WATER RESOURCES RESEARCH, VOL. 33, NO. 7, PAGES 1731–1741, JULY 1997

Experimental design for estimating parameters of rate-limited mass transfer: Analysis of stream tracer studies

Brian J. Wagner
U.S. Geological Survey, Menlo Park, California

Judson W. Harvey
U.S. Geological Survey, Reston, Virginia

Abstract. Tracer experiments are valuable tools for analyzing the transport characteristics of streams and their interactions with shallow groundwater. The focus of this work is the design of tracer studies in high-gradient stream systems subject to advection, dispersion, groundwater inflow, and exchange between the active channel and zones in surface or subsurface water where flow is stagnant or slow moving. We present a methodology for (1) evaluating and comparing alternative stream tracer experiment designs and (2) identifying those combinations of stream transport properties that pose limitations to parameter estimation and therefore a challenge to tracer test design. The methodology uses the concept of global parameter uncertainty analysis, which couples solute transport simulation with parameter uncertainty analysis in a Monte Carlo framework. Two general conclusions resulted from this work. First, the solute injection and sampling strategy has an important effect on the reliability of transport parameter estimates. We found that constant injection with sampling through concentration rise, plateau, and fall provided considerably more reliable parameter estimates than a pulse injection across the spectrum of transport scenarios likely encountered in high-gradient streams. Second, for a given tracer test design, the uncertainties in mass transfer and storage-zone parameter estimates are strongly dependent on the experimental Damkohler number, DaI, which is a dimensionless combination of the rates of exchange between the stream and storage zones, the stream-water velocity, and the stream reach length of the experiment. Parameter uncertainties are lowest at DaI values on the order of 1.0. When DaI values are much less than 1.0 (owing to high velocity, long exchange timescale, and/or short reach length), parameter uncertainties are high because only a small amount of tracer interacts with storage zones in the reach. For the opposite conditions ($DaI \gg 1.0$), solute exchange rates are fast relative to stream-water velocity and all solute is exchanged with the storage zone over the experimental reach. As DaI increases, tracer dispersion caused by hyporheic exchange eventually reaches an equilibrium condition and storage-zone exchange parameters become essentially nonidentifiable.

Introduction

The dynamics of solute transport play a critical role in determining the fate of pollutants in rivers and streams. Numerous studies have shown that there can be significant exchange between the active channel and storage zones in pools and eddies near the sides of the channel or in subsurface hyporheic flow paths [e.g., *Bencala*, 1984; *Wallis et al.*, 1989; *D'Angelo et al.*, 1993]. Storage processes increase the solute retention time in channels and the contact of stream-water solutes with sediment, which stimulates biotic and geochemical processes that affect solute reaction during downstream transport [*Grimm and Fisher*, 1984; *Kim et al.*, 1992; *Rutherford et al.*, 1995; *Runkel et al.*, 1996].

The stream tracer experiment has become a widely used tool for analyzing the transport characteristics of complex stream systems. In a typical stream tracer experiment a tracer-labelled solution is injected into the stream and solute concentrations are sampled at downstream locations. The tracer-experiment data are then combined with solute transport simulation to quantify the physical process parameters that characterize solute advection, dispersion, lateral inflow of groundwater, and exchange with storage zones [*Stream Solute Workshop*, 1990]. The parameters estimated from stream tracer simulations are not always reliable for two reasons. First, tracer experiments are not uniquely sensitive to surface or subsurface (hyporheic) storage processes [*Harvey et al.*, 1996], which limits physical interpretations and transferability of results; second, difficulties are sometimes encountered quantifying storage processes with acceptable precision. In this paper we address the second issue of designing stream tracer experiments to produce more precise estimates of storage parameters.

The techniques of experimental design [see *Steinberg and Hunter*, 1984] have previously been used to design aquifer tests to reliably estimate the physical process parameters defining groundwater flow and contaminant transport [e.g., *Nishikawa and Yeh*, 1989; *McCarthy and Yeh*, 1990; *Cleveland and Yeh*, 1990, 1991]. Of particular interest to the present study is the work of *Cleveland and Yeh* [1990, 1991], who present two

Paper number 97WR01067.

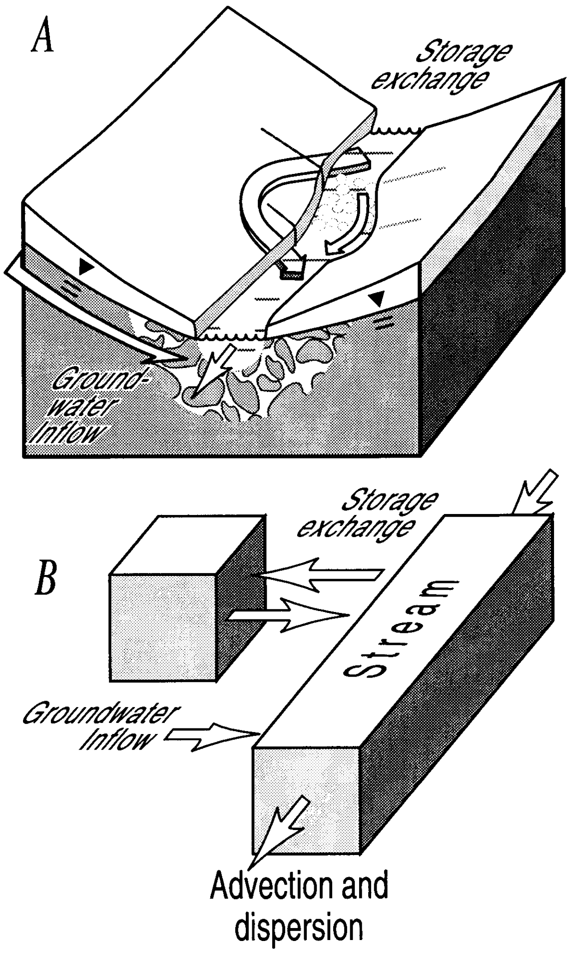

Figure 1. Hydrologic interactions between stream and subsurface. (a) Physical system indicating stream-water flow, groundwater inflow, and exchange between stream and storage zones. (b) Reach-scale modeling of stream tracer injection subject to advection, dispersion, groundwater inflow, and exchange with storage zones.

optimization algorithms for designing a monitoring network for an aquifer tracer test. The optimization models select the monitoring strategy that provides minimum parameter uncertainty. The resulting designs are intuitively reasonable and can be related to the idea of model sensitivity analysis: The informative data for parameter estimation are in general those data that exhibit high sensitivity to the model parameters. Although techniques of optimization and statistics have been applied to estimate stream transport parameters [*Wagner and Gorelick*, 1986; *Hart*, 1995; *Green et al*., 1994; *Harvey et al*., 1996], to our knowledge there have been no studies that formally apply the techniques of experimental design to stream tracer test design.

One aspect common to many tracer experiments is that they rely on prior estimates of the system parameters which, in fact, the experiment is being designed to estimate. This is sometimes referred to as the paradox of experimental design [*Moss*, 1979]. The results of *Cleveland and Yeh* [1990, 1991] illustrate the design paradox. Through postoptimization Monte Carlo analysis, Cleveland and Yeh demonstrate that the design of an "optimal" sampling strategy for their groundwater tracer experiment can be sensitive to poor prior estimates of the unknown transport parameters. For the design of many tracer experiments, the best we can hope to do is define broad ranges for parameters that bracket the likely values. Therefore a successful approach to experimental design must be able to account for a wide range of scenarios that might be encountered in the field.

We present here a methodology for evaluating and comparing alternative tracer experiment designs for use in high-gradient stream systems (slope greater than 1%) where exchange between the active channel and storage zones is important. The goal is to develop a robust experimental design that considers the limits of parameter reliability and parameter identifiability that are attainable given different combinations of actual system parameters. The methodology is based on the concept of global parameter uncertainty analysis, which combines solute transport simulation with parameter uncertainty analysis in a Monte Carlo framework. On the basis of prior information, we generate many realizations of the physical process parameters, with each realization representing a possible "model" of the true stream-aquifer system. Then, for a series of alternative tracer experiment designs (i.e., a combination of tracer injection and tracer sampling strategies) a global parameter uncertainty analysis is performed to analyze parameter reliability for each parameter realization and for each alternative design. The result is a suite of parameter covariance matrices that are used to analyze and compare tracer experiment designs over the spectrum of possible transport scenarios. Here we use the experimental Damkohler number to quantify the limitations of parameter identifiability and experimental design.

Methodology

In this study we focus on designing tracer experiments to characterize the physical transport properties of a stream. In the basic stream tracer study, a conservative tracer-labelled solution is injected into the stream. As the tracer mass moves downstream, it is acted upon by the four basic transport processes described in Figure 1: (1) advection, which describes the rate at which the tracer body moves downstream; (2) dispersion, which accounts for the mixing processes in the stream that cause the tracer body to spread; (3) groundwater inflow, which serves to increase the rate of flow and to dilute the tracer; and (4) storage-zone exchange, which describes the movement of solute between the active channel and stagnant or slowly moving zones in the stream or in the subsurface. At some point downstream the tracer is sampled, providing a history of in-stream tracer concentrations. The goal is to characterize the transport properties of the stream on the basis of the observed tracer concentration history.

Modeling Stream Tracer Experiments

A solute transport simulation model provides a means to quantitatively link the observed tracer concentrations to the four transport processes described above. In this conceptualization the hydrologic regime is divided into two coupled systems: a system of flowing water in the main stream channel and a system of storage zones at the margins of the stream channel or in the subsurface that contain slowly moving or immobile water (Figure 1). The two systems are coupled by a simple mass-transfer formulation that exchanges solutes between the main channel and storage zones [*Bencala and Walters*, 1983]. The model for one-dimensional (1-D) advective-dispersive transport with inflow and storage-zone exchange is

$$\frac{\partial C}{\partial t} = -\frac{Q}{A}\frac{\partial C}{\partial x} + \frac{1}{A}\frac{\partial}{\partial x}\left(AD\frac{\partial C}{\partial x}\right) + \frac{q_L}{A} \cdot (C_L - C) + \alpha(C_s - C) \quad (1)$$

$$\frac{\partial C_s}{\partial t} = -\alpha\frac{A}{A_s}(C_s - C) \quad (2)$$

where

- C solute concentration in the stream (mg L^{-1});
- Q volumetric flow rate (m^3 s^{-1});
- A cross-sectional area of stream channel (m^2);
- D dispersion coefficient (m^2 s^{-1});
- q_L lateral volumetric groundwater inflow rate (per length) (m^3 s^{-1} m^{-1});
- C_L solute concentration in the lateral inflow (mg L^{-1});
- C_s solute concentration in the storage zone (mg L^{-1});
- A_s cross-sectional area of the storage zone (m^2);
- α stream-storage exchange coefficient (s^{-1});
- t time (s);
- x distance (m).

The finite difference method described by *Runkel and Chapra* [1993] was used to solve (1) and (2).

To simulate solute transport using (1) and (2), the model parameters, Q, A, D, q_L, C_L, α, and A_s, must be specified. Since direct measurement of these parameters (other than Q) is difficult or impossible, the parameters must be estimated by "fitting" the model to solute concentration data obtained in stream tracer experiments. The "best-fit" parameter estimates can be obtained through manual calibration [see *Stream Solute Workshop*, 1990], or a more rigorous statistics and optimization approach can be used [*Wagner and Gorelick*, 1986; *Hart*, 1995; *Green et al.*, 1994].

If the parameters are estimated precisely, the simulation model can be used with confidence to analyze solute migration and redistribution in stream-aquifer systems. In reality, there will always be a degree of uncertainty associated with the parameter estimates obtained from experimental data. Consequently, any model-based analyses and interpretation of transport processes will also be uncertain. The usefulness of a stream tracer experiment largely depends on the reliability of the model parameter estimates.

Estimating Parameter Reliability

The parameter estimate covariance matrix provides a quantitative measure of the reliability of model parameters and forms the basis of the methodology for evaluating and comparing alternative tracer experiment designs. The first-order approximation to the parameter estimate covariance matrix, $\mathbf{V}_p$, is [*Draper and Smith*, 1981]

$$\mathbf{V}_p = (\mathbf{J}^t\mathbf{V}_c^{-1}\mathbf{J})^{-1} \quad (3)$$

where $\mathbf{V}_c$ is the covariance that defines the uncertainty in the concentration data and $\mathbf{J}$ is the Jacobian, the matrix of sensitivities of modeled concentrations with respect to changes in the model parameters. The parameter estimate covariance matrix can be used to identify parameters that are well (or poorly) estimated and, as we will demonstrate here, to evaluate data worth and data needs when designing stream tracer experiments.

The concentration sensitivities found in the Jacobian matrix, $\mathbf{J}$, in (3) play an important role in defining the worth of data for tracer experiment design. The sensitivity is the partial derivative of modeled stream tracer concentration with respect to a change in the value of a parameter

$$J_{ij} = \frac{\partial C_i}{\partial p_j} \quad (4)$$

where J_{ij} is the sensitivity of modeled stream tracer concentration C_i to the parameter p_j. The concentration sensitivities for a hypothetical stream system are shown in Figure 2. Shown are the sensitivities $\partial C/\partial D$, $\partial C/\partial A$, $\partial C/\partial q_L$, $\partial C/\partial A_s$, and $\partial C/\partial \alpha$ for the cases injection to plateau and pulse injection. Superimposed on the sensitivity plots is the concentration breakthrough curve.

From inspection of (3) and the concentration sensitivity plots in Figure 2, the following principles of tracer experiment design emerge. First, information about a parameter is, in general, most efficiently gained by sampling at points with high sensitivity to the parameter. Concentration sensitivities define the elements of the Jacobian matrix in (3). By increasing the sensitivities, we can reduce the parameter covariance. Second, the concentration history is sensitive to the parameters over relatively narrow time spans that are associated with specific segments of the concentration history; therefore, data that provide information for estimating one parameter may provide little or no information for estimating other parameters. Third, the choice of experimental design can affect the amount of information contained in the data. For the stream cross-sectional area, A, and dispersion coefficient, D, the maximum sensitivities are associated with the concentration fronts, and they appear to be of similar magnitudes for both injection to plateau and pulse injection. Likewise, for groundwater inflow, q_L, maximum sensitivities for pulse injection are associated with the concentration fronts. However, for the case of plateau injection, additional information for estimating q_L can be gained from the plateau segment of the concentration history where the sensitivities are nonzero. In the case of the storage zone cross-sectional area, A_s, and exchange coefficient, α, the informative data are found on both the shoulder and tail portions of the concentration history. The sensitivity plots suggest the pulse injection strategy will have less information than plateau injection for estimating A_s.

From the above discussion it might appear that determining the "best" tracer experiment design is simply a process of choosing the injection/sampling strategy that provides high sensitivity to the parameters. Although high sensitivities are desirable, there are three reasons why this approach will not always be successful. First, the concentrations defined by (1) and (2) are nonlinear with respect to the parameters. The sensitivity plots presented in Figure 2 represent local derivatives of concentration with respect to these parameters. Therefore an analysis of the covariance matrix (3) is dependent on the parameter values used to calculate the sensitivities (4). If the parameter values change, the sensitivities and covariance will also change. An injection/sampling strategy that provides reliable parameter estimates for one stream scenario (i.e., one set of transport properties) may be incapable of identifying the parameters in another scenario. Consequently, the design of an "optimal" tracer test will be based on assumptions regarding the parameter values that, paradoxically, the tracer test is being designed to estimate. Second, as we will show later, the issue of parameter nonidentifiability is a concern when designing tracer experiments. Simply stated, nonidentifiability occurs when different parameter values result in the same model output. In this case, data that exhibit seemingly high sensitivity to the

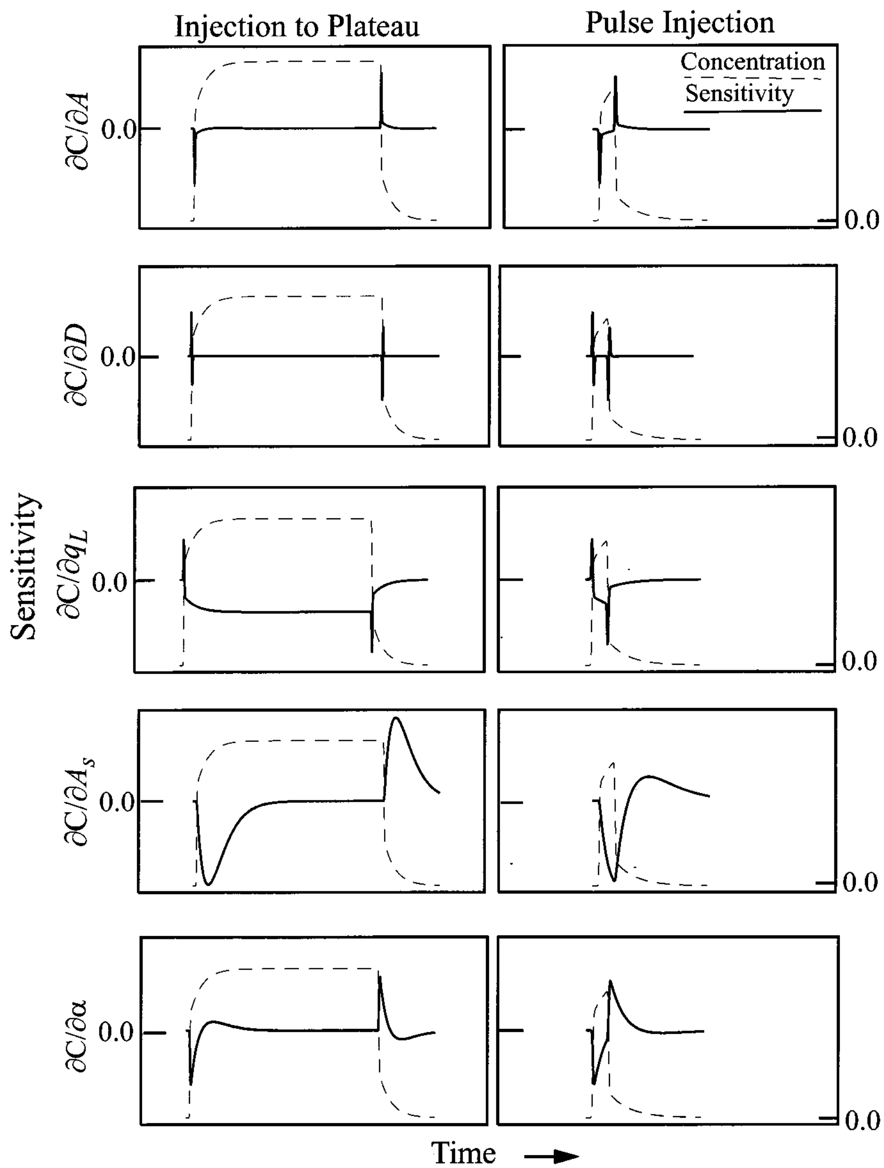

Figure 2. Sensitivity plots for injection to plateau and pulse injection. These figures show the segments of the concentration history that provide information for estimating stream cross-sectional area, $\partial C/\partial A$; dispersion coefficient, $\partial C/\partial D$; groundwater inflow, $\partial C/\partial q_L$; storage zone cross-sectional area, $\partial C/\partial A_s$; and stream-storage exchange coefficient, $\partial C/\partial \alpha$.

parameters will nonetheless be incapable of identifying those parameters, regardless of how much data are collected. In this study we adopt the approach of *McLaughlin and Townley* [1996] and assume that parameters are identifiable if they can be estimated with acceptable reliability. Finally, parameter uncertainty is a function not only of sensitivity but also of the number of data and the precision with which those data can be measured. Therefore there is a three-way trade-off between sensitivity, data precision, and number of data that must be analyzed when evaluating tracer study designs.

Evaluating and Comparing Alternative Tracer Experiment Designs

A natural question to ask when analyzing stream tracer data is, Are the parameter estimates sufficiently reliable to permit the analysis of contaminant transport? This question could be rephrased to ask, How can a tracer experiment be designed to ensure reliable parameter estimates? The difference between these two questions is the difference between analyzing stream tracer data and designing stream tracer experiments. Whereas data analysis evaluates a tracer experiment after it has been performed, experimental design evaluates the efficiency of alternative tracer experiment designs prior to performing the experiment. In this study we employ techniques of experimental design to evaluate and compare tracer test designs for high gradient stream systems. The methodology consists of five stages and is diagrammed in Figure 3.

The first stage involves defining the prior information regarding the stream transport parameters. Inspection of the governing mathematical model (equations (1) and (2)) indicates there are seven stream properties that must be determined in order to analyze the transport of solutes in stream-aquifer systems. They are stream discharge, Q; stream cross-sectional area, A; dispersion coefficient, D; groundwater inflow, q_L; groundwater concentration, C_L; stream-storage exchange coefficient, α; and storage zone cross-sectional area, A_s. These seven transport properties can vary widely from stream to stream. Moreover, there can be considerable variation along a stream and even seasonally for a given stream reach. In practice, the prior knowledge of these parameters will be limited, particularly for the storage-zone exchange parameters α and A_s. The first stage of the design methodology involves quantifying the prior parameter information; for example, lower and upper bounds on the parameter values can be set on the basis of prior investigations of the system and/or experience with similar systems.

The prior parameter information defines a spectrum of possible stream transport scenarios. The next step is to generate many sets of the unknown parameter values based on the prior information. Each parameter set will define a possible model of the true stream-aquifer system, with specific values for Q, A, D, q_L, C_L, α, and A_s. The idea is to generate a body of parameter realizations that encompasses all of the allowable (as defined by the prior parameter information) variations and interrelationships of the seven transport properties. Tracer experiment design must then consider each possible scenario.

In stage three of the design procedure the tracer test design is specified. There are three fundamental aspects of the stream tracer test that can be controlled by the experimenter: the tracer injection strategy, the tracer sampling strategy, and the experimental stream reach length. By varying the injection strategy, sampling strategy, or reach length, we can, to some extent, control the information contained in the data collected as part of the experiment. As we will see later, the injection strategy can be designed to increase the information content of the data, the sampling strategy can be designed to ensure that a sufficient number of informative data are collected, and the experimental reach length can be selected to enhance parameter identifiability.

The next stage, global parameter uncertainty analysis, involves evaluating the efficiency of the tracer test design for the spectrum of possible stream transport scenarios. In this study we focus on the ability of the design to provide reliable parameter estimates. For each parameter realization generated in stage two, a parameter uncertainty analysis is performed. Parameter uncertainty is defined by (3), which is a function of the Jacobian, $\mathbf{J}$, and the concentration covariance, $\mathbf{V}_c$. Therefore, for each parameter realization, $\mathbf{J}$ and $\mathbf{V}_c$ must be determined for the specified injection/sampling configuration. Calculating $\mathbf{J}$ is a simple process that involves multiple solute transport simulations to define the sensitivities (4) based on finite difference calculations. Calculating $\mathbf{V}_c$ is also straightforward. Here we assume the concentration uncertainties are proportional to the magnitude of the true concentration values [see *Wagner and Gorelick*, 1986] which are easily defined by running the simulation model (1) and (2) for each parameter realization. Therefore each element of $\mathbf{V}_c$ can be defined as a function of the simulated (true) concentration values and the proportionality constant.

Global parameter uncertainty analysis provides a suite of parameter covariance matrices, one for each parameter realization. In the final stage we evaluate the efficiency of the injection/sampling design across the spectrum of parameter realizations. In this way we can determine the parameter realizations for which the design works well (poorly), and we can identify the stream characteristics that determine if a design is (un)successful. The coefficient of variation is used as the measure of parameter uncertainty

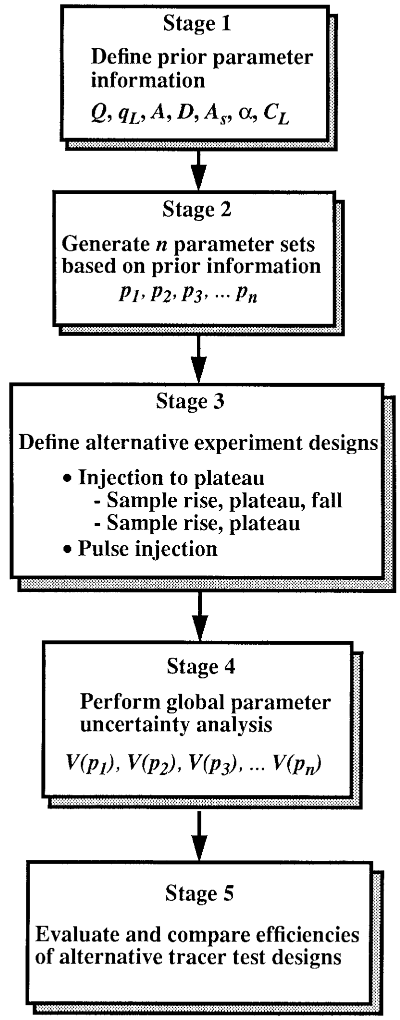

Figure 3. Flow chart describing the five-stage approach to analyzing and designing tracer experiments.

$$\text{c.o.v. } (p_i) = \frac{\text{std } (p_i)}{p_i} \tag{5}$$

where c.o.v. (p_i) is the coefficient of variation for parameter p_i, and std (p_i) is the standard deviation of p_i, which is defined as the square root of the ith diagonal element of the covariance matrix (3). The coefficient of variation is a unitless measure that defines the standard deviation as a fraction of the parameter value. Using the coefficient of variation as the measure of parameter uncertainty allows us to compare results for parameters of different magnitudes and to compare results across parameter realizations and injection/sampling strategies. *Steinberg and Hunter* [1984] present a detailed discussion of alternative criteria that can be used to evaluate parameter uncertainty when designing experiments.

Tracer Experiment Design for High-Gradient Streams

In this section we apply the procedure outlined in Figure 3 to analyze alternative tracer experiment designs for high-

Table 1. Prior Parameter Information for High-Gradient Stream Analysis

Parameter	Range
Discharge, Q (m^3/s)	0.005–0.2
Groundwater inflow, q_L (m^3/s m)	0.0–0.0001
slope, s (m/m)	0.01–0.15
stream width, b (m)	0.5–5.0
Stream area, A (m^2) = $f(Q, q_L, s, b)$	0.02–0.6
Dispersion coefficient, D (m^2/s)	0.025–0.8
Storage area, A_s (m^2)	0.01–2.0
Exchange coefficient, α (1/s)	0.000005–0.001
Inflow concentration, C_L (mg/L)	1.0

gradient streams. The analysis considers three basic injection/sampling alternatives (see Figure 2): (1) injection to plateau, with sampling of the concentration rise, plateau, and fall; (2) injection to plateau with sampling of the concentration rise and plateau only; and (3) pulse injection (no plateau) with sampling of the concentration rise and fall.

As outlined in Figure 3, the first stage of the analysis is to define the prior parameter information. High-gradient stream systems include streams with widely varying transport properties. Stream discharge and velocity, groundwater inflow, dispersion coefficient, and storage-zone exchange properties can vary by orders of magnitude from one system to another, from one stream reach to another, and seasonally for a specific reach. Here we define the prior parameter information on the basis of the parameter values that have been reported in the literature. The parameter ranges that define the prior information are listed in Table 1.

The next stage is to generate many realizations of the stream transport parameters. For this study we generated 800 sets of stream transport parameters. Referring to Table 1, a single set of parameter values was generated as follows. First, values of stream discharge Q, groundwater inflow q_L, stream slope s, and stream width b were randomly generated assuming they were uniformly distributed between the values listed in Table 1. For each set of Q, q_L, s, and b values, the associated stream cross-sectional area, A, was determined using Manning's equations for open channel flow [*Hwang*, 1981]. This means that the values of discharge and cross-sectional area are physically consistent, with a strong positive correlation between discharge and area. The remaining parameters, D, A_s, and α, were generated assuming they were independent and uniformly distributed between the upper and lower limits in Table 1. This procedure was repeated 800 times to produce a suite of parameter sets that represents a wide range of stream transport scenarios. The only condition placed on the parameter values was that the ratio of A/A_s could not be less than 0.2, which is consistent with the A/A_s ratios that have been reported in the literature. Otherwise the parameter sets can have any combination of parameter values that is consistent with the ranges in Table 1 and with Manning's equation. In general, prior parameter information can take any distributional form. We believe that the uniform distribution was appropriate for this study because the existing data, in particular those for the storage exchange parameters, are insufficient to justify a more complex distributional model. As the number of tracer experiments performed increases and the body of physical parameter estimates grows, it might become possible at some later date to develop more complex distributional models for these parameters.

In stage three, we specified the alternative tracer experiment designs that were to be compared. Here we considered three injection/sampling strategies. The first strategy involved injection to plateau with sampling of the rise, plateau, and fall of the concentration history. This strategy provided information for estimating A, D, q_L, A_s, and α on both the rising and falling limbs of the concentration history (see Figure 2). The second design involved injection to plateau with sampling of the concentration rise and plateau only. In this case there is only the rising limb of the concentration history from which to gain information about the parameters. Finally, we considered a pulse injection strategy in which the concentrations at the sampling site did not reach plateau. This strategy has both the rising and falling limb information but generally to a lesser degree than the rise-plateau-fall strategy.

The final step is to analyze the relative efficiencies of the three alternative designs considered here. For each of the three injection/sampling strategies and for each of the 800 parameter sets, the parameter covariance (3) was calculated. In total, approximately 13,000 solute transport simulations were performed as part of the 2400 parameter uncertainty analyses. The suite of 2400 covariance matrices provided the basis for evaluating and comparing the alternative tracer test designs.

The Basic Tracer Experiment

Before discussing the results it is necessary to define the basic tracer study that was used in this study. There are many variables that must be considered when designing a tracer experiment, such as the length of the reach over which the experiment is performed, the amount and duration of tracer injection, and the tracer-sampling schedule. In order to normalize the comparison of the three injection/sampling strategies, these design variables have been standardized. For every case analyzed the experiment was assumed to take place over a 150-m reach of stream. Furthermore, for the cases involving injection to plateau, tracer injection was assumed to last 6 hours; for the pulse strategy, injection was assumed to last 20 min. It was further assumed that the tracer injection would give a plateau concentration at the sampling site that was 25 times the background concentration, and that the stream water would be sampled in 30-s intervals. Finally, it was assumed that the concentration data errors, which define $\mathbf{V}_c$ in (3), have standard deviations equal to 15% of the true concentration value [see *Wagner and Gorelick*, 1986]. In later sections we will investigate the effects of these assumptions.

As noted earlier, there are seven unknown parameters in the governing model (1) and (2) that must be estimated. However, tracer experiments normally include sampling controls that can be used to reduce the number of parameters that must be estimated. One form of sampling control is to measure in-stream solute concentrations prior to tracer injection and upstream of the injection point during the tracer study. This provides data that can be used to define the inflow concentration C_L. Another form of control is to measure in-stream solute concentrations during the experiment directly downstream of the injection point. In this way the discharge, Q, can be determined by the dilution gaging method [*Kilpatrick and Cobb*, 1985]. For this study we assume that the tracer experiments included the two sampling controls described above. Therefore the model parameters that the experiment must be designed to estimate are A, D, q_L, A_s, and α. The remainder of this paper will evaluate and compare the ability of the

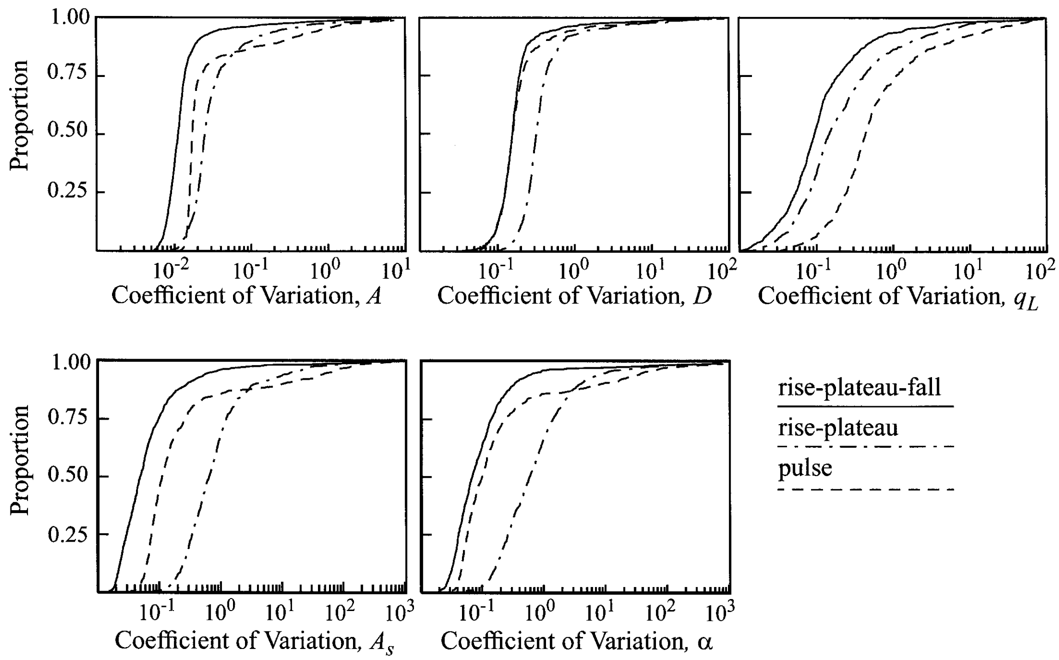

Figure 4. Results of the global parameter uncertainty analysis for three solute injection and sampling strategies. Shown are the cumulative distributions of the coefficients of variation for A, D, q_L, A_s, and α. Each point on a curve represents the fraction of parameter sets with equal or smaller coefficients of variation.

alternative tracer experiment designs to reliably estimate these five parameters.

Results

The results for the high-gradient stream analysis are presented in Figure 4, which summarizes 2400 parameter uncertainty analyses, 800 for each of the three tracer injection/sampling strategies. Plotted in these figures are the cumulative distributions of the coefficients of variation (5) for the parameters A, D, q_L, A_s, and α. Each point on a curve represents the fraction (vertical axis) of parameter sets with equal or smaller coefficients of variation (horizontal axis). Note the log scale of the horizontal axis. These plots highlight several important characteristics of stream tracer experiment design.

First, injection to plateau with sampling through the rise, plateau, and fall is clearly superior in all cases. The reasons for this can be explained by referring to the sensitivity plots in Figure 2. The rise-plateau-fall strategy outperforms the rise-plateau strategy because it provides information for estimating A, D, q_L, A_s, and α on both the rising and falling limbs of the concentration history. The effects of this additional information are significant. The rise-plateau-fall curves for the storage-zone exchange parameters A_s and α are shifted approximately an order of magnitude to the left of the rise-plateau curves. As for the pulse strategy, the sensitivity plots in Figure 2 suggest this strategy will also gain information from both the rising and falling limbs. However, the rising- and falling-limb data of the pulse strategy are cumulatively less informative than those of the rise-plateau-fall strategy. The differences appear to be insignificant for the dispersion coefficient, but for q_L, A_s, and α the coefficient of variation can be considerably smaller for the rise-plateau-fall strategy. For q_L the superiority of the rise-plateau-fall and rise-plateau strategies is a direct result of allowing the concentrations to reach plateau at the tracer sampling site. The plateau concentrations provide an additional source of information for estimating q_L that is not available in the pulse strategy (see Figure 2). For A_s the effect of tracer plateau is again evident. When the tracer is injected as a short duration pulse, the falling-limb data is less informative than the falling-limb data found in the rise-plateau-fall strategy (see Figure 2).

The results presented in Figure 4 suggest that sampling both the rising and falling limbs of the concentration history, as in the rise-plateau-fall and pulse designs, will more likely result in reliable estimates of the storage zone parameters. It should be noted that the rise-plateau-fall strategy is a particular type of pulse design. As the time of pulse injection increases, the ability of the pulse design to reliably estimate the parameters approaches that of the rise-plateau-fall design. That is, as the time of pulse injection increases, the curves in Figure 4 associated with pulse injection will shift to the left, eventually coinciding with the rise-plateau-fall curves. Similarly, as the time of pulse injection decreases, the "pulse" curves will shift to the right with the pulse injection strategy becoming a generally less reliable design.

Although the rise-plateau-fall strategy consistently outperforms the two other strategies across the spectrum of scenarios considered here, there is a considerable number of scenarios for which none of the three designs performs adequately. (Here we define "adequately" to be a coefficient of variation of 0.1 or less. Although somewhat arbitrary, the choice of 0.1 is derived from the concept of a 95% confidence interval for a normally distributed random variable. Parameters with coefficients of variation equal to 0.1 will have 95% confidence intervals that are approximately plus or minus 20% of the parameter value.) For example, approximately 25% of the parameter sets will have c.o.v. $(A_s) > 0.1$ for the rise-plateau-fall strategy; for pulse injection approximately 50% will have c.o.v. $(A_s) > 0.1$; for the rise-plateau strategy, approximately

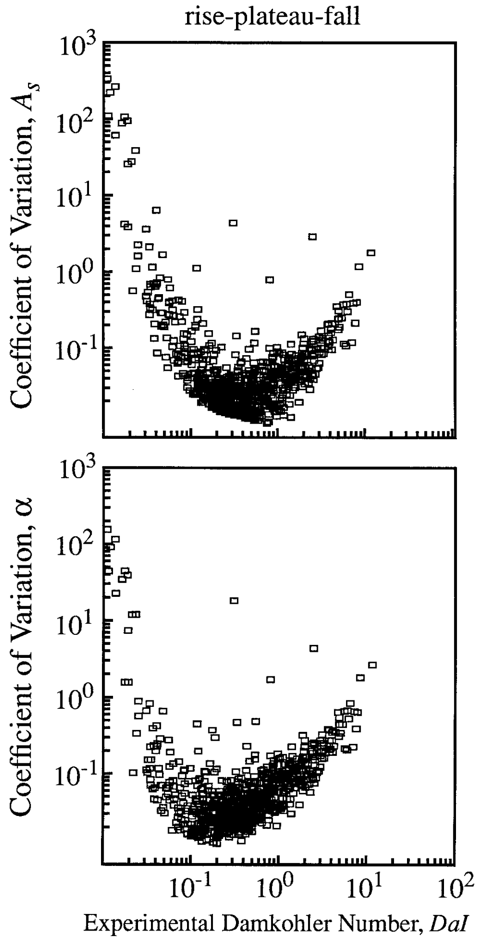

Figure 5. Plot of coefficient of variation versus the experimental Damkohler number for storage zone cross-sectional area, A_s, and stream-storage exchange coefficient, α.

100%. For the exchange coefficient α, the percentages are approximately 35%, 50%, and 100%, respectively.

The natural question to ask is, Why is it possible to reliably estimate the storage-zone exchange parameters for some scenarios, whereas for other scenarios these parameters have very high uncertainty? A partial answer to this question is found in Figure 5, which plots the coefficients of variation for A_s and α versus the experimental Damkohler number, DaI [*Bahr and Rubin*, 1987]:

$$DaI = \frac{\alpha(1 + A/A_s)L}{v} \tag{6}$$

where L is the length of the stream reach over which the experiment is performed, and v is the average stream water velocity over that reach. Recall that the stream reach length, L, is the same for every scenario.

The data in Figure 5 show a strong link between parameter uncertainty and the experimental Damkohler number DaI. In general, for the storage zone cross-sectional area, A_s, minimum parameter uncertainty is obtained when DaI is approximately 1.0, and there is a strong trend of increasing parameter uncertainty when the Damkohler number decreases below or increases above 1.0. There is a similar linkage between DaI and c.o.v. (α), with minimum parameter uncertainty occurring when DaI is approximately 0.1. Inspection of (6) shows that small DaI values may occur for three reasons: (1) stream-water velocity, v, is high; (2) exchange timescales are long, as indicated by small values of α and A/A_s; and/or (3) reach length, L, is short. Under these conditions, parameter uncertainties are high because only a small amount of tracer interacts with the storage zone in the reach. For high DaI values, solute exchange rates are fast relative to stream-water velocity and/or the experimental stream reach length is long. For these cases, all solute is exchanged with the storage zone over the experimental reach. As DaI increases, tracer dispersion caused by storage zone exchange eventually reaches an equilibrium condition (see discussion on page 2445 of *Harvey et al.* [1996]) and storage zone exchange parameters become effectively nonidentifiable (i.e., estimated with very high uncertainties).

In a study of kinetically influenced groundwater solute transport, *Bahr and Rubin* [1987] demonstrated the use of the dimensionless factor DaI for identifying those cases where non-equilibrium transport cannot be distinguished from equilibrium transport. Their results show that as DaI approaches 100, the effects of rate-limited mass transfer become indistinguishable from equilibrium transport. Our analyses suggest that the parameters can become effectively nonidentifiable at much lower DaI values. The reason for this can be traced back to the Jacobian matrix, **J**, and its role in defining the covariance (3). Strict non-identifiability occurs when two or more columns of **J** are linearly related [see *Carrera and Neuman*, 1986; *Knopman and Voss*, 1987]. In that case the parameters cannot be estimated regardless of how much data are collected. Our study shows that although the parameters are not strictly nonidentifiable, they are approaching nonidentifiability, as indicated by a general trend of increasing parameter correlation as DaI increases above or decreases below 1.0 for A_s and 0.1 for α. Moreover, when DaI is considerably smaller or larger than approximately 1.0, the storage-zone exchange parameters cannot be reliably estimated.

Improving the Efficiencies of Tracer Experiment Designs

The above analysis indicates that tracer experiment designs with very high or very low DaI values can present difficulties for estimating the storage zone exchange parameters and therefore for designing tracer experiments. In particular, streams that have very fast or very slow rates of exchange relative to the stream-water velocity present potential difficulties when estimating the storage zone exchange parameters. Streams with these characteristics require special consideration to ensure the exchange parameters can be reliably estimated based on tracer data. Recall that the analysis above was based on a standardized tracer test with assumptions regarding data uncertainty, frequency of data collection, and length of study reach. In this section we investigate options for changing the basic tracer test to improve parameter reliability.

The design of experiments for parameter estimation requires weighing the trade-offs between three elements that define the parameter covariance matrix (3): (1) the concentration sensitivities that compose the Jacobian matrix; (2) the number of data, particularly the number of high-sensitivity data; and (3) the uncertainty associated with those data. Therefore attempts to improve the design of an experiment will focus on increasing concentration sensitivities, increasing the number of data collected, and/or decreasing data uncertainty.

Reducing data uncertainty. The effect of data uncertainty is the easiest to analyze. In this study we assumed the standard deviation of concentration errors increased with concentration and was equal to 15% of the true concentration value. The plots in Figures 4 and 5 are based on this assumption. A value of 15% for the data uncertainty factor was selected because it represents the upper limit of uncertainty factors that we have encountered in our analyses of stream tracer experiments. Given that the coefficient of variation is proportional to the data uncertainty factor, assessing variations in this factor is simply a matter of shifting the curves in Figure 4 based on the relative change in the factor's value. For example, if the uncertainty factor is doubled, the curves shift to the right by a factor of 2.0; if the uncertainty factor is halved, the curves shift to the left by a factor of 0.5. There is no change in the relative positions of the curves in Figure 4 or the data points in Figure 5.

Varying the sampling frequency. An alternative to decreasing data uncertainty is to increase the number of data collected as part of the tracer experiment. As described by Figure 2, the informative data for each parameter are associated with specific parts of the concentration history. By increasing the number of data collected in these high-sensitivity zones, we can reduce the uncertainty in the parameter estimates. In this case the reduction in parameter uncertainty is not as easily predicted as the reduction resulting from a decrease in data uncertainty. To determine the effects of additional data, the parameter uncertainty analyses must be redone. The results summarized in Figures 4 and 5 are based on the assumption that concentration data are collected in 30-s intervals. To analyze the effect of varying the sampling frequency, the parameter uncertainty analyses were recalculated for sampling intervals of 15 and 90 s. The results are presented in Figure 6 for the exchange coefficient α for the rise-plateau-fall strategy. As we would expect, the uncertainty in the estimate of α decreases with increasing sampling frequency and vice versa. Recall that with a 30-s sampling frequency 35% of the parameter sets had c.o.v. $(\alpha) > 0.1$. When the sampling interval is reduced to 15 s, 25% of the parameter sets have c.o.v. $(\alpha) > 0.1$; for a 90-s sampling interval, the number is 70%. The relative changes in parameter uncertainty with changes in sampling frequency are similar for the other parameters. It should be noted that parameter uncertainty does not in general vary linearly with sampling frequency as it does with

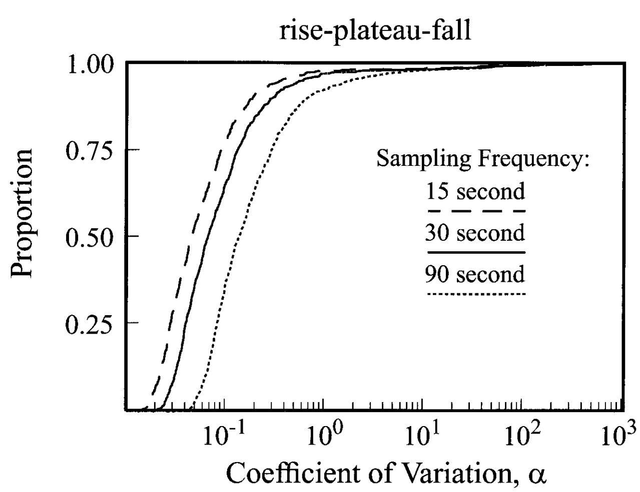

Figure 6. Sensitivity of parameter uncertainty to variations in the concentration sampling frequency. Results shown are for the stream-storage exchange coefficient, α, under the rise-plateau-fall injection/sampling strategy.

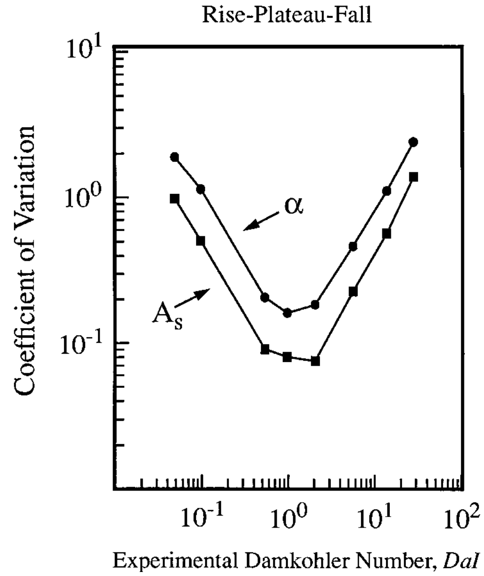

Figure 7. Sensitivity of parameter uncertainty to variations in the experimental stream reach length. Results shown are for a single parameter realization where the stream reach length was varied in order to vary DaI from approximately 0.05 to approximately 25.0.

the uncertainty factor. If the sampling interval is further increased, the parameters will eventually become nonidentifiable, as the sampling strategy will no longer include the data needed to independently estimate each of the unknown parameters. This was the case for a small number of parameter realizations when the sampling interval reached 120 s.

Varying the reach length. Reducing data uncertainty and increasing sampling frequency do not affect the concentration sensitivities that comprise the Jacobian matrix in (3). A third option for improving a tracer test design is to revise those elements of the design that control concentration sensitivities in order to increase the information content of the concentration data. In this study we have focused on the ability of a tracer experiment to reliably estimate the hyporheic exchange parameters, and we have identified, on the basis of the experimental Damkohler number DaI, those properties that limit our ability to reliably estimate the exchange parameters. Moreover, the one defining property of DaI that we can control as part of the design of a tracer experiment is the experimental reach length L. This suggests that the choice of the experimental reach length could be an important factor in determining if the storage zone exchange parameters can be reliably estimated (recall that the results presented thus far are for a stream reach length of 150 m). To investigate the influence of reach length on parameter uncertainty, we selected a single parameter realization and performed a series of parameter uncertainty analyses, varying the reach length in each analysis in order to vary the experimental Damkohler number from approximately 0.05 to approximately 25.0. The results are presented in Figure 7 for the rise-plateau-fall injection and sampling strategy. Again, the results indicate a strong dependency between the uncertainty in the hyporheic exchange parameters and the value of the experimental Damkohler number DaI, and again the minimum parameter uncertainty is obtained when the Damkohler number is on the order of 1.0.

The data plotted in Figures 5 and 7 suggest that the uncertainty associated with the storage-zone exchange parameters will be minimized if the stream reach length is selected so as to have an experimental Damkohler number on the order of 1.0. This does not mean, however, that reliable estimates cannot be obtained for other reach lengths. In fact, it is obvious from inspecting Figures 5 and 7 that reliable parameter estimates can be obtained for Damkohler numbers considerably smaller or larger than 1.0. It should be noted, however, that Figures 5 and 7 present the results for the rise-plateau-fall strategy, which we have shown to be the most informative injection/sampling strategy for parameter estimation. Moreover, these results were obtained assuming that in-stream solute concentrations would be sampled at 30-s intervals. Because of cost considerations, it may be desirable to use a less informative injection/sampling strategy (i.e., the rise-plateau or pulse strategy) and/or less frequent sampling. In those cases, Figures 5 and 7 would retain the same general shape but would shift upward. The conclusion that we can draw from this is that as we move to less informative tracer test designs, the range of stream reach lengths that provide reliable parameter estimates becomes more tightly constrained, and the choice of stream reach length becomes more critical.

There have been a small number of stream tracer studies reported in the literature that have been analyzed using formal techniques of parameter estimation and uncertainty analysis [*Wagner and Gorelick*, 1986; *Wagner and Harvey*, 1996; *Harvey et al*., 1996]. Although these studies comprise only four tracer tests, the parameter estimates suggest they represent a set of high-gradient streams with large differences in α, A_s, A, and v. The parameter estimation results presented in these studies indicate the storage-zone exchange parameters are reasonably well estimated, with coefficients of variation of approximately 0.30 or less. In addition, the experimental Damkohler numbers for these tracer experiments (calculated using the optimal parameter estimates) range from approximately 0.07 to approximately 2.0, which is consistent with our conclusion that "well-estimated" parameters are likely to be obtained when the Damkohler number is on the order of 0.1–1.0.

We have shown that a tracer experiment's Damkohler number is indicative of the experiment's ability to reliably identify the storage zone exchange parameters. We also believe that the Damkohler number is useful for designing tracer studies, although this step is complicated by the paradox of experimental design stated earlier, namely that the parameters needed to define DaI are in fact the parameters to be estimated by the experiment. One scenario for using DaI in designing experiments is when relatively reliable values of the parameters α, A_s, A, and v are available from earlier tracer experiments in the stream of interest. In this scenario it would be straightforward to use DaI to determine the appropriate stream reach length for minimizing parameter uncertainty. This could also be taken a step further by using uncertainty analysis (based on the available parameter estimates) to weigh the trade-offs between alternative injection strategies, sampling strategies, and reach lengths. A second scenario is when the parameters are not well defined, and the range of DaI values is such that there is no single experimental reach length that ensures reliable parameter estimates for every possible outcome (as defined by the prior parameter information). In this case, one can use the concepts of decision theory [*Lindley*, 1978] to, for example, design an experiment that has a "high" likelihood of succeeding. This topic is beyond the scope of this report.

Summary

The value of the stream tracer experiment for analyzing the transport properties of stream-aquifer systems has been demonstrated in numerous studies. In this paper we presented a methodology for systematically analyzing stream tracer study designs. The methodology uses the concept of global parameter uncertainty analysis to evaluate and compare tracer test designs over the spectrum of possible transport scenarios that may be encountered in the field. Parameter uncertainty analysis provides a quantitative framework for analyzing the various trade-offs that are encountered when designing tracer experiments, such as the trade-offs between different injection and sampling strategies, the gain (or loss) in information resulting from more (or less) frequent sampling, and the effect of lengthening (or shortening) the reach over which the experiment is performed.

The analyses presented in this paper focused on the design of tracer experiments for high-gradient stream systems. However, we believe the conclusions of this work are general and can be extended to cases outside those studied here. Our analyses highlighted two basic properties of stream tracer experiments. First, the choice of solute injection/sampling strategy is an important one. We found that constant injection with sampling through the rise, plateau, and fall was able to provide considerably more reliable parameter estimates than pulse injection for the spectrum of scenarios that were considered. Second, we found that the experimental Damkohler number DaI is a valuable indicator of the reliability with which the storage exchange parameters can be estimated using the stream tracer approach. Our analyses suggest that when the Damkohler number is on the order of 1.0, the storage-zone exchange parameters will be estimated with minimum uncertainty. The use of this criteria when designing stream tracer studies will be limited somewhat by the fact that evaluation of the Damkohler number for a given problem requires knowledge of the storage-zone exchange parameters as well as information on the physical transport parameters. However, placing bounds on these parameters may be sufficient to place bounds on the Damkohler number for assessing alternative tracer experiment designs.

Although this work analyzed tracer test design, and in general parameter estimation, for stream systems, the conclusions can be extended to groundwater contaminant transport subject to rate-limited mass transfer, which for many cases follows governing equations almost identical to those analyzed in this study [see *Bahr and Rubin*, 1987]. Rate-limited mass transfer of groundwater contaminants (such as sorption onto grain surfaces or diffusion into low-permeability layers) places significant limitations on groundwater remediation systems [*National Research Council*, 1994]. Consequently, considerable effort is directed at determining the mass transfer capacity of aquifers using field tracer studies and/or laboratory column experiments. Given the similarities in the governing equations for solute transport in streams and in groundwater, it is likely that the results presented here for stream tracer studies can be extended to aid the design of tracer tests in groundwater systems.

Acknowledgments. The authors gratefully acknowledge the funding of the U.S. Geological Survey's Toxic Substances Hydrology Program that supports research in Pinal Creek Basin, Arizona. The authors are also grateful for the helpful review comments provided by Tracy Nishikawa and Rob Runkel, U.S. Geological Survey, and the anonymous journal reviewer.

References

Bahr, J. M., and J. Rubin, Direct comparison of kinetic and local equilibrium formulations for solute transport affected by surface reactions, *Water Resour. Res.*, *23*(3), 438–452, 1987.

Bencala, K. E., Interactions of solutes and streambed sediment, 2, A dynamic analysis of coupled hydrologic and chemical processes that determine solute transport, *Water Resour. Res.*, *19*(3), 718–724, 1984.

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

Carrera, J., and S. Neuman, Estimation of aquifer parameters under transient and steady state conditions, 2, Uniqueness, stability, and solution algorithms, *Water Resour. Res.*, *19*(3), 211–227, 1986.

Cleveland, T. G., and W. W.-G. Yeh, Sampling network design for transport parameter identification, *J. Water Resour. Plann. Manage. Div. Am. Soc. Civ. Eng.*, *116*(6), 765–783, 1990.

Cleveland, T. G., and W. W.-G. Yeh, Optimal configuration and scheduling of ground-water tracer test, *J. Water Resour. Plann. Manage. Div. Am. Soc. Civ. Eng.*, *117*(1), 37–51, 1991.

D'Angelo, D. J., J. R. Webster, S. V. Gregory, and J. L. Meyer, Transient storage in Appalachian and Cascade mountain streams as related to hydraulic characteristics, *J. N. Am. Benthol. Soc.*, *12*, 223–235, 1993.

Draper, N. R., and H. Smith, *Applied Regression Analysis*, 2nd ed., 709 pp., John Wiley, New York, 1981.

Green, H. M., K. J. Beven, K. Buckley, and P. C. Young, Pollution incident prediction with uncertainty, in *Mixing and Transport in the Environment*, edited by K. J. Beven, P. C. Chatwin, and J. H. Millbank, pp. 113–137, John Wiley, New York, 1994.

Grimm, N. B., and S. G. Fisher, Exchange between interstitial and surface water: Implications for stream metabolism and nutrient cycling, *Hydrobiologia*, *111*, 219–228, 1984.

Hart, D. R., Parameter estimation and stochastic interpretation of the transient storage model for solute transport in streams, *Water Resour. Res.*, *31*(2), 323–328, 1995.

Harvey, J. W., B. J. Wagner, and K. E. Bencala, Evaluating the reliability of the stream tracer approach to characterize stream-subsurface water exchange, *Water Resour. Res.*, *32*(8), 2441–2452, 1996.

Hwang, N. H. C., *Fundamentals of Hydraulic Engineering Systems*, 367 pp., Prentice-Hall, Englewood Cliffs, N. J., 1981.

Kilpatrick, F. A., and E. D. Cobb, Measurement of discharge using tracers, in *Techniques in Water Resources Investigations*, book 3, chap. A-16, U.S. Geol. Surv., Reston, Va., 1985.

Kim, B. K. A., A. P. Jackman, and F. J. Triska, Modeling biotic uptake by periphyton and transient hyporheic storage of nitrate in a natural stream, *Water Resour. Res.*, *28*(10), 2743–2752, 1992.

Knopman, D. S., and C. I. Voss, Behavior of sensitivities in the one-dimensional advection-dispersion equation: Implications for parameter estimation and sampling design, *Water Resour. Res.*, *23*(2), 253–272, 1987.

Lindley, D. V., *Making Decisions*, 195 pp., John Wiley, New York, 1978.

McCarthy, J. M., and W. W.-G. Yeh, Optimal pumping test design for parameter estimation and prediction in groundwater hydrology, *Water Resour. Res.*, *26*(4), 779–791, 1990.

McLaughlin, D., and L. R. Townely, A reassessment of the groundwater inverse problem, *Water Resour. Res.*, *32*(5), 1131–1161, 1996.

Moss, E. M., Some basic considerations in the design of hydrologic data networks, *Water Resour. Res.*, *15*(6), 1673–1676, 1979.

National Research Council, *Alternatives for Ground Water Cleanup*, 315 pp., Natl. Acad. Press, Washington, D. C., 1994.

Nishikawa, T., and W. W.-G. Yeh, Optimum pumping test design for the parameter identification of groundwater systems, *Water Resour. Res.*, *25*(7), 1737–1747, 1989.

Runkel, R. L., and S. C. Chapra, An efficient numerical solution of the transient storage equations for solute transport in small streams, *Water Resour. Res.*, *29*(1), 211–215, 1993.

Runkel, R. L., D. M. McKnight, K. E. Bencala, and S. C. Chapra, Reactive solute transport in streams, 2, Simulation of a pH modification experiment, *Water Resour. Res.*, *32*(2), 419–430, 1996.

Rutherford, J. C., J. D. Boyle, A. H. Elliott, T. V. J. Hatherell, and T. W. Chiu, Modelling benthic oxygen uptake by pumping, *J. Environ. Eng. N. Y.*, *121*(1), 84–95, 1995.

Steinberg, D. M., and W. G. Hunter, Experimental design: Review and comment, *Technometrics*, *26*(2), 71–97, 1984.

Stream Solute Workshop, Concepts and methods for assessing solute dynamics in stream ecosystems, *J. N. Am. Benthol. Soc.*, *9*(2), 95–119, 1990.

Valett, H. M., A. A. Morrice, C. N. Dahm, and M. E. Campana, Parent lithology, surface-ground water exchange, and nitrate retention in headwater streams, *Limnol. Oceanogr.*, *41*(2), 333–345, 1996.

Wagner, B. J., and S. M. Gorelick, A statistical methodology for estimating transport parameters: Theory and applications to one-dimensional advective-dispersive systems, *Water Resour. Res.*, *22*(8), 1303–1315, 1986.

Wagner, B. J., and J. W. Harvey, Solute transport parameter estimation for an injection experiment at Pinal Creek, Arizona, in *U.S. Geological Survey Toxic Substances Hydrology Program—Proceedings of the Technical Meeting, Colorado Springs, Colorado, September 20–24, 1993*, edited by D. W. Morganwalp and D. A. Aronson, *U.S. Geol. Surv. Water Resour. Invest. Rep. 94-4015*, 1996.

Wallis, S. G., P. C. Young, and K. J. Beven, Experimental investigation of aggregated dead zone model for longitudinal transport in stream channels, *Proc. Inst. Civ. Eng.*, Part 2, *87*, 1–22, 1989.

J. W. Harvey, U.S. Geological Survey, 430 National Center, Reston, VA 20192.

B. J. Wagner, U.S. Geological Survey, 345 Middlefield Rd., Menlo Park, CA 94025. (e-mail: bjwagner@usgs.gov)

(Received September 20, 1996; revised April 1, 1997;
accepted April 9, 1997.)