Nevada System of Higher Education

TECHNICAL REPORT

Geostatistical and Stochastic Study of Flow and Tracer Transport in the Unsaturated Zone at Yucca Mountain			
Report Document Identifier: **TR-07-003** **Task ORD-FY04-016**			
REVISION: 0			
DRAFT: A			
Authors:	Ming Ye		8/13/2007
	Feng Pan		8/14/2007
	Xiaolong "Bill" Hu		8/13/2007
	Jianting Zhu		8/14/200
	(print name)	(signature)	date
PI:	Jianting Zhu		8/14/2007
	(print name)	(approval signature)	date
QA Manager:	Amy Smiecinski		8-14-07
	(print name)	(approval signature)	date

1.0 TABLE OF CONTENTS

1.1 SUMMARY

Yucca Mountain has been proposed by the U.S. Department of Energy as the nation's long-term, permanent geologic repository for spent nuclear fuel or high-level radioactive waste. The potential repository would be located in Yucca Mountain's unsaturated zone (UZ), which acts as a critical natural barrier delaying arrival of radionuclides to the water table. Since radionuclide transport in groundwater can pose serious threats to human health and the environment, it is important to understand how much and how fast water and radionuclides travel through the UZ to groundwater. The UZ system consists of multiple hydrogeologic units whose hydraulic and geochemical properties exhibit systematic and random spatial variation, or heterogeneity, at multiple scales. Predictions of radionuclide transport under such complicated conditions are uncertain, and the uncertainty complicates decision making and risk analysis.

This project aims at using geostatistical and stochastic methods to assess uncertainty of unsaturated flow and radionuclide transport in the UZ at Yucca Mountain. Focus of this study is parameter uncertainty of hydraulic and transport properties of the UZ. The parametric uncertainty arises since limited parameter measurements are unable to deterministically describe spatial variability of the parameters. In this project, matrix porosity, permeability and sorption coefficient of the reactive tracer (neptunium) of the UZ are treated as random variables. Corresponding propagation of parametric uncertainty is quantitatively measured using mean, variance, 5^{th} and 95^{th} percentiles of simulated state variables (e.g., saturation, capillary pressure, percolation flux, and travel time). These statistics are evaluated using a Monte Carlo method, in which a three-dimensional flow and transport model implemented using the TOUGH2 code is executed with multiple parameter realizations of the random model parameters.

The project specifically studies uncertainty of unsaturated flow and radionuclide transport caused by multi-scale heterogeneity at the layer and local scales. Typically, in studies of Yucca Mountain, the layer scale refers to hydrogeologic layers with layer-wise average properties, and the local scale refers to the spatial variation of hydraulic properties within a layer. While most studies of radionuclide transport in the UZ have been conducted at the layer scale, the uncertainty at the local scale within a layer is also important, since it affects flow path, velocity, and travel time of radionuclide. This report first presents the uncertainty caused by layer-scale heterogeneity of matrix permeability, porosity, and sorption coefficients of reactive tracers. Homogeneous fields of the parameters are generated at each hydrogeologic layer for Monte Carlo simulations. This study is referred to as the homogeneous case. To assess the uncertainty caused by local-scale heterogeneity, the sequential Gaussian simulator (SGSIM) of GSLIB (Deutsch and Journel, 1998) is used to generate heterogeneous parameter fields within each layer, and Monte Carlo simulations are conducted. This study is referred to as the heterogeneous cases. For the homogeneous and heterogeneous cases, the mean, variance, 5^{th} and 95^{th} percentiles of simulated state variables are estimated for uncertainty assessment. In addition, the statistics of the two cases are compared to investigate effect of local-scale heterogeneity on the unsaturated flow and radionuclide transport. It is found that the local-scale heterogeneity increased the predictive uncertainty of percolation flux and cumulative mass arrival for computational blocks below the footprint of proposed repository, whereas mean predictions are hardly affected. The local-scale heterogeneity significantly affects travel times to the water table for both conservative and reactive tracers. In the early simulation

period, tracer mean travel times are delayed, whereas the influence of local-scale heterogeneity diminishes during the late simulation period.

Simulated state variables in this project are more realistic than those of using one- or two-dimensional models, due to a three-dimensional numerical model used in the project to characterize hydrological conditions at the UZ. Therefore, we expect that results of this project can be used directly to facilitate DOE site performance analysis and decision making.

1.2 ACRONYMS

YMP	Yucca Mountain Project
UZ	Unsaturated Zone
CDF	Cumulative Distribution Function
LN	Lognormal
SB	Log ratio
SU	Hyperbolic arcsine
NO	Normal Distribution
LHS	Latin Hypercube Sampling
SGSIM	Sequential Gaussian Simulation
TCw	Tiva Canyon welded
PTn	Paintbrush nonwelded
TSw	Topopah Spring welded
CHn	Calico Hills nonwelded
CFu	Crater Flat undifferentiated

1.3 ACKNOWLEDGMENT

Funding for this project was provided by the U. S. Department of Energy through a cooperative agreement with the Nevada System of Higher Education (DE-FC28-04RW12232). The authors thank Dr. Markus Berli of Desert Research Institute for the technical review of the scientific notebooks and Dr. Yu-Shu Wu of Lawrence Berkeley National Laboratory for the technical review of the report. We also thank Drs. Yu-Shu Wu and Keni Zhang for their help on TOUHG2 and other technical issues. We are grateful to the Quality Assurance Group, Amy Smiecinski, Raymond Keeler, Barbara Roosa, Morrie Roosa, Julie Bertoia, Terry Mueller, Aishia Henderson, and Jim Voigt for the QA reviews, and for their helpful advice on QA issues.

1.4 LIST OF TABLES

1.5 LIST OF FIGURES

2.0 PURPOSE

2.1 Purpose. The purpose of this report is to document the work and results of the project entitled "Geostatistical and stochastic study of radionuclide transport in the unsaturated zone at Yucca Mountain" (Task **ORD-FY04-016)** supported by the U.S. Department of Energy through Desert Research Institute. This report describes the procedures of random field generation, the results of Monte Carlo flow and radionuclide transport simulations and associated uncertainty caused by parametric uncertainty in hydraulic parameters. The objective of this project is to investigate uncertainty of unsaturated flow and radionuclide transport in the UZ of Yucca Mountain.

2.2 Scope. The scope of this report includes the purpose, introduction, methods and materials, assumptions, results of unsaturated flow and radionuclide transport simulation, uncertainty analysis of the unsaturated flow and radionuclide transport, discussions, and conclusions.

2.3 Limitations. The inputs, statistical distribution of hydraulic parameters, TOUGH2 model, and references in this report are complete and accurate. However, estimation of correlation length of porosity and permeability depends on the validity of the assumptions described in Section 6.

3.0 QUALITY ASSURANCE

The report is written in accordance with the NSHE Quality Assurance Program. No conclusions in the main body of this report are based on unqualified data. However, the results and conclusions in the Attachment A are for a non-quality affecting sub-task. Therefore, anything reported in the Attachment A is intended for possible future use and for information purpose only.

4.0 INTRODUCTION

Hydrogeologic environments consist of natural soils and rocks that exhibit multi-scale spatial variability, or heterogeneity, in hydraulic and transport parameters. Although the parameters are intrinsically deterministic (i.e., they exist and are potentially measurable at all scales), our knowledge of them is usually limited. It is not uncommon that laboratory and field measurements are too sparse to describe heterogeneity of model parameters in a field-scale simulation. Parameter uncertainty arises in such situations, and renders predictions of flow and contaminant transport uncertain. Quantification of parametric uncertainty and its propagation in hydrogeological models has been studied for decades using stochastic methods, as reviewed in several books (e.g., Gelhar, 1989; Dagan, 1989; Dagan and Neuman, 1997; Zhang, 2002; Rubin, 2003). It has become common to quantify uncertainty in groundwater flow and contaminant transport models by treating model parameters as random variables and estimating probability distributions, or statistics, of state variables of interest. The uncertainty analysis gives not only optimum predictions (i.e., mean predictions) but also their associated uncertainty (usually measured with variance or uncertainty bounds), which provides information to facilitate science-based decision-making by regulatory agencies, decision-makers, stakeholders, and informed segments of the general public.

Quantifying uncertainty of flow and contaminant transport at the field scale is of particularly importance, because decisions need to be made based on field-scale predictions. Parameters of field-scale models exhibit heterogeneity at multiple scales, from core samples to layer structures and lithofacies. A long-lasting challenge in contaminant transport modeling is to characterize multi-scale heterogeneity of the parameters and integrate the multi-scale heterogeneity into field-scale models. On the other hand, to represent open and complex hydrogeologic environments, field-scale models are complicated. This is particular true for nonlinear and dual–continuum models developed for unsaturated, fractured media. An ongoing challenge in uncertainty assessment of unsaturated flow and contaminant transport at the field scale is to evaluate propagation of parametric uncertainty through the complicated field-scale models. These two challenges motivate this study.

This study aims at developing a geostatistical method of characterizing multi-scale heterogeneity of model parameters and analyzing predictive uncertainty of unsaturated flow and tracer transport by combining the developed method with a mountain-scale model developed for Yucca Mountain, USA. Yucca Mountain has been proposed by the U.S. Department of Energy as the nation's first permanent geologic repository for spent nuclear fuel and high-level radioactive waste. The potential repository would be located in Yucca Mountain's Unsaturated Zone (UZ). Since the UZ acts as an important natural barrier in delaying arrival of radionuclides at the water table, it is important to understand how much and how fast water and radionuclides travel through the UZ to groundwater. The UZ consists of various complex hydrogeologic units, whose hydraulic properties vary both systematically and randomly at multiple scales. Yet, only limited data are available to characterize the multiple-scale heterogeneity, which results in uncertainty in model parameters and subsequently model predictions. It is essential to incorporate uncertainty assessment into the process of total system performance assessment and science-based decision-making. Regulations of U.S. Environmental Protection Agency (EPA) and Nuclear Regulatory Commission (NRC) specifically acknowledge that uncertainty in radionuclide concentration (or dose) is a key issue and call for including uncertainty in order to develop a reasonable expectation of compliance.

At waste disposal facilities such as the potential Yucca Mountain geological repository, there are two other major sources of uncertainty: uncertainty about conceptual models used to evaluate tracer or radionuclide transport and uncertainty in model scenarios capturing all applicable features, events, and processes (FEPs) at the geological repository (BSC, 2003). Recently, multi-model averaging method has been advocated to assess conceptual model uncertainty (Beven and Binly, 1992; Neuman 2003; Ye et al., 2004, 2005; Poeter and Anderson, 2005; Beven 2006; Refsgaard et al., 2006; Meyer et al., 2007), whereas study of model scenarios is mainly focused on infiltration (Wu et al., 2002, 2004a; Faybishenko 2007), the major driving force of radionuclide transport to groundwater. Meyer et al. (2007) developed a general framework of multi-models and multi-scenarios to assess the three types of uncertainties. Parametric uncertainty of each model under each modeling scenario is first assessed using stochastic methods, and conceptual model uncertainty of multi-models for a given scenario is then estimated by combining parametric uncertainty of each model using the Bayesian model averaging method (Hoeting et al., 1999). Scenario uncertainty is assessed in similar manner using the scenario averaging method (Draper et al., 1999) for a single or multiple models. It is seen that assessing parametric uncertainty is the fundamental task in the developed framework. If this framework is used to assess uncertainty for the proposed Yucca Mountain geologic repository, this study of parametric uncertainty can be used directly to assess conceptual model and/or scenario uncertainty.

Heterogeneity of hydraulic properties at the Yucca Mountain UZ has been investigated by many researchers, among whom Zhou et al. (2003) categorized the heterogeneity for site, layer, and local scales. Typically, in studies of Yucca Mountain, *site scale* refers to the UZ model domain in numerical modeling studies, *layer scale* refers specifically to hydrogeologic layers with layer-wise average properties, and *local scale* refers to the spatial variation of hydraulic properties within a layer. In the last decade, the layer-scale heterogeneity has been well characterized and incorporated into the three-dimensional (3-D), site-scale numerical modeling used for site performance assessment and license application preparation (Wu et al., 2004a, b). A total of 33 hydrogeologic layers of five major units were delineated based on degree of welding (Montazer and Wilson, 1984) and core measurements of rock properties, state variables, and hydraulic properties (Flint, 1998, 2003; Liu et al., 2003; Flint et al., 2006). The hierarchical structure of heterogeneity may be characterized using hydrofacies-based geostatistical methods (e.g., Koltermann and Gorelick, 1996 and references therein; Ritzi, 2000; Dai et al., 2005). However, extending the methods developed for sedimentary architectures to the volcanic environments at Yucca Mountain is beyond the scope of this study.

Uncertainty and sensitivity analysis of tracer or radionuclide transport in the Yucca Mountain UZ has been conducted mainly at the layer scale (Nicholes and Freshley, 1993; Illman and Hughson, 2005; Zhang et al., 2006). Local-scale heterogeneity of the model parameters within a layer is also important, since it affects flow path, velocity, and travel time of tracer or radionuclide (Bodvarsson et al., 2001; Haukwa et al., 2003; Zhou et al., 2003; Viswanathan et al., 2003; Illman and Hughson, 2005; Zhang et al., 2006). Although Zhou et al. (2003) conducted a sensitivity analysis that considers the layer- and local-scale heterogeneity, uncertainty analysis in this project with consideration of the multi-scale heterogeneity at the Yucca Mountain UZ has not been conducted. The hierarchical structure renders flow and transport properties nonstationary and highly

heterogeneous with large variance. In this case, although Monte Carlo methods can still be used for quantifying uncertainty, traditional methods of moment equations become unsuitable, since they require small variance of model parameters so that perturbation techniques used in the methods are valid. To resolve this problem, new stochastic methods have been developed, with boundaries of hydrogeologic layers treated either known (Zhang et al., 2000; Wu et al., 2003; Lu and Zhang, 2007) or unknown (Winter and Tartakovsky, 2000, 2002; Winter et al., 2003; Lu and Zhang, 2002; Hu et al., 2003). These methods however were developed for saturated media and only applied to simple synthetic cases. To our knowledge, their extension to unsaturated media and applications to real-world cases have not been reported in literature.

In this project, Monte Carlo simulation is used to investigate propagation of parametric uncertainty using a complicated, 3-D model developed by Wu et al. (2004a, b) to simulate unsaturated flow and transport at the Yucca Mountain UZ. The 3-D model, implemented using the dual-continuum modeling approach, is considered more physically meaningful than one-dimensional (1-D) and two-dimensional (2-D) models used in previous study, because it can simulate lateral flow, perched water, and capillary barriers. Since the numerical model used in this study represents real-world conditions at the Yucca Mountain UZ from the most current knowledge we have to this point, our uncertainty analysis is realistic, and the results from this study could be used directly to facilitate total system performance assessment and decision-making for the proposed geological repository. Hydraulic parameters studied in this project are matrix permeability and porosity, important to simulate water movement and contaminant transport velocity and residual time. Layer- and local-scale heterogeneity of the two parameters is characterized using statistical and geostatistical methods. Uncertainty of the unsaturated flow and radionuclide transport are first investigated separately at the two scales, and then compared to examine effect of local-scale heterogeneity on uncertainty of the unsaturated flow and tracer transport. Geochemical parameter considered in this report is sorption coefficients of a reactive tracer, neptunium (^{237}Np), for three types of rocks (devitrified, vitric, and zeolitic tuffs) in the UZ. Only layer-scale heterogeneity is exhibited for this parameter.

Other model parameters are considered as deterministic variables. Since the dual-continuum modeling approach is used for the numerical simulation, two sets of hydraulic properties for matrix and fracture media are required. While matrix permeability and porosity are treated as random variables, matrix van Genuchten α and n are treated as deterministic variables, since their distributions cannot be rigorously identified for the 3-D uncertainty analysis based on limited site measurements. For example, only two or three measurements of the matrix van Genuchten α and n are available in each hydrogeologic layer. Other matrix parameters (e.g., residual saturation) are also handled deterministically in the present study, due to their negligible spatial variability. Uncertainty of fracture properties is not assessed based on sensitivity analysis of Zhang et al. (2006), which shows that flow and transport simulations are not sensitive to fracture properties, because fracture flow dominates over the entire model domain.

In the remaining part of the report, we introduce in Section 5 the study site, the numerical model used to simulate the unsaturated flow and radionuclide transport, the statistical method of generating random parameters at the layer scale, and the geostatistical method of generating random parameters at the local scale. After listing assumptions involved in this study in Section 6, we discuss simulation

results and draw major conclusions in Section 7. At the layer scale, since each parameter realization is homogeneous within a layer, we refer to it as homogeneous case. In contrast, study at the local scale is referred to as heterogeneous case, since each parameter realization is heterogeneous within a layer. The non-quality affecting subtask "Stochastic analysis of transient flow in unsaturated heterogeneous porous media using KLME method" is discussed in attachment A as Non-Q for information purpose only.

5.0 METHODS AND MATERIALS

5.1 Study Site and Numerical Model

The study site and computer model used in this study is described briefly here, and more details can be found in Wu et al. (2004a). The study site of the UZ encompassing an area of approximate 20 km^2 at Yucca Mountain is shown in Figure 1 with the domain, grids, repository blocks and borehole locations (Wu et al., 2004a). The UZ overlies on a relatively flat water table, and thickness of the UZ is between 500 m and 700 m. Yucca Mountain is a structurally complex geologic system of Tertiary volcanic rocks and heterogeneous environment of layered, anisotropic fractured tuff. Primarily based on the degree of formation welding, the geologic formations at Yucca Mountain have been organized into five major units: Tiva Canyon welded (TCw) unit, Paintbrush nonwelded (PTn) unit, Topopah Spring welded (TSw) unit, Calico Hills nonwelded (CHn) unit, and Crater Flat undifferentiated (CFu) unit. A typical east-west cross section of subsurface geology is shown in Figure 2. The potential repository site will be located in the TSw unit of densely welded tuffs with a low porosity and below the PTn unit with a high porosity, low fractures and high storage capacity. These five major units are divided further into about 30 subunits, which are associated with the computer layers in the numerical model (Table 1). The 3-D model of the mountain-scale, unsaturated flow domain is discretized into a computational grid, which incorporates the layering heterogeneity at the site by representing each hydrogeologic subunit with several computational grid layers (at least one). The grid has an average of 45 vertical computational layers in the vertical direction and 980 columns (or gridblocks per layer) of both fracture and matrix continua, resulting in 86,400 gridblocks and 343,520 connections in a dual-permeability grid. It uses a refined mesh in the vicinity of the repository and includes every repository drift by taking account of orientations, lengths, elevations, and spacings of the drifts. This simulation domain is smaller than that of license application, since it is computationally more efficient for the uncertainty assessment.

The dual-continuum approach is used to simulate the fractured porous media in the UZ. Two sets of properties (i.e., relative permeability and capillary pressure curves), along with other intrinsic properties (e.g., permeability, porosity, density, fracture geometric parameters, and transport properties) are needed for the two media of fractured and matrix systems. Because the van Genuchten model of relative permeability and capillary pressure functions is used to describe variably saturated flow in both fracture and matrix continua, the basic rock and flow parameters used for each model layer include (a) fracture properties (frequency, spacing, porosity, permeability, van Genuchten α and n parameters, residual saturation, and fracture-matrix interface area); (b) matrix properties (porosity, permeability, van Genuchten α and n parameters, and residual saturation); (c) transport properties (grain density, diffusion, adsorption, and tortuosity coefficients); and (d) fault properties (porosity, matrix and fracture permeability, and active fracture-matrix interface area). As discussed above, only matrix permeability, porosity, and adsorption coefficient are treated as random variables in the present study, while the matrix permeability and porosity exhibit multi-scale heterogeneity.

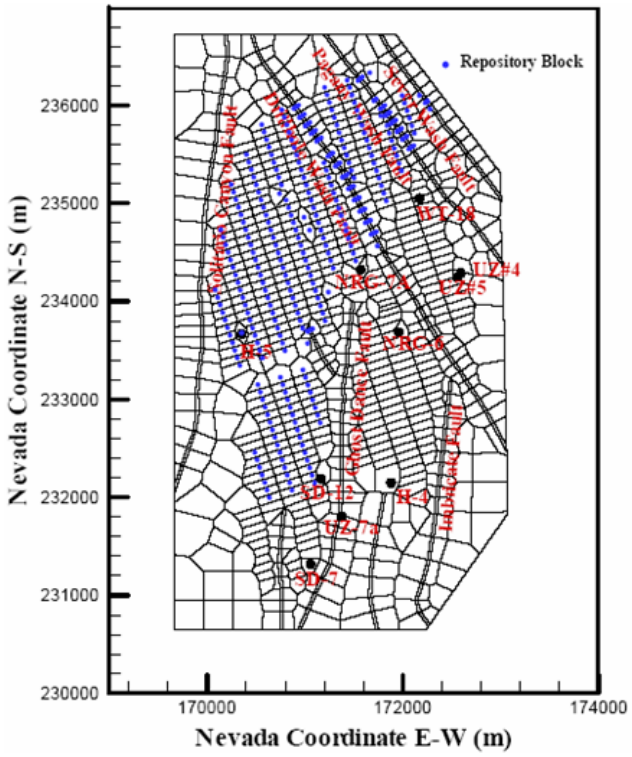


Figure 1. Plan view of the 3-D UZ numerical model grid shows the model domain, faults, proposed repository layout, and locations of several boreholes (Wu et al. 2004a)

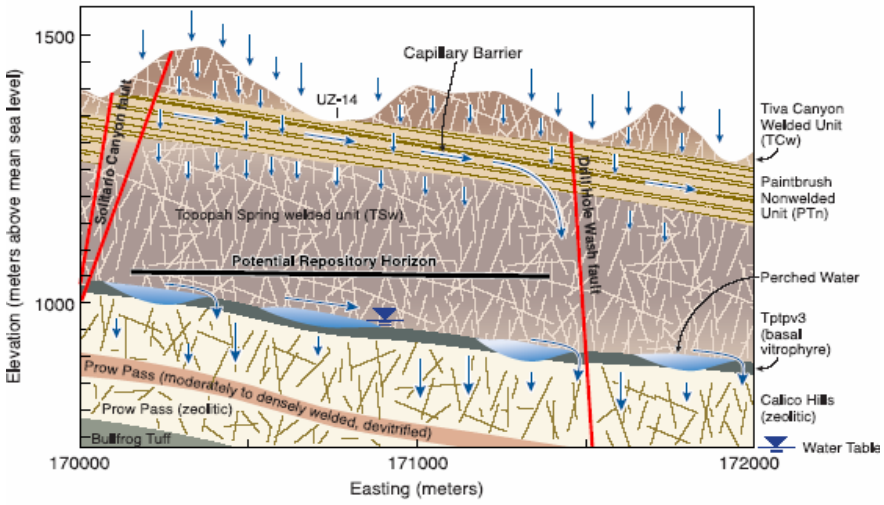


Figure 2. A typical cross section of geological profile at Yucca Mountain (Wu et al., 2004a)

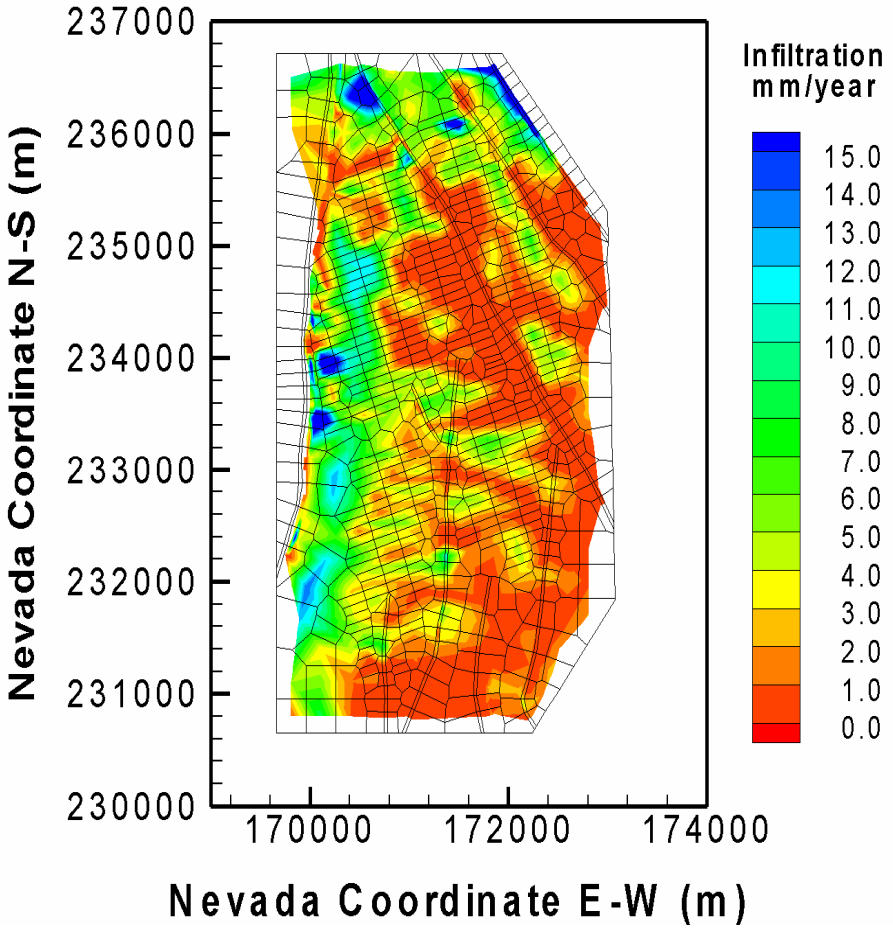
Table 1: Lithostratigraphy used in geological framework model, unsaturated zone model layer, and hydrogeological unit correlation used in this task (Wu et al., 2004a)

Major Unit	Lithostratigraphic nomenclature	Unsaturated Zone Model		Hydrogeological Unit (Flint, 1998)
		Unit/Layer	Number	
Tiva Canyon welded (TCW)	Tpcr	Tcw11	1	CCR, CUC
	Tpcp	Tcw12	2	CUL, CW
	TpcLD			
	Tpcpv3	Tcw13	3	CMW
	Tpcpv2			
Paintbrush nonwelded (PTn)	Tpcpv1	Ptn21	4	CNW
	Tpbt4	Ptn22	5	BT4
	Tpy (Yucca)			
		Ptn23	6	TPY
		Ptn24	7	BT3
	Tpbt3			
	Tpp (Pah)	Ptn25	8	TPP
	TPbt2	Ptn26	9	BT2
	Tptrv3			
	Tptrv2			
Topopah Spring welded (TSW)	Tptrv1	Tsw31	10	TC
	Tptrn			
		Tsw32	11	TR
	Tptrl, Tptf	Tsw33	12	TUL
	Tptpul, RHHtop			
	Tptpmn	Tsw34	13	TMN
	Tptpll	Tsw35	14	TLL
	Tptpln	Tsw36	15	TM2 (upper 2/3 of Tptpln)
		Tsw37	16	TM1 (lower 1/3 of Tptpln)
	Tptpv3	Tsw38	17	PV3
	Tptpv2	Tsw39 (vit, zeo)	18(zeo), 19 (vit)	PV2
Calico Hills nonwelded (CHn)	Tptpv1	Ch1 (vit, zeo)	20(zeo), 21(vit)	BT1 or BT1a (altered)
	Tpbt1			
	Tac (Calico)	Ch2 (vit, zeo)	22(vit), 26(zeo)	CHV (vitric) or CHZ (zeolitic)
		Ch3 (vit, zeo)	23 (vit), 27(zeo)	
		Ch4 (vit, zeo)	24(vit), 28(zeo)	
		Ch5 (vit, zeo)	25(vit), 29(zeo)	

Table 1: (Cont.) Lithostratigraphy used in geological framework model, unsaturated zone model layer, and hydrogeological unit correlation used in this task (Wu et al., 2004a)

Major Unit	Lithostratigraphic nomenclature	Unsaturated Zone Model		Hydrogeological Unit (Flint, 1998)
		Unit/Layer	Number	
Calico Hills nonwelded (CHn)	Tacbt (Calicobt)	Ch6 (vit, zeo)	30(zeo), 31(vit)	BT
	Tcpuv (Prowuv)	pp4	32	PP4 (zeolitic)
	Tcpuc (Prowuc)	pp3	33	PP3 (devitrified)
	Tcpmd (Prowmd)	pp2	34	PP2 (devitrified)
	Tcplc (Prowlc)			
	Tcplv (Prowlv)	pp1	35	PP1 (zeolitic)
	Tcpbt (Prowbt)			
	Tcbuv (Bullfroguv)			
Crater Flat undifferentiated (CFu)	Tcbuc (Bullfroguc)	bf3	36	BF3 (welded)
	Tcbmd (Bullfrogmd)			
	Tcblc (Bullfroglc)			
	Tcblv (Bullfroglv)	bf2	37	BF2 (nonwelded)
	Tcbbt (Bullfrogbt)			
	Tctuv (Tramuv)			
	Tctuc (Tramuc)	tr3	38	Not Available
	Tctmd (Trammd)			
	Tctlc (Tramlc)			
	Tctlv (Tramlv)	tr2	39	Not Available
	Tctbt (Trambt) and below			

The unsaturated flow module, EOS9 (solving Richards' equation), of TOUGH2 (Pruess et al., 1999) is used to simulate moisture movement in the UZ, which is approximated at a quasi-steady-state or steady-state condition. Another TOUGH2-family code, T2R3D (Wu et al., 1996), is used for modeling radionuclide transport through fractured tuffs. For the flow model, the ground surface and water table are taken as top and bottom boundaries, which are treated as Dirichlet-type conditions with specified pressure or saturation values. All lateral boundaries are treated as no-flow (closed) boundaries. A present-day, net infiltration estimate (Figure 3) is applied as a source term in the fracture gridblocks within the second grid layer from the top, since the first layer is treated as a Dirichlet boundary to represent average atmospheric conditions on the land surface. Net infiltration from precipitation is the major control on overall hydrologic and thermal-hydrologic conditions within the UZ.



Source: Infiltration data come from DTN: LB0303THERMSIM.001, file "th_pqm.dat"; see word file "Infiltration.doc" from DID: 016FP.001.

Figure 3. Plan view of present-day net infiltration distributed over the 3-D unsaturated zone flow model grid.

5.2 Generation of Layer-scale Random Fields of Matrix Permeability, Porosity, and Sorption Coefficient

Layer-scale heterogeneity is characterized in this section by first identifying probability distributions of the three random parameters and then generating homogeneous random fields of the parameters for each layer using the Latin Hypercube Sampling (LHS) method. Local-scale heterogeneity is characterized in Section 5.3 using a geostatistical method. The generated parameters are used as inputs of a Monte Carlo simulation described in Section 5.4.

5.2.1 Identification of Parameter Distributions

Several methods have been used to identify parameter probability distributions based on measurements and generate random parameter fields. For example, Carsel and Parrish (1988) employed three transformations of Johnson system to soil hydraulic parameters and used the Kolmogorov-Smirnov (K-S) test to determine which transform yields the best normality fitting. Random fields of the parameters with correlations were generated based on selected transformations, fitted distributions and Pearson correlations for the transformed variables (Carsel and Parrish, 1988). Mallants et al. (1996) applied seven transformations including three Johnson transformations and four classical re-expressions to transform the measured data of van Genuchten α, *n* and other parameters and the normality of the transformed data was judged by the Shapiro-Wilk test.

In general, parameter measurements are seldom adequate to describe the corresponding parameter distributions without proper transform and rigorous statistical test. In this task, we first apply three Johnson transformations and four classic re-expressions (Carsel and Parrish, 1988; Mallants et al., 1996) to the measured data of matrix porosity, saturated hydraulic conductivity and sorption coefficient, and then apply the Lilliefors Test to select the transform that yields the best normality fit. The Spearman rank correlations are calculated for the transformed measurements in each hydrogeologic layer. Subsequently, the random fields of the hydraulic parameters are generated by LHS method based on the fitted distributions and Spearman rank correlations for each layer. Three distribution types of transformations (lognormal, log ratio and hyperbolic arcsine) from Johnson system (Johnson and Lotz, 1970) and four classical re-expressions (1/X, $X^{1/2}$, $X^{1/3}$, X^2) (Mallants et al., 1996) are selected to transform the measured data. The lognormal (LN), log ratio (SB), and hyperbolic arcsine (SU) transforms are given as (Johnson and Lotz, 1970; Carsel and Parrish, 1988):

$$\text{LN:} \quad Y = \ln(X) \qquad \text{Eq. 1}$$

$$\text{SB:} \quad Y = \ln(U) = \ln\left(\frac{X - A}{B - X}\right) \qquad \text{Eq. 2}$$

$$\text{SU:} \quad Y = \sinh^{-1}(U) = \ln(U + \sqrt{1 + U^2}) \qquad \text{Eq. 3}$$

where X is the untransformed variable value with limits of variation from A to B (A < X <B) and U = (X-A)/(B-A).

Lilliefors Test, a variant of Kolmogorov-Smirnov (K-S) test, is used to test the goodness-of-fit of the transformed data to normal distributions. Lilliefors test is to test whether a certain set of data follow the normal distribution with unspecified parameters estimated from the observations. It differs from K-S test in that it does not require the hypothesized distribution with a completely specified cumulative distribution function (CDF) (Bowen and Bennett, 1988). The steps of Lilliefors test are as follows:

(1) Standardize the sample values with sample mean and standard deviation estimated from the samples:

$$z_i = \frac{x_i - \bar{x}}{s} \quad (i = 1,2,...,N) \qquad \text{Eq. 4}$$

where z_i and x_i are standardized and original sample values, respectively, $\bar{x}$ and s are the sample

mean and standard deviation, respectively, *N* is the sample size. Generally, the sample size is at least 4 for Lilliefors test.

(2) Calculate the empirical CDF G(z) of the standardized z_i and the standard normal CDF $F^*(z)$;
(3) Estimate the absolute maximum difference between the empirical CDF and the standard normal CDF for each z_i;

$$T = \max\left|F^*(z_i) - G(z_i)\right| \quad i = 1,2,...n \qquad \text{Eq. 5}$$

(4) Select Lilliefors test statistic T^* corresponding to a level of significance α from Lilliefors Test Statistical Table (Bowen and Bennett, 1988) and judge whether the hypothesis of normality is accepted or not. The test is rejected at α level of significance if T exceeds T^*.

One of the following eight transformations - Normal distribution (NO), LN, SB, SU, 1/X, $X^{1/2}$, $X^{1/3}$, and X^2 – is selected for random field generation if it gives the minimum T value of Lilliefors test for each layer. To get a better fitted distribution, outlying values are not used in estimating the mean and variance in few layers. Nevertheless, they are used in the Lilliefors test calculations to ensure objectivity (Carsel and Parrish, 1988). As discussed below, input parameters of the random field generator LHS are the values of the distribution at 1th and 99th percentiles. If the best fitted distributions of the parameters make the bound be non-physical meaning values in few layers, the second best fitted distributions of parameters are selected. For example, the best transform for matrix saturated hydraulic conductivity in BT4 layer was $X^{1/3}$ and its mean and variance are 0.0125 and 0.000069. The corresponding 1th and 99th percentiles are -0.00682 and 0.0318 and the negative 1th percentile of the saturated hydraulic conductivity has no physical meaning. Therefore, the best transform cannot be selected in this case and the second best one is selected to generate random field.

By applying Johnson transformations, classical re-expressions and Lilliefors Test, the best fits corresponding to mean and variance of the distribution of transformed values are determined for each layer. The procedures are as follows:
(1) Determine the limits of variation of matrix porosity based on the minimum and maximum values of measured data for SB and SU transformations;
(2) Transform the measured data according to seven types of transformations;
(3) Implement Lillifors Test to select the best transformation and its corresponding probability characteristics for each layer.

5.2.2 Latin Hypercube Sampling and Rank Correlation

LHS is used to generate random fields based on the distributions and their corresponding mean and variance determined above. LHS is one of sampling methods and can be used to address the need for uncertainty assessment (Swiler and Wyss, 2004; Helton and Davis, 2000; Mckay et al., 1979). LHS ensures that the generated random samples span the full coverage of a random variable even when the sample size is relatively small. This overcomes the disadvantage of the random sampling method that it possibly overemphasizes or omits the samples in some parts when the sample size is not large enough. This property of LHS reduces the computational cost of Monte Carlo simulations, since smaller number of random realizations is needed to represent parameter uncertainty.

The sampling procedure of LHS for multiple uncorrelated random variables is as follows (Helton and Davis, 2000; Swiler and Wyss, 2004):

(1) Divide the CDF of a random variable with equal probability into *N* intervals and obtain the corresponding range of CDF for each interval;
(2) Generate one random value from a uniform distribution on each interval of the CDF and identify the corresponding value of the random variable from the CDF;
(3) Pair the obtained *N* values for the first variable with the *N* values of the second variable randomly;
(4) Combine these *N* pairs in a random manner with the *N* values of the third variable to form n triplets and continue pairing until the last variable is combined with others.

If the random variables are correlated, it is necessary to incorporate the correlations into the samples because the random paring cannot represent the correlations. Iman and Conver (1982) proposed a method for restricting the pairing of variables based on a desired rank correlation matrix to generate correlated random samples. The properties of this technique are (Iman and Conver, 1982): (1) Distribution free; (2) Simple; and (3) generated original values are retained and only the pairing is affected by the desired rank correlations.

Correlation is measured in this study using the Spearman rank correlation Rx_jx_k between two variables x_j and x_k (Helton and Davis, 2003):

$$Rx_jx_k = \frac{\sum_{i=1}^{N}\left[R(x_{ij}) - \overline{R}(x_j)\right]\cdot\left[R(x_{ik}) - \overline{R}(x_k)\right]}{\left\{\sum_{i=1}^{N}\left[R(x_{ij}) - \overline{R}(x_j)\right]^2\right\}^{1/2}\left\{\sum_{i=1}^{N}\left[R(x_{ik}) - \overline{R}(x_k)\right]^2\right\}^{1/2}} \qquad \text{Eq. 6}$$

where $R(x_{ij})$ and $R(x_{ik})$ are the ranking index of x_{ij} and x_{ik} in ascending order, respectively; *N* is the sample size; and $\overline{R}(x_j) = \overline{R}(x_k) = (N+1)/2$. The Eq. 6 can also be written as:

$$r = 1 - 6\sum_{i=1}^{N}\frac{d_i^2}{N(N^2-1)} \qquad \text{Eq. 7}$$

where *r* is Spearman rank correlation coefficient and *d* is the difference in rank index of the corresponding variables. In this study, matrix porosity and saturated hydraulic conductivity are correlated and their Spearman rank correlation can be calculated in the following steps:

(1) Collect the available data for saturated hydraulic conductivity and corresponding porosity from Technical Data Management System (TDMS);
(2) Rank two variables in ascending orders and obtain the rank index of saturated hydraulic conductivity and corresponding porosity;
(3) Calculate the difference between the rank index of two variables;
(4) Calculate the Spearman rank correlation according to the above equation.

One of advantages of the rank correlation used in this study is that it can be quite meaningful with non-normal distributions of the input data (Iman and Conver, 1982; Helton and Davis, 2003). This can overcome the fact that the correlation coefficients may lose some meaning when the data are not normally distributed. The rank correlation can also capture the type of subjective information of correlation for uncertainty assessment (Helton and Davis, 2003). The subjective information means that the large or small value for one variable should correspond with large or small values for another variable.

By implementing LHS with Spearman rank correlation based on the estimations of the best transformations of matrix porosity, saturated hydraulic conductivity and sorption coefficient and their distributional characteristics, the parameter random fields of transformed data following normal distributions can be generated for each model layer by LHS V2.51 code. The numerical code LHS V2.51 requires 1th and 99th quintiles of a normal distribution, which can be calculated according to the mean (μ) and the standard deviation (σ) of the normal distribution. It is written as (Swiler and Wyss, 2004):

$$V_{0.01} = \mu - 2.326\sigma \; ; \; V_{0.99} = \mu + 2.326\sigma \qquad \text{Eq. 8}$$

where μ and σ of matrix porosity, saturated hydraulic conductivity, and sorption coefficient.

Since the generated random numbers are subject to the transforms, they need to be transformed back to the original scale for Monte Carlo simulation. The equations of inverse transformation for Johnson transformations are as follows (Carsel and Parrish, 1988):

$$\text{LN: } X = \exp(Y) \qquad \text{Eq. 9}$$

$$\text{SB: } X = [B\exp(Y) + A]/[1 + \exp(Y)] \qquad \text{Eq. 10}$$

$$\text{SU: } X = A + (B - A)[\exp(Y) - \exp(-Y)]/2 \qquad \text{Eq. 11}$$

where Y is the transformed value following a normal distribution with the estimated mean and variance generated by LHS. The equations of inverse transformation for another four classical re-expression (1/X, $X^{1/2}$, $X^{1/3}$ and X^2) transformations are as follows respectively:

$$X = \frac{1}{Y}; \; X = Y^2; \quad X = Y^3; \quad X = Y^{1/2} \qquad \text{Eq. 12}$$

The Spearman rank correlation between matrix porosity and saturated hydraulic conductivity is used as the input data of LHS V2.51 code to generate the random fields with correlation between matrix porosity and saturated hydraulic conductivity.

5.3 Generation of Local-scale Random Fields of Matrix Permeability and Porosity

The LHS method above generates multiple realizations of homogeneous parameters for each hydrogeologic layer. To consider heterogeneity of matrix permeability and porosity within each

layer, we use geostatistical method to generate heterogeneous parameter fields within each layer. Due to intensive characterization of the UZ, it is considered that boundaries of hydrogeologic layers are delineated with certainty. Nonstationarity of the matrix permeability and porosity is addressed by using applying the geostatistical method in each hydrogeologic layer. This differs from Istok et al. (1994), who generated trends of the parameters. Nonstationary behavior of state variables (e.g., matrix liquid saturation and water potential) is simulated using a Monte Carlo method, since it is theoretically straightforward and easy to implement.

5.3.1 Data of Matrix Permeability and Porosity

There are two types of data of the matrix permeability and porosity: their core measurements at the local scale and calibrated values at the layer scale. Total of 5,320 rock core samples from 33 boreholes were collected (Flint, 1998, 2003; Liu et al., 2003), yielding 546 measurements of matrix saturated hydraulic conductivity and 5,257 measurements of matrix porosity. Variogram analysis based on the measurements is conducted to characterize local-scale heterogeneity within each hydrogeologic layer. The Sequential Gaussian Simulator (SGSIM) of GSLIB (Deutsch and Journel, 1998) is then used to generate conditional, heterogeneous realizations of the parameters. To satisfy the requirement of SGSIM that conditional data need to be Gaussian (many studies simply assume that conditioning data are Gaussian), we adopt the transform method discussed in Section 5.2.1. At each layer, the measurements are transformed to follow normal distributions by using the best transform selected from the three Johnson transformations (Carsel and Parrish, 1988; Johnson and Kotz, 1970) and four classical re-expressions (Mallants et al., 1996). Random fields with local-scale heterogeneity is first generated with the transformed data, and then back-transformed to their real values.

The other type of parameter data are the layer-scale values of permeability obtained from calibrating the 3-D model (Wu et al., 2004a). The 3-D model calibration is based on the 1-D model calibration (BSC, 2004), and matrix permeability of layers BT3, BT2, CHV, and PP3 is increased one or two orders of magnitude. Such adjustment is not uncommon in inverse modeling, owing to scale disparity and model insufficiency. Since the calibrated permeability in these layers represents the optimum estimate of layer-scale UZ heterogeneity, the calibrated permeability values for these layers need to be retained in the random field generation discussed below.

5.3.2 Generation of Heterogeneous Parameter Fields

The first step of using SGSIM to generate heterogeneous parameter field is to determine parameter correlation length. In this study, since porosity measurements are abundant and widely spread in shallow boreholes, horizontal and vertical correlation lengths of porosity in each hydrogeologic layer of units TCw, PTn, and TSw are directly estimated from the measurements using variogram analysis. In deep units CHn and CFu, while the vertical variogram of porosity in each hydrogeologic layer can be calculated, the horizontal one in each layer is unavailable, due to the lack of measurements there. We note that, in units TCw, PTn, and TSw, the horizontal correlation length in each layer is similar to that of the unit to which the layer belongs. Based on this, we assume that horizontal correlation lengths of layers within the CHn unit are constant and the same as that of the CHn unit. The horizontal variogram of the CHn unit can be calculated from measurements, because

there are more boreholes in the unit than in each layer of the unit. Since only one borehole was drilled in the CFu unit (below the CHn unit), the horizontal correlation lengths of porosity in its two layers are assumed the same as that of the CHn. Measurements of permeability are only sufficient to estimate vertical correlation length in 14 layers. For the 14 layers, there appears a tendency that the permeability and porosity have similar vertical correlation length. The similarity may be attributed to correlation between permeability and porosity shown in Flint (2003) and the fact that the measurements of permeability and porosity were taken from the same boreholes. It appears reasonable to assume that vertical and horizontal correlation lengths of permeability are the same as those of porosity.

After obtaining the correlation lengths, SGSIM is used to generate 200 realizations of heterogeneous fields of matrix permeability and porosity, using the transformed parameters as conditioning data. Since SGSIM produces random fields on regular grids, the generation is conducted on a regular 3-D grid, which is designed specific to each hydrogeologic layer to cover the layer. The generated random fields are then interpolated to the 3-D irregular numerical grid using the nearest neighborhood method. Sample variograms of the interpolated fields are calculated for each layer to ensure that spatial correlation and variance are not affected by the interpolation. To honor the layer-scale values of permeability obtained from the 3-D model calibration, we first calculate for each numerical block sample mean (over the realizations) of permeability and then average them over each layer. The calibrated values are close to the resulting layer-average values and are within the range of minimum and maximum of generated values for most of the model layers, except for layers BT3, BT2, CHV, and PP3, where the layer-averaged permeability was increased one ore two orders of magnitude during model calibration. Since the calibrated permeability, as described earlier, needs be maintained in these layers, we adjust permeability in these layers to ensure that mean permeability of each realization is equal to the calibrated value. In this case, the local-scale core measurements are no longer conditioned in the four layers, despite that local-scale spatial correlation is still honored.

Our method of characterizing the multi-scale heterogeneity differs from that of Zhou et al. (2003), in which unconditional local-scale heterogeneous parameter fields were first generated using the Sequential Indicator Simulator (SISIM) of GSLIB and then imposed on layer-scale homogeneous parameter fields (the layer-scale parameter values were obtained from 2-D model calibration). One of the differences is that unconditional realizations at the local scale were generated in Zhou et al. (2003), while conditional parameter realizations are generated in this study by using the transformed core measurements as conditional data. As discussed later, the conditioning can significantly affect uncertainty of tracer or radionuclide transport. Another difference is that a 2-D model was used in Zhou et al. (2003), while a 3-D model is used in this study. As discussed before, a 3-D model is superior to a 2-D model in terms of simulating lateral flow, perched water, and capillary barriers. In addition, it is worth mentioning that Zhou et al. (2003) is focused on sensitivity analysis, while this paper on uncertainty assessment.

5.4 Monte Carlo Simulations

In this task, Monte Carlo simulation is used as a stochastic method to quantify uncertainty of the unsaturated flow and radionuclide transport fields in the UZ of Yucca Mountain. The general

procedure of Monte Carlo simulation is as follows:

(1) Generate numerous equally likely random fields for model parameter according to the parameter probabilistic distributions;
(2) Conduct numerical simulation to estimate the quantities of interest for each parameter random field;
(3) Calculate the statistics (e.g., mean and variance) of the quantities of interest to yield the optimum prediction and associated predictive uncertainty.

To complete the first step, 200 random realizations of matrix porosity, permeability and sorption coefficient are generated using the LHS method for homogeneous case or SGSIM method for heterogeneous case, as described in Sections 5.2 and 5.3, respectively. These random fields are then used to generate input files of the TOUGH2 code and unsaturated flow and radionuclide transport are simulated for each parameter realization, i.e., each TOUGH2 input file. A recently developed method by Ballio and Guadagnini (2004) is employed to examine the convergence of Monte Carlo Simulation. The means, variances and percentiles of simulated flow variables (e.g., saturation, capillary pressure, and percolation flux) over all realizations are estimated. The means are our optimum predictions and the variances measure the associated predictive uncertainty, which is further quantified by the 5th and 95th percentiles of the 200 realization of the simulated quantities.

Because there are no standard convergence criteria in Monte Carlo simulations, various methods for convergence analysis have been developed to evaluate the convergence (Bellin et al., 1992; Burr et al., 1994; Hassan et al., 1998; Ballio and Guadagnini, 2004). Ballio and Guadagnini (2004) proposed a new method for convergence analysis of Monte Carlo Simulation by estimating the ensemble mean and variance of random variables within a given confidence intervals. Therefore, the newly developed method for convergence analysis was implemented in this task. The properties of the sample mean and variance of random variable $\Re$ are as follows (Ballio and Guadagnini, 2004):

$$\left\langle \overline{\Re_n} \right\rangle = \mu; \quad \text{var}\left[\overline{\Re_n}\right] = \frac{1}{n}\sigma^2; \qquad \text{Eq. 13}$$

$$\left\langle S_n^2 \right\rangle = \sigma^2; \quad \text{var}\left[S_n^2\right] = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right) \qquad \text{Eq. 14}$$

where $\overline{\Re}$ is the sample mean of $\Re$; μ is the ensemble mean of $\Re$; S_n^2 is the sample variance; σ^2 is the ensemble variance; and μ_4 is ensemble kurtosis.

If $\Re$ is normal, then

$$\frac{\overline{R}_n - \mu}{\sigma/\sqrt{n}} \sim N_{0,1}(\ ); \quad \text{and} \quad \frac{\overline{R}_n - \mu}{S_n/\sqrt{n}} \sim t_{n-1}(\ ) \quad \forall n \qquad \text{Eq. 15}$$

$$\text{var}[S_n^2] = \frac{2}{n-1}\sigma^4 \qquad \text{Eq. 16}$$

$$\sum_{i=1,n}\frac{(\Re_i-\mu)^2}{\sigma^2}\sim\chi_n^2(\ );\quad \sum_{i=1,n}\frac{(\Re_i-\overline{\Re}_n)^2}{\sigma^2}=(n-1)\frac{S_n^2}{\sigma^2}\sim\chi_{n-1}^2(\ ) \qquad \text{Eq. 17}$$

where $N_{0,1}(\ )$ is the standard normal distribution; $t_{n-1}(\ )$ is the student distribution with $(n-1)$ degrees of freedom; and χ_{n-1}^2 is the chi-square distribution with $(n-1)$ degrees of freedom.

Therefore, according to the above the equations, the estimation of the uncertainty by the evaluation of confidence intervals are (Ballio and Guadagnini, 2004):

$$\Pr\left[\overline{\Re_n}-t_{n-1}(1-\frac{\alpha}{2})\frac{S_n}{\sqrt{n}}\le\mu\le\overline{\Re_n}+t_{n-1}(1-\frac{\alpha}{2})\frac{S_n}{\sqrt{n}}\right]=1-\alpha \qquad \text{Eq. 18}$$

$$\Pr\left[\frac{n-1}{\chi_{n-1}^2(1-\alpha/2)}S_n^2\le\sigma^2\le\frac{n-1}{\chi_{n-1}^2(\alpha/2)}S_n^2\right]=1-\alpha \qquad \text{Eq. 19}$$

where $1-\alpha$ is the probability that the value of μ lies within the confidence interval around sample mean $\overline{\Re}_n$.

According to Eq. 15 and 16, the confidence intervals (upper bound and lower bound) of the ensemble mean and variance can be obtained for a given number of realizations. Thus, the number of realizations that assures convergence can be determined.

6.0 ASSUMPTIONS

Assumptions involved in the simulation are listed below. While some assumptions (e.g., ergodicity) are related to stochastic theories, some are practical and related to selecting random model parameters and characterizing heterogeneity of the parameters based on limited number of parameter measurements.

1. Ergodicity assumption is used to infer probability distribution of the random parameters based on their measurements. The ergodicity assumption is commonly used in stochastic subsurface hydrology.

2. Fracture parameters are treated as deterministic variables, based on sensitivity analysis of Zhang et al. (2006), which shows that effect of variability of fracture parameter on simulated unsaturated flow and transport is insignificant.

3. Except matrix permeability, porosity, and sorption coefficient, other matrix parameters are treated as deterministic variables for two reasons. One is that measurements of some parameters (e.g., water retention parameters) are too sparse to infer their meaningful distributions. The other is that variation of other parameters (e.g., residual water content) is small and can be neglected.

4. For the layers in CHn unit, we assume that their horizontal correlation lengths of matrix porosity are the same as that of the unit. This assumption is based on the observation of the horizontal correlation length for layers in units TCw, PTn, and TSw.

5. For the layers in CFu unit, we assume that their horizontal correlation lengths of matrix porosity are the same as that of the unit CHn, since there is only one borehole in the CFu unit.

6. When estimating correlation length of matrix permeability, we assume that its vertical and horizontal correlation lengths are the same as those of matrix porosity. This assumption is based on the observation of 14 layers in which the vertical correlation length can be estimated for the matrix permeability and porosity.

7.0 RESULT DISCUSSION/CONCLUSIONS

In this section, we first present the generated random parameter fields of the layer- and local-scale heterogeneity. Subsequently, we present results of uncertainty analysis for the homogeneous and heterogeneous cases. In the former case, layer-scale parameter heterogeneity is considered, while the local-scale heterogeneity is considered in the latter case. Uncertainty in the two cases is also compared to investigate effect of local-scale heterogeneity on simulated unsaturated flow and radionuclide transport. For clarity of discussion, uncertainty of flow simulation of the two cases is discussed before that of transport simulation, although uncertainty analysis of flow and transport is conducted together for each case.

7.1 Generated Layer-scale Realizations of Matrix Porosity, Permeability, and Sorption Coefficient

We first discuss in Sections 7.1.1 – 7.1.4 the identified probability distributions of the three random parameters and then present the generated layer-scale parameter realizations in Section 7.1.5. Note that we first identify the distribution and generate random fields for saturated hydraulic conductivity, and then transform the saturated conductivity into permeability.

7.1.1 Probability Distribution of Matrix Porosity

Measurements of matrix porosity for each layer in the UZ at Yucca Mountain are collected from Yucca Mountain database (DTN: LB0207REVUZPRP.002) and their descriptive statistics (mean, standard variance, minimum, and maximum) is listed in Table 2. Spatial heterogeneity in matrix porosity is observed in each layer because of the large range of the measurements. For example, the matrix porosity varies from 0.228 to 0.633 in CNW layer and the range is from 0.137 to 0.578 in BT3 layer. Therefore, it is unreasonable to treat matrix porosity as deterministic values. The histograms of measured matrix porosity for each layer are shown in Figure 4, which also plots the histograms of the generated random fields for matrix porosity. Visually, matrix porosity follows a normal distribution in certain layers (e.g., CNW, TR, TUL and PP4) and lognormal distribution in some others (e.g., BT2, PV3, and BF3). However, the visual examination cannot quantify the matrix porosity distribution, which, instead, is determined by applying the transforms and Lilliefors Test discussed in Section 5.2.

The three Johnson transformations (Carsel and Parrish, 1988) and four classic re-expressions (Mallants et al., 1996) are applied to the matrix porosity measurements and the transform giving the best normality fit is selected according to the Lilliefors test. Table 3 lists the limits of variation (A and B) needed for the SB and SU transforms, selected best distribution based on the Lilliefors test, mean and variance of the transformed data, maximum absolute distribution difference (T), and Lilliefors test criteria (T*) corresponding to three significance levels (α = 0.1, 0.05, and 0.01) for matrix porosity in each hydrogeologic layer. The empirical and fitted CDF of the transformed matrix porosity measurements in each model layer are shown in Figure 5. While 28 distributions are accepted at various significance levels, no best distribution can be accepted for the four layers of CUL&CW, TMN, TM2&TM1, and CHZ. This may be attributed to the large sample sizes (ranging

from 277 to 694) in these layers, since the Lilliefors criteria, T^*, are inversely proportional to the square root of the sample size. The transform with the smallest T, however, is accepted in the study, because of the small difference between the empirical and theoretical CDFs, $G(Z)$ and $F^*(Z)$, shown in Figure 5. To yield a better distribution fit, the outlying values in the layers of CCR&CUC, BT2, CHV and BF3 are not used to calculate the mean and variance but are still included in the calculation of the maximum absolute distribution difference (T). For the layers of PV2a and BT1v, the selected distributions are not the best ones but the second best ones because the best ones make the generated random fields within an unreasonable range without physical meaning according to Eq.8. For example, the best fit of the transformations in BT1v layer is X^2 and the corresponding mean and variance are 0.112 and 0.00272, respectively. The 1^{th} and 99^{th} percentiles of the transformation are -0.00931 and 0.233 and are used as the minimum and maximum of random field generation. Because the values of transformation X^2 of matrix porosity cannot be negative, the best fit cannot be selected in this case and the second one is selected to generate random field.

Table 2: Descriptive statistics for matrix porosity and saturated hydraulic conductivity

HGU	Porosity (Φ)					Saturated Hydraulic Conductivity (K_S, m/s)				
	Mean	SD	Min	Maxi	N	Mean	SD	Min	Max	N
CCR&CUC	0.241	0.073	0.038	0.431	124	5.80E-08	6.53E-08	2.03E-08	1.33E-07	3
CUL&CW	0.088	0.032	0.032	0.213	694	7.68E-10	3.02E-09	2.49E-13	1.25E-08	17
CMW	0.200	0.055	0.100	0.452	96	1.89E-08	4.21E-08	3.34E-12	9.41E-08	5
CNW	0.387	0.069	0.228	0.633	104	2.90E-07	3.38E-07	5.12E-12	8.79E-07	10
BT4	0.428	0.100	0.134	0.669	58	4.56E-06	7.59E-06	1.80E-10	2.54E-05	11
TPY	0.233	0.057	0.073	0.309	39	1.38E-08	1.52E-08	3.00E-09	2.45E-08	2
BT3	0.413	0.082	0.137	0.578	73	1.77E-06	2.03E-06	1.90E-09	7.30E-06	11
TPP	0.498	0.041	0.388	0.623	159	1.17E-06	5.76E-07	9.00E-08	1.74E-06	11
BT2	0.490	0.095	0.104	0.614	176	7.10E-06	6.87E-06	1.24E-09	2.06E-05	21
TC	0.054	0.036	0.012	0.273	75	3.21E-08	6.72E-08	1.70E-11	1.68E-07	6
TR	0.157	0.030	0.062	0.267	449	2.03E-07	1.37E-06	1.70E-11	9.37E-06	47
TUL	0.155	0.030	0.076	0.250	438	3.94E-08	2.33E-07	4.20E-13	1.42E-06	37
TMN	0.111	0.020	0.055	0.192	277	4.18E-11	1.72E-10	4.76E-13	1.23E-09	74
TLL	0.131	0.031	0.088	0.263	502	4.11E-09	1.31E-08	1.39E-12	7.65E-08	52
TM2&TM1	0.103	0.025	0.053	0.341	300	4.28E-07	2.00E-06	5.33E-13	9.39E-06	22
PV3	0.043	0.040	0.011	0.340	125	1.66E-10	5.45E-10	8.63E-14	2.25E-09	17
PV2a	0.275	0.096	0.110	0.415	13	b	b	b	b	b
PV2v	0.243	0.122	0.048	0.470	49	3.23E-06	3.69E-06	5.03E-11	1.20E-05	16
BT1a	0.285	0.051	0.158	0.400	46	1.90E-08	3.21E-08	1.83E-13	8.70E-08	10
BT1v	0.324	0.085	0.031	0.500	80	3.76E-06	5.77E-06	1.04E-10	2.20E-05	35
CHV	0.341	0.048	0.038	0.490	130	1.48E-05	1.89E-05	1.68E-12	7.20E-05	47
CHZ	0.322	0.048	0.099	0.433	520	1.19E-09	9.62E-09	3.88E-13	9.54E-08	99
BTa	0.271	0.046	0.181	0.418	73	4.05E-11	6.96E-11	2.08E-13	2.10E-10	9
BTv	a	a	a	a	a	b	b	b	b	b
PP4	0.327	0.050	0.216	0.440	56	4.62E-08	1.08E-07	8.44E-13	3.08E-07	8
PP3	0.318	0.032	0.246	0.395	168	6.91E-08	6.72E-08	4.20E-12	3.65E-07	51
PP2	0.221	0.058	0.099	0.333	127	1.56E-09	3.01E-09	3.75E-12	1.15E-08	35
PP1	0.297	0.043	0.164	0.426	280	9.63E-08	3.88E-07	1.70E-12	1.94E-06	28
BF3	0.142	0.075	0.059	0.369	105	1.31E-08	2.01E-08	6.90E-11	5.58E-08	7
BF2	0.234	0.049	0.160	0.329	40	b	b	b	b	b

Source: Measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Statistic of measurements come from DID: 016FP.001, see word file "MeasurementStatistics.doc".

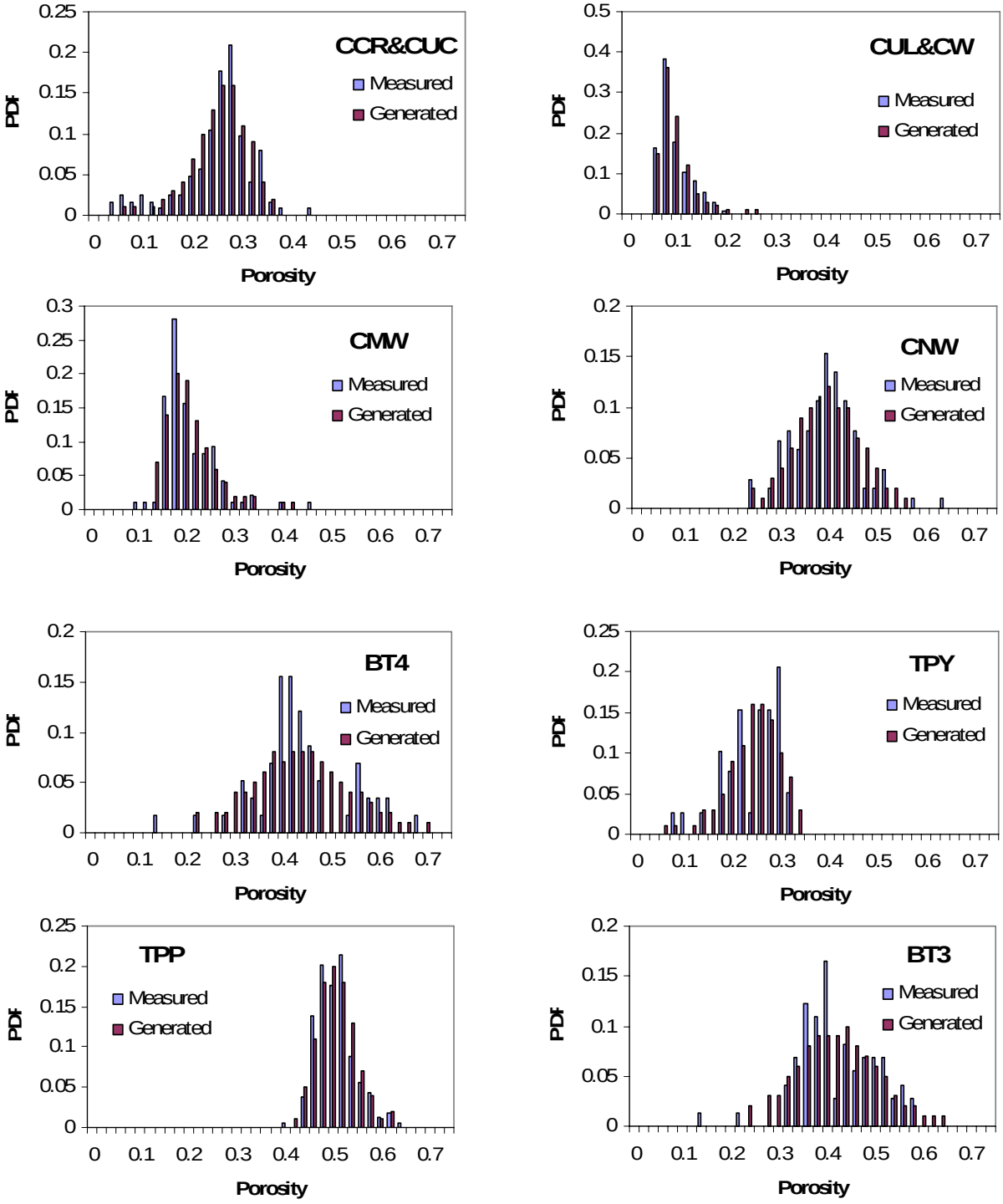
NOTE: (a) Only one porosity data point is available for BTv;
(b) Only one saturated conductivity data point is available for Pv2a, BTv and BF2 respectively;
(c) SD is standard deviation ; Min, Max are the minimum and maximum values; N is the sample size.

Table 3: Statistical parameters of matrix porosity for distribution approximation

HGU	Limits of Variation		Trans -form	Estimated Distribution			Critical Values (T^*)		
	A	B		Mean	Variance	T	α =0.10	α =0.05	α =0.01
CCR&CUC[b]	0.037	0.432	X^2	0.063	6.58E-04	0.083	0.072	0.080	0.093
CUL&CW	0.031	0.214	1/X	12.629	14.165	0.072	0.031	0.034	0.039
CMW	0.099	0.453	1/X	5.300	1.489	0.084	0.087	0.090	0.105
CNW	0.227	0.634	NO	0.387	4.74E-03	0.057	0.079	0.087	0.101
BT4	0.133	0.670	SU	0.520	2.63E-02	0.117	0.106	0.116	0.135
TPY	0.072	0.310	X^2	0.058	5.84E-04	0.106	0.129	0.142	0.165
BT3	0.136	0.579	SU	0.585	2.50E-02	0.084	0.094	0.104	0.121
TPP	0.387	0.624	1/X	2.021	2.56E-02	0.060	0.064	0.070	0.082
BT2[b]	0.103	0.615	SB	1.385	0.792	0.073	0.061	0.067	0.078
TC	0.011	0.274	$X^{1/3}$	0.365	4.66E-03	0.059	0.093	0.102	0.119
TR	0.061	0.268	NO	0.157	8.75E-04	0.048	0.038	0.042	0.049
TUL	0.075	0.251	NO	0.155	9.28E-04	0.044	0.038	0.042	0.049
TMN	0.054	0.193	LN	-2.218	3.29E-02	0.070	0.048	0.053	0.062
TLL	0.087	0.264	1/X	8.012	2.545	0.044	0.036	0.040	0.046
TM2&TM1	0.052	0.342	1/X	10.106	3.232	0.082	0.046	0.051	0.060
PV3	0.010	0.341	SB	-2.728	2.043	0.060	0.072	0.079	0.092
PV2a[c]	0.109	0.416	SB	0.181	6.769	0.191	0.214	0.234	0.268
PV2v	0.047	0.471	SB	-0.311	3.792	0.123	0.115	0.127	0.147
BT1a	0.157	0.401	$X^{1/3}$	0.656	1.58E-03	0.061	0.119	0.131	0.152
BT1v[c]	0.030	0.501	NO	0.324	7.20E-03	0.088	0.090	0.099	0.115
CHV[b]	0.037	0.491	1/X	2.944	0.079	0.061	0.071	0.078	0.090
CHZ	0.098	0.434	X^2	0.106	8.30E-04	0.068	0.035	0.039	0.045
BTa	0.180	0.419	1/X	3.791	0.400	0.064	0.094	0.104	0.121
BTv[a]									
PP4	0.215	0.441	NO	0.327	2.49E-03	0.049	0.108	0.118	0.138
PP3	0.245	0.396	NO	0.318	9.94E-04	0.054	0.062	0.068	0.080
PP2	0.098	0.334	NO	0.221	3.37E-03	0.066	0.071	0.079	0.091
PP1	0.163	0.427	X^2	0.090	6.21E-04	0.050	0.049	0.053	0.062
BF3[b]	0.058	0.370	1/X	8.573	9.402	0.078	0.079	0.086	0.101
BF2	0.159	0.330	1/X	4.451	0.805	0.095	0.127	0.140	0.163

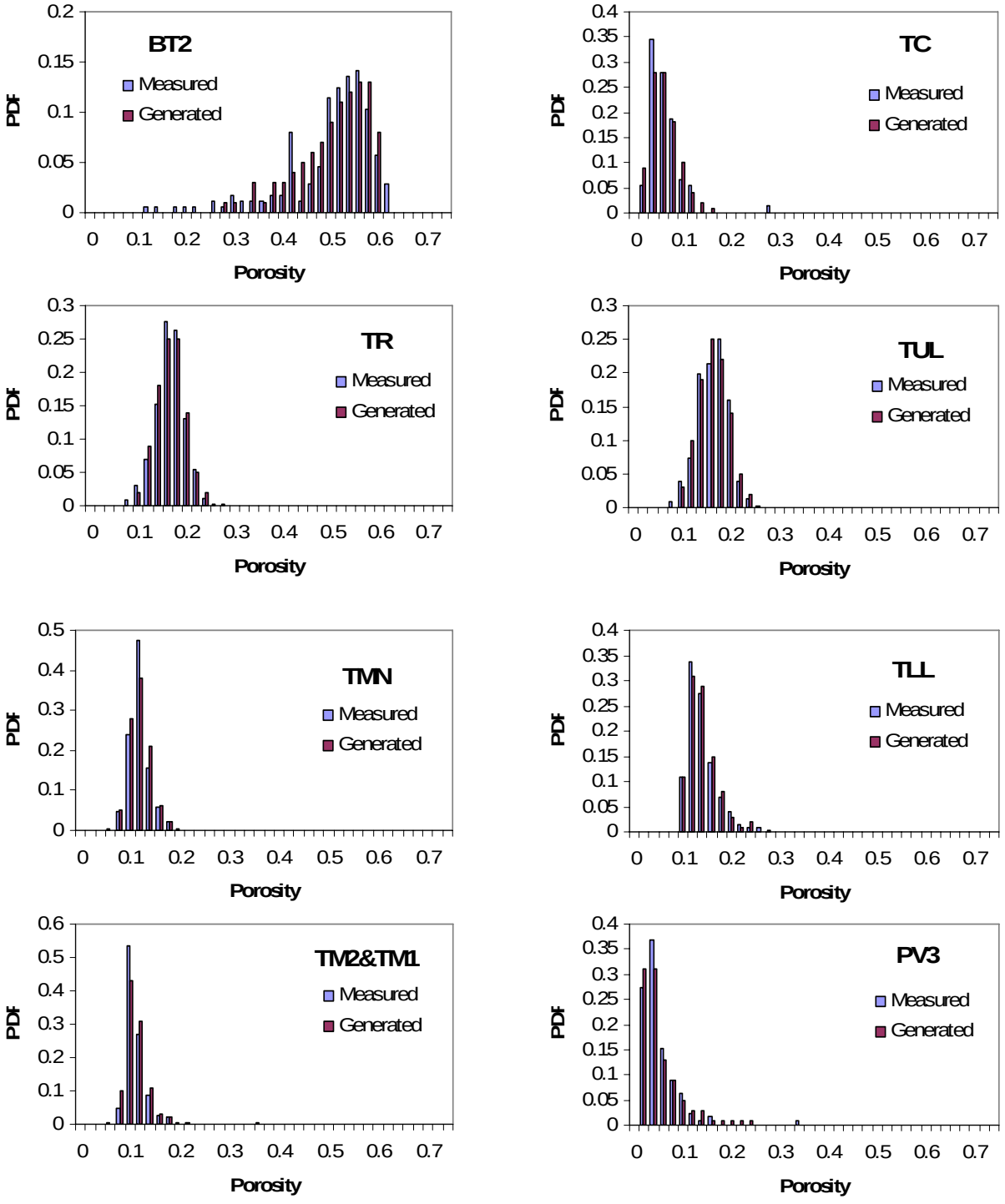
Source: Measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Statistical parameter data come from DID:016FP.001; see word file "transformation_porosity.doc".

Note: (a) The sample size is less than 4 and cannot fit the distribution using Lilliefors Test in the layers.
(b) The outlying values were discarding but were included for goodness of fit calculation in the layer.
(c) The distribution is not best fit in order to guarantee the reasonable ranges of random fields.



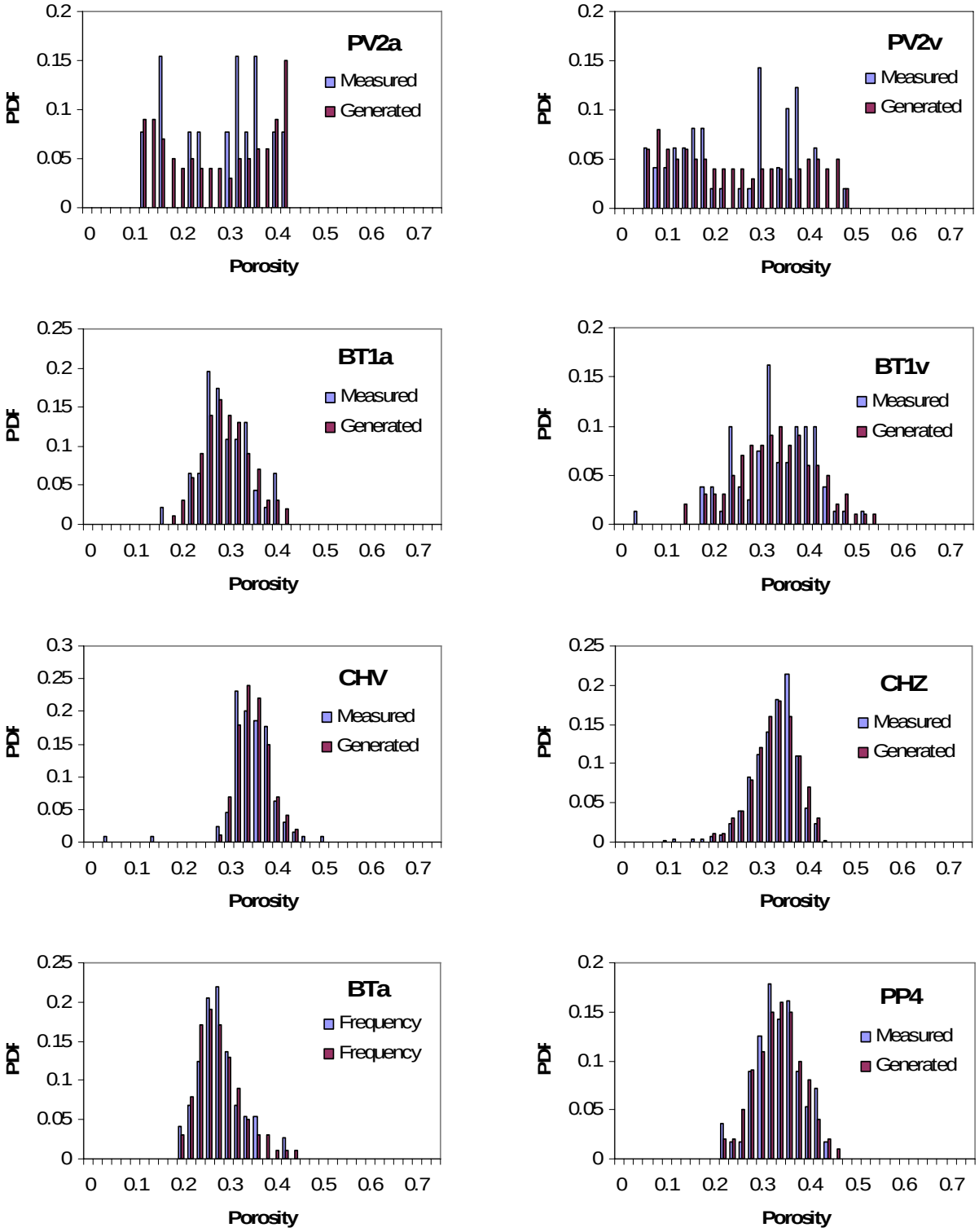
Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Generated data come from DID: 016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 4. Histograms of measured and generated data of matrix porosity for each layer



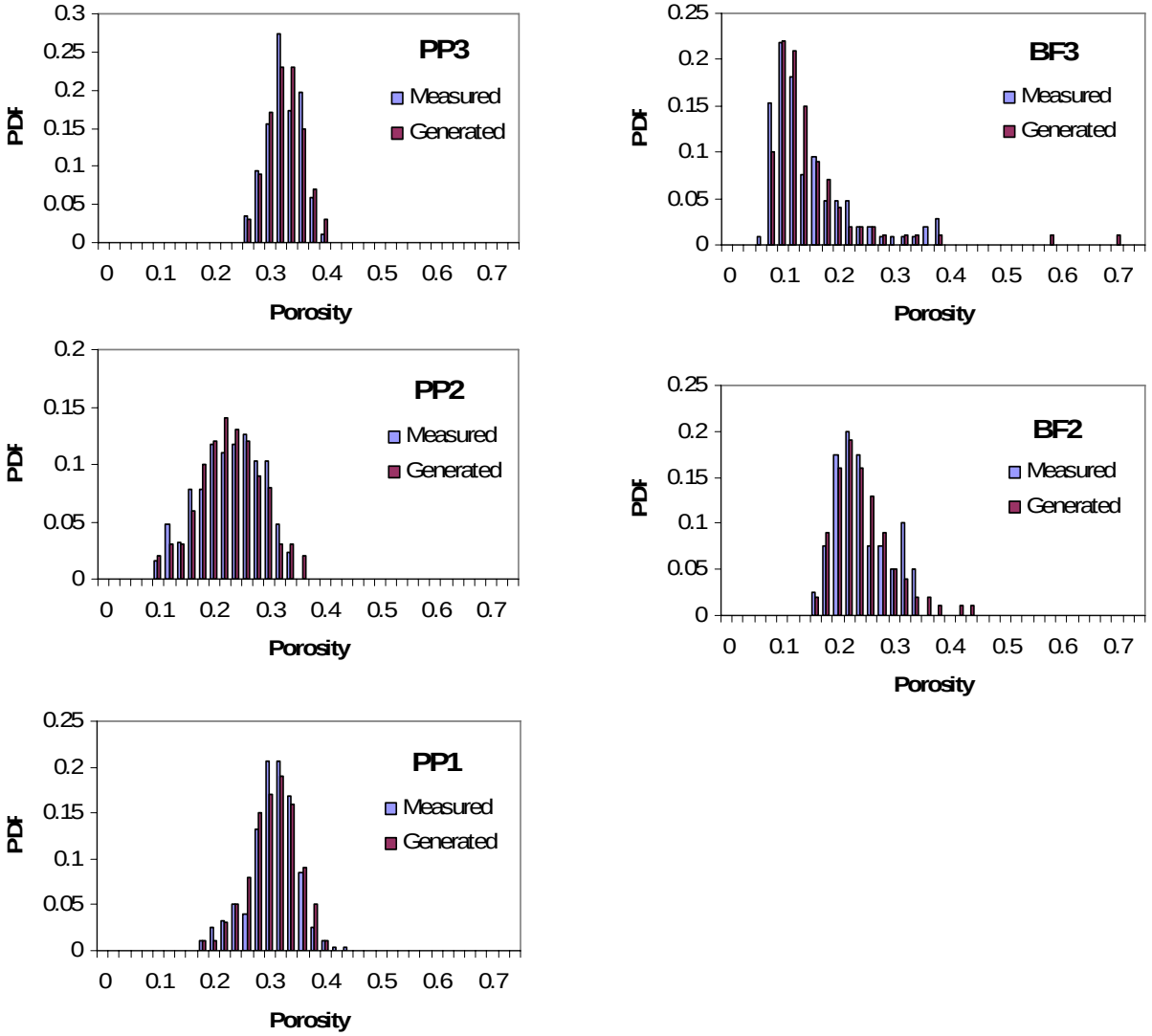
Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Generated data come from DID: 016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 4. (Cont.) Histograms of measured and generated data of matrix porosity for each layer



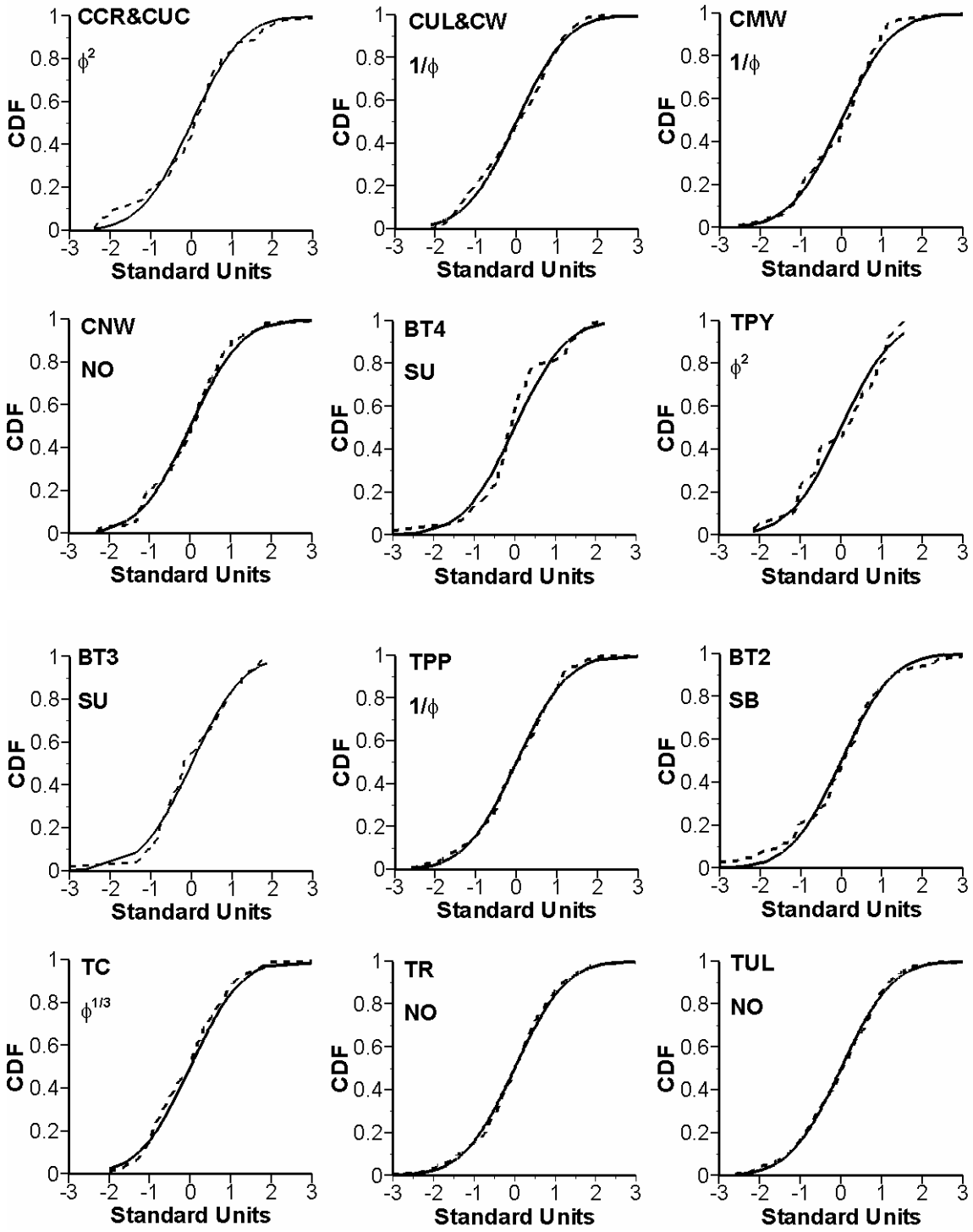
Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Generated data come from DID: 016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 4. (Cont.) Histograms of measured and generated data of matrix porosity for each layer



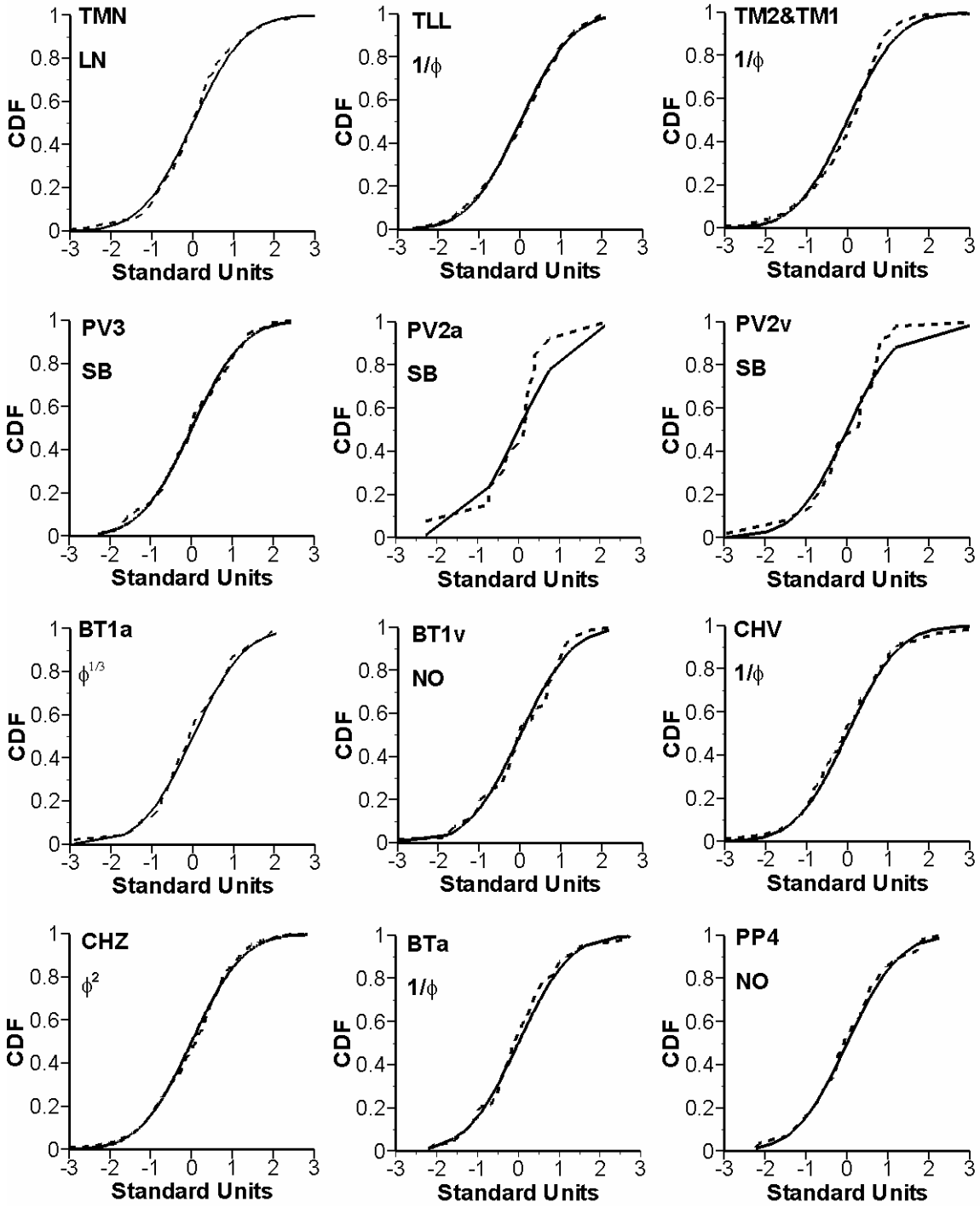
Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Generated data come from DID: 016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 4. (Cont.) Histograms of measured and generated data of matrix porosity for each layer



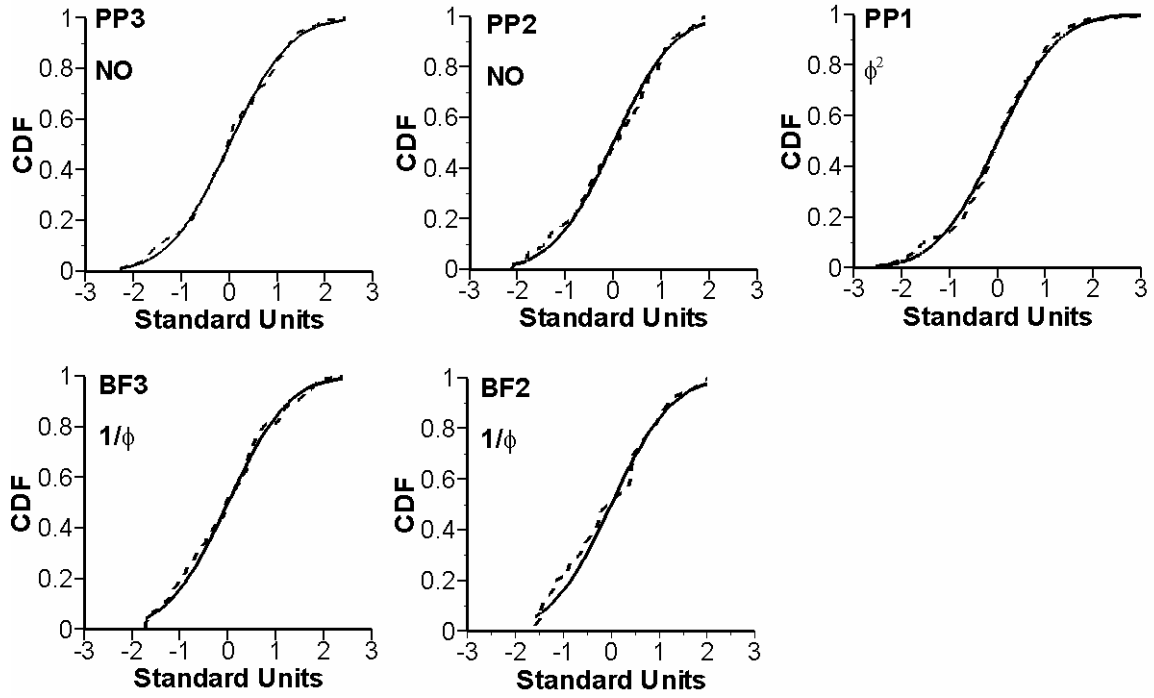
Source: Data come from DID016FP.001; see word file "Transformation_porosity.doc".

Figure 5. Empirical and theoretical distributions for transformed porosity in each layer



Source: Data come from DID016FP.001; see word file "Transformation_porosity.doc".

Figure 5. (Cont.) Empirical and theoretical distributions for transformed porosity in each layer



Source: Data come from DID016FP.001; see word file "Transformation_porosity.doc".

Figure 5. (Cont.) Empirical and theoretical distributions for transformed porosity in each layer

7.1.2 Probability Distribution of Matrix Saturated Hydraulic Conductivity

Measurements of matrix saturated hydraulic conductivity in the UZ at Yucca Mountain are also collected from Yucca Mountain database (DTN: LB0207REVUZPRP.002). The descriptive statistics of matrix saturated hydraulic conductivity are calculated in Table 2. The number of the measurements of matrix saturated hydraulic conductivity is significantly less than that of matrix porosity. It is not enough to apply Lillifors Test to select the best transformations in some layers such as CCR&CUC, TPY, PV2a, BTv and BF2 layers. From Table 2, one can see that there exists spatial heterogeneity in matrix saturated hydraulic conductivity in each layer because of the big difference of measured data. The conductivity has several orders of magnitude difference between the minimum and maximum measured values in all layers. For example it is from $5.12*10^{-12}$ m/s to $8.79*10^{-7}$ m/s in CNW layer. Therefore, the deterministic hydraulic conductivity cannot represent the real field and its heterogeneity should be assessed. The histograms of matrix saturated hydraulic conductivity for some layers are shown in Figure 6. These figures also plot the histograms of the generated random fields of matrix saturated hydraulic conductivity in corresponding layers. Due to the lacking of measured data, the histograms of some layers whose sample sizes are less than 20 such as CUL&CW, CMW, CNW and BT4 etc. are not plotted here. One can see from Figure 6 qualitatively that matrix saturated hydraulic conductivity has a lognormal distribution in TR, TUL and TMN layers. For most layers, the distributions cannot be determined from the histograms of the parameter.

The Lillifors Test and transformations are also used to fit the distributions of matrix saturated hydraulic conductivity for most layers. Table 4 lists the statistical parameters of distribution approximation for matrix saturated hydraulic conductivity in the same manner as Table 3. The empirical and fitted CDF of the best fitting distribution for the parameter in some layers are shown in Figure 7. The Lilliefors test is not applied to select the best distribution for the six layers of CCR&CUC, TPY, PV2a, BTv and BF2, since there are fewer than four measurements in each of the layers (matrix permeability in the six layers is thus fixed consequently in Monte Carlo simulations). For the units of PV2v, CHV, and CHZ, as suggested by Carsel and Parrish (1988), a few outliers are excluded from the estimations of mean and variance to yield a better distribution fit, whereas the outliers are still included in the calculation of the maximum absolute distribution difference (T). For the units of CNW, BT4, BT3, BT2, and BF3, the best distribution yields a negative 1st percentile, $V_{0.01}$, ca1culated by Eq.8 and thus is not selected. Instead, the second best fit is selected. Eighty percent of the selected best distributions are determined at the significance level of 0.05 in the Lilliefors normality test. Whereas the normality assumption is rejected in the two units of TMN and CHZ at all significance levels, the distribution with the smallest value of T is selected, which appears acceptable because of the small difference between the empirical and theoretical CDFs, $G(Z)$ and $F^*(Z)$, shown in Figure 7. The logarithm transform (LN) is the best distribution for the saturated hydraulic conductivity for 14 of the total 25 tested hydrogeologic layers.

7.1.3 Correlation of Porosity and Hydraulic Conductivity

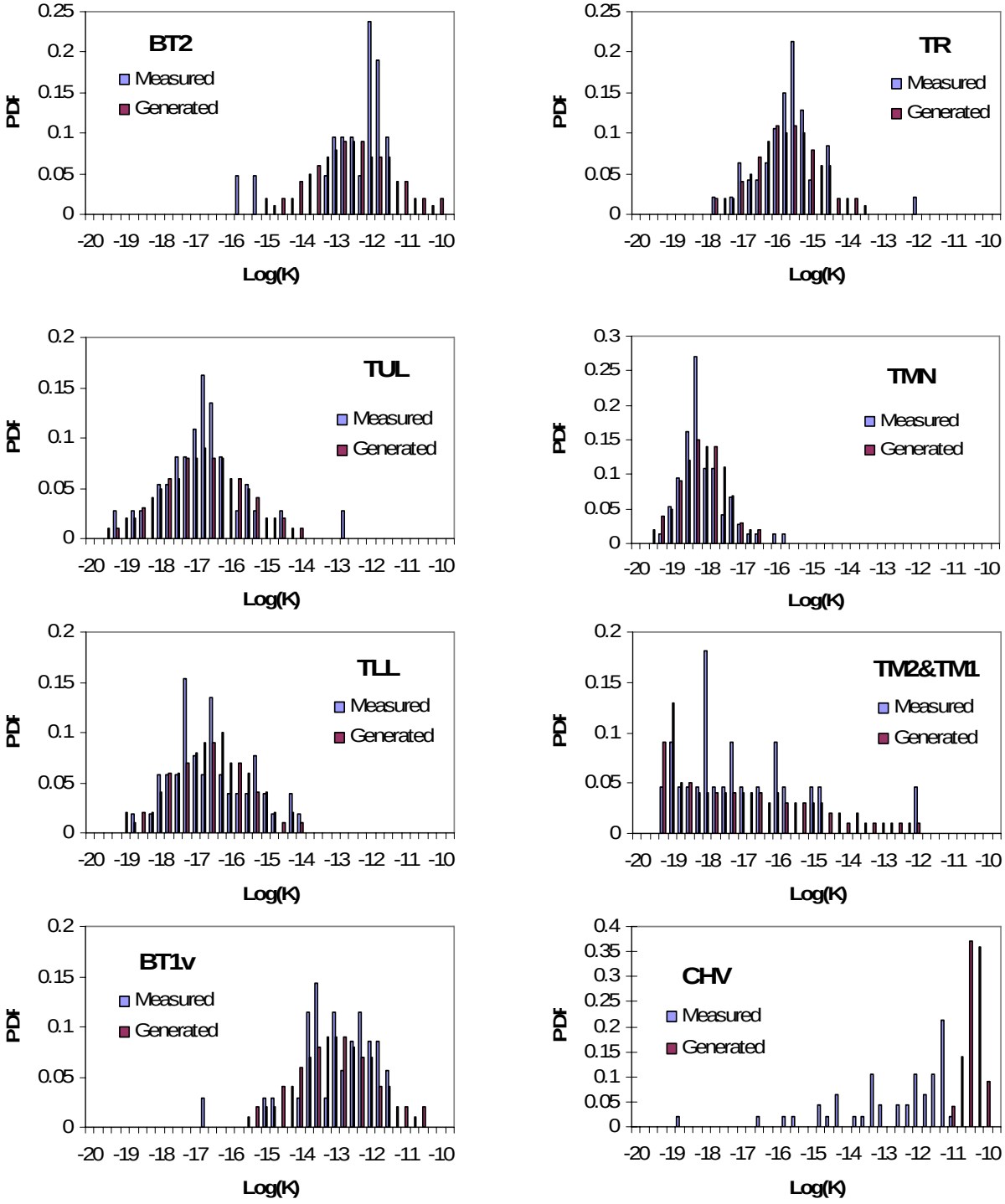
The Spearman rank correlation between matrix porosity and saturated hydraulic conductivity is used as the input data of LHS V2.51 code to generate the random fields with correlated matrix porosity and saturated hydraulic conductivity. The Spearman rank correlation is calculated according Eq. 7 described in Section 5.3. Table 5 displays the Spearman rank correlations between the transformed matrix porosity and saturated hydraulic conductivity. Note that the data pairs between the two parameters are significantly less than the measurements of individual parameters. In addition, the Spearman rank correlation cannot be estimated for hydrogeologic layers of CCR&CUC, TPY, PV2a, BTv, and BF2 because of the small number of data pairs.

Table 4: Statistical parameters of matrix saturated hydraulic conductivity for distribution approximation

HGU	Limits of Variation		Trans-form	Estimated Distribution			Critical value (T^*)		
	A	B		Mean	Variance	T	α =0.10	α =0.05	α =0.01
CCR&CUC[a]									
CUL&CW	2.48E-13	1.26E-08	LN	-25.445	8.373	0.140	0.189	0.206	0.245
CMW	3.33E-12	9.42E-08	SB	-6.061	67.470	0.323	0.315	0.337	0.405
CNW[c]	5.11E-12	8.80E-07	LN	-18.206	22.218	0.235	0.239	0.258	0.294
BT4[c]	1.79E-10	2.55E-05	LN	-14.173	10.558	0.207	0.230	0.249	0.284
TPY[a]									
BT3[c]	1.89E-09	7.31E-06	$X^{1/3}$	1.04E-02	2.40E-05	0.236	0.230	0.249	0.284
TPP	8.99E-08	1.75E-06	SB	0.754	16.897	0.209	0.230	0.249	0.284
BT2[c]	1.23E-09	2.07E-05	LN	-13.025	6.607	0.192	0.171	0.187	0.225
TC	1.69E-11	1.69E-07	SB	-5.231	42.758	0.172	0.294	0.319	0.364
TR	1.69E-11	9.38E-06	LN	-20.318	4.469	0.133	0.117	0.129	0.150
TUL	4.19E-13	1.43E-06	LN	-22.833	7.287	0.163	0.132	0.146	0.169
TMN	4.75E-13	1.24E-09	LN	-25.816	2.322	0.134	0.094	0.103	0.120
TLL	1.38E-12	7.66E-08	LN	-22.258	6.417	0.096	0.112	0.123	0.143
TM2&TM1	5.32E-13	9.40E-06	SB	-12.351	33.398	0.140	0.168	0.183	0.219
PV3	8.62E-14	2.26E-09	LN	-25.806	7.251	0.154	0.189	0.206	0.245
PV2a[a]									
PV2v[b]	5.02E-11	1.21E-05	$X^{1/3}$	1.26E-02	1.26E-05	0.211	0.195	0.213	0.250
BT1a	1.82E-13	8.71E-08	SB	-7.159	59.163	0.189	0.239	0.258	0.294
BT1v	1.03E-10	2.21E-05	LN	-14.147	6.481	0.131	0.136	0.150	0.174
CHV[b]	1.67E-12	7.21E-05	$X^{1/3}$	1.68E-02	1.34E-04	0.144	0.117	0.129	0.150
CHZ[b]	3.87E-13	9.55E-08	LN	-24.065	1.833	0.118	0.081	0.089	0.104
BTa	2.07E-13	2.11E-10	SB	-3.797	26.293	0.182	0.249	0.271	0.311
BTv[a]									
PP4	8.43E-13	3.09E-07	SB	-7.114	52.315	0.239	0.261	0.285	0.331
PP3	4.19E-12	3.66E-07	$X^{1/3}$	3.72E-03	1.62E-06	0.114	0.113	0.124	0.144
PP2	3.74E-12	1.16E-08	LN	-22.246	5.046	0.101	0.136	0.150	0.174
PP1	1.69E-12	1.95E-06	LN	-22.636	10.436	0.149	0.152	0.167	0.195
BF3[c]	6.89E-11	5.59E-08	LN	-20.018	7.319	0.191	0.276	0.300	0.348
BF2[a]									

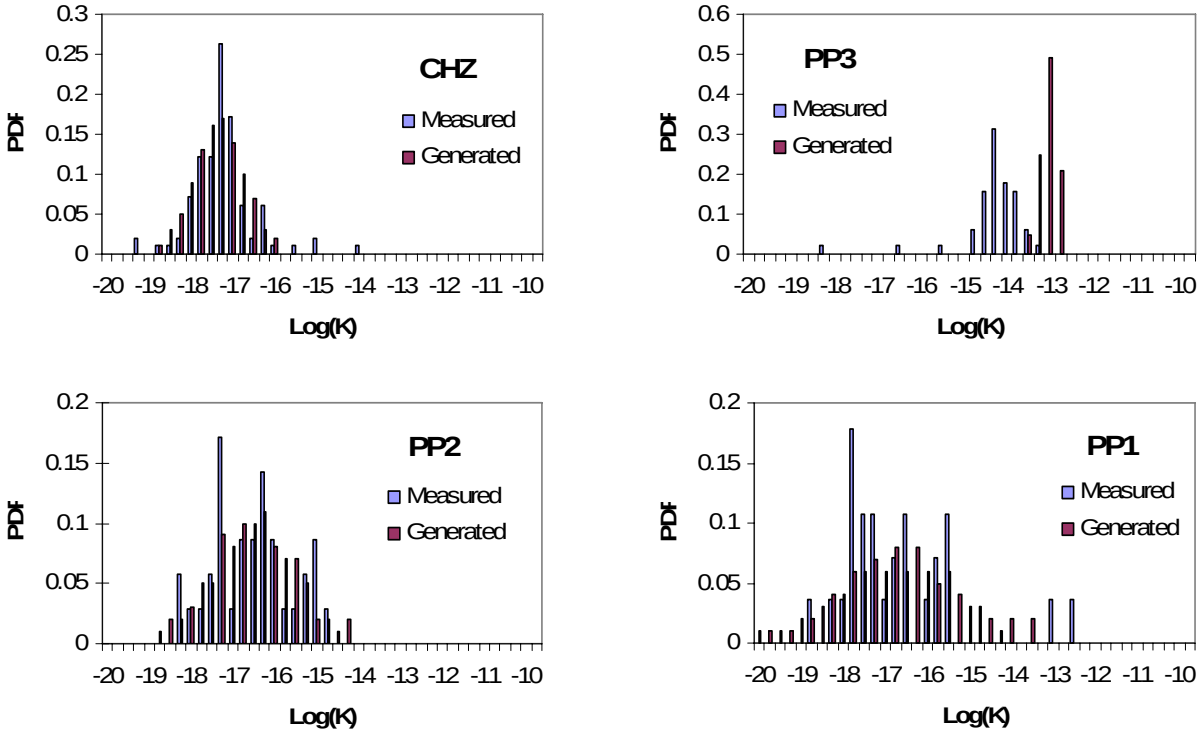
Source: Measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Statistical parameter data come from DID: 016FP.001; see word file "transformation_conductivity.doc".

Note: (a) the sample size is less than 4 and cannot fit the distribution using Lilliefors Test in the layers.
(b) The outlying values were discarding but were included for goodness of fit calculation in the layer.
(c) The distribution is not best fit in order to guarantee the reasonable ranges of random fields.



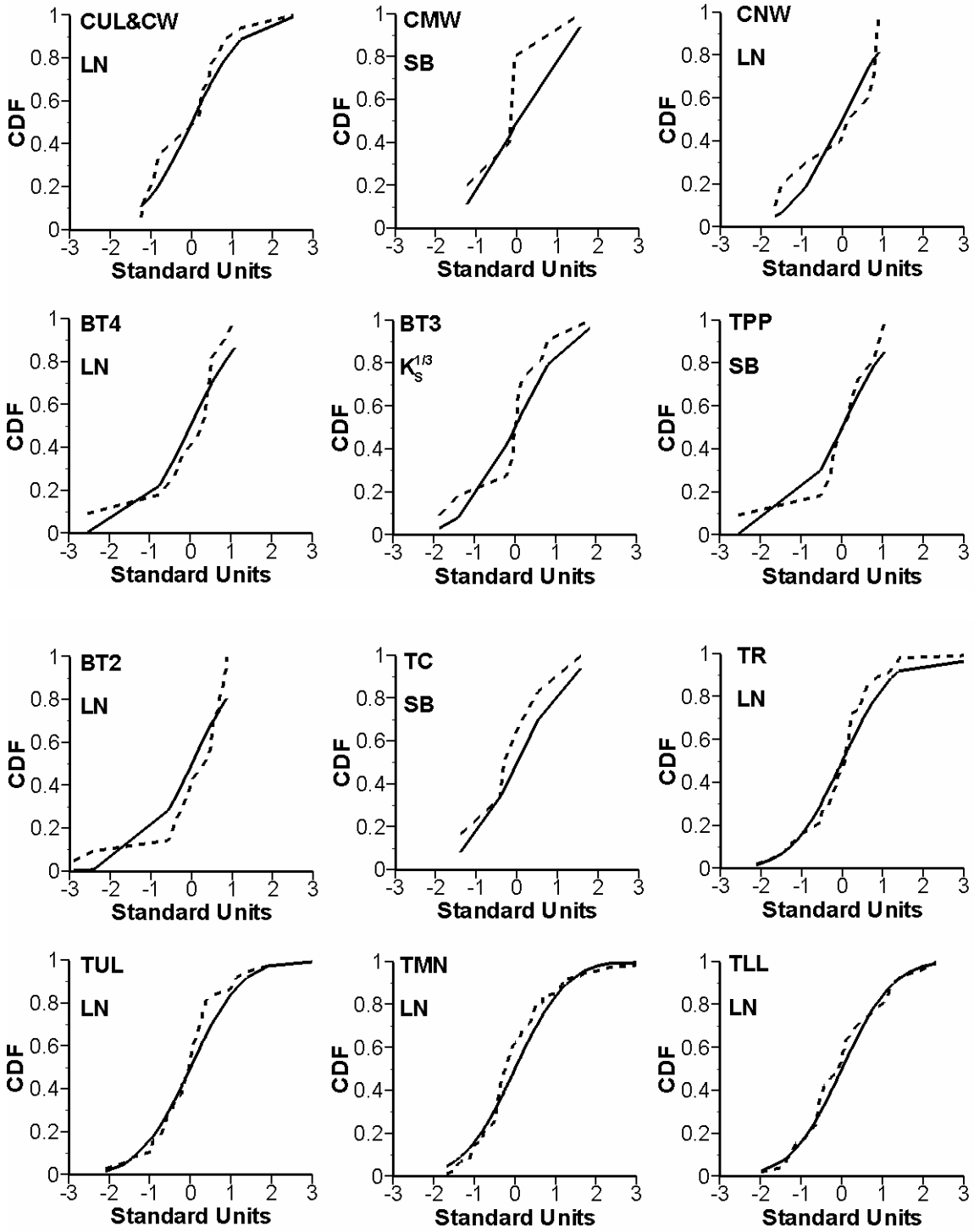
Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Generated data come from DID016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 6. Histograms of measured and generated data of matrix log permeability for the layers with sufficient measurements



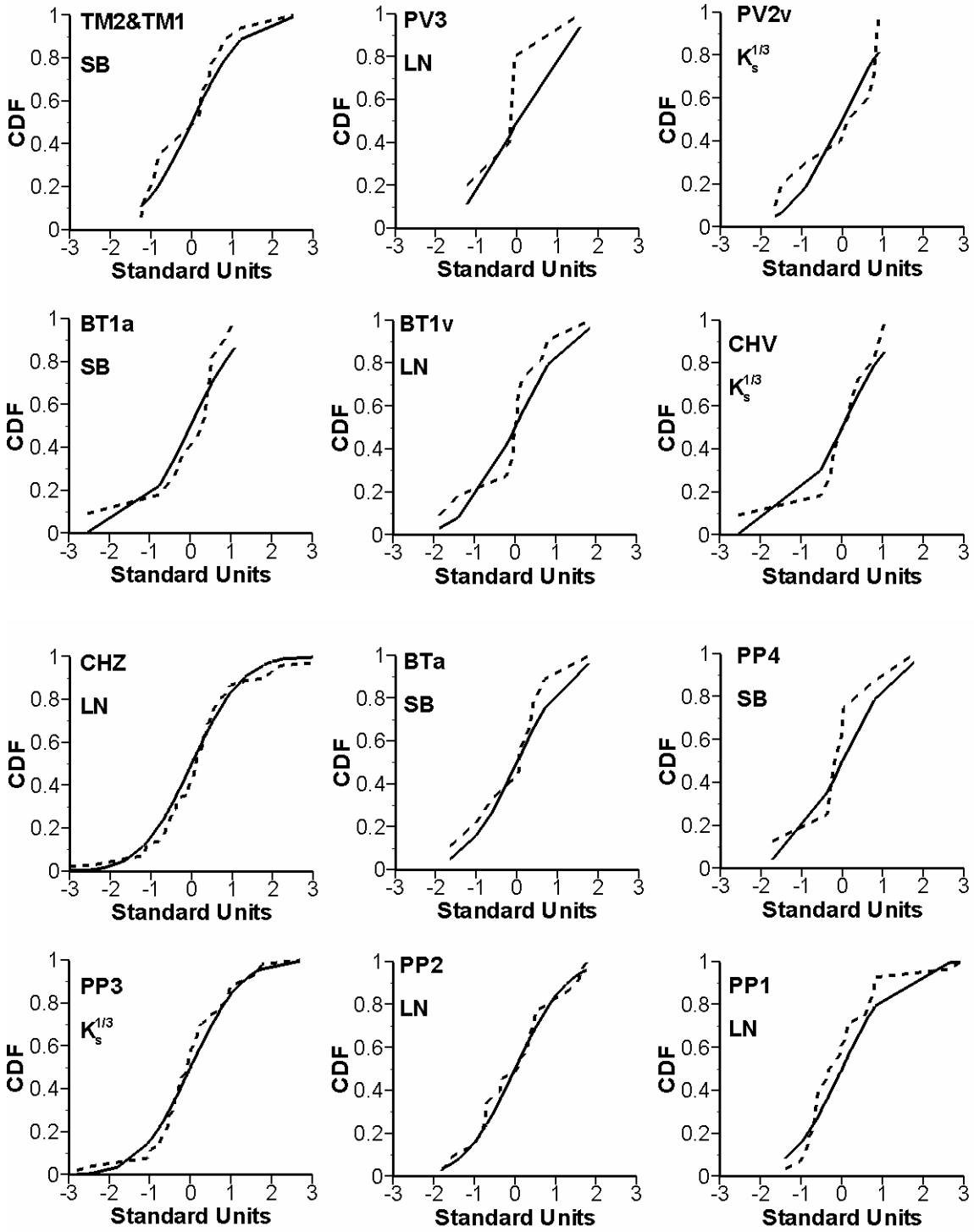
Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Generated data come from DID016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 6. (Cont.) Histograms of measured and generated data of matrix log permeability for the layers with sufficient measurements



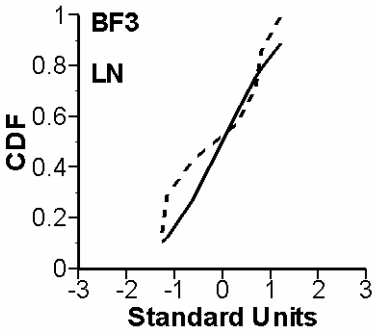
Source: Data come from DID016FP.001; see word file "Transformation_conductivity.doc".

Figure 7. Empirical and theoretical distributions for transformed hydraulic conductivity in each layer



Source: Data come from DID016FP.001; see word file "Transformation_conductivity.doc".

Figure 7. (Cont.) Empirical and theoretical distributions for transformed hydraulic conductivity in each layer



Source: Data come from DID016FP.001; see word file "Transformation_conductivity.doc".

Figure 7. (Cont.) Empirical and theoretical distributions for transformed hydraulic conductivity in each layer

Table 5: Spearman rank correlation between transformed data of matrix porosity and hydraulic conductivity

HGU	Spearman rank correlation	Sample size	HGU	Spearman rank correlation	Sample size
CCR&CUC[b]	N/A	3	PV3	-0.20	17
CUL&CW	-0.50	17	PV2a[a]	N/A	1
CMW	0.60	5	PV2v	0.06	16
CNW	0.61	10	BT1a	0.12	10
BT4	0.26	11	BT1v	0.37	35
TPY[b]	N/A	2	CHV	-0.19	47
BT3	0.03	11	CHZ	0.47	99
TPP	-0.47	11	BTa	0.22	9
BT2	0.42	21	BTv[a]	N/A	1
TC	-0.49	6	PP4	0.52	8
TR	0.39	47	PP3	0.45	51
TUL	0.40	37	PP2	0.68	35
TMN	0.48	9	PP1	0.24	28
TLL	-0.46	52	BF3	-0.71	7
TM2&TM1	-0.39	22	BF2[a]	N/A	1

Source: Measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Rank correlation data come from DID: 016FP.001; see word file "MeasurementStatistics.doc".
Note: (a) Only one measured data point is available for Pv2a, BTv and BF2 respectively;
(b) The measured data points are less than 5 in CCR&CUC and TPY.

7.1.4 Probability Distribution of Matrix Sorption Coefficient

The sorption coefficient (K_d) of the reactive tracer is the most important factor in transport simulations and has significant effects on the residence time of radionuclide in the UZ of Yucca Mountain. The measured sorption coefficients for the reactive tracer Neptunium (^{237}Np) in three types of rocks (Devitrified, Vitric and Zeolitc tuffs) are obtained from Yucca Mountain Database (DTN: LA0407AM831341.004). The descriptive statistics of sorption coefficient for ^{237}Np are tabulated in Table 6. From Table 6, one can see that the sample size is large enough to determine the distributions of the parameter for those three types of rocks. One can also know that the spatial heterogeneity in sorption coefficient exists in the three types of rocks because of significant difference of measured data. The measured K_d values have broad ranges from 0.008 to 8.235 mL/g in devitrified tuff, from 0.020 to 4.071 mL/g in vitric tuff and from 0.032 to 8.742 mL/g in zeolitic tuff. The heterogeneity of K_d can cause significant uncertainties on radionuclide transport and the effects of heterogeneous K_d should be assessed for uncertainty analysis. The histograms of sorption coefficient for three rock types are shown in Figure 8. One can observe that the sorption coefficient of neptunium has a lognormal distribution in Devitrified and Vitric Tuffs.

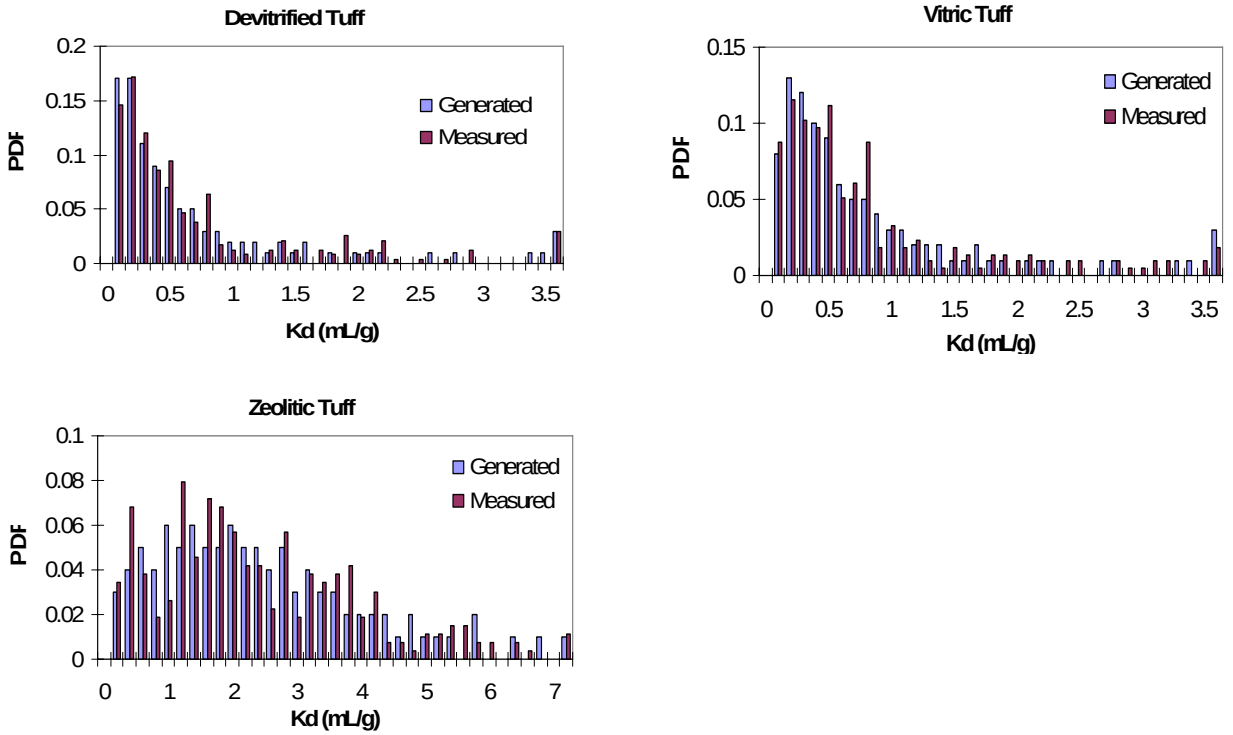
Table 7 presents the statistical parameters of distribution approximation for the sorption coefficient of the reactive tracer neptunium (^{237}Np) in the same manner as Tables 3 and 4. The empirical and fitted CDF of the best fitting distribution for K_d in the three rock types are shown in Figure 9. The best distributions of the sorption coefficients for the three rock types of devitrified, vitric, and zeolitic tuff are selected based on the Lilliefors test at a significance level of 0.05. For two of the three rock types, the logarithm transform is found to be the best distribution for the sorption coefficient.

The results of distributions determination of parameters are the basis of random field generation for Monte Carlo Simulation and also have important effects on final Monte Carlo Simulation results because the input of the simulations would be obtained based on the determined distributions.

Table 6: Descriptive statistics for sorption coefficient of neptunium

Materials	Sorption Coefficient (K_d, mL/g) of Neptunium (Np)				
	Mean	Stand Deviation	Minimum	Maximum	Sample Size (N)
Devitrified Tuff	0.720	1.006	0.008	8.235	233
Vitric Tuff	0.808	0.855	0.020	4.071	216
Zeolitic Tuff	2.333	1.589	0.032	8.742	264

Source: Measurements come from DTN: LA0407AM831341.004, file “Neptunium(RIT).xls”; Statistics of measurements come from DID: 016FP.001; see word file “MeasurementStatistics.doc”.



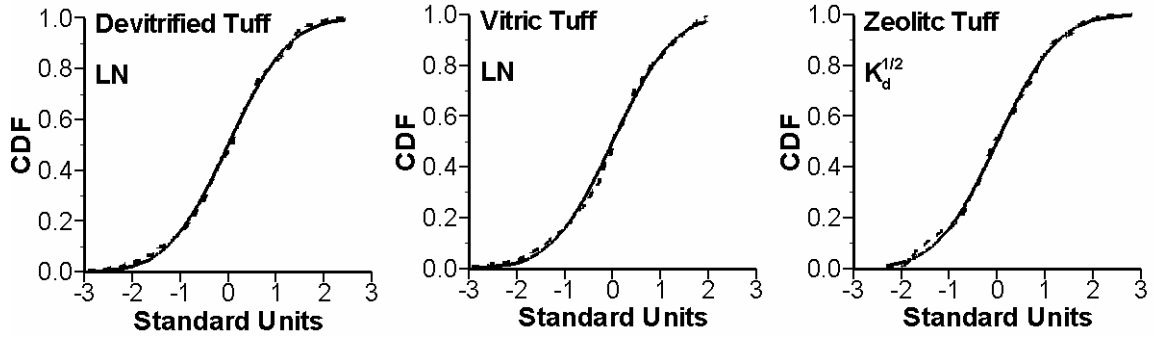
Source: Measured data come from DTN: LA0407AM831341.004, file "Neptunium(RIT).xls"; Generated data come from DID: 016FP.003; see word file "HomogeneousRandomFiled.doc".

Figure 8. Histograms of measured and generated data of sorption coefficient of neptunium for devitrified, vitric and zeolitic tuffs

Table 7: Statistical parameters of sorption coefficient of neptunium for distribution approximation

Materials	Limits of Variation		Trans -form	Estimated Distribution			Critical Values (T^*)		
	A	B		Mean	Variance	T	α =0.10	α =0.05	α =0.01
Devitrified Tuff	0.007	8.236	LN	-1.063	1.669	0.0396	0.0527	0.0580	0.0675
Vitric Tuff	0.019	4.072	LN	-0.730	1.182	0.0589	0.0548	0.0603	0.0702
Zeolitic Tuff	0.031	8.743	$X^{1/2}$	1.429	0.293	0.0382	0.0495	0.0545	0.0635

Source: Measurements come from DTN: LA0407AM831341.004, file "Neptunium(RIT).xls"; Statistical parameter data come from DID: 016FP.001; see word file "Transformation_Kd.doc".



Source: Data come from DID: 016FP.001; see word file "Transformation_Kd.doc".

Figure 9. Empirical and theoretical distributions for transformed sorption coefficient of neptunium in devitrified, vitric and zeolitic tuffs

7.1.5 Generated Realizations of Layer-scale Matrix Porosity, Permeability, and Sorption Coefficient

Two hundred realizations of saturated hydraulic conductivity, porosity, and sorption coefficients of neptunium (^{237}Np) are generated using the Latin Hypercube Sampling method for all hydrogeologic layers, except for the layers with less than four measurements. Figure 4 plots relative frequencies of measured and generated matrix porosity. Figure 4 shows the distributions of generated matrix porosity agree well with the corresponding measurements, even in the four layers of CUL&CW, TMN, TM2&TM1, and CHZ, where the normality assumption is rejected for all distributions.

Since measurements of saturated hydraulic conductivity are obtained from the Yucca Mountain database, random fields of saturated hydraulic conductivity (K, m/s) are generated and then converted to TOUGH2 required permeability (k, m^2) using the equation:

$$k = \frac{K\mu_w}{g\rho_w} \qquad \text{Eq. 20}$$

where μ_w is the viscosity of water (0.001 N s/m^2), g is the acceleration of gravity (9.81 m/s^2) and ρ_w is the density of water (998 kg/m^3) at 20 ºC.

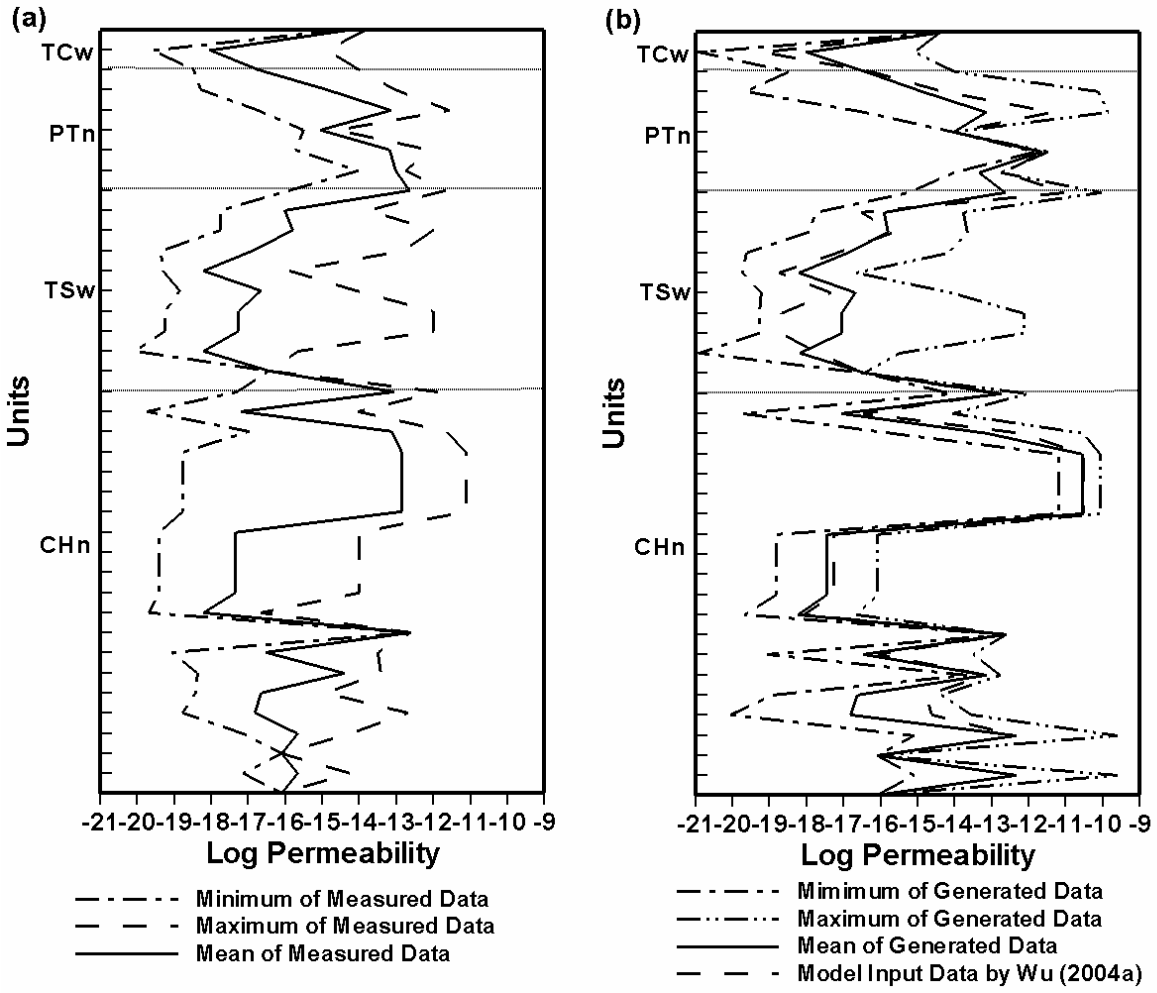
200-realization random fields of matrix saturated hydraulic conductivity are generated and these realizations are correlated with matrix porosity through the Spearman rank correlations for each hydrogeologic layer. Then they are transformed back to original scale according to the selected transforms in the layers. Based on Eq. 20, the random fields of matrix saturated hydraulic conductivity are converted to the matrix permeability. Figure 6 plots the histograms of the measured and generated matrix permeability. For the convenience of presenting, log matrix permeability, instead of matrix permeability, is plotted. Figure 6 shows good match between measured and generated data in some layers but not good match in few layers such as TM2&TM1, CHV and PP3

layers. One of the reasons is small sample size and some outliers in the layers. Another reason is that the model input data by Wu et al. (2004a) is assumed as the mean of generated random fields.

Previous modeling (e.g., Wu et al., 1999, 2004a, b) has shown that parameters measured in the field and laboratory, and/or parameters estimated by one-dimensional models, cannot be used directly by 3-D models to produce acceptable simulation results. This is due to model uncertainty and the different spatial-temporal scales between model input parameters and their measured and estimated values. Instead, the parameter set of permeability obtained from both parameter measurements and 3-D model calibrations (Bardurraga and Bardvarsson, 1999; Liu et al., 2003) were employed for the numerical simulations (Wu et al., 2004a, b). Their simulations appear reasonable in comparison with field measurements of water saturation and potential.

In this study, the model inputs of permeability used by Wu et al. (2004a, b) are included in the range (between the minimum and maximum) of measured permeability in almost all hydrogeologic layers. Nevertheless, in the units of BT3, TPP, CHV, PP3, and BF2, the model inputs are larger than the maximum measurements. To yield better simulations, the model inputs of Wu et al. (2004a, b) in these units are assumed to be the means of saturated hydraulic conductivity. The associated variances are adopted from those determined by the Lilliefors test after appropriate transforms, given in Table 4. Since the distributions of matrix saturated hydraulic conductivity cannot be identified in the hydrogeologic layers of CCR&CUC, TPY, PV2a, BTv, and BF2 because of their small sample size (fewer than four), matrix hydraulic conductivity in these layers is treated as a deterministic variable having the values of model inputs given in Wu et al. (2004a, b). Figures 10a and 10b plot the mean, maximum, and minimum logarithm of measured and generated permeability, respectively, for all hydrogeologic layers. Figure 10b shows that the model inputs from Wu et al. (2004a) are within the ranges of generated permeability but are not identical with their means. Mean permeability in Figure 10a differs from that in Figure 10b for certain layers due to the use of calibrated permeability as the mean, as discussed previously. In Figures 10a and 10b, the differences in range of permeability for certain hydrogeologic layers is caused by the exclusion of measurement outliers from estimates of mean and variance, as discussed in Section 7.1.2. For example, in the hydrogeologic layers of CHV and CHZ, deleting the outliers result in a smaller variance in permeability, which causes the smaller ranges of the generated data than the measurements.

200-realization random fields of sorption coefficient (K_d) for three types of rocks are also generated using LHS. Figure 8 shows the good agreement between the measured and generated data of sorption coefficients in three types of rocks, indicating that the random fields of sorption coefficient can respond the probabilistic distributions of the sorption coefficient.



Source: Measured data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Model input data by Wu (2004a) come from DTN: LB0303THERMSIM.001, file "th_pqm.dat"; and LB03013DSSCP3I.001, file "props_3d.doc"; Generated data come from DID: 016FP.003; see word file "HomogeneousRandomField.doc".

Figure 10. Comparison of measured data (mean, minimum and maximum) and generated data (mean, minimum and maximum) and model input data of Wu et al. (2004a) for 3-D model domain of matrix log permeability in each layer

7.2. Generated Local-scale Realizations of Matrix Porosity, and Permeability

In this section, after discussing the correlation coefficients of matrix porosity and permeability estimated from borehole measurements, we present the generated heterogeneous random fields of the two parameters.

7.2.1 Correlation Length of Matrix Porosity and Permeability

The measurements of porosity are sufficient to calculate the vertical sample variograms for all model layers, based on which vertical correlation lengths can be directly determined by fitting the variograms. However, the horizontal correlation length of porosity can only be determined by fitting their sample variograms in the certain layers with sufficient measurements. For permeability, only the vertical correlation length can be determined in several layers, and it is impossible to determine the horizontal one in any layer, due to lack of measurements. Figure 11a –bg show the vertical sample variogram for permeability and porosity and horizontal sample variogram for porosity for layers with sufficient data. Correspondingly, the fitted variograms using spherical model are also shown in Figure 11a –bg.

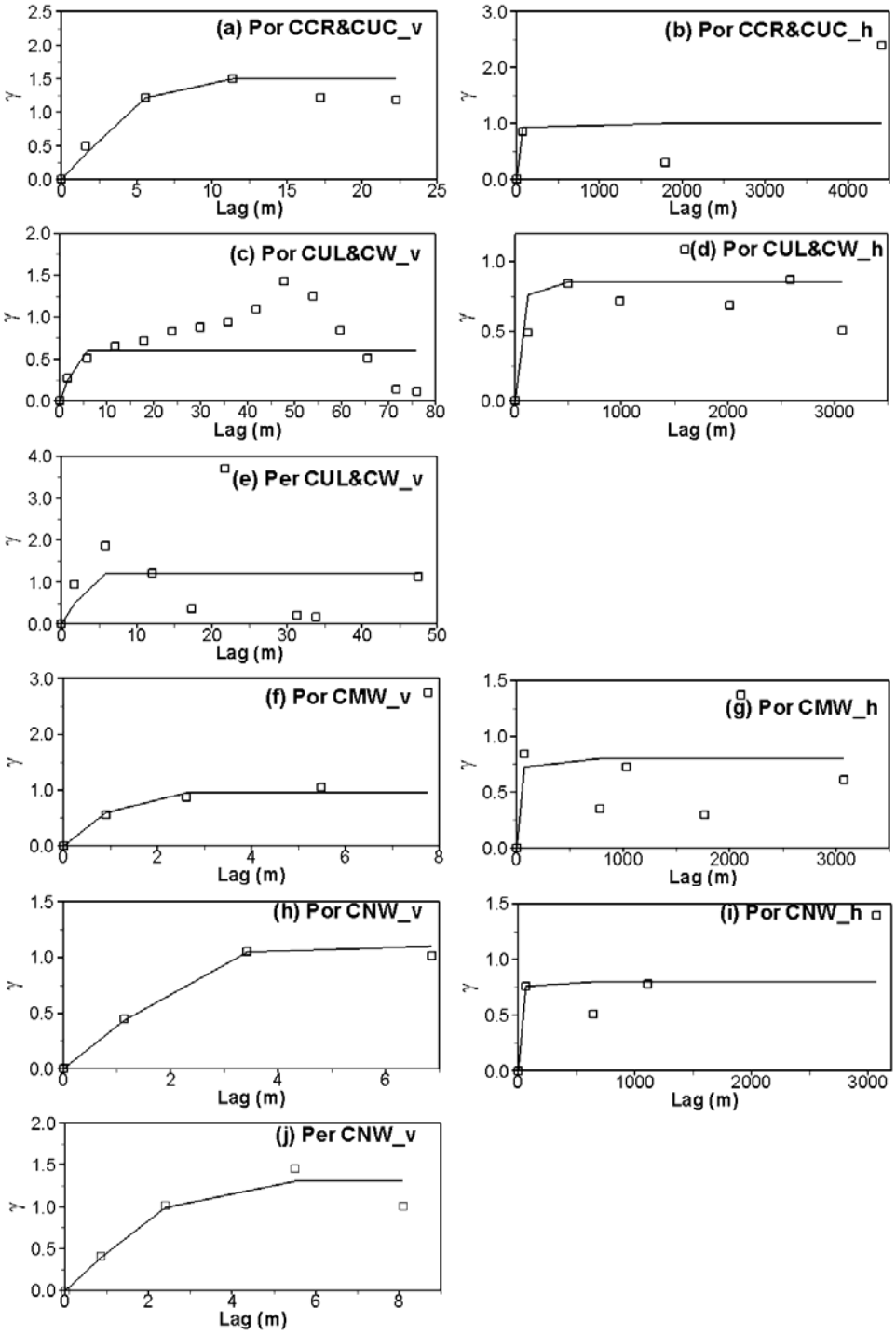
Since the horizontal correlation lengths cannot be estimated for deeper layers due to the lack of measurements, it is assumed that the horizontal correlation lengths are constant for all layers within the same major geologic unit. For BF3 and BF2 layers within CFu unit, only one borehole is drilled, and horizontal correlation lengths of porosity are assumed the same as those in the layers within CHn unit (above CFu). This assumption is based on the observation of the horizontal correlation length for layers in units TCw, PTn, and TSw shown in Figure 11.

Measurements of permeability are only sufficient to estimate vertical horizontal correlation lengths for 14 layers, as shown in Figure 11. We assume that the vertical and horizontal correlation lengths of permeability are the same as those of matrix porosity. This assumption is based on the observation of 14 layers in which the vertical correlation length can be estimated for the matrix permeability and porosity shown in Figure 11. Taking the hydrogeologic layer TLL as an example, Fig. 11ah illustrates that the vertical sample variograms of permeability can be fitted to a spherical model with correlation length of 1.8m, the same value of porosity correlation length as shown in Fig. 11af. It appears reasonable to assume that vertical and horizontal correlation lengths of permeability are the same as those of porosity.

7.2.2 Generated Local-scale Random Fields of Matrix Porosity and Permeability

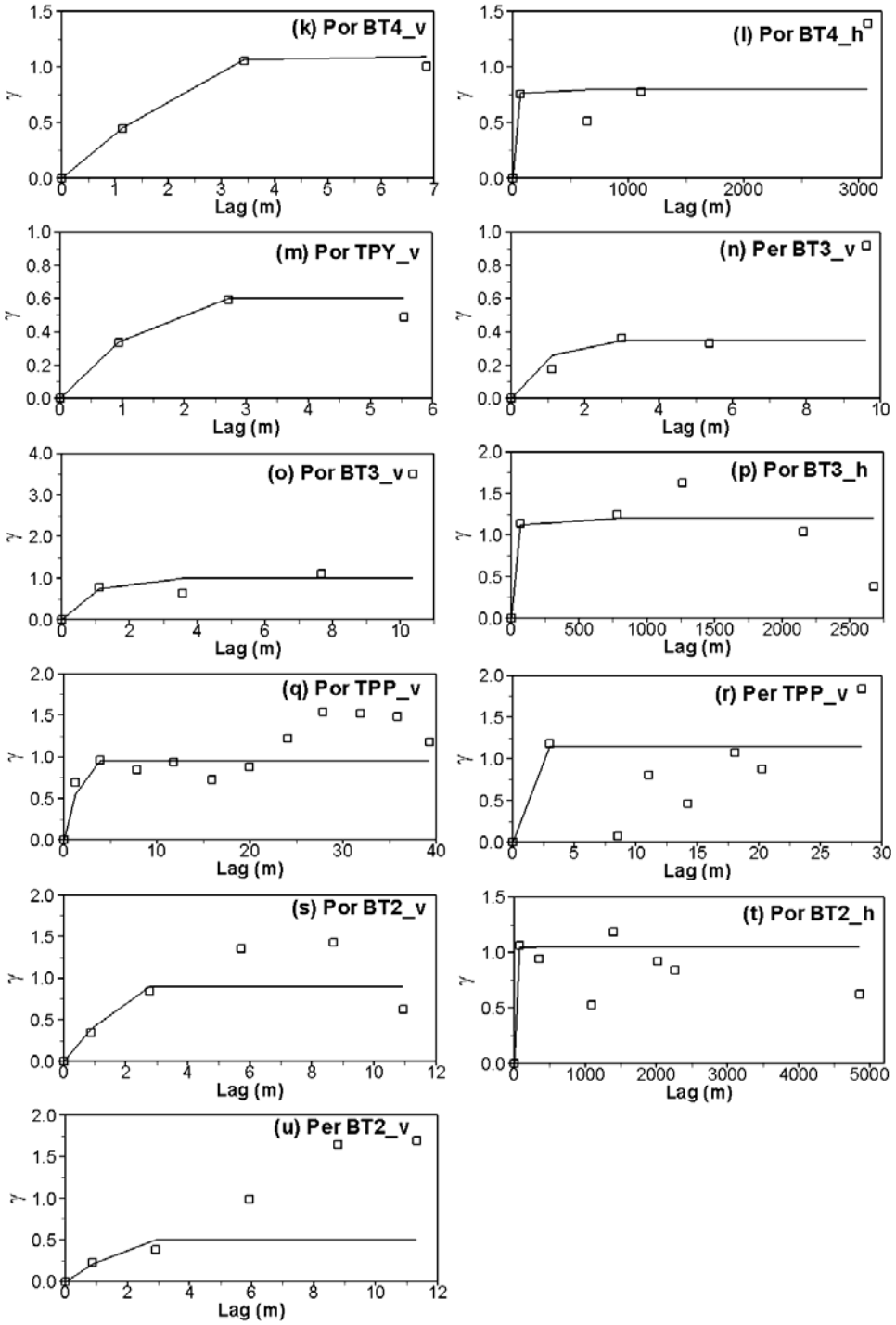
Two hundred random field realizations of matrix permeability and porosity are generated, following the procedures described in Section 5.3. Since SGSIM only supports random field generation on a regular grid, each of the generated random fields is interpolated using the method of "closest neighbor" to the 3-D irregular grid used for this study. Sample variograms of the interpolated fields are calculated for each layer to ensure that spatial correlation and variance are not affected by the interpolation. The sample variograms of the interpolated fields (figures not shown here) visually agree well with the theoretical variograms obtained from variogram analysis (Figure 11) and used to generate the random fields. After the interpolation, the random fields are back-transformed to their original values, as discussed in Section 5.2.

This procedure described in Section 5.3 is conducted for each hydrogeological layer, and the resulting random fields for all layers are combined to obtain the final parameter field used for the TOUGH2 simulation. Figure 12 plots sample mean of the 200 realizations of log permeability at the



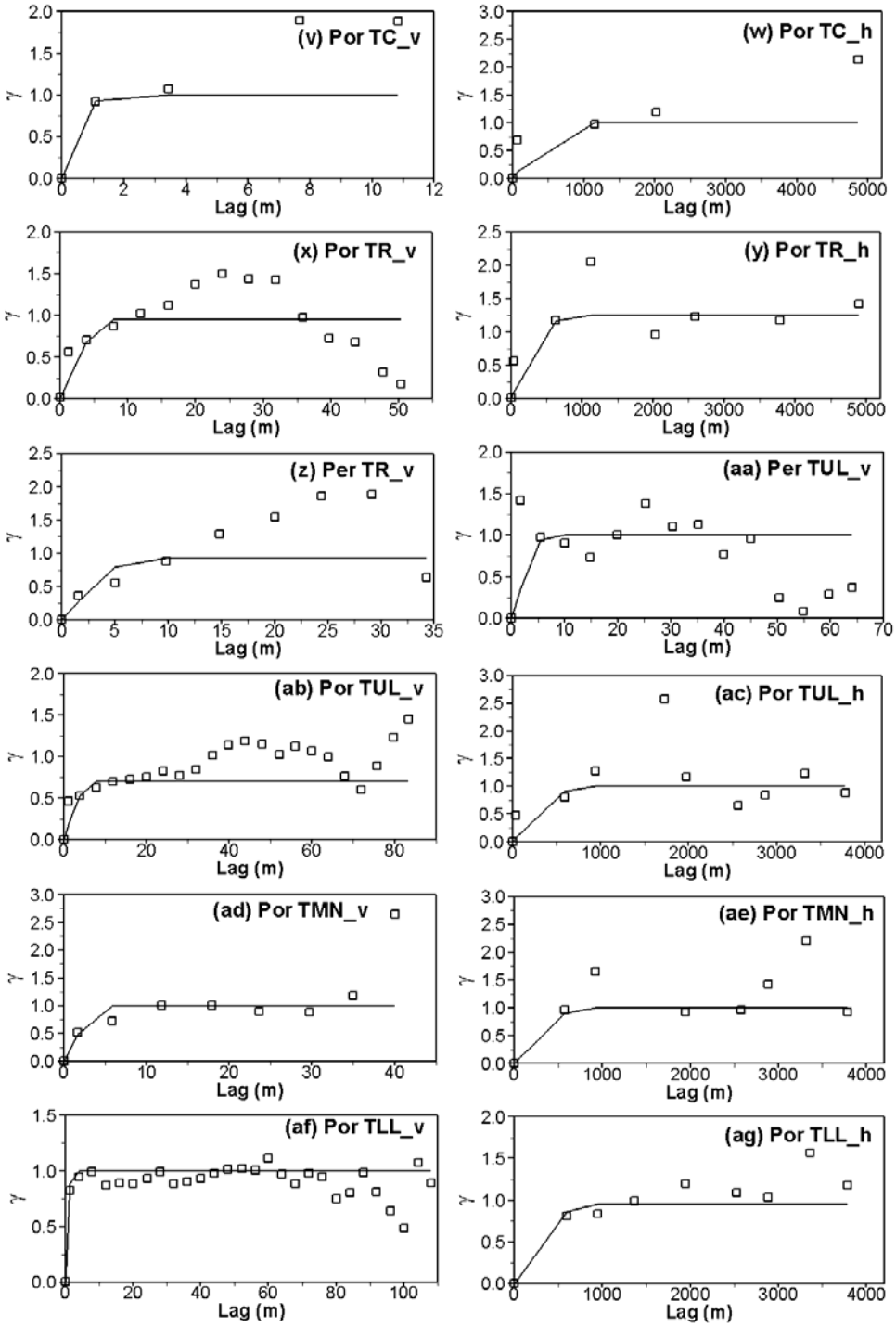
Source: The original measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Data come from DID: 016FP.003; see word file "variogram.doc".

Figure 11. Sample and fitted variograms of transformed measurements of matrix porosity in vertical and horizontal directions and matrix permeability in vertical direction for the layers with sufficient measurements



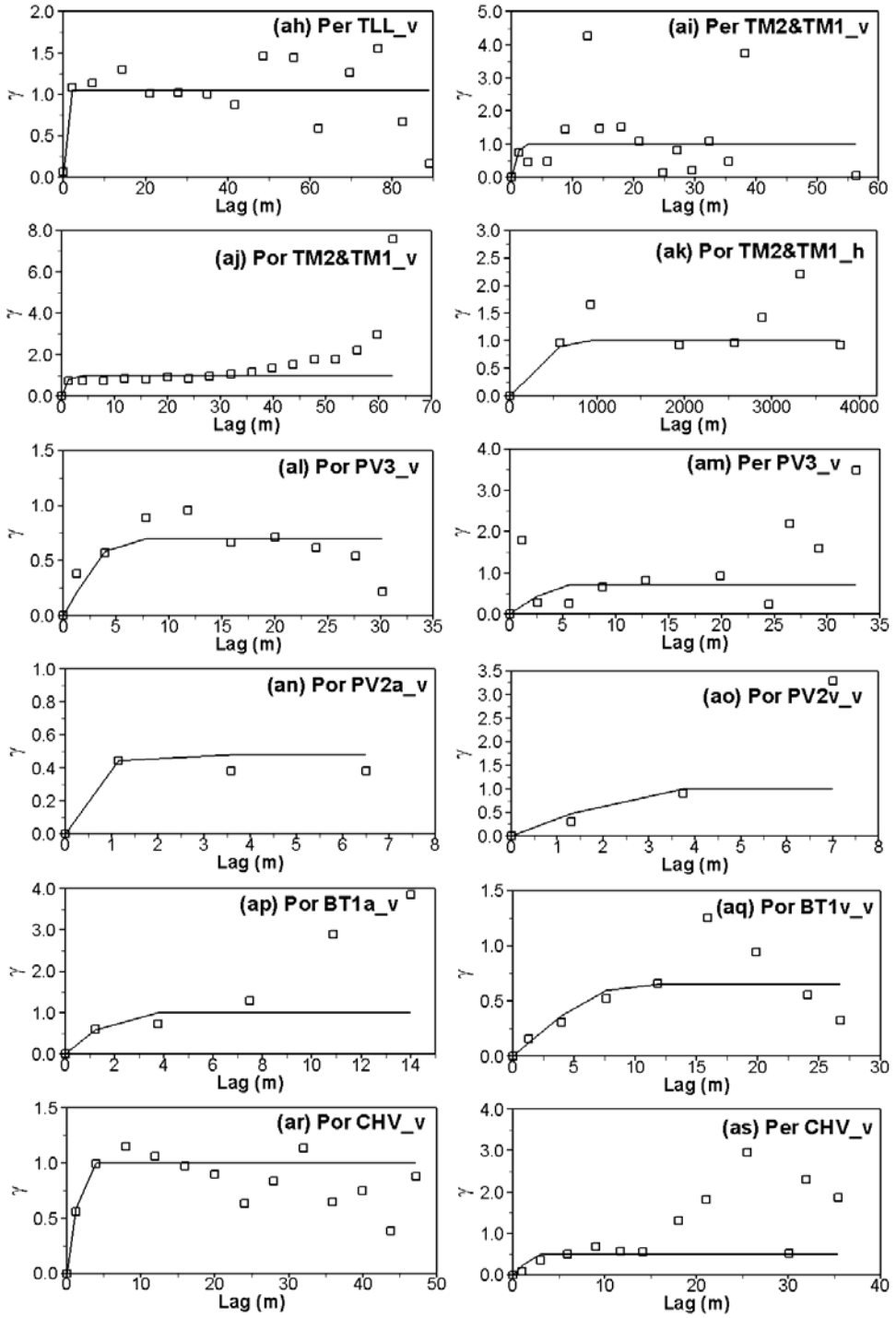
Source: The original measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Data come from DID: 016FP.003; see word file "variogram.doc".

Figure 11. (Cont.) Sample and fitted variograms of transformed measurements of matrix porosity in vertical and horizontal directions and matrix permeability in vertical direction for the layers with sufficient measurements



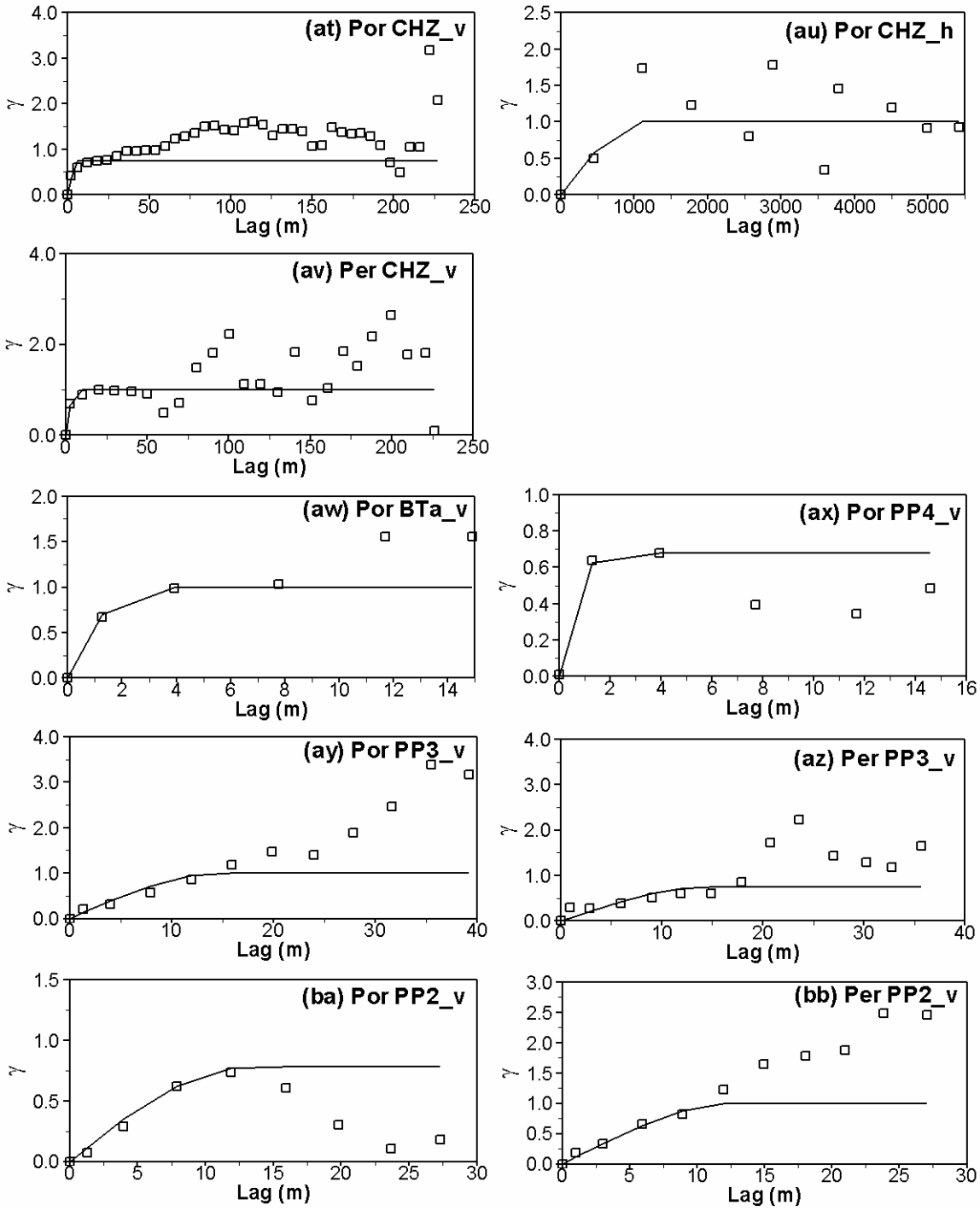
Source: The original measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Data come from DID: 016FP.003; see word file "variogram.doc".

Figure 11. (Cont.) Sample and fitted variograms of transformed measurements of matrix porosity in vertical and horizontal directions and matrix permeability in vertical direction for the layers with sufficient measurements



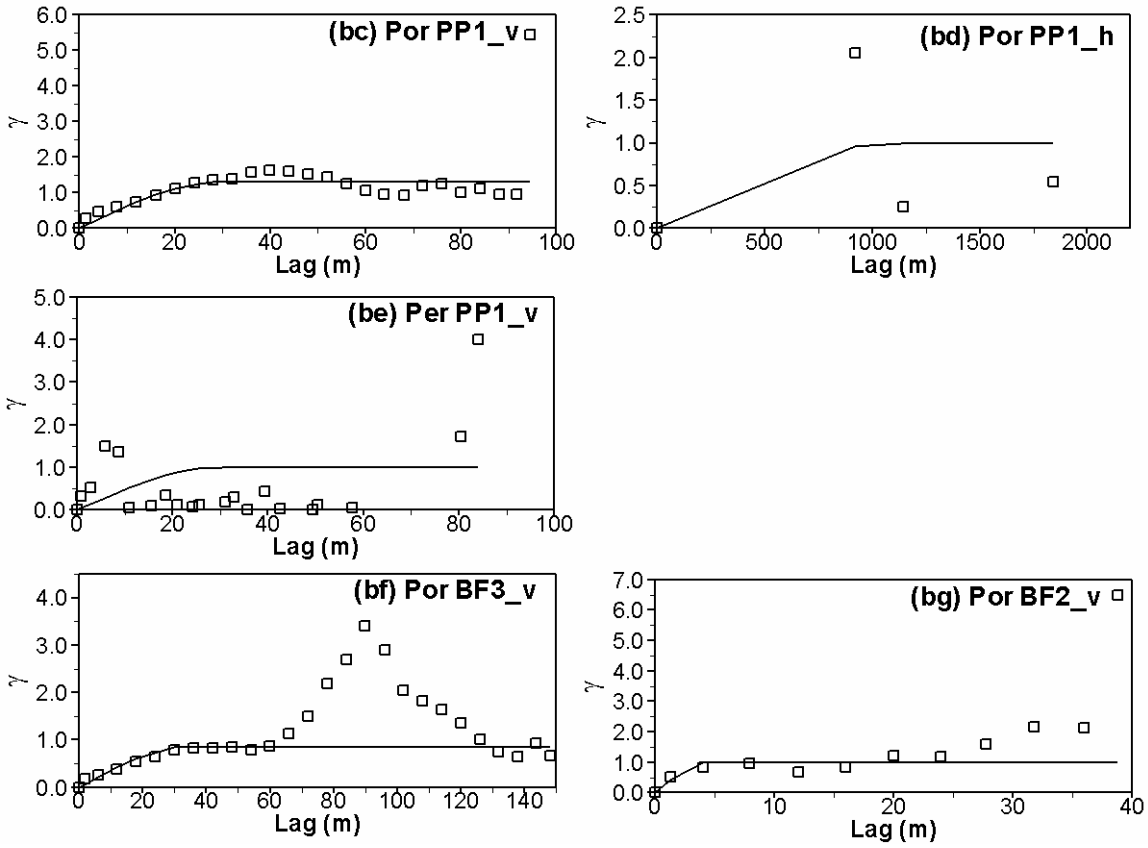
Source: The original measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Data come from DID: 016FP.003; see word file "variogram.doc".

Figure 11. (Cont.) Sample and fitted variograms of transformed measurements of matrix porosity in vertical and horizontal directions and matrix permeability in vertical direction for the layers with sufficient measurements



Source: The original measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Data come from DID: 016FP.003; see word file "variogram.doc".

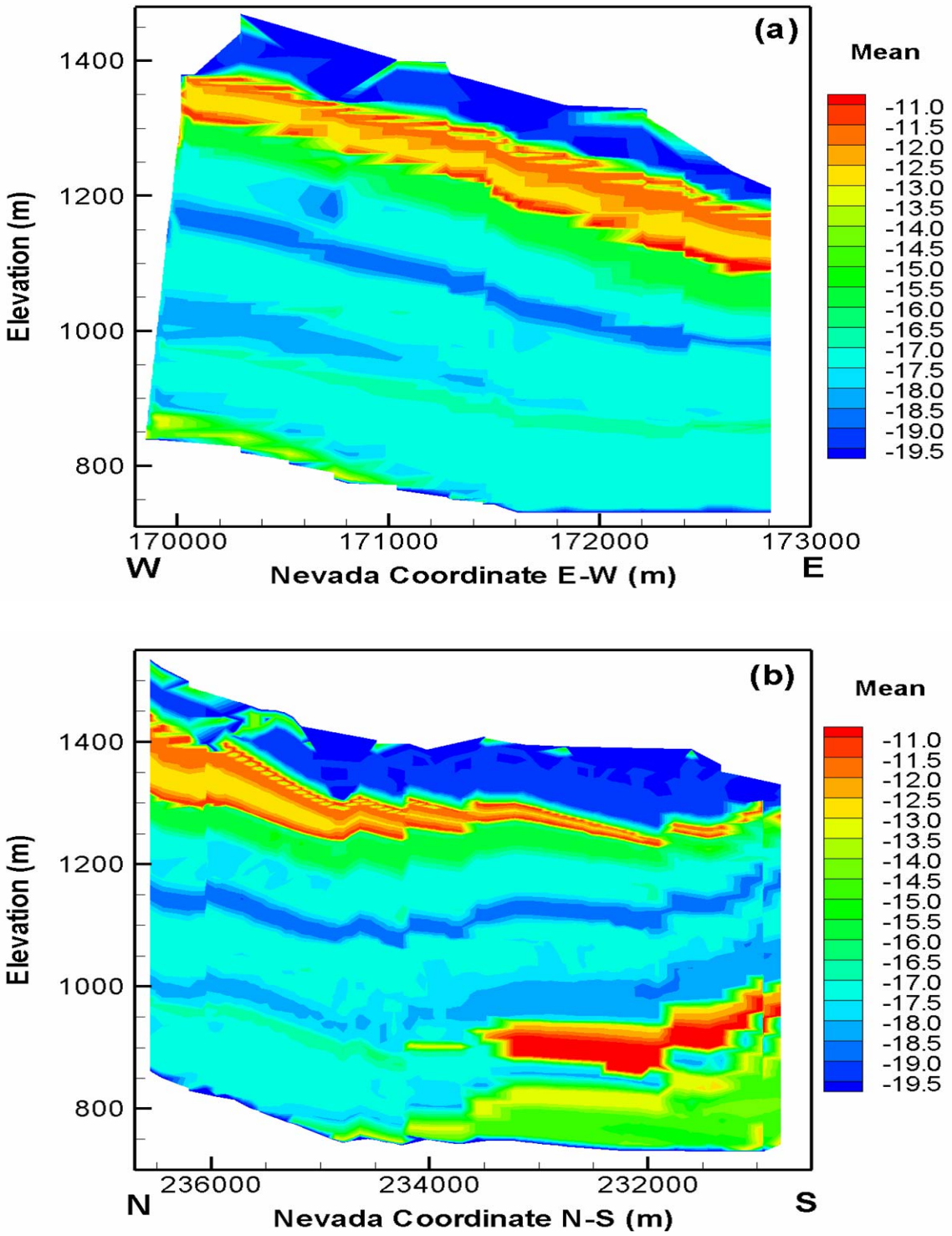
Figure 11. (Cont.) Sample and fitted variograms of transformed measurements of matrix porosity in vertical and horizontal directions and matrix permeability in vertical direction for the layers with sufficient measurements



Source: The original measurements come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Data come from DID: 016FP.003; see word file "variogram.doc".

Figure 11. (Cont.) Sample and fitted variograms of transformed measurements of matrix porosity in vertical and horizontal directions and matrix permeability in vertical direction for the layers with sufficient measurements

east-west (Figure 12a) and north-south (Figure 12b) cross sections through borehole UZ-14, located in the proposed repository area. The layer-scale heterogeneity is observed, and mean log permeability is significantly different in different layers. At the bottom layers, as shown in Figure 12b, the mean log permeability in the north part of the domain is significantly smaller than that in the south. The reason is that the CHn-unit zeolitic tuffs (with low permeability) are located in the north, while the vitric tuffs (with high permeability) are located in the south. Figure 12 also illustrates the local-scale heterogeneity of the mean log permeability within each layer. Sample variance of the generated realizations is calculated to evaluate spatial variability (figures not shown). Variance of log permeability varies significantly, from 0.5 to 8.0 in different layers, depending on density of measurements in each layer. In general, the variance is smaller for thinner layers with more measurements. Porosity spatial variability is similar to that of the log permeability, but with a smaller magnitude of variation (figures not shown).



Source: Data come from DID: 016FP.003; see word file "CrossSection_RandomField.doc".

Figure 12. Mean of generated random log permeability at east-west (a) and north-south (b) cross section through borehole UZ-14

7.3 Uncertainty Analysis of Unsaturated Flow

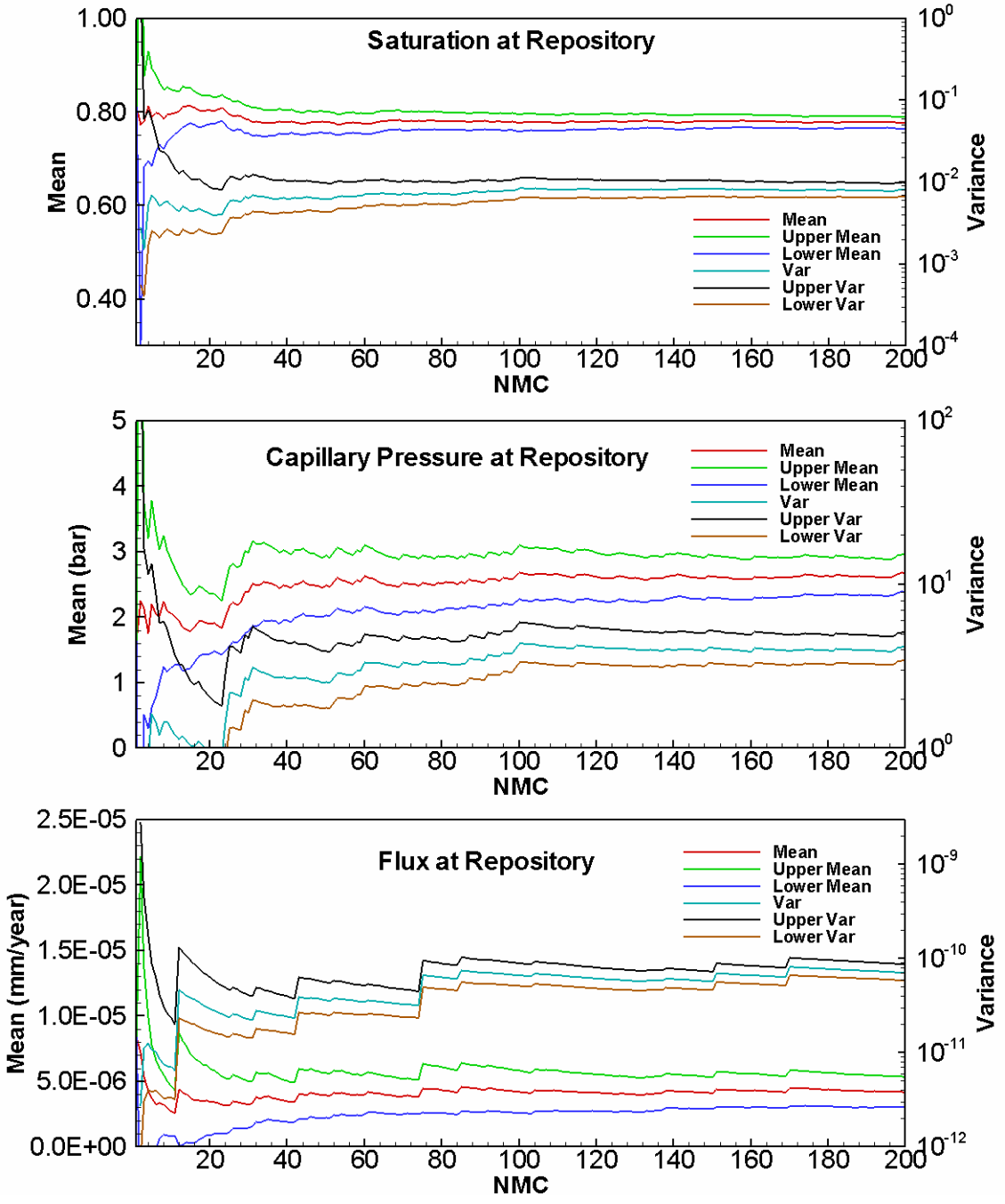
200-realization Monte Carlo simulations are conducted to investigate propagation of uncertainty of matrix permeability and porosity in the complicated numerical model of unsaturated flow. Other variables are treated deterministically, and their values are adopted from Wu et al. (2004a, b). Convergence of the Monte Carlo simulations is investigated by examining stabilization of mean and variance of simulated state variables. Mean, variance, and 5^{th}, 50^{th}, and 95^{th} percentiles of simulated state variables (e.g., saturation, water potential, and percolation flux) are evaluated based on the 200 realizations.

In this section, we first demonstrate in Section 7.3.1 convergence of the Monte Carlo simulation for the layer-scale parameters. Similar investigation is also conducted for the local-scale parameters, but results are not shown. Simulated flow and corresponding uncertainty analysis are presented in Sections 7.3.2 and 7.3.3 for the homogeneous and heterogeneous cases, respectively. Comparison of the uncertainty assessment for the homogeneous and heterogeneous cases is discussed in Section 7.3.4.

7.3.1 Convergence of Monte Carlo Flow Simulations

As there are no well-established convergence criteria, the convergence analysis is a major concern in Monte Carlo simulation to ensure that the sample statistics (e.g., mean and variance) obtained from multiple realizations are the ensemble ones. In this study, a recently developed convergence method by Ballio and Guadagnini (2004) is implemented to examine the convergence of groundwater flow simulations.

The saturation, capillary pressure, and vertical flux at the elements of Repository layer (TLL) are selected to check the convergence of Monte Carlo Simulation. Because the radionuclide would be released from Repository layer, the variables at Repository layer play important roles on overall repository performance. Figure 13 plots the sample means and variances of the three variables with 95% confidence intervals of simulated saturation, capillary pressure and vertical flux at repository layer for 200 realizations. Figure 13 shows that the sample means and variances of saturation at repository layer stabilized after 100 realizations. The 95% confidence intervals decrease with the increase of the realizations, but only vary at negligible level after 150 realizations, indicating convergence of our Monte Carlo simulations. Therefore, sample statistics obtained from the 200 realizations are considered the same as ensemble ones and used to present our optimum predictions and associated predictive uncertainty.



Source: Data come from DID: 016FP.002; see word file "ConvergenceAnalysis.doc".

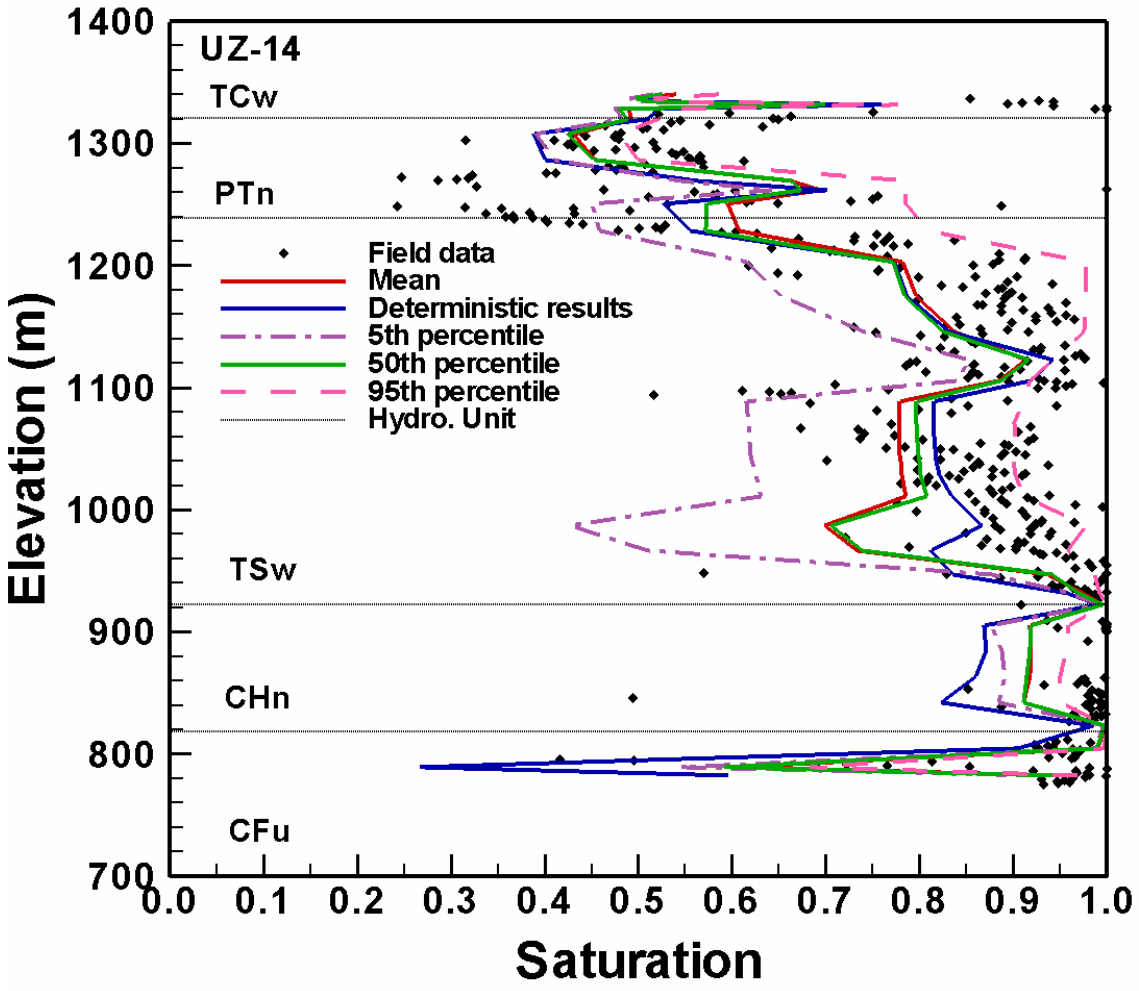
Figure 13. Sample mean and variance of simulated matrix satruation, capillary pressure, and vertical flux with 95% confidence interval at repository layer (TLL) for homogeneous case

7.3.2 Simulated Unsaturated Flow and Uncertainty Analysis for the Homogeneous Case

Based on results of the Monte Carlo simulation, mean, variance, and 5^{th}, 50^{th}, and 95^{th} percentiles of simulated state variables (e.g., saturation, water potential, and percolation flux) are evaluated. As a common practice of uncertainty assessment, mean is used as an optimum prediction and variances as a measure of associated predictive uncertainty. Assuming that simulated variables follow normal distributions, a 95% confidence interval can be calculated based on the mean and variance to quantify the predictive uncertainty. However, since the predicted quantities may not follow normal distribution, we used the 5^{th} and 95^{th} percentiles (also known as uncertainty bounds) to quantify predictive uncertainty. Therefore, for all figures presented below, in addition to the mean and variance, the 5^{th} and 95^{th} percentiles are also plotted. The deterministic simulation results of Wu et al. (2004a, b) are treated in this study as a baseline case for stochastic simulations. Note that only layer-scale heterogeneity was considered in the deterministic simulation. After comparing statistics of the two variables (saturation, and water potential) with corresponding site measurements, the spatial flow pattern of percolation flux and associated predictive uncertainty is assessed.

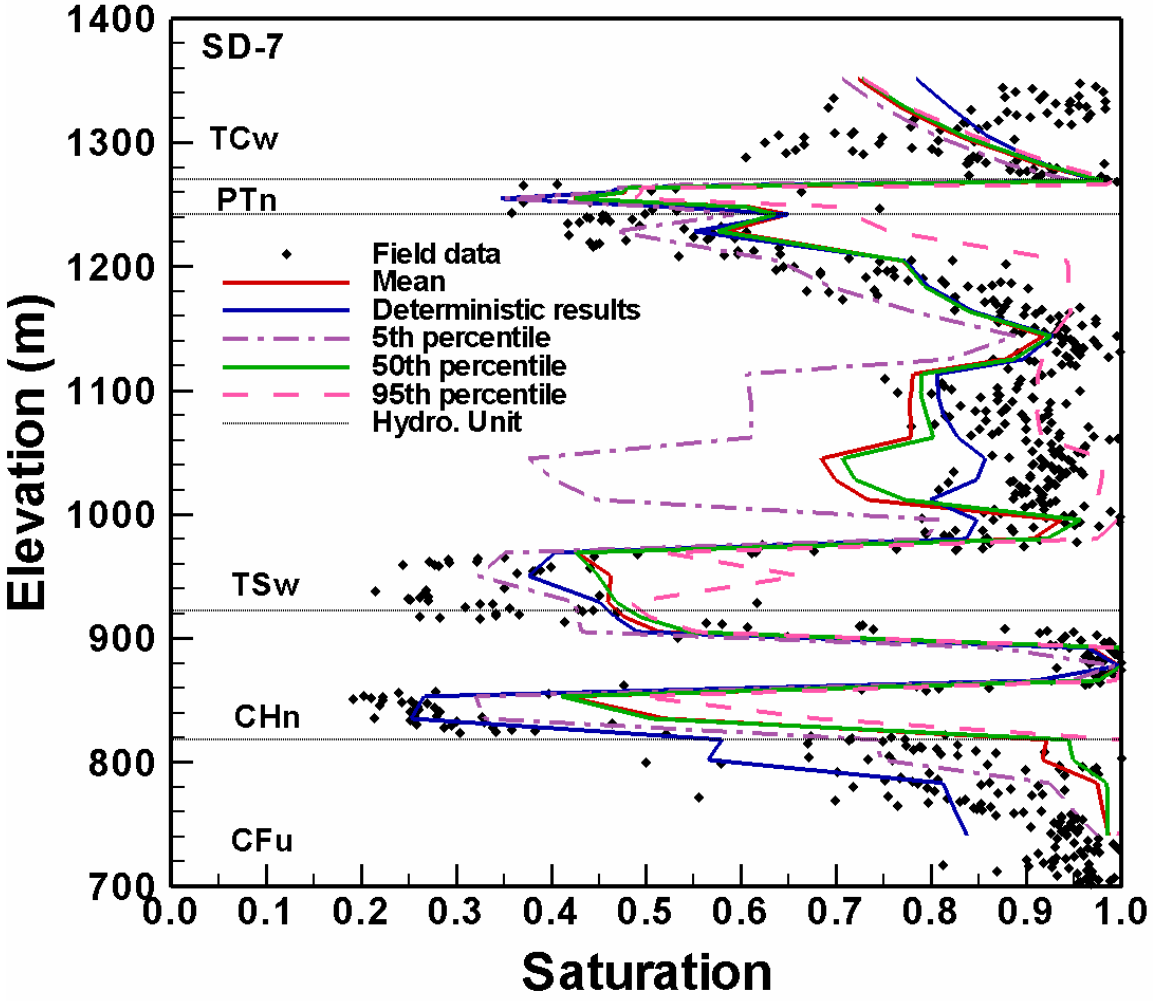
7.3.2.1 Comparisons of Simulated and Measured Data

Figures 14, 15 and 16 compares the observed and 3-D simulated matrix saturation along the vertical column of boreholes UZ-14, SD-12 and SD-7. Figure 17 does the same for water potential for borehole SD-12. In each figure, the mean and 50th-percentile predictions are close but not identical, suggesting that simulated variables may not follow normal distributions. Nonstationary behavior of matrix liquid saturation and water potential are observed in Figures 14, 15, 16, and 17, which can be simulated with nonstationary stochastic moment equations approaches (e.g., Hu et al., 2004; Wu and Hu, 2004; Zhang and Winter, 1998; Zhang, 1999). The mean and 50^{th}-percentile predictions deviate from the corresponding results of the deterministic case. This is not surprising because, as discussed in Section 7.1.5, the mean model parameters used for the stochastic analyses are different from the model inputs for the deterministic case. The 5^{th} and 95^{th} percentiles (also known as uncertainty bounds) of simulated matrix liquid saturation and water potential bracket a significant number of measurements, indicting the data variability can be partially explained by parametric uncertainty in the matrix permeability and porosity. In particular, certain measurements that cannot be caught by deterministic simulation (e.g., matrix liquid saturation at the bottom of TSw) are included in the uncertainty bounds. This is particularly true for the comparison of water potential shown in Figure 17. Nevertheless, the results of the deterministic case match the trend for measurements to be better at the bottom of hydrogeologic units TSw and CHn than at the mean and 50^{th} percentile in the stochastic predictions, suggesting that more calibration information should be included in stochastic simulations. Unbracketed measurements can be attributed to measurement error, conceptual model incompleteness, and different scales between model inputs and field and laboratory parameter measurements.



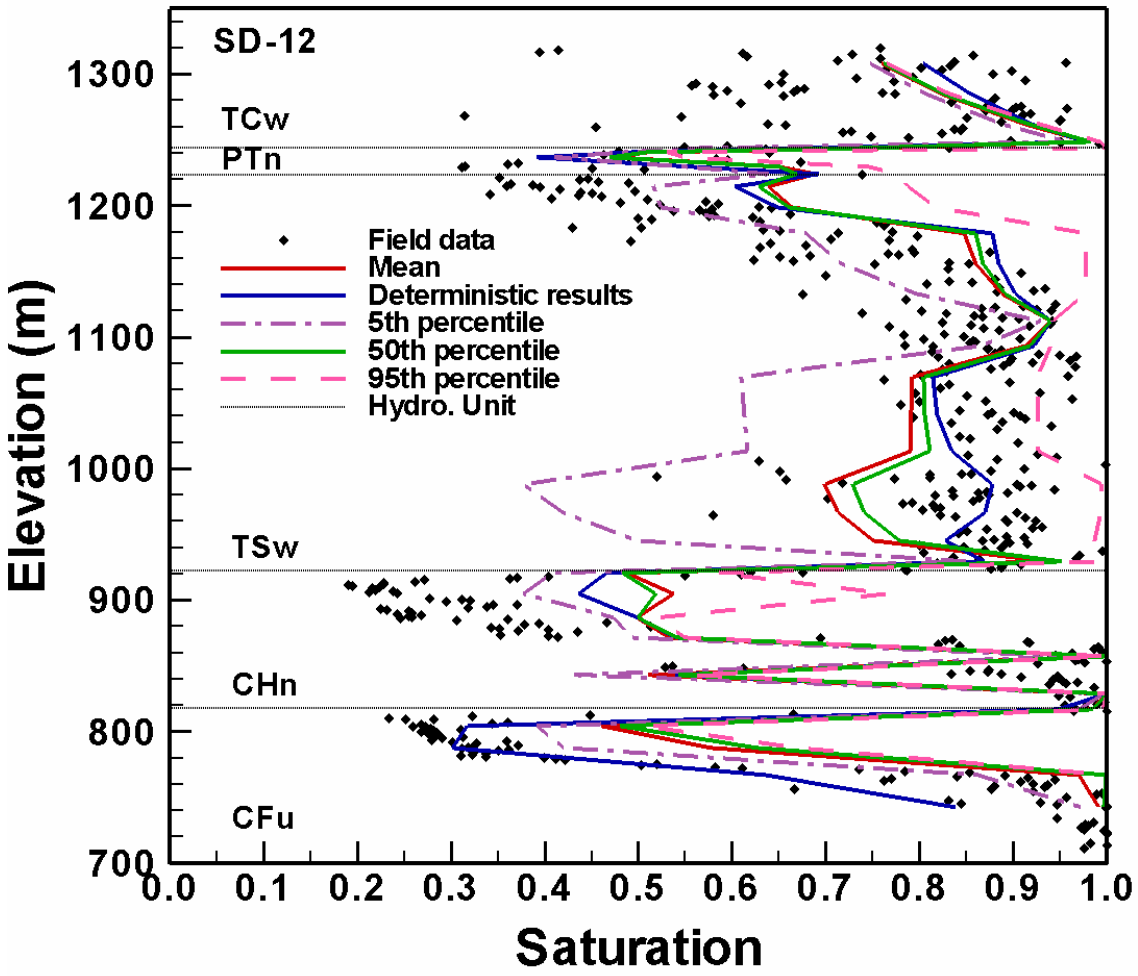
Source: Field data come from DTN: LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_homo.doc".

Figure 14. Comparison of observed and 3-D modle simulated matrix liquid saturation in borehole UZ-14 for homogeneous case



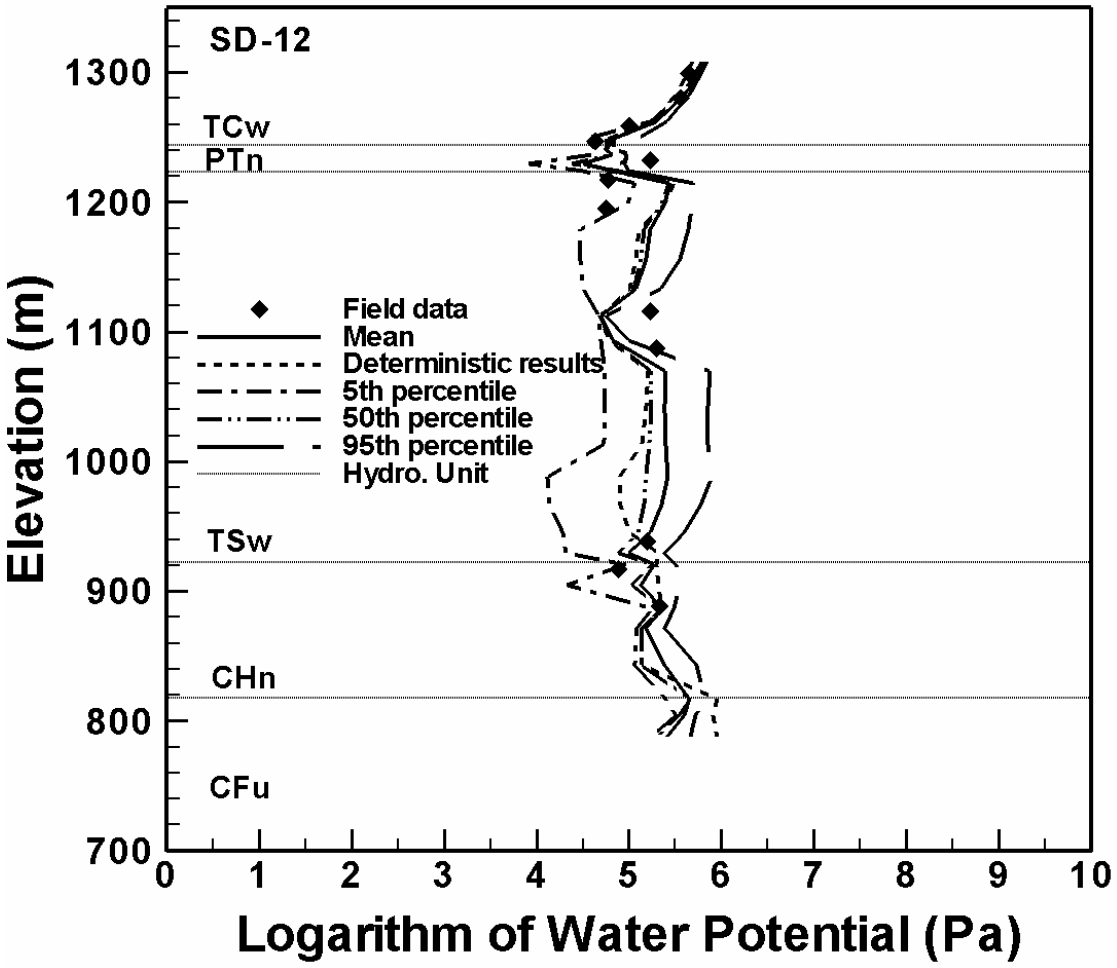
Source: Field data come from LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni..out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_homo.doc".

Figure 15. Comparison of observed and 3-D modle simulated matrix liquid saturation in borehole SD-7 for homogeneous case



Source: Field data come from LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_homo.doc".

Figure 16. Comparison of observed and 3-D modle simulated matrix liquid saturation in borehole SD-12 for homogeneous case



Source: Field data come from GS031208312232.003, file "waterpot.txt"; and GS970808312232.005, file "zz_sep_260947.txt "; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_homo.doc".

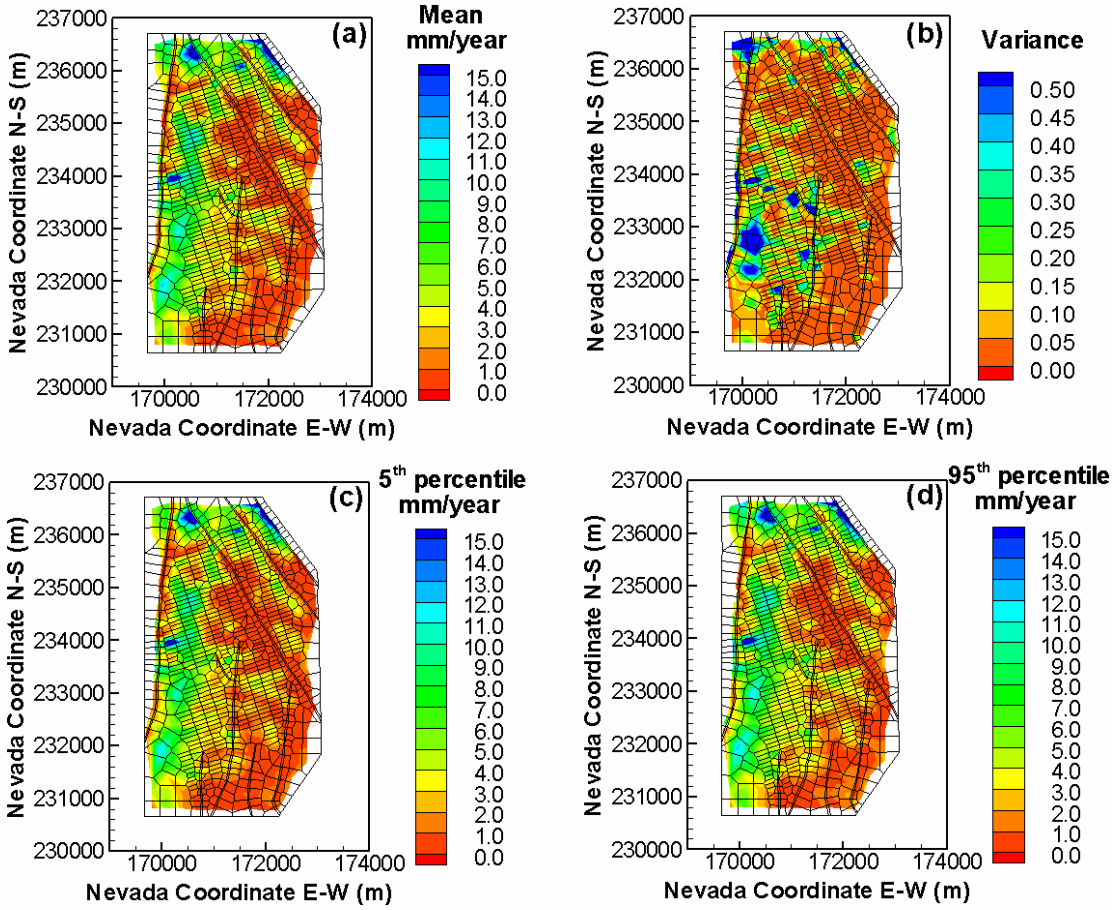
Figure 17. Comparison of observed and 3-D model simulated matrix water potential in borehole SD-12 for homogeneous case

7.3.2.2 Uncertainty Analysis of Unsaturated Flow

Percolation flux through the UZ is a key variable in evaluating the potential repository site, because percolation flux and its spatial variations could affect the amount of water flowing into waste emplacement drifts, radionuclide released from the repository, and radionuclide migration from the UZ to the groundwater table. The percolation flux is defined as the total vertical liquid mass flux through both fracture and matrix (Wu et al., 2004a, b), converted to millimeters per year using a constant water density for better presentation. Since the lateral boundaries of the model domain are impermeable, the percolation flux is driven by the surface infiltration shown in Figure 3. In this task, the predictions and uncertainties of percolation flux for each layer are calculated and quantified

based on the simulated unsaturated flow with 200 realizations. We select two layers, Repository layer and Water Table, to present the predictions and uncertainties here.
The source of percolation flux through the UZ is net infiltration from precipitation at the land surface at Yucca Mountain. The infiltration pattern is shown in Figure 3 and the average rate of net infiltration over the entire model domain is 3.583 mm/year.

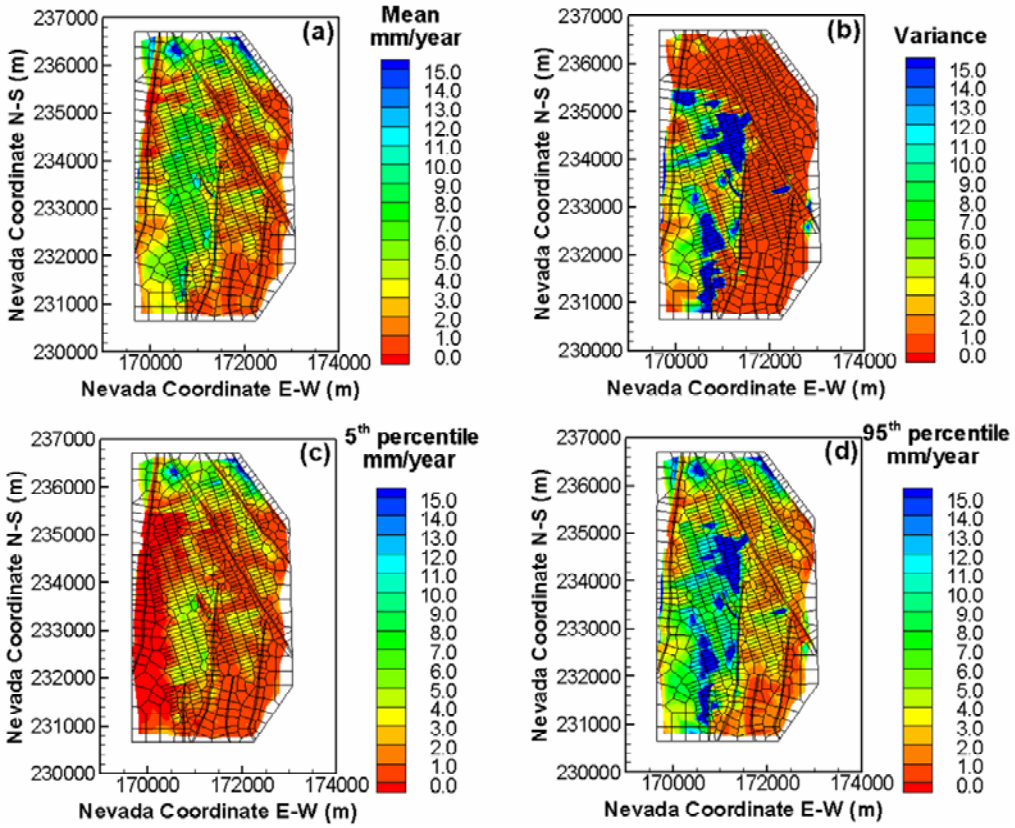
Figure 18 plots (a) mean, (b) variance, (c) 5th percentile, and (d) 95th percentile in the simulated percolation flux at the repository horizon. Figure 19 does the same at the water table. The mean percolation flux at the repository horizon (Figure 18a) is similar to the surface infiltration (Figure 3), indicating a small lateral movement of infiltrated water during the process of water traveling from the surface to the repository level. However, a comparison of Figures 19a and 3 shows that the high-infiltration zone (denoted by dark "green" areas) moves eastward, indicating a significant lateral movement of infiltrated water.



Source: The mesh file used for flow simulation comes from LB0303THERMESH.001, file "mesh_th.vf"; Simulated data come from DID: 016FP.002; see word file "Flux_homo.doc".

Figure 18. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of simulated percolation fluxes at the repository horizon for homogeneous case

Variance in percolation flux at the proposed repository horizon (Figure 18b) is small. The largest variance occurs in the high-infiltration zone and at its eastern edge where the eastward movement occurs. Consistent with the small variance, the difference between the 5th and 95th percentiles (Figures 18c and 18d) of the percolation flux is small, suggesting small predictive uncertainty in percolation flux caused by uncertainty in permeability. However, as shown in Figure 19b, predictive uncertainty in percolation flux at the water table is large. The largest uncertainty also occurs in the high-infiltration zone and at its edge. The 5th and 95th percentiles (Figures 19c and 19d) of percolation flux at the water table differ significantly, and this difference explains the variance distribution in Figure 19b. For example, the two zones of large variance in Figure 19b are attributed to the two zones (marked in blue) of large percolation flux appearing at the middle and bottom of the simulation domain in the 95th-percentile map but not the 5th-percentile map. Although the variance in percolation flux is significantly larger at the water table than at the repository horizon, Figure 10b shows that uncertainty in permeability near the water table is not significantly larger than near the repository zone. The large variance of percolation flux at the water table may be attributed to the accumulation of predictive uncertainty from the domain surface to its bottom, since infiltration rate is given as a deterministic condition at the domain surface.



Source: The mesh file used for flow simulation comes from LB0303THERMESH.001, file "mesh_th.vf"; Simulated data come from DID: 016FP.002; see word file "Flux_homo.doc".

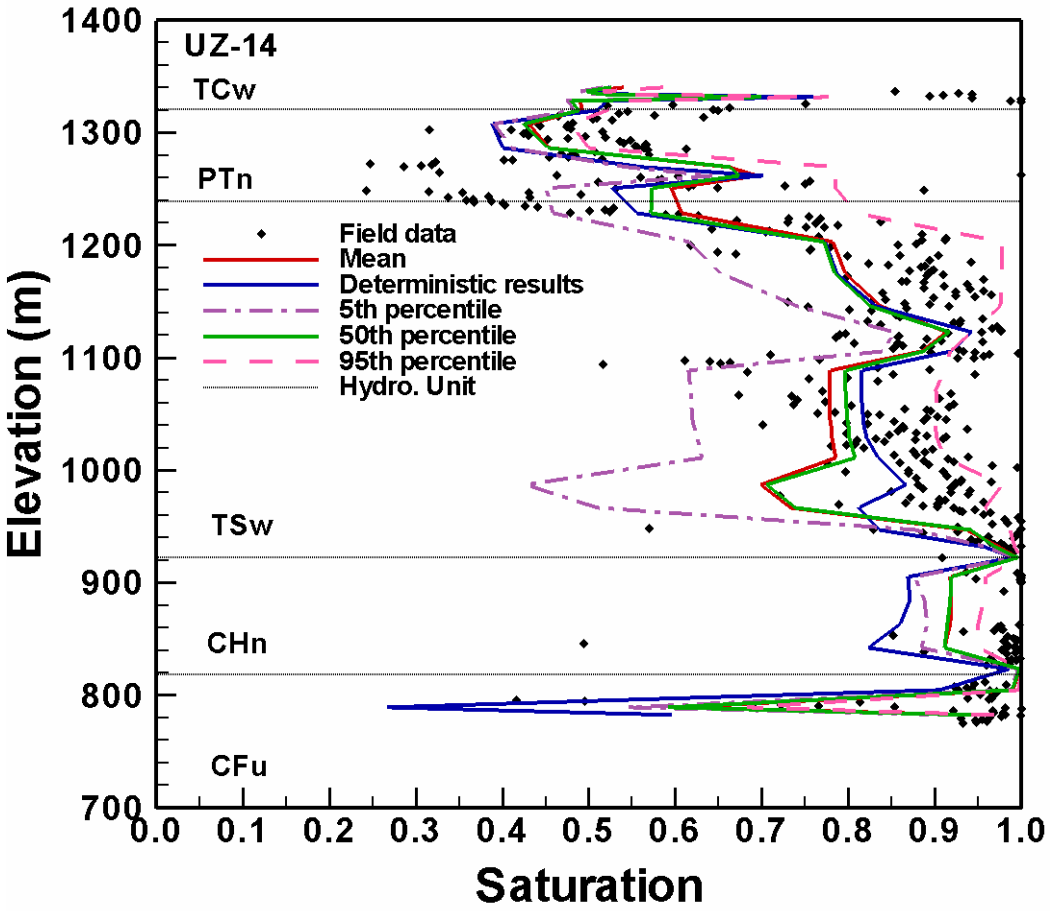
Figure 19. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of simulated percolation fluxes at the water table for homogeneous case

7.3.3 Simulated Unsaturated Flow and Uncertainty Analysis for the Heterogeneous Case

Similar to Section 7.3.2, we first compare the simulated and measured state variables, and then analyze uncertainty of the simulated unsaturated flow.

7.3.3.1 Comparisons of Simulated and Measured Data

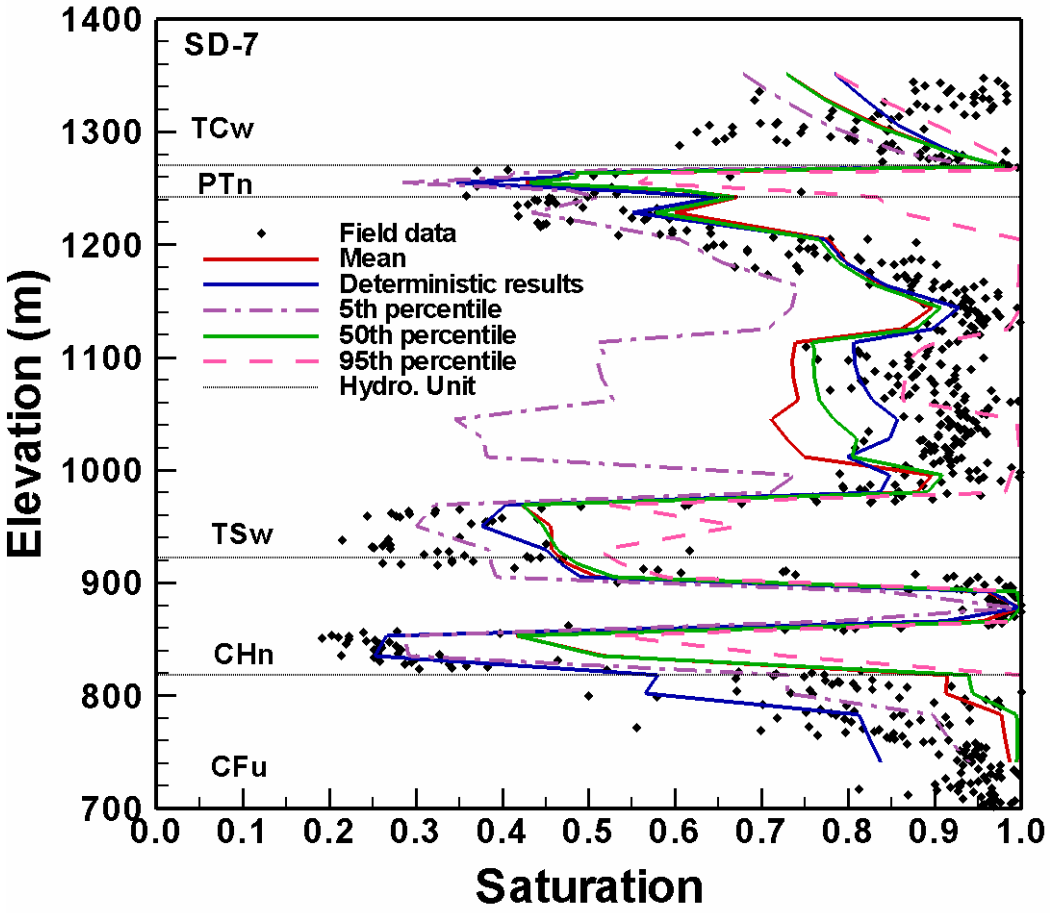
Figures 20-22 compare the observed and simulated matrix water saturations along boreholes UZ-14, SD-7, and SD-12. Figure 23 does the same for water potential for borehole SD-12. Nonstationary behavior of the observations is simulated, since the nonstationarity of model parameters is taken into account by generating random fields for each layer separately. In these figures, the means (as well as the 50th percentile) of simulated saturation and water potential are close to the corresponding results



Source: Field data come from LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_hete.doc".

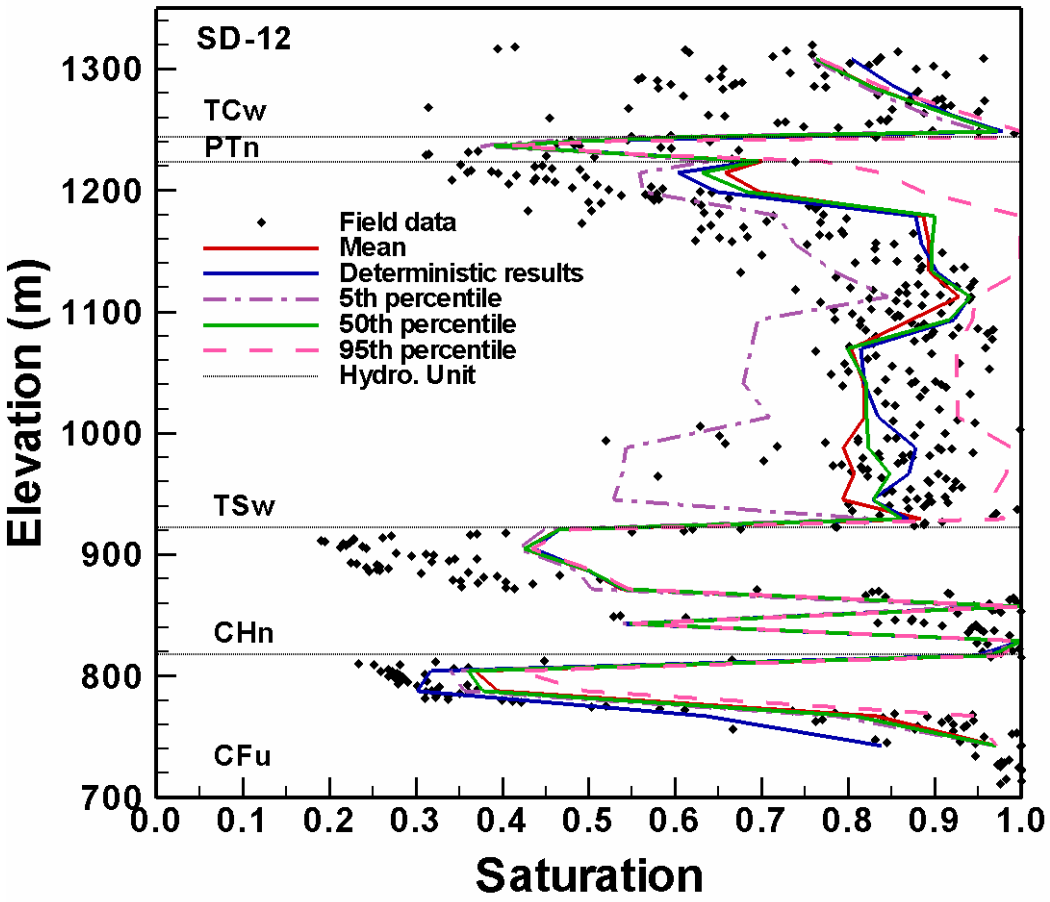
Figure 20. Comparison of observed and 3-D model simulated matrix liquid saturation in borehole UZ-14 for heterogeneous case

for the deterministic case (Wu et al., 2004a, b), indicating that the layer-scale heterogeneity of model parameters dominates over the local-scale heterogeneity in simulating the mean behavior of the unsaturated flow system. The means of simulated matrix liquid saturation and water potential are in reasonable agreement with the observed profiles, and the modeling results catch the patterns of variation. The 5th and 95th percentiles of simulated results bracket a large portion of the observations, indicting that data variability can be partially explained by parametric uncertainty in the matrix permeability and porosity. Unbracketed measurements can be attributed to measurement error, conceptual model incompleteness, and different scales between model inputs and field and laboratory parameter measurements. In Figure 23, almost all measurements of water potential are included in the uncertainty bound.



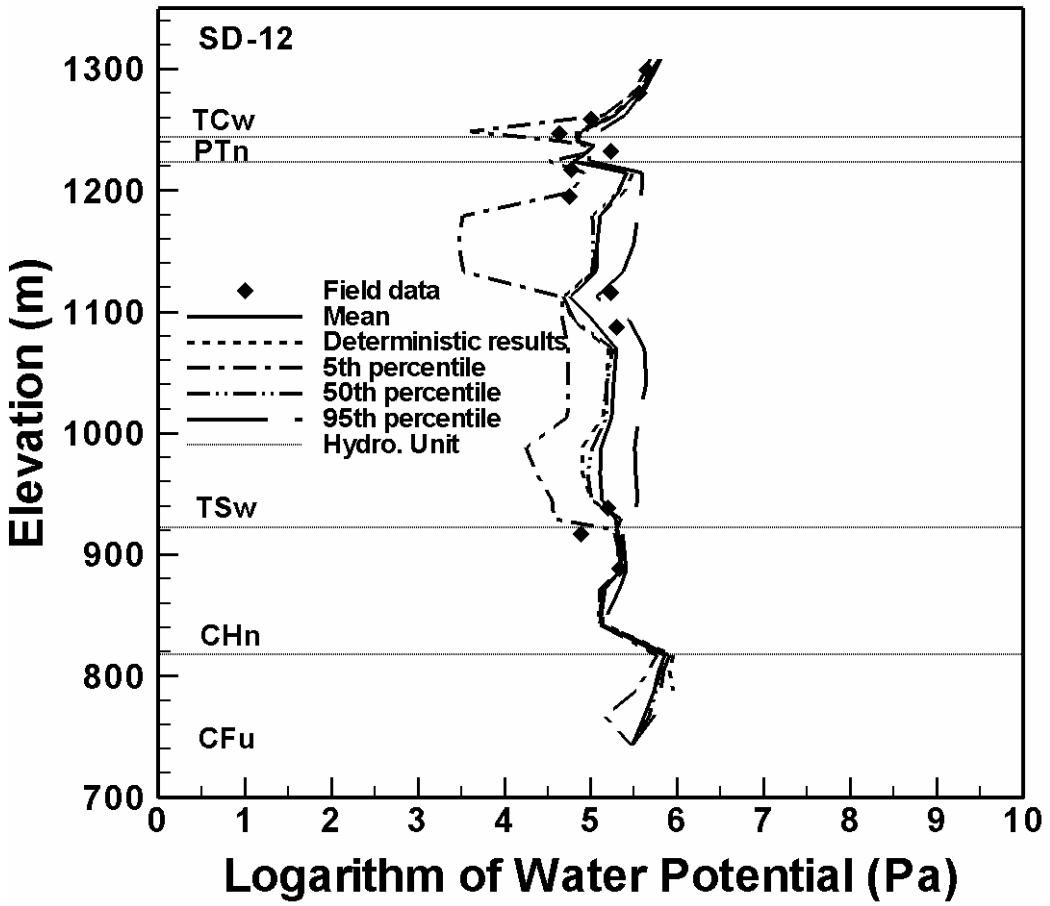
Source: Field data come from LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_hete.doc".

Figure 21. Comparison of observed and 3-D model simulated matrix liquid saturation in borehole SD-7 for heterogeneous case



Source: Field data come from LB0207REVUZPRP.002, file "hydroprops_fin.xls"; Deterministic results come from DTN: LB0303THERMSIM.001, , file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_hete.doc".

Figure 22. Comparison of observed and 3-D model simulated matrix liquid saturation in borehole SD-12 for heterogeneous case



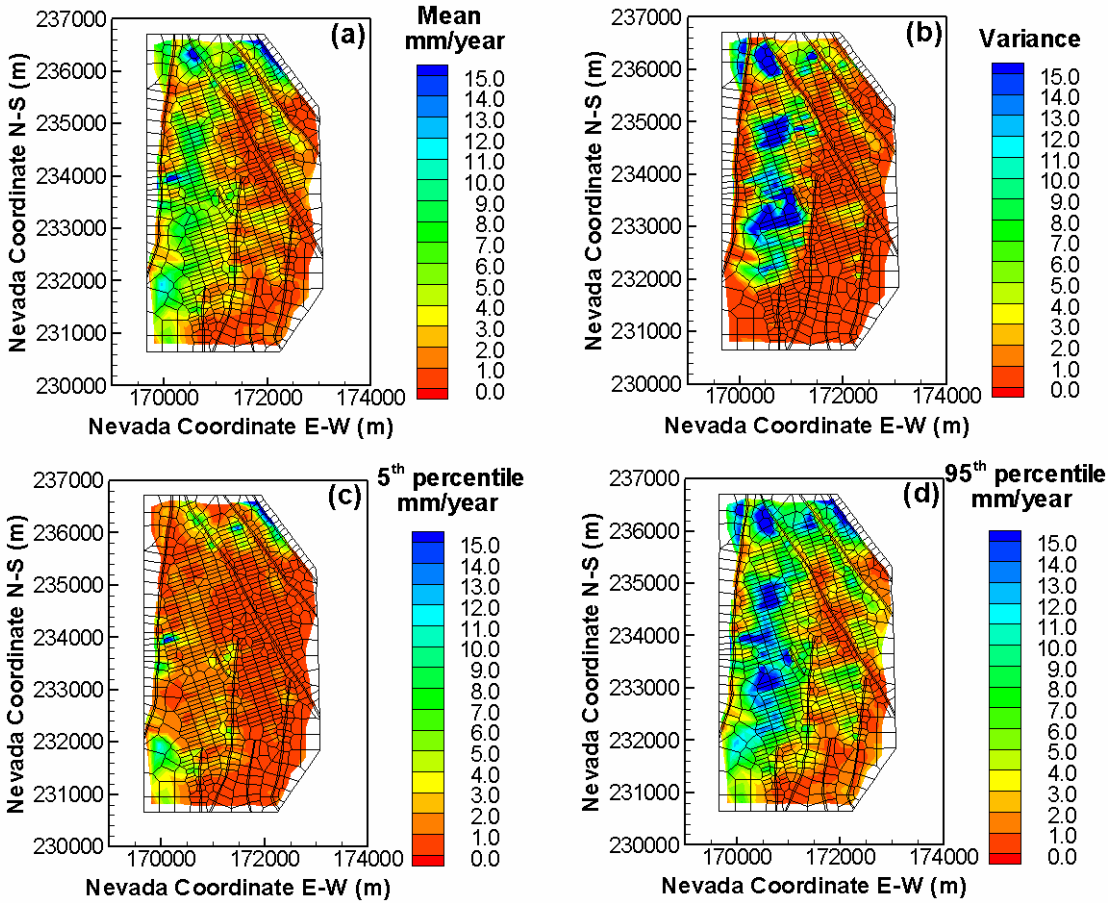
Source: Field data come from GS031208312232.003, file "waterpot.txt"; and GS970808312232.005, file "zz_sep_260947.txt"; Deterministic results come from DTN: LB0303THERMSIM.001, file "th_pqm_keni.out_2_20_03"; Simulated data come from DID: 016FP.002; see word file "BoreholeComparison_hete.doc".

Figure 23. Comparison of observed and 3-D model simulated matrix water potential in borehole SD-12 for heterogeneous case

7.3.3.2 Uncertainty Analysis of Unsaturated Flow Fields

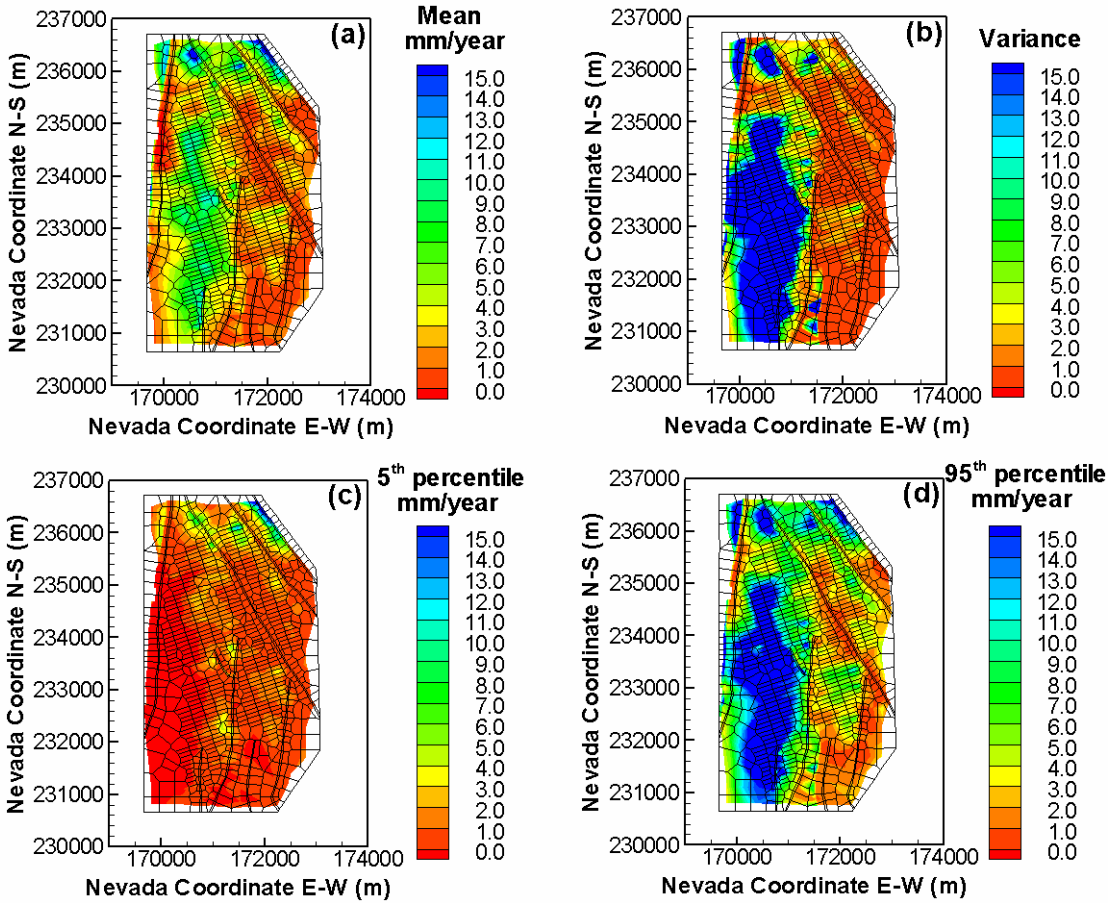
Figures 24 and 25 depict mean, variance, and the 5th and 95th percentiles of simulated percolation fluxes at the proposed repository horizon and water table, respectively. The pattern of mean percolation fluxes at the repository layer (Figure 24a) is similar to the surface infiltration pattern shown in Figure 3, indicating dominant vertical flow and negligible lateral movement from the land surface to the repository level. However, Figure 25a shows that, at the water table, the high percolation flux zone (denoted by dark "green" areas) shown west of the land surface (Figure 3) moves eastward, indicating significant lateral flow from the repository level to the water table. This is mainly attributed to dipping slope (around 5 to 10 degree) shown in Figure 12a and presence of CHn unit between the proposed repository and the water table shown in Figure 12b. Due to the high-

permeable CHn vitric tuffs (shown at south of Figure 12b), infiltration water moves toward east to the Ghost Dance Fault (Figure 1). Variance in the simulated percolation fluxes at the repository level (Figure 24b) is larger in the west of the model domain, associated with the high infiltration rate shown in Figure 3. In comparison with 24b, Figure 25b shows that, at the water table, a large variance also occurs at the west side of the domain, but covers a wider area that extends southward. This may be attributed to the larger spatial variation of matrix permeability at the bottom than at the top of the simulation domain (Figure 12b) and the accumulated effect of parameter uncertainty propagation downward to the water table. In Figures 24 and 25, the 5th and 95th percentiles of percolation fluxes (Figures c and d) are significantly different, indicating significant uncertainty regarding percolation flux, caused by the parametric uncertainty of matrix permeability.



Source: The mesh file used for flow simulation comes from LB0303THERMESH.001, file "mesh_th.vf"; Simulated data come from DID: 016FP.002; see word file "Flux_hete.doc".

Figure 24. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of simulated percolation fluxes at the repository horizon for heterogeneous case



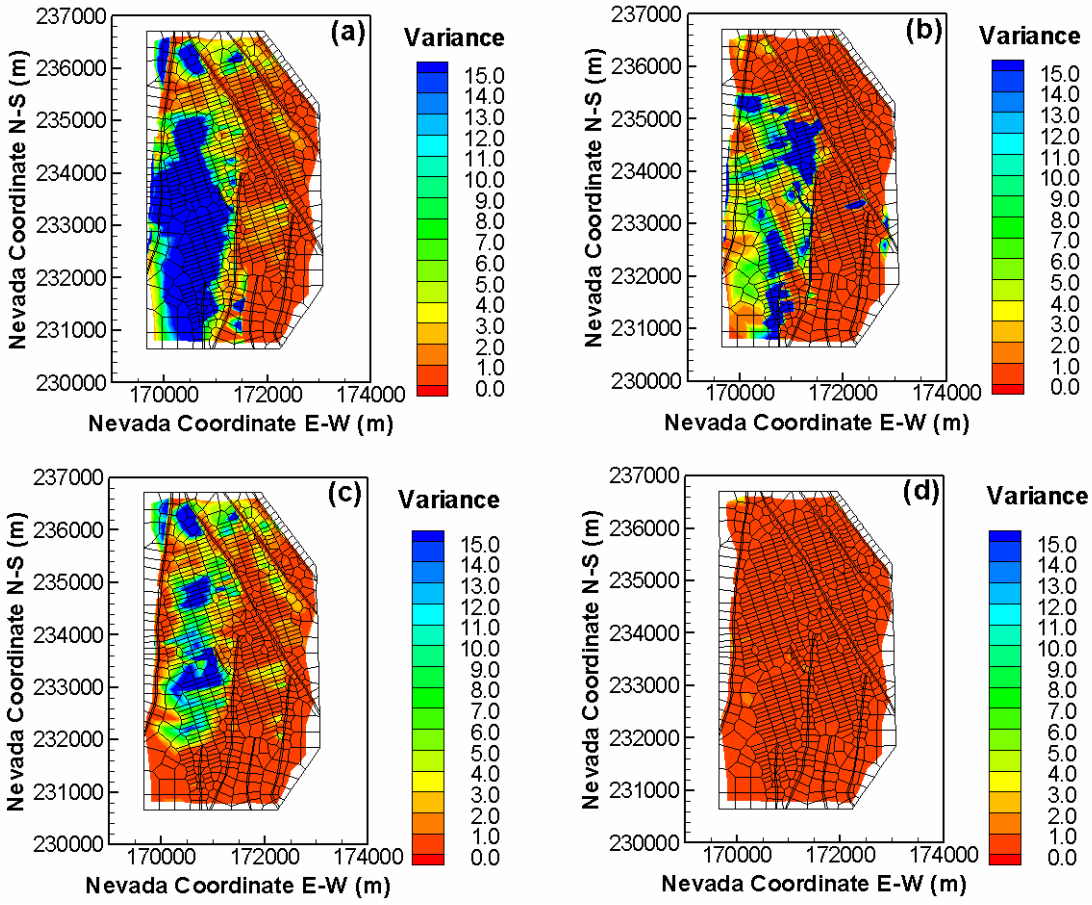
Source: The mesh file used for flow simulation comes from LB0303THERMESH.001, file "mesh_th.vf"; Simulated data come from DID: 016FP.002; see word file "Flux_hete.doc".

Figure 25. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of simulated percolation fluxes at the water table for heterogeneous case

7.3.4 Comparison of Uncertainty Flow Assessment in the Homogeneous and Heterogeneous Cases

Predictions of percolation flux are also compared for the homogeneous case and heterogeneous case. Comparing Figures 19 and 24, we observe that the mean predictions have a similar pattern and magnitude, whereas the variance in the heterogeneous case (Figure 24) is significantly larger than that of the homogeneous case (Figure 19), especially under the footprint of the repository area. This indicates that the local-scale heterogeneity of matrix permeability can result in more uncertainty for predicting percolation flux. Similarly, the 5th (95th) percentile of the heterogeneous case is smaller (larger) than that of the homogeneous case. These differences indicate that the local-scale heterogeneity creates flow paths that are more random and complicates the unsaturated flow in and between the matrix and fracture. This is also true for the percolation flux at the repository horizon, except that differences in variance for the heterogeneous and homogeneous cases are smaller.

Figure 26 plots the variances of percolation flux at the water table and repository horizon for heterogeneous and homogeneous (Figures 26b and d) cases. The variance at the water table and repository layer (Figure 26a and c) in the heterogeneous case is significantly larger than those of the homogeneous case (Figure 26b and d), especially under footprint of the repository area, indicating that the local-scale heterogeneous of matrix permeability can result in more uncertainty of predicting percolation flux, especially at the water table. This is not surprising, since the local-scale heterogeneity renders flow path more random and complicates the unsaturated flow in and between the matrix and fracture.



Source: The mesh file used for flow simulation comes from LB0303THERMESH.001, file "mesh_th.vf"; Simulated data come from DID: 016FP.002; see word file "Flux_hete.doc".

Figure 26. Variance of percolation flux (a) heterogeneous case at the water table, (b) homogeneous case at the water table, (c) heterogeneous case at repository horizon, (d) homogeneous case at repository horizon

7.4 Uncertainty Analysis of Unsaturated Radionuclide Transport

The radionuclide transport uncertainty is assessed with two tracers in this task: conservative (nonadsorbing) tracer, ^{99}Tc, and reactive (adsorbing) tracer, ^{237}Np. The 200 realizations of steady-state flow fields for homogeneous and heterogeneous cases simulated in Section 7.3 are used as the flow fields of the transport simulation in the UZ for 1,000,000 years. Multiple realizations of the sorption coefficient of ^{237}Np generated in Section 7.1.5 are used in the radionuclide simulations for the both homogeneous and heterogeneous cases. The other geochemical parameters (e.g., molecular diffusion coefficient, effective diffusion coefficient, and mechanical dispersion) for ^{99}Tc and ^{237}Np are deterministic, and their values are adopted from Wu et al. (2004a, b).

200 realizations of cumulative fractional mass breakthrough and tracer travel time are simulated, and transport uncertainty is evaluated based on the 200 realizations. In this task, the cumulative fractional mass breakthrough is defined as the cumulative mass of a tracer arriving at the water table over the entire bottom model boundary over time normalized by the total mass of the tracers released from the repository. The tracer travel time is the cumulative time of the tracers from repository to water table and is calculated from the cumulative fractional breakthrough curve.

In this section, uncertainty of the cumulative fractional mass and tracer travel time is discussed in Sections 7.4.1 and 7.4.2 for the homogeneous and heterogeneous cases, respectively. Comparison of the uncertainty for the two cases is given in Section 7.4.3.

7.4.1 Uncertainty Analysis of Transport Simulation for the Homogeneous Case

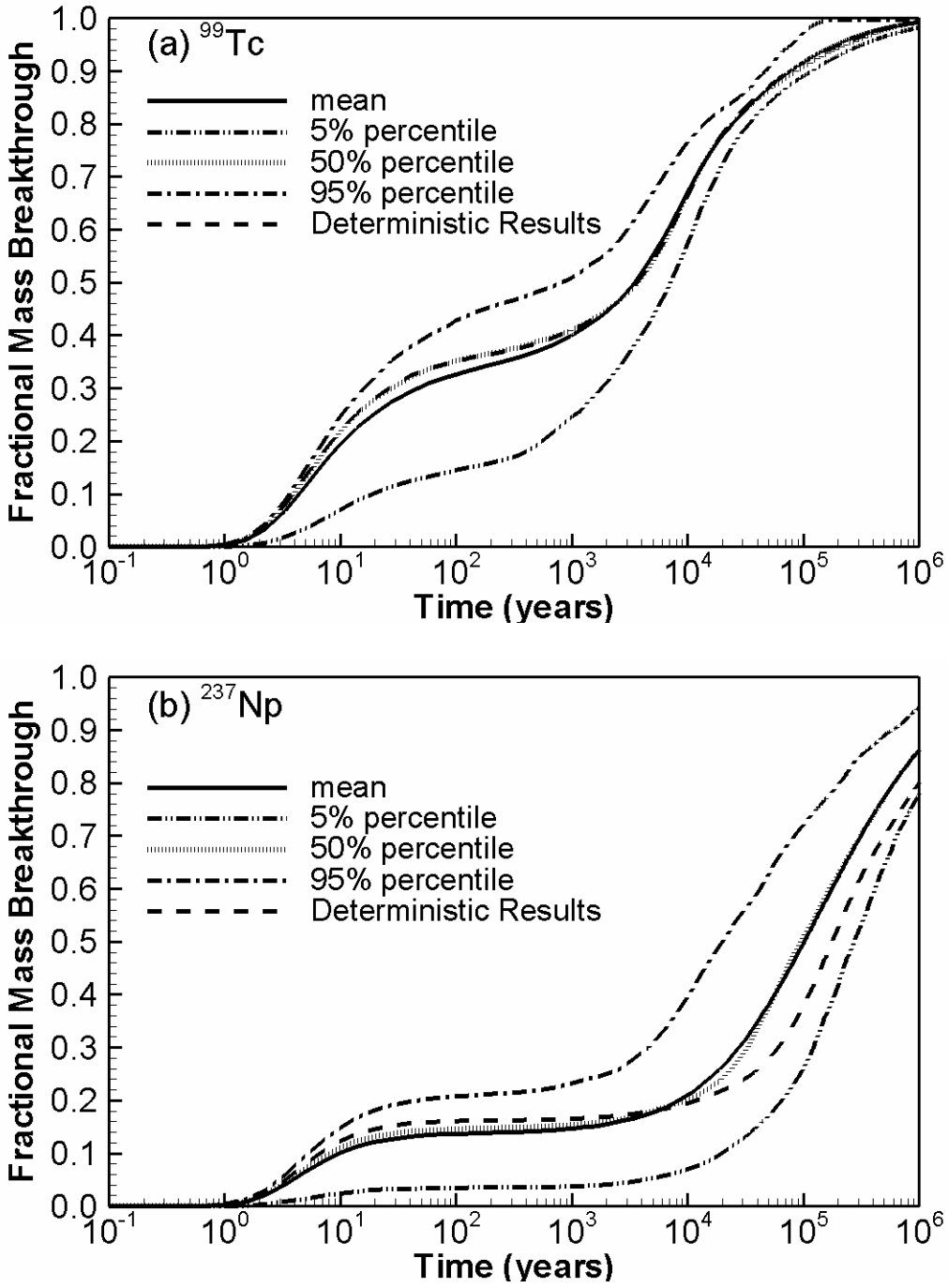
Random parameters generated for the homogeneous case are discussed in Section 7.1, and simulated flow and its uncertainty are discussed in Section 7.3.2.

7.4.1.1 Uncertainty Assessment of Cumulative Travel Time

The tracer travel time from repository to water table is analyzed using the breakthrough curve, which is obtained by calculating the cumulative fractional mass arriving at water table at each time step. The means, 5th, 50th and 95th percentiles of cumulative fractional breakthrough are used to evaluate the uncertainty of radionuclide.

Figures 27a and 27b plot the fractional breakthrough curves of cumulative mass arriving at the water table for the conservative (^{99}Tc) and reactive (^{237}Np) tracers, respectively. The effect of sorption is apparent in that the conservative tracer (^{99}Tc) travels about two orders of magnitude faster than the reactive tracer (^{237}Np). For example, the mean travel times corresponding to 50% mass fraction breakthrough are 3,500 and 100,000 years for ^{99}Tc and ^{237}Np, respectively. For both tracers, the mean and 50th-percentile simulated breakthrough curves are close to the deterministic curve, especially for the conservative tracer, which does not have the sorption coefficient.

The 5th- and 95th-percentiles (uncertainty bounds) breakthrough curves in Figure 27 show that the travel time of the reactive tracer (^{237}Np) is more uncertain than the travel time of the conservative



Source: The input files of deterministic results come from LB03033DUZTRAN.001, file "la_tc.dat", and "la_np.dat"; Simulated data come from DID: 016FP.004; see word file "Transport_homo.doc".

Figure 27. Simulated breakthrough curves of cumulative mass arriving at the water table for (a) the conservative tracer (^{99}Tc) and (b) the reactive tracer (^{237}Np) for homogenoeus case

tracer (^{99}Tc), owing to the uncertain sorption coefficient of ^{237}Np. For example, Figure 27a shows that at 1,000,000 years almost all ^{99}Tc flows out of the UZ into groundwater, while Figure 27b shows that at 1,000,000 years 78% and 94% of the total mass of ^{237}Np flows into groundwater at the 5th- and 95th-percentile levels, respectively. Table 8 lists the travel times of 10%, 25%, 50%, 75%, and 90% for mass fraction breakthrough obtained from the mean, 5th-, and 95th-percentile breakthrough curves. The ranges of travel time between the 5th and 95th percentiles are significantly larger for the reactive tracer than for the conservative tracer. For example, it takes the reactive tracer 2.0×10^4 years (95th percentile) to 2.75×10^5 years (5th percentile) to flow 50% of the mass into the groundwater, whereas it only takes the conservative tracer 8.22×10^2 years (95th percentile) to 7.17×10^3 years (5th percentile). This indicates that the travel time prediction is more uncertain for the reactive tracer than for the conservative tracer. Figure 27a shows that for the conservative tracer the range in uncertainty bounds first increases and then decreases with time. Whereas, Figure 27b shows that for the reactive tracer the travel time uncertainty is of the same magnitude during the entire simulation period of 1,000,000 years. This also is due to the effect of the random sorption coefficient, which retards the travel of the reactive tracer and renders the corresponding travel time prediction more uncertain. This information can be used directly for risk analysis and monitoring network design. For example, monitoring radionuclide transport needs to last longer for the reactive tracer than for the conservative tracer.

Table 8: Comparison of mean, 5th, and 95th percentiles of simulated travel time of the conservative (^{99}Tc) and reactive (^{237}Np) tracers arriving at water table at 10%, 25%, 50%, 75% and 90% mass fraction breakthrough for homogeneous case

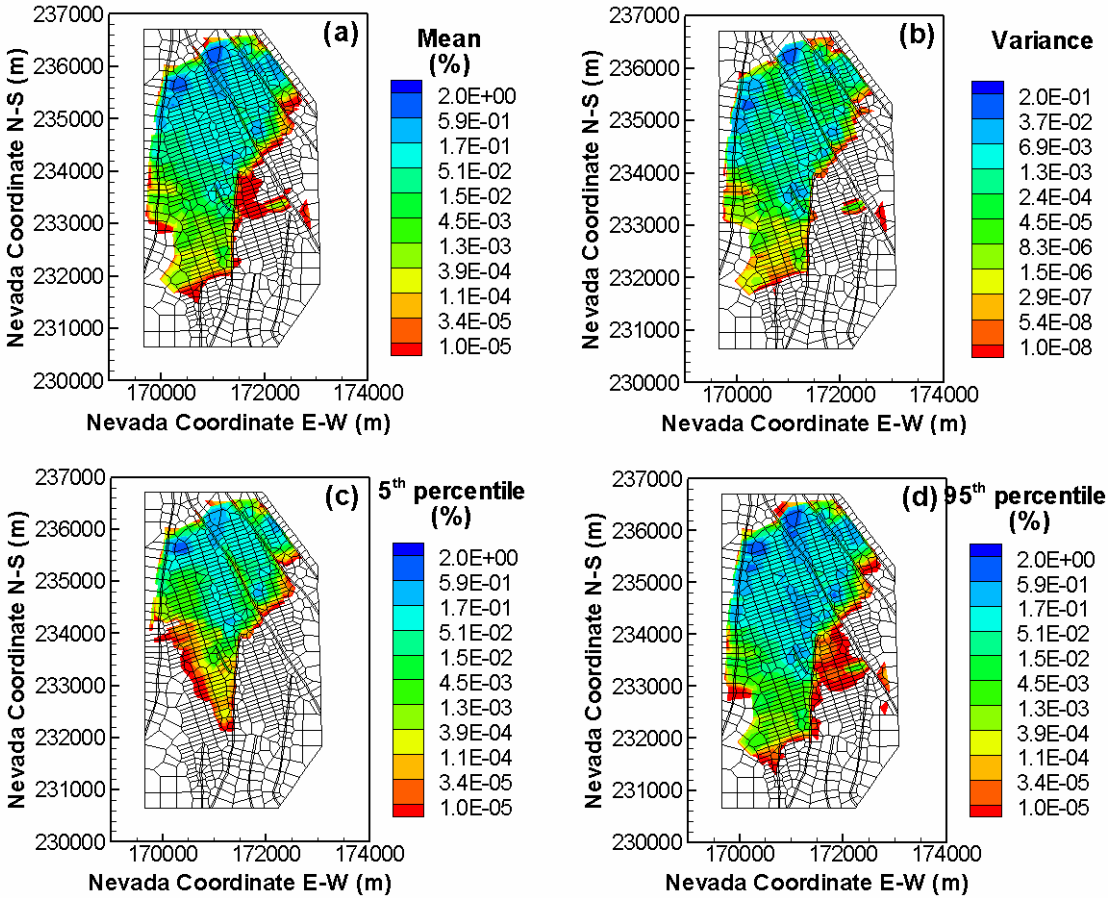
Breakthrough curves	Mass fraction	Travel time (years)	
		Conservative tracer (^{99}Tc)	Reactive tracer (^{237}Np)
5th percentile	10%	1.87E+1	1.99E+4
	25%	1.08E+3	9.40E+4
	50%	7.17E+3	2.75E+5
	75%	2.32E+4	8.38E+5
	90%	1.17E+5	>1.00E+6
Mean	10%	4.47	9.70
	25%	1.90E+1	1.72E+4
	50%	3.50E+3	1.02E+5
	75%	1.62E+4	4.26E+5
	90%	7.71E+4	>1.00E+6
95th percentile	10%	3.86	5.34
	25%	1.03E+1	1.98E+3
	50%	8.22E+2	2.00E+4
	75%	9.00E+3	1.29E+5
	90%	4.70E+4	5.80E+5

Source: Data comes from DID: 016FP.004; see the word file “transport_homo.doc”.

7.4.1.2 Uncertainty Assessment of Spatial Distribution in Radionuclide Plumes

Spatial distribution of normalized cumulative mass arrival (defined as cumulative mass arriving at each cell of the water table over time, normalized by the total mass of the initially released radionuclide from the repository) at the water table is an important variable with which to investigate transport patterns and help estimate the potential locations of high-radionuclide concentration.

Figure 28 plots that the mean, variance, 5^{th} and 95^{th} percentiles of cumulative normalized mass arrive at the water table at 1,000 years for the conservative tracer ^{99}Tc. The uncertainties of radionuclide transport could also be evaluated by these figures. Figure 28a shows that the mean of cumulative normalized mass arrival contour for ^{99}Tc at 1,000 years covers a large area below the repository footprint shown in Figure 1 and the large means of mass arrival are close to several faults in north of the model domain. The pattern of variances in Figure 28b is similar as that of the mean shown in Figure 28a. The largest variance appears in north boundary of repository footprint with the corresponding largest mean of mass fraction along several faults. The variances decrease from north to south of the model domain with corresponding decrease of the means. Because the repository is in north and west of the model domain (Figure 1) and only part of radionuclide has arrived at the water table at 1,000 years, the large mean of mass fraction at a grid cell can cause large variance at the grid. That is, the large variance appears in the north of model domain with large mean and no variance is in south-east of the model domain because no radionuclide releases from the locations at repository. The uncertainty and variance shown in Figure 28b can also be responded in 5^{th} and 95^{th} percentiles of mass arrival contours shown in Figures 28c and 28d. The 5^{th} percentile of mass arrival contour at 1,000 years covers much smaller area than the contour of mean; while 95^{th} percentile contour has a little larger covered area and values than those of the mean.

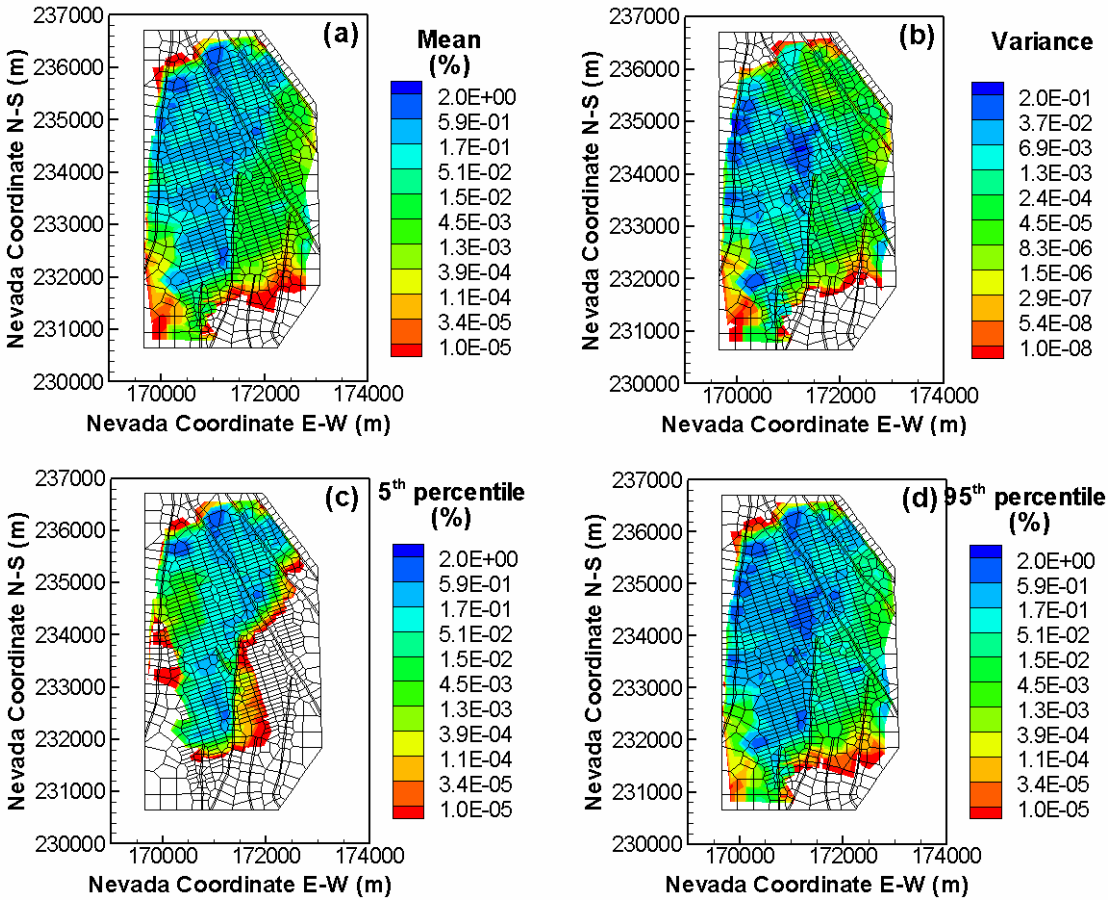


Source: Simulated data come from DID: 016FP.004; see word file "Transport_homo.doc".

Figure 28. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{99}Tc at the water table after 1,000 years for homogeneous case

Figure 29 plots the means, variances, 5th and 95th percentiles of cumulative normalized arrival at each grid of water table at 1,000,000 years for ^{99}Tc. The mean of mass arrival for ^{99}Tc at 1,000,000 years covers the entire area directly below the repository footprint, spreading to the east of the model domain but does not cover the area in south-east of the model domain. The mean of mass arrival for ^{99}Tc at 1,000,000 years shown in Figure 29a has larger value at the area below the repository footprint (Figure 1) than the one of other areas. The reason why the radionuclide spreads to east and covers almost the entire model domain is that the lateral flow at the water table affects the spread of transport. Because the simulated lateral flow at the water table in the east of repository footprint is much larger than the one in south-east of model domain, the radionuclide spreads to the east of the repository footprint and does not reach the south-east of model domain. Figure 29b shows the largest variance of mass arrival is in the west and center of the model domain and it is not the same as the largest mean appeared in the west and north of the model domain. The variance (Figure 29b) has a

similar pattern as the one of simulated fluxes at the water table shown in Figure 19b, indicating that the uncertainty of radionuclide mass arrival is correlated with the uncertainty of flow fields after 1,000,000 years. The uncertainty of radionuclide can also be quantified by the significant differences between the 5th and 95th percentile contours shown in Figures 29c and 29d.

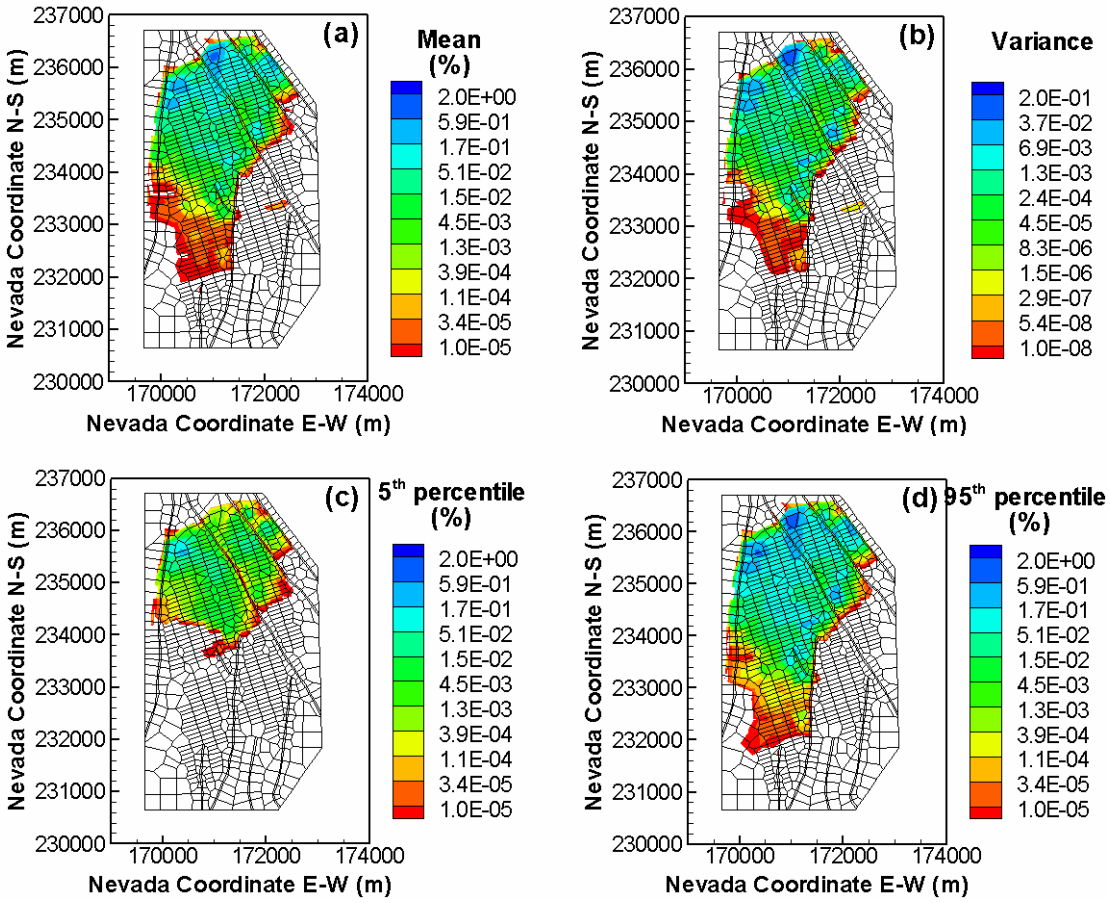


Source: Simulated data come from DID: 016FP.004; see word file "Transport_homo.doc".

Figure 29. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{99}Tc at the water table after 1,000,000 years for homogeneous case

Figure 30 show that the means, variances, 5th and 95th percentiles of cumulative normalized arrival at each grid of water table at 1,000 years for the reactive tracer (^{237}Np). These figures have similar patterns and characteristics as the ones for the conservative tracer (^{99}Tc). Figures 28a and 30a show a significant different spatial distribution of mean cumulative mass arrival at water table between ^{99}Tc and ^{237}Np at 1,000 years. This is also true for the 5th, 95th percentiles and variances. In general, the means, variances, 5th and 95th percentiles of cumulative mass arrival for ^{237}Np have much smaller covered area and values than the corresponding ones for ^{99}Tc, because the adsorption of the reactive

tracer slows down the movement of the tracer transport and the reactive tracer travels much longer time arriving at water table than the conservative one.



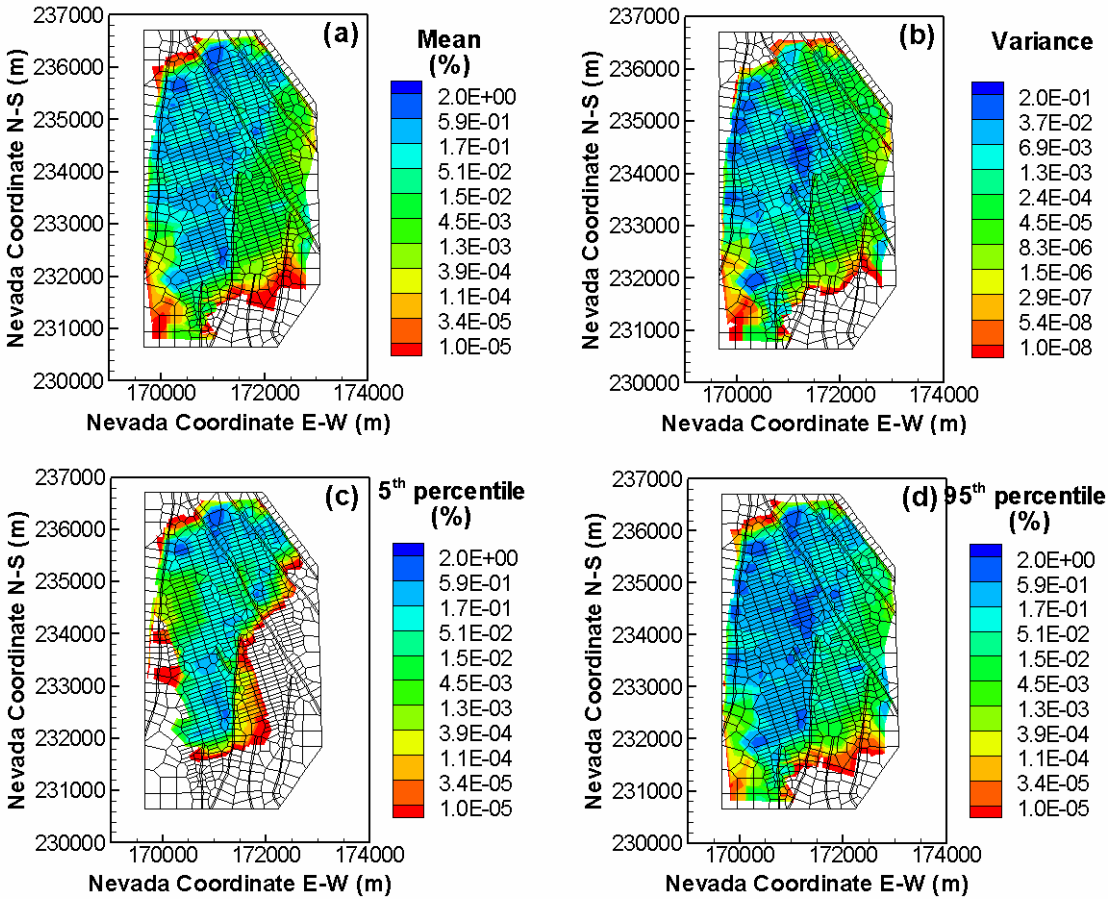
Source: Simulated data come from DID: 016FP.004; see word file "Transport_homo.doc".

Figure 30. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{237}Np at the water table after 1,000 years for homogeneous case

Figure 31 plots the means, variances, 5th and 95th percentiles of cumulative normalized arrival at each grid of water table at 1,000,000 years for ^{237}Np. The spatial distribution of mean cumulative mass arrival shown in Figure 31a is very similar to the one of ^{99}Tc shown in Figure 29a. Correspondingly, there are not much difference among their corresponding 5th, 95th percentiles and variances for ^{99}Tc and ^{237}Np. This is because more than 80 percent of both tracers have arrived at water table at this time.

From the above analysis, it can be concluded that the parameter uncertainty in porosity and sorption coefficient and the uncertainties of flow fields have significant effects on the simulated radionuclide transport in the UZ at Yucca Mountain. These uncertainties are evaluated by analyzing the means,

variances, 5th and 95th percentiles of cumulative mass arrival at water table and travel time from repository to water table.



Source: Simulated data come from DID: 016FP.004; see word file "Transport_homo.doc".

Figure 31. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{237}Np at the water table after 1,000,000 years for homogeneous case

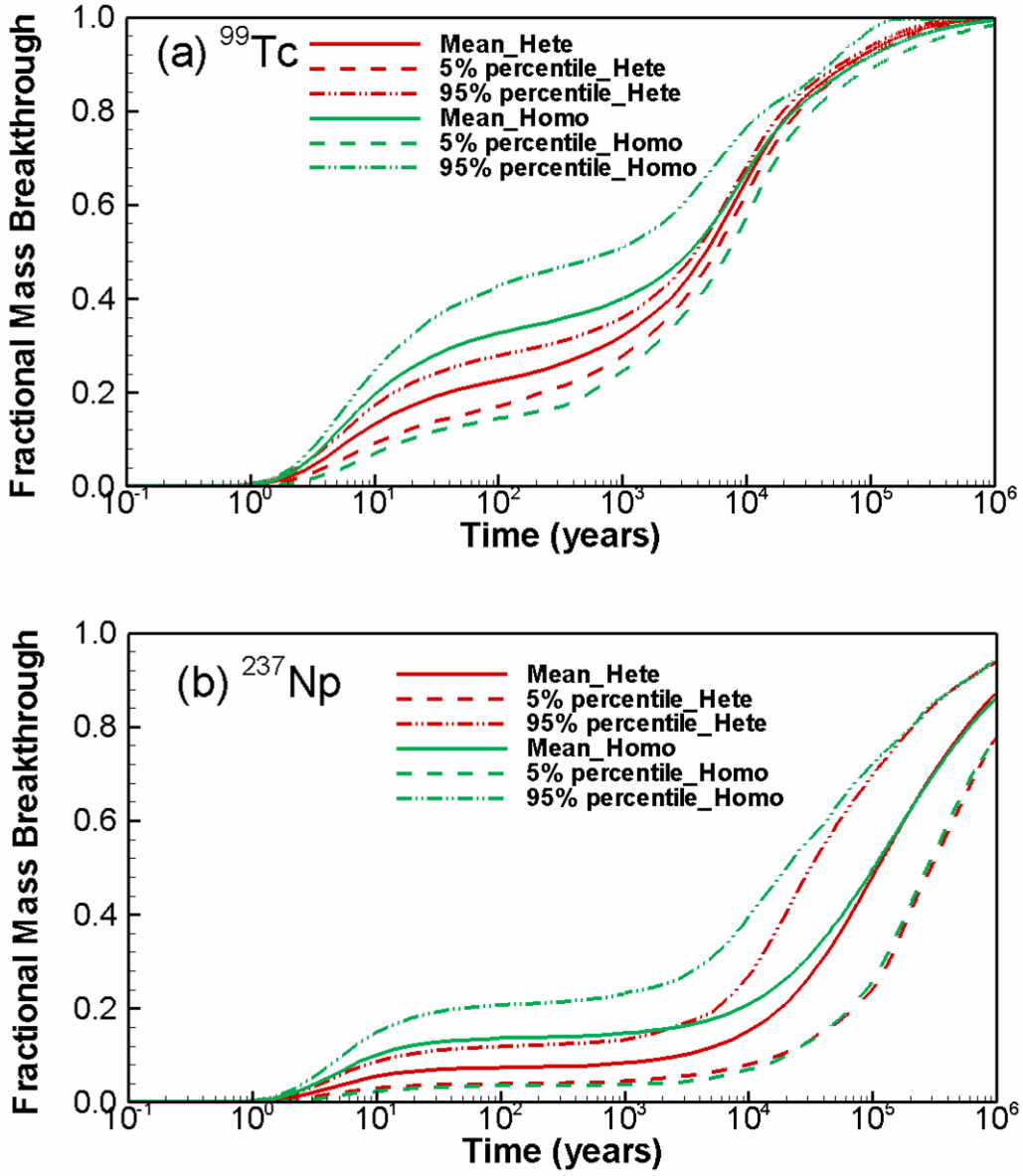
7.4.2 Uncertainty Analysis of Transport Simulations for the Heterogeneous Case

Random parameters generated for the homogeneous case are discussed in Section 7.2, and simulated flow and its uncertainty are discussed in Section 7.3.3. Note that rand sorption coefficients of the neptunium used for the heterogeneous case are the same as those of the homogeneous case.

7.4.2.1 Uncertainty Assessment of Cumulative Travel Time

Figure 32 plots the simulated fractional breakthrough curves of cumulative mass arriving at the

water table for the conservative (^{99}Tc) and reactive (^{237}Np) tracers. It also plots the homogeneous cases for comparison to investigate how the local-scale heterogeneity affects travel time prediction and its uncertainty. For each realization, the total cumulative mass is calculated by adding the cumulative mass of all blocks at the water table; the total cumulative mass of the 200 realizations are used to



Source: Simulated data come from DID: 016FP.004; see word file "Transport_hete.doc".

Figure 32. Simulated breakthrough curves of cumulative mass arriving at the water table for (a) ^{99}Tc, and (b) ^{237}Np

estimate uncertainties associated with groundwater travel times. Table 9 lists the travel times of 10%, 25%, 50%, and 75% of total mass estimated from the mean, 5th, and 95th percentiles. Travel-time uncertainty is significant, as measured by the 5th and 95th percentiles. For example, with ^{237}Np, it may take from 31,600 to 295,000 years for 50% of the total mass to enter the groundwater through the water table. Owing to the sorption effect of the reactive tracer, the reactive tracer (^{237}Np) travels about two orders of magnitude slower than the conservative tracer (^{99}Tc). For example, the mean travel times of the 50% mass fraction breakthrough is 4,760 years for ^{99}Tc, but 109,000 years for ^{237}Np. Uncertainties in fractional mass travel time of ^{237}Np are also much larger than those of ^{99}Tc, owing to the incorporation of parametric uncertainty in sorption coefficient of the reactive tracer.

Table 9: Comparison of mean, 5th, and 95th percentiles of simulated travel time of the conservative (^{99}Tc) and reactive (^{237}Np) tracers arriving at water table at 10%, 25%, 50%, 75% and 90% mass fraction breakthrough for heterogeneous case

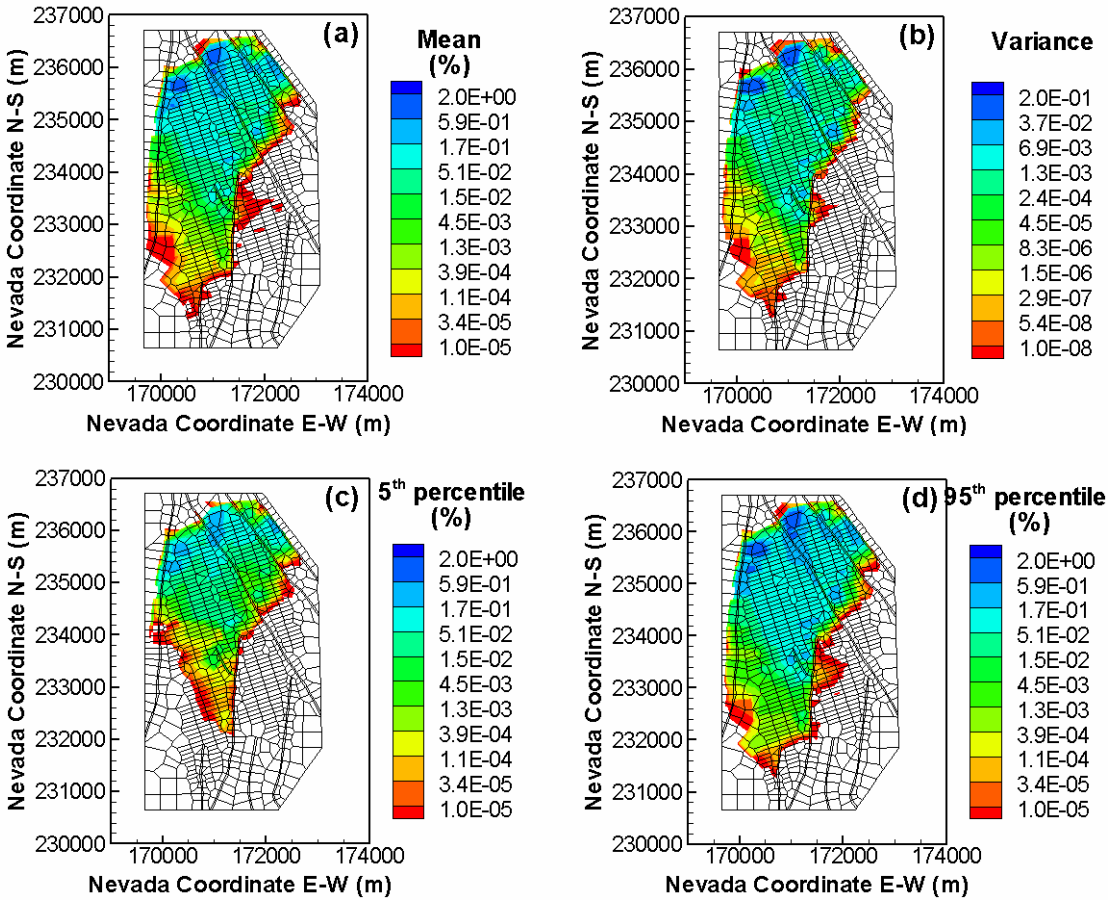
Breakthrough curves	Mass fraction	Travel time (years)	
		Conservative tracer (^{99}Tc)	Reactive tracer (^{237}Np)
5th percentile	10%	1.21×10^{1}	1.71×10^{4}
	25%	6.73×10^{2}	1.06×10^{5}
	50%	5.70×10^{3}	2.95×10^{5}
	75%	1.89×10^{4}	8.50×10^{5}
	90%	7.56×10^{4}	$>1.00\times10^{6}$
Mean	10%	6.59	2.75×10^{3}
	25%	2.37×10^{2}	2.73×10^{4}
	50%	4.76×10^{3}	1.09×10^{5}
	75%	1.64×10^{4}	4.10×10^{5}
	90%	6.59×10^{4}	$>1.00\times10^{6}$
95th percentile	10%	4.88	1.65×10^{1}
	25%	3.92×10^{1}	8.67×10^{3}
	50%	3.95×10^{3}	3.16×10^{4}
	75%	1.42×10^{4}	1.45×10^{5}
	90%	5.76×10^{4}	5.75×10^{5}

Source: Data comes from DID: 016FP.004; see the word file "transport_hete.doc".

7.4.2.2 Uncertainty Assessment of Spatial Distribution in Radionuclide Plumes

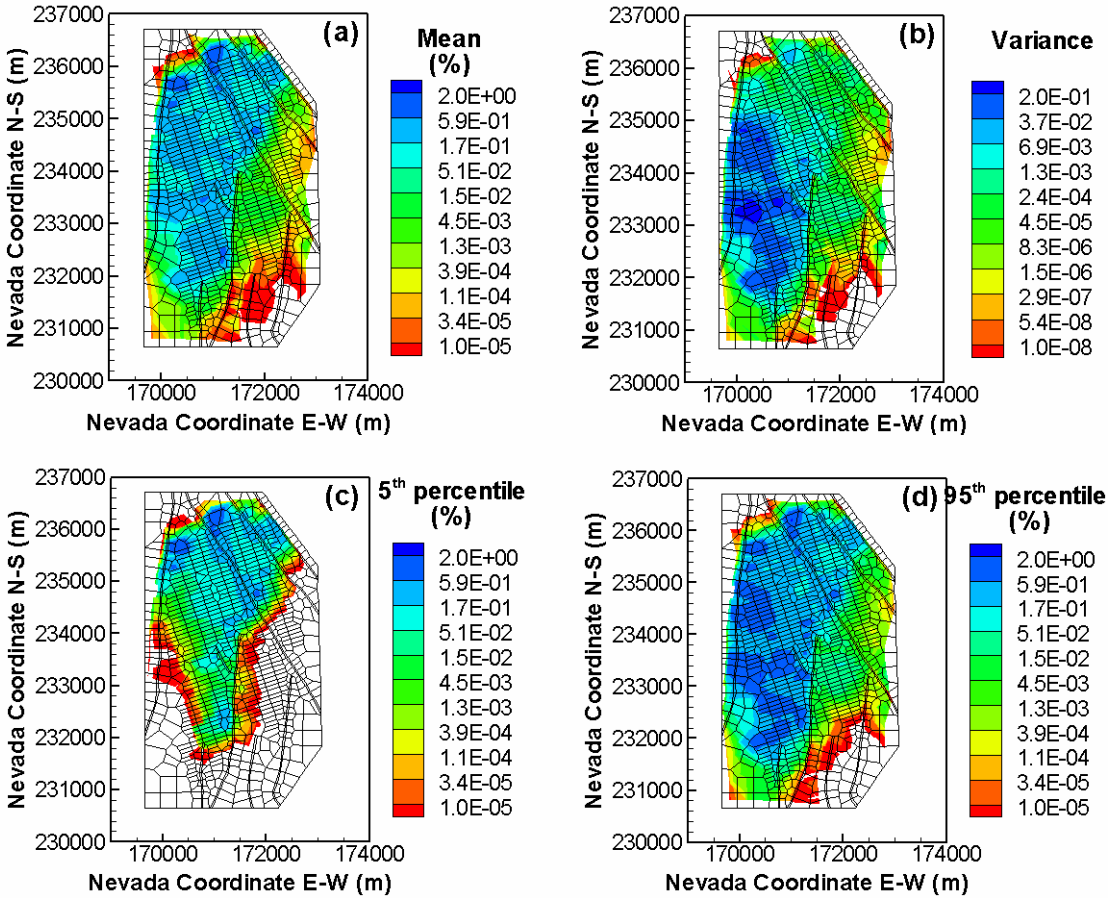
Figure 33 depicts that the mean, variance, 5th and 95th percentiles of cumulative normalized mass arrive at the water table at 1,000 years for the conservative tracer ^{99}Tc. Figure 34 does the same at 1,000,000 years. The mean of mass arrival at 1,000 years (Figure 33a) for ^{99}Tc covers a large area below the repository footprint shown in Figure 1 and the large means are close to several faults in north of the model domain. The pattern of variance (Figure 33b) is similar as that of the mean shown in Figure 33a. The mean of mass arrival at 1,000,000 years (Figure 34a) for ^{99}Tc covers the entire

area directly below the repository footprint, spreading to the east of the model domain and almost cover the whole model domain. The large mean values of mass arrival at 1,000,000 years are below the repository footprint (Figure 1). The pattern of variance at 1,000,000 years for ^{99}Tc (Figure 34b) is similar as the one of simulated percolation flux at water table shown in Figure 25b, indicating the uncertainty of radionuclide at 1,000,000 years is correlated with the uncertainty of flow fields. The uncertainty of radionuclide can also be quantified by the significant differences between the 5th and 95th percentile contours shown in Figures 34c and 34d.



Source: Simulated data come from DID: 016FP.004; see word file "Transport_hete.doc".

Figure 33. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{99}Tc at the water table after 1,000 years for heterogeneous case

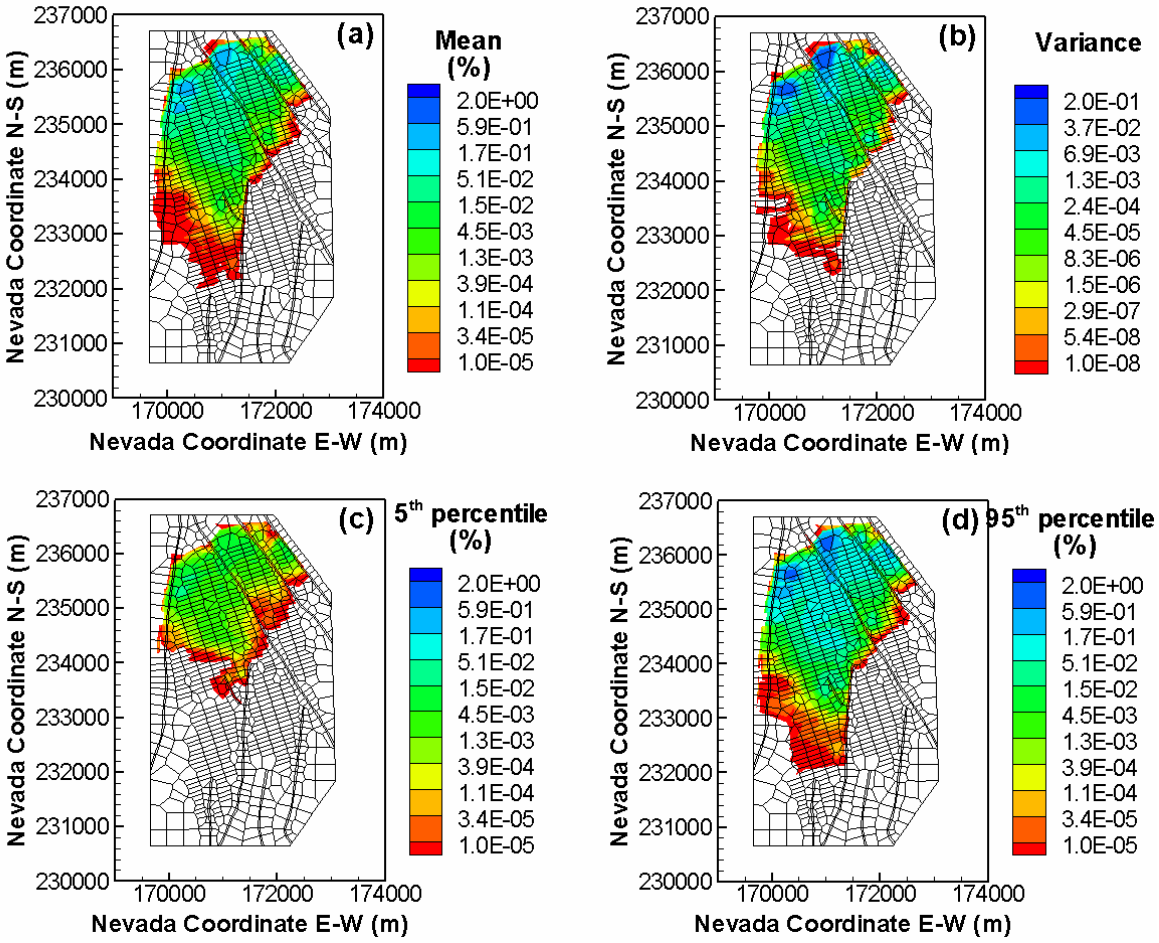


Source: Simulated data come from DID: 016FP.004; see word file "Transport_hete.doc".

Figure 34. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{99}Tc at the water table after 1,000,000 years for heterogeneous case

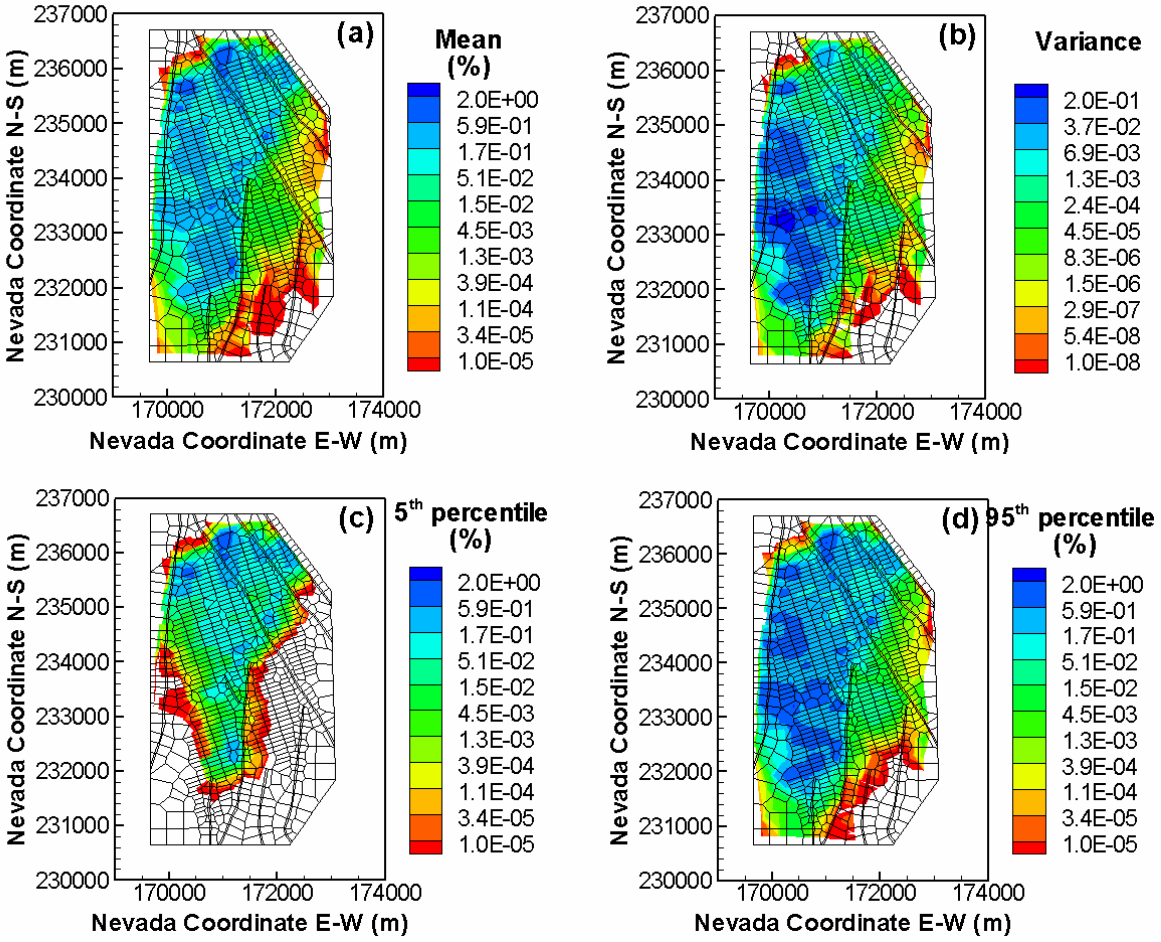
Figure 35 depicts mean, variance, and 5th and 95th percentiles of the normalized cumulative mass-arrival contours of the reactive tracer (^{237}Np) at 1,000 years. The mean of mass arrival at 1,000 years for ^{237}Np (Figure 35a) has a significant different spatial distribution from the one of ^{99}Tc (Figure 33a). This is also true for the 5th, 95th percentiles and variances. The means, variances, 5th and 95th percentiles of cumulative mass arrival for ^{237}Np have much smaller covered area and values than the corresponding ones for ^{99}Tc, because the adsorption of the reactive tracer slows down the movement of the tracer transport and the reactive tracer travels much longer time arriving at water table than the conservative one. Figure 36 does the same at 1,000,000 years. The mean of mass arrival at 1,000,000 years (Figure 36a) covers virtually the entire area, with higher values directly below the footprint of the proposed repository shown in Figure 1. The contour spreads widely, to the east of the model domain, but high values appear restricted to the west of the Ghost Dance Fault (the east boundary of the repository footprint, Figure 1), indicative of the dominant vertical movement for

radionuclide. The variance contour shown in Figure 36b has a similar pattern to Figure 36a, with higher values of variance below the repository footprint. In addition, the area of the higher variance corresponds to the area of high mean, except at the north end of the Drillhole Wash Fault (Figure 1), where the normalized cumulative is high in the 5th and 95th percentiles (Figures 36c and 36d). Patterns of variance (Figure 36b) are correlated with those of Figure 25b for the variance of percolation flux, indicating that the uncertainty of radionuclide mass arrival is related to that of the flow field.



Source: Simulated data come from DID: 016FP.004; see word file "Transport_hete.doc".

Figure 35. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{237}Np at the water table after 1,000 years for heterogeneous case



Source: Simulated data come from DID: 016FP.004; see word file "Transport_hete.doc".

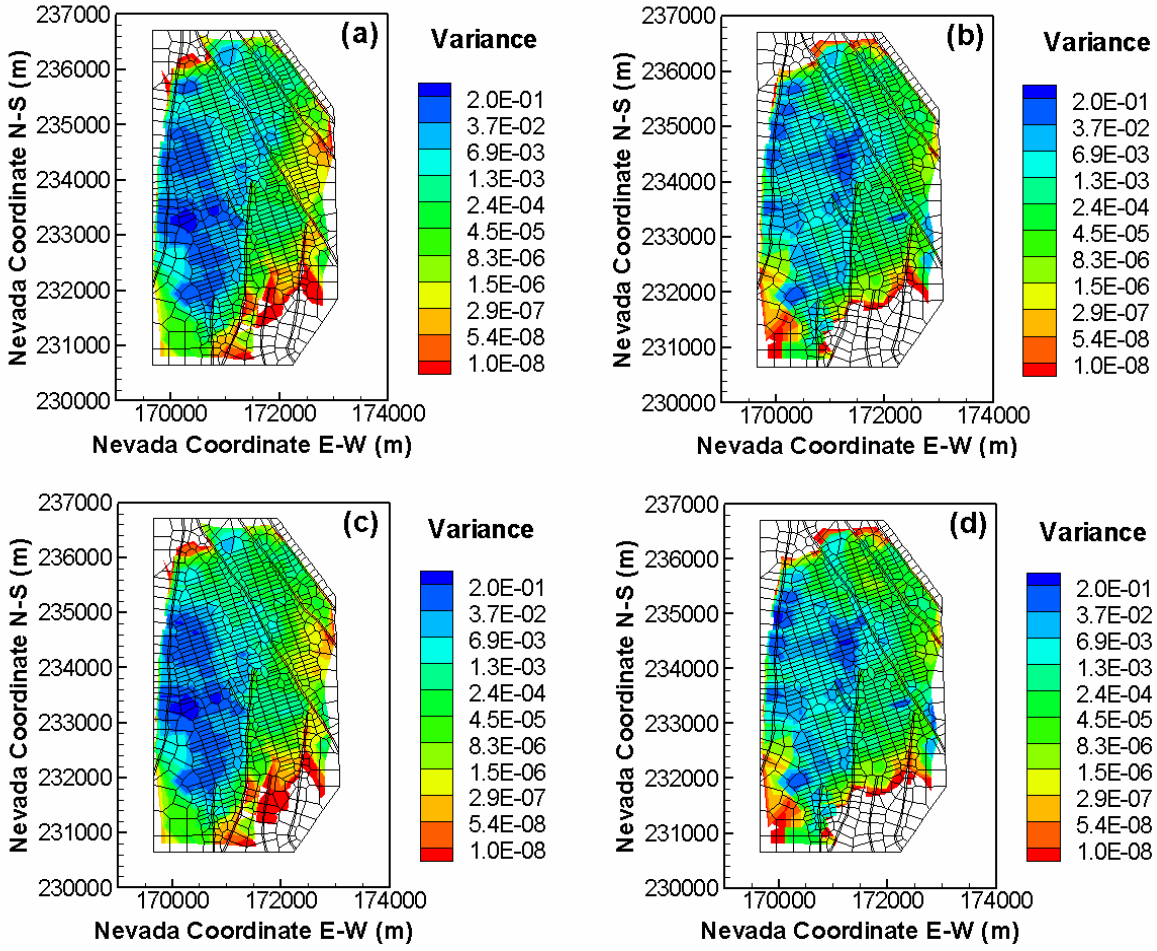
Figure 36. (a) Mean, (b) variance, (c) 5th percentile, and (d) 95th percentile of normalized cumulative mass arrival contours of ^{237}Np at the water table after 1,000,000 years for heterogeneous case

7.4.3 Comparison of Transport Uncertainty Assessment for the Homogeneous and Heterogeneous Cases

We compare the uncertainty of radionuclide transport in the UZ for the homogeneous and heterogeneous cases. Spatial patterns and magnitudes for the mean prediction of the homogeneous case (Figure 31a) are similar to those of the heterogeneous case (Figure 36a). The variance contours (Figures 36b and 31b) of the two cases have different spatial patterns, and the variance of the heterogeneous case is much larger than that of the homogeneous case. These differences show that incorporating the local-scale heterogeneity of permeability and porosity resulted in higher uncertainty for radionuclide transport. Comparing the 5th and 95th percentiles of the normalized cumulative mass arrival for the two cases shows that the 5th percentiles are similar, whereas the 95th

percentile of the heterogeneous case is significantly higher than that of the homogeneous case. This suggests that the large variance in the heterogeneous case is mainly caused by the higher value of normalized cumulative mass at the 95th percentile, resulting from local-scale heterogeneity.

By comparing the variances in cumulative mass arrival contours of ^{237}Np and ^{99}Tc at 1,000,000 years for heterogeneous and homogeneous cases shown in Figure 37, one can also see that the variances in cumulative mass arrival for heterogeneous case are much larger than the ones for homogeneous case in the west of model domain. This is similar as the comparisons of unsaturated flow fields for both cases, indicating that the parametric uncertainty in local-scale heterogeneity of hydrologic properties can result in much more uncertainty in unsaturated flow fields and tracer concentration contours than parametric uncertainty in layer-scale heterogeneity does.



Source: Simulated data come from DID: 016FP.004; see word file "Transport_hete.doc".

Figure 37. Variance after 1,000,000 years, (a) ^{237}Np for heterogeneous case; (b) ^{237}Np for homogeneous case; (c) ^{99}Tc for heterogeneous case; (d) ^{99}Tc for homogeneous case

To investigate how local-scale heterogeneity affects radionuclide travel time, the breakthrough curves of the fractional cumulative mass arrival at the water table in the homogenous case are also plotted in Figure 32. As mentioned before, the fractional cumulative mass at different times is calculated for each realization by summarizing cumulative mass over all blocks at the water table, with the statistics estimated based on results of the 200 realizations. Note that the cumulative mass shown in Figure 32 differs from that shown in Figures 31 and 36. For example, the 95th percentile of cumulative mass at 1,000,000 year in Figure 32 corresponds to one realization, while contour of the 95th percentile in Figure 36 (and 31) is a combination of the 200 realizations (i.e., the normalized cumulative mass arrival at different blocks of Figure 36 corresponds to different realizations). In other words, Figure 32 is used to evaluate site performance averaged over the domain, whereas Figure 36 (and 31) is focused on spatial variation.

Figure 32 shows that, for both tracers in the early travel period, mean travel time for the heterogeneous case is delayed compared to the homogeneous case. The slower travel in the heterogeneous case is not surprising. Since local-scale heterogeneity is incorporated, the flow path becomes more tortuous, and radionuclide travel between matrix and fracture becomes more complicated. With the downward movement of the tracers, preferential flow paths may develop along fractures of high permeability, and the effect of local-scale matrix-property heterogeneity on tracer transport diminishes. As a result, the travel time in the two cases becomes the same after 20,000 years, with 78% fractional mass breakthrough for ^{99}Tc, and 100,000 years, with 48% fractional mass breakthrough for ^{237}Np. Similar breakthrough behavior is also observed in Zhou et al. (2003).

Figure 32 also shows that, for both tracers, the uncertainty bound for the travel time prediction of the heterogeneous case is much smaller than that of the homogeneous case, indicative of reduced uncertainty. This can also be observed by comparing travel times listed in Table 9 with those in Table 8. For example, the variation in travel time when 75% of the mass for ^{99}Tc flows out of the UZ is between 9,000 and 23,200 years in the homogeneous case, but between 14,200 and 18,900 years in the heterogeneous case. This difference is attributed to the conditional random fields generated in the heterogeneous case. In the homogeneous case, without conditioning, difference of permeability and porosity between realizations is significant, which results in large uncertainty in travel time predictions. Whereas, in the heterogeneous case, due to the conditioning, variation of permeability and porosity between realizations is reduced, which reduces predictive uncertainty of travel time.

7.5 Main Conclusions

This project studies the uncertainty of unsaturated flow and radionuclide transport caused by layer- and local-scale spatial variability in hydraulic parameters (i.e., matrix permeability and porosity) in the UZ of Yucca Mountain. *Layer scale* was specific to hydrogeologic layers delineated based on geologic information and on-site measurements; *local scale* represented the spatial variation in hydraulic properties within a layer. Matrix permeability and porosity are treated as homogeneous and heterogeneous random variables, respectively. The sorption coefficient of reactive tracer (^{237}Np) is also treated as homogeneous random variables.

To generate homogeneous random fields of the parameters, the distribution of matrix permeability, porosity and sorption coefficient are identified based on site measurements and model calibration results. Seven transformations (including three transformations from the Johnson system and four classical re-expressions) are applied to the measurements, and the Lilliefors test is used to select the best transformation at a certain significance level. The mean permeability is further adjusted based on model calibration results for better model predictions. Distributions of matrix porosity and sorption coefficients are determined solely from site measurements. The statistical correlation (measured by the Spearman rank correlation coefficient) between the matrix permeability and porosity is incorporated in the random field generation. Random fields generated using the LHS method agree well with the measurements, and 200 realizations are sufficient to yield representative distribution functions for the three random parameters.

The heterogeneous random fields of matrix permeability and porosity are generated for each layer using SGSIM to account for local-scale heterogeneity, conditioned on core measurements. Because SGSIM requires that conditioning data follow a normal distribution, the measurements are transformed to be Gaussian according the determined transformations using Lilliefors Test. The generated heterogeneous random fields are adjusted so that their mean agree with the layer-scale parameters obtained from inverse modeling. As a result, the generated random fields represent layer-scale heterogeneity and local-scale spatial variability within each layer.

Monte Carlo simulations are conducted to simulate unsaturated flow and radionuclide transport in the UZ and the generated homogeneous and heterogeneous random fields are used as input data of the 3-D flow and transport model (TOUGH2 code), respectively. Mean, variance, and 5^{th} and 95^{th} percentiles of simulated variables (e.g., saturation, water potential, percolation flux, and normalized cumulative mass arrival) are calculated as optimum prediction and measures of associated predictive uncertainty. Predictions of unsaturated flow are investigated by comparing them with observations of matrix saturation and water potential. Mean predictions are in reasonable agreement with observations and match their spatial variation patterns for both cases. The 5^{th} and 95^{th} percentiles (also known as uncertainty bounds) bracket a large portion of the observations, indicating that the simulations are able to evaluate uncertainty in unsaturated flow.

The simulated unsaturated flow results for both cases indicate that the parameter uncertainty in layer- and local-scale heterogeneity in matrix porosity and permeability can cause significant uncertainty on percolation fluxes at water table. Comparing simulations of unsaturated flow for heterogeneous case with studies for homogeneous case, we find that the local-scale heterogeneity of permeability did not significantly affect mean flow predictions, but did affect predictive uncertainty as measured by the 5^{th} and 95^{th} percentiles as well as variance, especially under the footprint of the proposed repository.

Uncertainty of transport is assessed for conservative (^{99}Tc) and reactive tracers (^{237}Np) for homogeneous and heterogeneous cases, with focus on spatial distribution of the normalized cumulative mass arrival and fractional breakthrough curves of cumulative mass arriving at the water table. The results indicate the uncertainty in radionuclide mass arrival is related to that of the flow field, and sorption retards radionuclide transport. The parametric uncertainty in layer- and local-scale heterogeneity of matrix porosity, permeability and sorption coefficient of reactive tracer can

cause significant uncertainty of radionuclide transport in the UZ of Yucca Mountain. The conservative tracer without sorption could move one or two orders of magnitude faster than the reactive one with random sorption coefficient. Comparing simulation of radionuclide transport for heterogeneous case to that for homogeneous case, we find that the local-scale heterogeneity of permeability and porosity does not significantly affect the mean predictions of transport. However, it renders the variance in normalized cumulative mass arrival larger and more spatially variable. For both tracers, the local-scale heterogeneity delayed mean travel time during the early travel period, because of the more tortuous flow path and more complicated particle traveling between matrix and fracture. This effect diminished with the downward movement of the tracers, since preferential flow may develop along fractures of high permeability. Uncertainty in travel-time predictions is significantly reduced, as a result of the conditioning on local-scale measurements in the process of random field generation.

7.6 Future Studies

Although water retention parameters are treated as deterministic variables, variability of the parameters is observed in their limited measurements. Treating them as deterministic variables may underestimate predictive uncertainty. We suggest continuous study to evaluate uncertainty of the water retention parameters. On the other hand, a global sensitivity analysis is worthy based on the Monte Carlo results to identify parameters that are the most important to the unsaturated flow and radionuclide transport. This will help to concentrate limited resource to reduce predictive uncertainty to the maximum extent.

8.0 INPUTS AND REFERENCES

8.1 Inputs

Table 10: Input data source and data tracking numbers

Current DTN	Location in this report			Description/Remarks
	Text	Figure	Table	
LB0303THERMESH.001		18-19, 24-26		Three-dimensional UZ thermal model grids
LB0207REVUZPRP.002	7.1.1, 7.1.2	4, 6, 10-11, 14-16, 21-22	2-5	Matrix saturation, hydraulic conductivity, porosity, water retention parameters
GS970808312232.005 GS031208312232.003		17, 23		Water Potential data collected from boreholes USW NRG-7a, USW NRG-6, UE-25 UZ #4, UE-25 UZ #5, USW UZ-7a and USW SD-12
LA0407AM831341.004	7.1.4	8	6-7	Sorption coefficients of Neptunium
LB03033DUZTRAN.001		27		Tc and Np transport simulation scenarios, input/outputfies
LB03013DSSCP3I.001		10		Three-dimensional site scale model calibrated property set
LB0303THERMSIM.001		3, 10, 14-17, 20-24		Input and output files of deterministic results for flow simulation in thermal domain
DID: 016FP.001		3, 5, 7, 9	2-7	The results of statistics of measurements and distribution determination.
DID: 016FP.002		13-26		The results of unsaturated flow simulation.
DID: 016FP.003		4, 6, 8, 10-12		The results of homogeneous and heterogeneous random field generation.
DID: 016FP.004		27-37	8-9	The results of radionuclide transport simulation
DID: 016JZ.001		A1-A8		The results of stochastic analysis of transient flow.

8.2 Cited References

Ballio, F.; Guadagnini, A. 2004. "Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology." *Water Resources Research,* 10, (4), W04603. Washington D.C.: American Geophysical Union.

Bardurraga, T.M.; Bardvarsson, G.S.1999. "Calibrating hydrogeologic properties for the 3-D site-scale unsaturated zone model at Yucca Mountain, Nevada." Journal *of Contaminant Hydrology,* 38 (1-3), 5-46.

Bellin, A.; Salandin, P.; Rinaldo, A. 1992. "Simulation of dispersion in heterogeneous porous formations: Statistics, first-order theories, and convergence of computations." *Water Resources Research,* 28(9), 2211-2227.

Beven K.J; Binley A.M. 1992. "The future of distributed models: model calibration and uncertainty prediction." *Hydrological Processes*, 6, 279 – 298.

Beven, K. 2006. "A manifesto for the equifinality thesis." *Journal of Hydrology*, 320, 18-36.

Bodvarsson, G.S., Liu, H.H., Ahlers, R., Wu, Y.S., Sonnenthal, E., 2001. Parameterization and upscaling in modeling flow and transport at Yucca Mountain. Conceptual Models of Flow and Transport in the Fractured Vadose Zone, National Research Council. National Academy Press, Washington, DC, USA.

Bowen, W.M.; Bennett, C.A. 1988. *Statistical methods for nuclear material management*. NUREG/CR-4604. Richland, Washington: Battelle Pacific Northwest laboratory.

Burr, D.T.; Sudicky, E.A. 1994. "Nonreactive and reactive solute transport in three-dimensional heterogeneous porous media: mean displacement, plume spreading, and uncertainty." *Water Resources Research,* 30(3), 791-815.

BSC (Bechtel SAIC Company) 2003. *Total System Performance Assessment – License Application Methods and Approach.* TDR-WIS -PA-000006 REV 00 ICN 01. Bechtel SAIC Company, Las Vegas, Nevada, USA.

BSC (Bechtel SAIC Company) 2004. *Calibrated properties model*. MDL-NBS-HS 000003 REV 02. Las Vegas, Nevada: Bechtel SAIC Company.

Carsel, R.F.; Parrish, R.S. 1988. "Developing joint probability distributions of soil water retention characteristics." *Water Resources Research,* 24, (5), 755-769. Washington D.C.: American Geophysical Union.

Dagan, G. 1989. *Flow and Transport in Porous Formations.* Springer-Verlag, Berlin, Germany.
Dagan, G.; Neuman, S.P. 1997. *Subsurface flow and transport: A stochastic approach.* Cambridge University Press, International Hydrology Series Cup, Cambridge, British.

Dai, Z.; Ritzi, R.W.; Dominic, D.F. 2005. "Improving permeability semivariograms with transition probability models of hierarchical sedimentary architecture derived from outcrop analog studies." *Water Resources Research,* 41, W07032, doi:10.1029/2004WR003515.
Draper, D.; Pereira, A.; Prado, P.; Saltelli, A.; Cheal, R.; Eguilior, S.; Mendes, B.; Tarantola, S. 1999. "Scenario and parametric uncertainty in GESAMAC: A methodological study in nuclear waste disposal risk assessment." Computer *Physics Communication,* 117, 142-155.

Deutsch, C.V.; Journel, A.G. 1998. *GSLIB: Geostatistical software library and user's guide*. Oxford, New York: Oxford University Press.

Faybishenko, B. 2007. "Climatic forecasting of net infiltration at Yucca Mountain using analogue meteorological data." *Vadose Zone Journal,* 6(1), 77 – 92.

Flint, L.E. 1998. *Characterization of hydrogeologic units using matrix properties, Yucca Mountain, Nevada*. Water- Resources Investigations Report, 97-4243. Denver, Colorado: U.S. Geological Survey.

Flint, L.E. 2003. "Physical and hydraulic properties of volcanic rocks from Yucca Mountain, Nevada." *Water Resources Research,* 39, (5), 1-13. Washington, D.C.: American Geophysical Union.

Flint, L.E.; Buesch, D.C.; Flint, A.L. 2006. "Characterization of unsaturated zone hydrogeologic units using matrix properties and depositional history in a complex volcanic environment." *Vadose Zone Journal*, 5, 480–492. Madison, Wisconsin: Soil Science Society of America.

Gelhar, L.W. 1989. *Stochastic Subsurface Hydrology*. Englewood Cliffs, New Jersey: Prentice Hall.

Hassan, A.E.; Cushman, J.H.; Delleur, J.W. 1998. "A Monte Carlo assessment of Eulerian flow and transport perturbation models." *Water Resources Research,* 34(5), 1143-1163.

Haukwa, C.B.; Tsang, Y.W.; Wu, Y.-S.; Bodvarsson, G.S. 2003. "Effect of heterogeneity in fracture permeability on the potential for liquid seepage into a heated emplacement drift of the potential repository." *Journal of Contaminant Hydrology*, 62–63, 509–527.

Helton, J.C.; Davis, F.J. 2000. *Sampling-based methods for uncertainty and sensitivity analysis*. SAND99-2240. Albuquerque, New Mexico: Sandia National Laboratories.

Helton, J.C.; Davis, F.J. 2003. "Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems." *Reliability Engineering & System Safety*, 81, 23-69. New York, New York: Elsevier.
.

Hoeting, J.A.; Madigan, D.; Raftery, A.E.; Volinsky, C.T. 1999. "Bayesian model averaging: A tutorial." *Statistical Science,* 14(4), 382-417.

Hu, B. X.; Wu, J.; Panorska, A.K.; Zhang, D.; He, C. 2003. "Stochastic study on groundwater flow and solute transport in a porous medium with multi-scale heterogeneity." *Advances in Water Resources,* 26, 541-560.

Hu, B.X.; Wu, J.; Shirley, C.; Zhang, D. 2004. "A numerical method of moments for reactive solute transport in physically and chemically nonstationary formations: Linear equilibrium sorption with random Kd." *Stochastic Environmental Research and Risk Assessment*, 20, 46–63, doi:10.1007/s00589-003-0253-5. Heidelberg, Germany: Springer Berlin.

Illman, W.A.; Hughson, D.L. 2005. "Stochastic simulations of steady state unsaturated flow in a three-layer, heterogeneous, dual continuum model of fractured rock." *Journal of Hydrology,* 307, 17-37.

Iman, R.L.; Conover, W.J. 1982. "A distribution-free approach to inducing rank correlation among input variables." *Communications in Statistics: Simulations and computation.* 11, (3), 311-334. London, United Kingdom: Taylor & Francis.

Istok J.D.; Rautman C.A.; Flint L.E.; Flint A.L. 1994. "Spatial variability in hydrologic properties of a volcanic tuff." *Ground Water*, 32, (5): 751–760. Westerville, Ohio: National Ground Water Association.

Johnson, N.L.; Kotz, S. 1970. *Distributions in Statistics: Continuous Univariate Distributions, Vol. 1*. Boston, Massachusetts: Houghton Miffin Company.

Liu, H.H.; Zhang, M.; Persoff, P.; Gilkcrson, K.; Wu, Y.S.; Wang, J.S.Y. 2003. *Analysis of Hydrologic Properties Data. AMR Report MDL-NBS-HS-000014 REV00*. Berkeley, California: Lawrence Berkeley National Laboratory.

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

Lu, Z.; Zhang, D. 2007. "Stochastic simulations for flow in nonstationary randomly heterogeneous porous media using a KL-based moment-equation approach." *SIAM Multiscale Modeling and Simulation,* 6(1), 228 – 245.

Koltermann, C.E.; Gorelick, S.M. 1996. "Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches." *Water Resources Research,* 32(9), 2617–2658.

Mallants, D.; Jacques, D.; Vanclooster, M.; Diels, J.; Feyen, J. 1996. "A stochastic approach to simulate water flow in a macroporous soil." *Geoderma*, 10, 299-324. New York, New York: Elsevier.

Mckay, M.D.; Beckman, R.J.; Conover, W.J. 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. Alexandria, Virginia: The American Society for Quality and the American Statistical Association.

Meyer, P.D.; Ye, M.; Rockhold, M.L.; Neuman, S.P.; Cantrell, K.J. 2007. *Combined Estimation of Hydrogeologic Conceptual Model, Parameter, and Scenario Uncertainty.* Rep. NUREG/CR-XXXX, PNNL-16396, prepared for U.S. Nuclear Regulatory Commission, Washington, D.C., USA.
Montazer, P.; Wilson, W.E. 1984. *Conceptual hydrologic model of flow in the unsaturated zone, Yucca Mountain, Nevada.* Water Resour. Invest. Rep. 84-4345. US Geological Survey, Lakewood, Colorado, USA.

Neuman, S.P. 2003. "Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models." *Stochastic Environmental Research and Risk Assessment,* 17(5), 291-305, DOI: 10.1007/s00477-003-0151-7.

Nichols, W.E.; Freshley, M.D. 1993. "Uncertainty analyses of unsaturated zone travel time at Yucca Mountain." *Ground water,* 31(2), 293–301.

Poeter, E.P.; Anderson, D.A. 2005. "Multimodel ranking and inference in ground water modeling." *Ground Water,* 43(4), 597-605.

Pruess, K.; Oldenburg, C.; Moridis, G. 1999. *TOUGH2 user's guide, version 2.0*. LBL-43134. Berkeley, California: Lawrence Berkeley Laboratory.

Refsgaard, J.C.; van der Sluijs, J.P.; Brown, J.; van der Keur, P. 2006. "A framework for dealing with uncertainty due to model structure error." *Advances in Water Resources,* 29, 1586-1597.

Ritzi, R. W. 2000. "Behavior of indicator variograms and transition probabilities in relation to the variance in lengths of hydrofacies." *Water Resources Research,* 36(11), 3375– 3381.

Rubin, Y. 2003. *Applied Stochastic Hydrogeology.* Oxford University Press, New York, USA.

Swiler, L.P.; Wyss, G.D. 2004. *A user's guide to Sandia's Latin Hypercube Sampling software: LHS UNIX library/standalone version*. Technical Report SAND2004-2439. Albuquerque, New Mexico: Sandia National Laboratories.

Viswanathan, H.S.; Robinson, B.A.; Gable, C.W.; Carey, W.C. 2003. "A geostatistical modeling study of the effect of heterogeneity on radionuclide transport in the unsaturated zone, Yucca Mountain." *Journal of Contaminant Hydrology,* 62–63, 319–336.

Winter, C. L.; Tartakovsky, D.M. 2000. "Mean flow in composite porous media." *Geophysical Research Letters*, 27(12), 1759– 1762.

Winter, C.L.; Tartakovsky, D.M. 2002. "Groundwater flow in heterogeneous composite aquifers." *Water Resources Research*, 38(8), 1148, doi: 10.1029/2001WR000450.

Winter, C.L.; Tartakovsky, D.M.; Guadagnini, A. 2003. "Moment differential equations for flow in highly heterogeneous porous media." *Surveys in Geophysics*, 24, 81-106.

Wu, J.; Hu, B. X.; Zhang, D. 2003. "Applications of nonstationary stochastic theory to solute transport in multi-scale geological media." *Journal of Hydrology*, 275(3-4), 208-228.

Wu, J.; Hu, B.X. 2004. "Three-dimensional method of moments for linear equilibrium-adsorbing solute transport in physically and chemically nonstationary formations." *Mathematical Geology*, 36, (2), 261-287. Kingston, Canada: International Association of Mathematical Geology.

Wu, Y.S.; Ahlers, C.F.; Fraser, P.; Simmons, A.; Pruess, K. 1996. *Software qualification of selected TOUGH2 modules*. LBNL-39490. Berkeley, California: Lawrence Berkeley Laboratory.
Wu, Y.S.; Haukwa, C.; Bodvarsson, G.S. 1999. "A site-scale model for fluid and heat flow in the unsaturated zone of Yucca Mountain, Nevada." *Journal of Contaminant Hydrology,* 38 (1-3), 185-217.

Wu, Y.S.; Pan, L.; Zhang, W.; Bodvarsson, G.S. 2002. "Characterization of flow and transport processes within the unsaturated zone of Yucca Mountain, Nevada, under current and future climates." *Journal of Contaminant Hydrology,* 54, 215-247. New York, New York: Elsevier.

Wu, Y.S.; Lu, G.; Zhang, K.; Zhang, G.; Liu, H.H.; Xu, T.; Sonnenthal, E.L. 2004a. *UZ Flow Models and Submodels*. AMR Report MDL-NBS-HS-000006 REV02. Berkeley, California: Lawrence Berkeley National Laboratory.

Wu Y.S.; Lu, G.; Zhang, K.; Bodvarsson, G.S. 2004b. "A mountain-scale model for characterizing unsaturated flow and transport in fractured tuffs of Yucca Mountain." *Vadose Zone Journal,* 3, 796-805. Madison, Wisconsin: Soil Science Society of America.

Ye, M.; Neuman, S.P.; Meyer, P.D. 2004. "Maximum Likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff." *Water Resources Research,* 40, W05113, doi:10.1029/2003WR002557.

Ye, M.; Neuman, S.P.; Meyer, P.D.; Pohlmann, K.F. 2005. "Sensitivity analysis and assessment of prior model probabilities in MLBMA with application to unsaturated fractured tuff." *Water Resources Research,* 41, W12429, doi:10.1029/2005WR004260.

Zhang, D.; Winter, C.L. 1998. "Nonstationary stochastic analysis of steady-state flow through variably saturated, heterogeneous media." *Water Resources Research*, 34, (5), 1091-1100. Washington D.C.: American Geophysical Union.

Zhang, D. 1999. "Nonstationary stochastic analysis of transient unsaturated flow in randomly heterogeneous media." *Water Resources Research*, 35, (4), 1127–1141. Washington D.C.: American

Geophysical Union.

Zhang, D.; Andricevic, R.; Sun, A.Y.; Hu, B.X.; He, G. 2000. "Solute flux approach to transport through spatially nonstationary flow in porous media." *Water Resources Research,* 36, 2107–2120.

Zhang, D. 2002. *Stochastic Methods for Flow in Porous Media: Coping with Uncertainties*. San Diego, California: Academic Press.

Zhang K.; Wu Y.S.; Houseworth J.E. 2006. "Sensitivity analysis of hydrological parameters in modeling flow and transport in the unsaturated zone of Yucca Mountain." *Hydrogeology Journal*, 14, 1599-1619. Heidelberg, Germany: Springer Berlin.

Zhou, Q.; Liu, H.H.; Bodvarsson, G.S.; Oldenburg, C.M. 2003. "Flow and transport in unsaturated fractured rock: effects of multiscale heterogeneity of hydrogeologic properties." *Journal of Contaminant Hydrology,* 60, 1-30. New York, New York: Elsevier.

9.0 SOFTWARE

The following computer programs are used in this task and controlled according to QAP-3.2: *Software Management.*

Table 11: Lists of software used in this task

Program Name	Version	Tracking Number	Purpose	Computer
TOUGH2	V 1.6	NSHE-1.6-025	Conduct the flow simulation	Windows XP
T2R3D	V 1.4	NSHE-1.4-026	Conduct the transport simulation	Windows XP
LHS	V 2.51	NSH-2.51-023	Generate the homogeneous random fields	Windows XP
GAMV	V 2.0	10439-2.0-00	Calculate the sample variogram	Windows NT 4.0
GSLIB V1.4SGSIM	V 1.41	10110-1.4SGSIMV1.41-00	Generate the heterogeneous random fields	Windows 2000
Tecplot	V 10.0	N/A	Plot various figures to present the results	Windows XP
Intel Visual Fortran	V 8.0	N/A	Compile the source code to run the software	Windows XP

10.0 ATTACHMENTS

The work of "stochastic analysis of transient flow in unsaturated heterogeneous porous media" using the KLME method" in Attachment A is non-quality affecting subtask and the results and conclusions are as non-Q for information purpose only.

A. Stochastic Analysis of Transient Flow in Unsaturated Heterogeneous Porous Media Using the KLME Method

In this study, we develop a general nonstationary stochastic model for transient flow in unsaturated randomly heterogeneous media using the moment-equation approach based on Karhunen-Loeve decomposition (KLME). The widely used van Genuchten-Mualem constitutive relationship for the unsaturated media is used in developing the model. The KLME approach is developed on the basis of the Karhunen-Loeve (KL) decomposition, polynomial expansion, and perturbation method. The soil parameter $n(\boldsymbol{x})$, the log-transformed pore size distribution parameter $\beta(\boldsymbol{x}) = \ln \alpha(\boldsymbol{x})$, and the log-transformed saturated hydraulic conductivity $f(\boldsymbol{x}) = \ln K_s(\boldsymbol{x})$ are treated as random space functions, which are normally distributed with a separable exponential covariance model. We decomposed $f(\boldsymbol{x})$, $\beta(\boldsymbol{x})$ and $n(\boldsymbol{x})$ as infinite series in a set of orthogonal normal random variables by the Karhunen-Loeve expansion and expand the pressure head as polynomial chaos with the same set of orthogonal random variables.

We first derive a series of partial differential equations in which the dependent variables are the deterministic coefficients of the polynomial chaos expansion and then solve these equations with the method of finite differences. The random representation of pressure head is obtained by combining the deterministic coefficients obtained and the random variables from the Kahunen-Loeve expansion of the input random functions. Finally, the pressure head moments are determined directly from the random representation of the pressure head. We demonstrate the model with some two-dimensional examples of unsaturated flows, and compare the results with those from the moment-based stochastic model. The validity of the developed KL-based stochastic model is also confirmed through high-resolution Monte Carlo simulations. It is shown that the KL-based model is more computationally efficient than the conventional moment-based model and the Monte Carlo method. We also investigate the relative contributions of the soil variabilities (K_S, α and n) to the pressure head variance.

A.1 Introduction

Although geologic formations exhibit a high degree of spatial variability, medium properties, including fundamental parameters such as permeability and porosity, are usually observed only at a few locations due to the high cost associated with subsurface measurements. This combination of significant spatial heterogeneity with a relatively small number of observations leads to uncertainty about the values of medium properties and thus, and to uncertainty in predicting flow and solute transport in such media. It has been recognized that the theory of stochastic processes provides a powerful natural method for evaluating flow and transport uncertainties. Many stochastic theories have been developed to study the effects of spatial variability on flow and transport in saturated

zones (e.g., Dagan, 1989; Gelhar, 1993; Zhang, 2002) and in unsaturated zones (e.g., Dagan and Bresler, 1979; Bresler and Dagan, 1981; Andersson and Shapiro, 1983; Yeh et al., 1985a,b; Hopmans et al., 1988; Destouni and Cvetkovic, 1989; Polmann et al., 1991; Mantoglou, 1992; Indelman et al., 1993; Liedl, 1994; Russo, 1993, 1995a,b; Harter and Yeh, 1996a,b; Zhang and Winter, 1998; Zhang et al., 1998; Zhang, 1999; Tartakovsky et al., 1999; Foussereau et al., 2000; Lu et al., 2000; Lu et al., 2002, Yang et al., 2004). Early stochastic studies focused on steady-state, gravity-dominated unsaturated flow in unbounded domains (e.g., Yeh et al., 1985a,b; Russo, 1993, 1995a,b; Yang et al., 1996; Zhang et al., 1998; Harter and Zhang, 1999). Recently, some researchers investigated the effects of boundary conditions on steady-state flow and hence the effects of flow nonstationarity in one-dimensional semi-bounded domains (Andersson and Shapiro, 1983; Indelman et al., 1993) or two-dimensional bounded domains (Zhang and Winter, 1998), a number of studies looked at transient unsaturated flows (Protopapas and Bras, 1990; Unlu et al., 1990; Mantoglou, 1992; Liedl, 1994; Zhang, 1999; Foussereau et al., 2000) and transient unsaturated-saturated flow (Li and Yeh, 1998; Ferrante and Yeh, 1999; Zhang and Lu, 2002). More recently, some researchers study the saturated-unsaturated flow in heterogeneous media by using the conventional moment equation method (Lu and Zhang, 2003) or the KLME method (Yang et al., 2004).

To describe unsaturated flow, the constitutive relationships of unsaturated hydraulic conductivity K versus pressure head ψ and effective water content θ_e versus ψ must be specified. Three models are commonly used to describe these functional relationships: the van Genuchten--Mualem model (van Genuchten, 1980), the Brooks-Corey model (Brooks and Corey, 1964), and the Gardner-Russo model (Gardner, 1958; Russo, 1988). Most existing stochastic analyses utilize the Gardner-Russo model due to its simplicity (e.g., Yeh et al., 1985a,b; Yeh, 1989; Russo, 1993, 1995a,b; Yang et al., 1996; Harter and Yeh, 1996a,b; Zhang, 1999; Tartakovsky et al., 1999; Lu et al., 2000; Zhang and Lu, 2002). On the other hand, the more complex van Genuchten — Mualem and Brooks — Corey models usually fit measured $K(\psi)$ and $\theta(\psi)$ data better. However, due to its mathematical complexity, the van Genuchten-Mualem model is seldom used in stochastic modeling of unsaturated flow in heterogeneous media although it is the most commonly used model for deterministic numerical modeling. Zhang et al. (1998) investigated the impact of different constitutive models on the results of stochastic analyses of steady-state, gravity-dominated flow. On the basis of the van Genuchten-Mualem model, Hughson and Yeh (2000) have recently developed a geostatistical inverse approach to flow in variably saturated media, in which the flow covariances are derived with a space-state approach. Lu and Zhang (2003) develop a stochastic model based on the conventional moment equation (CME) method for transient flow in heterogeneous unsaturated-saturated media with the van Genuchten-Mualem constitutive model. The major problem with the CME method is substantial requirement on computation resources even only when the first order moment is evaluated. This requirement limits the application of the method to small-scale simulation problems. Zhang and Lu (2004) developed a method called KLME method, which combined Kahunen-Loeve decomposition with the polynomial chaos expansion, to perform the stochastic analysis of saturated flow. It has been demonstrate that KLME method is capable of evaluating higher-order approximations of the dependent variables (pressure head and flux) moments and is more efficient and accurate than CME and Monte Carlo approaches (Zhang and Lu, 2004; Lu and Zhang, 2004a, 2004b, 2005). Yang et al. (2004) applied KLME method to analysis of saturated-unsaturated flow on the basis of Gardner-Russo model. Chen et al. (2005, 2006) extended the KLME method to analysis of steady-state and transient water-oil two-phase flow system.

In this study, on the basis of the van Genuchten-Mualem constitutive relationship, we develop a general nonstationary stochastic model for transient flow in unsaturated randomly heterogeneous media using the KLME method. We first derive a series of partial differential equations in the zeroth- and first-order by Karhonen-Loeve expansion of independent variables and polynomial chaos expansion of dependentin variables and then solve these equations with the method of finite differences. The random representation of pressure head is obtained by combining the deterministic coefficients obtained and the random variables from the Kahunen-Loeve expansion of the input random functions. Finally, the moments pressure head is determined directly from the random representation of the pressure head. We demonstrate the model with some two-dimensional examples of unsaturated flows, and compared the results with those from the moment-based stochastic model. The validity of the developed KL-based stochastic model is also confirmed through high-resolution Monte Carlo simulations. We also investigate the relative contributions of the soil variabilities (K_S, α and n) to the pressure head variance. The stochastic model developed in this study is applicable to the entire domain of a bounded, multi-dimensional transient unsaturated flow system in the presence of deterministic recharge and sink/source as well as in the presence of multiscale, nonstationary medium features.

A.2 Stochastic Differential Equations

We consider transient flow in unaturated porous media satisfying the following continuity equation and Darcy's law:

$$C[\psi,\cdot]\frac{\partial\psi(\mathbf{x},t)}{\partial t}+\nabla\cdot\mathbf{q}(\mathbf{x},t)=g(\mathbf{x},t) \tag{A1}$$

$$q_i(\mathbf{x},t)=-K[\psi,\cdot]\frac{\partial}{\partial x_i}[\psi(\mathbf{x},t)+x_1], \tag{A2}$$

Subject to initial and boundary conditions:

$$\psi(\mathbf{x},0)=\Psi_0(\mathbf{x}) \qquad \mathbf{x}\in\Omega \tag{A3}$$

$$\psi(\mathbf{x},t)=\Psi(\mathbf{x},t), \qquad \mathbf{x}\in\Gamma_D \tag{A4}$$

$$\mathbf{q}(\mathbf{x},t)\cdot n(\mathbf{x})=Q(\mathbf{x},t), \quad \mathbf{x}\in\Gamma_N \tag{A5}$$

where $\boldsymbol{q}$ is the specific discharge (flux), $\psi(\boldsymbol{x},t)+x_1$ is the total head, ψ is the pressure head, $i=1,\cdots,d$ (where d is the number of space dimensions), $\Psi_0(\boldsymbol{x})$ is the initial pressure head in the domain Ω, $\Psi(\boldsymbol{x},t)$ is the prescribed head on Dirichlet boundary segments Γ_D, $Q(\boldsymbol{x},t)$ is the prescribed flux across Neumann boundary segments Γ_N, $\boldsymbol{n}(\boldsymbol{x})=(n_1,\cdots,n_d)^T$ is an outward unit vector normal to the boundary, $C[\psi,\cdot]=d\theta_e/\mathrm{d}\psi$ is the specific moisture capacity, and $K[\psi,\cdot]$ is the unsaturated hydraulic conductivity (assumed to be isotropic locally). Both C and K are functions of pressure head and soil properties at $\boldsymbol{x}$. For convenience, they will be written as $C(\boldsymbol{x},t)$ and $K(\boldsymbol{x},t)$ in the sequel. The elevation x_1 is directed vertically upward. In these coordinates, recharge has a negative sign. The seepage velocity at $\boldsymbol{x}$ is related to the specific flux q_i by

$$u_i(\mathbf{x},t)=\frac{q_i(\mathbf{x},t)}{\theta_e(\mathbf{x},t)} \tag{A6}$$

where $\theta_e \equiv \theta_e[\psi(\mathbf{x},t),\cdot]$ is the effective volumetric water content at $\mathbf{x}$, which depends on ψ and soil properties.

It is clear that some model is needed to describe the constitutive relationships of K versus ψ and θ_e versus ψ when the flow is unsaturated. No universal models are available for the constitutive relationships. Instead, several empirical models are usually used, including the Gardner-Russo model (Gardner, 1958; Russo, 1988), the Brooks-Corey model (Brooks and Corey, 1964), and the van Genuchten-Mualem model (van Genuchten, 1980). Most analytical solutions of the deterministic unsaturated flow equations and most previous stochastic analyses used the Gardner-Russo model because of its simplicity. However, it is generally accepted that the more complex van Genuchten-Mualem and Brooks-Corey models may perform better than the simple Gardner-Russo model in describing measured data of $K(\psi)$ and $\theta_e(\psi)$. In this study, we use the van Genuchten-Mualem model:

$$K(\mathbf{x},t) = K_s(\mathbf{x})\sqrt{S(\mathbf{x},t)}\{1-[1-S^{1/m}(\mathbf{x},t)]^m\}^2, \tag{A7}$$

$$S(\mathbf{x},t) = \{1+[-\alpha(\mathbf{x})\psi(\mathbf{x},t)]^n\}^{-m}. \tag{A8}$$

In the above, $S(\mathbf{x},t)= \theta_e/(\theta_s-\theta_r)$ is the effective saturation, θ_r is the residual (irreducible) water content, θ_s is the saturated water content, α and n are fitting parameters, and $m = 1 - 1/n$. With (8), $C(\mathbf{x},t) = d\theta_e/\mathrm{d}\psi$ can be expressed explicitly as

$$C(\mathbf{x},t) = \alpha(\mathbf{x})[n(\mathbf{x})-1](\theta_s-\theta_r)S^{1/m}(\mathbf{x},t)[1-S^{1/m}(\mathbf{x},t)]^m \tag{A9}$$

In this study, θ_S and θ_r are assumed to be deterministic as their variabilities are likely to be small compared to that of the effective water content θ_e (Russo and Bouton, 1992). The soil parameters $n(\mathbf{x})$, the log-transformed pore size distribution parameter $\beta(\mathbf{x}) = \ln \alpha(\mathbf{x})$, and the log-transformed saturated hydraulic conductivity $f(\mathbf{x}) = \ln K_s(\mathbf{x})$ are treated as random space functions. Although the distributional forms of the soil parameters need not be specified for the subsequent derivations of moment equations, they must be specified in the Monte Carlo simulations designed to verify the derived moment equations. Here the fitting parameter $n(\mathbf{x})$ is assumed to follow a normal distribution while the saturated hydraulic conductivity $K_S(\mathbf{x})$ and the pore size distribution $\alpha(\mathbf{x})$ to follow log-normal distributions. The particular distributional assumptions made are consistent with the finding of Russo and Bouton (1992) based on field data. In turn, the governing equations (A1)-(A5) become a set of stochastic partial differential equations whose solutions are no longer deterministic values but are probability distributions or related quantities such as statistical moments of the dependent variables.

In this study, the soil properties (i.e., f, β, and n) are generally treated as (spatially and/or temporally) nonstationary random space functions (random fields). Thus, the expected values may be space-time dependent and the covariances may depend on the actual points in space-time rather than only on their space-time lags. As discussed in Zhang (2002), multiscale medium features such as distinct soil layers, zones and facies may cause the soil properties $f(\mathbf{x})$, $\beta(\mathbf{x})$, and $n(\mathbf{x})$ to be spatially nonstationary.

In the next section, we derive equations governing the first two moments (means and covariances) of the flow quantities in an unsaturated system. For simplicity, we assume that the soil properties $f(\mathbf{x})$, $\beta(\mathbf{x})$, and $n(\mathbf{x})$ are independent of each other. The moment equation procedure given below can be easily extended to incorporate other correlations between the various random variables.

A.3 Moment Differential Equations

As is commonly done, we work with the log transformed unsaturated hydraulic conductivity $Y(\mathbf{x},t) = \ln K(\mathbf{x},t)$,

$$Y(\mathbf{x},t) = f(\mathbf{x}) + \frac{1}{2}\ln S(\mathbf{x},t) + 2\ln\{1-[1-S^{1/m}(\mathbf{x},t)]^m\} \tag{A10}$$

Let $C_S(\mathbf{x},t) = S_S H[\psi(\mathbf{x},t)] + H[-\psi(\mathbf{x},t)]C(\mathbf{x},t)$. As $C \equiv 0$ for $\psi \geq 0$, we have

$$C_S(\mathbf{x},t) = S_S H[\psi(\mathbf{x},t)] + C(\mathbf{x},t) \tag{A11}$$

It is seen from (A8) and (A10) that $Y(\mathbf{x},t)$ is a function of the random fields f, β, n, and ψ. As shown in the text, we decompose them as follows: $f(\mathbf{x}) = \langle f(\mathbf{x})\rangle + f'(\mathbf{x})$, $\beta(\mathbf{x}) = \langle\beta(\mathbf{x})\rangle + \beta'(\mathbf{x})$, $n(\mathbf{x}) = \langle n(\mathbf{x})\rangle + n'(\mathbf{x})$, and $\psi(\mathbf{x},t) = \psi^{(0)}(\mathbf{x},t) + \psi^{(1)}(\mathbf{x},t) + \cdots$. Expand $Y(\mathbf{x},t)$ by Taylor series around $\langle f\rangle$, $\langle\beta\rangle$, $\langle n\rangle$, and $\psi^{(0)}$,

$$\begin{aligned} Y(\mathbf{x},t) &= \langle f(\mathbf{x})\rangle + \frac{1}{2}\ln S^{(0)}(\mathbf{x},t) + 2\ln\{1-[1-S_0^{1/m_0}]^{m_0}\} \\ &+ f'(\mathbf{x}) + h_{100}(\mathbf{x},t)\psi^{(1)}(\mathbf{x},t) + h_{010}(\mathbf{x},t)\beta'(\mathbf{x}) + h_{001}(\mathbf{x},t)n'(\mathbf{x}) + \cdots \end{aligned} \tag{A12}$$

where $S_0 = S^{(0)}(\mathbf{x},t) = \left\{1+[-e^{\langle\beta(\mathbf{x})\rangle}\psi^{(0)}(\mathbf{x},t)]^{\langle n\rangle}\right\}^{-m_0}$, $m_0 = 1-1/\langle n(\mathbf{x})\rangle$, and $h_{ijk} = \partial^{i+j+k}Y(\mathbf{x},t)/\partial\psi^i\partial\beta^j\partial n^k$ evaluated at $\langle\beta\rangle$, $\langle n\rangle$, S_0, and $\psi^{(0)}$. The terms h_{ijk} can be evaluated with the aid of

$$\frac{\partial Y}{\partial\psi} = \frac{\partial Y}{\partial S}\frac{\partial S}{\partial\psi} \tag{A13}$$

$$\frac{\partial Y}{\partial\beta} = \frac{\partial Y}{\partial S}\frac{\partial S}{\partial\beta} \tag{A14}$$

$$\frac{\partial Y}{\partial n} = \frac{\partial Y}{\partial S}\frac{\partial S}{\partial n} - \frac{2}{n^2}\frac{(1-S^{1/m})^m}{1-(1-S^{1/m})^m}\left[\ln(1-S^{1/m}) + \frac{S^{1/m}}{1-S^{1/m}}\frac{\ln S}{m}\right] \tag{A15}$$

where

$$\frac{\partial Y}{\partial S} = \frac{1}{2S} + \frac{2(S^{-1/m} - 1)^{m-1}}{1 - (1 - S^{1/m})^m} \tag{A16}$$

$$\frac{\partial S}{\partial \psi} = -\frac{n-1}{\psi} S(1 - S^{1/m}) \tag{A17}$$

$$\frac{\partial S}{\partial \beta} = -(n-1) S(1 - S^{1/m}) \tag{A18}$$

$$\frac{\partial S}{\partial n} = -\frac{1}{n(n-1)} S \ln S - \frac{m}{n} S(1 - S^{1/m}) \ln(S^{-1/m} - 1) \tag{A19}$$

Substituting (A2) into (A1) and utilizing $Y(\mathbf{x},t) = \ln K(\mathbf{x},t)$ yields

$$\frac{\partial^2 \psi(\mathbf{x},t)}{\partial x_i^2} + \frac{\partial Y(\mathbf{x},t)}{\partial x_i}\left[\frac{\partial \psi(\mathbf{x},t)}{\partial x_i} + \delta_{i1}\right] = C(\mathbf{x},t) e^{-Y(\mathbf{x},t)} \frac{\partial \psi(\mathbf{x},t)}{\partial t} - g(\mathbf{x},t) e^{-Y(\mathbf{x},t)} \tag{A20}$$

$$\psi(\mathbf{x},0) = \Psi_0(\mathbf{x}), \quad \mathbf{x} \in \Omega \tag{A21}$$

$$\psi(\mathbf{x},t) = \Psi(\mathbf{x},t), \quad \mathbf{x} \in \Gamma_D \tag{A22}$$

$$n_i(\mathbf{x}) e^{Y(\mathbf{x},t)}\left[\frac{\partial \psi(\mathbf{x},t)}{\partial x_i} + \delta_{i1}\right] = -Q(\mathbf{x},t), \quad \boldsymbol{x} \in \Gamma_N \tag{A23}$$

where δ_{i1} is the Kronecker delta function. Summation for repeated indices is implied. Because the variability of $\psi(\boldsymbol{x},t)$ depends on the input variabilities, i.e., those of the soil properties (f, β, and n) and those of the initial/boundary and source/sink terms, and the variabilities of Y and C depend on those of ψ and the input variables, one may express these quantities as infinite series in the following form: $\psi(\boldsymbol{x},t) = \psi^{(0)} + \psi^{(1)} + \psi^{(2)} + \cdots$, $Y(\boldsymbol{x},t) = Y^{(0)} + Y^{(1)} + Y^{(2)} + \cdots$, and $C(\boldsymbol{x},t) = C^{(0)} + C^{(1)} + C^{(2)} + \cdots$. In these series, the order of each term is with respect to σ, which, to be clear later, is some combination of the variabilities of the input variables. By writing $Y(\boldsymbol{x},t) = Y^{(0)} + Y^{(1)} + \cdots$, we have from Eq. (A12)

$$Y^{(0)}(\mathbf{x},t) = \langle f(\mathbf{x}) \rangle + \frac{1}{2} \ln S_0 + 2 \ln\{1 - [1 - S_0^{1/m_0}]^{m_0}\} \tag{A24}$$

$$Y^{(1)}(\mathbf{x},t) = f'(\mathbf{x}) + h_{100}(\mathbf{x},t) \psi^{(1)}(\mathbf{x},t) + h_{010}(\mathbf{x},t) \beta'(\mathbf{x}) + h_{001}(\mathbf{x},t) n'(\mathbf{x}) \tag{A25}$$

Similarly, we may expand $C(\boldsymbol{x},t)$ in (A11) by Taylor series

$$\begin{aligned} C(\mathbf{x},t) &= e^{\langle \beta(\mathbf{x}) \rangle}\left[\langle n(\mathbf{x}) \rangle - 1\right](\theta_s - \theta_r) S_0^{1/m_0}\left[1 - S_0^{1/m_0}\right]^{m_0} \\ &+ p_{100}(\mathbf{x},t) \psi^{(1)}(\mathbf{x},t) + p_{010}(\mathbf{x},t) \beta'(\mathbf{x}) + p_{001}(\mathbf{x},t) n'(\mathbf{x}) + \cdots \end{aligned} \tag{A26}$$

where $p_{ijk} = \partial^{i+j+k} C(\boldsymbol{x},t) / \partial \psi^i \partial \beta^j \partial n^k$ evaluated at $\langle \beta \rangle$, $\langle n \rangle$, S_0, and $\psi^{(0)}$. The terms p_{ijk} can be evaluated with the aid of

$$\frac{\partial C}{\partial S} = \alpha(\theta_S - \theta_r)S^{1/m-1}(1-S^{1/m})^{m-1}[n-(2n-1)S^{1/m}] \tag{A27}$$

$$\frac{\partial C}{\partial \psi} = \frac{\partial C}{\partial S}\frac{\partial S}{\partial \psi} \tag{A28}$$

$$\frac{\partial C}{\partial \beta} = \frac{\partial C}{\partial S}\frac{\partial S}{\partial \beta} + C(\mathbf{x},t) \tag{A29}$$

$$\frac{\partial C}{\partial n} = \frac{\partial C}{\partial S}\frac{\partial S}{\partial n} + \frac{C(\mathbf{x},t)}{(n-1)^2}\left\{n-1+m^2\ln(1-S^{1/m}) + \frac{(m+1)S^{1/m}-1}{1-S^{1/m}}\ln S\right\} \tag{A30}$$

By writing $C(\boldsymbol{x},t) = C^{(0)} + C^{(1)} + \cdots$, we have from Eq. (A26)

$$C^{(0)}(\mathbf{x},t) = e^{\langle\beta(\mathbf{x})\rangle}\left\langle\left\langle n(\mathbf{x})\right\rangle - 1\right\rangle(\theta_S - \theta_r)S_0^{1/m_0}[1-S_0^{1/m_0}]^{m_0} \tag{A31}$$

$$C^{(1)}(\mathbf{x},t) = p_{100}(\mathbf{x},t)\psi^{(1)}(\mathbf{x},t) + p_{010}(\mathbf{x},t)\beta'(\mathbf{x}) + p_{001}(\mathbf{x},t)n'(\mathbf{x}) \tag{A32}$$

After substituting these and the following formal decompositions into (A20)-(A23) and collecting terms at separate order, we obtain

$$\frac{\partial^2\psi^{(0)}(\mathbf{x},t)}{\partial x_i^2} + \frac{\partial Y^{(0)}(\mathbf{x},t)}{\partial x_i}\left[\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial x_i} + \delta_{i1}\right] = \frac{C^{(0)}(\mathbf{x},t)}{K_m(\mathbf{x},t)}\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial t} - \frac{g(\mathbf{x},t)}{K_m(\mathbf{x},t)} \tag{A33}$$

$$\psi^{(0)}(\mathbf{x},0) = \Psi_0(\mathbf{x}), \quad \boldsymbol{x} \in \Omega \tag{A34}$$

$$\psi^{(0)}(\mathbf{x},t) = \Psi(\mathbf{x},t), \quad \boldsymbol{x} \in \Gamma_D \tag{A35}$$

$$n_i(\mathbf{x})\left[\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial x_i} + \delta_{i1}\right] = -\frac{Q(\mathbf{x},t)}{K_m(\mathbf{x},t)}, \quad \boldsymbol{x} \in \Gamma_N \tag{A36}$$

$$\begin{aligned}\frac{\partial^2\psi^{(1)}(\mathbf{x},t)}{\partial x_i^2} &+ J_i(\mathbf{x},t)\frac{\partial Y^{(1)}(\mathbf{x},t)}{\partial x_i} + \frac{\partial Y^{(0)}(\mathbf{x},t)}{\partial x_i}\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial x_i} \\ &= \frac{C^{(0)}(\mathbf{x},t)}{K_m(\mathbf{x},t)}\left[\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial t} - J_t(\mathbf{x},t)Y^{(1)}(\mathbf{x},t)\right] \\ &+ \frac{C^{(1)}(\mathbf{x},t)}{K_m(\mathbf{x},t)}J_t(\mathbf{x},t) + \frac{1}{K_m(\mathbf{x},t)}\left[g(\mathbf{x},t)Y^{(1)}(\mathbf{x},t)\right]\end{aligned} \tag{A37}$$

$$\psi^{(1)}(\mathbf{x},0) = 0, \ \boldsymbol{x} \in \Omega \tag{A38}$$

$$\psi^{(1)}(\mathbf{x},t) = 0, \ \boldsymbol{x} \in \Gamma_D \tag{A39}$$

$$n_i(\mathbf{x})\left[\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial x_i} + J_i(\mathbf{x},t)Y^{(1)}(\mathbf{x},t)\right] = 0, \ \boldsymbol{x} \in \Gamma_N \tag{A40}$$

where $K_m(\boldsymbol{x},t)= \exp[Y^{(0)}(\boldsymbol{x},t)]$, and $J_i(\boldsymbol{x},t) = \partial\psi^{(0)}(\boldsymbol{x},t)/\partial x_i + \delta_{i1}$ and $J_t(\boldsymbol{x},t) = \partial\psi^{(0)}(\boldsymbol{x},t)/\partial t$ are the respective spatial and temporal mean gradient of (total) head. It can be shown that $\langle\psi^{(0)}\rangle = \psi^{(0)}$, and $\langle\psi^{(1)}\rangle = 0$. Hence, the mean pressure head is $\langle\psi\rangle = \psi^{(0)}$ to zeroth- or first-order in σ. The head fluctuation is $\psi' = \psi^{(1)}$ to first-order. Therefore, the head covariance is $C_\psi(\boldsymbol{x},t;\boldsymbol{\chi},\tau) = \langle\psi^{(1)}(\boldsymbol{x},t)\psi^{(1)}(\boldsymbol{\chi},\tau)\rangle$ to first-order in σ^2 (or second-order in σ).
Substituting (A24) and (A31) into (A33)-(A36) yields

$$\frac{\partial^2\psi^{(0)}(\mathbf{x},t)}{\partial x_i^2} + a_i(\mathrm{x},t)\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial x_i} = e(\mathbf{x},t)\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial t} - \frac{g(\mathbf{x},t)}{K_m(\mathbf{x},t)} + d(\mathbf{x},t) \tag{A41}$$

$$\psi^{(0)}(\mathbf{x},0) = \Psi_0(\mathbf{x}), \quad \boldsymbol{x}\in\Omega \tag{A42}$$

$$\psi^{(0)}(\mathbf{x},t) = \Psi(\mathbf{x},t), \quad \boldsymbol{x}\in\Gamma_D \tag{A43}$$

$$n_i(\mathbf{x})\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial x_i} = -\frac{Q(\mathbf{x},t)}{K_m(\mathbf{x},t)} - \delta_{i1}n_i(\mathbf{x}), \quad \boldsymbol{x}\in\Gamma_N \tag{A44}$$

where

$$K_m(\mathbf{x},t) = e^{\langle f(\mathbf{x})\rangle}\sqrt{S_0}\left[1-(1-S_0^{1/m_0})^{1/m_0}\right]^2$$

$$a_i(\mathbf{x},t) = \frac{\partial Y^{(0)}}{\partial S_0}\frac{\partial S_0}{\partial x_i} + Y^{(0)'}_{\langle n\rangle}\frac{\partial\langle n(\mathbf{x})\rangle}{\partial x_i}$$

$$d(\mathbf{x},t) = -J_i(\mathbf{x},t)\frac{\partial\langle f(\mathbf{x})\rangle}{\partial x_i} - \delta_{i1}a_i(\mathbf{x},t)$$

$$e(\mathbf{x},t) = C^{(0)}(\mathbf{x},t)/K_m(\mathbf{x},t)$$

$$\frac{\partial S_0}{\partial x_i} = \frac{\partial S_0}{\partial\psi^{(0)}}\frac{\partial\psi^{(0)}(\mathbf{x},t)}{\partial x_i} + \frac{\partial S_0}{\partial\langle\beta\rangle}\frac{\partial\langle\beta(\mathbf{x})\rangle}{\partial x_i} + \frac{\partial S_0}{\partial\langle n\rangle}\frac{\partial\langle n(\mathbf{x})\rangle}{\partial x_i} \tag{A45}$$

Here $Y^{(0)'}_{\langle n\rangle}$ is the partial derivative of $Y^{(0)}$ with respect to $\langle n\rangle$ without considering S_0 as an implicit function of $\langle n\rangle$. It in fact equals to the second term in the right side of (A4), evaluated at $\langle\beta\rangle$, $\langle n\rangle$, and $\langle\psi\rangle$.

Substituting (A25) and (A32) into (A37)-(A40) yields

$$\frac{\partial^2\psi^{(1)}(\mathbf{x},t)}{\partial x_i^2} + b_i(\mathbf{x},t)\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial x_i} + c(\mathbf{x},t)\psi^{(1)}(\mathbf{x},t)$$
$$= e(\mathbf{x},t)\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial t}$$
$$- J_i(\mathbf{x},t)\frac{\partial f'(\mathbf{x})}{\partial x_i} - J_i(\mathbf{x},t)h_{010}(\mathbf{x},t)\frac{\partial\beta'(\mathbf{x})}{\partial x_i} - J_i(\mathbf{x},t)h_{001}(\mathbf{x},t)\frac{\partial n'(\mathbf{x})}{\partial x_i}$$

$$+ d_1(\mathbf{x},t) f'(\mathbf{x}) + d_2(\mathbf{x},t)\beta'(\mathbf{x}) + d_3(\mathbf{x},t) n'(\mathbf{x}) \quad \text{(A46)}$$

$$\psi^{(1)}(\mathbf{x},0) = 0, \quad \boldsymbol{x} \in \Omega \quad \text{(A47)}$$

$$\psi^{(1)}(\mathbf{x},t) = 0\,, \quad \boldsymbol{x} \in \Gamma_D \quad \text{(A48)}$$

$$n_i(\mathbf{x})\frac{\partial \psi^{(1)}(\mathbf{x},t)}{\partial x_i} + d_0(\mathbf{x},t)\psi^{(1)}(\mathbf{x},t) = d_4(\mathbf{x},t) f'(\mathbf{x}) + d_4(\mathbf{x},t) h_{010}(\mathbf{x},t)\beta'(\mathbf{x})$$
$$+ d_4(\mathbf{x},t) h_{001}(\mathbf{x},t) n'(\mathbf{x}), \quad \boldsymbol{x} \in \Gamma_N \quad \text{(A49)}$$

where

$$b_i(\mathbf{x},t) = J_i(\mathbf{x},t) h_{100}(\mathbf{x},t) + \partial Y^{(0)}(\mathbf{x},t)/\partial x_i$$
$$c(\mathbf{x},t) = -h_{100}(\mathbf{x},t) d_1(\mathbf{x},t) - p_{100}(\mathbf{x},t)\frac{J_t(\mathbf{x},t)}{K_m(\mathbf{x},t)} + J_i(\mathbf{x},t)\frac{\partial h_{100}(\mathbf{x},t)}{\partial x_i}$$
$$d_0(\mathbf{x},t) = n_i(\mathbf{x}) J_i(\mathbf{x},t) h_{100}(\mathbf{x},t)$$
$$d_1(\mathbf{x},t) = \left(g(\mathbf{x},t) - C^{(0)}(\mathbf{x},t) J_t(\mathbf{x},t)\right)/ K_m(\mathbf{x},t)$$
$$d_2(\mathbf{x},t) = h_{010}(\mathbf{x},t) d_1(\mathbf{x},t) - J_i(\mathbf{x},t)\partial h_{010}(\mathbf{x},t)/\partial x_i + J_t(\mathbf{x},t) p_{010}(\mathbf{x},t)/K_m(\mathbf{x},t)$$
$$d_3(\mathbf{x},t) = h_{001}(\mathbf{x},t) d_1(\mathbf{x},t) - J_i(\mathbf{x},t)\partial h_{001}(\mathbf{x},t)/\partial x_i + J_t(\mathbf{x},t) p_{001}(\mathbf{x},t)/K_m(\mathbf{x},t)$$
$$d_4(\mathbf{x},t) = -n_i(\mathbf{x}) J_i(\mathbf{x},t) \quad \text{(A50)}$$

A.4 KL-based Moment Equations (KLME)

Using the KL method (Zhang and Lu, 2004), we have

$$p'(\mathbf{x},\omega) = \sum_{m=1}^{\infty} \xi_m^p(\omega)\sqrt{\lambda_m^p} F_m^p(\mathbf{x}) \quad \text{(A51)}$$

$$\psi^{(1)}(\mathbf{x},t) = \sum_{p=f,\beta,n} \sum_{m=1}^{\infty} \xi_m^p(\omega)\ \psi_{\mathrm{m}}^{(1,p)}(\mathbf{x},t) \quad \text{(A52)}$$

where λ_m^p and $F_m^p(\boldsymbol{x})$ are eigenvalues and their corresponding eigenfunctions for parameters $p = f$, β, or n, ξ_m^p are orthogonal Gaussian random variables satisfying $\left\langle \xi_m^p \right\rangle = 0$ and $\left\langle \xi_m^p \xi_n^p \right\rangle = \delta_{mn}$, where δ_{mn} is the Kronecker delta function. The expansion in Eq. (A52) is called the Kahunen-Loeve expansion. In this study, we use a separable exponential covariance function of a 2-D illustrative example given by

$$C_p(\mathrm{h}) = \sigma_{\mathrm{p}}^2 \exp(-|\mathrm{h}|/\lambda_{\mathrm{p}}) \quad \text{(A53)}$$

where $p = f$, β, or n, σ_p^2 is the variance of p, λ_p is the correlation scale of p, and **h** is the separation vector. For this kind of covariance function, the analytical solution of eigenvalues and eigenfunctions can be found from ollowing Fredholm equation (Zhang and Lu, 2004):

$$\int_D C_p(\mathbf{x},\mathbf{y})F(\mathbf{x})d\mathbf{x} = \lambda F(\mathbf{y}) \tag{A54}$$

For the general case, the eigenvalues and eigenfunctions have to be solved numerically via iterative methods or a Galerkin-type method (Ghanem and Spanos, 1991).

For simplicity, the arguments $(\mathbf{x},t)$ are omitted in the following expression. Substituting (A52) into (A46)-(A49), we can obtain

$$\begin{aligned}
&\sum_{p=f,\beta,n}\sum_{m=1}^{\infty}\xi_m^p\left[\frac{\partial^2\psi_m^{(1,p)}}{\partial x_i^2}+b_i(\mathbf{x},t)\frac{\partial\psi_m^{(1,p)}}{\partial x_i}+c(\mathbf{x},t)\psi_m^{(1,p)}-e(\mathbf{x},t)\frac{\partial\psi_m^{(1,p)}}{\partial t}\right]\\
&= -J_i(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^f\sqrt{\lambda_m^f}\frac{\partial \mathrm{F_m^f}}{\partial \mathrm{x_i}} - J_i(\mathbf{x},t)h_{010}(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^\beta\sqrt{\lambda_m^\beta}\frac{\partial \mathrm{F_m^\beta}}{\partial \mathrm{x_i}}\\
&-J_i(\mathbf{x},t)h_{001}(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^n\sqrt{\lambda_m^n}\frac{\partial \mathrm{F_m^n}}{\partial \mathrm{x_i}}\\
&+d_1(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^f\sqrt{\lambda_m^f}\mathrm{F_m^f}+d_2(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^\beta\sqrt{\lambda_m^\beta}\mathrm{F_m^\beta}+d_3(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^n\sqrt{\lambda_m^n}\mathrm{F_m^n}
\end{aligned} \tag{A55}$$

$$\psi^{(1)}(\boldsymbol{x},0)=\sum_{p=f,\beta,n}\xi_m^p\psi_{\mathrm{m}}^{(1,p)}(x,0)=0 \qquad \boldsymbol{x}\in\Omega \tag{A56}$$

$$\psi^{(1)}(\boldsymbol{x},t)=\sum_{p=f,\beta,n}\xi_m^p\psi_{\mathrm{m}}^{(1,p)}(\boldsymbol{x},t)=0 \qquad \boldsymbol{x}\in\Gamma_D \tag{A57}$$

$$\begin{aligned}
&\sum_{p=f,\beta,n}\sum_{m=1}^{\infty}\xi_m^p\left[n_i(\boldsymbol{x},t)\frac{\partial\psi_m^{(1,p)}}{\partial x_i}+d_0(\boldsymbol{x},t)\psi_m^{(1,p)}\right]\\
&=d_4(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^f\sqrt{\lambda_m^f}\mathrm{F_m^f}+d_4(\mathbf{x},t)h_{010}(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^\beta\sqrt{\lambda_m^\beta}\mathrm{F_m^\beta}+d_4(\mathbf{x},t)h_{001}(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^n\sqrt{\lambda_m^n}\mathrm{F_m^n}\\
&\boldsymbol{x}\in\Gamma_N
\end{aligned} \tag{A58}$$

Multiplying (A55)-(A58) by ξ_m^f, ξ_m^β, ξ_m^n, respectively, taking the ensemble mean and considering the following properties of the orthogonality of sets ξ_m^f, ξ_m^β, ξ_m^n, m=1,2,…,

$$\left\langle\xi_m^f\xi_n^f\right\rangle=\delta_{mn},\ \left\langle\xi_m^\beta\xi_n^\beta\right\rangle=\delta_{mn},\ \left\langle\xi_m^n\xi_n^n\right\rangle=\delta_{mn},\ \left\langle\xi_m^f\xi_m^\beta\right\rangle=\left\langle\xi_m^f\xi_m^n\right\rangle=\left\langle\xi_m^n\xi_m^\beta\right\rangle=0,$$

We obtain

$$\begin{aligned}
&\frac{\partial^2\psi_m^{(1,f)}}{\partial x_i^2}+b_i(\mathbf{x},t)\frac{\partial\psi_m^{(1,f)}}{\partial x_i}+c(\mathbf{x},t)\psi_m^{(1,f)}-e(\mathbf{x},t)\frac{\partial\psi_m^{(1,f)}}{\partial t}\\
&= -J_i(\mathbf{x},t)\sqrt{\lambda_m^f}\frac{\partial \mathrm{F_m^f}}{\partial \mathrm{x_i}}+d_1(\mathbf{x},t)\sqrt{\lambda_m^f}\mathrm{F_m^f}
\end{aligned} \tag{A59}$$

$$\psi_{\mathrm{m}}^{(1,\mathrm{f})}(\mathbf{x},0)=0 \tag{A60}$$

$$\psi_{\mathrm{m}}^{(1,\mathrm{f})}(\mathbf{x},t)=0 \tag{A61}$$

$$n_i(\mathbf{x},t)\frac{\partial \psi_m^{(1,f)}}{\partial x_i}+d_0(\mathbf{x},t)\psi_m^{(1,f)}=d_4(\mathbf{x},t)\sqrt{\lambda_m^f}\,\mathrm{F}_{\mathrm{m}}^{\mathrm{f}} \tag{A62}$$

$$\frac{\partial^2 \psi_m^{(1,\beta)}}{\partial x_i^2}+b_i(\mathbf{x},t)\frac{\partial \psi_m^{(1,\beta)}}{\partial x_i}+c(\mathbf{x},t)\psi_m^{(1,\beta)}-e(\mathbf{x},t)\frac{\partial \psi_m^{(1,\beta)}}{\partial t}$$
$$=-J_i(\mathbf{x},t)h_{010}(\mathbf{x},t)\sqrt{\lambda_m^\beta}\,\frac{\partial \mathrm{F}_{\mathrm{m}}^\beta}{\partial \mathrm{x}_{\mathrm{i}}}+d_2(\mathbf{x},t)\sqrt{\lambda_m^\beta}\,\mathrm{F}_{\mathrm{m}}^\beta \tag{A63}$$

$$\psi_{\mathrm{m}}^{(1,\beta)}(\mathbf{x},0)=0 \tag{A64}$$

$$\psi_{\mathrm{m}}^{(1,\beta)}(\mathbf{x},t)=0 \tag{A65}$$

$$n_i(\mathbf{x},t)\frac{\partial \psi_m^{(1,\beta)}}{\partial x_i}+d_0(\mathbf{x},t)\psi_m^{(1,\beta)}=d_4(\mathbf{x},t)h_{010}(\mathbf{x},t)\sqrt{\lambda_m^\beta}\,\mathrm{F}_{\mathrm{m}}^\beta \tag{A66}$$

$$\frac{\partial^2 \psi_m^{(1,n)}}{\partial x_i^2}+b_i(\mathbf{x},t)\frac{\partial \psi_m^{(1,n)}}{\partial x_i}+c(\mathbf{x},t)\psi_m^{(1,n)}-e(\mathbf{x},t)\frac{\partial \psi_m^{(1,n)}}{\partial t}$$
$$=-J_i(\mathbf{x},t)h_{001}(\mathbf{x},t)\sqrt{\lambda_m^n}\,\frac{\partial \mathrm{F}_{\mathrm{m}}^n}{\partial \mathrm{x}_{\mathrm{i}}}+d_3(\mathbf{x},t)\sqrt{\lambda_m^n}\,\mathrm{F}_{\mathrm{m}}^n \tag{A67}$$

$$\psi_{\mathrm{m}}^{(1,n)}(\mathbf{x},0)=0 \tag{A68}$$

$$\psi_{\mathrm{m}}^{(1,n)}(\mathbf{x},t)=0 \tag{A69}$$

$$n_i(\mathbf{x},t)\frac{\partial \psi_m^{(1,n)}}{\partial x_i}+d_0(\mathbf{x},t)\psi_m^{(1,n)}=d_4(\mathbf{x},t)h_{001}(\mathbf{x},t)\sqrt{\lambda_m^n}\,\mathrm{F}_{\mathrm{m}}^n \tag{A70}$$

Up to the first-order in the σ_Y, the pressure head is approximated by

$$\psi(\boldsymbol{x},t)=\psi^{(0)}(\boldsymbol{x},t)+\psi^{(1)}(\boldsymbol{x},t) \tag{A71}$$

$$\langle\psi(\mathbf{x},t)\rangle=\psi^0(\mathbf{x},t) \tag{A72}$$

$$C_\psi(\boldsymbol{x},t;y,\tau)=\sum_{p=f,\beta,n}\sum_{m=1}^{\infty}\psi_m^{(1,p)}(\boldsymbol{x},t)\psi_m^{(1,p)}(y,\tau) \tag{A73}$$

$$\sigma_\psi^2(\boldsymbol{x},t)=\sum_{m=1}^{\infty}\left[\psi_m^{(1,f)}\right]^2+\sum_{m=1}^{\infty}\left[\psi_m^{(1,\beta)}\right]^2+\sum_{m=1}^{\infty}\left[\psi_m^{(1,n)}\right]^2 \tag{A74}$$

$$C_{f\psi}(\boldsymbol{x};y,\tau)=\sum_{m=1}^{\infty}\sqrt{\lambda_m^f}F_m^f(\boldsymbol{x})\psi_m^{(1,f)}(y,\tau) \tag{A75}$$

$$C_{\beta\psi}(\mathbf{x};y,\tau)=\sum_{m=1}^{\infty}\sqrt{\lambda_m^\beta}F_m^\beta(\mathbf{x})\psi_m^{(1,\beta)}(y,\tau) \tag{A76}$$

$$C_{n\psi}(\mathbf{x};y,\tau)=\sum_{m=1}^{\infty}\sqrt{\lambda_m^n}F_m^n(\mathbf{x})\psi_m^{(1,n)}(y,\tau) \tag{A77}$$

We now show how to derive the first two moments of flux. The flux in (A2) can be rewritten as

$$q_i(\mathbf{x},t)=-K_m(\mathbf{x},t)\left\{1+Y^{(1)}+Y^{(2)}+\frac{1}{2}\left[Y^{(1)}\right]^2+\cdots\right\}\left\{\frac{\partial}{\partial x_i}\left[\sum_{i=0}^{\infty}\psi^{(i)}\right]+\delta_{i1}\right\} \tag{A78}$$

Collecting terms at separate order, we have

$$q_i^{(0)}(\mathbf{x},t)=-K_m(\mathbf{x},t)J_i(\mathbf{x},t) \tag{A79}$$

$$\begin{aligned}
q_i^{(1)}(\mathbf{x},t)&=-K_m(\mathbf{x},t)\left\{J_i(\mathbf{x},t)Y^{(1)}(\mathbf{x},t)+\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial x_i}\right\}\\
&=-K_m(\mathbf{x},t)J_i(\mathbf{x},t)\left[f'(\mathbf{x})+h_{100}(\mathbf{x},t)\psi^{(1)}(\mathbf{x},t)+h_{010}(\mathbf{x},t)\beta'(\mathbf{x})+h_{001}(\mathbf{x},t)n'(\mathbf{x})\right]\\
&-K_m(\mathbf{x},t)\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial x_i}\\
&=-K_m(\mathbf{x},t)J_i(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^f\sqrt{\lambda_m^f}F_m^f(\mathbf{x})\\
&-K_m(\mathbf{x},t)J_i(\mathbf{x},t)h_{100}(\mathbf{x},t)\left[\sum_{m=1}^{\infty}\xi_m^f\,\psi_{\mathrm{m}}^{(1,\mathrm{f})}(\mathbf{x},t)+\sum_{m=1}^{\infty}\xi_m^\beta\,\psi_{\mathrm{m}}^{(1,\beta)}(\mathbf{x},t)+\sum_{m=1}^{\infty}\xi_m^n\,\psi_{\mathrm{m}}^{(1,\mathrm{n})}(\mathbf{x},t)\right]\\
&-K_m(\mathbf{x},t)J_i(\mathbf{x},t)h_{010}(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^\beta\sqrt{\lambda_m^\beta}F_m^\beta(\mathbf{x})-K_m(\mathbf{x},t)J_i(\mathbf{x},t)h_{001}(\mathbf{x},t)\sum_{m=1}^{\infty}\xi_m^n\sqrt{\lambda_m^n}F_m^n(\mathbf{x})\\
&-K_m(\mathbf{x},t)\frac{\partial\psi^{(1)}(\mathbf{x},t)}{\partial x_i}
\end{aligned} \tag{A80}$$

It can be shown that the mean flux is $\langle\mathbf{q}\rangle=\mathbf{q}^{(0)}=(q_1^{(0)},\cdots,q_d^{(0)})^T$ to zeroth- or first-order in σ, and the flux fluctuation is $\mathbf{q}'=\mathbf{q}^{(1)}=(q_1^{(1)},\cdots,q_d^{(1)})^T$ to first-order. Therefore, to first-order, the flux covariances are given as

$$C_{q\,q}(\mathbf{x},t;\boldsymbol{\chi},\tau)=\sum_{m=1}^{\infty}\left[q_m^{(1)}(\mathbf{x},t)q_m^{(1)}(\boldsymbol{\chi},\tau)\right] \tag{A81}$$

The moments of the effective water content can be derived similarly.
We may decompose the effective water content $\theta_e(\boldsymbol{x},t)=(\theta_s-\theta_r)S(\boldsymbol{x},t)$ into the zeroth-order mean and the first-order fluctuation,

$$\theta_e^{(0)}=(\theta_s-\theta_r)S^{(0)}(\mathbf{x},t)=(\theta_s-\theta_r)\left\{1+[-e^{\langle\beta(\mathbf{x})\rangle}\psi^{(0)}(\mathbf{x},t)]^{\langle n\rangle}\right\}^{-m_0} \tag{A82}$$

$$\theta_e^{(1)}(\mathbf{x},t)=(\theta_s-\theta_r)\left[s_{100}(\mathbf{x},t)\psi^{(1)}(\mathbf{x},t)+s_{010}(\mathbf{x},t)\beta'(\mathbf{x})+s_{001}(\mathbf{x},t)n'(\mathbf{x})\right] \tag{A83}$$

where $s_{ijk} = \partial^{i+j+k} S(\boldsymbol{x},t)/\partial\psi^i\partial\beta^j\partial n^k$ evaluated at $\langle\beta\rangle$, $\langle n\rangle$, and $\psi^{(0)}$.

A.5 Illustrative Examples

In this section, we attempt to demonstrate the applicability of the developed stochastic model to unsaturated flow in hypothetical soils. Although the general moment equations derived in Section 3 are applicable to any admissible stationary or nonstationary covariances with statistical anisotropy, in the examples we assume the log saturated hydraulic conductivity $f(\boldsymbol{x})$, the log pore size distribution parameter $\beta(\boldsymbol{x})$, and the fitting parameter $n(\boldsymbol{x})$ to be second-order stationary with an exponential covariance function (Eq.A53).

It is straightforward to extend the numerical moment equation approach to handle statistical nonstationarity and anisotropy. For simplicity, f, β, and n are further assumed to be uncorrelated in the examples.

A.5.1 Infiltration in Unsaturated Media

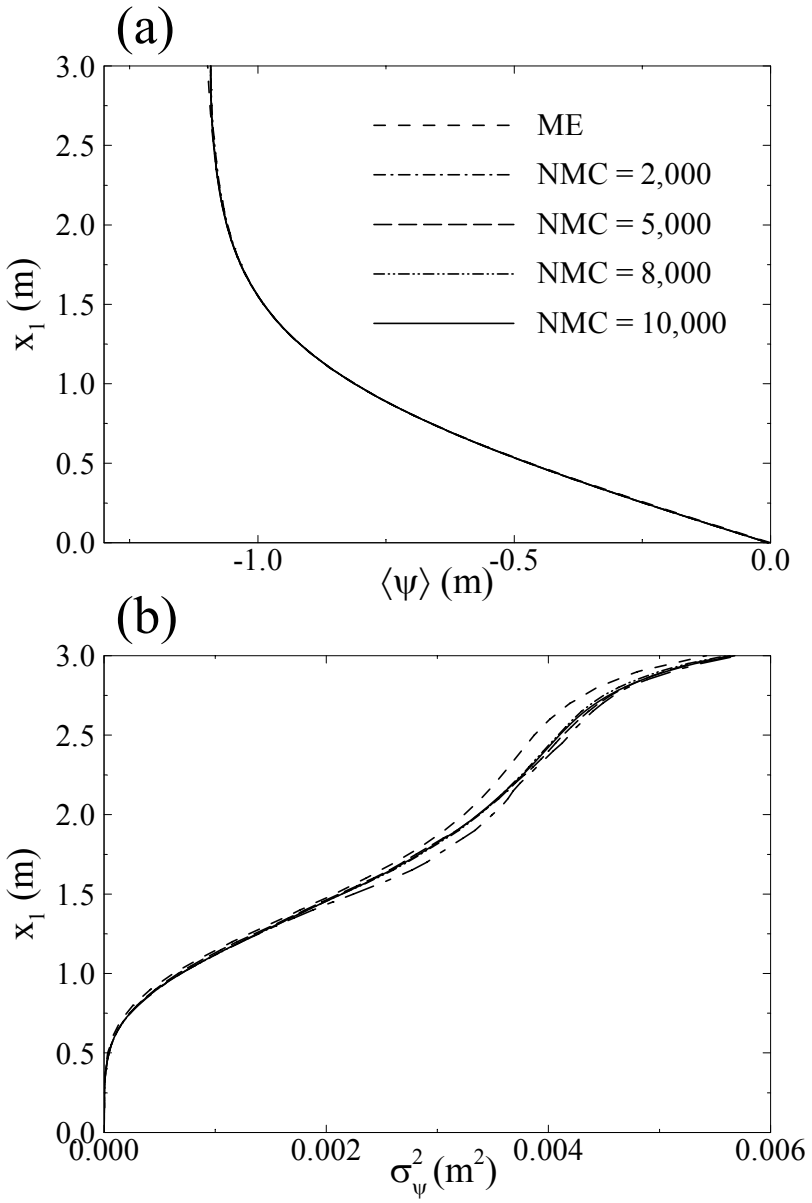
In this example, denoted as Case 1, we first try to show the validity of our mathematical derivation and numerical implementation by comparing our results with Monte Carlo simulations. We consider a square domain of 3m by 3m in a vertical cross-section, discretized into 30 × 60 rectangular elements of 0.1m by 0.05 m. The boundary conditions are specified as follows: a prescribed deterministic constant pressure head ψ = 0 (water table) at the bottom (x_1 = 0.0), a constant deterministic flux $Q = \langle Q\rangle$ at the top (x_1 = 3m), and no-flow boundary at the left and right sides. The input parameters are given as $\langle f\rangle$=0.0 (i.e., the geometric mean saturated hydraulic conductivity K_G = 1.0 *m/day*), the coefficient of variation $CV_{K_S} = \sigma_{K_S}/\langle K_S\rangle$ = 10.0%, $\langle\beta\rangle = \langle\ln(\alpha)\rangle$ = 0.6931, $CV_\alpha = \sigma_\alpha/\langle\alpha\rangle$ = 10%, $\langle n\rangle$ = 1.4, $CV_n = \sigma_n/\langle n\rangle$ = 5%, $\lambda_f = \lambda_\beta = \lambda_n$ = 0.5 m, θ_S = 0.4, θ_r =0.01, $\langle Q\rangle$ = -0.005 *m/day*, and σ_Q^2 = 0.0. For a log normally distributed variable p, the coefficient of variation of p is related to the variance of its log-transformed variable through the simple relation: $\sigma^2_{\ln p} = \ln[1 + CV_p^2]$. This example with relatively small variabilities in f, β, and n is chosen to ensure convergence of Monte Carlo simulations.

For Monte Carlo simulations, the input parameters f, β, and n are obtained based on 30,000 unconditional realizations with zero mean and unit variance. For each simulation, a log hydraulic conductivity $f(\boldsymbol{x})$ field, a log-transformed pore size distribution $\beta(\boldsymbol{x})$ field, and an $n(\boldsymbol{x})$ field are read from the unconditional realizations and then are scaled to the specified mean and variance of f, β, and n. The quality of random fields is then checked by comparing the sample covariance against the input, analytical covariance of (A53). The unsaturated flow equations (A1)-(A5) are solved for each set of $f(\boldsymbol{x})$, $\beta(\boldsymbol{x})$, and $n(\boldsymbol{x})$ realizations. A total of 10,000 simulations are conducted, on the basis of which sample mean and variance of flow quantities are calculated. The comparison between results from the KLME method, the moment-equation-based approach (ME) and Monte Carlo results (MC) is illustrated in Figure A1, which shows two vertical profiles passing through the center of the flow domain. It is seen that the mean pressure head derived from our model is almost identical to Monte Carlo results (Fig. A1a), while there is still slight discrepancy in the pressure head variance (Fig.

A1b). In addition, Figure A1 demonstrates that when the variabilities on f, β, and n are relatively small and the infiltration rate is low, the number of Monte Carlo simulations needed to obtain a convergent solution is low. For mean pressure head, 2,000 Monte Carlo simulations are enough to obtain a convergent solution, while about 5,000 simulations are needed for the pressure head variance. Monte Carlo simulations beyond 5,000 do not significantly affect the results. The discrepancy between pressure head variances computed from the moment-based approach and from the Monte Carlo simulations (NMC =10,000) is due to numerical errors in solving flow equations and due to neglecting higher-order terms in our moment-based approach. Nevertheless, the discrepancy is small, indicating the validation of the moment-based approach at least in the limit of relatively small variabilities on soil properties.

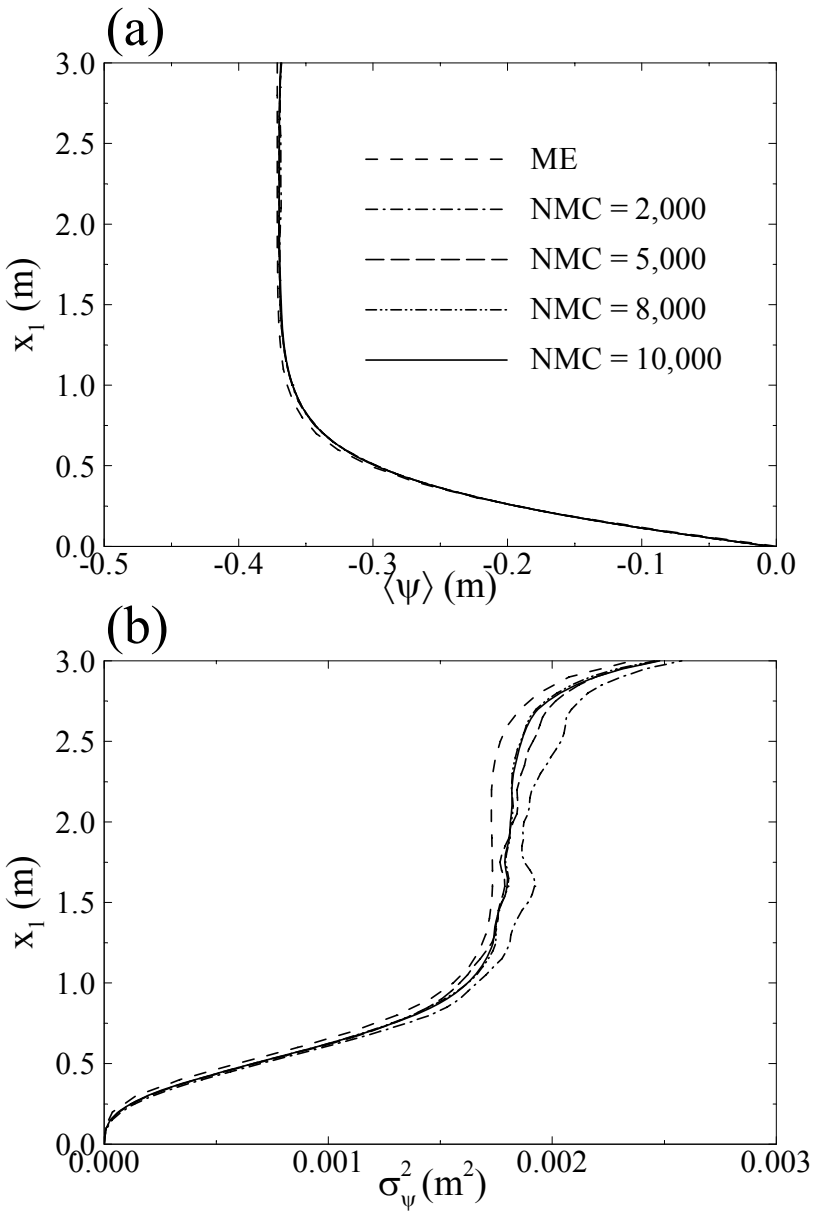
In our second example (Case 2), we increase the infiltration rate from $\langle Q \rangle = -0.005$ *m/day* to $\langle Q \rangle = -0.05$ *m/day*. The comparisons between Monte Carlo results and moment-based results are illustrated in Figure A2. Again, Figure A2a shows that 2,000 Monte Carlo simulations are enough for the mean pressure head. It is also indicated from Figure A2a that there is still a slight difference between the mean pressure head computed from the moment approach and Monte Carlo simulations (NMC = 10,000), which again is due to numerical errors and due to neglecting of higher-order terms in our moment solution. Unlike Case 1, due to a relatively large infiltration rate in Case 2, flow in the upper portion of the domain is mean gravity-dominated with a constant mean pressure head. For the pressure head variance (Fig. A2b), it is seen that about 8,000 Monte Carlo simulations are needed to achieve statistical convergence. In addition, the head variance experiences a quick increase in the capillary fringe, more or less stabilizes in the gravity-dominated region, and increases again near the upper flux boundary. The increase of pressure head variance near the upper flux boundary has been observed and explained previously (e.g., Zhang and Lu, 2002).

We also compared the mean of the log unsaturated hydraulic conductivity and its variance computed from the moment approach and Monte Carlo simulations (Fig. A3). The figure shows that there is an excellent agreement between the Monte Carlo results and the moment-based results. It is worthwhile to note that the profile of the variance of the log unsaturated hydraulic conductivity σ_Y^2 exhibits a quick increase right above the water table, as shown in both ME and MC results. It is found that this increase is due to a large gradient of $\langle Y \rangle$ with respect to $\langle n \rangle$, i.e., a large value of $\partial \langle Y \rangle / \partial \langle n \rangle$. The comparison of the effective water contents obtained from Monte Carlo simulations and the moment-equation-based approach is illustrated in Figure A4.



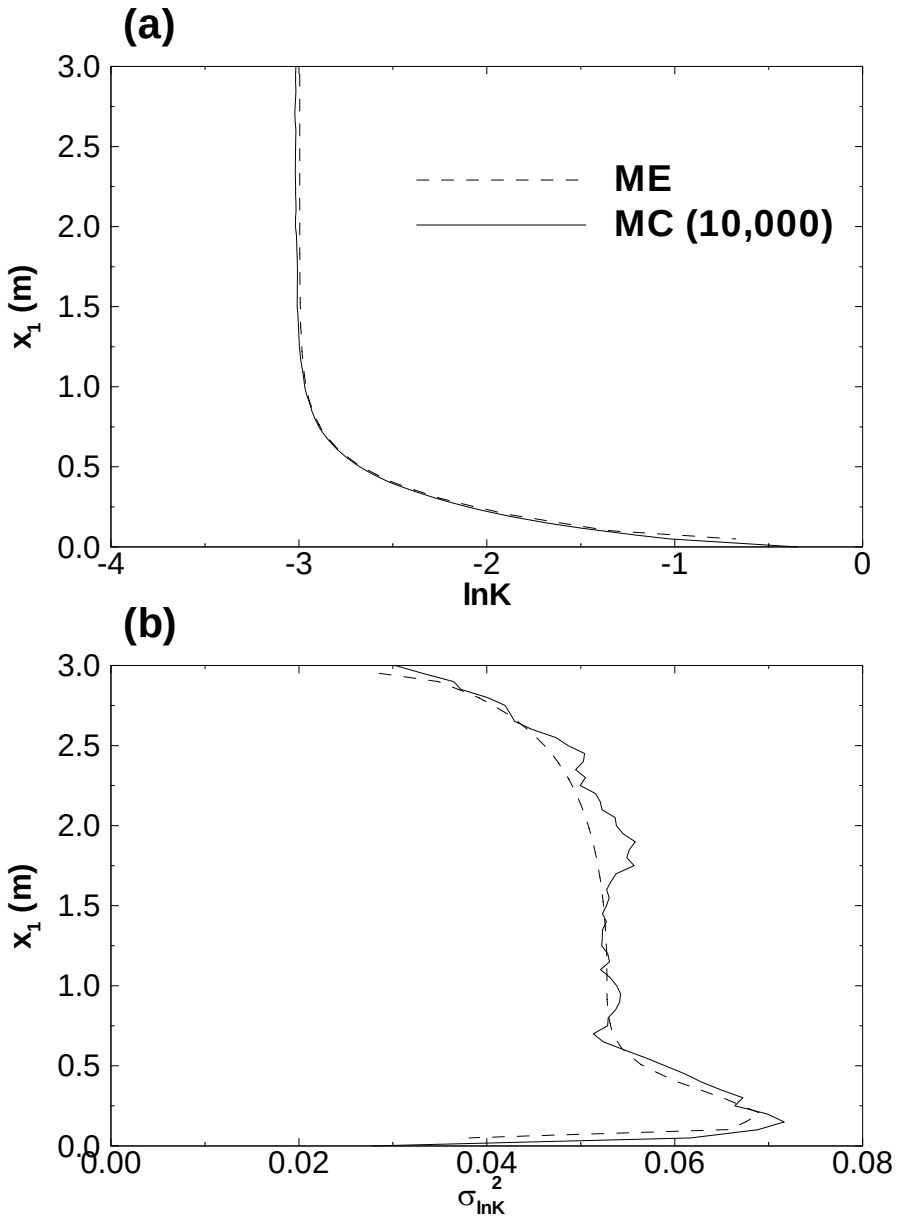
Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheets "fig1-a" and "fig1-b". For information only, not to be used for quality-affecting work.

Figure A1. Comparisons between moment-equation-based approach (ME) and Monte Carlo simulations (MC) for Case 1: CV_{K_S} = 10%, CV_α = 10%, CV_n = 5%, CV_Q = 0, and $\langle Q \rangle$ = –0.005 *m/day*. (a) Mean pressure head; and (b) head variance



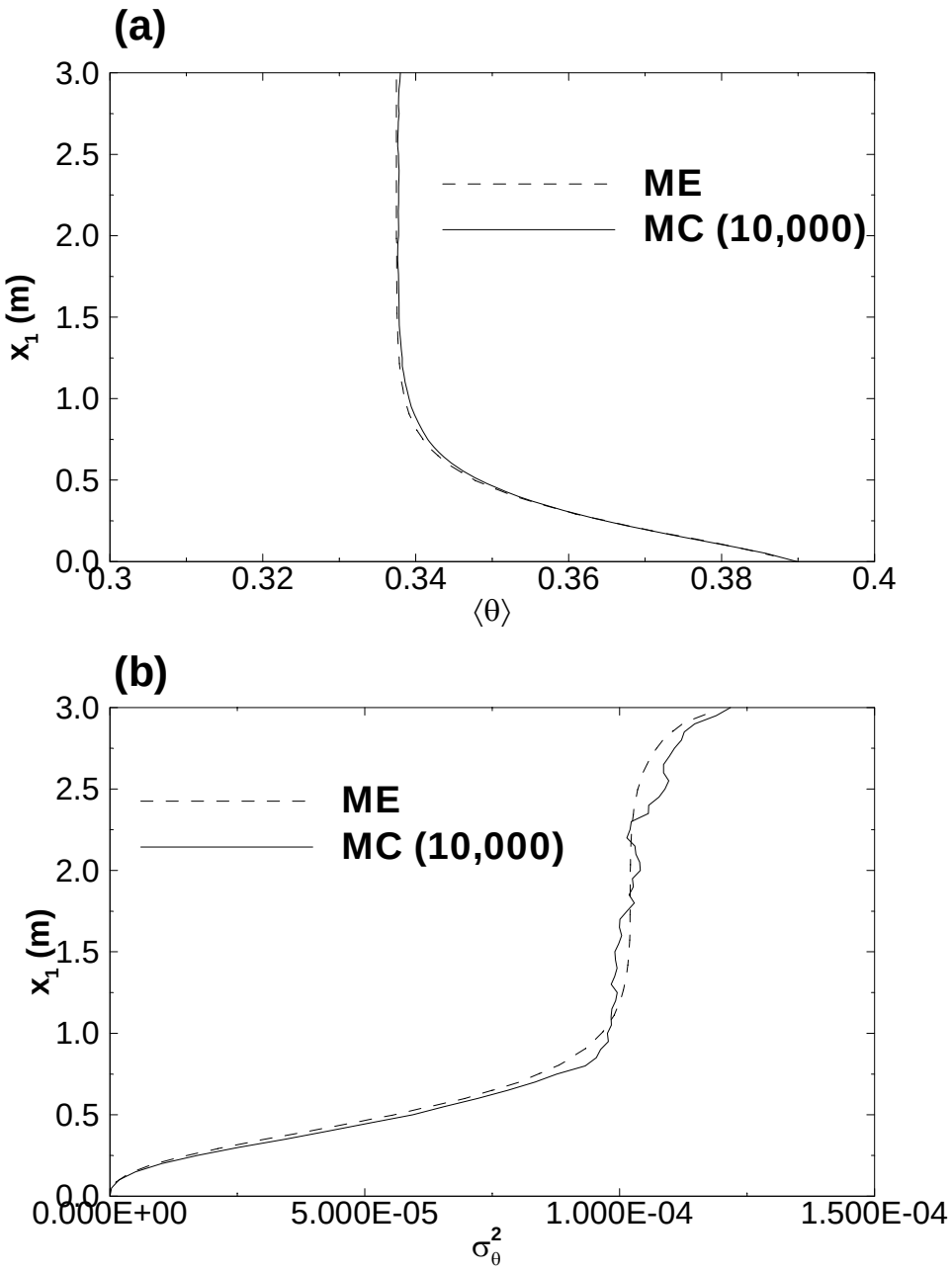
Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheets "fig2-a" and "fig2-b". For information only, not to be used for quality-affecting work.

Figure A2. Comparisons between moment-equation-based approach (ME) and Monte Carlo simulations (MC) for Case 2: CV_{K_S} = 10%, CV_α = 10%, CV_n = 5%, CV_Q = 0, $\langle Q \rangle$ = -0.05 *m/day*. (a) Mean pressure head; and (b) head variance



Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheets "fig3-a" and "fig3-b". For information only, not to be used for quality-affecting work.

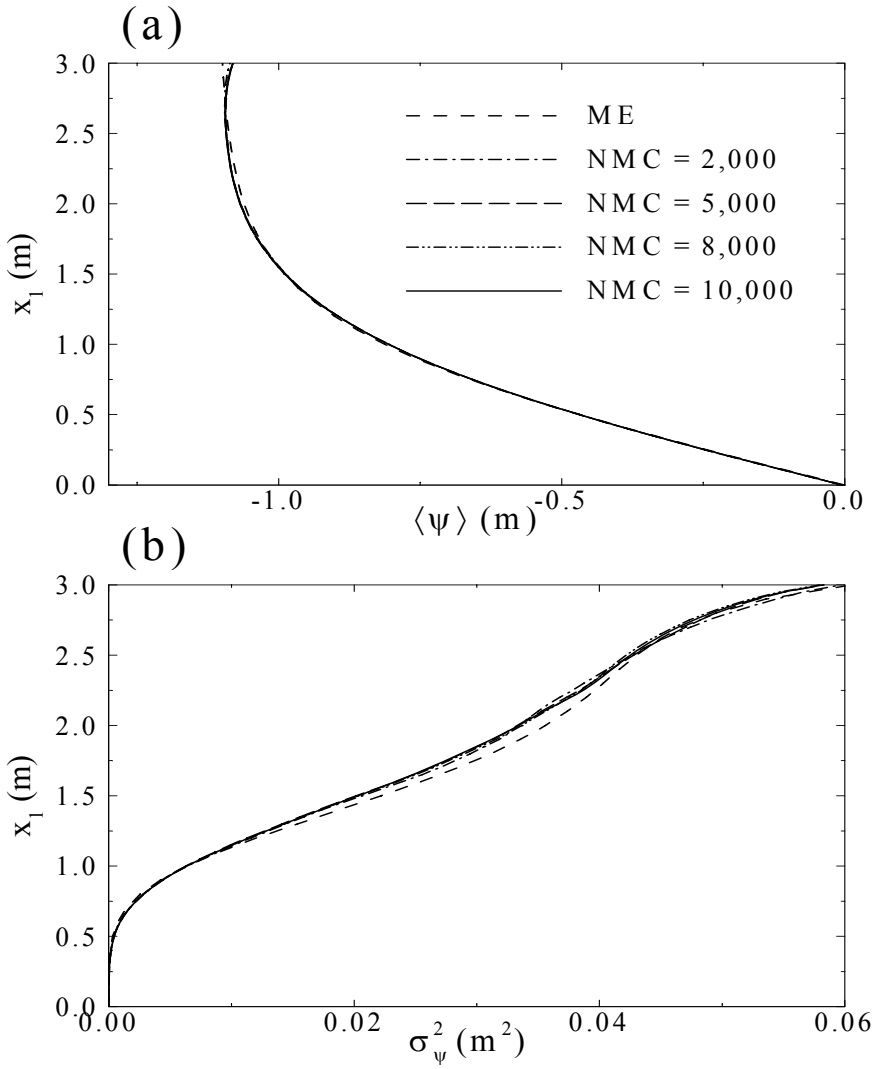
Figure A3. Comparisons between moment-equation-based approach (ME) and Monte Carlo simulations (MC) for the log unsaturated hydraulic conductivity Y in Case 2. (a) Mean $\langle Y \rangle$; and (b) variance σ_Y^2



Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheets "fig4-a" and "fig4-b". For information only, not to be used for quality-affecting work.

Figure A4. Comparisons between moment-equation-based approach (ME) and Monte Carlo simulations (MC) for the effective water content θ_e in Case 2. (a) Mean $\langle\theta_e\rangle$; and (b) variance $\sigma^2_{\theta_e}$

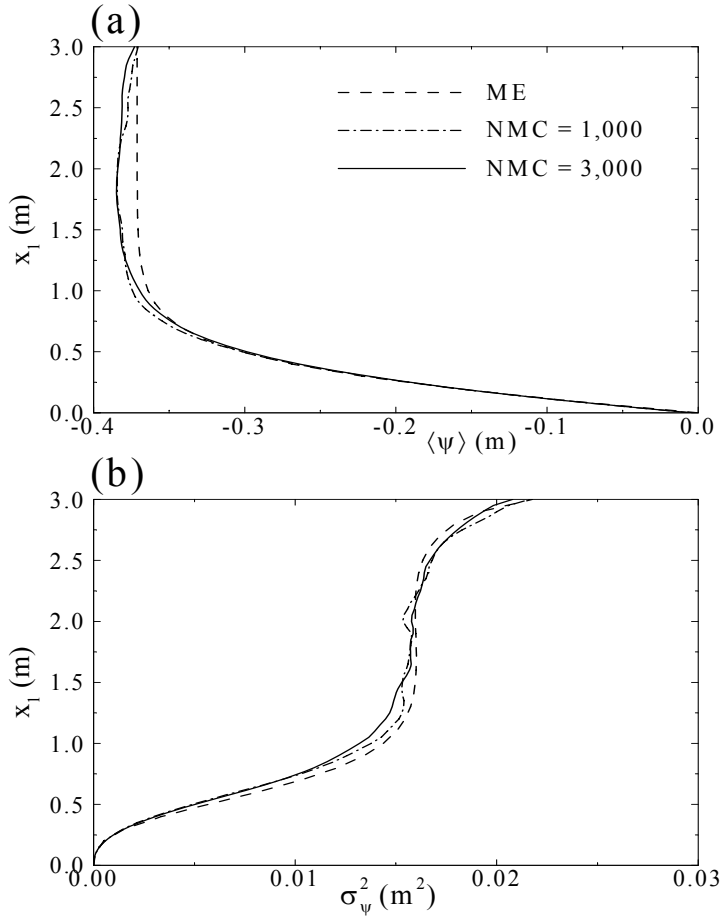
We next consider a case (denoted as Case 3) that has relatively large spatial variabilities on K_S and α: $C\mathrm{V}_{\mathrm{K}_S} = 100\%$, $C\mathrm{V}_{\alpha} = 20\%$. The infiltration rate is $\langle Q \rangle = -0.005$ *m/day*. The mean and correlation lengths for other parameters are the same as before. The results are depicted in Figure A5. The figure indicates that, even though the variabilities on K_S and α are large, 2,000 Monte Carlo simulations are enough for both mean pressure head and head variance, partially due to the relatively small infiltration rate and partially due to the small variability on n.



Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheets "fig5-a" and "fig5-b". For information only, not to be used for quality-affecting work.

Figure A5. Comparisons between moment-equation-based approach (ME) and Monte Carlo simulations (MC) for Case 3: $C\mathrm{V}_{\mathrm{K}_S} = 100\%$, $C\mathrm{V}_{\alpha} = 20\%$, $C\mathrm{V}_n = 5\%$, $C\mathrm{V}_Q = 0$, $\langle Q \rangle = -0.005$ *m/day*. (a) Mean pressure head; and (b) head variance

In the next example the infiltration rate in Case 3 is increased to $\langle Q \rangle$ = -0.05 *m/day* (Case 4). We ran 3,000 Monte Carlo simulations for this case, a few of which did not converge and have been removed from computing sample statistics. The results are illustrated in Figure A6. It is well known that flow in an unsaturated system poses an interesting numerical problem. Spatial variabilities in K_S, α, and n make it even more challenging. As a result, convergence may not be achieved for some of the realizations, especially in the case of large variabilities and a large infiltration rate. To efficiently simulate unsaturated or unsaturated-saturated flow in the presence of large material contrasts calls for robust numerical solvers. Without such a solver it would be very difficult to establish the upper limits of variabilities in soil properties above which the first-order stochastic model starts to break down because this effort would involve large sets of high-resolution Monte Carlo simulations with large variabilities on input variables. This is outside of the scope of the present study.

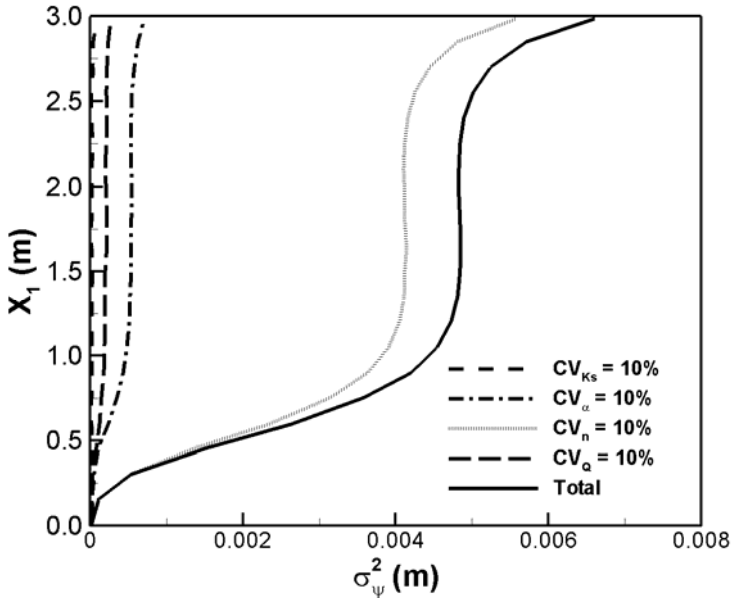


Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheets "fig6-a" and "fig6-b". For information only, not to be used for quality-affecting work.

Figure A6. Comparisons between moment-equation-based approach (ME) and Monte Carlo simulations (MC) for Case 4: CV_{K_S} = 100%, CV_α = 20%, CV_n = 5%, CV_Q = 0, $\langle Q \rangle$ = –0.05 *m/day*. (a) Mean pressure head; and (b) head variance

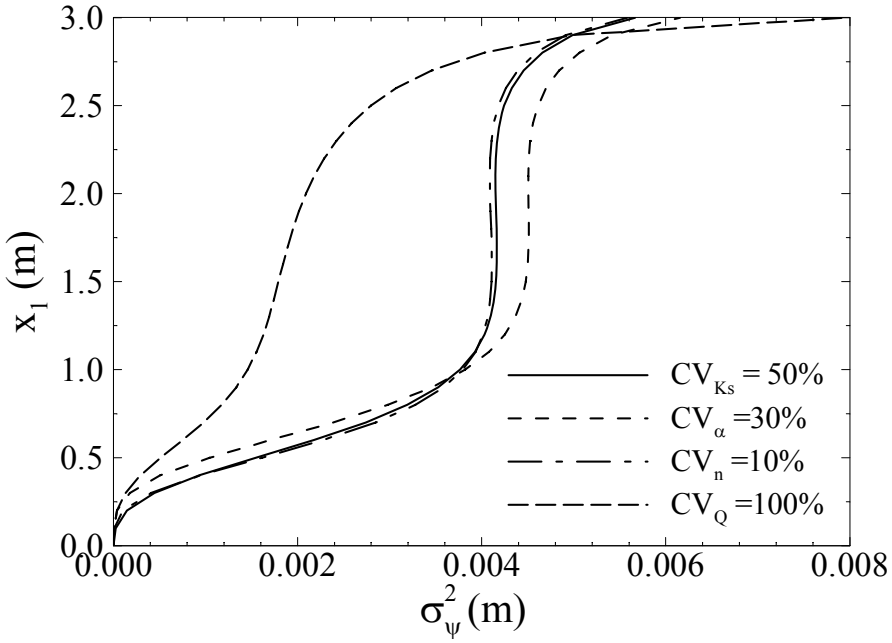
A.5.2 Contributions of Parameter Variances to Head Variance

We also conducted numerical simulations to investigate the relative contribution of the variability of f, β, and n to the pressure head variance. In each simulation, we only allow variation in one of these four parameters with a coefficient of variation $CV_p = 10.0\%$, where $p = K_S$, α, n, or Q, given $\langle f \rangle = 0.0$, $\langle \beta \rangle = 0.6931$, $\langle n \rangle = 1.4$, and $\langle Q \rangle = -0.05$ *m/day*. We then run one simulation with the coefficient of variation $CV = 10\%$ for all four parameters. The results are illustrated in Figure A7. It is seen that under the condition of mutually independent K_S, α, and n, the contribution of the variability in each parameter to the pressure head variance is additive, namely, the pressure head variance due to the variabilities of all four parameters equals the sum of the four pressure head variances due to the variability of each individual parameter. In addition, it seems that under the specific (unsaturated) condition, the variability in the fitting parameter n has the largest contribution to the pressure head variance, compared to other parameters with the same magnitude of coefficients of variation. The parameter α is of secondary importance in the pressure head variance. Of course, in reality, variabilities of K_S and α may be much larger than that of n. For this reason, we run more simulations with relatively high variabilities in K_S and α: $CV_{K_S} = 50\%$, $CV_\alpha = 30\%$, while keeping $CV_n = 10\%$. The head variances for these simulations are depicted in Figure A8. The figure shows that under these specific (unsaturated) conditions the contribution to the pressure head variance due to the variability $CV_n = 10\%$ is compatible to that due to the variability $CV_{K_S} = 50\%$ or that due to $CV_\alpha = 30\%$. These results indicate that under unsaturated conditions, the variability of n has the great impact on predictive uncertainty and should not be ignored in simulations.



Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheet "fig7". For information only, not to be used for quality-affecting work.

Figure A7. Contributions to head variance due to variabilities on individual parameters, $CV_p = 10\%$, where $p = K_S$, α, n, or Q



Source: Data come from DID: 016JZ.001, see Excel file "data-hu.xls", sheet "fig8". For information only, not to be used for quality-affecting work.

Figure A8. Contributions to head variance due to variabilities on individual parameters, CV_{K_S} = 50%, CV_α =30%, CV_n = 10%, or CV_Q = 100%

A.6 Summary and Discussion

With the method of KLME we developed a general first-order, nonstationary stochastic model for transient, unsaturated flow in randomly heterogeneous media on the basis of the van Genuchten-Mualem constitutive relationship. Due to its nonstationarity and nonlinearity, the model cannot generally be solved analytically. We solve it by the numerical technique of finite differences, which renders flexibility in handling different boundary conditions, medium multiscale, nonstationary features, and input covariance structures. The nonstationary stochastic model developed is applicable to the entire domain of bounded, multi-dimensional vadose zones in the presence of deterministic recharge and sink/source and in the presence of multiscale, nonstationary medium features. The results of the stochastic model are the first two moments (means and covariances) of the flow quantities such as pressure head and flux. The first moments estimate (or predict) the fields of pressure head and flux in a heterogeneous medium, and the corresponding (co)variances evaluate the uncertainty (error) associated with the estimation (prediction). These first two moments can be used to construct confidence intervals for the pressure and flux fields.

We demonstrated the KLME approach with several examples of transient flow in a two-dimensional rectangular domain and compared the results with those from the conventional moment method

(CME) as well as from Monte Carlo simulation (MC). The comparison results indicate that the developed stochastic model based on KLME approach produce very similar results, and the KLME approach is much more efficient than the CME approach and the MC approach.

The illustrative examples show the potential applicability of the proposed stochastic model to the complicated saturated-unsaturated cases. The examples indicate that under unsaturated conditions, the pressure head variance is sensitive to all the soil variabilities, in the order of n, α and K_S. Furthermore, although the variabilities of α and n are usually smaller than that of K_S, their effects on predicting uncertainty associated flow and transport in heterogeneous, unsaturated media should not be neglected.

The validity of the developed model was confirmed with high-resolution Monte Carlo simulations in the case of small variabilities ($C\mathrm{V}_{\mathrm{K}_S}$ = $C\mathrm{V}_{\alpha}$= 10% and $C\mathrm{V}_n$ = 5%) and relatively large ones ($C\mathrm{V}_{\mathrm{K}_S}$ = 100%, $C\mathrm{V}_{\alpha}$= 20%, and $C\mathrm{V}_n$ = 5%). To establish the upper limits of the variabilities in soil properties below which the first-order stochastic model is valid, however, would involve a large amount of high-resolution Monte Carlo simulation sets and require robust numerical solvers that handle large properties contrasts efficiently. This is outside of the scope of the present study.

A.7 References

Andersson, J.; Shapiro, A.M. 1983. " Stochastic analysis of one-dimensional steady state unsaturated flow: A comparison of Monte Carlo and perturbation methods." *Water Resources Research,* 19(1): 121-133.

Bresler, E.; Dagan, G. 1981. "Convective and pore scale dispersive solute transport in unsaturated heterogeneous fields." *Water Resources Research,* 17: 1683-1693.

Brooks, R. H.; Corey, A.T. 1964. *Hydraulic properties of porous media*. Fort Collins, Colorado: Hydrology Pap. 3, Colorado State Univ. Press.

Chen, M.; Zhang, D.; Keller, A. A.; Lu, Z. 2005. "A stochastic analysis of steady state two-phase flow in heterogeneous media," *Water Resources Research,* 41, W01006, doi: 10.1029/2004WR003412.

Chen, M.; Zhang, D.; Keller, A.; Lu, Z.; Zyvoloski, G.A. 2006. "A stochastic analysis of transient two-phase flow in heterogeneous porous media." *Water Resources Research*, 42, W03425, doi:10.1029/2005WR004257.

Dagan, G. 1989. *Flow and Transport in Porous Formations*. New York: Springer-Verlag.

Dagan, G.; Bresler, E. 1979. "Solute dispersion in unsaturated heterogeneous soil at field scale: I. Theory." *Soil Science Society of America Journal*, 43: 461-467.

Destouni, G.; Cvetkovic, V. 1989. "The effect of heterogeneity on large scale solute transport in the unsaturated zone." *Nordic Hydrology*, 20: 43-52.
Ferrante, M.; Yeh, T.-C. J. 1999. "Head and flux variability in heterogeneous unsaturated soils under transient flow conditions." *Water Resources Research,* 35(4): 1471-1479.

Foussereau, X.; Graham, W.D.; Rao, P. S. C. 2000. "Stochastic analysis of transient flow in unsaturated heterogeneous soils." *Water Resources Research.,* 36(4): 891-910.

Gardner, W.R. 1958. "Some steady state solutions of unsaturated moisture flow equations with application to evaporation from a water table." *Soil Science,* 85: 228-232.

Gelhar, L.W. 1993. *Stochastic Subsurface Hydrology*. Englewood Cliffs, New Jersey: Prentice-Hall.

Ghanem R.; Spanos, P.D. 1991. *Stochastic Finite Elements: a Spectral Approach.* New York: Springer-Verlag.

Harter, Th.; Yeh, T.-C. J. 1996a. "Stochastic analysis of solute transport in heterogeneous, variably saturated soils." *Water Resources Research,* 32(6): 1585-1595.

Harter, Th.; Yeh, T.-C. J. 1996b. "Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils." *Water Resources Research,* 32(6): 1597-1609.

Harter, Th.; Zhang, D. 1999. "Water flow and solute spreading in heterogeneous soils with spatially variable water content." *Water Resources Research,* 35(2): 415-426.

Hopmans, J. W.; Schukking, H.; Torfs, P. J. J. F. 1988. "Two-dimensional steady-state unsaturated water flow in heterogeneous soils with autocorrelated soil hydraulic properties." *Water Resources Research,* 24(12): 2005-2017.

Hughson, D.L.; Yeh, T.-C.J. 2000. "An inverse model for three-dimensional flow in variably saturated porous media." *Water Resources Research,* 36(4): 829-840.

Indelman, P.; Or, D.; Rubin, Y. 1993. "Stochastic analysis of unsaturated steady state flow through bounded heterogeneous formations." *Water Resources Research,* 29: 1141-1148.

Li, B.; Yeh, T.-C. J. 1998. "Sensitivity and moment analyses of head in variably saturated regimes." *Advances in Water Resources*, 21(6): 477-485.

Liedl, R. A. 1994. "Conceptual perturbation model of water movement in stochastically heterogeneous soils." *Advances in Water Resource*s, 17: 171-179.

Lu, Z.; Zhang, D. 2003. "Solute Spreading in Nonstationary Flows in Bounded, Heterogeneous Unsaturated-Saturated Media." *Water Resources Research,* 39(3): 1049, doi: 10.1029/2001WR000908.

Lu, Z.; Zhang, D. 2004a. "A comparative study on quantifying uncertainty of flow in randomly heterogeneous media using Monte Carlo simulations, the conventional and KL-based moment-equation approaches." *SIAM Journal on Scientific Computing*, 26(2), 558-577.

Lu, Z.; Zhang, D. 2004b. "Analytical Solutions to Unsaturated Flow in Layered, Heterogeneous Soils via Kirchhoff Transformation." *Advances in Water Resources*, 27, 775-784.

Lu, Z.; Zhang, D. 2005. "Analytical Solutions of Statistical Moments for Transient Flow in Two-Dimensional Bounded, Randomly Heterogeneous Media." *Water Resources Research*, 41, W01016, doi:10.1029/2004WR003389.

Lu, Z.; Neuman, S.P.; Guadagnini, A.; Tartakovsky, D.M. 2000. *Direct solution of unsaturated flow in randomly heterogeneous soils*. Proceedings of XIII International Conference of Computational Methods in Water Resources. L. R. Bentley, J. F. Sykes, C. A. Brebbia, W. G. Grey, and G. F. Pinder, eds., 785-792.

Lu, Z.; Neuman, S.P.; Guadagnini, A.; Tartakovsky, D.M. 2002. "Conditional moment analysis of steady state unsaturated flow in bounded randomly heterogeneous porous soils." *Water Resources Research,* 38(4): 10.1029/2001WR000278.

Mantoglou, A. 1992. "A theoretical approach for modeling unsaturated flow in spatially variable soils: Effective flow models in finite domains and nonstationarity." *Water Resources Research,* 28(1): 251-267.

Polmann, D. J.; McLaughlin, D.; Luis, S.; Gelhar, L.W.; Ababou, R. 1991. "Stochastic modeling of large-scale flow in heterogeneous unsaturated soils." *Water Resources Research,* 27: 1447-1458.

Protopapas, A.L.; Bras, R.L. 1990. "Uncertainty propagation with numerical models for flow and transport in the unsaturated zone." *Water Resources Research,* 26(10): 2463-2474.

Russo, D. 1988. "Determining soil hydraulic properties by parameter estimation: On the selection of a model for the hydraulic properties." *Water Resources Research,* 24: 453-459.

Russo, D. 1993. "Stochastic modeling of macrodispersion for solute transport in a heterogeneous unsaturated porous formation." *Water Resources Research,* 29: 383-397.

Russo, D. 1995a. "On the velocity covariance and transport modeling in heterogeneous anisotropic porous formations: 2. Unsaturated flow." *Water Resources Research,* 31: 139-145.

Russo, D. 1995b. "Stochastic analysis of the velocity covariance and the displacement covariance tensors in partially saturated heterogeneous anisotropic porous formations." *Water Resources Research,* 31: 1647-1658.

Russo, D.; Bouton, M. 1992. "Statistical analysis of spatial variability in unsaturated flow parameters." *Water Resources Research,* 28(7): 1991-1925.

Tartakovsky, D. M.; Neuman, S.P.; Lu, Z. 1999. "Conditional Stochastic Averaging of Steady State Unsaturated Flow by Means of Kirchhoff Transformation." *Water Resources Research,* 35(3): 731-745.

Unlu, K.; Nielsen, D.R.; Biggar, J.W. 1990. "Stochastic analysis of unsaturated flow: One-dimensional Monte Carlo simulations and comparisons with spectral perturbation analysis and field observations." *Water Resources Research,* 26(9): 2207-2218.

van Genuchten, M. Th.1980. "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils." *Soil Science Society of America Journal,* 44: 892-898.

Yang, J.; Zhang, R.; Wu, J. 1996. "An analytical solution of macrodispersivity for adsorbing solute transport in unsaturated soils." *Water Resources Research,* 32: 355-362.

Yang, J.; Zhang, D.; Lu, Z. 2004. "Stochastic analysis of saturated-unsaturated flow in heterogeneous media by combining Karhunen-Loeve expansion and perturbation method." *Journal of Hydrology,* 294: 18-38.

Yeh, T. -C.; Gelhar, L.W.; Gutjahr, A.L. 1985a. "Stochastic analysis of unsaturated flow in heterogeneous soils: 1. Statistically isotropic media." *Water Resources Research,* 21: 447-456.

Yeh, T. -C.; Gelhar, L.W.; Gutjahr, A.L. 1985b. "Stochastic analysis of unsaturated flow in heterogeneous soils: 2. Statistically anisotropic media with variable alpha." *Water Resources Research,* 21: 457-464.

Yeh, T.-C. J. 1989. "One-dimensional steady-state infiltration in heterogeneous soils." *Water Resources Research,* 25(10): 2149-2158.

Zhang, D. 1998. "Numerical solutions to statistical moment equations of groundwater flow in nonstationary, bounded heterogeneous media." *Water Resources Research,* 34(3): 529-538.

Zhang, D. 1999. "Nonstationary stochastic analysis of transient unsaturated flow in randomly heterogeneous media." *Water Resources Research,* 35(4): 1127-1141.

Zhang, D. 2002. *Stochastic Methods for Flow in Porous Media: Coping with Uncertainties.* San Diego, California: Academic Press.

Zhang, D.; Lu, Z. 2002. "Stochastic analysis of flow in a heterogeneous unsaturated-saturated system." *Water Resources Research,* 38(2): 10.1029/2001WR000515.

Zhang, D.; Lu, Z. 2004. "An Efficient, High-Order Perturbation Approach for Flow in Random Porous Media via Karhunen-Loeve and Polynomial Expansions." *Journal of Computational Physics*, 194(2), 773-794.

Zhang, D.; Winter, C.L. 1998. "Nonstationary stochastic analysis of steady-state flow through variably saturated, heterogeneous media." *Water Resources Research,* 34(5): 1091-1100.

Zhang, D.; Wallstrom, T.C.; Winter, C.L. 1998. "Stochastic analysis of steady-state unsaturated flow in heterogeneous media: Comparison of the Brooks-Corey and Gardner-Russo models." *Water Resources Research,* 34(6): 1437-1449.