A field comparison of multiple techniques to quantify groundwater–surface-water interactions

Ricardo González-Pinzón[1,10], Adam S. Ward[2,11], Christine E. Hatch[3,12], Adam N. Wlostowski[4,13], Kamini Singha[5,14], Michael N. Gooseff[4,15], Roy Haggerty[6,16], Judson W. Harvey[7,17], Olaf A. Cirpka[8,18], and James T. Brock[9,19]

[1]Department of Civil Engineering, University of New Mexico, Albuquerque, New Mexico 87131 USA
[2]School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana 47405 USA
[3]Department of Geosciences, University of Massachusetts, Amherst, Massachusetts 01003 USA
[4]Civil and Environmental Engineering, Colorado State University, Fort Collins, Colorado 80523 USA
[5]Department of Geology and Geological Engineering, Colorado School of Mines, Colorado 80401 USA
[6]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA
[7]National Research Program, US Geological Survey, Reston, Virginia 20192 USA
[8]University of Tübingen, Center for Applied Geoscience, Tübingen, Germany
[9]Division of Earth and Ecosystem Sciences, Desert Research Institute, Reno, Nevada 89512 USA

Abstract: Groundwater–surface-water (GW-SW) interactions in streams are difficult to quantify because of heterogeneity in hydraulic and reactive processes across a range of spatial and temporal scales. The challenge of quantifying these interactions has led to the development of several techniques, from centimeter-scale probes to whole-system tracers, including chemical, thermal, and electrical methods. We co-applied conservative and smart reactive solute-tracer tests, measurement of hydraulic heads, distributed temperature sensing, vertical profiles of solute tracer and temperature in the stream bed, and electrical resistivity imaging in a 450-m reach of a 3rd-order stream. GW-SW interactions were not spatially expansive, but were high in flux through a shallow hyporheic zone surrounding the reach. NaCl and resazurin tracers suggested different surface–subsurface exchange patterns in the upper ⅔ and lower ⅓ of the reach. Subsurface sampling of tracers and vertical thermal profiles quantified relatively high fluxes through a 10- to 20-cm deep hyporheic zone with chemical reactivity of the resazurin tracer indicated at 3-, 6-, and 9-cm sampling depths. Monitoring of hydraulic gradients along transects with MINIPOINT streambed samplers starting ~40 m from the stream indicated that groundwater discharge prevented development of a larger hyporheic zone, which progressively decreased from the stream thalweg toward the banks. Distributed temperature sensing did not detect extensive inflow of ground water to the stream, and electrical resistivity imaging showed limited large-scale hyporheic exchange. We recommend choosing technique(s) based on: 1) clear definition of the questions to be addressed (physical, biological, or chemical processes), 2) explicit identification of the spatial and temporal scales to be covered and those required to provide an appropriate context for interpretation, and 3) maximizing generation of mechanistic understanding and reducing costs of implementing multiple techniques through collaborative research.

Key words: stream–groundwater, hyporheic exchange, transient storage, resazurin, tracers, distributed temperature sensing, DTS, heat as a tracer, electrical resistivity, hydraulic head

Natural streams exchange water, solutes, and heat with ground water, providing habitats and facilitating biogeochemical processing (e.g., Peterson et al. 2001, Anderson 2005, Alexander et al. 2007, Constantz 2008, Harvey et al. 2013). Groundwater–surface-water (GW-SW) interactions occur at a wide range of spatial and temporal scales and involve complex, nested flow paths. These interactions affect the quality and quantity of water in streams and aquifers, and global cycling of C and nutrients (Stream Solute Workshop 1990, Dent et al. 2001, Battin et al. 2008, 2009, Kiel and Cardenas 2014). Despite numerous studies of GW-SW interactions in the last 3 decades and the broad

E-mail addresses: [10]gonzaric@unm.edu; [11]adamward@indiana.edu; [12]chatch@geo.umass.edu; [13]adam.wlostowski@coloradostate.edu; [14]ksingha@mines.edu; [15]mgooseff@rams.colostate.edu; [16]haggertr@geo.oregonstate.edu; [17]jwharvey@usgs.gov; [18]olaf.cirpka@uni-tuebingen.de; [19]jim.brock@dri.edu

DOI: 10.1086/679738. Received 2 March 2014; Accepted 13 August 2014; Published online 5 January 2015.
Freshwater Science. 2015. 34(1):000–000.

suite of technical and modeling tools available (e.g., Kalbus et al. 2006, Rosenberry and LaBaugh 2008, Healy 2010, Boano et al. 2014), quantifying these interactions remains challenging and highly uncertain.

Uncertainty in GW-SW interactions arises from the inherent spatial variability of hydraulic and reactive properties in streams and sediments. Large-scale observations, such as remote sensing applied to stream geomorphology, may miss dominant local-scale features (Legleiter et al. 2004, Adams and Spotila 2005, Wörman et al. 2007). This issue is even more critical in evaluation of subsurface properties, despite advances in hydrogeophysical exploration techniques (Kemna et al. 2002, Singha et al. 2008, Ward et al. 2010). Stream-tracer techniques yield integrated information about a stream and its transient-storage zones (surface and subsurface), but flow paths with greater residence times than the time scale of the experiment cannot be detected (e.g., Harvey and Wagner 2000, Ward et al. 2013a, Schmadel et al. 2013). Conversely, the relevance of point observations of hydraulic heads and concentrations for estimating effective processes at reach scales can be understood only when exchange rates, travel-time distributions, and reactivities of the hyporheic zone for the entire reach can be estimated. Fiber-optic distributed temperature sensing (FO-DTS) can detect groundwater inflows over a critical length between the reach scale and the bedform or local scale (e.g., Selker et al. 2006, Naranjo et al. 2013), but is insensitive to channel losses unless combined with measurements of the vertical hydraulic gradient (e.g., Krause et al. 2012) or vertical temperature profiles (e.g., Westhoff et al. 2011, Briggs et al. 2012a, b). Vertical temperature profiles in the stream bed can resolve both gaining and losing fluxes (e.g., Stallman 1965, Anderson 2005, Rau et al. 2014), but can do so only at individual points in space. Beyond the challenges of understanding the physical system, characterizing microbially mediated reactions to estimate biogeochemical transformations at the reach and network scales adds an additional layer of complication, e.g., relying on data from laboratory experiments or other sites may lead to systematic bias because these reactions are highly variable, and the decisive factors controlling the kinetics (including microorganisms) are not known decisively (Stewart and Franklin 2008, Flemming and Wingender 2010, Besemer et al. 2012).

Moreover, the models available for simulating transport and reactivity are either overly simplistic (e.g., assumed single immobile storage zone and 1st-order reactions) or are so complex that they cannot be uniquely calibrated. Simplistic models fail to represent all features of the observations (e.g., tails in breakthrough curves reflecting the longest storage times), and the inferred apparent properties may have no clear physical meaning. Increasing model complexity also puts the validity of the inferred parameters into question because of equifinality issues (Beven 2006). As a consequence, our ability to scale and predict hydrological and biogeochemical processes across stream ecosystems remains limited.

Each technique available to investigate GW-SW interactions provides key information about a single or a few processes of interest, and each technique has unique characteristics and problems. Thus, a potential approach to managing uncertainty and limitations inherent to each technique is to use multiple techniques in concert. We organized a workshop entitled "Techniques to Quantify Stream–Groundwater Exchange and Shallow Subsurface Transport: a Hands-on Workshop" in June 2012 to gather a community of researchers and practitioners interested in discussing the advantages and limitations of using conservative and smart reactive-solute-tracer tests measured in the stream and subsurface, measurement of hydraulic heads, distributed temperature sensing and vertical temperature profiles, and electrical resistivity imaging. We co-applied these techniques in a 3rd-order stream. Here, we describe the field study and the methods, present the results of the field experiment, and discuss 2 questions regarding the applicability and potential success of integrating techniques: 1) Do different techniques that are sensitive to processes on the same scale lead to similar interpretation? 2) What can we gain from incorporating multiple techniques in a single study of GW-SW interactions?

METHODS

Site description

We co-applied experimental techniques to investigate GW-SW interactions in Shaver Creek, a 3rd-order stream draining an 18.1-km^2 forested watershed near the Shale Hills Critical Zone Observatory (CZO) in Pennsylvania (USA). We worked in the former lake bed of Lake Perez, which was created in 1960 and drained in 2008. Total stream length is 13.4 km, mean annual precipitation is 980 mm, and mean elevation is 427 m asl.

We instrumented a 450-m stream reach with sampling stations for conservative and reactive (resazurin) tracers in the surface and subsurface, well transects to measure hydraulic heads, fiber-optic cable to measure in-stream temperature, temperature sensors in streambed piezometers to measure vertical temperature profiles, and electrical resistivity electrodes to image solute-tracer movement into the hyporheic zone (Fig. 1). In summary: 1) the tracer-injection site was at x = 15 m; 2) tracer sampling stations S1, S2, and S3 were at x = 125 m, x = 320 m, and x = 438 m, respectively; 3) 2 MINIPOINT (US Geological Survey [USGS]) subsurface water samplers were deployed 2 m upstream of S1 at x = 123 m; 4) the well transects were in the upper ½ of the reach; 5) the fiber-optic cable was deployed along the entire reach and in parts of the tributaries; 6) electrical resistivity was measured along 5 transects (J at x = 175 m, K at x = 165 m, L at x = 145 m, M at x =

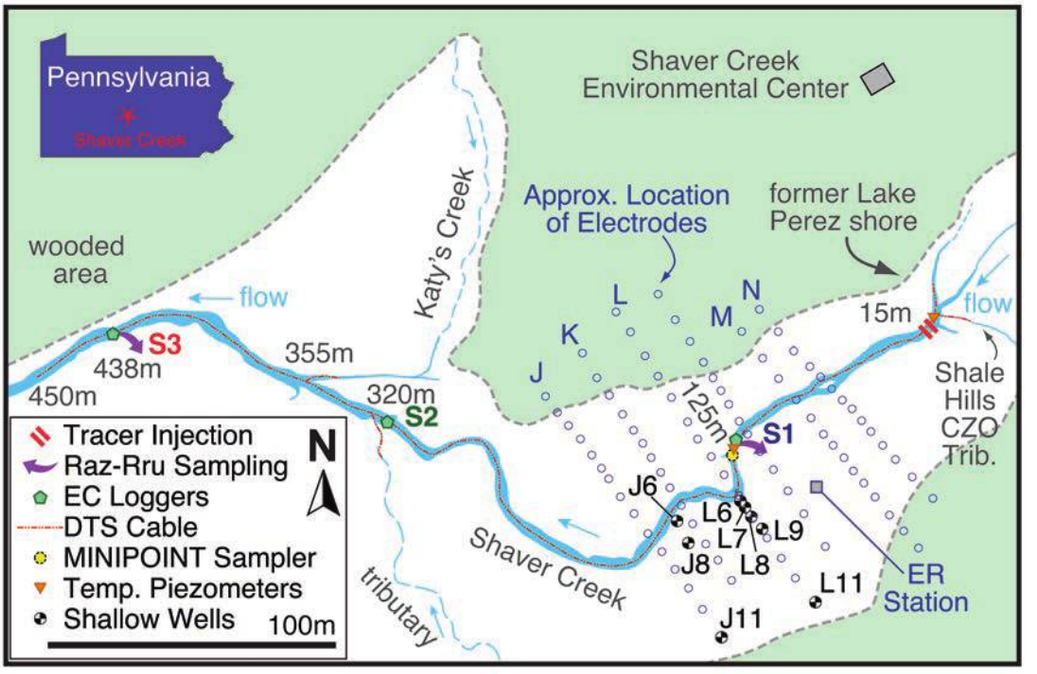

Figure 1. Field map of the 3rd-order stream Shaver Creek. Field methods included: in-stream injections of conservative (NaCl) and reactive (resazurin [Raz] and resorufin [Rru]) solute tracers monitored in the stream (EC loggers) and in the subsurface (MINIPOINT samplers), measurement of hydraulic head (shallow wells), subsurface measurement of heat (temp. piezometers), fiber-optic distributed temperature sensing along the study site (DTS cable), and electrical resistivity imaging in select cross-sections (ER station and electrodes). Trib. = tributary, approx. = approximate.

100 m, N at x = 65 m); and 7) significant tributary inflow of unlabeled water occurred at x = 355 m.

Conservative and reactive tracer injections

We co-injected NaCl as a conservative tracer and resazurin (Raz) as reactive tracer for 2.8 h. We dissolved the injectate in two 136.3-L containers filled with stream water. We tried to increase the specific conductivity in the stream water by ~70 µS/cm from its background signal (~90 µS/cm) and Raz to a maximum plateau concentration of ~150 µg/L. Starting at 1015 h on 13 June 2012, we injected the dissolved injectate from the 1st container at a constant rate of ~1620 mL/min into Shaver Creek (x = 15 m; Fig. 1). When the 1st container was empty, we quickly switched the pump intake hose to the 2nd container to maintain a nearly continuous mass flux of solute to the stream. The injection was terminated 2 h and 48 min after the injection began, and sampling lasted a total of 6 h and 15 min, at which time the conservative tracer in the stream reach had been indistinguishable from background conditions for >2 h. Discharge 50 min before the start of the constant-rate injection at S2 was ~206 L/s. We approximated discharge downstream of the tributary inflow with the NaCl breakthrough curve at S3 (tracer-dilution method) by assuming that most of the solute mass injected was recovered downstream.

We used Campbell Scientific (Logan, Utah) CS457 fluid electrical conductivity and temperature probes to log data at 10 s intervals at S1, S2, and S3. We took discrete, manual samples for Raz and its daughter product, resorufin (Rru), at stations S1 and S3. We filtered these samples immediately (0.7-µm pore size glass-fiber filter [GF/F]), refrigerated them (iced water in the field and 4°C in the laboratory), and read them within 24 h with a Fluorolog 3 laboratory spectrofluorometer (Horiba, Edison, New Jersey). Last, we manually sampled longitudinal plateau conditions along the study reach for conductivity and Raz beginning at ~1225 h just upstream of the injection site and ending at 1322 h at x = 490 m.

We sampled subsurface water for specific conductivity and Raz/Rru at 2 sites 2 m upstream of S1. We installed MINIPOINT samplers (Harvey and Fuller 1998, Duff et al. 1998) in the thalweg near the channel center and in an area of surface-water recirculation adjacent to the left bank of the channel (co-located with piezometers instrumented with temperature loggers). The MINIPOINT samplers allowed hyporheic water sampling via 0.32-cm stainless steel tubes with 1-cm slots forming a screen 0.5 cm behind a clamped tip. We pre-aligned the sampling tubes

for deployment at selected depths by passing each tube through fittings that gripped them in an acrylic disk that was lowered until the slotted ends of sample tubes reached the desired depths. We collected MINIPOINT samples at each site from surface water, and 3, 6, and 9 cm below the stream bed.

MINIPOINT sampling during the in-stream solute-tracer injection produced breakthrough curves for conservative and reactive tracers defined by 18 samples collected from each sampling depth. We pumped tubes simultaneously with a multihead pump that withdrew small-volume samples (6 mL) at low flow rates (1.5 mL/min) to minimize disturbance of natural subsurface fluxes and chemical gradients. Pump lines terminated at press-on luer fittings that were pushed onto 0.2-μm pore size (25-mm diameter) polyethersulfone filters (Pall, Port Washington, New York). Samples were collected into trays with prelabeled 20-mL low-density polyethylene (LDPE) plastic scintillation vials with Polyseal™ caps.

Conservative tracer analysis: transient-storage modeling The lumped transport equations describing advection, dispersion, nonequilibrium mass-exchange processes temporally retaining water and solutes (transient storage), and lateral exchange of a conservative tracer in a stream reach are (Bencala and Walters 1983):

$$\frac{\partial C}{\partial t} + \frac{Q}{A}\frac{\partial C}{\partial x} - \frac{1}{A}\frac{\partial}{\partial x}\left(AD\frac{\partial C}{\partial x}\right) = \frac{q_{LIN}}{A}(C_L - C) + \alpha(C_S - C) \qquad \text{(Eq. 1)}$$

$$\frac{\partial C_S}{\partial t} = \alpha\frac{A}{A_S}(C - C_S) \qquad \text{(Eq. 2)}$$

where Q (L^3/T) is the stream discharge; C, C_s, and C_L (M/L^3) are the solute concentrations in the main channel, the storage zone, and the lateral inflow, respectively; A (L^2) is the cross-sectional area of the channel; D (L^2/T) is the dispersion coefficient; q_{LIN} (L^2/T) is the discharge of lateral inflow per unit length; A_S (L^2) is the effective cross-sectional area of the storage zone; and α(/T) is the 1st-order mass-exchange coefficient between the main channel and the immobile storage zone. These transport equations are referred to as the transient-storage model. Once the model parameters (A, D, A_S, and α) are calibrated, the best-fit parameter set typically is used to estimate metrics for characterizing transient-storage processes (Bencala and Walters 1983, Harvey et al. 1996, Harvey and Wagner 2000, Runkel 2007).

We defined 3 reaches to simulate conservative transport with OTIS (Runkel 1998). Reach 1 extended from S1 (x = 125 m) to x = 350 m, shortly before the confluence of the major tributary (Katy's Creek) at ~x = 355 m. Reach 2 extended from x = 350 m to x = 360 m and covered the confluence zone. Reach 3 extended from x = 360 m to x = 450 m, 12 m downstream of the last measurement location.

We defined the upper boundary condition at S1 using the observed concentration–time data. We set the initial discharge along the main channel to 206 L/s and assumed flow was steady throughout the simulation. For reach 1, we set lateral inflows and outflows to 0 under the assumption that no net change in channel discharge or gross gains/losses occurred between S1 and x = 350 m. We modeled the confluence zone of reach 2 as 10 subreaches, each with a lateral inflow of 1.6 L $s^{-1}m^{-1}$, which simulated a 16 L/s lateral inflow from the tributary and increased discharge from 206 L/s (measured before injection) to 222 L/s (estimated from tracer-injection data). We assumed transport parameters at reach 2 (only 2% of the total length of the study site) were equal to those at reach 1 to reduce the number of parameters requiring calibration. We parameterized the channel area along reach 2 by a power-law hydraulic geometric relationship (Eq. 3) to simulate the increase in discharge and a corresponding increase in channel area:

$$A_{2,n} = c\,Q_{2,n}^{0.6}\; n = 1, 2 \ldots, 10 \qquad \text{(Eq. 3)}$$

where $A_{2,n}$ is the area of reach 2 at subreach n; c is a scaling coefficient; and $Q_{2,n}$ is the main channel discharge of reach 2 at subreach n (cf. Gooseff and McGlynn 2005). For reach 3, we set the main channel area to that of the last subreach of reach 2 ($A_{2,10}$). Similar to reach 1, we assumed lateral inflows and outflows were 0.

Our model conceptualization relies on 8 unknown parameters. We used 2 observed breakthrough curves at S2 (x = 320 m) and S3 (x = 438 m) to calibrate the model. We calibrated all 8 parameters simultaneously using the Shuffled Complex Evolutionary algorithm (Duan et al. 1992), which searched for a best-fit parameter set that minimized the sum of root mean squared error (RMSE) values at S2 and S3. The best simulations were assumed to have an objective function (f_{best}) within the best 0.1% of the total objective function range (Eq. 4):

$$f_{best} \leq f_{min} + 0.001(f_{max} - f_{min}) \qquad \text{(Eq. 4)}$$

where f_{min} is the lowest objective function (corresponding to the best simulation), and f_{max} is the greatest objective function.

Conservative tracer analysis: interpretation as residence time distributions For tracer injections done over finite time scales and assuming hydrological conditions remain constant during the study, the observed breakthrough

curve $c_{obs}(x, t)$ is a convolution of the input signal $c_{in}(t)$ with the so-called transfer function g(τ) (/T):

$$c_{obs}(x,t) = \int_0^\infty g(\tau) c_{in}(t-\tau) d\tau \qquad \text{(Eq. 5)}$$

where τ (T) is the residence time. Residence-time distributions are characterized by their statistical moments (mean, standard deviation [SD], skewness, and kurtosis). In Fickian transport, the residence-time distribution is the inverse Gaussian distribution (Kreft and Zuber 1978), in which the skewness equals the coefficient of variation (CV). Thus, a skewness larger than the CV may be indicative of a transient-storage process. Parametric models of g(τ) may be assumed, such as a log-normal distribution, a γ distribution, or the analytical solution for pulse-injection in solute transport, potentially undergoing kinetic mass exchange (Toride et al. 1993). Cirpka et al. (2007) presented a nonparametric approach to estimate g(τ) and uncertainties. This approach has been used to deconvolve pulse- and constant-rate injections (Payn et al. 2008).

We subtracted the background values of specific conductivity from the observed breakthrough curves at stations S1 (x = 125 m), S2 (x = 320 m), and S3 (x = 438 m) and deconvolved these signals with the approach of Cirpka et al. (2007) by using 200 conditional realizations for each input–output combination.

Reactive tracer analysis: reach-scale interpretation of reactive tracers Processing rate coefficients (reaction, decay, or uptake rates) in stream reaches can be estimated with algebraic relationships derived from the transient-storage-model equations adapted for reactive solutes following 1st-order processing (González-Pinzón and Haggerty 2013). Assuming uniform biogeochemical conditions and neglecting processing in the main channel, i.e., assuming that reactions preferentially take place within the hyporheic zone because of enhanced chemical gradients and larger volume of colonized sediments, the apparent processing rate coefficient λ_T (/T) in a stream reach is a function of exchange with the storage zone (González-Pinzón and Haggerty 2013, Runkel 2007):

$$\lambda_T = \frac{\alpha \lambda_{sz}}{\alpha \frac{A}{A_s} + \lambda_{sz}} \qquad \text{(Eq. 6)}$$

where λ_{sz} is the biochemical processing rate coefficient within the storage zone, and the other parameters are as defined for the transient-storage-model equations. This expression does not account for dispersion, but the error introduced by this assumption is negligible (González-Pinzón and Haggerty 2013). λ_T also can be estimated directly from tracer data without the need to calibrate model parameters that can be highly uncertain (A, A_s, λ_{sz}, α; see discussions on uncertain parameters in transport models by Wagner and Harvey 1997, Wagener et al. 2002, Wlostowski et al. 2013, González-Pinzón et al. 2013, Kelleher et al. 2013, Ward et al. 2013b) as:

$$\lambda_T = \frac{\ln\left[m_o^{up}/m_o^{dn}\right]}{\tau} \qquad \text{(Eq. 7)}$$

where $m_o^{up} = \int_0^\infty C^{up}(t)dt$ and $m_o^{dn} = \int_0^\infty C^{dn}(t)dt$ ($ML^{-3}T$) are the upstream and downstream 0th temporal moments (which can be substituted by plateau concentrations) of the solute of interest; $\tau = L/u$ (T) is the mean travel time of a conservative solute in the reach; L (L) is the length of the reach; and u (L/T) is the mean velocity in the reach ($u = Q/A$).

We estimated reach-scale apparent transformation (processing) rate coefficients of Raz (λ_T) in the same reaches 1 and 3 defined for the conservative transport modeling with OTIS. We used mean travel times (τ) and the plateau concentrations measured along the reach in place of the 0th temporal moments. We estimated mean travel times at selected locations along the study reach by fitting a trend line to the mean travel times estimated from specific conductivity data at x = 60 m, S1, S2, and S3.

Conservative and reactive tracer analysis: storage dynamics compared between reach-scale and individual (local-scale) geomorphic units The mass balance presented by Harvey et al. (2013) for the hyporheic zone can be used to express the measured concentration of a potentially reactive solute in hyporheic flow in terms of relative inputs from surface water and ground water, and the amount of solute that reacted during hyporheic-zone transport:

$$C'_{hz} = f_s C'_s - f_g C'_g + C^*_{hz} \qquad \text{(Eq. 8)}$$

where subscripts denote concentrations in hyporheic flow (hz) or end-member concentrations entering hyporheic flow from surface-water (s) or groundwater (g) inputs. Specifically, C'_{hz} (M/L^3) is the measured concentration of a potentially reactive solute at a given sampling depth; C'_s and C'_g (M/L^3) are the surface-water and groundwater end-member concentrations of the potentially reactive solute contributing to hyporheic flow at that depth; C^*_{hz}(M/L^3) is the mass concentration of the solute lost or gained by reaction during transport to that depth; and f_s and f_g are the associated mixing fractions contributed from surface water and ground water.

The mixing fractions of surface and ground waters contributing to hyporheic flow are estimated from measurements of the distribution of a conservative solute tracer as:

$$f_s = \frac{C_{hz} - C_g}{C_s - C_g} \text{ and } f_g = 1 - f_s \qquad \text{(Eq. 9)}$$

where C denotes the concentration of any conservative constituent, either a tracer that is naturally present and that differs between surface water and ground water, or concentration of an injected tracer after concentrations at all sampling locations have reached steady-state (usually several hours or more after the beginning of a constant-rate tracer injection into the stream).

For a constant-rate stream-tracer injection, the median subsurface residence time, τ_{hz} (T), is estimated for a given sampling depth in the hyporheic zone as the elapsed time between the time of arrival of the tracer in the stream to the time at which the subsurface tracer concentration reaches the 50th percentile relative to its plateau concentration. The estimate of hyporheic water flux (q_{hz} [L/T]), e.g., the vertical flux of water crossing the bed per unit area, is:

$$q_{hz} = \frac{d_{hz}\theta}{\tau_{hz}} \qquad \text{(Eq. 10)}$$

where θ is sediment porosity and d_{hz} (L) is the depth of the measurement below the stream bed. Results based on measurements from several sampling depths can be integrated to estimate average hyporheic-zone conditions using the flux-weighting approach (Harvey et al. 2013).

We applied the analysis in Eqs 8 and 9 specifically to Raz (i.e., $C'_g = 0$) to estimate the transformation of Raz in the hyporheic zone as normalized (nondimensional) quantities for different depths along the vertical profile (3, 6, and 9 cm) defined by MINIPOINT sampling. We estimated the normalized transformation for each depth as:

$$\Delta Raz_{hz} = \frac{(Raz_{surf} - Raz_{hz})}{Raz_{surf}} \qquad \text{(Eq. 11)}$$

where Raz_{surf} [M/L^3] represents surface plateau samples, and Raz_{hz} [M/L^3] represents hyporheic plateau samples corrected by specific conductivity. We estimated the magnitudes of Raz_{surf} at each depth from the average of 5 samples taken during plateau conditions.

To assess the extent of metabolic activity within the hyporheic zone at the reach scale, we compared the normalized hyporheic-zone transformation of Raz (ΔRaz_{hz}) with a normalized longitudinal transformation of Raz (ΔRaz_{long}) defined as:

$$\Delta Raz_{long} = \frac{(Raz_{x=50m} - Raz_x)}{Raz_{x=50m}} \qquad \text{(Eq. 12)}$$

where Raz_x is a plateau concentration corrected by specific conductivity. $Raz_{x=50m}$ is used to normalize longitudinal reactivity with respect to x = 50 m, which is the first well mixed data point seen from the longitudinal plateau sampling.

Hydraulic-head measurements

We installed 7 well transects by direct-push techniques in May 2012 (1 mo before the experiment) along the upper half of the experimental reach. Each well transect had 6 to 9 wells installed to a depth of 3 m when possible. Well casings were made of perforated (slotted) 3.8-cm outer diameter (3.2-cm inner diameter [ID]) schedule-80 polyvinyl chloride (PVC) pipe. The wells were left open at the bottom and screened to a height of 2.3 m above the bottoms of the wells. We deployed water-level data loggers (HOBO U20-001-01; Onset, Bourne, Massachusetts) to just above the full depth of each well. We calculated water-surface elevations in each well by measuring water-level depth from the top of the casing and subtracting this value from the elevation of the top of the well case as determined by global positioning system (GPS) survey in April 2013. Data were collected by water-level loggers at 10-min intervals and later corrected for atmospheric pressure changes by a separate record that was collected outside of a well.

Heat as a tracer

Distributed temperature sensing We used a Sensornet Limited® Oryx™ (Elstree, Hertfordshire, UK) fiber-optic distributed temperature sensing unit (FO-DTS), which provides temperature data at 2-m spatial resolution on 4 channels up to 2.5 km in length and can sample as often as every 10 s. We used a 4.5-mm diameter telecommunications cable (LLC™ Broadcast Deployable Tight Buffered Cable; AFL, Duncan, South Carolina) composed of central acrylite-coated glass fibers (consisting of a 50-μm-diameter glass core surrounded by a 37-μm-thick glass cladding with a lower index of refraction).

We deployed ~450 m of cable along the thalweg of the stream and secured it in place with large cobbles. We deployed additional coils of cable in the tributaries, and reserved coils for calibration (15 m upstream of the injection point and 19 m downstream of the end of the stream reach). During the deployment, the FO-DTS was contained in a weatherproof enclosure and powered by solar panels. We made independent temperature measurements for calibration of FO-DTS data in portable coolers filled with stream water or an ice-and-water slush and continually mixed with an aquarium bubbler to prevent stratification. We measured temperature continuously during FO-DTS deployments with Onset Computer Corporation WaterTemp Pro (ceramic thermistor) temperature loggers, with nominal resolution of 0.02°C and accuracy of 0.15°C over a temperature range of −4 to

37°C, or PT100 loggers (100-ohm resistance Pt thermocouples) incorporated into the FO-DTS unit, with reported resolution of 0.01°C and accuracy of 0.02°C. The cable passed through each bath twice, once at each end of the cable (27 m of cable were submerged in the baths on the outgoing end, and 19 m of cable were submerged in the baths on the returning end of the cable). The deployment was made from 13–21 June 2012 before the cable was removed from the stream. We collected data along the entire FO-DTS cable and from all independent temperature loggers every minute. Calibration of FO-DTS temperatures vs independent loggers was done every time for each measurement to account for any electronic artifacts or instrument drift (Hausner et al. 2011, Van de Giesen et al. 2012).

Vertical fluxes from temperature profiles in piezometers We installed PVC piezometers (0.03 m ID) to a depth of ~0.65 to 0.75 m into the stream bed at x = 13 m (upstream) and at mid-reach S1 (x = 125 m). We notched 0.60-m wooden rods (0.013 m ID) every 5 cm and inserted a Maxim Integrated iButton® (iButtonLink Technology, Whitewater, Wisconsin) temperature logger in each notch. We placed the rods at the bottom of the piezometers, making the effective logger depths every 0.05 m between 0.05 and 0.70 ± 0.05 m below the stream bed. We used iButton loggers to measure temperatures every 15 min in 2 piezometers (nominal resolution = 0.5°C, accuracy = 1°C over a temperature range of –55 to 100°C). We calculated a time series of daily average vertical fluxes for each pair of loggers for the duration of the record (from 13 June to 25 July 2012). The average depth between logger pairs was taken to characterize the depth represented by the flux vector calculated from the amplitude ratio (A_r = $A_{deep}/A_{shallow}$) between pairs of temperature time series.

Analysis of temperature measurements The governing equation for 1-dimensional heat transfer in a saturated fluid-sediment system is (Anderson 2005, Carslaw and Jaeger 1959, Stallman 1965, Goto et al. 2005):

$$\frac{\partial T}{\partial t} + \frac{q_{seepage}}{\gamma}\frac{\partial T}{\partial z} - \kappa_e \frac{\partial^2 T}{\partial z^2} = 0 \qquad \text{(Eq. 13)}$$

where T is temperature (°K); t is time; $q_{seepage}$ (L/T) is the vertical seepage rate into or out of the stream bed (positive values indicate upward flux and gaining conditions, negative values indicate downward flux and stream leakage); γ is the ratio of volumetric heat capacity of the saturated stream bed to that of the fluid; z is the vertical spatial coordinate (L); and κ_e (L^2/T) is the effective thermal diffusivity. The effective thermal diffusivity $\kappa_e = \lambda_e/\rho C$ is the effective thermal conductivity divided by the bulk heat capacity of the porous medium.

The approximate volume of shallow groundwater inflow to the surface can be quantified from observations of upstream discharge, in-stream temperature changes (FO-DTS), and groundwater temperature. FO-DTS applied to a cable on the stream bed senses only water in the stream, so the technique is not sensitive to seepage losses. However, temperature profiles at multiple depths in piezometers installed in the stream bed can be used to quantify upwelling and downwelling seepage rates (Stallman 1965, Stonestrom and Constantz 2003, Conant 2004, Hatch et al. 2006, 2010). This method requires inverting the 1-dimensional heat-transfer equation (Eq. 13) to estimate the seepage rate from differences (amplitude dampening and phase shift) in the diurnal temperature time series between 2 selected depths. Temperature data are paired, filtered, detrended, and processed via iterative algorithms to solve for streambed-seepage rates, yielding a daily mean value (Gordon et al. 2012, Swanson and Bayani Cardenas 2010).

Electrical resistivity imaging (ERI) of hyporheic transport

We installed 5 transects of 12 electrodes perpendicular to the stream channel, arranged as a grid (distance between electrodes was 4 m along each transect; transect J was at x = 175 m, K at x = 165 m, L at x = 145 m, M at x = 100 m, N at x = 65 m). The electrodes, constructed of stainless steel foil fixed to a 2.5-cm diameter PVC pipe, were driven ~20 cm into the subsurface (after Ward et al. 2012b). We selected a data-collection scheme using forward modeling via the R2 Generalized Inversion Code (version 2.7; Binley 2013) for a low-resistivity target approximately the size of the stream channel immediately below the stream bed, which was representative of a stream-sized hyporheic zone that would decrease in electrical resistivity because of transport of a saline tracer into the near-stream subsurface. The grid we used here and later for inversion was 2 m in the horizontal and 0.5 m in the vertical direction and included topographic data. We tested several data-collection schemes in an attempt to balance rapid collection of data with resolution. The selected scheme consisted of 145 electrode configurations (a current pair and a potential pair of electrodes) along each transect, including both dipole–dipole and Wenner configurations. Data collection cycled continuously through the transects, and a complete data set for the array was collected at intervals of ~30 min. We did not collect reciprocal data to maximize temporal resolution. We collected data continuously beginning with a complete data set before the tracer injection and continuing until ~7.5 h after the injection began. We averaged 2 measurements for each observation for an estimate of error. If the SD of the 2 measurements was >3% of the mean, we added a 3rd measurement to the average. The average background ap-

parent resistivity of all measurements made during the experiment was 117 ohm-m.

Analysis of electrical resistivity sampling ERI data acquisition involves establishing a difference in electrical potential between 2 source electrodes and measuring the resultant potential difference at 2 receiving electrodes. The physics underlying ERI is described by the Poisson equation subject to boundary conditions:

$$\nabla(\sigma\nabla\emptyset) = I\delta(x - x_s, y - y_s, z - z_s) \qquad \text{(Eq. 14)}$$

where σ (usually in S/m) is the electrical conductivity (the reciprocal of resistivity), a combined intrinsic property of both sediment matrix and the fluid; Ø ($ML^2T^{-2}Q^{-1}$; in V) is the electrical potential; I (Q/T; in amps) is the electrical current source; δ (reciprocal units of the argument) is the Dirac delta function; x, y, and z (L; in m) are the spatial position vectors; and x_s, y_s, and z_s (L; in m) are the spatial coordinates of the current source.

We solved Eq. 14 as the forward problem for potentials. We calculated resistance as the difference in potential at 2 locations divided by the applied current. Each measured resistance is a function of the electrical properties of both solids and liquids in the system (Keller and Frischknecht 1966). The resistance is not an intrinsic property of a rock or soil. Instead, it depends on the geometry of (or distance between) electrodes used for the measurement. The relevant intrinsic property is resistivity (ρ) or its reciprocal electrical conductivity (σ).

We inverted electrical data using the R2 Generalized Inversion Code (Binley 2013). First, we inverted background data with a homogeneous starting model. Next, we used the time-lapse inversion of LaBrecque and Yang (2001) to invert on differences from background measurements for all data collected after the injection began (cf. Ward et al. 2012a). For data collected during the injection, we assigned all observations in 1 sequence to a single time step. We acknowledge that this assumption leads to some temporal smearing, but we assumed it was minimal given the rapid data collection. For all inversions, we assumed an error model of 0.001 Ω absolute error and a relative error of 0.2%. We adjusted data weights iteratively during the inversion. Interpretation of these images must be completed in the context of the limitations of the method, which include temporal and spatial smearing, and out-of-plane effects for the 2-dimensional images (see Slater et al. 2002, Day-Lewis et al. 2005, 2007 for further discussion).

RESULTS

Conservative transport modeling at the reach scale

Transient-storage model The near-optimal transient-storage model parameters selected by the Shuffled Complex Evolution algorithm are shown in Table 1. The Shuffled Complex Evolution search algorithm generated 5776 parameter sets (same number of simulations) in total before it converged to the optimal parameter set. These simulations were used to generate confidence intervals (CIs) of the optimal parameter set. The range of a parameter within the best-performing parameter group provides a CI for that parameter. These values are reported in parentheses beside best-fit parameters in Table 1.

We estimated transient-storage metrics defined in Harvey and Wagner (2000) from the best-fit parameters (Table 2). The mean hyporheic residence time (t_s), hyporheic exchange flux (q_s), hydrologic retention factor (R_h), and the storage zone–main-channel area ratio (A_s/A) are more than 1 order of magnitude different along reach 3 compared to reach 1.

Deconvolution of conservative-tracer breakthrough curves

The resulting breakthrough curves of excess specific conductivity indicated nonuniform injection over the injection

Table 1. Calibrated transient-storage-model parameters and parameter confidence intervals (CIs).

Parameter	Range	Optimal value (CI)
Reach 1 (x = 125–50 m) Q = 206 L/s		
A (m^2)	0.5–1.0	0.76 (0.74–0.82)
D (m^2/s)	0.1–4.0	0.43 (0.26–0.85)
A_s (m^2)	1×10^{-5}–10	0.095 (0.04–0.11)
α (/s)	1×10^{-10}–1×10^{-2}	7.54×10^{-4} (1.16×10^{-4}–1.10×10^{-3})
Reach 2 (x = 350–360 m)		
c (cf. Eq. 3)	1.0–3.0	2.02 (1.96–2.10)
Reach 3 (x = 360–450 m) Q = 222 L/s		
D (m^2/s)	0.1–4.0	0.32 (0.10–0.80)
A_s (m^2)	1×10^{-3}–10	0.29 (3.33×10^{-3}–1.08)
α (/s)	1×10^{-10}–1×10^{-2}	1.67 x 10^{-5} (1.38×10^{-7}–6.09×10^{-5})

Table 2. Conservative transport metrics from the optimum set of parameters of the transient-storage model.

Metric	Reach 1 (x = 125–350 m)	Reach 3 (x = 360–450 m)
Mean hyporheic residence time t_s, (h)	0.043 (0.036–0.13)	22.59 (6.16–80.73)
Hyporheic exchange flux, q_s (m^2/s)	7.3×10^{-4} (8.27×10^{-5}–9.06×10^{-4})	4.2×10^{-5} (3.13×10^{-5}–5.63×10^{-5})
Hydrologic retention factor, R_h (s/m)	0.51 (0.15–0.56)	15.48 (5.28–45.02)
Storage zone–main-channel ratio, A_s/A	0.14 (0.037–0.16)	4.28 (1.46–12.00)

period because the plateau concentration was not constant (Fig. 2A). We deconvolved these signals with the approach of Cirpka et al. (2007). We used 200 conditional realizations for each input–output combination and a smoothness parameter $\theta = 1 \times 10^6/d^3$. The SD of specific conductivity, $\sigma_{EC} \approx 0.6$ μS/cm, indicates the goodness of the fit. It is smaller than the resolution of the loggers (1 μS/cm), which implies that smoothness of the transfer function enforced by the method did not deteriorate the quality of the fit. The resulting transfer functions are shown in Fig. 2B. The inset shows the same transfer functions as a double-logarithmic plot to highlight late-time tailing. Table 3 lists characteristic values of the median distributions.

We recovered 96% of the signal from observation points S1 to S2, indicating apparent gaining conditions. The recovery was 99% based on the mean transfer function averaged over 200 realizations. The mean travel time of 13.2 min for a travel distance of 195 m corresponds to an effective velocity of 24.6 cm/s. The CV of the residence-time distribution (0.20) was significantly smaller than the skewness (1.15), indicating a nonFickian contribution to transport.

We recovered only 90% of the signal from observation points S2 to S3. The recovery was 93% based on the mean transfer function. This decrease was caused mainly by dilution from the 2 tributaries, particularly Katy's Creek at x = 355 m. The mean travel time of 7.3 min for a travel distance of 118 m corresponds to an effective velocity of 27.0 cm/s, which is only slightly larger than that obtained between S1 and S2. The CV of the residence-time distribution (0.21) is comparable to that in the S1–S2 reach, but the skewness is considerably larger (6.46), indicating that transport is less Fickian than in the reach 1. The plot of the transfer function from S2 to S3 (Fig. 2B) does not look particularly skewed, but small contributions in the tail can cause large skewness. Because of the limited resolution offered by our conservative tracer, part of the skewness in the BTCs could be an artifact associated with high noise in the tail.

Quantifying hyporheic flow using subsurface conservative tracers

Specific conductivities differed between surface and subsurface water samples collected with the MINIPOINT sampler before the tracer experiment (surface water: 93 μS/cm, subsurface water: 306 μS/cm). We used this pre-injection difference in calculations of mixing between surface water and ground water in the hyporheic zone. After the injection began, we used the median breakthrough time of the conservative tracer at subsurface sampling points to calculate hyporheic water residence time and hyporheic exchange flux. Shallow subsurface water from beneath the channel thalweg at Shaver Creek consisted of 99.5, 96.3, and 32.1% surface water at 3, 6, and 9 cm depth, respectively, suggesting a hyporheic zone that is ~10 cm deep. In contrast, the subsurface water at 3 and 6 cm depth beneath a side cavity closer to the left margin of the channel was composed of 7.0 and 0.0% surface water, indicating that groundwater discharge to the stream dominated shallow subsurface flow beneath the channel margin.

Subsurface conservative tracer breakthrough curves identified hyporheic water residence times that were very short (~4 min) beneath the channel thalweg to a depth of 6 cm, which corresponds to hyporheic water flux (q_{hz}) of 6 m/d. At a depth of 9 cm, the hyporheic water residence time was much greater (1.2 h) and the corresponding q_{hz} was an order of magnitude lower (0.5 m/d at 9 cm). In contrast, beneath the surface-water side cavity at the channel margin, the hyporheic water residence time was undetectable even at the shallowest depth (3 cm), suggesting that hyporheic flow beneath the channel margin is small relative to the thalweg. An injection longer than 3 h probably would be required to quantify relatively slow hyporheic flow beneath the channel margin.

Reactive Raz–Rru system at the reach and local scales

The apparent transformation rate coefficient in reach 1 was λ_T = 0.35/h. In reach 3, λ_T = 1.01/h. For reach 1–3, λ_T = 0.59/h (Fig. 3). If we assume that Raz was primarily transformed in the hyporheic zone, as has been the case in other field studies (Haggerty et al. 2008, 2009, Argerich et al. 2011, González-Pinzón et al. 2012, 2014), the Raz–Rru system indicates more GW-SW interactions in reach 3 than in reach 1.

From the normalized transformation of Raz in the hyporheic zone (Eq. 11) beneath the thalweg near S1, ΔRaz_{hz} was 0.1 at 3 cm, 0.6 at 6 cm, and 1.0 at 9 cm depth. As was the case for specific conductivity, Rru (the daughter product of Raz) was detected at 9 cm, whereas

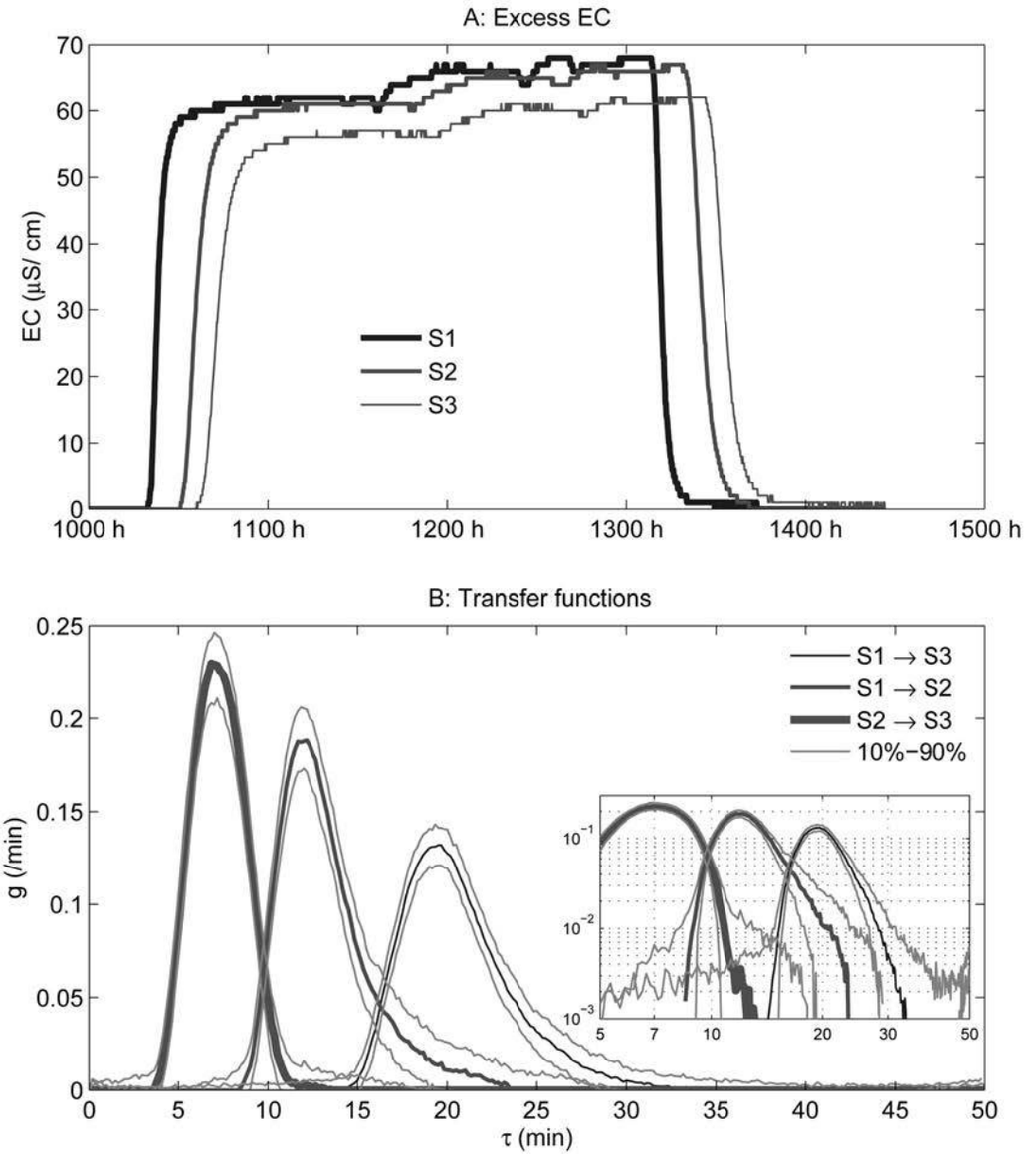

Figure 2. A.—Breakthrough curves of specific conductivity (EC) at observation points S1 (x = 125 m), S2 (x = 320 m), and S3 (x = 438 m) after subtracting the background values. B.—Transfer functions derived from the observed breakthrough curves. Dark lines used to describe transfer functions between sampling sites represent median values of 200 conditional realizations; lighter gray lines are 10th and 90th percentiles; inset: double-logarithmic plot of the transfer functions.

Raz was not detected. Because Raz was the only tracer that was not detected at this depth, we report ΔRaz_{hz} = 1.0. These estimates suggest that, on average, the entire Raz mass flowing upstream of S1 and entering the hyporheic zone would be transformed along the 9-cm-deep flow paths. On the other hand, from the normalized longitudinal transformation of Raz (Eq. 12), the maximum ΔRaz_{long} observed in the study site was $\Delta Raz_{x=430m}$ = 0.21. This result suggests that ~20% of the Raz mass crossing x = 50 m (the first well mixed location along the study reach) was transformed by x = 430 m.

Hydraulic-head measurements

Our ability to characterize hydraulic-head was limited by the small number of water-level loggers available, but we collected data from wells along 2 transects on the south side of the stream (3 wells on transect J [x = 180 m]

Table 3. Characteristic values of the transfer-functions g(τ) obtained by nonparametric deconvolution of NaCl breakthrough curves. μ_τ = mean residence time, σ_τ = standard deviation of residence time, $CV_\tau = \sigma_\tau/\mu_\tau$ = coefficient of variation, γ_τ = skewness of residence-time distribution. All values are calculated for the median of the conditional realizations.

			Residence-time distribution				
Subreach	Distance (m)	% recovery	μ_τ(min)	σ_τ (min)	CV_τ	γ_τ	γ_τ/CV_τ
S1–S3	313	89	20.7	3.12	0.15	0.94	6.26
S1–S2	195	96	13.2	2.59	0.20	1.15	5.75
S2–S3	118	90	7.3	1.55	0.21	6.46	30.76

and 5 wells on transect L [x = 145 m]; Fig. 1). Here we consider only the hydraulic gradients along the transects (perpendicular to the stream) to assess whether the stream was gaining or losing water. On transect J, our results indicate that water probably was flowing away from the stream between wells J8 and J6 (hydraulic gradient < 0), until ~1200 h noon on 11 June 2012 when flow was directed toward the stream. The latter conditions remained during our experiment (Fig. 4A). Further out along this transect, the hydraulic gradient strongly directed flow toward the channel. Along transect L, gradients were a bit more heterogeneous. Three of the 4 gradients calculated between each pair of the 5 wells along the transect were directed toward the stream, including the one adjacent to the channel (Fig. 4B). The gradient that was directed away from the stream corresponds to the 2nd and 3rd wells from the stream, possibly indicating a local preferential flow path within the aquifer.

FO-DTS

Groundwater temperatures were less variable than stream temperatures and cooler than surface temperatures most of the day. This setting provided reasonable, though not optimal conditions for detection using FO-DTS. Well monitoring at ~3 m depth showed that average groundwater temperature over the period of observation was 13.3°C and ranged from 12.5°C on 13 June to 14.3°C on 25 July 2012. Apparent air temperatures during FO-DTS deployment ranged from as low as 13.6°C at night to as high as 46°C during the day. Because the

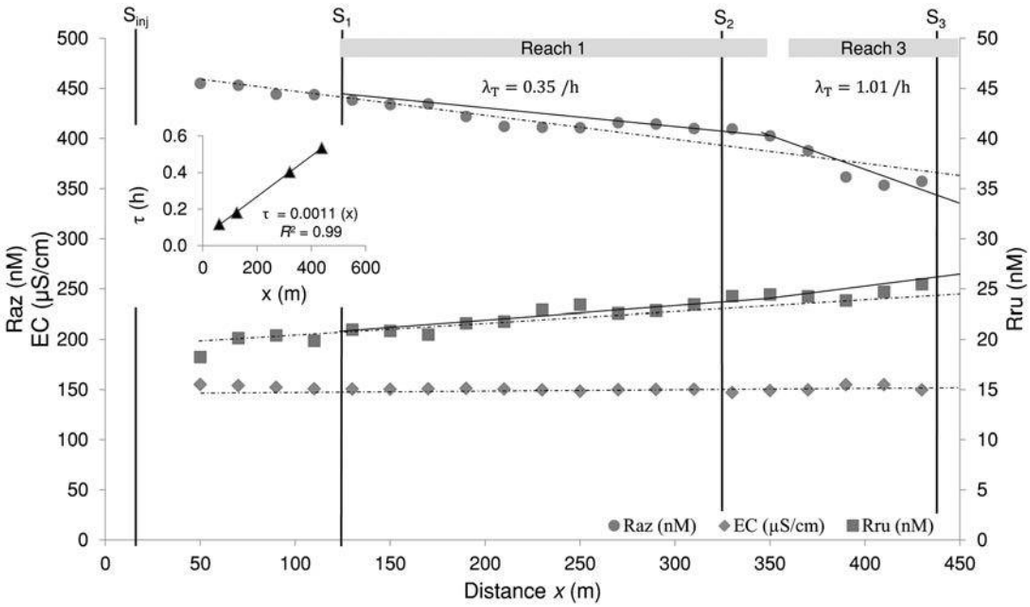

Figure 3. Longitudinal profile of resazurin (Raz) and resorufin (Rru) plateau concentrations along the experimental reach. Concentrations have been corrected for dilution. Differences in the slopes suggest spatial differences of in-stream processing. Discontinuous trend lines represent study site averages and continuous trend lines represent reach averages. Inset: mean travel times at selected locations were estimated by fitting a trend line to the mean travel times estimated from specific conductivity breakthrough curves at x = 60 m, S1, S2, and S3.

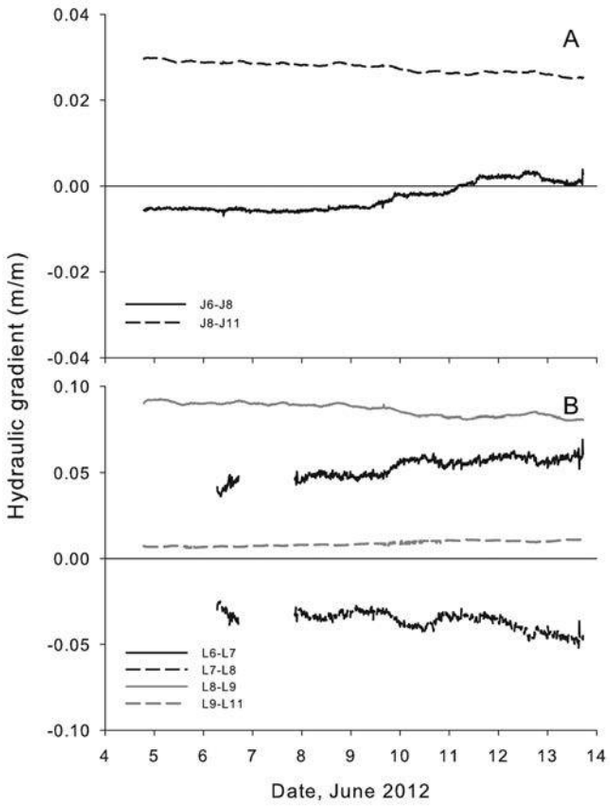

Figure 4. Time series of hydraulic gradients along the 2 well transects J (A) and L (B). J6 and L6 are adjacent to Shaver Creek, and J11 and L11 are at the distal edge of the transect, near the forest edge (Fig. 1). Gradients >0 are directed toward the stream, and <0 are directed away from the stream. Distances between wells are: J6–J8: 8.69 m, J8–J11: 34.48 m, L6–L7: 3.68 m, L7–L8: 3.19 m, L8–L9: 9.39 m, and L9–L11: 29.17 m.

fiber-optic cable was jacketed in black, the sun could heat it considerably above the true air temperature at sections where the cable was exposed to air (Neilson et al. 2010). Calibrated stream temperatures during deployment averaged ~17°C, with minimum water temperatures at night as low as 14°C and a midday maximum of 22.5°C (Figs 5, 6). Temperatures in most locations were similar at night, but solar radiation warmed the surface water bodies during the day, especially the small tributary that traveled a long distance across the unshaded dry lake bed (x = 320 m), the other tributaries (x = 5 m and x = 355 m), and near the downstream end of the reach (x = 438 m).

Temperatures increased in the downstream direction during the day because of solar radiation and interaction with warm air and decreased downstream at night as water exchanged heat with cooler air and lost heat by outgoing long-wave radiation. The same trend was present in the SDs of temperature along the cable (lowest variability upstream and progressively higher variability downstream; Fig. 6). Tributaries contained very small volumes of water relative to the main channel, so despite their significantly warmer temperature (Fig. 5), they did not contribute enough additional heat to the main channel to be distinguishable from other downstream heating processes (i.e., are not observable in Fig. 6).

Deviations from the general downstream heating trend might indicate groundwater upwelling or inflows, provided they were large enough in magnitude and distinct enough in temperature to change the main-stem temperature and to be detected above noise. One anomaly is present in the data at the downstream extreme of the reach, near x = 430–450 m. This location corresponds precisely to the location of the downstream calibration coil, which was held in place with a large rock. The weight of the rock probably impinged on the cable, resulting in a slight negative temperature anomaly at night and during the day. No other downstream cooling in the FO-DTS data was present that might indicate significant inflows from ground water. If such inflows existed, their magnitudes were not large enough to be detected by the resolution of FO-DTS. Based on calibration-bath temperatures over the study interval, the resolution and accuracy of FO-DTS temperatures for this deployment were estimated as 0.08 and 0.02°C, respectively.

Vertical streambed seepage fluxes from temperature profiles in piezometers

We used the average depth between pairs of temperature loggers as the representative depth for each flux

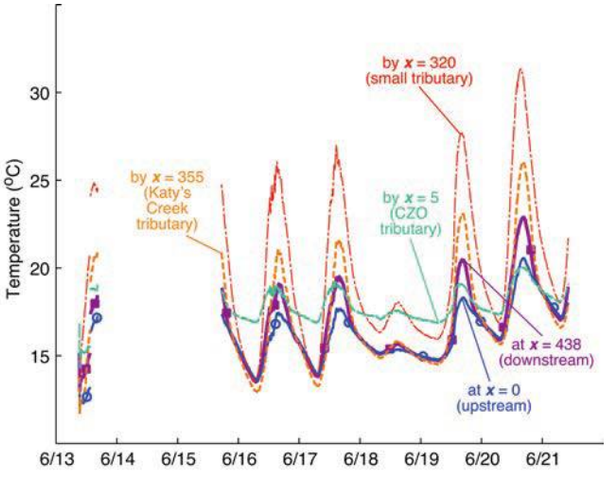

Figure 5. Average temperatures at specific locations in Shaver Creek and tributaries measured using fiber-optic distributed temperature sensing (FO-DTS) in June 2012. For improved accuracy at key locations, coils of cable were placed in individual locations from upstream to downstream at x = 0 m (upstream), x = 5 m (Critical Zone Observatory [CZO] tributary), x = 320 m (small tributary from south), x = 355 m (Katy's Creek tributary from north), and x = 438 m (downstream). Missing data from 13–15 June was a result of power loss.

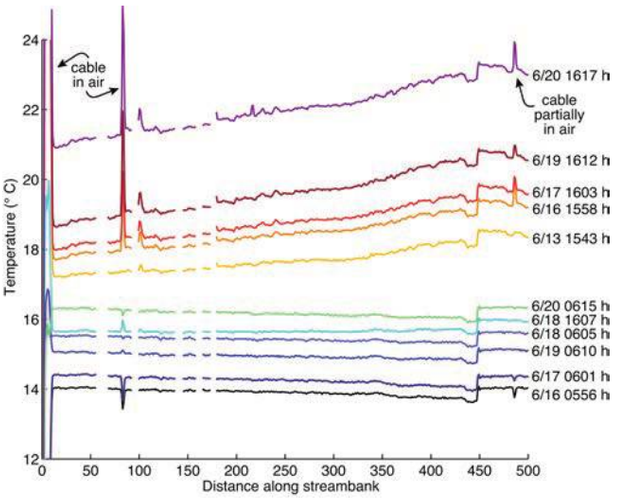

Figure 6. Hourly averaged longitudinal temperatures in Shaver Creek measured with fiber-optic distributed temperature sensing (FO-DTS) in June 2012. Temperature profiles were averaged temporally over the warmest and coolest hour of each day. The upstream end is x = 0 m, and downstream end is x = 438 m (the cable then loops back upstream). Gaps in data (from left to right) are locations where cable was out of the stream as it crossed electrical-resistivity-imaging transects N, M, L, K, and J. Spikes are locations where cable was out of the water completely or partially. Tributary temperatures are not shown.

(calculated from coupled heat and fluid transport), and plotted flux vs depth below the stream bed (Fig. 7A, B). The amplitude ratio provides direction and magnitude of flux (phase shift yields magnitude only), so we used the amplitude ratio to calculate fluxes for one piezometer installed upstream near the injection point (x = 13 m), and another installed mid-reach at S1 (x = 125 m) in the downstream pool of a pool–riffle–pool sequence.

Streambed seepage fluxes from vertical temperature profiles in piezometers agreed with FO-DTS observations near the S1 cross-section. No upward flux or inflow to the stream from ground water was observed along the streambed surface (represented by the top or shallowest portion of the seepage flux vs depth plot; Fig. 7A). Instead, all fluxes along the bed in the instrumented cross-section appeared to be negative or downward (i.e., losing stream conditions), and ranged from $q_{seepage} \approx -0.2$ m/d at S1 (riffle sequence, x = 125 m) to as high as $q_{seepage} \approx -1.0$ m/d just upstream of the injection point (x = 13 m). Toward the mid-point of the depth range of the vertical profiles, vertical fluxes decreased to near 0. At the deepest depths in the piezometers, we observed some upward fluid flow ranging from $q_{seepage} \approx 0.1$ m/d at S1 to as high as $q_{seepage} \approx 0.8$ m/d just upstream of the injection point. These differences in the flux direction along the vertical profile might be explained as sampling of different flow paths.

Electrical resistivity imaging of hyporheic transport

The electrical geophysical data were highly repeatable, and 90% of the data had an SD <2% of the mean. Inversions converged with final RMSE <1% for all background and time-lapse data sets. Our interpretation of the background images is that no notable lithologic boundaries or other geologic features can be identified in the images (Fig. 8A). As with all surface arrays, the resolution of the inversion is highest near the ground surface and decreases rapidly with depth (Fig. 8B). The shallow subsurface adjacent to the stream bed was highly resolved, and pixel values could be most uniquely determined in these locations. Evidence of hyporheic exchange was most clear in transect J (Fig. 1), the most upstream transect. Peak changes from background show a decrease of a few percent in electrical resistivity near the stream (Fig. 8C). The time-series of electrical resistivity at that pixel (highlighted in Fig. 8B) matched the temporal trend of in-stream solute tracer (Fig. 8D), and was interpreted as evidence of shallow hyporheic exchange. The peak cross-sectional hyporheic area was small relative to the total image, with an area <1 m^2

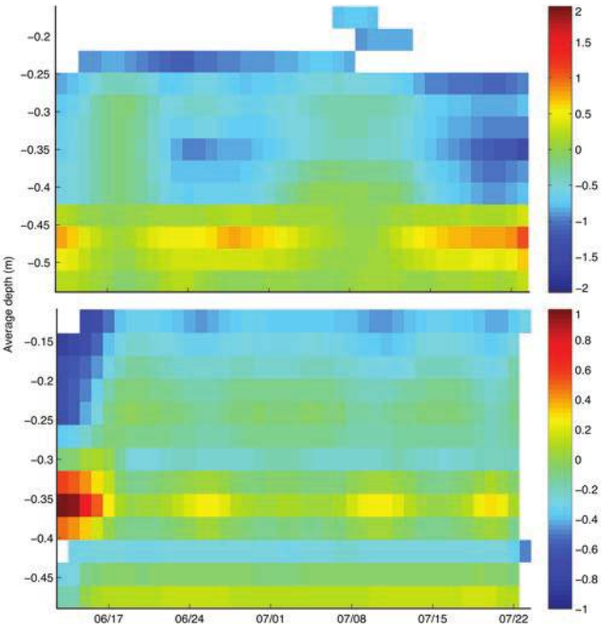

Figure 7. Estimates of vertical seepage rates and direction from amplitude and phase-shift analyses of time series of temperatures recorded at multiple depths in streambed piezometers at x = 13 m (A) and S1 (x = 125 m) (B). Amplitude ratios of deep/shallow temperatures from pairs of sensors were used to derive seepage rates (m/d), following the methods described by Hatch et al. (2006). Seepage rates were plotted vs average depth (m) between the 2 sensors (y-axis, positive up) and time (x-axis). Warm colors represent upwelling (m/d), green is 0 vertical flow, and blues indicate downward seepage. Areas missing data (in white) were flagged as having seepage values outside numerical or rational limits set by the fundamental equations.

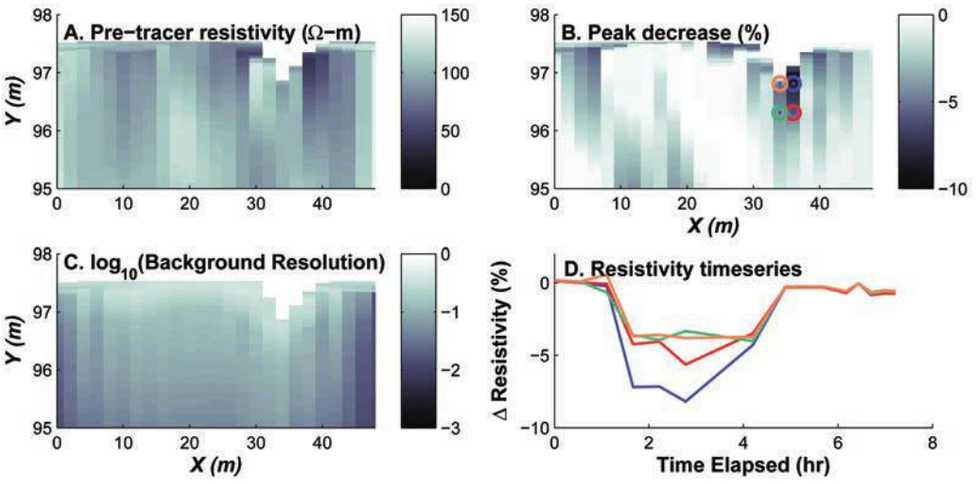

Figure 8. Electrical resistivity imaging results for transect J (single, representative transect results). A.—Background (pre-tracer) electrical resistivity. B.—Peak decrease in electrical resistivity during the tracer injection. C.—Resolution of the inversion on background (pre-tracer) data. D.—Time-series of electrical resistivity change for selected inversion pixels. *X* (horizontal or cross-sectional) and *Y* (vertical or into the ground) are dimensions of the 2-dimensional transect.

changing by >5%. Transect K was similar to transect J in pre-tracer resistivity, resolution, and hyporheic visualization. Transects L–N were visually similar in pre-tracer resistivity and resolution, but we did not observe decreasing resistivity near the stream.

DISCUSSION

Our field experiment included a host of techniques spanning spatial and temporal scales. Our study is, to the best of our knowledge, the first comparison of many techniques and their associated scales of observation in a single field campaign. The techniques that we co-applied and others can be used to investigate interactions at spatial scales ranging from 10^{-1} to 10^{5} m and can cover temporal scales of seconds to months. The techniques that can span more orders of magnitude in the spatial scale are: conservative tracers (10^{1}–10^{5} m) and distributed temperature sensing (10^{0}–10^{4} m). The techniques that are most sensitive to local scales are: mini-drive-point (10^{-1} m to 10^{0} m) sampling and streambed temperature vertical profile (10^{-1}–10^{0} m). The techniques that can span more orders of magnitude in the temporal scale are those involving temperature sampling (seconds–months). The techniques that span shorter temporal scales are: reactive Raz–Rru system (minutes–days) and electrical resistivity imaging (minutes–hours). In Table 4, we present a summary of the techniques we used and 2 others (seepage meter and differential discharge gauging) commonly used to investigate GW-SW interactions. We briefly describe the spatial and temporal scales covered and summarize their advantages and limitations. Below we discuss the interpretation of exchange processes based on our experimental results and explain the information gained from each method. We also evaluate the consistent or alternative interpretations supported by the results. The subsections are organized by spatial scale.

Local-scale analysis

We used MINIPOINT samplers and vertical temperature profiles close to S1 (at x = 123) to sample a subset of subsurface flow paths to quantify GW-SW interactions at the local (geomorphic unit or feature) scale. In combination, these techniques provided evidence of mass and heat fluxes across the stream bed. MINIPOINT samplers are ideally sized to capture fluxes at the uppermost interface between the stream and stream bed, where edge or boundary effects confound other methods (e.g., temperature or electrical resistivity surveys). Analysis of the MINIPOINT samples identified the existence of a shallow (~10 cm deep) hyporheic zone near the stream's thalweg, which decreases toward the banks of the stream because of groundwater discharge to the stream, as revealed by the analysis of hydraulic gradients (see below).

In this location, vertical temperature profiling provided independent estimates of vertical flux that generally agreed with MINIPOINT samples in delineating a shallow, but well connected hyporheic zone. Temperature profiling provided insight into complex flow geometries that may not have been apparent in net-flux or summed-exchange approaches (Fig. 7A, B).

Hyporheic flux was estimated at different depths for the MINIPOINT (3, 6, and 9 cm) and temperature-based

tracing (≥12.5 cm; Fig. 7B). Comparison of solute-tracing results at 9 cm and temperature-based tracing at 12.5 cm indicated hyporheic fluxes of the same order of magnitude (–0.5 and –0.2 m/d, respectively), suggesting agreement between techniques. The depth of hyporheic flow that was detected by shallower hyporheic sampling with MINIPOINT samplers (~10 cm) differed from depth of hyporheic flow detected by temperature profiling (~20 cm; Fig. 7B). The differences might be accounted for by: 1) the relatively short (3 h) injection, which may not have labeled the deeper hyporheic flow with solute tracer, whereas the longer-term temperature signal may have propagated more deeply into the subsurface; and 2) the natural variability in the stream (heterogeneity) because the measurements were not exactly co-located.

Last, the normalized transformation of Raz in the hyporheic zone (ΔRaz_{hz}, estimated from samples taken with the MINIPOINT samplers) provided insight on the relationship of hyporheic flow paths to biological processing. The magnitude of ΔRaz_{hz} increased with depth within the 9-cm vertical profile sampled, suggesting that the average particle of Raz that entered the hyporheic zone and traveled along the 9-cm-deep flow path was completely transformed to Rru. This local-scale finding has implications for the interpretation of reach-scale Raz–Rru results (see below).

Sub-reach-scale analysis

We monitored a subset of groundwater flow paths to identify interactions between the stream and the shallow aquifer occurring at scales larger than the MINIPOINT and vertical temperature profiles, but smaller than reach-scale studies. At this sub-reach scale, most hydraulic gradients were directed toward the stream, suggesting limited potential for hyporheic exchange caused by ground water pushing from the nearby aquifer (Fig. 4). Electrical geophysical transects J, K, M, and N were located near visible in-stream geomorphologic features, primarily in-channel bars and riffles, where hyporheic exchange was expected. However, images at these locations recorded limited evidence of hyporheic exchange (Fig. 8A, B). The interpretation of groundwater discharge to the stream, and limited hyporheic extent, was in agreement between hydraulic-gradient and electrical-geophysical methods.

The sub-reach-scale sampling techniques helped explain the shallowness of the hyporheic zone and its decrease toward the banks of the stream. These data provided a larger context within which the local-scale observations could be interpreted. This information was not provided by any other technique used. Well networks are expensive to install, but once in place, they can be used to monitor long-term GW-SW interactions (more easily than with other techniques), which is key to coupling flow dynamics and biogeochemical transformations.

Reach-scale analysis

We used NaCl as a conservative tracer and measured the breakthrough curves by logging specific conductivity over time. This technique is experimentally convenient, but conductivity provides limited analytical sensitivity because of its high background signal and natural temporal variability in most streams. Specific conductivity values are fairly high in streams and ground waters (~93 and ~306 μS/cm in Shaver Creek, respectively), so low tracer concentrations in the tail and transient-storage processes with long residence times may remain unnoticed (lost to "long-term storage" in the lexicon of Ward et al. 2013a, b). As a result, residence times estimated with tracers for which there is little dynamic range (e.g., NaCl) may be biased toward being too short. To better estimate long residence times, artificial pseudo-conservative tracers with low detection limits (e.g., fluorescent dyes) that allow a much greater dynamic range are preferred. The modeling techniques that we used for analyzing conservative transport (transient-storage model and deconvolution) can be applied to investigate GW-SW interactions regardless of the conservative tracer used. However, estimates of residence times that are biased short may result in apparently high processing rates, obscuring the realistic functioning of stream ecosystems (e.g., Frisbee et al. 2013). Our comparison of GW-SW interactions relies on the assumption that long-term residence times are not dominant. MINIPOINT sampling, vertical temperature analysis, and geophysical imaging all suggest that the hyporheic zone in our study reach was shallow (<20 cm, in the sampled sites). Thus, this assumption is reasonable.

The transient-storage-model analysis showed that reaches 1 (x = 125–350 m) and 3 (x =360–450 m) had different transient-storage dynamics (i.e., transient-storage metrics are ≥1 order of magnitude different along reach 3 than reach 1; Table 2). Transient-storage dynamics inferred from conservative tracers can be the result of surface or subsurface exchange processes. Thus, if we use only the conservative-tracer data at the reach scale we cannot objectively and decisively relate storage dynamics to GW-SW interactions. That is, we can link transient-storage dynamics and GW-SW interactions only by observing physical characteristics in the stream reaches. For instance, lack of surface recirculation zones because of side cavities, boulders, or woody debris complexes might indicate that subsurface processes may dominate storage dynamics. Rigorous analyses of transient-storage processes must be based on tracer injections with direct in situ sampling in putative storage zones.

We deconvolved the conservative-tracer breakthrough curves to estimate recovery rates and residence-time distributions between observation points. The inferred transfer functions exhibited only slight tailing beyond Fickian transport, but the ratio of the skewness of the residence-

Table 4. Review of techniques commonly used for estimating groundwater–surface-water interactions.

Method	Spatial scale	Temporal scale	Advantages	Limitations
Techniques used in our field study				
Conservative tracer tests (chemical)	10^1–10^5 m	min–mo	• Can assess flow loss and lateral inflow over an entire reach • Quantifies average storage volume and exchange rate at relatively large scales, although cannot necessarily distinguish type of storage (e.g. surface or subsurface)	• Tracer volume required may be restrictive for certain streams • Cannot distinguish subsurface flow from loss without additional measurements • May be affected by tracer adsorption • Point measurement in time (i.e., valid only for current stream flow) • Longer time-scale storage–exchange processes generally not observable • In-stream dispersion and hyporheic exchange are inseparable at long lengths
Reactive Raz–Rru tracer system	10^{-1}–10^3 m	min–d	• Distinct fluorescence properties with limits of detection as low as 10s of µg/L • Does not exist in nature: high sensitivity relative to natural background • Can be used to investigate metabolism in situ and in vivo at different spatial scales • Easy processing: field fluorometry, or lab-processed samples	• Reactive tracers are most reliable with a multihour stream tracer injection together with a conservative tracer • Organic compounds sorb to sediments • Fluorescence is quenched and sorption increased at pH < 6 • Costs of tracer may limit continuous tracer injections in streams with high discharge
Mini drive-point (USGS MINIPOINT)	10^{-1}–10^0 m	min–h	• Direct quantification of hyporheic tracer concentrations and chemical reactions in hyporheic flow • Provides fine vertical resolution and fast temporal resolution • Can distinguish and quantify stream-water and groundwater contributions to hyporheic flow	• Point measurement (both in space and time) • Labor intensive in field and often requires a simultaneous stream tracer injection for greatest value • Estimates of hyporheic fluxes assume vertical flow, which is not necessarily true
Piezometer (head)	10^0–10^2 m	h–mo	• Simple, robust assessment of hydraulic gradient • Allows estimates of temporal variability of seepage fluxes	• Point measurement (both space and time) • Labor intensive installation • Estimates of hyporheic fluxes assume vertical flow, which is not necessarily true
Streambed temperature (FO-DTS)	10^0–10^4 m	s–mo	• Distributed and rapid measurements over long spatial scales with a single calibration • When stream-water and groundwater temperatures are distinguishable, this method may be used to identify locations of groundwater inflows (gains)	• Generally cannot identify stream-water losses • Effective only when surface water and ground water are at different temperatures (time of day or seasonally)

	Streambed temperature (vertical drivepoint)	10^{-1}–10^{0} m	s–mo	• Relatively inexpensive for the accuracy and long-term records that can be collected • Combined with time-series or numerical analysis of thermographs, seepage rates and direction can be estimated • Measurements are independent of sedimentation and scour • Does not require knowledge or estimation of hydraulic conductivity • Provides time-series of seepage in systems with temporally varying seepage fluxes	• Point measurement (in space) • Generally does not distinguish groundwater fluxes from other types of subsurface flow such as hyporheic flow
	Electrical resistivity imaging	10^{-1}–10^{2} m	min–h	• Minimally invasive • Can collect continuous data in time automatically • Provides a 2- or 3-dimensional view of the subsurface • Data collection is relatively inexpensive	• Hardware is expensive • Averages over a larger area and time than co-located point measurements • Labor intensive to install initial electrodes and requires electric power • Requires multi-hour to multi-day stream tracer injection to quantify hyporheic flow • Inversion results are smoothed due to regularization • Conversion of bulk electrical conductivity to tracer concentration may be biased
	Other techniques not used in our field study				
	Seepage meter	10^{0}–10^{1} m	h–wk	• Inexpensive and easy to deploy • Can be configured for time-series measurements of seepage	• Flow can be induced around and into the meter • Not suitable for high-gradient, coarse-based stream beds • Point measurement (both in space and time) • Errors may be introduced by improper installation/deployment or if bag or meter outlet not properly shielded from stream shear stresses • Small-scale hyporheic fluxes (shallower than meter insertion depth) not quantified
	Differential discharge gauging (manual or automated)	10^{0}–10^{4} m	min–y	• Measures amount of water in stream directly • Simple water mass–balance calculation	• Does not measure hyporheic exchange, only net downstream flow • Does not work well in streams with complex morphologies • Labor intensive • Difficult when flows are low or turbulent • Must account separately for evapotranspiration and in/outflows

time distribution to the CV (γ_τ/CV_τ) can be used to quantify the strength of transient storage ($\gamma_\tau/CV_\tau = 1$ in Fickian transport) objectively. In our study, subreach S2–S3 (x = 320–438 m; $\gamma_\tau/CV_\tau = 30.76$) had greater transient storage than subreach S1–S2 (x = 125–320 m; $\gamma_\tau/CV_\tau = 5.75$) (Table 3). These results agree with the transient-storage-modeling analysis, even though the sites available for comparisons are slightly different.

Independent of the conservative-tracer analysis (i.e., transient-storage model and transfer functions), the Raz–Rru system showed a larger transformation of Raz along reach 3 than reach 1 (Fig. 3). If we assume that Raz was transformed primarily in the hyporheic zone because of the existence of enhanced chemical gradients and a larger volume of colonized sediments (the case in other field studies by Haggerty et al. 2008, 2009, Argerich et al. 2011, González-Pinzón et al. 2014), the Raz–Rru results suggest a larger extent of GW-SW interactions in reach 3 (x = 360–450 m) than in reach 1 (x = 125–350 m). In contrast to the limitations on using the conservative-tracer analysis to directly link storage dynamics and GW-SW interactions, the Raz–Rru system provides a direct measure of integrated, reach-scale metabolic reduction of material introduced from the stream. Therefore, it can be used to link local-scale observations of biogeochemical transformations to reach-scale transport and in-stream processing. In our study, the local transformation of Raz estimated from MINIPOINT samples showed that the potential to completely transform Raz existed in the shallow hyporheic zone, whereas $\sim 1/5^{th}$ of the Raz injected was transformed along the study site. This result highlights the importance of linking local- and reach-scale observations through estimates of processes across scales. Because Raz provides representative information on hyporheic fluxes at different scales along the study reach, it may help to address the question: when/where do GW-SW interactions matter for biogeochemical cycling and when/where they are less important? (Wondzell 2011, Briggs et al. 2013, Harvey et al. 2013).

FO-DTS also was sensitive to reach-scale GW-SW interactions. In our study, longitudinal FO-DTS suggested minimal locations of focused discharge of ground water (upwelling flow) or hyporheic water to the stream (Fig. 7A, B). This technique does not detect downwelling flow paths. FO-DTS did not detect upwelling flows, which indicates either that these flows were insufficient in magnitude to alter the stream water-column temperature within thermal detection limits (resolution and accuracy were estimated to be 0.08 and 0.02°C, respectively) or that these stream-directed flow paths mostly remained in the subsurface.

For the most part, transient-storage-model metrics, temporal moments from deconvolution analysis, apparent processing rates estimated with the Raz–Rru system, and results from FO-DTS agreed that differences exist in storage dynamics and GW-SW interactions along the reach.

Synthesis of all methods applied

A distinct limitation of local-scale measurements is their potential inability to represent conditions in the broader surrounding area. Limited sampling capacity systematically leads to biased understanding of processes. Therefore, a combination of methods may provide a more complete picture of stream complexity and, in turn, a better understanding of the extent of biochemical processing resulting from mass and heat exchange through GW-SW interactions. In this context, the techniques that we co-applied consistently indicated relatively limited exchange between surface water and ground water, except for a shallow zone in the stream bed. From the local to the reach scale, MINIPOINT sampling and vertical temperature profiles revealed that the hyporheic zone was confined to ~20 cm in the thalweg and that it vanished near the banks of the stream. Sub-reach-scale methods showed that the reason for the confined hyporheic zone was the hydraulic gradation toward the stream. Therefore, the tracer breakthrough curves recorded along the study reach did not exhibit long tails and the FO-DTS system did not detect substantial groundwater inputs along the reach. The lack of long tails in the breakthrough curves was detected with low- (specific conductivity [NaCl]) and high-resolution (fluorescence [Raz–Rru] with no background signal and a quantification limit <0.02 μg/L) tracer techniques.

Paradoxically, the increased information from multiple field methods generated both explanations and more questions. For instance, conservative- and reactive-tracer techniques detected differences in storage dynamics and probably the extent of GW-SW interactions between reaches 1 and 3. However, we did not instrument reach 3 with MINIPOINTs, vertical temperature sensors, wells, or electrical-resistivity equipment. Thus, even when the available data from tracers suggested that reach 3 may have more active mass exchange and biological processing than reach 1, we could not objectively identify the drivers of such exchanges. Many possible untracked drivers may explain the higher rates of exchange and processing along reach 3, for example: 1) the tributaries bring nutrients, O_2, and C fluxes that stimulate biological processing more strongly than stream geomorphology, 2) local hydraulic gradients along reach 3 favor extensive GW-SW interactions, 3) streambed sediments differ in ways that affect mass exchange and processing.

From the experience gained in our experiment, we recommend the following steps to decide which technique(s) should be implemented in a particular study: 1) clearly define the nature of the questions to be addressed (physical, biological, or chemical?), 2) identify the spatial and temporal scales to be covered explicitly and those required to provide an appropriate context for interpretation, and 3) engage in collaborative research efforts that maximize generation of mechanistic understanding and reduce the costs of implementing multiple techniques. Most studies

include only 1 or 2 techniques, often at different scales. We encourage collaborative work to enable application of multiple methods at field sites to the benefit of the research community.

Conclusions

We applied a number of field methods and data interpretation techniques in our study, seeking to address 2 primary questions: 1) *Do different techniques that are sensitive to processes on the same scale lead to similar interpretation?* We found this to be true at 3 scales. At the local scale, MINIPOINT sampling and vertical temperature profiles agreed that the hyporheic zone was shallow, and that the magnitudes of water fluxes through the stream bed were high. At the sub-reach scale, electrical resistivity imaging and hydraulic gradients showed that hydraulic gradation toward the stream constrained the hyporheic zone. At the reach scale, tracer methods and FO-DTS agreed that storage dynamics and GW-SW interactions differed among reaches. We conclude that different techniques sensitive to processes on the same scale lead to similar results. 2) *What can we gain from incorporating multiple techniques in a single study of GW-SW interactions*? In our study, the primary advance from using multiple methods was the context provided for the interpretation of results obtained by any method. Results from any method could be interpreted in isolation but carried a high level of uncertainty. Using multiple methods provided multiple lines of evidence to support our conceptual model of processes in the reach. Co-interpretation was useful across spatial scales where observed local-scale dynamics could be related to reach-scale results. Therefore, the combination of methods provided a more complete picture of stream complexity and, in turn, a better understanding of the extent of biochemical processing resulting from GW-SW interactions.

ACKNOWLEDGEMENTS

We are very thankful to all of the participants of the workshop "Techniques to Quantify Stream–Groundwater Exchange and Shallow Subsurface Transport: a Hands-on Workshop" for their enthusiastic engagement with the lectures, discussions, and field and laboratory activities. Financial support for this collaborative effort was provided by the National Science Foundation through grants EAR-0911435, EAR-0747629, EAR 08-38338, EAR-1331906, and IIA-1301346 and grant 07-53521 to the Consortium for the Advancement of Hydrologic Science (CUAHSI). OAC was also supported by a grant from the Ministry of Science, Research and Arts of Baden–Württemberg (AZ Zu 33-721.3-2), and the Helmholtz Centre for Environmental Research, Leipzig (UFZ). JH was also supported by the USGS Hydrologic Research and Development (HRD) and National Water Quality Assessment (NAWQA) programs. We sincerely appreciate CUAHSI for its support and help with the logistics, the Center for Transformative Monitoring Programs (CTEMPs) for providing the FO-DTS unit for this workshop free of charge, the participation of the USGS HRD and USGS NAWQA Programs, and Elizabeth Boyer and her students at Pennsylvania State University for in-kind support and use of laboratory facilities during the workshop. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. Last, we thank Aaron Packman, Scott Larned, and an anonymous referee for providing insightful comments that helped improve this manuscript.

LITERATURE CITED

Adams, R. K., and J. A. Spotila. 2005. The form and function of headwater streams based on field and modeling investigations in the southern Appalachian Mountains. Earth Surface Processes and Landforms 30:1521–1546.

Alexander, R. B., E. W. Boyer, R. A. Smith, G. E. Schwarz, and R. B. Moore. 2007. The role of headwater streams in downstream water quality. Journal of the American Water Resources Association 43:41–59.

Anderson, M. P. 2005. Heat as a ground water tracer. Groundwater 43:951–968.

Argerich, A., R. Haggerty, E. Martí, F. Sabater, and J. Zarnetske. 2011. Quantification of metabolically active transient storage (MATS) in two reaches with contrasting transient storage and ecosystem respiration. Journal of Geophysical Research 116:G03034.

Battin, T. J., L. A. Kaplan, S. Findlay, C. S. Hopkinson, E. Martí, A. I. Packman, J. D. Newbold, and F. Sabater. 2008. Biophysical controls on organic carbon fluxes in fluvial networks. Nature Geosciences 1:95–100.

Battin, T. J., S. Luyssaert, L. A. Kaplan, A. K. Aufdenkampe, A. Richter, and L. J. Tranvik. 2009. The boundless carbon cycle. Nature Geosciences 2:598–600.

Bencala, K. E., and R. A. Walters. 1983. Simulation of solute transport in a mountain pool-and-riffle stream: a transient storage model. Water Resources Research 19:718–724.

Besemer, K., H. Peter, J. B. Logue, S. Langenheder, E. S. Lindström, L. J. Tranvik, and T. J. Battin. 2012. Unraveling assembly of stream biofilm communities. ISME Journal 6:1459–1468.

Beven, K. 2006. A manifesto for the equifinality thesis. Journal of Hydrology 320:18–36.

Binley, A. 2013. Resistivity inversion software. (Available from: http://www.es.lancs.ac.uk/people/amb/Freeware/freeware.htm)

Boano, F., J. W. Harvey, A. Marion, A. I. Packman, R. Revelli, L. Ridolfi, and A. Wörman. 2014. Hyporheic flow and transport processes: mechanisms, models, and biogeochemical implications. Reviews of Geophysics. doi:10.1002/2012RG000417

Briggs, M. A., L. K. Lautz, D. K. Hare, and R. González-Pinzón. 2013. Relating hyporheic fluxes, residence times, and redox-sensitive biogeochemical processes upstream of beaver dams. Freshwater Science 32:622–641.

Briggs, M. A., L. K. Lautz, and J. M. McKenzie. 2012a. A comparison of fibre-optic distributed temperature sensing to traditional methods of evaluating groundwater inflow to streams. Hydrological Processes 26:1277–1290.

Briggs, M. A., L. K. Lautz, J. M. McKenzie, R. P. Gordon, and D. K. Hare. 2012b. Using high-resolution distributed temperature sensing to quantify spatial and temporal variability in vertical hyporheic flux. Water Resources Research 48: W02527.

Carslaw, H. S., and J. C. Jaeger. 1959. Conduction of heat in solids. Clarendon Press, Oxford, UK.

Cirpka, O. A., M. N. Fienen, M. Hofer, E. Hoehn, A. Tessarini, R. Kipfer, and P. K. Kitanidis. 2007. Analyzing bank filtration by deconvoluting time series of electric conductivity. Groundwater 45:318–328.

Conant, B. 2004. Delineating and quantifying ground water discharge zones using streambed temperatures. Groundwater 42:243–257.

Constantz, J. 2008. Heat as a tracer to determine streambed water exchanges. Water Resources Research 44:W00D10.

Day-Lewis, F. D., Y. Chen, and K. Singha. 2007. Moment inference from tomograms. Geophysical Research Letters 34:L22404.

Day-Lewis, F. D., K. Singha, and A. M. Binley. 2005. Applying petrophysical models to radar travel time and electrical resistivity tomograms: resolution-dependent limitations. Journal of Geophysical Research: Solid Earth 110(B8):B08206.

Dent, C. L., N. B. Grimm, and S. G. Fisher. 2001. Multiscale effects of surface–subsurface exchange on stream water nutrient concentrations. Journal of the North American Benthological Society 20:162–181.

Duan, Q., S. Sorooshian, and V. Gupta. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research 28:1015–1031.

Duff, J. H., F. Murphy, C. C. Fuller, F. J. Triska, J. W. Harvey, and A. P. Jackman. 1998. A mini drivepoint sampler for measuring pore water solute concentrations in the hyporheic zone of sand-bottom streams. Limnology and Oceanography 43:1378–1383.

Flemming, H.-C., and J. Wingender. 2010. The biofilm matrix. Nature Reviews Microbiology 8:623–633.

Frisbee, M. D., J. L. Wilson, J. D. Gomez-Velez, F. M. Phillips, and A. R. Campbell. 2013. Are we missing the tail (and the tale) of residence time distributions in watersheds? Geophysical Research Letters 40:1–5.

González-Pinzón, R., and R. Haggerty. 2013. An efficient method to estimate processing rates in streams. Water Resources Research 49:6096–6099.

González-Pinzón, R., R. Haggerty, and A. Argerich. 2014. Quantifying spatial differences in metabolism in headwater streams. Freshwater Science 33:798–811.

González-Pinzón, R., R. Haggerty, and M. Dentz. 2013. Scaling and predicting solute transport processes in streams. Water Resources Research 49:4071–4088.

González-Pinzón, R., R. Haggerty, and D. D. Myrold. 2012. Measuring aerobic respiration in stream ecosystems using the resazurin-resorufin system. Journal of Geophysical Research: Biogeosciences 117:G3.

Gooseff, M. N., and B. L. McGlynn. 2005. A stream tracer technique employing ionic tracers and specific conductance data applied to the Maimai catchment, New Zealand. Hydrological Processes 19:2491–2506.

Gordon, R. P., L. K. Lautz, M. A. Briggs, and J. M. McKenzie. 2012. Automated calculation of vertical pore-water flux from field temperature time series using the VFLUX method and computer program. Journal of Hydrology 420–421:142–158.

Goto, S., M. Yamano, and M. Kinoshita. 2005. Thermal response of sediment with vertical fluid flow to periodic temperature variation at the surface. Journal of Geophysical Research: Solid Earth 110:B1.

Haggerty, R., A. Argerich, and E. Martí. 2008. Development of a "smart" tracer for the assessment of microbiological activity and sediment-water interaction in natural waters: the resazurin-resorufin system. Water Resources Research 44:W00D01.

Haggerty, R., E. Martí, A. Argerich, D. von Schiller, and N. B. Grimm. 2009. Resazurin as a "smart" tracer for quantifying metabolically active transient storage in stream ecosystems. Journal of Geophysical Research 114:G03014.

Harvey, J. W., J. K. Böhlke, M. A. Voytek, D. Scott, and C. R. Tobias. 2013. Hyporheic zone denitrification: controls on effective reaction depth and contribution to whole-stream mass balance. Water Resources Research 49:1–19.

Harvey, J. W., and C. C. Fuller. 1998. Effect of enhanced manganese oxidation in the hyporheic zone on basin-scale geochemical mass balance. Water Resources Research 34:623–636.

Harvey, J. W., and B. J. Wagner. 2000. Quantifying hydrologic interactions between streams and their subsurface hyporheic zones. Pages 3–43 *in* J. A. Jones and P. J. Mulholland (editors). Streams and Ground Waters, Academic Press, San Diego.

Harvey, J. W., B. J. Wagner, and K. E. Bencala. 1996. Evaluating the reliability of the stream tracer approach to characterize stream-subsurface water exchange. Water Resources Research 32:2441–2451.

Hatch, C. E., A. T. Fisher, J. S. Revenaugh, J. Constantz, and C. Ruehl. 2006. Quantifying surface water–groundwater interactions using time series analysis of streambed thermal records: method development. Water Resources Research 42: W10410.

Hatch, C. E., A. T. Fisher, C. R. Ruehl, and G. Stemler. 2010. Spatial and temporal variations in streambed hydraulic conductivity quantified with time-series thermal methods. Journal of Hydrology 389:276–288.

Hausner, M. B., F. Suárez, K. E. Glander, N. van de Giesen, J. S. Selker, and S. W. Tyler. 2011. Calibrating single-ended fiber-optic Raman spectra distributed temperature sensing data. Sensors 11:10859–10879.

Healy, R. W. 2010. Estimating groundwater recharge. Cambridge University Press, Cambridge, UK.

Kalbus, E., F. Reinstorf, and M. Schirmer. 2006. Measuring methods for groundwater–surface water interactions: a review. Hydrology and Earth System Sciences 10:873–887.

Kelleher, C., T. Wagener, B. McGlynn, A. S. Ward, M. N. Gooseff, and R. A. Payn. 2013. Identifiability of transient storage model parameters along a mountain stream. Water Resources Research 49:5290–5306.

Keller, G. V., and F. C. Frischknecht. 1966. Electrical methods in geophysical prospecting. Pergamon Press, Oxford, UK.

Kemna, A., B. Kulessa, and H. Vereecken. 2002. Imaging and characterisation of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models. Journal of Hydrology 267:125–146.

Kiel, B. A., and M. B. Cardenas. 2014. Lateral hyporheic exchange throughout the Mississippi River network. Nature Geoscience 7:413–417.

Krause, S., T. Blume, and N. J. Cassidy. 2012. Investigating patterns and controls of groundwater up-welling in a lowland

river by combining fibre-optic distributed temperature sensing with observations of vertical head gradients. Hydrology and Earth System Sciences 16:1775–1792.

Kreft, A., and A. Zuber. 1978. On the physical meaning of the dispersion equation and its solutions for different initial and boundary conditions. Chemical Engineering Science 33:1471–1480.

LaBrecque, D., and X. Yang. 2001. Difference inversion of ERT data: a fast inversion method for 3-D in situ monitoring. Journal of Environmental and Engineering Geophysics 6:83–89.

Legleiter, C. J., D. A. Roberts, W. A. Marcus, and M. A. Fonstad. 2004. Passive optical remote sensing of river channel morphology and in-stream habitat: physical basis and feasibility. Remote Sensing of Environment 93:493–510.

Naranjo, R. C., G. Pohll, R. G. Niswonger, M. Stone, and A. Mckay. 2013. Using heat as a tracer to estimate spatially distributed mean residence times in the hyporheic zone of a riffle-pool sequence. Water Resources Research 49:3697–3711.

Neilson, B. T., C. E. Hatch, H. Ban, and S. W. Tyler. 2010. Solar radiative heating of fiber-optic cables used to monitor temperatures in water. Water Resources Research 46:W08540.

Payn, R. A., M. N. Gooseff, D. A. Benson, O. A. Cirpka, J. P. Zarnetske, W. B. Bowden, J. P. McNamara, and J. H. Bradford. 2008. Comparison of instantaneous and constant-rate stream tracer experiments through non-parametric analysis of residence time distributions. Water Resources Research 44:W06404.

Peterson, B. J., W. M. Wollheim, P. J. Mulholland, J. R. Webster, J. L. Meyer, J. L. Tank, E. Martí, W. B. Bowden, H. M. Valett, and A. E. Hershey. 2001. Control of nitrogen export from watersheds by headwater streams. Science 292:86–90.

Rau, G. C., M. S. Andersen, A. M. McCallum, H. Roshan, and R. I. Acworth. 2014. Heat as a tracer to quantify water flow in near-surface sediments. Earth-Science Reviews 129:40–58.

Rosenberry, D. O., and J. W. LaBaugh. 2008. Field techniques for estimating water fluxes between surface water and ground water. Techniques and Methods 4–D2. US Geological Survey, Reston, Virginia.

Runkel, R. L. 1998. One-dimensional transport with inflow and storage (OTIS): a solute transport model for streams and rivers. Water-Resources Investigations Report 98-4018. US Geological Survey, Reston, Virginia.

Runkel, R. L. 2007. Toward a transport-based analysis of nutrient spiraling and uptake in streams. Limnology and Oceanography: Methods 5:50–62.

Schmadel, N. M., B. T. Neilson, T. Kasahara. 2013. Deducing the spatial variability of exchange within a longitudinal channel water balance. Hydrological Processes 28:3088–3103.

Selker, J. S., L. Thevenaz, H. Huwald, A. Mallet, W. Luxemburg, N. Van de Giesen, M. Stejskal, J. Zeman, M. Westhoff, and M. B. Parlange. 2006. Distributed fiber-optic temperature sensing for hydrologic systems. Water Resources Research 42: W12202.

Singha, K., A. Pidlisecky, F. D. Day-Lewis, and M. N. Gooseff. 2008. Electrical characterization of non-Fickian transport in groundwater and hyporheic systems. Water Resources Research 44:W00D07.

Slater, L., A. Binley, R. Versteeg, G. Cassiani, R. Birken, and S. Sandberg. 2002. A 3D ERT study of solute transport in a large experimental tank. Journal of Applied Geophysics 49:211–229.

Stallman, R. W. 1965. Steady one-dimensional fluid flow in a semi-infinite porous medium with sinusoidal surface temperature. Journal of Geophysical Research: Solid Earth 70: 2821–2827.

Stewart, P. S., and M. J. Franklin. 2008. Physiological heterogeneity in biofilms. Nature Reviews Microbiology 6:199–210.

Stonestrom, D. A., and J. Constantz. 2003. Heat as a tool for studying the movement of ground water near streams. Circular 1260. US Geological Survey, Reston, Virginia.

Stream Solute Workshop. 1990. Concepts and methods for assessing solute dynamics in stream ecosystems. Journal of the North American Benthological Society 9:95–119.

Swanson, T. E.. and M. Bayani Cardenas. 2010. Diel heat transport within the hyporheic zone of a pool-riffle-pool sequence of a losing stream and evaluation of models for fluid flux estimation using heat. Limnology and Oceanography 55:1741–1754.

Toride, N., F. J. Leij, and M. T. van Genuchten. 1993. A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production. Water Resources Research 29:2167–2182.

Van de Giesen, N., S. C. Steele-Dunne, J. Jansen, O. Hoes, M. B. Hausner, S. Tyler, and J. Selker. 2012. Double-ended calibration of fiber-optic Raman spectra distributed temperature sensing data. Sensors 12:5471–5485.

Wagener, T., L. Camacho, and H. Wheater. 2002. Dynamic identifiability analysis of the transient storage model for solute transport in rivers. Journal of Hydroinformatics 4:199–211.

Wagner, B. J., and J. W. Harvey. 1997. Experimental design for estimating parameters of rate-limited mass transfer: analysis of stream tracer studies. Water Resources Research 33:1731–1741.

Ward, A. S., M. Fitzgerald, M. N. Gooseff, T. J. Voltz, A. M. Binley, and K. Singha. 2012a. Correction to "Hydrologic and geomorphic controls on hyporheic exchange during baseflow recession in a headwater mountain stream". Water Resources Research 48:W08903.

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

Ward, A. S., M. N. Gooseff, and K. Singha. 2010. Imaging hyporheic zone solute transport using electrical resistivity. Hydrological Processes 24:948–953.

Ward, A. S., M. N. Gooseff, T. J. Voltz, M. Fitzgerald, K. Singha, and J. P. Zarnetske. 2013a. How does rapidly changing discharge during storm events affect transient storage and channel water balance in a headwater mountain stream? Water Resources Research 49:5473–5486.

Ward, A. S., R. A. Payn, M. N. Gooseff, B. L. McGlynn, K. E. Bencala, C. A. Kelleher, S. M. Wondzell, and T. Wagener. 2013b. Variations in surface water-ground water interactions along a headwater mountain stream: comparisons between transient storage and water balance analyses. Water Resources Research 49:3359–3374.

Westhoff, M. C., M. N. Gooseff, T. A. Bogaard, and H. H. G. Savenije. 2011. Quantifying hyporheic exchange at high spatial resolution using natural temperature variations along a first-order stream. Water Resources Research 47: W10508.

Wlostowski, A. N., M. N. Gooseff, and T. Wagener. 2013. Influence of constant rate versus slug injection experiment type on parameter identifiability in a 1-D transient storage model for stream solute transport. Water Resources Research 49: 1184–1188.

Wondzell, S. M. 2011. The role of the hyporheic zone across stream networks. Hydrological Processes 25:3525–3532.

Wörman, A., A. I. Packman, L. Marklund, J. W. Harvey, and S. H. Stone. 2007. Fractal topography and subsurface water flows from fluvial bedforms to the continental shield. Geophysical Research Letters 34:L07402.