

## **Anonymous Referee #1**

Received and published: 6 May 2018

This paper presents some interesting simulations of a karst catchment in China. However, (at present) I cannot recommend publication, but after the following concerns are addressed. However, before I can recommend publication the following list of concerns need to be addressed.

We sincerely thank the reviewer for his/her comments and suggestions that significantly improved our manuscript. We have thoroughly revised our manuscript taking into account all suggestions and comments from the reviewer. Our point-to-point responses are detailed below.

### **Main comments**

From reading this paper, it is unclear what the real novel contribution is. Surely interesting results are presented, but what do we really learn? I cannot derive this from the abstract, nor the conclusions. Please make this MUCH more explicit. The specific aims tell you mostly “what” you do, instead of what you want to learn (and what is new about that). Only once I know what we aim to learn from this paper I can properly review the paper. Right now I mainly see a long list of results and statements. Sure I could comment on every detail of them, but that would not warrant a review which allows me to judge the scientific contribution of this paper well.

Reply: We have revised the manuscript to emphasize more clearly the novel contribution of the manuscript as follows:

- (1) For cockpit terrain in the southwest China karst area, hillslope runoff processes are mostly routed into depression aquifers prior to contributing to streamflow (“hillslope- to- depression- to-stream”). Identifying and quantifying the dynamics of water storage, hydrological connectivity between different stores and the associated ages of water fluxes is very important to understand how the unique landscape characteristics of karst affect flow transmission. This has applied significance for understanding water resource availability and flood hazard management.
- (2) Consequently, we developed a tracer-aided runoff model that disaggregates the cockpit karst terrain into the two dominant landscape units of hillslopes and depressions (further sub-dividing the depression into fast and slow reservoirs) extending an earlier model developed by the authors of the dual flow reservoirs for flow and solute (*Ca* and *Mg*) concentrations at the catchment scale. This tracer-aided model conceptualizes hydrological functions more comprehensively by estimating storage-flux dynamics and water ages in each unit. Such tracer-aided models enhance our understanding of the hydrological connectivity between different landscape units and the mixing processes between various flow sources.
- (3) Since the tracer-aided model increases model parametersisation for the tracer modules, we evaluated the uncertainty of the modelled results, including not only those of flow and stable isotopic values, but also water storage and flux ages at various landscape units. In particular, we found that the tracer-aided models can be used to characterize the uncertainty of modelled results at any units in the catchment.

The writing of this paper needs significant improvement. In its current format, the paper contains very awkward and confusing use of the English language, which makes it at times hard to read and review. I suggest a native speaker takes a critical look at the whole paper. That makes more sense than that

the reviewer does all this work for them. Nevertheless, I provide a long list of suggestions below, but addressing these will probably be not sufficient to tackle the language problems of this paper. Note that these problems with the writing do not only refer to grammar issues, but also to the plethora of statements, structure of reasoning, etc. that are unclear in the current format.

Reply: We appreciate the referee's suggestions, and the whole paper has been thoroughly revised to improve the clarity and grammar.

### Detailed comments

Line 9: "unique" does not seem appropriate since other studies have similar or higher temporal resolution isotope and hydrometric data. For example,

Floury, P., Gaillardet, J., Gayer, E., Bouchez, J., Tallec, G., Ansart, P., Koch, F., Gorge, C., Blanchouin, A., and Roubaty, J.-L.: The potamochemical symphony: new progress in the high-frequency acquisition of stream chemical data, *Hydrol. Earth Syst. Sci.*, 21, 6153-6165, <https://doi.org/10.5194/hess-21-6153-2017>, 2017.

von Freyberg, J., Studer, B., and Kirchner, J. W.: A lab in the field: high-frequency analysis of water quality and stable isotopes in stream water and precipitation, *Hydrol. Earth Syst. Sci.*, 21, 1721-1739, <https://doi.org/10.5194/hess-21-1721-2017>, 2017.

Reply: We have re-stated this. Whilst we recognize that others have such higher temporal resolution stream data, we wished to emphasize that our high resolution, extended isotope and hydrometric observations concurrently collected in hillslopes, depressions and streams of complex karst catchments are scarce.

Line 10: "flow-tracer model" is not really a clear term

Reply: This has been replaced by "tracer-aided model".

Line 10: the model represents "the movement of water" using "two main landscape ....". I suggest to add this, otherwise the sentence does not make much sense anymore.

Reply: We have revised this.

Line 11: "cock-pit": I think you can remove the hyphen.

Reply: We have removed the hyphen.

Line 12: "this inferred" is not logical. Something like "from these model results we inferred" would be much better.

Reply: We have revised this.

Line 13: or something like "had least water stored, whereas the slow reservoir has least water stored" (which makes the sentence more understandable, and it removes the redundant "intermediate" part.

Reply: We have revised this.

Line 14: specify that you talk about mean ages OF WATER.

Reply: We have revised this to "The estimated mean ages of the hillslope unit, fast and slow flow reservoirs during the study period ....."

Line 14: “marked” seems unclear and redundant to me

Reply: The “marked” has been replaced by “distinct”

Line 14-16: This statement is somewhat meaningless with its current explanation. “Connectivity can be defined in many ways” so I suggest that you describe what you physically found, rather than use an undefined buzzword. Actually, all the statements until sentence 18 are somewhat unclear. What do you mean by “reversible directionality”? I can guess, but please try to make the wording clearer to the reader.

Reply: We have revised the sentences from lines 14 to 19. We have clarified what we physically found.

Line 16-19: please revisit these sentences to make this an understandable abstract.

Reply: We have revised the sentences from line 16 to 19.

Line 32: “whole catchment” instead of “whole karst system” (the karst system may have a different scale).

Reply: We have revised this.

Line 33: “However, semi-distributed lumped models need to have hydrogeological units adequately represented, in order to relate water flow in different landscape units and model parameters that have physically meaningful concepts.” Is not logically connected to the previous statements. Where does the “however” come from?

Reply: We have revised this paragraph. In the revised manuscript, we have first described the lumped models and then the semi-distributed models. (see lines 48-65 in the revised manuscript)

Line 36: “Three main types of porosities – (a) micropores, (b) small fractures, and (c) large fractures and conduits – can be intuitively identified in karst systems.” Do would it not help to start a new paragraph here?

Reply: We have revised this and defined terms more precisely. “In karst aquifers, the solutional conduits connect with intergranular pores and small fractures (often termed as matrix porosity), showing dual or even triple porosity zones (Worthington, Jeannin et al., 2017). Thus, karst aquifers are often conceptualized as dual porosity systems as residence times in the matrix are often several orders of magnitude longer than those in the conduits (Goldscheider & Drew, 2007).” (see lines 52-58 in the revised manuscript)

Line 37: “can be intuitively identified” what do you mean here?

Reply: We have deleted the sentence in line 37 and replaced with the statement above.

Line 42: (Rimmer and Hartmann, 2012; Hartmann et al, 2014; Zhang et al, 2017). Include and “e.g. since many more examples will exist).

Reply: We have revised this.

Line 43-46: please rephrase “However, this kind of approach cannot disaggregate water storage and

flux dynamics within different landscape units, and may be inadequate for modelling when understanding known spatial differences in hydrogeological structure is important in terms of provisioning water supplies and understanding water quality issues (Fu et al, 2016; Zhang et al, 2013)”

Reply: We have revised this.

Line 59: I think what Kirchner said is that these tracers help to ‘highlight their differences’ rather than that they “resolve” anything really.

Reply: We have revised this to “..... have helped to resolve this celerity-velocity dichotomy known as the “old water paradox”.

Line 71: “Hydrological connectivity, which has been simply defined as the transfer of water from one part of the landscape to another (McGuire and McDonnell, 2010; Golden et al, 2014; Soulsby et al., 2015),” this statement suggests that hydrologic connectivity is about the transport of water (e.g. velocity) rather than the “celerity effects” it is used for to describe. I think you need to be more accurate in its description.

Reply: We have deleted the sentence which gave the different descriptions of hydrologic connectivity than what we used.

Section 2.1. Did you take this information from other (peer reviewed) publications? If yes, please cite these.

Reply: We have added the relevant publications.

Figure 1: please make it much more explicit in the caption what you display here.

Reply: We have added the relevant descriptions.

Table 1: the range is a redundant variable.

Reply: This table has been replaced by the flow duration curve.

Table 1: consider indicating how much of the time there is zero flow.

Reply: This table has been replaced by the flow duration curve. We have added description of the time there is zero flow in text. It occupies only a short period of time in our observation period.

Table 1: why not provide a flow duration curve instead. That will be WAY more informative than what you currently present.

Reply: The flow duration curve has been added.

Line 159: CalculationS

Reply: We have revised this.

Line 162: for each of the (not in each of the)

Reply: We have revised this.

Line 162-163: inconsistent with singular and plural. Check grammar.

Reply: We have revised this sentence and made corrections in the whole manuscript.

Line 168: fix superscript “rainfall (m<sup>3</sup> hour<sup>-1</sup>)” Equations 8-11: I presume you talk about some mean age for the box, please specify this. Equations 8-11 there equations are missing the “aging” term. (i.e. water gets older over time), please add this term and check if you calculations are correct: : :

Reply: We have revised this, and given a more complete description in the appendix.

We have considered the “aging” item. In the model procedure, each age item at the time t includes the age at the previous time step t-1. So, the results listed in this paper include the “aging effect” (this has been clarified in the appendix about the model descriptions).

Section 3.2 months spin up time may be sufficient spin up time for hydrometric fluxes, but will it be for modeling of ages?

Reply: The spin up time for the modeling of ages is sufficient given the young water dominance. Our two step calibration procedures show: as the mean value of the modelled water ages (meeting the target of KGE>0.3) in the first calibration was used as the initial water ages for the second calibration, the calibrated water ages for each conceptual store well matches the measured isotope values (the target of KGE increases to be higher than 0.5). It means that the selected initial period for “warm up” modeling of ages is reasonable.

Section 3.2: “First, different parameter combinations within the initial ranges in Table 3 were tested. And then, the parameter ranges were reduced according to the best models (KGE >0.3) for the second calibration. This resulted in a total of 10<sup>5</sup> tested different parameter combinations. I do not understand how you arrive at the second 10<sup>5</sup>.”

Reply: We have revised the descriptions and added the initial parameters from the first calibration (see lines 251-257 in the revised manuscript). A total of 10<sup>5</sup> different parameter combinations was given for the estimation of uncertainty of the modeled results from the random generation of the possible parameter combinations. The number of 10<sup>5</sup> different parameter combinations is believed to be sufficient according to uncertainty analysis in the literatures (Soulsby et al., 2015; Xie et al., 2017). In the first and second calibrations, the number of parameter combinations were set to 10<sup>5</sup>, but the range of the initial parameters are different (Initial range 1 and 2 for the 1st and 2nd calibration in revised Table 2).

After the first calibration when KGE >0.3, the range of each parameter was reduced. Then, the narrowed ranges (initial range 2 in Table 2) were used as the initial ranges for the second calibration.

**Table 2 Mean parameter values and fitness derived from the best 500 parameter sets after calibration**

For Flow	$K_s$ (hour <sup>-1</sup> )	$K_f$ (hour <sup>-1</sup> )	$K_e$ (hour <sup>-1</sup> )	$f$	$a$	$W$	$b$
Initial range 1	40-168	1-72	800-2200	0.005-0.025	0-1	0-0.015	0-1
Initial range 2	40-150	1-40	800-2200	0.008-0.025	0.47-1	0-0.015	0.48-1
Mean	92	11	1549	0.015	0.68	0.005	0.54
Range	48-120	5-18	1000-2000	0.01-0.02	0.51-0.9	0.003-0.01	0.5-0.62
For Isotope	$I_s$	$KK$ ( $\times 10^4$ )	$pp$	$con$	$fei$	Index	Mean(range)
						$KGE_a$	0.85 (0.81-0.87)

Initial range 1	0-1	0.8-1.6	0-1	0-1	0-1	<i>KGE<sub>i</sub></i>	0.56 (0.52-0.59)
Initial range 2	0-0.8	0.8-1.6	0-1	0-1	0.5-1	<i>KGE</i>	0.7 (0.72-0.66)
Mean	0.24	1.26	0.49	0.56	0.82		
Range	0.002-0.6	1-1.5	0.02-0.95	0.04-0.97	0.71-0.93		

Line 276: “rogue” ? what do you mean

Reply: It refers to some samples that are unusually high during the study period (in Fig.4b). These samples could be affected by the paddy water in the manuscript description.

Figure 10: these values cannot be correct since the areas under these curves do not add up to 1.

Reply: The figure gave the probability density functions (PDFs) of the flux ages from the three units. The sum of these values for the three units, e.g. the total areas covered by the three curves with different colors (conceptual stores), equals to 1.

Line 420: cannot instead of can't.

Reply: We have revised this.

Line 445-447 “Given the results on water storage dynamics and the relative contribution to the fast flow reservoir shown in Figures 7 and 8, it can be deduced that the storage change within each conceptual store is the main driver of hydrological connectivity between them.” Is this not just how you defined that the catchments functions yourself? So what did we really learn in the end? (also remove the “s” in stores)

Reply: We have revised the conclusions. The tracer-aided model supports general appropriateness of the model structure which related connectivity dynamics to storage change within different landscape units. During the dry period, there is weak hydrological connectivity between the hillslope and depression due to low storage. In contrast, during the wet period, hydrological connectivity between the hillslope and depression strengthens as water storage increases. In the early recession, after heavy rain, large fractures in the hillslope fill, leading to large water fluxes into the depression. Then, as storage declines, fluxes decrease and the hydrological connectivity weakens (lines 457-462 in the revised manuscript).

## Anonymous Referee #2

### Main comments

1. The study is well-written and concise. The model calibration and sensitivity analysis are detailed and well described. However, there are some remaining concerns that the authors may account for before their manuscript can be considered for publication:

We sincerely thank the reviewer for carefully reviewing our manuscript and for the thoughtful, constructive feedback.

2. For the reader, who is not familiar with the author's preceding work, it is not clear how the model works. The schematic description in Fig 2 indicates that ET is taking place from the slow and fast karst groundwater storages, which would be quite unusual. To avoid misconception, please provide a complete model description in appendix (the table A1 is hardly understandable).

Reply: In this model, the karst critical zone in the hillslope was conceptualized as one reservoir, but the water stored in the reservoir was further sub-divided into upper active and lower passive storage zones (Fig 3 in the revised manuscript) for the simulation of isotope ratios and estimation of water ages. This division follows our previous measurements of the vertical distribution of the rock fractures/conduits along hillslopes where the large rock fractures/conduits decrease exponentially in the vertical direction (Zhang et al., 2011).

The karst critical zone in the depression was conceptualized as two connected reservoirs, fast and slow flow, representing the solutional conduits in karst aquifers connecting with intergranular pores and fractures (often termed as matrix porosity).

The evapotranspiration could occur from the rich conduit/fracture areas by extended plant roots in the deep aquifer (Rong et al., 2011). Therefore, evapotranspiration is sourced from both the fast and slow reservoirs in the model.

We have revised the model descriptions in an appendix for clearer explanation of the module functions and meanings in the revised manuscript.

3. Some clarification on where the novelties of this work start is necessary. The authors inform the reader that in Zhang et al. (2017), the model was developed in previous work that used tracer data in addition to stream discharge to constrain the model structure, improve parameterization, and aid calibration. If this was done before, and the methods only describe how the isotope enabled model was parametrized and evaluated, what is the novelty of this particular study?

Reply: The model in the preceding work (Zhang et al., 2017), conceptualized the flow and the geochemical solute (Ca+Mg) routings using conceptualization of the dual flow system at the catchment scale. So, the original model had no basis for disaggregating the hydrological connectivity between different landscape units (e.g. "hillslope- to- depression- to- stream" in the study catchment). The hillslope-depression is a typical landform with variable hydrological connectivity in the karst catchments in southwest of China (Figure r1, Chen et al., 2018). Here, we



improved our previous model structure by conceptualizing the hillslope and depression units (the improved part is in the red dotted box in Figure r2), and then use the hourly discharge and isotope values to calibrate the model. In addition, the new model has the parameters to represent passive storage inferred by isotope damping and the function of estimating the water ages from various landscape units in the catchment.

Although the tracer-aided model enhanced our understanding of the hydrological connectivity between different landscape units and the mixing processes, it increased the model parameters in the tracer modules. Therefore, we also evaluated the uncertainty of the simulation results including flow discharges, isotopic values, storages and ages at the different landscape units in this study.

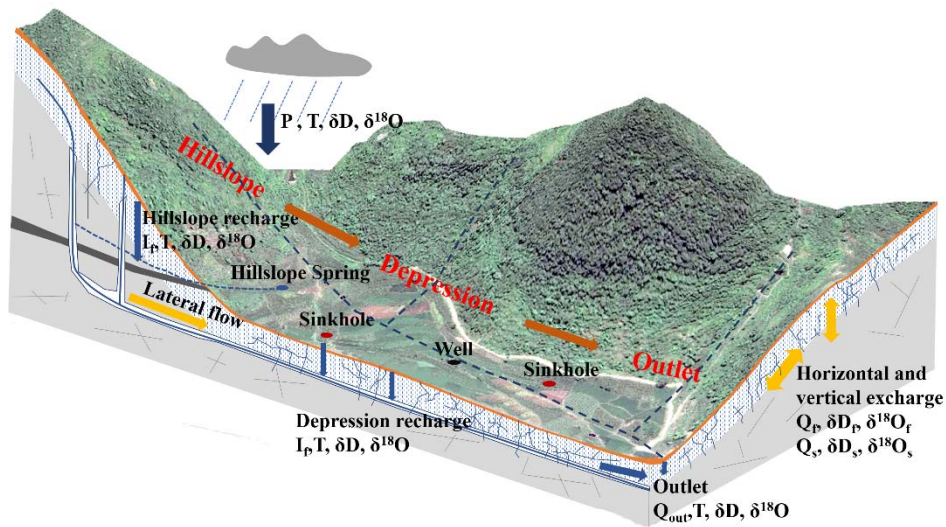


Figure r1 Sketch map of karst hydrological processes (Chen et al, 2018)

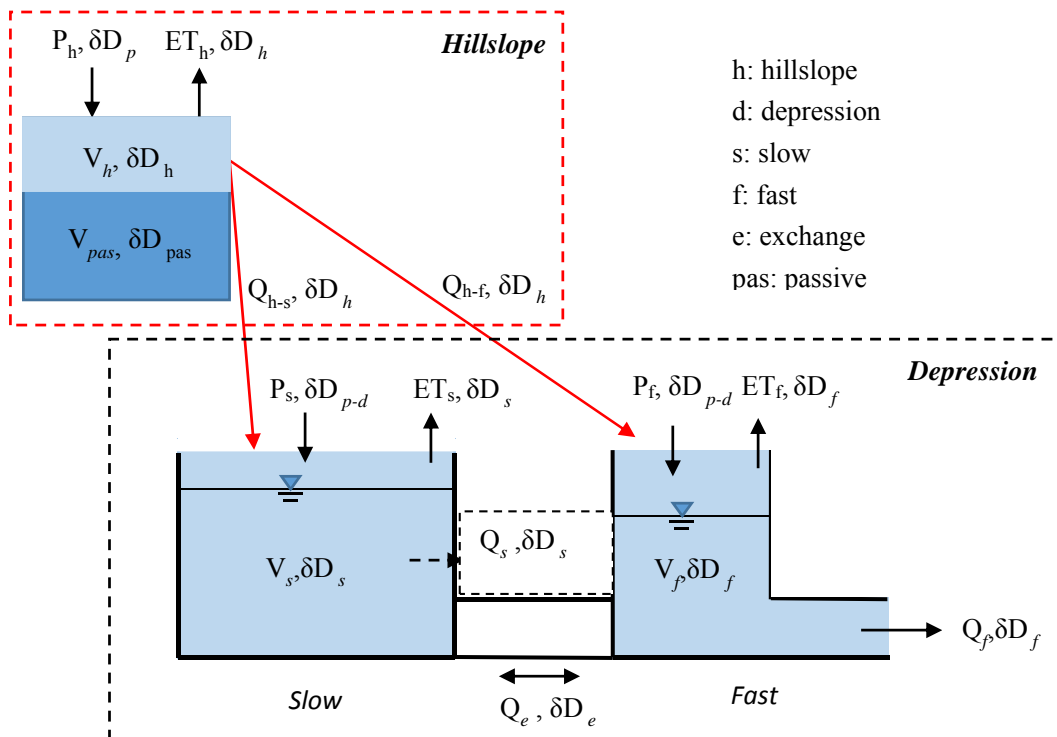


Figure r2 Structure of the improved model, and the improved part is in the red dotted box



4. Figure 4 shows that only 5 of 12 parameters are sensitive, which is quite a low number. Usually, discharge contains enough information to identify 4-6 parameters (Jakeman & Hornberger, 1993). Adding of additional information like isotopes should increase this number, if the model structure is well-chosen. To check the contribution of discharge data and isotopes, could the authors show the parameter sensitivities using discharge or water isotopes only?

Reply: The trace-aided model includes 12 parameters, seven for flow routing ( $Ks$ ,  $Kf$ ,  $Ke$ ,  $f$ ,  $a$ ,  $w$ , and  $b$ ) and five for isotope ratios and water ages ( $Is$ ,  $KK$ ,  $pp$ ,  $con$  and  $fei$ ). So, the overall model increased by five parameters in the isotopic module.

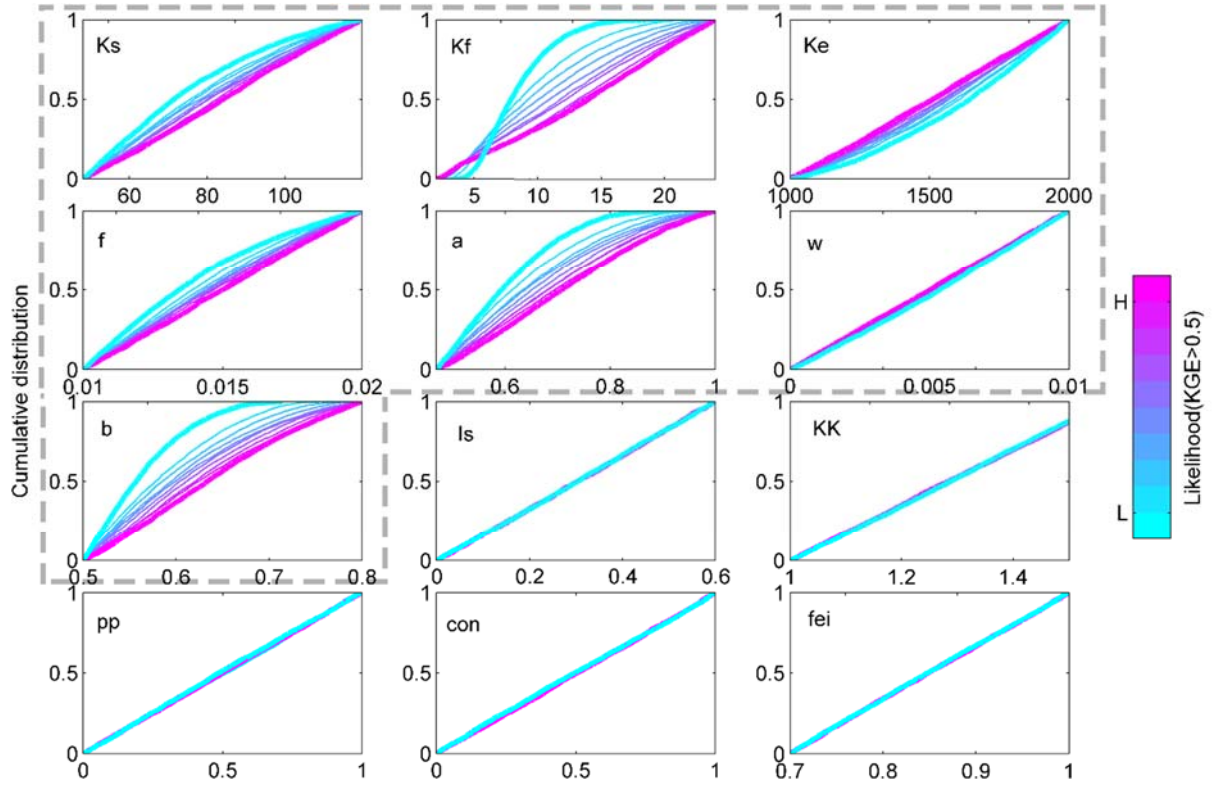
We analyzed the parameter sensitivities using either the outlet discharge and/or water isotopes. Targeting the discharge, six parameters (except  $w$ ) among the seven parameters in the flow routing module are sensitive and the parameters in the isotopic module are all insensitive (Fig r3 (a)). Targeting only isotopic values and both flow discharge and isotopic composition, the sensitive parameters are same, including  $Kf$ ,  $a$ , and  $b$  in the flow routing module, and  $Is$  and  $fei$  in the tracer module) (Fig r3 (b) and (c)). Using both flow discharge and isotopic composition as the target, these parameters were more sensitive than those using only isotopic values (see the wide ranges of the cumulative distributions in Fig r3 (c)).

Interestingly, increasing the two sensitive parameters in the isotopic module (the coefficient for evaporation fractionation  $Is$  and the weighted isotope composition of rainfall input by the parameter  $fei$ ) results in three parameters in the flow module becoming insensitive (slow reservoir constant ( $Kf$ ), the exchange constant between the two reservoirs  $Ke$  and the ratio of porosity of the quick to slow flow reservoir  $f$ ).

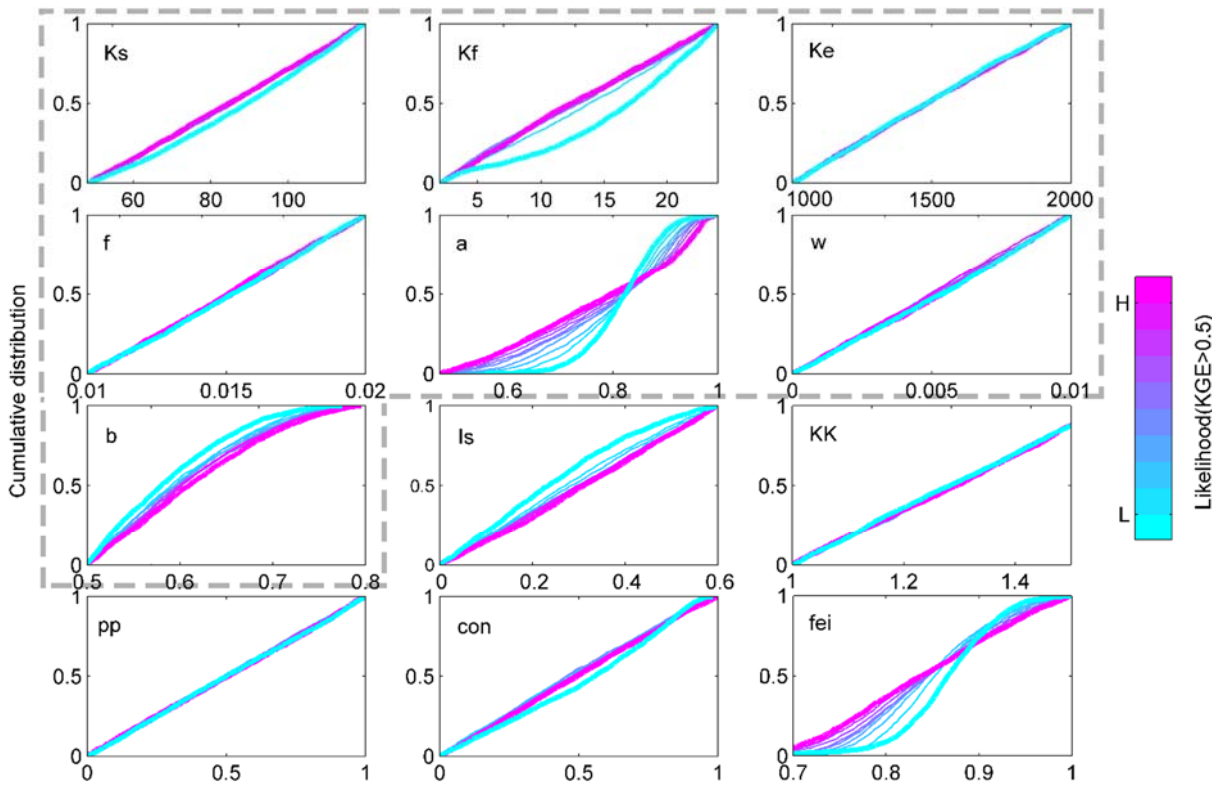
This can be explained as follows: the former two sensitive parameters in the isotopic module emphasize atmospheric effects on the outlet flow (being “old/new”). Larger  $Is$  indicates more evaporative effect on the stored water, leading to the stored and released water being older, particularly during the dry period. Larger  $fei$  indicates newer rainfall recharge (more negative isotopic values) into aquifer, leading to the stored and released water being newer during rainfall period. Alternatively, the latter three parameters in the flow module emphasize effects of fast (newer) and slow (older) flows in aquifer on the outlet flow (being “old/new”). More water release from the slow reservoir (larger  $Kf$ ) and greater release of the slow reservoir into the fast reservoir (larger  $Ke$ ) could lead to the released water being older in the dry season; a high proportion of the fast flow storage (larger  $f$ ) and a greater exchange between the fast reservoir and the slow reservoir (larger  $Ke$ ) could lead to the released water being newer in the wet season.

Consequently, there is equifinality for these parameters in the trace-aided model, which can be overcome only when we have additional data to constrain some of the parameters, e.g. knowing the evaporative effect on water  $Is$  and the weighted isotope composition of rainfall input by the parameter  $fei$ .

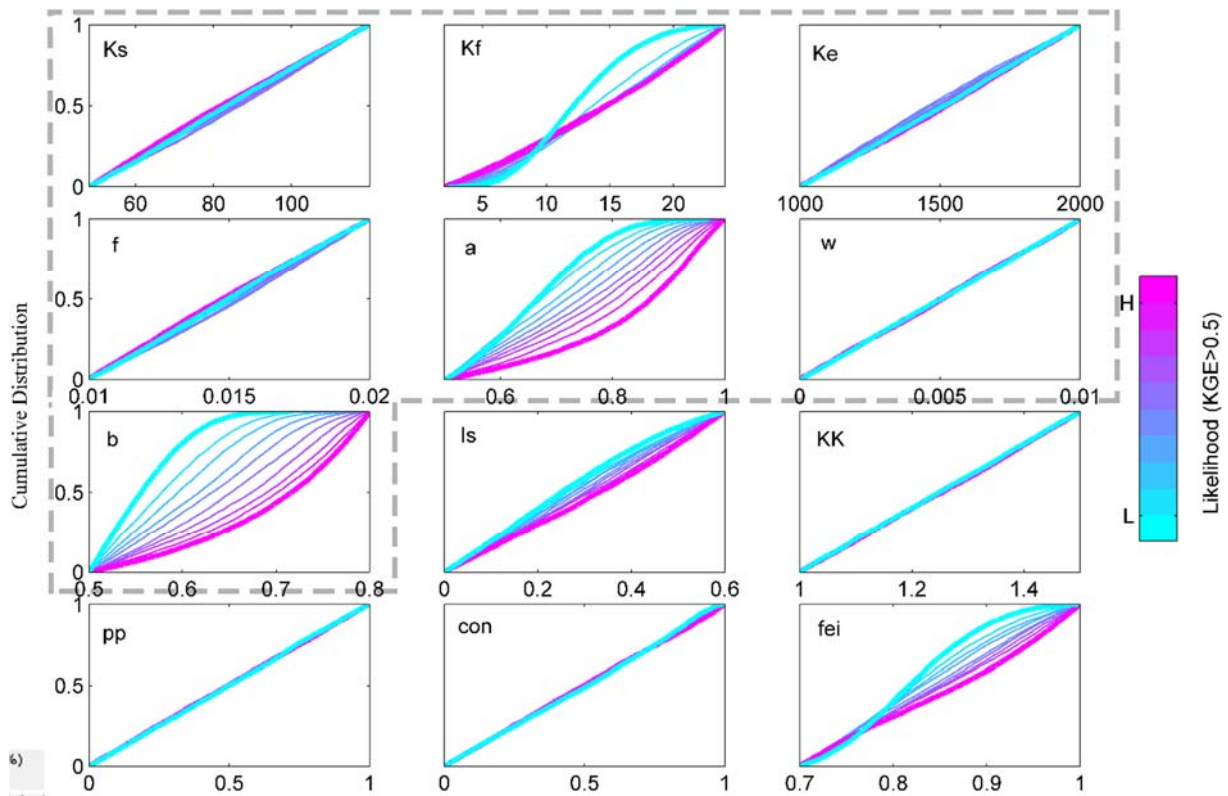
We have added the above reasoning in our discussion of the revised manuscript.



(a) Sensitive parameters include  $K_s$ ,  $K_f$ ,  $K_e$ ,  $f$ ,  $a$ , and  $b$



(b) Sensitive parameters include  $K_f$ ,  $a$ ,  $b$ ,  $l_s$  and  $fe_i$



(c) Sensitive parameters include  $K_f$ ,  $a$ ,  $b$ ,  $l_s$  and  $fe_i$

Figure r3 Sensitivity of 12 model parameters using (a) flow, (b) isotope composition and (c) combined simulation of flow and isotopic composition. (The parameters inside the gray dotted box are for flow routing, and the outside parameters are for isotope routing.)

5. With a large fraction of the model parameters insensitive, how conclusive are the interpretations on the model internal dynamics that the authors use to explain connectivity and water age distribution in the system? In some of the figures, uncertainty ranges are provided and they are quite wide. In other figures (e.g., Fig. 5), only the mean is provided although the parameters controlling the observed processes (“w” in case of Fig. 5) are insensitive.

Reply: In the revised manuscript, we have described the uncertainty of the modeled results for the various landscape units in the catchment. We reached the following conclusions:

The outlet hydrometric and isotope observations (consisting of mostly young and fast flows) were used as the calibration targets in this study. The outlet simulations had the least uncertainty, while uncertainty in the hillslope and depression units were highly related to their hydrological connectivity with the outlet. The simulated fast flows in the hillslope and depression units had lower uncertainty than the simulated slow flows in the depression since the two former units are highly connect with the outlet.

Although some parameters (e.g.  $w$  controlling hillslope flow dynamics) are insensitive, uncertainty bands of the hillslope flow (Fig r4) are narrow and the model captures quite well the hillslope

seasonality and event-based dynamics through targeting the best matching of outlet discharges and isotopic values. This indicates that the hillslope dynamics are closely linked to the outlet dynamic patterns (with strong connectivity between them), which is consistent with the ranges of  $\delta D$  and  $\delta^{18}O$  values at the hillslope spring being close to the ranges at the outlet discharge in Table 1 in the revised manuscript. (see lines 328-331 in the revised manuscript)

We discussed the modelled uncertainty in the discussions (see Section 5.4 in the revised manuscript)

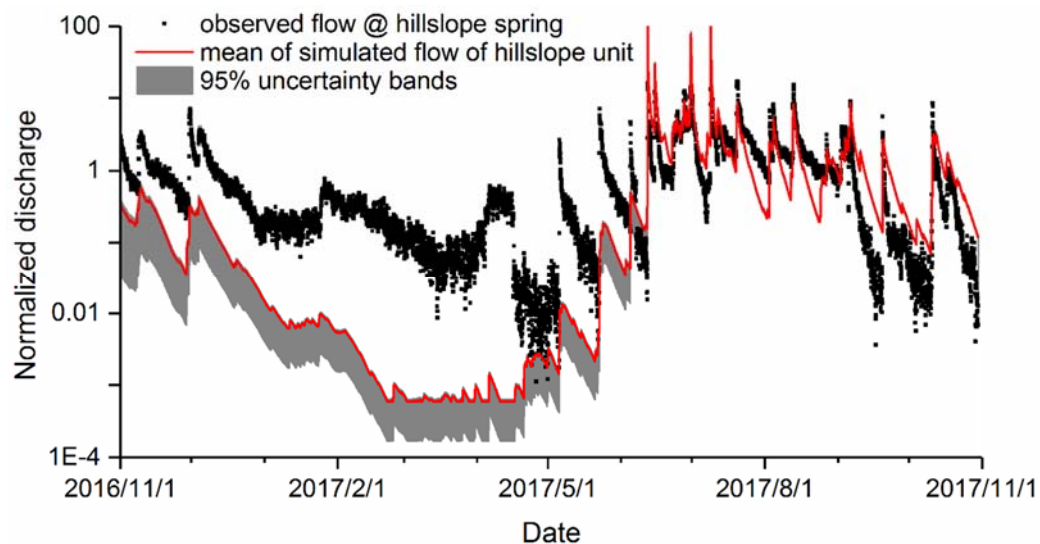


Figure r4 Observed discharge at hillslope spring against the simulated discharge of hillslope unit (values are normalized)

### Detailed comments

Line 66: There are a few studies on water storage, flux and age dynamics using tracers in karst environments.

Reply: we have revised this expression.

Line 132: If this was done before, what is the novelty of this study?

Reply: Please see response to Q3 in the main comments above.

Line 145: Is there are distinction between soil/epikarst and groundwater? What controls matrix-conduit exchange?

Reply: We have revised the descriptions. In this model, we conceptualized the groundwater aquifer in the depression by a dual flow system (involving fast and slow flow reservoirs), and the groundwater aquifer in the hillslope by an upper active storage (mostly from epikarst) mixing with a lower passive storage since the rock fractures/conduits reduce with depth from the ground surface in the hillslope profile according to our previous investigations.

The exchange between matrix and conduit is controlled by the water storage (relate to water level)

and the exchange constant between the two reservoirs ( $Ke$ ) in each reservoir.

Line 222: This is not correct - please remove

Reply: We have revised this.

Line 260: typo

Reply: We have revised this.

Line 296: only 5 of 12 parameters are sensitive, which is quite a low number. Usually, discharge contains enough information to identify 4-6 parameters (Jakeman & Hornberger, 1993). Adding of additional information like isotopes should increase this number, if the model structure is well-chosen. To check this, could the authors show the parameter sensitivities using discharge only?

Reply: Please see the reply to Q4 in the main comments above.

Line 299: In the text, a rejection limit of 0.3 is mentioned. Please clarify

Reply: Two step calibrations were carried out in this study. First,  $10^5$  different parameter combinations were selected with the broad ranges of initial parameter values. And then, we obtained the narrower ranges of the parameters according to the best modelled results (meeting the  $KGE > 0.3$  criteria). For the second calibration, the narrowed ranges of the parameters were used as the initial ranges of the parameter to search the next best modelled results ( $KEG > 0.5$ ). We have revised the descriptions in the new manuscript accordingly.

Line 308: The parameter  $w$  is completely insensitive meaning that this storage's dynamics are not well identifiable, right? Please provide all behavioral instead of the mean to show the precision of simulation of its discharge

Reply: Please see the reply to Q5 in main comments above.

Line 324: how much can you conclude from such wide uncertainty ranges?

Reply: The greater uncertainty of the modelled isotopic values in the depression arose from the insensitive parameters of  $Ke$  and  $Ks$  that affect the slow flow discharge and its exchange with the fast flow when the outlet hydrometric and isotope observations (consisting of mostly young and fast flows) used as metrics for the objective function for model calibration.

Here, the modelled isotope composition in the depression (see Figure 6b) refers to the release of water from the slow flow reservoir, representing a relatively constant source. The uncertainty bands can cover the limited variability of the measured values of  $\delta D$  at W1 and W5 (blue and yellow points in Figure 6b) where the aquifer has much lower permeability (W5) and is confined (W1) (cf the geophysical survey reported by Chen et al, 2018). This means that our tracer-aided model captures the slow flow dynamics in the depression even though the uncertainty is large.

The highly negative values of  $\delta D$  at W3 and W4 (red and black points in Figure 6b) are mostly below the uncertainty bands. This means that the stored water at W3 and W4 was younger than water from the slow flow reservoir, which is consistent with recent geophysical evidence (see Chen et al, 2018). Since W3 and W4 are located at high permeability areas, water at W3 and W4 was contributed mostly



by fast flows (mixing with the young water), particularly during rainfall events (e.g. 9/7, and 20/7 in Fig 6b). So the high negative values of  $\delta D$  at W3 and W4 below the uncertainty bands were reasonable. We have revised the descriptions accordingly in the new manuscript. (see lines 344-351 in the revised manuscript)

Line 368: KE is also quite insensitive. Can you also show the entire 500 ensemble (or confidence limits)?

Reply: The simulations from the entire 500 ensemble are shown in Fig r5 (or Fig 9 in the revised manuscript). Since the parameter of  $Ke$  that determines the exchange amount between the fast and slow flow reservoirs is insensitive, the simulated exchange flux is highly uncertainty, though much smaller, compared to the water fluxes from the rainfall recharge and hillslope flow.

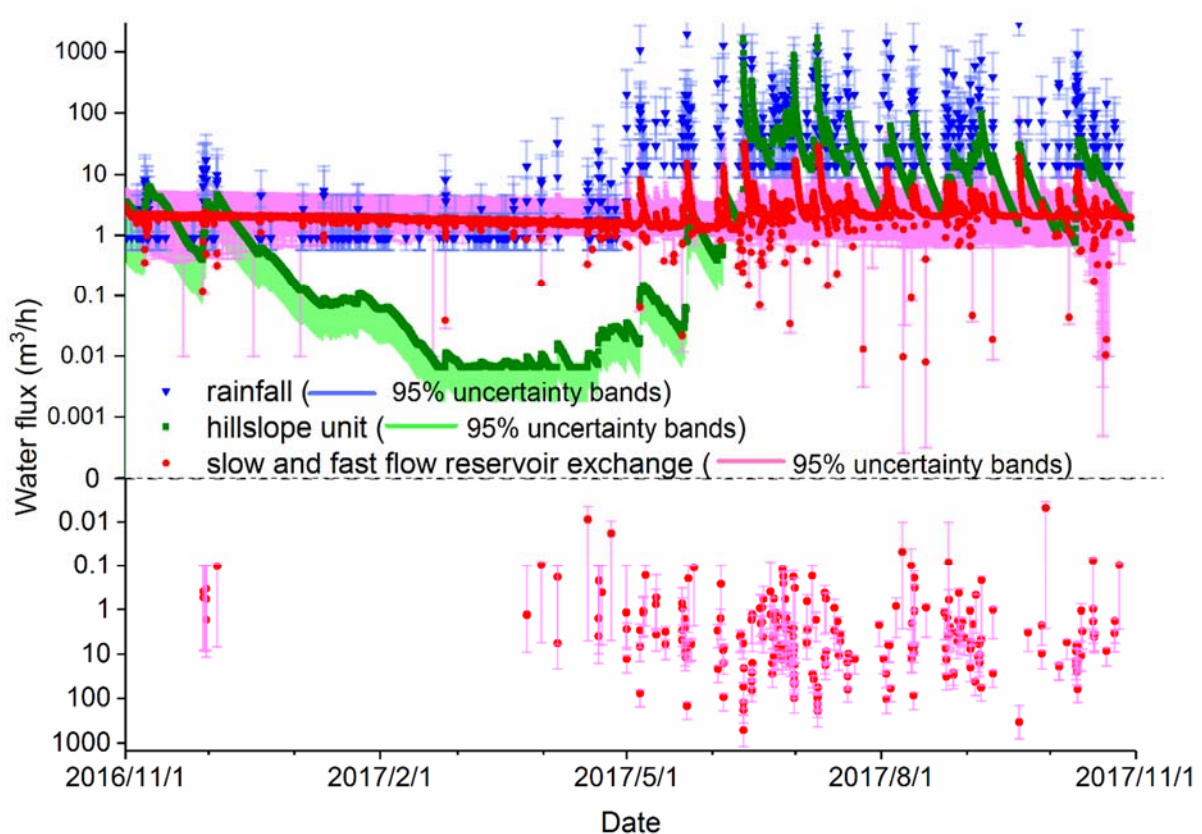


Figure r5 Source contributions to the underground stream flow (fast reservoir) at the catchment outlet. The red dots above and under the dotted line represent transient reverse water fluxes from the slow reservoir to fast reservoir and fast reservoir to slow reservoir, respectively.

Line 387: Please double-check this with literature values. Fast flow components in karst systems provide water with ages mostly between days or weeks (including temporal storage in the epikarst). Mostly, the ages found here are too large, even in the wet period.

Reply: We believe that the estimated ages are reasonable. Most models do not include mixing processes with stored water so tend to under-estimate water ages.

Here we listed  $\delta D$  values at the sampling points in this catchment for the two largest rainfall events

in 2017 (the details refer to Chen et al., 2018, <https://doi.org/10.1002/hyp.13232>).

Date	rainfall amount	rain water	outlet water	hillslope spring
12/6	86.6 mm	-85	-48 ~ -70	-62~-67
9/7	83.4mm	-80	-62~ -73	-59~ -70

It shows that  $\delta D$  values at outlet and hillslope spring are much less negative than rainwater. So there was strong mixing of the “new” rainwater with “old” stored water during and after the rainfall although the response of discharge to rainfall is fast.

Also, our estimated ages in the manuscript refer to the mean of the ages over a long period of time. For short-term (event based) responses to the rainfall, the ages of water from hillslope flow and fast reservoirs can be shortest as 4 and 2 days, respectively. There were 8 and 23 events for the fast flow with the ages of water less than 5 and 10 days, respectively (see the lowest values in Fig 9). So, the results are not inconsistent with previous work, rather capture the time-variance of water ages. We have added these explanations in the revised manuscript (see lines 412-414 in the revised manuscript).

Line 398: See comment above

Reply: The same response as for Line 387.

Line 409: Some recent example how this can be done with water quality data in karst:

Hartmann, A., Barberá, J. A., & Andreo, B. (2017). On the value of water quality data and informative flow states in karst modelling. *Hydrology and Earth System Sciences*, 21, 5971–5985.

<https://doi.org/10.5194/hess-2017-230>

Reply: We have added this relevant literature.

Line 441: Large fractions of the fast reservoir have ages larger than several months, which appears a bit slow. (see also comments above)

Reply: Please see response to Line 387.



## The list of all relevant changes in the revised manuscript

1. Some grammatical errors in the manuscript, related to the use of articles and plural/singular, have been corrected;
2. We have adjusted the Table and Figure numbers, and then revised these in the new manuscript accordingly;
3. We have deleted the Table 1 in original version and added a flow duration curve, according to the referee's suggestion;
4. Line 14-12 and Line 17-27: We have rephrased the abstract;
5. Line 52-58: We have revised this paragraph;
6. Line 61-65: We have revised these sentences;
7. Line 101: We have added the relevant literature;
8. Line 118-119: We have added the relevant descriptions;
9. Line 125-128: We have added the description of the time there is zero flow;
10. Line 130: We have added the flow duration curve;
11. Line 144-154: We have rephrased the Modelling approaches;
12. Line 177-179 and Line 193-194: We have added the equation descriptions;
13. Line 198: We have added the description of partial mixing;
14. Line 201-202 and Line 209-210: We have added the equation descriptions;
15. Line 217-218: We have added the description of the "aging effect";
16. Line 227-229: We have added the description of parameters;
17. Line 251-257: We have revised this paragraph;
18. Line 289: We have revised Fig.4b;
19. Line 316: We have revised Table 2;
20. Line 328-331: We have added the description of the uncertainty of the modeled results;
21. Line 332: The original figure have been replaced by a revised plot;
22. Line 344-351: We have revised the descriptions of the uncertainty ranges;
23. Line 369-373: We have revised this paragraph;
24. Line 399-401: We have revised this paragraph;
25. Line 402: The original figure have been replaced by a revised plot;
26. Line 412-414: We have added the description of water age;
27. Line 423-428: We have revised the descriptions of the uncertainty ranges;
28. Line 449-474: We have rephrased this paragraph about the hydrological connectivity.
29. Line 518-547: We have added the discussion of the equifinality of model parameters and uncertainty of the modelled results;
30. Line 565-568: We have revised the conclusions;
31. Line 586-587: We thanks the two anonymous reviewers and the editor for their constructive comments;
32. Line 641-642: We added the reference;
33. Line 652-653: We added the reference;
34. Line 713-714: We added the reference;
35. Line 743-746: We added the references;
36. Line 755: The Table A1 in Appendix have been replaced by a revised table;
37. We have added the supplementary material.

# Storage dynamics, hydrological connectivity and flux ages in a karst catchment: conceptual modelling using stable isotopes

Zhikai Zhang<sup>1,2,4</sup>, Xi Chen<sup>3</sup>, Qinbo Cheng<sup>1,4</sup>, Chris Soulsby<sup>2</sup>

<sup>1</sup>State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

<sup>2</sup>School of Geosciences, University of Aberdeen, Aberdeen AB24 3UF, United Kingdom

<sup>3</sup>Institute of Surface-Earth System Science, Tianjin University, Tianjin China

<sup>4</sup>College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China

*Correspondence to:* Zhikai Zhang (zhangzhikai\_0@hhu.edu.cn)

**Abstract:** We developed a new tracer-aided hydrological model that disaggregates cockpit karst terrain into the two dominant landscape units of hillslopes and depressions (with fast and slow flow systems). The new model was calibrated by using high temporal resolution hydrometric and isotope data in the outflow of the Chenqi catchment in Guizhou province of Southwest China. The model could track hourly water and isotope fluxes through each landscape unit, and estimate the associated storage and water age dynamics. From the model results we inferred that the fast flow reservoir in the depression had the smallest water storage and the slow flow reservoir the largest, with the hillslope intermediate. The estimated mean ages of water draining the hillslope unit, and the fast and slow flow reservoirs during the study period were 137, 326 and 493 days, respectively. Distinct seasonal variability in hydroclimatic conditions and associated water storage dynamics (captured by the model) were the main drivers of non-stationary hydrological connectivity between the hillslope and depression. During the dry season, slow flow in the depression contributes the largest proportion (78.4%) of flow to the underground stream draining the catchment, resulting in weak hydrological connectivity between the hillslope and depression. During the wet period, with the resulting rapid increase in storage, the hillslope unit contributes the largest proportion (57.5%) of flow to the underground stream due to the strong hydrological connectivity between the hillslope and depression. Meanwhile, the tracer-aided model can be used to identify the sources of uncertainty in the model results. Our analysis showed that the model uncertainty of the hydrological variables in the different units relies on their connectivity with the outlet when the calibration target uses only the outlet information. The model uncertainty was much lower for the “newer” water from fast flow system in the depression and flow from the hillslope unit during the wet season and higher for “older” water from the slow flow system in depression. This suggests that to constrain model parameters further, increased high resolution hydrometric and tracer data on the internal dynamics of systems (e.g. groundwater responses during low flow periods) could be used in calibration.

## 1 Introduction

Karst aquifers are characterized by complex heterogeneous and anisotropic hydrogeological conditions which are very different to most other geological formations (Bakalowicz, 2005; Ford and Williams, 2013). The hydrological function of the critical zone in cockpit karst landscapes is consequently dominated by the strong influence of this unique geomorphology and the structure of carbonate rocks. Subsurface drainage networks in karst aquifers form mixed-flow systems that integrate flow paths with markedly different velocities, ranging from low in the matrix and small fractures, to very high in large fractures and conduits (which often form subterranean channel networks), with associated transitions between states of laminar and turbulent flow (White, 2007; Worthington, 2009). Connectivity is particularly important in karst areas as the complex subsurface hydrogeology results in frequent and abrupt changes of hydrological connectivity. The system alternates through periods of varying strengths of connection and disconnection to create dynamic feedbacks, which in turn influence the systems function. Thus, understanding hydrological connectivity dynamics can provide key insights into the dominant processes governing water and solute fluxes (Lexartza-Artza and Wainwright, 2009). In the southwest karst area of China, the cockpit karst terrain encompasses flow paths sequencing in runoff generation from “hillslope to depression to stream”. The generation of hillslope runoff mostly drains into depression aquifers prior to contribution to underground streamflow. The hydrological connectivity between these units is related to not only the catchment topographic features that affect water transmission (including slope length, gradient and flow convergence e.g. Reaney et al., 2014), but also the subsurface flow connections between the fractures and conduits at any landscape units.

Due to the high spatial variability of the hydrodynamic properties of the karst critical zone, karst hydrological models are often conceptual, and are generally lumped at the catchment scale (e.g. Rimmer and Salinger, 2006; Fleury et al., 2007; Jukic and Denic-Jukic, 2009; Tritz et al., 2011, Hartmann et al., 2013; Ladouche et al., 2014). Such lumped approaches, mostly based on linear or nonlinear relationships between storage and discharge, conceptualize the physical processes at the scale of the whole catchment. In karst aquifers, the solutional conduits connect with intergranular pores and small fractures (often termed as matrix porosity), showing dual or even triple porosity zones (Worthington et al., 2017). Thus, karst aquifers are often conceptualized as dual porosity systems as residence times in the matrix are often several orders of magnitude longer than those in the conduits (Goldscheider and Drew, 2007). Accordingly, the behavior of karst spring hydrographs, is often conceptualized using a two-reservoir model to represent the dual flow system of the karst aquifer: a low permeability “slow flow” reservoir captures the function of fractured matrix blocks of the aquifer, whilst a highly permeable “fast flow” reservoir represents the larger karst conduits (Rimmer and Hartmann, 2012; Hartmann et al, 2014; Zhang et al, 2017). In addition, changing hydrological connectivity between different landscape units (e.g. hillslopes and depressions in cockpit karst areas) is often a key control on the non-linearity of the flow responses of karst systems, though this is usually not explicitly represented in most conceptual models. Consequently, developing semi-distributed lumped models is necessary to adequately

represent the hydrological function of different landscape units and the hydrological connectivity between them. For example, a conceptual model consisting of three regional phreatic aquifers (reservoirs) was proposed as the sources of three baseflow components at Mt. Hermon in Israel as the groundwater discharge patterns at three sites (the Dan, Snir and Hermon) are significantly different (Rimmer and Salingar, 2006).

The utility of tracers in karst hydrology is well-established and has given insights into advection-dispersion processes, physical exchange between conduits and smaller fractures/matrix, as well as identifying relevant contaminant transport parameters (e.g. Field and Pinsky, 2000; Goldscheider et al, 2008; Kübeck et al, 2013; Kogovsek and Petric, 2014). More generally, integration of tracers into rainfall-runoff models is becoming more common in hydrology and shows promise as such tracer-aided models can provide useful learning tools in hypothesis testing regarding water and solute transport (Birkel and Soulsby, 2015a). Indeed, McDonnell and Beven (2014) have argued that such models provide a basis for ensuring that both the celerity (i.e. the speed) of the hydrological response, along with the velocity of individual water particles (i.e. the travel times) can be captured. Moreover, they identify this as one of the fundamental challenges for contemporary hydrological modelling.

In many studies, such tracer aided models have helped to resolve this celerity-velocity dichotomy known as the “old water paradox” (Kirchner, 2003). Such integration has helped to understand the functional influence of heterogeneity in catchment landscapes; the importance of hydrological connectivity between different landscape units and the mixing processes that regulate solute transport and control water ages, as well as generating runoff responses (Jencso et al., 2010; Tetzlaff et al., 2014; Soulsby et al., 2015). Tracer-aided models that conceptualize the transport of tracers through karst systems via advection-dispersion, mixing, flow partitioning through different conduits, and exchange of tracer with the matrix have been widely used (Morales et al, 2010; Charlier et al., 2012; Mudarra et al, 2014; Dewaide et al, 2016). Using such models to simulate storage dynamics, transit times, and water ages can provide useful metrics to characterize the karst critical zone. Additionally, incorporating isotope tracers into such models facilitates multi-objective calibration, which provides the opportunity to improve the rigor of model evaluation, constrain parameter sets and potentially reduce uncertainty (Birkel et al, 2015b; Ala-Aho et al, 2017a).

The lumped models that use rather simple model structures and focus on key karst processes deemed to be dominant at particular study sites can avoid over-parameterisation (Perrin et al., 2001; Beven, 2006). Nevertheless, they have less skill in differentiating flow paths in a catchment. Whilst tracer-aided models enhance our understanding of the hydrological connectivity between different landscape units, and the associated mixing processes, they increase model parameterisation. Consequently, the effectiveness of tracer-aided models used for flow simulation and hydrological connection of the “hillslope-to- depression- to- stream” in the cockpit karst catchment needs to be evaluated.

The aim of this study is to develop a tracer-aided model that can simulate storage dynamics, hydrological connectivity and flux ages in a karst catchment and evaluate the model uncertainty of the simulation results along the “hillslope to depression

to outlet stream” continuum. The model was applied in the 1.25km<sup>2</sup> Chenqi catchment in Guizhou province of Southwest China. This catchment has typical cockpit karst landscape and associated karst critical zone architecture (Zhang et al., 2017).

95 There are detailed observations of hydrometrics and stable isotopes in hillslope springs, depression wells and at the catchment outlet, which offers an unusually rich catchment data set to evaluate model capability in karst areas.

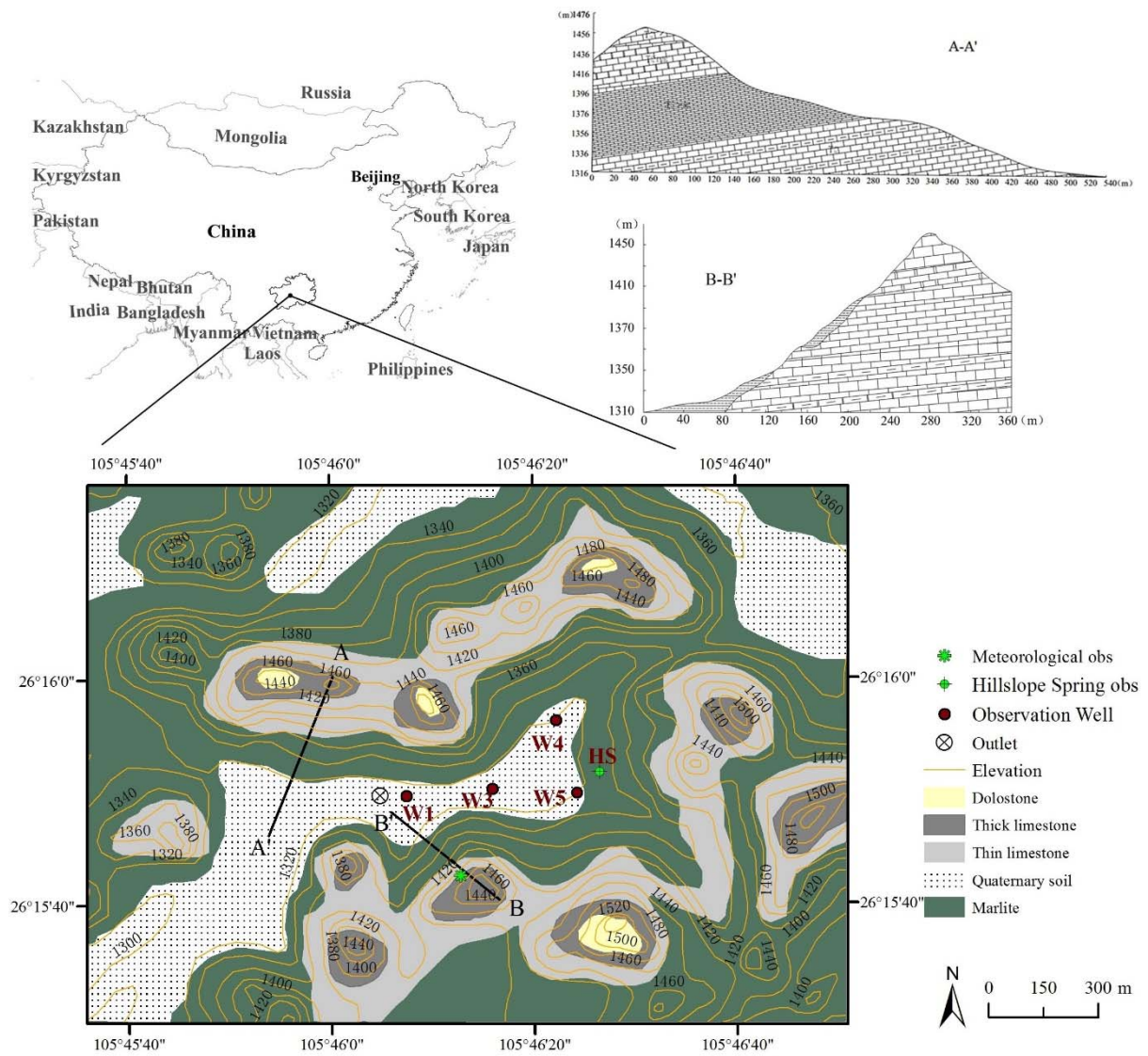
## 2 Study catchment and data

### 2.1 Study catchment

100 The study catchment of Chenqi, with an area of 1.25km<sup>2</sup>, is located in the Puding Karst Ecohydrological Observation Station in Guizhou Province of southwest China (Fig. 1). It is a typical cockpit karst landscape, with surrounding conical hills separated by star shaped valleys (Zhang et al., 2011; Chen et al., 2018). The catchment, which is drained by a single underground channel/conduit, can be divided into two units: depression areas with low elevation (<1340m) and steeper hillslopes with high elevation ranging from 1340~1500m. The spatial extent of the depression and hillslopes is 0.37 and 0.88 km<sup>2</sup>, respectively.

105 Geological strata in the basin include dolostone, thick and thin limestone, marlite and Quaternary soil profiles (see cross-sections of A-A' and B-B' in Fig 1). Limestone formations dominate the higher elevation areas with 150-200 m thickness, which lie above an impervious marlite formation. Therefore, precipitation recharging can be perched on the impervious marlite layers that discharges at the lower areas (mostly as hillslope springs). In the hillslopes, Quaternary soils are thin (less than 30cm) and irregularly developed on carbonate rocks. Outcrops of carbonate rocks cover 10-30% of the hillslope area. In the depression, soils are thick (> 2m deep). Dominant vegetation ranges from deciduous broad-leaved forest on the upper and  
110 middle parts of the steep hillslopes to corn and rice paddy in the lower gentle foot slopes and depressions, where soils are also thicker. The paddy fields are often flooded for the summer during the heavy rainfall period. Additionally, there are three sinkholes outcropping in the depression where the surface and subsurface runoff can be directly drained into the underground channel during heavy rainfall events.

115 The catchment is located in a region with a subtropical wet monsoon climate with mean annual temperature of 20.1°C, highest in July and lowest in January. Annual mean precipitation is 1140 mm, almost all falling in a distinct wet season from May to September and a dry season from October to April. Average monthly humidity is high, ranging from 74% to 78%.



**Figure 1** Map of location, geology, geomorphology and hydrological monitoring locations in the Chenqi catchment. Discharge was measured at the outlet and hillslope (HS). Water was sampled from the outlet, HS and four depression wells.

## 120 2.2 Hydrometric and isotopic data

In Chenqi catchment, the discharge of a hillslope spring (HS) located at the foot of the eastern steep hillslope, and the underground channel at the catchment outlet were measured with v-notch weirs (Fig.1). Their water levels were automatically recorded by HOB0 U20 water level logger (Onset Corporation, USA) with a time interval of 15 minutes. Additionally, an automatic weather station was established on the upper hillslope to record precipitation, air temperature, wind, radiation, air humidity and pressure. The data collection ran from 28 July 2016 to 30 October 2017. During the drought period, there were few times that flow discharges ceased (328 and 713 hours of zero flow for outlet and HS, respectively). Although the hillslope flow discharge was much smaller than the outlet discharge, their patterns in temporal variability were similar in terms of flow duration curves (Fig.2).



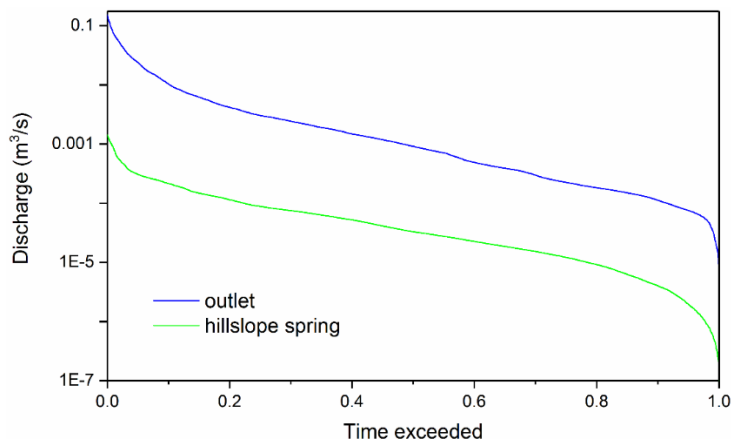


Figure 2 Flow duration curve at outlet and hillslope spring (HS) (28 July 2016 ~ 30 October 2017)

130

For isotope analysis, precipitation, the hillslope spring and catchment outlet flows were intensively sampled during eight rainfall events in the wet season (May ~ October) using an autosampler set to hourly intervals (from 12 June to 14 August 2017). Groundwater in the low elevation depressions was also sampled from four wells (Fig.1), with depth below the ground surface ranging from 13 to 35m, during four rainfall events. The well screening was installed over the whole depth for each of the wells to reflect local flow exchanges at various depths in the karst. In each event, groundwater samples were collected before, during and after rainfall at each well from multiple depths.

135

All water samples were collected in 5 ml glass vials. The stable isotope composition of  $\delta^2\text{H}$  ( $\delta\text{D}$ ) and  $\delta^{18}\text{O}$  ratios were determined using a MAT 253 laser isotope analyser (the instrument precision  $\pm 0.5\text{‰}$  for  $\delta^2\text{H}$  and  $\pm 0.1\text{‰}$  for  $\delta^{18}\text{O}$ ). Isotope ratios are reported in the d-notation using the Vienna Standard Mean Ocean Water standards. Statistical characteristics of isotope signature are summarized in Table 1.

140

Table 1 Statistical summary of isotope data for rainfall, hillslope spring (HS), catchment outlet and depression wells

Obs	$\delta\text{D}$ (‰)					$\delta^{18}\text{O}$ (‰)					lc-excess
	Max	Min	Range	Cv	Mean	Max	Min	Range	Cv	Mean	
Rainfall	-17.9	-120.2	102.3	0.3	-73.2	0	-16.4	16.4	0.29	-9.9	-0.59
Outlet	-46.9	-73.1	26.2	0.06	-61.9	-5.1	-10.6	5.5	0.09	-8.7	1.13
HS	-51.8	-77	25.2	0.04	-64.3	-5.9	-10.8	4.9	0.06	-9.3	2.91
W1	-50.7	-65.7	15	0.03	-60.8	-6.3	-9.6	3.3	0.05	-8.7	1.72
W3	-56.1	-73.6	17.5	0.06	-62.4	-7.4	-10	2.6	0.06	-8.7	0.66
W4	-55	-70.2	15.2	0.07	-62.5	-7.9	-10.1	2.2	0.07	-8.9	2.37
W5	-55.7	-67.5	11.8	0.03	-58.7	-7.9	-10.1	2.2	0.04	-8.5	2.39

### 3 Methodology

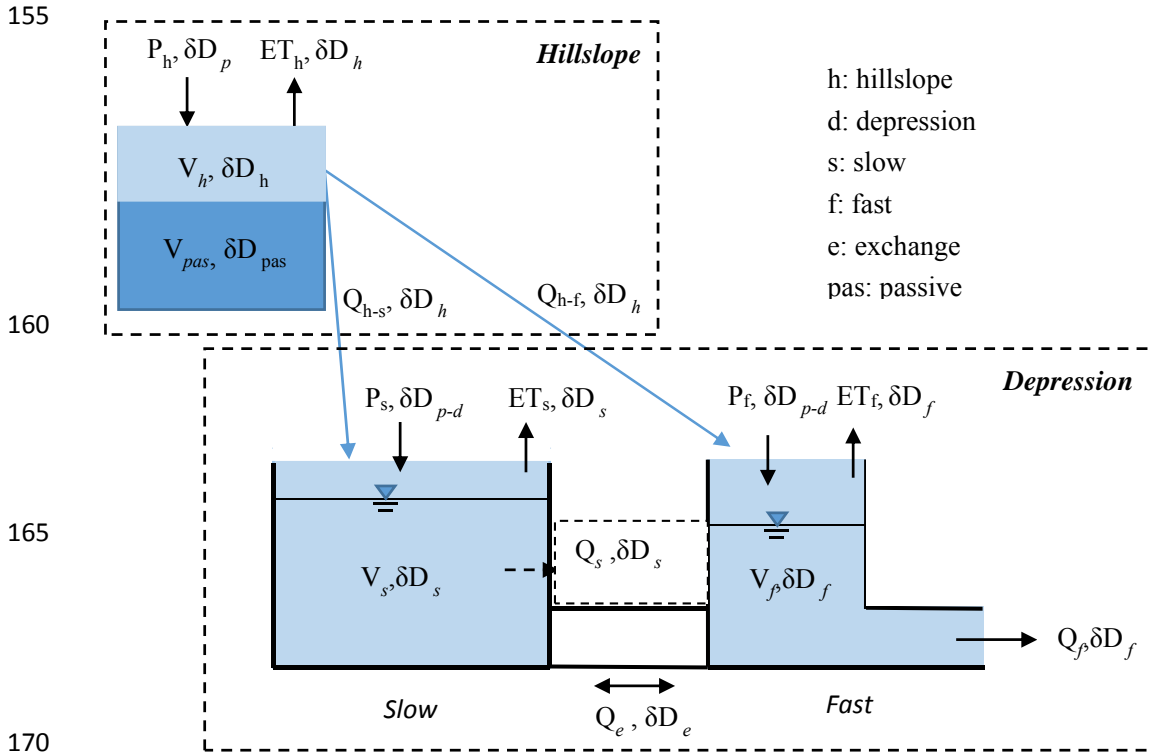
#### 3.1 Modelling approaches

The model used in this work was based on a simpler framework developed in a previous study to simulate the catchment-scale water and solute transport (Mg and Ca) in the dual flow system of the karst critical zone at daily time-steps (Zhang et al., 2017). The original model had no basis for spatially disaggregating differences in flow and tracer dynamics from different landscape

145



units. Here we significantly improved the model structure by separately conceptualizing the dominant hillslope and depression landscape units (Fig. 3), and then used the high resolution discharge and isotope data to drive the modelling at hourly time-steps. In addition, the new model has the parameters to represent passive storage inferred by isotope damping and the capacity to track the ages of water fluxes from various landscape units. As shown in Fig. 3, the Chenqi catchment was therefore subdivided into two spatially distinct units to represent the hillslope and depression. The depression unit was further conceptualized into two flow systems, represented by “fast” and “slow” flow reservoirs which could exchange water. In contrast, the hillslope unit was conceptualized as a single reservoir because of the dominant influence of the thin soil/epikarst on water movement.



**Figure 3** Structure of the coupled flow tracer model (modified from Zhang et al., 2017). Equations used to calculate state variables and storages are presented in Appendix A.

### 3.1.1 Hydrological simulation

The water balance for each of the three reservoirs (hillslope unit, fast flow and slow flow reservoirs in depression) in the catchment is expressed as follows:

$$\frac{dV_h}{dt} = P_h - ET_h - Q_{h-s} - Q_{h-f} \quad \text{for hillslope unit} \quad (1)$$

$$\frac{dV_s}{dt} = P_s - ET_s + Q_{h-s} - Q_e \quad \text{for slow flow in depression unit} \quad (2)$$

$$\frac{dV_f}{dt} = P_f - ET_f + Q_{h-f} + Q_e - Q_f \quad \text{for fast flow in depression unit} \quad (3)$$

180 where  $P$  is rainfall ( $\text{m}^3 \text{ hour}^{-1}$ ),  $ET$  is evapotranspiration ( $\text{m}^3 \text{ hour}^{-1}$ ),  $Q$  is flow discharge ( $\text{m}^3 \text{ hour}^{-1}$ ) and  $V$  is storage ( $\text{m}^3$ );  
subscripts of  $h$ ,  $s$  and  $f$  represent the hillslope, slow and fast flow reservoirs, respectively, the subscripts of  $h-s$  and  $h-f$  represent  
from hillslope reservoir to slow and fast flow reservoirs, respectively, and subscript of  $e$  represents flow exchange between  
fast and slow reservoirs. The hydrological connection and flow discharge routing for the dual flow system in the depression  
was derived by Zhang et al. (2017). Here, we further include the hydrological connectivity of the hillslope flow discharging  
185 into the depression reservoirs ( $Q_{h-s}$  and  $Q_{h-f}$  in Eqs. (1)~(3)).

### 3.1.2 Simulation of isotope ratios and estimation of water ages

The model tracks and simulates the isotope ratios for each reservoir separately, in which the isotope ratios can be completely  
or partially mixed. Experimental evidence suggests that the common complete mixing assumption is overly-simplistic, in  
particular for systems with pronounced switches between rapid shallow subsurface flow (e.g. macropores) or overland flow  
190 on the one hand and slow matrix flow on the other hand (Van Schaik et al., 2008; Legout et al., 2009). Since the depression  
unit was divided into the connected fast and slow reservoirs, complete mixing of the isotope ratios is assumed for both  
reservoirs. Thus, the isotope mass balance in the slow and fast flow reservoirs can be expressed as:

$$\frac{di_s(V_s)}{dt} = i_{p-d}P_s - i_sET_s + i_hQ_{h-s} - i_sQ_e \quad \text{for slow flow reservoir in depression} \quad (4)$$

$$\frac{di_f(V_f)}{dt} = i_{p-d}P_f - i_fET_f + i_hQ_{h-f} + i_sQ_e - i_fQ_f \quad \text{for fast flow reservoir in depression} \quad (5)$$

195 where  $i$  is the  $\delta^2\text{H}$  signature of the storage components (‰), the subscript of  $p-d$  represents rainfall infiltration in depression  
unit.

Hence, partial mixing was assumed for the hillslope (e.g. the upper active storage  $V_h$  mixing with the lower passive storage  
 $V_{pas}$  in Fig. 3 since the upper rock fractures/conducts reduce exponentially along the hillslope profile (Zhang et al., 2011))  
according to

$$200 \quad \frac{di_h(V_h)}{dt} = i_pV_{p-h} + i_{pas}V_{p-pas} - i_hET_h - i_hQ_{h-s} - i_hQ_{h-f} + i_{pas}V_{pas-in} - i_hV_{pas-in} \quad \text{for the upper active storage in} \\ \text{hillslope} \quad (6)$$

$$\frac{di_{pas}(V_{pas})}{dt} = i_pV_{p-pas} - i_{pas}V_{p-pas} + i_hV_{pas-in} - i_{pas}V_{pas-in} \quad \text{for the lower passive storage in hillslope} \quad (7)$$

The additional volumes  $V_{pas}$  ( $\text{m}^3$ ) is the storage of passive reservoir in hillslope which is available to determine isotope storage,  
mixing, and transport in a way that does not affect the dynamics of water flux volumes  $V_h$ .  $V_{pas-in}$  ( $\text{m}^3$ ) is water volume from  
205 the active store to the passive store.  $V_{p-h}$  and  $V_{p-pas}$  ( $\text{m}^3$ ) are the volume of rainfall into active and passive stores, respectively.  
To further quantify how catchment functioning affects water partitioning, storage and mixing, water ages are also tracked in  
the model. For water age estimation in the fast and slow flow reservoirs in the depression unit, complete mixing of the inputs  
is assumed and ages tracked according to determine the dynamic storage volumes on an hourly time step:

$$\frac{dAge_s(V_s)}{dt} = Age_p P_s - Age_s ET_s + Age_h Q_{h-s} - Age_s Q_e \quad \text{for slow flow reservoir in depression} \quad (8)$$

$$210 \quad \frac{dAge_f(V_f)}{dt} = Age_p P_f - Age_f ET_f + Age_h Q_{h-f} + Age_s Q_e - Age_f Q_f \quad \text{for fast flow reservoir in depression} \quad (9)$$

where  $Age$  is the water age.

For the age of the hillslope reservoir, the partial mixing is used:

$$\frac{dAge_h(V_h)}{dt} = Age_p V_{P_h} + Age_{pas} V_{P_{pas}} - Age_h ET_h - Age_h Q_{h-s} - Age_h Q_{h-f} + Age_{pas} V_{pas_{in}} - Age_h V_{pas_{in}} \quad \text{for the upper active storage} \quad (10)$$

$$215 \quad \frac{dAge_{pas}(V_{pas})}{dt} = Age_p V_{P_{pas}} - Age_{pas} V_{P_{pas}} + Age_h V_{pas_{in}} - Age_{pas} V_{pas_{in}} \quad \text{for the lower passive storage in hillslope} \quad (11)$$

where  $Age_{pas}$  is passive reservoir in hillslope. In the model implementation, each age item at the time  $t$  includes the age at the previous time step  $t-1$ . So, the results listed in this paper include the “aging effect”.

Details of the modules within the model and related equations and parameters (highlighting those calibrated) are given in Appendix A. In the equations of each module shown in Table A1, fast and slow flow reservoir storages in depression are drained by the calibrated linear rate parameters  $K_f$  and  $K_s$  ( $\text{hour}^{-1}$ ), and the exchange flow between them is calculated using the parameter  $K_e$  ( $\text{hour}^{-1}$ ) and  $f$  (Table A2). Hillslope storage is drained by the exponent parameter  $w$ ; precipitation recharging to the slow flow reservoir is calculated by the parameter  $a$  (and to the fast flow reservoir by  $1-a$ ); hillslope lateral flow to the slow reservoir is calculated by the parameter  $b$  (to fast flow reservoir by  $1-b$ ); estimation of the effects of evaporative fractionation is considered by the parameter  $Is$ ; rainfall recharge to active and passive stores in the hillslope is calculated by the parameters  $KK$  and  $pp$ ; exchange flow between active and passive stores in hillslope is calculated by the parameter  $con$ ; and the weighted isotope composition of rainfall input is calculated by the parameter  $fei$ . Therefore, the model includes 12 calibrated parameters, seven for flow routing ( $K_s$ ,  $K_f$ ,  $K_e$ ,  $f$ ,  $a$ ,  $w$  and  $b$ ) and five parameters ( $Is$ ,  $KK$ ,  $pp$ ,  $con$  and  $fei$ ) for simulation of isotope ratios and estimation of water ages. The initial range for each of the parameters is shown in Table 2.

230 Additionally, lateral surface flow can directly recharge into the fast reservoir through sinkholes in the depression in heavy rainfall events. According to research at Chenqi by Peng and Wang (2012), the mean surface runoff coefficient from the hillslopes is about 10% when the hourly rainfall amount exceeds 30mm. Hence, ten percent of rainfall infiltration of hillslope will recharge to fast flow reservoir via sinkholes in this situation (rainfall amount >30 mm/hr).

### 3.2 Modelling procedure

235 The modelling period started on 23 July 2016, but calibration was initiated using available discharge data from 1 November 2016. The preceding three months were therefore used as a spin-up period (the mean of precipitation isotope signatures over

the sampling period was used for the spin-up period) to fill storages, initialise storage tracer concentrations, and minimize the effects of initial conditions on water age calculations.

The modified Kling–Gupta efficiency (*KGE*) criterion (Kling et al., 2012) was used as the objective function for calibration.

240 The *KGE* breaks the goodness of fit into three components, so is more representative of the overall simulation than the traditionally used Nash–Sutcliffe metric which focuses on flow peaks. This overcomes some limitations of the latter (Schaeffli and Gupta, 2007) and balances how well the model captures the dynamics (correlation coefficient), bias (bias ratio) and variability (variability ratio) of the actual response. Using flow and isotopic composition as calibration targets, objective functions were combined to formulate a single measure of goodness of fit:  $KGE = (KGE_d + KGE_i) / 2$  (where  $KGE_d$  is discharge, 245 and  $KGE_i$  is isotopic composition).

The time series of discharge and isotope data were different in length. The high-resolution samples for stable isotope composition were collected over 8 events from 12 June to 13 August, 2017, giving a total of 589 samples. Hence, the  $KGE_d$  and  $KGE_i$  were each calculated using the all available data for the outlet discharge and isotope ratios, respectively. Additionally, available data such as the discharge and stable isotope signatures of hillslope spring and isotopes in the depression wells were 250 used as qualitative “soft” data to aid model evaluation. A Monte Carlo analysis was used to explore the parameter space during calibration (Table A2) and the modelling uncertainty. In order to derive a constrained parameter set, two iterations were carried out in the calibration. First, a total of  $10^5$  different parameter combinations within the initial ranges (initial range 1 in Table 2) was randomly generated as the possible parameter combinations (Soulsby et al., 2015; Xie et al., 2017). After the first calibration using  $KGE > 0.3$  as a threshold, the range of each parameter was narrowed. Then, the narrowed ranges (initial range 255 2 in Table 2) were used as for the second calibration. From the total of  $10^5$  tested different parameter combinations, only the best (in terms of the efficiency statistics) parameter populations (500 parameter sets) were retained and used for further analysis, which included calculation of simulation bounds representing posterior parameter uncertainty (Birkel et al., 2015b).

A regional sensitivity analysis (Freer et al., 1996) was further used to identify the most important model parameters. The parameter sets were split into 10 groups and ranked according to the selected objective function. For each group the likelihoods 260 were normalized by dividing by their total, and the cumulative frequency distribution was calculated and plotted. If the model performance is sensitive to a particular parameter there will be a large difference between the cumulative frequency distributions compared to a 1:1 line.

### 3.3 Line-conditioned excess

The lc-excess describes the deviation of a water sample from the Local Meteoric Water Line (LMWL) in dual-isotope space, 265 which indicates evaporation-driven kinetic fractionation of precipitation inputs (Sprenger et al., 2016; McCutcheon et al., 2017). With a known LMWL of  $\delta^2\text{H} = \alpha * \delta^{18}\text{O} + \beta$ , it was thus proposed by Landwehr and Coplen (2004) that: lc-excess =  $\delta^2\text{H} - \alpha * \delta^{18}\text{O} - \beta$ . As oxygen has a higher atomic weight, non-equilibrium fractionation during the liquid-to-vapour phase

change will preferentially evaporate (in terms of statistical expectation)  $^1\text{H}^2\text{H}^{16}\text{O}$  molecules. The isotopic signature of a water sample affected by evaporation thus shows negative  $\text{lc-excess}$  values, and plots under the LMWL in dual-isotope space (Landwehr et al., 2014). The LMWL of  $\delta^2\text{H} = 7.77 * \delta^{18}\text{O} + 4.88$  was defined based on a daily value set of isotope signature at precipitation from August 2016 to September 2017 in Chenqi catchment. The calculated  $\text{lc-excess}$  values were shown in Table 1.

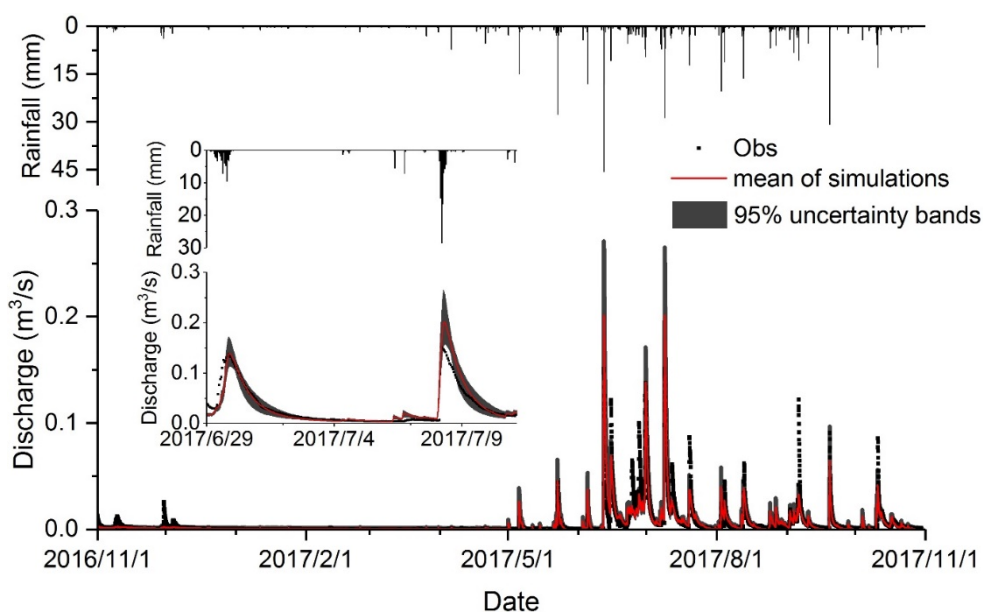
## 4 Results

### 4.1 Simulating flow and tracer dynamics

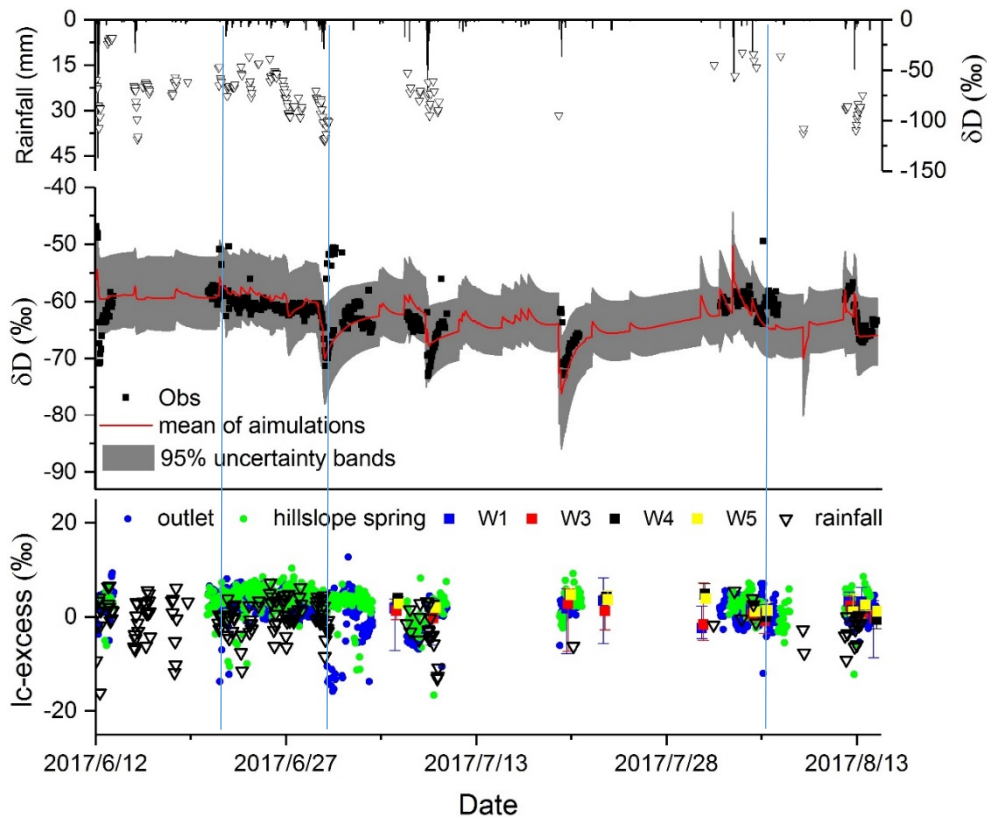
#### 4.1.1 The simulated flow and tracer at catchment outlet

The model results show that the discharge and isotope dynamics were mostly bracketed by the simulation ranges at the outlet though some peak discharges were underestimated (Fig. 4). The objective function values of the combined  $KGE$  for flow and isotopes at the outlet were all greater than 0.65 for the best 500 parameter sets (Table 2). As is common in coupled flow-tracer models, the performance in the simulation of isotopes was less satisfactory and more uncertain than for discharge;  $KEG_d \sim 0.8$  compared with  $\sim 0.5$  for  $KEG_f$ . In general isotope values in rainfall events depleted as the event progressed and this also depressed values in the underground stream, which the model generally reproduced (Fig. 4).

The sensitivity analysis results shown in Fig. 5 indicate that the fast flow reservoir constant ( $K_f$ ), the precipitation recharge coefficient for the slow flow ( $a$ ) and for fast flow reservoirs ( $1 - a$ ), the recharge coefficient of the hillslope to slow flow reservoir ( $b$ ) and to the fast flow reservoir ( $1 - b$ ), coefficient for evaporation fractionation ( $I_s$ ) and weighting constant ( $f_{ei}$ ) are generally most sensitive to the combined simulation of flow and isotopic composition.



(a) Observed and simulated stream discharge over the study period (inset shows higher resolution response over a 12 day period)



(b) Observed and simulated deuterium, and lc-excess values

**Figure 4** Observed stream discharge and deuterium during the study period, and discharge and deuterium simulations for the best 500 parameter sets; and lc-excess values of rainfall, outlet, hillslope spring and depression wells

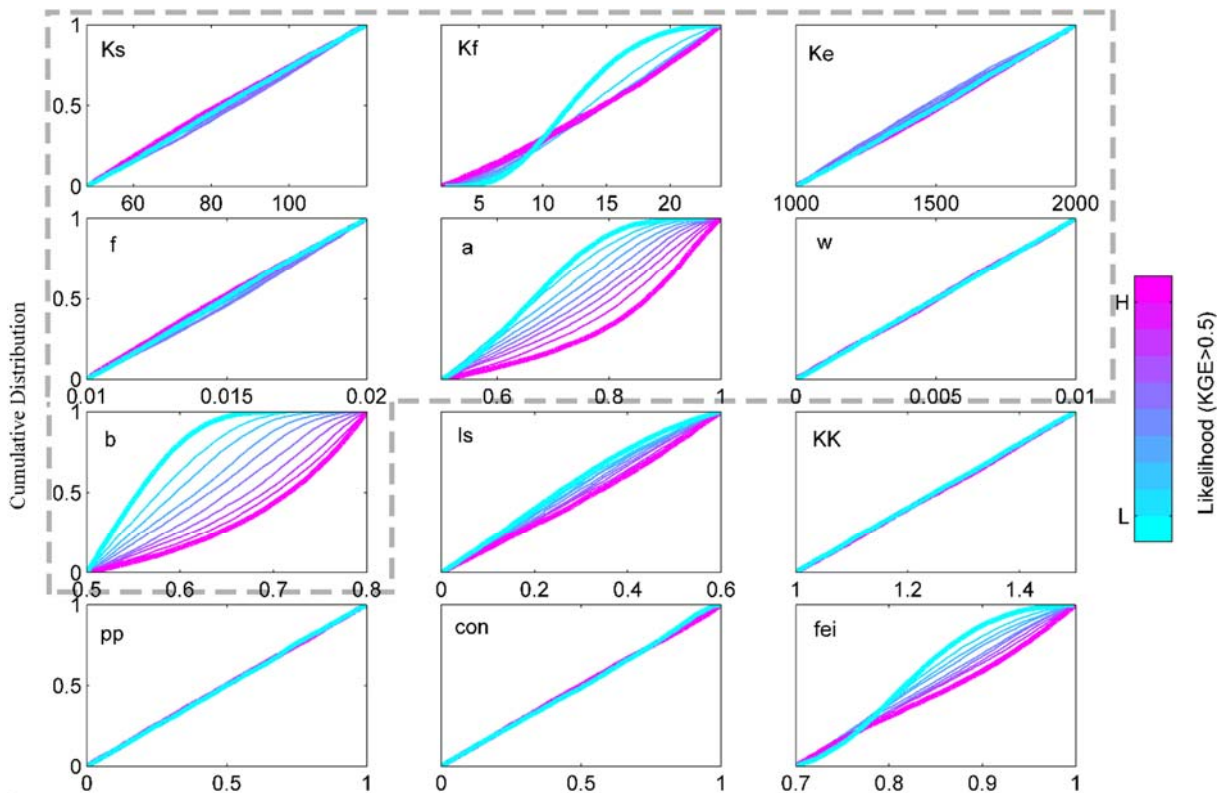
The results of isotope simulations showed that during events, the model generally reproduces the depletion of isotopes in the outlet stream in response to isotopically depleting rainfall inputs. However, the model sometimes fails to capture some high isotope values during event peaks where isotope values generally depleted (e.g. late June/early July 2017 in Fig.4b). In order to explore this further, the line-conditioned excess (lc-excess) of samples was calculated from samples. The results of lc-excess values are in Fig. 4b, with the mean of -0.59 and 1.13 ‰ for rainfall and the outlet, respectively (Table 1). There were a few samples which showed markedly negative lc-excess values around event peaks (e.g. 22/6, 1/7, and 5/8), indicating a strong fractionation effect. These outliers correspond to the unexpected clusters of enriched isotope values that the model fails to capture.

The lc-excess of the isotope time series of the hillslope spring and wells in the depression were also calculated (Fig. 4b). The mean lc-excess values for the hillslope spring, W1, W3, W4 and W5 for depression wells were shown in Table 1. While the mean is slightly positive, negative values are common indicating an evaporative fractionation effect on recharge water. However, the underground stream flow (mainly reflecting the response of “fast flow” reservoirs) with the unexpected “outliers” with high isotope values could not be attributed to the hillslope response or groundwater in the depression (the maximum of  $\delta D$  less than -50 ‰ in Table 1), because the lc-excess values of these sources were substantially less negative than the simultaneous values of the underground stream (Fig. 4b) and the maximum of  $\delta D$  (-46.9 ‰) at outlet was higher than that at hillslope spring and depression wells (less than -50‰) in Table 1. The most likely explanation relates to flooded paddy fields

which are extensively distributed in the depression during the growing season. Consequently, significant volumes of surface water are impounded in the paddy fields and exposed for evaporative fractionation. Therefore the markedly enriched isotope signals at the outlet around some event peaks would be consistent with fractionated water being displaced from the paddy fields and entering the fast flow system. This would explain the model's lack of skill in capturing such effects of evaporative fractionation.

**Table 2 Mean parameter values and fitness derived from the best 500 parameter sets after calibration**

For Flow	$K_s$ (hour <sup>-1</sup> )	$K_f$ (hour <sup>-1</sup> )	$K_e$ (hour <sup>-1</sup> )	$f$	$a$	$W$	$b$
Initial range 1	40-168	1-72	800-2200	0.005-0.025	0-1	0-0.015	0-1
Initial range 2	40-150	1-40	800-2200	0.008-0.025	0.47-1	0-0.015	0.48-1
Mean	92	11	1549	0.015	0.68	0.005	0.54
Range	48-120	5-18	1000-2000	0.01-0.02	0.51-0.9	0.003-0.01	0.5-0.62
For Isotope	$I_s$	$KK$ ( $\times 10^4$ )	$pp$	$con$	$fei$	Index	Mean(range)
Initial range 1	0-1	0.8-1.6	0-1	0-1	0-1	$KGE_a$	0.85 (0.81-0.87)
Initial range 2	0-0.8	0.8-1.6	0-1	0-1	0.5-1	$KGE_i$	0.56 (0.52-0.59)
Mean	0.24	1.26	0.49	0.56	0.82	$KGE$	0.7 (0.72-0.66)
Range	0.002-0.6	1-1.5	0.02-0.95	0.04-0.97	0.71-0.93		

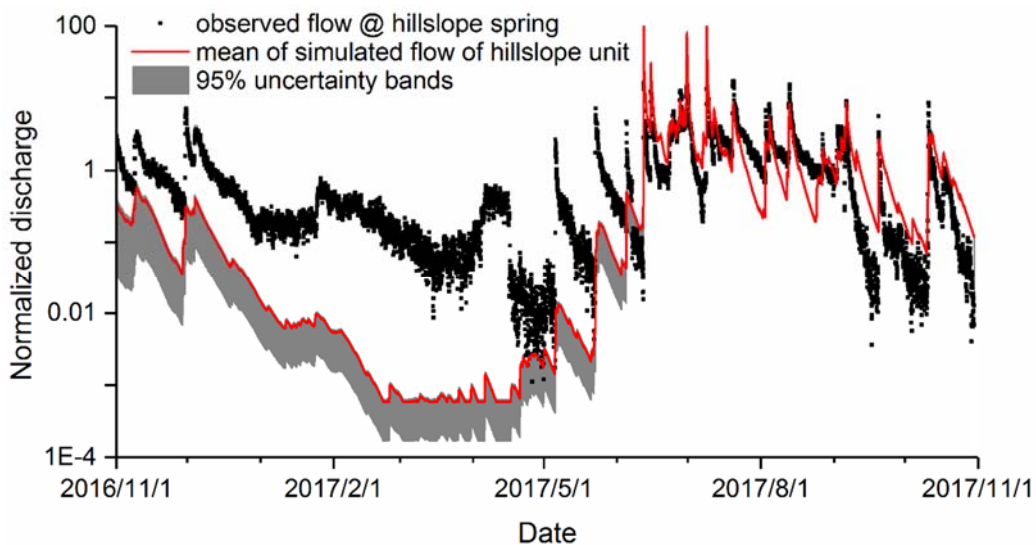


**Figure 5 Sensitivity of 12 model parameters expressed as cumulative distributions in ten levels of likelihood values for the model simulations from the lowest likelihood value (blue) to the highest likelihood value (purple). Likelihood based on  $KEG$  and rejection of values  $<0.5$ . (The parameters inside the grey dotted box are for flow routing, and the outside parameters are for isotope routing.)**



#### 4.1.2 The simulated flow and tracer for hillslope spring and depression wells

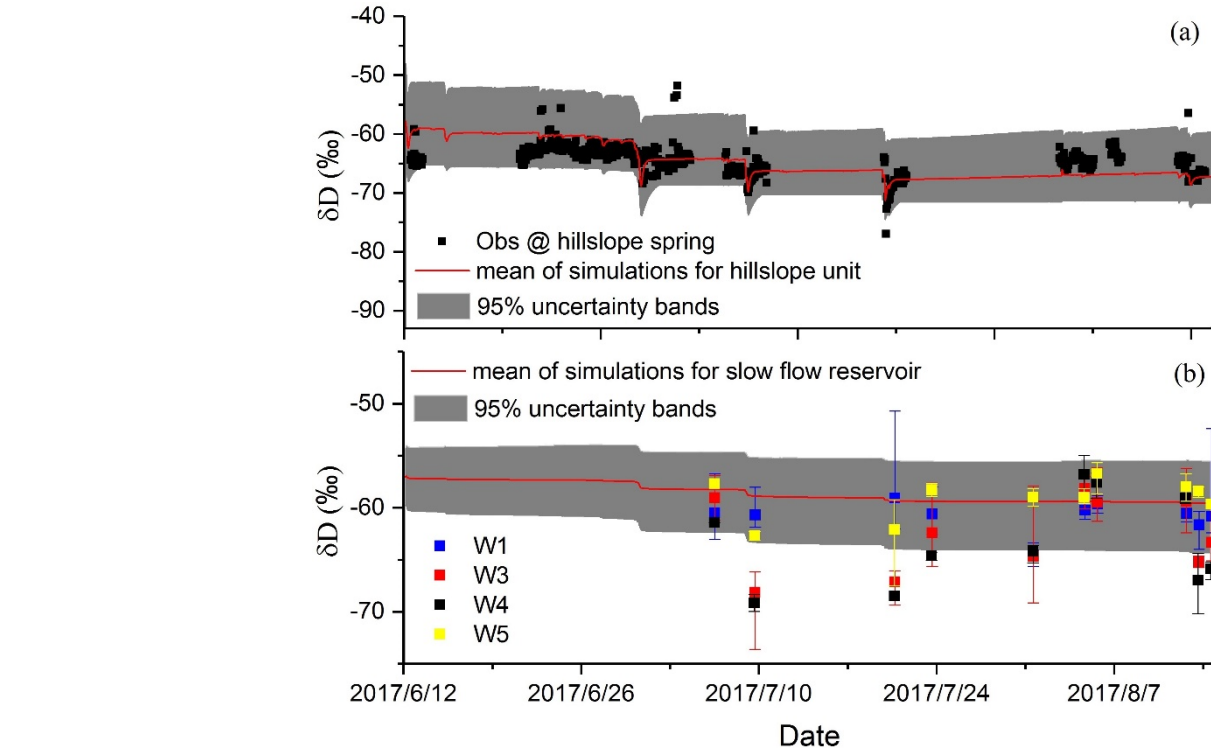
As a more qualitative indication of model performance, Fig.6 shows the normalized simulated discharge ( $Q_n=Q_i/Q_{mean}$ ) of the hillslope unit had very similar seasonality and event-based dynamics to the normalized observed discharge at the hillslope spring. The magnitude of the modelled discharge fluxes is, of course, different to those observed at the specific hillslope (e.g. HS at the east hillslope in Fig.1) because the simulation results represented the lumped outputs of the whole hillslope unit. However, as a “soft” validation of the model it adds confidence that the temporal dynamics of the hillslope response are appropriately captured. Additionally, the measured flow duration curves (Fig. 2) and isotopic values (the ranges of  $\delta D$  and  $\delta^{18}O$  values in Table 1) of the hillslope spring are broadly similar to those measured at the outlet. This, perhaps explains why the model, although using only flow and isotope data from the catchment outlet time series for calibration, is able to capture the dynamics of the hillslope, even though the hillslope drainage parameter  $w$  is insensitive.



**Figure 6** Observed discharge at hillslope spring (in Fig.1) against the simulated discharge of hillslope unit (mean of the simulations for the best 500 parameter sets). Note values are normalized.

This is further illustrated through an additional qualitative evaluation of the model given by comparing the internal tracking of isotope dynamics in the conceptual stores of the hillslope and slow flow reservoir with respective data available from measured isotope values collected for the hillslope spring and wells (Fig. 7). The sampling frequency of the hillslope spring was same as at the outlet; however, there were only 10 sampling occasions from the depression wells over the dry and wet season, and the water samples were collected across a range of depths of the well. Again, although these point measurements are not strictly comparable with the tracked isotope composition of conceptual stores, they do give an indication that the internal states of the model are being plausibly simulated in terms of the mixing volumes which damp the isotope inputs in precipitation. These results are again encouraging, showing that the model captures the general directions of changes in the isotope dynamics of the hillslope spring, albeit with a relatively high degree of uncertainty (Fig.7a).

345 The modelled isotope composition in the depression in Fig.7b shows the release of water from the slow flow reservoir, representing a relatively stable, well-mixed source. The uncertainty bands cover the limited variability of the measured values of  $\delta D$  at W1 and W5 (blue and yellow points in Fig.7b) where the aquifer has much lower permeability (W5) and is confined (W1) (cf. the geophysical survey reported by Chen et al, 2018). This indicates that our tracer-aided model captures the general slow flow dynamics in the depression even though the uncertainty is large. The highly negative values of  $\delta D$  at W3 and W4 (red and black points in Fig.7b) mostly plot below the uncertainty bands. This is reasonable as water at W3 and W4 has been shown by Chen et al. (2018) to be mostly contributed by faster flows (mixing with the young water) in high permeability areas, particularly during rainfall events (e.g. 9/7, and 20/7 in Fig.7b).



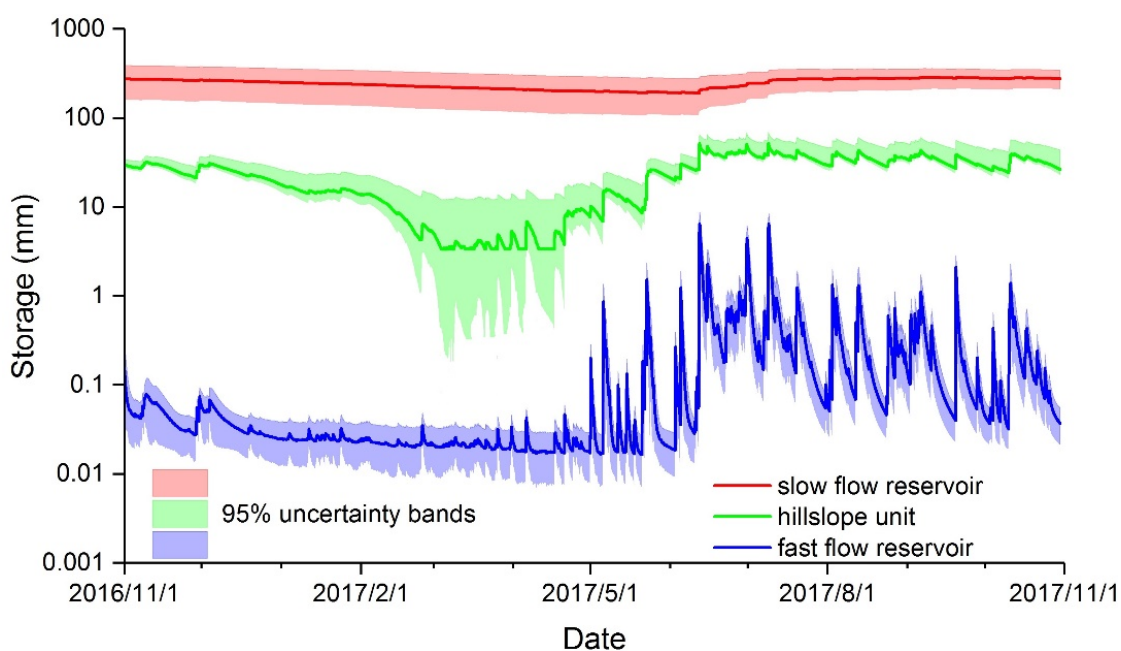
**Figure 7** Modelled isotope signature at hillslope unit and slow reservoir vs observation at hillslope spring and depression wells (the red line represents mean of the simulations for the best 500 parameter sets).

#### 355 4.2 Storage dynamics of different reservoirs and source contributions of underground stream flow

The storage dynamics of the catchment derived from the model in order to simulate the concurrent flow and tracer response can be disaggregated according to the conceptual stores (Fig. 8). The model structure dictates that the main variability in the runoff response to precipitation is driven by the storage dynamics, depending on hydrological connectivity between the hillslope ( $V_h$ ), slow ( $V_s$ ) and fast ( $V_f$ ) flow reservoirs (Fig. 3). The modeled storage results show that slow flow reservoir was the largest store in the catchment (>100mm with mean of 245 mm), consistent with the wide distribution of small fractures and matrix pores in the karst critical zone (Zhang et al., 2011, 2017). The fast flow reservoir had the smallest storage (the mean value was only 0.2mm) because the underground river/conduit volume represents only a very small proportion of the porosity of the entire aquifer. Although the hillslopes cover a larger area than the depression, the thin soil, shallow epikarst and rapid

drainage resulted in a relatively small dynamic storage reservoir, with a calibrated mean value of 23mm. The discharge over the study period showed clear seasonality, which reflects the uneven distribution of precipitation throughout the year (Fig.4a). This seasonality is mirrored somewhat differently in the storage dynamics of each reservoir (Fig.8). The storage change in fast flow reservoir was very rapid, especially in the wet season; reflecting rapid recharge and water release. The rapid response of storage to rainfall was also evident in hillslope reservoir because of the low capacity and short response time.

Using flow and isotopes at the catchment outlet as calibration targets, the uncertainty bands for the three storages increase in the order: fast flow in the depression < hillslope flow < slow flow in the depression (Fig. 8). Additionally, the uncertainty bands become narrower in the wetter period (Figs. 6 and 8). This indicates that the model structure along with the calibration targets emphasizes the rapid flow component, and the modelling uncertainty increases when flux components from storage units are less closely correlated with the outlet discharge used as the calibration target.

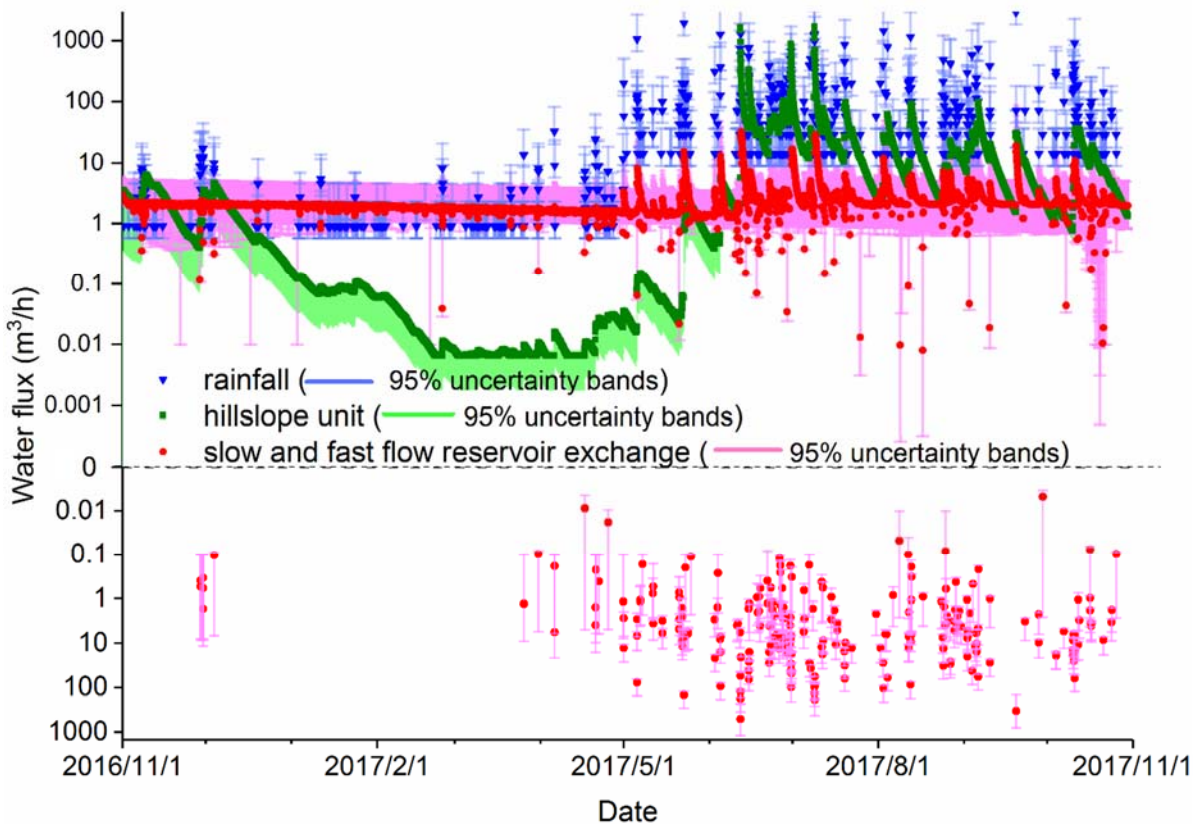


**Figure 8** Model-derived storage dynamics in hillslope ( $V_h$ ), slow ( $V_s$ ) and fast ( $V_f$ ) conceptual reservoirs for the best 500 parameter sets, and the lines represent the mean of simulations

The relative contributions of the different sources to stream flow changed with hydroclimatic conditions, and this could be estimated using the calibrated model. Fig.9 shows that during the dry period (November 2016 to April 2017), underground stream flow was mainly sustained from the small fractures (conceptualized as release from the slow flow reservoir). Overall, this provided the largest proportion (78.4%) of dry season flows, followed by the hillslope unit contribution (16.8%). During this period, direct rainfall infiltration contributed only limited water to the underground stream (4.8%) due to the low rainfall and limited storage, resulting in weak hydrological connectivity between the hillslope and depression. During the wet period, with the resulting rapid increase in storage, the hillslope unit contributes much more water to the underground stream, accounting for the largest proportion (57.5%) of overall flow, due to the strong hydrological connectivity between the hillslope and depression. Meantime, the contribution of direct rainfall infiltration to the underground stream flow also increased (with

an overall wet season contribution of 35.6%). This likely reflects the increased influence of sinkholes and larger fractures as the catchment becomes wetter. In such conditions, during storm events, overland flow and epikarst water were collected by sinkholes and large fractures and recharged to underground stream directly. The relative contribution of small fractures in the slow reservoir decreased substantially (7%) although the overall magnitude of the water flux to the underground stream increased during the wet period.

The bi-directional exchange between the underground conduit and the surrounding porous matrix is a unique feature of the karst critical zone (Zhang et al, 2017). During the dry period, as water table levels in the conduits drop more rapidly than in the matrix, water stored in the matrix drains into conduits and underground channels as baseflow. In the wet season, especially during the periods of highest flow, infiltrated water quickly fills conduits where the water table is higher than the adjacent matrix. Water is temporarily stored in the conduits, and hence induces recharge in the matrix. These bi-directional exchange flows between the underground channels and the matrix were captured by the model (represented by fast and slow reservoirs, respectively) and are shown in Fig. 9, where the negative values represent the flux from conduits to the adjacent matrix. This bi-directional flow was affected by the wetness conditions, being evident in the wet season and indicating both the seasonal and short-term temporal change of hydrological connectivity between the fast and slow flow reservoirs. Despite this, as the parameter  $K_e$  that determines the exchange between the fast and slow flow reservoirs is insensitive (Fig. 5), the simulated exchange flux is more uncertain (red lines in Fig. 9) compared with the water fluxes from the direct rainfall and hillslope flow.



**Figure 9** Source contributions to the underground stream flow (fast reservoir) at the catchment outlet (mean of the simulations for the best 500 parameter sets). The red dots above and under the dotted line represent transient reverse water fluxes from the slow reservoir to fast reservoir and fast reservoir to slow reservoir, respectively.

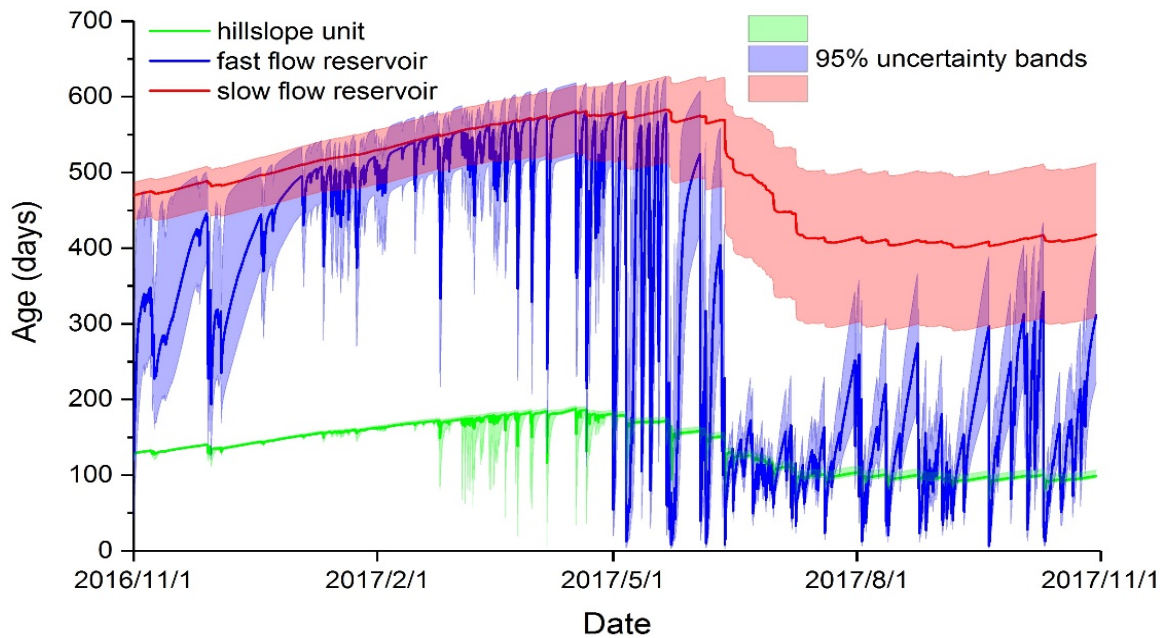
### 4.3 Simulated flux ages from different conceptual stores

The ages of water fluxes from the different landscape units were tracked using the model. The simulated ages were linked to the size of storage in each unit and the ages decreasing in the order of hillslope reservoir < fast flow reservoir < slow flow reservoir, with mean ages of 137, 326 and 493 days over the study period, respectively (Fig.10). The mean ages of water flux decreased between the dry and wet seasons: ranging from 159, 466 and 528 days for the dry season to 115, 187 and 458 days for the wet season, for the hillslope, fast flow and slow flow reservoirs, respectively. The ages of fluxes from hillslope flow and fast flow reservoirs change greatly for each of the rainfall events. For short-term (event based) responses to the rainfall, the ages of water from hillslope flow and fast reservoirs can be shortest as 4 and 2 days, respectively. There were 8 and 23 events for the fast flow when the ages of water were less than 5 and 10 days, respectively (see the lowest values in Fig. 10).

The ages of fluxes from the fast flow reservoir in the underground stream generally reflected the integration of younger water fluxes from the hillslope and older fluxes from the slow flow reservoir, as shown in the time variant flux ages shown in Fig.10. Consequently, the water age dynamics of the fast flow reservoir were relatively close to the slow flow reservoir in the dry season and close to hillslope reservoir in wet season as connectivity changed. This is consistent with the changing storage dynamics shown above. However, a distinct feature in Fig.10 is that the water ages in the fast flow reservoir were younger than which from hillslope reservoir during some events in the wet period. This again, most likely reflects the role of sinkholes in collecting water with a high proportion of new rainfall (young water) in intense wet season rain events and then recharging underground stream rapidly due to the direct, transient connectivity.

Fig. 10 also shows that uncertainty bands increase with age in the three water fluxes, i.e. narrowest for the youngest hillslope flow and widest for the oldest slow flow. However, the seasonal changes of uncertainty bands are different for the three water fluxes. For the hillslope flow and the fast flow in depression, the uncertainty bands reduce in the wet period as ages decrease. In contrast, for the slow flow reservoir, uncertainty bands increase during wet period (Fig. 10). This underlines the resulting uncertainty for the slow reservoir, reflecting the structural limitations with the model for conceptualising the flow dynamics of this heterogenous zone during the rainfall season.

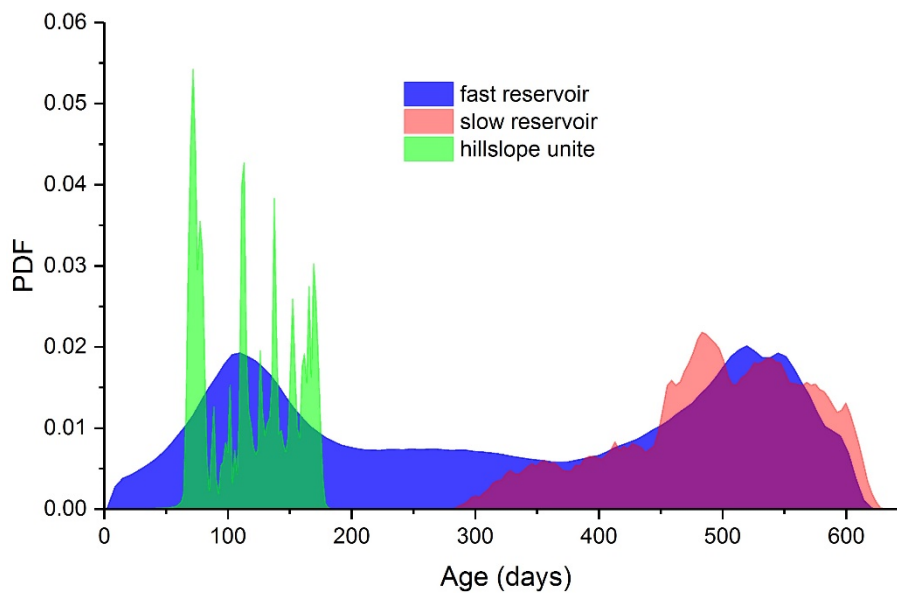




430 **Figure 10** Mean of water flux age of hillslope, fast and slow reservoir (for the 500 best parameter sets)

The probability density functions (PDFs) of the simulated flux ages from the three reservoirs are shown in Fig.11 (using the best 500 parameter sets). The ages of fluxes from the fast flow reservoir varied from a few days to over 600 days, and it is clearly evident that the PDF was bimodal with peaks corresponding to water ages of ~100 and ~550 days. From the water age dynamics in Fig.10, it is equally clear that the bimodal distribution of ages of underground stream flow reflected the seasonality of different water sources contributions in the wet and dry seasons. The underground stream flow was dominated by older water from the matrix and small fractures during the dry period and by younger hillslope fluxes during wet period, respectively.

435 The water age distributions for the hillslope also showed seasonal bimodality in flux ages, albeit less pronounced, though the model has also produced a less smooth distribution of more transient younger ages.



440 **Figure 11** Probability density functions of the simulated water ages for the best 500 runs in fluxes for all three reservoirs



## 5 Discussion

In karst areas, complex subsurface flow systems, with high spatial heterogeneity in porosity and structure, and marked temporal variations in hydrological connectivity, dictates that the karst critical zone is a particular challenge to hydrological modelling. Tracer-aided conceptual modelling is helpful in understanding karst regions, because using isotope tracers as “fingerprints” means that hydrological processes can be tracked in a way that provides insights into storage dynamics and can resolve “fast” and “slow” water fluxes and estimate their ages with different units of a catchment. This supports other recent studies that water quality data can help inform and constrain modelling in karst environments (e.g. Hartmann et al., 2017).

### 5.1 Hydrological connectivity between different landscape units

Since the hillslope-depression is a typical landform with variable hydrological connectivity in the karst catchments in southwest of China and elsewhere, the separation of the hillslope and depression in the new model structure improved model performance and yielded more informative results showing clearly the flow and tracer dynamics within different landscape units, as well as tracking spatially distributed storages and ages of water flux. The model was successfully calibrated to the flow and tracer dynamics at the catchment outlet; the results also showed more qualitative consistency performance in terms of the dynamics of the modelled hillslope fluxes compared with spring discharge and the simulated isotopic composition of fluxes with measurements in the spring and wells. Moreover, the modelling approach is potentially transferable to other cockpit karst catchments with similar landscape organization.

The tracer-aided model supports general appropriateness of the model structure which related connectivity dynamics to storage change within different landscape units. During the dry period, there is weak hydrological connectivity between the hillslope and depression due to low storage. In contrast, during the wet period, hydrological connectivity between the hillslope and depression strengthens as water storage increases. In the early recession, after heavy rain, large fractures in the hillslope fill, leading to large water fluxes into the depression. Then, as storage declines, fluxes decrease and the hydrological connectivity weakens. Moreover, in each of the units, there is hydrological connectivity and exchange between dual porosity systems that were conceptualized as the slow and fast flow reservoirs in this study. The hydrological connectivity and exchange between the slow and fast flow reservoirs is mainly controlled by the water level of each medium, rather than the storage. The flow directionality will change with the hydraulic gradient between the two reservoirs. The bidirectional water flux makes it fundamental to consider the directionality of connectivity within the karst critical zone. Direct hydrological connectivity between the surface and subsurface is also important in stream flow generation in karst catchments. Besides infiltration through fractures and the matrix, concentrated infiltration from surface to underground flow systems via sinkholes is a unique aspect of transient connectivity in karst catchments. This influence is captured by the model conceptualisation and shown in the contribution of rainfall to the underground stream in Fig. 9. Although this hydrological connection only occurs during heavy rain in the wet season, it is one of the most distinct hydrological functions of the karst critical zone. In this regard, flow paths

in urban areas, with transient connectivity of storm drains, have been compared to karst (Bonneau et al. 2017); and whilst this gives similar short response times and a dominance of young water (Soulsby et al., 2014) urban systems are simpler and bi-directional connectivity is less significant.

## 475 **5.2 Water ages of different conceptual stores**

Through characterizing the variable water ages of different landscape units, we can deepen our understanding of the non-linear water storage dynamics and runoff generation processes (Soulsby et al, 2015). Water ages reflect the time variance and non-linearities of how different runoff sources are connected and the dynamics of their relative contribution to runoff generation (Birkel et al., 2012). Recent work has demonstrated the controlling effect of hydrogeological conditions on water ages in karst areas (Mueller et al, 2013). The underground stream water ages at the catchment outlet can be viewed as the time-varying integration of spatially distributed water fluxes from the hillslope unit and small fractures in the depression aquifer, which each have their own age dynamics (Fig.10). There is a distinct pattern of bimodality in the age distribution of underground stream flow (represented by the fast reservoir in Fig.11), which reflects the seasonality of the different water sources. Younger waters mainly come from surface water recharge through sinkholes after heavy rain and drainage from the hillslopes, whilst the slow flow reservoir dominates low flows. According to the water age dynamics of different conceptual stores, it can be deduced that storage-driven changes in hydrological connectivity and associated mixing processes largely determine the nonstationary water age distribution of the underground stream. In this sense, karst catchments seem to be subject to the “inverse storage effect”, where periods of high storage facilitate release of younger water to drainage (Harman, 2014).

It should be noted that the ages derived from the modelling are based on stable isotope tracers, which whilst well-suited to characterizing the influence of younger waters, are less well-suited to constraining the age of older waters (>5 years) that may be present in deeper aquifers and fine pores that contribute to the slow flow reservoir (e.g. Jasechko et al., 2017). Thus further work is needed to assess the role of these older waters and quantify their influence on the ages of water in the underground channel (e.g. McDonnell and Beven, 2014). That said, the dominance of younger water in the outflow of responsive karst catchments is consistent with recent theoretical (Berghuijs and Kirchner 2017), larger scale (Jasechko et al., 2016) and more local studies (Ala-aho et al., 2017b) which show that deeper, oldest groundwater often makes insignificant or limited contributions to stream flow.

## 5.3 High temporal resolution isotope data for karst area

There is a marked shift in the isotopic composition of storm event rainfall which effects the short-term response of the catchment outlet, thus weekly or even daily isotope data would not adequately capture the variability of rainfall isotope signatures at a resolution appropriate to the response times of sub-tropical karst systems (Coplen et al., 2008). The assessment of water ages in the critical zone is highly dependent on the temporal resolution of tracer data in rainfall and stream flow for

model conceptualization (Birkel et al., 2012; McDonnell and Beven, 2014). The high-frequency measurements of tracer behaviour enhanced our understanding of catchments' hydrological function and the associated time scales of the celerity of the hydrological response to rainfall inputs and the velocity of water particles. Also, the high-resolution tracer data yielded novel insights into how the model integrates and aggregates the intrinsic complexity and heterogeneity of catchments, in order to reproduce behaviour adequately across a range of time scales (Kirchner et al., 2004). However, it should be noted that in much previous tracer-based modelling, the temporal resolution of hydrometric data (typically hourly) is at a much finer temporal resolution than tracer data, sampled more often at daily or even weekly resolution (Stets et al., 2010; Birkel et al., 2010, 2011; McMillan et al., 2012; Soulsby et al., 2015; Ala-Aho et al., 2017a). Here, due to the marked heterogeneity of flow paths in the karst critical zone, and the very rapid (i.e. sub-daily) stream flow responses to high intensity precipitation, the modelling of flow and tracer dynamics, as well as flux age estimates, need to account for the rapid flow velocities within the karst aquifer. The response time of stream/conduit flows or groundwaters level to rainfall is very short in karst catchments, e.g. typically a few hours in small catchments like Chenqi (Zhang et al., 2013; Delbart et al., 2014; Labat and Mangin, 2015; Rathay et al., 2017). Coarser resolution data would result in increased uncertainty in the short-term components of travel times (Seeger and Weiler, 2014) and a likely bias towards longer transit times (Heidbüchel et al., 2012). Thus a significant advance of the new model used in this study was that observation and model results captured the flashy (sub-daily) responses of flow and isotope signatures at hourly timescales.

#### 5.4 Equifinality of model parameters and uncertainty of the modelled results

The tracer-aided conceptual model used here provided an opportunity to improve the basis for model evaluation and constrain parameter sets potentially reducing such uncertainty (Beven, 1993). However, high uncertainty always accompanies modelling in such complex landscapes since the tracer-aided model increases model parametersisation for the tracer modules. When we further compared the parameter sensitivities when the model was separately calibrated against the outlet discharge and/or water isotopes, using  $KGE_d$  and  $KGE_i$  respectively, the trade-offs associated with different calibration strategies became evident. The sensitivity analysis (using plots similar to those Figure 5 are shown in the Supplement) showed that increasing the two sensitive parameters in the isotope module (the coefficient for evaporation fractionation  $I_s$  and the weighted isotope composition of rainfall input by the parameter  $fei$ ) results in three parameters in the flow module becoming insensitive when a combined objective function is used. These are the slow reservoir constant ( $Kf$ ), the exchange constant between the two reservoirs  $Ke$  and the ratio of porosity of the quick to slow flow reservoir  $f$ . Consequently, equifinality remain for parameters in the trace-aided model, as the former two sensitive parameters in the isotopic module take functions in the outlet flow (being "old/new") similar to the latter three parameters in the flow module. The former two sensitive parameters in the isotopic module emphasize atmospheric effects on the outlet flow (being "old/new"). A higher  $I_s$  indicates more evaporative effect on the stored water, leading to the stored and released water being older, particularly during the dry period. Increasing  $fei$  indicates newer rainfall

recharge (with more negative isotopic values) into aquifer, resulting in the stored and released water being newer during rainfall period. Alternatively, the latter three parameters in the flow module emphasize effects of fast (newer) and slow (older) flows in aquifer on the outlet flow (being “old/new”). More water release from the slow reservoir (larger  $Kf$ ) and greater release of the slow reservoir into the fast reservoir (larger  $Ke$ ) could lead to the released water being older in the dry season; a high proportion of the fast flow storage (larger  $f$ ) and a greater exchange between the fast reservoir and the slow reservoir (larger  $Ke$ ) could lead to the released water being newer in the wet season. The equifinality for these parameters might only be overcome when we have additional field data to better constrain them (e.g. knowing the evaporative effect on water  $I_s$  and the weighted isotope composition of rainfall input by the parameter  $fei$ ). Despite this, the using tracers in the model provide evidence on the mixing, flux and age relationships that would not be possible from flow-related calibration alone.

The modelling uncertainty of hydrological variables in different units relies on connectivity of the units with the outlet if only the flow and isotopes at the catchment outlet are used as calibration targets. For example, since hydrological connectivity between the outlet and the catchment units decreases in the order: fast flow in the depression < hillslope flow < slow flow in the depression, the uncertainty bands for the three storages increase in the same order. Also the modelling uncertainty increases with ages of water sources contributing into the catchment outlet due to the decrease in variability of the tracer signal in the larger stores. Some of the markedly enriched isotope signals at the outlet during the event peaks are most likely explained by fractionated water being displaced from the paddy fields during event peaks. Hence, the model skill in capturing the effects of evaporative fractionation need to be further investigated in the model; both in terms of process-based parameterisation of fractionation (e.g. Kuppel et al., 2018) and possibly differentiating paddy fields as a separate landscape unit. Though of course this would be a trade-off with increased parameterization and further equifinality.

## 6 Conclusions

We significantly enhanced a catchment-scale flow-tracer model for karst systems developed by Zhang, et al (2017) by conceptualizing two main hydrological response units: hillslope and depression each containing fast and slow flow reservoirs. With this framework, we could calibrate the model using high temporal resolution hydrometric and isotopic data to track hourly water and isotope fluxes through a 1.25 km<sup>2</sup> karst catchment in southwest China. The model captured the flow and tracer dynamics within each landscape unit quite well, and we could estimate the storage, fluxes and age of water within each. This inferred that the fast flow reservoir had the smallest storage, the hillslope unit was intermediate, and the slow flow reservoir had the largest. The estimated mean ages of the hillslope unit, fast and slow flow reservoirs were 137, 326 and 493 days, respectively. Marked seasonal variability in hydroclimate and associated water storage dynamics were the main drivers of non-stationary hydrological connectivity between the hillslope and depression. Meanwhile, the hydrological connectivity between the slow and fast slow reservoirs had variable directionality, which was determined by the hydraulic head within each

medium. Sinkholes can make an important hydrological connectivity between surface water and underground stream flow after heavy rain. New water recharges the underground stream via sinkholes, introducing younger water in the underground stream flow. [Such tracer-aided models enhance our understanding of the hydrological connectivity between different landscape units and the mixing processes between various flow sources.](#) Meanwhile, the tracer-aided model can be used to identify uncertainty sources of the modelled results, e.g., the modelling uncertainty of the hydrological variables in any units in relation to their connectivity with the outlet and ages of the flow components. Whilst the model here needs further development (e.g. the parameterization of isotopic fractionation in the paddy fields) and further assessment and testing requires longer and more detailed (e.g. better characterization of older waters) observation data, it is an encouraging step forward in tracer-aided modelling of karst catchments.

*Data availability.* The isotope data as well as rainfall and flow measurements used for this paper are not publicly accessible due to the constraints of governmental policy in China. The data were obtained through a purchasing agreement for this study. GIS data in this study are available.

*Author contributions.* ZZ, XC, and CS conducted the modelling work and data interpretation, ZZ and QC conducted the field and laboratory work, ZZ prepared the manuscript with contributions from co-authors.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgments.* This research was supported by The UK-China Critical Zone Observatory (CZO) Programme (41571130071), the National Natural Scientific Foundation of China (41571020, 41601013), National 973 Program of China (2015CB452701), the National Key Research and development Program of China (2016YFC0502602), the Fundamental Research Funds for the Central Universities (2016B04814) and the UK Natural Environment Research Council (NE/N007468/1). In addition, we thank Sylvain Kuppel, [the two anonymous reviewers and the editor for their constructive comments which significantly improved the manuscript.](#)

## References

- Ala-Aho, P., Tetzlaff, D., McNamara, J. P., Laudon, H. and Soulsby, C.: Using isotopes to constrain water flux and age estimates in snow-influenced catchments using the STARR (Spatially distributed Tracer-Aided Rainfall-Runoff) model, *Hydrol. Earth Syst. Sci.*, 21(10), 5089–5110, doi:10.5194/hess-21-5089-2017, 2017a.
- Ala-aho, P., Soulsby, C., Wang, H. and Tetzlaff, D.: Integrated surface-subsurface model to investigate the role of groundwater in headwater catchment runoff generation: A minimalist approach to parameterisation, *J. Hydrol.*, 547, 664–677, doi:10.1016/j.jhydrol.2017.02.023, 2017b.

- Bakalowicz, M.: Karst groundwater: A challenge for new resources, *Hydrogeol. J.*, 13(1), 148–160, doi:10.1007/s10040-004-0402-9, 2005.
- Berghuijs, W. R. and Kirchner, J. W.: The relationship between contrasting ages of groundwater and streamflow, *Geophys. Res. Lett.*, 44(17), 8925–8935, doi:10.1002/2017GL074962, 2017.
- 600 Beven, K.: Prophecy, reality and uncertainty in distributed hydrological modelling, *Adv. Water Resour.*, 16(1), 41–51, doi:10.1016/0309-1708(93)90028-E, 1993.
- Beven, K.: A manifesto for the equifinality thesis, in *Journal of Hydrology*, vol. 320, pp. 18–36., 2006.
- Birkel, C., Tetzlaff, D., Dunn, S. M. and Soulsby, C.: Towards a simple dynamic process conceptualization in rainfall-runoff models using multi-criteria calibration and tracers in temperate, upland catchments, *Hydrol. Process.*, 24(3), 260–275, 605 doi:10.1002/hyp.7478, 2010.
- Birkel, C., Tetzlaff, D., Dunn, S. M. and Soulsby, C.: Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models, *Water Resour. Res.*, 47(2), doi:10.1029/2010WR009547, 2011.
- Birkel, C., Soulsby, C., Tetzlaff, D., Dunn, S. and Spezia, L.: High-frequency storm event isotope sampling reveals time-variant transit time distributions and influence of diurnal cycles, *Hydrol. Process.*, 26(2), 308–316, doi:10.1002/hyp.8210, 610 2012.
- Birkel, C. and Soulsby, C.: Advancing tracer-aided rainfall-runoff modelling: A review of progress, problems and unrealised potential, *Hydrol. Process.*, 29(25), 5227–5240, doi:10.1002/hyp.10594, 2015a.
- Birkel, C., Soulsby, C. and Tetzlaff, D.: Conceptual modelling to assess how the interplay of hydrological connectivity, catchment storage and tracer dynamics controls nonstationary water age estimates, *Hydrol. Process.*, 29(13), 2956–2969, 615 doi:10.1002/hyp.10414, 2015b.
- Bonneau, J., Fletcher, T. D., Costelloe, J. F. and Burns, M. J.: Stormwater infiltration and the “urban karst” – A review, *J. Hydrol.*, 552, 141–150, doi:10.1016/j.jhydrol.2017.06.043, 2017.
- Charlier, J. B., Bertrand, C. and Mudry, J.: Conceptual hydrogeological model of flow and transport of dissolved organic carbon in a small Jura karst system, *J. Hydrol.*, 460–461, 52–64, doi:10.1016/j.jhydrol.2012.06.043, 2012. 620
- Chen, X., Zhang, Z.C., Soulsby, C., Cheng, Q.B., Binley, A., Jiang, R. and Tao, M.: Characterizing the heterogeneity of karst critical zone and its hydrological function: an integrated approach, *Hydrol. Process.*, 1-15, DOI: 10.1002/hyp.13232, 2018.
- Coplen, T. B., Neiman, P. J., White, A. B., Landwehr, J. M., Ralph, F. M. and Dettinger, M. D.: Extreme changes in stable hydrogen isotopes and precipitation characteristics in a landfalling Pacific storm, *Geophys. Res. Lett.*, 35(21), 625 doi:10.1029/2008GL035481, 2008.



- Delbart, C., Valdes, D., Barbecot, F., Tognelli, A., Richon, P. and Couchoux, L.: Temporal variability of karst aquifer response time established by the sliding-windows cross-correlation method, *J. Hydrol.*, 511, 580–588, doi:10.1016/j.jhydrol.2014.02.008, 2014.
- 630 Dewaide, L., Bonniver, I., Rochez, G. and Hallet, V.: Solute transport in heterogeneous karst systems: Dimensioning and estimation of the transport parameters via multi-sampling tracer-tests modelling using the OTIS (One-dimensional Transport with Inflow and Storage) program, *J. Hydrol.*, 534, 567–578, doi:10.1016/j.jhydrol.2016.01.049, 2016.
- Field, M. S. and Pinsky, P. F.: A two-region nonequilibrium model for solute transport in solution conduits in karstic aquifers, *J. Contam. Hydrol.*, 44(3–4), 329–351, doi:10.1016/S0169-7722(00)00099-1, 2000.
- 635 Fleury, P., Plagnes, V. and Bakalowicz, M.: 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, doi:10.1016/j.jhydrol.2007.07.014, 2007.
- Ford, D. and Williams, P.: *Karst Hydrogeology and Geomorphology.*, 2013.
- Foster, S., Hirata, R. and Andreo, B.: The aquifer pollution vulnerability concept: aid or impediment in promoting groundwater protection?, *Hydrogeol. J.*, 21(7), 1389–1392, doi:10.1007/s10040-013-1019-7, 2013.
- 640 Freer, J., Beven, K. and Ambroise, B.: 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, doi:10.1029/96WR03723, 1996.
- [Goldscheider, N. and Drew, D.: \*Methods in Karst Hydrogeology: IAHS: International Contributions to Hydrogeology\*, 26., 2007.](#)
- Goldscheider, N., Meiman, J., Pronk, M. and Smart, C.: Tracer tests in karst hydrogeology and speleology, *Int. J. Speleol.*, 37(1), 27–40, doi:10.5038/1827-806X.37.1.3, 2008.
- 645 Harman, C. J.: Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, *Water Resour. Res.*, 51(1), 1–30, 2015.
- Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H. and Weiler, M.: Testing the realism of model structures to identify karst system processes using water quality and quantity signatures, *Water Resour. Res.*, 49(6), 3345–3358, doi:10.1002/wrcr.20229, 2013.
- 650 Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. and Weiler, M.: Karst water resources in a changing world: Review of hydrological modeling approaches, *Rev. Geophys.*, 52(3), 218–242, doi:10.1002/2013RG000443, 2014.
- [Hartmann, A., Antonio Barberá, J. and Andreo, B.: \*On the value of water quality data and informative flow states in karst modelling\*, \*Hydrol. Earth Syst. Sci.\*, 21\(12\), 5971–5985, doi:10.5194/hess-21-5971-2017, 2017.](#)
- Heidbüchel, I., Troch, P. A., Lyon, S. W. and Weiler, M.: The master transit time distribution of variable flow systems, *Water Resour. Res.*, 48(6), doi:10.1029/2011WR011293, 2012.

- Jasechko, S., Kirchner, J. W., Welker, J. M. and McDonnell, J. J.: Substantial proportion of global streamflow less than three months old, *Nat. Geosci.*, 9(2), 126–129, doi:10.1038/ngeo2636, 2016.
- Jasechko, S., Perrone, D., Befus, K. M., Bayani Cardenas, M., Ferguson, G., Gleeson, T., Luijendijk, E., McDonnell, J. J., Taylor, R. G., Wada, Y. and Kirchner, J. W.: Global aquifers dominated by fossil groundwaters but wells vulnerable to modern contamination, *Nat. Geosci.*, 10(6), 425–429, doi:10.1038/ngeo2943, 2017.
- Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Bencala, K. E. and Wondzell, S. M.: Hillslope hydrologic connectivity controls riparian groundwater turnover: Implications of catchment structure for riparian buffering and stream water sources, *Water Resour. Res.*, 46(10), doi:10.1029/2009WR008818, 2010.
- Jukic, D. and Denić-Jukić, V.: Groundwater balance estimation in karst by using a conceptual rainfall – runoff model, *J. Hydrol.*, 373(3–4), 302–315, doi:10.1016/j.jhydrol.2009.04.035, 2009.
- Kirchner, J. W.: A double paradox in catchment hydrology and geochemistry, *Hydrol. Process.*, 17(4), 871–874, doi:10.1002/hyp.5108, 2003.
- Kirchner, J. W., Feng, X., Neal, C. and Robson, A. J.: The fine structure of water-quality dynamics: The (high-frequency) wave of the future, *Hydrol. Process.*, 18(7), 1353–1359, doi:10.1002/hyp.5537, 2004.
- Kling, H., Fuchs, M. and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, *J. Hydrol.*, 424–425, 264–277, doi:10.1016/j.jhydrol.2012.01.011, 2012.
- Kogovsek, J. and Petric, M.: Solute transport processes in a karst vadose zone characterized by long-term tracer tests (the cave system of Postojnska Jama, Slovenia), *J. Hydrol.*, 519(PA), 1205–1213, doi:10.1016/j.jhydrol.2014.08.047, 2014.
- Kübeck, C., Maloszewski, P. J. and Benischke, R.: Determination of the conduit structure in a karst aquifer based on tracer data-Lurbach system, Austria, *Hydrol. Process.*, 27(2), 225–235, doi:10.1002/hyp.9221, 2013.
- Kuppel, S., Tetzlaff, D., Maneta, M. and Soulsby, C.: EcH2O-iso: Water isotopes and age tracking in a process-based, distributed ecohydrological model, *Geoscientific Model Development*, In review, 2018.
- Labat, D. and Mangin, A.: Transfer function approach for artificial tracer test interpretation in karstic systems, *J. Hydrol.*, 529, 866–871, doi:10.1016/j.jhydrol.2015.09.011, 2015.
- Ladouche, B., Marechal, J. C. and Dorfliger, N.: Semi-distributed lumped model of a karst system under active management, *J. Hydrol.*, 509, 215–230, doi:10.1016/j.jhydrol.2013.11.017, 2014.
- Landwehr, J. and Coplen, T.: Line-conditioned excess: a new method for characterizing stable hydrogen and oxygen isotope ratios in hydrologic systems, in *Aquatic Forum 2004: International conference on isotopes in environmental studies*, p. 132–134. [online] Available from: <http://dx.doi.org/10.8/S>, 2004.
- Landwehr, J. M., Coplen, T. B. and Stewart, D. W.: Spatial, seasonal, and source variability in the stable oxygen and hydrogen isotopic composition of tap waters throughout the USA, *Hydrol. Process.*, 28(21), 5382–5422, doi:10.1002/hyp.10004, 2014.

- Legout, A., Legout, C., Nys, C. and Dambrine, E.: Preferential flow and slow convective chloride transport through the soil of a forested landscape (Fougères, France), *Geoderma*, 151(3–4), 179–190, doi:10.1016/j.geoderma.2009.04.002, 2009.
- Lexartza-Artza, I. and Wainwright, J.: Hydrological connectivity: Linking concepts with practical implications, *Catena*, 79(2), 146–152, doi:10.1016/j.catena.2009.07.001, 2009.
- 690
- McCutcheon, R. J., McNamara, J. P., Kohn, M. J. and Evans, S. L.: An evaluation of the ecohydrological separation hypothesis in a semiarid catchment, *Hydrol. Process.*, 31(4), 783–799, doi:10.1002/hyp.11052, 2017.
- McDonnell, J. J. and Beven, K.: Debates - The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph, *Water Resour. Res.*, 50(6), 5342–5350, doi:10.1002/2013WR015141, 2014.
- 695
- McMillan, H., Tetzlaff, D., Clark, M. and Soulsby, C.: Do time-variable tracers aid the evaluation of hydrological model structure? A multimodel approach, *Water Resour. Res.*, 48(5), doi:10.1029/2011WR011688, 2012.
- Morales, T., Uriarte, J. A., Olazar, M., Antigüedad, I. and Angulo, B.: Solute transport modelling in karst conduits with slow zones during different hydrologic conditions, *J. Hydrol.*, 390(3–4), 182–189, doi:10.1016/j.jhydrol.2010.06.041, 2010.
- 700
- Mudarra, M., Andreo, B., Marín, A. I., Vadillo, I. and Barberá, J. A.: Combined use of natural and artificial tracers to determine the hydrogeological functioning of a karst aquifer: the Villanueva del Rosario system (Andalusia, southern Spain), *Hydrogeol. J.*, 22(5), 1027–1039, doi:10.1007/s10040-014-1117-1, 2014.
- Mueller, M. H., Weingartner, R. and Alewell, C.: Importance of vegetation, topography and flow paths for water transit times of base flow in alpine headwater catchments, *Hydrol. Earth Syst. Sci.*, 17(4), 1661–1679, doi:10.5194/hess-17-1661-2013, 2013.
- 705
- Peng, T. and Wang, S.: Effects of land use, land cover and rainfall regimes on the surface runoff and soil loss on karst slopes in southwest China, *CATENA*, 90, 53–62, doi:10.1016/j.catena.2011.11.001, 2012.
- Perrin, C., Michel, C. and Andréassian, V.: 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, doi:10.1016/S0022-1694(00)00393-0, 2001.
- 710
- Rathay, S. Y., Allen, D. M. and Kirste, D.: Response of a fractured bedrock aquifer to recharge from heavy rainfall events, *J. Hydrol.*, doi:10.1016/j.jhydrol.2017.07.042, 2017.
- [Reaney, S. M., Bracken, L. J. and Kirkby, M. J.: The importance of surface controls on overland flow connectivity in semi-arid environments: Results from a numerical experimental approach, \*Hydrol. Process.\*, 28\(4\), doi:10.1002/hyp.9769, 2014.](#)
- 715
- Rimmer, A. and Salinger, Y.: Modelling precipitation-streamflow processes in karst basin: The case of the Jordan River sources, Israel, *J. Hydrol.*, 331(3–4), 524–542, doi:10.1016/j.jhydrol.2006.06.003, 2006.

- Rimmer, A. and Hartmann, A.: Simplified Conceptual Structures and Analytical Solutions for Groundwater Discharge Using Reservoir Equations, *Water Resour. Manag. Model.*, 2, 217–238, doi:10.5772/34803, 2012.
- Schaefli, B. and Gupta, H. V.: Do Nash values have value?, *Hydrol. Process.*, 21(15), 2075–2080, doi:10.1002/hyp.6825, 2007.
- 720 Seeger, S. and Weiler, M.: Lumped convolution integral models revisited: on the meaningfulness of inter catchment comparisons, *Hydrol. Earth Syst. Sci. Discuss.*, 11(6), 6753–6803, doi:10.5194/hessd-11-6753-2014, 2014.
- Soulsby, C., Birkel, C. and Tetzlaff, D.: Assessing urbanization impacts on catchment transit times, *Geophys. Res. Lett.*, 41(2), 442–448, doi:10.1002/2013GL058716, 2014.
- Soulsby, C., Birkel, C., Geris, J., Dick, J., Tunaley, C. and Tetzlaff, D.: Stream water age distributions controlled by storage
- 725 dynamics and nonlinear hydrologic connectivity: Modeling with high-resolution isotope data, *Water Resour. Res.*, 51(9), 7759–7776, doi:10.1002/2015WR017888, 2015.
- Sprenger, M., Leistert, H., Gimbel, K. and Weiler, M.: Illuminating hydrological processes at the soil-vegetation-atmosphere interface with water stable isotopes, *Rev. Geophys.*, 54(3), 674–704, doi:10.1002/2015RG000515, 2016.
- Stets, E. G., Winter, T. C., Rosenberry, D. O. and Striegl, R. G.: Quantification of surface water and groundwater flows to
- 730 open - and closed-basin lakes in a headwaters watershed using a descriptive oxygen stable isotope model, *Water Resour. Res.*, 46(3), doi:10.1029/2009WR007793, 2010.
- Tetzlaff, D., Birkel, C., Dick, J., Geris, J. and Soulsby, C.: Storage dynamics in hydrogeological units control hillslope connectivity, runoff generation, and the evolution of catchment transit time distributions, *Water Resour. Res.*, 50(2), 969–985, doi:10.1002/2013WR014147, 2014.
- 735 Tritz, S., Guinot, V. and Jourde, H.: Modelling the behaviour of a karst system catchment using non-linear hysteretic conceptual model, *J. Hydrol.*, 397(3–4), 250–262, doi:10.1016/j.jhydrol.2010.12.001, 2011.
- van Schaik, N. L. M. B., Schnabel, S. and Jetten, V. G.: The influence of preferential flow on hillslope hydrology in a semi-arid watershed (in the Spanish Dehesas), *Hydrol. Process.*, 22(18), 3844–3855, doi:10.1002/hyp.6998, 2008.
- White, W. B.: A brief history of karst hydrogeology: contributions of the NSS, *J. Cave Karst Stud.*, 69(1), 13–26, doi:PNR61,
- 740 2007.
- Worthington, S. R. H.: Diagnostic hydrogeologic characteristics of a karst aquifer (Kentucky, USA), *Hydrogeol. J.*, 17(7), 1665–1678, doi:10.1007/s10040-009-0489-0, 2009.
- Worthington, S. R. H., Jeannin, P.-Y., Alexander, E. C., Davies, G. J. and Schindel, G. M.: Contrasting definitions for the term “karst aquifer,” *Hydrogeol. J.*, 25(5), 1237–1240, doi:10.1007/s10040-017-1628-7, 2017.
- 745 Xie, Y., Cook, P. G., Simmons, C. T., Partington, D., Crosbie, R. and Batelaan, O.: Uncertainty of groundwater recharge estimated from a water and energy balance model, *J. Hydrol.*, doi:10.1016/j.jhydrol.2017.08.010, 2017.

Zhang, Z., Chen, X., Ghadouani, A. and Shi, P.: Modelling hydrological processes influenced by soil, rock and vegetation in a small karst basin of southwest China, *Hydrol. Process.*, 25(15), 2456–2470, doi:10.1002/hyp.8022, 2011.

750 Zhang, Z., Chen, X., Chen, X. and Shi, P.: Quantifying time lag of epikarst-spring hydrograph response to rainfall using correlation and spectral analyses, *Hydrogeol. J.*, 21(7), 1619–1631, doi:10.1007/s10040-013-1041-9, 2013.

Zhang, Z., Chen, X. and Soulsby C.: Catchment-scale conceptual modelling of water and solute transport in the dual flow system of the karst critical zone, *Hydrol. Process.*, 31 (19): 3421–3436, doi: 10.1002/hyp.11268, 2017.

## Appendix A

755 Table A1 Water/isotope/age flux equations of the model.

Balance equations	Calculation and explanation for the equation items	
<b>Flow routing</b>		
$\frac{dV_h}{dt} = P_h - ET_h - Q_{h-s} - Q_{h-f}$ (water balance in hillslope)	$P_h = P * Area_h / Area$	$a$ is coefficient of rainfall recharge into slow flow reservoir; $Area_h$ and $Area$ represent the hillslope and catchment area.
	$P_s = (P - P_h) * a$	
	$P_f = (P - P_h) * (1 - a)$	
$\frac{dV_s}{dt} = P_s - ET_s + Q_{h-s} - Q_e$ (water balance in slow flow reservoir)	$ET_h = ET * Area_h / Area$	$ET = \eta * ETP$ , where $ETP$ is potential evapotranspiration estimated by Penman formula, and $\eta$ is a conversion factor estimated by other study in this region.
	$ET_s = (ET - ET_h) * a$	
	$ET_f = (ET - ET_h) * (1 - a)$	
$\frac{dV_f}{dt} = P_f - ET_f + Q_{h-f} + Q_e - Q_f$ (water balance in fast flow reservoir)	$Q_{h,t} = w * exp(V_{h,t}/5000)$	$w$ is flow routing constant for the hillslope unit; $b$ is coefficient of hillslope lateral flow $Q_{h,t}$ to the slow reservoir $Q_{h-s}$ ; $K_s$ and $K_f$ are constant for slow and fast reservoirs, respectively
	$Q_{h-f,t} = b * Q_{h,t}$	
	$Q_{h-s,t} = (1 - b) * Q_{h,t}$	
	$Q_{s,t} = K_s V_{s,t}$	
	$Q_{f,t} = K_f V_{f,t}$	
$P$ is rainfall amount ( $m^3 \text{ hour}^{-1}$ ), $ET$ is evapotranspiration ( $m^3 \text{ hour}^{-1}$ ), $Q$ is flow discharge ( $m^3 \text{ hour}^{-1}$ ) and $V$ is storage ( $m^3$ ); the subscripts of $h$ , $s$ and $f$ represent the hillslope, slow and fast flow reservoirs, respectively; the subscripts of $h-s$ and $h-f$ represent hillslope to slow and to fast flow reservoirs, respectively, and the subscript of $e$ represents flow exchange between fast and slow reservoirs.	Consecutive routings for the time series: $Q_{s,t} = \Phi_{s,1}(P_{s,t} - ET_{s,t} + Q_{h-s,t}) + \Phi_{s,2}\Phi_{s,1}(P_{s,t-1} - ET_{s,t-1} + Q_{h-s,t-1}) + \Phi_{s,2}^2\Phi_{s,1}(P_{s,t-2} - ET_{s,t-2} + Q_{h-s,t-2}) + \Phi_{s,2}^2Q_{s,t-2} + \Phi_{s,3}Q_{f,t} + \Phi_{s,2}\Phi_{s,3}Q_{f,t-1} + \Phi_{s,2}^2\Phi_{s,3}Q_{f,t-2}$ $Q_{f,t} = \Phi_{f,1}(P_{f,t} - ET_{f,t} + Q_{h-f,t}) + \Phi_{f,2}Q_{f,t-1} + \Phi_{f,3}\{\Phi_{s,1}(P_{s,t} - ET_{s,t} + Q_{h-s,t}) + \Phi_{s,2}\Phi_{s,1}(P_{s,t-1} - ET_{s,t-1} + Q_{h-s,t-1}) + \Phi_{s,2}^2\Phi_{s,1}(P_{s,t-2} - ET_{s,t-2} + Q_{h-s,t-2}) + \Phi_{s,2}^2Q_{s,t-2} + \Phi_{s,3}Q_{f,t} + \Phi_{s,2}\Phi_{s,3}Q_{f,t-1} + \Phi_{s,2}^2\Phi_{s,3}Q_{f,t-2}\}$ where $\Phi_{s,1} = 1/(K_s + fK_s/K_e)$ $\Phi_{s,2} = K_s/(K_s + fK_s/K_e)$ $\Phi_{s,3} = K_f/(K_sK_e + fK_s)$ $\Phi_{f,1} = 1/(K_f + K_f/K_e + 1)$ $\Phi_{f,2} = K_f/(K_f + K_f/K_e + 1)$ $\Phi_{s,3} = fK_s/\{K_e(K_f + K_f/K_e + 1)\}$ The derivations for dual flow model in details refer to Zhang et al., 2017.	

## Isotope routing

$$\frac{di_h(V_h)}{dt} = i_p V_{p,h} + i_{pas} V_{p,pas} - i_h ET_h -$$

$$i_h Q_{h-s} - i_h Q_{h-f} + i_{pas} V_{pas,in} - i_h V_{pas,in}$$

(isotope balance in active store in hillslope)

$$\frac{di_{pas}(V_{pas})}{dt} = i_p V_{p,pas} - i_{pas} V_{p,pas} +$$

$$i_h V_{pas,in} - i_{pas} V_{pas,in}$$

(isotope balance in passive store in hillslope)

$$\frac{di_s(V_s)}{dt} = i_p P_s - i_s ET_s + i_h Q_{h-s} - i_s Q_e$$

(isotope balance in slow flow reservoir)

$$\frac{di_f(V_f)}{dt} = i_p P_f - i_f ET_f + i_h Q_{h-f} + i_s Q_e -$$

$$i_f Q_f$$

(isotope balance in fast flow reservoir)

$V_{pas}$  (m<sup>3</sup>) is storage of the passive reservoir in hillslope used to determine isotope storage, mixing, and transport in a way that does not affect the dynamics of water flux volumes.

$V_{pas,in}$  (m<sup>3</sup>) is water volume from the active store to the passive store.  $V_{p,h}$  and  $V_{p,pas}$  (m<sup>3</sup>) are the volume of rainfall recharge into the active and passive stores, respectively.

Consecutive routings for the time series:

$$\delta D_{h,t} = \{(\delta D_{h,t-1} * V_{h,t-1})(1 - con) + \delta D_{pas,t-1} * V_{h,t-1} * con + \delta D_{p,t} * P_{h,t} * Par_t + \delta D_{pas,t} P_{h,t} (1 - Par_t) - \delta D_{h,t-1} * ET_{h,t} - \delta D_{h,t-1} (Q_{h-f,t} + Q_{h-s,t})\} / V_{h,t}$$

$$\delta D_{pas,t} = \{\delta D_{pas,t-1} * V_{pas,t-1} + \delta D_{p,t} * P_{h,t} (1 - Par_t) + \delta D_{h,t-1} * V_{h,t-1} * con - \delta D_{pas,t-1} V_{h,t-1} * con - \delta D_{pas,t} P_{h,t} (1 - Par_t) - (1 + Is) * \delta D_{pas,t-1} * ET_{h,t}\} / V_{pas,t}$$

$con$  is coefficient of mass exchange between active and passive stores,  $Is$  is coefficient of evaporation fractionation.  $Par_t = pp * exp(V_{h,t}/kk)$ ,  $pp$  and  $kk$  are constants for calculation of rainfall recharge into the active store in hillslope.

$$\delta D_{s,t} = \{\Phi_{s,1}(\delta D_{p-d,t}(P_{s,t} - ET_{s,t}) + \delta D_{h,t} * Q_{h-s,t}) + \Phi_{s,2} * \Phi_{s,1}(\delta D_{p-d,t-1}(P_{s,t-1} - ET_{s,t-1}) + \delta D_{h,t-1} * Q_{h-s,t-1}) + \Phi_{s,2}^2 * \Phi_{s,1}(\delta D_{p-d,t-2}(P_{s,t-2} - ET_{s,t-2}) + \delta D_{h,t-2} * Q_{h-s,t-2}) + \Phi_{s,2}^2 * \delta D_{s,t-2} * Q_{s,t-2} + \Phi_{s,3} * \delta D_{f,t} * Q_{f,t} + \Phi_{s,2} * \Phi_{s,3} * \delta D_{f,t-1} * Q_{f,t-1} + \Phi_{s,2}^2 * \Phi_{s,3} * \delta D_{f,t-2} * Q_{f,t-2}\} / Q_{s,t}$$

$$\delta D_{f,t} = \{\Phi_{f,1}(\delta D_{p-d,t}(P_{f,t} - ET_{f,t}) + \delta D_{h,t} * Q_{h-f,t}) + \Phi_{f,2} * \delta D_{f,t-1} * Q_{f,t-1} + \Phi_{f,3}[\Phi_{s,1}(\delta D_{p-d,t}(P_{s,t} - ET_{s,t}) + \delta D_{h,t} * Q_{h-s,t}) + \Phi_{s,2} * \Phi_{s,1}(\delta D_{p-d,t-1}(P_{s,t-1} - ET_{s,t-1}) + \delta D_{h,t-1} * Q_{h-s,t-1}) + \Phi_{s,2}^2 * \Phi_{s,1}(\delta D_{p-d,t-2}(P_{s,t-2} - ET_{s,t-2}) + \delta D_{h,t-2} * Q_{h-s,t-2}) + \Phi_{s,3} * \delta D_{f,t} * Q_{f,t} + \Phi_{s,2} * \Phi_{s,3} * \delta D_{f,t-1} * Q_{f,t-1} + \Phi_{s,2}^2 * \Phi_{s,3} * \delta D_{f,t-2} * Q_{f,t-2}] - Is * \Phi_{f,1} * \delta D_{f,t-1} * ET_{f,t}\} / (Q_{f,t} - \Phi_{f,3} \Phi_{s,3} Q_{s,t-2})$$

$$\delta D_{p-d} = fei * \delta D_p, \quad fei \text{ is weighting constant.}$$



## Flux age

$$\frac{dAge_h(V_h)}{dt} = Age_p V_{P_h} + Age_{pas} V_{P_{pas}} -$$

$$Age_h ET_h - Age_h Q_{h-s} - Age_h Q_{h-f} +$$

$$Age_{pas} V_{pas\_in} - Age_h V_{pas\_in}$$

(age balance for active store in hillslope)

$$\frac{dAge_{pas}(V_{pas})}{dt} = Age_p V_{P_{pas}} -$$

$$Age_{pas} V_{P_{pas}} + Age_h V_{pas\_in} -$$

$$Age_{pas} V_{pas\_in}$$

(age balance for passive store in hillslope)

$$\frac{dAge_s(V_s)}{dt} = Age_p P_s - Age_s ET_s + Age_h Q_{h-s} -$$

$$Age_s Q_e$$

(age balance in slow flow reservoir)

$$\frac{dAge_f(V_f)}{dt} = Age_p P_f - Age_f ET_f +$$

$$Age_h Q_{h-f} + Age_s Q_e - Age_f Q_f$$

(age balance in fast flow reservoir)

Consecutive routings for the time series

$$Age_{h,t} = \{(Age_{h,t-1} + 1) * V_{h,t-1}(1 - con) + (Age_{pas,t-1} + 1) * V_{h,t-1} * con + Age_{p,t} * P_{h,t} * Par_t + (Age_{pas,t-1} + 1)P_{h,t}(1 - Par_t) - (Age_{h,t-1} + 1) * ET_{h,t} - (Age_{h,t-1} + 1)(Q_{h-f,t} + Q_{h-s,t})\}/V_{h,t}$$

$$Age_{pas,t} = \{(Age_{pas,t-1} + 1) * V_{pas,t-1} + Age_{p,t} * P_{h,t}(1 - Par_t) + (Age_{h,t-1} + 1) * V_{h,t-1} * con - (Age_{pas,t-1} + 1)V_{h,t-1} * con - (Age_{pas,t-1} + 1)P_{h,t}(1 - Par_t) - (Age_{pas,t-1} + 1) * ET_{h,t}\}/V_{pas,t}$$

$$Age_{s,t} = \{\Phi_{s,1}(Age_{p,t}(P_{s,t} - ET_{s,t}) + Age_{h,t} * Q_{h-s,t}) + \Phi_{s,2} * \Phi_{s,1}((Age_{p,t-1} + 1)(P_{s,t-1} - ET_{s,t-1}) + (Age_{h,t-1} + 1) * Q_{h-s,t-1}) + \Phi_{s,2}^2 * \Phi_{s,1}((Age_{p,t-2} + 2)(P_{s,t-2} - ET_{s,t-2}) + (Age_{h,t-2} + 2) * Q_{h-s,t-2}) + \Phi_{s,2}^2(Age_{s,t-2} + 2)Q_{s,t-2} + \Phi_{s,3} * Age_{f,t} * Q_{f,t} + \Phi_{s,2} * \Phi_{s,3} * (Age_{f,t-1} + 1) * Q_{f,t-1} + \Phi_{s,2}^2 * \Phi_{s,3} * (Age_{f,t-2} + 2) * Q_{f,t-2}\}/Q_{s,t}$$

$$Age_{f,t} = \{\Phi_{f,1}(Age_{p,t}(P_{f,t} - ET_{f,t}) + Age_{h,t} * Q_{h-f,t}) + \Phi_{f,2}(Age_{f,t-1} + 1)Q_{f,t-1} + \Phi_{f,3}[\Phi_{s,1}(Age_{p,t}(P_{s,t} - ET_{s,t}) + Age_{h,t} * Q_{h-s,t}) + \Phi_{s,2} * \Phi_{s,1}((Age_{p,t-1} + 1)(P_{s,t-1} - ET_{s,t-1}) + (Age_{h,t-1} + 1) * Q_{h-s,t-1}) + \Phi_{s,2}^2 * \Phi_{s,1}((Age_{p,t-2} + 2)(P_{s,t-2} - ET_{s,t-2}) + (Age_{h,t-2} + 2) * Q_{h-s,t-2}) + \Phi_{s,3} * Age_{f,t} * Q_{f,t} + \Phi_{s,2} * \Phi_{s,3} * (Age_{f,t-1} + 1) * Q_{f,t-1} + \Phi_{s,2}^2 * \Phi_{s,3} * (Age_{f,t-2} + 2) * Q_{f,t-2}]\}/(Q_{f,t} - \Phi_{f,3} \Phi_{s,3} Q_{s,t-2})$$

The age of rainfall,  $Age_{p,t}$ , equals to 0.

**Table A2 Description of the calibrated parameters**

Coefficient	Units	Descriptions
$K_s$	hour	The slow flow reservoir constant
$K_f$	hour	The fast flow reservoir constant
$K_e$	hour	Exchange constant between the two reservoirs
$f$	-	The ratio of porosity of the quick to slow flow reservoir
$a$	-	Precipitation recharge coefficient for slow flow reservoir
$w$	-	The hillslope unit constant
$b$	-	Recharge coefficient of Hillslope to slow flow reservoir
$Is$	-	Coefficient for evaporation fractionation
$KK$	-	Constant for calculation of rainfall recharging the active store in hillslope
$pp$	-	
$con$	-	Coefficient for exchange flow between active and passive stores in hillslope
$fei$	-	Weighting constant