Groundwater

Towards Improved Remedial Outcomes in Categorical Aquifers with an Iterative Ensemble Smoother

by Prashanth Khambhammettu[1], Philippe Renard[2], John Doherty[3], Jeremy White[4], Marc Killingstad[5], and Michael Kladias[5]

Abstract

Categorical parameter distributions consisting of geologic facies with distinct properties, for example, high-permeability channels embedded in a low-permeability matrix, are common at contaminated sites. At these sites, low-permeability facies store solute mass, acting as secondary sources to higher-permeability facies, sustaining concentrations for decades while increasing risk and cleanup costs. Parameter estimation is difficult in such systems because the discontinuities in the parameter space hinder the inverse problem. This paper presents a novel approach based on Traveling Pilot Points (TRIPS) and an iterative ensemble smoother (IES) to solve the categorical inverse problem. Groundwater flow and solute transport in a hypothetical aquifer with a categorical parameter distribution are simulated using MODFLOW 6. Heads and concentrations are recorded at multiple monitoring locations. IES is used to generate posterior ensembles assuming a TRIPS prior and an approximate multi-Gaussian prior. The ensembles are used to predict solute concentrations and mass into the future. The evaluation also includes an assessment of how the number of measurements and the choice of the geological prior determine the characteristics of the posterior ensemble and the resulting predictions. The results indicate that IES was able to efficiently sample the posterior distribution and showed that even with an approximate geological prior, a high degree of parameterization and history matching could lead to parameter ensembles that can be useful for making certain types of predictions (heads, concentrations). However, the approximate geological prior was insufficient for predicting mass. The analysis demonstrates how decision-makers can quantify uncertainty and make informed decisions with an ensemble-based approach.

Introduction

Groundwater contamination remains above cleanup goals at hundreds of thousands of hazardous waste sites across the United States (National Research Council 2013) and the rest of the world. According to the U.S. Environmental Protection Agency, expenditures for soil and groundwater cleanup at over 300,000 sites in the United

[1]Corresponding author: Arcadis Inc, USA, 7550 Teague Road, Hanover, MD 21076; 410 923-7823; prashanth.khambhammettu@arcadis.com
[2]University of Neuchatel, Neuchatel, Switzerland; philippe.renard@unine.ch
[3]Watermark Numerical Computing, Brisbane, Queensland, Australia; johndoherty@ozemail.com.au
[4]Intera Inc, Perth, Western Australia, Australia; jwhite@intera.com
[5]Arcadis Inc, Hanover, Maryland; marc.killingstad@arcadis.com; michael.kladias@arcadis.com

States may exceed $200 billion (not adjusted for inflation) by the year 2033 (USEPA 2004)—a significant amount. Heterogeneous formations often underlie contaminated sites with distinct geological facies of different hydraulic conductivities. The low-permeability facies in these formations can serve as secondary contaminant sources to higher permeability zones over time (e.g., Zheng and Gorelick 2003; Chapman and Parker 2005; Chapman et al. 2012; Sale et al. 2013; Farhat et al. 2020). Therefore, characterizing low-permeability facies is vital for accurately forecasting concentrations at contaminated sites.

Managers and stakeholders at these sites rely on various tools, including numerical models, to assess the impact of remedial alternatives, predict future conditions, and manage remediation costs. These numerical models consist of parameters (m) and calculate a state vector d (e.g., groundwater heads; solute concentrations) in response to initial and boundary conditions by solving the *forward problem* using partial differential equations representing groundwater flow and solute transport.

Traditionally, the numerical model is initially parameterized based on a conceptual site model that describes

the geological heterogeneity, and initial/boundary conditions. Subsequently, the parameter vector (m) is adjusted by solving the *inverse problem*, such that the model can match measured aquifer states. Once the match between simulated and measured aquifer states is deemed acceptable, this calibrated numerical model is used in a predictive context. Often, predictions are made using a single model; however, inaccurate predictions can have human health/ecological/monetary repercussions. The stochastic form of the inverse problem (Aster et al. 2013) is shown below.

$$q(\boldsymbol{m}|\boldsymbol{d}) = \frac{f(\boldsymbol{d}|\boldsymbol{m})\,p(\boldsymbol{m})}{\text{constant}} \tag{1}$$

In equation 1, $p(\boldsymbol{m})$ represents the prior probability distribution and contains parameters unconstrained by measurement states ($\boldsymbol{d}$) or site-specific information. The term $f(\boldsymbol{d}|\boldsymbol{m})$ represents the likelihood of simulating the measured state data $\boldsymbol{d}$ given model parameter $\boldsymbol{m}$. The posterior probability density function ($q(\boldsymbol{m}|\boldsymbol{d})$) represents the probability of occurrence of the parameter $\boldsymbol{m}$ given the measured state data $\boldsymbol{d}$.

Because the inverse problem does not have a unique solution (Tarantola 2005), predictions made with a calibrated numerical model do not always bear fruit (Moore and Doherty 2005; Ahmmed et al. 2020). Since it is nearly impossible to characterize heterogeneous subsurface environments, Rajaram 2016, suggests that predicting solute behavior in such settings for deterministic predictions at relevant scales necessitates a stochastic approach, which is in line with a typical decision-making agency's expectation that predictions should be accompanied by uncertainty measures that allow risk assessment.

Zhou et al. 2014 published a review article on the inverse problem and the associated methods and challenges. Aquifers at contaminated sites often consist of distinct geological facies/categories. The inverse problem at these sites is known as the categorical inverse problem. Several authors (e.g., Hu et al. 2001; Caers and Hoffman 2006; Mariethoz et al. 2010; Li et al. 2013) have presented methods to solve the categorical inverse problem and generated parameter ensembles that sample the posterior distribution. Linde et al. 2017 presented a summary of several methods to solve this class of problems.

Khambhammettu et al. 2020 developed the Traveling Pilot Point method (TRIPS) based on "traveling" pilot points to parameterize the categorical inverse problem. For a synthetic model, the TRIPS method was used in conjunction with linear subspace methods to solve the categorical inverse problem and estimate a categorical parameter ensemble.

A computationally intensive requirement of linear subspace methods is the estimation of a sensitivity (Jacobian) matrix $\boldsymbol{G}$, which contains the partial derivatives of the measured data with respect to the parameters in the model. The iterative ensemble smoother (IES) (Chen and Oliver 2013) offers a computationally efficient method for estimating an approximate form of the Jacobian matrix, $\boldsymbol{G}$,

and sample from the posterior distribution. A software package called pestpp-ies (White 2018) contains an implementation of the IES.

Ensemble smoothers are increasingly used to condition groundwater parameter fields to aquifer state measurements (e.g., Cao et al. 2018; Knowling et al. 2019; Lam et al. 2020; White et al. 2020a). The ensemble smoother's documented success and computational efficiency motivated us to explore their use at contaminated sites with discrete geological facies. This paper demonstrates a framework where the TRIPS method for parameterizing the categorical inverse problem and the pestpp-ies software for conditioning parameter ensembles could be used for evaluating predictive uncertainty.

The fate and transport of a conservative solute (no sorption/decay) in a synthetic aquifer with two distinct geological facies (permeable channels incised in a low-permeability matrix) is analyzed in this paper. Since groundwater heads alone are not adequate for accurately characterizing aquifer heterogeneity (Fiori et al. 2016; Rajaram 2016; Schilling et al. 2019), a measurement vector $\boldsymbol{d}$ comprising both groundwater solute concentrations and groundwater heads is considered. TRIPS and pestpp-ies are used to assimilate this measurement vector and generate posterior parameter ensembles under various assumptions of heterogeneity and size of the measurement vector $\boldsymbol{d}$.

Since the categorical parameter ensembles explicitly represent both permeable channels and low-permeability matrix, they are well suited to represent both the rapid advective transport through the permeable channels and the delayed diffusive transport from the solute stored in the low-permeability matrix. The applicability of these posterior parameter ensembles in predicting solute concentrations and mass at a synthetic contaminated site is evaluated. The evaluation also includes an assessment of how the number of measurements and the choice of the geological prior determine the characteristics of the posterior ensemble and the resulting predictions.

This paper is organized as follows. The various methods used in this paper are presented under the *Methods* section. The synthetic problem used in this paper is described in the section titled *Synthetic Problem*. Posterior parameter ensembles for the synthetic problem are presented in the *Results* section. The implications of the number of observations and the choice of the geological conceptual model on the accuracy of predictions are discussed. A summary of findings is presented in the *Summary* section.

Methods

In this section, we briefly summarize the IES method and the TRIPS method.

Iterative Ensemble Smoother

The Ensemble Smoother (ES) (van Leeuwen and Evensen 1996) "smooths" all the members in an ensemble by assimilating data in a single step and computing a

global member update. Skjervheim and Evensen 2011, demonstrated the applicability of ensemble smoothers in characterizing the posterior distribution for reservoir history-matching problems. To improve the fit to the measured data, Emerick and Reynolds 2013, developed the IES by assimilating the same data several times to iteratively improve the match obtained by the ES method. IES progressively smooths ensembles of prior parameter distributions so that they represent a posterior parameter ensemble. Chen and Oliver 2013 proposed an improved IES by starting with a prior parameter ensemble and repeatedly iteratively assimilating data and smoothing. This method, referred to as LM-EnRML, used the Levenberg–Marquardt method (Hanke 1997) to regularize the parameter update vector for an ensemble. The pestpp-ies software (White 2018) incorporates the LM-EnRML method into the PEST++ framework (White et al. 2020b) and is used in this paper.

The Traveling Pilot Point Method

Khambhammettu et al. 2020 proposed the TRIPS method where pilot points are not used for spatial interpolation but are instead used to define the geometry of discrete facies. The method can be used in conjunction with different geostatistical techniques like, for example, transition probabilities or object-based models. Providing the TRIP locations as conditioning data to the geological simulation algorithm allows changing the parameterization of the geological simulation, and solving the inverse problem in this manner means to search for the optimal locations of these TRIPSs. This approach modifies a discrete inverse problem into a continuous one and makes it more tractable.

Synthetic Problem

In this section, a synthetic problem is designed to illustrate the decision-making process for evaluating the uncertainty in forecasting/predicting solute concentrations and mass at contaminated sites. This problem has several elements in common with complex real world settings: a mix of low- and high-permeability facies, decade-long solute transit times, head and concentration measurements, sensitivity to the size of the measurement dataset, and the assumed geological model of heterogeneity. The problem is described below and contains two data sets used for the inversion, plus one data set to assess the quality of predictive uncertainty assessment. Posterior parameter ensembles are developed for this problem using TRIPS in conjunction with pestpp-ies. These parameter ensembles are applied in a predictive context as described in the *Predictive Evaluation* section.

In this problem, two-dimensional groundwater flow and transport are simulated in a synthetic cube-shaped confined aquifer $(100\,m \times 100\,m \times 1\,m)$ containing two facies, high-permeability channels incised in a low-permeability matrix. As in heterogeneous real-world contaminated sites, the low-permeability matrix facies in this problem can serve as a secondary contaminant source to the higher-permeability channel facies over time. A reference hydraulic conductivity distribution was previously developed for this aquifer using the TRIPS method (Khambhammettu et al. 2020) based on a training image. The reference aquifer was assumed to have a maximum of three channels with five TRIPS per channel. For simplicity, the channels were assumed to have a constant width of 6.5 m and a constant thickness of 1 m. A previously developed prior covariance matrix (Khambhammettu et al. 2020) for flow was augmented to represent the prior covariance for the transport parameters. An ensemble member with hydraulic conductivities of 8.7×10^{-3} m/s (meters per second) for the channel facies and 1.1×10^{-4} m/s for the matrix facies was selected as the reference hydraulic conductivity distribution (Figure 1).

Constant head boundaries of 1 and 0.95 m were enforced on the left and right edges of the model, respectively, to simulate steady-state two-dimensional groundwater flow from left to right. The low gradient across the model grid (5×10^{-4}) increases solute travel times through the model, mimicking travel times at a real-world site. Since the simulation of solute transport requires a finely spaced grid (Konikow 2011), a numerical grid with 1 layer (thickness of 1 m), 500 rows, 500 columns (row and column spacings of 0.2 m) was used in this model.

Two different monitoring scenarios are considered in this evaluation. In the first monitoring scenario, groundwater heads and concentrations are monitored at 13 monitoring locations in the Site. This scenario is assumed to represent a relatively sparse monitoring network. The second monitoring scenario considers a denser monitoring network with 25 monitoring locations—almost double the number of locations in the sparse scenario. The model grid, facies, boundary conditions, and monitoring locations are shown in Figure 1.

A dissolved solute source with a concentration of $10^4\,\mu g/L$ is assumed to be present in the second column of the model along all the rows for a period of 2 years. After 2 years, the source is assumed to be removed instantaneously, and the remaining solute is allowed to transit through the aquifer and exit through the downgradient specified head boundary.

Groundwater flow and transport are simulated using MODFLOW 6 (Langevin et al. 2021). For this problem, the solute is assumed to be conservative, and sorption and decay processes are not represented. Hydrodynamic dispersion is ignored for simplicity. An attempt was also made to explicitly incorporate diffusive transport, but very little difference was observed in the simulated concentrations with and without diffusion. This observation is consistent with the findings of Chapman et al. 2012, who demonstrate that simulation of diffusive mass transport required centimeter scale grid cells with very small time steps. Ideally, a more refined grid with very small timesteps would be better suited for representing the diffusion processes. However, that spatial and temporal refinement would have resulted in very long simulation times, complicating the task of uncertainty analysis, which requires several model evaluations.
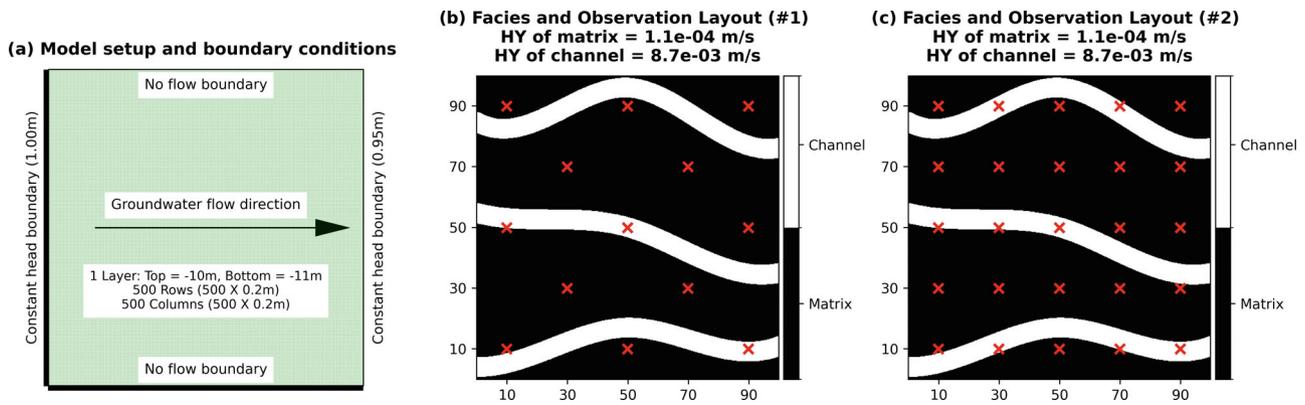
Figure 1. Model setup, boundary conditions, and monitoring well locations for the groundwater flow problem. The model setup and boundary conditions are shown in panel (a). The facies distribution of the reference model and the sparse (13 locations) and dense (25 locations) monitoring networks are shown in panels (b) and (c), respectively.

As a compromise, the dual-domain formulation, with upstream finite difference scheme and adaptive time stepping is used in this evaluation. In the dual-porosity/domain formulation (Deans 1963; Coats and Smith 1964; Van Genuchten et al. 1974), a heterogenous aquifer is assumed to comprise of a mobile domain where advective transport is dominant, and overlapping immobile domain where molecular diffusion is dominant. A mass-transfer coefficient (ζ_{im}) is used to control mass transfer between the mobile and immobile domains.

In the reference model, the total porosity of the channel facies (θ_{ch}) is assumed to be 35% and the total porosity of the matrix facies (θ_{mx}) is assumed to be 43%. These porosities are representative for high- and low-permeability materials (Payne et al. 2008). The mobile porosities for the channel ($\theta_{m,ch}$) and matrix ($\theta_{m,mx}$) were assumed to be 10% and 1%, respectively. The mass-transfer coefficient (ζ_{im}) between the mobile and immobile domains is assumed to be 10^{-3}/day.

In the reference model, the hydraulic conductivity contrast ratio between the channel and matrix facies is 79 ($8.7 \times 10^{-3}/1.1 \times 10^{-4}$). This ratio compares well with researchers (Zheng and Gorelick 2003) who noted that when the hydraulic conductivity in permeable channels approaches 100 times that of the matrix's hydraulic conductivity, calculated concentration distributions start exhibiting non-Gaussian patterns with pronounced tails which are incidentally observed in real-world contaminated sites.

Solute transport was simulated for a period of 30 years assuming steady-state groundwater flow. It was assumed that measurement data exists only for the first 4 years and that the remaining 26 years represent future (predictive) conditions. Because of the steady-state flow assumption, a single head measurement was taken at the monitoring locations (13 or 25, depending on the scenario). At each monitoring location, solute concentrations were measured quarterly over a 4-year period, resulting in a total of 16 measurements (4×4). Given measured heads and concentrations over the first 4 years, the inverse problem requires the estimation

of parameter ensembles that honor the measured data. The objective function, ($\varnothing_m$), minimized for the inverse problem is represented mathematically as shown below in equation 2.

$$\varnothing_m = \left[d_{h,mes} - d_{h,sim}\right]^T \left[d_{h,mes} - d_{h,sim}\right]$$
$$+ \left[d_{c,mes} - d_{c,sim}\right]^T \left[d_{c,mes} - d_{c,sim}\right] \quad (2)$$

$d_{h,mes}$ and $d_{h,sim}$ represent the measured and simulated heads, respectively. These vectors have sizes of 13 and 25 for the sparse and dense monitoring scenarios, respectively. $d_{c,mes}$ and $d_{c,sim}$ represent the $\log_{10}$ transformed measured and simulated concentrations. These vectors have sizes of 208 and 400 for the sparse and dense monitoring scenarios, respectively. Measurement noise was ignored in this evaluation for simplicity and hence the covariance matrix of the measurement noise error is not shown in the equation. Since aquifer heterogeneity is a key driver of solute transport (Zheng and Gorelick 2003; Konikow 2011), two different priors to model the geological heterogeneity for the inverse problem are evaluated: the continuous and the discrete cases.

Continuous Case

One of the aims of this paper is to investigate how the inverse problem and uncertainty quantification will perform if, instead of using the proper conceptual model for the geological heterogeneity, a simpler multi-Gaussian model is used. This mismatch of the geological priors is designed to illustrate real-world conditions where one often works with a limited understanding of subsurface heterogeneity and relies on multi-Gaussian models.

In this case, the hydraulic conductivity is evaluated at 250,000 (500×500) grid block locations and is assumed to vary between 5×10^{-5} m/s and 5×10^{-2} m/s. Since the number of model evaluations for calculating the parameter sensitivity matrix in the pestpp-ies implementation is not governed by the number of parameters; it is feasible to estimate the sensitivity matrix for many parameters efficiently.

P. Khambhammettu et al. Groundwater 62, no. 1: 60–74

The parameters of the multi-Gaussian model were obtained by seeking a match with the variance of the reference hydraulic conductivity field rather than the underlying conceptual heterogeneity. The hydraulic conductivity distribution for the reference model was log-transformed, and a spherical semivariogram with a sill of 0.583, a range of 140 along the X direction, and an anisotropy ratio of 4 was fitted to the empirical variogram computed from the reference hydraulic conductivity field. This variogram was then used to generate realizations of the logarithm of the hydraulic conductivity using the FIELDGEN utility from the PEST groundwater utility suite (Doherty 2018). FIELDGEN generates these parameter distributions based on the Sequential Gaussian Simulation technique (Deutsch and Journel 1998).

Three additional parameters are also estimated for the continuous case. These were the dual domain mass-transfer coefficient (ζ_{im}), the mobile porosity for the channel facies ($\theta_{m,ch}$), and a parameter identifying the transition between channel and matrix facies (K_{tran}). Model cells with estimated hydraulic conductivities greater than K_{tran} were assumed to be channel cells and assigned a total porosity of 35% (θ_{ch}) and the estimated mobile porosity for the channel facies ($\theta_{m,ch}$). Model cells with estimated hydraulic conductivities less than K_{tran} were assumed to be facies cells and assigned a total porosity of 43% (θ_{mx}) and the estimated mobile porosity for the matrix facies ($\theta_{m,mx}$). To simplify the problem, $\theta_{m,mx}$ was constrained to be 1/10th of $\theta_{m,ch}$.

Discrete (Categorical) Case

In the discrete case, a categorical hydraulic conductivity distribution is created using the TRIPS method, assuming three channels with five TRIPS per channel, resulting in 15 TRIPS. The number of channels was determined based on a training image and discussed in Khambhammettu et al. 2020. Besides the TRIPS locations, the channel and matrix hydraulic conductivities are also estimated using pestpp-ies. The channels are allowed to move in and out of the model domain to represent situations with fewer than three channels. Channel width and thickness are assumed to be constant.

Two additional parameters: the dual domain mass-transfer coefficient (ζ_{im}), and the mobile porosity for the channel facies $\theta_{m,ch}$ are also estimated. The mobile porosity for the matrix facies, $\theta_{m,mx}$, was constrained to be 1/10th of $\theta_{m,ch}$. The discrete approach has a clear advantage as it assumes a geological prior for the aquifer heterogeneity that is the same as the reference model.

Input Summary

In summary, four different cases are considered: two with a sparse data set and continuous or discrete hydraulic conductivity fields, and two with a dense data set with continuous or discrete hydraulic conductivity fields. Table 1 summarizes the parameters and observations for these four cases. In subsequent sections of this paper, parameter ensembles that honor the measured heads and concentrations in these different cases are presented.

Results

Posterior parameter ensembles for the continuous case and the categorical case, along with predictive uncertainty analyses, are presented in this section.

Results for the Continuous Case

pestpp-ies was used to estimate the posterior parameter ensemble. As shown in Table 1, this inverse problem required the estimation of 250,003 parameters ($250{,}000 + 3$) based on 13 (or 25) head observations and 208 (or 400) concentration observations. An initial ensemble of 100 realizations was progressively smoothed for a maximum of 15 iterations. Models with realizations that took over four times the average run time or whose objective function was two standard deviations outside the ensemble mean objective function (defining the misfit between the simulated and measured values) were discarded to keep the estimation stable.

The ensemble diversity and simulated heads and concentrations were evaluated at each iteration. Parameters corresponding to iterations 13 and 10 were selected to be the posterior ensembles for the sparse and dense monitoring scenarios, respectively. The process of choosing the iteration was a compromise between ensemble diversity and matching the measured data. The posterior ensemble for the sparse monitoring scenario had 45 members, while the posterior ensemble for the dense monitoring scenario had 65 members.

Estimated ensemble characteristics are shown in Figure 2. Hydraulic conductivity and simulated heads for the reference model are shown in the top row. Prior ensemble characteristics (mean and standard deviation of the hydraulic conductivity and simulated heads) are shown

Table 1
Parameters and Observations for Inverse Problem

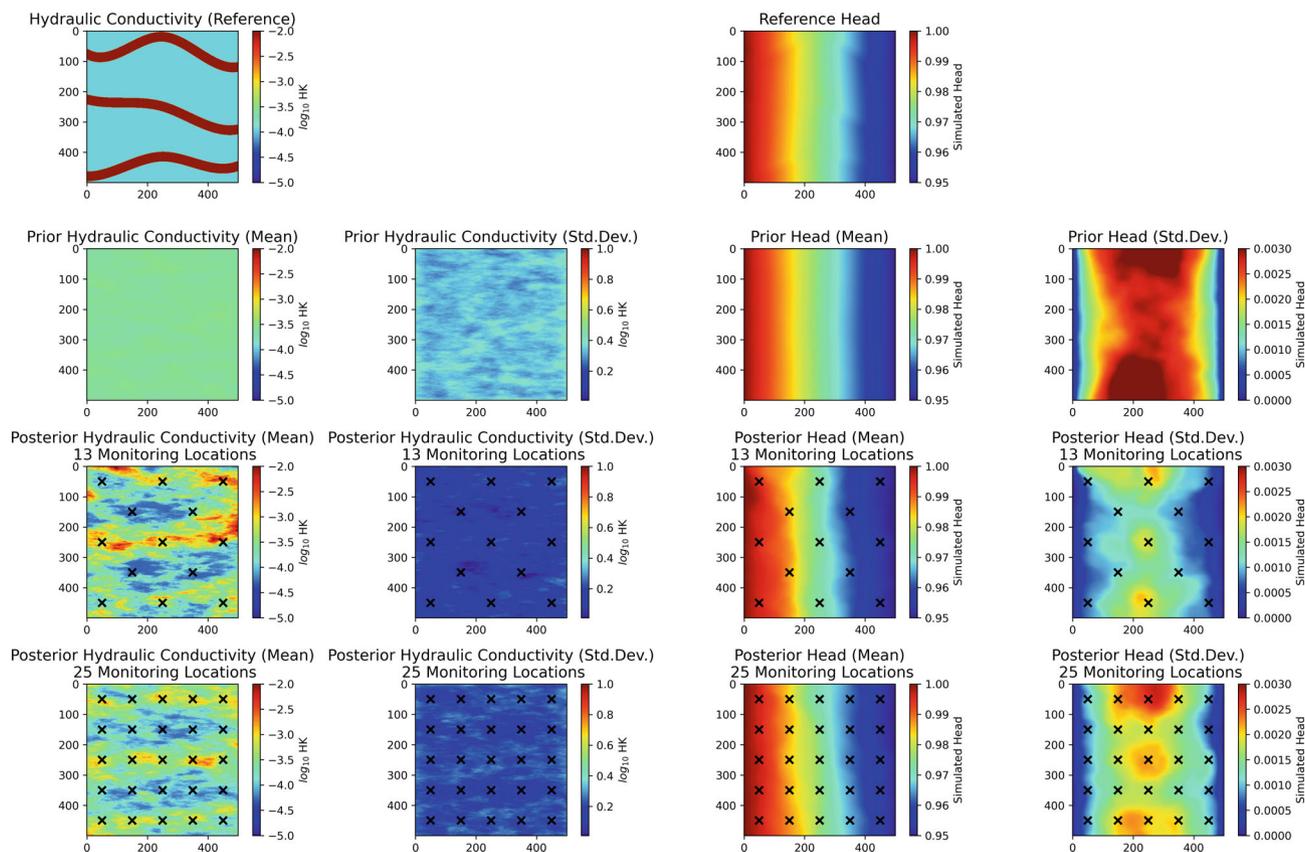| Heterogeneity Approach	Monitoring Network	Head Measurements	Concentration Measurements	Heterogeneity Parameters	Other Parameters
Continuous	Sparse	13	208 (13×16)	250,000	$\zeta_{im}, \theta_{m,ch}, K_{tran}$
Continuous	Dense	25	400 (25×16)	250,000	$\zeta_{im}, \theta_{m,ch}, K_{tran}$
Categorical	Sparse	13	208 (13×16)	17	$\zeta_{im}, \theta_{m,ch}$
Categorical	Dense	25	400 (25×16)	17	$\zeta_{im}, \theta_{m,ch}$

Figure 2. Ensemble hydraulic conductivity and heads for the continuous parameterization case. The hydraulic conductivity and simulated heads for the reference model are shown in the top row. The characteristics of the prior ensemble (mean and standard deviation of the hydraulic conductivity and simulated heads) are shown in the second row. The characteristics of the posterior ensemble for the sparse and dense monitoring scenarios are shown in the third and fourth rows, respectively.

in the second row. Posterior ensemble characteristics for the sparse and dense monitoring scenarios are shown in the third and fourth rows, respectively. While neither monitoring scenario could faithfully produce the original ensemble's characteristics, they could still produce specific characteristics of the reference model. The ensemble mean hydraulic conductivity for both monitoring scenarios has highs in the center and along the north and south edges and lows in the center—a pattern observed in the reference model. However, the mean hydraulic conductivity from the posterior ensemble estimated using the denser monitoring network seems to represent the orientation of the channel facies better. The ensemble standard deviation of hydraulic conductivity for the posterior ensemble was lower than that of the prior ensemble.

Assimilation of the measured data led to a better match between the measured and simulated heads for the posterior ensemble. The ensemble mean head exhibited some differences when compared to the reference head distribution mainly in the southern regions of the aquifer. The ensemble standard deviation of simulated head for the posterior ensembles is lowest near the left and right edges because of the specified boundary heads.

The simulated concentrations for the reference model are compared against the ensemble mean concentrations

of the posterior ensembles obtained from the sparse and dense monitoring networks in Figure 3. Concentrations are compared at times of 0, 2, 5, 10, 20, and 30 years. The reference concentrations are shown in the first column. The high-permeability regions allow the solute to quickly exit the domain, whereas the low-permeability regions retain solute mass for decades. This inability to match the reference concentrations is most likely a consequence of incorrectly assuming a continuous parameter distribution model. Compared to the reference model, the posterior ensembles underestimate the time at which solute reaches the boundary because the estimated permeable regions are disconnected. However, the estimated posterior ensembles exhibit similarities in where the high and low concentrations occur in the model. The concentration results from the denser monitoring network compare slightly better to the reference model because the percentage of the channel facies in this ensemble is closer to the reference distribution.

Results for the Categorical Case

As shown in Table 1, this inverse problem required the estimation of 19 (17 + 2) parameters based on 13 (or 25) head observations and 208 (or 400) concentration observations. An initial ensemble of 100 realizations was progressively smoothed for 10 iterations. Models with
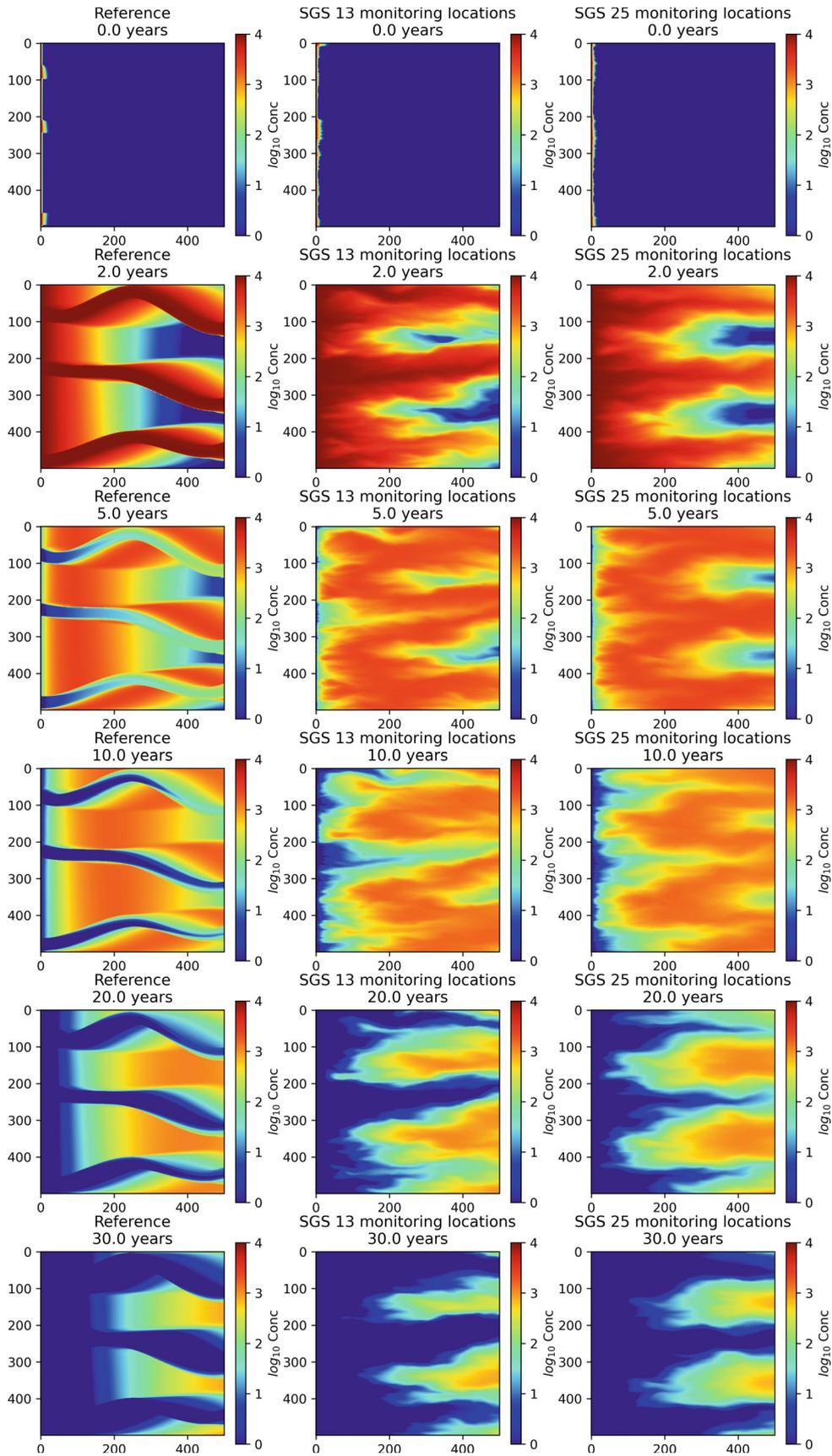
Figure 3. Simulated concentrations for the continuous parameterization case. The reference concentrations are shown in the first column. The ensemble average for the sparse monitoring scenario is shown in the second column. The ensemble average for the dense monitoring scenario are shown in the third column.
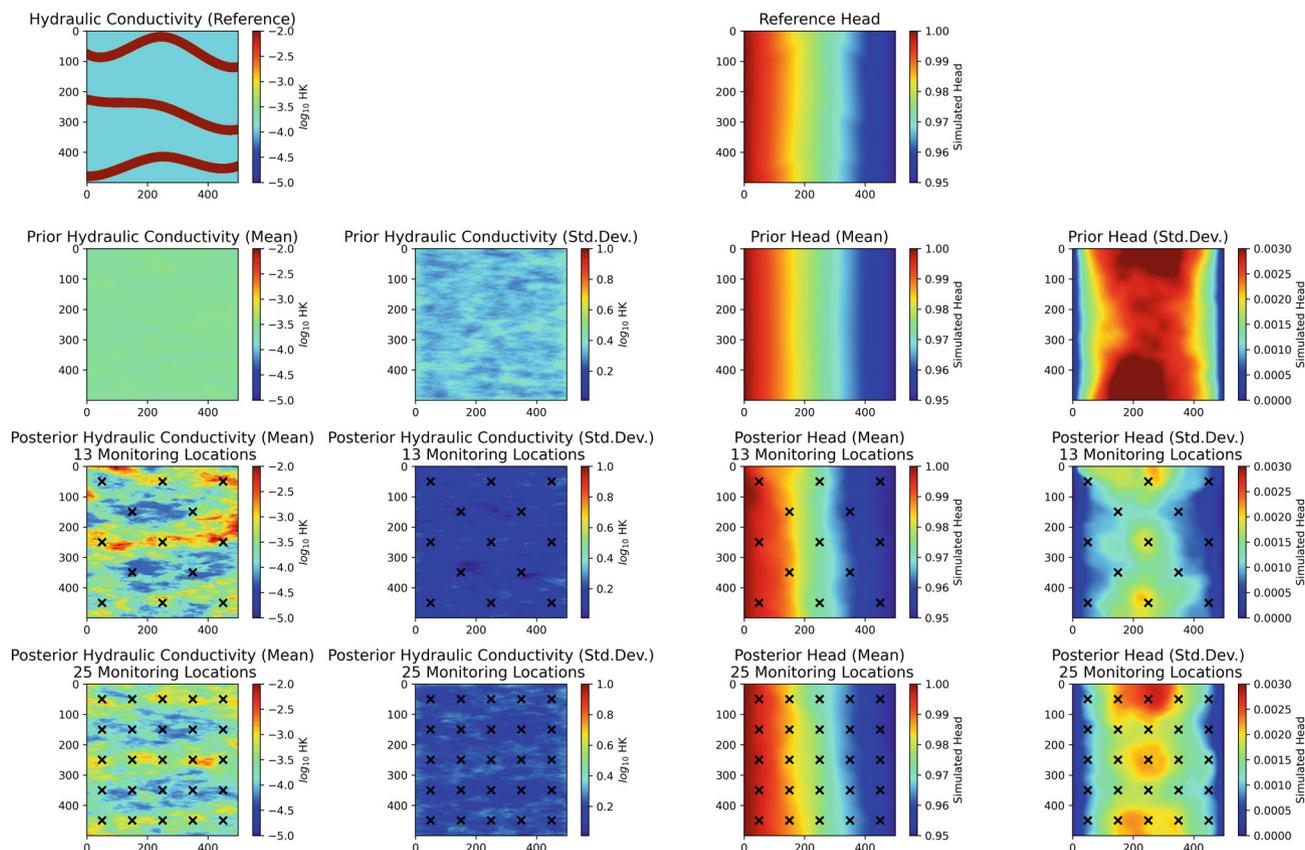
Figure 4. Ensemble hydraulic conductivity and heads for the discrete parameterization case. The hydraulic conductivity and simulated heads for the reference model are shown in the top row. The characteristics of the prior ensemble (mean and standard deviation of the hydraulic conductivity and simulated heads) are shown in the second row. The characteristics of the posterior ensemble for the sparse and dense monitoring scenarios are shown in the third and fourth rows, respectively.

realizations that took more than four times the average run time or whose objective function was two standard deviations outside the ensemble mean objective function were discarded to keep the estimation stable.

Parameters corresponding to iterations 6 and 7 were selected to be the posterior ensembles for the sparse and dense monitoring scenarios, respectively. As in the continuous case, this selection was performed to maintain diversity in the ensemble while also matching the measured data. The posterior ensemble for the sparse monitoring scenario had 41 members while the posterior ensemble for the dense monitoring scenario had 45 members.

The characteristics of the estimated ensembles are shown in Figure 4. The hydraulic conductivity and simulated heads for the reference model are shown in the top row. The features of the prior ensemble (mean and standard deviation of the hydraulic conductivity and simulated heads) are shown in the second row. The characteristics of the posterior ensemble for the sparse and dense monitoring scenarios are shown in the third and fourth rows, respectively. The ensemble mean hydraulic conductivities and simulated heads closely resemble the reference distribution for both the sparse and dense monitoring networks. The match for the discrete case is a lot better than the corresponding match for

the continuous case shown in Figure 3. The standard deviation for the posterior ensembles is highest at the boundaries of the channel facies suggesting that the ensemble members differ in the thickness of the channels. The ensemble standard deviation of simulated head is very small indicating similarity in the simulated heads.

The simulated concentrations for the reference model are compared against the ensemble mean concentrations of the posterior ensembles obtained from the sparse and dense monitoring networks in Figure 5. Concentrations are compared at times of 0, 2, 5, 10, 20, and 30 years. The reference concentrations are shown in the first column. The posterior ensembles generated from both the sparse and dense monitoring are very much like the concentrations from the reference model. The ensemble parameter statistics (mean and standard deviation) for the continuous and discrete cases are shown in Table 2. In both monitoring cases, ensemble smoothing reduced the variability of simulated concentrations and led to good matches at all the monitoring locations.

Predictive Evaluation

This section uses the posterior ensembles for the continuous and discrete cases to support the remedial decision-making process. The prior and posterior ensembles for both cases were used to make two kinds of
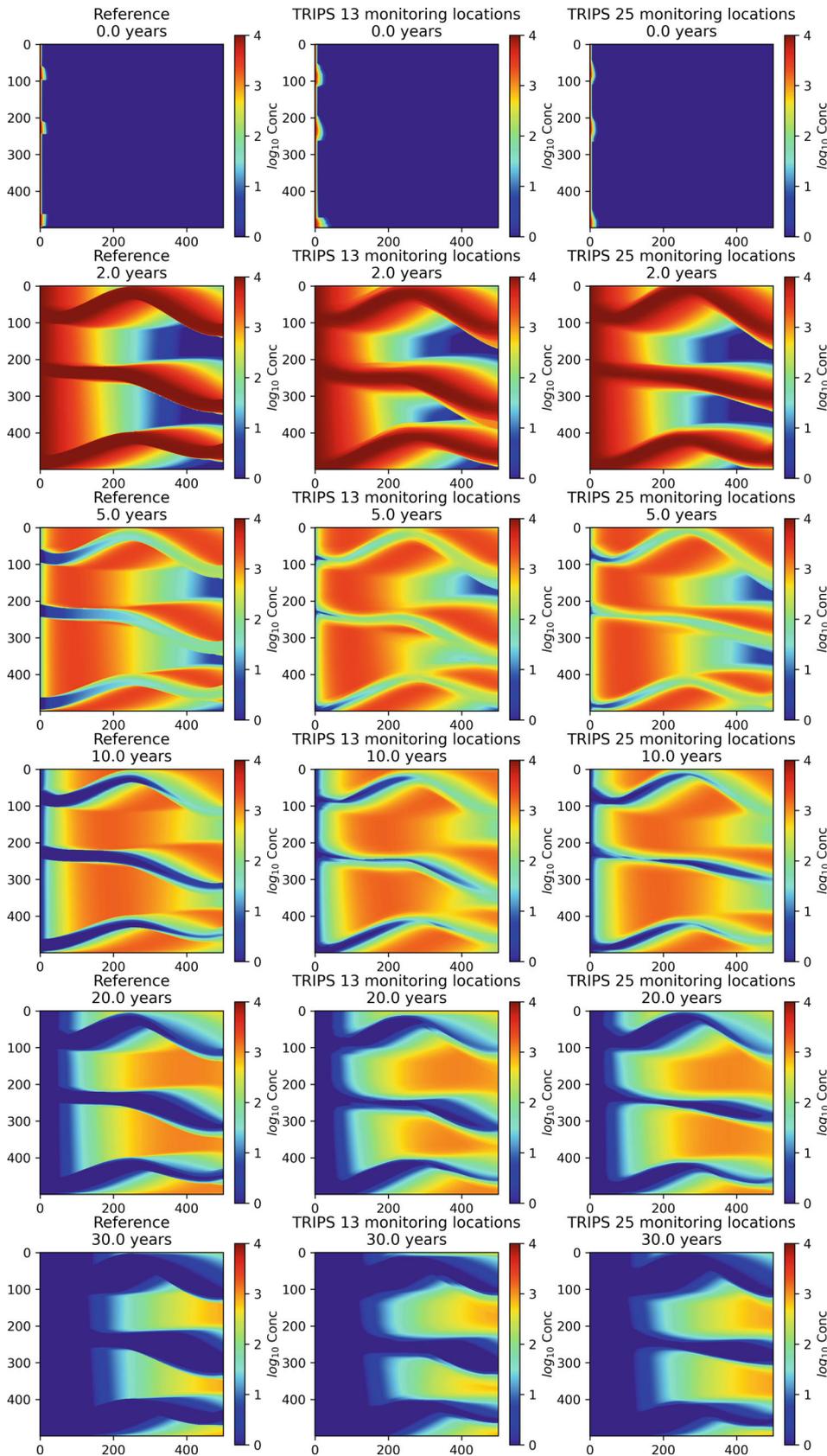
Figure 5. Simulated concentrations for the discrete parameterization case. The reference concentrations are shown in the first column. The ensemble average for the sparse monitoring scenario is shown in the second column. The ensemble average for the dense monitoring scenario is shown in the third column.

Table 2
Parameter Ensemble Statistics

Case	Parameter	Ensemble	Ensemble Mean	Ensemble Standard Deviation
Continuous	ζ_{im} ($\log_{10}$ day^{-1})	Reference	−3.00	—
		Prior	−2.94	0.6851
		Posterior (Sparse)	−3.00	0.0207
		Posterior (Dense)	−3.00	0.0379
	Mobile porosity of Channel facies ($\theta_{m,ch}$)	Reference	0.10	—
		Prior	0.10	0.0280
		Posterior (Sparse)	0.14	0.0090
		Posterior (Dense)	0.12	0.0157
Discrete	ζ_{im} ($\log_{10}$ day^{-1})	Reference	−3.00	—
		Prior	−3.00	0.5281
		Posterior (Sparse)	−3.01	0.0315
		Posterior (Dense)	−3.04	0.0254
	Mobile porosity of Channel facies ($\theta_{m,ch}$)	Reference	0.10	—
		Prior	0.10	0.0450
		Posterior (Sparse)	0.07	0.0072
		Posterior (Dense)	0.10	0.0092
	Hydraulic Conductivity ($\log_{10}$ m/s) of Channel facies	Reference	−2.06	—
		Prior	−1.95	0.0342
		Posterior (Sparse)	−2.11	0.0543
		Posterior (Dense)	−2.13	0.0465
	Hydraulic Conductivity ($\log_{10}$ m/s) of Matrix facies	Reference	−3.96	—
		Prior	−3.89	0.0363
		Posterior (Sparse)	−3.98	0.0222
		Posterior (Dense)	−3.98	0.0175

predictions. First, the ensembles were used to predict maximum concentrations at the down-gradient boundary (column 499 of the model). At a real site, this type of prediction is typically performed for areas with sensitive receptors where concentrations are compared against a regulatory threshold limit. Second, the ensembles were used to predict the total mass exiting the aquifer. At a real site, this prediction is typically used to design treatment systems and quantify the mass remaining in the system. The mass prediction is more challenging between the two predictions because it is a spatial aggregation of both the concentration distribution and the volumetric flux.

The predicted maximum concentrations for the various ensembles are shown in Figure 6. The maximum concentrations from the reference model are shown with a dashed line in this figure. The predictions from the prior ensembles for both the continuous and discrete cases are shown in the first column. The prior predictions vary several orders of magnitude while encompassing the reference prediction. The predictions with the posterior ensembles for the continuous case are shown in the top row and the predictions with the posterior ensembles for the discrete case are shown in the bottom row. The second column represents the ensembles based on the sparse monitoring network, while the third column represents the ensembles based on the denser monitoring network.

For the continuous case, the ensemble for the denser monitoring network does better at bounding the

reference prediction than the ensemble for the sparser monitoring network. This is interesting given the discrepancies between the posterior ensembles for the continuous case and the reference distribution. It is possible that for some predictions, the act of assimilating additional head/concentration data could reduce predictive uncertainty even when working with an imperfect geological prior.

For the discrete case, both sets of posterior ensembles (sparse and dense) performed well in bounding the reference prediction and reproducing the features of the reference prediction, like the kink observed between 10 and 15 years. The close correspondence between the ensembles for the two different monitoring networks indicates that when the geological prior is representative, additional head/concentration data did not significantly add value to the process of quantifying the predictive uncertainty. These findings should be considered in the context of the current example. At most real-world locations, the geological prior, contamination source locations/ release history, are unknown/uncertain, and additional data may help better quantify the predictive uncertainty.

The predicted mass estimates (cumulative over time) in kilograms (kg) for the various ensembles are shown in Figure 7. This figure shows the mass exiting the reference model with a dashed line. The predictions from the prior ensembles for both the continuous and discrete cases
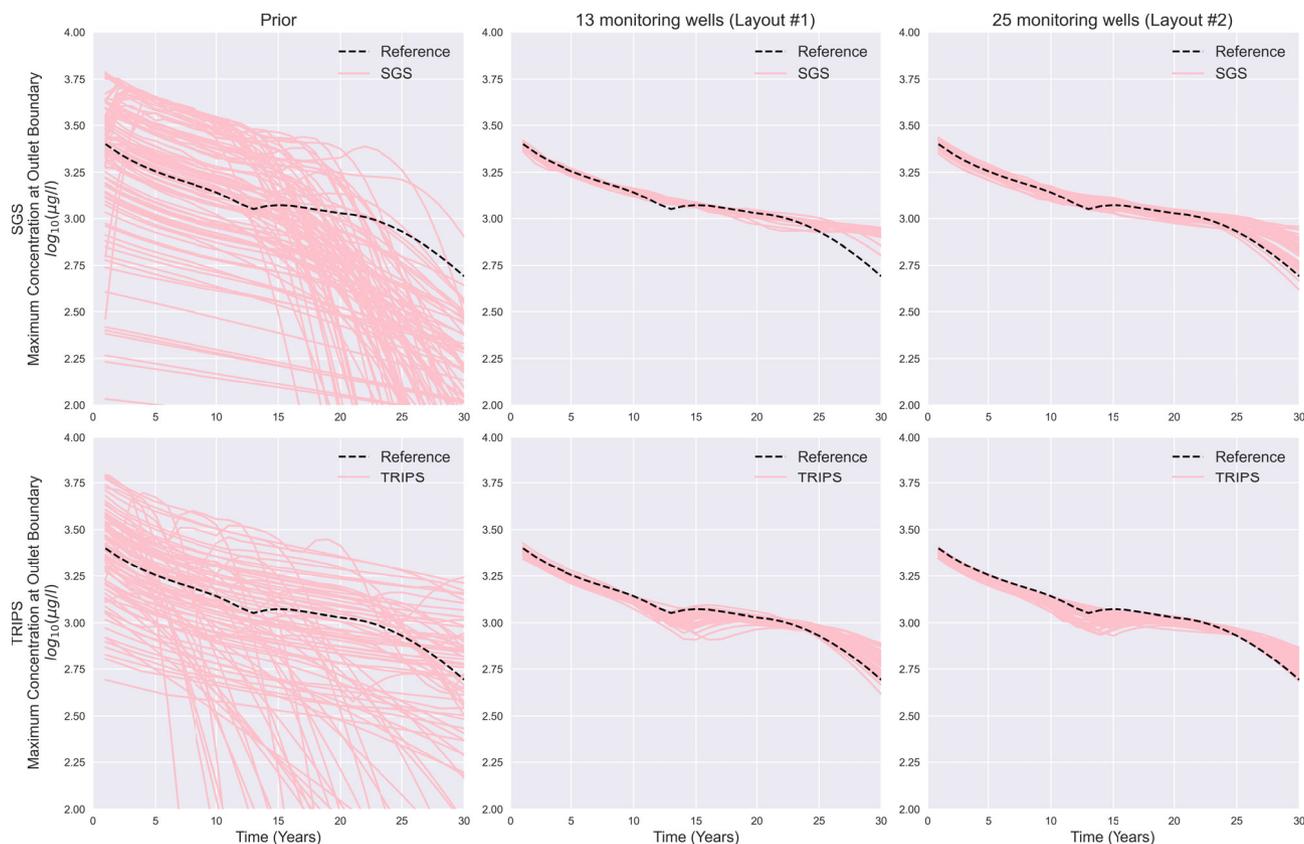
Figure 6. Predicted maximum concentrations at the downgradient boundary (column 499). Predictions from the ensembles for the continuous case are shown in the top row and predictions for the ensembles for the discrete case are shown in the bottom row. The first column depicts the concentrations from the prior ensembles. The second column depicts the concentrations from the scarce monitoring scenario. The last column depicts the concentrations from the dense monitoring scenario.

are shown in the first column. The predictions with the posterior ensembles for the continuous case are shown in the top row and the predictions with the posterior ensembles for the discrete case are shown in the bottom row. The prior predictions consistently underpredict the reference estimate. The second column represents the ensembles based on the sparse monitoring network while the third column represents the ensembles based on the denser monitoring network.

The reference model estimates that nearly 54 kg of solute would exit the aquifer at the end of 30 years. The reference model also predicts that most of the mass would have exited the aquifer within the first 5 years. Interestingly, none of the continuous prior ensembles (first row, first column) can match this behavior. The highest mass estimate from the continuous prior ensemble is about 10 kg—an underprediction by a factor of five. The primary reason for the discrepancy is that the high-permeability areas in the continuous case are not well connected. As a result, there is an accumulation of mass within the aquifer, implying that lesser amount of mass inflow (as compared to the discrete case) is required into the system to maintain similar concentrations. This discrepancy in the mass estimates is also seen in the continuous posterior ensembles for the two monitoring scenarios (second and third columns, first row). On the

other hand, the estimates from the discrete TRIPS prior ensemble encapsulate the reference estimate. Assimilation of head and concentration measurements lead to the posterior ensembles encapsulating the reference estimate (second and third columns, second row).

This analysis thus demonstrates that IES can assimilate head and concentration data to develop parameter ensembles that could be used for evaluating prediction uncertainty. More importantly, these results show that all predictions are not the same. For the concentration predictions, the continuous ensembles could "reasonably" simulate the spatial patterns of heads and concentrations and bound the reference prediction. However, the continuous ensembles were not suited to estimate the mass exiting the system. The discrete TRIPS ensembles on the other hand do an excellent job of simulating both heads and concentrations and are also well suited as a predictive tool for this problem.

Summary and Discussion

Decision makers at contaminated sites are often required to make complex decisions regarding the fate and transport of the contaminants at their sites. They often rely on a single calibrated groundwater model to make forecasts several years/decades into the future to
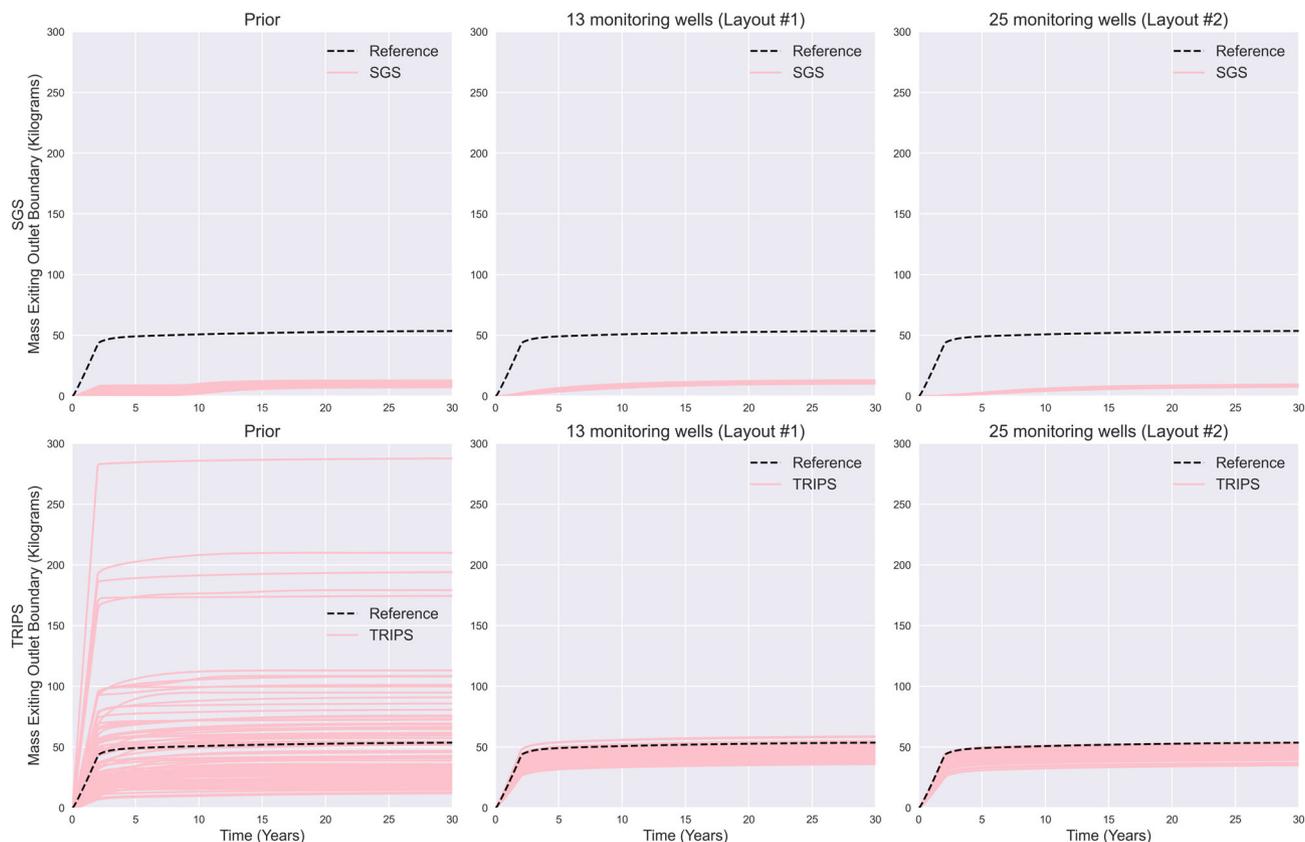
Figure 7. Predicted mass (kilograms) exiting the downgradient boundary (cumulative over time). Predictions from the ensembles for the continuous case are shown in the top row and predictions for the ensembles for the discrete case are shown in the bottom row. The first column depicts the mass from the prior ensembles. The second column depicts the concentrations from the scarce monitoring scenario. The last column depicts the concentrations from the dense monitoring scenario.

guide decision-making. However, the non-unique nature of the inverse problem can lead to predictions that may not occur. This paper presented an approach for developing an ensemble of models that honor measured aquifer states and can be used to develop multiple remedial forecasts to quantify the predictive uncertainty arising from parameter uncertainty.

The primary contribution of this paper is to demonstrate that TRIPS and pestpp-ies together can be used to develop categorical parameter ensembles that honor measured aquifer heads and concentrations simultaneously. Additionally, the analysis illustrates how multiple puzzle pieces (geological parametrization, history matching, and remedial forecasts) could be assembled to guide decision-makers at contaminated sites by quantifying the predictive uncertainty associated with parameter uncertainty. The results indicate that even with an approximate geological prior model, a high degree of parametrization and history matching can lead to parameter ensembles that can be useful for making certain predictions (e.g., heads/concentrations). However, an approximate geological prior may be inadequate for more demanding predictions (e.g., mass).

Uncertainty in predicting solute transport can arise from an imperfect understanding of model parametrization (e.g., heterogeneity), initial aquifer conditions, boundary

stresses, solute sources, and release history. This paper only focuses on the predictive uncertainty arising from the possible variations in parametrization. However, the uncertainty in initial/boundary conditions could also be incorporated into the uncertainty framework based on "soft" information like the site history and climatic variations.

Previous attempts by the authors at solving the categorical inverse problem with only head measurements (Khambhammettu et al. 2020) produced parameter ensembles that had much more variability than the ensembles described in this paper. By incorporating concentration data which are directly affected by permeability connections and contrasts in the subsurface, the posterior parameter uncertainty was decreased.

Despite using many parameters for the continuous case, the resulting hydraulic conductivity distribution(s) did not have the connectedness of the reference model's hydraulic conductivity distribution. This discrepancy arises from representing the non-Gaussian variance of hydraulic conductivity with a multi-Gaussian assumption (variogram). For example, Gómez-Hernández and Wen 1998, showed for a hypothetical site that using a multi-Gaussian model resulted in over-estimating source-to-receptor travel times because of incomplete permeability connections. Other researchers (Zinn

and Harvey 2003; Renard and Allard 2013) also convey similar warnings.

The observations related to the role of the number and location of monitoring data in sampling the posterior and quantifying model uncertainty are problem specific. In real-world scenarios, the geological prior, contamination source locations/release history, are unknown/uncertain, and additional data may more often help better quantify the predictive uncertainty.

The synthetic problem analyzed in this paper is simple—three channels of uniform width; two-dimensional groundwater flow and solute transport, etc. However, utility software like PLPROC (Doherty 2023) based on the TRIPS framework can be used to generate complex meandering and intersecting channels of nonuniform width and thickness in three dimensions to represent real world conditions by associating a width and thickness parameter to each TRIPS along the channel. The prior statistics on the number of channels can be derived from a set of initial simulations that are conditioned to site-specific geologic knowledge.

Despite its simplicity, the synthetic problem analyzed in this paper has several similarities to real world sites, namely facies with high and low permeability, decade-long solute transit times, and a partial understanding of the subsurface. Therefore, the approach presented here could be adopted at large contaminated sites where the cost of an incorrect prediction far outweighs the costs associated with sampling the posterior and generating an ensemble of likely models.

Site managers, regulators, stakeholders, and decision-makers can thus make informed decisions by evaluating the spectrum of likely predictions rather than relying on a single prediction. The framework should include periodic data collection efforts, which could then be used to update the predictions. Data worth analyses (Dausman et al. 2010; Wöhling et al. 2016) which have been shown to reduce predictive uncertainty, could be used to decide the locations of future monitoring wells. A pragmatic framework that integrates data acquisition efforts with an ensemble-based approach could lead to better future outcomes.

Acknowledgments

Authors' Note

The authors do not have any conflicts of interest or financial disclosures to report.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article. Supporting Information is generally *not* peer reviewed.

Figure S1. Training Image representative of the geology in a synthetic aquifer.

Figure S2. Development of object-based model from pilot points with (top) initial position of TRIPS and (bottom) updated positions of TRIPS.

Figure S3. Variogram of log-transformed hydraulic conductivity distribution in the reference model.

Figure S4. Ensemble Spread vs. mean objective function for the continuous parameterization case.

Figure S5. Ensemble Spread vs. mean objective function for the discrete parameterization case.

Figure S6. Measured vs. Simulated Heads for the continuous parameterization case

Figure S7. Measured vs. Simulated Heads for the discrete parameterization case.

Figure S8. Measured vs. Simulated concentrations (prior ensemble) for the continuous parameterization case.

Figure S9. Measured vs. Simulated concentrations (posterior ensemble for the sparse monitoring scenario) for the continuous parameterization case.

Figure S10. Measured vs. Simulated concentrations (posterior ensemble for the dense monitoring scenario) for the continuous parameterization case.

Figure S11. Measured vs. Simulated concentrations (prior ensemble) for the discrete parameterization case.

Figure S12. Measured vs. Simulated concentrations (posterior ensemble for the sparse monitoring scenario) for the discrete parameterization case.

Figure S13. Measured vs. Simulated concentrations (posterior ensemble for the dense monitoring scenario) for the discrete parameterization case.

References

Ahmmed, B., S.C. James, and J. Yelderman. 2020. Post-calibration uncertainty analysis for travel times at a naval weapons industrial reserve plant. *Water* 12, no. 12: 3428. https://doi.org/10.3390/w12123428

Aster, R.C., B. Borchers, and C.H. Thurber. 2013. *Parameter Estimation and Inverse Problems*, 2nd ed. Waltham, MA, USA: Academic Press.

Caers, J., and T. Hoffman. 2006. The probability perturbation method: A new look at Bayesian inverse modeling. *Mathematical Geology* 38, no. 1: 81–100. https://doi.org/10.1007/s11004-005-9005-9

Cao, Z., L. Li, and K. Chen. 2018. Bridging iterative ensemble smoother and multiple-point geostatistics for better flow and transport modeling. *Journal of Hydrology* 565: 411–421. https://doi.org/10.1016/j.jhydrol.2018.08.023

Chapman, S.W., B.L. Parker, T.C. Sale, and L.A. Doner. 2012. Testing high resolution numerical models for analysis of contaminant storage and release from low permeability zones. *Journal of Contaminant Hydrology* 136: 106–116. https://doi.org/10.1016/j.jconhyd.2012.04.006

Chapman, S.W., and B.L. Parker. 2005. Plume persistence due to aquitard back diffusion following dense non-aqueous phase liquid source removal or isolation. *Water Resources Research* 41, no. 12: 1–16. https://doi.org/10.1029/2005WR004224

Chen, Y., and D.S. Oliver. 2013. Levenberg-Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences* 17: 689–703. https://doi.org/10.1007/s10596-013-9351-5

Coats, K.H., and B.D. Smith. 1964. Dead-end pore volume and dispersion in porous media. *Society of Petroleum Engineers Journal* 4, no. 1: 73–84. https://doi.org/10.2118/647-PA

Dausman, A.M., J. Doherty, C.D. Langevin, and M.C. Sukop. 2010. Quantifying data worth toward reducing predictive uncertainty. *Ground Water* 48, no. 5: 729–740. https://doi.org/10.1111/j.1745-6584.2010.00679.x

Deans, H.A. 1963. A mathematical model for dispersion in the direction of flow in porous media. *Society of Petroleum Engineers Journal* 3, no. 71: 49–52. https://doi.org/10.2118/493-pa

Deutsch, C.V., and A.G. Journel. 1998. *GSLIB: Geostatistical Software Library and User's Guide*, 6th ed. New York: Oxford University Press.

Doherty, J.E. 2023. *PLPROC A Parameter List Processor*. Brisbane: Watermark Numerical Computing and Groundwater Model Decision Support Initiative (GMDSI).

Doherty, J.E. 2018. *PEST Model-Independent Parameter Estimation*, 7th ed. Brisbane: Watermark Numerical Computing.

Emerick, A.A., and A.C. Reynolds. 2013. Ensemble smoother with multiple data assimilation. *Computers & Geosciences* 55: 3–15. https://doi.org/10.1016/j.cageo.2012.03.011

Farhat, S.K., D.T. Adamson, A.R. Gavaskar, S.A. Lee, R.W. Falta, and C.J. Newell. 2020. Vertical discretization impact in numerical modeling of matrix diffusion in contaminated groundwater. *Groundwater Monitoring and Remediation* 40, no. 2: 52–64. https://doi.org/10.1111/gwmr.12373

Fiori, A., V. Cvetkovic, G. Dagan, S. Attinger, A. Bellin, P. Dietrich, A. Zech, and G. Teutsch. 2016. Debates—Stochastic subsurface hydrology from theory to practice: The relevance of stochastic subsurface hydrology to practical problems of contaminant transport and remediation. What is characterization and stochastic theory good for? *Water Resources Research* 52, no. 12: 9228–9234. https://doi.org/10.1002/2015WR017525

Gómez-Hernández, J., and X.H. Wen. 1998. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. *Advances in Water Resources* 21, no. 1: 47–61. https://doi.org/10.1016/S0309-17089600031-0

Hanke, M. 1997. A regularizing Levenberg – Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems* 13: 79–95. https://doi.org/10.1088/0266-5611/13/1/007

Hu, L.Y., G. Blanc, and B. Noetinger. 2001. Gradual deformation and iterative calibration of sequential stochastic simulations. *Mathematical Geology* 33, no. 4: 475–489. https://doi.org/10.1023/a:1011088913233

Khambhammettu, P., P. Renard, and J. Doherty. 2020. The traveling pilot point method. A novel approach to parameterize the inverse problem for categorical fields. *Advances in Water Resources* 138: 103556. https://doi.org/10.1016/j.advwatres.2020.103556

Knowling, M.J., J.T. White, and C.R. Moore. 2019. Role of model parameterization in risk-based decision support: An empirical exploration. *Advances in Water Resources* 128: 59–73. https://doi.org/10.1016/j.advwatres.2019.04.010

Konikow, L.F. 2011. The secret to successful solute-transport modeling. *Ground Water* 49, no. 2: 144–159. https://doi.org/10.1111/j.1745-6584.2010.00764.x

Lam, D.-T., P. Renard, J. Straubhaar, and J. Kerrou. 2020. Multiresolution approach to condition categorical multiple-point realizations to dynamic data with iterative ensemble smoothing. *Water Resources Research* 56, no. 2: e2019WR025875. https://doi.org/10.1029/2019WR025875

Langevin, C.D., J.D. Hughes, E.R. Banta, A.M. Provost, R.G. Niswonger, and S. Panday. 2021. *MODFLOW 6 Modular Hydrologic Model version 6.2.1: U.S. Geological Survey Software Release*. https://doi.org/10.5066/F76Q1VQV.

van Leeuwen, P.J., and G. Evensen. 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review* 124, no. 12: 2898–2913. https://doi.org/10.1175/1520-0493(1996)124<2898:DAAIMI>2.0.CO;2

Li, L., S. Srinivasan, H. Zhou, and J.J. Gómez-Hernández. 2013. A pilot point guided pattern matching approach to integrate dynamic data into geological modeling. *Advances in Water Resources* 62: 125–138. https://doi.org/10.1016/j.advwatres.2013.10.008

Linde, N., P. Renard, T. Mukerji, and J. Caers. 2017. Geological realism in hydrogeological and geophysical inverse modeling: A review. *Advances in Water Resources* 86: 86–101. https://doi.org/10.1016/j.advwatres.2015.09.019

Mariethoz, G., P. Renard, and J. Caers. 2010. Bayesian inverse problem and optimization with iterative spatial resampling. *Water Resources Research* 46, no. 11: 1–17. https://doi.org/10.1029/2010WR009274

Moore, C., and J. Doherty. 2005. Role of the calibration process in reducing model predictive error. *Water Resources Research* 41, no. 5: 1–14. https://doi.org/10.1029/2004WR003501

National Research Council. 2013. *Alternatives for Managing the Nation's Complex Contaminated Groundwater Sites*, 1–407. Washington, DC: The National Academies Press. https://doi.org/10.17226/14668

Payne, F.C., J.A. Quinnan, and S.T. Potter. 2008. *Remediation Hydraulics*, 1st ed., 439. Boca Raton, Florida: CRC Press.

Rajaram, H. 2016. Debates—Stochastic subsurface hydrology from theory to practice: Introduction. *Water Resources Research* 52, no. 12: 9215–9217. https://doi.org/10.1002/2016WR020066

Renard, P., and D. Allard. 2013. Connectivity metrics for subsurface flow and transport. *Advances in Water Resources* 51: 168–196. https://doi.org/10.1016/j.advwatres.2011.12.001

Sale, T. C., B.L. Parker, C.J. Newell, and J.F. Devlin. 2013. STATE-OF-THE-SCIENCE REVIEW Management of Contaminants Stored in Low Permeability Zones SERDP Project ER-1740. SERDP/ESTCP, 348. ISBN: 1120091020. Accessed January 15, 2021 https://www.serdp-estcp.org/Program-Areas/Environmental-Restoration/Contaminated-Groundwater/Persistent-Contamination/ER-1740/ER-1740-TR.

Schilling, O.S., P.G. Cook, and P. Brunner. 2019. Beyond classical observations in hydrogeology: The advantages of including exchange flux, temperature, tracer concentration, residence time, and soil moisture observations in groundwater model calibration. *Reviews of Geophysics* 57, no. 1: 146–182. https://doi.org/10.1029/2018RG000619

Skjervheim, J. A., and G. Evensen. 2011. An ensemble smoother for assisted history matching. Society of Petroleum Engineers – SPE Reservoir Simulation Symposium 2011. https://doi.org/10.2118/141929-ms.

Tarantola, A. 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, 358. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics.

USEPA. 2004. Cleaning up the nation's waste sites: Markets and technology trends. EPA 542-R-04-015, Washington, DC: United States Environmental Protection Agency, Office of Solid Waste and Emergency Response.

Van Genuchten, M.T., J.M. Davidson, and P.J. Wierenga. 1974. An evaluation of kinetic and equilibrium equations for the prediction of pesticide movement through porous media. *Soil Science Society of America Journal* 38, no. 1: 29–35. https://doi.org/10.2136/sssaj1974.03615995003800010016x

White, J.T. 2018. A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions. *Environmental Modelling and Software* 109: 191–201. https://doi.org/10.1016/j.envsoft.2018.06.009

White, J.T., M.J. Knowling, and C.R. Moore. 2020a. Consequences of groundwater-model vertical discretization in risk-based decision-making. *Groundwater* 58, no. 5: 695–709. https://doi.org/10.1111/gwat.12957

White, J.T., R.J. Hunt, M.N. Fienen, and J.E. Doherty. 2020b. PEST++, a software suite for parameter estimation, uncertainty analysis, management optimization and sensitivity analysis. *U.S. Geological Survey Techniques and Methods* 7C26: 52. https://doi.org/10.3133/tm7C26

Wöhling, T., A. Geiges, and W. Nowak. 2016. Optimal design of multitype groundwater monitoring networks using easily accessible tools. *Groundwater* 54, no. 6: 861–870. https://doi.org/10.1111/gwat.12430

Zheng, C., and S.M. Gorelick. 2003. Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale. *Ground Water* 41, no. 2: 142–155. https://doi.org/10.1111/j.1745-6584.2003.tb02578.x

Zhou, H., J. Gómez-Hernández, and L. Li. 2014. Inverse methods in hydrogeology: Evolution and recent trends. *Advances in Water Resources* 63: 22–37. https://doi.org/10.1016/j.advwatres.2013.10.014

Zinn, B., and C.F. Harvey. 2003. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. *Water Resources Research* 39, no. 3: 1–19. https://doi.org/10.1029/2001WR001146