Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-108-RC1, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Spatial and temporal variation in river corridor exchange across a 5th order mountain stream network" *by* Adam S. Ward et al.

Matt Cohen (Referee)

mjc@ufl.edu

Received and published: 3 June 2019

Adam and colleagues have developed a truly impressive data set from which they test a specific hypothesis about scaling of river corridor exchange. The topic is important, not least because it challenges some of the major pronouncements derived from steady-state models that assume network scaling rules. The technical treatment of the breakthrough curves is exemplary, spanning the full complement of modern approaches, and the writing is uniformly clear and compelling. In short, this is a paper that clearly merits publication. Below I document several areas where I found the paper in need of clarity, with one area in particular inviting at least some additional discussion if

not some new analysis (#2 below). My recommendation of minor revision is predicated on the former (discussion), recognizing that some additional statistical treatment of the responses would more accurately considered major revisions. I've also created a list of minor comments, provided in no particular order (typos, questions, comments).

1) Among the many technical strengths of this paper is the breadth of response metrics interpreting solute breakthrough curves. It is truly a smorgasbord of measures, consistent with the assembly of masters that comprise the author list. After a while, however, it ceased to be clear to me why so many metrics were necessary. The hypothesis is about predicting river corridor exchange with discharge, and while I would admit (and their results confirm) that we probably lack a singular measure of that exchange, the methods provided no specific rationale for the ones selected other than literature precedent, nor justify their independence from others selected. In figures 5 and 6, skewness finally emerged as the "response" and much of the paper would have been easier if the adequacy of this metric were proposed at the outset, justified theoretically, and supported empirically (e.g., as meaningfully covarying with other more complex response measures). Otherwise, despite an elegant hypothetico-deductive framework, the resulting effort feels a little like metric-fishing. I don't recommend removing metrics, but rather suggest making their selection strategic (rather than exhaustive) and supportive of general inference (rather than analyzed in parallel). And where that rationale is forced, then consider removing.

2) The setup for the research effort was exemplary. In the intro, the authors convey the existing conceptual model of river corridor exchange driven simultaneously by time- and space-varying discharge, as well as stream and valley geomorphic variation. A naïve view might be that these aspects act independently, but since changing discharge alters the head gradients that enable river-porewater exchange, and also the lateral and longitudinal geometry of the stream channel, the intro text points clearly to the plausibility (even primacy) of interactions. For this reason, the insistence on pairwise regression is confusing. There's a single passing acknowledgement (P20L21) that a

multivariate approach may be useful but no effort to explicitly consider contingency as a native feature of the question at hand. Framed as a question: is current theory consistent with interactions between geomorphology and discharge being important, or would such considerations be mostly a statistical contrivance? I believe it's the former, and that there's an opportunity with this data set to set the stage for future explorations of such interactions. If the authors agree, I think at least passing consideration of interaction terms is merited. If instead the authors feel conditional relationships are not implicitly supported by theory, say so explicitly. I'd note that the presentation of the Wondzell model in Fig. 6a implicitly suggests that the influences of watershed area and hyporheic potential are conditional (although in an additive sense); my contention is that there may indeed be informative interaction terms, and few data sets before this one are adequate to that challenge.

3) One core reason articulated (intro and discussion) for reduced river corridor exchange at high flow is that augmented hydraulic gradients to the stream compress hyporheic flowpaths. This is true when the hydraulic response in the stream and hillslope are synchronized. It seems demonstrably untrue otherwise, such as when flow generation is uneven (in small catchments) or when rainfall is uneven (in large catchments). Perhaps these are special cases, but the rivers where I've worked extensively exhibit significant "bank" storage during floods when storm-induced head changes are more rapid and pronounced in the stream than in the adjacent aquifer. The resulting hot moments of groundwater pumping into (and later out of) the hyporheic and bank sediments indicate that a simple monotonic association between instantaneous exchange and discharge is probably naïve. Only slightly less oversimplified might be to interrogate the river corridor exchange as a function of hydrograph position (or the time-rate of change of discharge) rather than discharge alone. I recommend the authors consider this. We did this for a setting where tidal variation created interesting hysteresis in hydraulic exchange (Hensley et al. 2015 WRR) and others (Audrey Sawyer among others) have seem similar dynamics. It's reasonable to rebut this comment by saying that explicit consideration of hydrograph position (or dQ/dt) invites an entirely different

paper, but the general critique of steady-state assumptions that underlies this work might be bolstered by avoiding the view of variable stream discharge as a sequence of steady-states. It is not.

Minor Comments: - P1L43. Should be "is" not "are". Or "exchange" should be "exchanges" - P2L4. The inclusion of the "and" between #2 and #3 underscores the interaction effects that may exist. - What does it mean (P6L19) for streams to change on annual to subannual time scales? Doesn't everything that changes at any time scale vary at all time scales? Do you mean that the streams change quickly? - I don't understand the rationale for stratifying by stream order (P7L5); more precisely, I don't understand why it was advantageous to bias the sampling to headwater sites over higher order reaches. The point here is not to characterize the network (where we might expect most of the variation to occur in the low order streams), but rather to explore geomorphic vs. discharge controls on river corridor exchange. To that end, a more balanced portfolio of sites makes more sense. I'll note that the resulting sample population (Fig. 3c) is pretty impressively distributed so this comment is more conceptual than operational. - It's been a while since I took a groundwater class, but why is the porosity term in the subsurface flow equation (P8)? Darcy's Law applies to the bulk cross section (here valley width times mean colluvium depth) and the Hvorslev K is for porous media. - I really appreciate the guidance on standardizing the reach length by wetted widths. I think this is an important standard operating procedure. - P9 refers to a companion manuscript. What/where is that? - The equations on P10L7-8 appear to have a typo. Doesn't the comparison for the conditional value have to be between CADE and COBS? I am confused how it could be CAD. - I really like the fMTS metric. It would be informative to consider how this compares with H (which I like less because I'm too dense to really understand it) and skewness (which I like a lot as well). For what it's worth, it was upon introduction of holdback (H) that the array of metrics started to seem excessive (or at least poorly defended). Some correlation among metrics (e.g., as a supplemental table) would be helpful. - For the SAS analysis, I was impressed by the explanation and by the utility of the metrics. I'd only note that the discharge used (to compare against storage) is only surface stream discharge. The subsurface discharge (downvalley groundwater flow) is not included, and the relative importance of this flow depends strongly on network position. - P20L5 should be "hold" - The criterion of statistical significance is, of course, defensible, but I don't find the associations compelling just because they meet the criterion of being non-zero. The authors aren't trying to hide behind statistical significance, but seeing Table 2 made me wonder if the real story of these data (namely that we are really very poor at prediction of the thing we care most about) aren't a little too softened by putting pluses and minuses in almost every box. - On the subject of Table 2, I wonder if the predicted sign might be included somehow. For example, I would have (admittedly naively) predicted that skewness is reduced with increasing Q, UAA, V, order, width, and stream power, but perhaps not sinuosity or K. - QHEF on page 23 has the "HEF" subscripted. Elsewhere it's just "QHEF". - It's a little incongruous to show the overarching concept (Fig. 5) using watershed area and hyporheic potential, but then only use discharge for the pairwise plots. They are (Fig. 3a) clearly correlated, but not perfectly so. - Among the most important points is P31L10-12. We are mostly measuring in-stream storage with these short-term pulse tests. Unless we suppose that these high turnover storages are where most of the reactivity occurs (and I don't believe they are), efforts to link pulse-based breakthrough curves in a reach to network scale retention seems doomed to failure. The inclusion of metrics of storage proportion labelled by tracer is really informative.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-108, 2019.