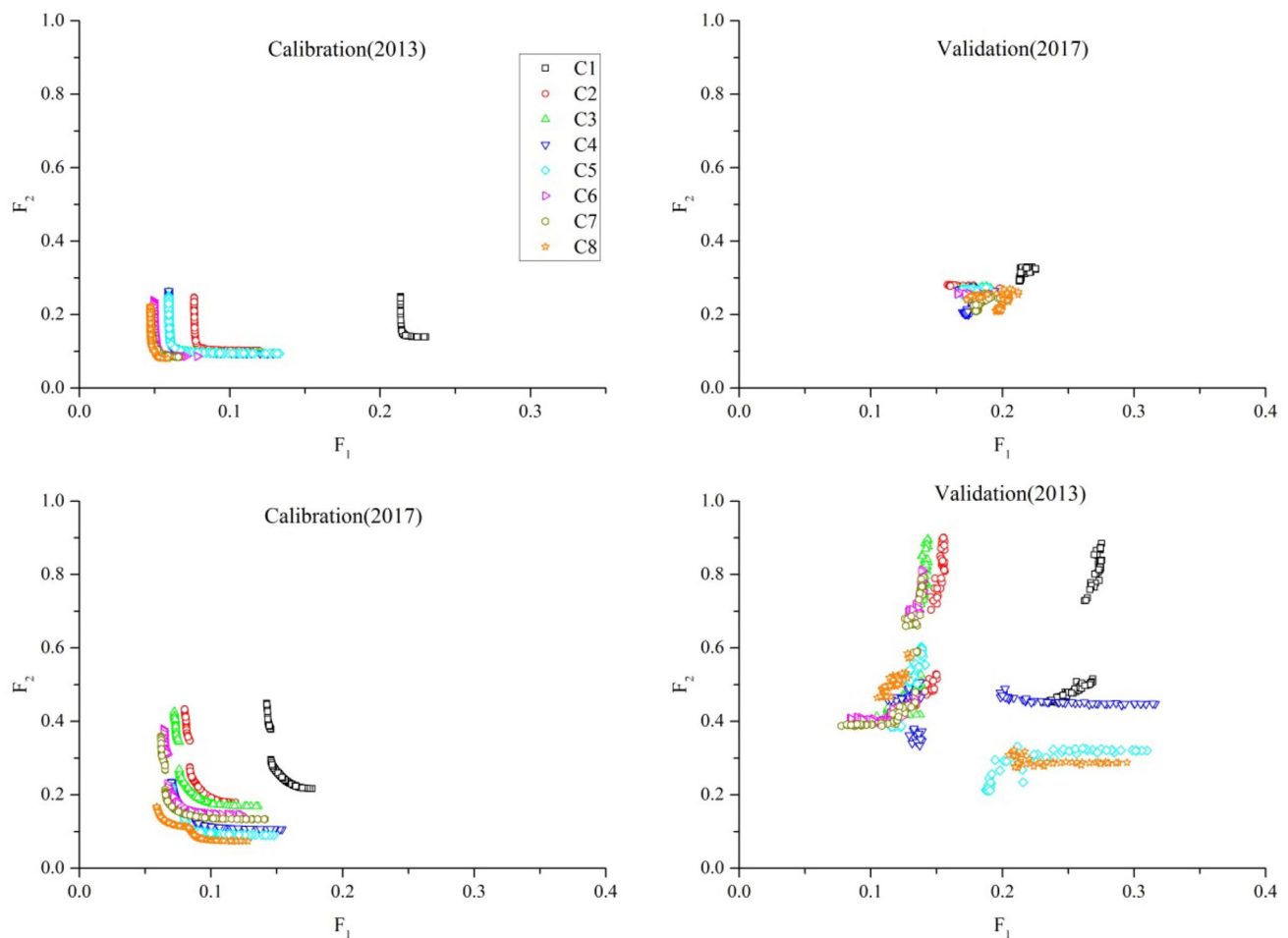


Fig. 5 Calibration and validation results of C1–C8 in two calibration–validation periods

conduit network. Meanwhile, the routing function used in this paper assumes that the celerity is always a constant, and in reality, this parameter may be different under different hydraulic conditions (Cholet et al. 2017). These factors may affect the comparison result between lumped and semi-lumped models.

Appropriate spatial complexity for different semi-lumped models

The appropriate spatial complexity of each semi-distributed model could be identified according to the comparison of model performance and parameter sensitivity. Theoretically, if the spatial parameter heterogeneity is truly needed, the model performance should be highly improved and the parameters in the model should be all identified by the available data (Chang et al. 2017).

For the semi-lumped model S3, its model performance is only highly improved when two partitions of routing reservoir are considered in S6. The sensitivity analysis indicates

that the parameters of S6 could all be finely identified. When further considering three partitions of routing reservoir or the partition of evapotranspiration reservoir, the model performance shows very limited improvement. Meanwhile, parameter sensitivity also indicates that the spatial partition of evapotranspiration reservoir is not needed. Therefore, the semi-distributed model S6 with two parameter partitions of routing reservoir is preferred among eight models (S1–S8) to simulate the discharge of this spring (Fig. 9).

For the semi-lumped model M3 or C3, the model performance only shows marginal or a very small improvement in the calibration period when the spatial parameter partition of routing or evapotranspiration reservoir is considered. The parameter sensitivity also pointed that there are always insensitive parameters when the spatial parameter partition is considered. Therefore, the spatial parameter partition is not necessary for the semi-lumped model M3 or C3. Given their similar performance and complexity (same parameter number) between lumped and semi-lumped models, M2 (or

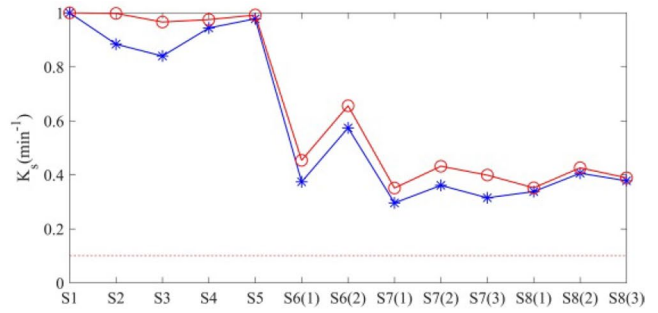
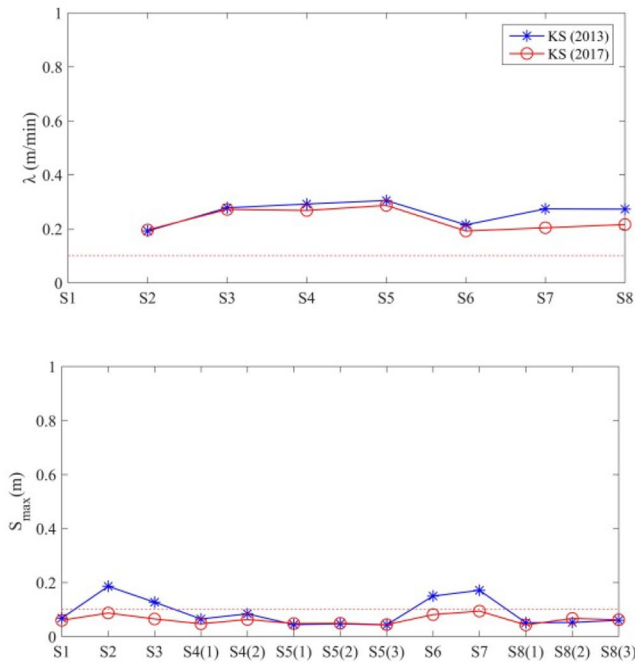


Fig. 6 Variations of parameter sensitivity (KS) from S1 to S8. When KS is lower than 0.1 (red dash line), the parameter is considered to be insensitive. The number in the parenthesis behind the model number

represents different spatial partitions, for example, S7(1), S7(2), and S7(3) represent the corresponding parameter in the first, second, and third partitions, respectively

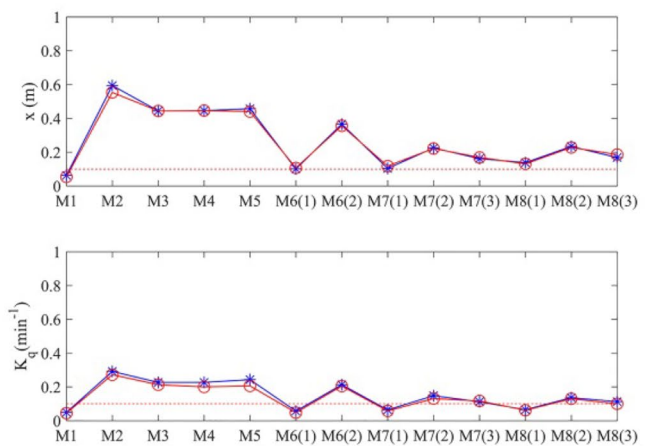
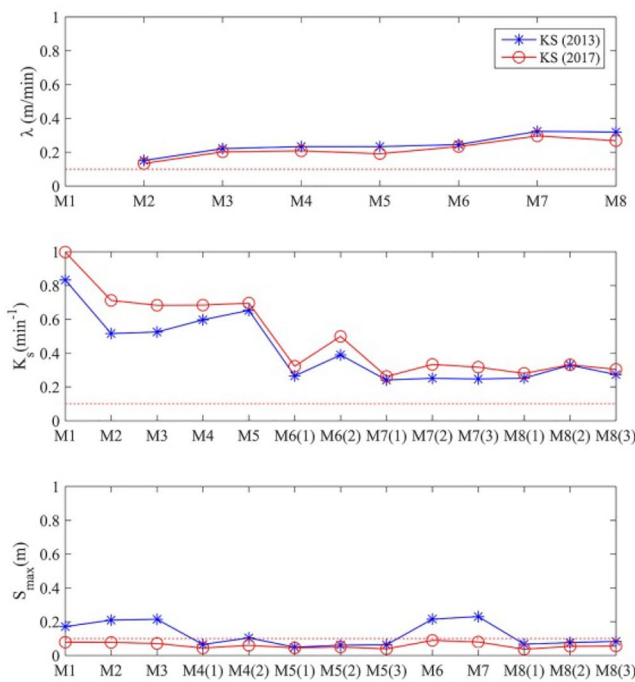


Fig. 7 Variations of parameter sensitivity (KS) from M1 to M8. When KS is lower than 0.1 (red dash line), the parameter is considered to be insensitive. The number in the parenthesis behind the

model number represents different spatial partitions, for example, M7(1), M7(2), and M7(3) represent the corresponding parameter in the first, second, and third partitions, respectively

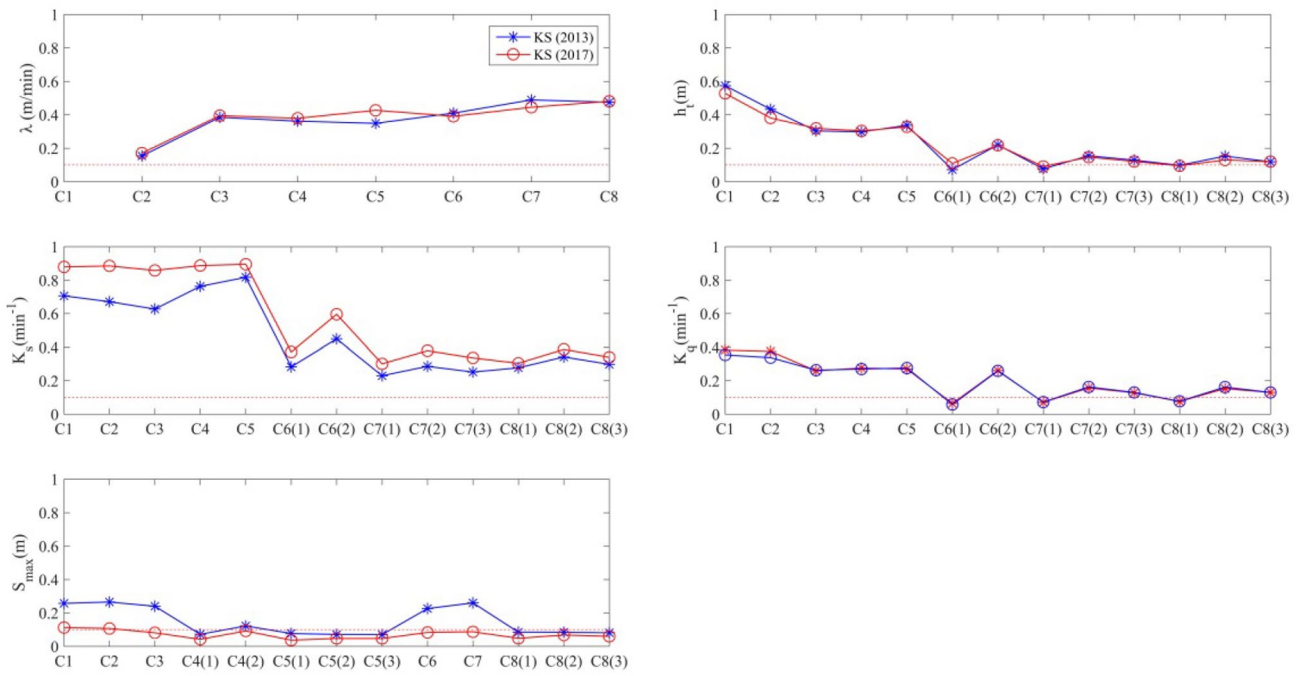
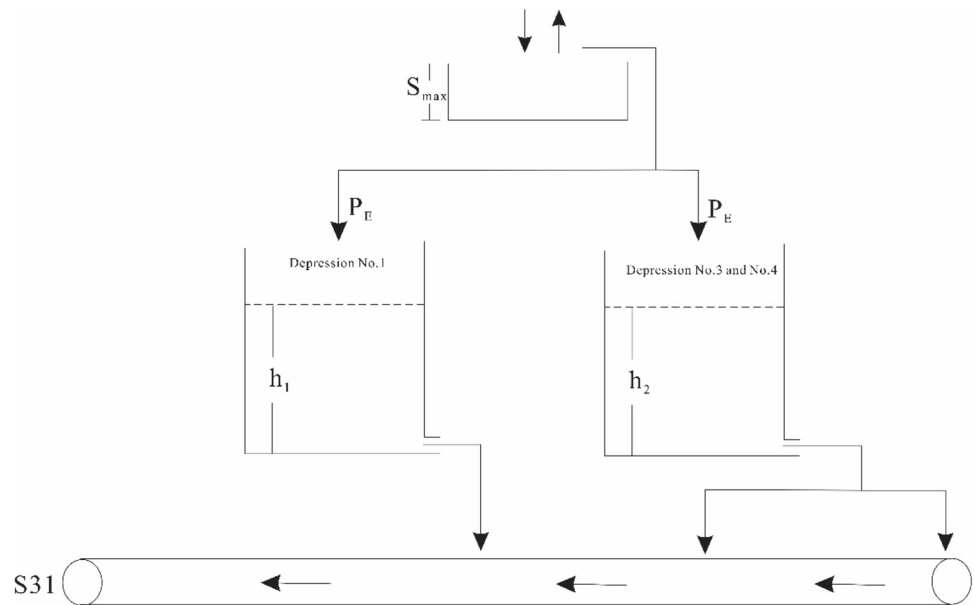


Fig. 8 Variations of parameter sensitivity (KS) from C1 to C8. When KS is lower than 0.1 (red dash line), the parameter is considered to be insensitive. The number in the parenthesis behind the model number

represents different spatial partitions, for example, C7(1), C7(2), and C7(3) represent the corresponding parameter in the first, second, and third partitions, respectively

Fig. 9 Model structure of S6



M3) and C2 (or C3) can both be considered as the appropriate complexity for each model structure.

In general, for three different semi-lumped models, only the performance of the simplest model S3 can be highly improved by the consideration of extra spatial parameter heterogeneity. The same situation was also pointed by Hellebrand and Bos (2008) when they used

two different models to simulate the 18 sub-basins of the Nahe basin. They found that only the performance of a simple model showed an improvement when the spatial parameter heterogeneity was introduced. The possible reason is that the simple model structure is too simple to capture the behavior of the catchment and the consideration of spatial parameter heterogeneity could add extra

flexibility to the model (Atkinson et al. 2003; Boyle et al. 2001). However, when the complex model structure is already flexible enough to simulate the spring behavior, extra consideration of spatial parameter heterogeneity can hardly highly improve the model performance (Das et al. 2008; Refsgaard and Knudsen 1996; Vansteenkiste et al. 2014). This indicates that whether extra spatial complexity for the lumped model is needed may have a strong relationship with the process complexity of the lumped model. If process complexity of the lumped model is sufficiently flexible to capture the behavior of catchment, extra consideration of spatial complexity in the model may be not needed, even for the karst aquifer, to simulate the spring discharge.

Comparisons among different model structures

The optimal pareto fronts of appropriate models for each model structure (S6, C2 and M2) are shown in Fig. 10 to compare their performances in two calibration–validation procedures. For simplicity, their simulated hydrographs

after calibration from 15 April to 14 June in 2013 are only represented in the paper (Fig. 11). It can be found that S6 provides the very similar performance to M2 after considering the spatial parameter heterogeneity of routing reservoir in S3 (Fig. 10; Table 3). Moreover, S6 provides a relatively shorter length of pareto front than M2, especially in the second calibration–validation procedure, indicating its less structure uncertainty. This indicates that the process deficiency of S3 relative to M2 could be effectively compensated by the increase of spatial complexity. It should be noted that although S6 has a similar structure to M2 (both contain two different parallel reservoirs), they have obviously different physical meanings. In S6, the different routing reservoirs are used to represent the behavior of different depressions (Fig. 9), whereas in M2 different routing reservoirs represent different hydrological processes (quick flow vs. slow flow). Moreover, S6 has one less parameter than M2. The distribution of recharge on two reservoirs is determined according to the area of depression in S6, whereas in M2 the recharge distribution is controlled by parameter x .

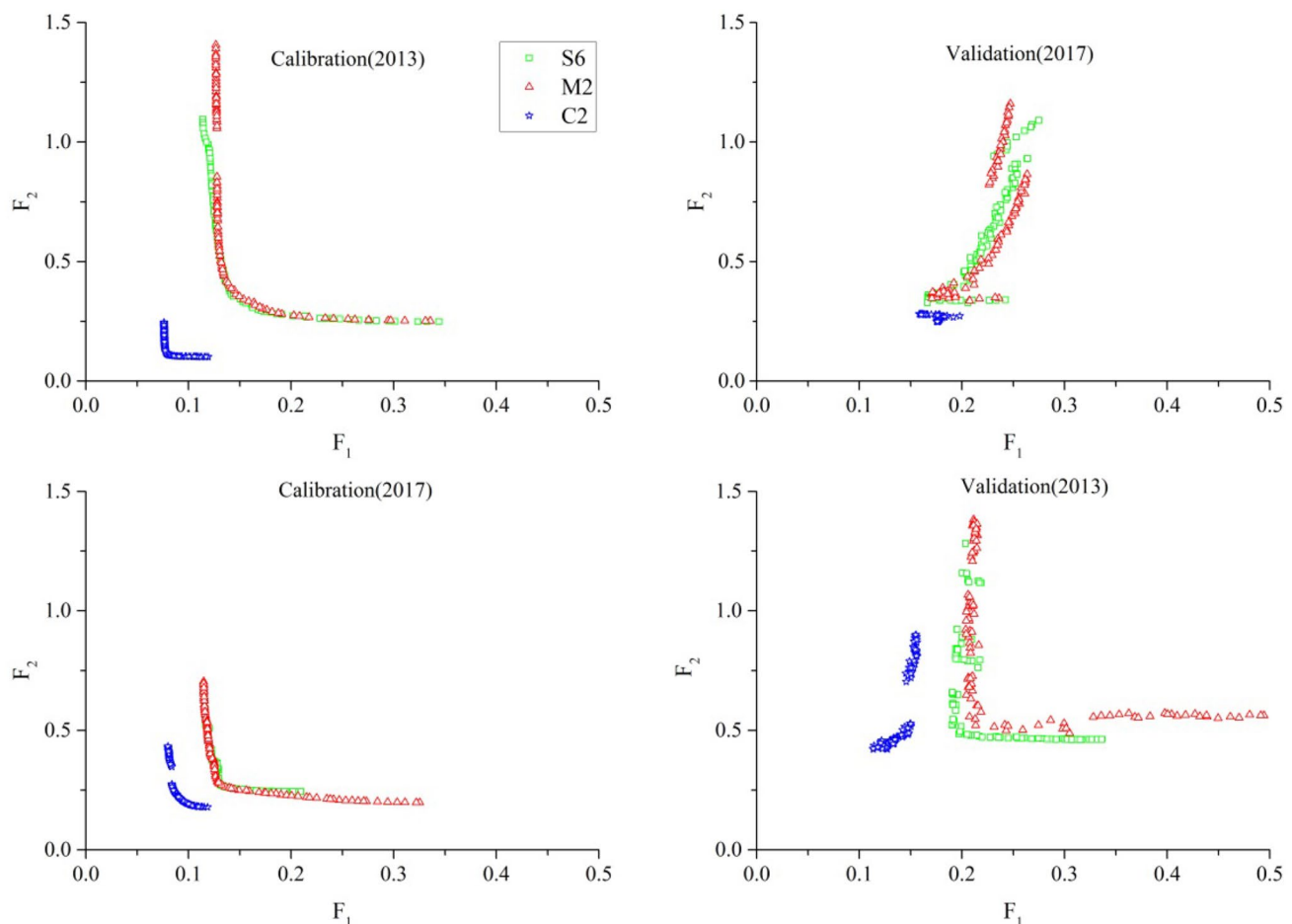


Fig. 10 Comparison of simulation results of S2, S6, M2, and C2 in two calibration–validation procedures

Fig. 11 Simulated hydrographs by three different models (S6, M2, and C2) from April 15 to June 14 in 2013 after the calibration

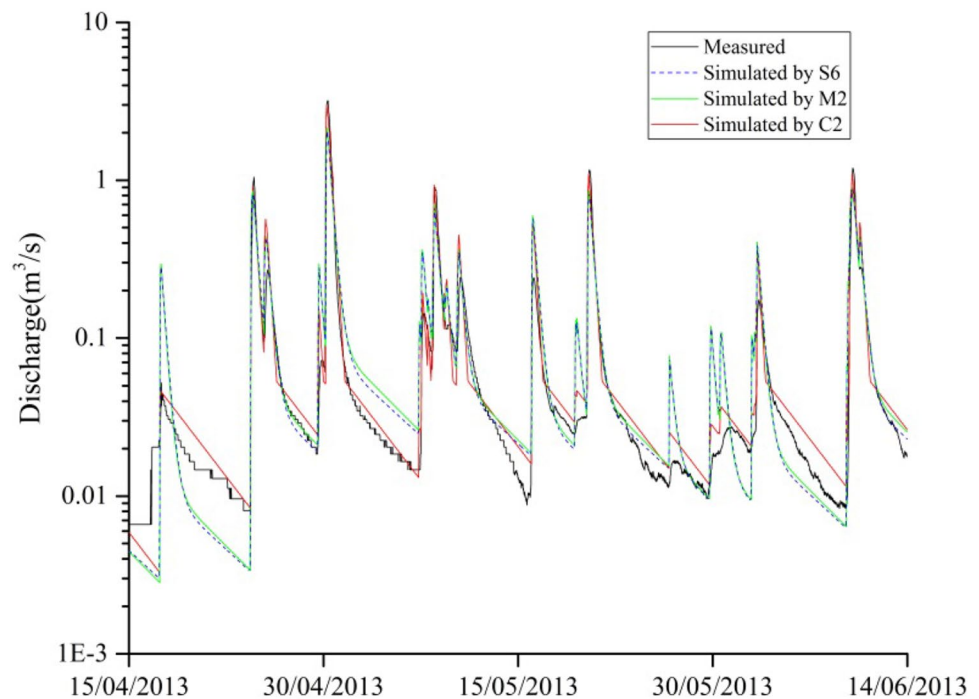


Figure 10 also shows that S6 or M2 has worse performance than C2. C2 can relatively finely capture the peak discharge and the recession curves especially the spring response under very low recharge events (Fig. 11). This result is consistent with the previous study (Chang et al. 2017) that the model structure of C2 is better than M2 to simulate this spring. The performance of M2 is not highly improved by the increase in spatial complexity. Meanwhile, even though the performance of S3 can be highly improved with the increase in spatial complexity, S6 still has worse performance than C2. The main process difference between M2 (or S6) and C2 is that upper outlet in C2 is threshold-driven. As pointed in Chang et al. (2017), two outlets in C2 should correspond to the discharge from point recharge and diffuse recharge, respectively, and the point recharge is threshold-driven. This structure is much closer to the actual hydrological process in the study site. This indicates that the increase of spatial parameter heterogeneity of S2 and M2 cannot effectively compensate this process deficiency which further supports the results of the previous study.

In general, C3 provides a better performance than other models. Given that C3 and C2 have similar performance and same complexity (parameter number), they can be both considered as the most appropriate model structure for this karst catchment. This result further supports the previous point that a simple reservoir with two different outlets is enough to capture the main behavior of spring discharge (Chang et al. 2017) and extra consideration of spatial parameter heterogeneity is not necessary. It also should be noted that there are no extra internal measurements in the study site. Therefore,

it is hard to diagnose whether these lumped or semi-lumped models are realistic or not to represent the internal hydrological process. For that reason, the lumped model might also be the model structures with the best performance but due to wrong reasons (Kirchner 2006).

Discussion of the relationship between process complexity and spatial complexity

The process complexity and spatial complexity are two different aspects that we should consider in the hydrologic models. Although they represent different physical meanings, they all, in essence, introduce new parameters (representing new hydrologic processes or spatial heterogeneity) in the model to improve the model performance. However, limited available data, such as rainfall-streamflow data, often support limited model complexity (parameter number) (Beven 1989; Chang et al. 2017; Jakeman and Hornberger 1993; Perrin et al. 2001). Therefore, there should be a trade-off between the process complexity and spatial complexity of a model theoretically when the available data are limited. If the process complexity of the lumped model is flexible enough, the extra consideration of spatial complexity should be not needed, such as M2 and C2.

When the process complexity of the model is not flexible enough to capture the behavior of catchment, the increase of spatial complexity could make up the deficiency of process complexity to improve the model performance to some degree, such as the comparison between S6 and M2.

However, it should be noted that the way to increase the spatial complexity of the model is often limited. They are often conducted by spatially distributing the runoff process based on different HRUs (hydrological response units) and connecting them in parallel or in series (Adinehvand et al. 2017; Barrett and Charbeneau 1997; Euser et al. 2015; Gao et al. 2014; Hartmann et al. 2012a; Savenije 2010; Seibert et al. 2003; Uhlenbrook et al. 2004). Therefore, these limited ways could not guarantee that the increase in spatial complexity could fully cover all the kinds of deficiencies of process complexity, such as the comparison between semi-distributed models of M2 and C2. If the spatial complexity cannot make up the deficiency of process complexity, even the distributed model may provide worse performance than the lumped model. This may be one possible reason that the lumped model shows a better overall performance than the distributed model in some studies (Reed et al. 2004). From this point of view, we should consider the process complexity of the model as a priority and the lumped model without consideration of spatial complexity should be flexible enough to simulate the discharge series of study catchment (even for the karst catchment). The distributed model should be only considered when more extra data, such as the internal measurements (heads), are available.

Conclusions

In this paper, we investigated the appropriate model structure for the simulation of a karst catchment from the aspect of spatial complexity. Three lumped models (S1, M1, and C1) from simple to complex structures were chosen to gradually increase their spatial complexities to establish the semi-distributed models by adding the routing function and considering spatial parameter heterogeneity. The performance comparison and parameter sensitivity were used to investigate appropriate spatial complexity for each model structure. And then, these different models were compared to explore the appropriate model structure for the simulation of a karst spring.

Our analysis results show that the performances of lumped-routing versions of three models (S2, M2, and C2) all highly improve by adding the routing function. However, when further considering the spatial complexity based on S2, M2, and C2, different models give different results. For the simplest model S2, its performance highly improves by considering two parameter partitions of the linear storage reservoir (S6). However, performances of M2 and C2 show very limited improvement when further considering extra spatial complexity. These results indicated that whether the model performance can be highly improved by the consideration of spatial complexity has a strong relationship with the process complexity. If the process complexity of the lumped

model is sufficiently flexible to capture the behavior of the catchment, the extra consideration of spatial complexity may be not needed.

The comparison results among different model structures indicate that S6 could provide the very similar performance to M2 after considering appropriate spatial parameter heterogeneity of S2. However, S6 and M2 still provide obvious worse performance than C2. The increase of spatial complexity of S2 or M2 cannot effectively make up its process deficiency relative to C2. Given that the semi-lumped model C3 has similar performance and complexity to C2, both models can be considered as the appropriate model structure for simulation of this karst catchment. There is no extra appropriate model which has relatively simpler process complexity and higher spatial complexity than C2 for the simulation of this spring. This result further verifies the previous point that a simple reservoir with two different outlets is enough to capture the main behavior of this karst spring (Chang et al. 2017).

Acknowledgements This research was supported by a grant from the National Natural Science Foundation of China (Nos.: 41730856, U1503282, 41602242, 41502260) and the National Key Research and Development Program of China (NO. 2016YFC0402809).

References

- Adinehvand R, Raeisi E, Hartmann A (2017) A step-wise semi-distributed simulation approach to characterize a karst aquifer and to support dam construction in a data-scarce environment. *J Hydrol* 554:470–481
- Atkinson SE, Sivapalan M, Woods RA, Viney NR (2003) Dominant physical controls on hourly flow predictions and the role of spatial variability: Mahurangi catchment, New Zealand. *Adv Water Resour* 26(3):219–235. [https://doi.org/10.1016/S0309-1708\(02\)00183-5](https://doi.org/10.1016/S0309-1708(02)00183-5)
- Bailly-Comte V, Borrell-Estupina V, Jourde H, Pistre S (2012) A conceptual semidistributed model of the Coulazou River as a tool for assessing surface water-karst groundwater interactions during flood in Mediterranean ephemeral rivers. *Water Resour Res*. <https://doi.org/10.1029/2010wr010072>
- Barrett ME, Charbeneau RJ (1997) A parsimonious model for simulating flow in a karst aquifer. *J Hydrol* 196(1–4):47–65. [https://doi.org/10.1016/S0022-1694\(96\)03339-2](https://doi.org/10.1016/S0022-1694(96)03339-2)
- Bentura PLF, Michel C (1997) Flood routing in a wide channel with a quadratic lag-and-route method. *Hydrol Sci J* 42(2):169–189. <https://doi.org/10.1080/02626669709492018>
- Beven K (1989) Changing ideas in hydrology—the case of physically-based models. *J Hydrol* 105(1–2):157–172. [https://doi.org/10.1016/0022-1694\(89\)90101-7](https://doi.org/10.1016/0022-1694(89)90101-7)
- Beven K (2012) Rainfall-runoff modelling: the primer. 15(1): 84–96
- Boyle DP et al (2001) Toward improved streamflow forecasts: value of semidistributed modeling. *Water Resour Res* 37(11):2749–2759
- Chang Y (2015) Analysis and simulation of the hydrological process of the karst aquifer with fracture-conduit dual structure. Phd Thesis, Nanjing University, China
- Chang Y, Wu J, Jiang G (2015) Modeling the hydrological behavior of a karst spring using a nonlinear reservoir-pipe model.

- Hydrogeol J 23(5):901–914. <https://doi.org/10.1007/s10040-015-1241-6>
- Chang Y, Wu J, Jiang G, Kang Z (2017) Identification of the dominant hydrological process and appropriate model structure of a karst catchment through stepwise simplification of a complex conceptual model. *J Hydrol* 548:75–87. <https://doi.org/10.1016/j.jhydrol.2017.02.050>
- Cholet C, Charlier JB, Moussa R, Steinmann M, Denimal S (2017) Assessing lateral flows and solute transport during floods in a conduit-flow-dominated karst system using the inverse problem for the advection–diffusion equation. *Hydrol Earth Syst Sci* 21(7):3635–3653. <https://doi.org/10.5194/hess-21-3635-2017>
- Clark MP et al., 2008. Framework for understanding structural errors (FUSE): a modular framework to diagnose differences between hydrological models. *Water Resour Res*. <https://doi.org/10.1029/2007WR006735>
- Coxon G, Freer J, Wagener T, Odoni NA, Clark M (2014) Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments. *Hydrol Process* 28(25):6135–6150. <https://doi.org/10.1002/hyp.10096>
- Das T, Bardossy A, Zehe E, He Y (2008) Comparison of conceptual model performance using different representations of spatial variability. *J Hydrol* 356(1–2):106–118. <https://doi.org/10.1016/j.jhydrol.2008.04.008>
- Eagleman, J.R., 1967. Pan Evaporation, Potential and Actual Evapotranspiration. *J Appl Meteorol* 6(3): 482–488. [https://doi.org/10.1175/1520-0450\(1967\)006%3C0482:PEPAAE%3E2.0.CO;2](https://doi.org/10.1175/1520-0450(1967)006%3C0482:PEPAAE%3E2.0.CO;2)
- Euser T et al (2013) A framework to assess the realism of model structures using hydrological signatures. *Hydrol Earth Syst Sci* 17(5):1893–1912. <https://doi.org/10.5194/hess-17-1893-2013>
- Euser T, Hrachowitz M, Winsemius HC, Savenije HHG (2015) The effect of forcing and landscape distribution on performance and consistency of model structures. *Hydrol Process* 29(17):3727–3743
- Fenicia F, McDonnell JJ, Savenije HHG (2008) Learning from model improvement: on the contribution of complementary data to process understanding. *Water Resour Res*. <https://doi.org/10.1029/2007WR006386>
- Fenicia F et al (2014) Catchment properties, function, and conceptual model representation: is there a correspondence? *Hydrol Process* 28(4):2451–2467. <https://doi.org/10.1002/hyp.9726>
- Fleury P, Plagnes V, Bakalowicz M (2007) Modelling of the functioning of karst aquifers with a reservoir model: Application to Fontaine de Vaucluse (South of France). *J Hydrol* 345(1–2):38–49. <https://doi.org/10.1016/j.jhydrol.2007.07.014>
- Ford DC, Williams P (2007) *Karst hydrogeology and geomorphology*. Wiley, Hoboken
- Freer J, Beven K, Ambrose B (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resour Res* 32(7):2161–2173. <https://doi.org/10.1029/95WR03723>
- Gao H, Hrachowitz M, Fenicia F, Gharari S, Savenije HHG (2014) Testing the realism of a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe, China. *Hydrol Earth Syst Sci* 18(5):1895–1915. <https://doi.org/10.5194/hess-18-1895-2014>
- Ghavidelfar S, Alvankar SR, Razmkhah A (2011) Comparison of the lumped and quasi-distributed Clark runoff models in simulating flood hydrographs on a semi-arid watershed. *Water Resour Manage* 25(6):1775–1790. <https://doi.org/10.1007/s11269-011-9774-5>
- Goldscheider N, Drew D (2007) *Methods in Karst hydrogeology: IAH: international contributions to hydrogeology*, 26. CRC Press, Cambridge
- Han J-C, Huang G-H, Zhang H, Li Z, Li Y-P (2014) Effects of watershed subdivision level on semi-distributed hydrological simulations: case study of the SLURP model applied to the Xiangxi River watershed, China. *Hydrol Sci J* 59(1):108–125. <https://doi.org/10.1080/02626667.2013.854368>
- Hartmann A, Lange J, Weiler M, Arbel Y, Greenbaum N (2012a) A new approach to model the spatial and temporal variability of recharge to karst aquifers. *Hydrol Earth Syst Sci* 16(7):2219–2231. <https://doi.org/10.5194/hess-16-2219-2012>
- Hartmann A et al (2012b) A multi-model approach for improved simulations of future water availability at a large Eastern Mediterranean karst spring. *J Hydrol* 468–469(0):130–138. <https://doi.org/10.1016/j.jhydrol.2012.08.024>
- Hartmann A et al (2013a) Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resour Res* 49(6):3345–3358. <https://doi.org/10.1002/wrcr.20229>
- Hartmann A et al (2013b) Process-based karst modelling to relate hydrodynamic and hydrochemical characteristics to system properties. *Hydrol Earth Syst Sci* 17(8):3305–3321. <https://doi.org/10.5194/hess-17-3305-2013>
- Hartmann A, Goldscheider N, Wagener T, Lange J, Weiler M (2014) Karst water resources in a changing world: review of hydrological modeling approaches. *Rev Geophys*, 52(3): 2013RG000443. <https://doi.org/10.1002/2013RG000443>
- Hellebrand H, Bos RVD (2008) Investigating the use of spatial discretization of hydrological processes in conceptual rainfall runoff modelling: a case study for the meso-scale. *Hydrol Process* 22(16):2943–2952
- Hornberger GM, Beven KJ, Cosby BJ, Sappington DE (1985) Shenandoah watershed study: calibration of a topography-based, variable contributing area hydrological model to a small forested catchment. *Water Resour Res* 21(21):1841–1850
- Hrachowitz M, Clark MP (2017) Opinions HESS: the complementary merits of competing modelling philosophies in hydrology. *Hydrol Earth Syst Sci Dis* 21(8):1–22
- Jakeman AJ, Hornberger GM (1993) How much complexity is warranted in a rainfall-runoff model? *Water Resour Res* 29(8):2637–2649. <https://doi.org/10.1029/93WR00877>
- Jiang GH, Yu S, Chang Y (2011) Identification of runoff in karst drainage system using hydrochemical method (in chinese). *J Jilin Univ Earth Sci Edn* 41(5):1534–1541
- Jukić D, Denić-Jukić V (2009) Groundwater balance estimation in karst by using a conceptual rainfall–runoff model. *J Hydrol* 373(3–4):302–315. <https://doi.org/10.1016/j.jhydrol.2009.04.035>
- Kirchner JW (2006) Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. *Water Resour Res*, 42(3). <https://doi.org/10.1029/2005wr004362>
- Kumar R, Samaniego L, Attinger S (2010) The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics. *J Hydrol* 392(1):54–69
- Lee G, Tachikawa Y, Takara K (2011) Comparison of model structural uncertainty using a multi-objective optimisation method. *Hydrol Process* 25(17):2642–2653. <https://doi.org/10.1002/hyp.8006>
- Lerat J et al (2012a) Do internal flow measurements improve the calibration of rainfall-runoff models? *Water Resour Res* 48:18. <https://doi.org/10.1029/2010wr010179>
- Lerat J, Perrin C, Andreassian V, Loumagne C, Ribstein P (2012b) Towards robust methods to couple lumped rainfall-runoff models and hydraulic models: a sensitivity analysis on the Illinois River. *J Hydrol* 418:123–135. <https://doi.org/10.1016/j.jhydrol.2009.09.019>
- Liu Z, Groves C, Yuan D, Meiman J (2004) South China Karst aquifer storm-scale hydrochemistry. *Ground Water* 42(4):491–499. <https://doi.org/10.1111/j.1745-6584.2004.tb02617.x>
- Moussa, R., 1996. Analytical hayami solution for the diffusive wave flood routing problem with lateral inflow. *Hydrol Process* 10(9):

- 1209–1227. [https://doi.org/10.1002/\(SICI\)1099-1085\(199609\)10:9%3C1209::AID-HYP380%3E3.0.CO;2-2](https://doi.org/10.1002/(SICI)1099-1085(199609)10:9%3C1209::AID-HYP380%3E3.0.CO;2-2)
- Nijzink RC et al (2016) The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological models. *Hydrol Earth Syst Sci* 20(3):1151–1176. <https://doi.org/10.5194/hess-20-1151-2016>
- Perrin C, Michel C, Andréassian V (2001) Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *J Hydrol* 242(3–4):275–301. [https://doi.org/10.1016/S0022-1694\(00\)00393-0](https://doi.org/10.1016/S0022-1694(00)00393-0)
- Pianosi F, Sarrazin F, Wagener T (2015) A Matlab toolbox for global sensitivity analysis. *Environ Model Softw* 70:80–85. <https://doi.org/10.1016/j.envsoft.2015.04.009>
- Reed S et al (2004) Overall distributed model intercomparison project results. *J Hydrol* 298(1):27–60. <https://doi.org/10.1016/j.jhydrol.2004.03.031>
- Refsgaard JC, Knudsen J (1996) Operational validation and intercomparison of different types of hydrological models. *Water Resour Res* 32(7):2189–2202. <https://doi.org/10.1029/96wr00896>
- Savenije HHG (2010) HESS Opinions “Topography driven conceptual modelling (FLEX-Topo)”. *Hydrol Earth Syst Sci* 14(12):2681–2692. <https://doi.org/10.5194/hess-14-2681-2010>
- Seibert J, Rodhe A, Bishop K (2003) Simulating interactions between saturated and unsaturated storage in a conceptual runoff model. *Hydrol Process* 17(2):379–390
- Singh VP, Woolhiser DA (2002) Mathematical modeling of watershed hydrology. *J Hydrol Eng* 7(4):270–292
- Smith MB et al (2004) The distributed model intercomparison project (DMIP): motivation and experiment design. *J Hydrol* 298(1–4):4–26. <https://doi.org/10.1016/j.jhydrol.2004.03.040>
- Tritz S, Guinot V, Jourde H (2011) Modelling the behaviour of a karst system catchment using non-linear hysteretic conceptual model. *J Hydrol* 397(3–4):250–262. <https://doi.org/10.1016/j.jhydrol.2010.12.001>
- Uhlenbrook S, Roser S, Tilch N (2004) Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. *J Hydrol* 291(3):278–296
- Vansteenkiste T et al., 2014. Intercomparison of five lumped and distributed models for catchment runoff and extreme flow simulation. *J Hydrol* 511(Suppl C): 335–349. <https://doi.org/10.1016/j.jhydrol.2014.01.050>
- Vrugt JA, Robinson BA (2007) Improved evolutionary optimization from genetically adaptive multimethod search. *Proc Natl Acad Sci USA* 104(3):708–711. <https://doi.org/10.1073/pnas.0610471104>
- Vrugt JA, Robinson BA, Hyman JM (2009) Self-adaptive multimethod search for global optimization in real-parameter spaces. *IEEE Trans Evolut Comput* 13(2):243–259. <https://doi.org/10.1109/tevc.2008.924428>
- Wagener T et al (2001) A framework for development and application of hydrological models. *Hydrol Earth Syst Sci* 5(1):13–26. <https://doi.org/10.5194/hess-5-13-2001>
- Wöhling T, Vrugt JA, Barkle GF (2008) Comparison of three multi-objective optimization algorithms for inverse modeling of vadose zone hydraulic properties. *Soil Sci Soc Am J* 72(2):305–319
- Yan C, Zhang W (2014) Effects of model segmentation approach on the performance and parameters of the hydrological simulation program—fortran (HSPF) models. *Hydrol Res* 45(6):893
- Yuan DX, Dai AD, Cai WT, Liu ZH, He SY, Mo XP, Zhou SY, Lao WK (1996) Karst water system of a peak cluster catchment in south china’s bare karst region and its mathematic model. Guangxi Normal University Publishing House, Guilin

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.