Identifying more realistic model structures by electrical conductivity observations of the karst spring

Yong Chang¹, Andreas Hartmann^{2,3}, Ling Liu⁴, Guanghui Jiang^{5,6}, and Jichun Wu⁷

¹ School of Earth Science and Engineering, Hohai University, Nanjing 210098, China.

² Chair of Hydrological Modeling and Water Resources, Freiburg University, Freiburg, 79098, Germany.

³ Department of Civil Engineering, University of Bristol, Bristol, BS8 1TR, United Kingdom.

⁴ Geological Survey of Jiangsu Province, Nanjing 210018, China.

⁵ Karst Dynamics Laboratory, Institute of Karst Geology, Guilin 541004, China.

⁶ International Karst Research Center auspices of UNESCO, Guilin 541004, China.

⁷ Department of Hydrosociences, School of Earth Sciences and Engineering, Nanjing University, Nanjing 210093, China.

Corresponding authors: Yong Chang (wwwkr@163.com) and Guanghui Jiang (bmnxz@126.com)

Key Points:

- A new framework to integrate EC observations into lumped karst hydrologic models is presented.
- Spring EC shows a high potential for more realistic modeling.
- Incorporation of spring EC in models can reduce parameter uncertainty concerning concentrated flow in the conduits.

Abstract

Electrical conductivity (EC) of karst spring discharge has always been a fundamental variable to characterize karst aquifers. However, to incorporate EC into the hydrologic modeling is challenging but has a huge potential since EC observations are widely collected. In this paper, we presented a new framework to integrate EC into lumped karst hydrological models for model structure identification and parameter uncertainty reduction. Our framework is tested in a small, well-instrumented karst catchment near Guilin city (China) where EC dynamics are mostly controlled by the dissolution of carbonate rock and dilution by event water. Four karst models with different structures were equipped to consider the linear growth of EC with the carbonate rock dissolution and its mixing within the karst system. Applying a parameter estimation framework that accounts for uncertainty in discharge and EC simulations, we find that all hydrologic models obtain similar performances concerning spring discharge. However, their performances in simulating spring EC strongly differ permitting us to identify the most realistic model out of the four model structures. Our analysis further shows that including EC into parameter estimation reduces parameter uncertainty concerning concentrated flow in the conduits and helps to estimate immobile water storage in the aquifer. Consequently, our new framework to include EC in karst models provides new directions for the more realistic simulation that can easily be transferred to other karst aquifers where EC dynamics are controlled by rock dissolution and dilution by event water.

Keywords: Karst spring, electrical conductivity, model structure identification, parameter uncertainty

1 Introduction

Karst aquifers are estimated to provide drinking water to around 10%-25% of the world population [Stevanović, 2019]. The most significant characteristic of karst aquifers is their high heterogeneity and anisotropy caused by the long-time dissolution of soluble rocks [Ford and Williams, 2007]. Large fractures or conduits often exist in karst aquifers which create rapid groundwater flow, whereas, at the same time, slow flow occurs in the small fissures of the rock matrix [Goldscheider and Drew, 2007]. This typical dual structure leads to complex internal groundwater flow characteristics. At the shallow subsurface, a highly weathered layer with larger permeability and porosity may develop beneath the surface which can temporarily store the water and strongly change the recharge process to the phreatic zone [Williams, 1983; 2008]. Due to their complex internal structures and hydrologic processes, simulating these karst aquifers still poses a challenge even for experienced modelers.

A problem we often face when applying the rainfall-runoff model in a specific site is to select an appropriate model structure. Generally, the model structure should be consistent with the modeler's perceptions from the field investigation to get the right answers for the right reasons [Kirchner, 2006]. The most ideal situation is that all perceptions of modelers can be incorporated into the model [Enemark *et al.*, 2019]. However, a fundamental paradox always exists between the incorporation of the more complex hydrologic process into the model and the need to define extra model parameters that in turn bring the problem of parameter identifiability when only discharge data are available [Chang *et al.*, 2017; Jakeman and Hornberger, 1993; Ye *et al.*, 1997]. Therefore, parsimonious model structures are often favored to simulate the discharge, which is often too simple to provide realistic simulations, especially for karst systems [Hartmann *et al.*, 2017a]. Consequently, they may not deliver reliable predictions outside the hydrologic conditions that they were calibrated for [Kuczera and Mroczkowski, 1998]. To establish more realistic model structures, one way is to exploit more information of spring discharge through extra evaluation objectives, such as flow duration curve or autocorrelation function [Hartmann *et al.*, 2013b; Moussu *et al.*, 2011] or using extra tools to further explore the internal process [Bittner *et al.*, 2020; Duran *et al.*, 2020]. Another promising way is to incorporate auxiliary data into the model that can provide extra internal

information, i.e. flow path, residence time or runoff components, which are not involved in the discharge, to constrain the internal behavior of models to be more realistic and reduce the risk of selecting wrong model structures [Fenicia *et al.*, 2008; Hartmann *et al.*, 2017a; Mudarra *et al.*, 2019; Son and Sivapalan, 2007; Vache and McDonnell, 2006]. Moreover, the auxiliary data can also be helpful to further identify some model parameters and reduce the uncertainty of simulated discharge since their strong links to the internal runoff process [Hartmann *et al.*, 2017a].

The spring outlet is often the main monitoring point for the karst aquifers. Apart from spring discharge, the hydrochemical fluctuations of karst springs under different recharge events also can give effective insights into the internal hydrological process of karst aquifers [Aquilina *et al.*, 2006; Dreiss, 1989; Liu *et al.*, 2004; Mudarra and Andreo, 2011; Ravbar *et al.*, 2011]. Several hydrochemical indexes, such as the stable water isotopes [Zhang *et al.*, 2019], DOC (dissolved organic carbon) [Charlier *et al.*, 2012], Cl^- , SO_4^{2-} and NO_3^- [Hartmann *et al.*, 2013a; Hartmann *et al.*, 2017a; Husic *et al.*, 2019], have already been incorporated into the hydrologic models to prove their abilities to further identify the model structure and reduce the parameter uncertainty. However, currently obtaining concentrations of these ions is still relatively labor-consuming and costly particularly in remote places because most of them cannot be measured automatically and manual sampling at coarser temporal resolutions (weekly or biweekly) is applied [Goldscheider and Drew, 2007, Mudarra and Andreo, 2011], which limits their value for model improvement even for well-equipped catchments. Although DOC can also be measured by natural fluorescence continuous measurements using a field fluorometer (GGUN-FL30, Albillia, Neuchatel, Switzerland) [Schnegg, 2003], a careful calibration for the fluorometer using the DOC measurements is needed which may not provide very stable and reliable estimation results [Goldscheider and Drew, 2007]. Hence, there is a need to find a more economical, stable and easily-measured alternative allowing continuous monitoring at a high temporal resolution that can also be used to promote the model development and evaluation.

Compared to the specific hydrochemical ions, electrical conductivity (EC) of karst springs can be measured easily at a high temporal resolution using the automatic instrument with a low cost and it has already been monitored in many springs as a fundamental index [Jourde *et al.*, 2018;

1107 *Meus et al.*, 2014]. Moreover, spring EC always exhibits high fluctuations under storms making
1108 it a potential indicator for the internal behavior of the aquifer. Typically, the behavior of spring
1109 EC is controlled by the dissolution of carbonate rock and dilution by stormwater which contains
1110 three phases [*Goldscheider and Drew*, 2007; *Ravbar et al.*, 2011]: (1) at the beginning of a
1111 storm, EC often stays constant or show a temporary rise caused by flashing of ‘old water’,
1112 which is usually denoted as ‘piston effect’; (2) later, EC would decrease sharply indicating the
1113 arrival of unsaturated ‘new rainwater’ which originates from the concentrated rapid flow; (3)
1114 finally, EC increases slowly because of an increasing proportion of water in the storage and
1115 increasing amount of time for carbonate dissolution until it returns to the previous state or
1116 interrupts by next storms. Undoubtedly, the EC dynamic is strongly related to transport paths
1117 and transit times of water through the aquifer. For instance, a flat EC-graph often indicates low
1118 karstification and good mixing of water from different sources with an absence of preferential
1119 flow passages (i.e. conduit flow); on the contrary, a sharp EC peak usually means the existence
1120 of rapid concentrated flows and high karstification [*Goldscheider and Drew*, 2007]. Therefore,
1121 the spring EC is widely used to distinguish different types of karst aquifers [*Torresan et al.*,
1122 2020], identify the runoff composition [*Massei et al.*, 2007] or investigate the internal behavior
1123 combining with other ions [*Liu et al.*, 2007; *Mudarra et al.*, 2014; *Toran and Reisch*, 2013;
1124 *Winston and Criss*, 2004]. Furthermore, *Duran et al.* (2020) used spring EC to check the
1125 consistency between the internal flow dynamic in the model and the real aquifer. However, to
1126 our best knowledge, no research effort has yet considered the benefits of incorporating EC into
1127 the evaluation and parameter calibration of hydrologic models due to the relatively complex
1128 nature of EC. Given the advantage of easy measurement of spring EC at high temporal
1129 resolution with relatively little effort, exploring an effective method to integrate EC into the
1130 model and investigating its value to model evaluation has a huge potential to be widely applied
1131 in karst aquifers. In addition, the previously used environmental tracers (DOC, Cl^- , SO_4^{2-} and
1132 NO_3^-) in the models mainly focus on the infiltration process in the soil-unsaturated zone or the
1133 hydrological process in dry periods, whereas spring EC is expected to give more information
1134 about the fast flow. Understanding the value of spring EC for model identification and
1135 parameter estimation would therefore be very helpful for the modelers to obtain more realistic
1136 simulations.

Generally, EC is not related to a specific ion directly but represents a measure of the ability of water to pass an electrical current, which depends on the amount of total dissolved solids in the water [Massei *et al.*, 2007], and its changes could be either caused by the concentration of solutes or component ratios (water type) [Marandi *et al.*, 2013]. However, for most karst aquifers particularly without strong inferences of human activities and other non-carbonate strata, the hydrochemical composition of karst water is relatively stable and mainly dominated by the calcium carbonate equilibria. Therefore, we can assume that the fluctuation of spring EC is mainly caused by the dissolution process of carbonate rocks. This assumption is reasonable because in many karst aquifers Ca^{2+} , HCO_3^- and EC always exhibit a good correlation even a nice linear relationship in some aquifers (including the study site in this paper) [Liu *et al.*, 2004; Liu *et al.*, 2007; Torresan *et al.*, 2020; Zeng *et al.*, 2015]. We can consider the simulation of spring EC by adding the rock dissolving process and solute mixing process in the lumped rainfall-runoff model. In this paper, we presented a new framework to integrate EC into karst hydrologic modeling and investigate its contribution to obtaining more realistic simulations. We test this framework at a catchment with high purity of carbonate rock and limited human activities, where the fluctuation of EC is dominated by the calcium carbonate equilibria and has a good linear relationship between EC and calcium (or bicarbonate). The aim of this paper is twofold: (1) to identify realistic model structures with a multi-variate parameter estimation approach; (2) to quantify the value of spring EC to model structure identification and parameter uncertainty reduction.

2 Study site

The Yaji karst experimental site is located in the southeast of Guilin city, China. It is a typical karst aquifer developed in a thick formation of upper Devonian pure limestone. Its degree of karstification is high and many typical karst features, such as caves, shafts, sinkholes, karren fields, etc., are present across the study area. Karst development is mainly controlled by NEE-oriented fractures, and caves or conduits are all developed along this direction. Field investigations indicate that the epikarst zone is widely distributed beneath the surface with a depth of about 3-10 m. The climate in this region is a typical subtropical monsoon with an average temperature of 18.8 °C and annual precipitation of 1,915 mm. The main rainy season

is from April to August, during which 75% of annual precipitation occurs, and storms are frequent while snow events seldom occur. The vegetation in this area is mainly secondary shrub due to the massive deforestation in the last century and the coverage is more than 80%.

S31 is the biggest spring in this area and has been the focus of previous investigations [Chang *et al.*, 2017, 2019; Liu *et al.*, 2004a]. Its catchment is mainly recharged by direct infiltration of rainfall (no sinking streams). According to previous tracer tests, the spring catchment is about 1 km² (Fig. 1) and contains three depressions that accommodate preferential infiltration (Nos. 1, 3 and 4, similar to dolines with sinkholes in them) [Yuan *et al.*, 1996]. At each depression, there are several sinkholes at the bottom that directly connect to the spring. The epikarst zone is ubiquitous in this catchment and many epikarst springs drain towards the depression.

Fig. 1 a) Location of the Yaji experimental site; b) locations of the North railway station (NRS) and Institute of Karst Geology (IKG) from where rainfall data obtained for gap-filling of the rain gauge in the study site; c) catchment area of spring S31.

According to the field investigation and previous studies, the conceptual model of the hydrological process in this spring catchment can be described as follows [Chang *et al.*, 2019]. During the recharge events, the rainwater firstly stores in the vegetation and soil compartments, some of the water returns to the atmosphere by evapotranspiration and the extra rainwater recharges the epikarst zone directly. Due to much higher porosity and permeability of epikarst in contrast to the beneath carbonate rock, the water can be stored temporarily in the epikarst zone. A fraction of epikarst water slowly recharges the saturated fissured matrix below through small fractures (diffuse recharge) whereas the remaining fraction of the epikarst water can induce lateral flow towards sinkholes at the bottom of the depressions (point recharge). Parts of the lateral flow is discharged by epikarst springs and flow towards the sinkholes, as well. In the phreatic zone, the saturated fissured matrix and large conduits receive both diffuse recharge and rapid point recharge, and the water is finally drained by spring S31.

Spring discharge is measured at a time resolution of 15 minutes at a combination of rectangular weirs and there is a rain gauge located near spring S31 with a precision of 0.2 mm (Fig.1). A HOBO weather station also near spring S31 records meteorological data, such as air temperature, humidity and wind speed. During the periods from 2017 to 2019, there are two missing intervals of rainfall data (14.05.2018 – 31.07.2018 and 29.04.2019 – 31.07.2019) due to the instrument malfunction which were gap-filled by two nearby climatic stations, the North railway station (NRS, Fig.1), about 11 km away, and the Institute of Karst Geology (IKG), about 5.5 km away. Note that there is a recorded storm with 157.9 mm/day on June 21, 2018 in NRS, but the measured discharge peak under this storm is just 0.19 m³/s indicating a severe overestimation of this precipitation using the gap-filled data from NRS (Fig.2). Spring EC is monitored continuously by HOBO salt conductivity data logger with a temporal resolution of 15 minutes (corrected for 25°C) starting from May 11, 2017. Some data gaps of EC are mainly caused by drying-out of the spring (grey boxes in Fig.2), transducer malfunction (09.08.2017 – 10.09.2017), or unreliable data because the EC transducer was above the water level in dry periods.

In general, this spring exhibits a typical EC dynamic often observed in high-karstified aquifers with a sharp plunge during storms and a slow recovery afterward. The initial EC rising before it drops ('piston effect') is not obvious during the monitoring period. A higher discharge peak leads to a higher drop of EC indicating more unsaturated rapid infiltration water. For the small recharge events with a discharge peak lower than $0.50 \text{ m}^3/\text{s}$, EC shows a marginal change or no change. The obvious seasonal variation of spring EC is not observed since most of the recharge events happen in the rainy season. We find the initial spring EC is always much higher in the first storm after a long dry period than the following time and EC cannot return to this high value after the rainfall. This phenomenon is obvious in 2018 and 2019 which is possibly caused by the flushing of the 'old water', similar to the 'piston effect' often observed at the beginning of storms. During the rainy season in 2018 and 2019, the maximal value of EC is close and around $475 \text{ }\mu\text{S}/\text{cm}$, however, its maximal value in 2017 is much higher and around $498 \text{ }\mu\text{S}/\text{cm}$. This discrepancy may be mainly caused by the instrument drift because the maximal value of EC always close to $475 \text{ }\mu\text{S}/\text{cm}$ according to the previous measurements [Jiang, 2010].

Due to the high purity of limestone and little human activities in this catchment, the dissolved ions are mainly calcium and bicarbonate, and the concentration of other ions, such as K^+ , Na^+ , Mg^{2+} , Cl^- and SO_4^{2-} are relatively low, so that EC always has a good linear relationship with the concentration of Ca^{2+} and HCO_3^- [Liu *et al.*, 2004a, 2004b; Jiang, 2010]. The concentration of NO_3^- in this spring was just 0.088 mg/L according to a test result in 2012 [Wang *et al.*, 2014]. Therefore, it is reasonable to assume that the dissolution process of carbonate rock dominates EC behavior in the study site.

Fig. 2 The observed spring discharge and EC from January 1, 2017 to December 31, 2019. Due to the malfunction of rain gauge in the study site, the data gap between 14 May and 31 July in 2018 is from the North railway station (NRS) (about 11km away, Fig. 1) and between 29 April and 31 July in 2019 is from Institute of Karst Geology (IKG, about 5.5km away, Fig. 1). The grey boxes with ‘No discharge’ and ‘No EC’ mean dying-out of the spring in these periods.

3 Methodology

3.1 Workflow

To investigate the value of spring EC for model structure identification, four typical model structures, which were already applied to other karst modeling studies, were chosen to simulate spring discharge. To account for uncertainty in model structure definition and observations, we developed a Monte Carlo framework. Firstly, our framework performs a performance test of the different models to identify those model structures that can simulate both discharge and EC

simultaneously and adequately. Secondly, by iteratively applying different performance criteria the framework allows for quantifying the value of spring EC to parameter uncertainty reduction (Fig 3). Finally, the screened parameter sets were used in separate validation periods to test the stability of model performances.

Fig. 3 Schematic representation of the workflow in this study.

3.2 Different model structures for simulating discharge

Chang et al. (2017) discussed the best model structures to simulate this spring but only using the discharge for model evaluation. Through model performance comparison and parameter sensitivity analysis, their results indicated that a simple linear reservoir with two outlets is enough to simulate the spring discharge while avoiding over-parameterization. Based on *Chang et al. (2017)*'s research, we chose the best model structure from *Chang et al. (2017)* and three more complex model structures representing the heterogeneity of karstic groundwater flow in different ways [*Fleury et al., 2007; Geyer et al., 2008; Hartmann et al., 2012; Moussu et al., 2011*]. All models were equipped with transport routines accounting for EC dynamics through carbonate rock dissolution and solute mixing.

Fig. 4 Structures of the proposed different models. (P: precipitation, ET: evapotranspiration, EVP: evapotranspiration reservoir, SLOW: slow reservoir, FAST: fast reservoir, EPI: Epikarst reservoir, Storage: passive storage)

Figure 4 shows the structures of four selected models (M1~M4). All models have the same evapotranspiration reservoir (EVP) to calculate the effective precipitation. The potential evapotranspiration is calculated according to *Eagelman's* method (1967) and the actual ET is assumed to be linear to potential ET and saturation degree of the EVP reservoir [*Chang et al.*, 2015] (Table 1). When the evapotranspiration reservoir is full, extra precipitation converts to effective precipitation to recharge the reservoir below.

Among these four models, M1 is the most appropriate model structure pointed by *Chang et al.* (2017) to simulate the spring discharge. This model contains one linear reservoir but with two outlets of different height to represent the fast flow and slow flow respectively. The upper outlet is triggered when the water level is higher than a threshold. The spring discharge is the sum of two outlets. In M2, two separate linear reservoirs are used to represent the behavior of the fast and slow flow. The effective precipitation first recharges the slow reservoir, and when the slow reservoir is full, extra effective precipitation recharges the fast reservoir directly. Besides, the passive storage below the outlet is considered in the fast reservoir just for the EC simulation (introduced in the next section). It should be mentioned that M2 has some differences with the

classic two-parallel-reservoirs model proposed by *Fleury et al.* (2007) or the similar structure in *Chang et al.* (2017)'s work, which uses a partitioning parameter to divided the effective recharge into two parts recharging the slow and fast reservoirs respectively. However, for M2 in this paper, the fast reservoir receives the extra recharge directly at each time step when the slow reservoir is full without using the partitioning parameter. This structure is chosen because it works much better than the classic two-parallel-reservoirs model in the study site. In contrast to M2, M3 also considers two separate reservoirs but with different connections that the slow reservoir recharges the fast reservoir.

M4 adds an extra linear epikarst reservoir (EPI) above the fast and slow reservoirs to represent the behavior of epikarst zone. The epikarst reservoir has a limited capacity and the discharge of the lower outlet recharges the slow reservoir. When the epikarst reservoir is full, the extra effective precipitation recharges the fast reservoir directly. Due to the inclusion of EPI, one more parameter to simulate the slow flow is needed in contrast to the other three models. However, we found adding this extra parameter cannot improve the simulation of spring discharge and baseflow significantly, which is consistent with *Chang et al.* (2017)'s study that one linear coefficient is sufficient to get a good simulation of low flows of this spring. To avoid over-parameterization, we do not consider the routing ability of the slow reservoir in M4 and the reservoir drains all the recharge it receives at each time step. That is, no extra hydrologic parameter is considered in the slow reservoir. This slow reservoir is just used for EC simulation. After the adjustment, M4 also has four parameters for the discharge simulation. It should be noted that M3 and M4 have very similar structures (two effective routing reservoirs), but with different physical meanings and hypothesizes. Apart from the fast reservoir and evapotranspiration reservoir, another reservoir in M3 represents the behavior of slow flow, whereas in M4 it indicates the epikarst zone. Therefore, when they are used for the hydrochemical simulation, the excess input in the slow reservoir in M3 cannot mix with the water in the slow reservoir before it recharges the fast reservoir because the slow and fast reservoirs are two completely separated systems, however, the excess input in EPI experiences the hydrochemical process in the epikarst reservoir at each time step before it flows out.

From M1 to M4, they all contain four parameters. For the convenient analysis after, the same symbols were used in different models by merely considering their physical meanings. S_{max} is the maximal storage of the evapotranspiration reservoir. K_f and K_s control the discharge of fast and slow flow respectively in different models. h_1 is a threshold to control the occurrence of fast flow. Elaborations of the models' parameters and governing equations are all summarized in Table 1 and their feasible ranges are shown in Table 2, which are determined according to *Chang et al. (2017)*'s work.

[Table 1 and Table 2]

3.3 Simulation of spring EC

The EC enrichment through the dissolution of carbonate rock and mixing processes was added to the four models. Given the evapotranspiration reservoir is mainly used to calculate the effective precipitation, we assume that it is not involved in the rock dissolution and all dissolution processes occur in other reservoirs that represent subsystems within the carbonate rock. For each reservoir, we assume EC increases linearly with the rock dissolution and the solute in the reservoir can completely mix at each model step. Similar to previous work including environmental traces in lumped hydrological models, complete and instantaneous mixing is assumed in each reservoir [*Fenicia et al., 2010; Hartmann et al., 2013b; Vache and McDonnell, 2006*]. Then, the change of EC in the reservoir can be described by the following equation:

$$\begin{cases} \frac{d(C_{EC}(t) \times Vs(t))}{dt} = \sum C_{EC}^{in}(t) \times Q^{in}(t) + D_{EC} \times Vs(t) - \sum C_{EC}(t) \times Q^{out}(t) \\ C_{EC}(t) = \max(C_{EC}(t), EC_{max}) \end{cases} \quad (1)$$

where C_{EC} and C_{EC}^{in} are the EC value and input EC in the reservoir ($\mu\text{S}/\text{cm}$), D_{EC} is the average changing rate of EC ($(\mu\text{S}/\text{cm})/\text{min}$) caused by the rock dissolution, Q^{in} and Q^{out} are the input and output of the reservoir (m^3/s), Vs is the water level in the reservoir (m). EC has a maximum value, EC_{max} ($\mu\text{S}/\text{cm}$), when the dissolution of carbonate rock reaches its equilibrium state. As pointed before, EC_{max} in the rainy season of 2017 is higher than the other two years, which is probably caused by the instrument drift. However, given that it is still possible that some

discrepancies of EC_{max} between different years may be caused by the variation in soil CO_2 relating to the temperature [Liu *et al.*, 2007], it was set to a calibrated parameter in the models to fully consider its uncertainty. Its range is set to be a little wider than its variation range in the monitoring period (Table 2).

In M2, M3 and M4, the passive storage, referring to immobile storage that does not affect the hydrologic response but can strongly change the solute dynamic [Birkel *et al.*, 2011; Fenicia *et al.*, 2010; Hrachowitz *et al.*, 2013], was considered in the fast reservoir. Besides the mixing process, we assume that the passive storage is also involved in the dissolution process of carbonate rocks. Then, the change of EC in the fast reservoir can be described as

$$\frac{d(C_{EC}(t) \times (Vs(t) + Vp))}{dt} = \sum C_{EC}^{in}(t) \times Q^{in}(t) + D_{EC} \times (Vs(t) + Vp) - \sum C_{EC}(t) \times Q^{out}(t) \quad (2)$$

Where V_p is the passive storage in the reservoir (m).

For the slow reservoir in M2, M3 or M4, it represents the behavior of slow flow in the entire aquifer (M2 and M3) or the aquifer below the epikarst zone (M4). The slow flow always has a relatively slow velocity and high reaction time with the rock, such that the rock dissolution (most carbonate rock) in groundwater can always reach its equilibrium state. Therefore, we assume that EC in the slow reservoir can always reach its maximum value at each time step in the model irrespective of input EC. We had also tried to add the passive storage to M1 or the epikarst reservoir in M4, but it cannot highly improve their simulation results of spring EC. So the passive storage is not considered in M1 and only one passive storage is added to the fast reservoir in M4. In 2017, we measured the EC value of rainwater in three separate storms ranging from 7.9 $\mu S/cm$ to 10.7 $\mu S/cm$, so the average value 9.3 $\mu S/cm$ is set to the input EC of rainwater.

After adding the EC simulation process in each model, M1 has one extra parameter D_f for considering the dissolution process. M2 and M3 have two extra parameters, D_f for considering the dissolution process in the fast reservoir and V_p for the passive storage. Compared to M3, M4 has another one extra parameter D_{epi} for considering the dissolution process in EPI. Their feasible ranges are shown in Table 2. For M1, M2 and M3, the maximal value of D_f is determined by assuming that the EC in the reservoir can reach its maximal value at each time

step. For V_p in M3 and M4, its maximal value is decided by assuming that spring EC can reach the lowest EC in the calibration period by only considering the solute mixing process. For M4, there are two parameters, D_{epi} and D_f , considering the dissolution process. The range of D_{epi} is set to the same range as D_f in the other three models, whereas the range of D_f in M4 is mainly determined according to the recovery speed of EC after storms.

The models are run at the 12-hours temporal resolution, which was roughly determined according to the value of $1/K_f$ after calibration by spring discharge. Observed data from 2017 to 2019 were used for the model performance test. Given the data interruption and high EC values after the long dry period possibly caused by ‘piston effect’, the longest data series from March 1 to September 26 in 2019 was used for calibration and data in other two periods, from May 11 to August 8 in 2017 (Validation 1) and May 1 to August 7 in 2018 (Validation 2), were used for the validation (subsection 3.5, Fig. 9). The time from January 1 to May 10 in 2017 was set to the spin-up period for each model.

3.4 Performance evaluation

To take account of the various sources of uncertainty including the model structural errors and observational errors in precipitation, discharge and EC observations, we use a Monte Carlo based performance and uncertainty evaluation framework. For each model, 1 000 000 parameter sets were sampled uniformly from realistic parameter ranges (obtained from previous work, e.g., *Chang et al. (2017)*) to investigate their performances in the calibration. Only the model showing an obvious performance improvement for a substantial number of parameter sets compared to other models is regarded as a better model. We evaluate discharge performance by the square root transformed Nash-Sutcliffe that was shown to give a relatively balanced assessment on the spring discharge concerning high and low flows [*Oudin et al., 2006*]:

$$NSE_Q = 1 - \frac{\sum (\sqrt{Q_m} - \sqrt{Q_s})^2}{\sum (\sqrt{Q_m} - \sqrt{Q_m})^2} \quad (3)$$

where Q_m and Q_s are the measured and observed spring discharge respectively (m^3/s). Two criteria were used to assess the model performance on spring EC. The correlation coefficient

aims to evaluate whether the model can capture the temporal dynamics of EC (quick plunge and slow recovery process) while disregarding absolute bias of simulations and observations:

$$r_{EC} = \frac{Cov(EC_m, EC_s)}{\sqrt{EC_m} \sqrt{EC_s}} \quad (4)$$

Where EC_m and EC_s are measured and simulated EC. Also, we use again the square root transformed Nash-Sutcliffe Efficiency to account for correlation and bias between simulations and observations.

$$NSE_{EC} = 1 - \frac{\sum (\sqrt{EC_m} - \sqrt{EC_s})^2}{\sum (\sqrt{EC_m} - \sqrt{EC_m})^2} \quad (5)$$

3.5 Parameter uncertainty analysis and validation

We use the same 1,000,000 parameter sets for parameter uncertainty analysis. Similar to that [Hartmann et al., 2017a], we iteratively confine the number of parameters using the performance measures defined in subsection 3.4. We compare how the distribution of each of the parameters changes in each confinement step to understand the contribution of different signatures to parameter identification. Three confinement steps are performed:

- (1) All parameter sets were only confined by NSE_Q . The parameter sets with NSE_Q lower than 0.85 were discarded and we can get first parameter sets (PS1).
- (2) All parameter sets with NSE_Q lower than 0.85 and r_{EC} lower than 0.75 were discarded and the second parameter sets can be obtained (PS2).
- (3) All parameter sets with NSE_Q lower than 0.85, r_{EC} lower than 0.75 and NSE_{EC} lower than 0.60 were discarded and we can get the third parameter sets (PS3).

The thresholds of NSE_Q , r_{EC} , and NSE_{EC} were defined by a trial and error procedure aiming at the best trade-off between achieving acceptable performance and avoiding over-fitting during the calibration period. Because of the relatively good performance of four models in simulating spring discharge (NSE_Q) and EC dynamics (r_{EC}), higher thresholds were defined compared to the more challenging overall EC simulations (NSE_{EC}). In addition, the confined parameter sets

were used to test the performance stability of each model with independent observations in the two validation periods [Klemeš, 1986].

4 Results

4.1 Model performances in the calibration period

Fig. 5 shows the performances of four models of NSE_Q and r_{EC} in the calibration period. All models show similar simulation results of spring discharge despite their different structures. Hence, the four models cannot be distinguished merely by the spring discharge. The highest NSE_Q is around 0.875 with its lowest value all above 0.30. Four models present very similar simulations of spring discharge and the simulated results of M4 modeling by the parameter sets confined by $NSE_Q > 0.85$ are shown in Fig. 9. The model captures the observed discharge dynamics and most discharges are within the interval except some overestimated peaks.

Fig.5 Scatter plots of four models between NSE_Q and r_{EC} (NSE_{EC}). Four models have similar optimal performances of spring discharge. M3 and M4 both can get good simulation results of r_{EC} , but only M4 shows much better performance of NSE_{EC} . The y-axis is capped to -1.0 for better visualization.

For the simulation of spring EC, the models exhibit distinct performances. M1 has the worst performance with its highest r_{EC} just a little more than 0.40 (Fig 5). M2 gets a little better performance than M1 and the average optimal r_{EC} is a little lower than 0.6 but with a small peak where the highest r_{EC} is 0.84. The peak of r_{EC} appears with a low value of NSE_Q (around 0.8) indicating a trade-off between these two objectives. A better simulation of EC dynamics is obtained at the expense of a worse simulation of spring discharge. In contrast, M3 and M4 both obtain high values of NSE_Q and r_{EC} simultaneously with a limited trade-off between them. Their highest value of r_{EC} is close to 0.89 indicating their much better performances on the EC dynamic than M1 and M2.

When further inspecting their abilities to reproduce spring EC values (Fig. 5), we find that only M4 gets a good result with the highest NSE_{EC} of 0.75. M2 and M3 exhibit similar performance but most parameter sets do not exceed NSE_{EC} values above zero despite a good simulation result of M3 in r_{EC} . M1 obtains the worst performance with NSE_{EC} below -0.5. Fig.9 and Fig. 10 show the simulated spring EC of M4 by PS2 ($NSE_Q > 0.85$ and $r_{EC} > 0.75$) and PS3 ($NSE_Q > 0.85$, $r_{EC} > 0.75$ and $NSE_{EC} > 0.6$) and M3 by PS2 ($NSE_Q > 0.85$ and $r_{EC} > 0.75$) in the calibration period. Although M3 can reproduce the drop-recovery process of spring EC, the simulated EC often increases more quickly than the observed EC. Meanwhile, for the deepest drop in the calibration period, most EC values are severely underestimated. For M4, its simulation result is much better and most simulated EC values are around the observations except two drop peaks (Fig.9).

4.2 Evaluation of parameter uncertainty

Different number reductions of parameter sets were found for different criteria and models. By a high threshold of NSE_Q (0.85), there are 25,400 ~ 60,000 parameter sets left for all models (Fig. 6). After confinement by $r_{EC} > 0.75$, there were no parameter sets left for M1 and M2. As for M3, the number of parameter sets was reduced sharply to 423 parameter sets, whereas 11,513 parameter sets were still left for M4. This indicates a more flexible ability of M4 to simulate the behavior of spring EC by considering the epikarst zone. Finally, after further confinement by $NSE_{EC} > 0.60$, 151 parameter sets were left in M4 while all other models could not exceed this threshold.

Fig. 6 Reduction in the numbers of parameter sets of each model by three objectives. There is no parameter set left by $NSE_Q > 0.85$ and $r_{EC} > 0.75$ for M1 and M2. After further confinement by NSE_{EC} , there is also no parameter left in M3.

Fig. 7 Variations of parameter distributions of four models confined by three different criteria. The parameters all were normalized by the maximal and minimal values within their feasible ranges. The light gray bars indicate that there are no parameter sets left after confinement. 'NAN' means the parameter does not exist in the model.

For all models, the distributions of four hydrologic parameters (S_{max} , K_s , K_f , h_I) in PS1 show an obvious difference from the initial uniform distribution (Fig. 7). Both parameter ranges and distances between 25th and 75th percentiles were reduced indicating they can be constrained by spring discharge observations. Among these parameters, K_f shows an obvious uncertainty reduction implying its increased sensitivity. The distributions of the chemical parameters (D_{epi} , D_f , V_p , EC_{max}) do not change because they are just used for EC simulation.

Further confined by $r_{EC} > 0.75$, the distributions of the chemical parameters in M3 and M4, except EC_{max} , show obvious alterations compared to their previous uniform distributions, especially the parameter D_f indicating its significant role for the EC simulation. The EC_{max} 's distribution is unchanged because r_{EC} only takes account of the EC dynamic, not the specific value of EC. For the hydrologic parameters, just K_s and h_I exhibit small differences from their preceding distributions in M3, while in M4 only h_I has an altered distribution. When the parameter sets were further confined by $NSE_{EC} > 0.60$ in M4, the distributions of the hydrochemical parameters including EC_{max} are further confined. However, only the distribution of h_I exhibits a further noticeable change, whereas K_s and K_f only show minor changes.

4.3 Model validation

The confined parameter sets were tested for their prediction skills in two validation periods (Fig. 8, Fig. 9 and Fig.10). For the spring discharge, in contrast to the calibration period, a slightly worse performance compared to the calibration period is found in the first validation period and the performances are much worse in the second validation period, which is mainly due to the relatively large error by using the observed rainfall data in RSS to represent the actual precipitation in the study area (Fig.2). It is obvious that one discharge peak is severely overestimated by M4 in the second validation period (Fig. 9). For both validation periods, M3 and M4 both show a slight reduction of their NSE_Q ranges when considering EC performances.

For the spring EC, M3 and M4 show a better performance of r_{EC} in the first validation period with just a relatively large variation range and similarly poor values of r_{EC} in the second validation period. The spring EC is still highly underestimated by M3 in the first validation period (Fig. 10). M4 has a wide range of NSE_{EC} changing from 0.74 to -0.55 with a median

value of 0.1 and still exhibits much better performance than M3. Note that this wide range mainly results from the differences of observed EC_{max} between the first validation period and the other two periods (Fig. 2 and Fig.9). If the observed EC curve in the first validation is slightly moved downward, most observed values would be within the interval except the two peaks. We had tested that if all the observed EC values were subtracted from the drifting value (about 23 $\mu\text{S}/\text{cm}$), M4 would have similar validation performances to the calibration period. The range of NSE_{EC} is much lower in the second validation period with a median value of 0.12 due to the overestimated precipitation. In general, M3 and M4 both provide relatively robust simulation results of EC dynamics or values in the validation periods.

Fig. 8 Performance boxplots of four models simulated by different confined parameters sets. a) simulation results of spring discharge of different models using different parameter sets (PS1, PS2 and PS3); b) simulation results of spring EC of M3 and M4 using PS2 and PS3. The white line in the middle of each box represents the median value. The light gray bars means there is no parameter set available for the model validation.

Fig. 9 The simulated spring discharge and EC of M4 in the calibration and validation periods using the confined parameter sets by different criteria.

Fig. 10 The simulated spring EC by M3 using PS1 ($NSE_Q > 0.85$) and PS2 ($NSE_Q > 0.85$ & $r_{EC} > 0.75$).

5 Discussion

5.1 Model structure identification by EC

The performance tests indicate that four models have similar simulation results of spring discharge despite their different model structures. For M1 and M2, the model performance does not present an obvious difference by considering the whole aquifer as one reservoir or two

separate reservoirs. However, through the comparison of parameter distributions between two models confined by $NSE_Q > 0.85$, we found that the different model structure leads to different distributions of two parameters (K_f and h_l , particularly h_l) (Fig. 7). This implies the structure discrepancy between M1 and M2, to some degree, is compensated by the adjustment of parameter values. Regarding M2 and M3, different links between two separated reservoirs do not strongly alter the model performance and also their parameter distributions (Fig. 7). This is probably due to the much larger value of K_f than K_s in models that the fast reservoir cannot strongly regulate the input from the slow reservoir. Generally, these four structures cannot be effectively distinguished by the spring discharge alone.

In contrast to the spring discharge, we found the spring EC is quite valuable to distinguish four different model structures. The optimal model structure proposed by *Chang et al. (2017)*, which considers the whole aquifer as one reservoir, is proven to be too simple to simulate the behavior of EC. When the aquifer is conceptualized as two separate reservoirs (representing fast and slow flow), their different links can lead to different EC simulations. The most common model structure for karst systems (M2) with two parallel reservoirs [*Fleury et al., 2007*] has worse simulation results of spring EC than the structure that considers the slow reservoir recharging the fast reservoir (M3). A similar structure to M3 been used by *Geyer et al. (2008)* to analyze the temporal distribution of recharge in karst systems. Theoretically, the structure of M3 is closer to reality because the karst springs always connect to preferential passages (i.e. conduits) especially in highly karstified aquifers that drain the slow flow in the matrix. This is also true in the study site because the previous tracer tests indicate the spring connects to the sinkholes in the depressions directly through the highly developed conduits [*Yuan et al., 1996*]. It should be mentioned that it is more realistic to consider exchanging relationship between the fast reservoir and the slow reservoir in the model [*Hartmann et al., 2014*] because the inverse flow from the matrix to the conduit may also happen in karst aquifers under storms [*Bailly-Comte et al., 2010*]. However, this inverse flow often exists in a relatively short time which is much smaller than the time step in which the completely solute mixing can reach. Therefore, it does not make much sense to consider this process in the lumped hydrochemical models in our study site.

Further comparing M3 and M4, we found that M3 can only capture the temporal dynamics of spring EC but severely underestimates the absolute EC values. The epikarst reservoir that considers the additional dissolution process of carbonate rock in the epikarst zone before it recharges the fast reservoir, as included in M4, is truly needed in the model to get a good simulation of EC value. This is also realistic because the epikarst zone is the most active upper place receiving fresh rainwater with abundant CO₂ where a mass of dissolution of carbonate rock takes place, which leads to the higher permeability and porosity of the epikarst zone compared to the beneath layer [Williams, 2008]. The final parameter sets of M4 confined by the three criteria (PS3) show the median value of $Depi$ is 0.35 (μs/cm)/min in the epikarst reservoir which is extremely higher than D_f (median value : 7.52×10^{-04} (μs/cm)/min). This further proves the consistency between the internal process in M4 and our perception. Generally, M4 is the optimal model structure to simultaneously simulate the spring discharge and EC based on the available data and our proposed selection of model structures.

5.2 Information provided by spring EC

Previous studies revealed that hydrochemical signals are useful to identify model parameters controlling simulated hydrodynamics [Fenicia et al., 2008; Hartmann et al., 2017a; Vache and McDonnell, 2006]. Our study also shows that spring EC in M4 has a noticeable contribution to further confine the hydrologic parameter h_I , but shows a limited contribution to other hydrologic parameters (S_{max} , K_f and K_s). For M4, although h_I represents the maximal of epikarst storage, it is mainly used to control the occurrence of the point recharge (fast flow) and its quantity because in the model we assume that the extra precipitation exceeding h_I recharges the fast flow directly at each time step (Table 1). So h_I , to some degree, determines the ratio of fast to slow flow at each time step. The high sensitivity of spring EC to h_I is reasonable since the fast flow with a large amount of unsaturated water is the main driving force for the EC drop during storms which mainly derives from the fast infiltration through the sinkholes within depressions in the study site [Liu et al., 2004a; Yuan et al., 1996]. For other hydrologic parameters, the parameter K_f mainly controls the residence time of the active water in the fast reservoir (not including the passive storage) to affect the spring EC dynamics. According to the parameter distributions confined by three criteria in M4, the median value of V_p is 222.4 mm

(25th: 169.4mm, 75th: 288.0mm, Fig. 7), which is larger than the total amount of the largest storm (191.2 mm, 13 July 2019) in the calibration period. The active water in the fast reservoir should be much smaller than the passive storage at each time step during the EC rising period, so the EC rising process mainly depends on the EC dynamics of passive storage controlled by D_f . Similarly, K_s mainly controls the discharge of slow flow, which is also much lower than the passive storage, it is no doubt that it cannot strongly affect the spring EC dynamics. S_{max} is a hydrologic parameter related to the water balance of the catchment, so it, undoubtedly, has a limited influence on the spring EC.

Our analysis shows that the fluctuation of spring EC can provide useful information about the occurrence of the fast flow and its quantity during storms. This finding indicates that spring EC can be used as a tracer to separate simulated slow flow and fast flow if the uncertainty is well considered. For example, *Dreiss* (1989) had used the calcium concentration (similar to EC) to roughly estimate the amount of slow and fast flow (pre-storm and new water in their paper) of a spring. However, due to the existence of the passive storage in M4, the rising process of spring EC is mainly controlled by the hydrochemical parameter D_f which leads to the very limited contribution of spring EC to identify other hydrologic parameters. It should be mentioned that the relatively low contribution of spring EC to parameter identification may also be related to the structural limitations of model structure or inadequate consideration of hydrologic-hydrochemical process interactions. Compared to the stable isotope data (deuterium) [*Zhang et al.*, 2019], spring EC may have a similar function in the models because both can provide information about mixing process between different flow sources, however, in contrast to other used hydrochemical indexes (DOC , Cl^- , SO_4^{2-} and NO_3^-) [*Charlier et al.*, 2012; *Hartmann et al.*, 2017a; *Hartmann et al.*, 2013a], incorporating spring EC into the model may provide extra new information about the fast flow during the storm which can further improve the model evaluation. Besides, our study indicates that the passive storage is recommended in the fast reservoir to get a good simulation of spring EC. This can be seen from the well-confined distribution of V_p confined by $r_{EC} > 0.75$ and $\text{NSE}_{EC} > 0.60$ in M4 (Fig. 7). This passive storage is often considered in non-karst models to better simulate the hydrochemical behavior of the non-karst catchments [*Birkel et al.*, 2010; *Fenicia et al.*, 2010] but used very limited in the karst

models. *Charlier et al.* (2012) had utilized a similar structure to represent the arrival of pre-event flows when they used a lumped model to simulate the transport behavior of dissolved organic carbon in a small karst system. *Zhang et al.* (2019) used the passive storage in a lumped model to simulate the discharge and stable isotope dynamics of a small karst catchment. *Hrachowitz et al.* (2013) pointed out that this passive storage can be interpreted as soil moisture below field capacity or groundwater below the dynamic storage in non-karst catchments. In fact, this passive storage also widely exists in karst aquifers, i.e. the storage in the phreatic conduits, or some storage below the spring outlet. This storage does not actively contribute to the spring flow but can be gradually replaced by the new input water. These water resources may be very valuable for the local people especially in long dry periods when the spring dries out. It should therefore be included in models that are used to support local water management. Besides, parts of the permanently stored water in the unsaturated zone, especially the epikarst zone [*Perrin et al.*, 2003], can also act as passive storage to mix the new input water. Normally, by the observed spring discharge, we can only estimate the active storage in the aquifer which may severely underestimate the total water storage of a system. The spring EC can give us a precious chance to get a first order estimate of the quantity of the total storage in the aquifer to better manage the water resources.

5.3 Limitations of simulating the spring EC

In this paper, we assumed that the variation of spring EC is mainly caused by the dissolution of carbonate rocks and dilution by the unsaturated event water. The simple linear dissolving (EC increases linearly with the rock dissolution) and completely mixing processes were added to the lumped rainfall-runoff model to simulate the spring EC. It is worth noting that, in addition to the carbonate rock, the dissolution of some non-carbonate rocks in the strata or the soil (a small proportion in the study site) may also create other ions contributing to the values and dynamics of spring EC. This contribution was, to some degree, also considered by the parameter D_f or D_{epi} during the model calibration that may lead to a certain uncertainty of spring EC. This uncertainty may increase with the proportion of other non-carbonate-rock related ions in the spring water. Besides, the assumption of linear increasing EC with rock dissolution and complete mixing in the model may still deviate from the real internal processes. Inside actual

karst aquifers, the dissolution process of carbonate rock may be closer to a nonlinear process especially in the late dissolution stage [Buhmann and Dreybrodt, 1985, Hartman and Baker, 2017b]. For the solute mixing process, the volume of the passive storage involving in the mixing may change under different hydrologic conditions. Normally, the involved passive storage should increase with the amount of rainfall. The partial mixing process may be more realistically represented by the solute exchange between active storage and passive storage [Fenicia et al., 2010]. Nevertheless, the performances of our model in the study site indicate that our simplifications were acceptable considering all these limitations or uncertainties.

5.4 Transferability of the approach

In this paper, we proposed an effective and simple method to simulate the EC of a karst spring. This approach can be easily applied with other existing hydrologic models at karst aquifers with similar dissolution dynamics of carbonate rock. In our study site, because most discharge events happen during the rainy season, an obvious seasonal variation of spring EC is not apparent in our data. However, for other karst springs, especially the epikarst springs, a strong seasonal variation of spring EC may exist [Liu et al., 2007]. This variation may be considered through seasonal variations of the parameter EC_{max}, i.e. setting different EC_{max} in different seasons. Furthermore, the initial rising of EC caused by the ‘piston effect’ after storms [Goldscheider and Drew, 2007], which is also not pronounced in the study site, may also be abundant in other karst aquifers. However, this phenomenon often lasts just a short time which is much shorter than the simulation time step of most models. If the time scale of this process substantially exceeds the simulation time step, the observed EC during this interval should not be involved in the model evaluation or more complex dissolution processes have to be introduced into the model. Our method is based on the assumption that the hydrochemical composition of water in the karst aquifer is relatively stable which is dominated by the calcium carbonate equilibria. Therefore, it is not suitable in karst aquifers (1) that are strongly influenced by human activities that affect the EC dynamics or (2) that have massive non-carbonate rocks, where the component ratios of solute in the water may strongly change with time.

6. Conclusions

EC has always been a fundamental variable to characterized karst aquifers as it is easily obtained by affordable and automatic monitoring instruments. Although spring EC is widely used to analyze the internal behavior of karst aquifers, its implementation in hydrologic modeling to assist model evaluation is still difficult. In this paper, we proposed an effective method to incorporate EC into hydrologic modeling and investigate its value for model structure identification and parameter estimation. EC is assumed to increase linearly with the dissolution of carbonate rocks, and the complete solute mixing is used in the lumped rainfall-runoff models to simulate the spring EC. After a performance comparison of four typical model structures, we found that spring EC can be used efficiently as auxiliary data in addition to spring discharge to identify realistic model structures. In addition, we show that EC exhibits a potential value to further constrain the range of one hydrologic parameter related to the occurrence and quantity of point recharge. Consequently, the incorporation of spring EC shows a high potential value for more realistic modeling reducing the risk of using oversimplified model structures for the simulation of karst aquifers.

Our approach is independent of specific model structures and particular test sites. It is therefore easily transferrable to other modeling studies with similar characteristics to the study site. Although our study does not present a large contribution of spring EC to lower parameter uncertainty, spring EC can strongly constrain model structure selection towards more model structural realism. Furthermore, a promising way is to use EC together with other environmental tracers in the model to further improve the model selection and evaluation. On one hand, the spring EC can be combined with δD to compensate for the low temporal resolution of δD . On the other hand, it can also be used together with other hydrochemical indexes to provide extra information about fast flow to further promote structure and parameter identification [Hartmann *et al.*, 2017]. At almost every karst system, EC is traditionally monitored in addition to spring discharge [Jourde *et al.*, 2018] and can, therefore, be a potentially useful source for further identifying the most adequate model structure. Consequently, our new framework to consider EC in karst model application opens new doors for more realistic karst modeling. Further improvements of our method may be obtained by future research to involve nonlinear

dissolution processes and partial mixing in hydrologic models to consider the simulation of EC dynamics.

Acknowledgments

We thank Dr. Yin Jianjun for providing the precipitation data in NRS and IKG. The authors would like to thank the Associate Editor and three anonymous reviewers to provide constructive comments that greatly improve the quality of this manuscript. This research was supported by the National Key Research and Development Program of China (NO. 2016YFC0402809), the National Natural Science Foundation of China (No. 41602242 and No. 41702243) and the Fundamental Research Funds for the Central Universities (2018B05114). Support to Andreas Hartmann was provided by the Emmy Noether-Programme of the German Research Foundation (DFG; grant number HA 8113/1-1; project ‘Global Assessment of Water Stress in Karst Regions in a Changing World’). All the model codes and the data for model evaluation are available for download: <https://zenodo.org/record/3554764#.XzGJJdVKjIU>.

References

- Aquilina, L., B. Ladouche, and N. Dörfliger (2006), Water storage and transfer in the epikarst of karstic systems during high flow periods, *Journal of Hydrology*, 327(3–4), 472–485.
- Bailly-Comte, V., J. B. Martin, H. Jourde, E. J. Screaton, S. Pistre, and A. Langston (2010), Water exchange and pressure transfer between conduits and matrix and their influence on hydrodynamics of two karst aquifers with sinking streams, *Journal of Hydrology*, 386(1–4), 55–66.
- Birkel, C., S. M. Dunn, D. Tetzlaff, and C. Soulsby (2010), Assessing the value of high-resolution isotope tracer data in the stepwise development of a lumped conceptual rainfall–runoff model, *Hydrological Processes*, 24(16), 2335–2348.
- Birkel, C., C. Soulsby, and D. Tetzlaff (2011), Modelling catchment-scale water storage dynamics: reconciling dynamic storage with tracer-inferred passive storage, *Hydrological Processes*, 25(25), 3924–3936.
- Bittner, D., M. T. Parente, S. Mattis, B. Wohlmuth, and G. Chiogna (2020), Identifying relevant hydrological and catchment properties in active subspaces: An inference study of a lumped karst aquifer model, *Advances in Water Resources*, 135, 103472.
- Buhmann, D., and W. Dreybrodt (1985), The kinetics of calcite dissolution and precipitation in geologically relevant situations of karst areas: 1. Open system, *Chemical Geology*, 48(1), 189–211.
- Chang, Y., J. Wu, and G. Jiang (2015), Modeling the hydrological behavior of a karst spring using a nonlinear reservoir-pipe model, *Hydrogeology Journal*, 23, 901–914.
- Chang, Y., J. Wu, G. Jiang, and Z. Kang (2017), Identification of the dominant hydrological process and appropriate model structure of a karst catchment through stepwise simplification of a complex conceptual model, *Journal of Hydrology*, 548, 75–87.

- Chang, Y., J. Wu, G. Jiang, L. Liu, T. Reimann, and M. Sauter (2019), Modelling spring discharge and solute transport in conduits by coupling CFPv2 to an epikarst reservoir for a karst aquifer, *Journal of Hydrology*, 569, 587-599.
- Charlier, J. B., C. Bertrand, and J. Mudry (2012), Conceptual hydrogeological model of flow and transport of dissolved organic carbon in a small Jura karst system, *Journal of Hydrology*, 460, 52-64.
- Dreiss, S. J. (1989), Regional scale transport in a Karst Aquifer: 1. Component separation of spring flow hydrographs, *Water Resources Research*, 25(1), 117-125.
- Duran, L., N. Massei, N. Lecoq, M. Fournier, and D. Labat (2020), Analyzing multi-scale hydrodynamic processes in karst with a coupled conceptual modeling and signal decomposition approach, *Journal of Hydrology*, 583, 124625.
- Eagleman, J. R. (1967), Pan Evaporation, Potential and Actual Evapotranspiration, *Journal of Applied Meteorology*, 6(3), 482-488.
- Enemark, T., L. J. M. Peeters, D. Mallants, and O. Batelaan (2019), Hydrogeological conceptual model building and testing: A review, *Journal of Hydrology*, 569, 310-329.
- Fenicia, F., J. J. McDonnell, and H. H. G. Savenije (2008), Learning from model improvement: On the contribution of complementary data to process understanding, *Water Resources Research*, 44, W06419.
- Fenicia, F., S. Wrede, D. Kavetski, L. Pfister, L. Hoffmann, H. H. G. Savenije, and J. J. McDonnell (2010), Assessing the impact of mixing assumptions on the estimation of streamwater mean residence time, *Hydrological Processes*, 24(12), 1730-1741.
- Fleury, P., V. Plagnes, and M. Bakalowicz (2007), Modelling of the functioning of karst aquifers with a reservoir model: Application to Fontaine de Vaucluse (South of France), *Journal of Hydrology*, 345(1-2), 38-49.
- Ford, D. C., and P. Williams (2007), *Karst hydrogeology and geomorphology*, John Wiley & Sons.
- Geyer, T., S. Birk, R. Liedl, and M. Sauter (2008), Quantification of temporal distribution of recharge in karst systems from spring hydrographs, *Journal of Hydrology*, 348(3-4), 452-463.
- Goldscheider, N., and D. Drew (2007), *Methods in Karst Hydrogeology: IAH: International Contributions to Hydrogeology*, 26, CRC Press.
- Hartmann, A., J. A. Barber á and B. Andreo (2017a), On the value of water quality data and informative flow states in karst modelling, *Hydrology & Earth System Sciences*, 21(12), 5971-5985.
- Hartmann, A., & Baker, A. (2017b). Modelling karst vadose zone hydrology and its relevance for paleoclimate reconstruction. *Earth-Science Reviews*, 172, 178-192.
- Hartmann, A., N. Goldscheider, T. Wagener, J. Lange, and M. Weiler (2014), Karst water resources in a changing world: Review of hydrological modeling approaches, *Reviews of Geophysics*, 52, 218-242.
- Hartmann, A., J. Lange, À. Viv ó Aguado, N. Mizyed, G. Smiatek, and H. Kunstmann (2012), A multi-model approach for improved simulations of future water availability at a large Eastern Mediterranean karst spring, *Journal of Hydrology*, 468-469, 130-138.
- Hartmann, A., T. Wagener, A. Rimmer, J. Lange, H. Brielmann, and M. Weiler (2013a), Testing the realism of model structures to identify karst system processes using water quality and quantity signatures, *Water Resources Research*, 49, 3345-3358.
- Hartmann, A., et al. (2013b), Process-based karst modelling to relate hydrodynamic and hydrochemical characteristics to system properties, *Hydrology and Earth System Sciences*, 17(8), 3305-3321.
- Hrachowitz, M., H. Savenije, T. A. Bogaard, D. Tetzlaff, and C. Soulsby (2013), What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, *Hydrology and Earth System Sciences*, 17(2), 533-564.

- 771 Husic, A., J., Fox, E., Adams, W., Ford, C., Agouridis, J., Currens, and J., Backus (2019), Nitrate Pathways,
772 Processes, and Timing in an Agricultural Karst System: Development and Application of a Numerical
773 Model. *Water Resources Research*, 55, 2079–2103.
- 774 Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model?,
775 *Water Resources Research*, 29(8), 2637-2649.
- 776 Jiang, G. H. (2010). Solute Transport in Karst Hydrological System (doctoral thesis), *Nanjing University*. (in
777 chinese)
- 778 Jourde, H., et al. (2018), SNO KARST: A French Network of Observatories for the Multidisciplinary Study of
779 Critical Zone Processes in Karst Watersheds and Aquifers, *Vadose Zone Journal*, 17(1).
- 780 Klemeš, V. (1986). Dilettantism in hydrology: Transition or destiny? *Water Resources Research*, 22(9S), 177S-
781 188S.
- 782 Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and
783 models to advance the science of hydrology, *Water Resources Research*, 42(3).
- 784 Kuczera, G., and M. Mroczkowski (1998), Assessment of hydrologic parameter uncertainty and the worth of
785 multiresponse data, *Water Resources Research*, 34(6), 1481-1489.
- 786 Liu, Z., C. Groves, D. Yuan, and J. Meiman (2004a), South China Karst Aquifer Storm-Scale Hydrochemistry,
787 *Ground Water*, 42(4), 491-499.
- 788 Liu, Z.H., Q. Li, W.J. Liang, K.Y. Wu, H.L. Sun, and Y.P. Liang (2004b), Interpretation of borehole storm-scale
789 hydrochemical variations at the Guilin in Karst experimental site, *Carsological Sinica*, 23(3), 169-176. (in
790 chinese)
- 791 Liu, Z., Q. Li, H. Sun, and J. Wang (2007), Seasonal, diurnal and storm-scale hydrochemical variations of
792 typical epikarst springs in subtropical karst areas of SW China: Soil CO₂ and dilution effects, *Journal of*
793 *Hydrology*, 337(1-2), 207-223.
- 794 Marandi, A., M. Polikarpus, and A. Jõelet (2013), A new approach for describing the relationship between
795 electrical conductivity and major anion concentration in natural waters, *Applied Geochemistry*, 38, 103-
796 109.
- 797 Massei, N., B. J. Mahler, M. Bakalowicz, M. Fournier, and J. P. Dupont (2007), Quantitative interpretation of
798 specific conductance frequency distributions in karst, *Ground Water*, 45(3), 288-293.
- 799 Meus, P., P. Moureaux, S. Gailliez, J. Flament, F. Delloye, and P. Nix, (2014). In situ monitoring of karst springs
800 in Wallonia (southern Belgium). *Environmental Earth Sciences*, 71(2), 533–541.
- 801 Moussu, F., L. Oudin, V. Plagnes, A. Mangin, and H. Bendjoudi (2011), A multi-objective calibration
802 framework for rainfall-discharge models applied to karst systems, *Journal of Hydrology*, 400(3-4), 364-
803 376.
- 804 Mudarra, M., and B. Andreo (2011), Relative importance of the saturated and the unsaturated zones in the
805 hydrogeological functioning of karst aquifers: The case of Alta Cadena (Southern Spain), *Journal of*
806 *Hydrology*, 397(3-4), 263-280.
- 807 Mudarra, M., B. Andreo, J. A. Barber á and J. Mudry (2014), Hydrochemical dynamics of TOC and NO₃⁻
808 contents as natural tracers of infiltration in karst aquifers, *Environmental Earth Sciences*, 71(2), 507-523.
- 809 Mudarra, M., A. Hartmann, and B. Andreo (2019), Combining Experimental Methods and Modeling to Quantify
810 the Complex Recharge Behavior of Karst Aquifers, *Water Resources Research*, 55(2), 1384-1404.
- 811 Oudin, L., V. Andreassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging of rainfall-runoff
812 model simulations from complementary model parameterizations, *Water Resources Research*, 42(7),
813 W07410.

- Perrin, J., P.-Y. Jeannin, and F. Zwahlen (2003), 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.
- Ravbar, N., I. Engelhardt, and N. Goldscheider (2011), Anomalous behaviour of specific electrical conductivity at a karst spring induced by variable catchment boundaries: the case of the Podstenjšek spring, Slovenia, *Hydrological Processes*, 25(13), 2130–2140.
- Schnegg, P.-A. (2003). A new field fluorometer for multi-tracer tests and turbidity measurement applied to hydrogeological problems. In: Proceedings of the Eight International Congress of the Brazilian Geophysical Society, Rio de Janeiro.
- Son, K., and M. Sivapalan (2007), Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, *Water Resources Research*, 43(1), W01415.
- Stevanović, Z. (2019). Karst waters in potable water supply: a global scale overview. *Environmental Earth Sciences*, 78(23), 1–12.
- Toran, L., and Reisch, C. E. (2013). Using Stormwater Hysteresis to Characterize Karst Spring Discharge. *Ground Water*, 51, 1745–6584.
- Torresan, F., P. Fabbri, L. Piccinini, N. Dalla Libera, M. Pola, and D. Zampieri (2020), Defining the hydrogeological behavior of karst springs through an integrated analysis: a case study in the Berici Mountains area (Vicenza, NE Italy), *Hydrogeology Journal*, 28, 1229–1247.
- Vache, K. B., and J. J. McDonnell (2006), A process-based rejectionist framework for evaluating catchment runoff model structure, *Water Resources Research*, 42, W02409.
- Wang, K. R., F. Guo, G.H. Jiang and H.Y. Bian (2014), Application of ^{15}N and ^{18}O to nitrogen pollution source in karst water in Eastern Guilin, *China Environmental Science*, 34(9), 2223–2230. (in chinese)
- Williams, P. W. (1983), The role of the subcutaneous zone in karst hydrology, *Journal of Hydrology*, 61(1–3), 45–67.
- Williams, P. W. (2008), The role of the epikarst in karst and cave hydrogeology: a review, *International Journal of Speleology*, 37, 1–10.
- Winston, W. E., and R. E. Criss (2004), Dynamic hydrologic and geochemical response in a perennial karst spring, *Water Resources Research*, 40, W05106.
- Ye, W., B. C. Bates, N. R. Viney, M. Sivapalan, and A. J. Jakeman (1997), Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments, *Water Resources Research*, 33(1), 153–166.
- Zeng, C., Z. Liu, J. Yang, and R. Yang (2015), A groundwater conceptual model and karst-related carbon sink for a glacierized alpine karst aquifer, Southwestern China, *Journal of Hydrology*, 529, 120–133.
- Zhang, Z., X. Chen, Q., Cheng, and C. Soulsby (2019), Storage dynamics, hydrological connectivity and flux ages in a karst catchment: Conceptual modelling using stable isotopes. *Hydrology and Earth System Sciences*, 23(1), 51–71.
- Yuan, D.X., A.D. Dai, W.T. Cai, Z.H. Liu, S.Y. He, X.P. Mo, S.Y. Zhou, and W.K. Lao (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, China. (in chinese)