

Hydrol. Earth Syst. Sci. Discuss., referee comment RC1
<https://doi.org/10.5194/hess-2021-355-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on hess-2021-355

Anonymous Referee #1

Referee comment on "Tandem use of transit time distribution and fraction of young water reveals the dynamic flow paths supporting streamflow at a mountain headwater catchment" by Ravindra Dwivedi et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-355-RC1>, 2021

Recommendation to Editor:

I recommend this paper be rejected for publication in *Hydrology and Earth System Science* (HESS) in its current form. I recommend the authors resubmit after major revision. The topic is certainly of great interest in scientific hydrology. The combination of data sets *might* be better leveraged to make clearer inferences about the true range of water residence times in small headwater catchments. I provide three general criticisms here that are reiterated in a numbered list of specific comments about the manuscript and graphics.

One major criticism is that the work relies heavily on antiquated methodologies. Major portions of the results are based on application of the so-called lumped-parameter-transport model based on *time-invariant* transit-time distributions (TTDs; equation 1). The TTDs have time-invariant parameters, which assumes that the distribution of flow pathways within the landscape and associated water velocities are constant in time. This assumption defies intuition, but was applied out of convenience for decades [e.g., most works reviewed by *McGuire and McDonnell*, 2006]. A theoretical basis for analysis based on *time-variable* TTDs was presented as early as Lewis and Nir [1978]. The theory has been advanced by many recent works [*Botter et al.*, 2010; 2011; *Harman*, 2015; *Rinaldo et al.*, 2011; *van der Velde et al.*, 2012]. That TTDs should be time-invariant was shown to be theoretically implausible for low-order watersheds with dynamic flow [*Botter et al.*, 2011]—a result that has been supported by empirical results from manipulative tracer experiments [e.g., *Kim et al.*, 2016b]. I don't think our premier disciplinary journals should continue publishing results based on this antiquated approach.

A second major criticism is that there are significant shortcomings in the data, especially the measurements of Tritium in surface waters. There appear to be only 6 data points representing Tritium abundance in stream water. That's not many, but the authors rely on that data to calibrate and compare a range of models. Also, the time interval over

which precipitation is sampled is coarse. Much of the temporal dynamics of tracer concentration in that inflow will be lost. For these reasons, most of the parameters for the various TTD models are not uniquely identifiable. The different models generate markedly different estimates of different water age metrics. There is inadequate guidance, or rationale, for the reader to understand which, if any, should be considered correct.

The third major criticism is that the article does not clearly convey what outstanding question/problem in scientific hydrology is likely to be resolved through the elaborate set of methodologies employed here. The discussion section does not convey any new insights about flow processes in headwater catchments. Rather, that section seems to emphasize intricacies of the various technical approaches that lead to order-of-magnitude differences in water age metrics such as the fraction of young water and mean transit time.

My first recommendation for revision would include omission of methods and results that are based on lumped-parameter transport modeling using time-invariant TTDs. My second recommendation for revision would be a deep consideration of what is the specific gap in knowledge that the research would address, and a deeper discussion about what all these (somewhat abstract) age metrics tell us about flow processes, or the linkage between water age and catchment structure, that we didn't already know. Emphasis should be placed on new knowledge that may be generalizable across watersheds. This is especially important since the study focuses on a single watershed where quite a lot of tracer-aided flow and transport studies have already been conducted [*Heidbuchel et al.*, 2013; *Heidbuchel et al.*, 2012; *Lyon et al.*, 2008; 2009], including multiple previous works by the lead author.

Specific comments on content in the text:

- Line 23: The phrase "single age tracer" is unclear and not conventional. Please omit or rephrase.
- Lines 42-51: The rationale provided here is not very strong. To say that the processes being studied are "still incompletely understood" is not a very effective way to communicate (1) what aspects of the processes *are* well understood (because certainly a lot of prior knowledge does exist), (2) what is/are the explicit knowledge gap/s, and (3) how this research is designed to specifically address that knowledge gap.
- Lines 59-61: Doesn't really make sense to say that an underestimated quantitative metric affects actual substrate weathering in Earth's crust. The conjunctive phrase "As a result" in the following sentence seems out of place. Consider rephrasing.
- Lines 70-71: I think you should temper the language here. Robust? That implies the result should be representative across a range of systems. Yet the papers you cite apparently rely on simulated experiments in synthetic landscapes. Any evidence from the real world that you could cite?
- Lines 74-76: To help emphasize the knowledge gap, I think you need to clarify what

exactly is meant by “only one period”.

- Line 90: The question “what is the appropriate TTD type” is somewhat unclear. Precipitation is episodic. There is not a continuous inflow of water volumes with different ages entering any watershed. Therefore, the distribution of transit times (i.e., exit time – entry time) must also be discontinuous. This fact is illustrated by real TTDs observed from active tracer introductions [e.g., *Kim et al.*, 2016a]. They are quite messy and not continuous distributions. Any continuous function that is chosen as a TTD for application in lumped-parameter-transport modeling is therefore just an approximation of reality. If that is accepted as true, then it seems your question could be restated as “what mathematical distribution yields simulation results that best fit the data from this particular watershed?”. That is not a question of great relevance for scientific hydrology in general, in my opinion.
- Lines 91-93: I have a very hard time interpreting this sentence. Please rephrase. Again, I would suggest carefully explaining, or omitting, the phrase “age tracers”. Are you meaning to distinguish stable isotopes from radioactive isotopes?
- Lines 93-94: Suggest deleting “...as determined by stable water isotope tracers”. It implies to the reader that the answer to your more general question (i.e., “what is the discharge sensitivity of F_{yw} ”) is somehow conditional on this particular data type? Is that in fact what you think? If so, it raises some concern about the generality of the results.
- Lines 95-100: Suggest deleting all of this. A prelude to the methods elaborated on the following pages is unnecessary. The concluding paragraph of the introduction should highlight the identified knowledge gap then state the objectives of this study and how they address that gap. The final sentence raises some concern that the current work is partially redundant.
- Lines 112-113: So it was a notably drier than average 9 years, or the PRISM results are biased high here?
- Lines 130-131: This is a very coarse sampling resolution for the intended application of the data. Undoubtedly there are tremendous temporal dynamics in the stable-isotope composition of precipitation within and among individual storms that occur during 5-7 day intervals. The range of stable-isotope abundances in precipitation observed during individual storms may be comparable or greater to the range observed among monthly-aggregated samples collected across years [e.g., *Rozanski et al.*, 1993]. The true temporal dynamics of tracer concentration in precipitation are lost in a lumped sample that aggregates over 5-7 days. Any quantitative model that uses those tracer concentrations as input will be very limited in its ability to accurately simulate the temporal dynamics of the same tracer in the stream. That limitation seems very germane to the stated objectives of this study. Passive, sequential sampling devices are easy to make and deploy. Analysis of stable isotope abundances by laser spectrometry for large sample numbers is relatively inexpensive. This data limitation is hard to excuse.
- Lines 149-151: I can’t quite understand what this means. Please consider rephrasing.
- Line 164-166: Simplify the headings and sub-headings. Here and elsewhere there are sub-headings with no content underneath. Suggest deleting.
- Line 171: When you say “thereafter”, do you mean over longer time increments than 1 month? Please rephrase to clarify.
- Line 176: Some formatting inconsistencies with citations here and throughout the manuscript. Uneven use of open and closed parentheses and lack of spacing between cited papers within in-line citations. Proofread carefully. Suggest using “[(“ instead of duplicate parentheses. Also, I cannot find the Dwivedi 2019b entry in your bibliography. Is it missing? Put spaces between entries in the bibliography. It is terribly difficult to read through single spaced.
- Line 177: “expand on these results” again seems to suggest this is somewhat redundant with the previous works from the same catchment.
- Lines 185-190: Pretty sure $h(\tau)$ is the specified functional form of the TTD, but that is not stated in the paragraph.

- Line 193: I am not familiar with the Downhill Simplex method. It is described in a single sentence, yet it is apparently the method for evaluating how appropriate is one versus the other TTD model. Could you please elaborate just a little bit on what this is for the unaffiliated reader? The KGE is used as the "model performance criteria" but you say that the Downhill Simplex was used to evaluate "the performance of each TTD". This is confusing to me. Equation 1 is the model, but the variable performance of the model is due only to the selection of different functional forms of the TTD. So the performance of the model is a direct reflection of the performance of the function selected as TTD, no? Please clarify.
- Line 228, equation 5: Use "C" with "Q" and "P" subscripted to indicate concentration in streamflow versus precipitation. You already adopted this notation in equation 1. Be consistent here and in subsequent equations.
- Lines 348-357: What about all the other models? You only discuss PF and Gamma. The KGE of the 1d-ADE falls exactly between the values for the Gamma and PF models, yet the mTT estimated by the 1d-ADE is factors of 8-9 less than the mTT from those models, respectively. Why do you ignore the other models and what do you conclude from this order-of-magnitude difference? If I understand Figure 4 correctly, then only the "ADE-nx" and Exponential function as TTDs seem to generate uniquely identifiable parameters. Is that correct? Neither model is discussed at all here.
- Lines 390-392: The data are also far too sparse to reliably fit the parameters for TTDs used in equation 1. Isn't this confirmed by (1) the lack of unique solutions illustrated in most cases shown in Figure 4 and (2) the generally poor accuracy of all model simulations shown in Figure 5? I would argue yes.
- Lines 400-414: The text makes no allusion at all to Figure 7D, which has unusual qualitative axes and cannot be easily interpreted by the reader. The figure caption only provides a citation to a previous work to explain the graphic. More explanation is needed in this section of the Results, or Figure 7B should be deleted.
- Lines 437-440: I am unclear what is the importance or relevance of this concept of "short-term storage". What is it and why does it matter? In any case, you present estimates of this metric based on three competing approaches that vary by a factor of approximately 125 (0.08, 0.22, and 10.7)! Which, if any, should we believe is correct, and why?
- Lines 472-477: The results are highly dependent on the temporal resolution of the input time series. As I noted in a comment above, if a temporal dynamic in the tracer concentration in precipitation is hidden within a sample that accumulated over 5-7 days, then the model can't possibly simulate the effects of that dynamic in streamflow. The results are entirely dependent on the resolution of sampling the tracer concentration in inflow, and the resolution used in this study is quite coarse.
- Lines 499-501: You make a sweeping assumption here that the bedrock at several research sites is "water tight". That seems quite speculative. What evidence supports this assumption? Most rocks are fractured and jointed to some extent. Even exposed, granitic plutons commonly have sufficient fracturing and water storage capacity to host woody-stemmed plant communities and support inter-storm flow from emergent springs. More support for this assumption is needed here, perhaps through more extensive synopsis of the geology of the sites used in these cited studies.
- Lines 538-539 and 544-546: So, you're saying the fraction of young water estimates are invalid when based on the use of Tritium as a tracer?
- Lines 555-564: Here there are a series of sentences elaborating some intricate details of methodology which seem misplaced in the discussion. They lead to the ultimate conclusion at the end of the paragraph that "infiltration may activate deeper groundwater flowpaths". That is not a novel conclusion in scientific hydrology, and it is not even stated definitively here (i.e., *may*..). This paper uses a wide ensemble of methodological approaches, which, from my view, has only created ambiguity in how the markedly contrasting results can be interpreted. I find no new insight into hydrological processes resulting from all this computational effort.

Comments on Figures and Tables:

- Figure 3: Does the inset have a linear scale? If not, please make it linear. If so, please add more tick marks to the vertical axis so we can approximate the numeric values of the data points. Are these six data points all you have to calibrate the parameters of the TTD models used with equation 1? If so, that seems inadequate.
- Figure 4: Is the vertical axis the KGE? If so, please label it that way. I am unclear what the variable "response surface" on the vertical axis indicates.
- Table 1: Words and numbers should not be split between rows. Use emboldened lines, or no lines, to better delineate the content. This is not acceptable for a journal article. Please make it more presentable.
- Figure 5: Y-axis labels should have "3" as a superscript preceding "H" to conform to established conventions of symbolizing isotopes. Would suggest compressing this into a single graph with a legend indicating the results from different TTD models. The gray dots are the same across all 5 subplots.
- Figure 7: Here and elsewhere the font size is illegible. Please enlarge font on axes and in legends.

Works Cited:

Botter, G., E. Bertuzzo, and A. Rinaldo (2010), Transport in the hydrologic response: Travel time distributions, soil moisture dynamics, and the old water paradox, *Water Resources Research*, 46, doi:W0351410.1029/2009wr008371.

Botter, G., E. Bertuzzo, and A. Rinaldo (2011), Catchment residence and travel time distributions: The master equation, *Geophysical Research Letters*, 38, doi:L1140310.1029/2011gl047666.

Harman, C. J. (2015), Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, *Water Resources Research*, 51(1), 1-30, doi:10.1002/2014wr015707.

Heidbuchel, I., P. A. Troch, and S. W. Lyon (2013), Separating physical and meteorological controls of variable transit times in zero-order catchments, *Water Resources Research*, 49(11), 7644-7657, doi:10.1002/2012wr013149.

Heidbuchel, I., P. A. Troch, S. W. Lyon, and M. Weiler (2012), The master transit time distribution of variable flow systems, *Water Resources Research*, 48, doi:W0652010.1029/2011wr011293.

Kim, M., L. A. Pangle, C. Cardoso, M. Lora, T. H. M. Volkmann, Y. Wang, C. J. Harman, and P. A. Troch (2016), Transit time distributions and StorAge Selection functions in a sloping soil lysimeter with time-varying flow paths: Direct observation of internal and external transport variability, *Water Resources Research*, 52(9), 7105-7129, doi:10.1002/2016WR018620.

Lewis, S., and A. Nir (1978), On tracer theory in geophysical systems in the steady and non-steady state. Part II. Non-steady state - theoretical introduction, *Tellus*, 30, 260-271.

Lyon, S. W., S. L. E. Desilets, and P. A. Troch (2008), Characterizing the response of a catchment to an extreme rainfall event using hydrometric and isotopic data, *Water*

Resources Research, 44(6), doi:W0641310.1029/2007wr006259.

Lyon, S. W., S. L. E. Desilets, and P. A. Troch (2009), A tale of two isotopes: differences in hydrograph separation for a runoff event when using delta D versus delta O-18, *Hydrol. Process.*, 23(14), 2095-2101, doi:10.1002/hyp.7326.

McGuire, K. J., and J. J. McDonnell (2006), A review and evaluation of catchment transit time modeling, *J. Hydrol.*, 330(3-4), 543-563, doi:10.1016/j.jhydrol.2006.04.020.

Rinaldo, A., K. J. Beven, E. Bertuzzo, L. Nicotina, J. Davies, A. Fiori, D. Russo, and G. Botter (2011), Catchment travel time distributions and water flow in soils, *Water Resources Research*, 47, doi:10.1029/2011wr010478.

Rozanski, K., L. Araguas-Araguas, and R. Gonfiantini (1993), Isotopic patterns in modern global precipitation, *American Geophysical Union Monographs*, 78, 1-36.

van der Velde, Y., P. J. J. F. Torfs, S. E. A. T. M. van der Zee, and R. Uijlenhoet (2012), Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, *Water Resour. Res.*, 48(6), W06536, doi:10.1029/2011wr011310.