Analysis of hydrogeological structure uncertainty by estimation of hydrogeological acceptance probability of geostatistical models

Dylan R. Harp *, Velimir V. Vesselinov

Los Alamos National Laboratory, MS T003, Los Alamos, NM 87544, USA

ABSTRACT

The following describes a proposed approach to account for the equifinality of solutions that result from comparing observations to flow simulations when using realizations of geostatistical models. We introduce hydrogeological acceptance probability to estimate the propensity of a geostatistical model to produce acceptable realizations with respect to the consistency of their simulations with observations. The estimation of hydrogeological acceptance probability is equivalent to the calculation of the sample mean of a Bernoulli distribution. This allows the estimation of the acceptance probability to be preemptively terminated based on the current estimate and subject to the desired confidence level and interval length. We propose a composite uncertainty analysis of the hydrogeological heterogeneity utilizing acceptable realizations from multiple geostatistical models collected during the estimation of their acceptance probability. In the case of a non-fuzzy definition of realization acceptance, this produces a facies probability map. If the definition of realization acceptance is imprecise, the analysis yields upper and lower bounds on the facies probability map in the form of facies plausibility and belief maps, respectively. These maps can provide indications of the information content of the data and provide guidance for the collection of additional data.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The collection of a sufficient number of direct observations to adequately characterize hydrogeological heterogeneity is a goal that is approached asymptotically at best. Use of limited observations is then required to determine where additional information should be collected, or to support decisions that cannot wait for additional data collection. This has led to approaches to infer hydrogeological heterogeneity on the basis of comparisons between output of flow and transport simulators and available observations. Typically, the only available information regarding the hydrogeological heterogeneity of a site are sparsely located direct point estimates and indirect (e.g. water level and/or contaminant concentration) observations. Considering reducible (epistemic) and irreducible (aleatoric) uncertainty [1], multiple models of heterogeneity will produce acceptable simulations of the system when compared to the available observations [2]. This situation has been called the equifinality of solutions, and can be due to multiple discrete acceptable solutions and/or a region of acceptable solutions (i.e. an area of attraction in a response surface with an ill-defined optimal solution). It is often the case that the prior information is not sufficient to reduce the set of solutions to a single solution, especially in the case of stochastic models, where each parameter set characterizes an infinite set of solutions.

Well known approaches to account for the equifinality of solutions (i.e. multiple equally acceptable simulations of a system [3]) include maximum likelihood Bayesian model averaging (MLBMA) [4], generalized likelihood uncertainty estimation (GLUE) [2], and the Bayesian approach of Gaganis and Smith [5]. MLBMA provides a relatively efficient approach to rank calibrated models; however, acceptable sub-optimal solutions may be ignored. GLUE provides a general framework for evaluating model uncertainty given a Monte Carlo sampling and parameters of a deterministic model (e.g. hydraulic conductivities of a distributed model). The Bayesian approach of Gaganis and Smith [5], also requiring a Monte Carlo sampling, provides a formal Bayesian approach to estimate model uncertainty. An approach by Rojas et al. [6] combines GLUE and MLBMA, utilizing the statistical approach of MLBMA without ignoring suboptimal solutions.

This paper presents an approach to evaluate hydrogeological heterogeneity uncertainty, including imprecision in the definition of an acceptable solution, utilizing a preemptive sampling scheme based on the estimation of hydrogeological acceptance probability. The acceptance probability has some similarities to the posterior probabilistic weight associated with Bayesian analyses, except that it is designed to evaluate the propensity of stochastic models to generate realizations consistent with observations, and in the

---

current methodology, does not intend to provide probabilistic evidence that a stochastic model represents the 'truth'. In the current methodology, its function is to define the termination of a geostatistical sampling based on the desired confidence and interval length. Except for some philosophical similarities to approaches developed in hydrology to deal with equifinality (e.g. [2]), the approach is unrelated and developed specifically for stochastic models.

For clarity, we will refer to a geostatistical functional form as a geostatistical framework, while a particular instance of a geostatistical framework, defined by the specification of its parameters, as a geostatistical model. We will refer to a particular instance of a geostatistical model, usually specified by the designation of a random seed, as a realization of the geostatistical model. A geostatistical model defines an infinite set of equally-probable realizations with respect to the statistical characteristics of the spatial variation (geostatistically-equally-probable). Realizations can be conditioned to honor values at particular locations while maintaining the global statistics of the field by using conditional simulation [7]. Various geostatistical frameworks have been developed to characterize and evaluate the structure of aquifer heterogeneity, including variogram [7], multiple-point [8], and Markov-chain geostatistics [9]. While alternatives to geostatistical modeling of heterogeneity exist as well [10–14], we explore a geostatistical approach here. We use a Markov-chain geostatistical model conditioning the hydraulic conductivity heterogeneity at the observation wells. These conditioning data could be available from pumping tests. The use of a Markov-chain geostatistical model is merely a selection, and is not required by the method, which could incorporate results from many models of heterogeneity in a single analysis.

Many of the current geostatistical inverse approaches utilized in hydrogeological studies explore uncertainty by evaluating residuals between simulated and observed values (e.g. water levels, contaminant concentrations) produced using geostatistical realizations from a single geostatistical framework that is inferred from the observed or assumed hydrogeology of a site (e.g. [15–19]). These approaches provide unrealistic estimates of uncertainty in cases where an appropriate statistical model of the heterogeneity is uncertain, a common scenario in hydrogeological studies with sparse data. Pardo-Igùzquiza et al. [20] quantify uncertainty in estimating semivariogram parameters from empirical semivariograms using a maximum likelihood estimation, and demonstrate how this uncertainty can be propagated to predictions of head. Nowak et al. [21] developed an approach for geostatistical design that they refer to as continuous Bayesian model averaging considering uncertainty in parameters of the Matèrn family of covariance functions. Other approaches evaluate the consistency of geostatistical models (conceptual model of the statistical properties of the heterogeneity) with observations in a statistical sense [22–24]. These approaches assume that the statistical characteristics of the 'true' heterogeneity can be inferred by the observed hydrogeology (hydrogeological observations from a single realization of the 'true' geostatistical model).

In general, water pressure residuals will vary substantially between realizations from the same geostatistical model, indicating that geostatistically-equally-probable realizations are not necessarily equally-consistent with respect to hydraulic observations. Conversely, it is also possible that realizations from various alternative geostatistical models will produce similar hydraulic responses. Therefore, evaluating modeling residuals in the unavoidable presence of uncertainty means that multiple geostatistical models will produce realizations resulting in acceptable simulations of the system given the available observations. As a result, the identification of the geostatistical model that is the most consistent with observations in a statistical sense does not neces-

sarily ensure that the true geostatistical model has been identified. This is an example of asking a question that cannot be answered with the information provided. The question that can be answered is what is the propensity of a geostatistical model to produce realizations resulting in simulated values consistent with observations given the ambient uncertainty. In the process of answering this question, a set of acceptable realizations can be collected. Combining acceptable sets of realizations from multiple models provides a basis for a composite uncertainty analysis, incorporating acceptable features from multiple models. This provides the possibility that the uncertainty analysis may uncover features not present in the individual models, but present in the composite uncertainty analysis.

One can evaluate stochastic realizations by how consistent their simulated values are to observed values. Researchers have proposed many functional forms to accomplish this [2], with perhaps the most common being the sum of squared residuals (SSR). Using a ranking suited to a particular application, criteria for the acceptibility of a realization can be established. A higher acceptance probability of a geostatistical model indicates its propensity to produce realizations resulting in simulated water levels (heads) consistent with observations accounting for measurement and conceptual uncertainty.

While the propensity of a geostatistical model to produce acceptable realizations does provide some indication that the model defines the characteristics of the 'true' heterogeneity, choosing the geostatistical model with the highest acceptance probability as the 'truth' would be a naive assumption. Basing an uncertainty analysis on such an assumption is ill-advised as the potential that the 'true' heterogeneity is most similar to an extreme realization of a geostatistical model with low acceptance probability always exists. Therefore, we advocate developing a composite uncertainty analysis based on acceptable realizations from multiple models, potentially from multiple geostatistical frameworks.

The fact that multiple geostatistical models can produce realizations with indistinguishable acceptability indicates that an uncertainty analysis must consider more than just the geostatistical model with the highest acceptance probability. A composite uncertainty analysis would require that all possible geostatistical models be evaluated. Of course, in practice this is typically not feasible. Therefore, such uncertainty analyses are relative, conditional on the considered geostatistical frameworks. We propose that the relative uncertainty associated with hydrogeological heterogeneity can be evaluated by computing the acceptance probability of a sampling of geostatistical models from one or several geostatistical frameworks that incorporate all acceptable realizations. Such an analysis can be utilized to produce facies probability maps of hydrogeological properties, providing indications of the information content of the data and providing guidance for siting further investigations.

As the distinction between an acceptable and unacceptable realization with respect to epistemic uncertainty will undoubtedly be imprecise, the use of fuzzy sets [25,26] is appropriate. Fuzzy set theory allows the assignment of grades of membership to elements and use of fuzzy set operations in place of classical set operations [27]. Therefore, fuzzy sets allow an expert to assign grades of membership to the acceptable and unacceptable sets of realizations. Fuzzy sets can also be constructed by integrating information from multiple experts [25]. Previous approaches using fuzzy sets in a hydrological context include mapping flood propagation model outputs to likelihoods in order to evaluate roughness coefficient uncertainty within the GLUE framework [28]; mapping imprecise remotely sensed observations of energy partition variability into model functional types of evapotranspiration [29]; approximation of the unsaturated Darcy law using a fuzzy rule-based model [30]; and a fuzzy least-squares regression approach to the Cooper–Jacob method [31].

Considering a dual relationship between the membership functions for the acceptable and unacceptable sets ($\mu_A = 1 - \mu_U$, $\mu \in [0,1]$, where μ_A and μ_U are the membership functions for the acceptable and unacceptable sets, respectively), implies that there is a lack of conflict in the information. As a result, it is possible to estimate the acceptance probability of a geostatistical model while using an imprecise definition of acceptance.

An uncertainty analysis of the hydrogeological heterogeneity in the case of an imprecise definition of acceptibility will produce upper and lower bounds on the facies probability map. These upper and lower probability bounds are consistent with Dempster–Shafer theory (DST) [32], and are referred to as plausibility and belief measures, respectively. The machinery of interest from DST for our methodology is the construct of an upper and lower probability, which can be attributed to Dempster [33]. Recently, Mathon et al. [34] demonstrated the use of DST to account for epistemic and aleatoric uncertainty of permeability measurements. In the limiting case where the fuzzy membership function becomes an indicator function, the gap between the belief and plausibility measures, defining the span of possible probability distributions, collapses to a single probability distribution.

This paper proposes the concept of hydrogeological acceptance probability of a geostatistical model to allow for an imprecise definition of realization acceptance, and demonstrates its ability to characterize the hydraulic conductivity heterogeneity and associated uncertainty. The following sections demonstrate the estimation of acceptance probability, mapping of acceptance probabilities over a geostatistical parameter space, and the mapping of facies probabilities, beliefs and plausibilities based on discrete samplings of geostatistical parameters (models).

2. Estimating hydrogeologic acceptance probability

The concept of the hydrogeological acceptance probability of a geostatistical model incorporates aleatoric and epistemic uncertainty. In the present case, the source of aleatoric uncertainty (i.e. uncertainty dependent on a random event and irreducible without the improvement of measuring techniques and/or collection of new measurements [1]) is the measurement of the water levels. The source of epistemic uncertainty (i.e. uncertainty related to the lack of information about the system) is in the definition of the conceptual model (e.g. restricted to the hydrogeological structure here). In the following, we define criteria for determining the membership of a realization in the set of acceptable realizations. It should be realized that other criteria could be defined for other applications to fit a particular scenario.

We define an acceptance metric (Φ) as

$$\Phi(\boldsymbol{\theta}) = \sum_{i=1}^{N_o} \begin{cases} (\bar{h}_{i,max} - \hat{h}_i(\bar{\theta}))^2 & \text{if } \hat{h}_i(\boldsymbol{\theta}) > \bar{h}_{i,max}, \\ (\bar{h}_{i,min} - \hat{h}_i(\boldsymbol{\theta}))^2 & \text{if } \hat{h}_i(\boldsymbol{\theta}) < \bar{h}_{i,min}, \\ 0 & \text{if } \bar{h}_{i,min} \leqslant \hat{h}_i(\boldsymbol{\theta}) \leqslant \bar{h}_{i,max}, \end{cases} \quad (1)$$

where $\boldsymbol{\theta}$ is a vector of parameters defining the realization (geostatistical model parameters and random seed), N_o is the number of observations, $\hat{h}_i(\boldsymbol{\theta})$ is the ith simulated value given $\boldsymbol{\theta}$, and $\bar{h}_{i,max}$ and $\bar{h}_{i,min}$ define an interval of values for the ith simulated value consistent with measurement uncertainty determined by the resolution and precision of the measuring instruments (e.g. pressure transducers). While methods to model measurement error using an unknown statistical parameter have been developed for various inverse approaches, the approach in Eq. (1) is deemed appropriate for this sampling scheme. For the remainder of the paper, the argument of $\Phi(\boldsymbol{\theta})$ will be dropped for notational convenience.

Accounting for epistemic uncertainty requires expert opinion of values considered to be consistent with observations given uncertainty in the conceptual model. In general, the fuzzy membership function defining membership in the set of acceptable realizations μ_A can be expressed as

$$\mu_A(\Phi) = \begin{cases} 1 & \text{if } \Phi \leqslant \Phi_1, \\ 0 & \text{if } \Phi \geqslant \Phi_2, \\ f(\Phi) & \text{if } \Phi_1 < \Phi < \Phi_2 \quad \text{where } f(\Phi) : \Phi \to (0,1), \end{cases} \quad (2)$$

where Φ_1 and Φ_2 define the limits of maximum and minimum membership, respectively, in the acceptable set of realizations. $f(\Phi)$ defines the grades of membership between the maximum and minimum and is required to decrease monotonically. For demonstration purposes, $f(\Phi)$ is assumed to be a linear function here, defined as

$$f(\Phi) = \frac{\Phi_2 - \Phi}{\Phi_2 - \Phi_1}, \quad \Phi_1 < \Phi < \Phi_2. \quad (3)$$

Other possible functional forms for $f(\Phi)$ include Gaussian, exponential, and step functions [25]. In the special case where $\Phi_1 = \Phi_2$, Eq. (2) reduces to an indicator function $\chi_A(\Phi)$ as

$$\chi_A(\Phi) = \begin{cases} 1 & \text{if } \Phi \leqslant \Phi_c, \\ 0 & \text{if } \Phi > \Phi_c, \end{cases} \quad (4)$$

producing sets consistent with classical set theory [26], where $\Phi_c = \Phi_1 = \Phi_2$ is used to denote the cutoff of acceptibility.

Selection of Φ_c or Φ_1, Φ_2, and $f(\Phi)$ will require the value judgment of an expert and may entail some investigation. This process is analogous to the selection of a likelihood function in a Bayesian analyses, however, with less restrictions and assumptions.

Considering a non-fuzzy definition of an acceptable realization (Eq. (4)), the acceptance probability p_A of a geostatistical model can be estimated as the sample mean of a Bernoulli distribution as

$$\hat{p}_A = \frac{1}{N_r} \sum_{i=1}^{N_r} \chi_A(\Phi_i), \quad (5)$$

where N_r is the total number of realizations and $\Phi_i = \Phi(\theta_i)$. Assuming that membership in the set of unacceptable realizations U, denoted by μ_U, has a dual relationship to μ_A (i.e. $\mu_A = 1 - \mu_U$), indicating that the information is not conflicting [26] (see Fig. 1), a similar equation can be used in the fuzzy case as

$$\hat{p}_A = \frac{1}{N_r} \sum_{i=1}^{N_r} \mu_A(\Phi_i). \quad (6)$$

Considering Eq. (6), it can be demonstrated that the information is not conflicting by verifying that $\hat{p}_A + \hat{p}_U = 1 (\hat{p}_U = 1/n \sum_{i=1}^{n} \mu_U(\Phi_i))$, which is consistent with the axioms of probability theory.

The two-sided confidence interval length L for the probability of a discrete random variable can be approximated using the current estimate as

$$L = 2z_{\alpha/2} \sqrt{\frac{\hat{p}_A(1 - \hat{p}_A)}{n}}, \quad (7)$$
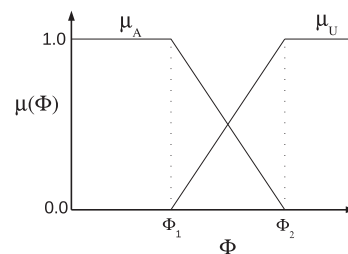

Fig. 1. Membership functions for the set of acceptable (μ_A) and unacceptable (μ_U) realizations.

where $z_{\alpha/2}$ is the z-score at $\alpha/2$ where $(1 - \alpha)100\%$ is the confidence level [35, page 485]. By inspecting Eq. (7), it is apparent that L depends on the current value of $\hat{p}_A$. An equation defining the required sample size to estimate $\hat{p}_A$ within L at the $(1 - \alpha)100\%$ confidence level can be determined by solving Eq. (7) for n as

$$n = \frac{4z_{\alpha/2}^2 \hat{p}_A(1 - \hat{p}_A)}{L^2}. \tag{8}$$

By considering Eq. (8) and inspecting Fig. 2, it is apparent that the largest number of samples will be required when $\hat{p}_A = 0.5$. Therefore, estimation of small or large values of $\hat{p}_A$ will be less computationally intensive and/or time consuming at a particular confidence and interval length. This property is used to preemptively terminate the estimation of an acceptance probability at a specified confidence level $((1 - \alpha)100)\%$ and interval length (L) based on the current value of $\hat{p}_A$. This provides a significant reduction in forward model runs required for the estimation of high or low acceptance probabilities. Values of $\hat{p}_A$ provide a statistic describing the propensity of geostatistical models to produce acceptable realizations.

Performing a multi-realization sampling to estimate acceptance probability is distinct from performing a multi-realization sampling with the intent to estimate convergent statistics of state variables, and does not imply that the statistical moments of state variables will be converged at a specified confidence and interval length of the acceptance probability estimate.

3. Mapping hydrogeological uncertainty using a set of acceptable realizations

The estimation of the acceptance probability of a geostatistical model produces a set of realizations deemed acceptably consistent with observations. The hydrogeological uncertainty within the context of this geostatistical model can be evaluated from this set. This analysis of uncertainty can be extended indefinitely by considering the union of acceptable sets from multiple geostatistical models as

$$\bigcup_{i=1}^{N_m} A_i, \tag{9}$$

where N_m is the number of geostatistical models and A_i is the acceptable set of realizations from the ith geostatistical model. The geostatistical models are not limited to a particular geostatistical framework. For instance, it would be possible to combine acceptable sets from Markov-chain [9] and indicator kriging [7] geostatistical frameworks in a single analysis.

In the following, we consider the case of geologic units with distinct hydraulic conductivities. A similar analysis can be performed for continuous hydraulic conductivities using discretized intervals. If A is considered non-fuzzy (Eq. (4)), it is possible to map the probability of a particular geologic unit (facies) by approximating the one-location marginal probabilities at each location as
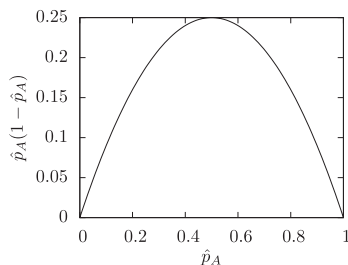

Fig. 2. Values of $\hat{p}_A(1 - \hat{p}_A)$ as a function of $\hat{p}_A$ demonstrating the influence of $\hat{p}_A$ on the calculation of the sample size n (Eq. (8)).

$$\hat{p}(l;\mathbf{x}) = \frac{\sum_{i=1}^{N_r} I_{l,i}(\mathbf{x})\chi_A(\Phi_i)}{\sum_{i=1}^{N_r} \chi_A(\Phi_i)}, \tag{10}$$

where $\hat{p}(l;\mathbf{x})$ denotes the probability of the lth facies at location $\mathbf{x}$, $I_{l,i}(\mathbf{x})$ is the lth facies indicator function at location $\mathbf{x}$ associated with the ith realization, and N_r is the total number of realizations (including acceptable and unacceptable realizations) from all geostatistical models from all geostatistical frameworks included in the analysis. By filtering the realizations to the union of sets of acceptable realizations (Eq. (9)), Eq. (10) can be simplified as

$$\hat{p}(l;\mathbf{x}) = \frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x}), \tag{11}$$

where i is now an index for the collection of acceptable realizations and N_A is the number of realizations in the union of acceptable sets ($N_A = \sum_{i=1}^{N_r} \chi_A(\Phi_i)$). The facies indicator functions are defined as

$$I_{l,i}(\mathbf{x}) = \begin{cases} 1 & \text{if unit } l \text{ occurs at location } \mathbf{x} \text{ in realization } i, \\ 0 & \text{otherwise}. \end{cases} \tag{12}$$

A facies probability map for the lth facies can be generated by mapping $p(l;\mathbf{x})$ for all $\mathbf{x}$.

If A is considered a fuzzy set (Eq. (2)), the calculation of the one-location marginal probability is not valid. In this case, Eq. (11) must be generalized by transferring the information about the acceptance of a realization to a measure that a facies exists at a location as

$$m(l;\mathbf{x}) = \frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x})\mu_A(\Phi_i), \tag{13}$$

where $m(l;\mathbf{x})$ is a basic probability assignment (bpa) consistent with Dempster–Shafer theory (DST) [32]. DST provides a generalization of probability theory, relaxing certain axiomatic constraints to allow for the representation of conflicting information.

According to DST, lower and upper bounds on a cumulative distribution function (CDF) can be defined as belief (Bel) and plausibility (Pl) measures, respectively. The belief and plausibility of an arbitrary event D can be defined a

$$Bel(D) = \sum_{E|E \subseteq D} m(E) \tag{14}$$

and

$$Pl(D) = \sum_{E|D \cap E \neq \varnothing} m(E), \tag{15}$$

respectively. The belief that facies l exists at location $\mathbf{x}$ has been derived here as

$$Bel(l;\mathbf{x}) = m(l;\mathbf{x}) = \frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x})\mu_A(\Phi_i). \tag{16}$$

According to the duality of belief and plausibility measures [26], the plausibility that facies l exists at $\mathbf{x}$ can be defined as

$$Pl(l;\mathbf{x}) = 1 - Bel(\bar{l};\mathbf{x}). \tag{17}$$

It is demonstrated in Appendix A that the derived bpa (Eq. (13)), belief measure (Eq. (16)), and plausibility measure (Eq. (17)) satisfy the axiomatic properties defined by DST.

Mapping Eqs. (16) and (17) for all $\mathbf{x}$ provides upper and lower bounds for the facies probability map in the form of plausibility and belief maps, respectively. The gap between the plausibility and belief measures indicates the range of possible values for the facies probability, indicating the epistemic uncertainty in defining the acceptability of a realization. In the case where $\Phi_1 = \Phi_2$, the gap be-

tween the plausibility and belief collapses to the facies probability map.

As the acceptance information (degree of membership) is applied to the entire realization, and not to individual cells (locations) within the realization, the gap between the belief and plausibility will be constant over the model domain, with its magnitude dependent on the imprecision in the definition of realization acceptance. In other words, the conflict in the information due to imprecision in the definition of the acceptance metric is constant over the model domain ($Pl(l;\mathbf{x}) - Bel(l;\mathbf{x}) = c, c \geqslant 0$).

4. Synthetic case study

The use of acceptance probability estimation is demonstrated on a numerically-generated synthetic pumping test where the 'true' hydrogeological property structure, boundary conditions, and initial conditions are specified. In this way, the feasibility of using acceptance probability estimation to perform an uncertainty analysis of the hydrogeological heterogeneity can be evaluated. The model (presented in Fig. 3) is a 2-dimensional representation of the horizontal plane of an aquifer composed of 2 distinct geologic units (facies) with uniform hydraulic conductivities (i.e. $K_1 = 10^{-2}$ m/s, $K_2 = 10^{-5}$ m/s). The 'true' mapping of the facies is generated as a realization of a Markov-chain geostatistical model [9,36] with facies volumetric proportions of $p_1 = p_2 = 0.5$ and mean facies lengths of $\bar{l}_{1,x} = \bar{l}_{1,y} = 20$ m where the numerical subscripts indicate the facies and letter subscripts indicate direction (refer to [9,36,24] for detailed discussions of Markov-chain geostatistics).

The flow simulations are performed on an orthogonal grid with 1 m spacing (10,000 nodes) using FEHM [37]. The geostatistical grid cells are concurrent with the flow simulation grid cells. An ambient gradient of 0.02 m/m is induced in the positive x-direction by imposing constant head boundaries of 100 m and 98 m at nodes along lines $x = 0$ m and 100 m, respectively. A pumping well is located in the approximate center of the model. The pumping well discharges at a constant rate of 5 kg/s beginning at $t = 0$ s. Fifteen observation wells are located throughout the model domain (black x's in Fig. 3). For demonstration purposes, it is assumed that the hydraulic conductivity is perfectly known at the pumping and observation wells and these values are used as conditioning points for the simulation of the hydraulic conductivity fields (realizations). Head measurements are collected at the pumping and observation wells at 11 times ($t = 0.01, 0.02, 0.04, 0.08, 0.16, 0.32, 0.64, 1.28, 2.56, 5.12$ and 10.0 days), resulting in 176 (16 wells × 11 times) observations. Gaussian noise with a mean of zero and standard deviation of

2.5×10^{-5} m was added to the simulated values for the 'true' case to produce observed values with measurement error.

5. Results and discussion

This paper introduces the application of acceptance probability estimation to perform structural hydrogeological uncertainty analysis with limited pressure and hydrogeologic observations. In order to investigate the acceptance probability of various geostatistical models within a Markov-chain geostatistical framework, discrete geostatistical parameter sets are evaluated, where each parameter set defines a geostatistical model. The parameters are the mean facies lengths in the x and y directions, $\bar{l}_x$ and $\bar{l}_y$ [9], respectively. A matrix of parameter combinations are considered, varying both parameters from 10 to 50 m at 5 m increments (81 parameter sets), where $\bar{l}_x = \bar{l}_y = 20$ m are the parameters used to generate the 'true' realization (Gaussian noise with zero mean and standard deviation equal to 2.5×10^{-5} m is added to the simulated values from the 'true' realization to produce observed values with measurement error). While generation of the 'true' mapping of the facies using the same geostatistical framework as used in the inversion presents an idealized case, in practice, there would be no restriction to the number or type of geostatistical frameworks that could be included in a single analysis (refer to Eq. 9 for details).

It is important to note that these stochastic geostatistical parameters do not define the 'true' spatial distribution of hydraulic conductivity, but do define the geostatistical model used to generate the 'true' realization of hydraulic conductivity. Other sampling schemes could have been chosen (e.g. Monte Carlo, Markov chain Monte Carlo, Latin Hypercube Sampling) in order to search the parameter space more thoroughly. This demonstration uses discrete parameter sets in order to decrease the computational requirements.

The results are presented stepwise to illustrate the accumulation of information through the sampling approach. It includes (1) example residual analyses using realizations from the 'true' geostatistical model (Fig. 4); (2) example calculations of acceptance probability estimation for three geostatistical models (Fig. 5); (3) mapping of estimated acceptance probabilities for all 81 geostatistical models at three interval lengths (Fig. 6) and three realization acceptance cutoff values (Fig. 7); (4) mappings of facies probability from the union of acceptable realizations collected during the estimation of the acceptance probability using non-fuzzy definitions of realization acceptance at three interval lengths (Fig. 8) and at four realization acceptance cutoff values (Fig. 9); and (5) facies plausibility and belief maps generated from the union of acceptable realizations collected during the estimation of acceptance probability using fuzzy definitions of realization acceptance (Fig. 10). In all cases, a value of 0.015 m is added and subtracted from the observed value to obtain values for $\bar{h}_{max}$ and $\bar{h}_{min}$, respectively. These bounds for measurement error are significantly larger than the Gaussian noise added to the 'true' simulated values (zero mean, standard deviation 2.5×10^{-5} m), providing a sufficiently large range of values considered consistent with respect to measurement error.

Fig. 4 presents plots of simulated versus observed heads generated from geostatistical realizations of the 'true' geostatistical model. These plots illustrate the extreme variability possible in the hydraulic response of realizations from a geostatistical model. In order to demonstrate that even realizations of the 'true' geostatistical model have significant hydrogeologic variability, the simulated values are from different realizations generated by the 'true' geostatistical model ($\bar{l}_x = \bar{l}_y = 20$ m). From Fig. 4, it is apparent that as the consistency of the realizations to the observed hydraulic response decreases (Φ increases), initially there is increased scatter both above and below the 1:1 line. However, in the plot in the low-
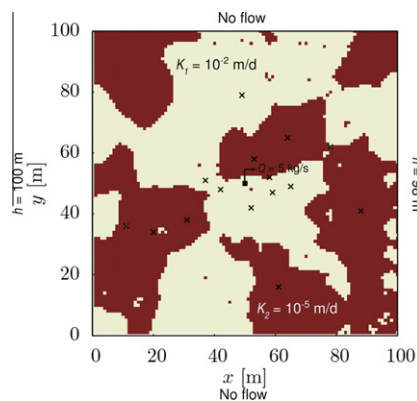

Fig. 3. Map of horizontal plane of synthetic model. Colored regions indicate geologic units with distinct hydraulic conductivities. Black x's indicate locations of water-level monitoring wells (15). The pumping well is indicated by a black square with its associated pumping rate at the approximate center of the model domain. Boundary conditions are noted on the four sides of the map.
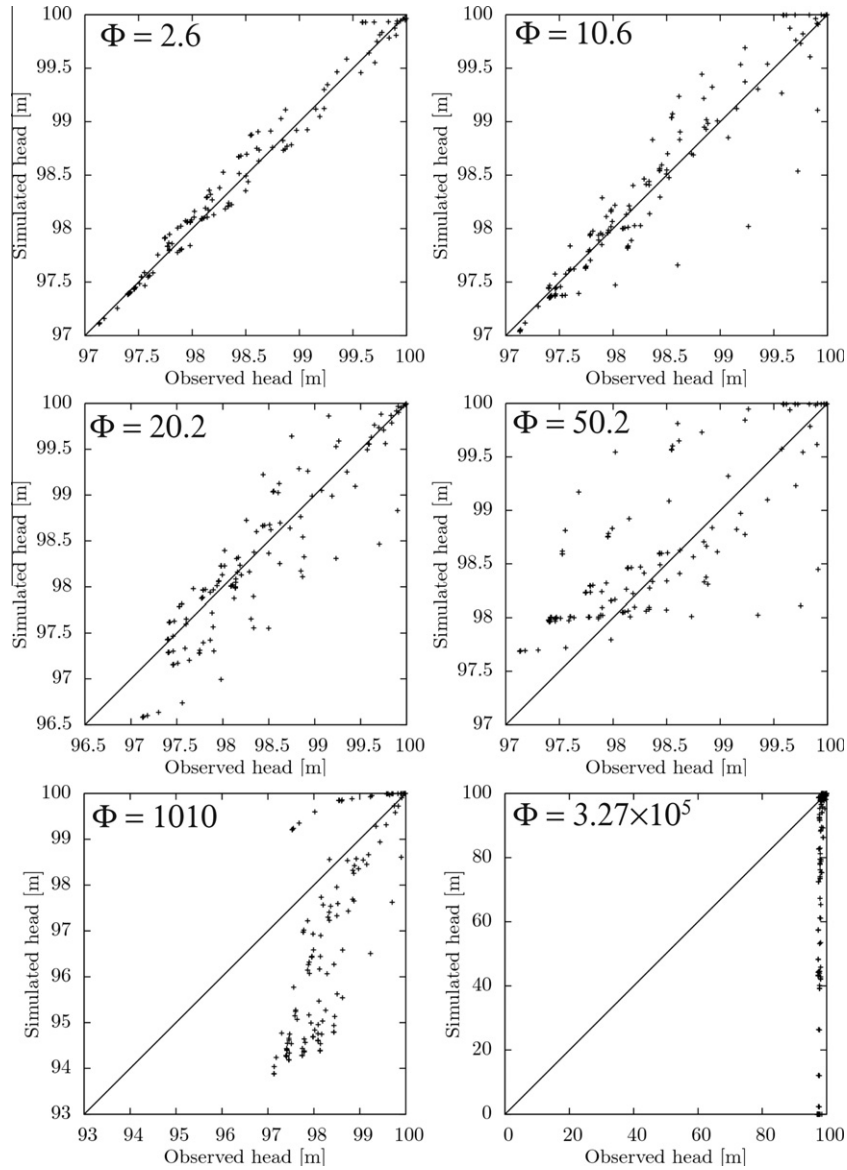
Fig. 4. Plots of simulated heads for realizations from the geostatistical model used to generate the 'true' heterogeneity versus observed heads from the 'true' heterogeneity. Cases with a range of acceptance metric Φ values are presented and noted on each plot. The diagonal line has a 1:1 slope. Note differences in scale between plots.
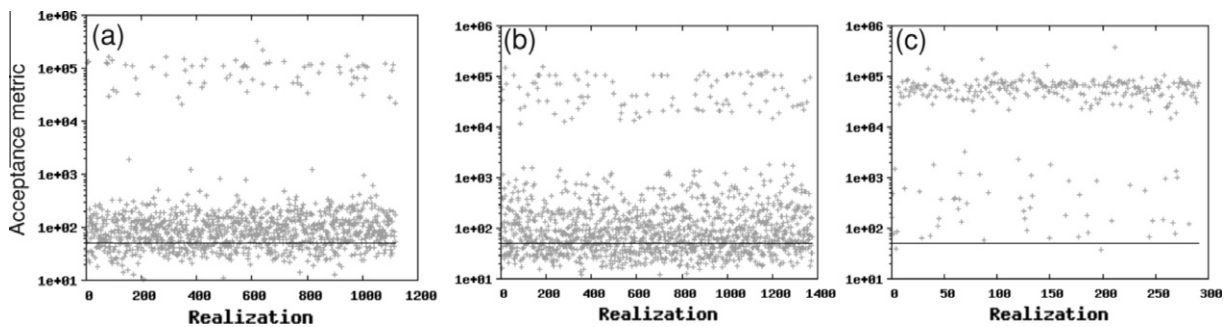

Fig. 5. Acceptance metric values for realizations from three geostatistical models. The horizontal line at $\Phi = 50$ indicates the values of Φ_c used to produce the samples of realizations ($\alpha = 0.05$, $L = 0.05$). The geostatistical models and resulting acceptance probabilities (rounded to their specified precision) are from the geostatistical model (a) used to generate the 'true' realization $\hat{p}_A = 0.25$, (b) with highest acceptance probability $\hat{p}_A = 0.35$, and (c) with lowest acceptance probability $\hat{p}_A = 0.00$.

er left with $\Phi = 1010$, there is a bias for increased simulated drawdown for locations with higher observed drawdown (lower observed head). This is due to changes in heterogeneity patterns at locations near the wells (i.e. simulating low permeability near

monitoring wells with high permeability in the 'true' heterogeneity). While all monitoring wells and the pumping well are conditioned to the 'true' facies, it is possible to produce significantly different heterogeneity near these locations. In the extreme
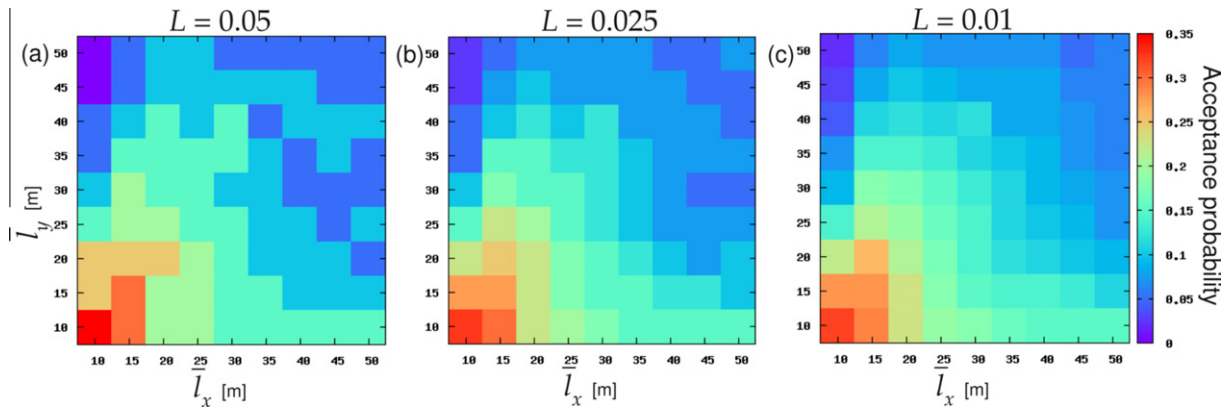
Fig. 6. Maps of hydrogeologic acceptance probability on the parameter space of geostatistical model parameters. Probabilities are estimated at the 95% confidence level for interval lengths L of (a) 0.05, (b) 0.025, and (c) 0.01. Φ_c is 50 for all three cases.
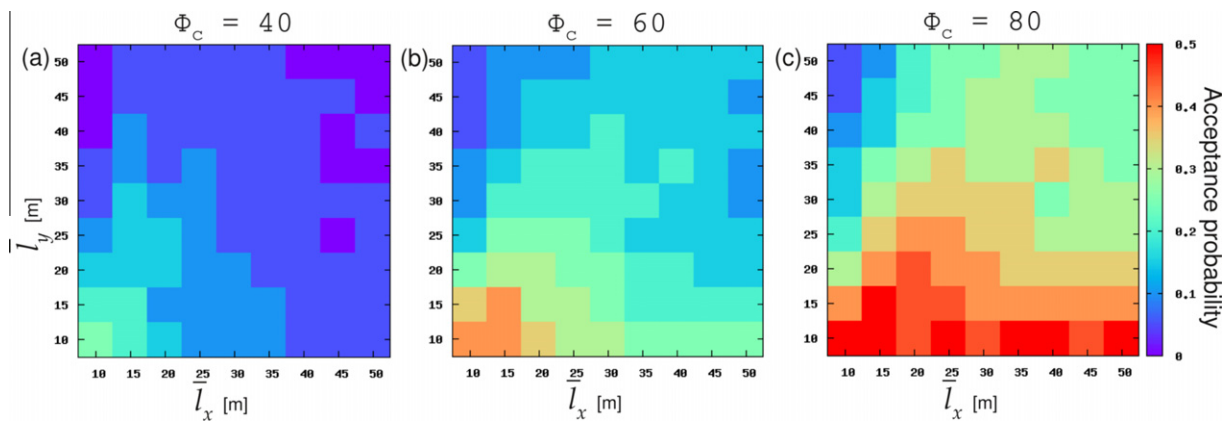

Fig. 7. Maps of hydrogeologic acceptance probability on the parameter space of geostatistical model parameters. The effect of perceived conceptual model uncertainty are demonstrated by presenting maps for Φ_c of (a) 40, (b) 60, and (c) 80. Probabilities are estimated at the 95% confidence level for an interval length L of 0.05.
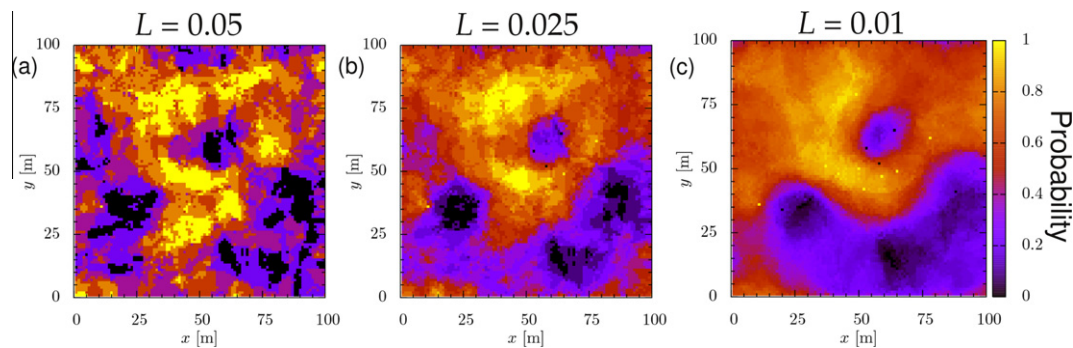

Fig. 8. Facies probability maps for the high conductivity geologic unit ($K_1 = 10^{-2}$ m/s). Acceptance probabilities used to generate these maps were estimated with confidence interval (a) 0.050, (b) 0.025, and (c) 0.010 with 95% confidence and $\Phi_c = 10$.

case ($\Phi = 3.27 \times 10^5$), the heterogeneity is so inconsistent with the observed water levels that the majority of simulated values are less than the observed values (greater simulated drawdown than observed). The plots in Fig. 4 illustrate the analysis of the consistency of flow simulations from a single realization to observations. Such analyses form the basis for the estimation of acceptance probability.

Fig. 5 presents acceptance metric values (Eq. (1)) for realizations from three geostatistical models: (a) the geostatistical model used to generate the 'true' realization ($\bar{l}_x = \bar{l}_y = 20$ m), (b) the geostatistical model with the highest acceptance probability ($\bar{l}_x = 10$ m, $\bar{l}_y = 10$ m), and (c) the lowest acceptance probability

($\bar{l}_x = 10$ m, $\bar{l}_y = 50$ m). Each cross in Fig. 5 signifies the results of one plot from Fig. 4.

For reference, a horizontal line is drawn through $\Phi = 50$, representing an example value for Φ_c. This horizontal line denotes the expert defined cutoff between acceptable and unacceptable realizations, separating the realizations into these two sets. A fuzzy definition of acceptance would require a range of values of Φ where the membership would transition monotonically from one set to the other. The acceptance probability of the geostatistical model is calculated by dividing the number of realizations in the acceptable set (points below Φ_c presented in Fig. 5) by the total number of realizations. The acceptance probabilities are (a) 0.25,
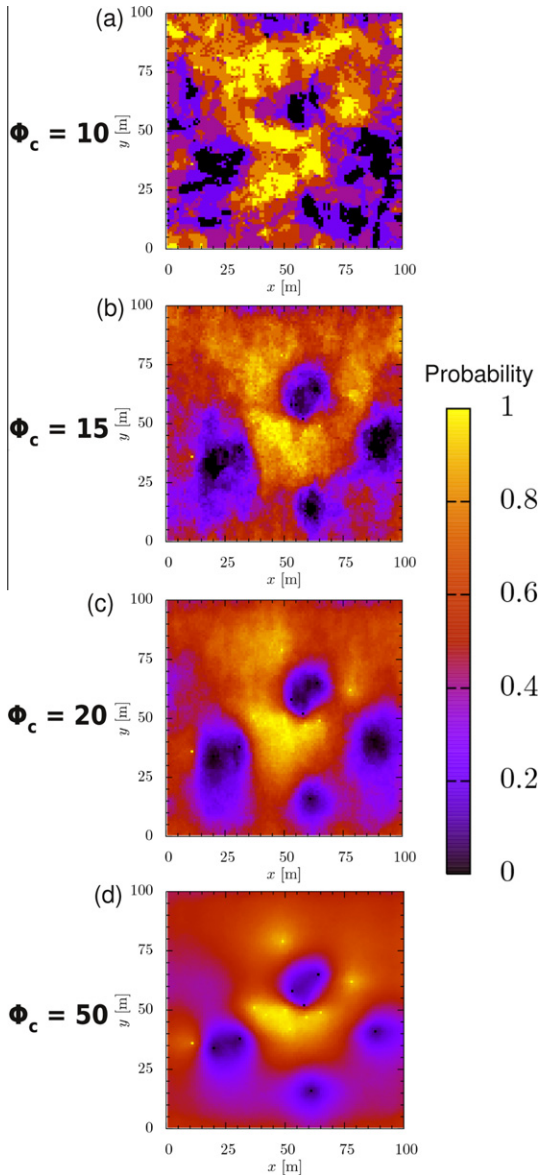
Fig. 9. Facies probability maps for the high conductivity geologic unit ($K = 10^{-2}$ m/s). The plausibilities used to generate these maps were estimated with confidence interval 0.050 with 95% confidence. The effect of conceptual model uncertainty is evaluated using values of Φ_c of (a) 10, (b) 15, (c) 20, and (d) 50.

(b) 0.35, and (c) 0.0 rounded to the specified resolution of the estimation ($L = 0.05$) at 95% confidence.

As illustrated in Fig. 2, probability estimates closer to 0.5 will require more samples n (Eq. (8)) for the same confidence level and interval length, as apparent in the number of realizations presented in the plots in Fig. 5, where (b), with probability closest to 0.5, requires the most realizations at 1377, and (c), with probability farthest from 0.5, requires the least at 292. Fig. 5 (a) requires 1120 realizations.

Fig. 6 presents acceptance probability maps on the geostatistical parameter space estimated at 95% confidence with $\Phi_c = 50$. Each 'cell' in the plots of Fig. 6 represents an estimation of the acceptance probability of a parameter set illustrated in Fig. 5. Confidence interval lengths (L) of (a) 0.05, (b) 0.025, and (c) 0.01 are presented. Acceptance probability estimates are rounded to increments spaced at the confidence interval length to avoid presenting greater resolution than the estimates warrant. This is evident as an increase in the number of colors in moving from Fig. 6(a) (b) (c). The geostatistical model (geostatistical parameter set) with the highest acceptance probability is $\bar{l}_x = \bar{l}_y = 10$ m (Fig. 5 (b)), not the model used to generate the 'true' realization, $\bar{l}_x = \bar{l}_y = 20$ m (Fig. 5 (a)). This is not a surprising result, as the head observations come from a single realization of this geostatistical model, and should not be expected to uniquely characterize the hydraulic response of the geostatistical model. In practice, this is the case as well, where the head observations are produced by a single spatial distribution of heterogeneity, where not only will the geostatistical parameters be unknown, but the appropriate geostatistical framework may be uncertain as well. It is therefore reasonable to expect that the observed hydraulic response may be more consistent with geostatistical models that are not the 'true' model. It is for this reason that the identification of an optimal geostatistical model from hydraulic observations is an ill-conceived strategy. This issue is circumvented in the current approach, as the analysis is based on the premise that all acceptable realizations are equally acceptable, irrespective of the magnitude of the estimated acceptance probability of their associated geostatistical model.

A similar presentation to Fig. 6 is made in Fig. 7, in this case demonstrating the effect of the perceived conceptual model uncertainty by modifying values of Φ_c ((a) 40, (b) 60, and (c) 80) holding $L = 0.05$. Once again, the confidence level of the estimation is 95%. By inspecting Fig. 7, it is apparent that increasing the perceived conceptual model uncertainty (increasing Φ_c) increases the acceptance probability values of the matrix of geostatistical models as a greater proportion of realizations become acceptable.

Facies probability maps generated from realizations collected during acceptance probability estimation at three interval lengths with a non-fuzzy definition of realization acceptance are presented in Fig. 8. The facies probability represents the probability of the existence of the high conductivity unit ($K_1 = 10^{-2}$ m/s) at a location. Probabilities near 1.0 or 0.0 can be interpreted as locations with high certainty of belonging to the high (K_1) or low (K_2) conductivity geologic unit, respectively. Locations with probabilities of 0.5 indicate that the geologic unit at that location is highly uncertain (i.e. the least amount of information is available for these locations). Complimentary information could be presented by mapping the probability of the low conductivity unit alternatively. The maps are developed using realizations from the union of acceptable sets from a matrix of geostatistical parameters (identical to the matrix of parameter values presented in Fig. 6 (81 parameter sets)) with $\Phi_c = 10$. The acceptance probabilities are estimated within confidence interval lengths L of (a) 0.05, (b) 0.025, and (c) 0.01 at 95% confidence. The number of realizations increases as L decreases, resulting in a smoothing effect of the probability map. Comparisons can be made between these maps and the 'true' hydrogeologic model in Fig. 3. These facies probability maps, and those in the subsequent discussions, can be used to site additional investigations by selecting locations that will reduce the uncertainty of the heterogeneity (locations with facies probability near 0.5 in the current scenario).

Fig. 9 presents a similar analysis to Fig. 8, in this case increasing Φ_c, effectively increasing the degree of uncertainty that the expert perceives in the conceptual model. For reference, Figs. 8(a) and 9(a) are the same. It is apparent that as Φ_c increases, indicating a greater level of epistemic uncertainty and that a greater number of realizations will be considered acceptable, the more that the information is restricted to the conditioning points, with locations far from the conditioning points having the greater uncertainty (facies probability ≈ 0.5).

Belief and plausibility maps generated from acceptable realizations collected during estimation of acceptance probability with a fuzzy definition of realization acceptance are presented in Fig. 10 for various membership functions. Two cases are presented, both starting with non-fuzzy acceptance metrics: (1) $\Phi_1 = \Phi_2 = 10$ and (2) $\Phi_1 = \Phi_2 = 20$, demonstrating that the facies belief and plausibil-
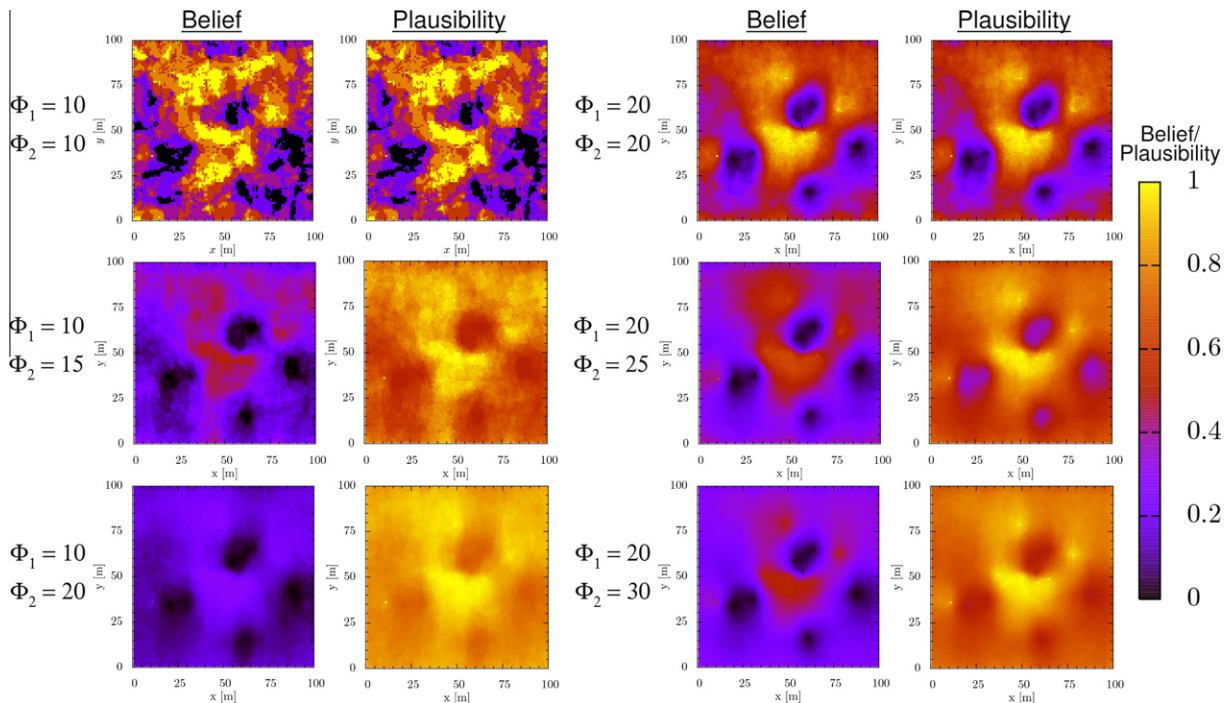
Fig. 10. Facies belief and plausibility maps for various definitions of realization acceptance defined by Φ_1 and Φ_2 (refer to Eqs. (2) and (3)).

ity maps are equal to the facies probability maps in these cases (refer to Fig. 9 (a) and (c), respectively). Below the plots with non-fuzzy acceptance metrics are facies belief and plausibility maps with increasing Φ_2 maintaining the same value for Φ_1. These plots indicate the increasing gap between $Pl(l;\mathbf{x})$ and $Bel(l;\mathbf{x})$ as the uncertainty in the definition of realization acceptance increases, demonstrating how the increased epistemic uncertainty in the definition of the acceptance metric is included in the analysis.

6. Conclusions

Accounting for the equifinality of solutions requires the inclusion of all solutions deemed acceptable by their consistency with observations. In geostatistical models, this means that model ranking can not be used to identify the 'true' model of statistical characteristics of heterogeneity, as the possibility always exists that a model that produces few acceptable realizations may be the most similar to the 'true' heterogeneity. Based on the potentially extreme variability of hydraulic responses from realizations from a single geostatistical model demonstrated here, it is apparent that identification of a single most likely geostatistical model may not improve hydrogeologic predications or provide an appropriate uncertainty analysis. We propose to develop uncertainty analyses including acceptable realizations from multiple geostatistical models. In this way, acceptable features from the included geostatistical models from multiple geostatistical models, and potentially multiple geostatistical frameworks, will be incorporated into a composite uncertainty analysis. This approach provides the potential to identify features that are not part of any of the single geostatistical models, but are present in the composite of the acceptable realizations, providing for a more comprehensive uncertainty analysis than the identification of the most likely geostatistical model.

A non-fuzzy definition of acceptance will allow the calculation of facies probability maps. A fuzzy definition of acceptance will produce bounds on the facies probability map in the form of facies belief and plausibility maps. The facies probability maps provide indications of the information content of the available data and

provide guidance for the selection of the location of future investigations intending to reduce uncertainty.

This paper has demonstrated that:

1. Geostatistically-equally-probable realizations from a geostatistical model can produce extremely varied flow simulations,
2. Considering epistemic and aleatoric uncertainty, many geostatistical models will produce acceptably consistent realizations with respect to observations,
3. Acceptance probability estimation can be used to collect acceptable realizations up to a specified confidence and interval length. Varying the confidence level and/or interval length can be used to obtain information at the level of detail and computational expense desired by increasing the precision of the acceptance probability estimates.
4. Estimation of acceptance probability at a specified confidence level and interval length can be preemptively terminated based on the current estimate reducing the computational demand for acceptance probabilities which deviate significantly from 0.5.
5. Expert opinion of epistemic uncertainty provides an appropriate uncertainty in facies distribution (increasing Φ_c increases the uncertainty in facies distribution as more realizations become acceptable). This step in the approach has similar implications as the selection of a likelihood function in a Bayesian analysis, requiring expert judgment and/or evaluation; however, with less restrictions and assumptions.
6. Imprecise (fuzzy) definitions of epistemic uncertainty can be accounted for by utilizing DST, providing bounds of the facies probability map in the form of facies belief and plausibility maps.

Acknowledgements

Appendix A. Satisfaction of axiomatic properties of the derived *bpa*, belief measure, and plausibility measure

We begin by demonstrating that the derived *bpa* (Eq. (13)),

$$m(l;\mathbf{x}) = \frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x}) \mu_A(\Phi_i), \tag{A.1}$$

satisfies the axiomatic properties of a basic probability assignment, defined here as

$$m(\varnothing) = 0, \tag{A.2}$$
$$m(B) \geqslant 0 \quad B \in P(\mathbf{X}), \tag{A.3}$$
$$\sum_{B \in P(\mathbf{X})} m(B) = 1, \tag{A.4}$$

where $\mathbf{X}$ is the universe of discourse (commonly referred to as the frame of discernment in DST) and $P(\mathbf{X})$ is the power set of $\mathbf{X}$ consisting of singletons (i.e. single components of $\mathbf{X}$) and collections of singletons [26]. Therefore, m does not adhere to the axioms of probability, as it is defined on $P(\mathbf{X})$, whereas probabilities must be defined solely on $\mathbf{X}$ (i.e. singletons).

By inspecting Eq. (A.1), it is apparent that the properties defined by Eqs. (A.2) and (A.3) are satisfied. The property defined by Eq. (A.4) is also satisfied due to the assumption that the facies are mutually-exclusive and exhaustively defined implicit in a Markov chain or indicator geostatistical framework [9]. This implies that, prior to analysis, the *bpa* that at least one of the facies exists at a location ($m(\mathbf{X};\mathbf{x})$) is equal to one. This is the case of total ignorance with respect to the frame of discernment (i.e. $m(\mathbf{X};\mathbf{x}) = 1$ and $m(l;\mathbf{x}) = 0$ for $l = 1,\ldots,N_f$, where N_f is the number of facies). During the analysis, a portion of this evidence is shifted onto the singletons of the power set, reducing the initial uncertainty as information is obtained through the acceptance probability estimation. Therefore, the evidence that is not assigned to a singleton during the analysis remains on $\mathbf{X}$ as

$$m(\mathbf{X};\mathbf{x}) = 1 - \sum_{l=1}^{N_f} m(l;\mathbf{x}), \tag{A.5}$$

indicating the remaining conflicting information. Rearranging Eq. (A.5) demonstrates that the property defined in Eq. (A.4) is satisfied.

Next, we demonstrate that the derived belief measure (Eq. (16)),

$$Bel(l;\mathbf{x}) = m(l;\mathbf{x}) = \frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x}) \mu_A(\Phi_i), \tag{A.6}$$

satisfies the axiomatic properties of a belief measure, defined here as

$$Bel(\varnothing) = 0, \tag{A.7}$$
$$Bel(\mathbf{X}) = 1 \tag{A.8}$$

and

$$Bel(D_1 \cup D_2 \cup \cdots \cup D_n) \geqslant \sum_j Bel(D_j) - \sum_{j<k} Bel(D_j \cap D_k) + \cdots$$
$$+ (-1)^{n+1} Bel(D_1 \cap D_2 \cap \cdots \cap D_n). \tag{A.9}$$

A fundamental property of belief measures can be derived by substituting $D = D_1$ and $\overline{D} = D_2$ ($\overline{D}$ represents "not D") for $n = 2$ in Eq. A.9 as

$$Bel(D) + Bel(\overline{D}) \leqslant 1, \tag{A.10}$$

where D_1 and D_2 can represent the acceptable and unacceptable sets of realizations, for example. It is apparent that Eq. (A.6) satisfies

the axiomatic properties of belief measure defined by Eqs. (A.7) and (A.8) considering the arguments related to the properties of a *bpa*.

To demonstrate that the derived belief measure (Eq. (A.6)) satisfies Eq. (A.10), consider that

$$Bel(\bar{l};\mathbf{x}) = \sum_{m=1|m \neq l}^{N_f} Bel(m;\mathbf{x}) \tag{A.11}$$

and

$$\sum_{l=1}^{N_f} Bel(l;\mathbf{x}) = \sum_{l=1}^{N_f} \left[\frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x}) \mu_A(\Phi_i) \right]$$
$$\leqslant \sum_{l=1}^{N_f} \left[\frac{1}{N_A} \sum_{i=1}^{N_A} I_{l,i}(\mathbf{x}) \right] = 1. \tag{A.12}$$

Therefore,

$$Bel(l;\mathbf{x}) + Bel(\bar{l};\mathbf{x}) = Bel(l;\mathbf{x}) + \sum_{m=1|m \neq l}^{N_f} Bel(m;\mathbf{x})$$
$$= \sum_{l=1}^{N_f} Bel(l;\mathbf{x}) \leqslant 1 \tag{A.13}$$

in accordance with Eq. (A.10).

Last, we demonstrate that the derived plausibility measure (Eq. (17)),

$$Pl(l;\mathbf{x}) = 1 - Bel(\bar{l};\mathbf{x}) \tag{A.14}$$

satisfies the axiomatic properties of a plausibility measure, defined here as

$$Pl(\varnothing) = 0, \tag{A.15}$$
$$Pl(\mathbf{X}) = 1 \tag{A.16}$$

and

$$Pl(D_1 \cap D_2 \cap \cdots \cap D_n) \leqslant \sum_j Pl(D_j) - \sum_{j<k} Pl(D_j \cup D_k) + \cdots$$
$$+ (-1)^{n+1} Pl(D_1 \cup D_2 \cup \cdots \cup D_n). \tag{A.17}$$

A fundamental property of plausibility measures can be derived by substituting $D = D_1$ and $\overline{D} = D_2$ for $n = 2$ in Eq. A.17 as

$$Pl(D) + Pl(\overline{D}) \geqslant 1. \tag{A.18}$$

The derived plausibility measure (Eq. (A.14)) satisfies Eqs. (A.15) and (A.16) as

$$Pl(\varnothing;\mathbf{x}) = 1 - Bel(\overline{\varnothing};\mathbf{x}) = 1 - Bel(\mathbf{X}) = 0 \tag{A.19}$$

and

$$Pl(\mathbf{X};\mathbf{x}) = 1 - Bel(\overline{\mathbf{X}};\mathbf{x}) = 1 - Bel(\varnothing) = 1, \tag{A.20}$$

respectively. Adherence of Eq. (A.14) to the axiomatic property defined by Eq. (A.18) can be demonstrated considering Eq. (A.10) as

$$Pl(l;\mathbf{x}) + Pl(\bar{l};\mathbf{x}) = 2 - (Bel(l;\mathbf{x}) + Bel(\bar{l};\mathbf{x})) \geqslant 1. \tag{A.21}$$

References

[1] Ross JL, Ozbek MM, Pinder GF. Aleatoric and epistemic uncertainty in groundwater flow and transport simulation, Water Resources Research 45, doi:10.1029/2007WR006799.
[2] Beven K, Binley A. The future of distributed models: model calibration and uncertainty prediction. Hydrol Process 1992;6:279–98.
[3] Beven K, Freer J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J Hydrol 2001;249:11–29.
[4] Neuman S. Maximum likelihood Bayesian averaging of uncertain model predictions. Stoch Environ Res Risk Assess 2003;17:291–305.

[5] Gaganis P, Smith L. A Bayesian approach to the quantification of the effect of model error on the prediction of groundwater models. Water Resour Res 2001;37(9):2309–22.

[6] Rojas R, Feyen L, Dassargues A. Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and Bayesian model averaging, Water Resources Research 44. doi:10.1029/2008WR006908.

[7] Deutsch C, Journel A. GSLIB geostatistical software library and user's guide. New York: Oxford University Press; 1992.

[8] Strebelle S. Conditional simulation of complex geologic structures using multiple-point statistics. Math Geo 2002;34:1–21. doi:10.1023/A:1014009426274.

[9] Carle S, Fogg G. Transition probability-based indicator geostatistics. Math Geo 1996;28(4):453–76.

[10] Matheron G. Eléments pour une théorie des milieux poreux [elements for a theory of porous media]. Paris: Masson; 1967.

[11] Ben Ameur H, Chavent G, Jaffre J. Refinement and coarsening indicators for adaptive parametrization: Application to the estimation of hydraulic transmissivities. Inv Prob 2002;18:775–94.

[12] Tsai F-C, Sun N-Z, Yeh W-G. Geophysical parameterization and parameter structure identification using natural neighbors in groundwater inverse problems. J Hydrol 2005;308:269–83. doi:10.1016/j.jhydrol.2004.11.004.

[13] Lu Z, Robinson B. Parameter identification using the level set method, Geophysical Research Letters 33, l06404, doi:10.1029/2005GL025541.

[14] Chang H, Zhang D, Lu Z. History matching of facies distribution with the EnKF and level set parameterization. J Comput Phys 2010;229:8011–30. doi:10.1016/j.jcp.2010.07.005.

[15] Ramarao B, Lavenue A, de Marsily G, Marietta M. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1. theory and computational experiments. Water Resour Res 1995;31(3):475–93.

[16] Kitanidis P. On the geostatistical approach to the inverse problem. Adv Water Resour 1996;19(6):333–42.

[17] Yeh T-C, Liu S. Hydraulic tomography: Development of a new aquifer test method. Water Resour Res 2000;36(8):2095–105.

[18] Vesselinov V, Neuman S, Illman W. Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff 2. equivalent parameters, high-resolution stochastic imaging and scale effects. Water Resour Res 2001;37(12):3019–41.

[19] Hernandez A, Neuman S, Guadagnini A, Carrera J. Conditioning mean steady state flow on hydraulic head and conductivity through geostatistical inversion. Stoch Environ Res Risk Assess 2003;17(5):329–38.

[20] Pardo-Igùzquiza E, Chica-Olmo M, Garcia-Soldado MJ, Luque-Espinar JA. Using semivariogram parameter uncertainty in hydrogeological applications. Ground Water 2009;47(1):25–34.

[21] Nowak W, de Barros F, Rubin Y. Bayesian geostatistical design: task-driven optimal site investigation when the geostatistical model is uncertain, Water Resources Research 46. doi:10.1029/2009WR008312.

[22] Meier PM, Medina A, Carrera J. Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. Ground Water 2001;39(1):10–7.

[23] Harp D, Dai Z, Wolfsberg A, Vrugt J, Robinson, B, Vesselinov V. Aquifer structure identification using stochastic inversion, Geophysical Research Letters L08404, doi:10.1029/2008GL033585.

[24] Harp D, Vesselinov V. Stochastic inverse method for estimation of geostatistical representation of hydrogeologic stratigraphy using borehole logs and pressure observations, Stochastic Environmental Research and Risk Assessment.

[25] Ross TJ. Fuzzy logic with engineering applications. second ed. West Sussex, England: John Wiley & Sons; 2004.

[26] Klir G. Uncertainty and information: foundations of generalized information theory. Hoboken, NJ: John Wiley & Sons; 2006.

[27] Zadeh L. Fuzzy sets. Inform Control 1965;8(3):338–53.

[28] Aronica G, Hankin B, Beven K. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Adv Water Resour 1998;22(4):349–65.

[29] Franks SW, Beven KJ. Estimation of evapotranspiration at the lanscape scale: a fuzzy disaggregation approach. Water Resour Res 1997;33(12):2929–38.

[30] Bàrdossy A, Bronstert A, Merz B. 1-,2- and 3-dimensional modeling of water movement in the unsaturated soil matrix using a fuzzy approach. Adv Water Resour 1995;18(4):237–51.

[31] Mathon BR, Ozbek MM, Pinder GF. Transmissivity and storage coefficient estimation by coupling the Cooper–Jacob method and modified fuzzy least-squares regression. J Hydrol 2008;253:267–74. doi:10.1016/j.jhydrol.2008.02.004.

[32] Shafer G. A mathematical theory of evidence. Princeton, NJ: Princeton University Press; 1976.

[33] A. Dempster, Upper and lower probability inferences based on a sample from a finite univariate population, Biometrika 54 (3 and 4) 515–528.

[34] Mathon BR, Ozbek MM, Pinder GF. Dempster–Shafer theory applied to uncertainty surrounding permeability. Math Geosci 2010;42:293–307. doi:10.1007/s11004-009-9246-0.

[35] Hayter AJ. Probability and Statistics for Engineers and Scientists. second ed. Duxbury; 2002.

[36] Carle S, Fogg G. Modeling spatial variability with one and multidimensional continuous-lag markov chains. Math Geol 1997;29(7):891–918.

[37] G. Zyvoloski, B. Robinson, Z. Dash, L. Trease, Summary of the models and methods for the fehm application—a finite-element heat- and mass-transfer code, Tech. Rep. LA-13307-MS, Los Alamos Natl. Lab., Los Alamos, NM (1997).