A Markov chain model for characterizing medium heterogeneity and sediment layering structure

Ming Ye[1] and Raziuddin Khaleel[2]

Received 15 February 2008; revised 12 June 2008; accepted 24 June 2008; published 20 September 2008.

[1] By leveraging use of "soft" data (e.g., initial moisture content, θ_i), this study applies the transition probability (TP) based Markov chain (MC) model to sediment textural classes for characterizing the medium heterogeneity and sediment layering structure. The TP/MC method is evaluated by simulating the vadose zone moisture movement at a field site, where the stratigraphy consists of imperfectly stratified soil layers. Soil heterogeneity is characterized via spatial variability of the geometry of soil textural classes. When the θ_i measurements, which carry signature about medium heterogeneity and stratigraphy, are not included in the TP/MC model, it is not possible to identify the horizontal TP. The θ_i measurements, when transformed into soil classes, are necessary in mapping the soil layering structure prevalent at the site. The soil hydraulic parameters for each soil class are treated deterministically and are estimated on the basis of core samples. To evaluate uncertainty in characterizing geometry of the soil classes, multiple conditional realizations of the soil classes are generated. A Monte Carlo simulation shows that the simulated mean moisture contents agree well with corresponding field observations. The observed splitting of the moisture plume in a coarse sand layer that is sandwiched between two fine-textured layers, the southeastward movement of the plume during the redistribution period, and the near-zero fluid flux below the bottom fine layer are adequately simulated. Spatial variability of the field-measured moisture content is sufficiently captured by the 95% confidence intervals calculated from the Monte Carlo simulations. Investigating the effect of data conditioning on the simulated results shows that a reduction of conditioning data does not necessarily deteriorate simulation results if other conditioning data exist within the mean length of the soil classes. The TP/MC method is flexible so that other types of site characterization data (e.g., geophysical data) can be incorporated as they become available.

Citation: Ye, M., and R. Khaleel (2008), A Markov chain model for characterizing medium heterogeneity and sediment layering structure, *Water Resour. Res.*, *44*, W09427, doi:10.1029/2008WR006924.

1. Introduction

[2] Accurate simulation of flow and contaminant transport in heterogeneous media is often hampered by the difficulty in characterizing medium heterogeneity and the inherent variability of sediment layering structure prevalent in sedimentary environments. The difficulty partly arises from lack of site characterization data (e.g., core samples, lithologic boring logs, and wireline logs). Because the data are typically obtained from widely spaced boreholes, they can provide relatively adequate information about heterogeneity in the vertical direction, but not necessarily in the horizontal direction. Interpolating information between boreholes is thus needed for characterizing the lateral heterogeneity, which, however, is a longstanding challenge in subsurface characterization. In order to resolve the problem of sparse characterization data, it is often suggested using multiple types of "hard" and "soft" data, which, according to *Koltermann and Gorelick* [1996], are, respectively, the direct and indirect measurements or observations of rock and sediment properties. This study characterizes medium heterogeneity via a description of spatial variability of sediment textual classes and by using a transition probability (TP) based Markov chain (MC) geostatistical method. The objectives of this study are (1) to investigate the feasibility of using "soft" data to improve the TP/MC model, especially in the horizontal direction to better characterize the sediment layering structure; (2) to evaluate the effect of the improvement in enhancing numerical simulation of moisture movement in heterogeneous media; and (3) to explore the effect of conditioning data (especially the "soft" data) on enhancement of characterization and numerical simulations.

[3] Geostatistical methods have been used extensively in characterizing subsurface heterogeneity and associated uncertainty, as reviewed in several books [e.g., *Journel and Huijbregts*, 1978; *Isaaks and Srivastava*, 1989; *Cressie*, 1991; *Kitanidis*, 1997; *Goovaerts*, 1997; *Deutsch and Journel*, 1998; *Chiles and Delfiner*, 1999; *Christakos*, 2000] and review articles [e.g., *Neuman*, 1980; *Haldorsen*

[1]Department of Scientific Computing, Florida State University, Tallahassee, Florida, USA.
[2]Fluor Government Group, Richland, Washington, USA.

0043-1397/08/2008WR006924$09.00

and Damsleth, 1990; *Tyler et al.*, 1994; *Koltermann and Gorelick*, 1996; *Journel et al.*, 1998]. Conventional geostatistical methods (e.g., kriging) use the two-point variogram (or covariance) to describe spatial structures; recently the three-point variogram has been advocated [*Strebelle*, 2002; *Krishnan and Journel*, 2003; *Journel and Zhang*, 2006; *Feyen and Caers*, 2006]. An alternate geostatistical method, one that does not use a variogram, uses transition probability and Markov chains (TP/MC) to describe spatial structures of categorical data [e.g., *Carle and Fogg*, 1996, 1997; *Fogg et al.*, 1998; *Carle*, 1999; *Weissmann et al.*, 1999; *Li et al.*, 1999; *Ritzi*, 2000; *Elfeki and Dekking*, 2001; *Lu and Zhang*, 2002; *Park et al.*, 2004; *Dai et al.*, 2005; *Sivakumar et al.*, 2005; *Maji et al.*, 2006; *Zhang et al.*, 2006; *Li*, 2007a, 2007b; *Lee et al.*, 2007; *Dai et al.*, 2007; *Sun et al.*, 2008]. An advantage of the TP/MC method relative to the conventional geostatistical methods is that it can incorporate geological information such as mean lengths and directional orientation of lithofacies [*Carle and Fogg*, 1996, 1997]. For example, in a study of tritium transport at a site in Rulison, Colorado, because the mean horizontal length of shale could not be identified from the borehole data, the mean horizontal length estimates for an outcrop were used [*Cooper et al.*, 2007]. Outcrop data were also used by *Dai et al.* [2005] for a better definition of transition probability. Recent studies [*Maji et al.*, 2006; *Lee et al.*, 2007] demonstrate that the TP/MC method outperforms conventional geostatistical methods by reconstructing the linkage between lithology and contaminant plumes.

[4] In spite of the above mentioned advantages, usefulness of the TP/MC method for characterizing heterogeneity is limited, like other geostatistical methods, by availability of data to accurately estimate TP/MC parameters: mean lengths for each category in the x, y, and z directions (or horizontal and vertical directions). Estimating the mean horizontal length is particularly difficult, when only a few boreholes are available. To resolve this problem, *Weissmann et al.* [1999] used a soil survey map to estimate mean horizontal lengths for gravel, sand, muddy sand, and mud at the King River alluvial fan. However, it was not investigated whether the improved estimation enhances the numerical simulation of groundwater flow and contaminant transport at the site. On the other hand, since the correlation between the TP/MC parameters and the soil survey or outcrop measurements is not straightforward, a geologic knowledge–based subjective adjustment is needed. This raises questions regarding the general applicability of the TP/MC method [*Zhang et al.*, 2006]. One thus needs other "soft" data that can better reveal subsurface structure (not just surficial exposure), as well as a more objective method of using the data.

[5] This study presents a statistical method wherein the initial moisture content (θ_i,%) is used as the "soft" data to improve the TP/MC parameter estimates; the effect of the improvement is evaluated via numerical simulation of vadose zone moisture flow at a field site. The study area is the Sisson and Lu site (*Sisson and Lu* [1984], hereinafter referred to as S-L site) at the U.S. Department of Energy's Hanford Site in Washington State. Field investigation and core sampling [*Last and Caldwell*, 2001; *Last et al.*, 2001] revealed the apparent sediment layering structure in site lithostratigraphy; the core samples are, however, insufficient to represent the layering structure throughout the S-L site. However, the site has an extensive θ_i database, which was measured about one month before a series of field injection experiments were conducted [*Gee and Ward*, 2001]. Since the θ_i measurements carry signature of medium heterogeneity, as discussed later, they play a key role in characterizing heterogeneity, especially the imperfectly stratified layers. The field injection experiments are well simulated using the TP/MC method, which, however, is not possible without incorporating the θ_i database. It is worth mentioning that the statistical method of incorporating the θ_i database can be applied to other types of site characterization data. One data type is geophysical data which is more readily accessible and often can reveal whether the medium heterogeneity is adequately captured [*Rubin and Hubbard*, 2005]. The geophysical data used in geostatistical applications may include gamma ray [*Miller et al.*, 2000; *Cassiani and Binley*, 2005], seismic [*Yao*, 2002], ground penetrating radar [*Moysey et al.*, 2003; *Kowalsky et al.*, 2005], and electrical resistance and hydraulic tomography [*Yeh and Liu*, 2000; *Zhu and Yeh*, 2005; *Yeh and Zhu*, 2007]. Different ways of using the geophysical data are warranted, depending on the relationships between the geophysical data and the variables of interest (e.g., the soil classes in this study).

[6] While the TP/MC method can handle any categorical data, it is applied specifically in this study to the vadose zone soil textural classes. To our knowledge, this is the first time that the TP/MC methodology is evaluated in simulating vadose zone moisture movement at a field site. Soil heterogeneity is characterized by representing the spatial variability of soil class geometry. There are several advantages of using the soil classes. First, because the soil classes have been rigorously defined on the basis of sediment particle size distributions, a requirement of the TP/MC method that all soil classes are mutually exclusive is naturally satisfied, thus avoiding the selection of cutoff values. The other theoretical requirement that soil classes for TP/MC consideration include all soil classes of a domain (comprehensive exhaustiveness) can be met provided that a reasonable number of boreholes are present. This is of particular significance, since a recent sensitivity study [*Zhang et al.*, 2006] indicates that the number of lithofacies is the most important TP/MC parameter. Another advantage of using the soil classes is that soil hydraulic parameters for various soil classes are available in several national and international databases [e.g., *Rawls et al.*, 1982; *Carsel and Parrish*, 1988; *Schaap and Leij*, 1998]. This is of particular use for numerical modeling, especially when only a small number of parameter estimates are available. There have been two ways of handling hydraulic parameters for lithofacies in the TP/MC method. One approach uses effective parameter values for each lithofacies [e.g., *Fogg et al.*, 1998; *Zhang et al.*, 2006; *Lee et al.*, 2007]; the other approach treats the parameters as random variables [e.g., *Haldorsen and Damsleth*, 1990; *Tyler et al.*, 1994; *Bi and Oliver*, 2000; *Feyen and Caers*, 2006]. Both approaches can be implemented for the soil classes in a straightforward manner on the basis of decades of vadose zone studies. By adopting the first approach, soil hydraulic parameters are estimated from core samples, as discussed later. The second approach can be easily implemented using the method of *Lu*

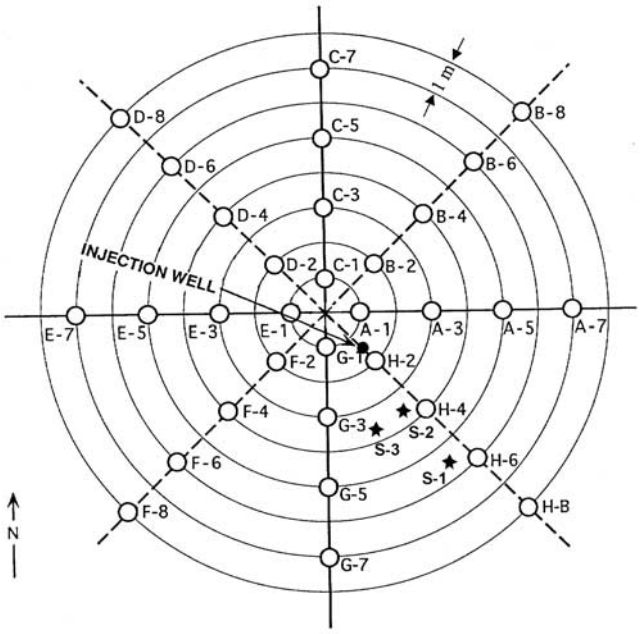

Figure 1. Plan view of the Sisson and Lu (S-L) injection test site and well numbering scheme [after *Sisson and Lu*, 1984; *Gee and Ward*, 2001].

and Zhang [2002], which considers the correlation between hydraulic parameters for different facies, or the methods of *Meyer et al.* [1997] and *Hou and Rubin* [2005], which estimate probability distributions of the soil hydraulic parameters. Although only the geometry of the soil classes (not corresponding soil hydraulic parameters) is treated as a random variable in this study, confidence intervals of predicted moisture content from Monte Carlo simulations sufficiently cover spatial variability of the observed moisture contents. This is consistent with the *Cooper et al.* [2007] finding that, in describing predictive uncertainty, uncertainty in geometry of lithofacies is more dominant than uncertainty of hydraulic parameters.

[7] The rest of this article is organized as follows. Section 2 describes the data sets used in this study, and briefly presents the TP/MC method to make this paper self contained. Section 3 discusses results of generating randomly heterogeneous soil class geometry and numerical simulation of the injection experiments, followed by concluding remarks in section 4.

Table 1. Breakdown of the Four Soil Textural Classes for the Seven Boreholes

Borehole Name	Sample Number	Coarse Sand	Sand	Loamy Sand	Sandy Loam
S-1	33	23	5	5	0
S-2	12	6	0	5	1
S-3	15	9	1	4	1
E-7	14	11	0	3	0
E-1	10	7	2	1	0
A-7	8	7	0	1	0
B-8	1	1	0	0	0
Total	93	64	8	19	2
Percentage (%)	100	69.82	8.60	20.43	2.15

2. Data and Methods

[8] The S-L injection site (Figure 1) was originally designed by *Sisson and Lu* [1984] within the 200 East Area of the U.S. Department of Energy's Hanford Site in southeastern Washington State. Two field injection experiments were conducted at the site in 1980 [*Sisson and Lu*, 1984] and 2000 [*Ward et al.*, 2000; *Gee and Ward*, 2001]. Several numerical simulations have been conducted, with mixed success, to simulate the 1980 and 2000 moisture plumes [*Fayer et al.*, 1995; *Rockhold et al.*, 1999; *Ward et al.*, 2000; *Gee and Ward*, 2001; *Zhang et al.*, 2004; *Ye et al.*, 2005, 2007; *Yeh et al.*, 2005; *Kowalsky et al.*, 2005; *Ward et al.*, 2006a, 2006b].

2.1. Soil Textural Classes and Initial Moisture Contents

[9] The particle size distribution (PSD) measurements for 93 samples at the S-L site are used to identify four soil textural classes (coarse sand, sand, loamy sand, and sandy loam) on the basis of the *U.S. Department of Agriculture* [1999] soil taxonomy. Among the PSD measurements, 60 are from boreholes S-1, S-2, and S-3 [*Schaap et al.*, 2003], 17 from A-7, E-1, and E-7 [*Khaleel and Freeman*, 1995], and 16 from E-1, E-7, and B-8 [*Fayer et al.*, 1993]. Locations of the boreholes are shown in Figure 1. Table 1 lists the breakdown on the number of the PSD samples and soil textural classes at each borehole as well as the percentage of each soil class for the entire data set. The vertical locations of the four soil classes are shown in Figure 2,

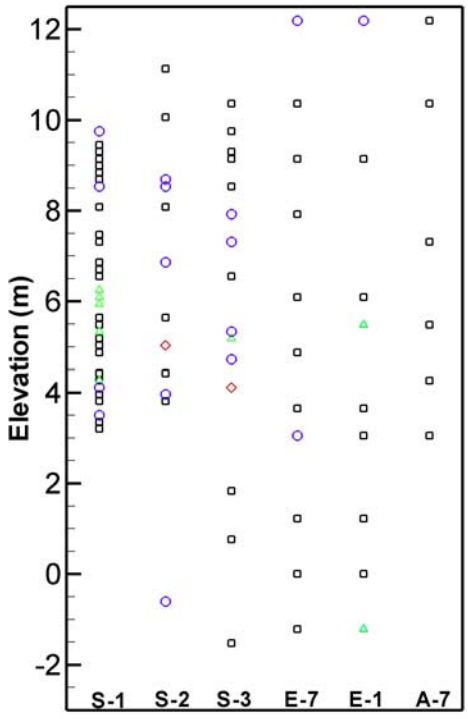

Figure 2. Sampling elevation for core samples for various soil textural classes at six boreholes. (squares, coarse sand; triangles, sand; circles, loamy sand; diamonds, sandy loam).

Table 2. Soil Hydraulic Parameter Values of the Four Soil Classes

	Coarse Sand	Sand	Loamy Sand	Sandy Loam
K_s (m/d)	12.329	5.679	1.906	1.041
θ_r (vol %)	3.175	3.640	3.500	4.000
θ_s (vol %)	32.064	36.307	39.500	35.000
α (1/m)	5.863	3.397	11.524	10.268
n	2.142	2.097	1.463	1.366

which does not include borehole B-8 because it has only one measurement. Borehole S-1 has the highest sampling density, with 33 core samples (length of each sample is 0.1524 m, i.e., 0.5 ft) between 5.4864 to 12.192 m (18 to 40 ft). The coarse sand is the dominant soil class, with 69.82% of the 93 samples being coarse sand. The sandy loam is only found at boreholes S-2 and S-3, indicating that this soil class occurs in thin layers of limited horizontal extent. Because of the existence of thin layers and relatively coarse sampling density, it is likely that this soil textural class is undersampled. Consequently, the percentage of soil class breakdown (Table 1) may not be representative, and adjustment may be warranted. It is worth noting that soil classes can be categorized using any classification method. A finer categorization of soil classes may lead to an improved characterization of medium heterogeneity.

[10] Table 2 lists the hydraulic parameters for each soil class. The van Genuchten water retention parameters (van Genuchten α and n, residual moisture content θ_r, and saturated moisture content θ_s) are fitted using the RETC software [*van Genuchten et al.*, 1991] for the retention data of all samples for each soil class. Although the core samples are from three different sources, they are considered equally reliable and accurate. Saturated hydraulic conductivity (K_s) is the geometric mean of individual K_s values for each soil class.

[11] Compared to the soil class measurements, the initial moisture content (θ_i) data (measured on 5 May 2000, about a month before the first injection on 1 June 2000) are more abundant and uniformly distributed. They were measured at the 32 radially arranged boreholes (Figure 1) at a depth interval of 0.3048 m (1 ft) starting from a depth of 3.9625 m (13 ft) and continuing to a depth of 16.764 m (55 ft), resulting in a total of 1,376 measurements. The θ_i data reflect medium heterogeneity, especially the imperfectly stratified layering structure [*Ye et al.*, 2005]. In particular, the θ_i database reveals two distinct zones of high θ_i values, corresponding to two fine-textured layers shown in lithostratigraphic cross sections of *Last et al.* [2001]. The two zones are depicted in Figure 3a, which is a two-dimensional (2-D) cross section of θ_i interpolated using the kriging method [*Ye et al.*, 2005]. The correspondence suggests that the θ_i distribution can be used as a surrogate for describing medium heterogeneity at the S-L site. This motivated the study of *Ye et al.* [2007], who demonstrated that using the θ_i database significantly improved simulation of the field injection experiments.

[12] This study thus uses the θ_i database to facilitate the heterogeneity characterization via geometry of the soil classes. Figure 4 illustrates the spatial correlation between the soil classes and the θ_i at the six boreholes. Since there

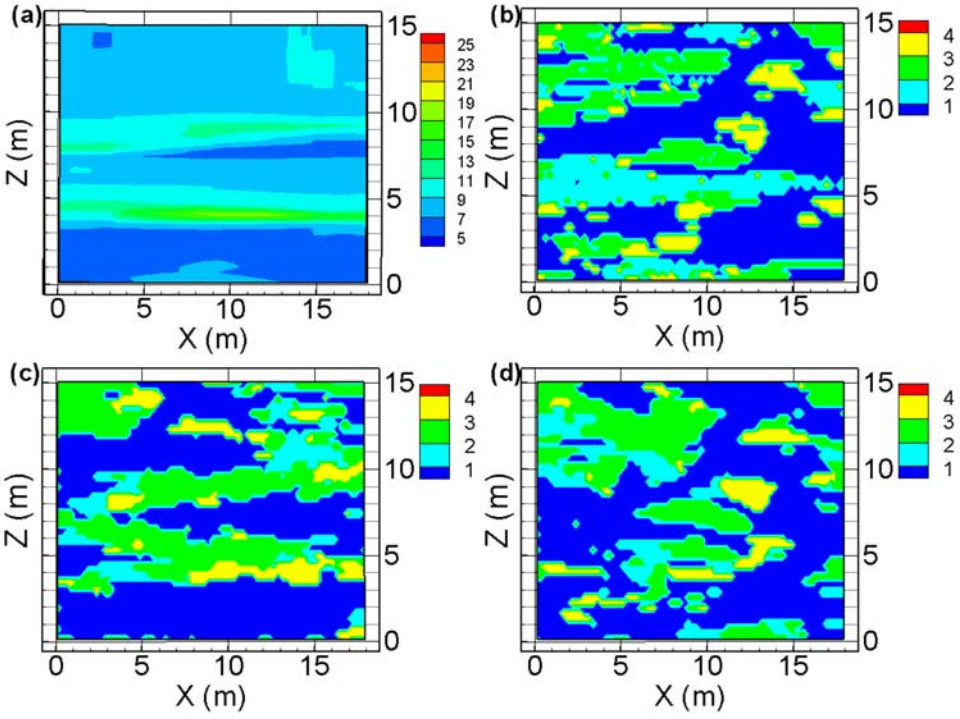

Figure 3. Two-dimensional contours passing through the injection well for (a) initial moisture content (θ_i) and (b and d) generated soil classes conditioned on measured soil classes only and for (c) both measured and transformed soil classes. Figures 3c and 3d use the same MC model but different conditioning data of soil classes. Indices 1–4 in Figures 3b–3d correspond to coarse sand, sand, loamy sand, and sandy loam, respectively.

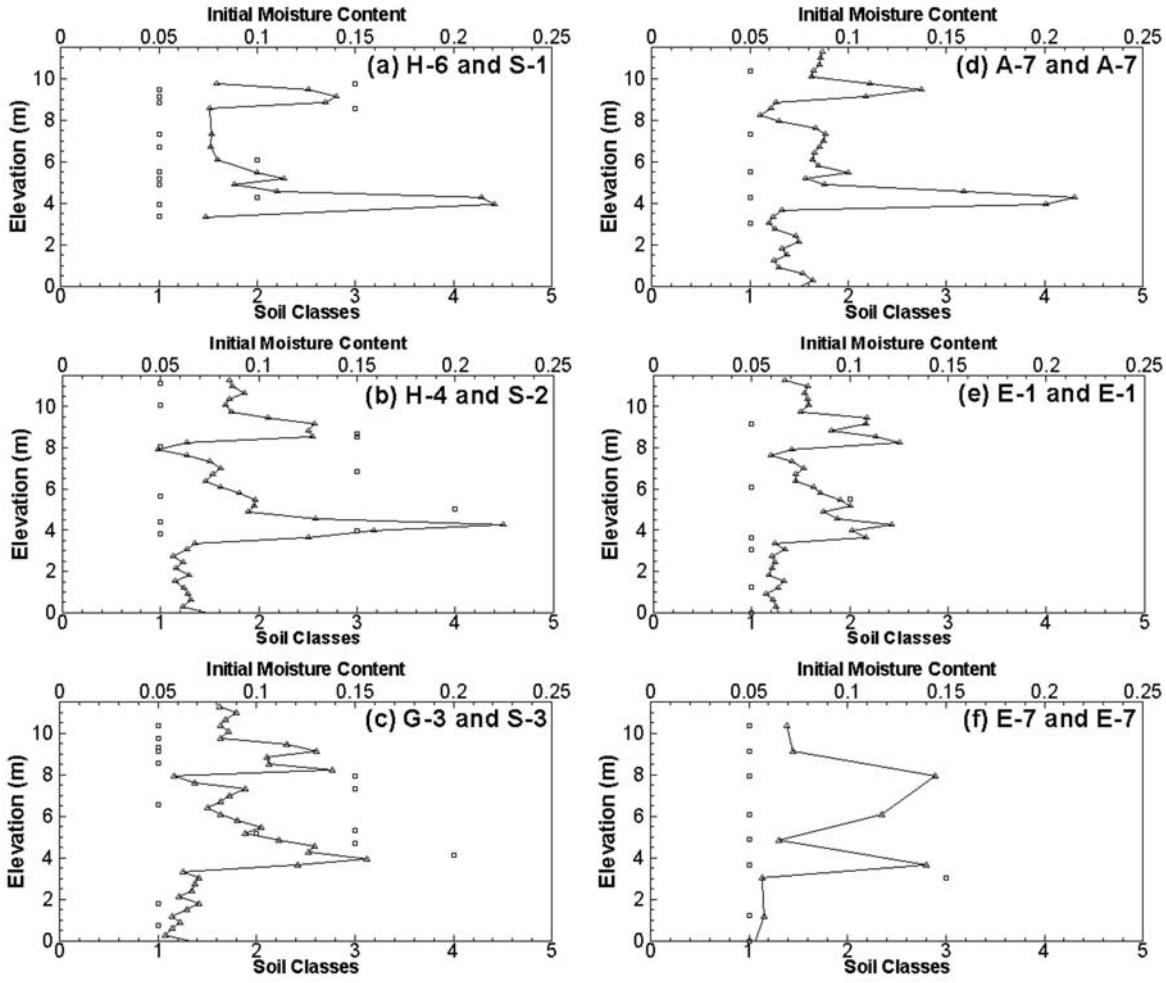

Figure 4. Spatial correlation between the soil classes (squares) and initial moisture content (triangles) at six boreholes. Soil classes 1–4 correspond to coarse sand, sand, loamy sand, and sandy loam, respectively.

was no θ_i measurement at boreholes S-1, S-2, and S-3, the θ_i values at the closest boreholes (H-6, H-4, and G-3, respectively) are used. Figure 4 shows an apparent correlation between the soil classes and the θ_i, indicative of a tendency that a higher θ_i corresponds to a finer soil texture. This justifies using the θ_i in section 3 for improving the estimation of mean horizontal length of soil classes. During the injection period, the moisture contents were sampled at all 32 boreholes on 2, 9, 16, and 23 June 2000; after the injections ceased, three additional sets of θ measurements were obtained on 7, 17, and 31 July. As described in section 3, these measurements are used to evaluate numerical simulation results for the injection experiment.

2.2. TP/MC Method

[13] The TP/MC method is briefly described here to make the paper self contained; for more details of the method, readers are referred to *Carle and Fogg* [1996, 1997] and *Carle* [1999]. Occurrence of a soil class, $k = \{1, \ldots, K\}$, at location $\mathbf{x}$ can be quantified using an indicator

$$I_k(\mathbf{x}) = \begin{cases} 1 & \text{if soil class } k \text{ occurs at location } \mathbf{x} \\ 0 & \text{otherwise} \end{cases}. \quad (1)$$

[14] The K soil classes should be mutually exclusive and exhaustively define all soil classes in a domain, the two requirements that can be easily satisfied as mentioned above. Because of the comprehensive exhaustiveness, summation of volume portion, p_k, for each soil class is one, i.e.,

$$\sum_{k=1}^{K} p_k = 1. \quad (2)$$

[15] Using the notation of *Carle* [1999], the transition probability from soil class j at $\mathbf{x}$ to k at $\mathbf{x} + h_\phi$ (h being lag) along direction ϕ is defined by

$$t_{jk}(h_\phi) = \Pr\{k \text{ occurs at } \mathbf{x} + h_\phi | j \text{ occurs at } \mathbf{x}\}. \quad (3)$$

For a one-dimensional stationary Markov chain model, the continuous-lag transition probability matrix, $\mathbf{T}$, satisfies the following conditions:

$$\sum_{k=1}^{K} t_{jk}(h_\phi) = 1 \qquad \forall j \quad (4)$$

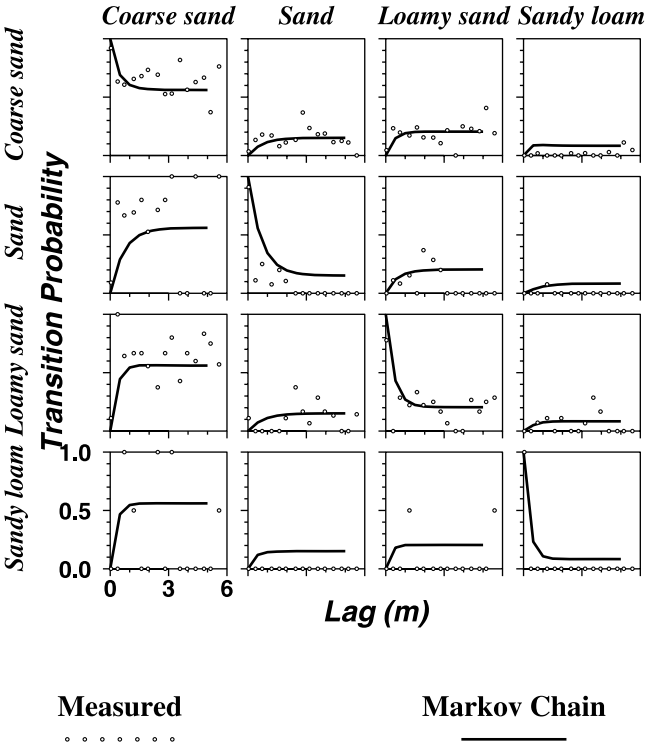

Figure 5. Transition probability matrix along the vertical direction based only on the measured soil classes.

$$\sum_{j=1}^{K} p_j t_{jk}(h_\phi) = p_k \qquad \forall k \tag{5}$$

$$\lim_{h_\phi \to \infty} t_{jk}(h_\phi) = p_k \qquad \forall j,k. \tag{6}$$

Equation (4) implies that soil class j must transition to one of the K soil classes at the lag distance of h_ϕ, whereas equation (5) indicates that the chance for soil class j transitioning to k depends on both the transition probability t_{jk} and its volume portion p_j. Equation (6) implies that, when the lag is large enough, the transition probability from any classes to class k will stabilize at its volumetric portion. Matrix $\mathbf{T}$ can be expressed using a transition rate via

$$\mathbf{T}(h_\phi) = \exp(\mathbf{R}_\phi h_\phi), \tag{7}$$

where $\mathbf{R}_\phi$ is a $K \times K$ transition rate matrix whose entry $r_{jk,\phi}$ represents the rate of change from j to k per unit length of j in the direction ϕ, and the relationship between t_{jk} and $r_{jk,\phi}$ is

$$\left.\frac{\partial t_{jk}(h_\phi)}{\partial h_\phi}\right|_{h_\phi=0} = r_{jk,\phi}. \tag{8}$$

Substituting (8) into (4) and (5) gives

$$\sum_{k=1}^{K} r_{jk,\phi} = 0 \qquad \forall j \tag{9}$$

$$\sum_{j=1}^{K} p_j r_{jk,\phi} = 0 \qquad \forall k. \tag{10}$$

A particularly useful application of (8) is that

$$\left.\frac{\partial t_{kk}(h_\phi)}{\partial h_\phi}\right|_{h_\phi=0} = r_{kk,\phi} = -\frac{1}{\bar{L}_{k,\phi}}, \tag{11}$$

where $\bar{L}_{k,\phi}$ is the mean length for soil class k along the direction ϕ, i.e., the total length occupied by k divided by the number of embedded occurrence of k. Equation (11) implies that the tangent of $t_{kk}(h_\phi)$ at $h_\phi = 0$ hits the lag at mean length of k along the direction ϕ. Estimating the mean horizontal length is always difficult in case of sparse borehole data. The use of an embedded TP/MC alleviates this problem somewhat by selecting a background soil class (the one with the largest volume proportion), because row and column entries involving the background soil class are not needed for the transition rate matrix [*Carle*, 1999]. However, the embedded TP/MC still requires estimating mean horizontal lengths for nonbackground soil classes. This paper develops a method for improving estimation of mean horizontal lengths based on the θ_i database.

3. Results and Discussion

3.1. TP/MC Using Measured Soil Classes Only

[16] Figure 5 illustrates the TP matrix in the vertical direction calculated using the Transition Probability Geostatistical Software (T-PROGS) [*Carle*, 1999]. The diagonal elements are auto-TP (t_{ii}) of soil class i, while the off-diagonal elements are cross-TP (t_{ij}) from soil class i at the row position to class j at the column position. Ideally, the auto-TP decreases from 1 when lag increases, and stabilizes at the volume proportion p_i of soil class i. The cross-TP is zero at lag zero, and increases with lag until stabilizing at the volume proportion of the tail soil class j (equation (6)). Although the experimental TP shown in Figure 5 is noisy, a reasonable MC model can be fitted using the embedded TP/MC with the coarse sand as the background soil class. Since the sand and sandy loam are likely undersampled, the volume proportions for the two classes are adjusted. On the basis of Table 1, the proportion, 5/33 = 15.15%, of sand at borehole S-1 and 1/12 = 8.34% of sandy loam at borehole S-2 are used. The percentage of coarse sand is reduced from 69.82% to 56.08% accordingly, while the percentage of loamy sand is left unchanged. The fitted mean vertical lengths are 0.9 m, 0.8 m, 0.5 m, and 0.3 m for the coarse sand, sand, loamy sand, and sandy loam, respectively. The fitted model matches the experimental TP reasonably well for both sand and loamy sand (Figure 5), while it is difficult to judge the fit for sandy loam because of lack of data. For the coarse sand, in spite of the underfit to the experimental TP from sand to coarse sand, the fit to other TP is acceptable. Because of the lack of data, it is not possible to quantify the experimental TP in the horizontal direction. Using the approach of *Cole and Cumella* [2004] and several unpublished geologic maps, the mean horizontal lengths are estimated as 4 m, 5 m, 3 m, and 1 m, respectively, for the coarse sand, sand, loamy sand, and sandy loam.

[17] Heterogeneous fields of soil class geometry are generated using T-PROGS and the fitted MC model, conditioned on the 93 core measurements. Figure 3b shows the 2-D cross section, passing through the injection well (Figure 1), for the first realization of soil classes. Although

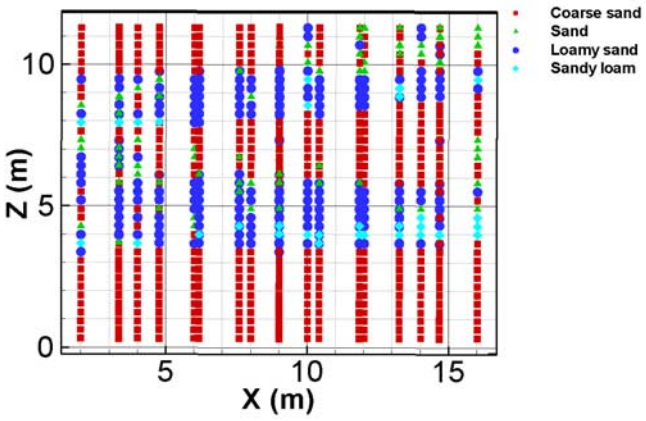

Figure 6. Four soil classes transformed from measurements of initial moisture content at the 32 boreholes.

Figure 3b shows a layer of sand (index 2) at about 6 m elevation, the two layers of fine-textured horizon (Figure 3a) are not captured, a behavior also observed in other T-PROGS realizations. This indicates that the two layers at 9 m and 4 m elevations cannot be characterized by the TP/MC using only the core measurements. Numerical simulation (results not shown) confirms that, without the two layers of fine-textured soils, the injected water moves too fast downward.

3.2. TP/MC After Incorporating the Initial Moisture Content

[18] The θ_i database is used to improve estimates of the mean lengths and characterization of the sediment layering structure. On the basis of the observation from Figure 4, the θ_i data are transformed into soil classes using a method similar to the normal score transform. With X being a discrete random variable, the probabilities that X takes the values of x_1 = coarse sand, x_2 = sand, x_3 = loamy sand, and x_4 = sandy loam are p_1, p_2, p_3, and p_4, which essentially are the volume proportions (equation (2)). The cumulative probability distribution $P_k = P(x_k)$ is

$$P(x_k) = \sum_{x_i \le x_k} p(x_i) \text{ or } P_k = \sum_{i \le k} p_i. \tag{12}$$

[19] The transformed soil classes for the θ_i data set are expected to follow the same cumulative probability distribution of the measured soil classes. Threshold values $\theta_{i,k}$ transforming θ_i to the soil classes can thus be determined via

$$F(\theta_{i,k}) = F(\theta_i \le \theta_{i,k}) = P_k, \tag{13}$$

where F is the probability distribution function for θ_i. With $\theta_{i,k-1} < \theta_i \le \theta_{i,k}$, the θ_i value is transformed into soil class k ($\theta_{i,0}$ and $\theta_{i,4}$ being the minimum and maximum θ_i). In this study, the volume proportions (p_1 *to* p_4) are determined, on the basis of Table 1, as 60%, 10%, 25%, and 5% with a slight increase for sand, loamy sand, and sandy loam (such adjustment is not uncommon [*Carle and Fogg*, 1996]). After the θ_i database is sorted in ascending order to obtain $F(\theta_{i,k})$, the three thresholds are 8.43% ($\theta_{i,1}$), 9.04% ($\theta_{i,2}$), and 12.88% ($\theta_{i,3}$) for the corresponding P_k of 60%, 70%, and 95%. Figure 6 shows the transformed soil classes at 1-ft intervals for the 32 boreholes in the x-z plane. Figure 6 illustrates the two layers consisting of mainly loamy sand and sandy loam.

[20] Using both the measured and transformed soil classes, the vertical TP is recalculated and shown in Figure 7. In comparison with Figure 5, which is based only on measured soil classes, Figure 7 shows that the TP results are significantly improved, especially for the TP associated with the sandy loam (Figures 5 and 7). The new fitted mean vertical lengths are 1.1 m, 0.4 m, 0.8 m, and 0.3 m for the coarse sand, sand, loamy sand, and sandy loam, respectively. Figure 8 shows the horizontal TP calculated from the measured and transformed soil classes. Mean horizontal lengths for the coarse sand, sand, loamy sand, and sandy loam are 4 m, 1 m, 2 m, and 1 m, respectively. Note that quantifying the horizontal TP would not be feasible without incorporating the transformed soil classes.

[21] Figures 3c and 3d illustrate the two-dimensional cross sections (the same as those in Figures 3a and 3b) of the heterogeneous geometry of soil classes generated using the same MC model but different conditioning data. With conditioning on both the measured and transformed soil classes, Figure 3c is able to capture the two layers consisting mainly of loamy sand and coincident with the location of higher θ_i shown in Figure 3a. However, the same two layers are not captured in Figure 3d by conditioning only on measured soil classes, although the same TP/MC model is used. The difference in presenting the layering structure illustrates the importance of conditioning data.

3.3. Numerical Simulation of the 2000 Injection Experiment

[22] Using T-PROGS, 50 realizations of soil class geometry are generated for a Monte Carlo simulation of the 2000

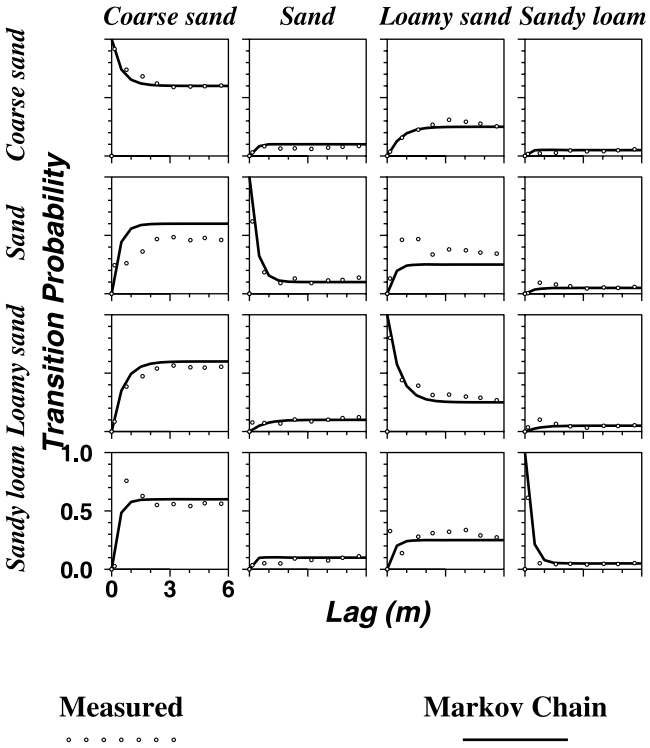

Figure 7. Transition probability matrix along the vertical direction based on both measured soil classes and transformed soil classes from the initial moisture content.

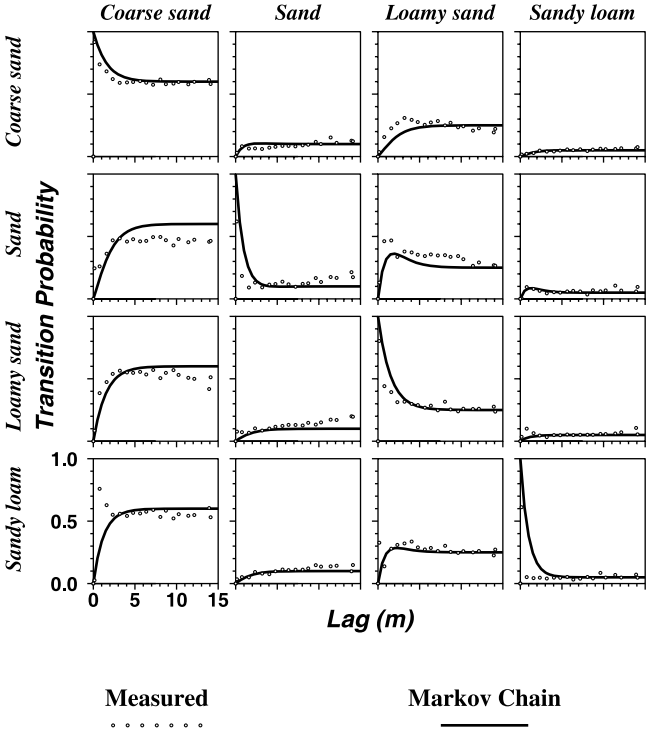

Figure 8. Transition probability matrix in the horizontal direction based on both measured soil classes and transformed soil classes from the initial moisture content.

injection experiment. The soil hydraulic parameters of the four classes are listed in Table 2. The anisotropy for the horizontal and vertical saturated hydraulic conductivity is assigned a value of 4, as it gives the best fit between the simulated and observed moisture contents. Anisotropy estimates ranging from 2 to 60 have been used by other investigators [*Sisson and Lu*, 1984; *Cole et al.*, 2001; *Pace et al.*, 2003; *Zhang et al.*, 2004; *Ward et al.*, 2006a; *Mayes et al.*, 2003]. The 2000 injection experiment is simulated using the MMOC code [*Srivastava and Yeh*, 1992] which was also used in our earlier work [*Yeh et al.*, 2005; *Ye et al.*, 2007]. The numerical grid and initial and boundary conditions are determined in the same manner as in the work by *Ye et al.* [2007].

[23] Simulated and observed moisture contents are compared to evaluate accuracy and robustness of our method of characterizing the medium heterogeneity. Figure 9 compares the observed moisture content (Figures 9a and 9b) and the simulated mean moisture content (Figures 9c and 9d) on 2 June 2000 (the first observation date) and 31 July 2000 (the last observation date). Results for other observation times are similar and thus not shown. Figure 9 illustrates good agreement between the observed and simulated mean moisture plumes in terms of their overall shape and local variation. In particular, the effect of imperfectly stratified layering structure on moisture movement is captured well. The injected water spreads in the top fine-textured layer (at an elevation of about 9 m); the vertical movement of the injected water is retarded by the bottom fine-textured layer (at an elevation of about 5 m). Between these two layers is a layer of coarse-textured horizon, where the plume splits. The simulated plume separation beneath the top fine-textured layer is comparable with similar results presented elsewhere [*Kowalsky et al.*, 2005]. However, unlike *Kowalsky et al.* [2005] who conducted inverse modeling and used both moisture content and cross borehole radar data, our study only uses the moisture content database and does not require solving the Richards' equation iteratively as in the inverse modeling. Figures 9e and 9f show the correlation between the observed and simulated mean moisture contents at the sixteen A, C, E, and G boreholes (Figure 1); the moisture contents are directly simulated at these borehole locations. The goodness of fit is measured by the Pearson's linear correlation coefficient

$$r = \left(\sum_{j=1}^{N}\left(\theta_j^* - \bar{\theta}^*\right)\left(\theta_j - \bar{\theta}\right)\right) \Bigg/ \cdot \left(\sqrt{\sum_{j=1}^{N}\left(\theta_j^* - \bar{\theta}^*\right)^2}\sqrt{\sum_{j=1}^{N}\left(\theta_j - \bar{\theta}\right)^2}\right), \quad (14)$$

where θ^* and $\bar{\theta}^*$ are observed moisture content and its mean, θ and $\bar{\theta}$ are simulated moisture content and its mean, and $N = 592$ is the number of observations. The r values of 0.84 and 0.74 suggest good correlation and indicate an acceptable simulation of the injection experiment.

[24] Spatial variability of the observed and simulated θ is also examined at individual boreholes to avoid the smoothing effect of contouring. Figure 10 compares the observed and simulated θ (mean and 95% confidence intervals) on 23 June 2000 (the last measurement date following injections) for the eight A and G boreholes (Figure 1). Similar plots are obtained for other boreholes at other observation times (results not shown). Figure 10 shows that the profiles of the observed moisture contents are well captured by the simulated mean, and that the confidence intervals sufficiently bracket the observed θ variability. Figure 10 also lists the root-mean-square error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(\theta_j^* - \theta_j\right)^2} \quad (15)$$

and Pearson's coefficient (r) (equation (14)) for each borehole. These statistics quantitatively confirm the good agreement between the observed and simulated mean θ. The average RMSE is 2.70%, comparable with that obtained via inverse modeling, e.g., 2.59% of *Zhang et al.* [2004], and 1.87% of *Ward et al.* [2006b].

[25] No fluid flux measurements were made at the S-L site. However, we investigate whether the simulated moisture flow honors the observed flow behaviors: the near-zero fluid flux below the bottom layer of fine material as well as the overall southeastward movement of the moisture plume. During the injection experiments, downward movement of the injected water through the bottom layer of fine material is small, since the observed moisture contents below the layer are nearly the same before and after the injection experiment [*Ye et al.*, 2005, Figure 3]. This behavior is correctly simulated, because our calculated vertical Darcy flux is about −0.001 m/d (negative value indicating downward direction) below the bottom fine layer, and essentially

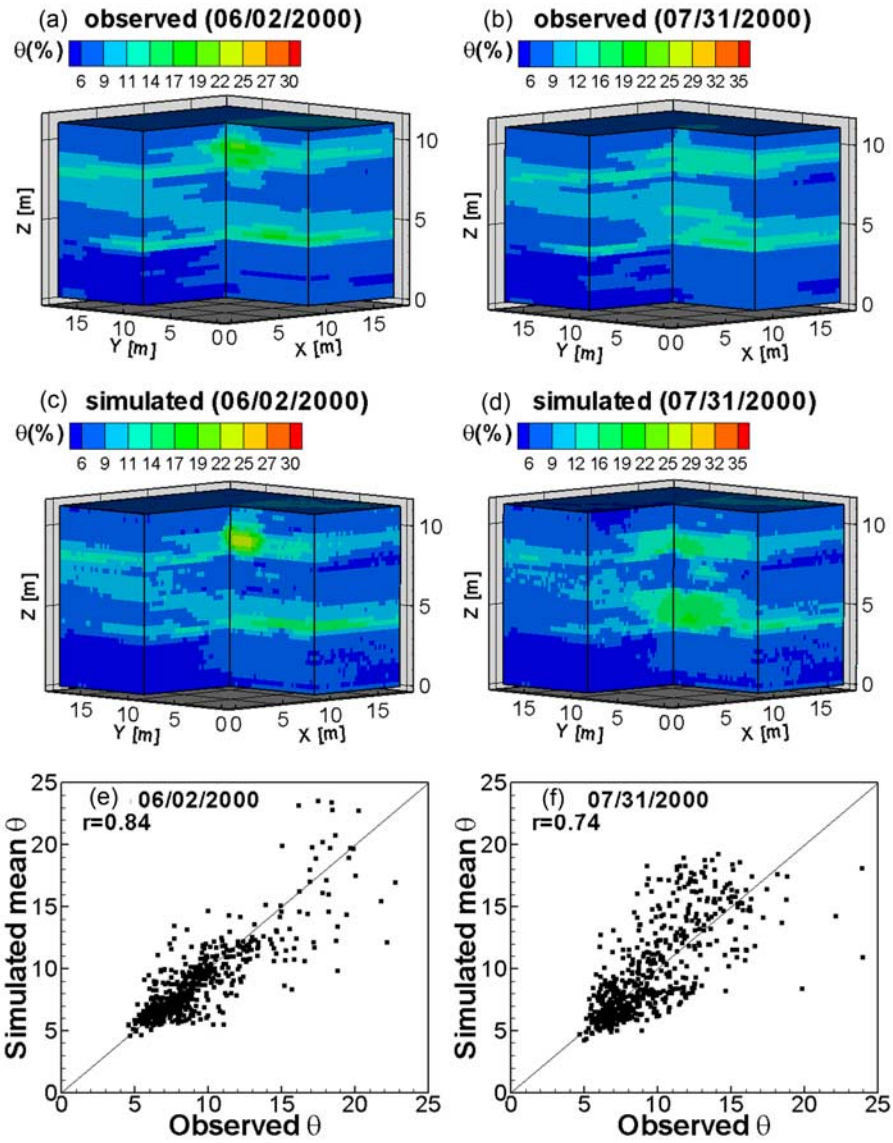

Figure 9. Three-dimensional contours of (a) observed and (c) simulated mean moisture content on 2 June 2000; three-dimensional contours of (b) observed and (d) simulated mean moisture content on 31 July 2000; and observed and simulated mean water content on (e) 2 June 2000 and (f) 31 July 2000. Pearson's linear correlation coefficients (r) are calculated for the two simulation times.

zero toward the domain bottom. The overall pattern of the moisture movement is evaluated by calculating the spatial moments (up to second order) of the simulated and observed moisture contents. Denoting $\theta_{diff} = \theta - \theta_i$ as the moisture content difference between moisture content (θ) at an observation time and the initial moisture content (θ_i), the spatial moments of the θ_{diff} are [*Ye et al.*, 2007]

$$M_{ijk}(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \theta_{diff}(x,y,z,t) x^i y^j z^k dxdydz. \quad (16)$$

[26] The zeroth, first, and second moments correspond to $i + j + k = 0$, 1, and 2, respectively. The zeroth moment (M_{000}) represents the changes in moisture storage within the domain. The normalized first moments,

$$X_c = M_{100}/M_{000} \quad Y_c = M_{010}/M_{000} \quad Z_c = M_{001}/M_{000}, \quad (17)$$

represent the location (X_c, Y_c, Z_c) of the mass center of the plume at a given time. The spread of the plume about its center in the x, y, and z directions is measured by the second moments

$$\sigma_{xx}^2 = \frac{M_{200}}{M_{000}} - X_c^2 \quad \sigma_{yy}^2 = \frac{M_{020}}{M_{000}} - Y_c^2 \quad \sigma_{zz}^2 = \frac{M_{002}}{M_{000}} - Z_c^2. \quad (18)$$

[27] The spatial moments are calculated for the observed and simulated mean moisture contents. Figure 11 illustrates the plume center and spreading (i.e., the first and second moments) for the seven observation times. Figure 11a suggests that, since X_c increases and Y_c decreases, the overall flow direction is southeast as observed in the field. The simulated X_c and Y_c agree reasonably well with the observed values. The simulated Z_c is larger than the observed Z_c during the injection period (for the first four

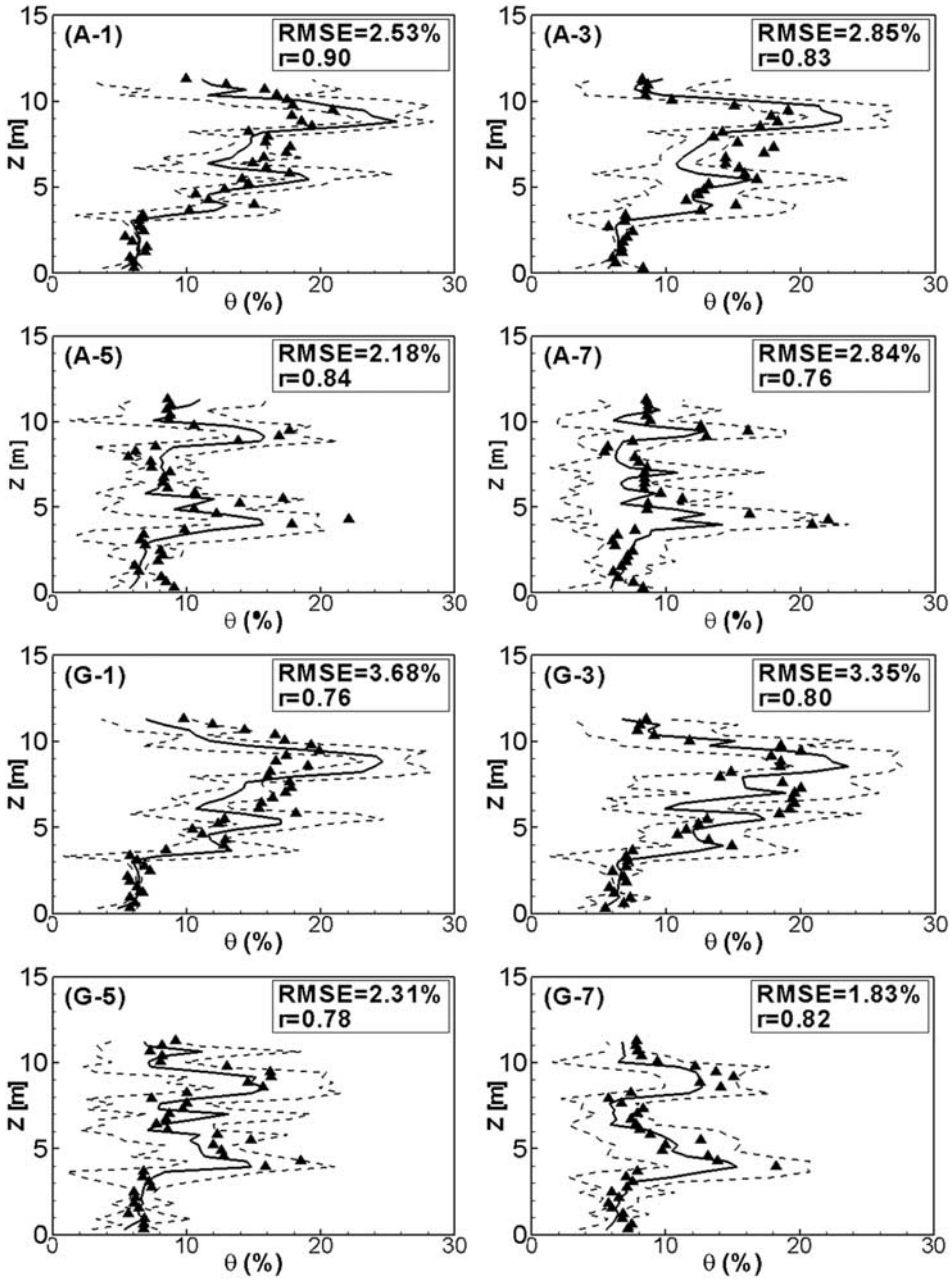

Figure 10. Mean predictions (solid lines), lower and upper 95% confidence intervals (dashed lines), and observations (triangles) of moisture content on 23 June 2000 for eight boreholes. Root-mean-square error (RMSE) and Pearson's linear correlation coefficients (r) are calculated for each borehole.

observation times), indicative of a slower vertical movement. During the redistribution period (the last three observation times), the simulated Z_c is smaller, indicative of a faster vertical movement. For the second moments, the simulated variances are smaller than the observed values during the injection period, suggesting that the simulated moisture movement is slower than the observed one. However, the difference between the simulated and observed spatial variances becomes smaller during the redistribution period.

[28] This study is an improvement over the simulation results of *Ye et al.* [2007] using cokriging. For example, the Pearson's correlation coefficient increased from 0.62 [*Ye et al.*, 2007] to 0.74 in this study for the 31 July 2000 simulation results; the average r (representing all the 16 borehole locations where the moisture contents are directly simulated) is 0.81 in this study, 35% larger than the average value of 0.60 in the work by *Ye et al.* [2007]. More importantly, the vertical moisture movement is simulated more accurately. The larger values of simulated Z_c and σ^2_{zz} relative to those of *Ye et al.* [2007] indicate a slower vertical movement and a larger spreading, agreeing better with the observed behavior. The improvement is attributed to the TP/MC method used in this study for heterogeneity characterization, because it captures the contrast between the fine and coarse soil textural classes. The cokriging method used by *Ye et al.* [2007] results in relatively smooth soil hydraulic

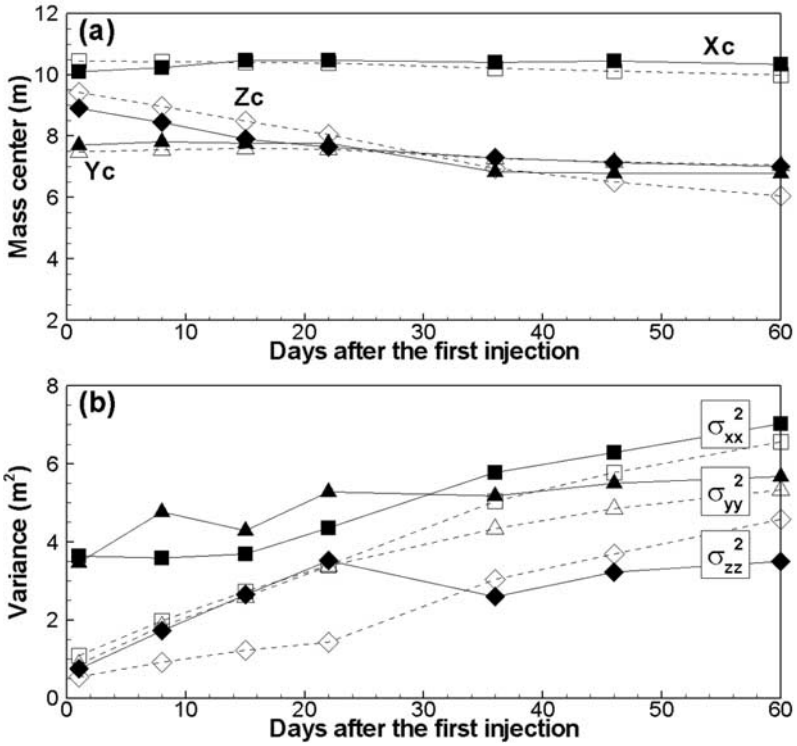

Figure 11. Comparison of observed (solid symbols and solid lines) and simulated (open symbols and dashed lines) (a) first moment and (b) second moment of moisture plume in the *x*, *y*, and *z* directions (represented by squares, triangles, and diamonds, respectively).

parameter fields thus yielding a less favorable simulation of the vertical movement of the injected water.

3.4. Effect of Conditioning Data on the Numerical Simulation

[29] The results presented in section 3.2 demonstrate that the sediment layering structure cannot be preserved without incorporating the transformed soil classes. This section further investigates the effect of conditioning data on characterization of the sediment layering structure and on accuracy of the simulated moisture contents. This is achieved by gradually reducing the conditioning data, first from the H boreholes, then the G, and finally the A boreholes (Figure 1). Figure 12 shows the 2-D contours of generated soil class geometry for the cross section passing through the injection well (the same cross section of Figure 3). Figure 12 also shows locations of boreholes A-1, G-1, and H-2, the three closest to the cross section. When the soil class data at the H boreholes are eliminated from conditioning, the sediment layering structure is still preserved (Figure 12b). The reason is that the random field generation is still conditioned on the soil classes data at the A, E, F, and G boreholes (especially the A and E boreholes). The layering structure disappears in the southwest corner away from the A and G boreholes (results not shown). When the soil class data at the G boreholes are further eliminated from conditioning, because of the conditioning data at the A, E, and F boreholes, the layering structure still exists but becomes less continuous (Figure 12c). When the soil class data at the A boreholes are excluded, the layering structure for the right half of the cross section disappears, whereas the layering structure is preserved at the left half (Figure 12d). The gradual change of the sediment layering structure with the amount of conditioning data is related to the mean horizontal length of the loamy sand, the major soil class within the two fine-textured layers (Figure 6). The mean length is 2 m, larger than the distance of 1.125 m between the cross section and the A and E boreholes. Therefore, when the conditioning data include the soil classes at the A and E boreholes, the layering structure is preserved at the cross section; when the conditioning data at the A boreholes are excluded, the layering structure disappears, because there essentially are no conditioning data within the mean length of 2 m of the loamy sand for the right half of the cross section. The layering structure is still preserved for the left half of the cross section because of the E boreholes.

[30] Figure 13 shows the observed and simulated moisture contents for the four conditioning cases at borehole A-7. This borehole is selected because Figure 12 suggests that this borehole should be most severely affected by the reduction of conditioning data at the A boreholes. The RMSE and Pearson's correlation coefficients in Figures 13a–13c indicate an almost identical predictive performance, suggesting that reducing conditioning data within the mean horizontal length does not necessarily deteriorate the simulated moisture profiles. However, when the soil classes at the A boreholes are excluded from conditioning, as shown in Figure 13d, the RMSE increases to 3.87% and the *r* value decreases to 0.23. In addition, Figure 13d shows that the mean predictions cannot capture the variability of the observed moisture profile and that the overall predictive uncertainty increases. These indicate that the simulated moisture profile is adversely affected

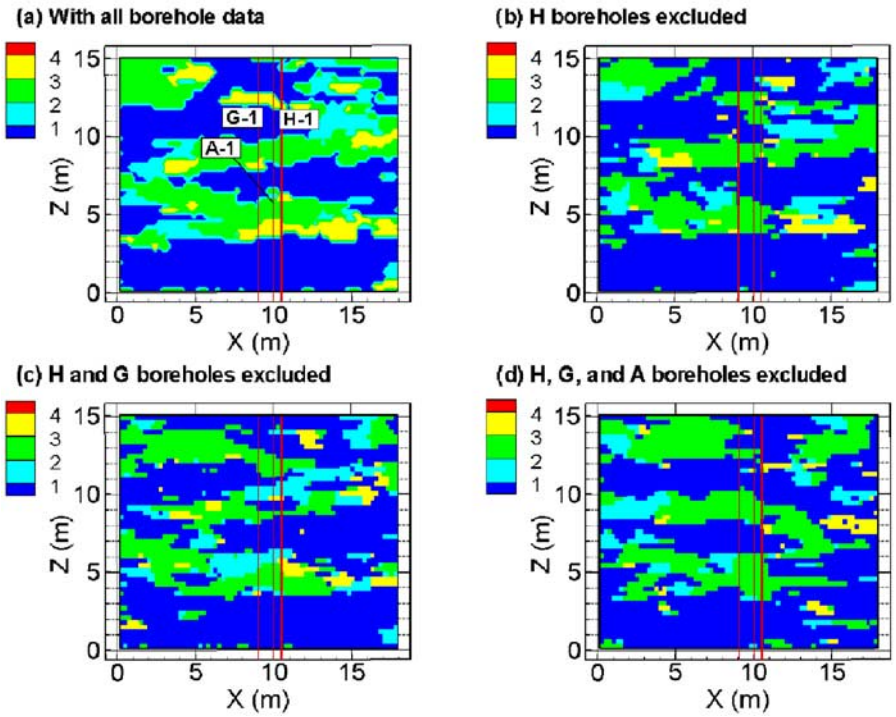

Figure 12. Two-dimensional contours passing through the injection well with conditioning on transformed soil classes (a) from all boreholes; (b) with H boreholes excluded; (c) with H and G boreholes excluded; and (d) with H, G, and A boreholes excluded. Red lines represent the A-1, G-1, and H-1 boreholes (Figure 1), the closest to the injection well.

by excluding the soil classes at the A boreholes from the conditioning.

4. Concluding Remarks

[31] The study leads to the following major conclusions and suggested future research.

[32] 1. Medium heterogeneity can be characterized by describing the spatial variability of geometry of soil textural classes using the TP/MC method. The characterization method meets the two theoretical requirements (mutual exclusiveness and comprehensive exhaustiveness) of the TP/MC. More importantly, the characterization method

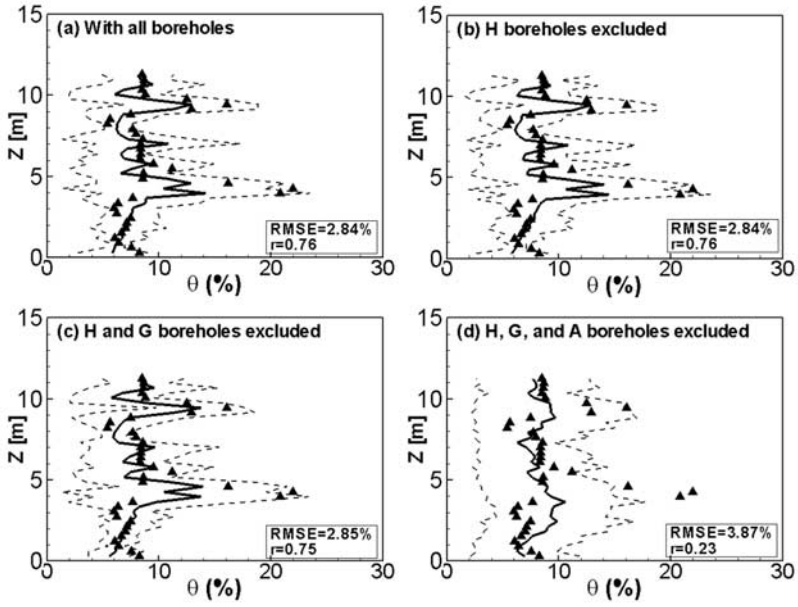

Figure 13. Mean predictions (solid lines), lower and upper 95% confidence intervals (dashed lines), and observations (triangles) of moisture content on 23 June 2000 at borehole A-7 for the four conditioning cases with the transformed soil classes (a) from all boreholes; (b) with H boreholes excluded; (c) with H and G boreholes excluded; and (d) with H, G, and A boreholes excluded.

helps preserve the contrast of soil hydraulic parameters among soil classes, which is important for accurately simulating vertical moisture movement. This is believed to be the major reason that the TP/MC method outperforms the cokriging method used by *Ye et al.* [2007] for simulating the field injection experiments at the S-L site.

[33] 2. The difficulty of estimating the mean horizontal length for the TP/MC method can be resolved by incorporating the "soft" data, e.g., the θ_i database used in this study. The procedure used in this work leads to an improved estimation of mean lengths in the vertical direction; it also renders it possible to estimate the mean horizontal lengths for the soil classes.

[34] 3. The θ_i data are transformed objectively into soil classes using the statistical method, ensuring that the transformed soil classes follow the same cumulative probability distribution of the soil classes measured from core samples. This alleviates reliance on subjective expert judgment based on geological knowledge. The statistical method is applicable to other types of data (e.g., geophysical). Future research will further investigate the method of transforming the initial moisture content and/or geophysical data into soil textural classes, with focus on its capability of representing the soil classes.

[35] 4. The heterogeneous geometry of soil classes generated by the TP/MC method is the cornerstone of successful simulation of the injection experiments at the S-L site. To our knowledge, this is the first time that the TP/MC method is applied in simulating vadose zone moisture movement. The RMSE and Pearson's correlation coefficients suggest that the overall pattern of the moisture movement and the spatial variability of observed moisture content are well captured. In addition, the TP/MC method adequately simulates the observed flow patterns, including the splitting of the moisture plume in the coarse layer which is sandwiched between the two fine layers, the southeastward movement of the plume, and the near-zero fluid flux below the bottom fine layer.

[36] 5. The conditioning data are critical for the TP/MC method to characterize the sediment layering structure. This study shows that using the same TP/MC model but different data sets of the soil classes may result in significantly different soil class geometry. More importantly, the desired layering structure cannot be preserved without using the transformed soil classes. Reducing conditioning data does not necessarily deteriorate the numerical simulation, if there are other conditioning data within the mean lengths of soil classes. This is useful for developing guidelines for collecting additional data for site characterization.

[37] 6. Uncertainty in geometry of soil classes appears more important than uncertainty of the soil hydraulic parameters. Spatial variability of the observed moisture contents is well simulated by only considering uncertainty in the soil class geometry. A study of the relative importance of soil texture distribution vis-à-vis hydraulic properties is another avenue for future research. In other words, if soil types and their textural distribution are important aspects for an accurate simulation of unsaturated flow and solute transport, more focus should be placed on collecting such data.

[38] **Acknowledgments.** The work reported was performed for the CH2M Hill Hanford Group Inc. and the U.S. Department of Energy Office of River Protection under contract DE-AC06-99RL14047. The first author is also supported by the DOE EPSCoR program under contract DE-FG02-06ER46265. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof or its contractors or subcontractors. The views and opinions of the authors do not necessarily state or reflect those of the United States government or any agency thereof.

References

Bi, Z., and D. S. Oliver (2000), Conditioning 3D stochastic channels to pressure data, *SPE J.*, *5*(4), 474–484.

Carle, S. F. (1999), T-PROGS: Transition Probability Geostatistical Software, version 2.1, Univ. of Calif., Davis, Calif.

Carle, S. F., and G. E. Fogg (1996), Transition probability-based indicator geostatistics, *Math. Geol.*, *28*(4), 453–476, doi:10.1007/BF02083656.

Carle, S. F., and G. E. Fogg (1997), Modeling spatial variability using one- and multi-dimensional continuous Markov chains, *Math. Geol.*, *29*(7), 891–918, doi:10.1023/A:1022303706942.

Carsel, R. F., and R. S. Parrish (1988), Developing joint probability distributions of soil water retention characteristics, *Water Resour. Res.*, *24*(5), 755–769, doi:10.1029/WR024i005p00755.

Cassiani, G., and A. Binley (2005), Modeling unsaturated flow in a layered formation under quasi-steady state conditions using geophysical data constraints, *Adv. Water Resour.*, *28*, 467–477, doi:10.1016/j.advwatres.2004.12.007.

Chiles, J.-P., and P. Delfiner (1999), *Geostatistics: Modeling Spatial Uncertainty*, John Wiley, New York.

Christakos, G. (2000), *Modern Spatiotemporal Geostatistics*, Oxford Univ. Press, New York.

Cole, C. R., M. P. Bergeron, S. K. Wurstner, P. D. Thorne, S. Orr, and M. I. McKinley (2001), *Transient Inverse Calibration of Hanford Site-Wide Groundwater Model to Hanford Operational Impacts—1943 to 1996*, *Rep. PNL-13447*, Pac. Northwest Natl. Lab., Richland, Wash.

Cole, R. D., and S. Cumella (2004), Fluvial sand-body dimensions in the Lower Williams Fork Formation (Upper Cretaceous), Southwestern Piceance Basin, Colorado, paper presented at Rocky Mountain Section Meeting, Am. Assoc. of Pet. Geol., Denver, Colo., 9–11 Aug.

Cooper, C., M. Ye, and J. Chapman (2007), Tritium transport at the Rulision site, a nuclear-stimulated low-permeability natural gas reservoir, *Publ. 45224*, Off. of Legacy Manage., U.S. Dep. of Energy, Washington, D. C.

Cressie, N. (1991), *Statistics of Spatial Data*, John Wiley, New York.

Dai, Z., R. W. Ritzi Jr., and D. F. Dominic (2005), Improving permeability semivariograms with transition probability models of hierarchical sedimentary architecture derived from outcrop analog studies, *Water Resour. Res.*, *41*, W07032, doi:10.1029/2004WR003515.

Dai, Z., A. Wolfsberg, Z. Lu, and R. Ritzi Jr. (2007), Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix, *Geophys. Res. Lett.*, *34*, L20406, doi:10.1029/2007GL031608.

Deutsch, C. V., and A. G. Journel (1998), *GSLIB—Geostatistical Software Library and User's Guide*, Oxford Univ. Press, Orlando, Fla.

Elfeki, A., and M. Dekking (2001), A Markov chain model for subsurface characterization: Theory and applications, *Math. Geol.*, *33*(5), 569–589, doi:10.1023/A:1011044812133.

Fayer, M. J., J. B. Sisson, W. A. Jordan, A. H. Lu, and P. R. Heller (1993), Subsurface injection of radioactive tracers, *Rep. NUREG/CR-5996*, Off. of Nucl. Regul. Res., U.S. Nucl. Regul. Comm., Washington, D. C.

Fayer, M. J., R. E. Lewis, R. E. Engelman, A. L. Pearson, C. J. Murray, J. L. Smoot, R. R. Randall, W. H. Wegener, and A. H. Lu (1995), Re-evaluation of a subsurface injection experiment for testing flow and transport models, *Rep. PNL-10860*, Pac. Northwest Natl. Lab., Richland, Wash.

Feyen, L., and J. Caers (2006), Quantifying geological uncertainty for flow and transport modeling in multi-modal heterogeneous formations, *Adv. Water Resour.*, *29*, 912–929, doi:10.1016/j.advwatres.2005.08.002.

Fogg, G. E., C. D. Noyes, and S. F. Carle (1998), Geologically based model of heterogeneous hydraulic conductivity in an alluvial setting, *Hydrogeol. J.*, *6*, 131–143, doi:10.1007/s100400050139.

Gee, G. W., and A. L. Ward (2001), Vadose zone transport field study: Status report, *Rep. PNNL-13679*, Pac. Northwest Natl. Lab., Richland, Wash.

Goovaerts, P. (1997), *Geostatistics for Natural Resources Evaluation*, Oxford Univ. Press, New York.

Haldorsen, H. H., and E. Damsleth (1990), Stochastic modeling, *J. Pet. Technol.*, *42*(4), 404–412, doi:10.2118/20321-PA.

Hou, Z., and Y. Rubin (2005), On minimum relative entropy concepts and prior compatibility issues in vadose zone inverse and forward modeling, *Water Resour. Res.*, *41*, W12425, doi:10.1029/2005WR004082.

Isaaks, E. H., and R. M. Srivastava (1989), *An Introduction to Applied Geostatistics*, Oxford Univ. Press, New York.

Journel, A. G., and C. J. Huijbregts (1978), *Mining Geostatistics*, Academic, San Diego, Calif.

Journel, A., and T. F. Zhang (2006), The necessity of a multiple-point prior model, *Math. Geol.*, *38*(5), 591–610, doi:10.1007/s11004-006-9031-2.

Journel, A. G., R. Gundeso, E. Gringarten, and T. Yao (1998), Stochastic modeling of a fluvial reservoir: A comparative review of algorithms, *J. Pet. Sci. Eng.*, *21*(1–2), 95–121, doi:10.1016/S0920-4105(98)00044-8.

Khaleel, R., and E. J. Freeman (1995), Variability and scaling of hydraulic properties for 200 area soils, Hanford site, *Rep. WHC-EP-0883*, Westinghouse Hanford Co., Richland, Wash.

Kitanidis, P. K. (1997), *Introduction to Geostatistics: Applications in Hydrogeology*, Cambridge Univ. Press, Cambridge, U. K.

Koltermann, C. E., and S. M. Gorelick (1996), Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches, *Water Resour. Res.*, *32*(9), 2617–2658, doi:10.1029/96WR00025.

Kowalsky, M. B., S. Finsterle, J. Peterson, S. Hubbard, Y. Rubin, E. Majer, A. Ward, and G. Gee (2005), Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data, *Water Resour. Res.*, *41*, W11425, doi:10.1029/2005WR004237.

Krishnan, S., and A. G. Journel (2003), Spatial connectivity: From variograms to multiple-point measures, *Math. Geol.*, *35*(8), 915–925, doi:10.1023/B:MATG.0000011585.73414.35.

Last, G. V., and T. G. Caldwell (2001), Core sampling in support of the vadose zone transport field study, *Rep. PNNL-13454*, Pac. Northwest Natl. Lab., Richland, Wash.

Last, G. V., T. G. Caldwell, and A. T. Owen (2001), Sampling of boreholes WL-3A through -12 in support of the vadose zone transport field study, *Rep. PNNL-13631*, Pac. Northwest Natl. Lab., Richland, Wash.

Lee, S.-Y., S. F. Carle, and G. E. Fogg (2007), Geologic heterogeneity and a comparison of two geological models: Sequential Gaussian and transition-probability-based geostatistical simulation, *Adv. Water Resour.*, *30*, 1914–1932, doi:10.1016/j.advwatres.2007.03.005.

Li, W. (2007a), A fixed-path Markov chain algorithm for conditional simulation of discrete spatial variables, *Math. Geol.*, *39*(2), 159–176, doi:10.1007/s11004-006-9071-7.

Li, W. (2007b), Markov chain random fields for estimation of categorical variables, *Math. Geol.*, *39*(3), 321–335, doi:10.1007/s11004-007-9081-0.

Li, W., B. Li, and Y. Shi (1999), Markov-chain simulation of soil textural profiles, *Geoderma*, *92*, 37–53, doi:10.1016/S0016-7061(99)00024-5.

Lu, Z., and D. Zhang (2002), On stochastic modeling of flow in multimodal heterogeneous formations, *Water Resour. Res.*, *38*(10), 1190, doi:10.1029/2001WR001026.

Maji, R., E. A. Sudicky, S. Panday, and G. Teutsch (2006), Transition probability/Markov chain analysis of DNAPL source zone and plumes, *Ground Water*, *44*(6), 853–863, doi:10.1111/j.1745-6584.2005.00194.x.

Mayes, M. A., P. M. Jardine, T. L. Mehlhorn, B. N. Bjornstad, J. L. Ladd, and J. M. Zachara (2003), Transport of multiple tracers in variably saturated humid region structured soils and semi-arid region laminated sediments, *J. Hydrol.*, *275*, 141–161, doi:10.1016/S0022-1694(03)00039-8.

Meyer, P. D., M. L. Rockhold, and G. W. Gee (1997), Uncertainty analysis of infiltration and subsurface flow and transport for SDMP sites, *Rep. NUREG/CR-6565*, Off. of Nucl. Regul. Res., U.S. Nucl. Regul. Comm., Washington, D. C.

Miller, R. B., J. W. Castle, and T. J. Temples (2000), Deterministic and stochastic modeling of aquifer stratigraphy, South Carolina, *Ground Water*, *38*(2), 284–295, doi:10.1111/j.1745-6584.2000.tb00339.x.

Moysey, S., J. Caers, R. Knight, and R. M. Allen-King (2003), Stochastic estimation of facies using ground penetrating radar data, *Stochastic Environ. Res. Risk Assess.*, *17*(5), 306–318, doi:10.1007/s00477-003-0152-6.

Neuman, S. P. (1980), A statistical characterization of aquifer heterogeneity: An overview, in *Recent Trends in Hydrogeology*, edited by T. N. Narasimhan, *Spec. Pap. Geol. Soc. Am.*, *189*, 81–102.

Pace, M. N., M. A. Mayes, P. M. Jardine, T. L. Mehlhorn, J. M. Zachara, and B. N. Bjornstad (2003), Quantifying the effects of small-scale heterogeneities on flow and transport in undisturbed cores from the Hanford formation, *Vadose Zone J.*, *2*, 664–676.

Park, Y.-J., E. A. Sudicky, R. G. McLaren, and J. F. Sykes (2004), Analysis of hydraulic and tracer response tests within moderately fractured rock based on a transition probability geostatistical approach, *Water Resour. Res.*, *40*, W12404, doi:10.1029/2004WR003188.

Rawls, W. L., D. L. Brakensiek, and K. E. Saxton (1982), Estimation of soil properties, *Trans. ASAE*, *25*, 1316–1320.

Ritzi, R. W. (2000), Behavior of indicator variograms and transition probabilities in relation to the variance in lengths of hydrofacies, *Water Resour. Res.*, *36*(11), 3375–3381, doi:10.1029/2000WR900139.

Rockhold, M. L., C. J. Murray, and M. J. Fayer (1999), Conditional simulation and upscaling of soil hydraulic properties, in *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media*, edited by M. T. van Genuchten, F. J. Leij, and L. Wu, pp. 1391–1401, Univ. of Calif., Riverside, Calif.

Rubin, Y., and S. Hubbard (2005), *Hydrogeophysics*, *Water Sci. Technol. Libr.*, vol. 50, Springer, Dordrecht, Netherlands.

Schaap, M. G., and F. J. Leij (1998), Using neural networks to predict soil water retention and soil hydraulic conductivity, *Soil Tillage Res.*, *47*, 37–42, doi:10.1016/S0167-1987(98)00070-1.

Schaap, M. G., P. J. Shouse, and P. D. Meyer (2003), Laboratory measurements of the unsaturated hydraulic properties at the vadose zone transport field study site, *Rep. NNL-14284*, Pac. Northwest Natl. Lab., Richland, Wash.

Sisson, J. B., and A. H. Lu (1984), Field calibration of computer models for application to buried liquid discharges: A status report, *Rep. RHO-ST*, 46 pp., Rockwell Hanford Oper., Richland, Wash.

Sivakumar, B., T. Harter, and H. Zhang (2005), Solute transport in a heterogeneous aquifer: A search for nonlinear deterministic dynamics, *Nonlinear Processes Geophys.*, *12*, 211–218.

Srivastava, R., and T.-C. J. Yeh (1992), A three-dimensional numerical model for water flow and transport of chemically reactive solute through porous media under variably saturated conditions, *Adv. Water Resour.*, *15*, 275–287, doi:10.1016/0309-1708(92)90014-S.

Strebelle, S. (2002), Conditional simulation of complex geological structures using multiple-point statistics, *Math. Geol.*, *34*(1), 1–21, doi:10.1023/A:1014009426274.

Sun, A. Y., R. W. Ritzi, and D. W. Sims (2008), Characterization and modeling of spatial variability in a complex alluvial aquifer: Implications on solute transport, *Water Resour. Res.*, *44*, W04402, doi:10.1029/2007WR006119.

Tyler, K., A. Henriquez, and T. Avanes (1994), Modeling heterogeneities in fluvial domains: A review of the influence on production profiles, in *Stochastic Modeling and Geostatistics*, *AAPG Comput. Appl. Geol.*, vol. 3, edited by J. M. Yarus and R. L. Chambers, chap. 8, pp. 77–89, Am. Assoc. of Pet. Geol., Tulsa, Okla.

U.S. Department of Agriculture (1999), *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Surveys*, *Agric. Handb. 436*, 2nd ed., Washington, D. C.

van Genuchten, M. T., F. J. Leij, and S. R. Yates (1991), The RETC code for quantifying the hydraulic functions of unsaturated soils, version 1.0, *Rep. EPA/600/2-91/065*, U.S. Salinity Lab., Agric. Res. Serv., U.S. Dep. of Agric., Riverside, Calif.

Ward, A. L., T. G. Caldwell, and G. W. Gee (2000), Vadose zone transport field study: Soil water content distributions by neutron moderation, *Rep. PNNL-13795*, Pac. Northwest Natl. Lab., Richland, Wash.

Ward, A. L., et al. (2006a), Vadose zone transport field study summary report, *Rep. PNNL-15543*, Pac. Northwest Natl. Lab., Richland, Wash.

Ward, A. L., Z. F. Zhang, and G. W. Gee (2006b), Upscaling unsaturated hydraulic parameters for flow through heterogeneous, anisotropic sediments, *Adv. Water Resour.*, *29*, 268–280, doi:10.1016/j.advwatres.2005.02.013.

Weissmann, G. S., S. F. Carle, and G. F. Fogg (1999), Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics, *Water Resour. Res.*, *35*(6), 1761–1770, doi:10.1029/1999WR900048.

Yao, T. (2002), Integrating seismic data for lithofacies modeling: A comparison of sequential indicator simulation algorithms, *Math. Geol.*, *34*(4), 387–403, doi:10.1023/A:1015026926992.

Ye, M., R. Khaleel, and T.-C. J. Yeh (2005), Stochastic analysis of moisture plume dynamics of a field injection experiment, *Water Resour. Res.*, *41*, W03013, doi:10.1029/2004WR003735.

Ye, M., R. Khaleel, M. G. Schaap, and J. Zhu (2007), Simulation of field injection experiments in heterogeneous unsaturated media using cokriging and artificial neural network, *Water Resour. Res.*, *43*, W07413, doi:10.1029/2006WR005030.

Yeh, T.-C. J., and S. Liu (2000), Hydraulic tomography: Development of a new aquifer test method, *Water Resour. Res.*, *36*(8), 2095–2105, doi:10.1029/2000WR900114.

Yeh, T.-C. J., and J. Zhu (2007), Hydraulic/partitioning tracer tomography for characterization of dense nonaqueous phase liquid source zones, *Water Resour. Res.*, *43*, W06435, doi:10.1029/2006WR004877.

Yeh, T.-C. J., M. Ye, and R. Khaleel (2005), Estimation of effective unsaturated hydraulic conductivity tensor using spatial moments of observed moisture plume, *Water Resour. Res.*, *41*, W03014, doi:10.1029/2004WR003736.

Zhang, H., T. Harter, and B. Sivakumar (2006), Nonpoint source solute transport normal to aquifer bedding in heterogeneous, Markov chain random fields, *Water Resour. Res.*, *42*, W06403, doi:10.1029/2004WR003808.

Zhang, Z. F., A. L. Ward, and G. W. Gee (2004), A combined parameter scaling and inverse technique to upscale the unsaturated hydraulic parameters for heterogeneous soils, *Water Resour. Res.*, *40*, W08306, doi:10.1029/2003WR002925.

Zhu, J., and T.-C. J. Yeh (2005), Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resour. Res.*, *41*, W07028, doi:10.1029/2004WR003790.

R. Khaleel, Fluor Government Group, P.O. Box 1050, Richland, WA 99352, USA.

M. Ye, Department of Scientific Computing, Florida State University, Tallahassee, FL 32306, USA. (mye@fsu.edu)