

Stochastic Inverse Method for Estimation of Geostatistical Representation of Hydrogeologic Stratigraphy Using Borehole Logs and Pressure Observations

Dylan R. Harp · Velimir V. Vesselinov

Received: date / Accepted: date

Abstract An approach is presented for identifying statistical characteristics of stratigraphies from borehole and hydraulic data. The approach employs a Markov-chain based geostatistical framework in a stochastic inversion. Borehole data provide information on the stratigraphy while pressure and flux data provide information on the hydraulic performance of the medium. The use of Markov-chain geostatistics as opposed to covariance-based geostatistics can provide a more easily interpreted model geologically and geometrically. The approach hinges on the use of mean facies lengths (negative inverse auto-transition rates) and mean transition lengths (inverse cross-transition rates) as adjustable parameters in the stochastic inversion. Along with an unconstrained Markov-chain model, simplifying constraints to the Markov-chain model, including (1) proportionally-random and (2) symmetric spatial correlations, are evaluated in the stochastic inversion. Sensitivity analyses indicate that the simplifying constraints can facilitate the inversion at the cost of spatial correlation model generality. Inverse analyses demonstrate the feasibility of this approach, indicating that despite some low parameter sensitivities, all adjustable parameters do converge for a sufficient number of ensemble realizations towards their “true” values. This paper extends the approach presented in Harp et al (2008) to (1) statistically characterize the hydraulic response of a geostatistical model, thereby incorporating an uncertainty analysis directly in the inverse method, (2) demonstrate that a gradient-based optimization strategy is sufficient, thereby providing relative computational efficiency

D.R. Harp · V.V. Vesselinov
Hydrology, Geochemistry, and Geology Group, Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, U.S.A.
Tel.: 1-505-667-5532
Fax: 1-505-665-8737
E-mail: dharp@lanl.gov

D.R. Harp
Department of Civil Engineering, University of New Mexico, Albuquerque, NM, U.S.A.
E-mail: dharp@unm.edu

compared to global optimization strategies, (3) demonstrate that the approach can be extended to a 3-D analysis, and (4) introduce the use of mean facies lengths and mean transition lengths as adjustable parameters in a geostatistical inversion, thereby allowing the approach to be extended to greater than 2 category Markov-chain models.

Keywords Markov-chain geostatistics · stratigraphy · model inversion

1 Introduction

Aquifers are often comprised of stratigraphic units with distinct sedimentological characteristics as well as lithological and mineralogical composition. In cases where the effective hydrogeologic properties of these stratigraphic units are distinct as well (e.g. LaBolle and Fogg (2001), Gégo et al (2001), and Fienen et al (2004)), information pertaining to the stratigraphy can be inferred from the hydrogeologic behavior of the aquifer. If the stratigraphy is known at some locations (e.g. based on bore-hole logs or outcrops), this information can be used to constrain the set of possible aquifer structures using various geostatistical techniques (e.g. covariance-based (Deutsch and Journel, 1992) and Markov-chain-based (Carle and Fogg, 1996) indicator co-kriging). However, given the typically sparse hydraulic and geologic observations available from most aquifers, the number of possible structures concurrent with the available data is generally large. Therefore, estimation of the stratigraphy must be posed in a stochastic framework. A set of geologically-equally-probable stratigraphies can be represented by a single geostatistical model. This allows the problem of aquifer structure identification to be simplified to the identification of a set of structures exhibiting highly-probable hydraulic characteristics conditioned on the available data.

In order to model the aquifer structure as a composition of distinct stratigraphic units, we employ a multidimensional continuous-lag Markov-chain model of the spatial variability of categorical variables. In this case, aquifer heterogeneity is assumed to be composed of spatially-discontinuous structures characterized by distinct stratigraphic units with uniform properties. This approach is appropriate not only for cases where the hydrogeologic properties of the stratigraphic units are distinct and approximately uniform, but also for watershed-to-basin scale models where delineations of large-scale changes in properties are the most important characteristics (Cardiff and Kitanidis, 2009). The Markov-chain model is represented in a transition-probability framework describing the auto- and cross-transition probabilities of the categorical variables at specific lag distances (Carle and Fogg, 1997).

A given geostatistical model defines spatial characteristics of a stratigraphy using statistical parameters. Therefore, any geostatistical model defines an infinite set of stratigraphies exhibiting defined statistical spatial characteristics. Equally-probable realizations of this infinite set of stratigraphies can be generated using conditional simulation. While the equally-probable realizations of a geostatistical model may exhibit deviations from the defined spatial characteristics, statistical inferences of these characteristics are expected to converge to the defined parameter values at a sufficiently large set of realizations. Similarly, we demonstrate that convergent statistical

inferences can be obtained that define the hydraulic response of an aquifer given a sufficient number of stochastic realization from a geostatistical model. This is fundamental to the structure identification approach presented here, providing a necessary link between stratigraphic spatial correlation and aquifer hydraulic response. We utilize this link to explore the plausibility of geostatistical models by comparing hydraulic responses inferred from sets of realizations from geostatistical models to observed hydraulic responses.

If the available geological data from an aquifer is obtained from vertical boreholes, characteristics of the vertical spatial correlation of stratigraphic units (e.g. juxtapositional tendencies and mean lengths) can be estimated directly (Zhang and Li, 2008). However, given the sparse and discontinuous nature of such data in the horizontal plane, direct computation of the horizontal juxtapositional tendencies and mean lengths is not always possible (Weissmann and Fogg, 1999). While mean lengths can be estimated indirectly by inference based on discontinuous geologic observations, these estimates can be highly uncertain. Recognizing this situation, we have conducted an analysis assuming that the vertical spatial correlation model is known, focusing on demonstrating the ability to complete the 3-D spatial-correlation model by inferring the horizontal spatial correlations using hydraulic data.

The use of covariance-based geostatistics have been developed and demonstrated to be useful in many applications (Gómez-Hernández et al, 2003; Neuman, 2004). Hendricks Franssen and Gómez-Hernández (2002) demonstrate a stochastic continuum approach to estimate hydraulic conductivity at a fractured site. Serre et al (2003) and Orton and Lark (2007) demonstrate the use of a Bayesian Maximum Entropy (BME) approach to estimate heterogeneous property fields. Recently, Zanini and Kitanidis (2009) proposed the use of a Markov-Chain Monte Carlo (MCMC) approach to generate conditional realizations of high-contrast transmissivity fields using a quasilinear approach and Huang et al (2009) explored the use of data assimilation using the Ensemble Kalman filter (ENKF) to estimate hydraulic conductivity to improve solute transport prediction.

However, the use of Markov-chain geostatistics has seen limited development. Carle et al (1998) demonstrate the use of a trial-and-error approach to modify cross-transition rates for horizontal Markov-chain spatial correlation models to obtain geologically-plausible juxtapositional relationships. The calibration of a conditional Markov-chain model to hydraulic data was proposed by Zhenxue Dai in a personal communication in 2007. This proposal provided the impetus for the research presented in Harp et al (2008), where a synthetic 2-D 2-unit (2 stratigraphic unit) aquifer was analyzed utilizing the concept of a representative realization of a spatial correlation model. Harp et al (2008) inverted structural parameters from 2-unit Markov-chain models in vertical and lateral directions (e.g. facies mean lengths and volumetric facies proportions) obtaining head predictions from a single stochastic realization for each geostatistical model. It was assumed that the hydraulic response from the realization is representative of the hydraulic response for the infinite set of geologically-equally-probable realizations for the geostatistical model.

In contrast, the current research considers a 3-D 3-unit aquifer utilizing the concept of representative hydraulic sample statistics of a spatial correlation model. In this case, flow simulations are performed on a hydraulically convergent set of geologically-

equally-probable stratigraphies from the current geostatistical model and averaging is performed on the simulated hydraulic characteristics (i.e. pressures and fluxes). This provides a characteristic hydraulic response of the current geostatistical model considering the inherent hydraulic response variability. This improves the inversion performance as averaging the hydraulic response of a set of realizations provides smoother variations with changes in structural parameters than the use of a single representative realization. We demonstrate that a gradient-based optimization strategy is sufficient with this approach, thereby providing computational efficiency compared to global search optimization strategies. In order to consider Markov-chain models with greater than two categories (stratigraphic units), the concept of mean facies and mean transition lengths are introduced and utilized as adjustable parameters to modify the horizontal Markov-chain models (mean transition lengths do not need to be explicitly specified in a 2-category Markov-chain geostatistical model as a single mean facies length is sufficient, see discussion in Section 3.1 and equation (25)). This paper demonstrates the feasibility of the proposed structure identification approach utilizing a 3-D 3-unit synthetic model by (1) a sensitivity analysis of the adjustable parameters and (2) model inversions demonstrating the performance of a gradient-based optimization on this inverse framework.

2 Theoretical Discussion of Stochastic Representations of Stratigraphy

Boolean models present the earliest approach for modeling discontinuous spatial variability (Matheron, 1967), providing the ability to model adjacently-located distinct stratigraphic units. While Boolean models presented an advantage over continuous geostatistical approaches in considering aquifer connectivity, it is difficult to honor conditioning data in this framework and the Boolean object geometries must be defined *a priori* (de Marsily et al, 2005). The introduction of the indicator kriging approach by Journel (1983) provided an alternative to Boolean models, modeling distinct stratigraphic units using indicator functions in a geostatistical framework. Indicator kriging is easily conditioned to observed geologic data and produces more realistic structures based on an inferred or imposed covariance model as opposed to assumed geometric objects utilized in Boolean models.

Carle and Fogg (1997) extended 1-D Markov-chain models, previously demonstrated to successfully model embedded occurrences of individual strata (Vistelius, 1949), to multidimensional continuous-lag Markov-chain models by interpolation of transition-rate matrices defining the 1-D Markov-chain models. The multidimensional Markov-chain model provides an alternative representation of spatial correlation to the indicator cross-covariance (or cross-variogram) conventionally used in indicator geostatistics.

Transition probabilities can be inferred from a sufficient number of geologic observations by determining the relative frequency of stratigraphic unit transitions at various lag distances. A transition-probability representation of a Markov-chain model can then be empirically derived by determining transition rates that optimally relate transition probabilities to distance (h [L]) according to the following matrix exponential functional form:

$$\mathbf{T}(h) = \exp(\mathbf{R}h), \quad (1)$$

where $\mathbf{T}(h)$ is a transition-probability matrix denoted as $\mathbf{T}(h) = (t_{ij}(h))$, $i, j = 1, \dots, N$ and \mathbf{R} is a transition-rate matrix denoted as $\mathbf{R} = (r_{ij})$, $i, j = 1, \dots, N$, where N is the number of mutually-exclusive, exhaustively-defined stratigraphic units (categories) considered in the Markov-chain model. Utilizing an eigensystem analysis of \mathbf{R} , the individual transition probability functions can be defined as

$$t_{ij}(h) = \sum_{k=1}^N z_{ij,k} \exp(\lambda_k h), \quad (2)$$

where $z_{ij,k}$ is the ij th component of the k th spectral component matrix and λ_k is the k th eigenvalue.

By inspecting equation (2), it is apparent that the functional form for the transition-probability is a summation of N exponentials. As a result, a Markov-chain model is not restricted to purely “exponential-looking” spatial correlation functions, allowing complex juxtapositional patterns to be modeled.

This empirical derivation of a Markov-chain geostatistical model is analogous to the inference of auto- and cross-variograms in covariance-based indicator geostatistics, differing in that a variogram relates a statistical tendency of deviation in property or indicator values to separation distance, while the Markov-chain geostatistical model relates probabilistic stratigraphic juxtapositioning to separation distance (i.e. relative-locational tendencies of stratigraphic units). Therefore, a Markov-chain model can provide a more interpretable stochastic representation of a stratigraphic distribution than the covariance-based approach.

As the Markov-chain approach directly models juxtapositional tendencies, asymmetric spatial correlations, such as fining or coarsening cycles, are easily represented. Additionally, given the necessary adherence to laws of probability concerning the stratigraphic transitions requires the simultaneous calibration of the components of the full transition-probability matrix, ensuring that the Markov-chain-model is internally consistent. This is not necessarily the case for the covariance-based approach as individual auto- and cross-variograms are calibrated independently (de Marsily et al, 2005).

Considering the previous discussion, the Markov-chain model presents certain advantages to modeling categorical spatial variability, while still being amenable to implementation into an indicator co-kriging interpolation scheme. While much effort has been focused on developing inversion strategies for covariance-based spatial-correlation models, little has been applied to the Markov-chain approach (e.g. Harp et al (2008)). This paper presents a strategy that implements a Markov-chain spatial-correlation framework into a stochastic inversion.

3 Methodology

Computation of the model evaluated by the stochastic inverse method presented here requires a multi-step approach: (1) generation of a hydraulically-convergent set of

geologically-equally-probable realizations of stratigraphy from a geostatistical model, (2) flow simulation on the stratigraphic fields assigning distinct uniform hydraulic properties to the stratigraphic units, and (3) inference of hydraulic-response characteristics of the geostatistical model. These steps are discussed in detail below along with a discussion of the applied stochastic inverse approach.

3.1 Generation of Stratigraphic Realizations of a Markov-Chain Geostatistical Model

A 1-D continuous-lag Markov-chain model can be represented by a transition-rate matrix (introduced in equation (1)) composed of auto- and cross-transition rates as

$$\mathbf{R} = (r_{ij}), \quad i, j = 1, \dots, N, \quad (3)$$

where r_{ij} is an auto-transition rate when $i = j$ and a cross-transition rate when $i \neq j$. As cross-transition rates are not as readily inferred from geologic observations as cross-transition probabilities (Carle and Fogg, 1997), it is generally desirable to compute the transition-rate matrix from an inferred discrete-lag transition-probability matrix $\mathbf{T}(\Delta h)$ as

$$\mathbf{R} = \frac{\ln(\mathbf{T}(\Delta h))}{\Delta h}, \quad (4)$$

where Δh is a discrete lag (Carle and Fogg, 1997). If a stratigraphic pattern has been identified in the Markov-chain model (i.e. in the transition probabilities), similar results will be obtained for arbitrary selections of Δh . As presented in equation (1), \mathbf{R} defines the continuous-lag transition-probability matrix $\mathbf{T}(h)$.

Adherence to probability theory constrains the Markov-chain model as

$$\sum_{i=1}^N p_i = 1; \quad (5)$$

$$r_{ii} \leq 0 \quad \forall i \quad \text{and} \quad r_{ij} \geq 0 \quad \forall i, j \neq i, \quad (6)$$

$$\sum_{j=1}^N r_{ij} = 0 \quad \forall i, \quad (7)$$

$$\sum_{i=1}^N p_i r_{ij} = 0 \quad \forall j, \quad (8)$$

where p_i is the one-location marginal probability or volumetric proportion of the i th stratigraphic unit (Carle and Fogg, 1997).

Adherence to the equations and inequalities numbered (5) through (8) ensures that the Markov-chain model is internally consistent and ergodic (i.e. $\lim_{h \rightarrow \infty} t_{ij}(h) = p_j$ (Ross, 1993)). As a result, the auto- and cross-transition rates associated with one of the stratigraphic units (i.e. rates along the row and column associated with a stratigraphic unit), usually denoted as the background category, can be automatically

determined from the other transition rates. Therefore, it is only necessary to determine $(N - 1)^2$ transition rates to define a transition-probability representation of a Markov-chain model.

A Markov-chain model can be completely modified by adjusting the transition rates comprised in a transition-rate matrix \mathbf{R} (equation (3)). A 2-category Markov-chain model is completely defined by specifying a single transition rate, assuming that the proportions are available, utilizing equations (7) and (8) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & -r_{11} \\ -p_1/p_2 r_{11} & p_1/p_2 r_{11} \end{bmatrix}, \quad (9)$$

where category two has been set as the background and r_{11} is the only transition rate requiring specification. As proportions are independent of direction, proportions may be available with reasonable certainty from continuous bore logs (assuming stationarity and lack of bias in borehole locations) for use in directions with discontinuous data.

A 3-category transition-rate matrix (designating category three as background) can be defined similarly as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & -r_{11} - r_{12} \\ r_{21} & r_{22} & -r_{21} - r_{22} \\ -p_1 r_{11} - p_2 r_{21} & -p_1 r_{12} - p_2 r_{22} & p_1(r_{11} + r_{12}) + p_2(r_{21} + r_{22}) \end{bmatrix}, \quad (10)$$

where it is apparent that four transition rates (i.e. r_{11} , r_{12} , r_{21} , and r_{22}) require specification.

It is possible to impose simplifying constraints on the spatial correlations within a Markov-chain model. For instance, it is possible to require that the probability of making a particular transition is proportional to making the opposite transition as

$$t_{ij}(h) = \frac{p_j}{p_i} t_{ji}(h). \quad (11)$$

where it is apparent that the proportionality constant is the ratio of the volumetric facies proportions. The assumption may be justified for some geologic scenarios, and is in fact an intrinsic constraint of the variogram analysis (Carle and Fogg, 1996). Using the definition of the transition rate,

$$r_{ij} = \frac{\partial t_{ij}(0)}{\partial h} \quad \forall i, j, \quad (12)$$

equation (11) can be expressed as

$$r_{ij} = \frac{p_j}{p_i} r_{ji}. \quad (13)$$

Using the assumption of symmetry between non-background categories, a 3-category Markov-chain model requires the specification of three transition rates ($N(N - 1)/2$, in general), reducing the number of rates needing specification by one ($(N - 1)/2$,

$1)(N - 2)/2$, in general). This is illustrated, utilizing equation (13), to impose symmetric spatial correlation in the non-background categories (categories 1 and 2) in equation (10) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & -r_{11} - r_{12} \\ p_1/p_2 r_{12} & r_{22} & -p_1/p_2 r_{12} - r_{22} \\ -p_1 r_{11} - p_1 r_{12} & -p_1 r_{12} - p_2 r_{22} & p_1 r_{11} + 2p_1 r_{12} + p_2 r_{22} \end{bmatrix}, \quad (14)$$

where it is apparent that it is only necessary to specify r_{11} , r_{12} , and r_{22} (alternatively, r_{21} could have been specified as opposed to r_{12}).

An alternative spatial correlation constraint is to assume that the probability of a transition from the i th to j th facies is solely dependent on the random probability of encountering the j th facies without juxtapositional spatial correlations. This is similar to assuming that the juxtapositional tendencies are random and solely dependent on the volumetric proportions of the facies (we call this proportionally-random spatial correlation). If the occurrence of stratigraphic units is assumed to be proportionally random for non-background categories, specification of the Markov-chain model is simplified by assuming that cross-transition probabilities of the non-background categories can be determined by weighting the collective probability of any cross-transition (not including an auto-transition) from a particular category by the relative proportion of the considered transition category (Carle and Fogg, 1997) as

$$t_{ij}(h) = (1 - t_{ii}(h)) \frac{p_j}{1 - p_i} \quad \text{for } j \neq i. \quad (15)$$

This can be expressed with respect to transition rates utilizing equation (12) as

$$r_{ij} = -r_{ii} \frac{p_j}{1 - p_i} \quad \text{for } j \neq i. \quad (16)$$

The proportionally-random assumption for non-background category spatial correlations will reduce the number of transition rates needing specification for a 3-category Markov-chain model to two ($N - 1$, in general), as it is only necessary to specify the auto-transition rates, excluding the background auto-transition rate. This is illustrated utilizing equation (16) to impose proportional randomness in the non-background categories in equation (10) as

$$\mathbf{R} = \begin{bmatrix} r_{11} & -r_{11} \frac{p_2}{(1-p_1)} & r_{11} \left(\frac{p_2}{1-p_1} - 1 \right) \\ -r_{22} \frac{p_1}{1-p_2} & r_{22} & r_{22} \left(\frac{p_1}{1-p_2} - 1 \right) \\ p_1 \left(r_{22} \frac{p_2}{1-p_2} - r_{11} \right) & p_2 \left(r_{11} \frac{p_1}{1-p_1} - r_{22} \right) & p_1 r_{11} + p_2 r_{22} - p_1 p_2 \left(\frac{r_{11}}{1-p_1} + \frac{r_{22}}{1-p_2} \right) \end{bmatrix}, \quad (17)$$

where it is apparent that it is only necessary to specify r_{11} and r_{22} .

It is important to note that proportional randomness does not imply symmetric correlations (Turk, 1982). In fact, spatial correlations are proportionally random and symmetric only if

$$r_{ii} = \frac{1 - p_i}{1 - p_j} r_{jj} \quad \text{for } j \neq i, \quad (18)$$

which can be derived by substituting equation (16) in equation (13). Equation (18) indicates that symmetric proportionally-random spatial correlations require that transition rates are related by their volumetric proportions. This is an intrinsic property of a 2-category Markov-chain model as $1 - p_1 = p_2$ and $1 - p_2 = p_1$. Using these relations for a 2-category Markov-chain model and setting $i = 2$ and $j = 1$ in equation (18) results in $r_{22} = p_1/p_2 r_{11}$, which is equivalent to the formula for r_{22} presented in equation (9), demonstrating that 2-category Markov-chain models are necessarily composed of symmetric, proportionally-random spatial correlations.

As indicated by Dai et al (2007), the eigenvalues of a transition rate matrix representing symmetric proportionally-random spatial correlations will be $\eta_1 = 0$ and $\eta_i = r_{11}/(1 - p_1)$ for $i = 2, \dots, N$. The numerical eigenvalue decomposition utilized in the current research requires the computation of spectral component matrices \mathbf{Z}_i (Carle and Fogg, 1997) as

$$\mathbf{Z}_i = \frac{\prod_{j \neq i} (\eta_j \mathbf{I} - \mathbf{R})}{\prod_{j \neq i} (\eta_j - \eta_i)} \quad \forall i = 1, \dots, N \quad (19)$$

where \mathbf{I} is the $N \times N$ identity matrix. As noted by Dai et al (2007), it is apparent that cases where $\eta_i \approx \eta_j$ will cause numerical instability in equation (19). We avoid symmetric proportionally-random spatial correlations here as this case results in 2 repeating eigenvalues for a 3-category Markov-chain model ($(N - 1)$ in general). Note that this issue is irrelevant for 2-category Markov-chain models as there are no repeating eigenvalues (i.e. $N - 1 = 1$). Dai et al (2007) present an analytical expression for transitional probability matrix coefficients t_{ij} as a function of the lag distance. This expression is derived under the assumption of symmetric, proportionally-random spatial correlations and can be applied only for this special case.

To facilitate a geometric representation of structures generated from a Markov-chain model, facies mean lengths have been defined (Carle and Fogg, 1996) as

$$\bar{l}_i = -\frac{1}{r_{ii}} = \left[\frac{\partial t_{ii}(0)}{\partial h} \right]^{-1} \quad (20)$$

where \bar{l}_i is the facies mean length of the i th stratigraphic unit in a particular direction. We utilize inverse transition rates as the adjustable parameters in the stochastic inversion as we have found they scale more appropriately in the inversion than transition rates. In order to facilitate the use of inverse transition rates, we extend equation (20) to define mean facies lengths as

$$\bar{l}_{ii} = -\frac{1}{r_{ii}} = -\left[\frac{\partial t_{ii}(0)}{\partial h} \right]^{-1} \quad (21)$$

and mean transition lengths as

$$\bar{l}_{ij} = \frac{1}{r_{ij}} = \left[\frac{\partial t_{ij}(0)}{\partial h} \right]^{-1} \quad i \neq j \quad (22)$$

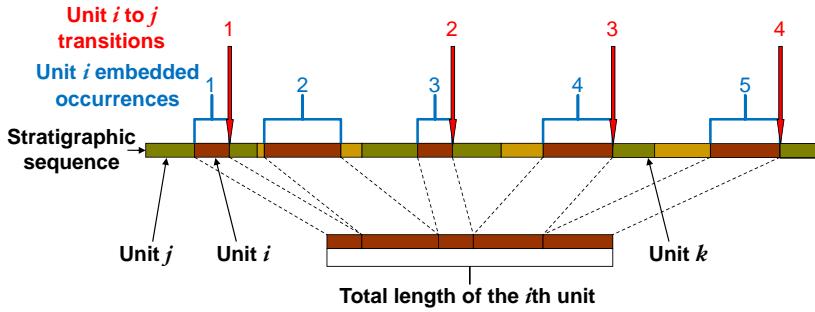


Fig. 1 Illustration identifying the difference between the determination of the mean facies length and mean transition length with respect to equation (23) and (24).

where \bar{l}_{ii} is equivalent to \bar{l}_i in equation (20) and \bar{l}_{ij} is the mean length of the *i*th stratigraphic unit between transitions from the *i*th to the *j*th stratigraphic unit. To clarify this linguistically,

$$\bar{l}_{ii} = \frac{\text{total length of the } i\text{th unit}}{\text{number of embedded occurrences of the } i\text{th unit}} \quad (23)$$

while

$$\bar{l}_{ij} = \frac{\text{total length of the } i\text{th unit}}{\text{number of transitions from the } i\text{th to the } j\text{th unit}} \quad i \neq j. \quad (24)$$

Collectively, we will refer to mean facies lengths and mean transition lengths as mean lengths.

In a 2-category Markov-chain model, \bar{l}_{ii} will be equivalent to \bar{l}_{ij} as the number of embedded occurrences (i.e. the number of transitions from the *i*th unit to any other unit) will equal the number of transitions from the *i*th to *j*th unit. This is also apparent by inspecting equation (9) considering equations (21) and (22), which can be expressed as

$$\mathbf{R} = \begin{bmatrix} -1/\bar{l}_{11} & 1/\bar{l}_{11} \\ p_1/(p_2\bar{l}_{11}) & -p_1/(p_2\bar{l}_{11}) \end{bmatrix}. \quad (25)$$

However, in a Markov-chain model with more than two categories, this will not necessarily be the case, as the number of embedded occurrences of the *i*th unit will not necessarily equal the number of transitions from the *i*th to the *j*th unit. This is illustrated in Figure 1 for a 3-unit stratigraphic sequence. The mean facies length represents the mean length of the *i*th unit, considering transitions to all other units (embedded occurrences), while the mean transition length only considers the *i* to *j* transitions, representing the mean length of the *i*th unit between *i* to *j* transitions. For further details, refer to Appendix A for derivations of \bar{l}_{ii} and \bar{l}_{ij} .

Utilizing equations (21) and (22), we can express equations (10), (14), and (17) in terms of mean lengths as

	Inverse Approach		
	A	B	C
Non-background spatial correlation constraint	proportionally random	symmetric	unconstrained
Governing equations	(17)/(28)	(14)/(27)	(10)/(26)
“True” model for inversion	1	2	2

Table 1 Information on inverse approaches. The “true” models are discussed in section 4.

$$\mathbf{R} = \begin{bmatrix} -\frac{1}{l_{11}} & \frac{1}{l_{12}} & \frac{1}{l_{11}} - \frac{1}{l_{12}} \\ \frac{1}{l_{21}} & -\frac{1}{l_{22}} & \frac{1}{l_{22}} - \frac{1}{l_{21}} \\ \frac{p_1}{l_{11}} - \frac{p_2}{l_{21}} & \frac{p_2}{l_{22}} - \frac{p_1}{l_{12}} & \frac{p_1}{l_{12}} - \frac{p_2}{l_{11}} + \frac{p_2}{l_{21}} - \frac{p_1}{l_{22}} \end{bmatrix}, \quad (26)$$

$$\mathbf{R} = \begin{bmatrix} -\frac{1}{l_{11}} & \frac{1}{l_{12}} & \frac{1}{l_{11}} - \frac{1}{l_{12}} \\ \frac{p_1}{p_2 l_{12}} & -\frac{1}{l_{22}} & \frac{1}{l_{22}} - \frac{p_1}{p_2 l_{12}} \\ \frac{p_1}{l_{11}} - \frac{p_1}{l_{12}} & \frac{p_2}{l_{22}} - \frac{p_1}{l_{12}} & -\frac{p_1}{l_{11}} + \frac{2p_1}{l_{12}} - \frac{p_2}{l_{22}} \end{bmatrix}, \quad (27)$$

and

$$\mathbf{R} = \begin{bmatrix} -\frac{1}{l_{11}} & \frac{p_2}{l_{11}(1-p_1)} & \frac{1}{l_{11}} \left(1 - \frac{p_2}{1-p_1} \right) \\ \frac{p_1}{l_{22}(1-p_2)} & -\frac{1}{l_{22}} & \frac{1}{l_{22}} \left(1 - \frac{p_1}{1-p_2} \right) \\ \frac{p_1}{l_{11}} - \frac{p_1 p_2}{l_{22}(1-p_2)} & \frac{p_2}{l_{22}} - \frac{p_1 p_2}{l_{11}(1-p_1)} & -\frac{p_1}{l_{11}} - \frac{p_2}{l_{22}} + \frac{p_1 p_2}{l_{11}(1-p_1)} + \frac{p_1 p_2}{l_{22}(1-p_2)} \end{bmatrix}, \quad (28)$$

respectively.

In the research presented here, we utilize equations (26), (27), and (28) to demonstrate the performance of the stochastic inverse approach with different levels of complexity using mean facies lengths and mean transition lengths as adjustable parameters. We will denote inverse approaches constrained by equations (26), (27), and (28) as inverse approaches A, B, and C, respectively (refer to Table 1).

Geologically-equally-probable realizations of the Markov-chain models are generated using simulated quenching (Carle et al, 1998) utilizing initial fields (images) generated by transition-probability based indicator co-kriging. The generation of initial fields by transition-probability based indicator co-kriging is analogous to the process of utilizing covariance-based indicator co-kriging in sequential indicator simulation (SIS) (Deutsch and Journel, 1992), except that the weighting coefficients are determined using a transition-probability based co-kriging set of equations as opposed to a covariance-based co-kriging set of equations (Carle, 1999). It is important to note that while the methods for computation and inference of Markov-chain and indicator cross-covariance models are fundamentally different, Carle and Fogg (1996) demonstrate that they are related as $C_{ij}(h) = p_i(t_{ij}(h) - p_j)$, where $C_{ij}(h)$ is the indicator cross-covariance of the i th and j th indicator variables at lag h .

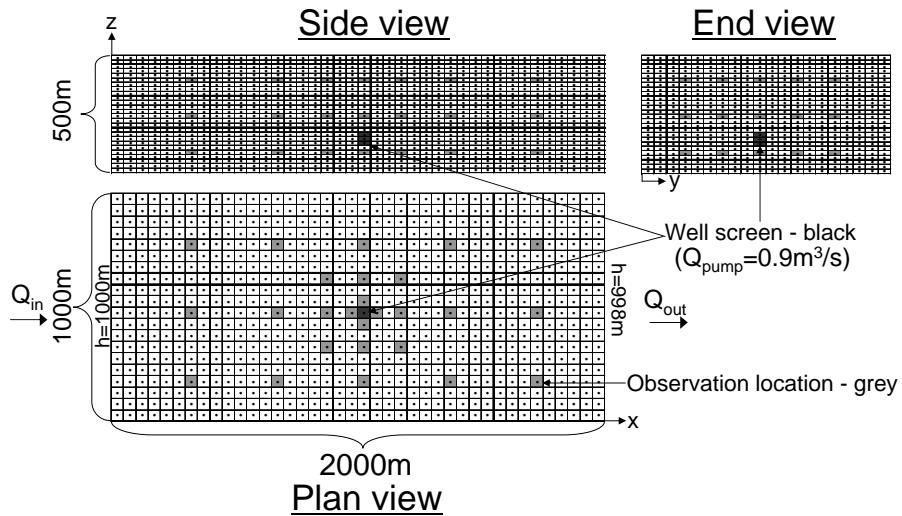


Fig. 2 Block-centered finite difference grid with points representing nodes and lines indicating the connectivity of the grid. Light gray cells denote observation locations and black cells denote the screen of the pumping well. Constant head boundaries are indicated on the plan view.

3.2 Steady-State Flow Simulation

We evaluate the feasibility of the proposed stochastic inverse approach utilizing 3-D synthetic models. The models contain three substantially-contrasting stratigraphic units with uniform hydraulic conductivities of $K_1 = 10^{-2}$ m/s, $K_2 = 10^{-5}$ m/s, and $K_3 = 10^{-8}$ m/s, where the subscript denotes the stratigraphic unit. Note the significant contrast in the facies conductivity; such a contrast is expected for facies with significantly different hydrogeological properties such as, for example, gravels, silts, and clays. The flow equations are solved on a block-centered finite-difference grid, presented in Figure 2, using FEM (Zyvoloski et al, 1997) with $40 \times 20 \times 25$ nodes in the x , y , and z directions, respectively, where x and y are horizontal directions and z is vertical. Node spacing is $50\text{ m} \times 50\text{ m} \times 20\text{ m}$, resulting in model dimensions of $2000\text{ m} \times 1000\text{ m} \times 500\text{ m}$ in the x , y , and z directions, respectively.

The flow simulations are performed under steady-state conditions. The flow is driven by a gradient in the x direction and a well pumping at the approximate horizontal center of the model. Constant-head boundaries are prescribed with a head drop of 2 m (hydraulic gradient of $i = 0.1\%$) from $x = 0\text{ m}$ to $x = 2000\text{ m}$. No-flow conditions are assigned for the rest of the model domain boundaries. The well is located at $(x = 1025\text{ m}, y = 525\text{ m})$ with a well screen depth, with respect to the top of the model, from 320 m to 380 m . The well pumping rate, Q_{pump} , is $0.9\text{ m}^3/\text{s}$ distributed evenly among the three grid cells designated as the well screen. Pressure-head data are collected at three depths from 26 observation wells for a total of 78 observation points within the model domain (refer to Figure 2). The head data is defined as a drawdown from an initial hydrostatic water level (initial head is $h_{init} = 1000\text{ m}$ at all

locations, except where $x = 2000\text{ m}$, where $h_{init} = 998\text{ m}$ (see Figure 2), maintaining confined conditions).

It should be noted that since the flow direction is predominantly along the x -axis and the model domain is elongated along the x -axis (which allows for better spatial representation of the facies), it can be expected that in the sensitivity and inverse analyses performed below, the transitional properties of the facies will be better characterized along the x -axis than the y -axis.

3.3 Inference of Characteristic Hydraulic Response of a Geostatistical Model

The proposed approach hinges on the assumption that alternate geostatistical models of stratigraphy will produce distinct hydraulic-response characteristics in the simulation model. To characterize the hydraulic response of a given Markov-chain model, we generate a series of geologically-equally-probable realizations and perform a flow simulation on each realization. Then we compute representative statistics of hydraulic responses. This requires that the representative statistics converge by a reasonable number of realizations. Given this consideration, we use the sample mean of the simulated drawdowns Δh and the sample median of the model output flux Q_{out} (refer to Figure 2 for its definition). We use the sample median instead of the sample mean for Q_{out} as the median converges more quickly due to its insensitivity to outlier values. We found that outlier values are not an issue with the Δh values, and that the sample mean converges as fast or faster than the sample median in most cases for Δh .

4 Facies Geostatistical Models

Two models were developed as “true” stratigraphic distributions for the sensitivity analysis and model inversion runs, including (1) a model with proportionally-random non-background spatial correlations (model 1) to demonstrate inverse approach A and (2) a model with symmetric non-background spatial correlations (model 2) to demonstrate inverse approaches B and C. Both models include 3 facies with volumetric proportions $p_1 = 0.2$, $p_2 = 0.4$, and $p_3 = 0.4$. An unconstrained model was not developed, as model 2 was used as the “true” stratigraphy for inverse approach C. It is assumed that the behavior of an unconstrained model inversion will behave similarly if the “true” stratigraphy contains symmetric or non-symmetric non-background spatial correlations, as symmetry is merely a special case for unconstrained spatial correlation. Stratigraphic conditioning data are assumed to be available from the pumping well and the 26 observation wells (Figure 2). This produces 675 conditioning data points for both models (27 wells \times 25 cells). The support scales of the conditioning data and the finite-difference grid are equal. At the same 26 observation wells, drawdown data are collected from 3 depths in each observation well (Figure 2); note that no drawdown is collected at the pumping well. The conditioning data were collected from a representative realization produced from a geostatistical model with the prescribed characteristics (e.g. proportionally-random (model 1) and symmetric (model 2) non-background spatial correlations) and a reasonable hydraulic response. In order to ensure that the pumping well screen is located in the high conductivity unit,

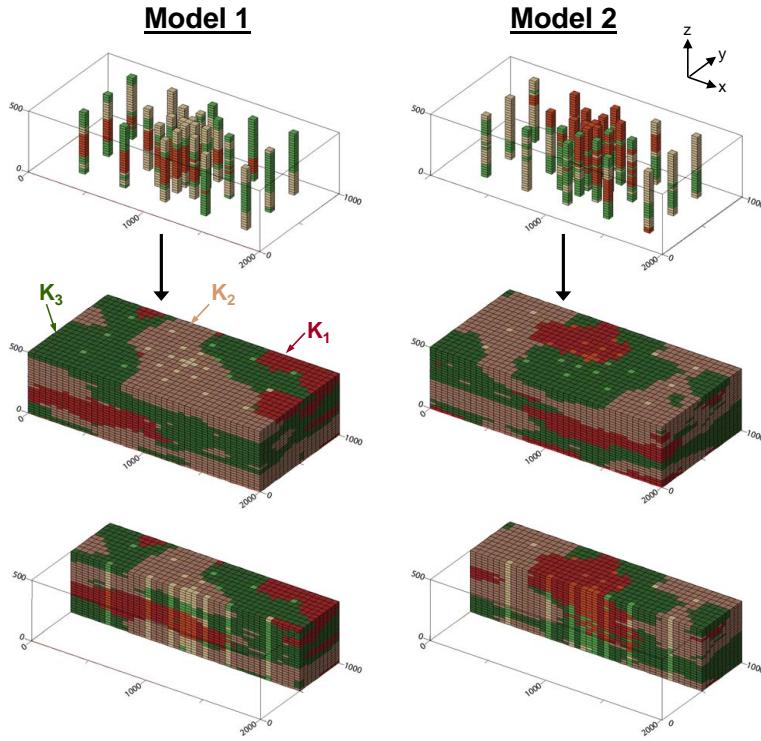


Fig. 3 Model 1 (proportionally-random non-background spatial correlations) and model 2 (symmetric non-background spatial correlations) conditioning data and example realization produced from the given geostatistical model and the conditioning data. Stratigraphic units are identified by color as unit 1 ($K_1 = 10^{-2} \text{ m/s}$) green, unit 2 ($K_2 = 10^{-5} \text{ m/s}$) tan, and unit 3 ($K_3 = 10^{-8} \text{ m/s}$) red. Volumetric proportions for units 1, 2, and 3 are 0.2, 0.4, and 0.4 respectively. Conditioning data are presented in a slightly lighter color than non-conditioning data. Axis labels are in meters.

the well screen was conditioned to K_1 for the development of the representative realization. Figure 3 presents the conditioning data for the models along with example realizations produced from the given geostatistical model and the conditioning data in each case. A cutout at $y = 500\text{m}$ of both models is also included to display the model interior where the conditioning data are highlighted, where it is apparent that the conditioning are honored in the realization.

The transition rate matrix for model 1 used as the “true” model to demonstrate inverse approach A, in both the x - and y -directions in inverse meters is

$$\mathbf{R}_{x,y;rnd} = \begin{bmatrix} -2.00 \times 10^{-3} & 1.00 \times 10^{-3} & 1.00 \times 10^{-3} \\ 3.33 \times 10^{-4} & -1.83 \times 10^{-3} & 1.50 \times 10^{-3} \\ 6.67 \times 10^{-4} & 1.33 \times 10^{-3} & -2.00 \times 10^{-3} \end{bmatrix}. \quad (29)$$

Equation (29) defines a Markov-chain model with proportionally-random non-background spatial correlations in accordance with equation (17) and (28), and is therefore within

the set of possible spatial correlations for inverse approach A. The corresponding mean facies and transition lengths in meters for model 1 are

$$\bar{\mathbf{L}}_{x,y;rnd} = \begin{bmatrix} 500 & 1000 & 1000 \\ 3000 & 545 & 667 \\ 1500 & 750 & 500 \end{bmatrix}, \quad (30)$$

where $\bar{\mathbf{L}} = (\bar{l}_{ij})$, $i, j = 1, \dots, N$. Due to the designation of category two as the background category and the constraint of proportional randomness for the non-background categories, equation (30) can be completely specified from two mean lengths (refer to equation (28)).

The transition rate matrix for model 2 used as the “true” model to demonstrate inverse approaches B and C, in both the x - and y -directions in inverse meters is

$$\mathbf{R}_{x,y;sym} = \begin{bmatrix} -2.00 \times 10^{-3} & 1.00 \times 10^{-3} & 1.00 \times 10^{-3} \\ 5.00 \times 10^{-4} & -2.00 \times 10^{-3} & 1.50 \times 10^{-3} \\ 5.00 \times 10^{-4} & 1.50 \times 10^{-3} & -2.00 \times 10^{-3} \end{bmatrix}. \quad (31)$$

Equation (31) defines a Markov-chain model with symmetric non-background spatial correlations in accordance with equation (14) and (27), and is therefore within the possible set of spatial correlations for inverse approach B. It can also be verified that equation (31) is in accordance with equations (10) and (26), equations defining unconstrained spatial correlations utilized by inverse approach C. The corresponding mean facies and transition lengths in meters for model 2 are

$$\bar{\mathbf{L}}_{x,y;sym} = \begin{bmatrix} 500 & 1000 & 1000 \\ 2000 & 500 & 667 \\ 2000 & 667 & 500 \end{bmatrix}. \quad (32)$$

Due to the designation of category two as the background category and the constraint of symmetry for the non-background categories, equation (32) can be completely specified from three mean lengths (refer to equation (27), while in the unconstrained case, four mean lengths must be specified (refer to equation (26)).

It is important to note the mean facies lengths (diagonal values in equations (30) and (32)) are equal and approximately 1/4 of the x -direction model domain (2000 m) and 1/2 of the y -direction model domain (1000 m). As a result, the x -direction spatial correlations are more fully represented than the y -direction.

The vertical, or z -direction mean facies and transition lengths are not included as adjustable parameters here, as it is assumed that the vertical Markov-chain model can be determined directly from the borehole data. As the volumetric proportions are independent of direction, it is assumed that they are available from the vertical Markov-chain model, and are not included as adjustable parameters. While in reality the vertical spatial correlation model will not be precisely known from the borehole data, and the uncertainty of this estimation would have to be considered, we perform the analysis assuming it can be in order to draw decisive conclusions with regard to the feasibility of constraining horizontal spatial correlation models using hydraulic data. For reference, the z -direction Markov-chain model transition rates for both models in inverse meters are

$$\mathbf{R}_z = \begin{bmatrix} -1.00 \times 10^{-2} & 5.00 \times 10^{-3} & 5.00 \times 10^{-3} \\ 1.67 \times 10^{-3} & -9.17 \times 10^{-3} & 7.50 \times 10^{-3} \\ 3.33 \times 10^{-3} & 6.67 \times 10^{-3} & -1.00 \times 10^{-2} \end{bmatrix}, \quad (33)$$

while the corresponding mean facies and transition lengths in meters are

$$\bar{\mathbf{L}}_z = \begin{bmatrix} 100 & 200 & 200 \\ 600 & 109 & 133 \\ 300 & 150 & 100 \end{bmatrix}. \quad (34)$$

4.1 Stochastic Inverse Approach

The goal of the inversion is to identify model parameters that minimize residuals between calibration targets and simulated values. The simulated values are inferred from a set of equally-probable geostatistical realizations, and include the sample mean of pressure responses at the observation locations (\bar{h}) and the sample median of the down-gradient model output flux (\bar{Q}_{out}). The adjustable parameters are the required mean facies lengths and mean transition lengths ($\bar{l}_{ij,\phi}$ $i, j = 1, \dots, N$, where ϕ indicates direction) defining the geostatistical properties of the realizations. Therefore, the inversion attempts to identify the most plausible geostatistical model constrained by observed geology concurrent with the available hydraulic data. Note that in the inverse method presented here, we assume that the facies conductivities are perfectly known. In practice, the uncertainty associated with hydraulic conductivity will have to be evaluated, and it may be desirable to include conductivity parameters in the inversion. This paper does not include conductivities as adjustable parameters so that we can focus on the feasibility of constraining horizontal stratigraphic uncertainty using hydraulic observations without other sources of uncertainty.

The stochastic inversion described here utilizes a Levenberg-Marquardt optimization strategy (Doherty, 1994) where the objective function is defined as

$$\Phi(\theta) = \sum_{i=1}^M (w_i * (s_i - \hat{s}_i(\theta)))^2, \quad (35)$$

where M is the number of calibration targets (78 mean drawdown observations (\bar{h}) and the median output flux (\bar{Q}_{out})), θ is a vector containing the current adjustable parameter values, w_i is the weight associated with the i th calibration target, s_i is the i th calibration target, and \hat{s}_i is the i th predicted value given θ . Weights are set proportional (scaled between 0 and 1) to the inverse of the variance of the calibration target values from the n realizations for the “true” parameter values in order to assign greater weight to calibration targets with lower variance; given a particular model, pressure responses from certain locations will converge more slowly than others, thereby introducing greater uncertainty in the model. Maps of the observation weights for models 1 and 2 are presented in Figure 4. By inspecting these maps, it is apparent that the weights are controlled by the boundary conditions and heterogeneities. This is apparent as some of the higher weights are located closer to the constant head boundaries

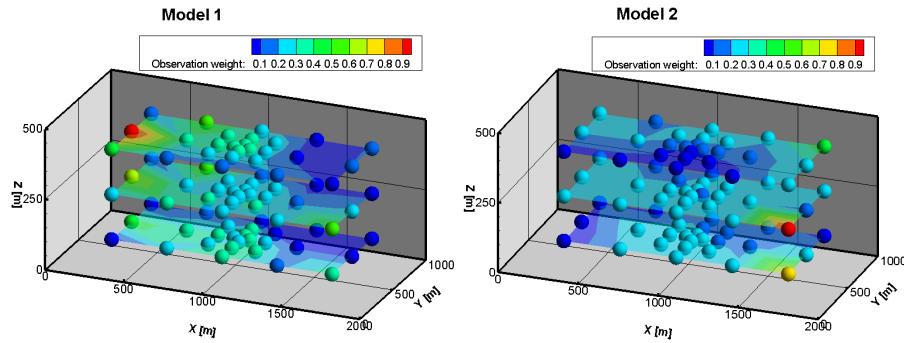


Fig. 4 Maps of inversion observation weights for models 1 and 2. The weights have been interpolated in between observation locations by triangulation in the three observation tiers to aid in visualization. Vertically exaggerated $\times 1.8$.

at $x = 0$ m and $x = 2000$ m, while the remaining distributed variability in the weights are assumed to be due to heterogeneities. The observation weights for \bar{Q}_{out} for models 1 and 2 are 9.43×10^{-7} and 6.49×10^{-7} , respectively, much smaller than the weights associated with the $\bar{\Delta}h$, indicating the greater variability of this calibration target. The small value for this weight indicates that residuals of \bar{Q}_{out} are not allowed to influence the inversions significantly. In practice, it may be desirable to omit \bar{Q}_{out} from the set of calibration targets as it is often difficult or impossible to measure or reasonably estimate its value.

In order to evaluate the mathematical feasibility of the proposed inversion, the calibration targets are collected as the sample mean and sample median of $\bar{\Delta}h$ values and \bar{Q}_{out} , respectively. This facilitates an analysis of the feasibility of the approach as a global minimum objective function value of zero exists at the “true” parameters. This differs from an actual application, where the calibration targets will be measurements from a single realization of a random field and conclusive knowledge of the global minimum is not available without extensive sampling.

5 Results and Discussion

5.1 Aquifer Response Statistical Convergence

The generation of stochastic realizations of a geostatistical model by conditional simulation constitutes a Monte Carlo sampling from a uniform distribution. While the realizations generated with respect to a given model are equally probable geostatistically, this is not necessarily the case hydraulically. Therefore, we use sample statistics from a set of realizations to characterize the aquifer response of a geostatistical model. An analysis of the convergence of these sample statistics is necessary in order to determine the number of realizations (n) required for a desired or necessary precision. Figure 5 presents $\bar{\Delta}h$ at the 78 observation locations as n increases for model 1, where n is presented on a log scale as suggested by Ballio and Guadagnini (2004).

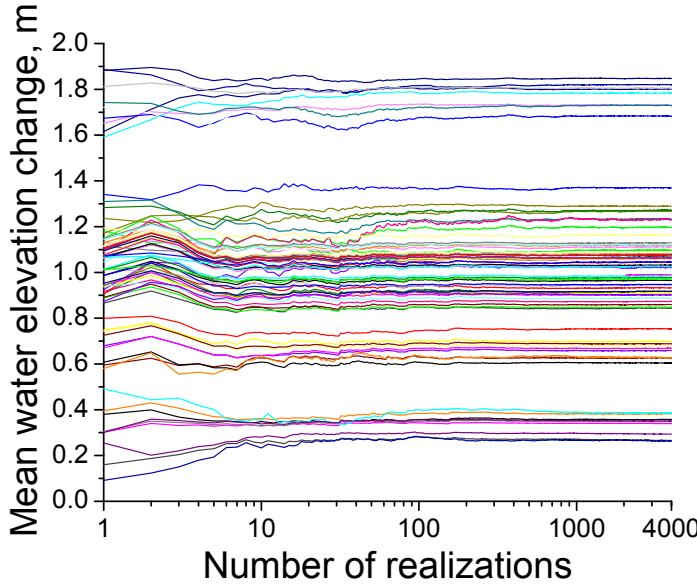


Fig. 5 Convergence of drawdown from undisturbed conditions (water elevation change) for the 78 observation locations as a function of the number of realizations plotted on log scale.

Presenting n on a log scale as opposed to a linear scale provides a better visualization of convergence as the uncertainty associated with inferring a statistic decreases at a decreasing rate as n increases. This is apparent by inspecting the root mean square (RMS) of Δh defined as

$$\overline{\Delta h}_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^n \overline{\Delta h}_i^2}, \quad (36)$$

which implies that the uncertainty associated with $\overline{\Delta h}$ scales with $1/\sqrt{n}$, a trend that is more easily visualized on a log scale.

We find that values of Δh collected at a location from a set of stochastic realizations obtained from model 1 fail the Shapiro-Wilks test for normality (R Development Core Team, 2005) at the 5% significance level at all observation locations. Therefore, a conservative estimate (over-estimate) of the confidence interval length for the expected value of Δh , $E[\Delta h]$, from n realizations can be obtained considering the Chebyshev inequality as

$$\Pr \left[\overline{\Delta h}_n - k \frac{\sigma}{\sqrt{n}} \leq E[\Delta h] \leq \overline{\Delta h}_n + k \frac{\sigma}{\sqrt{n}} \right] \geq 1 - \frac{1}{k^2} \quad (37)$$

where σ is the standard deviation of Δh , and k is the number of standard deviations. The right hand side of inequality (37) underestimates the probability that $\overline{\Delta h}_n$ is k standard deviations from the $E[\Delta h]$. Therefore, an overestimate of the confidence

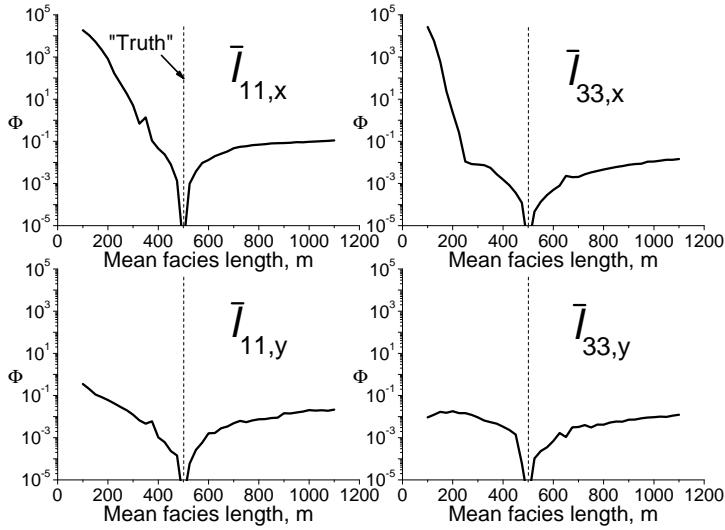


Fig. 6 Sensitivity analysis for inverse approach A where $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$ are plotted versus the objective function while holding the remaining adjustable parameters fixed at their “true” values. The objective function has been evaluated at 25 m intervals, half the model grid spacing.

limit interval length $|CL|$ can be obtained using the square root of the sample variance $\sqrt{S_n}$ to approximate σ as

$$|CL| = 2k \sqrt{\frac{S_n}{n}}. \quad (38)$$

Based on an approximation of σ by $\sqrt{S_{4000}}$, the average 95% $|CL|$ ($k = 4.472$) for Δh for the 78 observation locations at $n = 1000$ is 0.029 m with a standard deviation of 0.011 m. This indicates that we are inferring Δh within precision typical for field measurements of Δh . This ensures that the stochastic inversion will be limited by the precision of field measurements and not by the statistical inference of the hydraulic response. A larger number of realizations will provide smoother gradients for the objective function at the cost of computational efficiency. We find $n = 1000$ provides an appropriate compromise between inference precision and computational efficiency for the models 1 and 2.

5.2 Sensitivity analyses

Sensitivity analyses for inverse approaches A, B, and C are presented in Figures 6, 7, and 8, respectively. These figures contain plots of each of the adjustable parameters versus the objective function Φ (equation (35)) while holding the other adjustable parameters constant at their “true” values. The “true” value of the parameter of interest is indicated by a vertical dashed line in each plot, indicating the parameter value

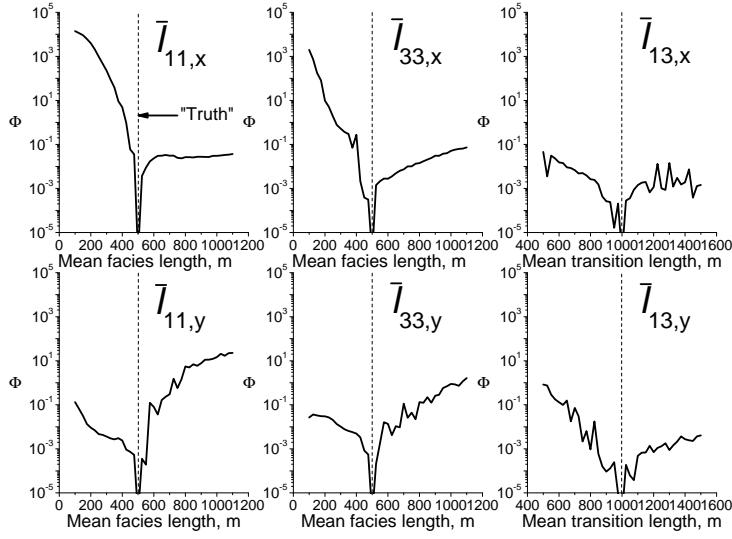


Fig. 7 Sensitivity analysis for inverse approach B where $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, and $\bar{l}_{13,y}$ are plotted versus the objective function while holding the remaining adjustable parameters fixed at their “true” values. The objective function has been evaluated at 25 m intervals, half the model grid spacing.

used to generate the calibration targets. A subscript ϕ will be appended to the mean lengths to indicate the direction of interest as $\bar{l}_{ij,\phi}$, where ϕ can be either x or y .

Figure 6 contains four plots, presenting the sensitivities of $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$ for inverse approach A. Due to the designation of category two as the background category and the constraint of proportional randomness for the non-background categories, the 14 remaining mean lengths can be determined from these four mean lengths (refer to equation (28)). By inspecting the plots in Figure 6, it is apparent that the x -direction mean lengths are more sensitive than the y -direction mean lengths. As discussed in the previous section this behavior is expected. This is due to the fact that the x -direction is parallel to the flow direction and has better representation in the model due to the domain size. It is also apparent that the sensitivities appear smooth at this discretization (half the grid spacing), implying that a gradient-based optimization strategy will function well on this problem.

Figure 7 presents the parameter sensitivities for inverse approach B. Plots are included for all adjustable parameters: $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, and $\bar{l}_{13,y}$. Due to the designation of facies two as the background category and the constraint of symmetry on the non-background categories, the twelve remaining mean lengths can be determined from these six mean lengths (refer to equation (27)). By comparing the sensitivities for inverse approach B (Figure 7) with those from inverse approach A (Figure 6), it is apparent that while the mean facies lengths in the x -direction are still relatively smooth, the y -direction mean facies lengths are significantly less smooth, and the mean transition length sensitivities are the least smooth.

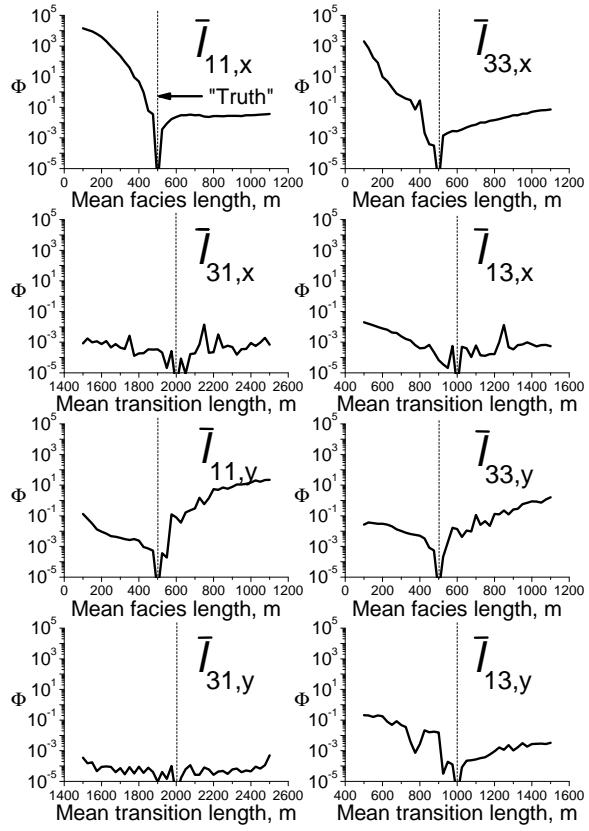


Fig. 8 Sensitivity analysis for inverse approach C where $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{31,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, $\bar{l}_{31,y}$, and $\bar{l}_{13,y}$ are plotted versus the objective function while holding the remaining adjustable parameters fixed at their “true” values. The objective function has been evaluated at 25 m intervals, half the model grid spacing.

Figure 8 presents the sensitivities for inverse approach C. Plots are included for all adjustable parameters: $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{13,x}$, $\bar{l}_{31,x}$, $\bar{l}_{11,y}$, $\bar{l}_{33,y}$, $\bar{l}_{13,y}$, and $\bar{l}_{31,y}$. As model 2 is used as the “truth” for inverse approaches B and C, the sensitivity plots of $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$ (mean facies lengths) in Figure 8 are the same as in Figure 7 and are repeated solely for reference. By inspecting the other plots (mean transition lengths) in Figure 8, it is apparent that the mean transition lengths in inverse approach C are relatively insensitive, especially $\bar{l}_{31,y}$.

The sensitivity analyses presented in Figures 6, 7, and 8 indicate that model inversions will benefit from constraints on the Markov-chain model (inverse approaches A and B). In cases where these constraints are not appropriate requiring the use of an unconstrained Markov-chain model (inverse approach C), the mean facies lengths are still relatively sensitive, while the mean transition lengths exhibit low non-smooth sensitivities. In general the shape of sensitivity curves are convex with relatively well defined minima at the “true” values. This suggests that a gradient-based method for

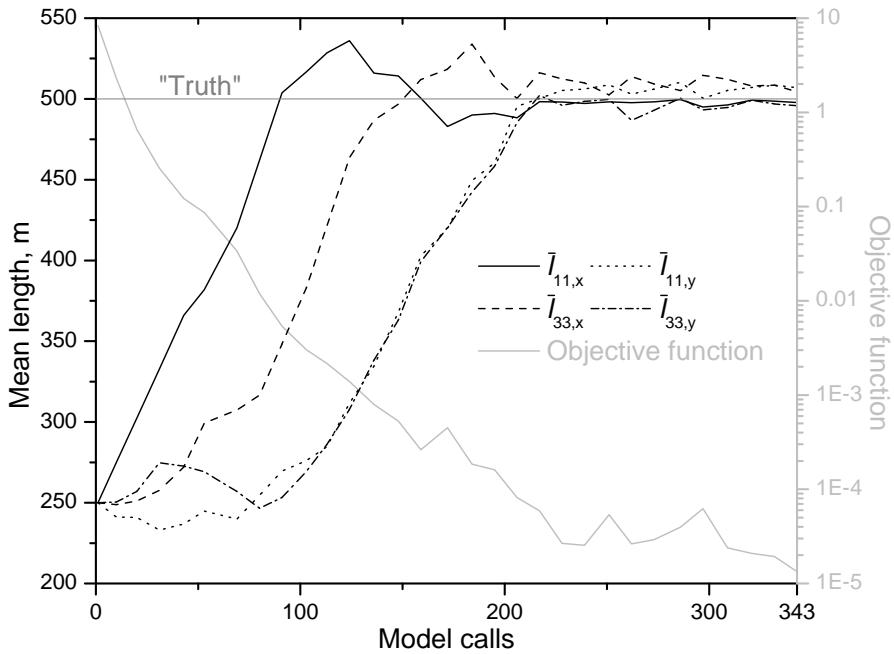


Fig. 9 Inverse approach A results. Parameter values and objective function are plotted as a function of the number of model calls.

model inversion can be successfully applied in these cases. In the next section, we demonstrate the effect of these sensitivities on the estimation of the mean lengths.

5.3 Inverse analyses

Model inversion results for inverse approach A are presented in Figure 9 and Table 2. The initial values of the adjustable parameters were set to 250 m from their “true” values of 500 m. However, the sensitivities in Figure 6 imply that the gradient-based optimization utilized here is expected to perform similarly if the initial values are defined within explored sensitivity ranges. Underestimates of the “true” mean facies lengths are used here. While not as significant in inverse approach A (utilizing model 1), we propose that choosing low estimates of mean facies lengths can aid the model inversion in inverse approaches B and C discussed below. Future work will address the effect of the initial values in the inverse process. As expected, the mean facies length of the high permeability category in the x -direction ($\bar{l}_{11,x}$) approaches the “true” value more quickly than the other mean facies lengths, followed by the mean facies length of the low permeability category in the x -direction ($\bar{l}_{33,x}$). This corresponds with the sensitivity analysis results presented in Figure 6, which indicate that the x -direction mean facies lengths are more sensitive than the y -direction mean facies lengths. To reiterate previous discussions, this is likely due to a combination of

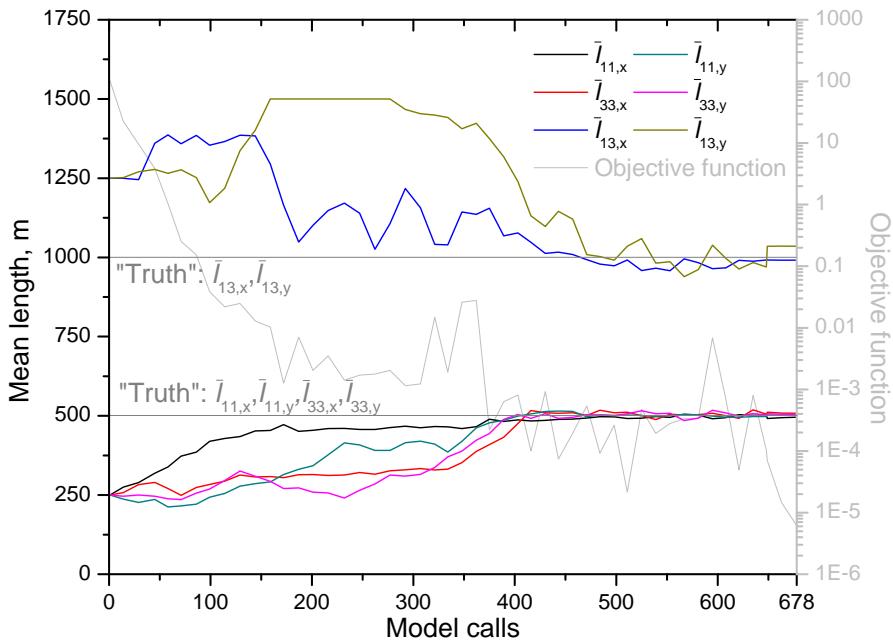


Fig. 10 Inverse approach B results. Parameter values and objective function are plotted as a function of the number of model calls.

factors including the longer x-direction distance compared with the y-direction and the fact that the x-direction mean facies lengths are parallel to the direction of flow, thereby affecting the hydraulic response of the aquifer by controlling the connectivity in the flow direction.

The lowest objective function was achieved at 343 model calls for inverse approach A. This corresponds to 31 Levenberg-Marquardt optimization iterations, which include multiple model calls for the calculation of derivatives and evaluation of Marquardt parameters. It was determined that no further progress would be achieved by 447 models calls, or 41 Levenberg-Marquardt optimization iterations.

Figure 10 and Table 2 presents the model inversion results for inverse approach B, where the mean transition lengths $\bar{l}_{13,x}$ and $\bar{l}_{13,y}$ are included. The transition rates associated with the corresponding diagonally-opposing mean transition lengths $\bar{l}_{31,x}$ and $\bar{l}_{31,y}$ are determined by the symmetry constraint. The initial values of the mean facies lengths were set to 250 m from their “true” value of 500 m, and the mean transition lengths were set to 1250 m from their “true” value of 1000 m. Based on Figure 7, it is apparent that mean facies length sensitivities (e.g. $\bar{l}_{11,x}$, $\bar{l}_{33,x}$, $\bar{l}_{11,y}$, and $\bar{l}_{33,y}$) are more smooth for values less than the “truth” in our synthetic example. This indicates that using a low estimate of the mean facies lengths will benefit the model inversion. Using this information, we have set the initial values of the mean facies lengths less than their “truth”. Future work will address the effect of the initial values in the inverse process.

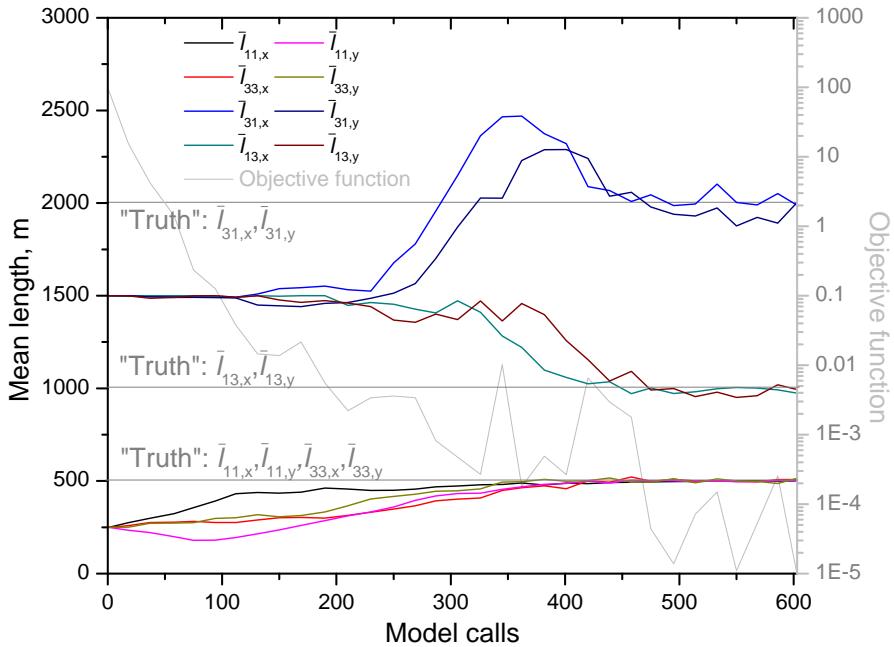


Fig. 11 Inverse approach C results. Parameter values and objective function are plotted as a function of the number of model calls.

The lowest objective function for the inverse approach B was achieved at 678 model calls, corresponding to 49 Levenberg-Marquardt optimization iterations. It was determined that no further progress would be achieved at 820 model calls, or 59 Levenberg-Marquardt optimization iterations.

Figure 11 and Table 2 presents the model inversion results for inverse approach C. As indicated by equation (26), in this case it is necessary to specify 4 mean lengths for the spatial correlation model in each direction. Therefore, there are eight mean length parameters plotted in Figure 11. The initial values of the mean facies lengths are set to 250 m from their “true” values of 500 m, while the mean transition lengths are set to 1500 m from their “true” values of 2000 m ($\bar{l}_{31,x}$ and $\bar{l}_{31,y}$) and 1000 m ($\bar{l}_{13,x}$ and $\bar{l}_{13,y}$). It is apparent that $\bar{l}_{11,x}$ converges towards its “true” value more quickly than the other parameters, as in inverse approaches A and B. It is also apparent that the mean transition lengths from the low conductivity to the high conductivity facies ($\bar{l}_{31,x}$ and $\bar{l}_{31,y}$) converge to their “true” values slower than the other parameters. This behavior is expected considering the relatively low, erratic sensitivities for these parameters in Figure 11. It is important to note that despite having poor sensitivities, these parameters do converge towards their “true” values during the inversion.

The lowest objective function for inverse approach C was achieved at 603 model calls, corresponding to 33 Levenberg-Marquardt optimization iterations. It was determined that no further progress would be achieved at 780 model calls, or 43 Levenberg-Marquardt optimization iterations.

	True Model		Inverse Approach		
	1	2	A	B	C
$\bar{l}_{11,x}$ m	500	500	497.8	491.9	500.5
$\bar{l}_{33,x}$ m	500	500	504.4	511.8	507.7
$\bar{l}_{13,x}$ m	1000	1000	—	991.7	974.3
$\bar{l}_{31,x}$ m	1500	2000	—	—	1988.0
$\bar{l}_{11,y}$ m	500	500	507.0	502.4	500.3
$\bar{l}_{33,y}$ m	500	500	495.8	506.6	514.9
$\bar{l}_{13,y}$ m	1000	1000	—	1035.5	993.5
$\bar{l}_{31,y}$ m	1500	2000	—	—	2004.0
Obj. Function	—	—	1.34×10^{-5}	2.14×10^{-5}	9.68×10^{-6}
Model calls	—	—	343	678	603
L-M iterations	—	—	31	49	33
DOF	—	—	75	73	71
AIC	—	—	105.7	179.0	193.7
BIC	—	—	96.3	152.7	174.7

Table 2 “True” parameter values for models 1 (equation (29)) and 2 (equation (31)) and the optimized parameter values for inversions utilizing inverse approaches A, B, and C (refer to equations (28), (27), and (26), respectively). Model 1 is the “true” model for inverse approach A, while model 2 is the “true” model for inverse approaches B and C. Corresponding minimum objection function values, number of model calls and Levenberg-Marquardt (L-M) optimization iterations for convergence, degrees of freedom (DOF), Akaike identification criterion (AIC), and the Bayesian information criterion (BIC) are also listed.

Table 2 lists the degrees of freedom (DOF) (number of calibration targets are 79 in all three cases), Akaike information criterion (AIC), and Bayesian information criterion (BIC) (Carrera and Neuman, 1986). It is apparent that the degrees of freedom decreases from inverse approach A to B to C, as the number of adjustable parameters increases. As inverse approach A is demonstrated using model 1 and inverse approaches B and C are demonstrated using model 2, the values of AIC and BIC cannot be compared directly between these inverse approaches. However, as inverse approaches B and C both use the same model as the “truth”, differences in values of AIC and BIC are meaningful. It is apparent that inverse approach B is considered the more appropriate approach for both AIC and BIC. This indicates that the inclusion of additional adjustable parameters in inverse approach C is not warranted even though it has reduced the objective function lower than inverse approach B in this case. It should be noted, however, that the constraint of symmetric spatial correlations in inverse approach B limits its ability to explore geostatistical characteristics, while inverse approach C is unconstrained with respect to the Markov-chain model.

As the approach presented here utilizes a statistical characterization of the hydraulic response of a geostatistical model, as opposed to a single realization as in Harp et al (2008), the results present a stochastic analysis, where an uncertainty analysis is incorporated in the inverse method. Figure 12 presents facies probability maps for each stratigraphic unit based on the inversion results for inverse approach C. The facies probabilities are calculated as the one-location marginal probabilities of the set of realization from the optimized Markov-chain geostatistical model. In order to visualize the locations where the inversion has identified that the probability of a particular facies has equal or greater probability than the other facies, maps are also presented displaying model cells with probability equal or greater than 0.5 only.

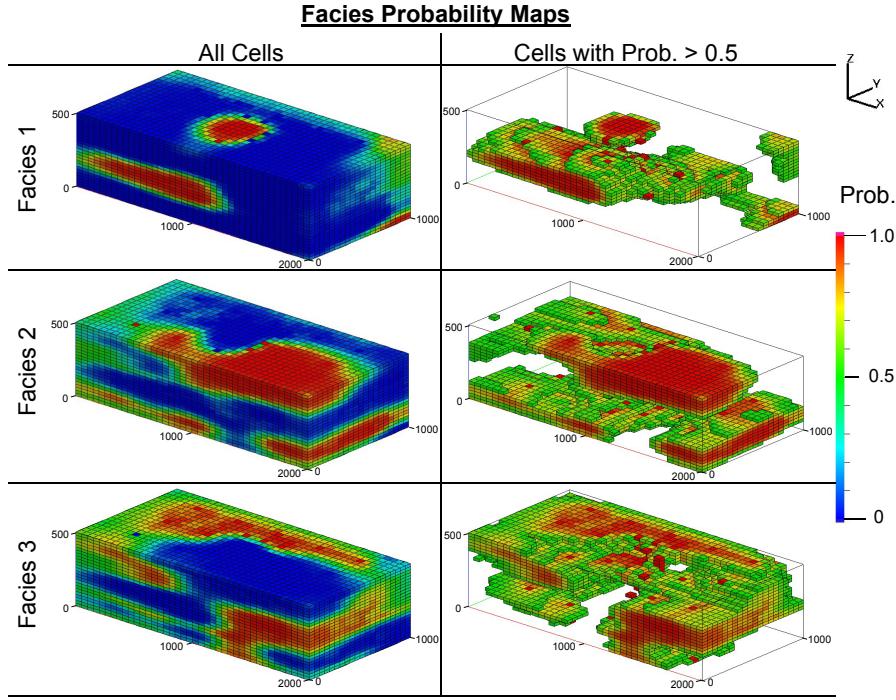


Fig. 12 Facies probability maps resulting from applying inverse approach C to model 2. The left column presents the complete probability maps for each facies, while the right column presents the cells with probability of 0.5 or greater.

6 Conclusions

While much effort has been directed towards geostatistical inversion methods utilizing covariance-based geostatistics, little effort has been spent on similar approaches employing Markov-chain geostatistics. This paper demonstrates the feasibility of a stochastic inversion utilizing a Markov-chain model of spatial variability. The approach hinges on the use of inverse transition rates as adjustable parameters, where inverse auto-transition rates are defined as mean facies lengths and inverse cross-transition rates are defined as mean transition lengths.

Simplifying constraints to the Markov-chain model allowing reductions in the number of adjustable parameters are explored within the stochastic inversion. These simplifying constraints include proportionally-random spatial correlations (inverse approach A) and symmetric correlations (inverse approach B) in non-background categories. In addition to these constrained models of spatial correlation, an unconstrained Markov-chain model was evaluated as well (inverse approach C). Inverse approaches A and B performed well, indicating that these spatial correlation constraints can be useful in cases where these simplifications to a Markov-chain model are expected to be valid, or as references to be used in conjunction with an uncon-

strained analysis (inverse approach C) to determine the level of deviation from these simplified models. The results using inverse approach C demonstrate that it is not necessary to constrain the Markov-chain model to obtain a successful inverse analyses. AIC and BIC both indicate that inverse approach B is more appropriate than C, however, this neglects to recognize that the “true” model in this case has symmetric non-background spatial correlations. If this analysis were repeated with unconstrained spatial correlations, we should expect different results. In fact, in this case, the “true” model would not be in the set of possibilities for inverse approach B. Given the fact that stratigraphic spatial correlations are not symmetric in general, we advocate the use of inverse approach C to identify plausible stratigraphic characteristics and inverse approaches A and B to evaluate the level of deviation from simplified spatial correlations. The less constraints on the geostatistical model in the inverse process allows for characterization of wider ranges of aquifer heterogeneities.

Using the Chebyshev inequality to provide an overestimate of the length of the confidence limits of the sample mean of the simulated drawdowns, we determine that 1000 realizations provide adequate convergence at the scale of our synthetic models. In this way, we ensure that limitations are due to measurement imprecision, not sample statistic inference imprecision. It is important to note that a larger number of realizations is expected to produce smoother sensitivities at the cost of computational efficiency.

A single forward model run (including facies structure generation and flow simulation) takes approximately 30 seconds on a 2.4 GHz AMD Opteron processor. Therefore, a single evaluation of the objective function, or model call (1000 forward model runs), takes approximately 20 wall clock hours on the same processor. The sensitivity analyses were performed at half the horizontal grid spacing (25 m) over a range of 1000 m, resulting in 41 model calls. Therefore, a sensitivity analysis for a single parameter requires approximately 850 wall clock hours. As the forward model runs can be performed independently, the model call computation time can be greatly reduced utilizing multiple processors to evaluate individual forward model runs concurrently.

Sensitivity analyses indicate that inverse approach A provides parameters with well-behaved sensitivities. The inverse approach B produces decent sensitivities for all adjustable parameters, with the poorest sensitivities for the mean transition lengths. The mean facies length sensitivities are decent for inverse approach C, while significantly poorer sensitivities are observed for the mean transition length sensitivities. Nevertheless in general the shape of sensitivity curves are convex with relatively well defined minima at the “true” values. This suggests that a gradient-based optimization strategy for model inversion can be successfully applied in these cases. In general, the x -direction mean length parameters are more sensitive than their corresponding y -direction mean length parameters. This is likely the result that the x -direction is parallel to the flow direction, where the x -direction mean lengths will provide more connectivity information than y -direction mean lengths, and that the model domain is longer in the x -direction than the y -direction, allowing a better representation of x -direction mean lengths than y -direction mean lengths.

The example inversion runs demonstrate that despite some low sensitivities, all parameter do converge towards their “true” values for the three inverse approaches. This demonstrates the feasibility of a gradient-based optimization strategy on the

proposed stochastic inversion framework. In contrast, previous analyses (Harp et al, 2008) were performed using a global search method called AMALGAM (Vrugt and Robinson, 2007) which allows for more detailed search for an optimal solution, but at a significantly greater computational cost in general. Further research is needed to extend this framework to allow the concurrent estimation of hydraulic conductivity, variable and/or non-uniform hydraulic conductivity, and the inclusion of transport observations as calibration targets.

APPENDIX

A Derivation of mean facies length and mean transition lengths

As demonstrated by Carle and Fogg (1996), mean facies length can be derived considering two categories. In the following, we review their derivation and present the derivation of the mean transition length for a 1-D Markov-chain model.

Considering the definition of an auto-transition rate, presented in equation (12), and the definition of the mean facies length, presented in equation (21), the auto-transition probability can be related to the mean facies length as

$$\frac{\partial t_{ii}(0)}{\partial h} = r_{ii} = -\frac{1}{l_{ii}}. \quad (39)$$

Using the definition of a derivative, equation (39) can be expressed as

$$\lim_{h \rightarrow 0} \frac{t_{ii}(h) - t_{ii}(0)}{h} = -\frac{1}{l_{ii}}. \quad (40)$$

Recognizing that $t_{ii}(0) = 1$ (i.e. probability of an auto-transition at lag 0 is 1) and assuming a 2-category Markov-chain model, $t_{ij}(h)$ can be expressed as

$$t_{ij}(h) = 1 - t_{ii}(h) = t_{ii}(0) - t_{ii}(h). \quad (41)$$

To consider cases with greater than two categories, the j th category can represent the union of all categories other than the i th category. Substituting equation (41) into equation (40) produces

$$\lim_{h \rightarrow 0} \frac{t_{ij}(h)}{h} = -\frac{1}{l_{ii}}. \quad (42)$$

The transition probability can be defined using the one-location marginal probability p_i (facies volumetric proportion) and two-location joint probability $p_{ij}(h)$ (Carle and Fogg, 1996) as

$$t_{ij}(h) = \frac{p_{ij}(h)}{p_i}. \quad (43)$$

The two-location joint probability $p_{ij}(h)$ can be defined as

$$p_{ij}(h) = \frac{T_{ij}(h)}{N(h)} \quad (44)$$

where $T_{ij}(h)$ is the number of transitions encountered from the i th to the j th category at lag h in a particular direction and $N(h)$ is the total number of lag intervals. Substituting equation (44) into equation (43) and dividing by h produces

$$\frac{t_{ij}(h)}{h} = \frac{T_{ij}(h)}{p_i N(h) h}. \quad (45)$$

As $N(h) * h$ is the total length, $p_i N(h) h$ is the total length of the i th category. Since categories i and j are assumed to be mutually-exclusive exhaustively-defined categories, $T_{ij}(h)$ is equivalent to the number of embedded occurrences of the i th category. Therefore, substituting equation (45) into equation (42) produces

$$\lim_{h \rightarrow 0} \frac{T_{ij}(h)}{p_i N(h) h} = \frac{\text{Number of embedded occurrences of the } i\text{th unit}}{\text{Total length of the } i\text{th unit}}. \quad (46)$$

Which leads to a linguistic definition of the mean facies length as

$$\bar{l}_{ii} = \frac{\text{Total length of the } i\text{th unit}}{\text{Number of embedded occurrences of the } i\text{th unit}}. \quad (47)$$

The mean transition length can be derived similarly starting from the cross-transition rate definition presented in equation (12) and the definition of a mean transition length presented in equation (22) as

$$\frac{\partial t_{ij}(0)}{\partial h} = r_{ij} = \frac{1}{\bar{l}_{ij}}. \quad (48)$$

Using the definition of a derivative, equation (48) can be expressed as

$$\lim_{h \rightarrow 0} \frac{t_{ij}(h) - t_{ij}(0)}{h} = \frac{1}{\bar{l}_{ij}}. \quad (49)$$

Recognizing that $t_{ij}(0) = 0$ (i.e. probability of a cross-transition at lag 0 is 0) equation (49) can be expressed as

$$\lim_{h \rightarrow 0} \frac{t_{ij}(h)}{h} = \frac{1}{\bar{l}_{ij}}, \quad (50)$$

where $t_{ij}(h)/h$ has been defined in equation (45). This demonstrates that in a 2-category Markov-chain model, $\bar{l}_{ii} = \bar{l}_{ij}$, as presented in equation (25). However, in the derivation of the mean transition length, we are not restricted to 2-category Markov-chain models. Recalling that $T_{ij}(h)$ is the number of transitions from the i th to j th unit at lag h , \bar{l}_{ij} can be defined linguistically as

$$\bar{l}_{ij} = \frac{\text{Total length of the } i\text{th unit}}{\text{Number of transitions from the } i\text{th to the } j\text{th unit}}. \quad (51)$$

This indicates that \bar{l}_{ij} is the mean length of the i th category between transitions from the i th to the j th category. This should not be confused with the total length between transitions from the i th to the j th category as the lengths of intervening segments are not considered in the mean transition length (refer to Figure 1).

Acknowledgements The first author is supported as a Graduate Research Assistant by Los Alamos National Laboratory. The authors are grateful for the ideas provided by Zhenxue Dai during the early stages of development of the presented methodologies. Christophe Frippiat, Toshihiro Sakaki, and Tissa Illangasekare generously provided data, specifications, and analyses from a synthetic aquifer experiment utilized in the early stages of development of the presented methodologies. The authors would also like to thank Zhenxue Dai, Zhiming Lu, and members of the first author's Ph.D. advisory committee (Bruce Thomson, Gary Weissmann, and John Stormont) for comments and suggestions on early drafts of this paper.

References

- Balio F, Guadagnini A (2004) Convergence assessment of numerical Monte Carlo simulations in ground-water hydrology. *Water Resources Research* 40:W04603, DOI 10.1029/2003WR002876
- Cardiff M, Kitandis P (2009) Bayesian inversion for facies detection: An extensible level set framework. *Water Resources Research* 45:W10416, DOI 10.1029/2008WR007675
- Carle S (1999) T-PROGS: Transition Probability Geostatistical Software, Version 2.1. Hydrological Sciences Graduate Group, University of California, Davis
- Carle S, Fogg G (1996) Transition probability-based indicator geostatistics. *Mathematical Geology* 28(4):453–476
- Carle S, Fogg G (1997) Modeling spatial variability with one and multidimensional continuous-lag markov chains. *Mathematical Geology* 29(7):891–918

- Carle S, LaBolle E, Weissmann G, VanBrocklin D, Fogg G (1998) Conditional simulation of hydrofacies architecture: A transition probability/markov approach. In: Fraser O, Davis J (eds) Hydrogeological Models of Sedimentary Aquifers SEPM Concepts in Hydrogeol. Environ. Geol., Soc. for Sediment. Geol., Tulsa, OK, pp 147–170
- Carrera J, Neuman S (1986) Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resources Research* 22(2):199–210
- Dai Z, Wolfsberg A, Lu Z, Ritzi Jr R (2007) Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix. *Geophysical Research Letters* 34:L20406, DOI 10.1029/2007GL031608
- Deutsch C, Journel A (1992) *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press, New York
- Doherty J (1994) *PEST Model-Independent Parameter Estimation*. Watermark Computing, Corinda, Australia
- Fienen M, Kitanidis P, Watson D, Jardine P (2004) An application of Bayesian inverse methods to vertical deconvolution of hydraulic conductivity in a heterogeneous aquifer at Oak Ridge National Laboratory. *Mathematical Geology* 36(1):101–126
- Gégo E, Johnson G, Hankins M (2001) An evaluation of methodologies for the generation of stochastic hydraulic conductivity fields in highly heterogeneous aquifers. *Stochastic Environmental Research and Risk Assessment* 15:47–64
- Gómez-Hernández J, Hendricks Franssen HJ, Sahuquillo A (2003) Stochastic conditional inverse modeling of subsurface mass transport: A brief review and the self-calibrating method. *Stochastic Environmental Research and Risk Assessment* 17:319–328, DOI 10.1007/s00477-003-0153-5
- Harp D, Dai Z, Wolfsberg A, Vrugt J, Robinson B, Vesselinov V (2008) Aquifer structure identification using stochastic inversion. DOI 10.1029/2008GL033585
- Hendricks Franssen H, Gómez-Hernández J (2002) 3d inverse modelling of groundwater flow at a fractured site using a stochastic continuum model with multiple statistical populations. *Stochastic Environmental Research and Risk Assessment* 16:155–174, DOI 10.1007/s00477-002-0091-7
- Huang C, Hu B, Li X, Ye M (2009) Using data assimilation method to calibrate a heterogeneous conductivity field and improve solute transport prediction with an unknown contamination source. *Stochastic Environmental Research and Risk Assessment* 23:1155–1167, DOI 10.1007/s00477-008-0289-4
- Journel A (1983) Nonparametric estimation of spatial distribution. *Mathematical Geology* 15(3):445–468
- LaBolle E, Fogg G (2001) Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system. *Transport in Porous Media* 42(1-2):155–179
- de Marsily G, Delay F, Gonçalvés J, Renard P, Teles V, Violette S (2005) Dealing with spatial heterogeneity. *Hydrogeology Journal* 13(1):161–183
- Matheron G (1967) *Éléments pour une théorie des milieux poreux [elements for a theory of porous media]*. Masson, Paris
- Neuman S (2004) Stochastic groundwater models in practice. *Stochastic Environmental Research and Risk Assessment* 18:268–270, DOI 10.1007/s00477-004-0192-6
- Orton T, Lark R (2007) Accounting for the uncertainty in the local mean in spatial prediction by bayesian maximum entropy. *Stochastic Environmental Research and Risk Assessment* 21:773–784
- R Development Core Team (2005) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org>, ISBN 3-900051-07-0
- Ross S (1993) *Introduction to Probability Models*, 5th edn. Academic Press, San Diego, California
- Serre M, Christakos G, Li H, Miller C (2003) A BME solution of the inverse problem for saturated groundwater flow. *Stochastic Environmental Research and Risk Assessment* 17, DOI 10.1007/s00477-003-0156-2
- Turk G (1982) Markov chains: Letters to the editor. *Mathematical Geology* 14(5):539–542
- Vistelius A (1949) On the question of the mechanism of formation of strata. *Dokl Akad Nauk SSSR* 65(2):191–194
- Vrugt J, Robinson B (2007) Improved evolutionary optimization from genetically adaptive multimethod search. *PNAS* 104(3):708–711, DOI 10.1073/pnas.0610471104
- Weissmann G, Fogg G (1999) Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. *Journal of Hydrology* 226:48–65
- Zanini A, Kitanidis P (2009) Geostatistical inversions for large-contrast transmissivity fields. *Stochastic Environmental Research and Risk Assessment* 23, DOI 10.1007/s00477-008-0241-7

Zhang C, Li W (2008) A comparative study of nonlinear Markov chain models for conditional simulation of multinomial classes from regular samples. *Stochastic Environmental Research and Risk Assessment* 22:217–230, DOI 10.1007/s00477-007-0109-2

Zyvoloski G, Robinson B, Dash Z, Trease L (1997) 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