Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Contaminant Hydrology 80 (2005) 107–129
JOURNAL OF Contaminant Hydrology
www.elsevier.com/locate/jconhyd

Parameter and observation importance in modelling virus transport in saturated porous media—investigations in a homogenous system

Gilbert R. Barth [a,*], Mary C. Hill [b]

[a] *S.S. Papadopulos and Associates, Inc, 1877 Broadway, Suite 703, Boulder, CO 80302-5245, United States*
[b] *U.S. Geological Survey, 3215 Marine Street, Boulder, CO 80303-1066, United States*

Received 30 April 2004; received in revised form 24 June 2005; accepted 28 June 2005
Available online 3 October 2005

Abstract

This paper evaluates the importance of seven types of parameters to virus transport: hydraulic conductivity, porosity, dispersivity, sorption rate and distribution coefficient (representing physical–chemical filtration), and in-solution and adsorbed inactivation (representing virus inactivation). The first three parameters relate to subsurface transport in general while the last four, the sorption rate, distribution coefficient, and in-solution and adsorbed inactivation rates, represent the interaction of viruses with the porous medium and their ability to persist. The importance of four types of observations to estimate the virus-transport parameters are evaluated: hydraulic heads, flow, temporal moments of conservative-transport concentrations, and virus concentrations. The evaluations are conducted using one- and two-dimensional homogeneous simulations, designed from published field experiments, and recently developed sensitivity-analysis methods. Sensitivity to the transport-simulation time-step size is used to evaluate the importance of numerical solution difficulties. Results suggest that hydraulic conductivity, porosity, and sorption are most important to virus-transport predictions. Most observation types provide substantial information about hydraulic conductivity and porosity; only virus-concentration observations provide information about sorption and inactivation. The observations are not sufficient to estimate these important parameters uniquely. Even with all observation types, there is extreme parameter correlation between porosity and hydraulic conductivity and between the sorption rate and in-

* Corresponding author. Fax: +1 303 939 8877.
E-mail address: gbarth@sspa.com (G.R. Barth).

0169-7722/$ - see front matter. Published by Elsevier B.V.
doi:10.1016/j.jconhyd.2005.06.012

solution inactivation. Parameter estimation was accomplished by fixing values of porosity and in-solution inactivation.
Published by Elsevier B.V.

Keywords: Reactive transport; Sorption; Virus; Sensitivity; Parameter estimation

1. Introduction

Viral contamination of groundwater can occur in a variety of ways including landfills, open dumps, broken sewer pipelines, leaking septic tanks, urban runoff, and crop irrigation with treated sewage effluent (Zelikson, 1994; Yates et al., 1985; Keswick and Gerba, 1980) resulting in serious health problems (Craun, 1992, 1989; Parsonnet et al., 1989; Weissman et al., 1976). Most urban centers in the United States expend considerable effort to treat drinking-water supplies for a wide range of pathogenic contaminants, but similar precautions usually are not taken for many rural drinking-water sources and even some urban sources that rely on groundwater as their primary source of drinking water (Yates et al., 1985). The EPA's proposed Ground Water Rule (Code of Federal Regulations, 40 CFR 141 and 142, 2000) addresses this issue by requiring that the potential for contamination of groundwater drinking supplies be assessed. The objective is to keep the risk of infection below one per 10,000 people per year. Maximum allowable concentrations for this level of risk can be calculated based on drinking water consumption and dose response which, for viruses, translates to a maximum allowable concentration of 1.8×10^{-7} plaque-forming units per liter (pfu L^{-1}) (Regli et al., 1991), or 1.0 pfu per 5.5×10^{6} L. Given that source concentrations have been reported in the range of 0.02-10 pfu L^{-1} (Schijven et al., 1999), with higher values expected for incidental storm-water overflow, concentrations ranging over eight orders of magnitude should be considered.

Virus transport simulation can serve as an effective tool for evaluating contamination potential but requires accurate model construction and parameter values. Variability of site conditions, and the typically short distances over which viruses remain a potential threat to water quality, suggest that a site-specific assessment of the mechanisms controlling virus transport is necessary to produce accurate transport predictions (Bales et al., 1997).

Most simulations of virus transport have been one-dimensional (1D), and have represented 1D laboratory experiments or simple field experiments (e.g., Johnson et al., 1996; Harvey and Garabedian, 1991; Bales et al., 1997; Schijven et al., 1999). Here we consider 1D and 2D simulations of a homogeneous system, though, as is shown, the 2D simulation can be closely approximated by a 1D simulation.

The published virus transport simulations have used a variety of mechanisms, including advection, dispersion, sorption, reaction, and velocity enhancement (e.g., Johnson et al., 1996). Campbell-Rehmann and Welty (1999), among others, also considered colloid filtration. Velocity enhancement refers to virus transport at rates faster than the conservative solute. Velocity enhancement has been observed (e.g., Becker et al., 1999) or predicted (e.g., Campbell-Rehmann and Welty, 1999). However, field and laboratory investigations by Harvey et al. (1993), Bales et al. (1997), and Schijven et al. (1999) detected minimal or no velocity enhancement. Here we consider all of the

mechanisms listed except for velocity enhancement and colloid filtration. This is reasonable considering the absence of consistent results for velocity enhancement, and that, due to the small size of viruses, potential for significant colloid filtration and velocity enhancement would typically be anticipated only in porous media with very small pore sizes. While additional mechanisms are important at some sites, the mechanisms considered here are common to nearly all sites and often dominate virus transport.

An important use of a model is to perform sensitivity analysis to determine the relative importance of parameters to predictions of interest, and the ability of available observations to estimate those parameters. Sensitivity analysis can be used to help identify field efforts most likely to improve the simulation of system dynamics and predictions of interest, and to justify these efforts to resource managers.

Here we consider the sensitivity-analysis methods of Hill (1998) and Barth and Hill (2005). These methods are based on sensitivities defined as the derivatives of simulated values with respect to parameters. The simulated values considered in this work are either the simulated equivalents to observations or the predictions of interest. Sensitivities can be calculated using sensitivity-equation, adjoint-state, and perturbation-based methods (e.g., Yeh, 1986). Perturbation-based methods are easy to apply to any model and are used in this work. They are calculated using UCODE (Poeter and Hill, 1998). The sensitivity-analysis methods use scaled sensitivities and parameter correlation coefficients, and are very versatile. For example, the same methods have been applied to groundwater flow problems by Poeter and Hill (1997) and Hill (1998); groundwater transport problems by Anderman et al. (1996) and Barlebo et al. (1998), and surface-water transport by Scott et al. (2003).

Sensitivity analysis of virus transport has been addressed in a few publications. Yates and Jury (1995) used a one-dimensional, analytical, virus-transport model that included inactivation and equilibrium sorption to evaluate the sensitivity of peak simulated virus concentrations to system parameters. Stochastic analysis by Campbell-Rehmann and Welty (1999) examined a wide range of virus-transport mechanisms and demonstrated the effect of individual parameters on virus-transport breakthrough curves, over distances large enough for the stochastic interpretation to be valid. Sensitivity analysis of conservative transport has been more common; for example Knopman and Voss (1987) and Barlebo et al. (1998).

This work differs from previous investigations in four ways. (1) The sensitivity analysis methods used are substantially different. (2) This work focuses on travel distances on the order of tens of meters, which are much shorter than those considered by Campbell-Rehmann and Welty (1999). This is important because in many field situations viruses maintain dangerous concentrations only over short travel distances. (3) This work considers many more types of typically available field observations than have been considered in previous works, including hydraulic heads, flow, conservative-transport concentrations, and virus concentrations. (4) This work takes advantage of four methods identified in Barth and Hill (2005): (a) conservative-transport breakthrough curve moments are used instead of individual conservative-transport observations, (b) the sensitivity of the transport step size (*TSS*) is used to indicate the minimum significant parameter sensitivity, (c) simulated-value weighting is designed such that virus-transport

observations and predictions properly account for concentrations that vary over many orders of magnitude, and (d) precautions are taken to produce accurate calculated virus-concentration sensitivities despite the very wide range of concentrations that need to be considered.

This work investigates a homogeneous system for a select range of parameter values, mimicking conditions from a published field experiment (Schijven et al., 1999). Similar experiments in a heterogeneous system may produce different spatial trends in parameter importance, relative parameter importance, parameter correlation and the potential for parameter estimation. Future work on the scale of interest, on the order of tens of meters, may address some basic heterogeneous configurations but this investigation's focus on a homogeneous system produces results which might otherwise be masked or misinterpreted by the effects of heterogeneity. The analysis performed in this work is repeated for several parameter sets reflecting slightly more than the range of conditions in the field experiments by Schijven et al. (1999). While the range of conditions evaluated does not encompass the entire range of possible virus transport conditions, it provides insight to some of the variability that can be expected for different parameter values.

The objectives of this paper are to, (1) determine which parameters are most important to virus transport, and how that importance changes over time to determine the utility of early-time virus-transport observations for predicting late-time concentrations, (2) identify changes in relative parameter importance with distance from the source, (3) investigate the effects of transport-equation nonlinearity with respect to parameters on evaluations of parameter importance to simulated virus transport and observation importance to parameter estimation, and (4) demonstrate the potential for parameter estimation.

2. Methods

This section provides detailed information about (1) the mechanisms used to represent virus transport, (2) the virus-transport parameter values used in the sensitivity analysis and parameter estimation, (3) details of the numerical simulations, and (4) sensitivity-analysis statistics and weighting.

2.1. Virus-transport mechanisms

This work considers advection, dispersion, sorption (to represent physical–chemical filtration), and reaction mechanisms (to represent virus inactivation) to simulate virus transport. Each mechanism is discussed briefly in the following paragraphs. While advection and dispersion are common to virtually all subsurface transport, the combination of sorption and reaction mechanisms are used in this work to represent the interaction of viruses with the porous medium and their ability to persist.

Advective transport is calculated using the groundwater flow equation, as would be typical of a field problem. This differs from many works that estimate advection directly. In field problems, accurate characterization of groundwater flow is critical for predicting virus transport in the saturated subsurface because it determines advective transport and is an integral factor in determining potential for sorption (McCarthy et al., 1996).

Dispersion can smooth a concentration front and reduce the peak concentration. Even in the absence of sorption and reaction, dispersion is important to early breakthrough preceding the advection front and the tendency for a prolonged breakthrough curve (BTC) tail.

Physical–chemical filtration is represented as sorption. Virus size precludes other particle filtration mechanisms, such as surface caking, straining or, as previously mentioned, colloid filtration, from being a significant factor under most conditions. Sorption of two viruses commonly used in field and laboratory studies, MS2 and PRD1, has been shown to be reversible and kinetically limited with different attachment and detachment time scales (Bales et al., 1997; Kinoshita et al., 1993). Two sorption terms are used in this work, and can be interpreted or rearranged so that they provide a direct indication of different rates of sorption attachment and detachment.

Blanc and Nasser (1996) observed virus inactivation to be a first-order irreversible reaction. Schijven et al. (1999) simulated virus inactivation as a first-order irreversible reaction. In this work virus inactivation is represented as a first-order irreversible reaction.

Using the mechanisms identified above, virus transport can be simulated as a sorptive, reactive solute using the advective–dispersive equation (e.g., Corapcioglu and Haridas, 1984; Tim and Mostaghimi, 1991; Schijven et al., 1999).

$$\frac{\partial C}{\partial t} = \underbrace{\frac{\partial}{\partial x_i}\left(D_{ij}\frac{\partial C}{\partial x_j}\right)}_{\text{dispersion}} - \underbrace{\frac{\partial}{\partial x_i}(v_i C)}_{\text{advection}} - \underbrace{\frac{\beta}{\theta}\left(C - \frac{\overline{C}}{K_d}\right)}_{\text{non-equilibrium sorption}} - \underbrace{(\lambda 1)C}_{\text{inactivation}} \tag{1a}$$

$$\frac{\partial \overline{C}}{\partial t} = \underbrace{\frac{\beta}{\rho_b}\left(C - \frac{\overline{C}}{K_d}\right)}_{\text{non-equilibrium sorption}} - \underbrace{(\lambda 2)\overline{C}}_{\text{inactivation}} \tag{1b}$$

$$v_i = -\frac{K}{\theta}\frac{dh}{dx_i} \tag{1c}$$

The following list defines the terms used in Eqs. (1a,b,c,d).

C	Concentration in solution [ML^{-3}]	$\overline{C}$	Adsorbed concentration [MM^{-1}]
t	Time [T]	ρ_b	Bulk density [ML^{-3}]
x_i	Spatial dimension i [L]	K	Hydraulic conductivity [LT^{-1}]
D_{ij}	Dispersion tensor [L^2T^{-1}]	α_l	Dispersivity [L]
v_i	Interstitial velocity [LT^{-1}]	θ	Porosity
β	Sorption rate [T^{-1}]	K_d	Sorption distribution coef. [L^3M^{-1}]
$\lambda 1$	In-solution inactivation rate [T^{-1}]	$\lambda 2$	Adsorbed inactivation rate [T^{-1}]
h	Hydraulic head [L]	D^*	Molecular diffusion [L^2T^{-1}]

Conceptually, the relationship between the D_{ij}, α, v_i and D^*, is most easily summarized using a curvilinear coordinate system along the direction of a flowline (Freeze and Cherry, 1979):

$$D_l = \alpha_l \upsilon + D^* \tag{1d}$$

where the subscript l indicates the curvilinear coordinate direction taken along the flowline. A full expansion of the fourth order dispersion tensor can be found in many subsurface transport texts (e.g., Zheng and Bennett, 2002). With respect to dispersivity, this work focuses on longitudinal dispersivity, α, for systems where the primary longitudinal axis is aligned with the principal direction of flow.

This investigation focuses on the seven parameters (K, θ, α, β, K_d, $\lambda 1$, $\lambda 2$). This approach is useful because in most systems much of the model uncertainty is associated with uncertainty in these seven parameter values.

2.2. Parameter data sets

The base set of parameter values, boundary conditions and system stresses used in this work closely approximate conditions from a series of field experiments by Schijven et al. (1999). Conservative- and virus-transport experiments, with observations collected for 125 days, were performed in a homogeneous water-supply aquifer in Castricum, the Netherlands. Results and analysis focused on four wells, W1–W4, ranging from 2.4 (W1) to 10 (W4) m down gradient of the source. Schijven et al. (1999) simulated transport using a one-dimensional form of (1a) and (1b) and estimated all but one parameter for each well, W1–W4, using a modified version of CXTFIT (Toride et al., 1995). The value of parameter $\lambda 1$ was determined from laboratory experiments.

In the present work, simulations are conducted using the sets of parameter values shown in Table 1. Set A is the base set of parameter values and consists of the average sorption and inactivation parameters estimated by Schijven et al. (1999) for the four wells. Parameter set B includes the much lower sorption rates estimated using observations from well W4 (Schijven et al., 1999). In parameter set C, the inactivation coefficients are increased by one order of magnitude to examine sensitivities under conditions more hostile to viruses. The other parameter sets, D and E, are used as

Table 1
Sets of parameter values used in the simulations

Parameter	Parameter Set				
	A	B	C	D/D2[a]	E[a]
K, hydraulic conductivity (m day^{-1})	12[b]	12[b]	12[b]	**6.0**[d]	**6.0**[d]
α, dispersivity, (m)	0.032[b]	0.032[b]	0.032[b]	**0.06**[d]	**0.06**[d]
θ, porosity	0.35[b]	0.35[b]	0.35[b]	0.35[b,e]	0.35[b,e]
K_d, sorption distribution coefficient (m^3 kg^{-1})	0.238[b]	**0.040**[d]	0.238[b]	**0.040**[d]	**0.040**[d]
β, sorption rate (day^{-1})	0.747[b]	**0.245**[d]	0.747[b]	**0.245**[d]	**0.245**[d]
$\lambda 1$, in-solution inactivation (day^{-1})	0.075[c]	0.075[c]	**0.750**[d]	0.075[c,e]	0.075[c,e]
$\lambda 2$, adsorbed inactivation (day^{-1})	0.07[b]	0.07[b]	**0.70**[d]	0.07[b]	**0.30**[d]
TSS, transport step size (day)	0.01	0.01	0.01	0.01	0.01

[a] Parameter sets used to start regressions.
[b] Average of the values estimated for wells W1–W4 by Schijven et al. (1999) using CXTFIT (Toride et al., 1995).
[c] Measured in laboratory experiments (Schijven et al., 1999).
[d] **Bold** values differ from Set A values.
[e] Underlined values were fixed to obtain successful regression results.

starting values for the parameter-estimation runs. Parameter sets D and E combine the low sorption and high inactivation of B and C, respectively, with incorrect initial values of K and α.

2.3. Virus-transport simulation

Virus transport is simulated using the groundwater flow model MODFLOW (McDonald and Harbaugh, 1988) with the transport model MT3DMS (Zheng, 1998). Double precision versions are used to reduce the effects of round-off error on the sensitivity-analysis and regression results described by Poeter and Hill (1998) and Hill and Østerby (2003). The MODFLOW PCG2 solver (Hill, 1990) was used. MT3DMS uses operator splitting so that the advective term is solved separately from the other terms. The Total Variation Diminishing (TVD) method was used to solve the advection term, and the standard explicit method was used to solve the other terms. The implicit method was not used because most of the simulations included a sharp front, requiring a longer implicit-method execution time for comparable solution accuracy.

2.3.1. Finite-difference grid and boundary conditions

The two-dimensional grid consists of 40 rows and 160 columns. Each finite-difference grid cell measures 0.0254 by 0.0635 m, producing a domain just over 1 m tall and 10.16 m long (Fig. 1). The system is saturated and does not experience any dewatering, so it is simulated as a confined system. The upstream boundary was represented as a constant head and the downstream boundary as a general head boundary with a loss of about 0.33 m producing a gradient of about 0.03. The source is centered along the upgradient boundary, and spans the central 0.254 m. For ease of interpretation, an arbitrary initial concentration equal to 1.0 was used, which is analogous to performing the analysis on concentrations normalized to the initial concentration.

2.3.2. Generating head, flow, and conservative-transport concentration observations

The analysis in this work uses only simulated values. Although the system investigated is based on a real site, the "observed" and "simulated" values are both generated by numerical simulation. As discussed below, the distinction between simulated and observed is the noise added to the observed values. The added noise therefore represents the entire uncertainty between simulated and observed in this work. Knowing the uncertainty, the weights can be assigned correctly allowing the analysis to focus on: relative parameter

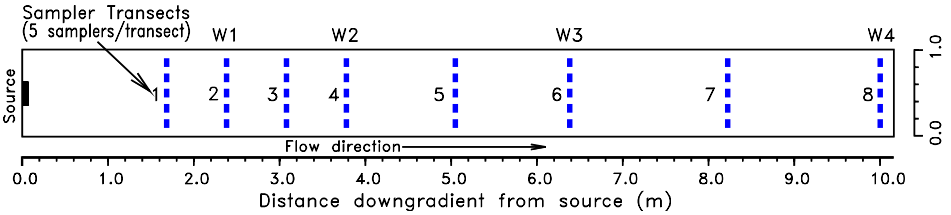

Fig. 1. Two-dimensional model domain, source and observation locations. Even-numbered sampler transects correspond to wells sampled by Schijven et al. (1999) and include the W1–W4 identifiers used in that paper.

importance, spatial trends in parameter importance, parameter correlation, and the potential for parameter estimation. Although beyond the scope of this work, future work could evaluate the impact of decreasing accuracy in the knowledge of the uncertainty by using several sets of weights, reflecting a range of error in the uncertainty of the observed values.

Following the methods of Barth and Hill (2005), forward runs using parameter set A (Table 1) were used to produce heads, flows, and conservative and virus concentrations, $\hat{y}^k$, where k indicates the observation type: flow $k=f$, head $k=h$, or concentration $k=c$. Noise was added to these values to create the observations. A brief overview of the methods associated with generating noise is provided in this work, details of the process can be found in Barth and Hill (2005). Noise with a standard deviation of about 0.003 m, intended to represent a typical pressure-transducer resolution, was added to the simulated heads to create the observations (y^h); noise with a standard deviation that was about 1% of the flow through the system was added to the flow to create the flow observations (y^f). Noise that tended to be larger for larger concentrations was added to the forward-run simulated transport values ($\hat{y}^c$), to create transport observations (y^c). In general, the coefficient of variation for concentrations was 0.08, but was modified for small concentrations, as described by Barth and Hill (2005).

Eight observation transects are located downgradient from the source (Fig. 1). Each transect has five observation locations for head and transport. Each transport observation is comprised of the flux-averaged concentration from four finite-difference cells, analogous to sampling from a well screened over a finite interval. In each transect the third observation is on the longitudinal centerline.

The simulated flow system is steady state, so there is one head observation for each observation location. Bulk flow through the system is used as the single flow observation. Concentration observations were defined at half-day intervals for the first 37 days and then with increasing intervals, as large as 12 days towards the end of the 125-day simulation period.

2.3.3. Temporal moments of conservative-transport observations

Temporal moments are used in both the sensitivity analysis and parameter estimation as an alternative to using conservative-transport observations directly. Temporal moments provide two benefits (Barth and Hill, 2005), (1) sensitivities are less susceptible to minor asymmetries and variability associated with numerical errors that occur for conservative transport in highly advective systems, and (2) the BTC moments remain sensitive even when the predicted conservative-transport BTC does not overlap the observed BTC. In this paper, the normalized first moment (m_1) calculated at each observation location is used. m_1 is calculated as

$$m_1 = \frac{M_1}{M_0} \quad \text{where} \quad M_j = \int_0^\infty t^j y^c(x,t)dt \quad \text{and} \quad j = 0, 1 \tag{2}$$

where t is time and $y^c(x,t)$ is concentration as a function of space and time. m_1 provides a lumped indicator of the arrival time of the center of mass. There are 8 temporal moment observations in the one-dimensional simulations, and 40 in the two-dimensional simulations.

2.3.4. Virus transport as observations or predictions

The purpose of most models of virus transport is to predict virus concentrations for conditions or times for which observations are not available. The term "prediction", as used in this work, refers to numerical-simulation estimates. Predictions must be interpreted with regard for the inherent limitations associated with any numerical representation of a physical process. These limitations are due to factors including assumptions in the governing equations, physical-property parameterization, and the numerical solution. Despite these limitations the predictions provide a significant contribution, and most importantly provide the opportunity to examine how changes to the system may affect virus transport. Virus transport observations that exist for other conditions or times can be used as observations in model development. Thus, simulated virus transport can be used in two ways—as predictions and as simulated equivalents to observations. In this work, simulated virus transport is used both ways to investigate two issues: the importance of model parameters to virus-transport predictions, and the importance of observations to estimating the parameters. In either case, as previously mentioned, there are always the inherent uncertainties associated with simulated values.

Virus-transport concentrations are simulated in the same manner and for the same times as conservative-transport observations. When used as observations, the same method is used to add noise to the simulated values. Temporal moments are not used for virus concentrations because of their long BTC tails.

2.3.5. Observation weighting

Weighting is needed to include all types of observations in a single objective function for parameter estimation, and is used to calculate the statistics used in the sensitivity analysis. Weighting is determined based on the observation errors, as is consistent with theoretical considerations (Hill, 1998). In this work, the errors in the observations are generated, and the weighting equals one divided by the variance of the errors specified above. For virus-transport concentrations, the standard deviation generally equals the coefficient of variation times the observed concentration, with an adjustment made for very small concentration, as described by Barth and Hill (2005). Weights on the moments of conservative transport observations were set equal to one divided by 10% of the typical moment value.

2.4. Statistics of the sensitivity analysis

Prior to attempting parameter estimation, sensitivity analysis can be used to, (1) identify the parameters most important to predictions of interest, (2) identify the information observations contain for estimating parameters and identify parameters for which so little information is available that they are unlikely to be estimated by regression, and (3) identify parameters that are so correlated given the available observations that they cannot be uniquely estimated. In practice, the results of 1–3 are used to design a model development strategy that best serves the issues of concern (Hill, 1998).

In this work, observations, y_i^k, and parameters, b_l, represent a variety of quantities and predictions, z_m, represent virus concentrations that can vary over many orders of magnitude. Sensitivity analysis is performed using sensitivities defined as the derivatives

$\partial \hat{y}_i^k / \partial b_l$ or $\partial z_m / \partial b_l$. The sensitivities are calculated using a centered finite-difference approximation by UCODE (Poeter and Hill, 1998). The sensitivities can have very different units and magnitudes caused by the characteristics of the quantities involved. To obtain quantities that can be used to compare the importance of different observations to different parameters or the importance of different parameters to different predictions, the sensitivities need to be scaled. Here, dimensionless and prediction scaled sensitivities are used, as defined below. To obtain statistics that summarize the information provided for each parameter, the scaled sensitivities have to be combined in a meaningful way. Here, composite scaled sensitivities are used.

The sensitivities of hydraulic heads to the parameters considered in this work are small because of system boundary conditions, a constant head at the upgradient end and a general head boundary at the downgradient boundary, and the homogeneous hydraulic conductivity field. In many circumstances, factors such as heterogeneity will increase the sensitivity of hydraulic heads. Despite their limited contribution for the conditions simulated, head observations are included to emphasize the method: incorporating many observation types, and presenting the approach to the readers.

2.4.1. Dimensionless scaled sensitivities

Sensitivity-analysis statistics used in this work include dimensionless scaled sensitivities (dss), composite scaled sensitivities (css), prediction scaled sensitivities (pss), and parameter correlation coefficients (ρ), as suggested by Hill (1998). The dss_{ij} are used to evaluate the importance of the ith observation in the estimation the jth parameter. For a diagonal weight matrix ($\underline{\omega}$), the dimensionless scaled sensitivities for observation y_i and parameter b_j are calculated as,

$$\mathrm{dss}_{ij} = \left(\frac{\partial y_i}{\partial b_j}\right) b_j \omega_i^{1/2} \tag{3}$$

where $\omega_i^{1/2}$ is the square root of the weight of observation y_i, and in this work equals one divided by the standard deviation of the observation error. The scaling provided by b_j and $\omega_i^{1/2}$ allows the scaled sensitivities to be compared. Each dss indicates the information available from each observation for estimating each parameter.

2.4.2. Composite scaled sensitivities

Composite scaled sensitivities (css) are calculated for each parameter as,

$$\mathrm{css}_j = \left[\sum_{i=1}^{\mathrm{ND}} \left(\mathrm{dss}_{ij}\right)^2 / \mathrm{ND}\right]^{1/2} \tag{4}$$

where ND is the number of observations and can include all head, flow, conservative-tracer, and virus transport observations, or a subset that could be based on observation type, location, and/or timing. css_j can be thought of as the average change in the simulated value, expressed as a percent of the observation error standard deviation, caused by a 1% change in the parameter value—it is a measure of the total information provided by the ND

observations for the estimation of the jth parameter. For this work, css values are calculated for (1) all types of observations for all times and locations, including virus transport; (2) all types of observations for all times and locations except for virus transport; and (3) all head and flow observations, and the conservative- and virus-transport observations from each observation location for all times. The last option allows evaluation of a common circumstance, obtaining concentrations from a single site, and provides insight to changes with distance from the source.

2.4.3. Prediction scaled sensitivities

In this work prediction scaled sensitivities are calculated as

$$\mathrm{pss}_{ij} = \left(\frac{\partial z_i}{\partial b_j}\right) b_j \omega_i^{1/2} \tag{5}$$

where for consistency the weight is used whether the virus concentration is considered as an observation or a prediction. From a practical viewpoint, the resulting prediction scaled sensitivity represents the amount the prediction would change, expressed as a percent of the accuracy with which the predicted quantity could be measured, given a 1% change in the parameter value. In some results, a summary statistic is produced by using prediction scaled sensitivities for virus transport for all times and locations in Eq. (5). The resulting statistics are called a composite of prediction scaled sensitivities (cpss), and can be used to evaluate the overall importance of each parameter to the simulation of virus transport.

2.4.4. Parameter correlation coefficients

Parameter correlation coefficients (ρ_{ij}) close to ± 1.00 indicate that the observations may not be sufficient to uniquely estimate parameters b_i and b_j. Parameter correlation coefficients should not be confused with the spatial correlation of stochastic hydraulic-conductivity fields. While parameter correlation coefficients only measure the correlation between parameter pairs, correlation between sets that include more than two parameters is evident because the parameter correlation coefficients between all possible pairs will be close to ± 1.00. Extreme parameter correlation easily confounds parameter-estimation attempts in both simple and complex systems (e.g., Carrera and Neuman, 1986; Poeter and Hill, 1997; Hill and Østerby, 2003).

2.5. Parameter estimation

To estimate parameters, the fit between observed and simulated values was quantified using a weighted least-squares objective function, $S(\underline{b})$, to measure the fit between simulated values $\underline{\hat{y}}(\underline{b})$ and $\underline{y}$ observations.

$$S(\underline{b}) = \left[\underline{y} - \underline{\hat{y}}(\underline{b})\right]^T \underline{\omega} \left[\underline{y} - \underline{\hat{y}}(\underline{b})\right] \tag{6}$$

The weight matrix, $\underline{\omega}$, accounts for the different units and accuracy of the observations. The objective function is minimized using a modified Gauss–Newton method (e.g., Sun, 1994) as described in Hill (1998)) and implemented in UCODE.

3. Results and discussion

Results are presented in three sections: (1) comparison of simulated conservative- and virus-transport concentrations, (2) relative observation and parameter importance as determined by sensitivity analysis, and (3) parameter-estimation results. It is important to consider that the results discussed in each section reflect the homogeneous system evaluated at the parameter values listed in Table 1. Different parameter values, of for that matter heterogeneity, will obviously produce different results. For all circumstances considered in this work, similar results were obtained using either the one- or two-dimensional models, so the results from the one- and two-dimensional models are used interchangeably.

3.1. Comparison of simulated conservative and virus transport concentrations

Except for the long trailing concentrations, simulated virus-transport BTCs, such as those shown in Fig. 2, are similar in shape to conservative-transport BTCs. This is advantageous because it suggests that the time of maximum virus-transport concentrations at any location can be determined based on conservative-transport tracer tests. Unfortunately, the enduring health risk of low concentrations of viruses means that in most circumstances additional information about the virus transport is required.

3.2. Sensitivity analysis

Fig. 3 shows composite scaled sensitivities for the system. If virus concentrations are used as observations, Fig. 3i shows that the virus concentrations would provide the most information toward estimating hydraulic conductivity, porosity, and the sorption parameter that represents physical–chemical filtration. If the virus concentrations are predicted, the composite scaled sensitivities of Fig. 3i can be thought of as composites of prediction scaled sensitivities (cpss).

Fig. 3ii shows how much information is available for estimating the parameters if all types of observations, including virus-transport observations, are available for model development. Based on these css values, more information is provided for K and θ than any of the sorption or inactivation parameters. The two primary reasons are (1) both conservative and virus concentrations provide information for K and θ, and (2) the simulated system is advection dominated, the Peclet number is equal to 2. Substantial information also is available to estimate sorption (β) and dispersivity (α). The importance of α suggests that even though there is little dispersion in this advection-dominated system, even a small increase in dispersion would be important. Progressively less information is available for estimation of the absorbed inactivation rate ($\lambda 2$), distribution coefficient (K_d), and in-solution inactivation rate ($\lambda 1$).

Fig. 3iii shows that without the rarely available virus-transport observations, there is, of course, no information available from the observations on any of the reactive-transport parameters, including sorption (β), the distribution coefficient (K_d), and the two inactivation rates ($\lambda 1$, $\lambda 2$). When efforts are taken to reproduce in situ solutions, laboratory-based evaluations can provide reasonable estimates of $\lambda 1$. However, the other

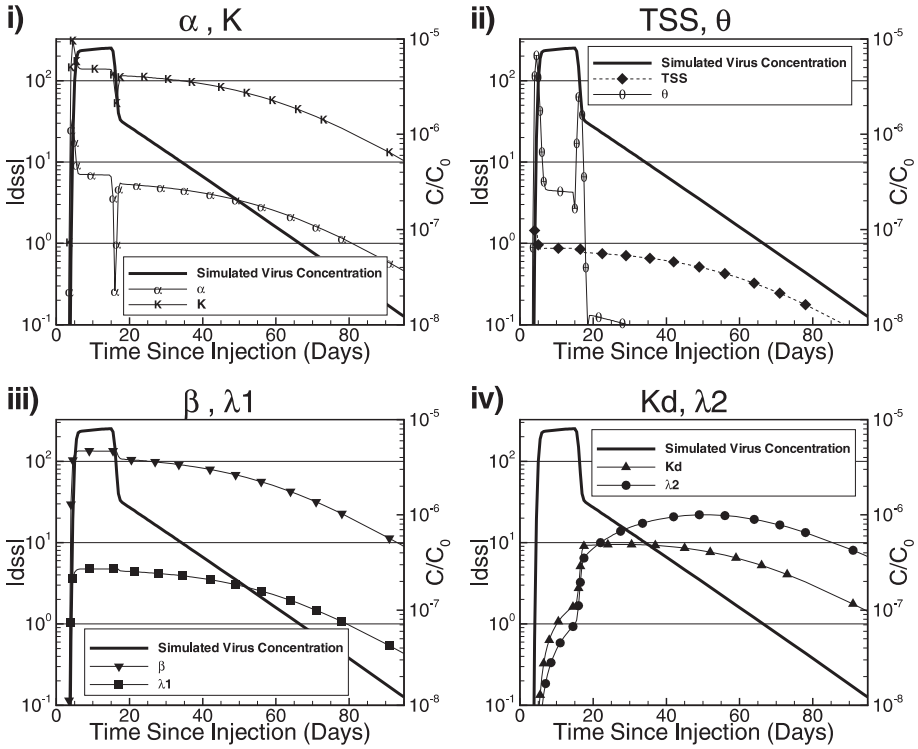

Fig. 2. The absolute value of dimensionless scaled sensitivities (|dss|) or prediction scaled sensitivities (pss) calculated for virus transport using parameter set A (Table 1) at transect six of a one-dimensional model for parameters representing (i) dispersivity (α) and hydraulic conductivity (K), (ii) transport step size (TSS) and porosity (θ), (iii) sorption rate coefficient (β) and in-solution inactivation ($\lambda 1$), and (iv) sorption distribution coefficient (K_d) and sorbed inactivation ($\lambda 2$). If the simulated values are associated with observations, the resulting dss values indicate the information provided by the virus-transport observations for estimating the parameters. If the simulated values are predictions, the resulting pss values indicate how important the parameter is to the predicted value.

parameter values (β,K_d, $\lambda 2$) are dependent on various porous media characteristics, so that laboratory-based estimates of these parameters have considerably more uncertainty.

Including TSS in Fig. 3 allows evaluation of whether the information provided by the observations is sufficient to overcome typical numerical inaccuracies (Barth and Hill, 2005). In Fig. 3ii, $css_{\lambda 1} < css_{TSS}$, indicating that for the observations considered in this graph, the system simulated, and the numerical methods used, estimates of in-solution inactivation, $\lambda 1$, need to be regarded with some suspicion. Thus, even when virus concentrations are available it can be difficult to obtain reliable estimates of $\lambda 1$ through model calibration.

BTCs and dimensionless scaled sensitivities (dss) for virus transport observations over time at a single sampling location in the one-dimensional model are shown in Fig. 2. At later times, in the long tail of the virus-concentration distribution, the values of dss are small for some parameters, and large for others. Without weighting that reflects the

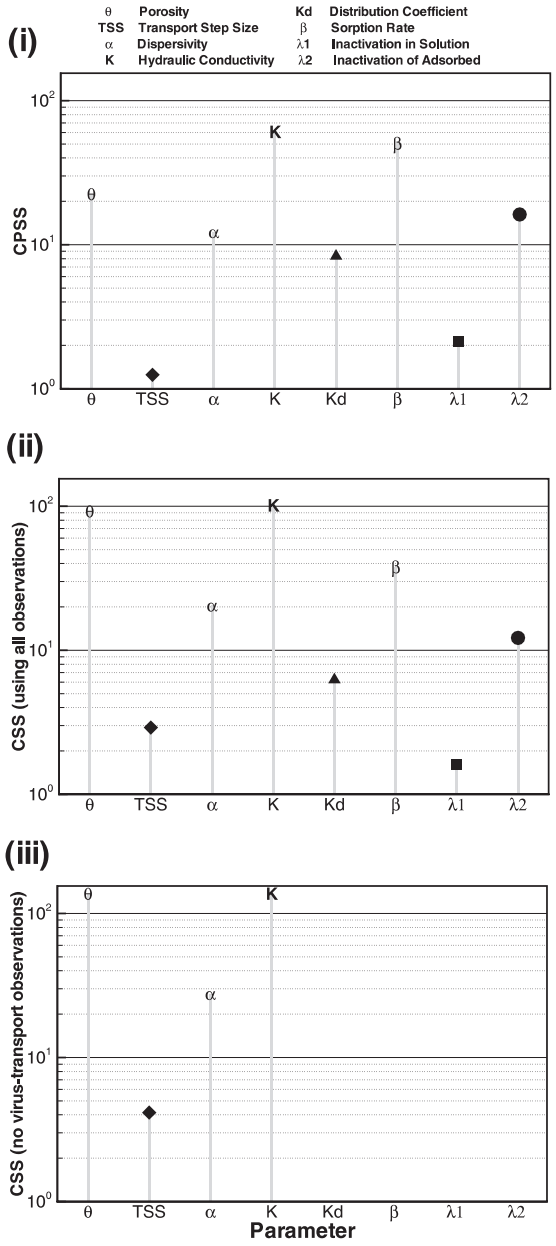

Fig. 3. Composite-scaled sensitivities (css) of the seven system parameters and the simulation transport step size, *TSS*, evaluated using parameter set A. Results are shown for (i) composited prediction scaled sensitivities (cpss) which are based only on simulated virus transport, (ii) css using all observations, and (iii) css without virus-transport observations. cpss evaluate the overall importance of each parameter to the simulation of virus transport. css indicate the amount of information that the observations provide. Hydraulic conductivity (K) and porosity (θ) are the most important parameters, *TSS* is more important than the rate of inactivation in solution ($\lambda 1$).

concentrations involved, the sensitivity analysis and regression would virtually ignore the low-concentration observations.

The graphs in Fig. 2 show the following for each of the defined parameters. For K, observations from the entire BTC provide significant information (Fig. 2i). For θ, most of the information is provided by virus-transport observations at the peak of the BTC (Fig. 2ii). To understand why the long tail provides so little information, consider Eqs. (1a,c) and processes in the tail. Eq. (1c) indicates that if θ increases, the fluid velocity slows. Eq. (1a) indicates that slower fluid velocities result in slower decreases in concentration over time caused by advective processes: virus concentrations do not drop off as rapidly. At the same time, the larger value of θ results in a smaller non-equilibrium sorption term. For the tail portion of a BTC, when the adsorbed concentrations may be higher than those in solution, the non-equilibrium sorption term will be negative. Under these conditions the impact of porosity changes to the non-equilibrium sorption term can potentially offset changes to the advection term. This is apparently the case for the results presented in Fig. 2 because the dss values for θ in the tail of the BTC are small.

For β and $\lambda 1$, both the BTC peak and tail provide significant information (Fig. 2iii). For K_d and $\lambda 2$, most of the information comes from the tail (Fig. 2iv).

In typical field-site investigations, virus concentrations included as observations are obtained at early times and short distances from the source to predict transport at greater time and distances. The shapes of the dss curves in Fig. 2 (with results at transect 6 using parameter set A) indicate that early-time observations of virus-transport at this location provide information on most parameters. However, early observations at this location provide minimal information for K_d and $\lambda 2$.

If virus concentrations are the predictions of interest instead of observations, the dimensionless scaled sensitivities of Fig. 2 can be thought of as prediction scaled sensitivities and Fig. 2 would show which parameters are important to predicting different parts of the BTCs. From the viewpoint of predicted virus transport, the results are most discouraging for K_d and $\lambda 2$, in that they are important only to the concentrations in the tail. This makes their accurate estimation problematic in most practical situations in which, at best, there might be virus concentrations at early times at this location. The possibility of virus-transport observations being available closer to the source is investigated using Fig. 4.

Fig. 4 shows css values from two-dimensional simulations using parameter sets A, B, and C. The css values are produced using Eq. (4) and include dss values for all observation types except flow for all times and for all five samplers in each transect. The css values for K, α, θ, and TSS are identical or similar for all three parameter sets; only results from set A are shown. The css for β, K_d, $\lambda 1$, and $\lambda 2$, differ between parameter sets A, B, and C, and results for all three are shown. The differences are mostly the result of model nonlinearity. Although the patterns shift substantially, the relative css values are remarkably stable. As will be shown, most of the conclusions drawn from the sensitivity analysis methods used in this work are robust in the presence of the nonlinearities present in this problem.

The differences in results for parameter sets A, B, and C for the reactive-transport parameters are easily understood. With less sorption (set B), viruses are transported to greater distances in larger numbers and the sorption and inactivation terms remain

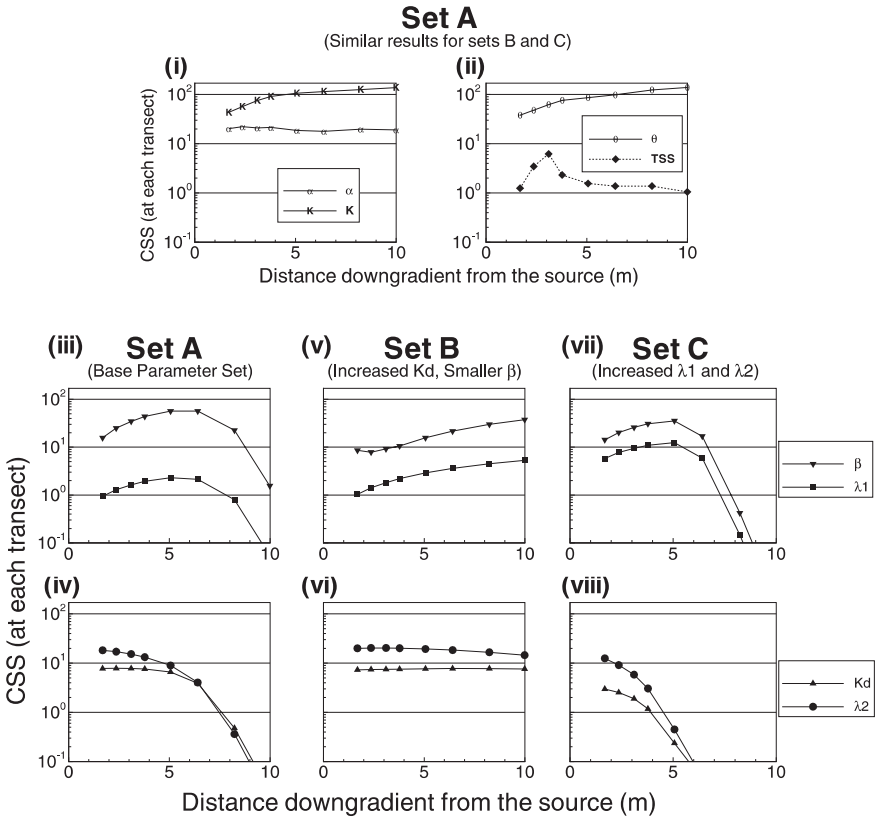

Fig. 4. The importance of observation location, as measured by composite-scaled sensitivities calculated for all observations over time at each distance from the source using parameter sets A, B, and C (Table 1) for (i) hydraulic conductivity (K) and dispersivity (α), (ii) porosity (θ) and transport step size (TSS), (iii, v, and vii) sorption rate (β) and in-solution inactivation ($\lambda 1$), and (iv, vi, and viii) sorption distribution coefficient (K_d) and adsorbed inactivation rate ($\lambda 2$). Transport-equation nonlinearity with respect to the parameters and scaling using the parameter value result in the parameter sets producing significantly different css values for the sorption/inactivation parameters.

important to greater distances. Higher inactivation (set C) decreases the numbers of viruses at larger distances, leading to a decline in sorption and inactivation sensitivity further from the source.

The observations provide the most information for parameters K and θ at all distances. Nearly as much information is provided for parameters α and β. The information provided for K_d and $\lambda 2$ is greatest for observations near the source. css_{TSS} is larger than $css_{\lambda 1}$ at many distances for parameter set A. css_{TSS} is larger than css_{K_d} and $css_{\lambda 2}$ at many distances for parameter set C.

The css in Fig. 4 indicate that as long as subsurface conditions are spatially consistent, observations near the source provide substantial information on parameters important to more distant transport. For data set A, the reactive parameters become less important at

distance because the viruses do not survive very long. For set B, decreasing K_d from 0.238 to 0.040 (an 83% reduction) and decreasing β from 0.747 to 0.245 (a 67% reduction), results in maintained virus concentrations and reactive-transport parameter importance. For set C, increasing $\lambda 1$ and $\lambda 2$ by an order of magnitude results in a more hostile environment for the viruses. Under these conditions parameter $\lambda 1$ has a larger css value than K_d and $\lambda 2$ for virtually all distances, indicating that the in-solution inactivation is the more important process.

The value of css_{TSS} varies with distance and has a maximum about 3 m from the source (Fig. 4ii). The cause is not clear, but may be related to grid size and simulated concentration gradients. While the graphs in Fig. 2 indicate that $dss_{\lambda 1} > dss_{TSS}$, Fig. 4ii demonstrates that this does not occur consistently and that when using simulated results from all of the transects, which is the case for Fig. 3, $css_{\lambda 1} < css_{TSS}$. In Fig. 4, the css results based on data from transect 6 indicate that at that distance from the source $css_{\lambda 1} > css_{TSS}$. However, at most distances, $css_{\lambda 1} < css_{TSS}$, which explains the results of Fig. 3ii. For the system simulated, the numerical inaccuracies indicated by the magnitude of css_{TSS} (Barth and Hill, 2005) in Fig. 4 suggest that the contribution of $\lambda 1$ to improving simulation results is quite limited at distances where, (1) conservative solute concentrations are significant, or more obviously, (2), where the virus concentrations become insignificant.

For parameters β, K_d, $\lambda 1$, and $\lambda 2$, sensitivities are non-zero for only virus-transport observations, so Fig. 4iii to viii can be interpreted as composited prediction scaled sensitivities if virus transport is considered to be a prediction.

For the homogeneous two-dimensional system with the source configuration considered, vertical-transverse observation-importance variability is small compared to longitudinal variability so that the transversely averaged values shown in Fig. 4 capture most significant trends and a full two-dimensional assessment is not necessary for these simple conditions. For example, Fig. 5 shows the longitudinal and transverse variations in css_β. The longitudinal trends are the same as seen in Fig. 4iii and v. Transverse variability is limited. The magnitude of transverse variability of other parameters was similar. For β and all other parameters, except dispersivity, sensitivity decreases with transverse distance from the domain centerline. For dispersivity, which can be a critical factor in determining plume spreading, the observations at the edge of the plume are very important. For all other parameters the observations along the plume centerline provide the most information for parameter estimation.

The extreme parameter correlation coefficients shown in Table 2 result when using observations of head, flow, moments of conservative transport, and virus transport, and parameter set A of Table 1. The high correlation between β and $\lambda 1$ is expected given their role in Eq. (1a). The high correlation is also evident in Figs. 2iii and 4iii, v, and vii, in which the patterns for β and $\lambda 1$ are translated but otherwise nearly identical. Considering the coupling between C and $\overline{C}$ in Eqs. (1a) and (1b), for some sets of parameters the correlation probably depends on other parameter values, with K_d being likely.

The high correlation between θ and K is expected given their role in (1c): the non-equilibrium sorption term of (1a) provides the only opportunity for θ and K to be identified uniquely. Again, the high correlation is evident in Figs. 2i,ii and 4i and ii; after the peak has passed the dss for θ resumes at much lower values, and this explains why the

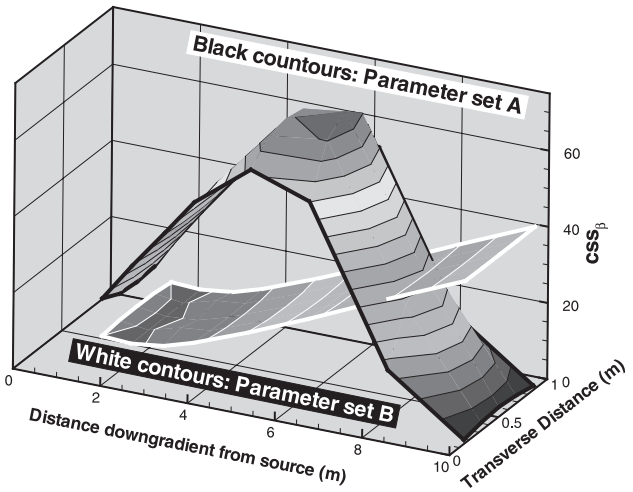

Fig. 5. Contours of composite scaled sensitivities evaluated for all observations at each of the 40 sampler locations for the sorption rate (β) for the two-dimensional system using parameter sets A and B. The flow observation is not included in the calculation. As with all other parameters, for β parameter-importance varies most in the direction of flow, demonstrating that for this problem the transversely averaged values shown in Fig. 4 capture the most significant trends.

correlation coefficient is not closer to 1.00. Eqs. (1a,b) suggest that the strength of the correlation between θ and K depends on the importance of the non-equilibrium sorption process and, therefore, on the value of β/θ. If β is small, the correlation between K and θ is expected to be larger. Eq. (1a) and the correlation between β and $\lambda 1$ also suggest that when $\lambda 1$ is large, inactivation dominates sorption and the correlation between K and θ is expected to be large. These results affect parameter estimation, as discussed below.

3.3. Estimating parameters

Fig. 3ii shows that the range in css across all the parameters is less than two orders of magnitude and the values are all greater than one, so that, in the absence of extreme correlation between any parameters, it should be possible to estimate all parameters (Hill, 1998, p. 38). However, Table 2 shows that parameter correlations are large enough to be potentially problematic. Indeed, preliminary attempts to estimate all parameters simultaneously from parameter set D in Table 1, and other sets not listed in this work,

Table 2
Parameter correlation coefficients calculated for parameter set A

Parameter 1	Parameter 2	Correlation coefficient[a] (ρ)
Porosity (θ)	Hydraulic conductivity (K)	0.91
Sorption rate (β)	Inactivation in solution ($\lambda 1$)	−0.97

[a] Only parameter correlations >0.8 are reported. Typically, only parameter correlation coefficients of 0.95 or larger cause problems with regression.

Table 3
Percent error in estimated parameter values

Regression run		Percent error in parameter value[a]							$S(\underline{b})$
		K	α	θ	K_d	β	$\lambda 1$	$\lambda 2$	
D	Initial	–50.0	89.0	0.0	−83.2	−67.2	0.0	0.0	5.2×10^9
	Final[b]	0.2	−0.2	0.0	−68.5	−43.74	1260	−2.6	3.56×10^3
D2[c]	Initial	–50.0	89.0	*0.0*	−83.2	−67.2	*0.0*	0.0	5.2×10^9
	Final	0.2	−0.2	*0.0*	−0.8	−0.1	*0.0*	−0.4	3.56×10^3
E	Initial	–50.0	89.0	*0.0*	−83.2	−67.2	*0.0*	900	5.9×10^9
	Final	0.2	−0.2	*0.0*	−0.8	−0.1	*0.0*	−0.4	3.56×10^3

Shaded cells indicate parameter values that were fixed.
[a]Calculated as $\frac{(\hat{b}_i - b_i)}{b_i} 100$, where $\hat{b}_i$ and b_i are the estimated and true parameter values, respectively.
[b]At these parameter values the parameter correlation coefficient between θ and K equals 1.00.
[c]Same as D except that θ and $\lambda 1$ are fixed.

were unsuccessful in that the estimated value of in-solution inactivation was unrealistically large and the other parameter estimates, although improved, were still not correct (Table 3). Successful parameter estimation required that two parameter values be assigned prior information or be fixed at assigned values. As suggested by Hill (1998), the two approaches produce virtually the same parameter values. Fixing the values resulted in shorter execution times and that approach is used in this work. First, one of the highly correlated parameters β or $\lambda 1$, was fixed. $\lambda 1$ was chosen because it is much less sensitive and independent determination of $\lambda 1$ from laboratory data is more likely. A second parameter needed to be set because if θ and K and either β or $\lambda 1$ are estimated, β increased and the correlation between θ and K approached 1.00 as the regression proceeded. Here we set θ because in unconsolidated deposits field data commonly constrain its value within a narrower range than can be achieved for K. This is consistent with Gelhar (1993), who noted that the field variability of θ in unconsolidated deposits is quite small so that typical variations in θ have a relatively minor impact on advective transport.

Table 3 shows that the same small value of the objective function, $S(\bar{b})$, is obtained in regression runs D, and D2 and E for a different set of parameter values. This demonstrates the existence of more than one minimum when all parameters are estimated. This is consistent with the parameter correlation coefficients of Table 2. The same parameters values are estimated for D2 and E, and this situation provides some indication that with the two parameter values fixed a unique minimum is defined. It is interesting, and not coincidental, that in run D the ratio of the percent errors in β and $\lambda 1$ is about 30, and that in Fig. 3ii the ratio of composite scaled sensitivities of 46 for these two parameters is of similar magnitude. Basically, the composite scaled sensitivities indicate that for an equal effect $\lambda 1$ needs to change much more than β, and the negative correlation coefficient indicates that an increase in one can offset a decrease in the other. The values of 30 and 46 probably differ for two reasons. First, model nonlinearity can make the composite scaled sensitivity change somewhat. Second, as demonstrated by Poeter and Hill (1998), model nonlinearity can make parameter correlation coefficients change significantly for different parameter values. In run D the final parameter values suggest that K_d is also involved in the correlation.

If the assigned values of θ and $\lambda 1$ had been in error in either of these runs, the estimated values of the associated correlated parameters would also be in error. This is most problematic when predictions of interests are sensitive to the individual fixed parameters, because smaller percent errors in the fixed parameter values can have a larger affect on estimated parameter values and, thereby, the simulated results.

4. Conclusions

Sensitivity-analysis and parameter-estimation results from the simulated system lead to the following set of conclusions.

1. As would be expected for the highly advective system simulated and the observations used, the observations provided the most information to hydraulic conductivity (K) and

porosity (θ) parameters. These were followed by, in order of decreasing relative importance, virus sorption-rate (β), dispersivity (α), the absorbed inactivation rate ($\lambda 2$), the distribution coefficient (K_d), and the in-solution inactivation rate ($\lambda 1$).
2. Even for the advective system examined, information in the tail is important to determining how long viruses remain a potential health threat.
3. The time and location at which observations are important depends on the plume characteristics and therefore the values of the parameters, but most of the conclusions drawn from the sensitivity analysis were robust.
4. For circumstances when virus-transport observations are available, the simulations performed indicate that virus-transport parameters β, K_d, and $\lambda 2$ can be estimated from site data.
5. Near-source concentrations that include observations of the BTC tail can be adequate, while early-time observations at locations further from the source may not provide sufficient information to estimate K_d and $\lambda 2$.
6. As demonstrated using sensitivities to the transport time-step size (*TSS*), it is unlikely that the inactivation in solution ($\lambda 1$) can be characterized even with virus-transport observations because of numerical inaccuracies involved with solving the equations for highly advective systems.

Parameter estimation results demonstrated the effects of extreme parameter correlation and its resolution using fixed parameters. Parameter insensitivity and extreme parameter correlation did not allow the simultaneous estimation of all seven parameters investigated. Values of θ and $\lambda 1$ were fixed. They were chosen because it is expected that values for these parameters can be obtained from independent information, and they were the least sensitive of each of the extremely correlated parameter pairs. For any pair of extremely correlated parameters, error in the fixed-parameter value will affect the estimation of the other parameter. Since $\lambda 1$ is insensitive, the affect of any error will be small. Error in θ would affect K, but the likely percent error in θ, and therefore K, is expected to be small. Methods for quantifying the impact of prior information (e.g., Weiss and Smith, 1997) can be helpful in selecting parameters to fix. However, the circumstances investigated in this work clearly demonstrate the need to fix at least two parameters, and that, even without quantifying parameter-fixing impacts, the choice of parameters is well defined by at least two of the following three factors, (1) parameter insensitivity, (2) the potential for independent parameter-value assessment, and (3) typical uncertainty of the parameter.

Acknowledgements

Funding for this research was provided through the National Academy of Sciences', National Research Council's Research Associateship Program in collaboration with the United States Geological Survey's National Research Program. Comments and suggestions from Steffen Mehl, Emil Frind and two anonymous reviewers helped to strengthen the manuscript and are greatly appreciated.

References

Anderman, E.R., Hill, M.C., Poeter, E.P., 1996. Two-dimensional advective transport in ground-water flow parameter estimation. Ground Water 34 (6), 1001–1009.

Bales, R.C., Li, S., Yeh, T.-C.J., Lenczewski, M.E., Gerba, C.P., 1997. Bacteriophage and microsphere transport in saturated porous media: forced-gradient experiment at Borden, Ontario. Water Resour. Res. 33 (4), 639–648.

Barlebo, H.C., Hill, M.C., Rosbjerg, Dan, Jensen, K.H., 1998. Concentration data and dimensionality in groundwater models, evaluation using inverse modelling. Nord. Hydrol. 29, 149–178.

Barth, G.R., Hill, M.C., 2005. Numerical methods for improving sensitivity analysis and parameter estimation of virus transport simulated using sorptive–reactive processes. J. Contam. Hydrol. 76, 251–277.

Becker, M.W., Reimus, P.W., Vilks, P., 1999. Transport and attenuation of caboxylate-modified latex microspheres in fractured rock laboratory and field tracer tests. Ground Water 37 (3), 387–395.

Blanc, R., Nasser, A., 1996. Effect of effluent quality and temperature on the persistence of viruses in soil, water reclamation and reuse 1995. Water Sci. Technol. 33 (10-11), 237–242.

Campbell-Rehmann, L.L., Welty, C., 1999. Stochastic analysis of virus transport in aquifers. Water Resour. Res. 35 (7), 1987–2006.

Carrera, J., Neuman, S.P., 1986. Estimation of aquifer parameters under transient and steady-state conditions: 1. Maximum likelihood method incorporating prior information. Water Resour. Res. 22 (2), 199–210.

Code of Federal Regulations (CFR), 2000. Environmental Protection Agency, National primary drinking water regulations: ground water rule; proposed rules, 40 CFR parts 141 and 142.

Corapcioglu, M.Y., Haridas, A., 1984. Transport and fate of microorganisms in porous media: a theoretical investigation. J. Hydrol. 72, 149–169.

Craun, G.F., 1989. Types and Effects of Microbial Contamination of Groundwater, EPA Report EPA/600/D-89/067. 32 pp.

Craun, G.F., 1992. Waterborne disease outbreaks in the United States of America: causes and prevention. World Health Stat. Q. 45 (2/3), 192–199.

Freeze, A.R., Cherry, J.A., 1979. Groundwater. Prentice-Hall Inc., New Jersey.

Gelhar, L., 1993. Stochastic Subsurface Hydrology. Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Harvey, R.W., Garabedian, S.P., 1991. Use of colloid filtration theory in modeling movement of bacteria through a contaminated sandy aquifer. Environ. Sci. Tech. 25 (1), 178–185.

Harvey, R.W., Kinner, N.E., MacDonald, D., Metge, D.W., Bunn, A., 1993. Role of physical heterogeneity in the interpretation of small-scale laboratory and field observations of bacteria, microbial-sized microsphere, and bromide transport through aquifer sediments. Water Resour. Res. 29 (8), 2713–2721.

Hill, M.C., 1990. Preconditioned conjugate-gradient 2 (PCG2), a computer program for solving ground-water flow equations, U.S. Geological Survey Water-Resources Investigations Report 90-4048. 43 p.

M.C., Hill, 1998. Methods and guidelines for effective model calibration, U.S. Geological Survey Water Resources Investigation Report 98-4005. 90 p.

Hill, M.C., Østerby, O., 2003. Determining extreme parameter correlation in ground water models. Groundwater 41 (4), 420–430.

Johnson, P.R., Sun, N., Elimelech, M., 1996. Colloid transport in geochemically heterogeneous porous media: modeling and measurements. Environ. Sci. Technol. 30 (11), 3284–3293.

Keswick, B.H., Gerba, C.P., 1980. Viruses in groundwater. Environ. Sci. Technol. 14 (11), 1290–1297.

Kinoshita, T., Bales, R.C., Maguire, K.M., Gerba, C.P., 1993. Effect of pH on bacteriophage transport through sandy soils. J. Contam. Hydrol. 14 (1), 55–70.

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

McCarthy, J.F., Gu, B., Liang, L., Mas-Pla, J., Williams, T.M., Yeh, T.-C.J., 1996. Field tracer tests on the mobility of natural organic matter in a sandy aquifer. Water Resour. Res. 32 (5), 1223–1238.

McDonald, M.G., Harbaugh, A.W., 1988. A modular three-dimensional finite-difference ground-water flow model. U.S. Geological Survey Techniques of Water Resources Investigations, book 6, chap. A1 586 pp.

Parsonnet, J., Trock, S.C., Bopp, C.A., Wood, C.J., Addiss, D.G., Alai, F., Gorelkin, L., Hargett-Bean, N., Gunn, R.A., Tauxe, R.V., 1989. Chronic diarrhea associated with drinking untreated water. Ann. Intern. Med. 110 (12), 985–991.

Poeter, E.P., Hill, M.C., 1997. Inverse models: a necessary next step in ground-water modeling. Ground Water 35 (2), 250–260.

Poeter, E.P., Hill, M.C., 1998. Documentation of UCODE, a computer code for universal inverse modeling, U. S. Geological Survey Water-Resources Investigations Report 98-4080. 116 p.

Regli, S., Rose, J.B., Haas, C.N., Gerba, C.P., 1991. Modeling the risk from Giardia and viruses in drinking water. J. AWWA 213, 76–84.

Schijven, J.F., Hoogenboezem, W., Hassanizadeh, S.M., Peters, J.H., 1999. Modeling removal of bacteriophages MS2 and PRD1 by dune recharge at Castricum, Netherlands. Water Resour. Res. 35 (4), 1101–1111.

Scott, D.T., Gooseff, M.N., Bencala, K.E., Runkel, R.L., 2003. Automated calibration of a stream solute transport model: implications for interpretation of biogeochemical parameters. J. North Am. Benthol. Soc. 22 (4), 492–510.

Sun, N-Z, 1994. Inverse Problems in Groundwater Modeling. Kluwer Academic Publishers, The Netherlands.

Tim, U.S., Mostaghimi, S., 1991. Model of predicting virus movement through soils. Ground Water 29 (2), 251–259.

Toride, N., Leij, F.J., van Genuchten, M.Th., 1995. The CXTFIT Code for Estimating Transport Parameters from Laboratory or Field Tracer Experiments, U.S. Salinity Laboratory Research Report No. 137, Riverside, CA.

Weiss, R., Smith, L., 1997. Efficient and responsible use of prior information in inverse methods. Ground Water 36 (1), 151–163.

Weissman, J.B., Craun, G.F., Lawrence, D.N., Pollard, R.A., Saslaw, M.S., Gangarosa, E.J., 1976. An epidemic of gastroenteritis traced to a contaminated public water supply. Am. J. Epidemiol. 103 (4), 391–398.

Yates, M.V., Jury, W.A., 1995. On the use of virus transport modeling for determining regulatory compliance. J. Environ. Qual. 24, 1051–1055.

Yates, M.V., Gerba, C.P., Kelley, L.M., 1985. Virus persistence in groundwater. Appl. Environ. Microbiol. 49 (4), 778–781.

Yeh, W-G., 1986. Review of parameter identification procedures in groundwater hydrology: the inverse problem. Water Resour. Res. 22 (2), 95–108.

Zelikson, R., 1994. Microorganisms and viruses in groundwater. Environ. Sci. Pollut. Control Ser. 11, 425–436.

Zheng, C.M., 1998. A modular three-dimensional multispecies transport (MT3DMS) model, Release DoD_3.00A, Prepared for the Waterways Experiment Station, U.S. Army Corps of Engineers.

Zheng, C., Bennett, G.D., 2002. Applied Contaminant Transport Modeling, 2nd ed. Wiley, New York. 621 pp.