On the Estimation of Solute Transport Parameters for Rivers

S. G. Wallis

Heriot-Watt University, Edinburgh, UK

M. Osuch

Institute of Geophysics, Warsaw, Poland

J. R. Manson

Richard Stockton College, New Jersey, USA

R. Romanowicz

Institute of Geophysics, Warsaw, Poland

B. O. L. Demars

James Hutton Institute, Aberdeen, UK

Abstract The modelling of solute transport in rivers is usually based on simulating the physical processes of advection, dispersion and transient storage, which requires the modeller to specify values of corresponding model parameters for the particular river reach under study. In recent years it has become popular to combine a numerical solution scheme of the governing transport equations with a parameter optimisation technique. However, there are several numerical schemes and optimisation techniques to choose from. The paper addresses a very simple question, namely, do we get the same, or do we get different, parameter values from the application of two independent solute transport models/parameter optimisation techniques to the same data? Results from seven different cases of observed solute transport suggest the latter, which implies that parameter values cannot be transferred between modelling systems.

1 Introduction

Environmental engineers and scientists frequently use mathematical models that simulate the physical processes of solute transport in rivers. Typical applications include assessing the impact of pollution incidents and assessing issues relating to stream ecology. In order to ensure reliable predictions from such models it is preferable to calibrate and validate them using observations of solute transport over a wide range of flow rates for the river of interest. Having achieved this, model parameters can be evaluated for the hydraulic conditions pertaining to the issue of concern, which should result in a reliable prediction. Often, however, the data required to calibrate and validate the model are not available, e.g. in the case where a pollution incident takes place in a river that has not been studied before, in which case parameter values need to be evaluated in another way. In such a situation the only realistic approach is to transfer parameter values from similar rivers, often using empirical equations that have been established using tracer experiments.

During the last fifty years several methods of analysing tracer data in order to evaluate solute transport parameters have been proposed, e.g. the method of moments (Fischer, 1967), routing procedures (Fischer, 1968; Singh and Beck, 2003) and inverse modelling (Wagner and Gorelick, 1986; Scott et al, 2003). Clearly, if these parameter values are to be used to establish empirical equations the parameter values so estimated need to be reliable, i.e. they need to be an accurate reflection of the physical processes being modelled. If not, for example estimated parameter values may be biased away from their true values, then significant inaccuracies in empirical equations that rely on these values will ensue.

In this paper we focus on the third of the above approaches because it is increasingly replacing the other two. The essential nature of inverse modelling is that it involves optimising objectively the parameters of a solute transport model such that a best fit is obtained between simulated and observed solute concentration data. Alternatively, but not considered herein, properties of observed concentration profiles may be objectively estimated using moment matching techniques (Seo and Cheong, 2001; Worman and Wachniew, 2007).

Errors in the parameter values obtained by inverse modelling may come from several sources including: inapplicability of the mathematical model to the physical situation; limitations of the numerical scheme used to implement the model; identification of locally optimised rather globally optimised parameters; non-uniqueness of the optimised parameters; and noisy tracer data. Of particular interest here is the case where an empirical equation might be derived from parameter values obtained from several different sources, in each of which different numerical schemes and/or optimisation algorithms have been used.

This particular situation is explored in this paper, with the aim of comparing optimised parameter values obtained from a range of solute concentration data using two different numerical scheme/optimisation approaches. A very simple ques-

tion is posed: do we get the same, or do we get different, parameter values from the application of two independent solute transport models/parameter optimisation techniques to the same data?

2. Solute Transport Modelling

The overwhelming majority of all solute transport modelling in rivers has been undertaken using one of two one-dimensional models, namely the advection-dispersion equation (ADE) and the transient storage model (TSM). The ADE predates the TSM and describes solute transport in terms of advection and longitudinal dispersion. In simple terms, solute is carried downstream at a velocity determined by location within the cross-section (moving faster in the channel centre and moving slower at the channel sides) and solute is continually mixed (by transverse turbulent diffusion and secondary currents) within the cross-section. The differential velocities act to stretch a cloud of solute longitudinally and the cross-sectional mixing acts to reduce the longitudinal stretching. Once a solute cloud has been evolving for long enough in such a flow field the ADE applies to the cross-sectional average solute concentration (Taylor, 1954; Fischer, 1967; Rutherford, 1994). Often, but not always, observed tracer concentration-time profiles are more skewed than those simulated by the ADE. Typically, observed profiles have more steeply rising limbs and less steeply falling limbs, sometimes with more elevated tails, than simulated ones.

It is recognised that the ADE does not explicitly include the role played by storage areas in rivers, of which there are several types. For example, dead zones are parts of the flow in which no net downstream of water occurs but into which solute migrates, moving back into the main flow sometime later. Thus there is a transient storage of solute in them. The water in some of these dead zones is stagnant but in some others it circulates. Solute transfer between main flow and dead zone and back again occurs by diffusion. In another type of dead zone there is exchange of water and solute with the main flow, in which case solute exchange occurs by a combination of diffusion and local advection. Indeed, some of these more dynamic zones might be considered to be part of the main flow. Generally, but not exclusively, dead zones are located around the channel sides. However, similar, but larger and more dynamic, semi-isolated storage areas exist in river bends, at river junctions, around obstructions such as bridge piers and fallen trees, in pool-riffle structures and amongst in-river vegetation. The hyporheic zone (where there is interaction between the river and the surrounding groundwater) also provides locations where solute can be dynamically stored.

As a result of including transient storage in the ADE, converting it to the TSM, simulated concentration-time profiles often, but not always, mimic observed concentration-time profiles much better. The TSM is described by the following two equations (Bencala and Walters, 1983):

$$\frac{\partial c}{\partial t} + \frac{Q}{A}\frac{\partial c}{\partial x} = \frac{1}{A}\frac{\partial}{\partial x}\left(AD\frac{\partial c}{\partial x} \right) + \alpha(s - c) \tag{1}$$

$$\frac{\partial s}{\partial t} = -\alpha\frac{A}{A_s}(s - c) \tag{2}$$

where c is the solute concentration in the main flow, t is time, Q is the flow rate, A is the cross-sectional area of the main flow, x is the longitudinal co-ordinate, D is the dispersion coefficient, α is the exchange rate between the main channel and the storage zones, s is the solute concentration in the storage zones and A_s is the cross-sectional area of the storage zones. In the absence of transient storage, the ADE is recovered by setting the exchange rate to zero (equation 2 becoming redundant).

Note that although equation (1) allows for spatially variable parameters, the TSM model is usually implemented in a reach-by-reach basis with parameters being constant within an individual reach. Steady flow is assumed, also.

Although the TSM was originally formulated to include the effect of stagnant dead zones in the ADE, it is often now used to include the effects of all the transient storage processes present in a river reach, which may result in a more complicated formulation and/or interpretation of the transient storage terms in recognition of the wide range of transient storage time scales that may exist (see, e.g., Deng et al 2010; Bottacin-Busolin et al, 2011). In the current work two TSM models were used: these are referred to as the Polish model and the UK model.

2.1 The Polish Model

This was based on an implicit Eulerian finite difference numerical solution of equations (1) and (2), as described in Runkel and Broshears (1991) and Runkel and Chapra (1993). Overall, the scheme is of the Crank-Nicolson type in which the first three terms in equation (1) are approximated by implicit central differences, in time or space, as appropriate, and the final term is treated in a central implicit fashion. Equation (2) is also approximated in the Crank-Nicolson way. The upstream boundary condition, applied at the upstream end of the reach of interest, was provided by an observed main channel concentration-time profile, a zero dispersive flux downstream boundary condition, applied further downstream than the downstream end of the reach of interest, was used and the initial condition was zero solute concentration in the main channel and in the storage zones.

The model parameters (D, A, A_s and α) were estimated by optimising the numerical solution to an observed concentration-time profile at the downstream end of the reach of interest, which involved minimising an objective function based on

the residuals between simulated and observed solute concentrations. This was achieved using a global optimisation approach based on the differential evolution (DE) technique, which is an example of an evolutionary algorithm (Storn and Price, 1997; Price et al. 2005). In this an initial population of N randomly selected individuals (N-dimensional vectors of parameters) evolves as a result of mutation, recombination and selection. Mutation involves adding the weighted difference between two random vectors to a third vector, and recombination is a crossover between the new vector and its parents. Then the performance of the old and new generations is compared, and only the better one survives to the next generation.

Differential evolution is an effective, robust and simple global optimisation technique which has only a few control parameters: the population size, the scaling factor used for calculating the weighted differences and the crossover parameter, which govern the probability that a particular element from the parent is passed to the offspring. The following values were assumed for these parameters, respectively: 40, 0.85 and 0.80. Ranges of parameter values for the DE method were chosen as follows: (0,10000) for A, D and A_S and (0,0.001) for α.

2.1 The UK Model

This was based on a different formulation of the model equations, a different numerical solution scheme and a different optimisation algorithm. The model equations are:

$$\frac{\partial(cA)}{\partial t}+\frac{\partial(cAU)}{\partial x}=\frac{\partial}{\partial x}\left(AD\frac{\partial c}{\partial x}\right)+k_1A(s-c) \tag{3}$$

$$\frac{\partial(sA_s)}{\partial t}=-\alpha A(s-c) \tag{4}$$

where U is the cross-sectional average flow velocity in the main channel (= Q/A), k_1 is a solute exchange rate parameter and the other symbols are as previously defined. Clearly, k_1 in equation (3) is equivalent to α in equation (1). When implemented in the usual steady flow, reach constant parameter form all the terms in equation (3) are equivalent to the corresponding terms in equation (1), and equation (4) can be re-written as

$$\frac{\partial s}{\partial t}=-k_2(s-c) \tag{5}$$

where k_2 is equivalent to the term $\alpha A/A_s$ in equation (2). The numerical solution of equations (3) and (5) was based on DISCUS (Manson et al, 2001), a semi-Lagrangian finite volume numerical solution. In DISCUS the (explicit) high-order QUICKEST scheme for advection is implemented in a semi-Lagrangian fashion and is combined with implicit backward finite differencing of the dispersion and transient storage. The boundary conditions and initial conditions were the same as for the Polish model.

The model parameters (D, U, k_1 and k_2) were estimated by optimising the numerical solution to an observed concentration-time profile at the downstream end of the reach of interest using a global optimisation technique based on a genetic algorithm (GA) approach. In general GA techniques are similar to DE techniques. However, in the GA technique used here the objective function was defined as the inverse of the sum of the squared residuals between observed and simulated downstream concentrations. The objective function was then maximised by adjusting the parameters according to the micro-genetic algorithm described in Yang et al (1998). The control parameters of population size, mutation probability and crossover rate were defined as 5, 0.2 and 0.5, respectively. 2500 generations were allowed to evolve, yielding a parameter space of 33554432^4 potential solutions. Ranges of parameter values for the GA method were chosen as follows: (0,2) for U, (0,2000) for D, (0,0.01) for k_1 and (0,0.1) for k_2.

3 Data Preparation and Analysis

The data required for parameter estimation consisted of solute concentration-time profiles at the upstream and downstream ends of a river reach, the corresponding reach length and the corresponding river flow rate. The parameters of both models were optimised by minimising or maximising an objective function as previously described. Seven cases were analysed, taken from a variety of sources (several tracer experiments in streams and rivers and one analytical solution). To ensure that exactly the same concentration-time data were used with both models both upstream and downstream concentration-time profiles for a particular case were supplied at the same, constant, sampling interval. This removed the need to interpolate the data, which might have been done in different ways by each modelling group. However, the time step varied between the cases. The cases are summarised in Table 1.

For both models, a river reach was treated as a uniform channel so that all model parameters were constants. To investigate the role of spatial resolution, which has been rarely even considered in previous work, each case was analyzed four times using space steps of L/200, L/100, L/50 and L/25, where L is the reach length. In addition, some cases were analysed over a wider range of space step.

The treatment of the downstream boundary condition was also investigated briefly. As with the spatial resolution issue there is little published guidance on

where the downstream boundary should be located – far enough so as not to influence the solution in the reach of interest being the best advice available, see e.g. Runkel and Broshears (1991). So for one or two cases results were compared from a series of runs in which the downstream boundary was located progressively further away from the downstream end of the reach of interest.

Table 1. Summary of cases

Case	River	Reach length (m)	Flow rate (l/s)	Timestep (s)
1	Murray Burn	184	144.8	30
2	Rhine	47150	1169000	600
3	Rhine	73800	1169000	1200
4	Cairn Burn	91	10.2	10
5	Uvas Creek	152	13.6	600
6	Uvas Creek	186	14.0	600
7	Analytical	250	1000	1

Results and Discussion

Figure 1 shows the variation of the optimised parameters from Case 7 over a wide range of space step values, and shows how the optimised parameter values converge as the magnitude of the space step decreases. In the figure the parameter values are expressed as a percentage of the value found for 500 space steps and the results are from the Polish model. Similar behavior was found in the UK model and for other Cases. It is particularly evident that A_s is the parameter that is most sensitive to this issue. There is no published guidance on what degree of spatial resolution is required for these sorts of analyses. Indeed, it is not clear that the issue has been studied before. The most likely reason for this is that, because the optimisation uses temporal data, the issue has been overlooked. Nevertheless, as one might expect, Figure 1 suggests that the spatial resolution needs to be sufficiently good to prevent it being a source of error. Of course, what really matters is the number of space steps over which the spatial concentration profile, corresponding to the problem in hand, is resolved rather than how well the reach length is resolved. This can be evaluated by dividing the length of the spatial concentration profile (estimated here as the product of the centroid velocity and the time difference between the trailing and leading edges of the upstream concentration-time profile) by the space step. For Case 7 this gives 392, 196, 98 and 49 for reach resolutions of L/200, L/100, L/50 and L/25, respectively, suggesting that resolving the spatial concentration profile over about 100 space steps is sufficient for parameters A, D and α, but maybe as many as about 400 space steps are required for A_s. For the other six Cases, the number of space steps covering the spatial concentration profile length for the L/200 reach resolution varied between about 100 and

1800. So based on the analysis of Case 7, the other Cases seem to have been adequately resolved in space.

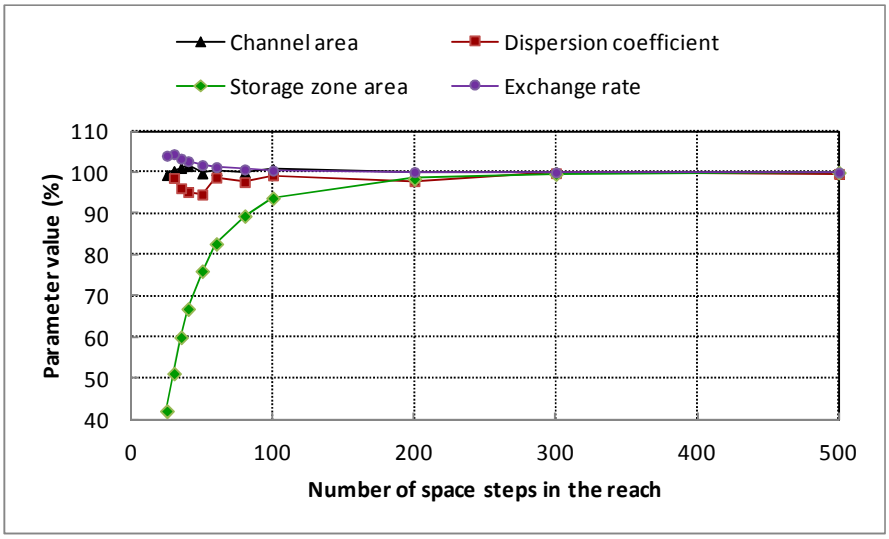

Fig. 1. Sensitivity of optimised parameters to spatial resolution (Case 7; Polish model)

Results from runs with both models where the location of the application of the downstream boundary condition was varied showed that only for the lowest spatial resolution case (L/25) was there any significant sensitivity to this issue. This seems reasonable because the zero dispersive flux boundary condition is more consistent with shallow spatial concentration gradients than with steep ones, with the former and latter being associated, respectively, with high and low spatial resolutions.

Table 2 compares the optimised parameter values from the two models for all seven cases. In view of the discussion above, it seems reasonable to base the comparison on the results obtained with the highest reach resolution, i.e. 200 space steps. It is clear that there are only a few examples of close agreement between the models and there are several examples of very poor agreement between them. Clearly, this is a worrying outcome since it suggests that the parameter values are not independent of the model/optimisation system used, indicating that they are not transferable. If this state of affairs were found to be generally true it would seriously threaten the credibility of empirical equations based on published parameter values that had been derived in the way used in this paper.

Of the four parameters, A shows the best agreement between the models with percentage differences being < 10% for all cases. The percentage errors in the other parameters vary widely (even exceeding 100%), and show no apparent pattern.

Table 2. Optimised parameter values from the two models (reach resolution of L/200)

Case	Model	A (m^2)	D (m^2/s)	α (1/s)	A_s (m^2)	DaI (-)
1	Polish	0.688	0.361	0.00100	0.083	>5
	UK	0.737	0.564	0.00025	0.039	>5
2	Polish	1418	22.5	0.00002	86.3	>5
	UK	1367	46.0	0.00062	112.8	>5
3	Polish	1284	299	0.00001	380	>1, <5
	UK	1336	743	0.00013	16.8	>5
4	Polish	0.137	0.187	0.00018	0.017	>1, <5
	UK	0.145	0.235	0.00031	0.661	>0.2, <1
5	Polish	0.451	0.204	0.00003	0.919	>0.2, <1
	UK	0.460	0.250	0.00003	1.892	>0.2, <1
6	Polish	0.513	0.153	0.00005	4.710	>0.2, <1
	UK	0.694	1.383	0.00075	0.053	>5
7	Polish	2.020	1.110	0.00035	3.722	>0.2, <1
	UK	2.019	1.097	0.00038	3.230	>0.2, <1

The good agreement for A is consistent with the idea that the response of a solute cloud to being transported through a river reach is most sensitive to A because it controls both the main channel travel time and the time available for dispersion and transient storage to occur. It is also consistent with the results of, e.g., Mrokowska and Osuch (2011) and Wagener et al (2002). In these studies the results of Monte Carlo simulations showed that A was the most easily identifiable parameter of the TSM.

Indeed, these and other studies highlight the problem of equifinality with the TSM, namely that many different combinations of the parameters can give equally good simulations. A corollary of this is that we can't have much confidence in the values of the optimised parameters. This is often expressed in terms of estimates of uncertainty in the parameter values or in terms of the identifiability of the parameters (Wagner and Harvey, 1997; Wagener et al, 2002). Clearly, if there isn't a unique parameter set that gives an optimum fit to observations, then the differences between the results of the Polish and UK models may simply be a reflection of differences between the two optimisation methods.

Several authors have discussed the importance of the Damkohler number, DaI, in relation to the identifiability of the TSM parameters (e.g. Wagner and Harvey, 1997; Scott et al, 2003; Worman and Wachniew, 2007). DaI expresses a dimensionless ratio of main channel to storage zone residence times and is defined as

$$DaI = \frac{\alpha(1 + A/A_s)L}{U} \tag{5}$$

When DaI is small the solute exchange between the main channel and the storage zones takes place so slowly that observed concentration-time profiles do not reflect the interaction between them. When DaI is large the solute exchange between the main channel and the storage zones takes place so quickly that it is impossible to differentiate between dispersion caused by the main channel shear flow and by the transient storage. Wagner and Harvey (1997) suggest that reliable parameters are only identifiable for $0.1 < DaI < 1.0$. Few of the Cases analysed herein fall within this range (see Table 2) so this may be contributing to the differences between the results of the Polish and UK models. However, there is great uncertainty in the values of DaI because they are derived from the optimised parameter values. What we really need, of course, are the true DaI values, i.e. those corresponding to the physical situations, but these remain elusive unless we can be certain that the optimised parameter values are robust.

Irrespective of the above, it is interesting to note that Case 7 shows the best agreement between the two models. The data used were derived from an analytical solution to the problem (Hart, 1995) so these data (unlike the rest) are not subject to experimental errors in field work, laboratory analysis or data analysis. Indeed a comparison of the optimised parameter values and the values used to generate the data ($A = 2$ m^2, $D = 1$ m^2/s, $\alpha = 0.00025$ 1/s, $A_s = 0.1$ m^2) is very encouraging for A and D, even if less so for α and A_s.

Similarly, not with-standing the equifinality and identifiability issues, the extent to which other issues might be contributing to the differences between the parameter values from the two models is considered below. Four obvious issues stand out. Firstly, the equations on which the two models are based are not identical, secondly the numerical schemes used to solve the equations are different, thirdly different optimisation algorithms are used and fourthly different parameters are optimised.

In regard to the model equations, the Polish model is based on a so-called non-conservative formulation whereas the UK model is based on a so-called conservative one. The distinction between these formulations is only usually important if there are discontinuities in parameter values or solutions of the equations. So for the steady, uniform flows in question here, in which all parameters are constants, we would not expect this to be a serious issue. The different numerical schemes can be expected to be an influence because they introduce different errors, although we might expect a reduction in this effect as the spatial resolution of the problem increases. Of course, numerical errors might also be introduced through the temporal resolution of the problem, but here this has been fixed at an identical level in both models and has not been studied. Both models employ a global optimisation approach (although the techniques are different), so it is unlikely that either is getting stuck in local minima. Finally, it is not known what differences might be introduced by optimising A, D, α and A_s directly in the Polish model whilst optimising U, D, k_1 (= α) and k_2 in the UK model and then calculating A and A_s from these values. Although U and A are simply related through Q (which

is known), the calculation of A_s in the UK model relies on using two optimised parameter values (k_1 and k_2), which may amplify errors in A_s.

Conclusions

The application of two different versions of the TSM to the same solute concentration data has yielded different optimum parameter values. Although it is comforting that the main channel cross-sectional area shows the best agreement between the two models, the differences in the values of the other three parameters are a source of concern. Indeed, if optimised parameter values are not independent of the modeling/optimisation systems being used, this raises serious concerns over the robustness and transferability of the values. It has not been possible to establish the cause of these differences but it is likely to be mainly associated with equifinality and identifiability issues of the problem under study. Nevertheless, various modeling errors derived from the form of the governing equations used, the numerical schemes employed and the combination of parameters being optimised cannot be ruled out. The study has posed more questions than it has answered.

References

Bencala KE, Walters RA (1983) Simulation of solute transport in a mountain pool-and-rifle stream: a transient storage model. Water Resour Res 19:718-724.

Bottacin-Busolin A, Marion A, Musner T, Tregnaghi M, Zaramella M (2011) Evidence of distinct contaminant transport patterns in rivers using tracer tests and a multiple domain retention model Adv Water Resour 34;737-746

Deng Z-Q, Jung H-S, Ghimire B (2010) Effect of channel size on solute residence time distributions in rivers. Adv Water Resour 33:1118-1127

Fischer HB (1967) The mechanics of dispersion in natural streams. J Hydraul Div Proc Am Soc Civ Eng 93:187-216

Fischer HB (1968) Dispersion predictions in natural streams. J Sanit Eng Div Proc Am Soc Civ Eng 94:927-943

Hart DR (1995) Parameter estimation and stochastic interpretation of the transient storage model for solute transport. Water Resour Res 31:323-328

Mrokowska MM, Osuch M (2011) Assessing validity of the dead zone model to characterize transport of contaminants in the River Wkra. In: Rowinski P, (ed) Experimental Methods in Hydraulic Research, Springer 235-245

Price KH, Storn RM, Lampinen, JA (2005) Differential Evolution: A Practical Approach to Global Optimization. Springer.

Runkel RL, Broshears RE (1991) One-dimensional transport with inflow and storage (OTIS): A solute transport model for small streams. CADSWES Dept Civ Eng Univ Colorado

Runkel RL, Chapra SC (1993) An efficient numerical solution of the transient storage equations for solute transport in small streams. Water Resour Res 29:211-215

Rutherford JC (1994) River Mixing. Wiley

Scott DT, Gooseff MN, Bencala KE, Runkel RL (2003) Automated calibration of a stream solute transport model: implications for interpretation of biogeochemical parameters. J N Am Benthol Soc 22:492-510

Seo IW, Cheong TS (2001) Moment-based calculation of parameters for the storage zone model for river dispersion. J Hydrol Eng Am Soc Civ Eng 127: 453-465

Singh SK, Beck MB (2003) Dispersion coefficient of streams from tracer experiment data. J Environ Eng Proc Am Soc Civ Eng 129:539-546

Storn RM, Price KH (1997) Differential Evolution – A simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11:41-359

Taylor GI (1954) The dispersion of matter in turbulent flow through a pipe. Proc R Soc Lond A 233:446-468

Wagener T, Camacho LA, Wheater HS (2002) Dynamic identifiability analysis of the transient storage model for solute transport in rivers. J Hydroinform 4:199-211

Wagner BJ, Gorelick SM (1986) A statistical methodology for estimating transport parameters: theory and applications to one-dimensional advective-dispersive systems. Water Resour Res 22:1303-1315

Wagner BJ, Harvey JW (1997) Experimental design for estimating parameters of rate-limited mass transfer: analysis of stream tracer studies. Water Resour Res 33:1731-1741

Worman A, Wachniew P (2007) Reach scale and evaluation methods as limitations for transient storage properties in streams and rivers. Water Resour Res 43: W10405 doi:10.1029/2006WR005808

Yang G, Reinstein LE, Pai S, Xu Z, Carroll DL (1998) A new genetic algorithm technique in optimization of permanent 125-I prostate implants. Med Phys 25:2308-2315