Stream solute tracer timescales changing with discharge and reach length confound process interpretation
Schmadel, Noah M.; Hannah, David; Blaen, Phillip; Klaar, Megan; Romeijn, Paul; Folegot, Silvia; Krause, Stefan

DOI: 10.1002/2015WR018062

License: None: All rights reserved

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 03/06/2016. Copyright: 2016. American Geophysical Union. All Rights Reserved.

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 15. Sep. 2019
Stream solute tracer timescales changing with discharge and reach length confound process interpretation


1School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana, USA, 2Department of Hydrogeology, Helmholtz Centre for Environmental Research-UFZ, Leipzig, Germany, 3Department of Geological Sciences, Michigan State University, East Lansing, Michigan, USA, 4School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, Edgbaston, UK, 5GFZ German Research Centre for Geosciences, Helmholtz Centre Potsdam, Telegrafenberg, Potsdam, Germany, 6Center for Applied Geoscience, University of Tübingen, Tübingen, Germany, 7IRSTEA, UR-MALY, Centre de Lyon-Villeurbanne, Villeurbanne, France, 8Department of Ecohydrology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

Abstract Improved understanding of stream solute transport requires meaningful comparison of processes across a wide range of discharge conditions and spatial scales. At reach scales where solute tracer tests are commonly used to assess transport behavior, such comparison is still confounded due to the challenge of separating dispersive and transient storage processes from the influence of the advective timescale that varies with discharge and reach length. To better resolve interpretation of these processes from field-based tracer observations, we conducted recurrent conservative solute tracer tests along a 1 km study reach during a storm discharge period and further discretized the study reach into six segments of similar length but different channel morphologies. The resulting suite of data, spanning an order of magnitude in advective timescales, enabled us to (1) characterize relationships between tracer response and discharge in individual segments and (2) determine how combining the segments into longer reaches influences interpretation of dispersion and transient storage from tracer tests. We found that the advective timescale was the primary control on the shape of the observed tracer response. Most segments responded similarly to discharge, implying that the influence of morphologic heterogeneity was muted relative to advection. Comparison of tracer data across combined segments demonstrated that increased advective timescales could be misinterpreted as a change in dispersion or transient storage. Taken together, our results stress the importance of characterizing the influence of changing advective timescales on solute tracer responses before such reach-scale observations can be used to infer solute transport at larger network scales.

1. Introduction

Meaningful comparison of stream solute transport processes across different discharge conditions and spatial scales (e.g., morphologic unit to reach to network) is necessary to accurately represent and predict solute transport through stream networks [e.g., Covino et al., 2011; Kelleher et al., 2013]. Such comparison remains a persistent methodological and conceptual problem due to the uncertain distinction between spatial variability of solute transport processes, such as dispersion and transient storage, and changes caused by varying discharge conditions [e.g., González-Pinzón et al., 2013]. At the reach scale, on the order of tens to thousands of meters, solute transport processes are commonly interpreted through stream solute tracer tests [e.g., Stream Solute Workshop, 1990]. Direct comparison of observed tracer responses, however, is impossible without separating the spatially variable solute transport process from the unique timescale (commonly transit time distribution) of each test [Harvey et al., 1996]. Unfortunately, this separation is not straightforward because the tracer response changes primarily as a function of the downstream advective transport time (commonly modal transport time, hereafter advective time), which shifts with discharge and reach length selection [e.g., Ward et al., 2013a].
Traditionally, comparison of solute tracer responses is performed by standardizing reach lengths through dimensionless numbers that relate physical processes to advective times \citep{runkel2002wagner, wagner1997}. This standardization is thought to yield the appropriate “window of detection” inherent in tracer studies (i.e., the time from tracer first arrival to last detection) \citep{harvey2000wagner} and allow for assessment of rapid exchanges between the stream and connected subsurface—the solute transport process often deemed most important to stream ecosystem functions \citep[e.g.,][]{boulton2010hetero, hester2010gooseff}. Subsurface and surface tracer exchange flow paths that return to the stream within the window of detection, but have residence times slower than the advective time, are defined as short-term storage (commonly “transient storage”) \citep[e.g.,][]{harvey1996}. Conversely, tracer flow paths that return outside of the window of detection (i.e., not recovered) are considered long-term storage and reflected in a tracer-based channel water balance \citep{payn2009schmadel, payn2014}. However, the window of detection changes with discharge and, consequently, defines an arbitrary boundary between short and long-term storage regardless of the actual physical flow paths \citep[e.g.,][]{ward2013, wondzell2006}. The window of detection may also be influenced by the tracer type and associated resolution of observations. For example, some tracers can provide more late-time tailing information than others due to differences in detection limit sensitivity \citep[e.g., fluorescent tracers in comparison to salt tracers][]{drummond2012}. While the tracer selection and interaction between the advective time and window of detection can influence the interpretation of short-term storage, the storage flow paths themselves can change with discharge \citep{wondzell2011, zarnetske2007}, further complicating meaningful comparison of tracer responses and thus impeding accurate conceptualization of solute transport processes.

Recent studies have shown, with the availability of high-frequency discharge observations, that analyzing solute transport processes across different discharge conditions is essential to improve conceptualization of how transient storage and dispersion change or compete with advection \citep{dudley2015, ward2013, zimmer2014}. A more refined conceptual understanding of stream solute transport across a range of discharges can better facilitate the upscaling of reach-scale observations to infer solute transport at larger network scales. Because solute transport processes are dynamic and spatially variable along a stream network, current upscaling strategies use relationships between replicate reach-scale processes \citep[e.g.,][]{covino2011, stewart2011}. While there are reported relationships between solute transport processes and discharge—both for long-term \citep{covino2011, ward2013, ward2013b} and short-term storage \citep{schmid2010, ward2013a, ward2013b}—varying advective times reflected in reach-scale observations can lead to misinterpreting a change in discharge as an apparent change in transient storage or dispersion \citep[e.g.,][]{gooseff2007}. For example, when fixing study reach lengths throughout a watershed \citep[e.g., standard 100 m or 200 m reaches][]{payn2009, ward2013b}, headwater reaches with lower discharge typically have larger windows of detection and advective times compared to higher discharge, downstream reaches.

An apparent process relationship with discharge may be incorrectly inferred because a common assumption is that tracer studies across fixed reach lengths are directly comparable \citep{covino2011, payn2009, ward2013b}, despite preexisting knowledge that different discharges in those fixed reaches will yield different windows of detection. The applicability of this assumption is uncertain as few studies have compared stream solute tracer studies across different discharges to characterize the influence of changing advective times and windows of detection on interpreting the processes of interest.

In this study, we directly examine the extent to which solute transport process information assessed from observed conservative solute tracer responses is comparable between different advective times and windows of detection. The objectives of this study are (1) to characterize relationships between physical solute transport processes and discharge in reach segments with distinct channel morphologies, and (2) to determine how study reach length influences the interpretation of these processes. To achieve these objectives, we conducted seven conservative solute tracer injections during a storm discharge period with in-stream responses recorded at seven downstream locations—enabling the comparison of up to 21 reach segments of different spatial scales, windows of detection, and advective times for each injection. We coinjected salt and fluorescein dye tracers for each injection, which also allowed us to test how tracer type and the associated data resolution influences the interpretation of solute transport.
2. Background on the Challenge of Interpreting Physical Solute Transport Processes From Stream Solute Tracers

The stream solute transport timescale, defined here as the advective time and window of detection, can be directly estimated from an observed in-stream conservative solute tracer time series (hereafter breakthrough curve, or BTC, Figure 1a). Following an instantaneous tracer injection, the elapsed time from injection to peak concentration is commonly interpreted as the downstream advective time ($t_{ad}$) [e.g., Haggerty et al., 2002]. The window of detection ($t_{w}$) of an observed BTC is often quantified as the elapsed time from tracer first arrival ($t_{1}$) to 99% of recovered signal above background ($t_{b99}$) [e.g., Mason et al., 2012; Ward et al., 2013a]. The elapsed time from $t_{ad}$ to $t_{b99}$ describes the persistence of tailing in an observed BTC, which is an indicator of transient storage (hereafter the transient storage index, or TSI) [Mason et al., 2012].

The unique shape of a BTC is controlled by complex interactions of solute transport processes. For example, dispersive processes that contribute to spreading of the BTC include molecular diffusion and turbulent mixing [e.g., Fischer, 1975]. However, a BTC typically provides a reach-average representation of solute transport processes. Therefore, the processes most commonly interpreted from the shape of a BTC are downstream advection, longitudinal dispersion, and transient storage [e.g., Ward et al., 2013a].

The extent to which downstream advection or processes of interest (i.e., dispersion and transient storage) control the shape of the BTC remains unclear as different combinations can result in similar shapes. At one extreme, if dispersion and transient storage are constant between high and low discharges in the same reach, differences between the BTCs are controlled solely by differences in advective times (Figure 1b). For example, during lower discharges that result in smaller advective velocities, there will be more time for dispersion and transient storage to act on the tracer. At the other extreme, if advective times are constant (e.g., discharges are equivalent) between two different reaches, variation in the shape of the BTC and, therefore, the transport timescale is controlled by differences in dispersion and transient storage (Figure 1d). In practice, in-stream tracer observations reflect interacting advective times with dispersive and transient storage processes, which can yield BTCs with similar shapes for different reasons (Figure 1c). Therefore, we anticipate that an
understanding of the proportional controls of advective timescales relative to dispersion and transient storage is necessary to compare stream solute tracer studies across discharges, reaches, and spatial scales.

3. Methods

3.1. Site Description and Discharge Measurements

The Selke River originates in the Harz Mountains in central Germany, flows through steep gradient, deeply incised valleys, and transitions to lower gradients underlain by alluvial deposits up to 10 m thick [Trauth et al., 2015] (Figure 2a). The study reach is a ~1 km section situated immediately downstream of this transition in largely agricultural lowlands (51°43’21.4”N, 11°18’17.6”E). Continuous stream discharge was generated from stage—measured at 10 min intervals (LTC Levellogger Junior M10, Solinst, Georgetown, Canada, accuracy ±/− 0.1% of reading) about 15 m upstream of site E (Figure 2b)—applied to a site-specific power function rating curve developed by researchers at Helmholtz Centre for Environmental Research. The confidence bounds of this rating curve were produced following Schmadel et al. [2010] to provide the 95% confidence intervals of the discharge estimates (see supporting information Figure S1 for details). Discharge within the study reach increased from 0.27 to 2.35 m$^3$ s$^{-1}$ with stage remaining below bankfull during the July 2014 storm event used in this study (Figure 2c). The reach-specific rating curve was corroborated with continuous stream discharge recorded 4.6 km upstream at the Meisdorf federal gaging station. This station reports that discharge can range from 0.2 to 16 m$^3$ s$^{-1}$ during storm events and seasonal snowmelt, and the long-term (1921 to 2014) annual mean discharge is 1.5 m$^3$ s$^{-1}$.

Figure 2. (a) The Selke River catchment located in central Germany and the location of the study reach and Meisdorf federal gaging station. (b) In-stream monitoring sites and general channel morphologic characteristics of the study reach. (c) Stream discharge estimated within the study reach from an established rating curve (15 m upstream of site E) and the associated 95% confidence bounds (CB), timing of tracer test durations (elapsed time from injection to last detection at site G), and discharge at the Meisdorf gaging station located ~4.6 km upstream of the injection site. (d) The monitoring sites delineate six reach segments, which were recombined into every possible unique combination (grouped by the number of individual segments) to produce segments of various lengths and transport timescales.
The study reach has general channel morphologic characteristics of meandering glides and shallow riffles, pool-riffle sequences, point bars, and in-stream gravel bars (Figure 2b). The streambed primarily consists of medium sand to coarse gravel. Visually different morphologic characteristics are expected to manifest as distinct solute transport controls across changing discharge conditions [e.g., Gostner et al., 2013]. Therefore, seven in-stream monitoring sites were selected within the study reach to bracket sections of visually distinct morphologies while preserving similar lengths. Labeled moving downstream from A to G (Figure 2b), these sites delineate the six individual reach segments (100–152 m). The three upstream-most segments, AB, BC, and CD, consist primarily of meandering glides connected by shallow riffles visible under low-flow conditions. Moving downstream, the morphology becomes more complex: segment DE contains one in-stream bar, two point bars, and a pronounced pool-riffle sequence; segment EF contains a pool-riffle sequence at the downstream end; and segment FG contains two pool-riffle sequences and a large in-stream and point bar. The segment-wise average streamed slopes are about 0.7, 0.4, 0.1, 0.4, 0.4, and 0.8% for segments AB, BC, CD, DE, EF, and FG, respectively, estimated from a 1 m digital elevation model acquired from the Land Survey Administration Saxony-Anhalt dated April 2009. The reach average streamed slope is 0.5% and active channel width is 9 m.

The individual segments can be combined into a total of 21 unique segments of different lengths (Figure 2d). These 21 combinations were used to test if combining shorter segments into longer reaches integrates differences in processes or only generates apparent differences due to changing advective timescales and windows of detection. Similar to the varying reach length analysis of Gooseff et al. [2013], we recombined individual segments into every possible unique combination based on available tracer data.

3.2. In-Stream Solute Tracer Observations

Seven instantaneous conservative solute tracer injections were performed over the storm discharge period (7–11 July 2014). For all injections, dissolved salt (NaCl) and fluorescein dye tracers were coinjected at the same location and measured at sites A through G, resulting in 49 observed BTCs each. In-stream responses of the salt tracer were observed at 30 s intervals (CTD diver 10 m, Schlumberger Water Services, Delft, Netherlands, accuracy 1% of reading) using fluid specific conductance as a surrogate for salt concentration. Background-specific conductance was corrected to zero. In some cases, the background-specific conductance drifted during the measurement period, which was observed by an additional sensor placed about 10 m upstream of the injection site. This background signal was shifted downstream based on the advective time observed from the corresponding BTC and subtracted to correct for background (hereafter salt BTC). Because instantaneous and constant rate tracer injection techniques are expected to provide similar transit time distributions [Payn et al., 2008], instantaneous tracer injections were chosen over constant rate injections to minimize the influence of dynamic discharge conditions on measurements, similar to Ward et al. [2013a]. In-stream responses of the fluorescein dye tracer (hereafter fluorescein BTC) were measured at 10 s intervals using flow-through fluorometers (GGUN-FL30, Albillia Sarl, Switzerland) at the same monitoring sites. No background correction was performed for the fluorescein BTCs because the instruments were calibrated in the field with stream water to reduce background interference and no dye was present prior to injection.

The observed tracer BTC from these types of experiments is an incomplete temporal signal due to detection limits and signal to noise interferences at late times [e.g., Drummond et al., 2012], but can provide meaningful dispersion and short-term storage information [e.g., Bellin et al., 2015; González-Pinzón et al., 2013]. In this study, we truncated each observed BTC to \( t_{99} \) to reduce the subjective selection of a response due to the injected tracer above background from noise at late times, limiting interpretation to observed short-term storage [after Mason et al., 2012; Ward et al., 2013a; Ward et al., 2013b]. An approximation to the window of detection \( t_w = t_{99} - t_1 \) was quantified by solving equation (1) for \( t_{99} \):

\[
0.99 = \frac{\int_{t=t_1}^{t=t_{99}} C(t)dt}{\int_{t=t_1}^{t=t_{CLP}} C(t)dt},
\]

where \( C \) is the observed, background corrected BTC (g m\(^{-3}\)), \( t \) is the time after injection (s), \( t_1 \) is the time from injection to tracer first arrival (s), and \( t_{CLP} \) is the time at which the observed BTC was initially clipped based on visual inspection. All BTCs (i.e., salt and fluorescein) used hereafter were truncated to \( t_{99} \).
We compared the salt and fluorescein BTCs to examine how much temporal information, such as \( t_w \), varies given differing tracer sensitivities and resolutions. In this comparison, we computed the associated temporal metrics (including those in Figure 1a and statistical moments as described in section 3.4.1) and examined the slope of the trend-line constructed through linear regression between those metrics corresponding to the fluorescein and salt BTCs. If a slope of unity and intercept of zero were within the associated 95% \( t \)-based confidence intervals (i.e., the ratio of fluorescein metrics to salt metrics is 1:1), the two BTCs were considered to provide similar temporal information. For this trend-line and those presented hereafter, we also tabulated the coefficient of determination (\( R^2 \)).

### 3.3. Net Change in Discharge and Unrecovered Tracer Mass

Sections of the study reach have been reported as losing during base flow conditions, but the hydraulic gradients between the stream and adjacent alluvial aquifer may alternate seasonally \cite{Schmidt et al., 2012}. We used the salt BTCs and rating curve estimates to examine the general pattern of net changes to discharge and mass losses through the study reach during the storm discharge period. Specifically, we estimated discharge at site A via dilution gaging \cite{Kilpatrick and Cobb, 1985},

\[
Q = \frac{M}{\int_{t_1}^{t_99} C(t)\,dt},
\]

where \( Q \) is the stream discharge (m\(^3\) s\(^{-1}\)) and \( M \) is the tracer mass injected (g), which ranged from 9 to 16 kg of NaCl in this study. The section of the study reach between the injection and site A was treated as a mixing length where we assume no tracer mass was lost. If some mass was lost over this section, \( Q \) will be slightly overestimated. We quantified the net change in discharge (downstream minus upstream) using estimates at sites A (dilution gaging) and E (rating curve), normalized by the upstream discharge to express as percent change. Assuming that the error in dilution gaging is roughly 8\% \cite{Schmadel et al., 2010} and using the 95% confidence intervals of the site E discharge estimates (Figure 2c), the 95% confidence intervals of these net changes were approximated. A net change in discharge is considered significant if zero is outside of these intervals. Likewise, tracer mass recoveries were estimated from the discharge estimates near site E,

\[
M_R = Q_D \int_{t_1}^{t_99} C_D(t)\,dt,
\]

where \( M_R \) is the tracer mass recovered (g), \( Q_D \) is the downstream stream discharge (i.e., site E) (m\(^3\) s\(^{-1}\)), and \( C_D \) is the observed tracer concentration at site E from the tracer released at the injection site (g m\(^{-3}\)). The percent unrecovered mass was quantified as the difference \( M_R - M \) normalized by \( M \). Again, because 95% confidence intervals were approximated for discharge estimates (Figure 2c), the 95% confidence intervals of unrecovered tracer mass. These confidence intervals do not include uncertainty due to \( \int C_D(t)\,dt \) in equation (3). Although this uncertainty is expected to be small relative to the uncertainty in \( Q_D \), a potential artifact of objectively truncating the BTC to \( t_{99} \) is artificially low \( M_R \) estimates.

### 3.4. Temporal Metrics

We estimated temporal metrics (including those in Figure 1a and statistical moments as described below) of each fluorescein and salt BTC to compare field-based tracer observations and interpret relationships between solute transport processes and discharge. Following this reach-scale analysis, we used the temporal metrics for each individual reach segment to examine how segment-wise processes changed with discharge. Lastly, we used the same metrics of each segment combination (Figure 2d) to test whether combining shorter segments into longer reaches influenced the interpretation of how solute transport processes change with discharge.

#### 3.4.1. Relationships Between Reach-Scale Temporal Metrics and Discharge

Using the \( t_{99} \) and \( t_{99} \) estimates, \( TSI \) was estimated (Figure 1a). To compare metrics across discharges, \( t_w \) and \( TSI \) were normalized by \( t_{ad} \) to remove the variation in the BTC caused by advection \cite{Gooseff et al., 2007; Ward et al., 2013a},

\[
t_w, \text{norm} = \frac{t_w}{t_{ad}}.
\]
where \( t_{\text{w,norm}} \) defines the window of detection relative to advective time and \( TSI_{\text{norm}} \) reflects the persistence of tailing relative to advective time. These metrics provide the overall influences of dispersion and short-term storage on the shape of the BTC independent of advective time assuming that \( t_{\text{w}} \) and \( TSI \) are linearly related to \( t_{\text{ad}} \). While this assumption is appropriate in advective-dominated stream systems [Gooseff et al., 2007; Ward et al., 2013a], other types of normalization may be necessary to better account for potential non-linearities [e.g., Gelhar et al., 1992]. Each observed BTC was normalized to express only the available temporal signature,

\[
c(t) = \frac{C(t)}{\int_{t=t_1}^{t_t} C(t) \, dt},
\]

where \( c \) reflects the recovered transit time distribution and the integral of \( C(t) \) with respect to time represents the zeroth temporal moment. We estimated the statistical moments of the transit time distributions by empirically calculating temporal moments of the recovered BTCs. Specifically, we estimated the first temporal moment \( (M_1) \),

\[
M_1 = \int_{t=t_1}^{t=t_2} tc(t) \, dt,
\]

which provides an estimate of mean arrival time. The \( n \)th-order temporal moment centered about \( M_1 \) (hereafter central moment) was estimated as

\[
\mu_n = \int_{t=t_1}^{t=t_2} (t-M_1)^n c(t) \, dt \quad \text{for } n > 1.
\]

The second central moment \( (\mu_2) \) provides the temporal variance, a metric of symmetrical spreading; the third central moment \( (\mu_3) \) reflects the temporal extent of late-time tailing. We computed the normalized metrics of the coefficient of variation \( (CV) \) and skewness \( (\gamma) \) to further compare BTCs across discharges,

\[
CV = \frac{\mu_2^{1/2}}{M_1},
\]

\[
\gamma = \frac{\mu_3}{\mu_2^{3/2}},
\]

where \( CV \) expresses the rate of symmetrical spreading relative to mean arrival time and \( \gamma \) reflects the extent of late-time tailing relative to symmetrical spreading (see supporting information Table S1 for a summary of all temporal metrics used). Therefore, these metrics should provide an indication of how dispersion and short-term storage processes change with discharge, respectively.

Relationships between discharge and the normalized reach-scale metrics \( (t_{\text{w,norm}}, TSI_{\text{norm}}, CV, \text{and } \gamma) \) for each monitoring site were constructed through linear regression. The discharges used for this relationship were those estimated from the rating curve near site E. Significant slopes were those whose 95% \( t \)-based confidence interval did not include zero.

### 3.4.2. Relationships Between Segment-Wise Temporal Metrics, Discharge, and Length

The transfer function, \( g(t) \), reflects the unique temporal signature of a solute traveling through a given reach segment independent of the input signal. When a single tracer injection is observed at different locations within a reach, the transfer function can be related to the upstream and downstream transit time distributions by convolution, assuming linear, time-invariant transport from one segment to the next,

\[
c_{ds}(t) = \int_0^\infty g(\tau) c_{ui}(t-\tau) \, d\tau,
\]

where \( c_{ds} \) is the observed transit time distribution at a downstream monitoring site and \( c_{ui} \) is the observed transit time distribution at the upstream monitoring site (input signal). To isolate segment-wise responses to discharge, we quantified metrics of the temporal characteristics of segment-specific transfer functions [after Ward et al., 2016] rather than apply a more complicated deconvolution scheme [e.g., Cirpka et al., 2007].
We estimated the change in window of detection between consecutive downstream and upstream BTCs ($\Delta t_w$) as an estimate of the transfer function window of detection,

$$\Delta t_w = t_{w,ds} - t_{w,us}. \quad (12)$$

To allow for comparison between segments, segment lengths, and discharges, we normalized this metric by the change in the advective time,

$$\Delta t_{w,\text{norm}} = \frac{\Delta t_w}{\Delta t_{ad}}. \quad (13)$$

where $\Delta t_{ad} = t_{\text{peak,ds}} - t_{\text{peak,us}}$, and $t_{\text{peak,ds}}$ and $t_{\text{peak,us}}$ are the times to peak concentrations of the downstream and upstream BTCs, respectively. Next, we estimated the change in TSI and normalized by the corresponding change in advective time,

$$\Delta TSI_{\text{norm}} = \frac{(TSI_{ds} - TSI_{us})}{\Delta t_{ad}}. \quad (14)$$

We used the change in $t_{ad}$, $t_w$, and TSI between downstream and upstream BTCs as surrogates to the corresponding metrics of the actual transfer function. Because the segment-wise temporal moments are theoretically linearly additive to produce the overall reach estimate [e.g., Riml and Wörman, 2011], and have been interpreted as such in existing field studies [Ward et al., 2014], we isolated statistical moments of the transfer function while eliminating the need to solve equation (11),

$$\Delta M_1 = M_{1,ds} - M_{1,us}. \quad (15)$$

$$\Delta \mu_n = \mu_{n,ds} - \mu_{n,us} \text{ for } n > 1. \quad (16)$$

Again, to compare between segments, segment lengths, and discharges, we computed the coefficient of variation and skewness of the transfer function,

$$\Delta CV = \frac{\Delta \mu_2^{1/2}}{\Delta M_1}. \quad (17)$$

$$\Delta \gamma = \frac{\Delta \mu_3}{\Delta \mu_2^{3/2}}. \quad (18)$$

We used the normalized segment metrics ($\Delta t_{w,\text{norm}}$, $\Delta TSI_{\text{norm}}$, $\Delta CV$, and $\Delta \gamma$) to investigate (1) relationships with discharge, (2) the spatial variability of segment response to discharge, and (3) whether estimates of these metrics were dependent on segment length.

Relationships with discharge were investigated by constructing linear trend-lines between the normalized metrics calculated for each segment (and combination of segments) and discharge near site E. Discharge was anticipated to vary slightly between segments (i.e., potentially a net losing stream), but if a relationship is present, discharge near site E will suffice because the large magnitude of change in discharge due to the storm event overwhelms the potential spatial change over the study reach. The metrics estimated for each of the seven discharges that reflect a single segment provided the sample population for the associated trend-line. Additionally, we pooled all normalized segment-wise metrics and examined the corresponding slope with discharge to provide an indication of the overall stream system response to changing discharge.

Spatial variability between segment responses to discharge was examined by comparing each metric ($\Delta t_{w,\text{norm}}$, $\Delta TSI_{\text{norm}}$, $\Delta CV$, and $\Delta \gamma$) sample population that reflected the individual segments (i.e., not the combined segments). This comparison was completed through a parametric test of comparing means in a one-way analysis of variance (ANOVA). We additionally completed a nonparametric test of comparing medians in a Kruskal-Wallis one-way ANOVA. The nonparametric test does not require the assumption of normally distributed residuals like the parametric version, so it provides additional information to prevent overinterpreting results from small sample sizes where the parent distribution is unknown. A 95% confidence level where $p < 0.05$ indicates means (denoted by $p_{\text{ANOVA}}$) or medians (denoted by $p_{\text{KW}}$) are different and segments are a significant source of variability. While these $p$-values indicate whether means and medians of at least one segment are different from all of the others, these tests were also repeated to compare each segment distribution to another on a paired basis. We selected these statistical tests because they are among the simplest to apply, acknowledging that the small
sample size limits the ability to identify underlying distributions or satisfy the assumptions of additional statistical tests.

To determine how the normalized transfer function metrics ($D_{tw,norm}$, $D_{TSI,norm}$, $D_{CV}$, and $D_c$) depended on segment length, we grouped corresponding estimates according to the number of combined segments. Each group provided the sample populations that were compared using the parametric and nonparametric tests outlined above. We only considered the sample populations for one, two, and three combined reach segments (Figure 2d) that contained more than 20 values representing different spatial scales and discharges to prevent overinterpretation of results. In this case, a $p_{ANOVA}$ and $p_{KW}$ of less than 0.05 indicates that the normalized transfer function metrics are dependent on the transport timescale (i.e., $\Delta t_{ad}$ and $\Delta t_{w}$) due to reach length selection. Hence, if this condition is recognized, interpreting how solute transport processes change with discharge is influenced by reach length selection.

4. Results

4.1. In-Stream Solute Tracer Observations

The influence of time-varying discharge conditions on the observed BTCs was assumed negligible due to the relatively short measurement periods. For the seven tracer injections measured at sites A through G, the duration of each corresponding measurement period from tracer injection to last detection at G ranged from 40–150 min (see supporting information Figure S2 for all observed BTCs). The possible change in discharge was between 1% and 6% over these measurement periods (red lines in Figure 2c).

Through the storm discharge period, some of the salt and fluorescein BTCs (49 each possible) were deemed erroneous based on initial quality control (e.g., debris potentially blocking the sensor, power failure, or data with visually high noise compared to other observations). This quality control resulted in a total of 29 colocated salt and fluorescein BTCs. In addition to these colocated BTCs, there were 12 instances where only salt BTCs were obtained and 4 instances where only fluorescein BTCs were obtained. A total of five fluorescein BTCs were omitted based on visually high noise. Data availability is summarized in supporting information Figure S2 and Table S2.

The colocated salt and fluorescein BTCs provided similar reach-scale temporal metrics ($t_w$, $TSI$, $M_1$, $\mu_2$, and $\mu_3$) (Figure 3, left column). All the slopes between these metrics estimated from salt BTCs and those estimated from fluorescein BTCs were statistically the same as 1 (i.e., 1 falls within the associated 95% confidence interval). The intercepts were also statistically the same as zero for all these metrics. The fluorescein tracer produced longer detectable tailing, resulting, mostly, in higher $t_w$ and $TSI$ estimates (above the 1:1 line in Figure 3, left column). The higher-order central moment estimates began to deviate further from the 1:1 line at higher values (i.e., those associated lower discharges). All slopes and intercepts of the normalized metrics ($t_{w,norm}$, $TSI_{norm}$, $CV$, and $c$) were statistically less than 1 and greater than zero, respectively (Figure 3, right column). However, the largest average difference between these metrics estimated from salt and fluorescein BTCs was $\sim$10%, indicating that both tracers provide similar temporal metric estimates following objective truncation to $t_{99}$.

As both tracers provided similar temporal metric estimates yet incomplete data sets, we used fluorescein BTCs to fill in gaps in the salt BTC data set, resulting in a total of 45 BTCs (41 salt and 4 fluorescein) to allow for greater coverage of advective times and windows of detection, and thus a more complete analysis. Fluorescein BTCs were selected to fill in gaps in the salt BTC data set, and not vice versa, because there were more salt BTCs available and salt is more readily available and commonly used in stream tracer studies.

4.2. Net Change in Discharge and Unrecovered Tracer Mass

The study reach was generally net losing through the storm discharge period (Figure 4a). The net change in discharge from site A to E ranged from $-15$ to $-20\%$ and was significant for every injection with the exception of injection 3 (where the discharge estimate at site E was highest) and injection 1 (where there was no salt BTC available at site A). A similar pattern occurred for the unrecovered tracer mass estimates (Figure 4b). For injections 4–7, unrecovered mass was significant and ranged from $-15$ to $-25\%$, providing further evidence that this study reach was a net losing stream.
Regardless of the tracer type, the normalized metrics were mostly insensitive to changing discharge where statistically significant relationships were the exception rather than the norm. The reach-scale metrics of the transit time distributions ($t_w$, $TSI$, $M_1$, $\mu_2$, and $\mu_3$) all decreased with increasing discharge (Figure 5, left column, estimated from 41 salt and 4 fluorescein BTCs). There was an increasing trend between these metrics and downstream location (bottom-to-top order of lines in the left column of Figure 5, representing increased length). The normalized metrics ($t_w,norm$, $TSI,norm$, $CV$, and $\gamma$) along the study reach were not significantly related to discharge in most cases (Figure 5, right column). The exceptions are that $CV$ was significantly negatively related to discharge only at sites C and D and $\gamma$ was significantly positively related to discharge at site A (Table 1).

Repeated analyses using only the salt BTCs and using only the fluorescein BTCs produced similar results that most relationships with discharge were not significant (see supporting information Figures S3 and S4 and Tables S3 and S4 for the results of these analyses).

### 4.4. Relationships Between Segment-Wise Temporal Metrics, Discharge, and Length

#### 4.4.1. Segment-Wise Responses to Changing Discharge

The normalized transfer function metrics ($\Delta t_w,norm$, $\Delta TSI,norm$, $\Delta CV$, and $\Delta \gamma$) indicated that some reach segments were more sensitive to changing discharge than others where more variability in the distribution is interpreted as a higher sensitivity (Figure 6). Segment EF was consistently the most sensitive to discharge based on the interquartile range. For $\Delta t_w,norm$ there was no significant difference between the segment-wise means and medians ($\rho_{ANOVA} = 0.30$ and $\rho_{KW} = 0.12$). For $\Delta TSI,norm$ the means were statistically the same ($\rho_{ANOVA} = 0.30$), but there
was evidence that the medians were different ($p_{KW} = 0.04$). The segment-wise means and medians of $\Delta T$ were statistically the same ($p_{ANOVA} = 0.11$ and $p_{KW} = 0.06$). Note that some negative $\Delta T$ estimates resulted from isolating the central moments of the transfer functions (equation (16)). The variation of $\Delta CV$ means and medians between segments was significant ($p_{ANOVA} < 0.001$ and $p_{KW} = 0.01$). When each segment distribution was compared to another on a paired basis, $\Delta T_{\text{norm}}$ of segment AB was different from segments BC, CD, and DE (also visualized by nonoverlapping notches in Figure 6); $\Delta T_{\text{norm}}$ of segment AB was different from segments BC, CD, and DE. The mean and median $\Delta CV$ of segment AB was different from all other segments. The mean and median $\Delta CV$ between all segments other than AB were the same. The means and medians of $\Delta c$ were statistically the same between each segment with the exception of segments AB and FG. See supporting information Table S5 for $p$-values of all these pairwise comparisons.

An apparent relationship between the normalized transfer function metrics and discharge was not clearly recognized. The slopes of the relationship between $\Delta T_{\text{norm}}$ and discharge for each segment and combination were generally negative (15 out of 21 segments; see supporting information Table S6 for slopes and $R^2$ values). However, the slope between this metric and discharge was only significant for segments AC and AD. The slopes of the relationship between $\Delta T_{\text{norm}}$ and discharge were generally negative (15 out of 21 segments), but none were significant. The slopes of the relationship between $\Delta CV$ and discharge were also generally negative (19 of 21 segments) with only two significant within segments AD and BD (negative slopes). A relationship between $\Delta c$ and discharge was unclear, with variable slopes (9 out of 21 segments with positive slopes, 12 out of 21 with negative slopes) that were not significant. Pooling all 21 segments indicated that $\Delta T_{\text{norm}}$, $\Delta T_{\text{norm}}$, and $\Delta CV$ overall had a slight downward trend with discharge while the relationship between $\Delta c$ and discharge was not significant (see supporting information Table S6 for slopes).

4.4.2. Segment-Wise Metrics Dependence on Length

We found a general pattern of decreasing variability in the transfer function metric estimates for longer segments (Figure 7). This decrease was due to both increased averaging of the heterogeneity in longer segments and an overall decrease in the number of samples available (i.e., where fewer extremes in the underlying distribution are reflected in longer segments). We statistically compared only the distributions associated with one to three combined segments where the numbers of individual observations in the sample populations were greater than 20. The means and medians of the $\Delta T_{\text{norm}}$ and $\Delta T_{\text{norm}}$ distributions did not significantly change when reach segments were combined (i.e., $p_{ANOVA} = 0.53$ and $p_{KW} = 0.44$ and $p_{KW} = 0.76$ and 0.96, respectively). The means and medians of the $\Delta CV$ distributions were significantly different (i.e., $p_{ANOVA}$ and $p_{KW} < 0.001$). By combining segments in this study reach under the seven discharge conditions, the mean and median of $\Delta CV$ significantly increased over longer lengths. While the medians of $\Delta c$ distributions decreased slightly over longer lengths ($p_{KW} = 0.04$), the means were statistically the same ($p_{ANOVA} = 0.81$).
5. Discussion

5.1. A Changing Transport Timescale Complicates Comparison of Stream Solute Tracer Studies

The extent to which dispersive and transient storage processes change with discharge can be easily misinterpreted from in-stream solute tracer tests based on different transport timescales (defined as the advective time and window of detection). For example, similar to findings from other tracer studies [e.g., Schmid et al., 2010], magnitudes of several BTC metrics decreased with increasing discharge (Figure 5, left column). A possible interpretation of this result is that physical processes have changed (e.g., activation or...
deactivation of storage flow paths), which is to be expected as varying discharge alters the turbulent energy, wetted geometry, hydraulic gradients, and connections to storage flow paths [Leopold and Maddock, 1953]. Yet through the suite of transport timescales of this study, the normalized temporal metrics ($t_{w,norm}$, $T_{S,norm}$, CV, and, $\gamma$) were mostly insensitive to changes in discharge (Figure 5, right column and Table 1). This lack of a strong relationship indicates that determining how differences in advective times and windows of detection control the shape of the BTC is needed when interpreting changes in physical solute transport processes. However, we recommend that similar tracer tests should be repeated across more reaches of different stream systems and discharge conditions to test the robustness of this conclusion.

We anticipate general changes to some transport processes due to varying discharge based on previous studies. The relative influence of dispersive processes can rise with discharge due to increased turbulent mixing [D’Angelo et al., 1993]. Likewise, we expect CV to increase with discharge. Subsurface short-term storage volume can increase with discharge in stream systems with relatively coarse (e.g., sand to gravel) streambeds [Dudley-Southern and Binley, 2015; Schmid et al., 2010; Zimmer and Lautz, 2014], which would likely cause an increase in prolonged tracer tailing or $\gamma$. Conversely, streams with low subsurface short-term storage potential (i.e., low-streambed hydraulic conductivity and valley slope) are less likely to undergo substantial storage flow path changes with discharge [Wondzell, 2011]. In this latter case, an increase in stream discharge should result in reduced tailing and a corresponding decrease in $\gamma$ because transport of the tracer is expected to be more sensitive to advective timescales than changing subsurface short-term storage processes. Similarly, based on a fluid-mechanics classification scheme after Jackson et al. [2013], we expect surface short-term storage to generally increase with discharge. The Selke study reach is not constrained by its valley and is set in highly permeable gravels [Schmidt et al., 2012]; therefore, we expected the reach-scale extent of short-term storage to increase with discharge, which would likely cause an increase in $\gamma$. Contrary to expectation, our estimates of CV were mostly insensitive to changes in discharge with some evidence of a negative relationship with discharge (sites C and D, Table 1). Although $\gamma$ was mostly insensitive to changes in discharge, there was some indication of a positive relationship at site A, suggesting that the extent of storage increased with discharge. The lack of a clear relationship with discharge indicates that these metrics either scaled directly with the transport timescale, or short-term storage and dispersion did not change appreciably despite the order of magnitude change in observed discharge. Direct scaling with the transport timescale is the more likely conclusion because a negative response in CV suggests that this metric is still influenced by changing transport timescales rather than changes in dispersive processes.

As expected, the advective time of the transfer function was inversely proportional to discharge and directly proportional to reach length—the longest advective time occurred at the lowest discharge and longest segment length (Figure 8a). Similarly, the window of detection expanded with an increase in advective time (Figure 8b). Theoretical temporal moments (i.e., those derived from commonly used transient storage models) suggest that both the coefficient of variation and skewness decrease nonlinearly with increasing reach length (and advective time) as the predicted BTC becomes more symmetrical [Schmid, 2002]. Contrary to theory, we observed an increasing trend in $\Delta CV$ with advective time (Figure 8c), which corresponds to an

<table>
<thead>
<tr>
<th>Site</th>
<th>$x$ (m)</th>
<th>Number of Q</th>
<th>$t_{w,norm}$</th>
<th>$T_{S,norm}$</th>
<th>CV</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>194</td>
<td>6</td>
<td>2.0E-02</td>
<td>0.07</td>
<td>--1.4E-02</td>
<td>0.02</td>
</tr>
<tr>
<td>B</td>
<td>294</td>
<td>7</td>
<td>-5.4E-02</td>
<td>0.40</td>
<td>-1.5E-02</td>
<td>2.2E-02</td>
</tr>
<tr>
<td>C</td>
<td>428</td>
<td>7</td>
<td>-8.4E-02</td>
<td>0.53</td>
<td>-6.2E-02</td>
<td>0.48</td>
</tr>
<tr>
<td>D</td>
<td>559</td>
<td>7</td>
<td>-8.6E-02</td>
<td>0.42</td>
<td>-7.2E-02</td>
<td>0.30</td>
</tr>
<tr>
<td>E</td>
<td>667</td>
<td>5</td>
<td>-1.2E-01</td>
<td>0.62</td>
<td>-9.5E-02</td>
<td>0.54</td>
</tr>
<tr>
<td>F</td>
<td>819</td>
<td>6</td>
<td>3.9E-02</td>
<td>0.09</td>
<td>5.1E-02</td>
<td>0.11</td>
</tr>
<tr>
<td>G</td>
<td>928</td>
<td>7</td>
<td>-7.6E-02</td>
<td>0.24</td>
<td>-5.3E-02</td>
<td>0.10</td>
</tr>
</tbody>
</table>

*Note that 41 salt and 4 fluorescein breakthrough curves were used to estimate these metrics and relationships. A positive or negative slope indicates the metrics increase or decrease with an increasing $Q$, respectively. A slope is considered significant (bold text) if zero is outside of the associated 95% confidence interval.

**Bold** indicates significant value.
increase in reach length (Figure 7). Consistent with theory, $D_c$ did have a slight decreasing trend with advective time (Figure 8d), but did not show a significant change with increased reach length (Figure 7). While Ward et al. [2013a] found that comparison of tracer observations across different discharges and reaches requires normalizing by associated advective times, the increasing trend in the coefficient of variation indicates that an expanding window of detection associated with longer advective times still limits direct comparison. A larger window of detection allows more time for processes like dispersion and short-term storage to act on the tracer. Consequently, tracer observations made at short advective times are more sensitive to the influence of truncation, or lack of an ability to measure the entire tail, than those made at longer advective times. An estimate of the coefficient of variation is clearly more sensitive to truncation than an estimate of skewness. This issue of truncation may also partially explain why dispersion in aquifers assessed from tracers typically increases with the spatial scale of observations [e.g., Gelhar et al., 1992]. Our results provide evidence that when tracer observations are compared, the influence of differences in the advective time and window of detection must be explicitly considered to prevent misinterpreting changes in solute transport processes. However, any change in the coefficient of variation or skewness at a given advective time is likely due to the influence of heterogeneity in channel morphology and process change associated with discharge.

In addition, other sources of uncertainty may restrict comparisons of tracer observations between different discharge conditions. A potential source of uncertainty is that short-term storage flow paths may change differently on the rising and falling limbs of the hydrograph [e.g., Ward et al., 2013a; Zimmer and Lautz, 2014]. For example, storage flow paths within the Selke study reach might be activated during the rising limb and subsequently deactivated during the falling limb. To better understand this influence, more tracer tests along the rising limb would be needed (Figure 2c). Furthermore, segment-wise tracer mass loss, as the study reach is anticipated to be net losing (Figure 4b), might change with discharge. Fortunately, if the stream only loses water without having significant gains, mass losses will not have substantially influenced the temporal information contained in the BTC because instream concentrations remain unchanged. Testing this mass loss influence would require tracer injections at

**Figure 6.** Segment-wise distributions (box-and-whisker plots of the quartiles where the notches are the 95% confidence intervals about the medians) of the transfer function metrics that express the window of detection ($\Delta W_{\text{norm}}$) and transient storage index ($\Delta TSI_{\text{norm}}$) normalized by the change in advective time, coefficient of variation ($\Delta CV$), and skewness ($\Delta \gamma$) over changing discharge. Note that 41 salt and 4 fluorescein breakthrough curves were used to estimate these metrics, resulting in 35 unique values (i.e., segment-wise estimates corresponding to number of discharges). Distribution means were compared through a one-way analysis of variance (ANOVA) and medians were compared through a Kruskal-Wallis one-way ANOVA. For mean and median comparisons, a $p_{\text{ANOVA}} < 0.05$ or $p_{\text{KW}} < 0.05$ indicates the mean or median of at least one segment is statistically different from all others, respectively. See supporting information Table S5 for segment-to-segment pairwise comparisons.
the limits of each segment (such as those by Payn et al. [2009]). Considerations of dynamic mixing length and data resolution would be necessary to perform finer-scale dilution gaging and mass recovery techniques during high discharges.

5.2. Sensitivity of Tracer Detection is Less Important at High Discharges

Detection of fluorescein was slightly more sensitive than detection of salt in late-time tailing based on generally larger windows of detection ($t_w$). Accordingly, the higher-order moments ($l_2$ and $l_3$) were generally larger for the fluorescein tracer (Figure 3, left column). These differences also caused the normalized metrics ($t_w$norm, TSI norm, CV, and $\gamma$) to be generally larger than those estimated from the salt tracer. Despite these differences, however, similar conclusions were drawn from both tracers: changes to the transport timescales confound the ability to observe changes in other processes because deviations between the metrics of the two different tracer types were small relative to changes in advective time. The metrics provided by the fluorescein and salt tracers were in close agreement at high discharges, indicating that the ability to detect short-term storage processes through tracer tests is limited at high discharges regardless of the tracer type selected.

An artifact of incomplete tracer response due to detection limits or signal to noise interferences in late-time tailing is artificially low temporal moment estimates [e.g., Drummond et al., 2012]. Consequently, objective truncation limits this study to observable short-term storage and may partially explain the lack of a robust trend between $D_{CV}$ and $D_{c}$ and discharge. Although we expect dispersion and short-term storage to change with discharge, this truncation may have reduced the influence of these processes relative to advection in the observations. One potential way to address this late-time detection issue is to apply statistical techniques to better isolate the tracer signal from noise [e.g., Aquino et al., 2015; Drummond et al., 2012]. Other sensing techniques and tracers may also help address this truncation issue. For example, highly sensitive tracers (e.g., synthetic DNA (deoxyribonucleic acid) molecules) have been shown to reflect much longer residence time storage flow paths [Foppen et al., 2011, 2013]. The use of such tracers...
thereby allows for a better determination of which portion of mass loss may be long-term storage and, in turn, which portion is groundwater recharge. Otherwise, modeling techniques to approximate the late-time tail may be necessary to provide a more complete representation of the actual transit time distribution and support a better understanding of relationships between short and long-term storage processes and discharge [e.g., Drummond et al., 2012; Stonedahl et al., 2012].

5.3. Temporal Metrics Do Not Reflect Observed Morphologic Patterns

Heterogeneity in stream morphologic characteristics is expected to cause differences in short-term storage [e.g., Gostner et al., 2013; Wondzell and Gooseff, 2013]. Likewise, this heterogeneity can cause dispersive processes to respond differently to changes in discharge. For example, increased turbulence due to higher discharge over features like riffles can increase dispersion, but higher discharge through slower moving sections like pools can reduce dispersion [Dyer and Thoms, 2006]. Based on this understanding, dispersion in sections with riffles (e.g., segment BC, Figure 2b) should increase with discharge compared to sections with pool-riffle sequences (e.g., segment DE, Figure 2b). The response of the window of detection ($D_{tw,norm}$) to changing discharge was different between segments BC and DE where DE was consistently higher (Figure 6 and supporting information Table S5), which corresponds to the more expected spreading of the tracer over riffle features. However, the symmetrical spreading relative to mean arrival time ($\Delta CV$) between segments BC and DE did not respond significantly differently to changing discharge, suggesting that the influence of spatially variable morphology was muted relative to changing advective timescales. The estimates of $\Delta CV$ within segment AB were consistently the largest, the least sensitive to changing discharge, and different from all other reach segments (Figure 6 and supporting information Table S5). Segment AB is a glide section with the steepest streambed slope (0.7%). We believe that the relatively uniform straight character and steep gradient of this segment (Figure 2b) caused the highest symmetrical spreading relative to mean arrival time while remaining the least sensitive to changing turbulent energy. Still, the means and medians of the normalized segment metrics were statistically equivalent between most segments, suggesting that the influence of segment spatial heterogeneity was muted or obscured by changes to transport timescales inherent in tracer observations. Only the most drastic differences in morphology appear to be reflected by tracer observations in our study.

Figure 8. (a) The advective time of the transfer function ($\Delta t_{ad}$) relative to discharge and reach segment length. A surface was linearly interpolated between these estimates to illustrate the general trend. (b) The window of detection of the transfer function ($\Delta t_{w}$) relative to $\Delta t_{ad}$ for each segment combination. A linear trend-line (solid line) and associated 95% confidence bounds (dashed lines) were approximated. The slope of this trend-line and coefficient of determination ($R^2$) is shown. (c) The coefficient of variation ($\Delta CV$) and (d) skewness ($\Delta \gamma$) for each segment combination relative to $\Delta t_{ad}$ and corresponding trend-line and 95% confidence bounds. Bold text indicates a significant value (i.e., zero is outside of the associated 95% confidence interval).
5.4. Reach-Scale Tracer Study Limitations May Complicate Upscaling

 Appropriately representing the heterogeneity and dynamics of solute transport processes across a network is critical to predict solute transport to receiving water bodies [e.g., Kiel and Cardenas, 2014; Wollheim et al., 2008]. The approach of many previous studies to examine solute transport processes throughout a network has been to conduct solute tracer tests over fixed reaches of equal length regardless of differences in transport timescales or morphologic characteristics [e.g., Covino et al., 2011; Gooseff et al., 2013; Stewart et al., 2011], with some calling for a length-normalized metric to compare processes between reaches or discharges [Runkel, 2002]. This study indicates that a change in the transport timescale (i.e., set by the interacting advective time and window of detection) resulting from a change in discharge or reach length selection can artificially manifest as spatial heterogeneity or process change. Therefore, we argue that fixing reaches of equal length or using length-normalized metrics to compare stream solute tracer observations may not be the best approach. Rather, since we observed variability between segment morphologies and that combining segments influences the interpretation of some transport processes, setting reach lengths and monitoring sites for tracer experiments should focus more on an understanding of transport timescales. One possible approach would be to fix the advective timescale between reaches, adjusting reach length accordingly. Ultimately, design of comparable tracer experiments may require preliminary studies or modeling efforts to initially quantify the range of transport timescales. Only comparable tracer observations will clarify relationships between solute transport processes and discharge. Improving the ability to compare tracer observations will likely lead to more accurate process-discharge relationships.

 Spatial patterns of stream morphology (e.g., streambed slope and channel width) throughout a network may also be an important consideration when representing solute transport processes within a network model [e.g., Zametske et al., 2007]. Fortunately, it may be possible to represent the heterogeneity of advective velocity at long reach scales by identifying the spatial pattern of the stream channel from imagery [Schmadel et al., 2014b]. Advances in upscaling from reach-based studies to networks will require consideration of within-reach spatial heterogeneity, spatial patterns in discharge and advective velocity, more sophisticated techniques to map the interconnected surface and subsurface waters, and a comprehensive analysis of tracer tests conducted over a range of discharges, spatial scales, and geologic settings. For example, tracer studies could be paired with independent and complementary measures of transport, such as through geophysics [e.g., Toran et al., 2013; Ward et al., 2012] or well networks in the adjacent aquifer [e.g., Voltz et al., 2013; Zametske et al., 2011], to better develop relationships between transient storage processes and discharge. However, from this study, one clear way forward is to better acknowledge and quantify the influence of changing advective timescales and windows of detection on the interpretation of processes such as dispersion and transient storage in our future reach-scale solute tracer studies.

6. Conclusions

 Using a suite of transport timescales (advective times and windows of detection) reflected in conservative stream solute tracer responses observed during a storm discharge period, we demonstrate that changes in advective times dominate and, consequently, mask other transport processes like dispersion and transient storage. While a possible interpretation of the tracer data could be that solute transport processes changed with discharge, we found through further analysis that the differences in the tracer data were generated primarily by variation in advective times. Furthermore, most individual segments within the original study reach did not respond differently to changes in discharge, suggesting that the influence of distinct channel morphologies was muted by the differences in advective times between tracer tests. Only the largest differences in morphologies were reflected in the tracer observations. We also found that through combining segments into longer reaches, differences of transport timescales could manifest as an incorrect interpretation of how processes change with discharge.

 Based on these findings, this study provides general recommendations for future tracer studies. First, while a high-sensitivity tracer (fluorescein) provided more late-time tailing information than that obtained from a lower-sensitivity tracer (salt), both independently provided the same conclusion that the impact of changing advective times obscured other solute transport process changes. Furthermore, differences between these tracer types were smaller at higher discharges, indicating that tracer selection is less important than considering the influence of changing timescales on the interpretation of changing processes. Second, the
influence of changing advective times and windows of detection must be established before tracer tests can be compared. Otherwise, a change in the transport timescale could be misinterpreted as a change in dispersion or transient storage. We show that this influence can be approximated by normalizing temporal metrics directly assessed from the observations by characteristics of the transport timescale including advective time and mean arrival time. However, the normalized metrics did not provide a sufficient correction to completely isolate dispersion and transient storage. Therefore, development of more complete relationships between processes and discharge may require better approximations of late-time tailing through other methods. Last, we recommend that selection of study reaches for design of tracer studies should be less influenced by length scales, such as fixed reaches of equal length proposed in previous studies, and more so by the transport timescales, which can be approximated from preliminary tests or modeling efforts. This recommendation is based on evidence that differences between windows of detection still limit the comparison of tracer responses despite normalization by characteristics such as the advective time. Improving the ability to compare tracer observations will clarify relationships between solute transport processes and discharge necessary to use reach-scale observations to infer solute transport at larger network scales.

Acknowledgments

The data collection campaign was funded by The Leverhulme Trust through the project “Where rivers, groundwater and disciplines meet: A hyporheic research network” with additional support from EU FP7-ITN “Interfaces: Ecohydrological interfaces as critical hotspots for transformations of ecosystem exchange fluxes and biogeochemical cycles” grant 607150, and from the authors’ institutions. Tools for solute tracer time series analyses were developed by Ward and others with support provided in part by the National Science Foundation (NSF) grant EAR-1331906 for the Critical Zone Observatory for Intensively Managed Landscapes (IML-CZO), a multi-institutional collaborative effort. A portion of Ward’s time in CZO, a multi-institutional collaborative Critical Zone Observatory for Tools for solute tracer time series analyses were developed by Ward and others with support provided in part by the National Science Foundation (NSF) grant EAR-1331906 for the Critical Zone Observatory for Intensively Managed Landscapes (IML-CZO), a multi-institutional collaborative effort. A portion of Ward’s time in CZO, a multi-institutional collaborative Critical Zone Observatory for

References


