

# Using Atmospheric Tracers to Reduce Uncertainty in Groundwater Recharge Areas

by J. Jeffrey Starn<sup>1,2</sup>, Amvrossios C. Bagtzoglou<sup>2</sup>, and Gary A. Robbins<sup>3</sup>

---

## Abstract

A Monte Carlo-based approach to assess uncertainty in recharge areas shows that incorporation of atmospheric tracer observations (in this case, tritium concentration) and prior information on model parameters leads to more precise predictions of recharge areas. Variance-covariance matrices, from model calibration and calculation of sensitivities, were used to generate parameter sets that account for parameter correlation and uncertainty. Constraining parameter sets to those that met acceptance criteria, which included a standard error criterion, did not appear to bias model results. Although the addition of atmospheric tracer observations and prior information produced similar changes in the extent of predicted recharge areas, prior information had the effect of increasing probabilities within the recharge area to a greater extent than atmospheric tracer observations. Uncertainty in the recharge area propagates into predictions that directly affect water quality, such as land cover in the recharge area associated with a well and the residence time associated with the well. Assessments of well vulnerability that depend on these factors should include an assessment of model parameter uncertainty. A formal simulation of parameter uncertainty can be used to delineate probabilistic recharge areas, and the results can be expressed in ways that can be useful to water-resource managers. Although no one model is the correct model, the results of multiple models can be evaluated in terms of the decision being made and the probability of a given outcome from each model.

---

## Introduction

Agencies at many levels of government have statutes and regulations to protect the quality of drinking water by managing land uses that overlie aquifers. The principle behind this approach is that it is less costly to prevent than to remediate groundwater contamination. Land use and shallow groundwater quality are related (Grady 1994), so it makes sense to manage land use in areas that contribute

water to a pumping well. Areas contributing recharge to a well hereinafter will be referred to as “recharge area” for brevity. Numerical models of groundwater flow are commonly used to delineate recharge areas, and although models often do not lead to accurate predictions (Konikow 1986), they can lead to insight that can be used to protect groundwater.

Many studies have delineated recharge areas using techniques from simple analytical models to complex numerical models. Some of these studies, discussed in the following section, addressed the role of data uncertainty on model predictions such as recharge areas. The present work extends those efforts by demonstrating the effect of different types of information (atmospheric tracers and prior information) on simulated recharge areas. We also consider the effect of uncertainty in recharge area on the estimated percentages of land cover and residence time associated with the recharge area. Sensitivity-based Monte Carlo simulations using a numerical model are

---

<sup>1</sup>Corresponding author: U.S. Geological Survey, 101 Pitkin Street, East Hartford, CT 06108; (860) 291-6746; jjstarn@usgs.gov

<sup>2</sup>Department of Civil and Environmental Engineering, University of Connecticut, 261 Glenbrook Road, UNIT-2037, Storrs, CT 06269-2037; (860) 486-4017; acb@enr.uconn.edu

<sup>3</sup>Department of Natural Resources and the Environment, University of Connecticut, 1376 Storrs Road, Unit 4087, Storrs, CT 06269-4087; (860) 486-2448; gary.robbs@uconn.edu

Received May 2009, accepted December 2009.

Copyright © 2010 The Author(s)

Journal compilation © 2010 National Ground Water Association.

doi: 10.1111/j.1745-6584.2010.00674.x

used in this study to assess the effects of (1) atmospheric tracer observations and (2) prior information on parameter values that are poorly constrained by available data. The atmospheric tracer used in this study was tritium, but the analysis applies to other atmospheric tracers such as sulfur hexafluoride and chlorofluorocarbons.

### Previous Studies

The groundwater-modeling literature describes many approaches for simulation of probabilistic recharge areas. For example, Bair et al. (1991) used a Monte Carlo approach to assess uncertainty in recharge areas to wells. They used assumed parameter distributions (with covariance) in an analytical model, and they did not consider complex cases. Varljen and Shafer (1991) used conditioned random fields of hydraulic conductivity in a Monte Carlo simulation. This approach accounts for geologic heterogeneity but not for parameter uncertainty. Tiedemann and Gorelick (1993) used a Monte Carlo approach as the basis on which to test their first-order uncertainty model. Their work was done in the context of a small-scale plume and was not done at the scale of a recharge area. Cooley (1997) compared four methods of estimating confidence intervals on model predictions of head and flow, one of which was similar to the Monte Carlo approach applied in this article. He found that the Monte Carlo method did not perform well compared to the other methods in all cases, but his model was highly nonlinear and the parameter estimates were not conditioned on a goodness-of-fit criterion. Feyen et al. (2003), among others, used a general likelihood estimator with spatially random fields of hydraulic conductivity. Morse et al. (2003) used a similar technique but without spatially random fields of hydraulic conductivity. The general likelihood estimator requires that a prior distribution of all parameters be estimated. One drawback to this method is that the results are dependent on the choice of prior distribution. Another approach that has been used is the adjoint-state method. Kunstmann and Kastens (2006) used a first-order second-moment simulation, which assumes that the model is linear and that higher order terms in the error model contribute significantly to the overall error. Neupauer and Wilson (2004) and Frind et al. (2006) used the adjoint-state method with a numerical solution to the advection-dispersion equation (ADE) to simulate the backward dispersion of probability. The dispersion term accounts for small-scale uncertainty in aquifer material but does not account for uncertainty in parameter values.

Many of the previously discussed studies assess the effect of random, spatially varying hydraulic conductivity on recharge areas, but they do not address uncertainty caused by estimates of parameter values. Parameter value uncertainty has been addressed in several studies in which the variance-covariance matrix generated during model calibration was used as the basis for Monte Carlo simulation. Starn et al. (2000), Starn (2001), and Hunt et al. (2001) delineated the uncertainty in recharge areas to a well and to a spring, respectively, using a Monte Carlo approach. Similar approaches were used by Lindsey

(2005) to delineate recharge areas to multiple wells, by Starn (2006) to site future monitoring wells near a pumping well, and by Starn and Brown (2007) to relate uncertainty in the recharge area to areas of land use that might affect well water quality.

The Monte Carlo method, as applied here, is different from previous efforts because parameter sets were derived from the variance-covariance matrix calculated from a model sensitivity analysis and because the Monte Carlo runs were subjected to a series of acceptance criteria to ensure that reasonable model runs were obtained. This method accounts for uncertainty in parameter values and as such accounts for the spatial variability of parameter values to the extent that spatial variability is defined in the model, as well as the variability of boundary coefficients that are defined as parameters. This method, like most other methods, assumes that the structure of the model is correct, that is, that the modeler has defined the model geometry and boundaries in accordance with known hydrogeology, and that this knowledge is sufficient to describe the behavior of the system. In reality, the actual hydrogeology is never completely known, and changes in information about the hydrogeology can lead to changes in the model, which are not accounted for here.

Moore and Doherty (2005) assessed the sensitivity of model predictions to calibration data. They show mathematically that it is critical to calibrate a model using the types of data for which predictions will be made; otherwise the calibration data may not contain much information about the parameters that govern the processes that produce the predictions of interest. For example, Sheets et al. (1998) used atmospheric tracers to improve the calibration of the groundwater flow model. In their study, the information provided by atmospheric tracer observations led to changes that improved model fit. For this reason, atmospheric tracer observations are used in this study to predict probabilistic recharge areas.

### Description of the Calibrated Model

The methods described in this article were applied to an existing calibrated model of groundwater flow in the aquifer system near Woodbury, Connecticut (Starn and Brown 2007). Details describing the geologic setting, model development, and model calibration are available at <http://pubs.usgs.gov/sir/2007/5210/> and are summarized in Table 1. The model was developed to simulate flow paths from the recharge area of a pumping well to the well as part of the transport of anthropogenic and natural compounds topical study within the USGS National Water Quality Assessment. The model simulates groundwater flow in part of a small New England watershed. The principal aquifer consists of valley-fill glacial sand and gravel deposits. The simulation also includes groundwater flow in uplands adjacent to the valley through glacial till and (or) fractured rock into the valley-fill aquifer. Water in the glacial aquifer discharges to surface streams and to one pumping well in the modeled area. The flow system is characterized by extreme heterogeneity on a large scale

**Table 1**  
**Model Grid and Calibration Data**

Active Modeled Area (km <sup>2</sup> )	15
Pumped Well Discharge (m <sup>3</sup> /d)	392
Groundwater Inflow (m <sup>3</sup> /d)	16,604
Grid Cell Row Dimension (m)	15
Grid Cell Column Dimension (m)	15
Grid Thickness (m)	20
Number of Layers	7
Number of Rows	241
Number of Columns	322
Number of Head Observations	97
Number of Streamflow Observations	1
Number of Atmospheric Tracer Observations	14
Number of Parameters, Excluding Porosity	11
Number of Porosity Parameters	4

(between crystalline rock and sand and gravel), but within each of the hydrogeologic domains, small-scale heterogeneity is not simulated. The model uses 11 hydraulic conductivity parameters, listed with their parameter abbreviations in square brackets: the horizontal hydraulic conductivity of two facies within the glacial stratified deposits (fluvial [Kfluv] and deltaic [Kdelt]), two tills (compact till [Ktcom] and noncompact till [Ktnon]), two types of bedrock (Paleozoic crystalline [Kxln] and Mesozoic sedimentary [Kmeso]), fine-grained glaciolacustrine deposits (Kfine), and streambed vertical hydraulic conductivity (Ktrib); and the vertical hydraulic conductivity of combined glacial stratified deposits (Kvgsd), combined till (Kvtill), and combined bedrock (Kvrock). Recharge was estimated independently using a calibrated rainfall-runoff model (D.M. Bjerklie, USGS, written communication, 2007) and was not adjusted in model calibration. Four additional porosity parameters were used in the advective particle tracking model: glacial stratified deposits (Pgds), noncompact till (Ptill), compact till (Ptt), and bedrock (Prock).

The model originally was calibrated to hydraulic head and streamflow observations using nonlinear regression (Cooley and Naff 1990; Aster et al. 2005; Hill and Tiedeman 2007). Some advantages of nonlinear regression over other calibration methods are that it produces optimal parameter estimates and the variance-covariance matrix of those estimates. Groundwater residence times were simulated using particle tracking (Pollock 1994), and porosity was adjusted manually to match tritium-helium apparent ages (Starn and Brown 2007). Nonlinear regression was not used to calibrate the model to atmospheric tracer observations. All parameters were transformed for the regression by taking their base 10 logarithm. The model was sensitive to 6 of the 11 model parameters and their values were estimated in the regression (Ktcom, Kfluv, Kdelt, Kmeso, Kxln, and Ktrib); however, the model was not sensitive to the remaining five parameters, and their values were fixed at values cited in the literature (Starn and Brown 2007).

## Methodology

The effects of adding a set of atmospheric tracer data and prior information on uncertainty in simulated recharge areas were assessed by running a suite of Monte Carlo simulations that either included or omitted the full set of tracer data and (or) prior information. Particle tracking was used in each parameter set in the Monte Carlo simulation to identify the recharge area and the associated residence-time distribution. The effect of parameter uncertainty on predictions of the recharge area is shown on maps of the probabilistic recharge area and on graphs showing the probability distributions of land cover and residence-time distribution in the recharge area associated with the well.

In this study, the original model was re-run using MODFLOW-2005 (Harbaugh 2005), and advective particle tracking was done using MODPATH (Pollock 1994). Sensitivity calculations and variance-covariance matrices were computed using UCODE (Poeter et al. 2005). MODPATH simulates advective transport based on groundwater velocities interpolated from cell faces between finite-difference cells. This method does not account for diffusion or dispersion caused by small-scale advection. Dispersion causes smearing of the residence-time distribution, but the effect of this is beyond the scope of this article and will be addressed in future studies. This method also does not account for transport across the unsaturated zone, which should be minimal at this site because of the shallow depth to water.

### Simulated Equivalents of Atmospheric Tracers

To use atmospheric tracer observations in the model, simulated equivalents to the observations needed to be constructed from the model output. The atmospheric tracer observations consisted of tritium concentrations measured in samples collected in 2003 and 2004 from 15 monitoring wells near the pumping well (Brown et al. 2009). Monitoring well screen lengths varied but were on the order of 3 m, and the wells were minimally pumped prior to sampling, so that the samples were considered to represent the resident concentration in the well. Flow paths to monitoring well screens were simulated by reverse tracking 10 evenly spaced particles from the vertical interval of the screen. The arithmetic mean of the 10 particle travel times was used to generate the simulated equivalent concentration for comparison to atmospheric tracer observations. The date of recharge was calculated by first subtracting the mean advective travel time from the date of sampling and then assigning a concentration to the well equal to the concentration in equilibrium with the atmosphere for the computed recharge date. Atmospheric inputs for tritium were estimated using a program that calculates tritium concentrations for specified latitude-longitudes by extrapolating results from stations where tritium in precipitation has been measured (Michel 1989). The simulated tritium concentration was decayed using the half-life of tritium. An arithmetic mean, converted to a concentration, may be valid for particle travel times with a small amount of variability, but the input

function of atmospheric tracers generally is nonlinear. The greater the variability of travel times, the less accurate this method would be. In the model, the average range of simulated residence times over the length of monitoring well screens was about 0.8 years (Starn and Brown 2007). Because of the small range, the arithmetic mean was considered valid.

### Use of Prior Information

Although a variance-covariance matrix can be calculated that includes poorly constrained parameters, the variances and covariances that result may be unrealistically large. In order to “steer” the calculation toward realistic values, prior information can be applied to model parameters. As described by Hill and Tiedeman (2007), “prior information” is a form of regularization wherein the weighted difference between the parameter value and the prior information value is added to the sum of squared weighted residuals objective function. To simulate the effect on the simulated recharge area of parameter values that are poorly constrained by available data, prior information was applied to those parameters. Prior information estimates can be obtained by collecting field measurements; however, in this study, the optimal values from the original model were used, which were based on literature values (Starn and Brown 2007).

Confidence in prior information values are reflected in weights assigned to them by the modeler. In UCODE, weights are the inverse of the measurement variance of the observation and are calculated based on statistics (standard deviation, variance, or coefficient of variation) input by the user (Poeter et al. 2005). In this study, the coefficient of variation (standard deviation of the parameter estimate divided by its mean value) was used because the statistics are then independent of the parameter value. The weight is calculated in UCODE as a function of the prior information value and the coefficient of variation, and the same coefficient of variation was used to calculate weights for all prior information within a given Monte Carlo simulation. High coefficients of variation indicate parameter values that are not well-known and hence produce low weights. The coefficients of variation tested in this study were 1.0 (low confidence) and 0.0001 (high confidence), and the simulations produced using them are identified by the terms “prior information” and “reduced parameter set,” respectively. A coefficient of variation equal to 0.0001 essentially fixed the parameter value at its optimal value, thus in effect reducing the number of parameters that can vary in the sensitivity analysis. Prior information was assigned to the parameters that were not estimated using nonlinear regression in the original model because the model was relatively insensitive to them (Starn and Brown 2007). Porosity was not included as an estimated parameter in the original model, and prior information was included on two of the four porosity parameters because of their relatively low sensitivity.

### Calculating the Variance-Covariance Matrices

Variance-covariance matrices were calculated as a function of the sensitivity to changes in parameter values of the original calibration data combined with new information consisting of atmospheric tracer concentrations and prior information values. Variance-covariance matrices were calculated with UCODE (Poeter et al. 2005), which uses a perturbation method wherein sensitivities are approximated by discrete differences in model parameters. Sensitivities were calculated for each simulation at the optimal parameter values determined in the original model calibration. The sensitivity matrix (also called the Jacobian matrix) was calculated by assembling the individual sensitivities of the system as follows

$$J = \begin{vmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{vmatrix} \quad (1)$$

where  $J$  is the sensitivity, or Jacobian, matrix;  $X_{ij} = \partial y_i / \partial x_j$ ;  $y_i$  is the  $i$ th observation of the system; and  $x_j$  is the  $j$ th parameter of the system.

In simulations where prior information is included, the sensitivity matrix is augmented such that the sensitivities to the prior information are included as additional rows at the bottom of the Jacobian matrix. The variance-covariance matrix was calculated using the sensitivity matrix and the weights assigned to parameter values and prior information as

$$V = s^2 (J^T \omega J)^{-1} \quad (2)$$

where  $V$  is the variance-covariance matrix;  $s^2$  is the calculated error variance of the regression (the expected value of the weighted sum of squared differences between observed and simulated values);  $J$  is the Jacobian matrix (superscript T indicates transpose); and  $\omega$  is the observation weight matrix.

From the preceding equation, it is seen that the variance-covariance matrix is a function of the sensitivities  $X_{ij}$ ; therefore, adding atmospheric tracer observations and prior information to the sensitivity calculations alters the variance-covariance matrix, even without recalibrating the model. Parameter variances are the diagonal elements of this matrix, and the off-diagonal elements are the covariances between parameters. The parameter standard deviations discussed herein are the square roots of the diagonals of the variance-covariance matrix.

### Generating Parameter Realizations

The next step in the analysis was to generate a large number (10,000 in this study) of parameter sets, the ensemble of which has the same mean parameter values as the optimal (calibrated) values and the same parameter variance-covariance matrix as the calibrated model. Parameter sets were generated by taking the Cholesky decomposition of the variance-covariance matrix and multiplying the results by randomly generated numbers from a normal distribution. True parameter values probably are

not from a normal distribution (Cooley 1997), and this remains an unexplored source of error. Restricting the parameter sets to those that reproduce the observation data to within a specified tolerance mitigates this source of error to some extent. Parameter sets were generated using the definition of a standard normal variable:

$$R = V^{1/2}N + B \quad (3)$$

where  $R$  is a set of parameter values;  $V^{1/2}$  is the matrix square root of the variance-covariance matrix, calculated using Cholesky decomposition;  $N$  is a vector of normally distributed random numbers with a mean of 0 and standard deviation of 1; and  $B$  is the vector of the optimal parameter values.

In order to ensure that parameter distributions were being sampled completely and efficiently, a stratified random sampling method known as the Latin hypercube sampling (LHS) was used (Hossain et al. 2006). To implement LHS, parameter distributions were divided into 20 equally probable regions, which were then sampled randomly. A rank correlation process was used to ensure that the random parameter values approximate the variance-covariance matrix from calibrated model (Iman and Conover 1982). The Monte Carlo simulation used parameter sets in groups of 20, so that each region of the input distribution was sampled before the closure criterion was deemed met.

In this study, 10,000 sets of parameters were initially generated and stored because: (1) it was computationally inexpensive to generate this number of sets, (2) the normal distributions of the parameters were accurately represented, (3) specific parameter sets could be re-run, and (4) it was unknown at the outset how many parameters sets would be needed. Only a small number of the 10,000 parameter sets needed to be run before the probability completion criterion was met.

### Running the Monte Carlo Simulations

Probabilistic recharge areas were delineated using a suite of Monte Carlo simulations comprising six variations of tracer data and prior information. Three of the six variations included tracer observations and three omitted them. Within each group of three, one simulation was run using the full set of parameters, one was run using prior information, and one was run using a reduced parameter set. Monte Carlo simulation involves running each variation for a large number of parameter sets and expressing the ensemble of the results in probabilistic terms. For each model run, one particle was placed on the top face of the top model layer (representing land surface) and tracked forward to its discharge point. The probability that a given particle started in the recharge area to the well was the number of model runs in which the particle reached the well divided by the total number of model runs. The Monte Carlo simulation was considered complete when the maximum change in probability for any particle from the previous run was less than 0.01 (1%).

A given set of parameter values is not guaranteed to produce a good model fit, so acceptance criteria were used to ensure that the model runs were reasonable. These criteria were that the model had to: (1) converge, (2) have a mass balance error of less than 0.05 (5%), (3) have a standard error less than 20, and (4) have at least one particle that ended in the cell that simulated the pumping well. To save simulation time, the criteria were applied sequentially, so that particle tracking was not done for runs with an unacceptable mass balance error. The last two criteria warrant further explanation. The standard error criterion of 20 was based on a subjective assessment of the calculated standard error from the original model. The standard error (dimensionless; the square root of the calculated error variance) depends on the observation weights. If the observation weights reflected only measurement error, the standard error should equal 1.0, and deviations from 1.0 indicate possible model error (Hill and Tiedeman 2007). The standard error was 6.2 in the model, indicating that model error was present. The standard error criterion was chosen arbitrarily to be larger than standard error of the model to accept models that did not fit the observed data well while rejecting those with particularly poor fits, and more work could be done to explore this criterion.

The criterion relating to particles ending in the pumping well cell was necessary because the model cell containing the well became a weak sink for some parameter realizations. If the flow rate of water into the cell equals the discharge of the well (in a steady-state model), the cell is a strong sink and all particles that enter the cell stop at the well. The model was designed such that the well cell was a strong sink. However, in the Monte Carlo simulations, a small number of parameter sets (Table 2) caused flow into the cell to be greater than the discharge of the well (a weak sink). MODPATH requires the user to specify what happens to particles in a weak sink, and in this study, particles were set to pass through weak sinks. This criterion eliminates parameter sets that create a weak sink at the well.

The general steps in the Monte Carlo simulations employed in this study are summarized as follows.

1. Calculate the sensitivities of each observation, including the new atmospheric tracer observations and (or) prior information, using UCODE.
2. Calculate the variance-covariance matrix from the sensitivities using UCODE.
3. Generate parameter realizations using LHS and the method of Iman and Conover (1982).
4. Substitute a parameter realization into MODFLOW-2005 input files and run the model.
5. If the model converged and the mass balance error was less than 5% go to the next step, otherwise go back to step 4.
6. If tracer observations were included, substitute a porosity realization into MODPATH input files.
7. Run MODPATH.
8. Calculate the simulated equivalents for the observations and calculate the standard error using UCODE.

**Table 2**  
**Summary of Monte Carlo Simulations**

	Atmospheric Tracer Omitted			Atmospheric Tracer Included		
	Full Parameter Set	Prior Information	Reduced Parameter Set	Full Parameter Set	Prior Information	Reduced Parameter Set
	<b>Total Number of Monte Carlo Runs</b>					
	260	240	120	200	180	120
<b>Sequentially Applied Acceptance Criterion</b>	<b>Percentage of Total Monte Carlo Runs Meeting Criterion</b>					
1. MODFLOW-2005 convergence	77	90	100	98	99	100
2. Mass balance error less than 5%	73	66	88	74	76	87
3. Standard error less than 20	40	40	88	53	62	87
4. No weak sinks	39	39	88	52	61	87
	<b>Number of Monte Carlo Runs Meeting All Criteria</b>					
	101	94	105	104	110	104

9. If the standard error was less than 20, and the well cell was not a weak sink, go to the next step, otherwise go back to step 4.
10. Save the residence time of each particle that reached the well.
11. Calculate the probability that each particle reached the well.
12. If the maximum difference in the probability for all particles was less than 0.01, go to the next step, otherwise go back to step 4.
13. Summarize the probability, land cover, and residence time for the recharge area.

#### **Effect of Recharge Area Uncertainty on Land Cover and Residence-Time Distribution**

Uncertainty in model parameters translates into uncertainty in the recharge area of a well, which in turn translates into uncertainty on estimates of the types of land cover in the recharge area and the flow-weighted mean residence time at the well. To calculate the flow-weighted mean residence time, the fraction of flow supplied by each particle was calculated by dividing the recharge rate associated with the particle by the sum of the recharge rates for all the particles that enter the well. The fractional flow-weighted residence time was calculated by multiplying the fractional flow rate of each particle by its residence time (the “age mass” described by Goode 1996). The residence-time distribution of water at the well was given by the set of flow-weighted fractional ages on all particles, and the mean residence time was the sum of the flow-weighted fractional ages. The mean residence time was used in this study to show the effect of parameter uncertainty on the residence time from particles starting in a particular land use; however, the mean residence time obscures the fact that water from a pumping well is a mixture of water with a potentially wide range of residence times.

#### **Results of Simulations**

All Monte Carlo simulations met the probability closure criterion in not more than 260 runs (Table 2). In each of the simulations, about 100 runs met all of the acceptance criteria, as can be seen on the last row of Table 2. Simulations with the reduced parameter set met the acceptance criteria most of the time whether tracer observations were included (87%) or omitted (88%). For the simulations with prior information and with full parameter sets, including tracer observations was advantageous in the sense that a higher percentage of runs met all acceptance criteria. The most restrictive acceptance criterion was the standard error criterion for the full parameter set and prior information simulations when tracer observations were omitted. The weak sink criterion had little effect on the number of accepted runs.

#### **Effect of Atmospheric Tracer and Prior Information on Model Predictions**

The addition of atmospheric tracer observations and prior information reduced the standard deviation of each parameter relative to the “tracer omitted/full parameter” simulation (Table 3). For example, in the “tracer included/full parameter simulation,” the standard deviation decreased between 9% (parameter Ktrib) and 67% (parameter Ktcom). For the same simulation, the standard deviation of the five parameters (last five rows of Table 3) that were poorly constrained by the model also decreased (by 24 to 62%). Decreasing the standard deviation of these parameters generally is the goal of adding prior information; however, adding prior information on poorly constrained parameters also reduced the standard deviations of other parameters, and a similar effect was achieved by including tracer observations.

**Table 3**  
**Reduction in Parameter Standard Deviation**

Parameter	Atmospheric Tracer Omitted			Atmospheric Tracer Included		
	Full Parameter Set	Prior Information	Reduced Parameter Set	Full Parameter Set	Prior Information	Reduced Parameter Set
	Parameter Standard Deviation	Percentage Reduction in Standard Deviation Relative to the Simulation with Atmospheric Tracer Data Omitted and a Full Parameter Set				
Kdelt	0.160	13	25	13	19	52
Kfluv	0.071	3	28	35	51	55
Ktrib	0.220	5	18	9	14	80
Ktcom	1.900	16	87	67	74	91
Kmeso	0.100	4	58	19	36	58
Kxln	0.120	8	47	17	32	53
Ktnon	0.470	9 <sup>1</sup>	100 <sup>1</sup>	23	36 <sup>1</sup>	100 <sup>1</sup>
Kvtill	2.100	14 <sup>1</sup>	100 <sup>1</sup>	60	63 <sup>1</sup>	100 <sup>1</sup>
Kfine	0.800	6 <sup>1</sup>	99 <sup>1</sup>	24	26 <sup>1</sup>	99 <sup>1</sup>
Kvgsd	7.700	39 <sup>1</sup>	100 <sup>1</sup>	62	69 <sup>1</sup>	100 <sup>1</sup>
Kvrock	0.420	10 <sup>1</sup>	100 <sup>1</sup>	40	43 <sup>1</sup>	100 <sup>1</sup>

<sup>1</sup>Indicates a parameter to which prior information was applied.

Correlation between parameters leads to model nonuniqueness and can cause problems in model calibration; however, correlation coefficients between pairs of parameters less than 0.85 often are not significant. Two pairs of parameters had correlation coefficients greater than 0.85—Ktcom/Kvtill and Prock/Ptt. Adding tracer observations decreased correlation between Ktcom/Kvtill (Prock/Ptt were porosity parameters that were not used when tracer observations were omitted) (Table 4). For the reduced parameter sets, all correlation coefficients were less than 0.85.

One concern with using acceptance criteria was that by eliminating parameter sets that produced model runs not meeting the criteria, the original parameter distributions would not be maintained. For each Monte Carlo simulation, plots were made of parameter distributions using all 10,000 parameter sets and only those parameter sets included in the accepted Monte Carlo runs. An example for the Kxln parameter is shown on Figure 1. Some deviations from the original distributions were expected because of the lower number of parameters, but in general, the parameter distributions for 10,000 parameter sets (Figure 1B) are similar to the parameter distributions derived using only model runs that met the acceptance criteria (Figure 1A). This result indicated that bias caused by the acceptance criteria was minimal.

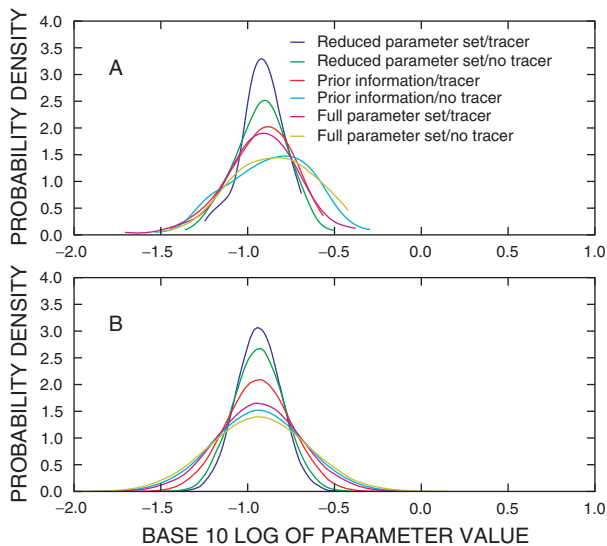
The effects of better (narrower) parameter distributions are manifested in the shape and size of the probabilistic recharge areas (Figures 2A to 2D). The largest recharge area was produced with the simulation with the full parameter set and no atmospheric tracer observations (Figure 2A). Adding either atmospheric tracer observations or prior information reduced the overall size of the recharge area similarly (Figures 2B and 2C); however, with the reduced parameter set, the probabilities

**Table 4**  
**Parameter Correlation Coefficients**

Parameter Pair	Correlation Coefficients, Shown Where Greater Than 0.85	
	Atmospheric Tracer Omitted	Atmospheric Tracer Included
Full parameter set		
Ktcom-Kvtill	0.98	0.87
Prock-Ptt	—	0.90
Prior information		
Ktcom-Kvtill	0.98	0.89
Reduced parameter set		
No parameter correlations greater than 0.85 were found for the reduced parameter sets.		

were higher in the center of the recharge area than with atmospheric tracer observations alone. Adding both prior information and atmospheric tracer observations led to the smallest recharge area having the highest probabilities (Figure 2D).

Estimates of the area of a given land cover in the recharge area depended on the spatial distribution of that land cover (shown in Starn and Brown 2007) as well as the choice of whether to include tracer observations or prior information. Uncertainty in the area of a given land cover could be important if, for example, well vulnerability were assessed based on the type of land cover in the recharge area. For example, simulations that included



**Figure 1. Parameter distributions for model parameter  $K_{xln}$  for: (A) Monte Carlo runs meeting acceptance criteria and (B) 10,000 parameter sets.**

prior information showed that the most probable area of developed land cover in the recharge area was zero, but the simulation with the full parameter set showed that the most probable area was 16 hectares (Figure 3). In both cases, adding tracer observations did not change the result significantly.

Uncertainty in the size and shape of the recharge area affected the distributions of flow-weighted mean residence time of water that enters the well (Figure 4). The simulations that included tracer observations also included porosity parameters, and the additional uncertainty in porosity led to more uncertainty in the mean residence time. This additional uncertainty is “real” because estimates of porosity are uncertain. Not including the effects of this uncertainty underestimates the overall uncertainty in residence time. The distribution of mean residence times, when considering only the part of the recharge area overlain by developed land cover, is less affected by porosity uncertainty. This is because the developed areas are in less uncertain parts of the overall recharge area. This could be of interest, for example, if the residence time at the well, in addition to the land cover in the recharge area, is considered to be a factor in well vulnerability.

## Discussion

Reductions in standard deviations on individual parameters seemed to suggest that atmospheric tracer observations had at least as strong an influence on parameter value uncertainty as prior information because adding atmospheric tracer observations and prior information individually reduced the size of the recharge areas about the same. The probabilities within the recharge area were higher in the simulation with prior information, suggesting that the model prediction of recharge area is sensitive to model parameters that are not well constrained by the available data. In this study, the model was not

recalibrated, and different results might be obtained if it had been. In particular, the calculated error variance would be expected to become smaller, although with models that are well calibrated, the reduction might be small.

Recharge areas are the basis for calculating many important factors that contribute to well vulnerability. It makes intuitive sense that using direct observations of water quality (in this study, atmospheric tracer observations) in model calibration will lead to better predictions of recharge areas. Atmospheric tracer observations contain information about the length and position of groundwater flow paths, flow rates through the aquifer, and residence-time distributions. The inclusion of atmospheric tracer observations, prior information, and the choice of weights can be a somewhat arbitrary undertaking. Although both types of information can reduce variance in model predictions, the effects need to be considered carefully.

Well vulnerability is, in part, a function of recharge rate, residence time, land cover, and uncertainty in the simulated recharge area. Combinations of this information can be shown graphically. A subset of the probabilistic recharge area shown in Figure 2D, for example, shows the probability that water with a residence time of 5 years or less reaches the well (Figure 5). This probability was the number of times, out of all the runs in a Monte Carlo simulation, those particles whose residence time was 5 years or less ended in the pumped well, divided by the total number of accepted runs. Five years was selected for this example as a minimal time frame for responding to aquifer contamination. This type of map may be a useful tool for prioritizing areas within an aquifer for protection and for placement of monitoring wells.

Although no one model is the “correct” model, the results of multiple models can be evaluated in terms of the decision being made and the probability of a given outcome from each model. In this example (Figure 6), the results of two models, one with atmospheric tracer observations and one without atmospheric tracer observations, are evaluated in terms of the percentage of water produced from the well. In both models, there is about a 20% probability that 100% of the water is young (arbitrarily defined as less than 5 years old for this example). At 40% probability, the percentage of young water predicted by the two simulations is the same (about 83%). There is a 100% probability that the percentage of young water was 30% and 60% for the simulations with atmospheric tracer observations and without atmospheric tracer observations, respectively. If different models agree that there is a high probability that young water is a large percentage of water from the well, then the well might be deemed vulnerable. If it is critical to know what percentage of well water is young, but the probabilities are low, there are several remedies, including improving the model structure and (or) using more and different data to calibrate the model. The result of both these actions is that parameter values could be estimated with more certainty (Figure 6).

One of the perceived drawbacks of the Monte Carlo method is that it can be computationally intensive; however, the number of runs needed depends on the prediction



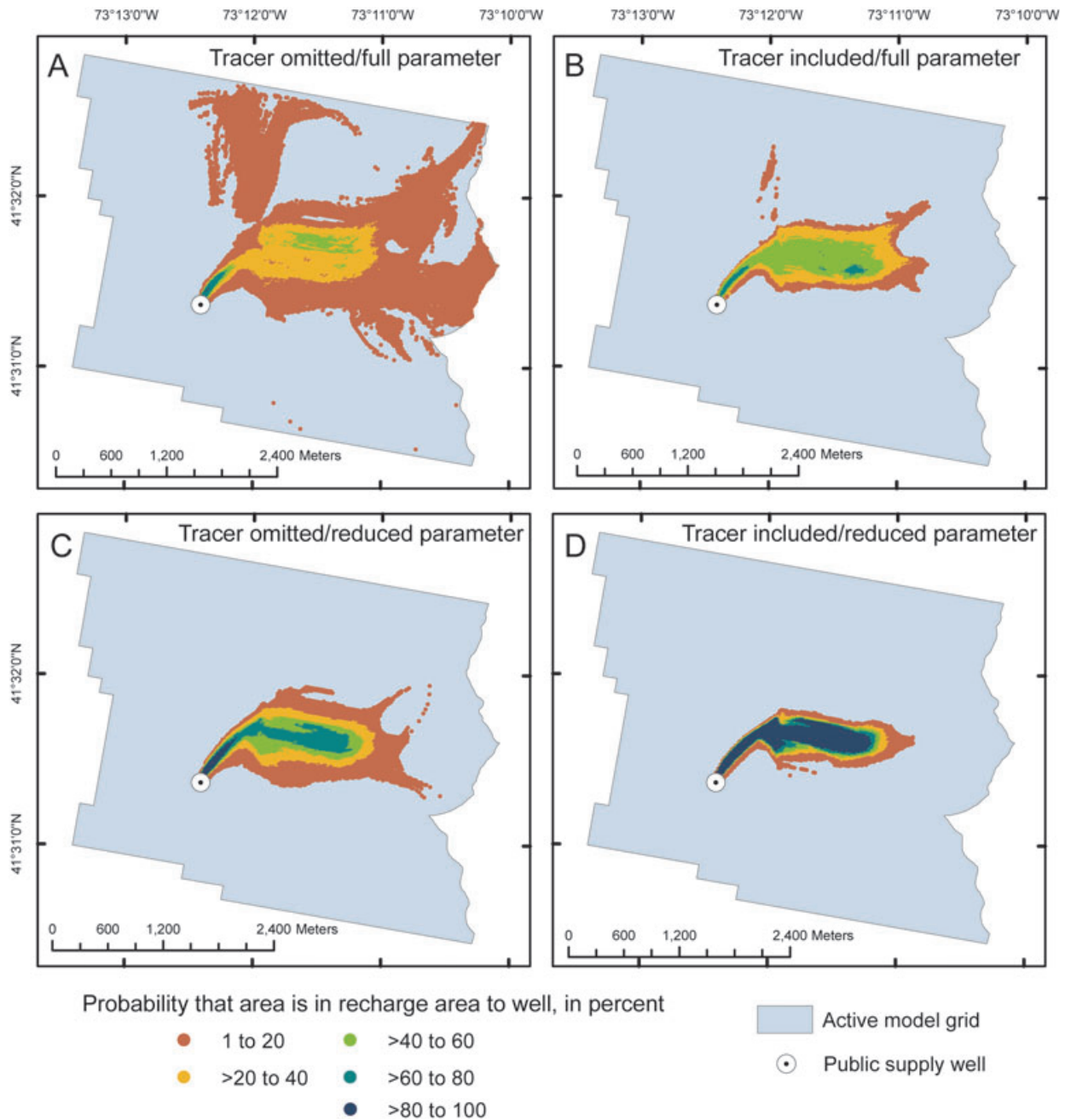


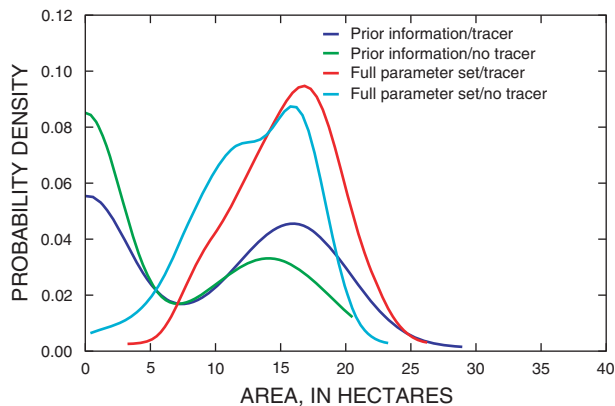
Figure 2. Probabilistic contributing recharge areas from Monte Carlo simulations.

of interest. If reasonable predictions can be made with a realistic number of runs, the technique is viable compared to first-order or ADE-based methods. Monte Carlo simulation has the additional benefits of being easy to understand, applicable to a wide variety of flow conditions and models, and is trivially parallelizable to reduce overall run times.

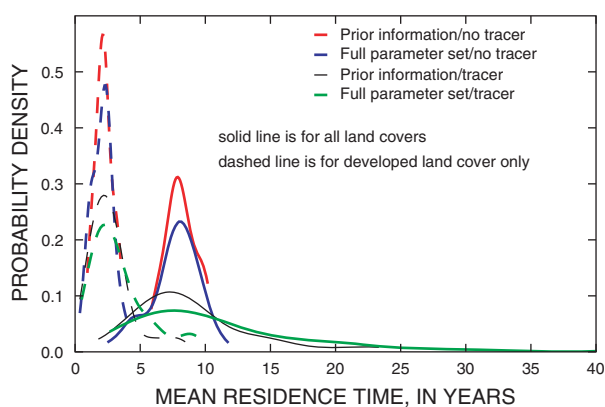
## Conclusions

The addition of information reduces the standard deviation and correlation within and among parameter distributions. In particular, using atmospheric tracer observations and prior information on model parameters improves

model predictions of recharge areas. Atmospheric tracer observations are an effective means to reduce standard deviation of parameter estimates, similar to the reductions achieved by adding prior information on parameter estimates. Additional information also reduces correlation among parameters. The use of acceptance criteria in the Monte Carlo simulations generally did not introduce bias in mean parameter values. Although the addition of atmospheric tracer observations and prior information had similar changes in the extent of predicted recharge areas, prior information had the effect of increasing probabilities within the recharge area to a greater extent than atmospheric tracer observations. This could be because the



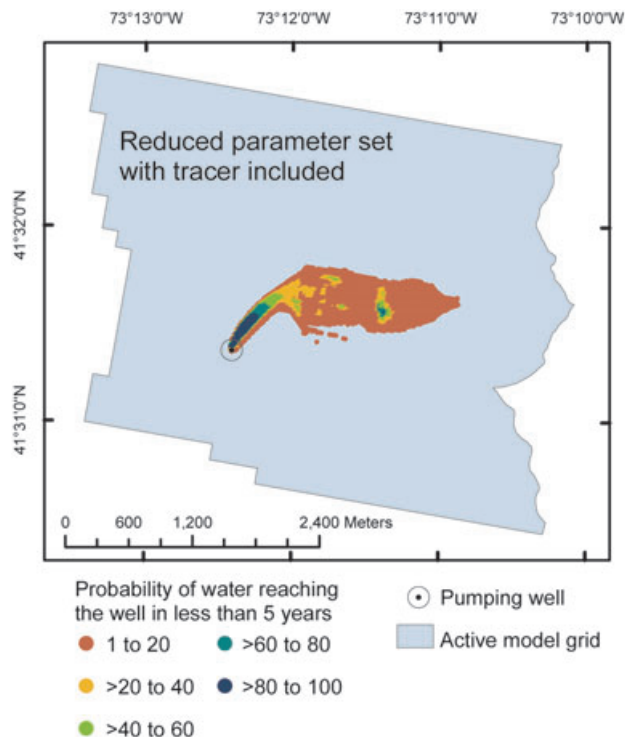
**Figure 3. Probability distributions of areas of developed land cover in the contributing recharge area from Monte Carlo simulations.**



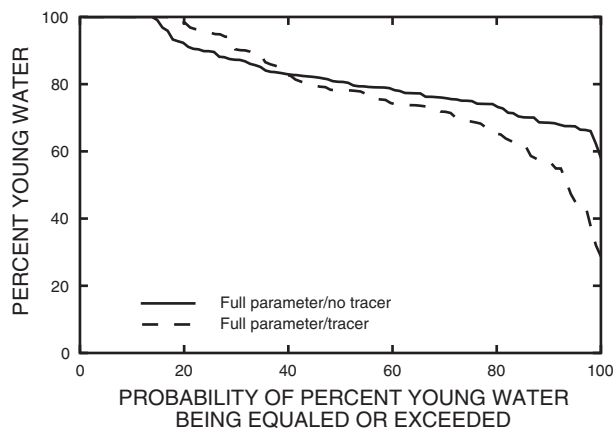
**Figure 4. Mean residence-time distributions in the recharge area from Monte Carlo simulations for the developed land cover and for all land covers.**

recharge area was sensitive to parameters whose distributions were not improved by the addition of atmospheric tracer observations. Prior information has to be applied carefully. Excluding insensitive parameters from the analysis by fixing their values is akin to placing high (and possibly unwarranted) weights on their estimates.

Uncertainty in the recharge area propagates into predictions that directly affect water quality predictions, such as land cover in the recharge area associated with a well and the residence time at the well. Assessments of well vulnerability that depend on these factors should include an assessment of model parameter uncertainty. A formal simulation of parameter uncertainty can be used to delineate probabilistic recharge areas, and the results can be expressed in ways that might be useful to water-resource managers. Examples include examining the probability that a given land cover is in the recharge area, mapping the probability that a given area within the recharge contributes water of a given residence time, and combining probability, residence time, and percentage of well discharge in graphs that depict the cumulative probability that water of a given age makes up a given percentage of the well discharge. Although no one model is



**Figure 5. Probability that residence time is less than 5 years for individual particles that were started at the land surface and tracked to the well.**



**Figure 6. Probability that water with a residence time of less than 5 years ("young water") provides a given percent of water pumped from the well.**

the correct model, the results of multiple models can be evaluated in terms of the decision being made and the probability of a given outcome from each model.

## Acknowledgments

The authors would like to thank their colleagues in the USGS NAWQA program for contributing their ideas and assistance, especially Sandra Eberts and Leon Kaufmann. Don Walter (USGS), Bruce Lindsey (USGS), and two anonymous reviewers provided helpful reviews of the manuscript.

## References

- Aster, R.C., B. Borchers, and C.H. Thurber. 2005. *Parameter estimation and inverse problems*. Burlington, Massachusetts: Elsevier Academic Press.
- Bair, E.S., C.M. Safreed, and E.A. Stasny. 1991. A Monte Carlo-based approach for determining travel time-related capture zones of wells using convex hulls as confidence regions. *Ground Water* 29, no. 6: 849–855.
- Brown, C.J., J.J. Starn, K. Stollenwerk, R.A. Mondazzi, and T.J. Trombley. 2009. Aquifer chemistry and transport processes in the zone of contribution to a public-supply well in Woodbury, Connecticut, 2002–06. US Geological Survey Scientific Investigations Report 2009–5051. Reston, Virginia: USGS.
- Cooley, R.L. 1997. Confidence intervals for ground-water models using linearization, likelihood, and bootstrap methods. *Ground Water* 35, no. 5: 869–880.
- Cooley, R.L., and R.L. Naff. 1990. *Regression Modeling of Ground-water Flow*, Book 3, Chap B4. US Geological Survey Techniques of Water-Resource Investigations, Reston, Virginia: USGS.
- Feyen, L., P.J. Ribeiro, J.J. Gomez-Hernandez, K.J. Beven, and F. De Smedt. 2003. Bayesian methodology for stochastic capture zone delineation incorporating transmissivity measurements and hydraulic head observations. *Journal of Hydrology* 271, no. 1–4: 156–170.
- Frind, E.O., J.W. Molson, and D.L. Rudolph. 2006. Well vulnerability: a quantitative approach for source water protection. *Ground Water* 44, no. 5: 732–742.
- Goode, D.J. 1996. Direct simulation of groundwater age. *Water Resources Research* 32, no. 2: 289–296.
- Grady, S.J. 1994. Effects of land use on quality of water in stratified-drift aquifers in Connecticut—Chapter B, Analysis of nonpoint-source ground-water contamination in relation to land use. U.S. Geological Survey Water-Supply Paper 2381.
- Harbaugh, A.W. 2005. MODFLOW-2005, the U.S. Geological Survey modular ground-water model—the Ground-water Flow Process. U.S. Geological Survey Techniques and Methods 6-A16.
- Hill, M.C., and C.R. Tiedeman. 2007. *Effective Groundwater Model Calibration with Analysis of Data, Sensitivities, and Uncertainty*. Hoboken, New Jersey: John Wiley and Sons.
- Hunt, R.J., J.J. Steuer, T.C. Mansor, and T.D. Bullen. 2001. Delineating a recharge area for a spring using numerical modeling, Monte Carlo techniques, and geochemical investigation. *Ground Water* 39, no. 5: 702–712.
- Hossain, F., E.N. Anagnostou, and A.C. Bagtzoglou. 2006. On Latin Hypercube Sampling for efficient uncertainty estimation of satellite rainfall observations in flood prediction. *Computers and Geosciences* 32, no. 6: 776–792.
- Iman, R.L., and W.J. Conover. 1982. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics—Simulation and Computation* 11, no. 3: 311–324.
- Konikow, L.F. 1986. Predictive accuracy of a ground-water model—lessons from a post audit. *Ground Water* 24, no. 2: 173–184.
- Kunstmann, H., and Kastens, M. 2006. Determination of stochastic well head protection zones by direct propagation of uncertainties of particle tracks. *Journal of Hydrology* 323, no. 1–4: 215–229.
- Lindsey, B.D. 2005. Hydrogeology and simulation of source areas of water to production wells in a colluvium-mantled carbonate-bedrock aquifer near Shippensburg, Cumberland and Franklin Counties, Pennsylvania. U.S. Geological Survey Scientific Investigations Report 2005–5195.
- Michel, R.L., 1989. Tritium deposition in the continental United States, 1953–83. U.S. Geological Survey Water-Resources Investigation Report 89–4072. Reston, Virginia: USGS.
- Moore, C., and Doherty, J. 2005. Role of the calibration process in reducing model predictive error. *Water Resources Research* 41, W05020. DOI: 10.1029/2004WR003501
- Morse, B.S., G. Pohll, J. Huntington, and R.R. Castillo. 2003. Stochastic capture zone analysis of an arsenic-contaminated well using the generalized likelihood uncertainty estimator (GLUE) methodology. *Water Resources Research* 39, no. 6: HWC21–HWC29.
- Neupauer, R.M. and J.L. Wilson. 2004. Numerical implementation of a backward probabilistic model of groundwater contamination. *Ground Water* 42, no. 2: 179–189.
- Poeter, E.P., M.C. Hill, E.R. Banta, S. Mehl, and S. Christensen. 2005. UCODE\_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation. U.S. Geological Survey Techniques and Methods 6-A11.
- Pollock, D.W. 1994. User's guide for MODPATH/MODPATH-PLOT, Version 3—a particle tracking post-processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model. U.S. Geological Survey Open-File Report 94–464. Reston, Virginia: USGS.
- Sheets, R.A., E.S. Bair, and G.L. Rowe. 1998. Use of  $^3\text{H}/^3\text{He}$  ages to evaluate and improve ground-water flow models in a complex buried-valley aquifer. *Water Resources Research* 34, no. 5: 1077–1089.
- Starn, J.J. 2006. Evaluating uncertainty in areas contributing recharge to wells for water-quality network design [abs.]. In *National Water Quality Meeting Proceedings, San Jose, California, May 7–11, 2006*.
- Starn, J.J. 2001. Analysis of uncertainty of contributing areas to wells using parameter estimation and Monte Carlo simulation. MODFLOW 2001 and Other Modeling Odysseys, September 11–14, 2001, Golden, Colorado. In *Proceedings of International Groundwater Modeling Center, Colorado School of Mines, and the U.S. Geological Survey* 1, 305–311.
- Starn, J.J., J.R. Stone, and J.R. Mullaney. 2000. Delineation and analysis of uncertainty of contributing areas to wells at the Southbury Training School, Southbury, Connecticut. U.S. Geological Survey Water-Resources Investigations Report 00–4158.
- Starn, J.J., and C.J. Brown. 2007. Simulations of ground-water flow and residence time near Woodbury, Connecticut. U.S. Geological Survey Scientific Investigations Report 2007–5210. Reston, Virginia: USGS.
- Tiedemann, C.R., and S.M. Gorelick. 1993. Analysis of uncertainty in optimal ground-water contaminant capture design. *Water Resources Research* 29, no. 7: 2139–2153.
- Varljen, M.D., and J.M. Shafer. 1991. Assessment of uncertainty in time-related capture zones using conditional simulation of hydraulic conductivity. *Ground Water* 29, no. 5: 737–748.