



Quantification of uncertainty in pedotransfer function-based parameter estimation for unsaturated flow modeling

Hailin Deng,¹ Ming Ye,¹ Marcel G. Schaap,² and Raziuddin Khaleel³

Received 24 September 2008; revised 20 January 2009; accepted 2 February 2009; published 9 April 2009.

[1] While pedotransfer functions (PTFs) have long been applied to estimate soil hydraulic parameters for unsaturated flow and solute transport modeling, the uncertainty associated with the estimates is often ignored. The objective of this study is to evaluate uncertainty of the PTF-estimated soil hydraulic parameters and its effect on numerical simulation of moisture flow. Contributing to the parameter estimation uncertainty are (1) the PTF intrinsic uncertainty caused by limited data used for PTF training and (2) the PTF input uncertainty in pedotransfer variables (i.e., PTF inputs). The PTF intrinsic uncertainty is assessed using the bootstrap method by generating multiple bootstrap realizations of the soil hydraulic parameters; the realizations follow normal or lognormal distributions. The PTF input variables (i.e., bulk density and soil texture) are obtained using the cokriging technique. The PTF input uncertainty is quantified by assuming that the cokriging estimates follow a normal distribution. Our results show that the PTF input uncertainty dominates over the PTF intrinsic uncertainty and determines the spatial distribution of the PTF parameter estimation uncertainty. When the parameter estimation uncertainty is included, the spatial variability of the measured soil hydraulic parameters is better captured. This is also the case for the observed moisture contents, whose spatial variability is well bracketed by the prediction intervals. However, this is only possible after the PTF input uncertainty is considered. These results suggest that additional sample acquisition for the PTF input variables would have a more favorable impact on reduction of the parameter estimation uncertainty than collecting additional soil hydraulic parameter measurements for PTF development.

Citation: Deng, H., M. Ye, M. G. Schaap, and R. Khaleel (2009), Quantification of uncertainty in pedotransfer function-based parameter estimation for unsaturated flow modeling, *Water Resour. Res.*, 45, W04409, doi:10.1029/2008WR007477.

1. Introduction

[2] Available data on soil hydraulic parameters (e.g., saturated hydraulic conductivity and water retention parameters) are often inadequate for an accurate simulation of unsaturated flow and contaminant transport. Statistical methods have been developed to estimate the parameters using data that can be easily obtained such as moisture content, soil texture, bulk density, and geophysical data. Pedotransfer functions (PTFs) are statistical methods that estimate the soil hydraulic parameters from the pedotransfer input variables such as bulk density, soil texture, and organic carbon content [e.g., Rawls *et al.*, 1991; Tamari *et al.*, 1996; Schaap and Bouting, 1996; Schaap *et al.*, 1998; Pachepsky *et al.*, 1996, 1999; Minasny *et al.*, 1999; Wosten *et al.*, 2001; Pachepsky and Rawls, 2004]. Typically, the PTF-estimated parameters are used directly in numerical modeling, while uncertainty of the parameter estimates and its effect on modeling results are often not investigated or

are simply ignored. The parameter estimation uncertainty can be significant, especially when multiple kinds of data are used for the estimation. We hypothesize that quantifying the uncertainty can improve the parameter estimation and subsequent modeling predictions and that the uncertainty quantification can help evaluate usefulness of PTF-based parameter estimation methods for unsaturated flow modeling.

[3] The objective of this study is to investigate uncertainty of the PTF-based parameter estimates and its effect on moisture flow simulation. The research is conducted for a field site where a cokriging method is used to generate pedotransfer variables and an artificial neural network (ANN) based PTF is used to estimate soil hydraulic parameters for numerical modeling [Ye *et al.*, 2007a]. Uncertainty of the estimated soil hydraulic parameters is attributed to two sources: the PTF intrinsic uncertainty due to limited data used to train the PTF and PTF input uncertainty in the pedotransfer variables. The relative effect of the two kinds of uncertainty on the parameter estimation uncertainty is investigated. The Monte Carlo simulation is used to propagate the parameter estimation uncertainty in modeling flow in unsaturated media. Two sets of Monte Carlo simulations are conducted. The first set addresses only the PTF intrinsic uncertainty, while the second one considers both PTF intrinsic and input uncertainty. The relative contribution of the two kinds of uncertainty on the predictive uncertainty is

¹Department of Scientific Computing and Department of Geological Sciences, Florida State University, Tallahassee, Florida, USA.

²Department of Soil Water and Environmental Sciences, University of Arizona, Tucson, Arizona, USA.

³Fluor Government Group, Richland, Washington, USA.

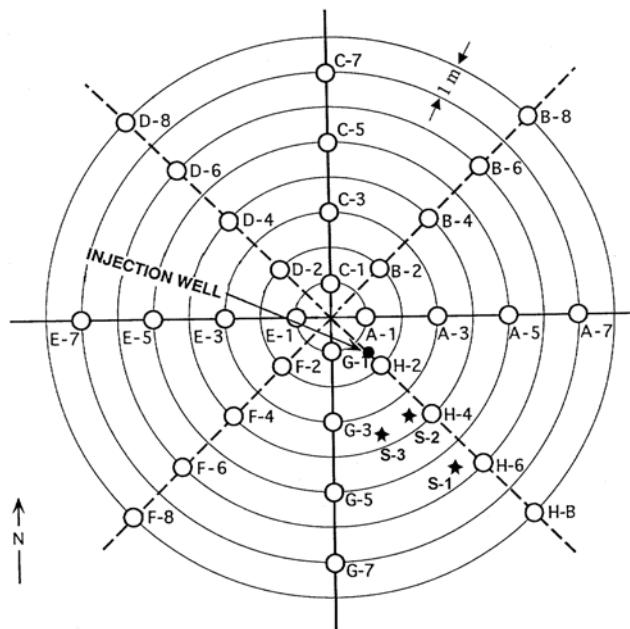


Figure 1. Plan view of the Sisson and Lu [1984] injection test site and well numbering scheme [after Gee and Ward, 2001].

also investigated. Exploring the relative contribution is important for uncertainty reduction, because limited resources can then be optimized to gather information on the most important uncertainty source. This study appears to provide the first comprehensive investigation on the relative contribution of PTF intrinsic and input uncertainties and their effects on parameter estimation uncertainty for unsaturated flow modeling.

[4] The uncertainty assessment is conducted for an ANN-based PTF developed by Ye *et al.* [2007a] for the Sisson and Lu (SL) field injection site [Sisson and Lu, 1984] located within the U.S. Department of Energy Hanford Site in southeastern Washington. The SL site was used for a field infiltration experiment from June to July in 2000 [Gee and Ward, 2001; Ward *et al.*, 2000, 2006a]. Initial moisture content distribution was measured on 5 May 2000 at the 32 radially and symmetrically arranged cased boreholes (Figure 1). Injections began on 1 June, and 4000 L of water were metered into an injection point 5 m below the land surface over 6 h. Similarly, 4000 L of water were injected in each subsequent injection on 8, 15, 22, and 28 June. During the injection period, neutron logging for the moisture content (θ) in 32 wells took place within a day (i.e., 2, 9, 16, and 23 June) following each of the first four injections. A wildfire burned close to the field site preventing immediate logging of the θ distribution for the fifth injection on 28 June. Three additional readings of the 32 wells were subsequently completed on 7, 17, and 31 July. During each neutron logging, moisture contents were monitored in each well at a depth interval of 0.3048 m (1 foot) starting from a depth of 3.9625 m (13 feet) and continuing to a depth of 16.764 m (55 feet), resulting in a total of 1376 measurements in each of the eight observation days over a 2-month period. The moisture content measurements, especially

those of initial moisture contents, reflect soil heterogeneity at the site, which is the basis for developing the ANN-based PTF for estimating heterogeneous soil hydraulic parameters.

[5] The ANN-based PTF is able to estimate the three-dimensional (3-D) distribution of soil hydraulic parameters from 3-D distributions of PTF input variables (i.e., bulk density and soil texture) obtained using the cokriging method. While the cokriged PTF input variables are uncertain (the uncertainty being measured by cokriging variance), the PTF input uncertainty was not considered by Ye *et al.* [2007a], and the cokriged PTF inputs were treated as being deterministic in the PTF. On the other hand, Ye *et al.* [2007a] also ignored the PTF intrinsic uncertainty in the estimated soil hydraulic parameters, and used only mean parameter estimates to simulate the SL site field injection experiment. Although the simulated moisture contents agreed reasonably well with the corresponding field measurements and the agreement is comparable with previous modeling studies of the same experiment [e.g., Zhang *et al.*, 2004; Yeh *et al.*, 2005; Kowalsky *et al.*, 2005; Ward *et al.*, 2006b], a mismatch was observed at each borehole [Ye *et al.*, 2007a]. This raises several questions such as: does the mismatch indicate that the PTF parameter estimates are inadequate to simulate the observed moisture content variability? Will consideration of the parameter estimation uncertainty improve the simulation in the sense that the observed moisture content variability can be sufficiently captured? The answers to these questions are explored in this study.

[6] According to Chirico *et al.* [2007], PTF parameter estimation uncertainty is attributed to three major sources: (1) PTF model parameters, (2) PTF input variables, and (3) PTF model structures. The PTF model parameter uncertainty is referred to in this paper as the PTF intrinsic uncertainty to distinguish it from PTF parameter estimation uncertainty of the soil hydraulic parameters. Exploring the intrinsic uncertainty is not a common practice in PTF development. Schaap and Leij [1998] estimated such uncertainty using the nonparametric bootstrap method [Efron, 1992]. In the work by Schaap and Leij [1998], the bootstrap method randomly resamples with replacement of an original data set of size N into a bootstrap realization also of size N . To simulate the repeated collection of data, the selection procedure is repeated M times for the original data set, with each bootstrap realization containing approximately 63% of the original data. An ANN-based PTF is calibrated for each of these M realizations and validated for approximately 37% of the data not contained in the specific bootstrap realization. The multiple realizations of the calibration and validation data set represent the PTF intrinsic uncertainty; the uncertainty is quantified by assigning a probabilistic distribution for the M realizations of parameter estimates. Although Schaap and Leij [1998] developed the theoretical basis for assessing the PTF intrinsic uncertainty, typically only the mean values for the parameter estimates are used in numerical modeling. This ignores the PTF intrinsic uncertainty; the mean estimates, however, may not be able to explain the observed physical behavior. For a hydrologic model, Srivastav *et al.* [2007] reported that use of the bootstrap mean failed to capture the hydrograph peak flow characteristics. This is also observed by Ye *et al.* [2007a], wherein the bootstrap mean parameter estimates cannot

adequately capture the observed moisture content variability. To our knowledge, investigating the effect of bootstrap parameter estimation uncertainty on unsaturated flow modeling has not yet been reported.

[7] Another source of PTF parameter estimation uncertainty is the PTF input uncertainty, which can be due to data error and spatial or temporal variability of PTF inputs. This uncertainty is usually ignored in PTF applications, because it is either unknown or difficult to propagate through the modeling. This study presents a unique opportunity of assessing the PTF input uncertainty caused by spatial variability, because the PTF input variables are estimated using the cokriging method and the input uncertainty is measured by the cokriging variance. In addition to factors contributing to (co)kriging variance as discussed by Isaaks and Srivastava [1989], the cokriging variance can also be due to the error in estimating cross variogram [Lark, 2003] and the error in primary and secondary data [Abbaspour et al., 1998]. As shown below, these PTF inputs, estimated largely from the initial moisture content (θ_i , cm^3/cm^3 , %), are subject to significant uncertainty; this uncertainty dominates the total PTF parameter estimation uncertainty, and increases predictive uncertainty of simulated moisture content.

[8] PTF structure is another source of uncertainty [Chirico et al., 2007]; that is, different PTFs implement different conceptualizations of physical or statistical realities and thus yield different estimations. One way of addressing this uncertainty is to compare the parameter estimates with corresponding measurements and to select the optimal model that gives the minimum difference [Williams et al., 1992; Tietje and Tapkenhinrichs, 1993; Kern, 1995; Minasny et al., 1999; Cornelis et al., 2001]. This method, however, does not guarantee that the selected model will give satisfactory predictions for the variables of interest. An alternate approach is to conduct numerical modeling using parameter estimates for various PTFs, and to select the best PTF model that gives the minimum difference between the simulated and observed state variables [Espino et al., 1996; Finke et al., 1996; Minasny and Field, 2005; Mermoud and Xu, 2006; Dai et al., 2008]. Because a single model is used for both methods, neither method completely addresses the PTF model uncertainty. Chirico et al. [2007] found that the PTF model structure uncertainty can contribute to PTF parameter estimation uncertainty more than the PTF input uncertainty. Pachepsky et al. [2006] suggested that the recently developed method of Bayesian model averaging [Neuman, 2003; Ye et al., 2004, 2005a, 2008] is a promising tool for addressing the PTF model uncertainty. However, addressing the model uncertainty is beyond the scope of this study.

2. PTF Intrinsic and Input Uncertainty

[9] The uncertainty assessment is conducted for the ANN-based PTF [Ye et al., 2007a] for estimating the soil hydraulic parameters: saturated hydraulic conductivity (K_s , m/d), saturated moisture content (θ_s , cm^3/cm^3), residual moisture content (θ_r , cm^3/cm^3), and van Genuchten parameters α (1/m) and n (–) [van Genuchten, 1980]. The parameters are estimated from the bulk density (BD, kg/m^3) and soil texture: gravel content ($>2000 \mu\text{m}$) (GR, %), coarse sand content (2000 to 200 μm) (CS, %), fine sand content

(200 to 50 μm) (FS, %), silt content (50 to 5 μm) (SI, %), and clay content ($<5 \mu\text{m}$) (CL, %). The ANN-based PTF was trained using laboratory measurements of the input and output variables for 70 core samples, 17 samples are from boreholes E-7, E-1 and A-7 [Khaleel and Freeman, 1995; Khaleel et al., 1995], and 53 from boreholes S-1, S-2 and S-3 [Schaap et al., 2003]. The location of the boreholes is shown in Figure 1. The calibrated PTF was used to estimate the soil hydraulic parameters for the 16 m \times 16 m modeling domain shown in Figure 1. The PTF inputs for the domain were estimated using the cokriging method, in which the initial moisture content (θ_i , %) was used as the secondary variable, because there are 1344 θ_i observations for 32 wells (Figure 1) at a vertical interval of 30.5 cm (1 foot). Ye et al. [2005b] found that the initial moisture content data carry the signature of soil heterogeneity, which was used by Ye et al. [2007a] as the secondary information for cokriging. A detailed description of the field site and the injection experiments can be found in work by Ward et al. [2000], Gee and Ward [2001], and Ward et al. [2006a].

2.1. PTF Intrinsic Uncertainty

[10] To quantify the PTF intrinsic uncertainty, we denote the true soil hydraulic parameters as β^* and the corresponding PTF estimates as $\hat{\beta}$. The $\hat{\beta}$ estimates can be obtained from a training data set containing measurements of the PTF inputs and outputs. Because of limited samples for the data set, the estimated parameters, $\hat{\beta}$, are random and, in general, are not equal to β^* . Evaluating the uncertainty for β requires knowing the distribution of $\hat{\beta} - \beta^*$, which however is unknown because β^* is unknown. Estimating the distribution invokes certain assumptions, i.e., linearity of the PTF and normal distribution of $\hat{\beta} - \beta^*$ as in linear and nonlinear regressions. The nonparametric bootstrap method [Efron and Tibshirani, 1993] is ideally suited for estimating the distribution because it does not invoke these particular assumptions. When implementing the bootstrap method, each bootstrap realization gives a bootstrap estimate, β_i . For a sufficiently large number of realizations, the mean of β_i is the parameter estimate, $\hat{\beta}$; the variance of β_i approximates the variance of the unknown parameter estimation uncertainty $\beta^* - \hat{\beta}$ [Efron and Tibshirani, 1993; Chernick, 1999]. This assumes that the bootstrap realizations obtained from the data set represent the data population, which, however, cannot be confirmed in reality since the population is unknown. To achieve convergence of the bootstrap mean and variance, Efron and Tibshirani [1993] suggest that the number of Monte Carlo realizations be between 50 and 200. Schaap and Leij [1998] used 100; this study uses 10,000 realizations to ensure convergence in variance.

[11] Figure 2 illustrates the histograms for $\log_{10}(K_s)$, $\log_{10}(\alpha)$, $\log_{10}(n)$, and θ_s for a sandy sample (plot of θ_r not shown); the histograms fit well to normal distribution. This is also true for other samples having different soil textures (i.e., coarse sand, loamy sand and sandy loam types for the 70 core measurements) (figures not shown). This suggests that the uncertainty of each PTF output parameter (or its log transform) can be described using a normal distribution with the mean and variance estimated from the bootstrap realizations.

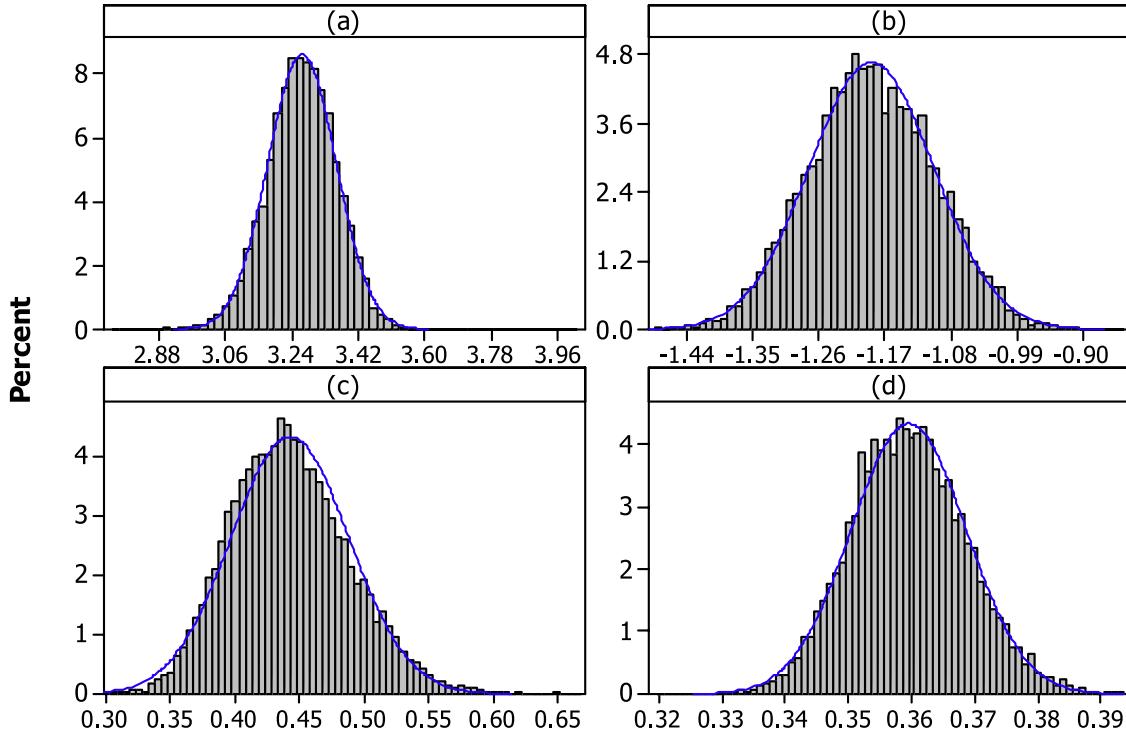


Figure 2. Histograms and fitted normal distributions for (a) $\log_{10}(K_s)$, (b) $\log_{10}(\alpha)$, (c) $\log_{10}(n)$, and (d) θ_s for a sandy sample based on 10,000 bootstrap realizations.

2.2. PTF Input Uncertainty

[12] When a PTF is used for predictions, the PTF inputs are often different from the data used for PTF calibration. For example, *Ye et al.* [2007a] calibrated the PTF using data based on core measurements, while the PTF inputs used for predictions were estimated using cokriging, which has a location-dependent uncertainty. The cokriging estimates the primary variable, Z , at point $\mathbf{x}_{Z,0}$ on the basis of a linear combination of measurements for the primary variable and the secondary variable, Y , via

$$Z^*(\mathbf{x}_{Z,0}) = \sum_{i=1}^I u_i(\mathbf{x}_{Z,0}) Z(\mathbf{x}_{Z,i}) + \sum_{j=1}^J v_j(\mathbf{x}_{Z,0}) Y(\mathbf{x}_{Y,j}), \quad (1)$$

where u_i and v_i are weights of sample Z at $\mathbf{x}_{Z,i}$ and Y at $\mathbf{x}_{Y,j}$, respectively [Journel and Huijbregts, 1978]. The bulk density and soil textures (GR, CS, FS, SI, and CL) are the primary variables; the initial moisture contents (θ_i) are used as the secondary variable, since they are abundant and evenly distributed in the domain. More importantly, the θ_i data carry signature about site heterogeneity, especially the alternating layering structure of fine and coarse materials [*Ye et al.*, 2005b]. To ensure that the cokriged value is unbiased, the following requirements must be satisfied:

$$\sum_{i=1}^I u_i(\mathbf{x}_{Z,0}) = 1 \quad \text{and} \quad \sum_{j=1}^J v_j(\mathbf{x}_{Z,0}) = 0. \quad (2)$$

The weights are obtained by solving the cokriging equations, which can be found in geostatistics textbooks [e.g., *Journel and Huijbregts*, 1978; *Isaaks and Srivastava*,

1989]. The uncertainty of estimating Z^* is measured by the cokriging variance,

$$\begin{aligned} \text{Var}(Z^*(\mathbf{x}_{Z,0})) &= \text{Cov}\{Z(\mathbf{x}_{Z,0}), Z(\mathbf{x}_{Z,0})\} + \mu_1 \\ &\quad - \sum_{i=1}^n u_i(\mathbf{x}_{Z,0}) \text{Cov}\{Z(\mathbf{x}_{Z,i}), Z(\mathbf{x}_{Z,0})\} \\ &\quad - \sum_{j=1}^m v_j(\mathbf{x}_{Z,0}) \text{Cov}\{Y(\mathbf{x}_{Y,j}), Z(\mathbf{x}_{Z,0})\}, \end{aligned} \quad (3)$$

where Cov stands for covariance, and μ_1 is the Lagrangian multiplier due to the constraint $\sum_{i=1}^I u_i(\mathbf{x}_{Z,0}) = 1$. The first covariance (i.e., the first term on the right side of equation (3)) is the variance of the primary variable, and the third covariance is the cross covariance between the primary and the secondary variables. Since the cokriging of equation (1) may limit the influence of the secondary variable Y on the primary variable Z , a standardized cokriging is always used by rewriting (1) as [Isaaks and Srivastava, 1989]

$$Z^*(\mathbf{x}_{Z,0}) = \sum_{i=1}^I u_i(\mathbf{x}_{Z,0}) Z(\mathbf{x}_{Z,i}) + \sum_{j=1}^J v_j(\mathbf{x}_{Z,0}) [Y(\mathbf{x}_{Y,j}) - m_Y], \quad (4)$$

where m_Z and m_Y are stationary mean of Z and Y . The constraints (equation (2)) ensure that the unbiased nature of Z^* become [Isaaks and Srivastava, 1989]

$$\sum_{i=1}^I u_i(\mathbf{x}_{Z,0}) + \sum_{j=1}^J v_j(\mathbf{x}_{Z,0}) = 1. \quad (5)$$

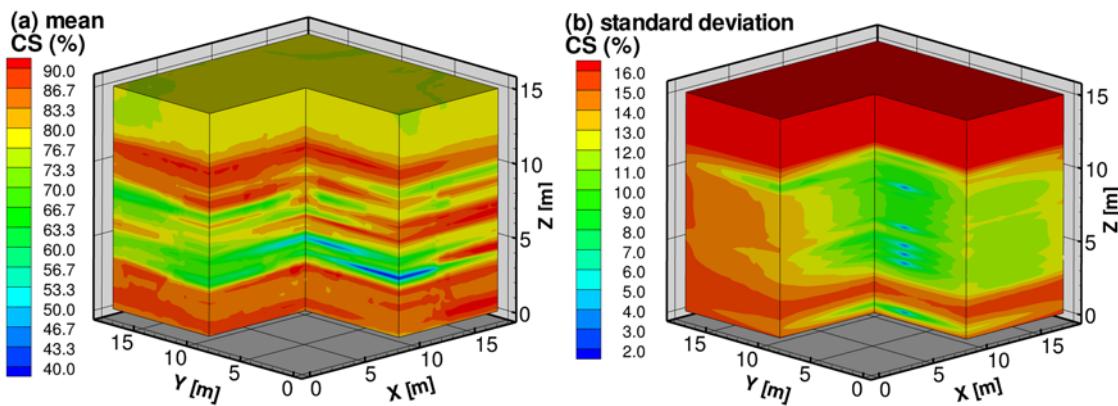


Figure 3. Three-dimensional contours of cokriging (a) mean estimate and (b) standard deviation (square root of cokriging variance) for the coarse sand percentage.

In this case, the cokriging variance is still expressed via equation (3). Although there is no theoretical proof about distribution of the cokriged variables, it appears reasonable to assume that they are Gaussian, whose mean and variance are the cokriged value and cokriging variance [e.g., Kitanidis, 1997; Cressie, 1991; Deutsch and Journel, 1998]. This distribution is used below for random sampling for the Monte Carlo simulations.

[13] Abbaspour *et al.* [1998] developed an alternative to cokriging for situations with small sample sizes; the method provides a means to quantify uncertainty of the estimated primary variable without having to assume the normal distribution. On the other hand, Lark [2003] developed two robust estimates of the cross variogram (equivalent to the cross covariance in equation (3)) for more reliable estimates of the cokriging variance, especially when outliers exist in data. Although not tested in this study, these two methods are expected to improve estimation of cokriging variance.

[14] While many applications use the cokriged values directly for numerical modeling, the cokriging variance may be too large to be ignored, especially when primary data are sparse. Figure 3 illustrates the three-dimensional (3-D) contours of cokriged coarse sand percentage (CS, %) and its standard deviation. The CS standard deviation can be as large as 1/5 of the cokriged value, which is also true for other textures. The smaller standard deviation regions (i.e., the blue spots in Figure 3b) correspond to locations where the CS measurements are available. As most of the measurements are on right side of the domain, the standard deviation is smaller for the right side ($x > 9$ m) than for the left. A larger standard deviation for the top and bottom of the domain is due to lack of measurements for the primary and secondary variables. The same pattern is also observed for other pedotransfer variables. Figure 3 suggests that, when the cokriged PTF input variables are used, the cokriging uncertainty cannot be ignored.

3. Uncertainty of PTF-Based Parameter Estimation

[15] Ye *et al.* [2007a] found that the PTF-estimated parameter values cannot explain the spatial variability inherent in soil properties measurements; this study demonstrates that this problem can be partially resolved by considering the uncertainty, especially the PTF input

uncertainty. This section presents the propagation of the parameter estimation uncertainty through numerical simulation of the SL field experiment [Ye *et al.*, 2007a] and the subsequent comparison to field measurements.

3.1. Parameter Estimation Uncertainty due to PTF Intrinsic Uncertainty

[16] Using K_s and α as examples, Figure 4 depicts the 3-D contours of their mean and standard deviation calculated from the 10,000 realizations of PTF-estimated parameters. Because large mean values correspond to large values of the standard deviation, the coefficient of variation (CV) (standard deviation divided by mean) is used to measure the PTF intrinsic uncertainty. The CV ranges for the entire modeling domain are 15–40%, 15–30%, 5–15%, 2–5% and 6–13% for K_s , α , n , θ_s and θ_r , respectively. This indicates that the PTF intrinsic uncertainty is relatively smaller for θ_s and θ_r , which is not surprising since the variability for these two parameters is known to be smaller than the variability for the other three parameters.

[17] Figure 5 displays the spatial variability of the measured soil hydraulic parameters at borehole S-1 (location shown in Figure 1), as well as the mean estimates and 95% prediction intervals (mean \pm 1.96 standard deviation) of the PTF estimates. The mean parameter estimates, as observed by Ye *et al.* [2007a], cannot fully capture the spatial variability. While more measurements are included in the prediction intervals after the PTF intrinsic uncertainty is considered, a large portion of the measurements is still outside of the intervals. Ye *et al.* [2007a] hypothesized that the smoothing effect of cokriging may be a reason for the small variability of the PTF estimates, because only the mean cokriging estimates of the pedotransfer variables are used as the PTF inputs. This is explored below.

3.2. Parameter Estimation Uncertainty due to PTF Intrinsic and Input Uncertainty

[18] In order to explore the effect of the PTF input uncertainty (i.e., the cokriging uncertainty), multiple realizations of the cokriged PTF input variables are sampled from their distributions, which are assumed to be Gaussian (section 2.2). Since each element of the computation grid has a cokriging estimate and a cokriging variance, the sampling is conducted for each element using the Latin hypercube sampling (LHS) method [McKay *et al.*, 1979].

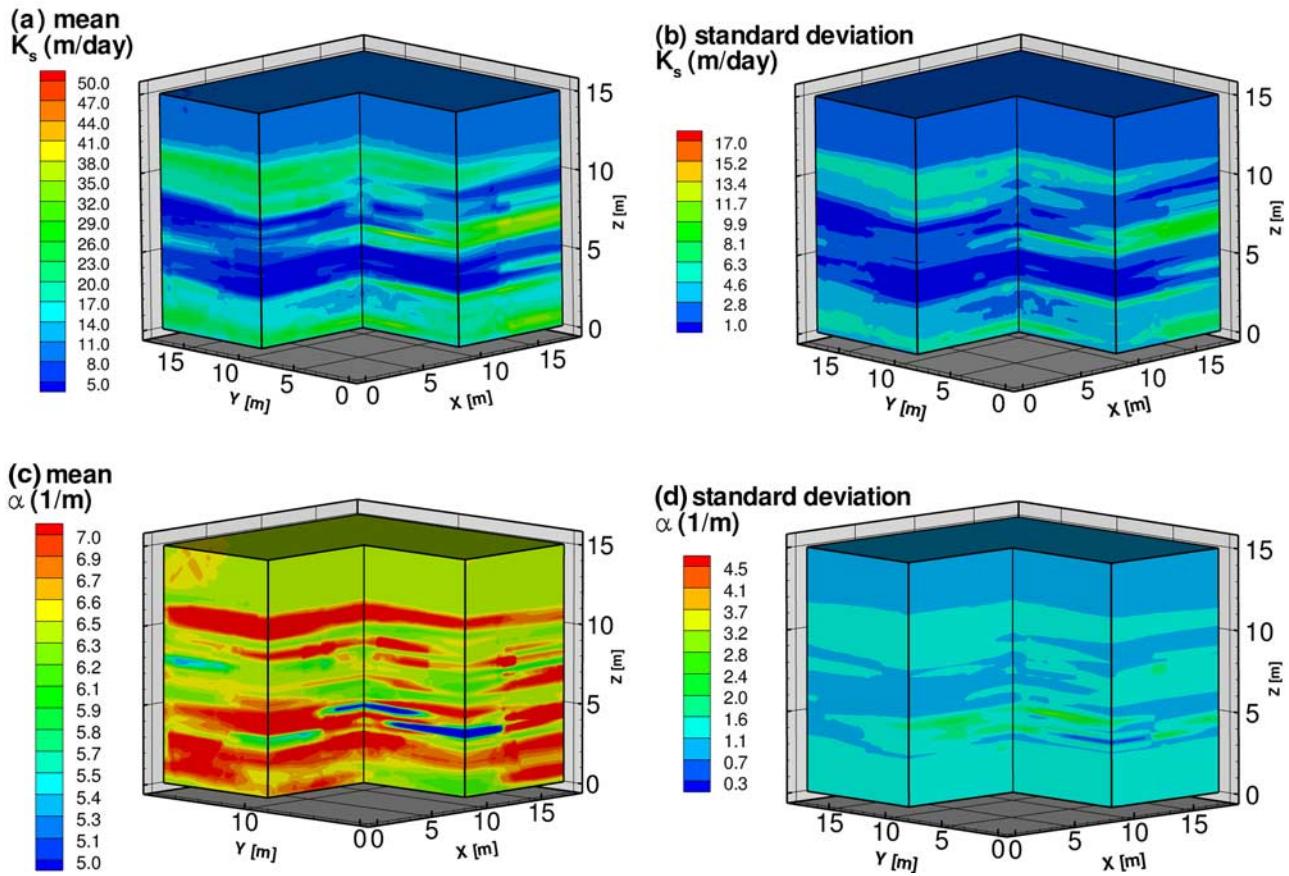


Figure 4. Three-dimensional fields of the saturated hydraulic conductivity (K_s) (a) mean and (b) standard deviation and the van Genuchten α (c) mean and (d) standard deviation based on 10,000 bootstrap realizations. The parameters are estimated by the ANN-based PTF using the cokriged pedotransfer variables as inputs; the cokriging uncertainty is not considered.

Subsequently, the random samples at the elements are combined to form the random fields of the grid. The statistical correlation between the input variables is considered using the Spearman rank correlation coefficient; the spatial correlation of the samples is not considered during the sampling. To ensure that the random samples of the pedotransfer variables are physically meaningful (e.g., not being negative), the bounded normal distribution implemented in the LHS code of the DAKOTA software (<http://www.cs.sandia.gov/DAKOTA/>) [Swiler and Wyss, 2004] is used; the minimum and maximum values of pedotransfer variables for the 70 samples are used as the bounds. After running the PTF with each random field of the PTF input variables, the LHS is used to sample a number of soil hydraulic parameters from the bootstrap distributions obtained in section 2.1. A total of 400 realizations of the soil hydraulic parameters are generated for 20 LHS samples of the soil hydraulic parameters, in addition to 20 LHS samples of PTF input variables. The 400 realizations (i.e., 20 for hydraulic parameters times 20 for PTF inputs) thus incorporate PTF intrinsic as well as the PTF input uncertainty.

[19] Figure 6 shows the estimated soil hydraulic parameters and the mean, minimum, and maximum estimates of the parameters along the S-1 borehole. A comparison of Figures 5 and 6 shows that, whereas the mean parameter

estimates are similar, the parameter estimation uncertainty increases dramatically. This indicates that the PTF parameter estimation uncertainty is primarily caused by the PTF input uncertainty, which dominates over the PTF intrinsic uncertainty. Incorporating the PTF input uncertainty improves the parameter estimation, because most of the measured parameters, except some extreme values, are within ranges of the estimated values. For the θ_s , α , n , and K_s parameters shown in Figure 6, 79%, 75%, 82% and 94% of measurements are encompassed in the ranges. This suggests that, if the PTF intrinsic and input uncertainty had been considered, the ANN-based PTF [Ye et al., 2007a] would have been able to capture the observed soil hydraulic parameter variability. This differs from the finding of Gutmann and Small [2005, 2007] that using soil texture alone is inadequate to describe the soil hydraulic parameters. The finding of Ye et al. [2007a] may have resulted from their use of bulk density as an additional PTF variable.

[20] Figure 7 shows additional evidence that the increase of uncertainty and the improvement of parameter estimates is attributed to consideration of the PTF input uncertainty. Figure 7 illustrates the measured pedotransfer variables along borehole S-1 and their corresponding cokriging estimates for the mean and 95% prediction intervals. A comparison of Figures 5–7 shows that the spatial pattern of the parameter estimation uncertainty (Figure 6) is deter-

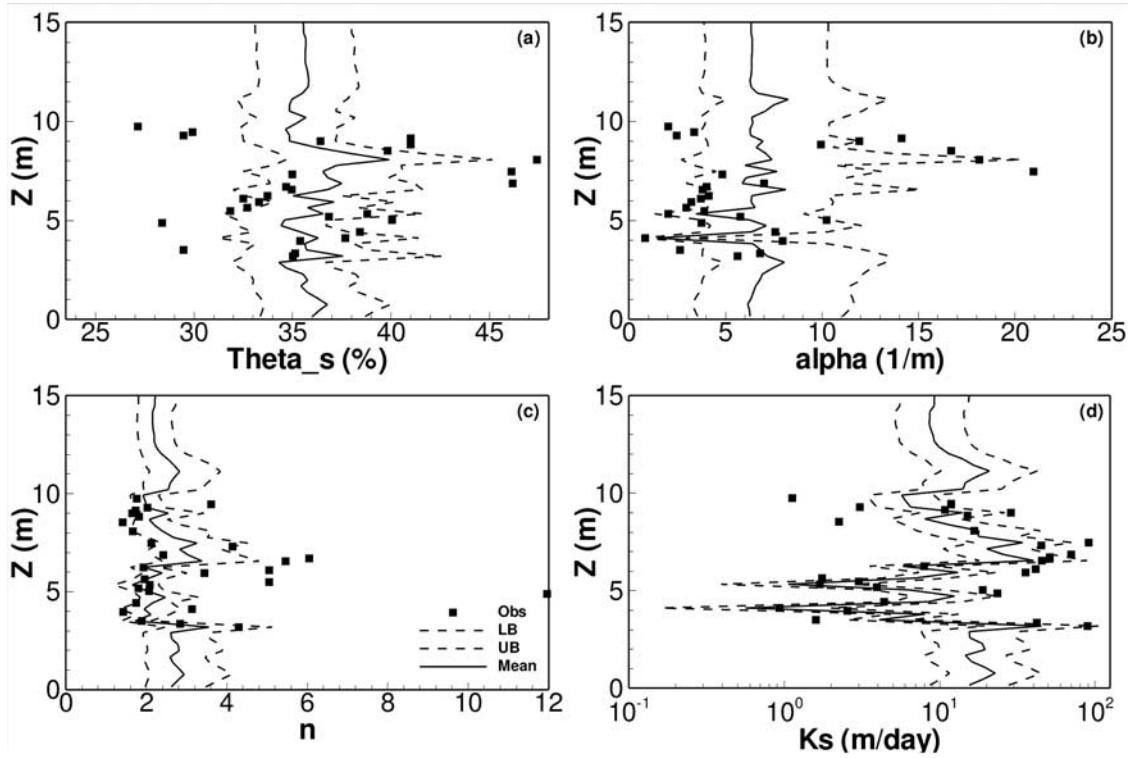


Figure 5. Comparison of the measured (squares) and PTF-estimated (a) saturated water content (θ_s), (b) van Genuchten α , (c) van Genuchten n , and (d) saturated hydraulic conductivity (K_s) at borehole S-1 (location shown in Figure 1) when PTF input uncertainty (i.e., cokriging uncertainty) is not included. The mean PTF estimates are shown as solid lines, and the 95% prediction intervals are shown as dashed lines.

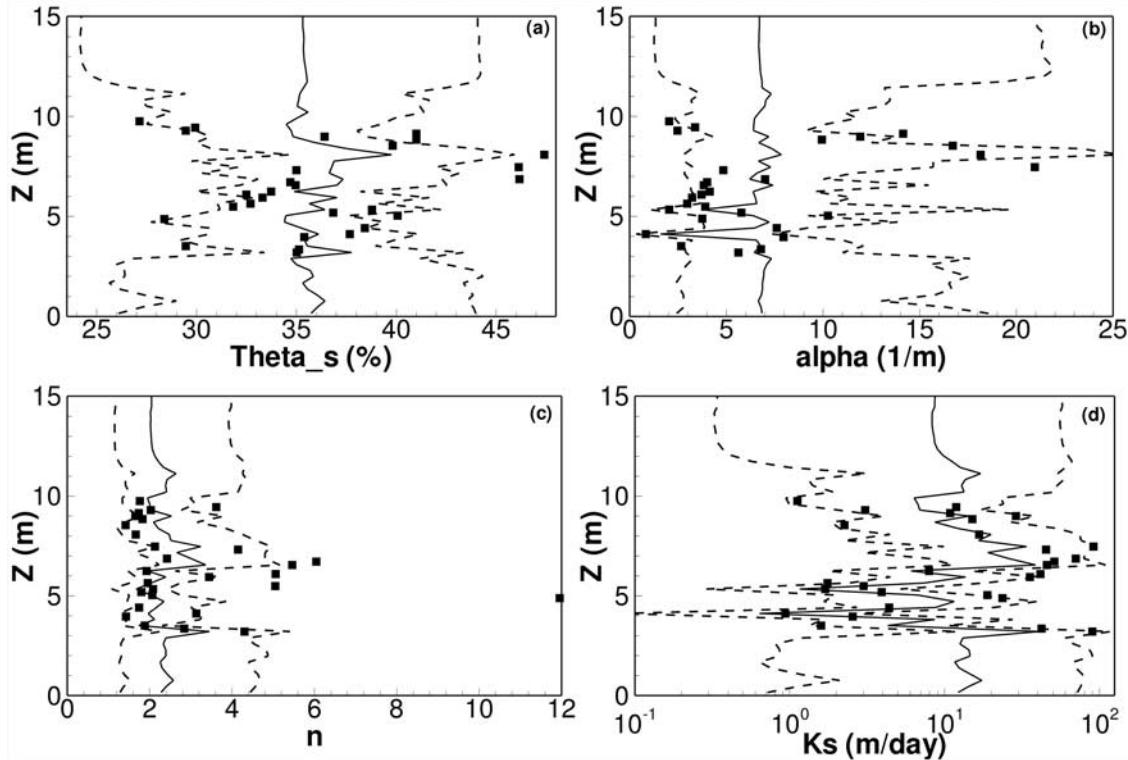


Figure 6. Comparison of the measured (squares) and PTF-estimated (a) saturated water content (θ_s), (b) van Genuchten α , (c) van Genuchten n , and (d) saturated hydraulic conductivity (K_s) at borehole S-1 (location shown in Figure 1) following incorporation of PTF input uncertainty (i.e., cokriging uncertainty). The mean PTF estimates are shown as solid lines, and the 95% prediction intervals are shown as dashed lines.

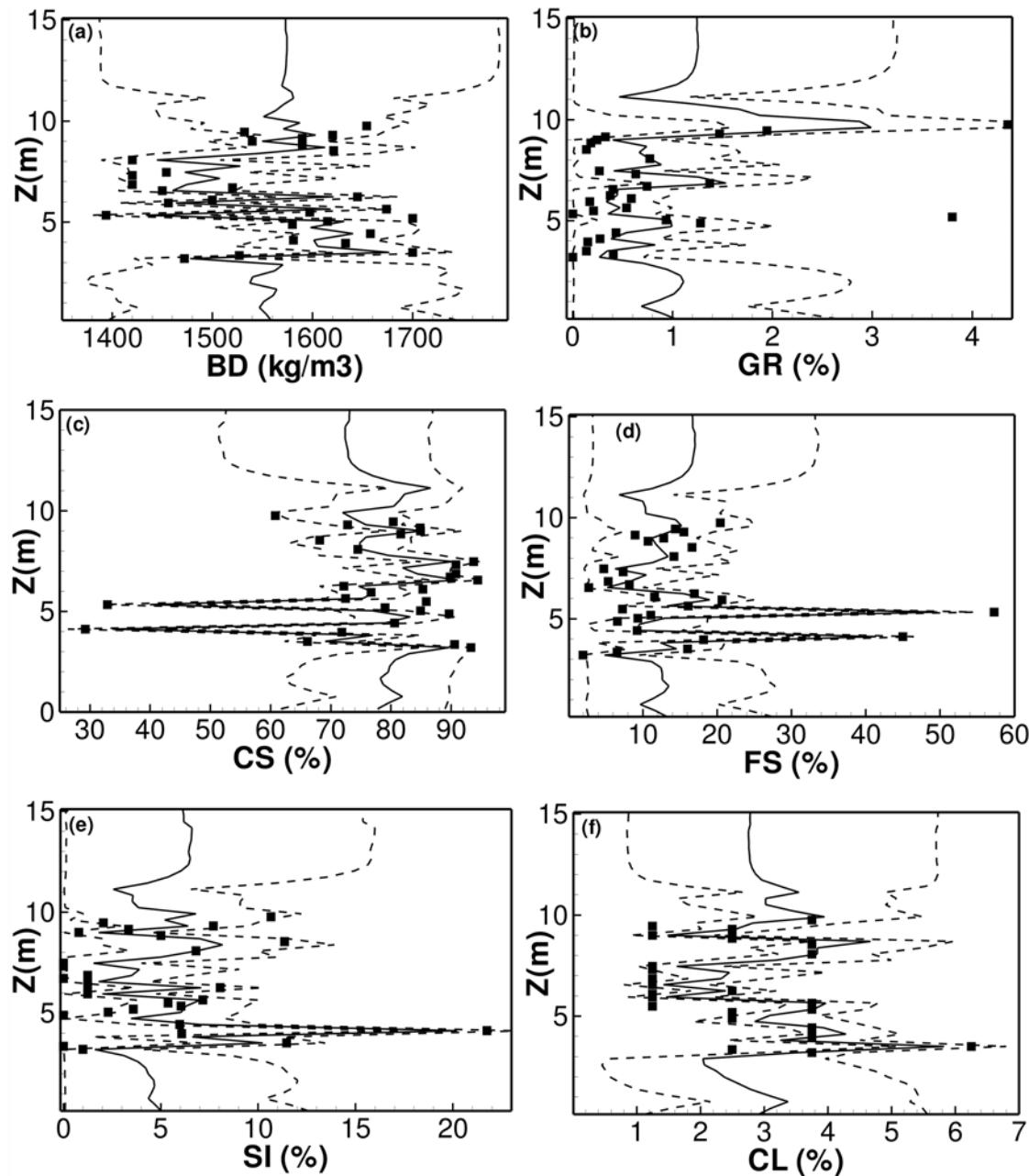


Figure 7. (a–f) Comparison of the observed (squares) and cokriging-estimated pedotransfer variables at borehole S-1 (location shown in Figure 1). The cokriged values are shown as solid lines, and the 95% prediction intervals are shown as dashed lines.

mined by the PTF input uncertainty (Figure 7) rather than by the PTF intrinsic uncertainty (Figure 5). This is not surprising since the PTF intrinsic uncertainty is smaller and more uniformly distributed in the 3-D domain.

3.3. Propagation of the PTF-Based Parameter Estimation Uncertainty

[21] This section investigates propagation of the parameter estimation uncertainty through numerical simulation of the 2000 injection experiment at the SL site. As discussed earlier, despite the fact that the overall agreement between the *Ye et al.* [2007a] observed and simulated moisture content was reasonably good, observed a mismatch of the

observed and simulated moisture contents at each borehole. On the basis of the simulation results presented in this study, we argue that the mismatch resulted because the parameter estimation uncertainty was ignored, and not because of the use of an inadequate parameter estimation method via the cokriging and ANN-based PTF.

[22] As in the work by *Ye et al.* [2007a], the numerical simulation of the 2000 injection experiment is conducted using the MMOC code [Srivastava and Yeh, 1992]. The planar area of the simulation domain is $18 \text{ m} \times 18 \text{ m}$ to include the entire sampling area depicted in Figure 1. Although moisture contents were measured to a depth of 16.775 m, the vertical dimension of the simulation domain

is 15.24 m. This is due to the fact that movement of the injected water is hindered by the second layer of fine material, and the moisture content below the 15 m depth essentially remains unchanged during the 2000 injection experiment [Ye *et al.*, 2005b]. The simulation domain is discretized into a grid with 259,200 uniform elements, each element being 0.25 m (Δx) \times 0.25 m (Δy) \times 0.3048 m (Δz). For each realization, the initial pressure head is estimated from the initial moisture content using the van Genuchten retention model and the soil hydraulic parameters. Constant head boundary conditions are assumed for all sides of the simulation domain with pressure head equal to the estimated initial head. Therefore, the initial and boundary conditions are different for different realizations because of the randomly assigned soil hydraulic parameters.

[23] Two sets of Monte Carlo simulations are conducted to investigate propagation of the parameter estimation uncertainty. The first one considers only the PTF intrinsic uncertainty; the second one considers both the PTF intrinsic and input uncertainty. For the first set of MC simulations, in order to save computational time, instead of using all the 10,000 bootstrap realizations, 100 parameter realizations are generated using the LHS method on the basis of the Gaussian distributions obtained from the 10,000 realizations. For the second set of MC simulations, as discussed in section 3.2, 400 parameter realizations are generated on the basis of the cokriging and bootstrap distributions. The use of more parameter realizations for the second set MC simulation is considered appropriate, since the uncertainty is larger. After the MC simulations are completed, mean and 95% prediction intervals (mean \pm 1.96 standard deviation) of the simulated moisture content are calculated for each element of the computational grid. The goodness of fit for the simulations is evaluated using the Pearson correlation coefficient (r):

$$r = \frac{\sum_{i=1}^n (\theta_i^* - \bar{\theta}^*) (\theta_i - \bar{\theta})}{\sqrt{\sum_{i=1}^n (\theta_i^* - \bar{\theta}^*)^2} \sqrt{\sum_{i=1}^n (\theta_i - \bar{\theta})^2}} \quad (6)$$

and the root-mean-square error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\theta_i^* - \theta_i)^2}, \quad (7)$$

where θ^* and θ are the observed and simulated mean moisture contents, respectively, $\bar{\theta}^*$ and $\bar{\theta}$ are the mean values for the entire data set (n).

[24] On the basis of results for the first set of MC simulations (only PTF intrinsic uncertainty is considered), Figure 8 illustrates the observed moisture contents, the simulated mean and the 95% prediction intervals for four boreholes. These boreholes are selected because they are most affected by the injection and the simulated and the observed water contents are for the same locations. Figure 8 is for the data observed on 23 June 2000 during the injection process; the comparison for other dates is similar and is not shown. The Pearson correlation coefficient, r , and RMSE values shown in Figure 8 are about the same as those

reported by Ye *et al.* [2007a]; the 95% prediction intervals are narrow, and only a small number of observations are included in the intervals. This indicates that considering only the PTF intrinsic uncertainty is insufficient to explain the spatial variability of the observed moisture content.

[25] Similar to Figure 8, Figure 9 illustrates the results for the second set of Monte Carlo simulations considering the PTF intrinsic as well as the input uncertainty. A comparison of Figures 8 and 9 shows that the mean predictions are improved, as indicated by the smaller RMSE values and larger Pearson correlation coefficients (Figure 9). More importantly, the 95% prediction intervals increase significantly, because the PTF parameter estimation uncertainty increases dramatically after the PTF input uncertainty is considered (Figures 5 and 6). As a result, almost all the observed moisture contents are included in the intervals, indicative of an improvement of the numerical simulation. The good quality of the coverage of the 95% prediction intervals suggests that, if the parameter estimation uncertainty had been considered, the parameter estimation method of Ye *et al.* [2007a] would have been able to explain the spatial variability of the field observations. This conclusion is not drawn solely on the basis of the observation that increasing the 95% prediction intervals captures more of the observed θ ; it is also supported by the fact that the mean predictions capture the overall spatial variability of the observed θ (Figures 8 and 9). It is generally true that considering more uncertainty sources (e.g., the PTF input uncertainty in addition to the PTF intrinsic uncertainty) will increase the predictive uncertainty. However, if the parameter estimation method (or a model) is inadequate, an increase in predictive uncertainty does not guarantee that the mean predictions will satisfactorily mimic the overall variability of field observations. As a result, if mean predictions deviate significantly from the majority of observations, the field data do not typically fall into the uncertainty bounds with an increase in bounds, unless the predictive uncertainty is unreasonably large. For example, in simulating the matrix moisture content for Yucca Mountain site [Ye *et al.*, 2007b], bulk of the field observations for matrix θ fell outside of uncertainty bounds, when the matrix permeability was treated as a random variable. This occurred because the mean predictions deviated severely from majority of field observations. Following treatment of matrix van Genuchten α and n as random variables [Pan *et al.*, 2009], the uncertainty bounds increased, whereas most of the field observations were still outside of the uncertainty bounds. This is not the case in this study, because Figures 8 and 9 show that the mean predictions capture reasonably well the overall spatial variability of the observed θ . This is believed to be one of the main reasons that almost all the observed θ fall within the 95% predictive uncertainty when the PTF input uncertainty is considered. In other words, the mismatch between the observed and simulated θ by Ye *et al.* [2007a] is the result of ignoring parameter estimation uncertainty, especially the PTF input uncertainty due to the cokriging of soil texture and bulk density.

4. Discussion

[26] Recalling that the PTF inputs are the bulk density and soil textures (GR, CS, FS, SI, and CL), the importance

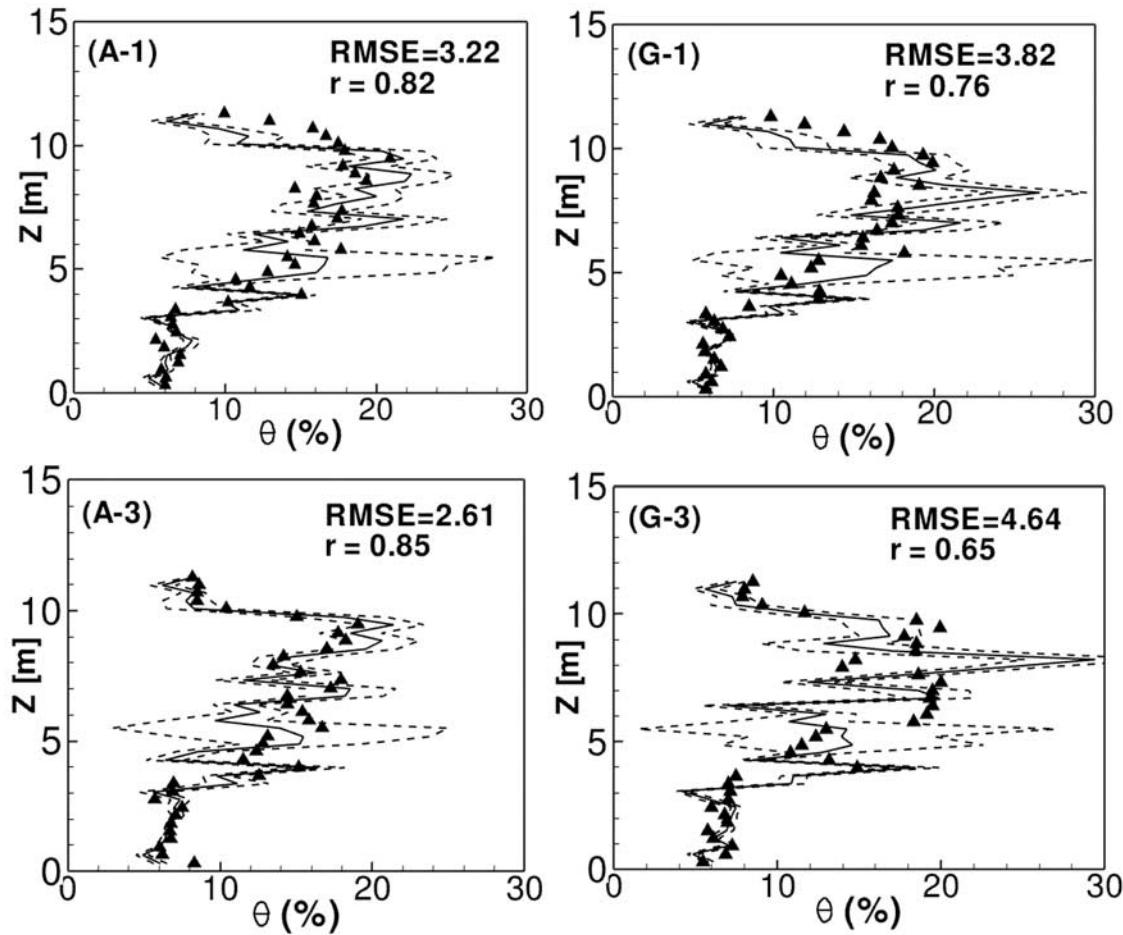


Figure 8. Comparison of the observed (triangles) and simulated moisture contents on 23 June 2000 at four boreholes. The mean and 95% prediction intervals of the simulated moisture content are shown as solid and dashed lines, respectively. Only for this set of simulations.

of the PTF input uncertainty shown above suggests that an accurate simulation of unsaturated flow largely depends on characterization of the soil textural structure. According to *Ye et al.* [2005b], the alternating layers of fine and coarse material control the moisture plume dynamics observed during the injection experiment. The alternating layering structure was characterized by *Ye et al.* [2007a] via cokriging of the initial moisture contents, which carry signature about the site heterogeneity, especially the layering structure. This is believed to be the reason that the overall spatial trend of the observed moisture contents is well simulated by *Ye et al.* [2007a]. However, because of well-known smoothing effect of the cokriging method, the cokriged textures cannot fully characterize the site heterogeneity reflected in the soil texture. Therefore, ignoring the PTF input uncertainty (i.e., cokriging variance) resulted in a mismatch between the observed and simulated moisture content [*Ye et al.*, 2007a]. After the uncertainty in the soil textures is incorporated in this study, the spatial variability of the observed moisture content is well captured by the predictive interval.

[27] This is consistent with the recent work of *Ye and Khaleel* [2008], in which the soil textural classes were directly simulated using the transition probability Markov chain (TP/MC) method [Carle and Fogg, 1996, 1997]. The heterogeneity induced by the four soil classes (coarse sand,

fine sand, loamy sand, and sandy loam) was directly characterized using a Markov chain model obtained from their transition probabilities. In the work by *Ye and Khaleel* [2008], the initial moisture contents also played an important role in the Markov chain model development. *Ye and Khaleel* [2008] did not consider uncertainty in the soil hydraulic parameters, but considered uncertainty in spatial distribution of the soil classes. Although only 50 realizations were considered by *Ye and Khaleel* [2008, Figure 10], their simulations were equally as good as the results of this study with 400 realizations (Figure 9). From a theoretical point of view, soil textural classes can characterize soil heterogeneity more directly than the soil hydraulic parameters. As a result, uncertainty due to soil heterogeneity can be reduced more efficiently by including more data for the soil textures. However, such an investigation is beyond the scope of this study.

5. Conclusions

[28] This study draws the following main conclusions.

[29] 1. When the uncertainty of PTF-estimated soil hydraulic parameters is considered, the ANN-based PTF developed by *Ye et al.* [2007a] is able to capture the spatial variability of the measured soil hydraulic parameters. This

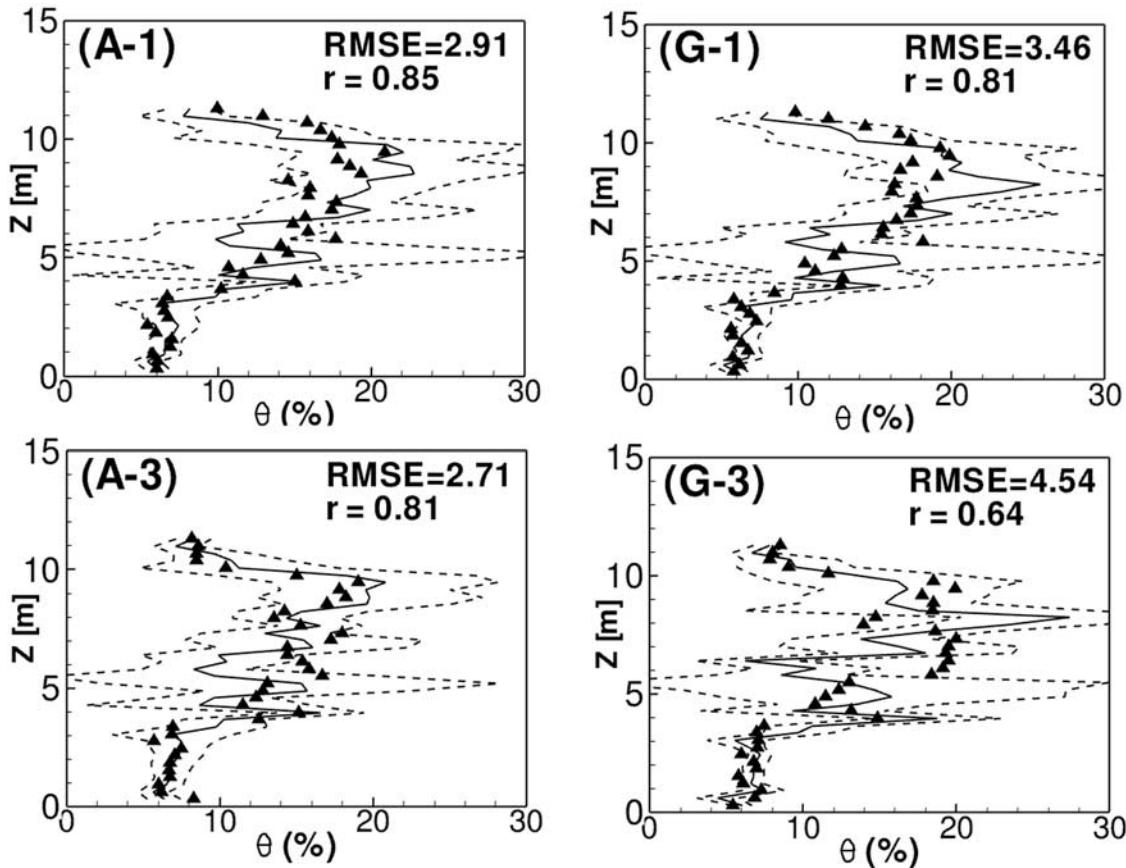


Figure 9. Comparison of the observed (triangles) and simulated moisture contents on 23 June 2000 at four boreholes. The mean and 95% prediction intervals of the simulated moisture content are shown as solid and dashed lines, respectively. Both PTF intrinsic and input uncertainty are considered.

however is impossible without considering the PTF input uncertainty. This is also the case for the spatial variability of the observed moisture content. The mismatch of the simulated and observed moisture content profiles of *Ye et al.* [2007a] is not a deficiency of the PTF itself, but because of ignoring the PTF parameter estimation uncertainty.

[30] 2. The PTF intrinsic uncertainty, caused by limited data used for training the PTF, can be assessed using the bootstrap method and by generating multiple bootstrap realizations of the soil hydraulic parameters. These realizations follow normal distribution for the $\log(K_s)$, $\log(\alpha)$, $\log(n)$, θ_s , and θ_r for the four soil classes (coarse sand, sand, loamy sand, and sandy loam) considered by *Ye and Khaleel* [2008]. The PTF intrinsic uncertainty is more or less uniformly distributed in the domain.

[31] 3. The PTF input uncertainty considered in this study is the cokriging variance, which is caused by the spatial variation of the parameters and varies in space depending on measurement locations of the primary and secondary variables.

[32] 4. For the PTF parameter estimation uncertainty, the PTF input uncertainty dominates over the PTF intrinsic uncertainty. The spatial pattern of the PTF parameter estimation uncertainty is also determined by the cokriging uncertainty.

[33] 5. The current analysis is based on 70 samples, 60 of which are obtained from three S boreholes at the SL site. Because the cokriging uncertainty of bulk density and soil

texture dominates the parameter estimation uncertainty, additional sample acquisition for PTF variables (soil texture and bulk density) would more likely lead to a reduction of uncertainty than collecting additional samples for hydraulic property measurements at the SL site. Therefore, it appears that a more complete characterization of a site for texture and layering and a judicious positioning of sample measurement locations would lead to an optimal site characterization.

[34] **Acknowledgments.** This study was supported by CH2M Hill Hanford Group Inc. and the U.S. Department of Energy Office of River Protection under contract DE-AC06-99RL14047. The second author is also supported by the DOE EPSCoR program under contract DE-FG02-06ER46265. The third author was supported in part by NSF-EAR grant 0737945. The authors are grateful for inspiring discussions with Shlomo Neuman and Jim Yeh and constructive comments from three anonymous reviewers.

References

- Abbaspour, K. C., R. Schulin, M. T. van Genuchten, and E. Schlappi (1998), An alternative to cokriging for situations with small sample sizes, *Math. Geol.*, 30(3), 259–274, doi:10.1023/A:1021724830427.
- 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.
- Chernick, M. R. (1999), *Bootstrap Methods: A Practitioner's Guide*, John Wiley, New York.
- Chirico, G. B., H. Medina, and N. Romano (2007), Uncertainty in predicting soil hydraulic properties at the hillslope scale with indirect methods, *J. Hydrol.*, 334(3–4), 405–422, doi:10.1016/j.jhydrol.2006.10.024.

- Cornelis, W. M., J. Ronsyn, M. van Meirvenne, and R. Hartmann (2001), Evaluation of pedotransfer functions for predicting the soil moisture retention curve, *Soil Sci. Soc. Am. J.*, 65(3), 638–648.
- Cressie, N. (1991), *Statistics of Spatial Data*, John Wiley, New York.
- Dai, Z., J. Samper, A. Wolfsberg, and D. Levitt (2008), Identification of relative conductivity models for water flow and solute transport in unsaturated bentonite, *Phys. Chem. Earth*, 31, S177–S185.
- Deutsch, C. V., and A. G. Journel (1998), *GSLIB-Geostatistical Software Library and User's Guide*, Oxford Univ. Press, Orlando, Fla.
- Efron, B. (1992), *The Jackknife, the Bootstrap, and Other Resampling Plans*, CBMS-NSF Reg. Conf. Ser. Appl. Math., vol. 38, SIAM, Philadelphia, Pa.
- Efron, B., and R. J. Tibshirani (1993), *An Introduction to the Bootstrap, Monogr. Stat. Appl. Probab.*, vol. 57, Chapman and Hall, New York.
- Espino, A., D. Mallants, M. Vanclooster, and J. Feyen (1996), Cautionary notes on the use of pedotransfer functions for estimating soil hydraulic properties, *Agric. Water Manage.*, 29(3), 235–253, doi:10.1016/0378-3774(95)01210-9.
- Finke, P. A., J. H. M. Wosten, and M. J. W. Jansen (1996), Effects of uncertainty in major input variables on simulated functional soil behaviour, *Hydrol. Processes*, 10(5), 661–669, doi:10.1002/(SICI)1099-1085(199605)10:5<661::AID-HYP346>3.0.CO;2-R.
- Gee, G. W., and A. L. Ward (2001), Vadose zone transport field study: Status report, *Rep. PNNL-13679*, Pac. Northwest Natl. Lab., Richland, Wash.
- Gutmann, E. D., and E. E. Small (2005), The effect of soil hydraulic properties vs. soil texture in land surface models, *Geophys. Res. Lett.*, 32, L02402, doi:10.1029/2004GL021843.
- Gutmann, E. D., and E. E. Small (2007), A comparison of land surface model soil hydraulic properties estimated by inverse modeling and pedotransfer functions, *Water Resour. Res.*, 43, W05418, doi:10.1029/2006WR005135.
- 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.
- Kern, J. S. (1995), Evaluation of soil water retention models based on basic physical properties, *Soil Sci. Soc. Am. J.*, 59(4), 1134–1141.
- Khaleel, R., and E. J. Freeman (1995), *Variability and Scaling of Hydraulic Properties for 200 Areas Soils, Hanford Site*, Westinghouse Hanford Co., Richland, Wash.
- Khaleel, R., J. F. Relyea, and J. L. Conca (1995), Evaluation of van Genuchten-Mualem relationships to estimate unsaturated hydraulic conductivity at low water contents, *Water Resour. Res.*, 31(11), 2659–2668, doi:10.1029/95WR02309.
- Kitanidis, P. K. (1997), *Introduction to Geostatistics: Applications in Hydrogeology*, Cambridge Univ. Press, New York.
- 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.
- Lark, R. M. (2003), Two robust estimators of the cross-variogram for multivariate geostatistical analysis of soil properties, *Eur. J. Soil Sci.*, 54(1), 187–202, doi:10.1046/j.1365-2389.2003.00506.x.
- McKay, M. D., R. J. Beckman, and W. J. Conover (1979), A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, 21(2), 239–245, doi:10.2307/1268522.
- Mermoud, A., and D. Xu (2006), Comparative analysis of three methods to generate soil hydraulic functions, *Soil Tillage Res.*, 87(1), 89–100, doi:10.1016/j.still.2005.02.034.
- Minasny, B., and D. Field (2005), Estimating soil hydraulic properties and their uncertainty: The use of stochastic simulation in the inverse modeling of the evaporation method, *Geoderma*, 126(3–4), 277–290, doi:10.1016/j.geoderma.2004.09.015.
- Minasny, B., A. B. McBratney, and K. Bristow (1999), Comparison of different approaches to the development of pedotransfer function for water-retention curves, *Geoderma*, 93(3–4), 225–253, doi:10.1016/S0016-7061(99)00061-0.
- Neuman, S. P. (2003), Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models, *Stochastic Environ. Res. Risk Assess.*, 17(5), 291–305, doi:10.1007/s00477-003-0151-7.
- Pachepsky, Y. A., and W. J. Rawls (2004), *Development of Pedotransfer Functions in Soil Hydrology*, Elsevier, Amsterdam.
- Pachepsky, Y. A., D. Timmlin, and G. Varallyay (1996), Artificial neural networks to estimate soil water retention from easily measurable data, *Soil Sci. Soc. Am. J.*, 60(3), 727–733.
- Pachepsky, Y. A., W. J. Rawls, and D. J. Timlin (1999), The current status of pedotransfer functions: Their accuracy, reliability, and utility in field- and regional-scale modeling in *Assessment of Non-point Source Pollution in the Vadose Zone*, *Geophys. Monogr. Ser.*, vol. 108, edited by D. L. Corwin, K. Loague, and T. R. Ellsworth, pp. 223–234, AGU, Washington, D. C.
- Pachepsky, Y. A., A. K. Guber, M. T. van Genuchten, T. J. Nicholson, R. E. Cady, J. Simunek, and M. G. Schaap (2006), Model abstraction techniques for soil-water flow and transport, *Rep. NUREG/CR-6884*, Nucl. Regul. Comm., Washington, D. C.
- Pan, F., M. Ye, J. Zhu, Y. S. Wu, B. Hu, and Z. Yu (2009), Effect of water retention parameters on predictive uncertainty of unsaturated flow and contaminant transport, *Vadose Zone J.*, 8, 158–166, doi:10.2136/vzj2008.0092.
- Rawls, W. J., T. J. Gish, and D. L. Brakensiek (1991), Estimating soil water retention from soil physical properties and characteristics, in *Advances in Soil Science*, vol. 16, edited by B. A. Stewart, pp. 213–234, Springer, New York.
- Schaap, M. G., and W. Bouten (1996), Modeling water retention curves of sandy soils using neural networks, *Water Resour. Res.*, 32(10), 3033–3040, doi:10.1029/96WR02278.
- Schaap, M. G., and F. L. Leij (1998), Database-related accuracy and uncertainty of pedotransfer functions, *Soil Sci.*, 163(10), 765–779, doi:10.1097/00010694-199810000-00001.
- Schaap, M. G., F. L. Leij, and M. T. van Genuchten (1998), Neural network analysis for hierarchical prediction of soil water retention and saturated hydraulic conductivity, *Soil Sci. Soc. Am. J.*, 62(4), 847–855.
- 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. PNNL-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*, Rockwell Hanford Oper., Richland, Wash.
- Srivastav, R. K., K. P. Sudheer, and I. Chaubey (2007), A simplified approach to quantifying predictive and parametric uncertainty in artificial neural network hydrologic models, *Water Resour. Res.*, 43, W10407, doi:10.1029/2006WR005352.
- 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(5), 275–287, doi:10.1016/0309-1708(92)90014-S.
- Swiler, L. P., and G. D. Wyss (2004), A user's guide to Sandia's Latin hypercube sampling software: LHS UBIX library/standalone version, *Tech. Rep. SAD2004-2439*, Sandia Natl. Lab., Albuquerque, N. M.
- Tamari, S., J. H. M. Wosten, and J. C. Ruiz-Suarez (1996), Testing an artificial neural network for predicting soil hydraulic conductivity, *Soil Sci. Soc. Am. J.*, 60(6), 1732–1741.
- Tietje, O., and M. Tapkenhinrichs (1993), Evaluation of pedotransfer functions, *Soil Sci. Soc. Am. J.*, 57(4), 1088–1095.
- van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, 44(5), 892–898.
- Ward, A. L., T. G. Caidwell, and G. W. Gee (2000), Vadose zone transport field study: Soil water content distribution 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(2), 268–280, doi:10.1016/j.advwatres.2005.02.013.
- Williams, R. D., L. R. Ahuja, and J. W. Naney (1992), Comparison of methods to estimate soil water characteristics from soil texture, bulk density, and limited data, *Soil Sci.*, 153(3), 172–184, doi:10.1097/00010694-199203000-00002.
- Wosten, J. H. M., Y. A. Pachepsky, and W. J. Rawls (2001), Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics, *J. Hydrol.*, 251(3–4), 123–150, doi:10.1016/S0022-1694(01)00464-4.
- 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.
- Ye, M., S. P. Neuman, and P. D. Meyer (2004), Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff, *Water Resour. Res.*, 40, W05113, doi:10.1029/2003WR002557.

- Ye, M., S. P. Neuman, P. D. Meyer, and K. Pohlmann (2005a), Sensitivity analysis and assessment of prior model probabilities in MLBMA with application to unsaturated fractured tuff, *Water Resour. Res.*, *41*, W12429, doi:10.1029/2005WR004260.
- Ye, M., R. Khaleel, and T.-C. J. Yeh (2005b), 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 (2007a), Simulation of field injection experiments in heterogeneous unsaturated media using cokriging and artificial neural network, *Water Resour. Res.*, *43*, W07413, doi:10.1029/2006WR005030.
- Ye, M., F. Pan, Y. S. Wu, B. Hu, C. Shirley, and Z. Yu (2007b), Assessment of radionuclide transport uncertainty in the unsaturated zone at Yucca Mountain, *Adv. Water Resour.*, *30*(1), 118–134, doi:10.1016/j.advwatres.2006.03.005.
- Ye, M., P. D. Meyer, and S. P. Neuman (2008), On model selection criteria in multimodel analysis, *Water Resour. Res.*, *44*, W03428, doi:10.1029/2008WR006803.
- 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, 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.

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

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

M. G. Schaap, Department of Soil Water and Environmental Sciences, University of Arizona, Tucson, AZ 85721, USA.