

## ***Interactive comment on “Surface water and groundwater: Unifying conceptualization and quantification of the two “water worlds”” by Brian Berkowitz and Erwin Zehe***

**Brian Berkowitz and Erwin Zehe**

brian.berkowitz@weizmann.ac.il

Received and published: 23 November 2019

We sincerely thank Markus Hrachowitz (MH) for his detailed, constructive and very thoughtful assessment of our manuscript. We are confident that we can fully address his valuable recommendations in the revised manuscript, as detailed in our response to the individual points.

**MH:** In the manuscript “Surface water and groundwater: unifying conceptualization and quantification of the two water worlds” by Berkowitz and Zehe, the authors aim to raise awareness for the unjustified and unnecessary distinction between surface and groundwater hydrology. They make the case that both systems are controlled by

C1

largely the same principles – one of these principles being the energetically necessary development of preferential flow paths. The manuscript touches a highly interesting topic and it has, without doubt, the potential to become an important reference for future studies. Although I really like and appreciate the overall intention and objective of the manuscript, I also feel that the manuscript could strongly benefit from the authors taking a step back to re-reflect some parts of their work.

**RESPONSE:** We thank MH for constructive comments. In our revised manuscript, we will certainly modify the text to clarify the arguments and discussion, along the lines suggested by MH in the specific comments below. To expand on and re-emphasize the “philosophy” of our manuscript, we note the following: Scientific communication and cooperation between the two communities is generally very limited, largely because of perceptions that the research problems and natures of the system dynamics are distinctly different and even unrelated. Indeed, current theoretical frameworks and experimental methods to characterize, measure, quantify, and model fluid flow – and related transport of chemicals and energy – are generally distinct and separate. And yet, as these “worlds” are two perspectives of the same terrestrial hydrological system, we propose that their dynamics are governed by common laws and principles. Moreover, each community has developed sophisticated modelling and measurement capabilities – which have led to significant scientific advances over the last two decades – that could benefit the other community and address its outstanding, unsolved problems. We synthesize the methods and thinking of the two communities in an effort to spark sharing and integration of future research efforts.

**MH:** (1) Throughout the manuscript it remains somewhat unclear what the authors want to convey. While some parts of the text read like a very interesting, yet general reflection on the structural similarity of surface and groundwater systems and the general principles behind that, other parts are very technical descriptions of one specific aspect (i.e. power law transition distributions) of groundwater and potentially surface water systems. Both parts are very interesting, but there is little clear coherence between them in the text. It therefore remains somewhat vague, how and if power law

C2

transition distributions are in detail linked to the formation, evolution and function of preferential flow paths in surface water systems and how and if they can be expected to improve representation thereof in models.

**RESPONSE:** We agree that two these different parts of the manuscript differ with respect to level mathematical detail. In fact, we considered at length various ways to address mathematical aspects, and reflected on the whether or not to include mathematical expressions as done in section 3. At the time, we were concerned that this might make the manuscript too dense. However, in light of MH's comments, we will introduce several key revisions: (1) We will clarify in the Introduction that our manuscript synthesizes concepts and methods from the generally disparate surface water (catchment hydrology) and groundwater research communities, and is in this sense innovative. We are not simply reviewing material or offering an opinion. (2) In the context of (1), we focus in section 2 on showing – in phenomenological and conceptual terms – how the structure and patterning of fluid flow (and chemical species) in surface water and groundwater systems are similar, and thus amenable to similar quantitative treatments. In the revised manuscript, we will add basic equations relevant to the discussion, to offer a firmer quantitative background to the general principles we discuss. We will then clarify that section 3 (i) offers a first effort at defining specific conceptual and quantitative tools, (ii) introduces the CTRW framework in this concept, with a clear connection to microscale physics, and places it in perspective relative to the well-known advection-dispersion equation (so that presentation of 6, and (actually) 5) basic equations is necessary), (iii) offers a new development and insight, in terms of comparing and contrasting power law and inverse Gamma distributions (groundwater literature; more appropriate to describe long tailing in residence time distributions) with Gamma distributions (surface water literature), which includes 2 additional equations, and (iv) uses the preceding discussion and insights to show (with 4 additional equations) how surface water systems (catchment response to chemical transport) can be treated within the CTRW framework. We emphasize that we must retain essential mathematics, but will ensure that as the manuscript aims to bridge two communities, it

C3

has to be in a style that it is attractive for both audiences. (3) The reviewer asks how power law transition distributions are linked to the formation, evolution and function of preferential flow paths in surface water systems and how and if they can be expected to improve representation thereof in models. We clarify that power law transition time distributions are linked to the \*function\* of preferential flow paths, but formation and evolution of (surface water) preferential flow paths are a separate matter. To add more: A key point of the more general part of the manuscript, that we shall stress even more, is that we step clearly beyond a mere statement of what does not happen if preferential flow occurs (in the sense that preferential flow implies non-Gaussian residence times or travel distance; we have known this for a long time) — we provide a statement of what does happen. The fingerprint of preferential flow in the travel time distribution can be captured by a truncated power law (for the transition time distribution, which underlies the overall travel time distribution); and through the related exponent we can quantify the deviation from the well mixed Gaussian case. The work of Ederly et al. (2014), we present in section 2.3 suggests a further connection between the characteristics of the aquifer and the power law exponent in breakthrough curves. We argue that this should also hold for other complex media characteristics that relate to their spatial organization (such as the correlation length or topology of preferential flow paths). We are therefore suggesting that these insights offer opportunities to relate signatures of spatial organization in flow patterns to signatures of temporal organization in the breakthrough curves. For both perspectives, we can quantify organization using information entropy, as we will show in the revised manuscript. A specific example calculation will be included in section 2.3, based on Figure 3, to illustrate this point. These arguments also offer, ultimately, opportunities to test whether hydrological systems and their preferential flow networks co-evolve towards more energy efficient drainage, which can also be quantified.

**MH:** (2) Linked to comment (1), it is difficult to discern from the text what the original and novel contribution of this manuscript is. In other words, it remains unclear if the manuscript is intended to provide a review of the state-of-the-art together with guidance

C4

towards promising future research avenues or if the thermodynamic considerations and the proposed power law transition distributions are a novel development that is introduced here for the first time (which I do not suspect). I therefore strongly encourage the authors to more clearly distinguish between existing literature and potential novel

**RESPONSE:** Thank you for this very valid point. As explained in our response to comment (1), we will indeed better explain that the manuscript relies largely on the synthesis of work that has already published in both communities, and proposes a common conceptual framework (we avoid the term theory, as this would require even more mathematics). We will also more clearly point out the completely new and novel results (e.g., the comparison of power law and inverse Gamma distributions against Gamma distributions; the quantification of organization in preferential flow patterns using information entropy). The latter calculation was presented by one of the authors during the recent 2019 EGU Leonardo conference in Luxembourg (Zehe et al., 2019).

**MH:** (3) I strongly agree with the authors that eventually the surface- and groundwater communities need to converge towards “unified” conceptualizations – we are talking about the same physical system after all. However, and in contrast to the authors, I believe that different modelling approaches are not mutually exclusive. Rather, they should all be embraced and exploited to their fullest to learn about the system. In other words, I think, different modelling approaches are complementary in what can be learned from them. Here, the authors provide a beautiful example of how theoretical considerations and “physically-based” models can teach us about real world systems. Their findings that power law transition distributions may be more suitable than other, currently used distributions can be of considerable interest for other modelling approaches. I think it would therefore be very helpful to emphasize this complementary aspect.

**RESPONSE:** We really thank MH for pinpointing this and for the detailed thoughts below. It was not our purpose to return to the old and very unfruitful debate about “which model paradigm is better”, the conceptual one or the physically based one. So we will definitely revise this section (particularly stressing the valuable and highly relevant

C5

studies MH listed in his review) which corroborate that – regardless of the model type – modelling of chemical species transport must overcome the limitation of well-mixed assumptions. This can be achieved either as MH explained, using skewed transit time distributions (as the Gamma or inverse Gamma function), or the concept of SAS in conjunction with conceptual models for the mass balance. This can be also be done by using Darcy-Richards model in combination with particle tracking, and solving them in heterogeneous domains (as shown, e.g., by Edery et al. (2014)), and in partially saturated soils (e.g., Klaus and Zehe, 2010, 2011; Wienhoefer and Zehe, 2014) that represent preferential flow paths as connected, highly conductive pathways with close-to-zero retention properties.

**MH:** Now large parts of the manuscript read like as if typical top-down model approaches cannot deal with the celerity-velocity dualism, nor that these models could resolve the incomplete mixing. I strongly disagree with this notion. While it is true that simple, lumped convolution integral approaches have limited use, they are quite outdated and cannot be seen as state-of-the-art anymore, for reasons also highlighted by the authors in this manuscript. That they are still widely used in the community is a different problem. On the other hand, the authors claim that conceptual multi-box models similarly, cannot represent the system nor reproduce its dynamics in terms of both, water and chemistry. I disagree with this opinion in the strongest terms, as there are many papers, in particular over the last 5-10 years, in which the opposite was shown. I agree that standard conceptual box models cannot simultaneously reproduce water flows and solute concentrations (and thus water/solute age distributions).

However, as already shown almost 20 years ago by Seibert et al. (2003; HP), the addition of “mixing” assumptions and hydraulically “passive” mixing volumes has demonstrated great ability in resolving this problem (and thereby the “old water paradox” – which is not a paradox anyway, really). Of course, many of these papers (e.g. Shaw et al., 2007, JoH; Fenicia et al., 2008, WRR; Birkel et al., 2010, HP; Dunn et al., 2007, WRR; McMillan et al., 2012,) WRR) started with the assumption of complete mixing in

C6

the individual reservoirs. This was done not out of conviction but rather in absence of more detailed information.

However, soon it was realized that complete mixing in the presence of preferential flow paths is an unsuitable assumption. Subsequent work therefore adopted the use of incomplete mixing at least for the unsaturated root-zone, and which could either be time-invariant (e.g. Hrachowitz et al., 2015, HP) or time-varying as a function of wetness conditions (e.g. Hrachowitz et al., 2013; HESS). In a parallel development, similar considerations were made using the SAS-function approach (e.g. van der Velde et al., 2010,2012; WRR; Benettin et al., 2015a,b, WRR; Harman, 2015; WRR; Rinaldo et al., 2015, WRR and many others). Here, please note that the SAS-function approach is functionally and even mathematically (!) equivalent to the mixing coefficient approach above, as described in detail in Hrachowitz et al. (2016; Wires). The only difference is in the choice of the SAS-function (piece-wise linear for mixing coefficient approach and typically gamma or beta distributions for SAS approach) and in the way the involved processes are described semantically. It is true that many applications(!) of the SAS-function approach limit themselves to overly simplified representations of hydrological systems. However, this does not mean that the general concept behind it is invalid. The application of different individual(!) mixing coefficients (and thus SAS-functions) for different individual(!) system components has, on the contrary, already proven very useful (Fenicia et al., 2010, WRR; Hrachowitz et al. 2013, HESS; Hrachowitz et al. 2015, HP; Hrachowitz et al. 2016, Wires). With the new information provided in this manuscript, a logical next step should thus be to check if these types of models can reproduce power law transit time distributions and if power law distributions would be suitable as SAS-functions in different components of the system (e.g. in the unsaturated root zone for drainage and evaporation; in groundwater for drainage). In fact in our paper Hrachowitz et al. (2013; HESS), using incomplete, temporally varying mixing in the unsaturated zone, we found that the system overall transit time distributions (i.e. TTD of the modelled combined outflows) had longer than gamma distribution tails. An observation that we could not make too much sense of at that time. However, it could fit very well into

C7

the reasoning presented by the authors here. In that sense, this would be a beautiful example of how different modelling approaches could learn from each other. It would make the manuscript much stronger if the authors emphasized how their findings could be helpful for other modelling approaches and if the authors invested some more effort in being more accurate in their description of what different model approaches can do. In the current description the authors seem to equate what “is” done with these models with what “can” be done. This is not a valid assumption. In fact we can do much more than is typically done.

**RESPONSE:** As we stated above, we very much agree with MH about the complementary merits of both model worlds and will revise the related sections accordingly. There is nothing to argue against integral approaches for simulating transport. At the end of the day, the CTRW approach is also an integral approach. The beauty is that the latter can be linked – we think in a straightforward manner – to distributed models as acknowledged by MH.

Minor comments:

**MH:** p.3,l.96: what are “uncorrelated noise pattern”?

**RESPONSE:** An uncorrelated noise pattern is a variable field without a spatial correlation. This means that the value at any grid point bears no predictive information about the neighbor. The analogue is a white noise time series, i.e., a purely random signal without temporal autocorrelation.

**MH:** p.4,l.133: Hrachowitz et al. (2016) would fit nicely here as reference.

**RESPONSE:** Absolutely, we will include the reference.

**MH:** p.4,l.146: not sure why this is referred to as “weak interaction” here. Given its importance for flow velocities and sediment transport, should friction not cause a “strong interaction” between water and solids?

**RESPONSE:** These terms were introduced in Kleidon et al. (Kleidon et al., 2013). Of course, friction is in both cases, for open channel flow and for porous media flow, the

C8

main free energy sink. In an open channel, energy losses occur only at the wetted perimeter, and flow is slowed down in the boundary layer close to the contact line, but there is also a large cross-section where free flow occurs. In a porous medium, friction is everywhere, because the fluid is in contact with the solid at the entire inner surface - so boundary layer effects dominate the entire flow. Note that the maximum acceleration acting on surface water is one g ( $g \, dz/dz$ ); in partially saturated soil, it can be 50 g ( $g \, d\theta/dz$ ), but flow velocities are orders of magnitude slower, because dissipative losses are larger. This is what we mean by strong interaction. We will include this clarification in the revised manuscript.

**MH:** p.5,l.161 and elsewhere: why only chemicals? This is also true for the behaviour of individual water molecules. The difference is that we can normally not tag and distinguish them. But this does not make the general process valid for chemicals only. There is also imperfect mixing of waters of different ages/provenance. Please try to be more precise in these formulations. In addition, please note that this is exactly what incomplete mixing assumptions and/or non-uniform SAS-functions try to mimic.

**RESPONSE:** Thanks for this point, but here we need to be really precise. We cannot measure diffusion of H<sub>2</sub>O molecules in H<sub>2</sub>O, but we can measure diffusion of HDO and HTO in H<sub>2</sub>O. The velocity of these molecules will indeed differ from the fluid velocity, but it is not for the H<sub>2</sub>O fluid molecules. In the revised manuscript, we will add a note to clarify this and explain our focus on chemicals. This point is also addressed explicitly in the second paragraph of section 3.1. Additional note: Diffusive mixing of HDO and HTO is, however, a problem with regard to the existence of stable end-members. Assume two water bodies of largely different isotopic signatures, young and old water, in contact. Diffusive mixing among both water bodies will alter the apparent age of the probe to an intermediate age, even if there is no physical mixing.

**MH:** p.5,l.178: also depending on the pre-storm wetness conditions

**RESPONSE:** Of course, we will add this.

**MH:** p.6,l.203-204: I disagree with this statement. While it is true that conceptual mod-

C9

els are often applied in a careless way, there is a wide body of literature that describes the necessity and value of treating models as hypotheses that need to be tested (e.g. Fenicia et al., 2014; Kavetski et al., 2011; Clark et al., 2011;2015), which, in an iterative approach, allows us to learn about the system. In addition, in Hrachowitz and Clark (2017) we argue that ideally, these catchment-scale, effective parameters should reflect real world numbers. The actual challenge is to find these numbers, which is far from trivial for many parameters, but which was shown to be feasible for others, e.g. Master-Recession Curve (Lamb and Beven, 1997) or the storage capacity in the unsaturated root zone (e.g. deBoer-Euser et al., 2016; Nijzink et al., 2016)

**RESPONSE:** We agree that this is a sensible point. Allow us to expand on what we mean by the statement “that conceptual models provide an abstract explanation”. “Explaining” means to us to provide reasons that are consistent with the underlying theory and that can be tested independently from the data that have been used to train the model. When explaining the result of a mathematical model, it is of course sufficient to get a mathematical explanation, based on arithmetic or algebra. Hydrological systems are, however, physical systems, which means that we need to provide the physical causes for the observed effects. A “beta” of 9 in conjunction with the HBV beta may represent a mathematically feasible reason, but it is not a physical cause. Furthermore, this explanation cannot be tested by taking independent observations as beta is neither directly observable nor is there a clear, unique relation to observable catchment characteristics. For a reductionist model, we can explain the model performance using, e.g., the fill and spill idea and the related storage volume in bedrock pools. This can be observed using geophysics or augers and bedrock can be represented in the model (see(Loritz et al., 2017)) via the subsurface permeability field. The structure of the model is an image of the structure of the landscape, and depths to bedrock can be inferred from data or related assumptions can be tested in a fashion that is independent of the calibration exercise. To us, this is (by the way) the core idea of PUB - building models that can be informed by data which are independent of discharge.

In the revised manuscript, we will better stress that these points, specifically in terms of

C10

the representation of partially saturated soil in conceptual models. We do not question the value of a master recession analysis nor the concept of dynamic root zone storage. We personally also think that statements like “this area works like a plateau”, “slope or wetlands is a cause”, etc., are very informative, in cases when landscape entities can be represented by unique and typical conceptual model building blocks. We will stress this even more in the revised manuscript, although we are not entirely convinced that conceptual modelling has already reached this stage.

**MH:** p.6,l.210: I also disagree here. These relations remain specific for the spectrum of environments they have been developed for. If they relations were developed using a wide range of different landscapes, as done for example for the determination of global parameters in the mhM model (e.g. Samaniego et al., 2010) than they will also give us a more general picture.

**RESPONSE:** This is indeed a valid point, and we will revise this statement to stress that the difference between the HBV and the mhM is the more realistic representation of the partially saturated soil. The mhM uses the retention curve after Brooks-Corey (Brooks and Corey, 1964) for soil moisture accounting (not the beta store). This is a reasonable representation of capillary forces acting on soil water and the additional asset is that the entire knowledge of the soil physical community can be in the regionalization.

**MH:** p.6,l.219: No, I strongly disagree! Why should a parameter that describes the aggregated effects of heterogeneity be an abstract quantity? I think it rather provides the macroscale perspective and it can in some instances already be independently observed/estimated (see comment above about Master-Recession-Curve and storage capacity in root zone)

**RESPONSE:** As we stated above, we mainly criticized the representation of the partially saturated zone and of hillslope scale processes. Given the success of the mhM it makes a difference whether the beta store concept of a retention function is used. We do not question the Master Recession curve, as it may be nicely related to the Darcy equation (de Rooij, 2013). Also storage root zone capacity is a valuable concept. But

C11

the latter does not imply that soil is unimportant, as is sometimes stated. There would not be any storage of water without capillarity. So the term “root zone storage” implies that capillarity exists, otherwise the roots have to tap groundwater.

Last but not least, a simple mathematical representation does not automatically mean that the model structures and the parameters are straightforward to interpret. This is evident not only in the fact that internal conceptual model states and parameters are difficult to compare and derive from field observations, but also in that model structure is neither simple nor intuitive to interpret. For example, the FLEXTOPO model structure is abstract, and it is not simple to infer the plateau, slope and wetland from the corresponding combination of storage elements. Physically-based models are not based on simple mathematical equations, but their spatial setup reflects much more intuitively our perception of how the system looks, particularly in the subsurface. And at the end of the day, at least finite differences based numerical solutions are founded on algebra and thus by no means more complex than the math underlying conceptual models.

**MH:** p.6,l. 223-224: no, although there is without doubt some tendency to believe that parameter search can help, I do not share this. I rather think that, for all the uncertainties involved, we need to start with step-by-step limiting the feasible parameter space by identifying and eliminating solutions that are inconsistent with our data and our understanding of the system. That is essentially the opposite of finding the “optimal” parameter set and relatively independent of parameter search algorithms.

**RESPONSE:** Agreed, but as we discuss elsewhere in the manuscript, it is preferable to constrain parameters that have physical meaning (such as root zone storage or saturated hydraulic conductivity).

**MH:** p.6,l.231-234: with all appreciation for the authors, but this is too bold a statement, which I need to consider as plainly wrong. First, these models \*can\* resolve the celerity/ velocity dualism, when implemented as described above (mixing coefficients/SAS function). And this, second, has already been demonstrated with a plethora of articles on combined modelling of water flow and (non-)conservative solute concentrations

C12

(e.g. chloride, nitrate, silica, ANC, EC, DOC and many others; see above references and references therein)

**RESPONSE:** In the context of these specific examples, we agree completely and will revise the manuscript accordingly. As an aside, though, we wonder whether the runoff components simulated with conceptual models can be used to constrain the average particle velocities in the respective flow domains. The latter would require being specific about the wetted cross section in these flow domains, with the mass balance used to partly constrain particle transit time distributions. In this respect, the reductionist model approach may be somewhat more consistent, but this might be seen as a matter of taste.

**MH:** p.7,l.247-263: while Darcy-Richards does indeed, and probably rightfully so, receive criticism, I think these problems can be fixed within the near future (as also suggested by the list of improvements given here by the authors). I rather consider another point why these models receive much criticism: the fact that the sheer number of parameters needed can never be observed at the spatial resolution and scale(!) of interest. Either we then need to use our scarce, existing observations to inform our model, in which case we will upscale homogeneity (as the spatial correlation fields of our system properties are unknown!), or we will need to calibrate these models, in which case we will obviously run into the problem of equifinality and our inability to meaningfully constrain our models.

**RESPONSE:** We think it is very important to critique Darcy-Richards models, as they rely on incomplete physics, just as it is important to comment on conceptual models. Some of the criticism is simply wrong. Darcy-Richards models can be informed using distributed data sets, at the catchment scale, although this is often claimed to be impossible. Darcy-Richards models can simulate preferential flow as well as conceptual models. We do not see that they contain so many more parameters; the Brooks and Corey model has 4 parameters, two of which are easily constrained, the beta store has 2 parameters, neither of which is easily constrained. We fully agree that they are subject to equifinality, but the parameter sets are also easier to be constraint. At the

C13

end of the day we can use hopefully infer these parameters from thermodynamic optimality. We note that the power law exponent in the breakthrough curve, as analyzed by means of the CTRW concept, has a clear relation to aquifer characteristics (e.g., mean, variance of conductivity) as has been shown by Edery et al. (2014). In the likely case that there is also a clear relation to the correlation length, it may be possible to use breakthrough curves to infer this information and parameterize distributed models. We will add a comment in this vein in the revised manuscript.

**MH:** p.7,l.264-265: why “alternative”? should these tracers not also be very informative and helpful to implement and test the above models?

**RESPONSE:** Agreed. We will modify the wording here.

**MH:** p.7,l.264-270: true, but mostly outdated, as these approaches are too simplistic.

**RESPONSE:** Given that this is “true”, we prefer to retain this material. However, in the revised manuscript, we will note that newer approaches have been advanced.

**MH:** p.7,l.271-283: in this paragraph the authors seem to be confusing different approaches. At least it reads in a rather incoherent way. For example, Rinaldo et al. (2015) and Botter et al. (2011) describe the same general concept. Further, the system overall SASfunction can also emerge from the choice of SAS functions from individual components, which can be a calibration parameter or informed by observations/theoretical considerations as demonstrated by the authors of this manuscript. Finally, yes, the gamma distribution is often used, but as often other choices, such as the beta-distribution (e.g. van der Velde et al., 2012) or some piece-wise linear distributions are made (Hrachowitz et al., 2013,2015). In any case, and as mentioned above, these choices are not necessarily made out of conviction but rather due to a lack of more information and can be easily adjusted. In particular, it will be very help to get a better understanding how the tails of the distributions should look like.

**RESPONSE:** We thank MH for pointing this out and will revise this section accordingly. We agree strongly that the inverse Gamma function might be a useful concept in this context.

C14

**MH:** p.10,l.364-370: unclear in how far this is different to what mixing coefficient/SASfunctions do.

**RESPONSE:** We agree with MH that in an integral context, age-ranked storage reflects this continuum of scales and water ages. This part of the discussion focuses on physical processes and dynamics, and how to conceptualize them, not how to model them. So we do not see any need to refer specifically to mixing coefficients (as designed for \*any\* of a variety of specific models) or SAS functions. We will add a note in the revised manuscript.

**MH:** p.10,l.375-378: I believe there is quite a good understanding in the community that there is no binary distinction between separated pools of water. The “two water worlds” need to be understood, prosaic literature terms, as a hyperbole – an exaggerated analogy, i.e. a pointed description of a concept.

**RESPONSE:** Thanks for pointing this out. While we acknowledge the usefulness of metaphors as a means of communicating science in an appealing way, this should be not overdone. There is evidence that pockets of mobile and immobile water ages can exist in very close spatial proximity; these pockets will have strongly different ages. And it is clear that there will be diffusive exchange of, e.g., stable isotopes between these pockets (as discussed above). This suggests that a distinction of different water worlds based on isotopic signatures is not straightforward. Every distinction that goes beyond the well-established blue and green water concept of Marlin Falkenmark, which is in fact based on straightforward soil physics and can be inferred from the soil water retention curve, remains highly uncertain. Given the extensive debate and literature on this subject, what is a “good understanding” can be considered uncertain. We state clearly that this picture of a distinct separation is indeed a highly idealized interpretation, and include this discussion as part of our overall appraisal of the “critical zone” in section 2.2. However, in light of MH’s comment, we will clarify this point further in the revised manuscript.

**MH:** p.10,l.379-380: perhaps better “drainage and evaporative fluxes”

C15

**RESPONSE:** Agreed. We will modify in the revised manuscript.

**MH:** p.11,l.431 and elsewhere: I am not sure if the term “chemical transport” is the best way to express what you want to express here. why not keeping it more general to “transport”, which then indeed is further divided into physical, bio-physical and chemical components in the transport processes.

**RESPONSE:** This term is used often in the groundwater literature. Here, we are referring explicitly and specifically to transport of chemical species, which is of course affected by physical and biogeochemical processes.

**MH:** p.12,l.454: conservative transport?

**RESPONSE:** Yes, the simulations here assumed that the chemical is conservative. We will clarify this in the revised manuscript.

**MH:** p.14,l.500: where can I see this?

**RESPONSE:** This inadequacy of the advection-dispersion equation is seen, for example, in Figure 4 (which is referred to in l.495); see the red dashed-dotted curve. We will clarify this in the revised manuscript.

**MH:** p.15,l.500-503: sure, but would be good to introduce what CTRW is before shown results.

**RESPONSE:** This section 2.3 focuses on a more general conceptual discussion, in keeping with the style of sections 2.1 and 2.2. We state clearly that the CTRW framework and governing transport equations are detailed in section 3.3; we prefer to retain the explanation of CTRW in there, which leads naturally into the quantitative analysis of various transition time distributions and long tailing of overall transit times.

**MH:** p.15,l.550ff: as also demonstrated by Hrachowitz et al. (2015 – sorry to again bring in one of my own papers here) using a conceptual model with mixing coefficients, where we showed that conservative (!) chloride transport is slower than water transport (evapoconcentration)

**RESPONSE:** Excellent. We will include here this reference and the result in the revised

C16



manuscript.

**MH:** p.17,l.594 ff: I really like this analogies here, but a stronger explanation of and link between spatial and temporal heterogeneity needed. Why are they different? What is going on there?

**RESPONSE:** We think that the key is to think in space-time and complementary manifestations of heterogeneity of preferential flow. Heterogeneous flow patterns that emerge in space (defined as deviation from uniform flow, which is the maximum entropy pattern), translate into signatures in the breakthrough curve, which are observed at a fixed location as a function of time. Again, this is through deviations from Gaussian travel times (which is the maximum entropy travel time distribution, reflecting well-mixed, advective-dispersive transport). When looking at the example of Ederly et al. (2014), it is important to stress that in the preferential flow paths, particles pass through “bottlenecks”, which would be excluded as possible flow path when using, for instance, percolation theory. This is because the higher local gradient in pressure head “pushes” the water through the bottleneck. As the gradient in pressure head reflects the potential energy difference, this corresponds to a large flow against a large gradient and thus a local power maximum. We will further emphasize this in the revised manuscript. The overall key messages of section 3 are that (a) CTRW is consistent with the ADE and advances beyond it, particularly in terms of capturing dispersion and tailing effects, (b) the power law exponent is related to porous media characteristics as well as the flow conditions, so that we can infer spatial signatures from temporal ones. In the revised manuscript, we will also clearly articulate these messages in the introductory discussion.

**MH:** p.17,l. 632ff, section 3.3: in how far is this different to SAS? This does not clearly come out here. I think it would be very interesting for the reader to clarify this. Is there a fundamental difference or is it only the choice of the transition time distributions?

**RESPONSE:** We realize that there is lack of clarity regarding “travel time distributions” and “transit time distributions”) in the manuscript. A “travel time distribution” refers

C17

to the overall response of a domain – catchment, soil column, or aquifer – from an input point/plane to an output (measurement) point/plane. We will certainly clarify this critical point in the revised manuscript, both when we first introduce these terms and throughout the manuscript, where appropriate. And we agree that this is linked strongly to age-ranked storage, which is the integral of the travel time distribution over all ages.

**MH:** p.17,l.646 and elsewhere: I noted the use of “transition” times instead of what is standard use in surface water hydrology: “transit” times. Was this made on purpose? Is there a difference? If yes, please specify. If no, please also clarify.

**RESPONSE:** Please see our response to the previous comment. We certainly agree that there is a lack of clarity and uneven use of terminology in our communities. We will clarify this key point in the revised manuscript.

**MH:** p.17,l.670: water molecules, too, are subject to both! It is just that its more difficult for us to observe.

**RESPONSE:** We pointed this out explicitly in the discussion in the second paragraph of section 3.1. In the revised manuscript, we will refer to that discussion here, too, for clarity.

**MH:** p.21,l.802-804: because that was the best information we had so far.

**RESPONSE:** Agreed. We were not criticizing anyone here, only stating a fact.

**MH:** p.22,l.821: not sure this is correct “. . .sum of a gamma function (which is also a gamma function). . .”. Please check if this makes sense. Do you mean the scaled incomplete gamma function? Should it not rather reader something like:”the sum of 2 random variables that follow a gamma distribution can also be described by a gamma distribution”?

**RESPONSE:** Thank you for noting this. We certainly agree (our wording was poor), and we will correct this in the revised manuscript.

**MH:** p.22,l.824: well, no. Not if suitable local (i.e. individual for each system component) mixing coefficients/SAS-functions are chosen.

C18

**RESPONSE:** We agree that the proper choice of an SAS goes beyond the assumption of perfect mixing. We will revise this part. The key point is, however, that we have been using similar concepts to characterize routing and transport; both phenomena are affected by preferential flow and we use the same distribution function to describe the related fingerprints of preferential flow. We think the CTRW concept is somewhat more general in capturing both the short tail and the long tails. We will stress this even more.

**MH:** p.22,l.839-841: I am not sure that that reference is a good example. First, the NSE is a very pessimistic performance metric in cases there is a low signal-to-noise ratio in the target variable (which is the fact in the Weiherbach: the stream isotope composition is strongly damped and essentially plots close to a straight line – any small deviation from that – error or real process – causes strong effects on NSE). Second, the choice of model also only allows a rather rudimentary partitioning and routing of water fluxes, which will have a considerable effect on the tracer composition.

**RESPONSE:** We agree and we will revise this section accordingly. The overall key messages of section 3 are that (a) CTRW is consistent with the ADE and advances beyond it, particularly in terms of capturing dispersion and tailing effects, (b) the power law exponent is related to porous media characteristics as well as the flow conditions, so that we can infer spatial signatures from temporal ones. We are pleased that we can make these points without criticizing the work of Rodriguez et al. (2019).

**MH:** p.23,l.868-873: it is difficult to judge for the reader if the CTRW framework is more physically justified than other models. What also remains unclear: is CTRW necessary to model long tails? Or can suitable distributions be used in other (conceptual) models to reproduce similar results?

**RESPONSE:** First, the CTRW framework is physically justified, and certainly is highly effective in capturing long tails. The references to CTRW cited throughout this manuscript provide the background (and, e.g., Figure 4 demonstrates long tailing phenomena); clearly, we cannot review all of the evidence for this in the current manuscript.

C19

Second, we are not claiming that CTRW is more physically justified than other \*physical\* models; we state only that CTRW is physically justified, as is incorporation of an inverse Gamma transition time distribution. The reviewer agrees, for example, that use of a Gamma distribution is convenient, but has not been given a physical basis for its use. Finally, we do not claim that other suitable distributions are unable to reproduce similar results, although we are unaware of any other \*physically-justified\* model that yields such results. In the revised manuscript, we will include further text clarifications on each of these points.

References: Brooks, R. H., and Corey, A. T.: Hydraulic properties of porous media, *Hydrology Paper*, 3, 22-27, 1964. de Rooij, G. H.: Aquifer-scale flow equations as generalized linear reservoir models for strip and circular aquifers: Links between the Darcian and the aquifer scale, *Water Resour. Res.*, 49, 8605-8615, doi:10.1002/2013WR014873, 2013. Klaus, J., and Zehe, E.: Modelling rapid flow response of a tile drained field site using a 2d-physically based model: Assessment of “equifinal” model setups, *Hydrological Processes*, 24, 1595 – 1609, DOI: 10.1002/hyp.7687., 2010. Klaus, J., and Zehe, E.: A novel explicit approach to model bromide and pesticide transport in connected soil structures, *Hydrology And Earth System Sciences*, 15, 2127-2144, 10.5194/hess-15-2127-2011, 2011. Kleidon, A., Zehe, E., Ehret, U., and Scherer, U.: Thermodynamics, maximum power, and the dynamics of preferential river flow structures at the continental scale, *Hydrology And Earth System Sciences*, 17, 225-251, 10.5194/hess-17-225-2013, 2013. Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J., and Zehe, E.: Picturing and modeling catchments by representative hillslopes, *Hydrol. Earth Syst. Sci.*, 21, 1225-1249, 10.5194/hess-21-1225-2017, 2017. Wienhoefer, J., and Zehe, E.: Predicting subsurface stormflow response of a forested hillslope - the role of connected flow paths, *Hydrology And Earth System Sciences*, 18, 121-138, 10.5194/hess-18-121-2014, 2014.

---

Interactive comment on *Hydrol. Earth Syst. Sci. Discuss.*, [https://doi.org/10.5194/hess-2019-](https://doi.org/10.5194/hess-2019-C20)

C20

