

Comment on hess-2021-249

Anonymous Referee #1

Referee comment on "Karst spring recession and classification: efficient, automated methods for both fast- and slow-flow components" by Tunde Olarinoye et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-249-RC1>, 2021

General comments

The authors propose an automatic approach to infer and classify the dynamic characteristics of karst systems based on the analysis of the recession hydrograph. The approach introduces variability through multiple recession extraction methods, fitting individual recession segments and optimization approaches with various degrees of freedom for a fast and slow flow recession model. Recognition of the duality of porosity and processes in karst systems leads to a vision of a generally accepted two-reservoir conceptual model. The desire to develop a method of recession analysis that is consistent with this view is well justified. The automated and multi-angle traits of the approach are indeed essential to cope with the known biases of single and visually supervised approaches to recession analysis. In my opinion, this work will improve the robustness of the comparison of the dynamic characteristics of karst hydrological systems. Furthermore, I believe that tools for comparing hydrological systems are an area of research that will continue to develop and contribute to a better understanding of hydrological systems. For these reasons, I consider that the article is following the scope of HESS.

In general, the paper is intelligible. The results are interpretable and make sense. The introduction correctly sets the context, and the discussion provides elements that facilitate the interpretation of the results. Nevertheless, I find that these two sections do not motivate well enough the choices that have been adopted and do not give a good account of the strengths and weaknesses of the approach. As a result, the modeling philosophy and epistemological approach are a bit blurry and sometimes inconsistent (see specific issues). Clearly stating these choices would allow the discussion section to be more structured, sharpened, and provide better perspectives and research avenues. Currently, I found the latter to be relatively weak or absent. Still, I enjoyed the approach and the content, particularly the idea of varying the degrees of freedom in the model optimization. I feel that this is a reliable piece of work that I could reuse myself and will certainly be useful to other researchers. Therefore, I consider the scientific quality to be good. Since it is a paper that proposes a method, I also regret the lack of additional material that would facilitate re-implementation (input/output and/or code recipes). I would also have appreciated having access to a bit more details about the results, not especially in the

paper itself but in the appendix or the supplementary materials, about the seasonality of parameters and the statistics of recession segments.

The figures are clear and good. The introduction and discussion sections are ok but could probably be improved and sharpened. But in my opinion, this is more a question of substance than form (see above and specific issues). However, I regret the long and complex sentence constructions, plus a large number of grammatical errors. I think they are below acceptable standards and could be easily removed with a good grammar/spellchecker.

Specific issues

As stated in my general comment, I have very few concerns about the approach itself (except perhaps SI-2), and I have a favorable opinion about the paper. Still, I am concerned about how the methodological choices are introduced and motivated and how the results are interpreted. It lacks coherence and proper motivation, which in some respects affects the clarity of the reading, might suggest that the method is inappropriate, and prevents the development of clear research perspectives. I present below the nature of my concerns and my suggestions on how they can be addressed in a motivated and coherent rationale as rooms for improvement through possible minor revisions of the introduction and discussion sections.

- SI-1: Motivation for individual recession event: seasonality.

There is a first inconsistency in the motivation for fitting individual recession events. It is said that fitting individual recession events would allow capturing the variability for better inference of the structural and dynamic traits of a system (L48-51) and would be more consistent because the parameters obtained from the entire set of recession points could not be actually representative of individual cases (L216-217). Yet, in the other section, the authors tend to interpret the mean of all parameters and indicators as representative of the karst system. In the meantime, the spread of the parameters is deemed both necessary and valuable as it reflects dynamic properties but, paradoxically, also considered as unwanted uncertainty that we shall reduce for having a set of "representative" parameters (L25-27, L450-460, L525-526). I discuss the author's suggestion for reducing uncertainties in SI-5. For now, I continue with the value of the individual segments approach.

Historically, recession analysis methods applied on all points pursue the same goal of inferring the dynamic and structural properties of the system. It is also possible to have uncertainty bounds, analytical or bootstrapped ones, for estimated parameters. So, why go for individual recession fitting if you can already depict the average and the variability of the estimated parameters? The answer is that those who have adopted this approach have learned something about the dynamics of the system. The parameters were

projected on the temporal/seasonal axis, and interesting patterns related to the seasonality or the antecedent conditions of hydrological variables appeared [See 1-5]. In doing so, the individual segment approaches proved themselves useful and showed that the low dimensional models that we commonly use (Mangin is no exception) are actually underfitting recession dynamics [6].

Therefore, I would encourage the authors to report the seasonal variation of the estimated parameters because it is a valuable insight gained through the individual segments approach and, moreover, a legitimate motivation in favor of the approach. They should capitalize on it and promote this feature as an additional strength of the approach. Furthermore, no paper, to my knowledge, reports the seasonal variability of η and ϵ for karst systems, same for i and K . For now, the importance of seasonality is barely mentioned despite that it is recognized as a source of variability (L435-437), and its contribution is not pictured or quantified. I understand that the paper is focused on proposing a method, but a graph about the seasonal dynamics of parameters or the classification is not expansive and could be placed in the supplementary materials.

- SI-2: Motivation for the Mangin model and framework.

The hypotheses behind the Mangin model are not discussed enough, and the classification framework is not sufficiently described in section 2.2. The hypothesis of the duality of porosity and processes in karst systems is well described. However, the hypothesis of the matrix flow following the linear recession equation of Maillet (L170, Eq. 3) is not. Nonlinear equations are common to describe baseflow, including in karst systems [7]. Some comments on why the Maillet approximation can be used would have been appreciated, if not theoretical, with an empirical justification of its application showing that Mangin's framework is widely used and will benefit from more robustness.

Besides, the authors paradoxically rely on the Askoy & Wittenberg (A&W) REM that explicitly states that baseflow is nonlinear (L141, Eq. 1). At first sight, it appears as a rationale problem. However, if the importance of the classification Mangin framework is more emphasized in the objective of the paper and better described in section 2.2 (or best, in its own section), using the linear model that is part of this classification framework of the authors' choice make sense. Also, even if the Mangin baseflow model is linear, extracting recession with the nonlinear method of A&W could be justifiable, because Mangin's model with both components and sufficient degrees of freedom, is in fact, a nonlinear recession model. Still, it would have been more consistent, in my opinion, to use A&W approach with the linear Maillet Eq.3, eventually, by choosing a higher CV to compensate the lack of degrees of freedom and the fear that the outcome would be too restrictive on the number of recession point.

- SI-3: Motivation/Discussion of the REMs

In the discussion (L411-412), the authors mention that: *"Overall, the adapted REMs and the introduced three POAs provide range of results that adequately represents the karst systems. This suggests that the modified REMs are well suited for application to karst spring recession analysis"*. I found the conclusion a little too optimistic and leaving little room for improvement. Plus, I feel that it is contradictory with the suggestion of focusing on different segment lengths to reduce uncertainty (L455-56). Why not say that REMs are unsuited then?

Three REMs are selected on the honorable basis of what previous authors have suggested (L160, Table 1). Despite my concerns about the A&W REMs (SI-2), it is quite common to refer to and implement these former REMs in recent papers [6, and 8-9 as referred by the authors]. Thus, it is justifiable. However, this is not a strong rationale. None of the authors of REMs had the primary objective of providing a method for recession extraction. They had to, but their focus was on recession analysis, not extraction. Also, they were not aware of the recent discovery inference problems associated with the extraction method (e.g., [8]). In fact, they look at recession from the narrow and specific scope of the criteria set in Table 1. Using them broaden the scope of the analysis, but still, these are three specific angles of approach to recession extraction, and it is delicate to affirm that they adequately represent the karst systems.

My point is that a noticeable perspective for improvement would be to rely on one or more generic method that allows varying the criteria of selections (beginning/end of recession and tolerance to anomalies). It would allow inferring the statistics for ranges of criteria instead of a few predetermined ones. Coupling a generic method to the POAs could certainly improve the framework in the future.

- SI-4: Degrees of freedom and data points

There is a concern about the parsimony of the approach and the limited number of points in the recession segments [6]. Instead of three models, one for each study site, more than two thousand models were fitted, in which the degrees of freedom is very close to the number of data points. So far, I believe that the results and their interpretation make some sense, and so, I am willing to believe that what is captured is meaningful.

I could have recommended relying on a nonlinear two parameters equation to add another degree of freedom to the late recession (the nonlinear exponent, SI-2), but should we? The natural question is how far we could go in terms of degrees of freedom and when equifinality will produce unreliable meaningless parameters? I understand that this is a separate question that would deserve another paper, but it is worth mentioning it.

Also, I miss some more statistics in Table 4 (L292), such as the average number of recession points per event, which would help appreciate this issue. Also, perhaps statistics about the flow Q and dQ/dt captured by the REM methods would also be interesting.

- SI-5: Uncertainties and the subsequent suggestions

I have experienced contradictory opinions about the uncertainties. On the one hand, they are considered essential because they reflect the actual dynamics of karst. On the other hand, the authors wish to reduce them and offer suggestions for doing so. Also, attention should be paid to using "reliable parameters" when we know that recession models are underfitting recession dynamics (SI-1). The term "robust parameters" seems more appropriate.

The authors refer to suggestions in the abstract and conclusion, but nothing is said explicitly about them in these sections. The suggestions are two, not very well evidenced and hidden in lines L455-460 : (i) focusing on segments of different recession lengths, (ii) using the master recession curve (MRC). I found none of them to be very relevant nor developed. With one (i), it is basically said that REMs are inappropriate (SI-3), and, by focusing on one type of segment length, some processes are dismissed, and you reduce uncertainty by occulting the natural variability that you aim at capturing. Perhaps, presenting the results per season would be more appropriate (SI-1) to reduce the uncertainty of sensitivity to initial conditions.

Regarding the use of MRC (ii), I am not sure I understand what the authors mean. I understand that since the MRC is an average behavior, fitting this average will produce a single estimate of the parameters, which will naturally reduce the uncertainty. Well, this is not reducing, it is still occulting, which is why I believe the authors ultimately reject their own suggestion in the last sentence (L459-60). Note that the MRC approaches with uncertainty quantification do exist (e.g. [10]), so what is said in L99-101 and L459-60 is not correct. It is entirely possible to consider an approach to MRC that takes variability into account. That said, I would not recommend it. I don't see the point of building an MRC, which is a delicate process when you have a model that you can calibrate directly on the whole hydrograph. I understand that some of the references cited do, but it doesn't make much sense to me. The MRC, in my opinion, is an empirical method that is applied when one wants to abstain from model fitting because of their too restrictive and inappropriate hypotheses on the dynamic (see the introduction of [11]). The authors' approach is hypothetical-deductive and relies on intelligible parameters to classify the dynamics. The references to the empirical MRC are, in my opinion, inappropriate and confusing.

I think better suggestions would most likely come from criticizing the approach, the Mangin model, and the classification framework. If one projects data from a complex, high-dimensional reality onto a lower-dimensional model and classification framework, there will be uncertainties that are both natural variability and projection artifacts. If they are too large, the message is that the model or framework is not helpful, and another projection must be found. Mangin developed his framework without taking uncertainties into account. Now that the authors have done so, relevant questions arise: should the dispersion of the parameter distribution, which reflects the sensitivity of the watershed, be taken into account in the watershed classification framework? How can this be done?

Furthermore, the i indicator was found to be a poor discriminant for the classification of the three study sites. Is this a poor indicator that should be removed from a classification framework, or can we expect that with other study sites, i may show more distinct and clear patterns? Should a nonlinear recession be considered in place of the Maillet equation? These are just a few interesting questions that arise when the authors allow themselves to critique the analytical framework.

Questioning it is not shooting oneself in the foot. I think the Mangin model and framework is a recognized and intelligible way of framing the dynamics of the karst system, and the authors have succeeded in providing a more robust way of doing this. The authors could stress this success while recognizing that they have also highlighted potential limits. Even so, the proposed automated framework could also be used more widely to assess the relevance of the Mangin model and classification framework in the future to propose relevant improvements. Finally, their approach is also transferable to other models and classification frameworks that will be developed in the future.

- SI-7: Reproducibility

This issue is not related to the introduction or the discussion section but to the lack of supplementary materials. The authors propose a technique. This technique is itself motivated by the idea of reducing user error and bias. Although the mathematics remains clear and straightforward, providing input/output files, code recipes, or examples will significantly help future users ensure that the technique is correctly applied and, therefore, be more in phase the paper's purpose.

- Comment on the title

The title is straightforward. However, "efficient" and "accurate" are not the best term. Efficient can mean many things, for instance, "computationally efficient" or well organized. Accurate is not the best given that the point is to depict and leverage the parameters' variability. The fact that the method is automated is stressed in the introduction and is probably more appropriate. The central classification task is missing in the title. I would suggest, for instance: "Karst spring recession analysis and classification: an automated method considering both fast and slow flow components."

References

- Wittenberg, H.; Sivapalan, M. Watershed Groundwater Balance Estimation Using Streamflow Recession Analysis and Baseflow Separation. *Journal of Hydrology* **1999**,

219, 20–33, doi:10.16/S0022-1694(99)00040-2.

- McMillan, H.K.; Clark, M.P.; Bowden, W.B.; Duncan, M.; Woods, R.A. Hydrological Field Data from a Modeller's Perspective: Part 1. Diagnostic Tests for Model Structure. *Hydrological Processes* **2011**, *25*, 511–522, doi:10.1002/hyp.7841.
- Shaw, S.B.; Riha, S.J. Examining Individual Recession Events Instead of a Data Cloud: Using a Modified Interpretation of $DQ/Dt-Q$ Streamflow Recession in Glaciated Watersheds to Better Inform Models of Low Flow. *Journal of Hydrology* **2012**, *434–435*, 46–54, doi:10.1016/j.jhydrol.2012.02.034.
- Karlsen, R.H.; Bishop, K.; Grabs, T.; Ottosson-Löfvenius, M.; Laudon, H.; Seibert, J. The Role of Landscape Properties, Storage and Evapotranspiration on Variability in Streamflow Recessions in a Boreal Catchment. *Journal of Hydrology* **2019**, *570*, 315–328, doi:10.1016/j.jhydrol.2018.12.065.
- Jachens, E.R.; Rupp, D.E.; Roques, C.; Selker, J.S. Recession Analysis Revisited: Impacts of Climate on Parameter Estimation. *Hydrology and Earth System Sciences* **2020**, *24*, 1159–1170, doi:10.5194/hess-24-1159-2020.
- Delforge, D.; Muñoz-Carpena, R.; Camp, M.V.; Vanclooster, M. A Parsimonious Empirical Approach to Streamflow Recession Analysis and Forecasting. *Water Resources Research* **2020**, *56*, e2019WR025771, doi:10.1029/2019WR025771.
- Hartmann, A.; Goldscheider, N.; Wagener, T.; Lange, J.; Weiler, M. Karst Water Resources in a Changing World: Review of Hydrological Modeling Approaches. *Geophys.* **2014**, *52*, 2013RG000443, doi:10.1002/2013RG000443.
- Stoelzle, M.; Stahl, K.; Weiler, M. Are Streamflow Recession Characteristics Really Characteristic? *Earth Syst. Sci.* **2013**, *17*, 817–828, doi:10.5194/hess-17-817-2013.
- Santos, A.C.; Portela, M.M.; Rinaldo, A.; Schaefli, B. Estimation of Streamflow Recession Parameters: New Insights from an Analytic Streamflow Distribution Model. *Hydrological Processes* **2019**, *33*, 1595–1609, doi:https://doi.org/10.1002/hyp.13425.
- Fiorotto, V.; Caroni, E. A New Approach to Master Recession Curve Analysis. *Hydrological Sciences Journal* **2013**, *58*, 966–975, doi:10.1080/02626667.2013.788248.
- Lamb, R.; Beven, K. Using Interactive Recession Curve Analysis to Specify a General Catchment Storage Model. *Earth Syst. Sci.* **1997**, *1*, 101–113, doi:10.5194/hess-1-101-1997.

Technical issues

Overall, the reading is understandable and straightforward as one feels the fear of not being understood correctly. However, as a result, the reading is sometimes wordy (lots of conjunctions) and redundant, with punctuation errors, past/present tenses getting mixed up, plural/singular mistakes, many missing articles, and other grammar errors. I recommend a few more iterations to correct, simplify and shorten the sentences. Accordingly, I have not reported all the errors and have gradually focused on the minor problems of interpretation that were not part of the SI sections.

- Abstract

L21: Mixing past and present. Missing word: ... parameters [that ?] reasonably represent ...

L22: The average KGE seems by far higher than 0.7 (see figure 3)

L25-27: On the recommendation for reducing uncertainties, see SP5 (Same remarks for the last sentence of the conclusion L525-526).

- Introduction

L40: missing word: time series [analysis?]

L43: a lumped attributes for karst systems

L49: describes

L69: I don't like the term "manually". It gives an impression of a "pen and paper" approach. I guess that recession fitting is still computationally assisted. I would replace manually with a more precise word such as visually supervised or validated by the user. If changed, run a find and replace to spot the word elsewhere in the paper.

L81: we can modify them

L85-99: the two first bullet items are redundant with the first sentence. I suggest removing the bullet list and keep the objective and method summarized in one paragraph.

L99-101: the? Master recession curve, the analysis. However, I recommend removing this last sentence. The motivation for fitting individual recession should be (and is already succinctly) discussed before so the last paragraph is entirely dedicated to what is shown in the paper.

- Data and Methods

L104: "generic", I would remove. These methods seem rather specific to me. They are called specific later on (L155).

L116: Do you apply the recession fitting on the smoothed points?

L120: "baseflow (non-influenced)" use the consistent term slow flow defined above (L108-110). Otherwise, it is confusing.

L123: a more strict

L128: Stoelzl --> Stoelzle

L133: the Askoy and Wittenberg iterative procedure is not described. How iterative? By truncating progressively the head of the segment?

L147: rule based and exclusion criteria --> rules and exclusion criteria

L148: I would use "aim at eliminating ..." instead of "ensure". By adding degrees of freedom, you acknowledge that they actually don't ensure anything.

L170-176-184: symbols can be simplified by using Q + subscript more often instead of Q , q , ϕ , ψ

L173: is referenced without being defined.

L188: requires + remove comma

L189: Eq.5 --> Eq.4

L192-193. The last sentence comes from nowhere. Put this last sentence below, after introducing K . Alternatively, you can start a new section 2.3 called "Karst classification".

L197: what five classes? I would definitely provide a more detailed description. Eventually

say that the classification is illustrated later on with Figure 5.

L203: I don't understand the definition of i with respect to Eq. 8. Should "the value of the quickflow recession" be "the proportion of remaining quickflow..."?

L208: move in the first paragraph.

L211: The figure is very clear but, perhaps, an illustration of the effect of η and ε would be nice as well.

L214-215: I will add some references about the inference of recession parameters from the MRC. Jachens et al. seems important enough to be placed in the introduction (see SI-1).

L221: "which is in reality untrue": I would not speak of reality or truthfulness. I would rather say that REMs are possibly not identifying the most relevant or optimal t_i .

L231: for clarity, "a predefined value" --> "the REMs predefined values".

L259: I would briefly precise why KGE is chosen (over Nash and Sutcliffe Efficiency, for instance).

L261-2: The Mangin classification is actually not described in section 2.2, only K and i . See comment above (L192-193). I would remove the last two sentences if the classification is correctly described in the appropriate section.

L262: This is used

- Test springs and data

L279 / Table 3 / Temporal res. : Since a reader first looks at the figures and tables in an article, specifying the time resolution as "hourly" and "sub-hourly" may mislead the reader into thinking that the methods have been applied at the time resolutions mentioned.

- Results

L292/Table 4: I miss some important statistics about the length of recession events, Q , and $-dQ/dt$. I believe that they could help in the interpretation of the results. If not, consider presenting these statistics in the supplementary materials.

L312: The modification ... produces ... interactions. To simplify ...

L314: Remove comma + a defined category

L316: The results show

L317: The estimation or estimations

L320-324: several missing articles. A much ...

L326: The infiltration rate, the curve concavity, ...

L329: "vice-versa" is confusing.

L331-335: I was frustrated about the patterns not being explained, but they are in the discussion section. Therefore, I suggest referring to the discussion section.

L337: The estimation .. the curve concavity ...

L347: What about Q_0 ? It is an important parameter to evaluate aquifer recharge and to perform hydrograph separation. I would recommend adding it or putting it in the supplementary materials. Besides, you mention the Q_0 compensation effect later on (L466).

L357: "The grouped mean" is confusing. Grouped how? Are the boxes min-max ranges?

L417-460: I miss a discussion about the role of spatio-temporal variability of rainfall and the full set of other factors affecting recession sensitivity to initial conditions

L431: I think that less permissive REMs will tend to include the early recession, hence, increase the recession coefficient α if the Maillet equation is used alone because the recession rate would be, on average, more pronounced when high flows are considered. However, I am not sure that this is true with the entire Mangin model that separates high from low recession flows, and I do not personally see in Figure 4 that "*permissive recession extraction tends to produce higher estimates of slow flow recession coefficients.*"

L441-442: what does "under" or "over-estimated" mean since there is no reference?

L443: higher infiltration rates

L486: a recent study ... categorises

L487: calculated

L491: methods involve

L658: Table A1 in the supplementary materials is unused and unexplained.