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 will help us improve our manuscript. We will thoroughly revise 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 will revise 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 it the current format.
Reply: We appreciate the referee's suggestions, and the whole paper will be thoroughly revised to improve the clarity and grammar.

Detailed comments

Line 9: "unique" does not seems 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 will re-state this. Whilst we recognize that others have such higher temporal resolution steam 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 will be 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 will revised this.

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

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

Line 14: specify that you talk about mean ages OF WATER.
Reply: We will revise 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” will be replaced by “highly”

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 will revise the sentences from lines 14 to 19. We will clarify that the connectivity in our study refers to fluxes from “hillslope- to- depression (fast and slow reservoirs)- to- outlet stream”.

Line 16-19: please revisit these sentences to make this an understandable abstract.
Reply: We will revise 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 will revise 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 will revise this paragraph. In the revised manuscript, we will first describe the lumped models and then the semi-distributed models.

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: Yes, we will revise this and define terms more precisely. In karst aquifers, the solutional conduits connect with intergranular pores and fractures (often termed as matrix porosity), showing dual or even triple porosity zones (Worthington, Jeannin, Alexander, Davies, & Schindel, 2017). Thus, karst aquifers are often conceptualized as dual porosity systems as residence times in the matrix are several orders of magnitude longer than those in the conduits (Goldscheider & Drew, 2007).

Line 37: “can be intuitively identified” what do you mean here?
Reply: We will delete the sentence in line 37 and replace 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 will revise 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 will revise 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 will revise this and replace the literature by the more precise descriptions from Birkel et al, 2015.

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, doi:10.1002/hyp.10414, 2015b.

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 will delete 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 will add the relevant publications.

Figure 1: please make it much more explicit in the caption what you display here.
Reply: We will add the relevant descriptions.

Table 1: the range is a redundant variable.
Reply: We will revise this.

Table 1: consider indicating how much of the time there is zero flow.
Reply: We will add this item. It occupies only a short period of time in our observation period.

Table 1 Statistical summary of flow discharge for hillslope spring (HS) and catchment outlet (m3/s)

Obs	Min	Max	Mean	Cv	Time with zero flow (h)
Outlet	0	0.15	4.7×10^{-3}	2.83	328
HS	0	1.4×10^{-3}	8.5×10^{-5}	1.73	713

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

Line 159: CalculationS
Reply: We will revise this.

Line 162: for each of the (not in each of the)
Reply: We will revise this.

Line 162-163: inconsistent with singular and plural. Check grammar.
Reply: We will revise this sentence and make corrections in the whole manuscript.

Line 168: fix superscript “rainfall (m3 hour-1)” 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 will revise this, and give 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 will be clarified in the appendix in an attached file 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 were 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 10ˆ5 tested different parameter combinations. I do not understand how you arrive at the second 10ˆ5.
Reply: We will revise the descriptions and add the initial parameters from the first calibration. A total of 10ˆ5 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 10ˆ5 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 10ˆ5, but the range of the initial parameters are different (Initial range 1 and 2 for the 1st and 2nd calibration in revised Table 3).
After the first calibration when KGE >0.3, the range of each parameter was reduced. Then, the narrowed ranges (initial range 2 in Table 3) were used as the initial ranges for the second calibration.

Table 3

For Flow	K_s (hour^{-1})	K_f (hour^{-1})	K_e (hour^{-1})	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	*Is*	*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_d	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		

Line 276: “rogue” ? what do you mean

Reply: It refers to some samples that are unusually high during the study period (in Fig.3). 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 will revise 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 will revise the conclusions. This description will be changed to: for the “hillslope-depression-outlet” connectivity, the weak/strong connectivity between hillslope and depression was inferred by only a small/large percentage (16.8% and 57.5%) of the outlet fluxes from the hillslope unit contribution during the dry/wet period (Fig 8). The seasonal decline and increase of the fast flow storage were primarily maintained by the slow flow while the short-term variability of the fast flow storage responds to the hillslope flow (Fig 7) as well as rainfall recharge (Fig 8).