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

Referees' comments in bold type. Authors responses below each comment.

Matt Cohen (Referee #1)
mjc@ufl.edu

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).

No response necessary as issues are detailed below.

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.

Accepted. We have restructured the methods section to now include a discussion of why these multiple approaches were implemented. Perhaps of most utility to the community, we have added a new table (Table 2 in the revised manuscript) that details the relative strengths and weaknesses of each approach. This table also include a summary of the key metrics that are interpreted from each approach. Finally, while the list of metrics presented here is large, we consider it far from exhaustive.

Importantly, we do not intend this manuscript to indicate that skewness is somehow the most important or otherwise "best" metric to describe river corridor exchange. We intended it as illustrative of patterns that were consistent across many metrics. We have modified Figures 5 and 6 to now include multiple response variables to decrease the emphasis on skewness. This change, combined with the modified section 2.2 and newly added Table 2 should clarify this for readers.

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 (P20, L21) 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.

Accepted. Our focus here was testing Wondzell's (2011) conceptual model, not conducting a robust multivariate assessment nor exploring interactions between geologic setting and hydrologic forcing as controls. While we do find merit in this, and we indeed believe this data set is one of the first that might support this effort, it is beyond the scope of our study. That said, we have revised the manuscript to clarify that we did fit simple multivariate relationships to each response metric considered (i.e., the planar surface shown in Fig. 5 of the original study). We now describe this in the methods section and show multiple fits in the revised figures, plus include a comparison of univariate and bivariate fits in a supplemental table.

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.

> Accepted. We have clarified that Wondzell (2011) focused on differences in steady-state discharge by modifying the introduction and discussion, which is our focus in this study. This was stated in the last paragraph of the introduction: "variation in discharge as a function of drainage area during a fixed baseflow condition", but could have been more clear throughout the manuscript. Edits in response to this comment are primarily in describing Wondzell's (2011) discharge axis as "steady-state discharge" or "baseflow". We also added the following text to the introduction to differentiate steady-state differences from unsteady (i.e., dQ/dt, or hydrograph position) studies: "Notably, most classical expectations are based on differing steady discharge conditions (e.g., high vs. low baseflow), though an emerging body of field studies (detailed above), modeling studies (e.g., Malzone et al., 2016; Schmadel et al., 2016b), and conceptual models (e.g., Fig. 8 in Ward et al., 2016) are beginning to actively address exchange during unsteady discharge conditions."

Minor Comments: -
P1L43. Should be "is" not "are". Or "exchange" should be "exchanges"

> Accepted. Modified as suggested.

-
P2L4. The inclusion of the "and" between #2 and #3 underscores the interaction effects that may exist. –

> Accepted. Point taken, Dr. Cohn (no direct edit required in response to this comment)

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? –

> Accepted. We have removed the text "on annual to subannual timescales"

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. –

> Acknowledged. Site selection stratification was an attempt to meet multiple objectives of the field campaign, which are described in a high level in the related ESSD manuscript. In short, the overarching objective of the campaign itself was, indeed, to characterize the network. Thus, the desire for added samples in lower order streams where you correctly note we would expect more variation. The network-scale patterns presented in this study take the data as opportunistic, as we did not execute a separate campaign solely for this publication. However, we do note this is precisely one of the use-cases that we hoped for with the ESSD data – a community resource with sufficient sampling that it could be used to support any number of questions. No modifications to the study were made in response to this comment.

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. –

> Accepted. Nice catch! The porosity term here was a typo. We confirmed that in the data analysis the porosity was not used, and have corrected the equation accordingly.

I really appreciate the guidance on standardizing the reach length by wetted widths. I think this is an important standard operating procedure. –

> Thanks!

P9 refers to a companion manuscript. What/where is that? –

> Accepted. This refers to the paired submittal in Earth Systems Science Data. We have added the full citation to the "ESSD-D" paper in this location.

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. –

> Acknowledged. We have confirmed that this formulation is correct and consistent with Wlostowski et al. (2017) where the approach is first published.

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. –

> Accepted. We have added a supplement to the manuscript that includes both tabular and visual representations of Pearson correlation and Spearman Rank Correlation.

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. –

> Accepted. We have revisited our analysis and confirmed that the denominator of the equation in question, Q (Page 14, Line 18) was used as the total down-valley discharge, not only the surface discharge. We have updated the denominator to now read "$Q+Q_{sub,cap}$" to clarify this point.

P20L5 should be "hold" –

> Accepted. Modified as suggested.

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. –

> Accepted. We wholeheartedly agree! Indeed, the reason we included r2 in the table was to demonstrate that while we may find trends that pass a statistical test making them likely to have one direction or the other, these offer very little predictive power. We discuss this in the paragraph that preceded Table 2 (page 20 lines 15+ in the original submittal). Our confidence bolstered by this comment, we have added the following text to the conclusions to underscore this point: "Importantly, we document consistent trends with discharge that have low explanatory power (low r^2) despite being statistically significant in their direction, indictaing that we have little predictive power"

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. –

> Acknowledged. We agree with this idea, in concept, but do not believe that there exists a consistent expectation for each of these metrics. Ward and Packman (2018) document

that conflicting predictions exist for nearly any outcome of interest (exchange flux, timescale, hyporheic geometry) as a function of any geologic or hydrologic input.

Ward AS, Packman AI. (2018). Advancing our predictive understanding of river corridor exchange. WIREs Water . 2018;e1327. https://doi.org/10.1002/wat2.1327

QHEF on page 23 has the "HEF" subscripted. Elsewhere it's just "QHEF". –

Accepted. Modified to use the subscript throughout.

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. –

Accepted. We have added a supplement showing the Pearson and Spearman's Rank correlations between all pairs of site descriptors and metrics, including both tabular data and a visualization. We have also included versions of Fig. 3 and Fig. 4 that include HYPPOT and UAA on the X-axis as these are the variables used by Wondzell (2011).

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.

Thanks!

<u>Anonymous Referee #2</u>
Received and published: 18 June 2019
The work presented by Ward et al. represents an incredible amount of analysis based on an extensive dataset presented in a companion article. I was very excited to read and review this paper and hope that my comments will help improve it. The companion piece lays out data from synoptic and baseflow sampling of fluid fluxes through a variety of low order streams and this paper describes the analyses the team took to understand how exchange varies in relation to streamflow in space and time. With these analyses they seek to in/validate the model set forth by Wondzell (2011) and show that exchange decreases with increasing discharge through space, but that exchange varies in response with time in fixed stream reaches. Ward suggests a number of best practices for future large-scale sampling excursions to improve on these findings and reach a more parsimonious conclusionâ˘A˘ Tfirst, control for advective time; second, control for storage volume. Finally, they note that a multivariate approach is likely necessary to improve the systematic understanding of exchange in response to spatiotemporal variations in stream discharge. This is an important contribution to the discipline, and I will be delighted to see it in print after some revisions.

No responses necessary to the comment above, as issues are addressed in more detail below.

The introductory section argues convincingly that many parameters affect the exchange in streams – channel width, K, hydraulic gradient, etc. The authors spend a lot of time walking us through the measurement and calculation of many of these values, and some discussion of what those values mean and why they do or do not correlate with exchange. While this discussion is useful, I had trouble following all of the methods, results, and discussion. I think discussion of these parameters could be streamlined somewhat. For instance, I'm not sure that all of the panels of tables 3 and 4 belong in the body of this paperâ˘A˘ Tseveral are not discussed and could be moved to the supplement. Additionally I spent a lot of time searching through the text to remind myself how each variable was defined. I think extra care could be taken when terms are defined, but I think most readers would find a list or table of variable definitions to be especially helpful.

Accepted. We have added a table to the manuscript that summarizes the various approaches and key metrics (Table 2 in the revised study). However, we have elected to keep the metrics all in the study for sake of completeness, and because we do not believe any of them to be redundant.

In the results and discussion sections there is a brief mention that a multivariate approach is likely necessary to understand these relationships more thoroughly, but no analyses to investigate and present any such multivariate relationships. The authors return to this topic in the conclusion and argue that future studies must focus on these higher-level statistics. I would suggest the authors pursue this topic further within or at least explicitly discuss why they did not pursue this approach further.

See response to major comment #2 for the first referee.

Ultimately, the authors reach the conclusion that skewness is the most predictive statistic. I think it is important to expand and further justify this conclusionâ˘A ˘Tespecially to explore a rationale for why skewness is a good indicator. I think it is also important to better support their claim with regards to skewness.

Acknowledged. We disagree with the reviewers statement "Ultimately, the authors reach the conclusion that skewness is the most predictive statistic." We selected skewness as a representative and easily understood variable to demonstrate our point in Figs. 5-6, but do not consider it to be a singular "Best" variable. We have taken care to clarify this by adding other metrics to Fig. 5 and Fig. 6 and emphasized this in a brief discussion of how metrics were selected (first paragraph in section 2.2 in the revised manuscript).

In particular, I had trouble understanding figures 5 and 6. Figure 5 was of low image quality, so an enhanced resolution image might have helped, but I had trouble seeing where the points were plotted in 3d space, and thus could not follow their argument. I found figure 6 unconvincing. The argument rests on best fit lines that don't seem supported by

the underlying data. I would suggest replacing the figure, removing the lines, or at least presenting some statistical treatment of why they believe the best fit lines are justified.

> Accepted. Figure 5 in the original version (Fig. 6 in the revised) has been revised to improve the visualization and interpretability of the data. Figure 6 in the original study has been moved to Fig. 5 in the revised form. The figure now depicts t99, holdback, and skewness as a function of advective time. The revised figure is described in the results (newly added section 3.3). We retain the linear trend-lines as a useful interpretive tool, but provide a quantitative comparison of the ranges of parameter values between the different approaches.

A last concern is the number of authorsâ˘A˘ TI am not used to seeing such a large author list on a data analysis paper. I think it is important to justify and define the contribution of each author toward the different tenets of authorship in a systematic mannerâ˘A˘ TI think it is important that the authors make an earnest attempt to do so. One approach would be the approach suggested by Clement (2014).

> Acknowledged. We respectfully note that author contributions were described, albeit briefly, in the acknowledgements section of the manuscript. The lead author hereby confirms that each co-author contributed at a level consistent with Clement's (2014) recommendations. This is perhaps best understood by the scope of the field campaign that was required to characterize these sites, the many approaches taken to interpret the data, a collaborative writing process where all co-authors were active participants, and a team that has been working together for several years on a series of collaborative projects.

Minor/general comments follow and are ordered chronologically.
Pp:line:comment
General: The paper would benefit greatly from a table/list of all variables at the start/end/supplement. I spent a lot of time flipping through the paper trying to remember what the variables and subscripts represented.

> Accepted. The newly added Table 2 include a summary of the key response variables that are used in this study.

2:5: The "more than 60 solute tracer studies" were conducted in a companion paper, not this article, it is probably worth clarifying here and elsewhere. Careful throughout that data from the companion paper are not presented as results of this paper.

> Accepted. We have modified this sentence to now read: "To test this conceptual model we conducted more than 60 solute tracer studies including a synoptic campaign in the 5th order river network of the H.J. Andrews Experimental Forest (Oregon, USA) and replicate-in-time experiments in four watersheds.". We have elected not to include a reference to the ESSD companion paper in the abstract, but make clear reference to this data set later in the study.

3: 13-14: is it expected that exchange volume will decrease or the ratio of Qex/Q?

3: 25: is it expected that exchange volume will decrease or the ratio of Qex/Q? Please clarify here and several other places.

Acknowledged. There is not a predominant expectation of this relationship. One could argue that if Q_{HEF} is constant (for example, due to some geologic feature that controls exchange flux and does not change with discharge), then increasing discharge would decrease Q_{HEF}/Q. However, other mechanisms (e.g., diffusion of turbulent momentum across the streamebed) may vary with discharge, changing both Q_{HEF} and Q simultaneously.

6:Table 1: I suggest you change the order of table items to match order they're presented in the text.

Accepted. Table order has been modified as suggested.

7: 15-25: The presented replicate falling head tests were all conducted at one location in the stream channel. Were tests conducted to understand the spatial variability of K within the channel and floodplain sediments? K varies widely over relatively short scales, is there any way to bracket the errors associated with this?
7: 25: K is typically log-normal, should this be the log-geometric mean?

Acknowledged. The tests to estimate K were conducted at a single location at each study site. The value reported is the geometric mean, taken from the published data set detailed in the ESSD manuscript.

8: 4-7: If Qsub,cap is volumetric and based on Darcy, I don't understand why porosity is included in the calculation of the "capacity of the subsurface to convey water down to the valley bottom" as porosity should impact velocity only, and not impact volumetric flux. If porosity is estimated as 30% for all sites, this shouldn't impact findings, but clarification would be helpful.

Accepted. Porosity was included in the equation as a typo. We confirmed it was not used in the calculations, and this reviewer is correct that it should not have been there. The equation and text have been updated accordingly.

8:6-7: You say, "hvalley is the valley colluvium depth (m; estimated as 50% of the wetted channel width)". To clarify, depth of colluvium is never independently determined, it's only estimated as 1/2 wetted channel width? If so, wetted at what stage (e.g. high discharge, mean discharge)? Please provide some references to support this as a valid approach.

Accepted. We have added several references here that have estimated depth of colluvium for several sites in the study basin. We have added the following text: "This estimate is consistent with depths used in past studies (Gooseff et al., 2006; Ward et al., 2012; Crook et al., 2008; Ward et al., 2018a; 2018c; Schmadel et al., 2017) and geophysical transects in the 4th and 5th order reaches of Lookout Creek (Wondzell, unpublished data)."

8:9: Suggest changing "nor" to "or"

Accepted. Modified as suggested.

8:20-29: Please define more thoroughly the term "mixing length." Is this the length required for advective mixing to result in a homogeneous surface water concentration of a released solute? How was this determined in cases without any tracer.

We have added the following text to clearly define the term "mixing length" at its first use in the manuscript: "(i.e., the distance required for the solute tracer to be well-mixed across the channel cross-section)". We have also added the following text to describe how mixing length was estimated in the field: "Mixing lengths were based on visual estimates in the field as empirical estimates are unreliable in mountain streams (Day et al., 1977). Moreover, field experience in a study system is recognized to be potentially more useful that theoretical estimates of mixing length (Kilbatrick and Cobb, 1985). Thus, we used visual estimates that are consistent with our past studies using these techniques and tracers in H.J. Andrews Experimental Forest (Ward et al., 2012; 2013a; 2013b; 2019; Voltz et al., 2013) and practices used in other mountain stream networks (e.g., Payn et al., 2009; Covino et al., 2010)."

9:1-2: The term "conflicting research" is unclear. Do you mean that you could not complete the test because other experiments meant that you could not do your own experiment, or that the findings of other experiments convinced you that your results were invalid, or something else?

Accepted. The sentence has been modified to more clearly explain, now reading "The differing number of replicates reflects either sensor failure or omission of a replicate due to conflicting research occurring at the same sites by other researchers (i.e., our replication would have negatively impacted their independent research campaigns, so we did not proceed with our injections)."

10:4: Please clarify how MREC was determined. Is "mass recovered" the total mass recovered during the entire tracer test, the tracer test up to time t, or the mass recovered during the current time step? Also, how was a tracer test duration determinedâ˘A ˘Twas it continued until 99% recovery or something similar?

Accepted. MREC was previously defined in section 2.5. We have moved that definition up into section 2.2 where it is first used.

10:8-9: I'm confused about this equation.
CAD (left hand side) is based on CAD (right hand side), which suggests CAD is known a priori? Should the RHS be CADE?

Accepted. The right-hand side "CAD" should have been "CADE", which has now been corrected.

10:10: "associated with" is confusing. Do you mean something more like the "total solute mass" moved downstream by advection and dispersion?
10:20: same comment as above about "Associated with" Pp 10 and

Acknowledged. We have elected to retain this language as it is consistent with the original publication of these techniques (Wlostowski et al., 2017).

10:11: 't' appears in some equation but not others that I expect to see it in. For instance, in all terms of "CTS=Cobs-CAD" I would expect the concentration to be a function of time.

Accepted. Several equations were missing "(t)" in this section, all of which have been updated.

11:3: Why 99% Is there some particular justification? Were you calculating this in the field to determine the length of time that tracer tests should be run?

Accepted. This truncation is performed post-hoc to minimize the disproportionally high impact of late-time noise on summary metrics calculated for short term storage, and is consistent with many past studies of solute tracer transport. We have added the following text to clarify this: "…consistent with common practices (e.g., Mason et al., 2012; Ward et al., 2013a; 2013b; Schmadel et al., 2016) and a community tool for interpretation of solute tracers (Ward et al., 2017a)."

13:7: What is tau?

Accepted. We have added the following text to define tau: "where τ is a random variable representing the age of a parcel of water (Harman, 2015)".

13:15: What is "P"? Should this be "PQ"? I never see "P" defined. This is one of many cases where a symbology sheet would help immensely.
13:30: Again, what is "P"?

Accepted. In both cases, "P" has been replaced with "P_Q".

16:10: You never define the subscript "ds" in Cobs,ds so far as I can tell, thought you do define QDS. Please make sure all symbology is explicitly defined to remove confusion. Also, should this be "Cobs,DS" with the DS capitalized to match other usage?

Accepted. We have dropped the "ds" convention as Cobs is always used in reference to the downstream solute tracer timeseries. We have clearly defined Cobs where it is first used in the study: "where $C_{obs,DS}$ (g m^{-3}) is the observed solute tracer concentration at the downstream location in response to the upstream solute tracer injection".

Pp 17: No reference to figure 3H, 3G is out of order.

Accepted. WE have added the reference to Fig. 3H in the paragraph. We have elected not to re-write the paragraph to address the order in which subplots are discussed.

Fig 3. This symbology is difficult to interpret. I cannot distinguish symbology for the 4 streams from one another because the blues and greens are too similar, especially with the poor-resolution image of the submitted pdf. I suggest making all points translucent and making the colors of the non-synoptic samples more dissimilar. Also, I would recommend adding a curly bracket around the non-synoptic samples in the legend and labeling them as the stream-reach samples. The caption begins "for synoptic data" – please clarify caption to make it clear that the figures also include the non-synoptic data. Also clarify whether the line of "best fit" is for all data in panel or only for the synoptic data. Figure 4: Same comments as in figure 3.

Acknowledged. We have used both shape and color to distinguish the sites, and have elected to retain this redundant differentiation to help readers. The colors are selected from the "Parula" colormap in Matlab, which is designed to retain contrast in greyscale and color prints and be accessible for color-impaired vision. We have clarified the symbology by adding the following text to the figure caption: "Data from unnamed creek (triangles, Cold creek (squares), WS03 (diamonds), and WS01 (stars) show the repeated injections through baseflow recession each headwater catchment."

20:4: "Hod") "Hold"

Accepted. Modified as suggested.

20:13: you say "most previous studies" but only cite one study. Please add more citations or remove statement.

Accepted. The one study cited is a notable exception to the "most" that we were referring to. We have modified the sentence to read: "Thus, while our selection of study reach lengths was imperfect to achieve identical advective timescales, we contend that we have adequately controlled for advective time."

20:22: You spent a lot of time showing and describing univariate values, but then say a multivariate approach is necessary to make sense of this data. Did you consider including some multivariate stats to explore these relationships?

Accepted. See response to Reviewer #1's "major comment 2"

22:27: “Sens slope was larger for the fixed reaches: : :” I don’t recall if this is explicitly discussed later.

Acknowledged. This point is discussed again in the conclusions of the study.

23:10-13: Is the decreased QHEF a volumetric decrease or a relative decrease as a fraction of stream discharge?

Acknowledged. The text in question reads “We did find an increasing fraction of total discharge sampled in higher discharge locations (Fig. 4C), but the overall trend indicates that Q_{HEF} does not grow as rapidly as Q, moving downstream along the network.”. We believe this text is sufficiently clear in its current form as an interpretation of changes in QHEF, not to the quantity QHEF/Q.

Fig 5: What are the vertical columns? The colored lines? The right-hand panel is very difficult to interpret. The lefthand panel benefits from the lines that extend to z=0, to show the footprint of each point, whereas I cannot tell where points in the righthand panel exist in XY space. Is there a better way to present this data? Same comment about the color scheme as in figure 3 and 4â˘Aˇ TI cannot differentiate between the points.

Accepted. We have revised the figure to highlight only the synoptic data, added “stems” to orient the data in the bottom X-Y plane of the figure, and included colored best-fit planar surfaces to aid in visualization.

28:21-2:Was this multi-sensor approach described in the methods of this paper? I did not see any previous mention.

Accepted. We have moved the appropriate portions of the text from the locations referenced here to the methods (newly added section 2.1.4) and results (newly added section 3.3).

28:25: Where were these data/results presented/discussed? I did not see previous mention in this paper. Fig 6: I do not trust the lines on these plotsâ˘A ˇ TI believe they are misleading and suggest they be removed.

See response to the third major comment from this reviewer.

31:10: Suggest removing “likely”

Accepted. Modified as suggested.

References: Clement, Prabhakar (2014). Authorship Matrix: A Rational Approach to Quantify Individual Contributions and Responsibilities in Multi-Author Scientific Articles. Sci Eng Ethics 20. 345–361. DOI 10.1007/s11948-013-9454-3

Spatial and temporal variation in river corridor exchange across a 5th order mountain stream network

Adam S. Ward[1], Steven M. Wondzell[2], Noah M. Schmadel[1,3], Skuyler Herzog[1], Jay P. Zarnetske[4], Viktor Baranov[5,6], Phillip J. Blaen[7,8,9], Nicolai Brekenfeld[7], Rosalie Chu[10], Romain Derelle[11], Jennifer Drummond[7,12], Jan Fleckenstein[13,14], Vanessa Garayburu-Caruso[15], Emily Graham[15], David Hannah[7], Ciaran Harman[16], Jase Hixson[1], Julia L.A. Knapp[17,18], Stefan Krause[7], Marie J. Kurz[13,19], Jörg Lewandowski[20,21], Angang Li[22], Eugènia Martí[12], Melinda Miller[1], Alexander M. Milner[7], Kerry Neil[1], Luisa Orsini[11], Aaron I. Packman[22], Stephen Plont[4,23], Lupita Renteria[24], Kevin Roche[25], Todd Royer[1], Catalina Segura[26], James Stegen[15], Jason Toyoda[10], Jacqueline Wells[24], Nathan I. Wisnoski[27]

[1] O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana, USA
[2] USDA Forest Service, Pacific Northwest Research Station, Corvallis, Oregon, USA.
[3] Earth Surface Processes Division, U.S. Geological Survey, Reston, Virginia, USA
[4] Department of Earth and Environmental Sciences, Michigan State University, East Lansing, Michigan, USA
[5] LMU Munich Biocenter, Department of Biology II, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany
[6] Department of River Ecology and Conservation, Senckenberg Research Institute and Natural History Museum, 63571 Gelnhausen, Germany
[7] School of Geography, Earth & Environmental Sciences, University of Birmingham, Edgbaston. Birmingham. B15 2TT. UK
[8] Birmingham Institute of Forest Research (BIFoR), University of Birmingham, Edgbaston. Birmingham. B15 2TT. UK
[9] Yorkshire Water, Halifax Road, Bradford, BD6 2SZ
[10] Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA
[11] Environmental Genomics Group, School of Biosciences, the University of Birmingham, Birmingham B15 2TT, UK
[12] Integrative Freshwater Ecology Group, Centre for Advanced Studies of Blanes (CEAB-CSIC), Blanes, Spain
[13] Dept. of Hydrogeology, Helmholtz Center for Environmental Research - UFZ, Permoserstraße 15, 04318 Leipzig, Germany
[14] Bayreuth Center of Ecology and Environmental Research, University of Bayreuth, 95440 Bayreuth, Germany
[15] Earth and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA
[16] Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, Maryland, USA
[17] Department of Environmental Systems Science, ETH Zürich, Zurich, Switzerland
[18] Center for Applied Geoscience, University of Tübingen, Tübingen, Germany
[19] The Academy of Natural Sciences of Drexel University, Philadelphia, Pennsylvania, USA
[20] Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Department Ecohydrology, Müggelseedamm 310, 12587 Berlin, Germany
[21] Humboldt University Berlin, Geography Department, Rudower Chaussee 16, 12489 Berlin, Germany
[22] Department of Civil and Environmental Engineering, Northwestern University, Evanston, Illinois, USA
[23] Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA
[24] Pacific Northwest National Laboratory, Richland, WA, USA
[25] Department of Civil & Environmental Engineering & Earth Sciences, University of Notre Dame, Notre Dame, IN
[26] Forest Engineering, Resources, and Management, Oregon State University Corvallis, OR, USA
[27] Department of Biology, Indiana University, Bloomington, Indiana, USA

Correspondence to: Adam S. Ward (adamward@indiana.edu)

Formatted: Font: (Default) Times New Roman

Abstract. Although most field and modeling studies of river corridor exchange have been conducted a scales ranging from 10's to 100's of meters; results of these studies are used to predict their ecological and hydrological influences at the scale of river networks. Further complicating prediction, exchanges are expected to vary with hydrologic forcing and the local geomorphic setting. While we desire predictive power, we lack a complete spatiotemporal relationship relating discharge to the variation in geologic setting and hydrologic forcing that are expected across a river basin. Indeed, Wondzell's (2011) conceptual model predicts systematic variation in river corridor exchange as a function of (1) variation in baseflow over time at a fixed location, (2) variation in discharge with location in the river network, and (3) local geomorphic setting. To test this conceptual model we conducted more than 60 solute tracer studies including a synoptic campaign in the 5th order river network of the H.J. Andrews Experimental Forest (Oregon, USA) and replicate-in-time experiments in four watersheds. We interpret the data using a series of metrics describing river corridor exchange and solute transport, testing for consistent direction and magnitude of relationships relating these metrics to discharge and local geomorphic setting. We confirmed systematic decrease in river corridor exchange space through the river networks, from headwaters to the larger mainstem. However, we did not find systematic variation with changes in discharge through time, nor with local geomorphic setting. While interpretation of our results is complicated by problems with the analytical methods, they are sufficiently robust for us to conclude that space-for-time and time-for-space substitutions are not appropriate in our study system. Finally, we suggest two strategies that will improve the interpretability of tracer test results and help the hyporheic community develop robust data sets that will enable comparisons across multiple sites and/or discharge conditions.

Commented [KTV1]: Exchanges are" plural, or "Exchange is"

Commented [n2R1]: Good catch. Changed.

Deleted: [

Deleted:]

Deleted: discharge

Deleted: collected in

Commented [n4R3]: Good catch. Changed.

Deleted: are

1 Introduction

Ecological functions and processes in the river corridor are influenced by the exchange of water, solutes, and energy between the surface stream and its catchment, and thus regulate downstream water quality (e.g., Brunke and Gonser, 1997; Krause et al., 2011; Wondzell and Gooseff, 2014; Ward, 2015). These exchange fluxes are collectively termed river corridor exchange and integrate the stream, hyporheic zone, and riparian zone along the river network (Harvey and Gooseff, 2015). Several recent studies have extended feature- and reach-scale findings to predict ecological functions of river corridors at basin scales relevant to resource management (e.g., Gomez-Velez and Harvey, 2014; Kiel and Cardenas, 2014; Gomez-Velez et al., 2015; Bertuzzo et al., 2017; Helton et al., 2018). These approaches require a scaling relationship to predict river corridor exchange across space and through time. Discharge is a logical scaling factor and has been studied as a control on river corridor exchange in both space (i.e., along a network) and time (i.e., under different hydrologic conditions at a fixed location). However, discharge integrates forcing at different scales and may not lead to consistent predictions of river corridor exchange (Ward & Packman, 2018). For example, increases in discharge have been found to cause increases, decreases, or no change in river corridor exchange (Morrice et al., 1997; Butturini and Sabater, 1999; Hart et al., 1999; Jin and Ward, 2005; Wondzell, 2011, 2006; Zarnetske et al., 2007; Schmid, 2008; Karwan and Saiers, 2009; Schmid et al., 2010; Fabian et al., 2011; Ward et al., 2013a). Clearly, to use discharge as a scaling factor to predict river corridor exchange, a more complete description of the exchange-discharge relationship is required.

Deleted: [

Deleted:]

Deleted: [

Deleted:]

Deleted: [

Deleted:]

Deleted: [

Deleted:]

Deleted: [

Deleted:]

River corridor exchange is broadly understood to be controlled by interactions between hydrologic forcing and geomorphic setting (Kasahara and Wondzell, 2003; Ward et al., 2012). First, hydrologic forcing encompasses variation in the catchment wetness and storage during storms (Ward et al., 2013a; Dudley-Southern and Binley, 2015; Malzone et al., 2016), seasonal baseflow recession (Payn et al., 2009; Voltz et al., 2013; Ward et al., 2013c; Schmadel et al., 2017), and diurnal fluctuations arising from natural (e.g., Harman et al., 2016; Musial et al., 2016) or anthropogenic (e.g., Sawyer et al., 2009; Gerecht et al., 2011) controls. While hydrologic forcing reflects a variation in the temporal domain, the geomorphic setting is typically assumed static during river corridor exchange studies. Thus, repeated studies under different discharge conditions are focused on predicting river corridor exchange as a function of hydrologic forcing and used to develop exchange-discharge relationships at individual study reaches (e.g., Rana et al., 2017). This strategy yields a fixed-in-space, varied-in-time exchange-discharge relationship. Notably, most classical expectations are based on differing steady discharge conditions (e.g., high vs. low baseflow), though an emerging body of field studies (detailed above), modeling studies (e.g., Malzone et al., 2016; Schmadel et al., 2016b), and conceptual models (e.g., Fig. 8 in Ward et al., 2016) are beginning to actively address exchange during unsteady discharge conditions. It is also important to note that, in some cases, changes in discharge can also change the effective geomorphic setting. For example, increases in discharge can flood pool-riffle sequences (e.g., Storey et al., 2003; Church and Zimmerman, 2007) or activate secondary channels (e.g., Ward et al., 2016). Exchange-discharge relationships during steady flow conditions have been examined in many studies with repeated studies over time at a single site resulting in both positive and negative correlations between river corridor exchange and discharge (Ward and Packman, 2018), though one classic expectation is decreased exchange with increased discharge due to compression of hyporheic flowpaths by toward-stream hydraulic gradients (e.g., Hakenkamp et al., 1993; Hynes, 1983; Palmer, 1993; Vervier et al., 1992; White 1993).

The second primary control on river corridor exchange is the geomorphic setting, including differences attributable to tectonics (e.g., Valett et al., 1996; Payn et al., 2009). Over geologic timescales the geomorphic setting has co-evolved with hydrologic forcing. For example, as drainage area and discharge accumulate through mountain stream networks, we expect predictable spatial patterns including lower slopes, smaller grain size, larger channel width-to-depth ratios, and increased valley bottom widths (e.g., Leopold and Maddock, 1953; Wohl and Merritt, 2005; 2008; Brardinoni and Hassan, 2007). The evolution of geologic setting occurs over extremely long timescale, allowing the common simplification of assuming geologic setting as static in hyporheic studies. As a result of this assumption, researchers commonly conduct experiments across a spatial gradient to describe patterns in river corridor exchange (Payn et al., 2009; Covino et al., 2011; Mallard et al., 2014). This approach provides a fixed-in-time, varied-in-space river corridor exchange-discharge relationship that describes a network under a fixed hydrologic condition, most commonly baseflow. Wondzell (1994) suggested that exchange should decrease with increasing watershed size based on first principles. For example, the potential maximum exchange is limited by the streambed area, indicating that the ratio of wetted perimeter to discharge (Q) should be correlated to the maximum possible exchange per unit length of stream channel. As Q increases more rapidly than wetted perimeter as watersheds increase in size, the amount of exchange should be expected to decrease. In fact, most studies have identified a decreasing

Deleted: […Kasahara and Wondzell, 2003; Ward et al., 2012]…. First, hydrologic forcing encompasses variation in the catchment wetness and storage during storms […Ward et al., 2013a; Dudley-Southern and Binley, 2015; Malzone et al., 2016]…, seasonal baseflow recession […Payn et al., 2009; Voltz et al., 2013; Ward et al., 2013c; Schmadel et al., 2017]…, and diurnal fluctuations arising from natural […e.g., Harman et al., 2016; Musial et al., 2016]… or anthropogenic […e.g., Sawyer et al., 2009; Gerecht et al., 2011]… controls. While hydrologic forcing reflects a variation in the temporal domain, the geomorphic setting is typically assumed static during river corridor exchange studies. Thus, repeated studies under different discharge conditions are focused on predicting river corridor exchange as a function of hydrologic forcing and used to develop exchange-discharge relationships at individual study reaches […e.g., Rana et al., 2017 ... [1]

Formatted: Not Highlight

Deleted: However, …i… is also important to note that, in some cases, changes in discharge can also change the effective geomorphic setting. For example, increases in discharge can flood pool-riffle sequences […e.g., Storey et al., 2003; Church and Zimmerman, 2007]… or activate secondary channels […e.g., Ward et al., 2016]…. These …e…change-discharge relationships during steady flow conditions have been examined in many studies with repeated studies over time at a single site resulting in both positive and negative correlations between river corridor exchange and discharge […see summary in …ard and Packman, 2018]…, though the …ne classic predominant …xpectation is decreased exchange with increased discharge due to compression of hyporheic flowpaths by toward-stream hydraulic gradients […e.g., Hakenkamp et al., 1993; Hynes, 1983; Palmer, 1993; Vervier et al., 1992; White 1993 ... [2]

Deleted: […e.g., Valett et al., 1996; Payn et al., 2009]…. Over geologic timescales the geomorphic setting has co-evolved with hydrologic forcing. For example, as drainage area and discharge accumulate through mountain stream networks, we expect predictable spatial patterns including lower slopes, smaller grain size, larger channel width-to-depth ratios, and increased valley bottom widths […e.g., Leopold and Maddock, 1953; Wohl and Merritt, 2005; 2008; Brardinoni and Hassan, 2007]…. The evolution of geologic setting occurs over extremely long timescale, allowing the common simplification of assuming geologic setting as static in hyporheic studies. As a result of this assumption, researchers commonly conduct experiments across a spatial gradient to describe patterns in river corridor exchange […Payn et al., 2009; Covino et al., 2011; Mallard et al., 2014]…. This approach provides a fixed-in-time, varied-in-space river corridor exchange-discharge relationship that describes a network under a fixed hydrologic condition, most commonly baseflow. Wondzell […1994] ... [3]

Commented [n6R5]: Thanks for catching this confusion. This text has been edited for clarity following your suggestion.

Deleted: Consider that

Deleted: then

Deleted: (P)

Formatted ... [4]

Deleted: P

role of river corridor exchange as river basins increase in size, attributable to less exchange flux relative to stream flow (Stewart et al., 2011; Mallard et al., 2014; Gomez-Velez and Harvey, 2014; Kiel and Cardenas, 2014; Gomez-Velez et al., 2015; Ward et al., 2018a).

To explain spatiotemporal patterns in river corridor exchange from headwaters to large rivers, Wondzell (2011) developed a conceptual framework describing the relative importance of river corridor exchange to reach-scale transport (i.e., hyporheic exchange flow normalized by river discharge, Q_{HEF}/Q), spanning three primary dimensions. First, Q_{HEF}/Q would be largest under the lowest steady-state discharge conditions, where subsurface flow may reflect a larger proportion of total down-valley flow. Second, Q_{HEF}/Q would be largest in the headwaters and decrease moving toward larger river segments as described above. Lastly, Wondzell (2011) characterized the local geomorphic setting at an individual study site as "hyporheic potential," combining valley slope and hydraulic conductivity to reflect local controls on exchange at the reach scale that might vary locally within the systematic spatial and temporal dimensions. Larger hyporheic potential was associated with larger Q_{HEF}/Q. Subsequently, Harvey et al. (2018) suggested that hydrologic connectivity (i.e., Q_{HEF}/Q) is a primary water quality regulator. Ward et al. (2018) further extended this concept to account for changes in valley bottom width and depth of colluvium, describing the down-valley capacity of the valley bottom to transmit water estimated via Darcy's Law. Unlike the first two dimensions, hyporheic potential may not have a predictable trend as one moves down a river continuum, because decreasing slopes and hydraulic conductivities may be offset by larger hyporheic cross-sections.

Efforts to predict river corridor exchange and associated ecosystem processes as a function of geomorphic setting and hydrologic forcing have been implemented in large-scale remotely sensed test cases. However, this method still lacks field validation across varying discharge and across a range of stream types with varying morphologic features. For example, Gomez-Velez and Harvey (2014) and Gomez-Velez et al., (2015) used the Networks with Exchange and Subsurface Storage (NEXSS) model to describe spatial patterns in exchange in low-gradient alluvial river networks. NEXSS is based on steady-state discharge and bed sediment grain size as a proxy for local morphologic control. While this modeling approach has demonstrated the importance of river corridor exchange in large river basins, it is built on scaling relationships derived from idealized mechanistic and conceptual models that may not be representative of headwater streams. Further, the model results have yet to be confirmed in field trials.

To our knowledge, only Payn et al's (2009) field study explicitly considered both spatial and temporal dimensions of the exchange-discharge relationship. The results of that study were broadly consistent with Wondzell's (2011) conceptual model. However, we now understand that fixed reach lengths cause systematic decreases in the "window-of-Detection" (the timescale of exchange flowpaths that are measurable with tracer studies (Harvey et al., 1996; Wagner and Harvey, 1997; Harvey and Wagner, 2000)). The systematic decrease of window of detection with increasing discharge along the study stream would have interacted with the fixed reach lengths, likely leading to the underestimation of Q_{HEF} at high discharges. As a result, it is difficult to separate the observed process from limitations of the

Deleted: [
Deleted:]
Deleted: [
Deleted:]
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Deleted: [
Deleted:]
Formatted: Font: Italic
Formatted: Font: Italic, Subscript
Formatted: Font: Italic
Deleted: [
Deleted:]
Deleted: QHEF/Q
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Formatted: Font: Italic
Formatted: Font: Italic, Subscript

measurement instrument (see discussion of Payn et al.'s data in Ward et al., 2013b, and similar studies by Schmadel et al., 2016a).

Several other studies have found general agreement with Wondzell's (2011) prediction of decreasing Q_{HEF}/Q with increasing baseflow through space and at individual study reaches (Kelleher et al., 2013; Patil et al., 2013; Ward et al., 2013c). Thus, Wondzell's (2011) conceptual model might provide an organized framework to extend reach-scale results across space and time in mountain river basins. However, the studies cited above were limited to headwater networks, whereas Wondzell (2011) suggested patterns should hold across much larger scales and geomorphic settings. To date, Wondzell's (2011) conceptual model lacks validation across large river basins studied with a systematic field approach. Given the variability of reach-scale river corridor exchange trends documented in the literature (see summary in Ward and Packman, 2018), it is critical to test Wondzell's (2011) conceptual model with field data that cover much more of the space-time parameter space.

In this study, we seek to characterize river corridor exchange in a mountain stream network as a function of (1) variation in baseflow at a fixed location through seasonal recession; (2) variation in discharge as a function of drainage area during a fixed baseflow condition; and (3) local geomorphic setting (quantified here as hyporheic potential). This study will directly test the conceptual model posed by Wondzell (2011) for mountain stream networks. If the conceptual relationships can be confirmed, this would enable transferability of findings from feature- and reach-scale studies to entire networks of high-gradient mountain streams, paralleling recent advances in low-gradient river networks (e.g., Gomez-Velez and Harvey, 2014; Kiel and Cardenas, 2014; Gomez-Velez et al., 2015). Further, confirmation of the conceptual model would provide a simple scaling relationship for time-variable discharge, which has not been possible to-date. In this study, we conducted a series of solute tracer studies to construct temporal exchange-discharge relationships (i.e., a fixed study reach with observations spanning a range in discharge) and spatial exchange-discharge relationships (i.e., a synoptic campaign to measure exchange at many locations under summer baseflow discharge) for a fifth-order mountain river network, together with physical observations (including hydraulic conductivity, drainage area, slope, valley bottom width, sinuosity) to also characterize hyporheic potential. We interpret the data using a series of metrics describing river corridor exchange and their relationships to discharge.

2 Methods

2.1 Field site and Solute Tracer Experiments

2.1.1 Site Description

The H.J. Andrews Experimental Forest (HJA) is a 5th order basin draining about 6,400 ha in the Western Cascade Mountains, Oregon, U.S.A. with elevations ranging from about 410 to 1,630 m a.m.s.l. The basin is heavily forested and includes stands of old growth Douglas fir trees as well as smaller areas that have been logged to study the effects of forest management practices. Additional

Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: QHEF/Q
Deleted: discharge
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Commented [KTV7]: Delete "the"
Commented [n8R7]: corrected
Deleted: the
Deleted: discharge
Deleted: in
Deleted: [
Deleted:]
Deleted: ¶

detail about the climate, morphology, geology, and ecology of the site are well described by others (Dyrness, 1969; Swanson and James, 1975; Swanson and Jones, 2002; Jefferson et al., 2004; Cashman et al., 2009; Deligne et al., 2017). The synoptic sampling spanned the entire HJA basin to characterize basin scale valley bottom conditions, while additional more detailed sampling occurred in three distinct landform types.

Headwater sites in the HJA generally fall into one of three landform types associated with underlying geology and geomorphic processes (Table 1). We selected four 2nd order basins to establish fixed stream reaches for replication through the summer baseflow recession period, one in each landform type plus one replicate. The first landform type occurs in the lower elevations of the HJA where geology is dominated by Upper Oligocene - Lower Miocene basaltic flows. These volcanoclastic rocks were weakened by hydro-thermal alteration from subsequent volcanic activity, enabling rapid downcutting and formation of a highly dissected landscape. Hillslopes are steep; valleys are v-shaped and tend to be narrow with steep longitudinal gradients. Valley bottom colluvium is typically shallow but variable, being emplaced by hillslope mass wasting and debris flows. Exposed bedrock is visible in many locations, while deeper deposits form behind individual large logs or larger log jams. We selected the well-studied Watersheds 1 and 3 (WS01 and WS03) for two of our fixed reaches (Figure 1). Briefly, WS01 and WS03 valley bottoms reflect different time periods in this landform. In 1996, WS03 was scoured to bedrock along 100s of meters of the valley bottom (Johnson, 2004). Since that time no debris flows have been recorded, resulting in a study reach nearly free of colluvium in the upper half of the study reach. WS01 is a paired catchment to WS03, reasonably representing a pre-scour and less-constrained comparison to WS03. WS01 has a wood-forced step-pool morphology (Montgomery and Buffington, 1997; 1998) over most of its mainstem length, representative of many steep mountain streams. River corridor exchange in the two catchments have been broadly studied using a paired catchment approach (e.g., Wondzell, 2006; Voltz et al., 2013; Ward et al., 2017b).

Deep-seated earth flows provide a second contrasting landform type in the HJA. These are emplaced on the Upper Oligocene - Lower Miocene basaltic flows and are characterized by a poorly developed channel network (many parallel channels), a general lack of lateral contributing area to the river corridor, little lateral constraint, and extensive colluvial deposits with no bedrock exposure. Based on visual inspection, channels on these earthflows are actively meandering, braiding, and downcutting. Characteristic geomorphic features include meander bends and cut-banks (visually similar to lower-gradient alluvial systems of the region) in addition to step-pool features. We selected an unnamed 2nd order reach on a large earth flow adjacent to WS03 for this study (Figure 1).

The third landform type occurs in high elevation headwater catchments with U-shaped valleys characteristic of glacial cirques, which formed in plieocascade volcanics. Valley bottoms are filled with compacted glacial tills. Large wood atop the till forms pools and steps with intermediate gravel and cobble riffles. Lateral tributary area is relatively uniform along the valley with few hollows or tributary valleys (in contrast to the highly dissected landforms in WS01 and WS03). Bedrock is rarely visible along the study site. We selected a 2nd order reach of Cold Creek to represent this landform (Figure 1).

Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: on annual to subannual timescales

Table 1. Summary of key characteristics of the fixed-reach sites. See detailed descriptions for further information (*Dyrness*, 1969; *Swanson and James*, 1975; *Swanson and Jones*, 2002; *Jefferson et al.*, 2004; *Cashman et al.*, 2009; *Deligne et al.*, 2017)

Site	Important Hydrologic Controls	Important Geologic Controls
WS01	• Highly-dissected landscape • Focused lateral inflows • Diurnal discharge fluctuations due to evapotranspiration	• Colluvium deposited by debris flows from hillslopes forms extensive deposits in the valley bottom • V-shaped, rapidly downcutting valley
WS03	• Highly-dissected landscape • Focused lateral inflows • Diurnal discharge fluctuations due to evapotranspiration	• Scoured to bedrock in 1996 leaving only small colluvial deposits • Highly constrained, low colluvium analogue to WS01 • V-shaped, rapidly downcutting valley
Unnamed Cr.	• Surficial aquifer on earthflow connects several parallel channels • Minimal lateral tributary area	• Deep-seated earthflow • No defined valley; parallel stream channels down hillslope
Cold Cr.	• Extensive aquifer provides high discharge, cold baseflow year-round • Diffuse lateral inflows	• Compressed glacial tills • U-shaped valley (glacial cirque) • Uniform lateral tributary area

2.1.2 Synoptic study

We conducted a synoptic study at 46 sites within the HJA during late summer baseflow conditions (Figure 1) that included solute tracer experiments. Site selection was stratified by by stream order so that more headwater sites were sampled than higher order reaches, as suggested by other synoptic investigations of sediment-water interfaces at the basin scale (Ruhala et al., 2017; Lee-Cullin et al., 2018). We selected low baseflow conditions to maximize our ability to measure Q_{HEF}/Q, which is expected to be largest under low discharge conditions (Wondzell, 2011). Study sites were selected to achieve coverage across stream orders, landforms, and on the basis of accessibility from roads in the basin. The data described here are documented and field methods described in detail by Ward et al. (2019), but we provide an overview below.

At each site we measured mean stream width and depth, valley width, and collected GPS coordinates. Subsequently, a modified version of TopoToolbox 2.0 (Schwanghart and Kuhn, 2010; Schwanghart and Scherler, 2014) and a 1-m LiDAR derived, digital elevation model (DEM) was used to extract upslope accumulated area (UAA; ha), valley slope (S_{val}; m m^{-1}), and a stream centerline that was used to calculate sinuosity (Sinuosity; m/m) . Our methods were identical to those previously used in the basin (Corson-Rikert et al., 2016; Schmadel et al., 2017; Ward et al., 2018c, 2018a).

At each synoptic site, we drove a Solinst 615N well point into the streambed so that the top of the 0.15 m screened interval was 50-cm below the streambed. After developing the well with a peristaltic pump, we conducted 3 to 6 replicate falling head tests, measuring head change through time using a down-well

Commented [KTV9]: In the text these sites are introduced from low elevation to high elevation (WS01, WS03, Unnamed Creek, Cold Creek), which is different from the order in this table. I suggest reordering this table to match the order in the text unless there are gradients in the controls that are important to represent in this table.

Commented [n10R9]: Thank you for the suggestion and detailed focus of the text. However, we feel that listing them as is in the table is more appropriate for a stand-alone table because WS01 and WS03 appear in many studies in the literature and will connect with broader readership, whereas Cold Cr. And Unnamed Creek are essential new watersheds in the present study. As described in the text, we feel is makes more sense to discuss these watersheds in order of gradients.

Deleted: Unnamed Cr. ... [5]

Formatted: Font: (Default) Times New Roman

Deleted: done to stratify

Deleted: sampling

Commented [KTV11]: The formatting of in text citations switches to *italics* here, when they are not italicised above. Suggest adjusting for consistency.

Commented [n12R11]: Thank you for catching this inconsistency in formatting. We well correct all referencing for consistency prior to submitting a revised version to the journal.

Deleted: [

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted:]

Deleted: [

Deleted:]

Formatted: Font: Not Italic

Deleted: [

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted:]

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted: [

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted:]

Van Essen Micro-Diver logging at a 0.5-s interval. Falling head tests were interpreted using the Hvorslev (1951) method:

$$K = \frac{r^2 \ln\left(\frac{L_e}{R}\right)}{2 L_e T_0}$$

where K is hydraulic conductivity (m/s), r is the radius of the well casing (0.025 m), R is the radius of the well screen (0.005 m), L_e is the screened length of the well (m), and T_0 is the time for the head to fall to about 37% of its original value (i.e., the e-folding time; s). We took the geometric mean of the replicate tests as the representative value of K at each site.

We calculated the capacity of the subsurface to convey water down the valley bottom ($Q_{sub,cap}$, sometimes termed "underflow"; m^3 s^{-1}) as:

$$Q_{sub,cap} = b_{valley} h_{valley} K S_{val}$$

following *Ward et al.* (2018a), where b_{valley} is the valley width, h_{valley} is the valley colluvium depth (m; estimated as 50% of the wetted channel width. This estimate is consistent with depths used in past studies (Gooseff et al., 2006; Ward et al., 2012; Crook et al., 2008; Ward et al., 2018a; 2018c; Schmadel et al., 2017) and geophysical transects in the 4th and 5th order reaches of Lookout Creek (Wondzell, unpublished data). We calculated hyporheic potential (HYP_{POT}; m s^{-1}) after Wondzell (2011), a similar metric that does not account for valley width, depth, or porosity, as:

$$HYP_{POT} = S_{val} K$$

We also calculated stream power (Ω; W m^{-2}) at each tracer release location as:

$$\Omega = \rho g Q S$$

where ρ is the density of water (kg m^{-3}), g is the gravitational constant (9.81 m s^{-2}), Q is the average discharge in the study segment (m^3 s^{-1}), and S is the DEM-derived slope along the stream channel in the study segment (m m^{-1}).

Finally, at each site, we established a stream-tracer study reach with length approximately 20 times the wetted channel width that would be representative of reach-scale morphologic variation (MacDonald et al., 1991; Montgomery and Buffington, 1997; Rot et al., 2000; Martin, 2001; Anderson et al., 2005). We instantaneously released a known mass of NaCl (assumed conservative), dissolved in stream water, one mixing length (i.e., the distance required for the solute tracer to be well-mixed across the channel cross-section) from the downstream end of the study reach, where we monitored in-stream specific conductance (Onset Computer Corporation, Bourne, MA, USA). Mixing lengths were based on visual estimates in the field as empirical estimates are unreliable in mountain streams (Day et al., 1977).

Deleted: [

Deleted:]

Formatted: Font: Not Italic

Commented [n13]: Adam – please confirm dimensions for r and R

Commented [KTV14]: What are the values for r and R in this study?

Deleted: $\frac{b_{valley} h_{valley} K}{\theta}$

Deleted: [

Deleted:]

Formatted: No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Deleted: e

Formatted: Font: Times New Roman

Formatted: Superscript

Formatted: Superscript

Deleted: , and θ is porosity, estimated as 30% for this site [after Domenico and Schwartz, 1990; Ward et al., 2018a]

Formatted: Font: Not Italic

Deleted: .

Formatted: Font: Not Italic

Deleted: [

Deleted:]

Formatted: Font: Not Italic

Deleted: n

Formatted: Font: Do not check spelling or grammar

Deleted: [

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted:]

Commented [KTV15]: Were the injections near instantaneous? Were the interjections distributed across the stream width? Did the authors consider/mitigate potential "plunging" effects of the salt slug?

Commented [n16R15]: Yes, injections were instantaneous. Injection were not distributed across the channel. We selected appropriate mixing lengths to ensure the tracer was well mixed both laterally and vertically at the measurement sites.

... [6]

Moreover, field experience in a study system is recognized to be potentially more useful that theoretical estimates of mixing length (Kilbatrick and Cobb, 1985). Thus, we used visual estimates that are consistent with our past studies using these techniques and tracers in H.J. Andrews Experimental Forest (Ward et al., 2012; 2013a; 2013b; 2019; Voltz et al., 2013) and practices used in other mountain stream networks (e.g., Payn et al., 2009; Covino et al., 2010). Next, we released a second known mass of NaCl one mixing length above the upstream end of the study reach. We monitored in-stream specific conductance at both the up- and downstream ends of the study reach. Mixing lengths were visually estimated in the field; small amounts of a fluorescent dye were used to assess mixing lengths where they could not be readily determined by surface hydraulic conditions. All in-stream specific conductance measurements were converted to concentrations of NaCl mass added using a 4-point calibration curve developed from standards made by mixing varying amounts of NaCl with stream water that encompassed the range of observations during the tracer tests. Results from all sensors were composited into a single linear regression ($r^2 > 0.99$).

2.1.3 Fixed-reach studies

We established 11 fixed reaches of about 50-m of valley length in the four headwater catchments. We conducted identical site characterizations as described above for the synoptic study. However, for each study reach, solute tracer injections were conducted 2-6 times through baseflow recession. The differing number of replicates reflects either sensor failure or omission of a replicate due to conflicting research occurring at the same sites by other researchers (i.e., our replication would have negatively impacted their independent research campaigns, so we did not proceed with our injections). These sites parallel the common approach of replication of a study at a fixed reach with varied discharge to relate river corridor exchange to discharge conditions (after Payn et al., 2009).

Commented [KTV17]: What equipment was used to monitor in-stream specific conductance? What is the expected level of accuracy with these sensors? Were multiple sensors deployed across the cross-section to estimate if completely mixed conditions were met?

Commented [n18R17]: This analysis paper was submitted as a companion data collection paper that contains more details, like you point out, regarding the data collection. However, we agree and have repeated some of the pertinent information here.

Commented [KTV19]: Do the authors have a sense if the sodium ions are as conservative in the study reaches as they are in the stream water matrix used to generate the standard curves?

My thought process here is that the tracer is exposed to a wider range of environmental conditions in the stream, including contact with substrate, than the tracer used to develop the standard curves.

As an end member, consider a case where all the sodium is removed from the stream via biogeochemical processes. In this case, the specific conductance signal produced from the chloride that was not lost would be lower than if the sodium were still present, and it would appear that a portion of the tracer (and therefore water) was lost in the study reach. This is clearly not the case, as evidenced by the chloride remaining in the water column, but it could result in an overestimation of M_{rec} and the dependant equations.

Determining sodium loss is beyond the scope of this paper, and I suggest that this issue could be addressed with a discussion of the assumption of the conservative nature of both sodium and chloride across the study reaches.

Commented [n20R19]: Thank you for the comment. Yes, we assume that NaCl is conservative and added text to point that assumption out to the reader. Given the short travel times we do not suspect biogeochemical losses of Na. Also, to prevent introducing a bias in the observations due to assuming Na is also conservative, we build calibration curves using stream water.

Commented [KTV21]: Did these varying amounts cover the full range of specific conductance measured in the breakthrough curves?

Commented [n22R21]: Thanks for the important question. Yes, calibration curves were cover the full range of specific conductance measured in the field. There were no specific conductance breakthrough curves with values larger than that high end of the calibration curve.

While it is important to make sure measurements are within the range of the constructed calibration curve, we do not think adding this detail here is necessary as this method of using salt tracers is well established.

Deleted: ¶

Deleted: The differing number of replicates reflects either sensor failure or omission of a replicate due to conflicting research occurring at the same sites by others.

Deleted: [

Formatted: Font: Not Italic

Deleted:]

Formatted: Font: (Default) Times New Roman

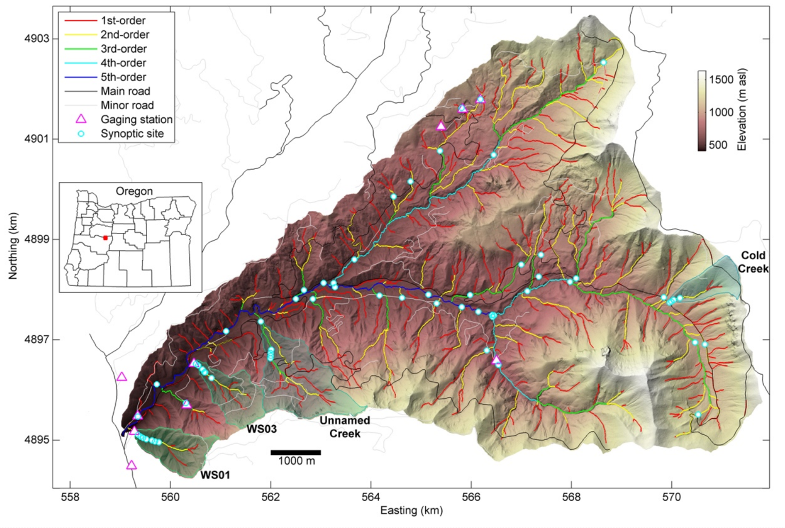


Fig. 1. Synoptic study sites and LiDAR-derived stream network for the H.J. Andrews Experimental Forest. Reprinted with permission from *Ward et al.* (2019).

2.1.4 Reach length and study design

In the synoptic campaign, we scaled our tracer reach lengths by wetted channel width in an effort to control for the advective timescales of the study. To demonstrate how this decision, or conversely the decision to fix our study reach in headwaters, may have biased our data collected, we conducted a series of four tracer injections in first through fourth stream orders in the study basin. For each study we fixed a single location for the injection and placed sensors downstream at three distances: (1) a fixed reach of 150 m; (2) an estimated 10-min of advective time downstream, based on timing debris floating along approximately 5-m of stream; and (3) a distance of 20 times the wetted channel width, which was identified as a length scale for a representative study reach in the HJA (Anderson et al., 2005; Gooseff et al., 2006). All injection protocols were consistent with synoptic and replicate injections described above.

2.2 Analysis of stream solute tracer injections

There exists no single, widely agreed upon, robust framework for describing river corridor exchange based based on stream solute tracer experiments. Instead, a host of approaches have been successfully used to interpret experimental data. In this section we detail the interpretation of stream solute tracers

Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted: In Review
Deleted: – companion manuscript]
Formatted: Font: (Default) Times New Roman
Formatted: Heading 3
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman

using several established approaches. Notably, the interpretations here were selected because they most directly interpret the observed solute tracer timeseries, in contrast to other strategies that focus on inverse model parameterization (e.g., Bencala and Walters, 1983; Haggerty and Reeves, 2002) and may be prone to parameter uncertainty and identifiability challenges (e.g., Ward et al., 2017a; Kelleher et al., 2013; Rana et al., 2019; Rana et al., 2019). The suite of approaches implemented here were selected because the provide complimentary interpretations that may be informative when jointly considered (Table 2). We emphasize here that we do not seek a singular, “best” metric to describe river corridor exchange, but instead seek to interpret a suite of metrics to provide a comprehensive understanding of our study system.

Table 2. Summary of solute tracer interpretation strategies

Approach	Strengths	Limitations	Key Metrics	Method Documentation & Example Studies
Separation of advection-dispersion from transient storage	• Quantifies the relative importance of 1-D advection-dispersion in the stream channel in comparison to transient storage	• Only considers recovered mass; no interpretation of lost mass. • Enforces separation of in-stream advection-dispersion from other processes. • Relies upon idealized model of advection-dispersion.	• Fraction of recovered mass primarily involved in advection-dispersion (f_{MAD}) • Fraction of recovered mass primarily involved in transient storage (f_{MTS})	Wlostowski et al. (2017); Ward et al. (2018b); Harms et al. (2019)
Short-term storage analysis	• Summary metrics for the time-integrated, recovered mass. • Relates timescales of storage to length-scales of subsurface flowpaths	• Only considers recovered mass; no interpretation of lost mass. • Assumes steady discharge (as formulated in this study, but can be modified) • Sensitive to late-time noise and errors in mass recovery	• Median arrival time (M_1) • Coefficient of variation (CV) • Skewness (γ) • Holdback (H) • Longest detectable subsurface flowpath (L_{detect}) • Time at which 99% of recovered mass is recovered (t_{99}) • Transient storage index (TSI, not presented in this study. $TSI = t_{99} - t_{peak}$)	Gupta and Cvetkovic (2000); Gooseff et al. (2007); Ward et al. (2010b); Schmadel et al. (2016)
StorAge Selection (SAS) analysis	• Relates outflow timeseries to age of water. • Does not require mechanistic (e.g., advection-dispersion) formulation. • No artificial division into stream and storage domains • Estimates volume of storage sampled by tracer	• Only considers recovered mass; no interpretation of lost mass. • Metrics do not map directly to classical transport mechanisms	• Storage volume sampled by tracer ($S_{comp}(T)$) • Discharge sampled by tracer ($Q_{comp}(T)$) • Fraction of storage volume sampled ($f_{VTOT}(T)$) • Fraction of discharge sampled ($fQ_{labelled}(T)$) • Fraction of stream volume sampled (f_{VSTR})	Harman (2015); Harman et al. (2016)
Long-term storage analysis	• Characterizes the fate of mass beyond the window of detection	• No interpretation of recovered tracer mass • Bounds a plausible range of gross gains and losses	• Maximum gross losses of stream water from the study reach ($Q_{LOSS,MAX}$) • Maximum gross gains of stream water to the study reach ($Q_{GAIN,MAX}$) • Maximum gross losses of stream water from the study reach ($Q_{LOSS,MIN}$, not presented in this study) • Maximum gross gains of stream water to the study reach ($Q_{GAIN,MIN}$, not presented in this study)	Payn et al. (2009); Covino et al. (2011); Payn et al. (2012); Mallard et al. (2014)

Deleted: Table X. Summary of solute tracer interpretation strategies
Approach ... [7]

Formatted: Left: 0.39", Right: 0.93", Top: 0.65", Bottom: 0.65", Width: 9.45", Height: 8.27", Not Different first page header

2.2.1 Separation of advection-dispersion from transient storage

We separated the recovered solute tracer mass into fractions that were primarily related to advection-dispersion and to short-term transient storage (after Wlostowski et al., 2017). Briefly, stream velocity (v; m s^{-1}) is estimated as $v = L/t_{peak}$, where L is the length of the reach along the centerline, and t_{peak} is the time at which the peak breakthrough curve concentration is observed, interpreted as the advective timescale of the study reach. The stream cross-sectional area (A; m^2) is estimated by $A = Q_{DS}/v$, where Q_{DS} is an estimate of discharge at the downstream end of the study reach based on dilution gauging. The mass of solute tracer recovered from the upstream injection at the downstream end of the study reach (M_{REC}; g) is calculated as:

$$M_{REC} = Q_{DS} \int_{0}^{t_{99}} C_{obs}(t)dt$$

where C_{obs} (g m^{-3}) is the observed solute tracer concentration at the downstream location in response to the upstream solute tracer injection. Using these estimates, the analytical solution to the advection-dispersion equation given the instantaneous tracer addition method is:

$$C_{ADE}(t) = \frac{M_{REC}}{A(4\pi Dt)^{1/2}} exp\left[\frac{(L - vt)^2}{4Dt}\right]$$

where C_{ADE} (g m^{-3}) is the concentration time-series predicted for the recovered mass transported via advection and dispersion only, M_{REC} is mass recovered (g), and D is the best-fit longitudinal dispersion coefficient (m^2 s^{-1}). Following this approach, the concentration time-series for a solute that is predominantly transported by advection and dispersion (C_{AD}) can be estimated as:

$$C_{AD}(t) = \begin{cases} C_{obs}(t); \; C_{ADE}(t) > C_{obs}(t) \\ C_{ADE}(t); \; C_{ADE}(t) < C_{obs}(t) \end{cases}$$

The total mass associated with advection and dispersion (M_{AD}) can be calculated as:

$$M_{AD} = \int_{0}^{t_{99}} C_{AD}(t) Q_{DS} dt$$

where t_{99} (s) is the time at which 99% of the recovered tracer signal has passed by the monitoring location. The component of C_{obs} that is primarily impacted by transient storage (C_{TS}, g m^{-3}) can be calculated as:

$$C_{TS} = C_{obs} - C_{AD}$$

Formatted: Heading 3

Formatted: Not Different first page header

Commented [KTV23]: Limiting the analysis to the mass recovered in the following equations applies a definition to transient storage as being within the window of detection, and excludes transient storage processes that are longer than the window of detection of these sites and flow conditions. I suggest clarifying here that the transient storage fraction represents storage that falls within the window of detection of these methods, and observations.

Commented [n24R23]: Yes, we absolutely agree. Any storage flowpath that is longer in time than the window of detection is not measured with this tracer method. This limitation has led to us ... [8]

Deleted: [

Deleted:]

Commented [KTV25]: Variables were *italicized* in previous ... [9]

Commented [n26R25]: Thanks again for your detailed ... [10]

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Deleted: , described below

Moved (insertion) [1]

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Formatted: Superscript

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Commented [KTV27]: Should this be C_{ADE}? ... [11]

Commented [n28R27]: You are correct. This has been ... [12]

Formatted: Font: Cambria Math

Deleted: where C_{obs} (g m^{-3}) is the observed in-stream ... [13]

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Similar to M_{AD}, the mass associated with transient storage (M_{TS}, g) can be calculated as:

$$M_{TS} = \int_{0}^{t_{99}} C_{TS}(t) Q_{DS} dt$$

Finally, we calculate the fraction of recovered mass primarily involved in advection-dispersion (f_{MAD}) or transient storage (f_{MTS}) as:

$$f_{MAD} = \frac{M_{AD}}{M_{REC}}$$

and

$$f_{MTS} = \frac{M_{TS}}{M_{REC}}$$

2.2.2 Short-term storage analysis

Observations of stream solute tracer releases were analyzed using a host of time-series metrics. We calculated the time at which 99% of the total mass recovery was achieved (t_{99}; s). To minimize the impacts of late-time noise on calculated metrics, C_{obs} was truncated at the downstream end to only include times bounded by the injection time and t_{99} (hereafter $C_{obs}(t)$), consistent with common practices (e.g., Mason et al., 2012; Ward et al., 2013a; 2013b; Schmadel et al., 2016a) and a community tool for interpretation of solute tracers (Ward et al., 2017a). The truncated time-series was normalized to isolate the features of the data in the temporal domain and minimize effects of different concentration magnitudes between injections. The normalized breakthrough curve ($c(t)$) was calculated as:

$$c(t) = \frac{C_{obs}(t)}{\int_{t=0}^{t_{99}} C_{obs}(t) dt}$$

We calculated the median arrival time (M_1; equivalent to the first temporal moment; s) as:

$$M_1 = \int_{t=0}^{t_{99}} t c(t) dt$$

Next, we calculated the 2nd and 3rd order moments about M_1 (μ_2 and μ_3) as:

$$\mu_n = \int_{t=0}^{t_{99}} (t - M_1)^n c(t) dt$$

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Heading 3

Deleted: 3

Deleted: t

where n represents the n^{th} order moment, and μ_2 and μ_3 contain information about symmetrical and asymmetrical spreading of the time-series, respectively. The central moments were normalized to provide information that could be compared between sites and injections by calculating the coefficient of variation (CV) and skewness (γ) as:

$$CV = \frac{\mu_2^{1/2}}{M_1}$$

$$\gamma = \frac{\mu_3}{\mu_2^{3/2}}$$

Finally, we calculated the holdback of the system (H), which describes transport in a continuum ranging from piston flow ($H = 0$) to no movement of the solute ($H = 1$) (Danckwerts, 1953). Ward et al. (2018b) interpret higher values of H to indicate greater influence of transient storage on reach-scale transport. Holdback is calculated as:

$$H = \frac{1}{M_1}\int_{t=0}^{M_1} F(t)\,dt$$

where

$$F(t) = \int_{\tau=0}^{t} c(\tau)\,d\tau$$

Finally, we estimated the maximum detectable flowpath length (L_{detect}) as:

$$L_{detect} = t_{99}\frac{K}{\theta}S_{val}$$

which is based on Darcy's Law, but uses the valley slope (S_{val}) as an estimate of the hydraulic gradient (after Wondzell, 2011; Ward et al., 2017b).

2.2.3 StorAge Selection (SAS) analysis

We interpreted the transport of tracer through the study reach using the StorAge Selection (SAS) approach (Harman, 2015; Harman et al., 2016). Briefly, this approach can be used to describe the composition of outflowing water from a study reach as a combination of water sampled from different

Deleted: [
Deleted: [
Formatted: Font: Not Italic
Deleted:]
Formatted: Font: Not Italic
Deleted:]
Commented [KTV29]: This is a new subscript, and should be defined.
Commented [n30R29]: Good catch. Thanks. M1 above is mean of the normalized breakthrough and this subscript is a mistake here.
Commented [KTV31]: Previous integrals did not include the variable name in the lower integration bound. Suggest editing for consistency.
Commented [n32R31]: Agreed. Edited throughout for consistency.
Deleted: [
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Deleted:]
Formatted: Font: (Default) Times New Roman
Formatted: Heading 3
Deleted: 4
Deleted: [
Formatted: Font: Not Italic
Formatted: Font: Not Italic
Deleted:]

ages within the study reach. The approach is closely related to transit time distributions, but isolates the contribution to the transit time of storage turnover from that of inflow and outflow variability. Although physically-based, in the sense of conforming to conservation of mass and describing physically meaningful properties, this approach describes the higher-level emergent effects of mechanisms like advection, dispersion and other processes (Harman et al., 2016). Instead, the approach provides a description of the reach as a zero-dimensional, integrated control volume (i.e., no arbitrary division of surface vs. subsurface or mobile vs. less mobile storage).

Here, we closely follow the adaptation of the general formulation of the SAS framework to interpret stream solute tracer results (Harman et al., 2016). Notably, we are able to further simplify the approach by assuming discharge was steady-state during each injection and having only a single release of tracer that did not overlap with other tracer signals. Under the assumption of steady flow the forward and backward transit time distributions are equal. First, we calculated the probability density of the (forward) transit time distribution ($p_Q(T)$) as:

$$p_Q(T) = \frac{Q C_{obs}}{M_{us}}$$

where M_{us} is the mass of the upstream tracer injection (g). Note that, due to the steady state assumption, $p_Q(T)$ is only a function of water age T and does not depend on time t. Next, we calculated the cumulative form of the transit time distribution ($P_Q(T)$) as:

$$P_Q(T) = \int_{\tau=0}^{T} p_Q(\tau)\,d\tau$$

where τ is a random variable representing the age of a parcel of water (Harman, 2015). This allows us to determine the age-rank discharge ($Q_T(T)$):

$$Q_T(T) = Q P_Q(T)$$

and the age-ranked storage ($S_T(T)$) as:

$$S_T(T) = Q\left(T - \int_{\tau=0}^{T} P_Q(\tau)\,d\tau\right)$$

The age-rank storage can be interpreted to determine the volume of reach storage that was sensed by the tracer. If the total storage in the study reach can be estimated, the fraction of total storage that was sensed by the tracer can also be determined. A perfect tracer study would be sensitive to the entirety of the storage volume. However, due to limitations arising from the window of detection and truncation of

Deleted: [

Deleted:]

Formatted: Font: Not Italic

Deleted:

Deleted: [

Formatted: Font: Not Italic

Deleted:]

Deleted: $C_{obs,ds}$ is the observed in-stream solute tracer concentration from the upstream injection at the downstream site (g m^{-3}) and

Deleted: P

the breakthrough curve, only a fraction of the storage is actually measured (e.g., Drummond et al., 2012). The knowledge of measured volume is important and is one advance enabled by using this interpretation framework.

Plotting the age-rank discharge as a function of the corresponding age-rank storage reveals the SAS function (Harman 2015; Harman et al. 2016). This relationship shows how discharge is composed of water drawn from storage of different ages. Flipping this plot along each axis to plot the complements is advantageous to interpret the results (Harman et al., 2016). Thus, we plot the age rank discharge complement

$$Q_{comp}(T) = Q\left(1 - P_Q(T)\right)$$

as a function of the age-rank storage complement

$$S_{comp}(T) = S_{ref} - S_T(T)$$

where S_{ref} is the total storage in the study reach (m^3). We estimated S_{ref} as the volume of the surface water (mean width × mean depth × length along centerline) plus the subsurface storage volume (valley width × valley segment length × depth × porosity). We estimated porosity as 30% for all locations (after Domenico and Schwartz, 1990; Ward et al., 2018a).

The SAS analysis can be interpreted to yield an understanding of how storage and discharge are related for the study. The minimum value of the age-rank discharge complement (Y-axis of Fig. 2) gives the discharge of outflowing water in the channel that was not labeled by the tracer at the upstream end of the study reach within the window of detection. In practice, unlabeled discharge represents some combination of (1) down-valley flow entering the segment from upstream and then upwelling, and (2) discharge originating from parts of storage that retain tracer for very long periods of time. Finally, while both the discharge and volume sampled will scale through the network, each can be normalized to a reference value as:

$$f_{VTOT}(T) = \frac{S_{comp}(T)}{S_{ref}}$$

$$f_{Q,labeled}(T) = \frac{Q_{comp}(T)}{Q + Q_{sub,cap}}$$

where f_{VTOT} is the fraction of the total storage volume that was sampled with the tracer and $f_{Q,labeled}$ is the fraction of the total down-valley discharge that was labeled with the tracer. We also calculated the fraction of the in-stream volume sampled (f_{VSTR}) as:

Deleted: [

Formatted: Font: Not Italic

Deleted:]

Deleted: [

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted:]

Deleted: [

Formatted: Font: Not Italic

Deleted:]

Commented [n34R33]: Yes. Has been corrected.

Deleted: $\mathcal{P}$

Formatted: Font: Cambria Math

Deleted: outflowing

$$f_{VSTR} = \frac{S_{comp}(T)}{AL}$$

The SAS approach requires a physically plausible bounding by input values. In practice, this means that errors in discharge can cause overestimations of mass recovery (i.e., greater than the mass that was injected), leading to physically impossible $Q_T(T)$ values. As a result, we assumed a typical error of 10% for dilution gauging (Schmadel et al., 2010). Within that range of discharge values, we calculated the range of physically plausible discharges (i.e., those which yield physically meaningful SAS calculations), and analyzed the midpoint of the plausible range. In the first study using the SAS approach to interpret solute tracers, Harman et al., (2016) found that a similar discharge adjustment was required to define the feasible parameter space.

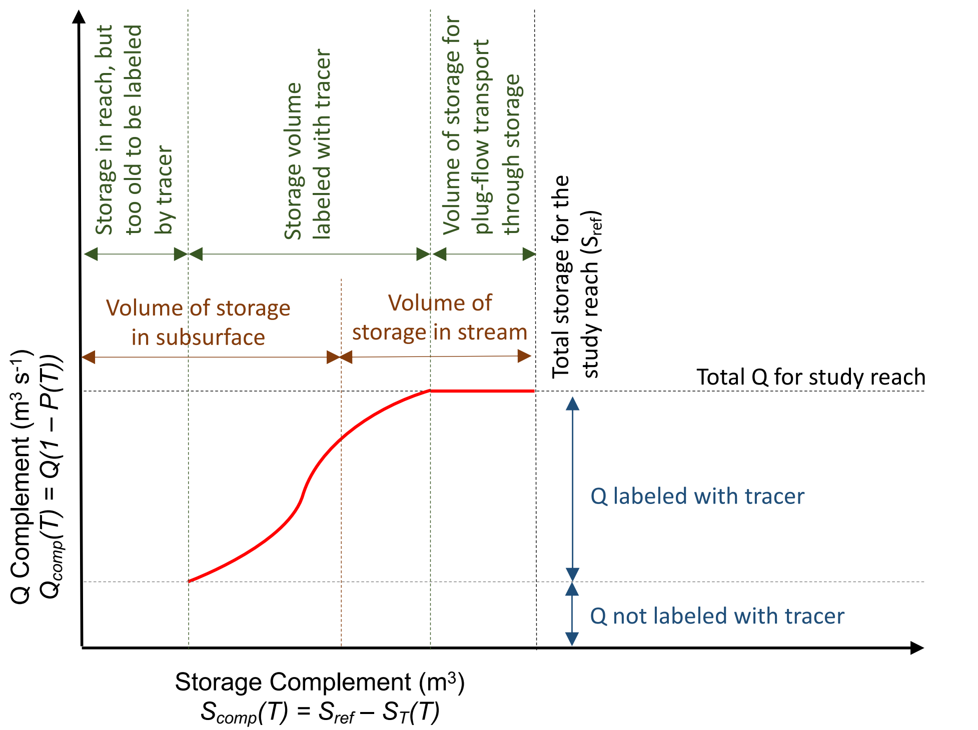


Fig. 2. Graphical representation and interpretation the SAS function. Note that the volume of storage in the stream vs. subsurface (orange above) is independent of the SAS analysis and is provided here as an example of integrating the SAS metrics with other knowledge about the system.

Formatted: Centered

Deleted: [

Deleted:]

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Deleted: [

Deleted:]

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

2.2.4 Long-term storage analysis

Long-term storage characterized the fate of mass beyond the window of detection (i.e., unrecovered mass that did not contribute to the analysis of short-term storage; Payn et al., 2009; Ward et al., 2013c). Dilution gauging at the up- and downstream ends of each study reach was used to estimate discharge (Q_{US} and Q_{DS}, respectively; $m^3\ s^{-1}$). $M_{REC} = Q_{DS}\int_0^{t_{99}} C_{obs,ds}(t)dt$Mass loss along the study reach can

$$M_{REC} = Q_{DS}\int_0^{t_{99}} C_{obs,ds}(t)dt\text{Mass}$$

be calculated by the difference of the mass injected (M_{INJ}; g) and M_{REC}:

$$M_{LOSS} = M_{INJ} - M_{REC}$$

Finally, Payn et al. (2009) demonstrate how M_{LOSS}, Q_{US}, and Q_{DS} can be used to bound the gross gains and losses of water to the channel through the study reach. We focus here on the case of all losses occurring before all gains, which is the end-member that yields the largest estimates for gross losses ($Q_{LOSS,MAX}$) and gains ($Q_{GAIN,MAX}$) respectively, calculated as:

$$Q_{LOSS,MAX} = \frac{M_{LOSS}}{\int_0^{t_{99}} C_{obs,ds}(t)dt}$$

$$Q_{GAIN,MAX} = Q_{DS} - Q_{US} - Q_{LOSS,MAX}$$

The net change in discharge along the study reach (ΔQ) is represented by the terms Q_{DS}-Q_{US} in the equation above. To compare between reaches, we normalized M_{LOSS} by M_{INJ} and normalized the gross gains and losses by Q_{US}. We also calculate gross gains and gross losses, $f_{QGAIN,MAX}$ and $f_{QLOSS,MAX}$, as a fraction of the inflow at the upstream end of the reach.

2.3 Statistical Tests

We applied a Mann-Kendall (MK) test to examine relationships between the metrics of river corridor exchange and characteristics of geologic setting and hydrologic forcing. The MK test is a non-parametric test used to assess the likelihood of a monotonically increasing or decreasing trend in a data set, which we interpret as the presence of a systematic trend through the river network. The MK test only provides an indication of a relationship's existence and does not characterize the direction nor magnitude of the relationship. Thus, we also calculated Sen's slope, a non-parametric test to fit a robust linear slope to a data set by choosing the median of slopes connecting all potential pairs of points. This

Deleted: 5
Formatted: Heading 3
Deleted: [
Deleted:]
Deleted: ¶
¶
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Moved up [1]: The mass of solute tracer recovered from the upstream injection at the downstream end of the study reach (M_{REC}; g) is calculated as:¶
¶
$M_{REC} = Q_{DS}\int_0^{t_{99}} C_{obs,ds}(t)dt$
Formatted: Font: Italic
Formatted: Font: Italic
Deleted: [
Deleted:]
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Font: Italic
Deleted: 6

metric was selected because it is less sensitive to outliers than a traditional linear regression and more robust for skewed or heteroskedastic data. Thus, we use the MK test to define the presence or absence of a statically significant trend ($p<0.05$) and Sen's slope to indicate the direction of that trend (positive or negative). We also compare the magnitude of Sen's slope among and within datasets to estimate the relative sensitivity of selected dependent variables to the same independent variable.

For the synoptic data we also report the coefficient of determination (r^2) for univariate best-fit power-law regression as an indicator of the predictive power of a parsimonious model fit. The coefficient of determination is commonly interpreted as the percent of variance explained by the model. We selected a power law regression because most independent and dependent variables span orders of magnitude. We did not test other functional forms as the purpose of this fit is to assess the explanatory power of a simple regression-model -- comparable to those commonly used to interpret field data for identifying relationships between two variables -- rather than identify an optimal predictive equation that relates the two variables. Finally, we fit a planar surface to each metric as a function of log-transformed baseflow and HYP_{POT} to approximate the conceptual model proposed by Wondzell (2011). We selected a planar surface in log-space as the simplest representation of a relationship. We also fit univariate linear relationships to the log-transformed Q and HYP_{POT} data for each metric. We emphasize here our foucs was on attesting Wondzell's (2011) conceptual model, not an exhaustive curve- nor surface-fitting exercise.

3 Results

3.1 Spatial patterns in hydrologic and geomorphic controls

Overall, all landscape metrics exhibited statistically significant monotonic trends with one another (MK test; $p < 0.05$). We found expected trends of increasing *UAA* (Fig. 3A) velocity (Fig. 3B) and stream order (Fig. 3C) with discharge. Moving from the headwaters to the outlet, we found increasing sinuosity (Fig. 3I), stream power (Fig. 3G), and flattening and widening of the valley with increasing discharge and *UAA* along the network (Figs. 3E, 3F). We also found an increasing hydraulic conductivity in the down-network direction (Fig. 3D), which is indicative of sediment size and sorting in high-relief headwater landscapes (Brummer and Montgomery, 2003), but opposite to typical low-relief alluvial systems (e.g., Gomez-Velez et al., 2015). This trend reflects the prevalence of fine material in the upper reaches emplaced by debris flows and coarsening in the downstream direction where stream power increases thus exporting fines from the system. The result of these trends in valley morphology and hydraulic conductivity is an increasing trend in $Q_{sub,cap}$ in lower network positions (Fig. 3H), indicating the increasing width and K are sufficient to overcome the decreases in slope in generating this relationship. Pairwise Pearson correlation coefficients and Spearman Rank correlation coefficients are summarized in supplemental figures S3 and S4, and table S1 and S2.

Deleted: Finally, for

Deleted: a

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Deleted: ¶

Formatted: Font: (Default) Times New Roman

Deleted: [

Formatted: Font: (Default) Times New Roman, Not Italic

Formatted: Font: (Default) Times New Roman

Deleted:]

Deleted: [

Formatted: Font: (Default) Times New Roman

Deleted:]

Formatted: Font: (Default) Times New Roman, Not Italic

Commented [KTV35]: Delete duplicate "where".

Commented [n36R35]: corrected

Deleted: where

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

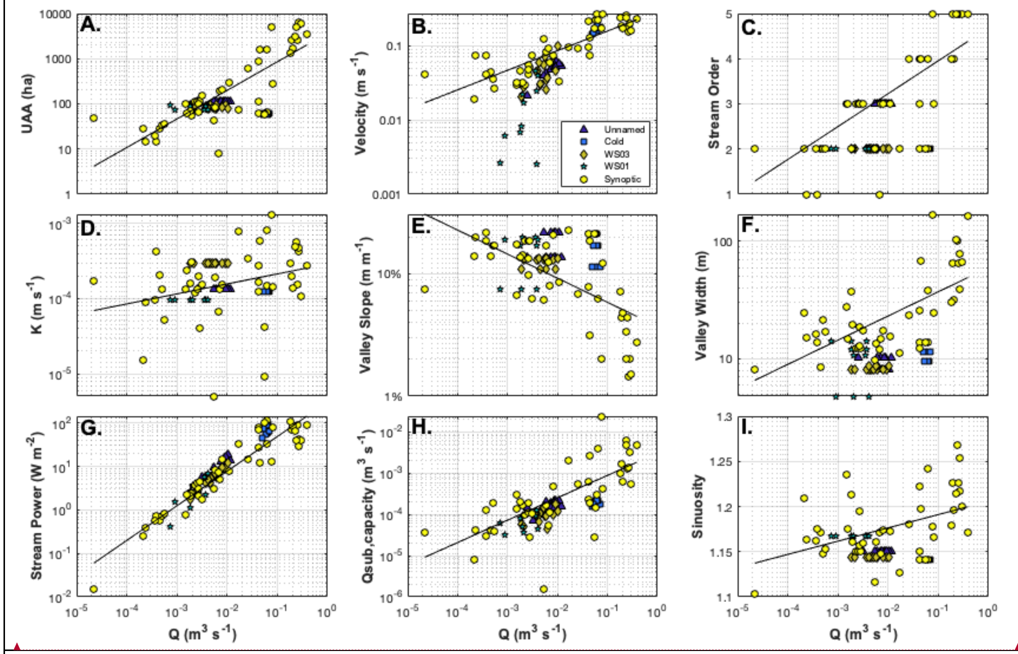


Fig. 3. For synoptic data (yellow circles), discharge exhibits a significant, monotonic trend with all other site variables considered (Mann-Kendall test; $p < 0.05$). Pairwise MK test results for all site characteristic pairs (i.e., all y-axis variables presented above) exhibit significant trends for all combinations ($p < 0.05$). The solid black line shows the best-fit power law regression for each panel. Data from unnamed creek (triangles, Cold creek (squares), WS03 (diamonds), and WS01 (stars) show the repeated injections through baseflow recession each headwater catchment. See supplemental Fig. S1 and S2 for similar plots with *HYP_{POT}* and *UAA* on x-axis.

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Commented [KTV37]: "test", singular.

Commented [n38R37]: corrected

Deleted: s

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Formatted: Font: Italic

Deleted: Page Break

Formatted: Font: (Default) Times New Roman

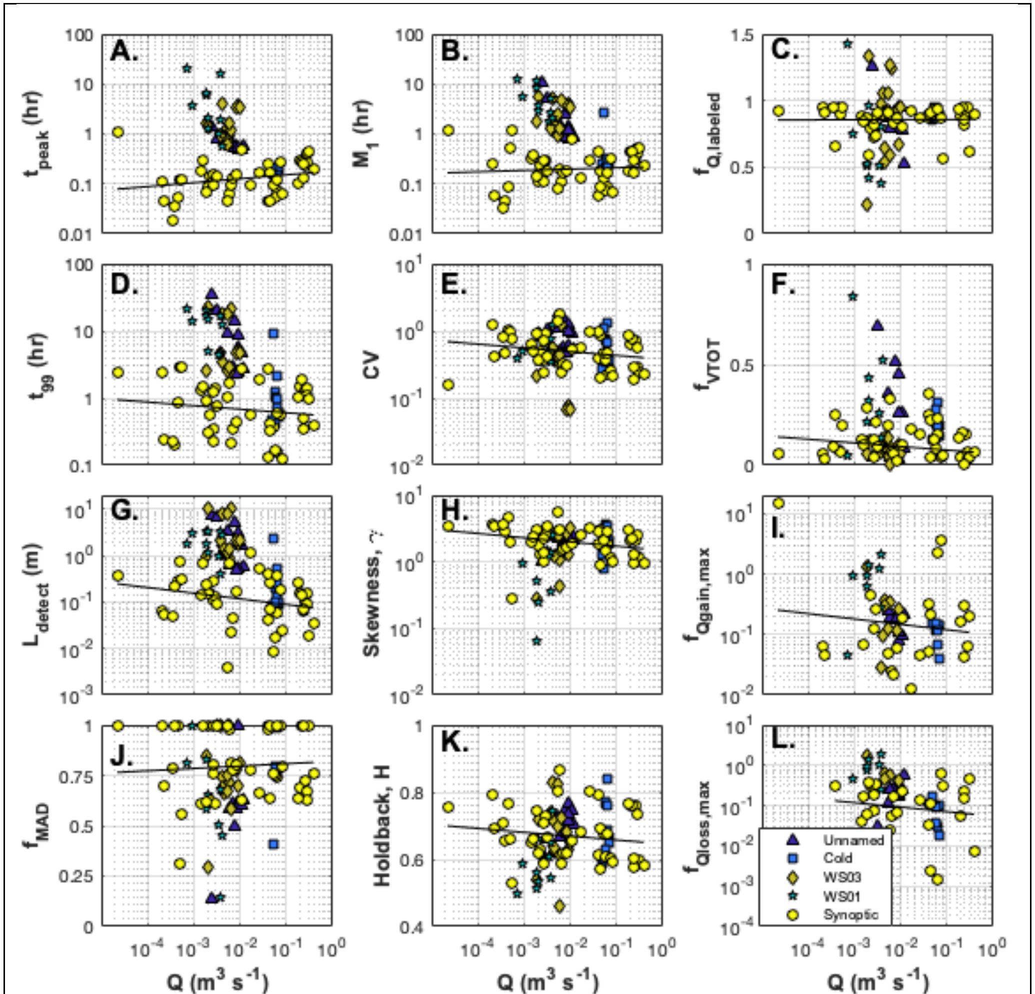


Fig. 4. Fixed reach and synoptic data as a function of stream discharge. Statistical likelihood of significant relationships (Mann-Kendall test) and their direction (Sen's slope) are detailed for all sub-reaches and the synoptic data in Table 3. All trends shown here are significant (MK test, $p < 0.05$). The coefficient of determination for power law best-fits to synoptic data (black lines) are reported in Table 3. Data from unnamed creek (triangles, Cold creek (squares), WS03 (diamonds), and WS01 (stars) show the repeated injections through baseflow recession each headwater catchment. See supplemental Fig. S5 and S6 for similar plots with *HYP$_{POT}$* and *UAA* on x-axis.

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Deleted: Table 2

Deleted: Table 2

Formatted: Font: (Default) Times New Roman

3.2 River corridor exchange trends with site characteristics

3.2.1 Basin-scale trends from synoptic campaign

An important element in our synoptic study design was the dynamic reach length, intended to minimize bias associated with the well-documented relationship between advective timescale and transient storage (e.g., Ward et al., 2013b; Schmadel et al., 2016a). Despite our efforts to hold advective travel time constant, we still found a trend of increasing t_{peak} with increasing discharge in our synoptic study (Fig. 4A). Clearly, scaling reach length relative to the wetted channel width (20 wetted channel widths) is not a perfect solution. A perfect experimental design would have resulted in no trend in advective time and provided a window of detection of constant size. While a trend was present, we also note that travel time based on t_{peak} exhibits less variation than discharge (coefficient of variation 1.00 for travel time compared to 1.49 for discharge). For context, a recent study by Ward et al. (2018b) attempted to control for experiments with 20-min of advective time and accepted a range from 17 to 50 minutes as comparable. Thus, while our selection of study reach lengths was imperfect to achieve identical advective timescales, we contend that we have adequately controlled for advective time.

Overall we found significant trends (MK test; $p < 0.05$) between nearly all site characteristics and metrics describing river corridor exchange. Of the 130 pairings investigated, only three (stream order vs. L_{detect}; stream order vs. f_{MAD}; Sinuosity vs. $f_{Qlabeled}$) were not significant (Table 3). However, while network-scale trends do exist, we note high site-to-site variation in the data set as evidenced by the low r^2 for the power-law fits (see trendlines in Fig. 4), representative of the range of explanatory power observed. Across all 130 pairings investigated, we found very little explanatory value in the model fits, with a median r^2 of less than 0.03 (i.e., the variance in the model errors are about 3% less than the variance in the dependent variable itself). The lack of explanatory power for individual variables may indicate that fits based on more complex functional forms and/or multivariate approaches would increase predictive power. We did observe improved r^2 for all fits using both Q and HYP_{POT} compared to univariate regressions (Table S3).

Formatted: Font: (Default) Times New Roman

Commented [KTV39]: "hold"

Commented [n40R39]: corrected

Deleted: [

Formatted: Font: (Default) Times New Roman

Deleted:]

Formatted: Font: (Default) Times New Roman, Not Italic

Formatted: Font: (Default) Times New Roman, Not Italic

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, Not Italic

Formatted: Font: (Default) Times New Roman

Deleted: [

Deleted:]

Deleted: in a more rigorous way than has been done in most previous studies.

Deleted: Table 2

Formatted: Font: (Default) Times New Roman

Deleted: indicates

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Deleted: modeling

Formatted: Superscript

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Deleted: may be necessary to increase predictive power.¶

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Table 3. Mann-Kendall tests indicate significant ($p < 0.05$), monotonic trends relating almost all site characteristics and metrics of river corridor exchange for the synoptic survey locations. The direction of the trend is indicated as increasing ("+") or decreasing ("(-)"). Three relationships that lacked a significant trend are denoted "?" in the table below. Additionally, the magnitude of the coefficient of determination (r^2) for univariate power-law fit is presented as an indicator of the power of a simple regression.

Metric		Experimental Design			AD vs. TS	Short term storage				rSAS			Long-term storage	
		t_{99}	t_{peak}	L_{detect}	f_{MAD}	M_1	CV	γ	Holdback	$f_{Q,labeled}$	f_{VSTR}	f_{VTOT}	$f_{QGAIN,MAX}$	$f_{QLOSS,MAX}$
Sen's Slope	Q	(-)	+	+	+	+	(-)	(-)	(-)	+	(-)	(-)	(-)	(-)
	UAA	+	+	+	(-)	+	(-)	(-)	(-)	(-)	(-)	(-)	+	(-)
	V	(-)	+	+	+	(-)	(-)	(-)	(-)	+	(-)	(-)	+	+
	Stream Order	+	+	?	?	+	(-)	(-)	(-)	(-)	(-)	(-)	+	(-)
	K	(-)	+	(-)	+	+	(-)	(-)	(-)	(-)	(-)	(-)	+	(-)
	Down-Valley Slope	(-)	(-)	(-)	(-)	(-)	+	+	+	(-)	+	+	(-)	(-)
	Valley Width	+	+	+	+	+	(-)	(-)	(-)	+	(-)	(-)	(-)	(-)
	Stream Power	(-)	+	+	+	+	(-)	(-)	(-)	+	(-)	(-)	(-)	(-)
	$Q_{sub,cap}$	(-)	+	(-)	+	+	(-)	(-)	(-)	(-)	(-)	(-)	+	(-)
	Sinuosity	+	+	(-)	+	+	+	(-)	+	?	(-)	(-)	+	(-)
r^2 for power law fit	Q	0.07	0.00	0.37	0.00	0.04	0.05	0.15	0.02	0.00	0.15	0.00	0.92	0.23
	UAA	0.00	0.11	0.51	0.01	0.03	0.06	0.13	0.01	0.00	0.18	0.05	0.00	0.02
	V	0.18	0.00	0.29	0.02	0.08	0.08	0.03	0.01	0.02	0.19	0.01	0.01	0.00
	Stream Order	0.01	0.10	0.03	0.00	0.04	0.02	0.10	0.00	0.00	0.14	0.04	0.00	0.03
	K	0.02	0.03	0.10	0.02	0.00	0.13	0.10	0.04	0.01	0.07	0.01	0.03	0.11
	Down-Valley Slope	0.01	0.19	0.03	0.04	0.07	0.05	0.04	0.01	0.00	0.10	0.13	0.01	0.02
	Valley Width	0.00	0.02	0.01	0.01	0.01	0.06	0.16	0.03	0.00	0.06	0.09	0.92	0.34
	Stream Power	0.07	0.00	0.74	0.00	0.04	0.05	0.15	0.02	0.00	0.15	0.00	0.92	0.23
	$Q_{sub,cap}$	0.02	0.03	0.62	0.00	0.00	0.19	0.30	0.07	0.00	0.17	0.03	0.02	0.11
	Sinuosity	0.01	0.00	0.01	0.01	0.00	0.00	0.03	0.00	0.00	0.03	0.01	0.92	0.22
	MAX	0.18	0.19	0.74	0.04	0.08	0.19	0.30	0.07	0.02	0.19	0.13	0.92	0.34
	MEDIAN	0.02	0.03	0.20	0.01	0.04	0.05	0.11	0.02	0.00	0.14	0.02	0.02	0.11
	MEAN	0.04	0.05	0.27	0.01	0.03	0.07	0.12	0.02	<0.01	0.12	0.04	0.37	0.13
	MIN	<0.01	<0.01	0.01	<0.01	<0.01	<0.01	0.03	<0.01	<0.01	0.03	<0.01	<0.01	<0.01

3.2.2 Fixed reach vs. synoptic results

We found decreasing t_{99} with increasing discharge for the synoptic study (Fig. 4D), which in turn resulted in a systematic reduction in the possible length of flowpaths that could be detected by tracer (Fig. 4G). Note that this ranges, on average, from 0.35 m at the lowest discharge to only 0.09 m at the highest discharge and the reach with the largest L_{detect} was only 2.0 m. In contrast, reach lengths used in the fixed reach studies were much longer relative to stream size than the synoptic reaches, thus t_{peak}, M_1, t_{99}, and L_{detect} were all much larger in the fixed reach studies (Table 4). These metrics all exhibited significant trends with discharge (Table 3), but the trends were not regularly consistent in their direction with the synoptic results. Overall, we found predominantly decreasing t_{peak} with discharge in the fixed reaches - opposite to the synoptic finding - for 9 of 11 fixed reaches (and steeper Sen's slope in 9 of 11 fixed reaches). We also found decreasing t_{99} with discharge in 9 of 11 fixed reaches (all with steeper Sen's slope than the synoptic), and decreasing L_{detect} with discharge in 9 of 11 fixed reaches (all with steeper Sen's slope than the synoptic). Even with the longer reach lengths, relative to stream size, used

Deleted: Table 2

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Deleted: 3.3

Formatted: Font: (Default) Times New Roman

Commented [KTV41]: It would be helpful to have a table of reach lengths and widths for each injection to illustrate this point. This table could go in an SI, as it would likely be quite large. If the authors decide to produce such a table, it would also be a good place to include NaCl injection masses, travel times, etc.

Commented [n42R41]: All the data used in this paper will be published and made publicly available, including all reach lengths, tracer masses, etc. We feel that including such a table as part of this manuscript may introduce confusion because we present metrics derived from that data that provide meaningful process understanding whereas lengths themselves do not.

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, Italic

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, Italic

Formatted: Font: (Default) Times New Roman, Italic

Formatted: Font: (Default) Times New Roman

Deleted: Table 3

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, Italic

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Deleted: Table 2

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, Not Italic

Formatted: Font: (Default) Times New Roman

in the fixed reach studies, L_{detect} averaged only ~2.0 m, and ranged from a maximum of 10 m to a minimum of 0.10 m.

With respect to short-term storage, we found increasing M_1 with increasing discharge in the synoptic study, but this direction was reflected in only 2 of 11 fixed reaches. Sen's slope was larger in magnitude for 10 of the 11 fixed reaches, indicating M_1 interpreted from the fixed reach approach is more sensitive to discharge than the synoptic approach. We found overall decreasing CV, γ, and H with increasing discharge in the synoptic study, indicating a decreasing importance of non-advective processes in the downstream direction along the network. The direction of this trend is consistent with 7 fixed reaches for CV, 2 sites for γ, and 3 sites for H. Regardless of the direction of the relationship, the magnitude of Sens slope was larger for all fixed reaches compared to the synoptic study, indicating increased sensitivity to discharge relative to the synoptic sites.

For long-term storage and mass involved in advection-dispersion, we again found fixed-reach trends were steeper and often opposed the direction of the trend for the synoptic data. For the synoptic study we found decreasing $f_{Qgainmax}$ (Fig. 4I) and $f_{Qlossmax}$ (Fig. 4L) with increasing discharge, which is consistent with 5 and 6 of the 11 fixed reaches, respectively. For the synoptic study we found an overall decreasing f_{MAD} with increasing discharge, consistent with 7 of the 11 fixed reaches. The magnitude of Sens slope was larger for the fixed reaches than the synoptic study for f_{MAD}, $f_{Qgainmax}$, and $f_{Qlossmax.}$

The SAS analysis revealed decreasing sampling of the total storage zone (f_{Vtot}) with increasing discharge, but increasing $f_{Q,labeled}$ with discharge for the synoptic study. Together, these results indicate that increasing discharge in synoptic experiments resulted in sampling a larger fraction of the water exiting the reach, but smaller total volume of storage. Put another way, experiments in locations with higher discharge were more likely to measure storage in (or proximal to) the stream channel at the expense of measuring more distal flowpaths and less-connected storage. For the fixed reach studies, we found decreasing f_{Vtot} and $f_{Qlabeled}$ in 7 and 6 of the 11 reaches, respectively. In all cases, the magnitude of Sens slope was larger for the fixed reaches than the synoptic study.

Commented [KTV43]: This seems to be a new term. Perhaps it should be "$f_{Q,labelled}$"?

Commented [n44R43]: Wow! Thanks for your detailed attention. Yes, you are correct and has been corrected.

Deleted: $_{tot}$

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Table 4. Sen's slope for all discharge-metric relationships across fixed reach study sites and the synoptic site data. All relationships were significant ($p < 0.05$) using the Mann-Kendall test. The values shown indicate the direction of the relationship based on a Sen's slope estimator ("+" indicates a direct relationship with discharge, and "(-)" indicates inverse relationship with discharge). Slopes were larger in magnitude for the fixed reaches in all cases except Cold Creek sites 12 and 23 for t_{peak} and Cold Creek site 12 for M_1, denoted with "*".

			Experimental Design			AD vs. TS	Short term storage				rSAS			Long-term storage	
Catchment	Study Reach	n	t_{99}	t_{peak}	L_{detect}	f_{MAD}	M_1	CV	γ	Holdback	$f_{Q,labeled}$	f_{VSTR}	f_{VTOT}	$f_{QGAIN,MAX}$	$f_{QLOSS,MAX}$
Unnamed	Site12	6	(-)	(-)	+	+	(-)	(-)	+	+	(-)	(-)	(-)	+	(-)
	Site23	6	(-)	(-)	+	(-)	(-)	(-)	+	+	(-)	(-)	(-)	+	(-)
Cold	Site12	5	(-)	(-)*	+	(-)	(-)*	(-)	(-)	(-)	+	+	+	(-)	+
	Site23	5	+	(-)*	(-)	+	+	+	+	+	+	(-)	(-)	+	+
WS03	Site12	5	+	(-)	(-)	(-)	+	(-)	+	+	+	(-)	+	(-)	+
	Site23	5	(-)	(-)	+	+	(-)	(-)	+	(-)	(-)	(-)	+	+	(-)
	Site34	5	(-)	+	+	+	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)
WS01	Site12	3	(-)	(-)	+	(-)	(-)	(-)	+	+	(-)	+	+	+	(-)
	Site23	2	(-)	+	+	(-)	(-)	+	+	+	+	+	(-)	(-)	+
	Site34	2	(-)	(-)	+	(-)	(-)	+	+	+	+	(-)	(-)	(-)	+
	Site45	3	(-)	(-)	+	(-)	(-)	+	+	+	+	+	(-)	(-)	+
HJA	Synoptic	46	(-)	+	+	(-)	+	(-)	(-)	(-)	+	(-)	(-)	(-)	(-)

3.3 Selection of study reach length across the network

For the injections that specifically tested the study reach length, we found the most consistent advective timescales were obtained by scaling reach length to 20 times wetted channel width (Fig. 5). Ranges of advective timescales were 25.2 minutes for the fixed-length approach, 27.2 minutes for the fixed timescale approach, and 4.8 minutes for the 20× wetted channel width approach (Fig. 5A). It is notable that our estimates of a 10-min advective time were reasonably accurate for the three highest-discharge reaches, but the lowest-discharge replicate primarily drives the visually steep trend. We hypothesize that a better estimate of advective velocity -- such as using a dye tracer rather than following debris or a longer length-scale of integration -- may have improved that estimate. For t_{99}, ranges for the 10-min and 150-m approaches are about 29% and 22% larger, respectively, than the 20× wetted channel width approach (Fig. 5B). Differences are even more striking for other parameters, with the 10-min and 150-m study designs, yielding 147% and 93% larger ranges for H compared to the 20× wetted channel width approach (Fig. 5C). Similarly, the 10-min and 150-m approaches result in ranges of γ that are 96% and 101% larger than the ranges using the 20× wetted channel width approach (Fig. 5D).

Deleted: Table 3

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

Formatted: Font: Italic

Formatted: Font: Italic

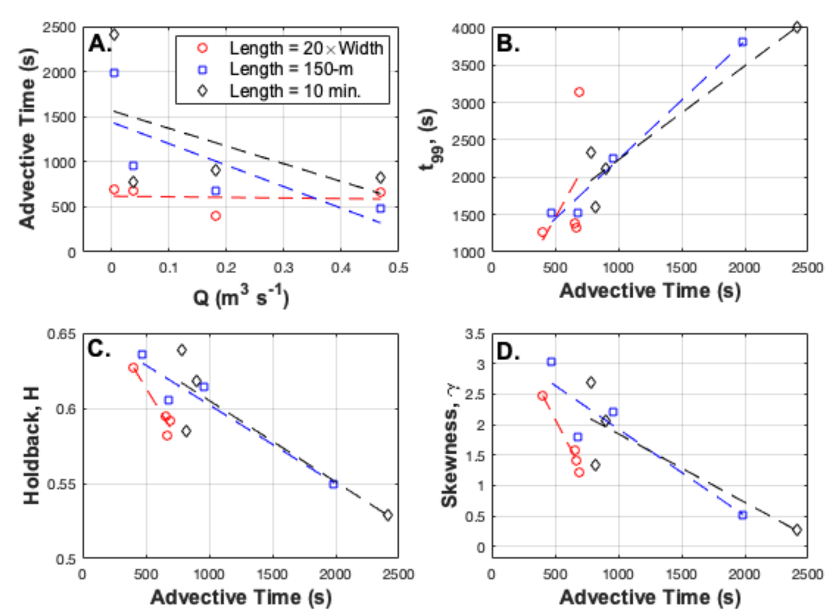


Fig. 5. Comparison of fixed reach (150-m), adaptive reach length (20 wetted channel widths), and fixed advective time (10-min) approaches for standardization of stream solute tracer studies. (A) Control of advective time across four stream orders. Additional panels show the observations and a best-fit linear regression for (B) longest detection timescale, (C) holdback, and (D) skewness in relation to the advective time of the study. Best-fit linear regressions are shown as dashed lines in each panel.

4 Discussion

4.1 How do discharge and local geomorphic setting modulate river corridor exchange?

Our overarching objective in this study was to test the conceptual model of Wondzell (2011), which predicted systematic changes in river corridor exchange as a function of changing baseflow and geomorphic setting (Fig. 6A). We found a generally decreasing influence of river corridor exchange with increasing steady-state discharge through space for most metrics considered (Fig. 4; Fig. 6). This finding is in agreement with the conceptual model of Wondzell (2011), who predicted Q_{HEF}/Q would decrease as drainage area increased. We did find an increasing fraction of total discharge sampled in higher discharge locations (Fig. 4C), but the overall trend indicates that Q_{HEF} does not grow as rapidly as Q, moving downstream along the network. This is consistent with findings of decreased river corridor exchange in network locations with larger discharge (e.g., Covino et al., 2011; Ward et al., 2013c).

Formatted: English (UK)
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: hydrologic forcing
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Deleted: [
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted:]

Two explanations have been posed relating river corridor exchange to time-variable baseflow in a given study reach, both of which result in less exchange under higher discharge conditions. First, many conceptual models would predict that increasing baseflow is associated with increasing groundwater discharge to the stream, resulting in compression of hyporheic zones and decreased river corridor exchange (Hakenkamp et al., 1993; Hynes, 1983; Palmer, 1993; Vervier et al., 1992; White 1993). Second, exchange may change little during storm events because, under a wide range of discharge conditions, the effect of the geomorphic features driving exchange flows may be relatively static (Ward et al., 2017b). Thus, if Q_{HEF} is relatively static, as Q increases the relative amount of relative exchange (Q_{HEF}/Q) will decrease. Both explanations appear logical and suggest that river corridor exchange should change systematically with discharge. However, we did not find a consistent pattern in our synoptic field study. Rather, of the diverse array of metrics used to characterize river corridor exchange in the synoptic study, some increased and some decreased with increasing discharge. We found similarly contradictory results among our fixed reach studies. For example, only 2 of 11 fixed reaches exhibited the expected negative relationship based on skewness (one indicator of Q_{HEF}/Q) and discharge (Table 4).

4.2 Heterogeneity in the river network

Wondzell's (2011) conceptual model followed general predictions about systematic changes in channel morphology with increasing stream size, predicting channel width, channel depth, and flow velocity will all increase with discharge, both over time at a fixed cross section or with location at a given time within a stream network. Further, bed sediment size distributions would generally decrease in a downstream direction (see, for example, Leopold and Maddock, 1953). While the physical attributes we measured at our synoptic sites did show systematic variation, the pattern in saturated hydraulic conductivity (K) was contrary to expectations, as we found K increased in the downstream direction. This change was so large that it overwhelmed the effect of decreasing longitudinal gradient so that the hyporheic potential actually increased in a downstream direction. We note, however, that our studies only spanned about 4 orders of magnitude in hyporheic potential while Wondzell's (2011) model visualizes a range that spans 14 orders of magnitude. Our study is also limited to the upper end of the range in hyporheic potential depicted by Wondzell (2011).

Our dataset also showed substantial spatial heterogeneity in all metrics along the river corridor. While Wondzell's (2011) conceptual model does not expressly disallow such heterogeneity, the data points he used to develop the conceptual model suggest very uniform changes with watershed area and little change in hyporheic potential from 2nd- to 5th-order reaches within the same mountain stream network studied here. Our results suggest that the influence of reach scale heterogeneity among sites may be as large as, or even larger than, the expected systematic changes with watershed size. We also note that our results may differ from those of Wondzell (2011) for methodological differences. First, Wondzell (2011) based his estimates of K from extensive well networks at each of his sites, using the geometric mean of all wells – including many wells on the floodplain adjacent to the stream as well as piezometers

Deleted: discharge
Formatted: Font: (Default) Times New Roman
Deleted: discharge
Formatted: Font: (Default) Times New Roman
Deleted: [
Formatted ... [14]
Deleted:]…. Second, exchange may change little during storm events because, under a wide range of discharge conditions, the effect of the geomorphic features driving exchange flows may be relatively static [... [15]
Formatted ... [16]
Deleted:]…. Thus, if Q_{HEF} is relatively static, as Q increases the relative amount of relative exchange (Q_{HEF}/Q) will decrease. These ... [17]
Formatted: Font: (Default) Times New Roman
Deleted: seem
Formatted: Font: (Default) Times New Roman
Deleted: changes in …ischarge. However, we did not find a consistent pattern in our synoptic field study. Rather, of the diverse array of metrics used to characterize river corridor exchange in the synoptic study, some increased and some decreased with increasing ... [18]
Formatted: Font: (Default) Times New Roman
Deleted: an
Formatted: Font: (Default) Times New Roman
Deleted: Table 3
Formatted: Font: (Default) Times New Roman
Deleted:
Formatted ... [19]
Deleted: […2011]… conceptual model followed general ... [20]
Formatted ... [21]
Deleted:]
Formatted ... [22]
Deleted: […2011]… model visualizes a range that is ... [23]
Formatted: Font: (Default) Times New Roman
Deleted: three times larger (…4 orders of magnitude) ... [24]
Formatted ... [25]
Deleted: […2011] ... [26]
Formatted ... [27]
Deleted: […2011] ... [28]
Formatted ... [29]
Deleted: […2011] ... [30]
Formatted ... [31]
Deleted: […2011] ... [32]

installed through the streambed. This study estimated K from a single, 50-cm deep piezometer located in the channel thalweg and Wondzell's (2011) data show that K is higher in piezometers inserted into the shallow streambed than in floodplain sediment adjacent to the stream. Second, Wondzell (2011) used numerical simulations from groundwater flow models to calculate Q_{HEF}, whereas exchange metrics in this study were derived from stream solute tracer injections. Solute injections are sensitive to both surface (in-stream) and subsurface transient storage, and metrics derived from these studies have a known bias toward the shortest transit times (Harvey et al., 1996; Wagner and Harvey, 1997; Harvey and Wagner, 2000), a bias that is clearly evident in our data. For example, the longest timescale flowpath detectable, interpreted from t_{99}, in our study reaches ranged from about 8 minutes to 2.8 hours. In contrast, Wondzell's (2011) simulations included flowpaths with up to 10 day transit times. However, cell sizes in the finite-difference grids used in his models limited the shortest flow paths that could be simulated, so his estimates of Q_{HEF} should underrepresent the very shortest flow paths present within the reach.

Transient storage in the surface (in-stream) channel is known to influence tracer breakthrough in solute injection experiments and more specifically, has been documented in our study basin (*Jackson et al.*, 2012, 2013). Thus, our data represent a combination of surface and hyporheic transient storage, but we expect the hyporheic component will be most sensitive to hydraulic conductivity. Thus, deviation from the expected trend with hyporheic potential may simply indicate that our tracer studies were not solely representative of Q_{HEF} between a stream and its hyporheic zone as defined and assumed by Wondzell (2011) (Fig. 6). Our SAS analyses indicate we measured storage volumes larger than the stream in most reaches, but it is unclear what the mechanisms or timescales of exchange were for the storage locations measured. Overall, this unique basin scale dataset does not appear to support Wondzell's (2011) conceptual model with respect to hyporheic potential, but it does not disprove it either due to the limitations in methods and clustering on only the highest end of the axis likely biased our results. Still, we suggest local-scale processes specific to individual sites may overwhelm basin-scale trends and limit the ability of continuum based conceptual models, such as Wondzell (2011), to predict local-scale hyporheic and river corridor exchange dynamics.

Formatted: Font: (Default) Times New Roman, Not Italic
Deleted: [
Formatted: Font: (Default) Times New Roman
Deleted:]
Deleted:]
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted: [
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Deleted:]
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Deleted: under represent
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Deleted: Fig. 5
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]

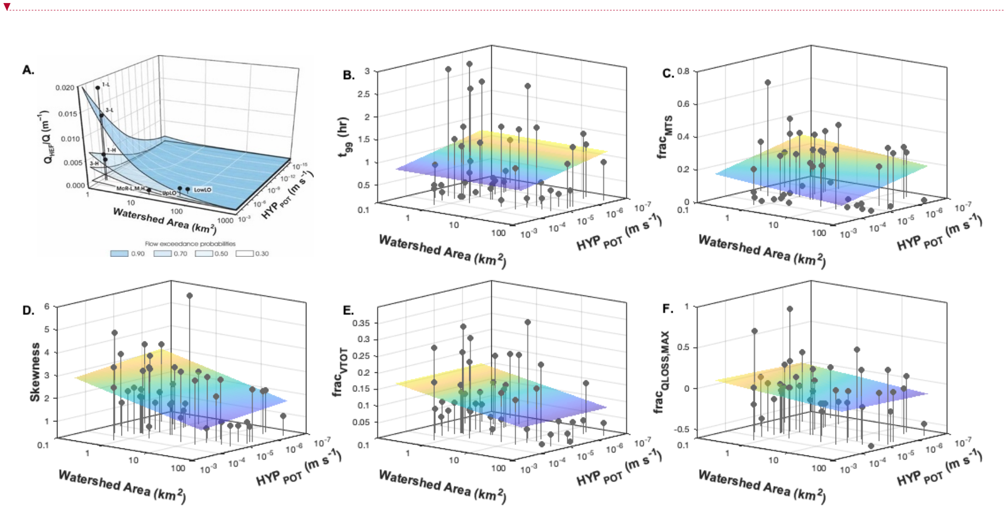
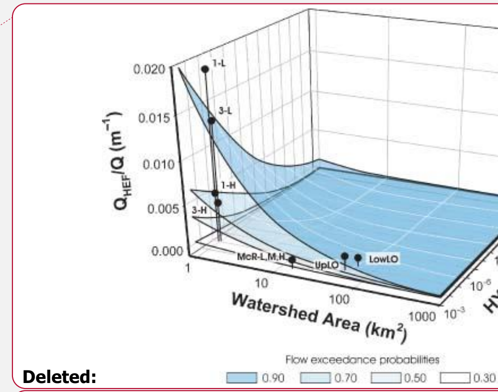


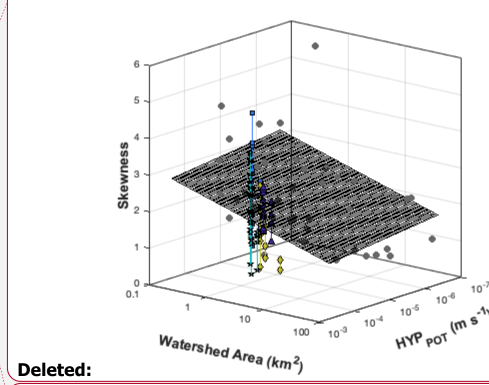
Fig. 6. **Comparison of (a) conceptual model of river corridor exchange (reprinted from *Wondzell* (2011) with permission) and findings from this study including a best-fit planar surface fit to the synoptic data for each panel (dots show the data points, stems extend to the bottom X-Y plane to aid in visualization; planar surface light-to-dark shading indicate high-to-low for the Z-axis variable). Panels show trends for a sub-set of variables representing (B) experimental design, (C) separation of advection-dispersion from transient storage, (D) short-term storage, (E) StorAge Selection, and (F) long-term storage. Goodness of fit and slopes for each fit are summarized in Table S3.**

4.3 Can space-for-time or time-for-space relationships be used to transfer findings based on reach-scale characteristics?

Transferability of findings in space or time relies upon two assumptions, both of which are necessary conditions for reliable prediction. First, transferability requires that the process of interest varies systematically with at least one observable variable at the study and predicted sites. In our case, this requires the relationship between discharge and river corridor exchange to be measurable and robust, commonly judged on the basis of a goodness-of-fit metric for a regression. Transferability also requires that the functional form established from the observations holds for the conditions that are being predicted. In the temporal domain this is most commonly interpolation in time to predict river corridor exchange under a discharge condition that was not actually observed (e.g., Harman et al., 2016; Ward et al., 2018a). In the spatial domain, this transferability strategy may manifest as interpolation between observed sites (e.g., Covino et al., 2011; Mallard et al., 2014) or extrapolation to sites that are morphologically similar, such as extending findings from one headwater site to make predictions in an adjacent basin or another stream reach (e.g., Jencso et al., 2011; Covino et al., 2011; Stewart et al., 2011). This approach assumes that the relationship holds because the observational and predicted sites



Deleted:



Deleted:

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Deleted: Fig. 5…ig. 6. Comparison of (a) conceptual model of river corridor exchange (reprinted from *Wondzell* […2011]… with permission) and (b) …indings from this study including a ... [33]

Deleted: ¶

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Deleted: [

Formatted ... [34]

Deleted:]…. In the spatial domain, this transferability strategy may manifest as interpolation between observed sites [... [35]

Formatted ... [36]

Deleted:]… or extrapolation to sites that are morphologically similar, such as extending findings from one headwater site to make predictions in an adjacent basin or another stream reach [... [37]

Formatted ... [38]

Deleted:]

are similar. However, we find that there is substantial variation among sites, particularly when reaches of similar size yield opposing relationships with explanatory variables (Tables 2, 3).

Overall, we conclude that discharge alone is a poor predictor of river corridor exchange in mountain stream networks due to heterogeneity in reach-scale geomorphic setting and should not be used as the sole basis for spatial or temporal extrapolation of findings. We found opposing relationships between river corridor exchange and discharge through space (synoptic approach) and time (fixed reach approach). For all metrics considered, at least 18% (2 of 11) of the intensively studied fixed reaches had trends opposite of that what would be predicted from the one-time sampling of the synoptic study. Moreover, the opposing trends were always located across at least two different landform types, and there were examples of within-landform type disagreement for every metric considered. Furthermore, the regressions we developed indicated that there was substantial inter-site heterogeneity overriding the observed network-scale trends. These findings are useful for identifying best practices to ultimately develop better scaling relationships to predict river corridor exchange as a function of hydrologic forcing and geomorphic setting from headwaters to oceans. For example, intensively studying a small number of study reaches is not indicative of the conditions occurring across an entire basin, even at the scale of our 5th order basin. We further develop suggestions for best practices and considerations in the next section.

4.4 Best practices to measure and interpret exchange-discharge relationships

Stream solute tracers are perhaps the empirical method most frequently used to measure river corridor exchange. Given the relative ease and low cost of this method, it is unsurprising that many studies have used solute tracer studies under different discharge conditions to assess relationships between discharge and river corridor exchange. For example, some studies repeat solute injections in a fixed reach under range of discharge conditions during different seasons (e.g., Zarnetske et al., 2007; Ward et al., 2018b), during baseflow recession (e.g., Payn et al., 2009; Ward et al., 2012), or during storm events (e.g., Ward et al., 2013b; Dudley-Southern and Binley, 2015). Still others use spatial replication at multiple sites within a network to construct a relationship that can be used to predict behavior for unstudied reaches during a single discharge condition (e.g., Jencso et al., 2011; Covino et al., 2011; Stewart et al., 2011). However, limitations of stream solute tracers are well documented in the literature as mentioned above (Harvey et al., 1996; Wagner and Harvey, 1997; Harvey and Wagner, 2000; Drummond et al., 2012; Kelleher et al., 2013; Ward et al., 2017a).

The ability to detect late-time tailing of the tracer (e.g., Drummond et *al.*, 2012) and parameter dependence on advective timescales of transport (e.g., Schmadel et al., 2016a) limit the interpretability of solute tracer studies. However, armed with a seemingly straightforward tool (e.g., stream solute tracers) and the expectation to find trends with discharge, it is logical that many studies have concluded discharge (or its tightly correlated proxy of drainage area) is a meaningful predictor of river corridor exchange. However, we argue this may be a self-fulfilling prophecy as it is often unclear exactly what is being measured by the tracer observations. For fixed-reach studies repeated under different discharge

Formatted: Font: (Default) Times New Roman

Deleted: [

Deleted:]

Formatted ... [39]

Formatted: Font: (Default) Times New Roman

Deleted: [

Deleted:]..., or during storm events [... [41]

Formatted ... [40]

Formatted ... [42]

Deleted:].... Still others use spatial replication at multiple sites within a network to construct a relationship that can be used to predict behavior for unstudied reaches during a single discharge condition [... [43]

Formatted ... [44]

Deleted:].... However, limitations of stream solute tracers are well documented in the literature as mentioned above [... [45]

Formatted ... [46]

Deleted:]

Deleted: [

Formatted ... [47]

Deleted:]... and parameter dependence on advective timescales of transport [... [48]

Formatted ... [49]

Deleted:]

Formatted: Font: (Default) Times New Roman

conditions, the observed trends between river corridor exchange and discharge can be plausibly explained by either physical transport processes or simply limitations of the tracer method. Indeed, this unfortunate conclusion was clearly illustrated by recent studies focused on solute tracer studies across a range of discharge conditions (e.g., Wondzell, 2006; Schmadel et al., 2016a). Thus, we contend that it is unknown if reported trends in the literature reflect mechanistic understanding of the river corridor or suffer from confirmation bias. Therefore, we detail two best practices for conducting and interpreting stream solute tracer tests for those seeking to do as we have attempted in this study.

4.4.1 Best practice 1: Control for advective timescales instead of reach length.

The most common paradigm in stream solute tracer studies is to use a fixed-length study reach, and hold length constant to compare different reaches (e.g., Payn et al., 2009; Covino et al., 2011) or to compare different discharge conditions at a single reach of fixed length (e.g., Schmadel et al., 2016a; Ward et al., 2013a). The implicit logic is that by fixing the reach length, the same morphologic features interact with the tracer and allow the researcher to measure changes in the same processes. However, this is only true in the case where the same suite of flowpaths can be detected. When advective timescales decrease, the window of detection (i.e., the longest timescale flowpath that can be detected) should decrease in response (e.g., Schmadel et al., 2016a). As a result, the fixed reach causes systematic bias in the tracer experiment. Higher discharges will have smaller windows of detection, biasing the results toward shorter timescale flowpaths compared to low-discharge injections.

Based on our findings (Fig. 5), plus the well-documented interaction of advective timescale with river corridor exchange measured with solute tracers, we strongly recommend experimental designs that control for advective timescale. We suggest that an upstream location be established and fixed in space. Then the length of the study reach should be determined, either by scaling by channel width (e.g., 20 times the wetted channel width) or by using a dye tracer to measure advective velocity over a length equal to perhaps 10 wetted channel widths, and then using advective velocity to calculate a study reach length that provides uniform advective travel times in all reaches studied.

When tracer injections are designed to provide uniform advective travel times, the resulting study reach lengths will be longest in the largest streams and/or at times of high discharge; reaches will be shortest under low discharge conditions. It is critical that the shortest reach length still encompass a length of stream that is sufficient to integrate representative variation in morphology of the study system. If reaches are too short, high reach-to-reach variability will be generated by one or a few morphologic features and these local conditions are likely to dominate comparisons among reaches and make it difficult to discern the influence of changing hydrologic conditions. It will be difficult to determine a length-scale long enough to integrate the full range of morphologic features present in any given stream. Schmadel et al. (2014) suggested that a morphologically representative reach could be determined by knowing the length of spatial autocorrelation of morphologic features, but this requires substantial effort to survey or map the study reach prior to conducting a tracer test. A less effort-intensive, but more

Deleted: [
Formatted ... [50]
Deleted:]
Formatted ... [51]
Formatted ... [52]
Formatted ... [53]
Formatted ... [54]
Formatted ... [55]
Deleted: [
Formatted ... [56]
Formatted ... [57]
Formatted ... [58]
Formatted ... [59]
Deleted:]
Deleted: [
Formatted ... [60]
Formatted ... [61]
Formatted ... [62]
Formatted ... [63]
Formatted ... [64]
Deleted:]
Deleted: [
Formatted ... [65]
Formatted ... [66]
Deleted:]
Formatted ... [67]
Deleted: In the synoptic campaign, we scaled our tracer reach ... [68]
Formatted ... [69]
Formatted ... [70]
Formatted ... [71]
Formatted ... [72]
Formatted ... [73]
Formatted ... [74]
Formatted ... [75]
Formatted ... [76]
Deleted: here
Formatted ... [77]
Formatted ... [78]
Formatted ... [79]
Deleted: [
Deleted:]

equipment-intensive approach would be to place multiple sensors in the study reach (perhaps 10, 20, 35, 50, 75, and 100 wetted channel widths) and select most appropriate downstream breakthrough curves to compare based on similarity of advective timescales after conducting the tracer test.

It is also essential that measures of the advective timescale and window of detection be reported for each tracer test. For slug injections these would include t_{peak} and t_{99}. For constant rate injections these would be time to the steepest point on the rising limb, time to median arrival (M_1), and time to achieve plateau. The L_{detect} estimates should also be reported and these should be based on time to achieve plateau as that indicates when tracer has traveled the full length of all measurable flowpaths and only tracer-labeled water is being returned to the stream. These metrics describing the advective timescale are necessary both to confirm that comparisons among reaches in any given study are valid and to facilitate comparisons of results among published studies.

We acknowledge here that the steps we've recommended above will require substantial time and analysis to design a stream tracer experiment. However, we contend this additional work is necessary to maximize the interpretability of the data and enable meaningful comparison across space and time.

4.4.2 Best Practice 2: Critical evaluation of which flowpaths may have been measured by the experiment

One persistent limitation of interpreting stream solute tracers is the inability to know which flowpaths and features were actually measured in the study reach. While additional observations in storage zones have been attempted via monitoring wells or geophysical imaging, multiple studies show that solute observed in the storage zone itself is not necessarily meaningful, as the stream breakthrough curve integrates only a sub-set of flowpaths (Ward et al., 2010a, 2017b, Toran et al., 2012, 2013). Briggs et al. (2009) suggest additional measurements in the surface storage domain may allow for parsing surface from subsurface transient storage. However, this approach relies upon measurement of a representative in-stream storage zone and interpretation via the transient storage model, which is known to be limited in identifiability of parameters and transferability to other sites (e.g., Kelleher et al., 2013; Ward et al., 2017a).

One simple approach to estimate the spatial and temporal scales of the measured flowpaths is to consider the truncation of the breakthrough curve itself. The window of detection describes the longest flowpath timescale that may have been measured. Several studies have converted this timescale to a length scale using Darcy's Law, parameterized it with representative values for hydraulic conductivity, porosity, and valley slope as a proxy for hydraulic gradient (after Ward et al., 2017b; 2018a). While imperfect, this interpretation at least indicates a spatial scale of flowpaths that may have been observed. For example, in previous studies of a small stream in the HJA basin (WS01; Fig. 1), where extensive penetration of the tracer into the subsurface was documented across a 10+ m wide valley bottom (Voltz et al., 2013; Ward et al., 2017b), the longest flowpaths detected by a tracer returning to the stream still only averaged 0.21 m (range 0.004 to 1.2 m) compared to overall reach lengths of tens of meters. This

Deleted: l
Deleted: ¶
¶ ... [80]
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Deleted: [
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted:]
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Deleted: [
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted:]
Deleted: [
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted:]
Commented [KTV49]: "flowpaths" plural.
Commented [n50R49]: corrected
Formatted: Font: (Default) Times New Roman
Deleted: [
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted:]

means that these studies were measuring in-stream storage and only the shortest and fastest subsurface flowpaths -- not integrating all the exchange in the valley bottom.

The SAS approach implemented in this study provides some valuable additional, contextual information about the storage volume and discharge that inform interpretation of findings. For example, our synoptic study labeled an average of 86% of the outflowing discharge in the surface channel (range 57% to 95%). Still, this equated to having only sampled an average of 12% of the total storage volume in the reach (range 0.3% to 35%), suggesting a bias toward in-stream storage. This bias is confirmed by the realization that, on average, only 18% of tracer mass was involved in transient storage (range 0% to 69%). Hence, the SAS approach gives us additional insights and reveals biases in the tracer methods. Altogether, this study clearly indicates that multiple data collection, analysis, and modeling techniques are needed to develop scaling relationships representative of river corridor exchange across varying hydrologic forcing and geomorphic settings.

5 Conclusions

We set out to leverage novel data sets collected across a 5th order basin to test the existence of systematic relationships linking river corridor exchange with temporal variation in discharge, spatial patterns in discharge, and local geomorphic setting. We specifically intended to use these data to critically test Wondzell's (2011) conceptual model (Fig. 6A). We found systematic patterns, namely decreases in several indicators of river corridor exchange with increasing discharge in space (i.e., moving downstream in the network), confirming this part of the Wondzell (2011) conceptual model. Wonzell's (2011) model predicts the same trend for increasing baseflow discharge in time, but we found both direct and inverse relationships between river corridor exchange and discharge at fixed reaches under varied baseflow conditions. These findings reflect a high degree of heterogeneity on a reach-to-reach basis in space, likely overwhelming or obscuring river corridor exchange patterns that might emerge in more spatially continuous and larger scale assessments, which would be a better test of the Wondzell (2011) model. Importantly, we document consistent trends with discharge that have low explanatory power (low r^2) despite being statistically significant in their direction, indicating that we have little predictive power. Moreover, our findings reveal the challenges that must be addressed to design and interpret stream solute data among sites or discharge conditions. Finally, we did not confirm Wondzell's (2011) predicted pattern with respect to local hyporheic potential at a site, which may have been confounded by integration of both surface and hyporheic storage by the stream solute tracers or by local-scale heterogeneity not captured in our reach-scale site characterization. Collectively, the larger Sen's slopes for the fixed reaches, when compared across variable hydrologic conditions, may indicate more temporal variation at a site through the season than there is through the network under a the single baseflow condition. This means that caution is needed in applying synoptic sampling approaches across time when studying river corridor exchange conditions in a river network.

Deleted: likely
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Formatted: Font: (Default) Times New Roman, Not Italic
Deleted: Fig. 5
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Deleted: [
Deleted:]
Deleted: discharge
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]
Formatted: Font: (Default) Times New Roman, Superscript
Formatted: Font: (Default) Times New Roman
Commented [KTV51]: "interpret"
Commented [n52R51]: corrected
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman, Not Italic
Formatted: Font: (Default) Times New Roman
Deleted: [
Deleted:]

This study documented the interaction between advective travel times and measurement of river corridor exchange with solute tracers. Our synoptic study design controlled for this complication by scaling study reach lengths based on wetted channel width. For future studies focused on exchange-discharge relationships, we suggest two best practices. First, controlling for advective time to measure consistent timescales of storage processes and limit artifacts that are due to limitations of solute tracer studies. Second, we suggest analyses that focus on the fractions of storage volume and outflow that were labeled with tracer to provide context for interpreting recovered timeseries. We also note that many previous studies have relied upon small sample sizes and focused on singular explanatory variables of interest considered in isolation. We suggest this is primarily descriptive, and conclude that consideration of multiple, interacting controls will be necessary to achieve predictive understanding of river corridor exchange across varying hydrologic forcing and geomorphic setting from headwaters to large river networks.

Finally, we underscore that a one-time synoptic sampling campaign does not address local-scale variability that is created by variable discharge conditions, nor does extensive study of a single reach provide data that are reflective of variation in space in the river network. In short, space-for-time and time-for-space substitutions based on the methods used in our study are not a reliable basis for transferability nor prediction.

Acknowledgements

Funding for this research was provided by the Leverhulme Trust (*Where rivers, groundwater and disciplines meet: a hyporheic research network*), the UK Natural Environment Research Council (Large woody debris – A river restoration panacea for streambed nitrate attenuation? NERC NE/L003872/1), and the European Commission supported *HiFreq: Smart high-frequency environmental sensor networks for quantifying nonlinear hydrological process dynamics across spatial scales* (project ID 734317), the US Department of Energy (DOE) Office of Biological and Environmental Research (BER) as part of Subsurface Biogeochemical Research Program's Scientific Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). Data and facilities were provided by the HJ Andrews Experimental Forest and Long Term Ecological Research program, administered cooperatively by the USDA Forest Service Pacific Northwest Research Station, Oregon State University, and the Willamette National Forest and funded, in part, by the National Science Foundation under Grant No. DEB-1440409. Ward's time in preparation of this manuscript was supported by the University of Birmingham's Institute of Advanced Studies. Additional support to individual authors is acknowledged from National Science Foundation (NSF) awards EAR 1652293, EAR 1417603, and EAR 1446328. Ward and Wondzell were also supported by Department of Energy award DE-SC0019377. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors. Finally, the authors acknowledge this would not have been possible without support from their home institutions. All data used in this study are archived in the Consortium of Universities for the

Deleted: suggest

Formatted: Font: (Default) Times New Roman

Deleted: reveal

Formatted: Font: (Default) Times New Roman

Commented [KTV53]: Are the data used in this study available in accordance with the journal's data policy?

Commented [n54R53]: We added a statement here indicating that all data used in this analysis are made publicly available.

Deleted: e

Advancement of Hydrologic Science, Inc. (CUAHSI) HydroShare data repository, publicly accessible as http://www.hydroshare.org/resource/f4484e0703f743c696c2e1f209abb842

Deleted: . In addition, timeseries for solute tracer experiments are archived.

Author Contributions

Formatted: Heading 1

All co-authors participated in the field collection, data analysis, and/or writing of this manuscript. ASW was primarily responsible for data analysis and preparation of this manuscript. ASW and JPZ conceived of the study design with input from all co-authors. The authors report no conflicts of interest.

Deleted: ¶
¶

References

Anderson, J. K., Wondzell, S. M., Gooseff, M. N. and Haggerty, R.: Patterns in stream longitudinal profiles and implications for hyporheic exchange flow at the H. J. Andrews Experimental Forest, Oregon, USA, Hydrol. Process., 19(15), 2931–2949, 2005.

Bencala, K. E. and Walters, R. A.: Simulation of solute transport in a mountain pool-and-riffle stream: a transient storage model, Water Resour. Res., 19(3), 718–724, doi:10.1029/WR019i003p00718, 1983.

Bertuzzo, E., Helton, A. M., Hall, Jr., R. O., Battin, T. J.: Scaling of dissolved organic carbon removal in river networks. *Adv. in Water Res.*, 110, 136-146, 2017.

Brardinoni, F. and Hassan, M. A.: Glacially induced organization of channel-reach morphology in mountain streams, J. Geophys. Res. Earth Surf., 112(3), 1–18, doi:10.1029/2006JF000741, 2007.

Briggs, M. A., Gooseff, M. N., Arp, C. D. and Baker, M. A.: A method for estimating surface transient storage parameters for streams with concurrent hyporheic storage., Water Resour. Res., 2009.

Brummer, C. J., and D. R. Montgomery, Downstream coarsening in headwater channels, *Water Resour. Res.,* 39, 1294, doi: 10.1029/2003WR001981, 2003.

Brunke, M. and Gonser, T.: The ecological significance of exchange processes between rivers and groundwater, Freshw. Biol., 37, 1–33, 1997.

Butturini, A. and Sabater, F.: Importance of transient storage zones for ammonium and phosphate retention in a sandy-bottom Mediterranean stream, Freshw. Biol., 41(3), 593–603, 1999.

Cashman, K. V., Deligne, N. I., Gannett, M. W., Grant, G. E. and Jefferson, A.: Fire and water: Volcanology, geomorphology, and hydrogeology of the Cascade Range, central Oregon, F. Guid., 15, 539–582, doi:10.1130/2009.fl, 2009.

Church, M. and Zimmermann, A.: Form and stability of step-pool channels: Research progress, Water Resour. Res., 43(3), 1–21, doi:10.1029/2006WR005037, 2007.

Corson-Rikert, H. A., Wondzell, S. M., Haggerty, R. and Santelmann, M. V: Carbon dynamics in the hyporheic zone of a headwater mountain streamin the Cascade Mountains, Oregon, Water Resour. Res., 52, 7556–7576, doi:10.1029/2008WR006912.M, 2016.

Covino, T., McGlynn, B., and Baker, M., Separating physical and biological nutrient retention and quantifying uptake kinetics from ambient to saturation in successive mountain stream reaches, *J. Geophys. Res.*, 115, G04010, doi:10.1029/2009JG001263, 2010.

Covino, T. P., McGlynn, B. L. and Mallard, J.: Stream-groundwater exchange and hydrologic turnover at the network scale, Water Resour. Res., 47(12), W12521, 2011.

Crook, N. *et al.* Electrical resistivity imaging of the architecture of substream sediments. *Water Resour. Res.* **44**, W00D13 (2008).

Day, T. J., Field procedures and evaluation of a slug dilution gauging method in mountain streams, J. Hydrol. N. Z., 16(2), 113–133, 1977.

Deligne, N. I., Mckay, D., Conrey, R. M., Grant, G. E., Johnson, E. R., O'Connor, J. and Sweeney, K.: Field-trip guide to mafic volcanism of the Cascade Range in Central Oregon—A volcanic, tectonic, hydrologic, and geomorphic journey, Sci. Investig. Rep., 110, doi:10.3133/sir20175022H, 2017.

Danckwerts, P.: Continuous flow systems. Distribution of residence times, *Chem. Eng. Sci.*, *2*, 1–13, 1953.

Domenico, P.A., Schwartz, F.W., Physical and Chemical Hydrogeology. John Wiley and Sons, NY. p. 824., 1990.

Drummond, J. D., Covino, T. P., Aubeneau, a. F., Leong, D., Patil, S., Schumer, R. and Packman, A. I.: Effects of solute breakthrough curve tail truncation on residence time estimates: A synthesis of solute tracer injection studies, J. Geophys. Res. Biogeosciences, 117(3), 1–11, doi:10.1029/2012JG002019, 2012.

Dudley-Southern, M. and Binley, A.: Temporal responses of groundwater-surface water exchange to successive storm events, Water Resour. Res., 51(2), 1112–1126, doi:10.1002/2014WR016623, 2015.

Dyrness, C. T.: Hydrologic properties of soils on three small watersheds in the western Cascades of Oregon, USDA For. SERV RES NOTE PNW-111, SEP 1969. 17 P., 1969.

Fabian, M. W., Endreny, T. A., Bottacin-Busolin, A. and Lautz, L. K.: Seasonal variation in cascade-driven hyporheic exchange, northern Honduras, Hydrol. Process., 25(10), 1630–1646, doi:10.1002/hyp.7924, 2011.

Gerecht, K. E., Cardenas, M. B., Guswa, A. J., Sawyer, A. H., Nowinski, J. D. and Swanson, T. E.: Dynamics of hyporheic flow and heat transport across a bed-to-bank continuum in a large regulated river, Water Resour. Res., 47(3), W03524, doi:10.1029/2010wr009794, 2011.

Gomez-Velez, J. D. and Harvey, J. W.: A hydrogeomorphic river network model predicts where and why hyporheic exchange is important in large basins, Geophys. Res. Lett., 41, 6403–6412, doi:doi:10.1002/2014GL061099, 2014.

Gomez-Velez, J. D., Harvey, J. W., Cardenas, M. B. and Kiel, B.: Denitrification in the Mississippi River network controlled by flow through river bedforms, Nat. Geosci., 8(October), 1–8, doi:10.1038/ngeo2567, 2015.

Gooseff, M. N., Anderson, J. K., Wondzell, S. M., LaNier, J. and Haggerty, R.: A modelling study of hyporheic exchange pattern and the sequence, size, and spacing of stream bedforms in mountain stream networks, Oregon, USA, Hydrol. Process., 20(11), 2443–2457, 2006.

Gooseff, M. N., Hall, R. O., Tank, J. L. and Jr, R. O. H.: Relating transient storage to channel complexity in streams of varying land use in Jackson Hole, Wyoming, Water Resour. Res., 43(1), 1–10, doi:10.1029/2005WR004626, 2007.

Gupta, A. and Cvetkovic, V.: Temporal moment analysis of tracer discharge in streams: combined effect of physicochemical mass transfer and morphology, Water Resour. Res., 36(10), 2000.

Deleted: ¶

Formatted: Indent: Left: 0", First line: 0", Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: 12 pt

Formatted: Font: 12 pt

Haggerty, R. and Reeves, P.: STAMMT-L Version 1.0 User's Manual, Sandia Natl. Lab. [ERMS# 520308], 1–76, 2002.
Hakenkamp, C. C., Valett, H. M. and Boulton, A. J.: Perspectives on the Hyporheic Zone: Integrating Hydrology and Biology. Concluding remarks, J. North Am. Benthol. Soc., 12(1), 94–99, doi:10.2307/1467683, 1993.
Harman, C. J.: Time-variable transit time distributions and transport: Theory and application to storage-dependent transport of chloride in a watershed, Water Resour. Res., 51(1), 1–30, doi:10.1002/2014WR015707.Received, 2015.
Harman, C. J., Ward, A. S. and Ball, A.: How does reach-scale stream-hyporheic transport vary with discharge? Insights from rSAS analysis of sequential tracer injections in a headwater mountain stream, Water Resour. Res., 52, 7130–7150, doi:10.1002/2016WR018832.Received, 2016.
Harms, T.K., Cook, C.L., Wlostowski, A.N., Gooseff, M.N., and Godsey, S.E. Ecosystems. Spiraling Down Hillslopes: Nutrient Uptake from Water Tracks in a Warming Arctic, Ecosystems, 15 pp, doi: 10.1007/s10021-019-00355-z, 2019.
Hart, D. R., Mulholland, P. J., Marzolf, E. R., DeAngelis, D. L. and Hendricks, S. P.: Relationships between hydraulic parameters in a small stream under varying flow and seasonal conditions, Hydrol. Process., 13(10), 1497–1510, 1999.
Harvey, J. W. and Gooseff, M. N.: River corridor science: Hydrologic exchange and ecological consequences from bedforms to basins, Water Resour. Res., Special Is, 1–30, doi:10.1002/2015WR017617.Received, 2015.
Harvey, J. W. and Wagner, B. J.: Quantifying hydrologic interactions between streams and their subsurface hyporheic zones, in Streams and Ground Waters, edited by J. B. Jones and P. J. Mulholland, pp. 3–44., 2000.
Harvey, J. W., Wagner, B. J. and Bencala, K. E.: Evaluating the reliability of the stream tracer approach to characterize stream-subsurface water exchange, Water Resour. Res., 32(8), 2441–2451, 1996.
Harvey, J. W., Gomez-Velez, J. D., Alexander, R. B., Eng, K., Moore, R., Pizzuto, J., Schmadel, N., Scott, D., Golden, H., Kettner, A., Schwarz, G., Soulsby, C., Boyer, E., Konrad, C. and Choi, J.: How Hydrologic Connectivity Regulates Water Quality in River Corridors, J. Am. Water Resour. Assoc., 1–13, doi:10.1111/1752-1688.12691, 2018.
Helton, A. M., Hall, Jr., R. O., and Bertuzzo, E.: How network structure can affect nitrogen removal by streams. *Freshwater Biology,* 63(1), 128-140. 2018.
Hynes, H. B. N.: Groundwater and stream ecology, Hydrobiologia, 100(1), 93–99, 1983.
Hvorslev, M. J.: Time lag and soil permeability in ground-water observations, Bull No 36, 50 PP. U.S. Army Waterways Experiment Station, Vicksburg, MS, 1951.
Jackson, T. R., Haggerty, R., Apte, S. V., Coleman, A. and Drost, K. J.: Defining and measuring the mean residence time of lateral surface transient storage zones in small streams, Water Resour. Res., 48(10), 1–20, doi:10.1029/2012WR012096, 2012.Jackson, T. R., Haggerty, R., Apte, S. V. and O'Connor, B. L.: A mean residence time relationship for lateral cavities in gravel-bed rivers and streams: Incorporating streambed roughness and cavity shape, Water Resour. Res., 49(6), 3642–3650, doi:10.1002/wrcr.20272, 2013.
Jefferson, A., Grant, G. E. and Lewis, S. L.: A River Runs Underneath It: Geological Control of Spring and Channel Systems and Management Implications, Cascade Range, Oregon, in Advancing the

Fundamental Sciences Proceedings of the Forest Service: Proceedings of the Forest Service National Earth Sciences Conference, vol. 1, pp. 18–22., 2004.

Jencso, K. G. and McGlynn, B. L.: Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation, Water Resour. Res., 47(11), W11527, doi:10.1029/2011wr010666, 2011.

Jin, H. S. and Ward, G. M.: Hydraulic characteristics of a small Coastal Plain stream of the southeastern United States: Effects of hydrology and season, Hydrol. Process., 19(20), 4147–4160, doi:10.1002/hyp.5878, 2005.

Swanson, F. J., Johnson, S. L., Gregory, S. V. and Acker, S. a.: Flood Disturbance in a Forested Mountain Landscape: Interactions of land use and floods, Bioscience, 48(9), 681–689, 1998.

Karwan, D. L. and Saiers, J. E.: Influences of seasonal flow regime on the fate and transport of fine particles and a dissolved solute in a New England stream, Water Resour. Res., 45(11), W11423, 2009.

Kasahara, T. and Wondzell, S. M.: Geomorphic controls on hyporheic exchange flow in mountain streams, Water Resour. Res., 39(1), 1005, 2003.

Kelleher, C. A., Wagener, T., McGlynn, B. L., Ward, A. S., Gooseff, M. N. and Payn, R. A.: Identifiability of transient storage model parameters along a mountain stream, Water Resour. Res., 49(9), 5290–5306, doi:10.1002/wrcr.20413, 2013.

Kiel, B. and Cardenas, M.: Lateral hyporheic exchange throughout the Mississippi River network, Nat. Geosci., 7(May), 413–417, doi:10.1038/ngeo2157, 2014.

Kilpatrick, F. A. and Cobb, E. D.: Measurement of discharge using tracers, U.S. Geol. Surv., Techniques of Water-Resources Investigations, Book 3, Chapter A-16, 1985.

Krause, S., Hannah, D. M., Fleckenstein, J. H., Heppell, C. M., Kaeser, D. H., Pickup, R., Pinay, G., Robertson, A. L. and Wood, P. J.: Inter-disciplinary perspectives on processes in the hyporheic zone, Ecohydrology, 4(4), 481–499, 2011.

Lee-Cullin, J. A., Zarnetske, J. P., Ruhala, S. S., and Plont, S.: Toward measuring biogeochemistry within the stream-groundwater interface at the network scale: an initial assessment of two spatial sampling strategies, *Limnology and Oceanography: Methods*, 16(11), 722-733, 2018.

Leopold, L. B. and Maddock, T.: The hydraulic geometry of stream channels and some physiographic implications, 1953.

MacDonald, L. H., Smart, A. W., and Wissmar, R. C.. Monitoring guidelines to evaluate effects of forestry activities on streams in the Pacific Northwest and Alaska, 177. http://doi.org/EPA 910/9-91-001, 1991.

Mallard, J., McGlynn, B. L. and Covino, T. P.: Lateral inflows, stream-groundwater exchange, and network geometry influence streamwater composition, Water Resour. Res., 50, 4603–4623, doi:10.1002/2013WR014222.Received, 2014.

Malzone, J. M., Lowry, C. S. and Ward, A. S.: Response of the hyporheic zone to transient groundwater fluctuations on the annual and storm event time scales, *Water Resour. Res.*, *52*, 1–20, doi:10.1002/2014WR015716, 2016.

Martin, D. J.: The influence of geomorphic factors and geographic region on large woody debris loading and fish habitat in Alaska coastal streams, North Am. J. Fish. Manag., 21(3), 429–440, doi:10.1577/1548-8675(2001)021<0429:TIOGFA>2.0.CO;2, 2001.

Deleted: ¶

Formatted: Font: 12 pt

Formatted: Font:

Mason, S. J. K., McGlynn, B. L. and Poole, G. C.: Hydrologic response to channel reconfiguration on Silver Bow Creek, Montana, J. Hydrol., 438–439, 125–136, 2012.
Montgomery, D. R. and Buffington, J. M.: Channel-Reach Morphology in Mountain Basins, GSA Bull., 109(5), 596–611, 1997.
Montgomery, D. R. and Buffington, J. M.: Chapter 2: Channel processes, classification, and response, in River Ecology and Management - Lessons from the Pacific Coastal Ecoregion, edited by R. Naiman and R. Bilby, pp. 13–42, Springer Verlag, New York, NY., 1998.
Morrice, J. A., Valett, H. M., Dahm, C. N. and Campana, M. E.: Alluvial characteristics, groundwater-surface water exchange and hydrological retention in headwater streams, Hydrol. Process., 11(3), 253–267, 1997.
Musial, C., Sawyer, A. H., Barnes, R., Bray, S. and Knights, D.: Surface water–groundwater exchange dynamics in a tidal freshwater zone, Hydrol. Process., 30, 739–750, doi:10.1002/hyp.10623, 2016.
Palmer, M. A.: Experimentation in the Hyporheic Zone: Challenges and Prospectus, J. North Am. Benthol. Soc., 12(1), 84–93, 1993.
Patil, S., Covino, T. P., Packman, A. I., McGlynn, B. L., Drummond, J. D., Payn, R. A. and Schumer, R.: Intrastream variability in solute transport: Hydrologic and geomorphic controls on solute retention, J. Geophys. Res. Earth Surf., 118(2), 413–422, doi:10.1029/2012JF002455, 2013.
Payn, R. A., Gooseff, M. N., McGlynn, B. L., Bencala, K. E. and Wondzell, S. M.: Channel water balance and exchange with subsurface flow along a mountain headwater stream in Montana, United States, Water Resour. Res., 45, 2009.
Payn, R. A., Gooseff, M. N., McGlynn, B. L., Bencala, K. E. and Wondzell, S. M.: Exploring changes in the spatial distribution of stream baseflow generation during a seasonal recession, Water Resour. Res., 48(4), W04519, 2012.
Rana, S. M. M., Boccelli, D. L., Scott, D. T. and Hester, E. T.: Parameter uncertainty with flow variation of the one-dimensional solute transport model for small streams using Markov chain Monte Carlo, J. Hydrol., 575(June), 1145–1154, doi:10.1016/j.jhydrol.2019.06.003, 2019.
Rot, B. W., Naiman, R. J. and Bilby, R. E.: Stream channel configuration, landform, and riparian forest structure in the Cascade Mountains, Washington, Can. J. Fish. Aquat. Sci., 57(4), 699–707, doi:10.1139/f00-002, 2000.
Ruhala S. S., Zarnetske, J.P., Long, D.T., Lee-Cullin, J. A., Plont, S. and Wiewiora, E. R.: Exploring dissolved organic carbon cycling at the stream-groundwater interface across a third-order, lowland stream network. *Biogeochemistry*, 137(1-2), 105-126, 2017.
Sawyer, A. H., Cardenas, M. B., Bomar, A. and Mackey, M.: Impact of dam operations on hyporheic exchange in the riparian zone of a regulated river, Hydrol. Process., 23(15), 2129–2137, 2009.
Schmadel, N. M., Neilson, B. T. and Stevens, D. K.: Approaches to estimate uncertainty in longitudinal channel water balances, J. Hydrol., 394(3–4), 357–369 (online) Available from: http://www.sciencedirect.com/science/article/B6V6C-513F93T-1/2/58bcbaea6ebd00023f3c25bb2bac75c4, 2010.
Schmadel, N. M., Neilson, B. T., Heavilin, J., Stevens, D. K. and Worman, A.: The influence of spatially variable streamhydraulics on reach scale transient storage modeling, Water Resour. Res., 50, 9287–9299, doi:10.1002/2015WR017273.Received, 2014.

Deleted: [
Deleted:]

Schmadel, N. M., Ward, A. S., Kurz, M. J., Fleckenstein, J. H., Zarnetske, J. P., Knapp, J. L. A., Klaar, M. J., Romeijn, P., Datry, T., Keller, T. and Folegot, S.: Stream solute tracer timescales changing with discharge and reach length confound process interpretation, Water Resour. Res., 52, 3227–3245, doi:10.1002/2015WR018062.Received, 2016a.
Schmadel, N. M., Ward, A. S., Lowry, C. S. and Malzone, J. M.: Hyporheic exchange controlled by dynamic hydrologic boundary conditions, Geophys. Res. Lett., 43, 4408–4417, doi:10.1002/2016GL068286.Received, 2016.
Schmadel, N. M., Ward, A. S. and Wondzell, S. M.: Hydrologic controls on hyporheic exchange in a headwater mountain stream, Water Resour. Res., doi:10.1002/2017WR020576, 2017.
Schmid, B. H.: Can Longitudinal Solute Transport Parameters Be Transferred to Different Flow Rates?, J. Hydrol. Eng., 13(6), 505–509, 2008.
Schwanghart, W. and Kuhn, N. J.: TopoToolbox: A set of Matlab functions for topographic analysis, Environ. Model. Softw., 25(6), 770–781, doi:10.1016/j.envsoft.2009.12.002, 2010.
Schwanghart, W. and Scherler, D.: Short Communication: TopoToolbox 2 - MATLAB-based software for topographic analysis and modeling in Earth surface sciences, Earth Surf. Dyn., 2(1), 1–7, doi:10.5194/esurf-2-1-2014, 2014.
Stewart, R. J., Wollheim, W. M., Gooseff, M. N., Briggs, M. a., Jacobs, J. M., Peterson, B. J. and Hopkinson, C. S.: Separation of river network–scale nitrogen removal among the main channel and two transient storage compartments, Water Resour. Res., 47(1), W00J10, 2011.
Storey, R. G., Howard, K. W. F. and Williams, D. D.: Factors controlling riffle-scale hyporheic exchange flows and their seasonal changes in a gaining stream: A three-dimensional groundwater flow model, Water Resour. Res., 39(2), 1034, 2003.
Swanson, F. J. and James, M. E.: Geology and geomorphology of the H.J. Andrews Experimental Forest, western Cascades, Oregon, Portland, OR., 1975.
Swanson, F. J. and Jones, J. A.: Geomorphology and hydrology of the H.J. Andrews Experimental Forest, Blue River, Oregon, in Field guide to geologic processes in Cascadia., 2002.
Toran, L., Hughes, B., Nyquist, J. E. and Ryan, R. J.: Freeze Core Sampling to Validate Time-Lapse Resistivity Monitoring of the Hyporheic Zone., Ground Water, 1–6, doi:10.1111/j.1745-6584.2012.01002.x, 2012.
Toran, L., Nyquist, J. E., Fang, A. C., Ryan, R. J. and Rosenberry, D. O.: Observing lingering hyporheic storage using electrical resistivity: Variations around stream restoration structures, Crabby Creek, PA., 2013.
Valett, H. M., Morrice, J. A., Dahm, C. N. and Campana, M. E.: Parent lithology, surface-groundwater exchange, and nitrate retention in headwater streams, Limnol. Oceanogr., 333–345, 1996.
Vervier, P., Gibert, J., Marmonier, P., Journal, S., American, N., Society, B., Mar, N. and Dole-Olivier, M. J.: A perspective on the permeability of the surface freshwater-groundwater ecotone, J. North Am. Benthol. Soc., 11(1), 93–102, 1992.
Voltz, T. J., Gooseff, M. N., Ward, A. S., Singha, K., Fitzgerald, M. and Wagener, T.: Riparian hydraulic gradient and stream-groundwater exchange dynamics in steep headwater valleys, J. Geophys. Res. Earth Surf., 118(2), 953–969, doi:10.1002/jgrf.20074, 2013.
Wagner, B. J. and Harvey, J. W.: Experimental design for estimating parameters of rate-limited mass transfer: Analysis of stream tracer studies, Water Resour. Res., 33(7), 1731–1741, 1997.

Ward, A. S.: The evolution and state of interdisciplinary hyporheic research, Wiley Interdiscip. Rev. Water, 3(1), 83–103, doi:10.1002/wat2.1120, 2015.

Ward, A. S. and A.I. Packman: Advancing our predictive understanding of river corridor exchange. *WIREs-Water*, 6, 2019.

Ward, A. S., Gooseff, M. N. and Singha, K.: Imaging hyporheic zone solute transport using electrical resistivity, Hydrol. Process., 24(7), doi:10.1002/hyp.7672, 2010a.

Ward, A. S., Gooseff, M. N. and Singha, K.: Characterizing hyporheic transport processes-Interpretation of electrical geophysical data in coupled stream-hyporheic zone systems during solute tracer studies, Adv. Water Resour., 33, 1320–1330, doi:doi:10.1016/j.advwatres.2010.05.008, 2010b.

Deleted: ¶

Ward, A. S., Fitzgerald, M., Gooseff, M. N., Voltz, T. J., Binley, A. M. and Singha, K.: Hydrologic and geomorphic controls on hyporheic exchange during base flow recession in a headwater mountain stream, Water Resour. Res., 48(4), W04513, 2012.

Ward, A. S., Gooseff, M. N., Voltz, T. J., Fitzgerald, M., Singha, K. and Zarnetske, J. P.: How does rapidly changing discharge during storm events affect transient storage and channel water balance in a headwater mountain stream?, Water Resour. Res., 49(9), 5473–5486, doi:10.1002/wrcr.20434, 2013a.

Ward, A. S., Gooseff, M. N. and Singha, K.: How does subsurface characterization affect simulations of hyporheic exchange?, Ground Water, 51(1), 14–28, doi:10.1111/j.1745-6584.2012.00911.x, 2013b.

Ward, A. S., Payn, R. A., Gooseff, M. N., McGlynn, B. L., Bencala, K. E., Kelleher, C. A., Wondzell, S. M. and Wagener, T.: Variations in surface water-ground water interactions along a headwater mountain stream: Comparisons between transient storage and water balance analyses, Water Resour. Res., 49(6), 3359–3374, doi:10.1002/wrcr.20148, 2013.

Ward, A. S., Kelleher, C. A., Mason, S. J. K., Wagener, T., McIntyre, N., McGlynn, B., Runkel, R. L. and Payn, R. A.: A software tool to assess uncertainty in transient-storage model parameters using Monte Carlo simulations, Freshw. Sci., 36(1), doi:10.1086/690444, 2017a.

Ward, A. S., Schmadel, N. M., Wondzell, S. M., Gooseff, M. N. and Singha, K.: Dynamic hyporheic and riparian flow path geometry through base flow recession in two headwater mountain streamcorridors, Water Resour. Res., 53, doi:10.1002/2016WR019875, 2017b.

Ward, A. S., Schmadel, N. M. and Wondzell, S. M.: Simulation of dynamic expansion, contraction, and connectivity in a mountain stream network, Adv. Water Resour., 114, 64–82, doi:10.1016/j.advwatres.2018.01.018, 2018a.

Ward, A. S., Morgan, J. A., White, J. R. and Royer, T. V.: Streambed restoration to remove fine sediment alters reach-scale transient storage in a low-gradient 5 th order river, Indiana, USA, Hydrol. Process., 1–15, doi:10.1002/hyp.11518, 2018b.

Ward, A. S., Schmadel, N. M. and Wondzell, S. M.: Time-Variable Transit Time Distributions in the Hyporheic Zone of a Headwater Mountain Stream, Water Resour. Res., doi:10.1002/2017WR021502, 2018c.

Deleted: ¶

Ward, A. S., Zarnetske, J. P., Baranov, V., Blaen, P. J., Brekenfeld, N., Chu, R., Derelle, R., Drummond, J., Fleckenstein, J., Garayburu-Caruso, V., Graham, E., Hannah, D., Harman, C., Hixson, J., Knapp, J. L. A., Krause, S., Kurz, M. J., Lewandowski, J., Li, A., Martí, E., Miller, M., Milner, A. M., Neil, K., Orsini, L., Packman, A. I., Plont, S., Renteria, L., Roche, K., Royer, T., Schmadel, N. M., Segura, C., Stegen, J., Toyoda, J., Wells, J., Wisnoski, N. I., and Wondzell, S. M.: Co-located contemporaneous mapping of morphological, hydrological, chemical, and biological conditions in a 5th

order mountain stream network, Oregon, USA, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2019-45, in review, 2019.
White, D. S., Journal, S., American, N. and Society, B.: Perspectives on defining and delineating hyporheic zones, J. North Am. Benthol. Soc., 12(1), 61–69, 1993.
Wlostowski, A. N., Gooseff, M. N., Bowden, W. B. and Wollheim, W. M.: Stream tracer breakthrough curve decomposition into mass fractions: A simple framework to analyze and compare conservative solute transport processes, Limnol. Oceanogr. Methods, 15(2), 140–153, doi:10.1002/lom3.10148, 2017.
Wohl, E. and Merritt, D.: Prediction of mountain stream morphology, Water Resour. Res., 41(8), 1–10, doi:10.1029/2004WR003779, 2005.
Wohl, E. and Merritt, D. M.: Reach-scale channel geometry of mountain streams, Geomorphology, 93(3–4), 168–185, doi:10.1016/j.geomorph.2007.02.014, 2008.
Wondzell, S. M.. Flux of Ground Water and Nitrogen Through the Floodplain of a Fourth-Order Stream. Ph.D. Dissertation. Oregon State University. 113 pp., 1994.
Wondzell, S. M.: Effect of morphology and discharge on hyporheic exchange flows in two small streams in the Cascade Mountains of Oregon, USA, Hydrol. Process., 20(2), 267–287, 2006.
Wondzell, S. M.: The role of the hyporheic zone across stream networks, Hydrol. Process., 25(22), 3525–3532, doi:10.1002/hyp.8119, 2011.
Wondzell, S. M. and Gooseff, M. N.: Geomorphic Controls on Hyporheic Exchange Across Scales: Watersheds to Particles, in Treatise on Geomorphology, vol. 9, edited by J. Schroder and E. Wohl, pp. 203–218, Academic Press, San Diego, CA., 2014.
Zarnetske, J. P., Gooseff, M. N., Brosten, T. R., Bradford, J. H., McNamara, J. P. and Bowden, W. B.: Transient storage as a function of geomorphology, discharge, and permafrost active layer conditions in Arctic tundra streams, Water Resour. Res., 43(7), W07410, 2007.

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [1]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [2]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [3]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [4]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [5]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [6]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [7]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [8]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [9]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [10]

Page 3: [1] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [11]

Page 3: [2] Deleted **Ward, Adam Scott** **8/17/19 1:12:00 PM**

... [12]

Page 3: [2] Deleted **Ward, Adam Scott** **8/17/19 1:12:00 PM**

... [13]

Page 3: [2] Deleted **Ward, Adam Scott** **8/17/19 1:12:00 PM**

... [14]

Page 3: [2] Deleted **Ward, Adam Scott** **8/17/19 1:12:00 PM**

... [15]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [20]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [21]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [22]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [23]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [24]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [25]

Page 3: [2] Deleted Ward, Adam Scott 8/17/19 1:12:00 PM

... [26]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [27]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [28]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [29]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [30]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [31]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [32]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [33]

Page 3: [3] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

... [34]

Yes, injections were instantaneous. Injection were not distributed across the channel. We selected appropriate mixing lengths to ensure the tracer was well mixed both laterally and vertically at the measurement sites.

While plunging is always a concern with using salt tracers, all locations were high-gradient, fast moving water with no deep, slowing moving pools, thus limiting potential concerns of plunging.

Page 12: [7] Deleted **Ward, Adam Scott** **8/17/19 12:38:00 PM**

Page 13: [8] Commented [n24R23] **nschmadel** **8/1/19 4:58:00 PM**

Yes, we absolutely agree. Any storage flowpath that is longer in time than the window of detection is not measured with this tracer method. This limitation has led to us conducting the SAS analysis and offering best practices for tracer experiments later in the manuscript. We added "short-term" of "recovered" mass for clarity.

Page 13: [9] Commented [KTV25] **King, Tyler Victor** **7/25/19 10:01:00 AM**

Variables were *italicized* in previous paragraphs. Suggest making formatting consistent throughout.

Page 13: [10] Commented [n26R25] **nschmadel** **8/1/19 4:56:00 PM**

Thanks again for your detailed attention. We will format all variables for consistency before this revision is resubmitted to the journal.

Page 13: [11] Commented [KTV27] **King, Tyler Victor** **7/25/19 9:58:00 AM**

Should this be C_{ADE}?

Same with the second row in this equation. Otherwise, C_{AD} would be a function of itself.

Page 13: [12] Commented [n28R27] **nschmadel** **8/1/19 4:57:00 PM**

You are correct. This has been corrected.

Page 13: [13] Deleted **Ward, Adam Scott** **8/16/19 10:17:00 PM**

Page 28: [14] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [14] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [14] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [14] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [14] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [14] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [40]
... [41]
... [42]
... [43]
... [44]
... [45]
... [46]
... [47]
... [48]
... [49]
... [50]
... [51]
... [52]
... [53]

Font: (Default) Times New Roman, Not Italic

Page 28: [15] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [15] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [16] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [16] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [17] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [17] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [18] Deleted **Ward, Adam Scott** **8/19/19 8:46:00 AM**

Page 28: [18] Deleted **Ward, Adam Scott** **8/19/19 8:46:00 AM**

Page 28: [19] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [19] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [20] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [20] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [20] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [21] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [21] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [58]
... [59]
... [60]
... [61]
... [62]
... [63]
... [64]
... [65]
... [66]
... [67]
... [68]
... [69]
... [70]
... [71]
... [72]
... [73]
... [74]
... [75]
... [76]
... [77]
... [78]
... [79]

... [87]
... [88]

Page 28: [23] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [24] Deleted **Ward, Adam Scott** **8/19/19 8:47:00 AM**

... [89]
... [90]

Page 28: [24] Deleted **Ward, Adam Scott** **8/19/19 8:47:00 AM**

... [91]
... [92]

Page 28: [25] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [25] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [93]
... [94]

Page 28: [26] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [95]
... [96]

Page 28: [26] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [97]
... [98]

Page 28: [27] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 28: [27] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [99]
... [100]
... [101]

Page 28: [28] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [102]

Page 28: [28] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [103]
... [104]

Page 28: [29] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [105]
... [106]

Page 28: [29] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [107]
... [108]

Page 28: [30] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 28: [30] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [109]
... [110]

Page 30: [33] Deleted	Ward, Adam Scott	8/16/19 11:04:00 PM

Page 30: [33] Deleted	Ward, Adam Scott	8/16/19 11:04:00 PM

Page 30: [33] Deleted	Ward, Adam Scott	8/16/19 11:04:00 PM

Page 30: [33] Deleted	Ward, Adam Scott	8/16/19 11:04:00 PM

Page 30: [33] Deleted	Ward, Adam Scott	8/16/19 11:04:00 PM

Page 30: [34] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 30: [34] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 30: [34] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 30: [34] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 30: [35] Deleted	Ward, Adam Scott	8/16/19 2:47:00 PM

Page 30: [35] Deleted	Ward, Adam Scott	8/16/19 2:47:00 PM

Page 30: [36] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 30: [36] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 30: [36] Formatted	Ward, Adam Scott	8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

... [117]
... [118]
... [119]
... [120]
... [121]
... [122]
... [123]
... [124]
... [125]
... [126]
... [127]
... [128]
... [129]
... [130]
... [131]
... [132]
... [133]
... [134]
... [135]
... [136]
... [137]
... [138]

Font: (Default) Times New Roman, Not Italic

Page 30: [38] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [145]

Page 30: [38] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [146]

Page 30: [38] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [147]

Page 30: [38] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [148]

Page 30: [38] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [149]

Page 31: [39] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [150]

Page 31: [39] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [151]

Page 31: [39] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [152]

Page 31: [40] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [153]

Page 31: [40] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [154]

Page 31: [40] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [155]

Page 31: [40] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [156]

Page 31: [41] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [157]

... [158]

Page 31: [41] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

... [159]

Page 31: [42] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [160]

Page 31: [43] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

Page 31: [44] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [44] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [44] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [44] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [44] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [44] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [45] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

Page 31: [45] Deleted Ward, Adam Scott 8/16/19 2:47:00 PM

Page 31: [46] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted Ward, Adam Scott 8/16/19 10:59:00 PM

Font: (Default) Times New Roman, Not Italic

... [166]
... [167]
... [168]
... [169]
... [170]
... [171]
... [172]
... [173]
... [174]
... [175]
... [176]
... [177]
... [178]
... [179]
... [180]
... [181]
... [182]

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 31: [46] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 31: [47] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 31: [47] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 31: [48] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 31: [48] Deleted **Ward, Adam Scott** **8/16/19 2:47:00 PM**

Page 31: [49] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 31: [49] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 32: [50] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 32: [51] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [52] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 32: [53] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [54] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [55] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [56] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

... [189]
... [190]
... [191]
... [192]
... [193]
... [194]
... [195]
... [196]
... [197]
... [198]
... [199]
... [200]
... [201]
... [202]
... [203]
... [204]

Font: (Default) Times New Roman, Not Italic

Page 32: [61] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [62] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [63] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 32: [64] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [65] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic

Page 32: [66] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [67] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman

Page 32: [68] Deleted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Page 32: [69] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Highlight

Page 32: [70] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic, Highlight

Page 32: [71] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Highlight

Page 32: [72] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Not Italic, Highlight

Page 32: [73] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Highlight

Page 32: [74] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Highlight

Page 32: [75] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

Font: (Default) Times New Roman, Highlight

Page 32: [76] Formatted **Ward, Adam Scott** **8/16/19 10:59:00 PM**

... [210]
... [211]
... [212]
... [213]
... [214]
... [215]
... [216]
... [217]
... [218]
... [219]
... [220]
... [221]
... [222]
... [223]
... [224]
... [225]
... [226]
... [227]
... [228]
... [229]
... [230]
... [231]

Page 33: [80] Deleted **Ward, Adam Scott** **8/16/19 11:05:00 PM**

... [238]

... [239]

... [240]

... [241]

... [242]