WATER RESOURCES RESEARCH, VOL. 38, NO. 11, 1232, doi:10.1029/2001WR000725, 2002

Infiltration in stratified, heterogeneous soils: Relative importance of parameters and model variations

Theodora Avanidou and Evan K. Paleologos

Department of Geological Sciences, University of South Carolina, Columbia, South Carolina, USA

Received 14 June 2001; revised 20 March 2002; accepted 3 April 2002; published 13 November 2002.

[1] This study presents a framework that accounts for the uncertainty, relative importance, and relative contribution of uncertain and spatially variable parameters and the influence of statistical model assumptions in problems of infiltration in unsaturated, heterogeneous formations. The relative importance of the saturated hydraulic conductivity K_s, the van Genuchten α, and β parameters was quantified toward the mean pressure head profile and the fluctuations about it. The broad implications of our study for site characterizations efforts are the following. Statistical moments were seen not to be sufficient for producing reliable modeling estimates. Significant information on the detailed shape of the probability distribution is required in order to produce meaningful predictions. When limited data are available, optimum allocation of resources is accomplished by concentrating on K_s. When more extensive sampling can be done, both K_s and β appear to be critical. Total system performance analyses may not provide the most appropriate modeling estimates. The results of the relative importance of the parameters for a specific layer did not coincide with the conclusions drawn from an analysis of the system as a whole. *INDEX TERMS:* 1829 Hydrology: Groundwater hydrology; 1869 Hydrology: Stochastic processes; 1875 Hydrology: Unsaturated zone; *KEYWORDS:* unsaturated zone, relative importance, stochastic hydrology

Citation: Avanidou, T., and E. K. Paleologos, Infiltration in stratified, heterogeneous soils: Relative importance of parameters and model variations, *Water Resour. Res.*, *38*(11), 1232, doi:10.1029/2001WR000725, 2002.

1. Introduction

[2] Unsaturated flow in soil formations is a complex process that even in relatively homogeneous systems is not easily understood [*Looney and Falta*, 2000]. One major factor that contributes to this complexity is the spatial variability of soil properties. Field observations have shown that the hydrologic properties of soils vary several orders of magnitude even in the same geologic formation [*Nielsen et al.*, 1973; *Freeze*, 1975; *Sudicky*, 1986; *Unlu et al.*, 1989, 1990a; *Wierenga et al.*, 1991; *Woodbury and Sudicky*, 1991; *Wilson et al.*, 1994] and their distributions are subject to uncertainty due to limitations in measurements. Thus the effect of the spatial variability on predictions of the flow of water and the transport of contaminants in unsaturated porous media has become one of the focal points of scientific investigations [e.g., *Warrick et al.*, 1977; *Dagan and Bresler*, 1983; *Russo*, 1983, 1984; *Yeh et al.*, 1985a; *Unlu et al.*, 1990a; *Wierenga et al.*, 1991; *Russo and Bouton*, 1992; *White and Sully*, 1992; *Russo et al.*, 1997; *Boateng and Cawlfield*, 1999]. Studies of unsaturated flow within porous media exhibiting heterogeneities have adopted a statistical framework employing either analytic approximations through linearizations and perturbation methods or Monte Carlo simulations that have focused, in their majority, on one parameter at a time.

0043-1397/02/2001WR000725

[3] Our study is concerned with one-dimensional infiltration in unsaturated, heterogeneous multilayered geologic formations. In particular we are interested in investigating the effect of assumptions made about the statistical structure of the data on the prediction of pressure head and saturation profile as well as the relative importance of each parameter that enters the *van Genuchten* [1980] relation. Stochastic studies by *Yeh et al.* [1985a, 1985b, 1985c] had investigated the spatial variability of the flow in an unsaturated multidimensional medium under steady state conditions. The one- and three-dimensional results were found to agree for large values of the autocorrelation length and for coarse-textured soils. The variance of the pressure head was found to be mean dependent, increasing with the mean soil water pressure head. *Yeh* [1989] used the exponential model [*Gardner*, 1958] for the unsaturated hydraulic conductivity to develop quasi-recursive analytical solutions for the pressure head profile that develops as a result of one-dimensional steady state infiltration in perfectly stratified media. The cross correlation between the Gardner model's parameters was found to affect flow behavior. When these parameters were assumed to be uncorrelated the variance in pressure head increased with a decrease in the infiltration rate, whereas the opposite behavior was observed for parameters that were perfectly correlated.

[4] The extent to which data collection influences our prediction of the saturation and pressure profiles has been investigated in several studies. *Smith and Diekkruger* [1996] considered one-dimensional vertical flow through various heterogeneous soils to determine the behavior of

average soil characteristics. These authors emphasized that a small number of soil samples cannot support conclusive interpretations of the statistical nature of the parameters' distributions and of the interdependencies that characterize heterogeneous porous media. Development of simple expressions that relate effective hydraulic properties with sample measurements or boundary conditions did not appear feasible. *Yeh and Zhang* [1996] developed a geostatistical inverse technique to estimate the saturated hydraulic conductivity and pore size distribution parameters of the Gardner model under steady state, unsaturated flow conditions, using data on moisture content and water pressure. These parameters were found to be identifiable when a large amount of information is available on soil water pressure and degree of saturation. These authors also showed that the cross correlation between hydraulic parameters varies with mean soil water pressure. *Hughson and Yeh* [1998] extended the model by *Yeh and Zhang* [1996] to three-dimensional transient flow, using the van Genuchten model to describe the pressure-hydraulic conductivity and pressure-water content relations. Their model provided good estimates of the saturated hydraulic conductivity and pore size distribution parameter, using pressure head and moisture content data. *Harter and Zhang* [1999] investigated the impact of the soil water content variability on water flow and solute transport. These authors found that in dry soils the coefficient of variation of the soil water content is sensitive to the pore size distribution parameter α and the tortuosity parameter m. At given water tension the coefficient of variation of the saturation increased nearly logarithmically with decreasing m. In wet soils the spatial correlation structure of the soil water content was found to be similar to that of the pressure head. This finding combined with the conditional simulation results by *Harter and Yeh* [1996], who demonstrated that significant uncertainty reduction in solute transport predictions can be accomplished through conditioning on pressure head, indicates that measurements of soil water content may, equally to soil water tension data, assist in solute flux predictions. *Tartakovsky et al.* [1999] developed a closed form alternative to conditional Monte Carlo simulations and used conditional second moments to predict pressure head, water content, and fluxes for steady state unsaturated flow. These authors used a Kirchhoff transformation of the Gardner model and treated the exponent α as a random constant. For variances of the logarithm of saturated hydraulic conductivity less than one the pressure head and its variance compared well with the results of Monte Carlo simulations.

[5] In this study we address the problem of one-dimensional infiltration in unsaturated, heterogeneous, multilayered geologic formations. Several types of distribution have been presented in the literature to describe the parameters of this problem. The first objective of our study is to assess what is the effect on the flow behavior of the system (as a whole and on each specific layer) of assumptions made about the type of parameter distribution. By flow behavior we designate here the mean pressure profile and the fluctuations about this mean profile. The assumption of the type of distribution is critical, because only a small amount of field data is usually available and even in relatively homogeneous formations, these data can be fitted by more than one type of probability distribution [*Smith*, 1981; *Unlu et al.*, 1990a, 1990b; *Woodbury and Sudicky*, 1991; *Russo and Bouton*, 1992; *White and Sully*, 1992]. The significance for site characterization and modeling efforts of this assumption lies in that, if the exact functional form of a probability distribution is required, then a very dense measurement network needs to be implemented in order to determine unequivocally the statistics of the flow field. The second objective of this study is to evaluate the importance (defined under specific criteria developed in subsequent sections) of each uncertain and spatial variable parameter of the flow problem toward the total system's uncertainty and the uncertainty in each individual layer's hydraulic behavior. A theoretical framework is supplied where relative contribution and relative importance of the parameters are clearly defined. This aspect has important implications in site characterization and modeling efforts because it allows the concentration of resources on those factors that dominate a system's uncertainty.

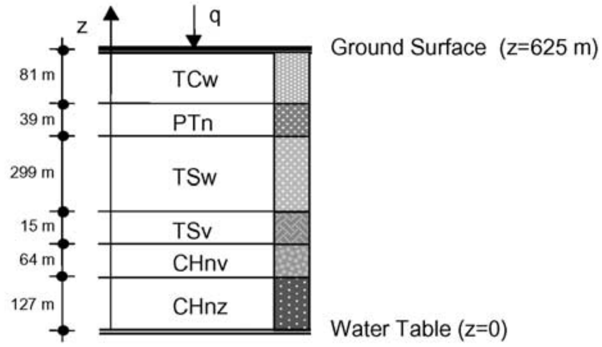

Figure 1. Schematic of the stratigraphy (not to scale).

2. Physical Problem and Conceptual Model

[6] Consider the physical problem of one-dimensional infiltration in unsaturated porous media under a constant infiltration rate. For this problem field data were utilized from the characterization of the flow field at the potential U.S. Department of Energy repository site of radioactive wastes at Yucca Mountain, Nevada. Figure 1 shows the six hydrogeologic units that were considered in our study which, starting from the ground surface to the water table, were: the Tiva Canyon welded (TCw), the Paintbrush nonwelded (PTn), the Topopah Spring welded (TSw), the Topopah Spring vitrophyre (TSv), the Calico Hills nonwelded-vitric (CHnv), and the Calico Hills nonwelded-zeolitic (CHnz) [*Buesch et al.*, 1996; *Rousseau et al.*, 1999]. These units exhibit significant differences in their properties and their hydraulic behaviors that are illustrated in Tables 1 and 2.

[7] If one assumes, as is commonly done [*Yeh*, 1989; *Wilson et al.*, 1994; *Zhang and Yeh*, 1997; *Harter and Yeh*, 1998; *Tartakovsky et al.*, 1999], that the infiltration rate is constant at the site, and that steady state conditions have been attained, then the hydrologic state between the ground surface and the water table (pressure head and saturation against depth) can be described by Darcy's equation (1) and the *van Genuchten* [1980] relation (2) for the unsaturated hydraulic conductivity $K(h_p)$. For one-dimensional, steady

Table 1. Statistics of the Hydrogeologic Parameters for Different Layers[a]

Rock Unit	Thickness, m	K_s, m/s		α, 1/m		β	
		E( )	SD( )	E( )	SD( )	E( )	SD( )
TCw	81	4.683E-11	1.68E-10	0.01373	1.89E-2	1.6122	1.31E-1
PTn	39	2.151E-07	4.22E-06	0.1337	2.03E-1	2.3562	8.27E-1
TSw	299	4.951E-11	1.16E-10	0.01737	1.53E-2	1.7395	3.24E-1
TSv	15	1.708E-11	2.37E-11	0.00297	2.21E-3	2.3265	6.78E-1
CHnv	64	3.503E-09	1.15E-08	0.03015	2.64E-2	2.5074	8.98E-1
CHnz	127	5.170E-11	1.56E-10	0.00841	1.01E-2	1.6992	3.13E-1

[a] Read 4.683E-11 as 4.683×10^{-11}.

state, vertical infiltration through unsaturated soils [*Stephens*, 1996] the specific discharge is given by

$$q = -K(h_p)\left(\frac{dh_p}{dz} + 1\right) \tag{1}$$

where h_p is the pressure head, which is negative when the soil is unsaturated, z is the vertical Cartesian space coordinate, taken to be positive upward, and q represents infiltration. The *van Genuchten* [1980] closed-form analytical solution for the unsaturated hydraulic conductivity is given by

$$K(h_p) = K_s\left(1 + |\alpha h_p|^{\beta}\right)^{-(1-\beta^{-1})/2}\left(1 - \left[\frac{|\alpha h_p|^{\beta}}{1+|\alpha h_p|^{\beta}}\right]^{1-\beta^{-1}}\right)^2. \tag{2}$$

Here K_s is the saturated hydraulic conductivity, α the van Genuchten air entry scaling parameter, and β the van Genuchten pore size distribution index parameter. The saturation-pressure head relation is given by

$$S_w = \frac{\theta - \theta_r}{\theta_s - \theta_r} = \left(1 + |\alpha h_p|^{\beta}\right)^{-(1-\beta^{-1})} \tag{3}$$

where S_w is the effective saturation, θ is the volumetric water content, θ_s is the saturated water content, and θ_r is the residual water content.

[8] Table 1 shows the statistical properties of the saturated hydraulic conductivity K_s, the van Genuchten α and β parameters for each individual unit, as tabulated by *Wilson et al.* [1994]. The notation E( ) denotes expected value and SD( ) standard deviation. The minimum and maximum values of these parameters are given in Table 2 for each hydrogeologic unit [*Wilson et al.*, 1994]. Tables 1 and 2 illustrate the high degree of variability of the parameters within each layer and between the different layers.

3. Theoretical Framework

3.1. Relative Contribution

[9] The parameters K_s, α, and β were considered as random variables that follow the same probability distribution. Three cases of cross correlation between K_s, α, and β were investigated: uncorrelated parameters (case I), perfectly correlated parameters (case II), and negatively correlated parameters (case III). For simplicity, each parameter was assumed in this study to be spatially uncorrelated. The distributions used for the parameters in the analysis were: the lognormal, the exponential, the uniform, and the triangular. The lognormal distribution was chosen because it has been shown to fit data at several sites [*Unlu et al.*, 1990a; *Russo and Bouton*, 1992; *Boateng and Cawlfield*, 1999]. The exponential distribution is appropriate for cases where low values of a parameter (relative to the mean) are expected to occur more often than larger values. This can be the result of material changes from coarser to finer texture, sands of the same grain size but of a stronger cementation in a part of a system than another, a sand-shale system with higher percentage of shale than sand etc. The uniform distribution was chosen because it describes the (common) situation where one might have knowledge of only the range within which a parameter lies, and the triangular because it represents the case where, in addition to the minimum and maximum values, one might have information about the most commonly occurring value. All distributions were generated following the methodology described by *Law and Kelton* [1982]. The lognormal distribution for each parameter in each layer was generated such that the simulated mean and the simulated standard deviation would match the values given in Table 1. Similarly, the simulated mean of the exponential distribution would match the values given in Table 1. The range of the uniform and triangular distributions was given by the minimum and maximum values in Table 2 while the most likely value of the triangular distribution was set to be the mean value given in Table 1. Truncated forms of the lognormal and exponential distributions [*Law and Kelton*, 1982] were utilized in this study with Table 2 providing the truncation limits.

[10] By selecting a triplet of values from a specific distribution i for the parameter set (K_s, α, β) for each point of the grid and by performing a series of Monte Carlo computations with N total selections one can create N h_p profiles. At each discretization point of the grid with a coordinate z one can then average the N equiprobable values of h_p to obtain the mean capillary head $\langle h_p(z)\rangle_i$ and

Table 2. Minima and Maxima of the Parameters for Different Layers

Rock Unit	K_s, m/s		α, 1/m		β	
	Min	Max	Min	Max	Min	Max
TCw	7.00E-13	4.83E-09	0.0003	0.1338	1.349	2.805
PTn	2.86E-12	2.35E-06	0.0104	1.6990	1.187	11.800
TSw	3.05E-13	5.23E-09	0.0021	0.4244	1.155	5.363
TSv	1.52E-12	6.95E-11	0.0002	0.0077	1.377	4.473
CHnv	5.13E-12	2.92E-07	0.0054	0.3752	1.249	9.888
CHnz	2.37E-14	3.14E-09	0.0004	0.2355	1.184	5.914

the variance of h_p, σ_i^2, that applies to this point for a specific distribution i:

$$\langle h_p(z)\rangle_i = \frac{1}{N}\sum_{j=1}^{N} h_{p_j} \tag{4}$$

$$\sigma_i^2 = \frac{1}{N-1}\sum_{j=1}^{N}\left(h_{p_j} - \langle h_p(z)\rangle_i\right)^2. \tag{5}$$

Here $\langle\ \rangle$ denotes ensemble averaging. Now one can calculate, at each point z, the global mean $\langle h_p(z)\rangle_G$, defined as the arithmetic mean of the expected values from the four distributions, as well as the global variance $\langle\sigma^2\rangle_G$, defined as the arithmetic mean of the variances obtained from each distribution:

$$\langle h_p\rangle_G = \frac{1}{4}\sum_{i=1}^{4}\langle h_p(z)\rangle_i \tag{6}$$

$$\langle\sigma^2\rangle_G = \frac{1}{4}\sum_{i=1}^{4}\sigma_i^2. \tag{7}$$

By calculating the quantity $\langle\rho^2\rangle_T$,

$$\langle\rho^2\rangle_T = \frac{1}{4}\sum_{i=1}^{4}\left(\langle h_p(z)\rangle_i - \langle h_p\rangle_G\right)^2 \tag{8}$$

one can obtain at each point the divergence of the distributions means from the global mean. Then the total uncertainty on the global mean $\langle h_p(z)\rangle_G$, can be obtained at each point by [*Paleologos and Lerche*, 1999]:

$$\sigma_T^2 = \langle\rho^2\rangle_T + \langle\sigma^2\rangle_G. \tag{9}$$

Here $\langle\rho^2\rangle_T$ is a measure of the uncertainty in the mean h_p-behavior because of the uncertainty in the type of distribution, and $\langle\sigma^2\rangle_G$ is the average fluctuation around the mean h_p-behavior irrespective of distribution.

[11] One can also examine, for every point, the relative contribution of each distribution i toward the global mean [*Thomsen and Lerche*, 1997], through the expression:

$$C_m(i) = \frac{\left(\langle h_p(z)\rangle_i - \langle h_p\rangle_G\right)^2}{\sum_{i=1}^{4}\left(\langle h_p(z)\rangle_i - \langle h_p\rangle_G\right)^2} \tag{10}$$

and also calculate the relative contribution of each distribution i toward the average variance:

$$C_{var}(i) = \frac{\sigma_i^2}{\sum_{i=1}^{4}\sigma_i^2} \tag{11}$$

Both $C_m(i)$ and $C_{var}(i)$, equations (10) and (11), respectively, lie in the range of 0 to 1. Finally, by calculating the ratios

$$\frac{\langle\rho^2\rangle_T}{\sigma_T^2} = 1 - \frac{\langle\sigma^2\rangle_G}{\sigma_T^2} \tag{12}$$

and

$$\frac{\langle\sigma^2\rangle_G}{\sigma_T^2}. \tag{13}$$

one can evaluate to what degree the total uncertainty at a point of the grid is dominated by the lack of knowledge in the type of distribution or by the fluctuations around the mean values, respectively. A large value of the first ratio (12) indicates that the choice of the probability distribution model is critical in total uncertainty and, hence more data need to be collected for a clear determination of the shape of distribution. In contrast, a large value of the second ratio (13) indicates that the fluctuations around the mean h_p-behavior are dominating the system's total uncertainty and, hence the parameters need to be defined more sharply.

3.2. Relative Importance

[12] Assume that out of the three parameters K_s, α, and β one holds two at their mean values, and varies the third according to a distribution i. By performing Monte Carlo simulations one can obtain $\langle h_p\rangle_{i,k}$, the mean pressure head, and $\sigma_{i,k}^2$, the variance of the pressure head, due to the fluctuations in the kth random parameter according to an ith distribution. By repeating the procedure for all parameters the relative importance [*Lerche*, 1994; *Paleologos and Lerche*, 1999] can be evaluated toward the mean pressure head of each parameter k for every distribution i:

$$RI_{i,k}^{\langle h_p\rangle} = \frac{\langle h_p\rangle_{i,k}}{\sum_{k=1}^{3}\langle h_p\rangle_{i,k}} \tag{14}$$

as well as the relative importance toward the variance of each parameter k and distribution i:

$$RI_{i,k}^{\sigma^2} = \frac{\sigma_{i,k}^2}{\sum_{k=1}^{3}\sigma_{i,k}^2}. \tag{15}$$

This way one can determine for a given range of variation of the parameters, which parameter controls uncertainty in the pressure head. Clearly, this process can be repeated for all distributions (i = 1,. . ., 4) and then, for each parameter k, one can calculate the relative importance toward the mean:

$$RI_k^{\langle h_p\rangle} = \frac{\sum_{i=1}^{4}\langle h_p\rangle_{i,k}}{\sum_{i=1}^{4}\sum_{k=1}^{3}\langle h_p\rangle_{i,k}} \tag{16}$$

and the variance:

$$RI_k^{\sigma^2} = \frac{\sum_{i=1}^{4}\sigma_{i,k}^2}{\sum_{i=1}^{4}\sum_{k=1}^{3}\sigma_{i,k}^2} \tag{17}$$

irrespective of distribution. Thus the above expressions can provide a ranking of the importance of each parameter in the

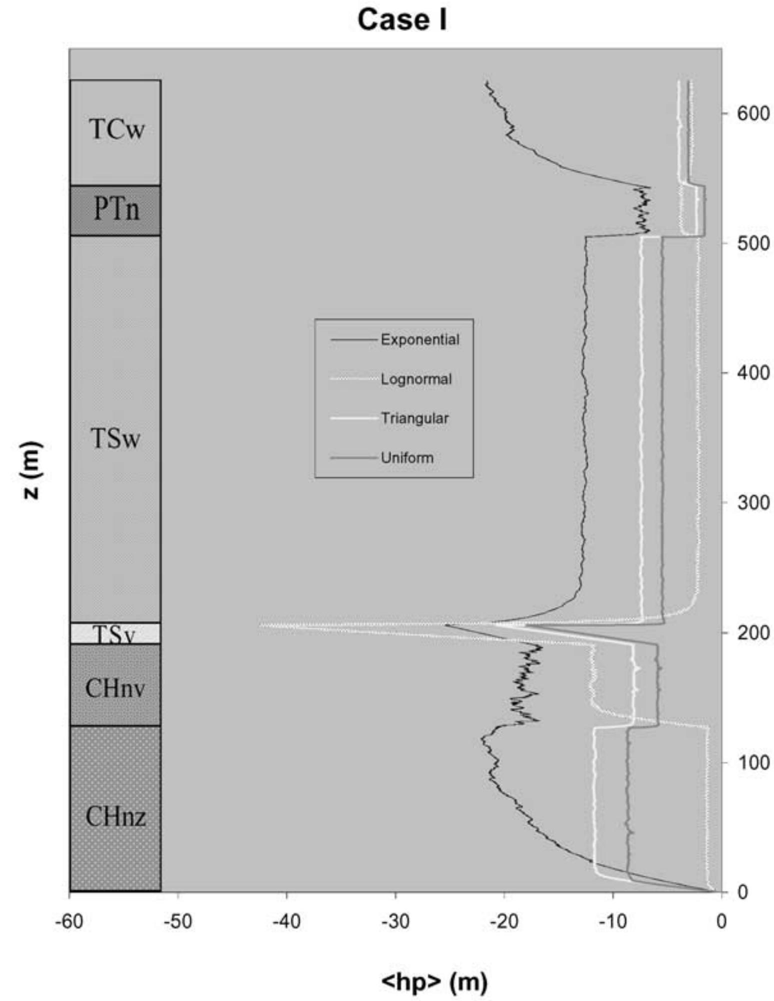

Figure 2. Case I: mean pressure head profiles.

evaluation of the mean and variance of the pressure head for a specific distribution, and irrespective of the choice of distribution.

4. Monte Carlo Results

[13] The one-dimensional flow domain of length of 625m was discretized into elements of length 0.1m, for a total of 6250 nodes. A constant infiltration rate of 0.1 mm/yr was assumed as in the studies by *Wilson et al.* [1994], *Mishra et al.* [1994], and *Reeves et al.* [1994]. The upper boundary was considered as a prescribed flux boundary, whereas the lower boundary was treated as a stationary water table. The three variables K_s, α, and β were assumed as random processes, with their statistics given by Tables 1 and 2. Equation (1) was transformed into:

$$\int_{h_{p_0}}^{h_{p_1}} \frac{K(h_p)dh_p}{(K(h_p)+q)} = -(z_1 - z_0) = -\Delta z \qquad (18)$$

which was solved iteratively through an adaptive automatic integration subroutine that was based on Newton's – Cotes 9-point rule.

[14] Three cases were analyzed. Case I, where the cross correlation between the hydraulic parameters was assumed to be negligible [*Russo and Bouton*, 1992; *Boateng and Cawlfield*, 1999; *Yeh et al.*, 1985b; *Yeh*, 1989; *Hughson and Yeh*, 1998; *Harter and Yeh*, 1998; *Tartakovsky et al.*, 1999]. Case II with the hydraulic parameters considered perfectly correlated random fields [*Yeh et al.*, 1985b; *Yeh*, 1989; *Harter and Yeh*, 1998] and case III, where a negative cross correlation between the hydraulic parameters was assumed [*Russo*, 1998]. For each node and for the parameter set (K_s, α, β) 500 Monte Carlo simulations were performed. Thus for each choice of distribution 500 equiprobable h_p profiles along the 6250 nodes were created. This process was repeated for the four types of distributions and for all three cases of cross correlation. The results are presented bellow.

4.1. Case I

[15] Figures 2 and 3 plot, for case I ($K_s, \alpha,$ and β uncorrelated random fields), the mean pressure head profile $\langle h_p \rangle$ and the variance around this profile, respectively. Figures 2 and 3 indicate that the exponential distribution function produces the largest (in absolute value) mean pressure head and head-variance profiles than any other statistical model

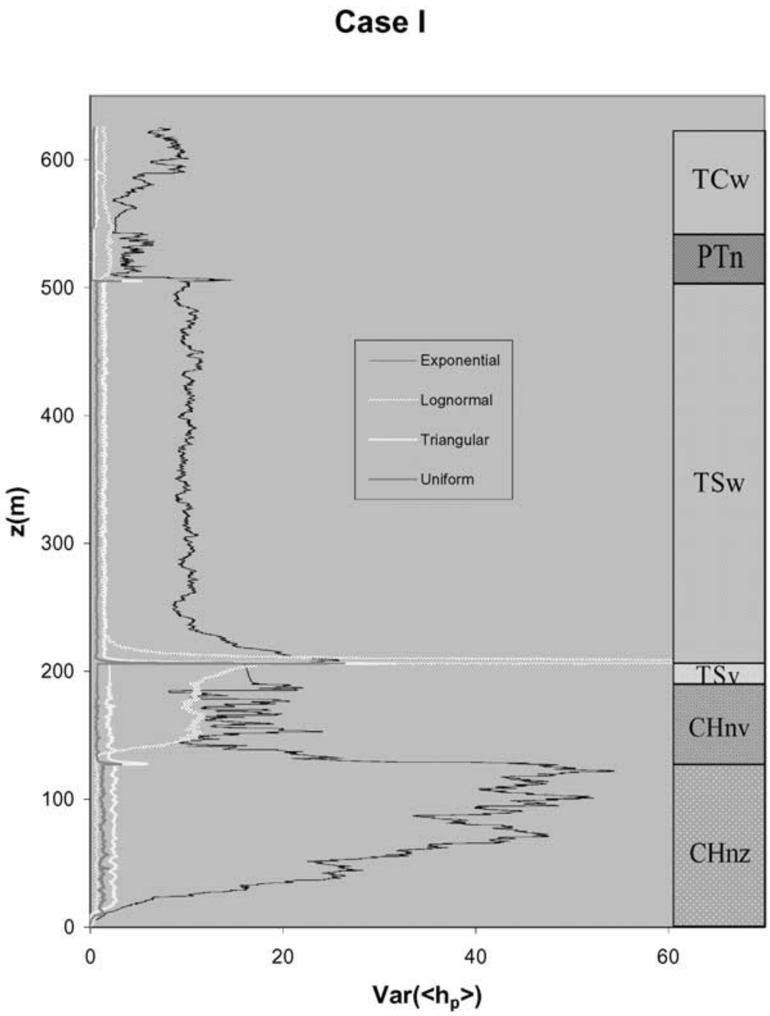

Figure 3. Case I: variance of the mean pressure head profiles.

in all layers except in the TSv layer where the lognormal distribution produced the largest profile and the largest uncertainty. The lognormal distribution produced the smallest (in absolute value) mean pressure head profiles in the CHnz, TSw, and TCw layers whereas the smallest mean pressure head profile in the remaining layers was generated by the uniform distribution. Our results in Figure 2 can be explained by considering the preference of the exponential model to select, within a fixed range, lower K_s values than any other probability distribution model, which given the inverse relation between pressure head and K_s leads (all other factors being equal) to the exponential model producing the largest (absolutely) pressure head profile in the majority of the layers. The uniform model resulted in the smallest fluctuations around the $\langle h_p \rangle$ profile in all layers except the CHnz layer where the uniform and lognormal gave close results. Figures 2 and 3 indicate that irrespectively of distribution type the variance of the mean pressure head increases with increasing (in absolute value) mean pressure head. This result is in agreement with the findings of stochastic studies by *Yeh et al.* [1985b], field observations by *Yeh et al.* [1986] and analytical results by *Yeh* [1989]. Additionally, as also noted by *Yeh* [1989], the variance in pressure head is seen (Figure 3) to be minimal near the water table, which can be attributed to the capillary fringe effect.

[16] Figures 4 and 5 plot for case I the mean saturation profile $\langle S \rangle$ and the variance around this profile, respectively. Figures 4 and 5 indicate that the use of a lognormal or an exponential distribution produces significantly larger mean saturation profiles (values near saturation) and significantly smaller fluctuations about this mean profile, than any other statistical model, in almost every layer. If one considers β fixed at about the value of two and the relatively small standard deviation of β (see Table 1) then at large $|\alpha h_p|$-values equation (3) indicates that the saturation is approximately proportional to $|\alpha h_p|^{-1}$. Because $|h_p|$ is inversely proportional to K_s then the saturation is approximately proportional to $|K_s/\alpha|$. Therefore the statistical distribution of saturation depends upon the relative distributions of K_s and $1/\alpha$. Figure 4 indicates that for uncorrelated parameters the dominance of the small values of α for an exponential distribution outweighs the fluctuations produced by K_s so that the typical saturation would be relatively high in comparison to the other distribution choices. For layer TSv the mean saturation profile and the variations around it are almost the same for all types of distribution. Figures 4 and 5 show that the saturation variance increases with

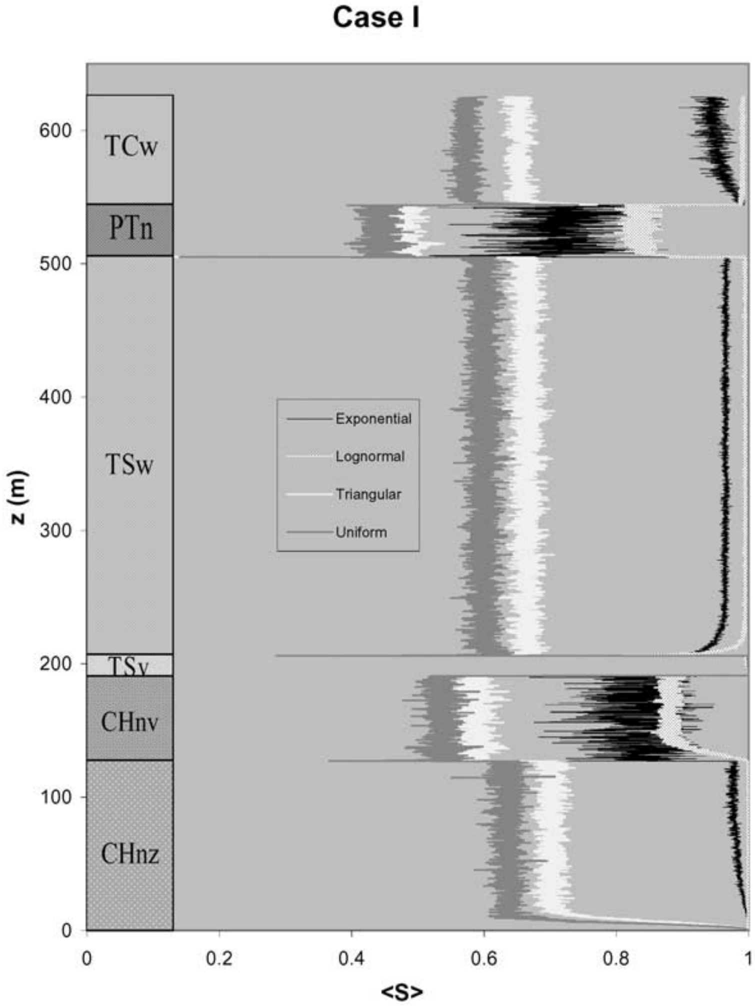

Figure 4. Case I: mean saturation profiles.

decreasing mean saturation values. The variance of mean saturation appears to increase in coarse-textured soils (soils that exhibit high α and K_s values), a result that agrees with the conclusions of *Harter and Zhang* [1999]. The variability in saturation is minimal near the water table irrespective of distribution type. The strong dependence of pressure head and saturation results on the shape of probability distribution supports the arguments of *Harter and Yeh* [1996] and *Harter and Zhang* [1999] to use data to condition numerical simulations.

[17] In order to simplify the depiction of our results for the relative contribution and importance of the parameters the detailed mean point profiles were averaged over the six layers for each distribution. Such lumping of results over all the formations is customarily done in studies of total system performance assessments [*Wilson et al.*, 1994] where the hydrologic behavior is only one component of analyzing a system's response that may include geochemical, structural, and other considerations. Figures 6 and 7 plot the relative contribution of each distribution toward the global mean and toward the average variance, respectively. Figures 6 and 7 demonstrate that the use of an exponential model for the parameters produces a mean pressure head profile, as well as fluctuations around this profile, that are significantly larger than of any other statistical model. The importance of assumptions made on the probability distribution is exhibited by the values of the ratios $\langle\rho^2\rangle_T/\sigma_T^2$ and $\langle\sigma^2\rangle_G/\sigma_T^2$, equations (12) and (13), that are presented in Table 3. A large value of the first ratio indicates that the choice of a particular model is critical in total uncertainty and more data need to be collected for a clear determination of the shape of distribution and that is observed for all the layers examined separately [*Avanidou*, 2000]. When the total system is considered however less than half of the total system's uncertainty is attributed to the distribution model assumption.

4.2. Cases II and III

[18] For case II parameters K_s, α, and β were considered as perfectly correlated random fields, whereas in case III α and β were considered negatively correlated with K_s. Only the graphs for case II are shown here; the graphs for case III have been omitted because of the similarity of the results with case I or case II, but they are given by *Avanidou* [2000]. Only a brief discussion is provided here that compares the outcome of negatively correlated parameters with the other cases. Figures 8 and 9 plot for all layers the mean pressure head $\langle h_p\rangle$ profile and the variations around

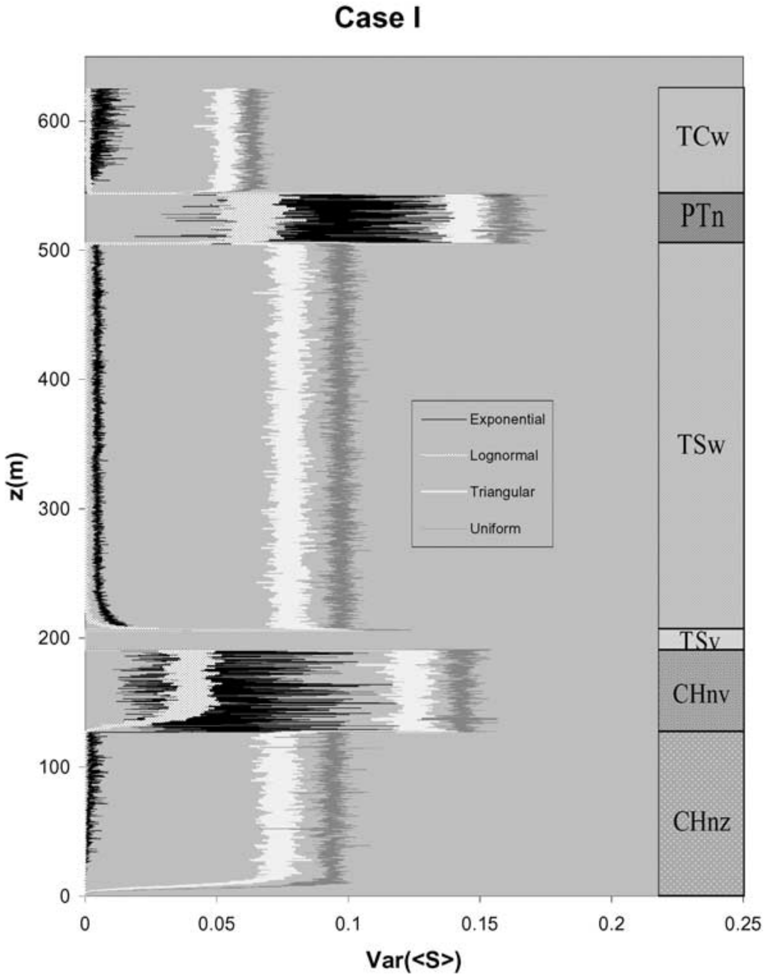

Figure 5. Case I: variance of the mean saturation profiles.

this profile, respectively. The use of an exponential distribution produces the largest $|\langle h_p \rangle|$ profile for all layers except for the TSv and PTn layers where the lognormal model produces the largest $|\langle h_p \rangle|$ values; the graph for case III closely resembles Figure 2 with the more pronounced difference being the dominance of the lognormal distribution in both the TSv and CHnv layers. Figure 8 also shows that the uniform distribution consistently underestimates,

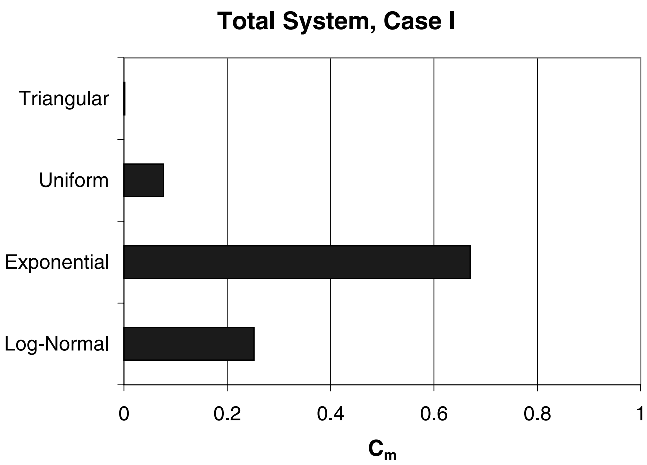

Figure 6. Total system, case I: uncorrelated parameters; contribution of each distribution toward the global mean, equation (10).

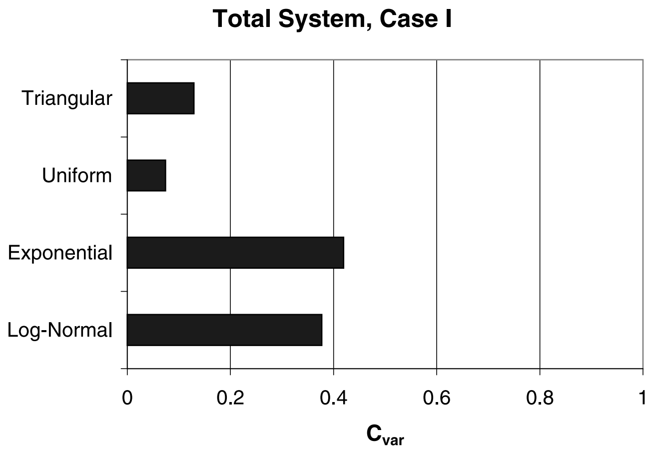

Figure 7. Total system, case I: uncorrelated parameters; contribution of each distribution toward the average variance, equation (11).

Table 3. Values for Expressions 12 and 13, Case I

Layer	CHnz	CHnv	TSv	TSw	PTn	TCw	Total
$\frac{\langle \rho^2 \rangle_T}{\sigma_T^2}$	0.6336	0.7089	0.5196	0.7177	0.7188	0.8829	0.4676
$\frac{\langle \sigma^2 \rangle_G}{\sigma_T^2}$	0.3664	0.2911	0.4804	0.2823	0.2812	0.1171	0.5324

relative to the other distributions, the mean pressure head for all layers, whereas in case I this was only true for the PTn, TSv, CHnv layers and for case III for the TCw, PTn, TSv, and CHnv layers. Figure 9 indicates that the lognormal distribution produces everywhere the largest uncertainty associated with the $\langle h_p \rangle$ profile except at the PTn–TSw, and CHnv–CHnz interfaces where the dominance of the exponential model may be explained by the large jumps in the mean pressure head values at these interfaces. In contrast with case I and with *Yeh et al.* [1985b] cases II and III show that the variability in pressure head is decreasing with increasing (absolutely) mean pressure head values. This discrepancy may be attributed to the value of infiltration rate that was used in our study since for perfectly correlated parameters the infiltration rate has been seen [*Yeh*, 1989] to influence the variability in pressure head. *Yeh* [1989, Figure 4] found that the pressure head variability decreased as the infiltration rate decreased up to a limiting infiltration rate value beyond which the pressure head variance increased. Figures 3 and 9 indicate that for the exponential distribution a positive (or negative) cross correlation between the parameters produces more uniform pressure head profiles than when parameters are uncorrelated. The lognormal distribution appears to give smaller variances in $\langle h_p \rangle$ when the parameters are uncorrelated (or negatively correlated) than when a positive cross correlation is used. The other two distribution choices appear to be insensitive to the form of cross correlation with regards to the variance of $\langle h_p \rangle$.

[19] Figures 10 and 11 plot for all layers the mean saturation $\langle S \rangle$ profile and the variance around it, respectively. Figure 10 shows that the lognormal distribution produces the largest mean saturation profile in every layer. Similar to case I, in layer TSv the mean saturation profile is almost identical for all choices of distributions (values near saturation). The lognormal model produced (Figure 11) the

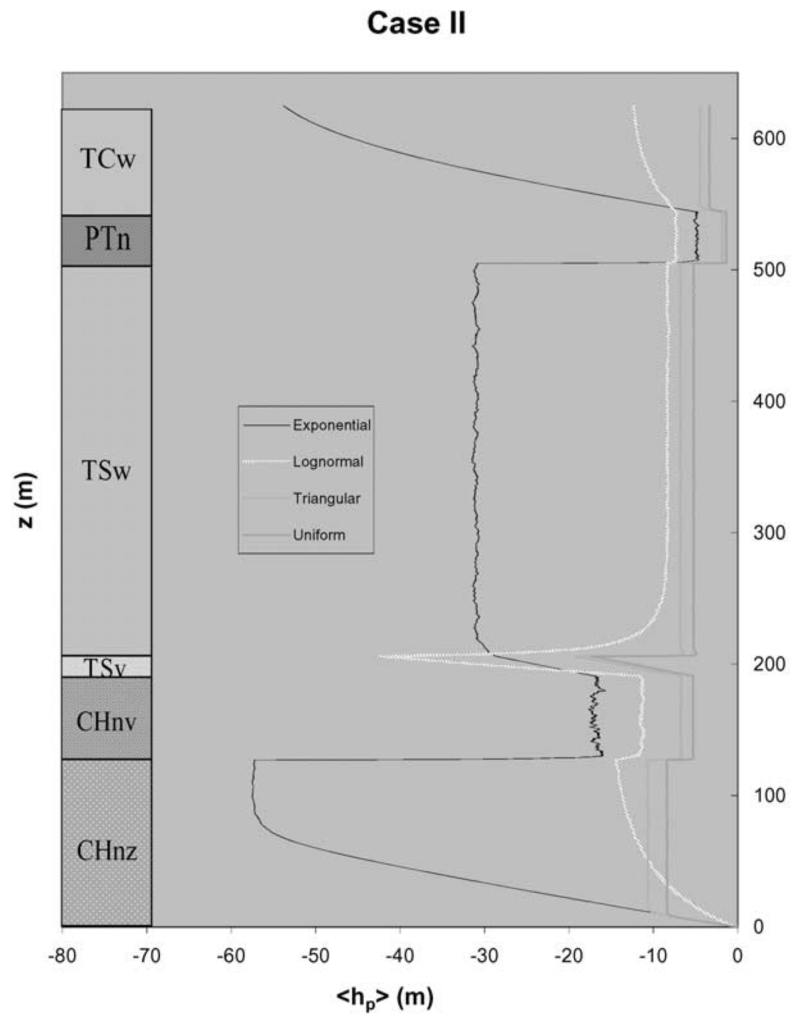

Figure 8. Case II: mean pressure head profiles.

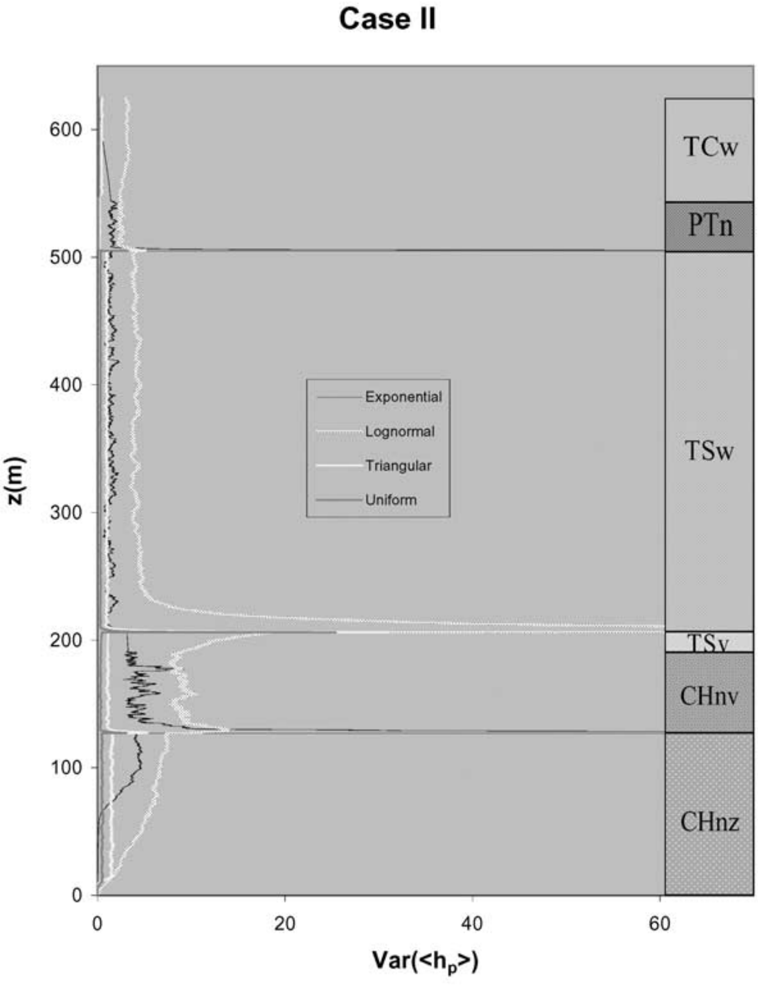

Figure 9. Case II: variance of the mean pressure head profiles.

smallest variations around the $\langle S \rangle$ profile for all layers except for the CHnz and PTn where the exponential model produced the smallest variance values. The corresponding figures for case III are not shown here due to the similarity with Figures 4 and 5. Figures 10 and 11 indicate that the saturation variance increases with a decrease in the mean saturation a conclusion shared with cases I and III. Our results agree with the conclusions of *Harter and Zhang* [1999] that the variance of mean saturation increases with increasing α and K_s values. Furthermore, for all cases I, II, and III, the variability in pressure head and saturation were seen to be minimal near the water table irrespectively of distribution type, which can be attributed to the capillary fringe effect.

[20] Figures 12 and 13 plot for the total system the relative contribution of each distribution model toward the global mean and global variance, respectively. These plots were produced by averaging the detailed mean point profiles over the six layers for each distribution. Similar to the results of case I, in case II the exponential model produced the largest $|\langle h_p \rangle|$ profiles and the largest fluctuations about this profile. In case III the exponential model also produced the largest $|\langle h_p \rangle|$ profiles (with similar values for C_m with case II) but the lognormal model produced the largest fluctuations around the $\langle h_p \rangle$ profile (in contrast to cases I and II).

4.3. Relative Importance

[21] In accordance with the procedure of section 3.2 K_s was considered a random variable following one distribution type and α and β were taken as constants equal to the mean values given for each layer in Tables 1 and 2. 500 Monte Carlo simulations were performed and the mean pressure head, $\langle h_p \rangle_{i,k}$, and the variance of the mean pressure head, $\sigma^2_{i,k}$, were obtained due to the fluctuations in the K_s parameter. This procedure was repeated for all parameters and for all four distributions. The mean point profiles were averaged over the six layers for each distribution.

[22] Figures 14 and 15 plot for the total system the relative importance toward the mean and the variance of the pressure head, equations (14) and (15), respectively, of each parameter and distribution model. Figures 14 and 15 show that the prediction of mean and variance in pressure head is dominated for the uniform and triangular models by K_s and β. The exponential appears to depend equally on all three parameters. The results for the lognormal distribution indicate a stronger dependence on β (primarily for the mean pressure head profile), followed by K_s. Similar graphs were produced

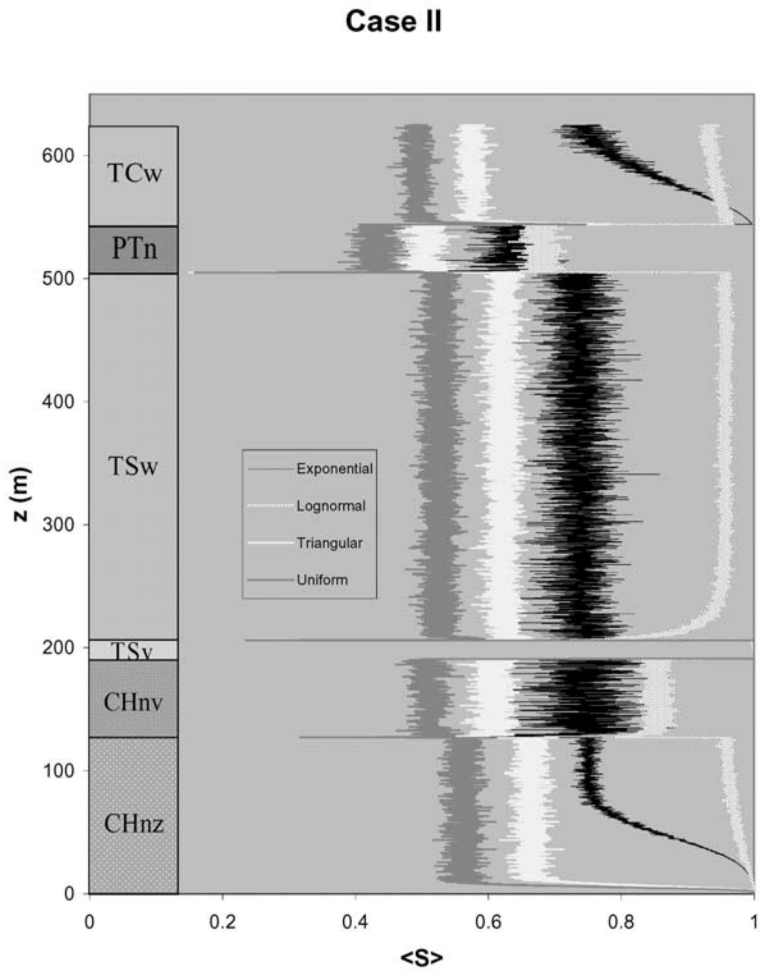

Figure 10. Case II: mean saturation profiles

for each individual layer and showed that the ranking (and relative percentage) of each parameter was not the same for all formations. For the lognormal distribution the results for the relative importance toward the mean in the CHnv, TSv, PTn, and TCw, layers qualitatively agree with the rankings shown in Figure 14. In contrast for the CHnz and TSw layers it was found that α, rather than K_s, is the second dominant parameter. For the lognormal distribution with respect to the variance the CHnv, TSw, PTn, and TCw qualitatively agree with Figure 15 whereas for the CHnz and TSv layers β and α were found to be dominant. *Mishra et al.* [1994] analyzed the marginal sensitivities at three depths that represented, respectively, the upper boundary of the TSw unit, the proposed repository horizon (within this unit), and the lower boundary of the TSw formation. These authors found that, for the same infiltration rate as ours and for a lognormal distribution, the sensitivities in water saturation at all three depths exhibited a stronger dependence on β and K_s rather than α.

[23] Figures 16 and 17 plot for each layer the relative importance toward the mean and the variance of each parameter irrespective of distribution. Figures 16 and 17 suggest that the saturated hydraulic conductivity was the most important parameter toward total system uncertainty, with $RI_k^{\langle h_p \rangle}$ values greater than 0.49 for all layers, with the exception of the CHnz layer where β was the one that dominated. This result agrees with studies by *Chen et al.* [1994a, 1994b] about the dominance of the saturated hydraulic conductivity in unsaturated flow predictions. For all layers (except the CHnz) the second most important parameter was β (with $RI_k^{\langle h_p \rangle}$ values always larger than 0.26), and the least important was α. This result is in agreement with *Boateng and Cawlfield* [1999] on the importance of the β parameter but do not support these authors' conclusion that the saturated hydraulic conductivity can be considered as a deterministic variable with no significant effect on the probability outcome.

5. Summary and Conclusions

[24] In this study a framework was presented that allows the analysis of the uncertainty, relative contribution, and relative importance of spatially variable parameters. This methodology was applied to the problem of one-dimensional infiltration through six hydrogeologic units that represent a typical cross section of the unsaturated flow domain at the Yucca Mountain site, Nevada. The parameters (K_s, α, and β) that enter the van Genuchten expression for the unsaturated hydraulic conductivity were modeled as random variables with means, coefficients of variation and ranges defined

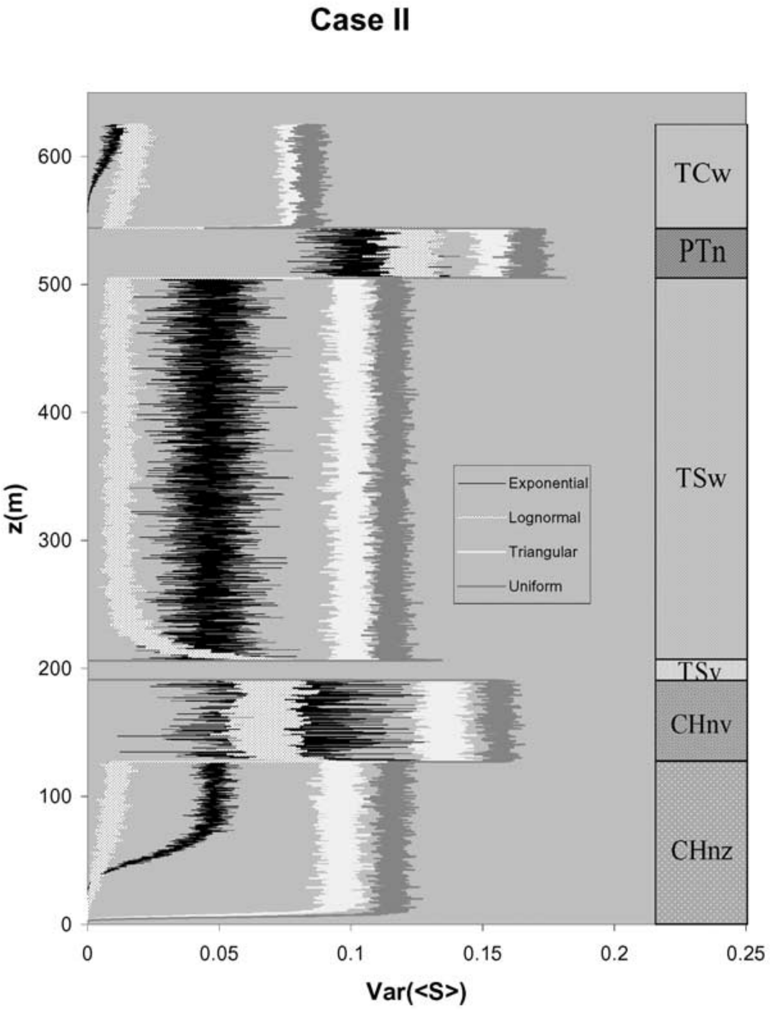

Figure 11. Case II: variance of mean saturation profiles.

from field data. Four different distributions were investigated: the lognormal, the exponential, the uniform and the triangular. Because of the lack of data regarding the cross correlation between the parameters three extreme situations were considered: case I, where the hydraulic parameters were treated as uncorrelated random variables, case II, where they were modeled as perfectly correlated, and case III, where a negative cross correlation was used.

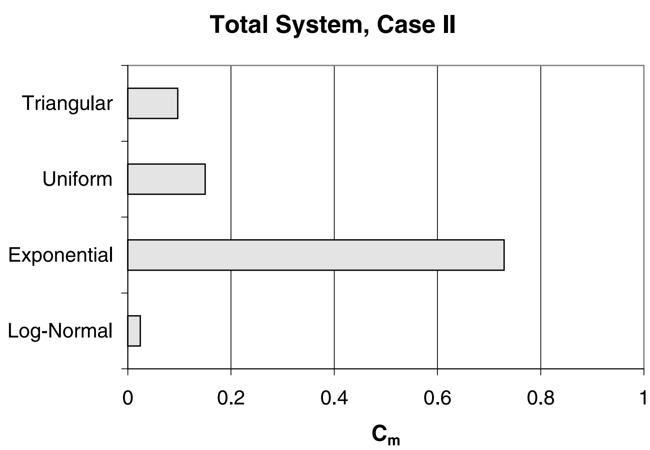

Figure 12. Total system, case II: perfectly correlated parameters; contribution of each distribution toward the global mean, equation (10).

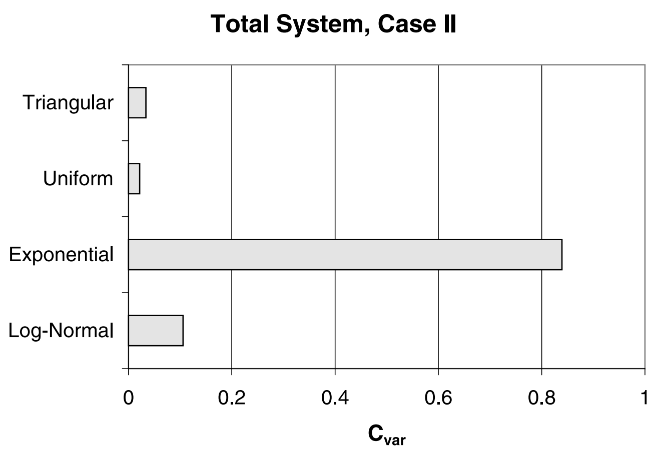

Figure 13. Total system, case II: perfectly correlated parameters; contribution of each distribution toward the average variance, equation (11).

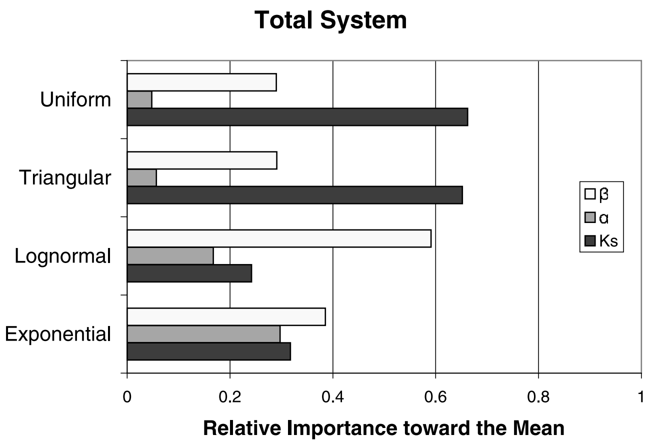

Figure 14. Total system: relative importance of each parameter for each distribution type toward the mean, equation (14).

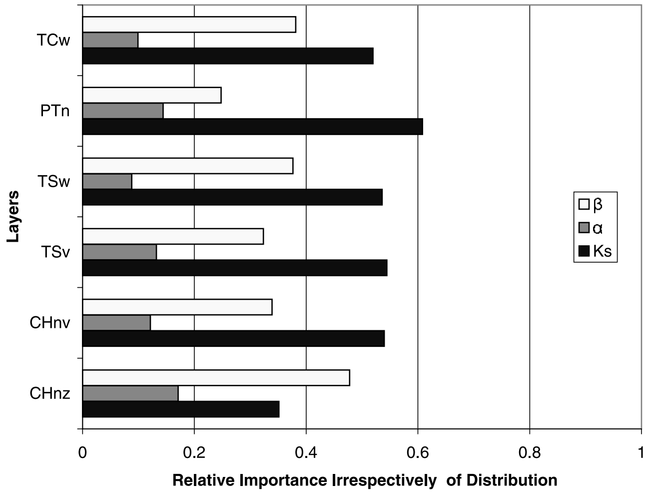

Figure 16. Relative importance of each parameter in each layer toward the mean irrespectively of choice of distribution type, equation (16).

[25] For all cross-correlation cases and for the majority of layers, when a total system behavior was considered, the exponential model produced a mean and standard error in the pressure head profile that were significantly larger than of any other statistical model. The lognormal distribution was the one that produced the largest mean saturation profile. Our results indicate that the assumption of probability distribution model for this problem is critical because for all cross-correlation cases more than 50% of the total uncertainty in each layer is due to the uncertainty in the type of distribution. The dependence of mean pressure head and saturation profiles on the distribution model indicates that definition of the parameters first moments only is not sufficient for accurate flow prediction but the functional form of the probability distribution should be better resolved. For all cross correlation cases the saturation variance increased with a decrease in the mean saturation. Our results agree with the conclusions of *Harter and Zhang* [1999] that the variance of mean saturation increases with increasing α and K_s values. Our results for uncorrelated parameters agree with *Yeh et al.* [1985b, 1986] and *Yeh* [1989] that the variance in pressure head increases with increasing mean pressure head values.

[26] This article also presents a methodology that accounts for the relative importance of each uncertain parameter. Monte Carlo simulations were used to calculate measures of the relative importance of the K_s, α, and β parameters for the four probability distribution models. For individual layers the relative importance measures in terms of mean and variance of the pressure head varied. Overall it appears that for all layers and models the saturated hydraulic conductivity and the pore size distribution index need to be treated as stochastic parameters with the exact statistical model clearly defined for each layer in order to establish the correct hydraulic behavior. Our results agree with those by *Chen et al.* [1994a, 1994b] about the dominance of the saturated hydraulic conductivity in unsaturated flow predictions. They also agree with those by *Boateng and Cawlfield* [1999] on the importance of the

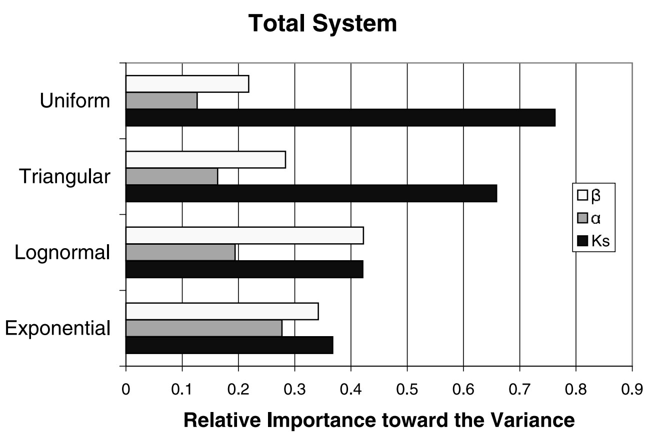

Figure 15. Total system: relative importance of each parameter for each distribution type toward the variance equation (15).

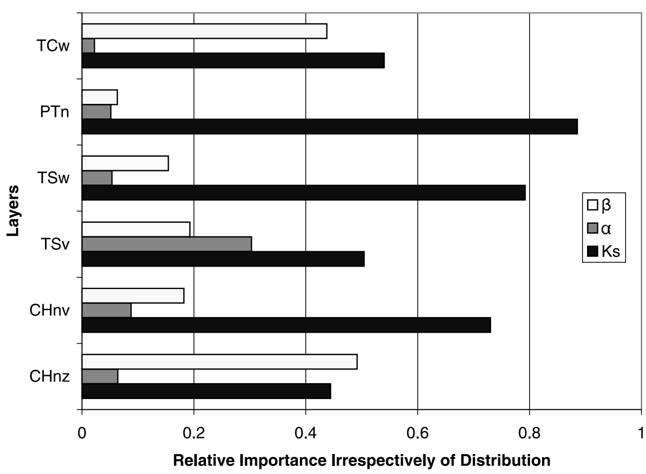

Figure 17. Relative importance of each parameter in each layer toward the variance irrespectively of choice of distribution type, equation (17).

β parameter but do not support these authors' conclusion that the saturated hydraulic conductivity can be considered as a deterministic variable with no significant effect on the probability outcome. Finally, our results for the TSw hydrogeologic unit, site of the potential nuclear waste repository, and for a lognormal distribution agree with the conclusions by *Mishra et al.* [1994] about the importance of β and K_s rather than α on water saturation in this unit.

[27] **Acknowledgments.** This work was supported in part by grant DE-FG02-97EW09999 from the U.S. Department of Energy, Office of Environmental Management, to the center for Water Research at the University of South Carolina. The authors would also like to acknowledge Alexandre Desbarats, of Geological Survey of Canada, and Ian Lerche, of the University of South Carolina, for their help at various stages of this study. Finally, we would like to thank the anonymous reviewers for bringing to our attention some related work and for their suggestions that improved the presentation of this manuscript.

References

Avanidou, T., Infiltration in stratified soils: Relative importance of parameters and model variations, M.S. thesis, 107 pp., Univ. of S. C., Columbia, 2000.

Boateng, S., and J. D. Cawlfield, Two-dimensional sensitivity analysis of contaminant transport in the unsaturated zone, *Ground Water*, *37*(2), 185–193, 1999.

Buesch, D. C., R. W. Spengler, T. C. Moyer, and J. K. Geslin, Proposed stratigraphic nomenclature and macroscopic identification of lithostratigraphic units of the Paintbrush Group exposed at Yucca Mountain, Nevada, *U.S. Geol. Surv. Open File Rep., 94-469*, 45 pp., 1996.

Chen, Z., R. S. Govindaraju, and M. L. Kavvas, Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields, 1, Development of models, *Water Resour. Res.*, *30*(2), 523–533, 1994a.

Chen, Z., R. S. Govindaraju, and M. L. Kavvas, Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields, 2, Numerical simulations, *Water Resour. Res.*, *30*(2), 535–548, 1994b.

Dagan, G., and E. Bresler, Unsaturated flow in spatially variable fields, 1, Derivation of models of infiltration and redistribution, *Water Resour. Res.*, *19*(2), 413–420, 1983.

Freeze, R. A., A stochastic conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, *11*, 725–741, 1975.

Gardner, W. R., Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table, *Soil Sci.*, *85*(4), 228–232, 1958.

Harter, T., and T.-C. J. Yeh, Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils, *Water Resour. Res.*, *32*(6), 1597–1609, 1996.

Harter, T., and T.-C. J. Yeh, Flow in unsaturated random porous media, nonlinear numerical analysis and comparison to analytical stochastic models, *Adv. Water Resour.*, *22*(3), 257–272, 1998.

Harter, T., and D. Zhang, Water flow and solute spreading in heterogeneous soils with spatially variable water content, *Water Resour. Res.*, *35*(2), 415–426, 1999.

Hughson, D. L., and T.-C. J. Yeh, A geostatistical based inverse model for three-dimensional variably saturated flow, *Stochastic Hydrol. Hydraul.*, *12*, 285–298, 1998.

Law, A. M., and W. D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill, New York, 1982.

Lerche, I., Ranges of probable uncertainty in gas resource estimates for the USA, *Energy Exp. Exploitation*, *12*(5), 381–392, 1994.

Looney, B. B., and R. W. Falta, *Vadose Zone Science and Technology Solutions*, Battelle, Columbus, Ohio, 2000.

Mishra, S., Y. Xiang, B. Dunlap, and R. W. Andrews, Analysis of steady-state infiltration into Yucca Mountain using 1-D and 2-D models: Preliminary results, *Rep. B00000000-01717-0200-00123*, Dep. of Energy, Washington, D. C., 1994.

Nielsen, D. R., J. W. Biggar, and K. T. Erh, Spatial variability of field-measured soil-water properties, *Hilgardia*, *42*, 215–260, 1973.

Paleologos, E. K., and I. Lerche, Multiple decision-making criteria in the transport and burial of hazardous and radioactive wastes, *Stochastic Environ. Res. Risk Assess.*, *13*(6), 381–395, 1999.

Reeves, M., N. A. Baker, and J. O. Duguid, Review and selection of unsaturated flow models, *Rep. B00000000-01425-2200-00001 Rev.00*, Dep. of Energy, Washington, D. C., 1994.

Rousseau, J. P., E. M. Kwicklis, and D. C. Gillies, Hydrogeology of the unsaturated zone, north ramp area of the exploratory studies facility, Yucca Mountain, Nevada, *U.S. Geol. Surv. Water Resour. Invest.*, 98–4050, 1999.

Russo, D., A geostatistical approach to the trickle irrigation design in heterogeneous soil, 1, Theory, *Water Resour. Res.*, *19*, 632–642, 1983.

Russo, D., A geostatistical approach to the solute transport in heterogeneous fields and its applications to salinity management, *Water Resour. Res.*, *20*, 1260–1270, 1984.

Russo, D., Stochastic analysis of flow and transport in unsaturated heterogeneous porous formation: Effects of variability in water saturation, *Water Resour. Res.*, *34*, 569–581, 1998.

Russo, D., and M. Bouton, Statistical analysis of spatial variability in unsaturated flow parameters, *Water Resour. Res.*, *28*, 1911–1925, 1992.

Russo, D., I. Russo, and A. Laufer, On the spatial variability of parameters of the unsaturated hydraulic conductivity, *Water Resour. Res.*, *33*, 947–956, 1997.

Smith, L., Spatial variability of flow parameters in a stratified sand, *Math. Geol.*, *13*(1), 1–21, 1981.

Smith, R. E., and B. Diekkruger, Effective soil water characteristics and ensemble soil water profiles in heterogeneous soils, *Water Resour. Res.*, *32*, 1993–2002, 1996.

Stephens, D. B., *Vadose Zone Hydrology*, CRC Press, Boca Raton, Fla., 1996.

Sudicky, E. A., A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity, *Water Resour. Res.*, *22*, 2069–2082, 1986.

Tartakovsky, D. M., S. P. Neuman, and Z. Lu, Conditional stochastic averaging of steady state unsaturated flow by means of Kirchhoff transformation, *Water Resour. Res.*, *35*, 731–745, 1999.

Thomsen, R. O., and I. Lerche, Relative contributions to uncertainties in reserve estimates, *Mar. Petrol. Geol.*, *14*(1), 65–74, 1997.

Unlu, K., M. L. Kavvas, and D. R. Nielsen, Stochastic analysis of field measured unsaturated hydraulic conductivity, *Water Resour. Res.*, *25*(12), 2511–2519, 1989.

Unlu, K., D. R. Nielsen, J. W. Biggar, and F. Morkoc, Statistical parameters characterizing the spatial variability of selected soil hydraulic properties, *Soil Sci. Soc. Am. J.*, *54*, 1537–1547, 1990a.

Unlu, K., D. R. Nielsen, and J. W. Biggar, Stochastic analysis of unsaturated flow: One-dimensional Monte Carlo simulations and comparison with spectral perturbation analysis and field observations, *Water Resour. Res.*, *26*, 2207–2218, 1990b.

van Genuchten, M. T., A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*, 892–898, 1980.

Warrick, A. W., G. J. Mullen, and D. R. Nielsen, Predictions of the soil water flux based upon field-measured soil-water properties, *Soil Sci. Soc. Am. J.*, *41*, 414–419, 1977.

White, I., and M. J. Sully, On the variability and use of the hydraulic conductivity alpha parameter in stochastic treatments of unsaturated flow, *Water Resour. Res.*, *28*, 209–213, 1992.

Wierenga, P. J., R. G. Hills, and D. B. Hudson, The Las Cruces trench site: Characterization, experimental results, and one-dimensional flow predictions, *Water Resour. Res.*, *27*, 2695–2705, 1991.

Wilson, M. L., et al., Total-system performance assessment for Yucca Mountain-SNL second iteration, (TSPA-1993), *Tech. Rep. SAND93-2675*, Sandia Natl. Lab., Albuquerque, N. M., 1994.

Woodbury, A. D., and E. A. Sudicky, The geostatistical characteristics of the Borden aquifer, *Water Resour. Res.*, *27*, 533–546, 1991.

Yeh, T.-C. J., One-dimensional steady state infiltration in heterogeneous soils, *Water Resour. Res.*, *25*, 2149–2158, 1989.

Yeh, T.-C. J., and J. Zhang, A geostatistical inverse method for variably saturated flow in the vadose zone, *Water Resour. Res.*, *32*, 2757–2766, 1996.

Yeh, T.-C. J., L. W. Gelhar, and A. L. Gutjahr, Stochastic analysis of unsaturated flow in heterogeneous soils, 1, Statistically isotropic media, *Water Resour. Res.*, *21*, 447–456, 1985a.

Yeh, T.-C. J., L. W. Gelhar, and A. L. Gutjahr, Stochastic analysis of unsaturated flow in heterogeneous soils, 2, Statistically anisotropic media with variable α, *Water Resour. Res.*, *21*, 457–464, 1985b.

Yeh, T.-C. J., L. W. Gelhar, and A. L. Gutjahr, Stochastic analysis of

unsaturated flow in heterogeneous soils, 3, Observations and applications, *Water Resour. Res.*, *21*, 465–472, 1985c.
Yeh, T.-C. J., L. W. Gelhar, and P. J. Wierenga, Observations of spatial variability of soil-water pressure in a field soil, *Soil Sci.*, *142*(1), 7–12, 1986.
Zhang, J., and T.-C. J. Yeh, An iterative geostatistical inverse method for steady state flow in the vadose zone, *Water Resour. Res.*, *33*, 63–71, 1997.

T. Avanidou and E. K. Paleologos, Department of Geological Sciences, University of South Carolina, Columbia, SC, USA. (epal@geol.sc.edu)