Interactive comment on "Quantifying new water fractions and transit time distributions using ensemble hydrograph separation: theory and benchmark tests" *by* J. Kirchner

J.W. Kirchner

kirchner@ethz.ch

I thank Riccardo Rigon (hereafter RR) for his comments, even if in some areas we do not agree. Below I reproduce his comments (in normal text), with my responses (in bold).

This paper presents some interesting ideas. The main topic, in my opinion, is the derivation by regression of the average of the backward probability distributions which can be used to infer the (mean) catchment behavior either related to transport or to the hydrologic response. Estimating the role of antecedent conditions by studying the mean shape of the distribution functions, which is the last topic treated in the paper, is

a great intuition that could become a classic.

Because I like the most the topic of the backward probabilities with respect to the one of new water/old water, I would change the title to put major emphasis on transit time distributions than on water fractions. I believe that finding a way to characterize the backward distributions by regression is a more general and better achievement than the hydrograph separation.

I suppose it is a matter of taste whether one thinks of new water fractions as special cases of transit time distributions, or whether one thinks of transit time distributions as generalizations of new water fractions, and transit time estimation as a generalization of hydrograph separation (as I do).

In the paper there is also section on the response time distributions, also known in literature as the forward distributions. Estimations made could be biased by the example used, i.e. of having just one outgoing flux. Since there is a simple relation between backward and forward probabilities. i.e. the Niemi relation (e.g. equation 34, and the whole section 7 in Rigon et al., 2016, or equation 1 in Botter, 2012), n outgoing fluxes require to make explicit n-1 partition coefficients, a fact that it is not so evident in this work because of the simple example used as benchmark. As Rigon et al. section 7 shows, there is also an empirical version of the Niemi's finding that works at any time time t, and this is the case treated. As not clearly stated in the paper, the empirical case does not match with a pdf but to a pdf divided by the partition coefficient.

The theoretical approaches of Rigon and Botter, including partition coefficients, are relevant to the theoretical modeling of forward transit time distributions, if those distributions measure the probability that a water drop that enters as precipitation today will leave the catchment today, tomorrow, or the day after (and so forth) *by all possible exit mechanisms combined* **(streamflow, evapotranspiration, etc.). To take an artificially simplified example: if all streamwater leaves the catchment after seven days and all evapotranspiration (ET) leaves the catch-**

ment after one day, then obviously the forward transit time could range from one day to seven days, depending on what fraction of the input ultimately leaves via streamflow vs. ET.

But these issues are really not relevant to the problem of *estimating forward transit time distributions from data using tracers*, because any such estimate will always be specific to *one particular exit mechanism* (whichever one is sampled). That is, the "forward" TTD's that are estimated in my approach are *necessarily* "forward" TTD's for the molecules that *enter as precipitation and eventually leave by streamflow*, since these are the fluxes that are sampled. We have *no information* about the life expectancies of the molecules that leave by ET, since we have no tracer data from the ET flux. This is not a special limitation to my approach; it is a general problem that will arise in every data-driven approach, unless and until we succeed in measuring tracers in catchment-scale ET fluxes. I will revise the text to make this clearer.

In this context, It should be noted by the "astute reader" that while the backward distributions assume knowledge of all the past, forward distributions assume knowledge of all the future, which cannot be the case of the analysis under my review (see also section 6 of Rigon et al., 2016).

Sure, if one wants to estimate the *infinite* tails of the transit time distributions, which my approach intentionally does not try to do, and which is *impossible in practice* anyhow (due to accumulation of errors, long before one reaches the end of the time series).

But as the equations in section 4 demonstrate (and as the benchmark test confirm), one can obtain ensemble estimates of both forward and backward distributions at time steps 0..m, by looking back only m+1 time steps from each discharge time step (and one does not need to look forward at all).

In any case the Author needs to be a little more explicative on these facts. A brief

C3

section is dedicated to talk about evapotranspiration and fractionation effects. Evapotranspiration is not present in the model used as benchmark and the way it is introduced is not clear. Reading the paper I will not be able to reproduce the Author's results and I suggest that this part, being unessential for the present study, could be omitted.

The point of this section is not to quantify evapotranspiration per se, but instead to test whether the estimates of new water fractions will be substantially affected by evaporative fractionation of the liquid that is left behind. While this section might appear "unessential" from RR's perspective, my perspective is different. It is too easy for skeptics to dismiss an approach like mine by saying "it doesn't account for evaporative fractionation so we don't trust any of these results." Therefore I think it is indeed essential to show that the results are not substantially biased by evaporative fractionation.

The paper is exceedingly long. Some technical parts on regressions, whilst important for research reproducibility, distract from the core topics and can be moved in my opinion to an Appendix or to some complimentary material.

Sure, the paper is long. But it also accomplishes a lot.

The "technical parts" can be skipped by readers who don't care about them (as the last sentence of the introduction clearly explains), but I think it is important that they are there. The reason is not just research reproducibility. It is also that one does not really understand what is going on without getting into how rainfree days are handled, and without seeing that the end result is not purely a multiple regression. Yet another reason is that these sections explain techniques, such as available-case analysis (Glasser's method) and Tikhonov-Phillips regularization, that are broadly useful but not widely known in our field.

In any case, moving these sections to appendices would actually make the paper *longer* in total, because of the need to mention all the key points both in the main text and again in the appendices. It would also make the presentation substan-

C4

tially less clear, due to the need for many "pointers" linking the main text and the appendices (and vice versa).

A further remark that embarrassed me a little. I believe that the coefficients beta of regressions, could be better understood in terms of the backward travel time distributions, and in my view, Rigon et al. 2016 can be a useful citation.

Equations 53 and 56, and the accompanying benchmark tests, demonstrate that the beta coefficients are _not_ , by themselves, the backward travel time distribution; instead they must be corrected for the fraction of time (or discharge, in the case of discharge-weighted TTDs) associated with time lags at which no rain fell.

Therefore, also looking to the minor notes I make below, I think the correct judgment is to go for major revisions (which I think will not require a lot of time though).

Riccardo Rigon

Detailed comments

Page 2 - Line 15 - Equation 1 - I think that an operational definition of new and old is required. In the subsequent text," new" is the discharge produced by the last rainfall interval (a day or a week) and " old" is the rest of the discharge. Specifying it here could be useful.

Of course in the specific context of _ensemble_ hydrograph separation, this is what "new" and "old" mean.

But Eqs. 1-3 describe _conventional_ hydrograph separation (not ensemble hydrograph separation). In this context it would be _factually incorrect_ to specify "new" and "old" in the way RR suggests, because this is not what they mean in conventional hydrograph separation.

Page 6- Line 20, Equation (10). Maybe saying that this is the mean backward distribution of travel times evaluated at lag 1 could be helpful. Probably I am biased to think

that way and does not correspond to the generic reader of this kind of papers. Anyway, it is my opinion.

I don't think this is helpful, except for readers who are already deep into the language and concepts of travel time distributions (who will already see this without having it explained to them).

Page 7 - Franky, I do not think sections 2.3 to 2.7 are so relevant. They probably reflect the genesis of the paper but they are full of technicalities and certainly scooped out by the more general section 4 of which these are just a particular case.

Although new water fractions may just be the m=0 case of a transit time distribution from a _theoretician's_ perspective, from a _practical_ perspective they are worth considering in their own right.

First of all, many catchments will have enough tracer data to reliably estimate new water fractions, but not nearly enough to estimate transit time distributions.

Second, the analyses presented in Figures 2-10 can be conducted easily on new water fractions, but not on transit time distributions (at least without huge masses of data).

Third, new water fractions will be much easier for readers to understand, because they do not require the full mathematical and statistical machinery that is involved in estimating transit time distributions. Readers will find it much easier to have an intuitive grasp of what is going on in these sections, than in the corresponding explanations of transit time distributions. Thus these sections are essential in building the reader's understanding.

Page 26 - Section 3.6 - Effects of evaporative fractionation - This section could be interesting but where is evapotranspiration in the model used as a benchmark ? So, how could have been it evaluated Figure 10 ? The indication given in the section are not exhaustive, and I suspect that going deeper in the subject would require major

work.

I suggest to take away it.

The description of the benchmark model makes clear that evapotranspiration is not explicitly simulated, but that instead the precipitation inputs can be considered as effective precipitation, net of evapotranspiration losses. This is a standard approach in conceptual hydrological modeling, and it is appropriate here because it eliminates the need to specify all the parameters that would be required for an explicit simulation of ET.

As the manuscript clearly explains, the purpose of this section is not to look at the effects of evaporation on the mass balance, but to look at how evaporative fractionation might affect the isotopic composition of discharge, and thus estimates of new water fractions. This is necessary and important, for the reasons that I outlined in my response when RR also raised this issue above in his general comments.

Page 28 - Definitions - All of it could be much more clearly explained in term of ager-anked functions (e.g. van der Velde, 2012, Rigon et al., 2016). I understand that the Author is a pioneer of the topic and derives everything from the scratch without being taught by anyone, but this is not useful for the general reader who will have great help from referring to those papers too.

Let's be clear: we are talking about just one paragraph of very simple definitions, where to cover the same ground using the van der Velde et al. (2012) or Rigon et al. (2016) formalisms would take significantly longer. Also, in this context age-ranked functions represent unnecessary complications, because they suggest that the system somehow is aware of water's age and selects or rejects it accordingly. More generally: it is important not to introduce terms and concepts that require specifying the age or age distribution of water in storage, because this is _unknowable in practice_ .

Page 29 - Equation 31 - q_{j,k} -> q_{ij}. Both notations are good but they should be used consistently throughout the paper.

Thanks for catching that typo (though I may switch to the comma-delimited notation throughout, so that cases like q_{i+k,k} and q_{j,k} are represented consistently).

Page 30 - Solution method - This is essential for results reproducibility, but, at the same time, not central for understanding the concepts. I think moving it as well as section 4.3 and 4.4 to an Appendix or to the complimentary material would make the paper more readable.

I think it makes more sense to highlight for the reader that if they don't care to know how this is really done, they can skip to section 4.5.

The point of including this material in the main text (it's only about 10% of the paper) is not just for purposes of documentation; it is also so that readers know that these techniques exist, because they are broadly applicable in other areas of environmental analysis as well. I have already heard from several individuals who have appreciated seeing this material.

Page 31 - Line 25. I would simply cut sentence 26 up to "missing values" at page 32

Sorry, this would make the presentation incoherent and misleading. It is important for readers to understand why complete case analysis will yield biased results. It is also important for them to understand why imputation methods are also a bad idea in this context.

Page 42 - Forward time distribution. I make my points in the general comments. I believe using the work by Niemi is easier and founded on literature.

The Niemi approach is _incorrect in this context_, because it requires assuming that the configuration, routing, and storage volume of the system _are the same_ at high and low flows. These are assumptions that I _do not_ make, because they

are generally not correct.

Page 43 - line 20 - Rewriting equation 57 into 58 seems to me a little pedantic.

I am _not_ being pedantic. Even though Eqs. 57 and 58 are _algebraically_ equivalent, they have _fundamentally different statistical behavior_, and will lead to _different_ estimates of the beta coefficients. That is because the variance introduced by Q_j will affect the results differently when it is factored into to the right-hand side (the explanatory variables) vs. the left-hand side (the response variable).

The point of showing this is precisely to caution users against thinking that they can rewrite the regression equations in this paper into algebraically equivalent forms, and get equivalent results. That's not what will happen. I will add something to the text to clarify this point.

Page 45 - Section 4.8 - I think it is sort of a Columbus' egg, a brilliant idea. The only possible objection is that results (not the method) can be biased by the use of the benchmark model.

I do not understand this objection. There is no good way to demonstrate the potential of the method without a benchmark model.

Obviously the specific results that are obtained will depend on the specific benchmark model that is used, as I point out in the paper. That's why I rather clearly state in the last paragraph of section 5.2 that the results presented here demonstrate that these analyses can yield accurate results (which cannot be demonstrated with real-world data because then we cannot know independently what the right answer is), but that they should not be taken as examples of what those real-world results would be.

Page 50 - Discussion - Should be shortened.

Even at its current length, the discussion is less than 10% of the whole paper

C9

(including appendices). In my view it is an essential aid for readers to understand several key points of context and interpretation.

References

Botter, G. (2012). Catchment mixing processes and travel time distributions. Water Resources Research, 48(5), n/a-n/a. http://doi.org/10.1029/2011WR011160

Rigon, R., Bancheri, M., & Green, T. R. (2016). Age-ranked hydrological budgets and a travel time description of catchment hydrology. Hydrology and Earth System Sciences, 20(12), 4929.

van der Velde, Y., Torfs, P., Zee, S., and Uijlenhoet, R.: Quantifying catchment-scale mixing and its effect on time-varying travel time distributions, Water Resour. Res., 48, W06536, doi:10.1029/2011WR011310, 2012.