An Effective Medium Approximation and Monte Carlo Simulation in Subsurface Flow Modeling

by

Surendra Prakash Shrestha

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in

Charles E. Via, Jr. Department of Civil Engineering

APPROVED:

G. V. Loganathan, Chairman

Theo A. Dillaha

Conrad D. Heatwole

John C. Parker

Tamim Younos

July 1993

Blacksburg, Virginia

An Effective Medium Approximation and Monte Carlo Simulation in Subsurface Flow Modeling

by

Surendra Prakash Shrestha
G. V. Loganathan, Chairman

(ABSTRACT)

An effective medium approximation and a refined Monte Carlo simulation procedure for solving the stochastic groundwater flow problem are presented. The effective medium approximation permits one to solve the stochastic groundwater flow problem in a single run to generate the expected pressure head field. The proposed effective hydraulic conductivity expression for the effective medium is of the same form as the local Gardner's equation and is easy to use.

The refined Monte Carlo simulation procedure uses analytical means to estimate the sample size by controlling the error incurred in using the sample average in place of its population mean at a chosen confidence level. This estimator consistently performs well. Also, a variance reducing estimator which is different from the simple average for pressure head is developed. This estimator takes advantage of the correlation between the saturated conductivity and the pressure head distribution to reduce the output variance and is unbiased. This reduced variance results in a smaller width of uncertainty about the predicted pressure head. Both the effective medium approximation and the Monte Carlo approaches perform well when applied to several problems.

In loving memory of my father

Ratna Kazi Shrestha

Acknowledgements

I express my deepest gratitude to Professor G. V. Loganathan whose guidance, encouragement, and support were invaluable to the progress and completion of this study. I extend my special thanks to Professor John C. Parker for his valuable suggestions, technical guidance, and careful review of this manuscript. I would also like to thank Dr. Tamim Younos for helpful comments and suggestions; and Professors Theo A. Dillaha and Conrad D. Heatwole for their suggestions, support, and careful review of the manuscript. I would also like to thank Professor Chin Y. Kuo for the encouragement and guidance during the early stages of my program of study.

I am thankful to Professors Marion R. Reynolds, Martin V. Day, and Keying Ye for their time and help with numerous questions during the mathematical development of this thesis. My warmest thanks to Dr. Parameshwar L. Shrestha and my fellow graduate students in Hydrosystems division for providing an enjoyable and fruitful working environment. I am grateful to Professors David F. Kibler and William E. Cox for providing the departmental support required for the successful completion of my graduate studies.

My sincere appreciation and love to my mother, sister, and brothers for their inspiration and support. Special thanks to my brother Bijaya for his encouragement and support.

My deepest gratitude and love to my wife, Sulekha and my daughters, Shweta and Srishti who have endured the numerous hours I spent in my office through patience, understanding, love and support.

Table of Contents

List of Figures

List of Tables

List of Symbols

$C(\psi)$	soil water capacity
$C_{X,Y}$	sample covariance of (X,Y)
e, exp	exponential operator
f	fluctuations in log-saturated hydraulic conductivity
F	mean log-saturated hydraulic conductivity
g	fluctuations in log-unsaturated hydraulic conductivity
g	acceleration due to gravity (chapter 2)
G	mean log-unsaturated hydraulic conductivity
h	phreatic surface level (chapter 5; section 6.6)
h	fluctuations in pressure head
H	mean pressure head
H_o	initial height of phreatic surface
J_i	hydraulic gradient in the i^{th} direction
k	intrinsic permeability of the porous medium
K, $K(\psi)$	unsaturated hydraulic conductivity of the porous medium
K_A	arithmetic mean of generated hydraulic conductivity
K_G	geometric mean of generated hydraulic conductivity
K_H	harmonic mean of generated hydraulic conductivity
K_{eff}	effective hydraulic conductivity of the porous medium
K_s	saturated hydraulic conductivity of the porous medium
K_s^*	effective saturated hydraulic conductivity of the porous medium
L	length of the flow domain
m	number of blocks
M	a positive constant

MC	*Monte Carlo*
n	number of Monte Carlo simulation runs
N	sample size using Hoeffding inequality
N	number of wells (chapter 5)
N_C	sample size using Chebyshev inequality
p	probability density function (chapter 3)
p	a constant
PDE	Partial Differential Equation
q	specific discharge
Q	pumping rate
q_i	specific discharge in the i-direction
r	radial distance between the pumping well and the observation well
R	radius of influence of the pumping well
R_N	standard normal random deviate
SP	Spectral Perturbation
S_X^2	sample variance of X
t	time
VRT	Variance Reduction Technique
x_i	cartesian coordinate axes
Y	log-saturated hydraulic conductivity of the porous medium
z	z-value from standard normal table
α	Gardner's soil parameter
β	weight for variance reduction
β^*	optimal weight for variance reduction
Γ	system boundary
δ	confidence coefficient
ϵ	absolute error (positive number)
η	dynamic viscosity of flowing fluid
θ	moisture content of the porous medium
μ_n	n-th central moment
μ_X	population mean of X

π	pi
ρ	density of flowing fluid
$\rho_{X,Y}$	correlation coefficient between X and Y
σ^2	population variance
σ_X	standard deviation of X
σ_X^2	variance of X
ϕ	hydraulic head
ϕ_L	hydraulic head at $x=L$
ϕ_{max}	maximum value of ϕ
ϕ_{min}	minimum value of ϕ
ϕ_o	hydraulic head at $x=0$
ψ	pressure head
$\Psi(x;t)$	pressure head random field
Ψ_{max}	maximum value of Ψ
Ψ_{min}	minimum value of Ψ
$Cov[\bullet]$	covariance function
$E[\bullet]$	expectation operator
$f(\bullet)$	stochastic process (chapter 3)
$g(\bullet)$	stochastic process (chapter 3)
$I[\bullet]$	indicator function
$L''[\bullet]$	second derivative of $L[\bullet]$
$L'[\bullet]$	first derivative of $L[\bullet]$
$P[\bullet]$	probability distribution function
$Var[\bullet]$	variance operator
∇	del operator

1. Introduction

1.1 Statistical Approach

Two broad methods of solving groundwater flow problems are the deterministic and stochastic approaches. In the deterministic approach, flow and soil-hydraulic parameters (e.g., hydraulic conductivity) are assumed to vary in a pre-defined manner. The procedure involves assigning input parameter values over the spatial domain with the data estimated from laboratory analyses of core samples and/or field tests (e.g., slug test, pump test). The groundwater problem is then solved for the dependent variable, namely, the pressure head. Sensitivity analyses may be performed to investigate the variability of model output for specific changes in input parameters, i.e., by observing head variability due to changes in soil-hydraulic parameters. The advantage of such an analysis is that the extreme variations in the dependent variable are assessed as a function of input variation. This analysis also identifies the criticality of input parameters. Unfortunately, this range of variation for the output variable can be quite wide to be of any practical use. Also, theoretically meaningful inferences about the dependent variable cannot be made without the needed statistical support for the input parameters; e.g., establishing confidence limits about the head

field. It is also not clear how one can assign a range of pointwise values for the input parameters, except to assign a gross estimate over a large region of the porous medium. Stochastic approaches attempt to overcome these difficulties.

It is economically and technically infeasible to provide the detailed input parameter distribution for a field problem [*Konikow and Bredehoeft*, 1992]. This then brings us to the question: "*is it possible to represent such spatially variable soil-hydraulic properties using some averaging mechanism that will provide a satisfactory representation of the heterogeneous system*?" This has been addressed by *de Marsily* [1984] and *Watt et al.* [1976] in the following way. What is available from field observations is a set of values (of soil-hydraulic properties) from which statistical moments can be estimated. Given this information, there exists a parent probability distribution function with parameters which are described by the statistical moments to characterize the variability of the input parameters over space. This implies that a variety of porous media configurations may be formulated to fit this statistical description. Therefore, when we assign a particular set of parameters as dictated by this underlying probabilistic structure, one particular realization of the porous medium is attained which is indeed the heart of the Monte Carlo approach. Under this framework, the heterogeneity of a porous medium is well addressed and a much needed theoretical support is created. *Dagan and Bresler* [1983] explicitly admit to the aforementioned approach while other authors [see *Gelhar*, 1986 for a review] have

implicitly assumed such a framework. A good review on different stochastic techniques is given by *El-Kadi* [1984] who groups these methods according to the mathematical approach adopted. Because of their popularity and relevance to the present study, attention is focussed on the following two stochastic approaches: (1) *Spectral-Perturbation* (SP) *analysis*, and (2) *Monte Carlo* (MC) *simulation.*

The basic objective of SP techniques is to obtain an analytically tractable solution for the groundwater problem in terms of the mean head variation. The spectral method of solving the groundwater flow equation uses the theory of perturbation analysis. The flow parameters are perturbed around their mean values. These zero mean fluctuations are assumed to be second order stationary, i.e., their variance is constant and the dependence between fluctuations at two spatial points is a function of the separation distance only and not on which specific locations are involved. It is further assumed that the mean values are slowly varying in space so that certain derivatives can be assumed small. The perturbed flow equations are then solved using the Fourier-Stieltjes representation to obtain a spectrum of head or flux in terms of the spectrum of the hydraulic conductivity field [*Lumley and Panofsky*, 1964]. Further manipulation yields the head variance in terms of the covariance of the input hydraulic conductivity. The theory relies upon Taylor series expansion in deriving analytical solutions. Higher order terms of fluctuations in input parameters are typically neglected rendering the method applicable to situations only when the

variances of input parameters are small. This need not be true in real field situations [*Yeh*, 1982]. The theoretical development also neglects the products of perturbations. Therefore, the method is valid only when the magnitude of perturbations are relatively small (the methodology is acceptable only when the absolute value of perturbations are much smaller than unity). In addition, the solution procedure using spectral density function requires the assumption of an infinite flow domain; all practical problems have bounded domains. This problem has been addressed to a certain extent by *Dikow* [1988] who used discrete spectral representation for solving flow in rectangular finite domains. But his method lacks from generalization and cannot be easily extended to different boundary shapes.

Another stochastic method of interest is the MC simulation technique. This method involves sequential generation of soil-hydraulic properties (input parameters) and solution of groundwater problems for each generation of input parameters. Usually the first two moments of the output variable, namely mean and variance of pressure heads, are estimated. However, the method is capable of yielding the entire probability distribution. Inferences are drawn based upon the estimated statistics. MC simulation is easy to use. The development of high speed digital computers have made the method very popular. The most widely used argument against MC simulation procedure is the sample size required to construct the probability distribution of the output variable.

1.2 Scope and Objectives of the study

As indicated by several researchers [see *Gelhar*, 1986; *Gómez-Hernández and Gorelick*, 1989 for review], methodologies to estimate the pressure head distribution within certain confidence limits are needed. Because specification of pointwise hydraulic conductivity values or other properties is not possible, it is being suggested that these methodologies focus on utilizing statistical moments of the input parameter values (such as mean conductivity and its variance) to estimate the mean pressure head distribution. If an expression for a field scale average conductivity, often called *effective conductivity*, can be found, then this effective conductivity can be used for the solution of the groundwater flow problem and the resulting pressure head distribution would be the expected head distribution if the groundwater flow equation had been treated as a stochastic partial differential equation (PDE).

As pointed out by *de Marsily* [1984] and *Freeze* [1975] usable effective conductivity expressions are not possible for complicated domain problems. MC approach is the recognized means for the solution of complicated problems. Also, to validate an effective conductivity approach, a performance comparison with the MC approach is needed. This thesis focuses on the following two major aspects:

(1) To derive a relationship for the field scale *effective hydraulic conductivity* that uses the statistical moments of the measured hydraulic conductivities over the domain. This effective hydraulic conductivity is designed to yield a mean head

distribution which would be the same as if the problem were solved using the random input parameters. The proposed effective conductivity involves various covariance terms. Upon simplifying these terms with a Taylor series approximation, the resulting equation can be cast in the form of Gardner's equation which is relatively simple to use. Suitability of the proposed effective conductivity expression is verified by using the relationship to solve various problems and comparing the results obtained with those obtained from MC simulation.

(2) A refined MC procedure is proposed. The method involves both the error criteria as well as the desired confidence coefficient to generate the required sample size. Also, a variance reducing estimator for the pressure head is proposed which exploits the correlation between the head and saturated hydraulic conductivity to achieve the reduction in variance.

1.3 Organization of the Thesis

Chapter 2 presents the boundary value problem describing flow through an unsaturated porous medium along with the necessary boundary and initial conditions. An effective medium approximation approach is presented in chapter 3. A relationship for computing the effective hydraulic conductivity, K_{eff}, based upon the statistics of the measured hydraulic conductivity values is developed. The relationship uses a perturbation analysis. Basic identities from stochastic processes are used to

derive an expression for effective hydraulic conductivity. This equation is then reduced to an equation similar to Gardner's equation with the aid of Taylor series expansion. This effective hydraulic conductivity function is easy to use and yields results close to that of the MC simulation.

Chapter 4 provides details on the MC simulation procedure. Previous work done in the area of groundwater flow modeling with MC simulation are delineated. Most of the chapter is devoted to developing the mathematical procedure for estimating the sample size with a pre-determined error and confidence limit criterion. This allows the modeler to determine the trade-off between the number of simulations and the error that will be encountered as well as the degree of confidence one can put on the results. The chapter also presents a description on the control variate technique for reducing the variance of the output variable.

Chapter 5 presents the development of an analytical estimator for the variance of the piezometric head in an unconfined aquifer. The sample size is estimated using this variance. The traditional method of using arithmetic, geometric, or harmonic means of hydraulic conductivity values to define effective medium is examined along with the proposed MC simulation procedure.

Chapter 6 contains the application of the methodology to five different problems. The first problem is the saturated flow problem solved by *Freeze* [1975]. The results show that the refined MC simulation procedure proposed here performs better in that

the same results are achieved with a smaller sample size. The second problem is a one-dimensional unsaturated flow problem. The results obtained by using the proposed effective hydraulic conductivity are quite close to those obtained from the MC simulation. The third problem is a two-dimensional, unsaturated flow problem. Comparison of the results obtained from MC simulation with that obtained from the proposed effective conductivity relationship shows good agreement. The fourth problem is a one-dimensional, transient, unsaturated flow. Both the proposed MC technique and the effective conductivity perform well. The last problem on well flow shows that the use of harmonic mean as the effective conductivity performs well in the case solved. Chapter 7 provides a general summary of the study. An appendix with basic definitions of the stochastic terms used in this thesis is provided at the end of the thesis.

2. Unsaturated Flow Theory

2.1 Governing Equation

The movement of water in the subsurface zone occurs through both saturated and unsaturated regions. Water from precipitation infiltrates the ground surface; moves downward, primarily under the influence of gravity; and accumulates, filling all the interstices of the porous medium above some impervious stratum. The rate at which water travels through a porous medium depends upon the potential difference between two points, the movement occurring from the point with higher potential to the point of lower potential.

Darcy's law provides the mathematical expression describing water movement through a porous medium. The law states that the flow rate through a porous medium is proportional to the potential gradient. Mathematically, the *specific discharge* or the superficial velocity, q, through a porous medium can be written as

$$q = - K \nabla\phi \tag{2.1}$$

where K is the hydraulic conductivity of the porous medium, ϕ is the hydraulic head, $\nabla\phi$ is the hydraulic gradient, $\nabla = (\partial/\partial x_1)\mathbf{i}+(\partial/\partial x_2)\mathbf{j}+(\partial/\partial x_3)\mathbf{k}$ is the *del* operator, and

x_i are the cartesian co-ordinate axes (i = *1, 2, 3*).

The constant of proportionality in Darcy's law, hydraulic conductivity (K), expresses the ease with which a fluid moves through a porous matrix. It depends on both the property of the porous medium and the property of the fluid flowing through it. The relevant porous medium properties are pore size distribution and porosity while the fluid properties of major concern are density and viscosity. This dependence of hydraulic conductivity can be expressed as [*Freeze and Cherry*, 1979]

$$K = \frac{k \rho g}{\eta} \tag{2.2}$$

in which k is the intrinsic permeability of the porous medium, ρ is the density of flowing fluid, g is the acceleration due to gravity, and η is the dynamic viscosity of flowing fluid.

The hydraulic head, ϕ, has two main components; the pressure head, ψ, and the gravitational head, x_3. ψ is positive in saturated soils and is negative in unsaturated soils. The relationship between ψ and ϕ can be expressed as

$$\phi = \psi + x_3 \tag{2.3}$$

in which x_3 is taken to be the vertical direction, positive upwards. Considering both the components of ϕ, Darcy's law takes the form, for flow in the vertical direction

$$q_3 = - K \left(\frac{\partial \psi}{\partial x_3} + 1 \right) . \tag{2.4}$$

In x_1 and x_2 directions, Darcy's law takes the form

$$q_i = - K \left(\frac{\partial \psi}{\partial x_i} \right) \qquad i = 1,2 . \tag{2.5}$$

In unsaturated soils, K depends on the capillary pressure head. Therefore, Darcy's law, in general, can be expressed as

$$q_i = - K_{ii}(\psi) \frac{\partial \phi}{\partial x_i} . \tag{2.6}$$

In writing Eq. (2.6), it has been assumed that x_1, x_2, and x_3 axes coincide with the principal axes. Combining the general Darcy law [Eq. (2.6)] with the continuity equation

$$\frac{\partial \theta}{\partial t} = \nabla \cdot q_i \tag{2.7}$$

and using the identity

$$C(\psi) = - \frac{\partial \theta}{\partial \psi} \tag{2.8}$$

we get Richards' equation describing flow through unsaturated porous media

$$C(\psi)\frac{\partial \psi}{\partial t} = \frac{\partial}{\partial x_i}\left[K_{ii}(\psi)\frac{\partial(\psi + x_3)}{\partial x_i}\right] . \qquad (2.9)$$

In Eqs. (2.7) and (2.8), θ is the moisture content which is a function of capillary tension. $C(\psi)$ in Eqs. (2.8) and (2.9) is called the soil water capacity. Eq. (2.9) describes the unsaturated flow through a non-deforming porous medium (constant porosity). The non-linear function of K-ψ has been approximated by various researchers with empirical relations [e.g., *Brooks and Corey*, 1964; *Gardner*, 1958; *van Genuchten*, 1980]. In the present study, a relatively simple relationship, Gardner's equation, is used. This aids in obtaining an analytically tractable solution. Gardner's equation is expressed as

$$K(\psi) = K_s\, e^{-\alpha\psi} \qquad (2.10)$$

where α is a non-negative soil parameter and K_s is the saturated hydraulic conductivity of the porous medium. Information on K_s and α for various soil types can be obtained from published literature.

2.2 Boundary Conditions

Boundary conditions specify the dependent variable or the derivative of the dependent variable at the boundaries of the problem domain [*Anderson and Woessner*, 1992]. These conditions are largely responsible for determination of the flow pattern.

Dirichlet, Neumann, and Mixed type boundary conditions are the three types of boundary conditions that normally occur.

In the Dirichlet boundary condition, the dependent function values are prescribed at all points on the system boundary. This occurs when there is ponded water on the soil surface. For the system boundary Γ, a Dirichlet-type boundary condition can be expressed as

$$\psi(x_i;t) = \psi_0(t) \qquad \text{for} \qquad x_i \in \Gamma \qquad (2.11)$$

where ψ_0 is a known function of time for all t over Γ.

When the flux rate is known at the boundary of the flow domain, Neumann type boundary condition exists. It is also possible to have cases when the boundary condition changes from one type to another. For example, the ground surface may be dry at the beginning of a storm event. Therefore, a Neumann-type boundary condition prevails with flux equal to the infiltration rate. As time progresses and the hydraulic conductivity of the porous medium reaches its saturated value, water will start ponding at the surface and Dirichlet-type boundary condition takes over. After cessation of the storm, as time progresses, a Neumann-type boundary condition once again will prevail. Such boundary conditions that change with time can be described by a mixture of dependent functions and their derivatives. These are known as mixed boundary conditions. These may be at boundaries where conditions are not known a priori.

2.3 Initial Conditions

Initial conditions provide the spatial distribution of head everywhere in the flow system at the beginning of the simulation. In addition to the boundary conditions discussed in the previous section, initial conditions also need to be specified to formulate a well defined boundary value problem. Initial conditions can be expressed as

$$\psi(x_i;0) = \psi_0(x_i) \tag{2.12}$$

where $\psi_0(x_i)$ denote the initial pressure head at all points in the flow domain.

Eqs. (2.9) through (2.12) describe the boundary value problem describing the flow through an unsaturated porous medium. The boundary value problem can now be solved to obtain $\psi(x_i;t)$ for all x_i and t. However, the non-linear partial differential equation (PDE) cannot be solved analytically. Therefore, numerical techniques must be used to solve the boundary value problem. These techniques transform the PDE of the groundwater flow system into systems of ordinary differential or algebraic equations. The solution of these equations determines the values of the dependent variables within the flow domain. Various numerical schemes are available for solving unsaturated subsurface flow problems. In the present work 3DFEMWATER [*Yeh*, 1987; 1990] is used.

3. Effective Medium Approximation

3.1 Introduction

The probabilistic treatment of groundwater flow equation involves the use of parameters which are considered to be random input. Because the pressure head depends on these random parameters, it is also random. *Mantoglou et al.* [1987a, b, c], *Yeh et al.* [1985a, b, c], and *Zhu* [1990] use perturbation approximations to express the random variables (both dependent and independent variables) in terms of their means and perturbations. The zero mean perturbations are assumed to be second-order stationary.

The effective hydraulic conductivity relationship proposed by *Mantoglou et al.* [1987a, b, c] and *Yeh et al.* [1985a, b, c] require a priori knowledge of the covariance relationship between the input parameters and the output variable. This is needed to estimate the covariance of the fluctuations. The methodology proposed in this thesis requires information only on the input parameters.

3.2 Theory

The general flow equation through an unsaturated porous medium is given by Richards' equation

$$-\frac{\partial\theta}{\partial t} = C(\psi)\frac{\partial\psi}{\partial t} = \frac{\partial}{\partial x_i}\left[K(\psi)\frac{\partial(\psi+x_3)}{\partial x_i}\right] \tag{3.1}$$

where $K(\psi)$ is the unsaturated hydraulic conductivity, ψ is the capillary tension head, θ is the moisture content, and $C(\psi)=-\partial\theta/\partial\psi$ is the specific moisture capacity, and t represents time. For brevity, here onwards $K(\psi)$ will be written as K. Taking expectations on both sides of Eq. (3.1), we get

$$-\frac{\partial\{E[\theta]\}}{\partial t} = \frac{\partial}{\partial x_i}\left\{E\left[K\,\frac{\partial(\psi+x_3)}{\partial x_i}\right]\right\}. \tag{3.2}$$

Let us define effective soil water capacity, C_{eff}, and effective hydraulic conductivity, K_{eff}, as

$$C_{eff} = -\frac{\partial E[\theta]}{\partial H} \tag{3.3}$$

and $$K_{eff_i} = \frac{E\left[-K\dfrac{\partial\phi}{\partial x_i}\right]}{E[J_j]} = \frac{E[q_i]}{\bar{J}_j}, \qquad i, j = 1,2,3,... \tag{3.4}$$

in which $J_j=-\partial\phi/\partial x_j$ and q_i is the Darcy flow. The field scale flow equation, from Eq. (3.2), can therefore be written as

$$C_{eff}\frac{\partial H}{\partial t} = \frac{\partial}{\partial x_i}\left[K_{effij}\frac{\partial(H+x_3)}{\partial x_j}\right] \tag{3.5}$$

where H is the mean of the pressure head. C_{eff} and K_{effij} are effective soil water capacity and effective hydraulic conductivity, respectively. In the above equation Einstein summation convention is used.

Unsaturated hydraulic conductivity, K, in Eq. (3.1) depends upon the pressure head, ψ. In the present study, the Gardner equation [*Gardner*, 1958]

$$K = K_s\, e^{-\alpha\psi} \tag{3.6}$$

is used. In Eq. (3.6), $\alpha>0$ is the soil parameter and $\psi>0$ is the capillary tension. It is customary to express the input parameters and output variables as the sum of their means and perturbations [*Mantoglou et al.*, 1987a; *Mishra et al.*, 1990; *Zhu*, 1990; *Russo*, 1992]. Therefore, we can write

$$\ln K_s = F+f \tag{3.7}$$

$$\ln K = G+g \tag{3.8}$$

$$\psi = H + h \tag{3.9}$$

such that

$$E[\ln K_s] = F, \qquad E[f] = 0 \tag{3.10}$$

$$E[\ln K] = G, \qquad E[g] = 0 \tag{3.11}$$

$$E[\psi] = H, \qquad E[h] = 0 \tag{3.12}$$

in which F, G, and H are the mean functions. The fluctuations f, g, and h have zero mean. Let us now define K_{effxi} as the effective hydraulic conductivity in the x_i-direction. Defining hydraulic gradient in the x_i-direction as J_{xi} and taking expectations, we get

$$\bar{J}_{x_i} = E\left[J_{x_i}\right] = E\left[-\frac{\partial \phi}{\partial x_i}\right], \qquad i=1,2,3,... \tag{3.13}$$

where $\bar{J}_{xi}$ is the mean hydraulic gradient in the x_i direction.

From the definition of *effective hydraulic conductivity*, the ratio of average flux to the average hydraulic gradient, given in Eq. (3.4), it is seen that evaluation of K_{effxi} necessitates evaluation of expected terms in the numerator and the denominator. First the denominator in Eq. (3.4) is evaluated.

From Eq. (3.6), we can write,

$$\ln K = \ln K_s - \alpha\psi \ . \tag{3.14}$$

Using Eqs. (3.8) and (3.9), we get

$$G+g = F+f-\alpha\psi \tag{3.15}$$

which upon differentiation with respect to x_i and upon rearrangement gives,

$$\frac{\partial\psi}{\partial x_i} = \frac{1}{\alpha}\left\{\left(\frac{\partial F}{\partial x_i} - \frac{\partial G}{\partial x_i}\right)+\left(\frac{\partial f}{\partial x_i} - \frac{\partial g}{\partial x_i}\right)\right\} \tag{3.16}$$

Noting that $\partial F/\partial x_i$ and $\partial G/\partial x_i$ are non-random (mean functions of space and time which are evaluated over the ensemble at each point in space and time), we can write from Eq. (3.16) and using the identity $\phi = -\psi - x_3$,

$$E\left[-\frac{\partial\phi}{\partial x_i}\right] = \frac{1}{\alpha}\left\{\left(\frac{\partial F}{\partial x_i} - \frac{\partial G}{\partial x_i}\right)\right\} + E\left[\frac{\partial x_3}{\partial x_i}\right]. \tag{3.17}$$

In obtaining Eq. (3.17), the following derivative properties are used:

$$E\left[\frac{\partial f}{\partial x_i}\right] = \frac{\partial E[f]}{\partial x_i} = 0 \tag{3.18}$$

and

$$E\left[\frac{\partial g}{\partial x_i}\right] = \frac{\partial E[g]}{\partial x_i} = 0 \ . \tag{3.19}$$

Eq. (3.17) facilitates evaluation of the denominator of Eq. (3.4) in terms of the mean saturated and the unsaturated hydraulic conductivity gradients and the potential

gradient. The next step is to evaluate the numerator of Eq. (3.4). For this purpose, from Eq. (3.8) and the identity $\phi = -\psi - x_3$, we can write

$$-K\frac{\partial \phi}{\partial x_i} = e^{G+g}\left(\frac{\partial \psi}{\partial x_i} + \frac{\partial x_3}{\partial x_i}\right) . \tag{3.20}$$

Taking expectations, we get

$$E\left[-K\frac{\partial \phi}{\partial x_i}\right] = E\left[e^{G+g}\left(\frac{\partial \psi}{\partial x_i} + \frac{\partial x_3}{\partial x_i}\right)\right] . \tag{3.21}$$

Using Eqs. (3.8) and (3.16) in Eq. (3.21) and re-arranging, we get

$$E\left[-K\frac{\partial \phi}{\partial x_i}\right] = e^{G}\left\{E\left[e^{g}\left\{\frac{1}{\alpha}\left[\left(\frac{\partial F}{\partial x_i} - \frac{\partial G}{\partial x_i} + \alpha\frac{\partial x_3}{\partial x_i}\right) + \left(\frac{\partial f}{\partial x_i} - \frac{\partial g}{\partial x_i}\right)\right]\right\}\right]\right\} . \tag{3.22}$$

Eq. (3.22) requires evaluation of the following two expectations,

$$E\left[e^{g}\frac{\partial f}{\partial x_i}\right] \qquad \text{and} \qquad E\left[e^{g}\frac{\partial g}{\partial x_i}\right] . \tag{3.23}$$

These expectations are evaluated in the following theorem.

Theorem 3.1

Let the variogram $\gamma_f(\xi) = \frac{1}{2} E\,[f(x+\xi) - f(x)]^2$ of f be a function $O_f(\xi^{\lambda})$ $1/\{2E[e^{2g}]\}$ for $\lambda > 1$. Then

$$E\left[e^{g} \frac{\partial f}{\partial x} \right] = 0 \ . \tag{3.24}$$

Proof: From the definition of stochastic differentiation

$$e^{g}\frac{\partial f}{\partial x} = e^{g} \lim_{\xi \to 0} \frac{f(x+\xi)-f(x)}{\xi} = \lim_{\xi \to 0} \frac{e^{g}f(x+\xi)-e^{g}f(x)}{\xi} \ , \tag{3.25}$$

and
$$E\left[e^{g} \frac{\partial f}{\partial x}\right] = \lim_{\xi \to 0} \frac{E\left[e^{g} f(x+\xi) - e^{g} f(x)\right]}{\xi} \ . \tag{3.26}$$

By the Schwarz inequality, we have

$$0 \le E\left[e^{g}(f(\xi)-f)\right]^2 \le E\left[e^{2g}\right] E\left[f(\xi)-f\right]^2 = O_f\left(\xi^{\lambda}\right) = 0 \quad \text{as} \quad \xi \to 0 \ , \tag{3.27}$$

which when used in Eq. (3.26) yields the required result.

It is noted that in general the covariance function $E[f(\xi)\, f]$ is continuous, and it is sufficient to show that $E[f(\xi)-f]^2 \to 0$ as $\xi \to 0$. The requirement on the variogram in Theorem 3.1 is needed to show that it goes to zero at a faster rate than ξ. *de Marsily* [1986] has recommended such a variogram. By imposing the same kind of restriction on the variogram of g, one can conclude $E\left[e^{g} \frac{\partial g}{\partial x} \right] = 0$. From this

result and Theorem 3.1, Eq. (3.22) is reduced to

$$E\left[-K\frac{\partial \phi}{\partial x_i}\right] = e^{G}\left\{\frac{1}{\alpha}\left[\left(\frac{\partial F}{\partial x_i} - \frac{\partial G}{\partial x_i} + \alpha\frac{\partial x_3}{\partial x_i}\right)\right] E[e^{g}]\right\}. \tag{3.28}$$

Eq. (3.28) is the numerator of Eq. (3.4). Using Eqs. (3.17) and (3.28) in the definition of K_{effxi}, Eq. (3.4), and re-arranging we get

$$K_{eff_{x_i}} = e^{G}\left\{E\left[e^{g}\right]\right\}. \tag{3.29}$$

Eq. (3.29) is the relationship for K_{effxi} expressed solely in terms of soil hydraulic parameters. It is important to note some of the major differences between Eq. (3.29) and expressions developed for effective hydraulic conductivity using Spectral Perturbation (SP) methods. While Eq. (3.29) expresses the effective conductivity solely in terms of soil hydraulic properties, the SP method requires an a priori knowledge of head perturbations. Furthermore, development of SP theory neglects third and higher order terms in Taylor series expansion. In the above development, no such approximations have been made. This has an advantage as it allows the proposed method to be used even when variability in input parameters are large; as opposed to SP method which is limited to small variances [*Yeh et al.*, 1985a]. For evaluating the expectation term within the braces in Eq. (3.29), the following strategy is adopted which reduces Eq. (3.29) to an expression similar to the Gardner equation.

3.3 Field Scale Gardner's Equation

Expanding the exponential function, e^g, using Taylor series, we get

$$e^g \approx 1 + g + \frac{g^2}{2} + O(g) \ . \tag{3.30}$$

This approximation is valid for small g, the log unsaturated hydraulic conductivity fluctuations. Taking expectations on both sides of Eq. [3.30] and neglecting third and higher order terms of g, we get

$$E\left[e^g \right] \approx 1 + E\left[g \right] + E\left[\frac{g^2}{2} \right] \tag{3.31}$$

which, when Eq. [3.11] is used, reduces to

$$E\left[e^g \right] \approx 1 + \frac{\sigma_g^2}{2} \ . \tag{3.32}$$

As recommended by *Zhu* [1990], the variance of log-unsaturated hydraulic conductivity is taken to be the same as the variance of log saturated hydraulic conductivity. Therefore, with use of Eq. (3.32), Eq. (3.29) may be rewritten as

$$K_{eff_{x_i}} \approx e^G \left(1 + \frac{\sigma_f^2}{2} \right) = e^F \left(1 + \frac{\sigma_f^2}{2} \right) e^{-\alpha H} \ . \tag{3.33}$$

Defining

$$K_s^* = e^F \left(1 + \frac{\sigma_f^2}{2}\right) \tag{3.34}$$

Eq. (3.33) can be re-written as

$$K_{eff_{x_i}} = K_s^* \, e^{-\alpha H} \tag{3.35}$$

Eq. (3.35) can be used to obtain an estimate of $K_{eff_{x_i}}$ which is then used in the field scale mean flow equation [Eq. (3.5)] to compute the mean head field. It is noted that Eq. (3.35) is of the same form as the local Gardner equation; the saturated hydraulic conductivity being replaced by its effective counterpart and the capillary tension replaced by mean pressure head. Because Eqs. (3.5) and (3.35) are of the same form as Richards' equation and Gardner's equation, the schemes used to solve the boundary value problem (with Richards' equation, Gardner's equation, initial and boundary conditions) can be used to solve the mean flow equation. In circumstances, when the above assumptions can not be justified, one may resort to the more powerful and less restrictive MC simulation procedure, which is presented in the next chapter.

4. Monte Carlo Simulation

4.1 Introduction

Monte Carlo (MC) methods comprise the branch of mathematics concerned with the use of random numbers. The procedure consists of generation of random numbers from a known probability distribution. The generated random numbers are used as inputs to governing equations, and the corresponding outputs are computed. This process is repeated a specified number of times to obtain a set of system outputs which are then analyzed to define the output statistics. Because the input numbers are generated from a known distribution, a priori knowledge of the distribution of these numbers is required. In groundwater flow analyses, the primary input is the saturated hydraulic conductivity and the hydraulic head is the output of main interest. Therefore, use of MC simulation procedure in groundwater flow analyses require an advance knowledge of the probability density function (*pdf*) of saturated hydraulic conductivity. In the present study, saturated hydraulic conductivity is assumed to be log-normally distributed. *Law* [1944] was the first researcher to propose log-normal distribution for saturated hydraulic conductivity, based upon core analysis from an oil field reservoir. This distribution is also implied by the observation that saturated

hydraulic conductivity can be estimated as an exponential function of porosity, which is generally taken to be normally distributed [*Bennion and Griffiths*, 1968]. The assumption of log-normally distributed saturated hydraulic conductivity can also be supported based upon small particle statistics [*Aitchison and Brown*, 1957]. Other researchers whose field studies confirm log-normal distribution of saturated hydraulic conductivity include *Nielsen et al.* [1973] and *Ünlü et al.* [1990]. *Carsel and Parrish* [1988] analyzed each of the 12 types of soils, based on the Soil Conservation Service classifications, to develop joint probability distributions of various soil water characteristics. They worked with an extensive database of cores. Various distributions were fitted for different soil-hydraulic parameters. One of their major conclusions is that the log-normal distribution fitted saturated hydraulic conductivity. *Wilson et al.* [1989] analyzed two forested watersheds and found saturated hydraulic conductivity to be log-normally distributed.

The set of soil-hydraulic parameter values such as hydraulic conductivity from field observations provides estimates of statistical moments of the variables which in turn define the underlying probability distribution. Using such a probabilistic description, one may postulate various porous media configurations as possible realizations of the underlying probability distribution. Therefore, when we assign a particular set of hydraulic conductivity values as dictated by this underlying probabilistic structure, one particular realization of the porous medium is attained. By

the same token, a second set of input parameters attains a different realization from the parent probability distribution. It follows that the sample moments from this ensemble of configurations would faithfully resemble the parent moments and therefore the output statistics can also be generated. The major advantages of MC simulation are:

- Practical adaptation of the algorithms for computer programming is simple.
- No additional difficulties arise in solving problems with discontinuous boundary functions and non-smooth boundaries.
- Problems with random coefficients, random boundary values, and other stochastic parameters can be solved without difficulty.

Warren and Price [1961] were among the first researchers to make use of MC procedure in groundwater flow analysis. They used a numerical scheme to investigate the effect of heterogeneous hydraulic conductivity on steady and transient flows. They assumed that the hydraulic conductivity was log-normally distributed. No objective scheme was given for computing the sample size.

Freeze [1975] carried out a detailed MC analysis of a saturated flow problem. He stated that:

- there is no simple way to define an equivalent effective porous medium in non-uniform porous formations when complicated boundary conditions are involved;
- the consideration of spatial variability leads to very large fluctuations in

hydraulic heads. Therefore, the values of hydraulic head computed with deterministic models may suffer from very large errors.

The above conclusions were the main reason for Freeze's questioning of traditional deterministic approaches. This criticism about a single deterministic run has been supported by various other researchers as well [*Smith and Freeze*, 1979a, b; *Ma et al.*, 1987; *Dagan*, 1989]. *Gómez-Hernández and Gorelick* [1989] provide a review, but they also do not offer a procedure for sample size estimate. However, it is implied in the traditional MC approach to groundwater flow problems that the sample size be based upon the reproduction of the input saturated hydraulic conductivity moments by the generated conductivity values. This method does not relate the sample size to the pressure head values. Because the concern is towards estimating the mean pressure head, naturally the question arises, "*how close is the sample average of the output pressure head to its population mean*?" Towards this end the following formulation is presented.

4.2 Estimation of Sample Size

Let $\Psi=\Psi(x;t)$ be the random field under consideration. For a given x and a fixed t, Ψ is a random variable. Let $\Psi_1,\Psi_2,...,\Psi_n$ be the realized values from the n number of MC simulations. We are interested in the absolute error {|sample mean - population mean|} being very small. The sample mean is given by

$$\bar{\Psi} = \frac{1}{n} \sum_{i=1}^{n} \Psi_i \tag{4.1}$$

which depends on the sample size n and is a random variable if the Ψ_i values are not yet observed. Let $E[\Psi]=E[\Psi_i]=\mu_\Psi$. We would like to have the probability of committing the error $\{|\bar{\Psi} - \mu_\Psi| > \varepsilon\}$, for small $\varepsilon > 0$, to be very small. That is,

$$P\left[|\bar{\Psi}-\mu_\Psi| > \varepsilon\right] \leq \delta \tag{4.2}$$

where $(1-\delta)$ is the level of confidence and ε is the error. In an attempt to compute the probability in Eq. (4.2) we use the following lemma 4.1 and theorem 4.1 due to *Hoeffding* [1963].

Lemma 4.1

Let Ψ be a random variable such that $\Psi_{min} \leq \Psi \leq \Psi_{max}$, then for any arbitrary $h>0$, the function $e^{h\Psi}$ is convex (second derivative is positive). From the property of the convex function that the chord lies above the curve, it follows that

$$e^{h\Psi} \leq \left\{ \left(\frac{\psi_{max}-\Psi}{\psi_{max}-\psi_{min}} \right) e^{h\psi_{min}} + \left(\frac{\Psi-\psi_{min}}{\psi_{max}-\psi_{min}} \right) e^{h\psi_{max}} \right\} \tag{4.3}$$

which upon taking expectation gives

$$E\left[e^{h\Psi}\right] \leq \left\{\left(\frac{\psi_{max}-E\left[\Psi\right]}{\psi_{max}-\psi_{min}}\right) e^{h\psi_{min}} + \left(\frac{E\left[\Psi\right]-\psi_{min}}{\psi_{max}-\psi_{min}}\right) e^{h\psi_{max}}\right\} \quad . \tag{4.4}$$

We will make use of this identity to prove theorem 4.1, which is presented next.

Theorem 4.1

Let $\Psi_1,\Psi_2,...,\Psi_n$ be independent and identically distributed random variables and $\Psi_{mini} \leq \Psi_i \leq \Psi_{maxi}$ $(i=1,2,...,n)$, then for any $\varepsilon > 0$,

$$P\left[\bar{\Psi}-\mu_{\Psi} \geq \varepsilon\right] \leq e^{-\frac{2n^2\varepsilon^2}{\sum_{i=1}^{n}(\Psi_{max_i}-\Psi_{min_i})^2}}, \tag{4.5}$$

which reduces to

$$P\left[\bar{\Psi}-\mu_{\Psi} \geq \varepsilon\right] \leq e^{-\frac{2n\varepsilon^2}{(\Psi_{max}-\Psi_{min})^2}} . \tag{4.6}$$

when the limits of Ψ_{mini} and Ψ_{maxi} are the same for all i.

Proof

Let $S = \Psi_1+\Psi_2+...+\Psi_n$ be the sum of random variables $\Psi_1,\Psi_2,...,\Psi_n$. Then,

$$\bar{\Psi} = \frac{S}{n} \tag{4.7}$$

and therefore

$$E\left[\,\bar{\Psi}\,\right] = \mu_{\Psi} = \frac{E\left[S\right]}{n} \tag{4.8}$$

and variance of Ψ is given by

$$Var\left[\,\bar{\Psi}\,\right] = \frac{Var\left[\Psi\right]}{n}\,. \tag{4.9}$$

Let the indicator function

$$I\left[S-E[S]-n\varepsilon\right] = 1 \qquad \text{if} \qquad S-E[S]-n\varepsilon > 0 \tag{4.10}$$

$$\text{and} \qquad I\left[S-E[S]-n\varepsilon\right] = 0 \qquad \text{otherwise}\,. \tag{4.11}$$

Then it follows

$$E\left[I\right] = \int_{-\infty}^{\infty} I\, f_I\, dI = \int_{0}^{\infty} 1\, f_I\, dI = P\left[\,S-E[S]-n\varepsilon > 0\right] \tag{4.12}$$

$$\text{or} \qquad E\left[I\right] = P\left[\,S-E[S] > n\varepsilon\,\right] \tag{4.13}$$

which is always less than or equal to 1. For an $h>0$, we obtain

$$0 < e^{h\left[S-E[S]-n\varepsilon\right]} < 1 \qquad \text{for} \qquad S-E[S]-n\varepsilon < 0 \tag{4.14}$$

$$\text{and} \qquad e^{h\left[S-E[S]-n\varepsilon\right]} > 1 \qquad \text{for} \qquad S-E[S]-n\varepsilon > 0\,. \tag{4.15}$$

Therefore, it is concluded that

$$e^{h[S-E[S]-n\varepsilon]} \geq P\left[S-E[S] \geq n\varepsilon\right] = P\left[\bar{\Psi}-\mu_{\Psi} \geq \varepsilon\right]. \tag{4.16}$$

Because Ψ_i are independent, we obtain

$$E\left[e^{h\,(S-E[S]-n\varepsilon)}\right] = e^{-hn\varepsilon} \prod_{i=1}^{n} E\left[\, e^{h(\Psi_i)-E[\Psi_i]}\right]. \tag{4.17}$$

In order to evaluate Eq. (4.16), we need to evaluate the terms on the RHS of the Eq. (4.17). To achieve this, we make use of Lemma 4.1 to write

$$E\left[e^{h(\Psi_i-\mu_i)}\right] \leq e^{-h\mu_i}\left\{\left(\frac{\psi_{\max_i}-\mu_i}{\psi_{\max_i}-\psi_{\min_i}}\right)e^{h\psi_{\min_i}} + \left(\frac{\mu_i-\psi_{\min_i}}{\psi_{\max_i}-\psi_{\min_i}}\right)e^{h\psi_{\max_i}}\right\} = e^{L(h_i)} \tag{4.18}$$

in which

$$L(h_i) = -h_i\, p_i + \ln\left(1 - p_i + p_i\, e^{h_i}\right) \tag{4.19}$$

with

$$h_i = h\left(\psi_{\max_i} - \psi_{\min_i}\right), \qquad p_i = \left[\frac{\mu_i-\psi_{\min_i}}{\psi_{\max_i}-\psi_{\min_i}}\right]. \tag{4.20}$$

Taking the first and the second derivatives of $L(h_i)$ with respect to h_i, we can write

$$L'(h_i) = -p_i + \left\{\frac{p_i}{p_i + (1-p_i)\, e^{-h_i}}\right\} \tag{4.21}$$

and

$$L''(h_i) = \left\{ \frac{p_i}{\left[p_i + (1-p_i)\, e^{-h_i}\right]} \right\} \cdot \left\{ \frac{(1-p_i)\, e^{-h_i}}{\left[p_i + (1-p_i)\, e^{-h_i}\right]} \right\} . \tag{4.22}$$

By considering the first term to be u, we have the second term $(1-u)$. Therefore,

$$L''(h_i) = u\,(1-u) . \tag{4.23}$$

For $0 < u < 1$, we must have the maximum of L"(hi)=½·½=¼, and therefore we can write

$$L''(h_i) \le \frac{1}{4} \tag{4.24}$$

The Taylor's formula with a remainder is given by

$$f(x) = f(a)+(x-a)f'(a)+\frac{(x-a)^2}{2!}f''(a)+\ldots+\frac{(x-a)^n}{n!}f^{(n)}(a)+R_n(x) \tag{4.25}$$

where
$$R_n(x) = \frac{(x-a)^{(n+1)}}{(n+1)!}\, f^{(n+1)}(c) \tag{4.26}$$

for some c between a and x. Also, if there are positive constants M and r such that

$$\left| f^{(n+1)}(t) \right| \le M\, r^{(n+1)} \tag{4.27}$$

for all t between a and x, inclusive. Then,

$$|R_n(x)| \leq \frac{M\, r^{n+1}\, |x-a|^{n+1}}{(n+1)!} \tag{4.28}$$

Therefore, we can write

$$L(h_i) = L(0) + h_i\, L'(0) + R_1(h_i) \tag{4.29}$$

and because

$$|L''(h_i)| \leq \frac{1}{4} = M \qquad \text{with} \qquad r = 1 \tag{4.30}$$

we have

$$R_1(h_i) \leq \frac{1}{4} \cdot \frac{h_i^2}{2!} \,. \tag{4.31}$$

Finally we obtain

$$L(h_i) \leq L(0) + h_i\, L'(0) + \frac{1}{8}\, h_i^2 = \frac{1}{8}\, h_i^2 \,. \tag{4.32}$$

Evaluating the function $L(h_i)$ and its derivatives about h_i=0, we also get

$$L(0) = L'(0) = 0 \,. \tag{4.33}$$

Now, using the Taylor series expansion about the origin, we can write $L(h_i)$ as

$$L(h_i) = L(0) + L'(0)\, h_i + L''(0)\, \frac{h_i^2}{2} + \ldots \,. \tag{4.34}$$

Substitution of $L(0)$, $L'(0)$, and $L''(0)$ in the above equation results in

$$L(h_i) \leq \frac{h_i^2}{8} = \frac{h^2 \left(\psi_{\max_i} - \psi_{\min_i}\right)^2}{8} . \tag{4.35}$$

Substituting this in Eq. (4.18), we get

$$E\left[e^{h(\Psi_i - \mu_i)}\right] \leq e^{\frac{h^2 \left(\psi_{\max_i} - \psi_{\min_i}\right)^2}{8}} \tag{4.36}$$

which when used in Eqs. (4.16) and (4.17) gives

$$P\left[\bar{\Psi} - \mu_\psi \geq \varepsilon\right] \leq e^{-h n \varepsilon + \frac{h^2 \sum_{i=1}^{n} \left(\psi_{\max_i} - \psi_{\min_i}\right)^2}{8}} . \tag{4.37}$$

The minimum value of the exponential function in the above equation can be obtained by differentiating the function with respect to h, equating the resulting equation to zero, and solving for h. The corresponding value of h for which the function is the minimum is thus obtained as

$$h = \frac{4n\varepsilon}{\sum_{i=1}^{n} \left(\psi_{\max_i} - \psi_{\min_i}\right)^2} \tag{4.38}$$

which when used in Eq. (4.37) gives

$$P\left[\bar{\Psi} - \mu_\psi \geq \varepsilon\right] \leq e^{-\frac{2n^2\varepsilon^2}{\sum_{i=1}^{n} (\psi_{\max_i} - \psi_{\min_i})^2}} \tag{4.39}$$

When the upper and the lower bounds, ψ_{maxi} and ψ_{mini}, are the same for all i's, we can

rewrite the above equation as

$$P\left[\bar{\psi}-\mu_{\psi}\geq\varepsilon\right] \leq e^{-\frac{2n\varepsilon^2}{(\psi_{max}-\psi_{min})^2}}$$

which proves theorem 4.1.

With the aid of theorem 4.1, we can estimate the sample size required to satisfy Eq. (4.2). This is accomplished in theorem 4.2.

Theorem 4.2

Consider the probability $P\left[|\bar{\Psi} - \mu_{\psi}| > \varepsilon\right] \leq \delta$.

The sample size N necessary to achieve the absolute error no larger than $\varepsilon > 0$ with a very small probability of excedence δ is

$$N \geq -\ln\left(\frac{\delta}{2}\right)\frac{\left[(\Psi_{max}-\Psi_{min})/\varepsilon\right]^2}{2} . \tag{4.40}$$

Proof

The sample size required is given by,

$$N = \min\left\{ n : P\left[|\bar{\Psi}(n)-\mu_{\psi}| > \varepsilon\right] \leq \delta\right\} \tag{4.41}$$

where $\varepsilon > 0$ is an absolute error, $(1-\delta)$ is the minimum confidence level, and $\bar{\psi}$ is the average given by Eq. (4.1).

By using the definition of absolute value, we obtain

$$P\left[|\bar{\psi}-\mu_{\psi}|\geq\varepsilon\right] = P\left[\bar{\psi}-\mu_{\psi}\geq\varepsilon\right]+P\left[-\bar{\psi}+\mu_{\psi}\geq\varepsilon\right] \tag{4.42}$$

which gives

$$P\left[|\bar{\psi}-\mu_{\psi}|\geq\varepsilon\right] = P\left[\bar{\psi}-\mu_{\psi}\geq\varepsilon\right]+P\left[\bar{\psi}-\mu_{\psi}\leq-\varepsilon\right] . \tag{4.43}$$

By using the theorem 4.1, we have

$$P\left[\ |\bar{\psi}-\mu_{\psi}|\ \geq\ \varepsilon\right] \leq 2e^{-\frac{2n\varepsilon^2}{(\psi_{max}-\psi_{min})^2}} \tag{4.44}$$

in which the right and the left tail probabilities have been taken to be the same. For minimum confidence level of (1-δ),

$$2e^{-\frac{2n\varepsilon^2}{(\psi_{max}-\psi_{min})^2}} \leq \delta \tag{4.45}$$

which upon manipulation and when used in Eq. (4.41) proves identity (4.40).

It is observed from Eq. (4.40) that the number of replications for the MC simulation requires a priori knowledge of the range of Ψ in addition to the desired error criteria and the level of confidence. It is also of interest to compare the performance of Eq. (4.40) with the Chebyshev inequality. The Chebyshev inequality [*DeGroot*, 1975] states that for any number ε > 0,

$$P\left[|\bar{\Psi}\ -\ \mu_{\psi}|\ >\ \varepsilon\ \right] \leq \frac{\sigma^2}{N_C\ \varepsilon^2} \tag{4.46}$$

where σ^2 is the variance of Ψ. Therefore, in order to satisfy the probability in Eq.

(4.41), the sample size required is given by

$$N_C \geq \frac{\sigma^2}{\delta \varepsilon^2} . \tag{4.47}$$

Crow et al. [1960] suggest that the range may be written as a constant multiple of the standard deviation, i.e., $\Psi_{max} - \Psi_{min} = p\sigma$. Therefore, Eq. (4.40) may be written as

$$N \geq - \ln \left(\frac{\delta}{2}\right) \frac{p^2 \sigma^2}{2 \varepsilon^2} \tag{4.48}$$

Comparing Eqs. (4.47) and (4.48), for N to be smaller than N_C, we need

$$- \ln \left(\frac{\delta}{2}\right) \left(\frac{p^2}{2}\right) \leq \frac{1}{\delta} \tag{4.49}$$

In general, the sample size estimate from Eq. (4.40) will be smaller than that given by Eq. (4.47). Because we want the probability of violation δ in Eqs. (4.40) and (4.47) to be extremely small, it is seen from Eq. (4.49) that the reciprocal of δ gets larger at a faster rate than $\ln(\delta/2)$ as δ gets smaller.

4.3 Variance Reduction

If the variance of the output variable of interest can be reduced without affecting its expected value, then greater precision can be achieved; i.e. either a pre-specified precision with lesser number of simulations or a narrower confidence interval around the expected value for the same number of simulations is achieved. When the correlation between the input parameter and the output variable is included in the

analysis, population mean can be estimated with a lesser number of simulations with the same deg. e of accuracy.

Let K_s be the saturated hydraulic conductivity of a porous medium with known mean $\mu_{Ks}=E[K_s]$. μ_{Ks} is a constant for a given type of soil; and varies from soil to soil. Values of μ_{Ks} can be obtained from published literature [*Carsel and Parrish*, 1988]. If the correlation between Ψ and K_s is $\rho_{\Psi,Ks}$, then from the definition of correlation coefficient,

$$\rho_{\Psi,K_s} = \frac{Cov[\Psi,K_s]}{\sqrt{Var[\Psi]}\sqrt{Var[K_s]}} \tag{4.50}$$

in which $Cov[\Psi,K_s]$ is the covariance of Ψ and K_s, and $Var[\Psi]$ and $Var[K_s]$ are the variances of Ψ and K_s, respectively.

Let Ψ be a random variable whose mean $\mu_\Psi=E[\Psi]$ is to be estimated. Then, for any real number β [*Law and Kelton*, 1991],

$$\Psi_c = \Psi - \beta(K_s - \mu_{K_s}) \tag{4.51}$$

is an unbiased estimator of μ_Ψ. In Eq. (4.51), the random variable K_s is called a *control variate* for Ψ as it controls Ψ. The purpose of the technique is to reduce variance of an estimate of Ψ through appropriate selection of β, a real number, called *weight*. From Eq. (4.51)

$$Var\left[\Psi_c\right] = Var\left[\Psi\right] + \beta^2\, Var\left[K_s\right] - 2\,\beta\, Cov\left[\Psi,K_s\right]. \tag{4.52}$$

In Eq. (4.52), for a variance reduction to be guaranteed we must have

$$\beta^2\, Var\left[K_s\right] - 2\beta\, Cov\left[\Psi,K_s\right] < 0 \tag{4.53}$$

which is not possible for every β. A particular value of β, called β^*, which guarantees not only a reduction in variance but also makes it the maximum reduction is given by

$$\beta^* = \frac{Cov\left[\Psi,K_s\right]}{Var\left[K_s\right]} \tag{4.54}$$

for which

$$Var\left[\Psi_c\right] = \left\{1 - \rho^2_{\Psi,K_s}\right\} Var\left[\Psi\right]. \tag{4.55}$$

It can be concluded from Eq. (4.55) that a variance reduction is always achieved with an optimal value of β as long as Ψ and K_s are correlated. Note that even when they are not correlated, nothing is lost

i.e., $$Var\left[\Psi_c\right] \le Var\left[\Psi\right]. \tag{4.56}$$

Evaluation of optimal value of β involves the use of sample equivalents of $Var[K_s]$ and $Cov[\Psi,K_s]$ estimated as

$$S_{K_s}^2 = \frac{1}{(n-1)} \left\{ \sum_{i=1}^{n} \left[K_{s_i} - \bar{K}_s(n) \right]^2 \right\} \tag{4.57}$$

and

$$C_{\Psi,K_s} = \frac{1}{(n-1)} \left\{ \sum_{i=1}^{n} \left[\Psi_i - \bar{\Psi}(n) \right] \left[K_{s_i} - \bar{K}_s(n) \right] \right\} \tag{4.58}$$

where $$\bar{\Psi}(n) = \frac{1}{n} \sum_{i=1}^{n} \Psi_i \,, \tag{4.59}$$

and $$\bar{K}_s(n) = \frac{1}{n} \sum_{i=1}^{n} K_{s_i} \,, \tag{4.60}$$

n being the number of simulations. The estimator of β is computed as

$$\beta^* = \frac{C_{\Psi,K_s}}{S_{K_s}^2} \,. \tag{4.61}$$

The estimator for μ_Ψ can, therefore, be computed as,

$$\Psi_c(n) = \bar{\Psi}(n) - \beta^* \left[\bar{K}_s(n) - \mu_{K_s} \right] \,. \tag{4.62}$$

It is recommended that the above estimator be used for the mean pressure head value.

It has been shown that the upper and the lower bounds of the pressure heads are necessary to obtain the desired sample size for MC simulation, with a pre-specified

level of confidence and a pre-determined error criteria. The methodology proposed here consists of performing simulations repeatedly and re-estimating the sample size based upon the simulation results. The procedure involves performing the simulation, with the use of random generation of hydraulic conductivity parameters required by the model, a pre-specified number of times. The initial number of simulations is just an arbitrary number. Results of the simulations are to be used to obtain the maximum and the minimum values of pressure heads. The pressure head range, obtained from simulation, is used in the sample size Eq. (4.40) to revise the number of simulations required. The simulation is performed for the revised number of simulations and the minimum and the maximum pressure head values are computed to re-estimate the sample size. The procedure is illustrated with the use of flowcharts in Figures 4.1 and 4.2.

4.4 Confidence Limits

Based upon the simulation results, the sample mean and the sample variance can be estimated. As the population mean is unknown, it is of interest to construct a confidence interval around the sample mean of the output variable. But, the construction of the confidence interval requires an apriori knowledge of the population variance, which is not known to begin with. Therefore, we will make use of the identity that the sample variance converges stochastically to the population variance when the sample size is large. Hence, $(\overline{\Psi} - \mu_{\Psi})/[S/(n\text{-}1)^{0.5}]$ has a limiting distribution

that is standard normal; where S^2 is the sample variance and n is the number of simulation runs [*Hogg and Craig*, 1978]. This fact enables us to construct a confidence interval around the sample mean. A 100(1-δ) percent (0<δ<1) confidence interval for the mean head, H, is then given by

$$H(n) \pm z \frac{S(n)}{\sqrt{(n-1)}} \tag{4.63}$$

where z is the cutoff value with area to the left of which is (1-δ/2) for the standard normal density curve. In the above equation, $H(n)$ and $S(n)$ are the mean and the standard deviation of head estimated from n number of simulation runs. It is important to note that as opposed to a deterministic approach which yields only a single pressure head value at a given location, the statistical method provides an interval which contains the sample mean along with the mean head value itself.

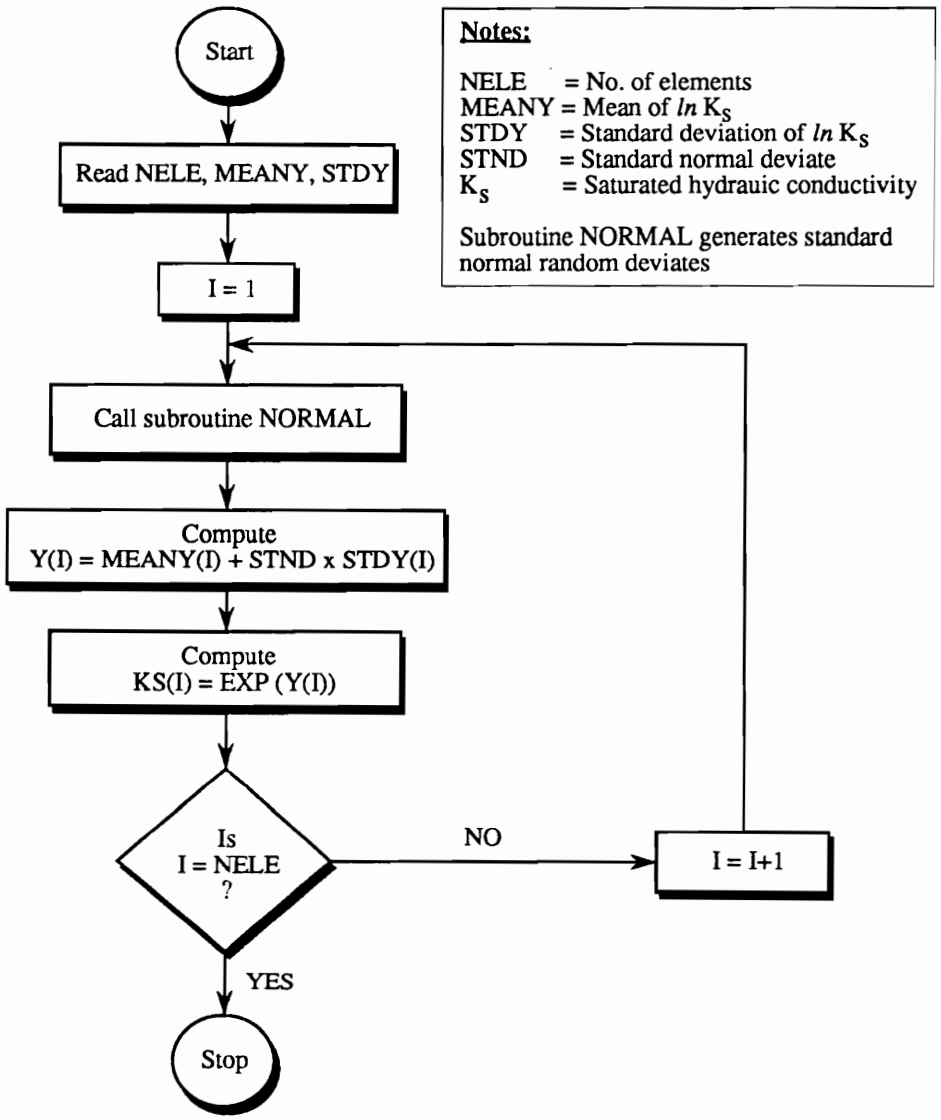

Figure 4.1 Flow Chart - Generation of Saturated Hydraulic Conductivity

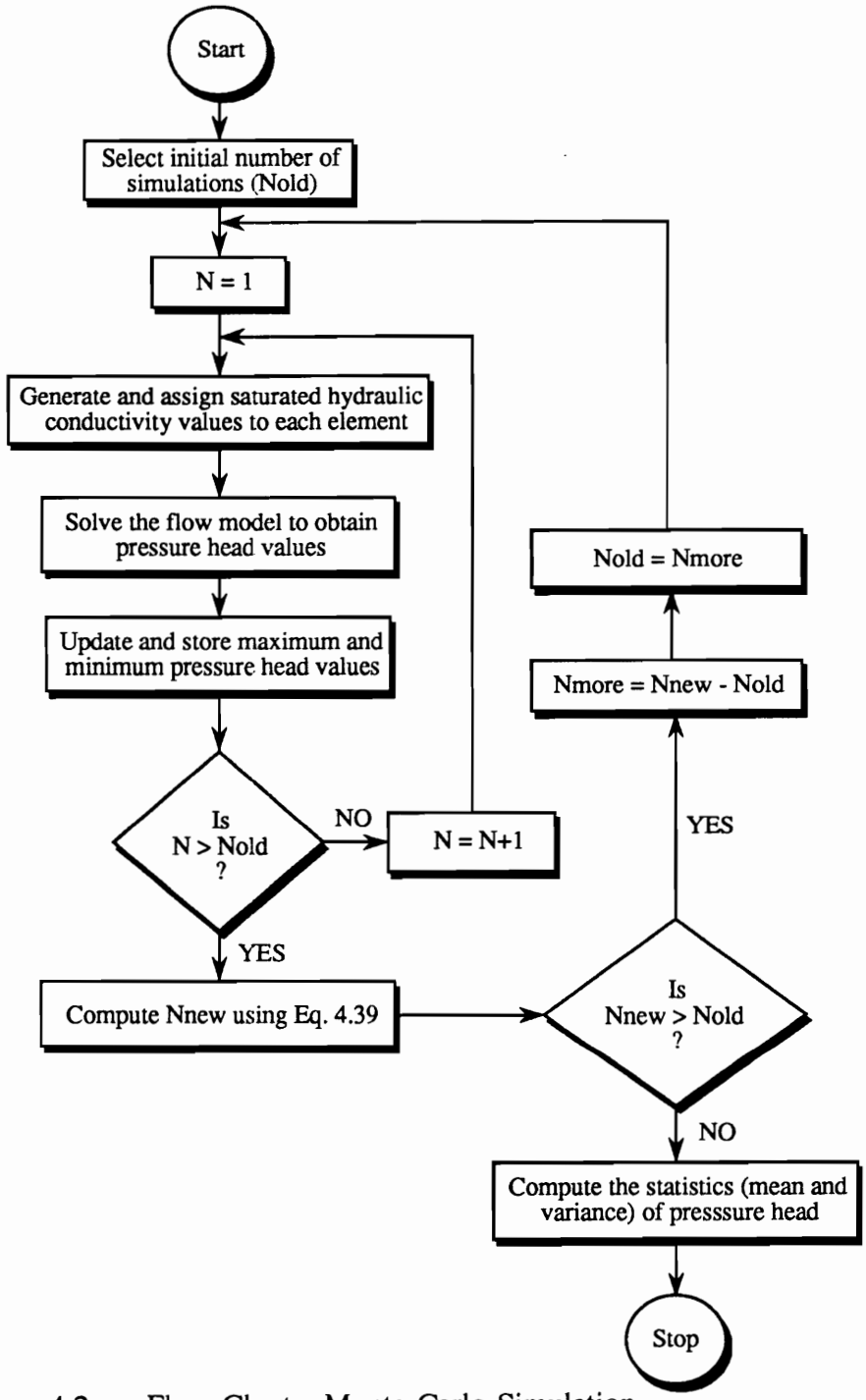

Figure 4.2 Flow Chart - Monte Carlo Simulation

5. Saturated Groundwater Flow Analysis

5.1 Introduction

Traditional methods of solving groundwater flow problems make use of arithmetic, geometric, or harmonic means of saturated hydraulic conductivity as the effective conductivity [*Binley et al.*, 1989]. As pointed out by *Freeze* [1975] and *de Marsily* [1984], it is important to note that the use of these effective values need not always yield correct results. They concluded that an effective medium definition may not be possible for non-uniform flows, e.g., converging radial flow. In the present study the above claim will be verified through a saturated well flow problem by using Monte Carlo (MC) simulation. This chapter deals with the choice of selecting an appropriate sample size for the analysis. A relationship for the variance of the piezometric head is developed. Advantages of the proposed MC approach and the use of control variate technique to reduce the variance of the piezometric head are highlighted.

5.2 Governing Equation

Let Q be the constant pumping rate of a well in an unconfined aquifer with an initial phreatic surface height of H_0. Let K_s be the saturated hydraulic conductivity of the aquifer medium. Then, under steady state conditions, the drawdown in an observation well located at a radial distance r can be obtained using the Dupuit approximation. If the radius of influence of the well is R and the phreatic surface level at the observation well is h (Figure 5.1), then h can be computed as [*Bear*, 1979]

$$h^2 = H_0^2 - \frac{Q}{\pi K_s} \ln\frac{R}{r} . \tag{5.1}$$

Although an assumption of a single well has been made above, Eq. (5.1) is valid for any number of wells as long as radii of influence of each of the pumping wells are the same and the observation well is located equidistant from each of the pumping wells. If the observation well is located at different radial distances from the pumping wells and if there are N number of pumping wells, each with flow rate Q/N, then the phreatic surface height at the observation well may be computed as

$$h^2 = H_0^2 - \frac{Q}{\pi K_s} \ln\left(\frac{R}{r^*}\right) \tag{5.2}$$

where

$$r^* = \left(r_1 r_2 \ldots r_N\right)^{1/N} . \tag{5.3}$$

Subsequent developments are based upon Eq. (5.1). From Eq. (5.1) we can write,

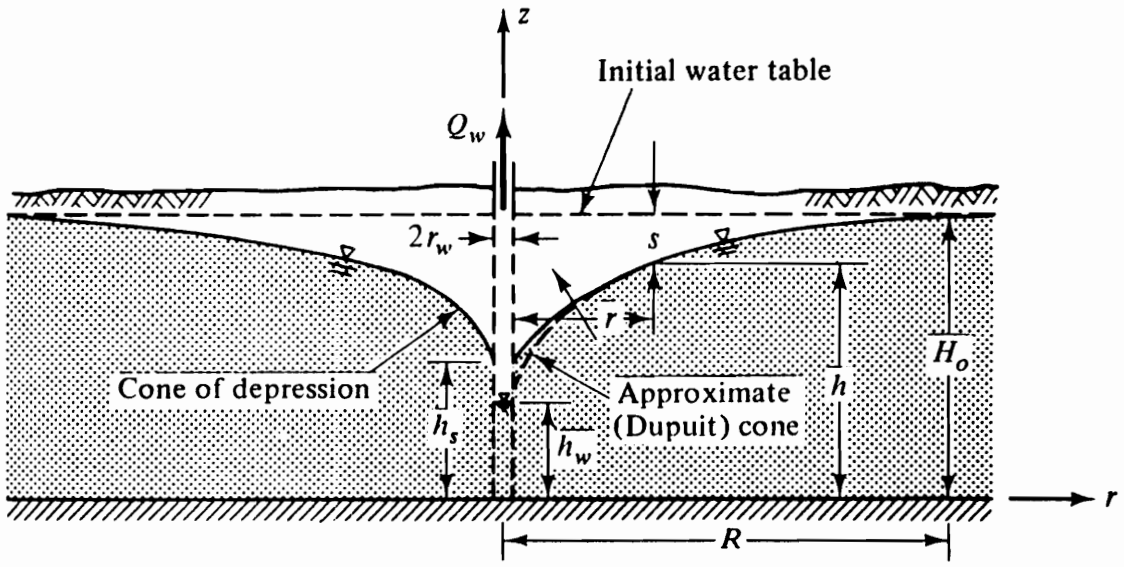

Figure 5.1 Radial Flow to a Well in a Phreatic Aquifer
(Source: *Bear*, 1979)

$$h = \left[H_0^2 - \frac{Q}{\pi K_s} \ln\left(\frac{R}{r}\right) \right]^{0.5} . \qquad (5.4)$$

Using Binomial expansion, Eq. (5.4) can be written as

$$h = H_0 - \frac{Q}{2\pi K_s H_0} \ln\left(\frac{R}{r}\right) - \left(\frac{Q^2}{8\pi^2 K_s^2 H_0^3}\right) \left[\ln\left(\frac{R}{r}\right) \right]^2 - \ldots \qquad (5.5)$$

which, when second and higher order terms are neglected, may be approximated as

$$h \approx H_0 - \left[\frac{Q}{2\pi K_s H_0}\right] \left[\ln\left(\frac{R}{r}\right)\right] . \qquad (5.6)$$

The next step is to derive an analytical expression for variance of h. This can then be compared with the estimated variance from the MC simulations. Before proceeding to the derivation for variance of h, a theorem required for this purpose is presented.

Theorem 5.1

If K is a random variable for which the n-th central moment $\mu_n = E[K_s - \mu]^n$ exists {where $\mu = E[K_s]$ and $\sigma^2 = E[K_s - \mu]^2$}, and if $g(\cdot)$ is a n-times differentiable function, then the variance of the function may be approximated as

$$Var\left[g(K_s)\right] = \left[g^{(1)}(\mu)\right]^2 \sigma^2 \qquad (5.7)$$

where $g^{(1)}(K_s)$ is the first derivative of $g(K_s)$.

Proof

Since $g(K_s)$ is n-times differentiable, use of Taylor series expansion about the mean μ gives

$$g(k_s) = g(\mu) + \frac{g^{(1)}(\mu)}{1!}(k_s-\mu) + \frac{g^{(2)}(\mu)}{2!}(k_s-\mu)^2 + \ldots + \frac{g^{(n)}(\xi)}{n!}(k_s-\mu)^n \qquad (5.8)$$

where $0<\xi<(k_s-\mu)$. Evaluating at K_s and taking expectation, we get

$$E[g(K_s)] = g(\mu) + g^{(2)}(\mu)\frac{\sigma^2}{2} + g^{(3)}(\mu)\frac{\mu_3}{6} + \ldots + E\left[g^{(n)}(\xi)\right]\frac{(K_s-\mu)^n}{n!}. \qquad (5.9)$$

When n=2, dropping higher order terms, we can approximate Eq. (5.9) as

$$E\left[g(K_s)\right] \approx g(\mu) + g^{(2)}(\mu)\frac{\sigma^2}{2}. \qquad (5.10)$$

Applying Eq. (5.10) to the function $g^2(.)$ whose first and second derivatives are $2g(.)g^{(1)}(.)$ and $2\{g(.)g^{(2)}(.) + [g^{(1)}(.)]^2\}$, we can write

$$E\left[g^2(K_s)\right] \approx g^2(\mu) + \left\{\left[g^{(1)}(\mu)\right]^2 + g(\mu)\,g^{(2)}(\mu)\right\}\sigma^2. \qquad (5.11)$$

From the definition of the variance of the function $g(K)$,

$$Var\left[g(K_s)\right] = E\left[g(K_s)\right]^2 - \left\{E[g(K_s)]\right\}^2 \qquad (5.12)$$

Using Eqs. (5.10) and (5.11) in Eq. (5.12), we get

$$Var\left[\, g(K_s)\right] = \left\{ g^{(1)}(\mu)\right\}^2\sigma^2 - \left\{ g^{(2)}(\mu)\right\}^2 \frac{\sigma^4}{4} = \left[\, g^{(1)}(\mu)\right]^2 \sigma^2 \qquad (5.13)$$

which proves theorem 5.1. This theorem is used in deriving a relationship to estimate an "*analytical*" variance of h which is presented next.

5.3 Variance of Piezometric Head

This section presents the development of the relationship that is used to analytically compute the variance of water table elevation. The mathematical derivation is based upon Eq. (5.6), which can be re-written as

$$h = H_0 - \left[\frac{Q}{2\pi H_0} \ln\left(\frac{R}{r}\right)\right]\frac{1}{K_s} \,. \qquad (5.14)$$

For fixed values of Q, H_0, R, and r, the phreatic surface level, h, depends only on the hydraulic conductivity, K_s; for a random K_s, h is also random. Using the variance operator on both sides of Eq. (5.14), we get

$$Var[h] = Var\left[H_0 - \frac{A}{K_s}\right] \qquad (5.15)$$

in which

$$A = \frac{Q}{2\pi H_0} \ln\left(\frac{R}{r}\right) . \tag{5.16}$$

Use of the property of a variance operator in Eq. (5.15) results in,

$$Var[h] = Var\ [\ H_0\ ] + A^2\ Var\left[\ \frac{1}{K_s}\ \right] - 2\ Cov\left[\ H_0\ ,\ \frac{A}{K_s}\ \right] . \tag{5.17}$$

Let us recall that, for any constant H_0,

$$Var\ [H_0] = E\ [H_0 - E[H_0]]^2 = E\ [H_0 - H_0]^2 = 0\ . \tag{5.18}$$

and

$$Cov\left[H_0, \frac{A}{K_s}\right] = E\left[\ (H_0 - E[H_0]\ )\left(\frac{A}{K_s} - E\left[\ \frac{A}{K_s}\ \right]\right)\ \right] = 0\ . \tag{5.19}$$

Therefore, we can write Eq. (5.17) as

$$Var[h] = \left[\ \frac{Q}{2\pi H_0} \ln\left(\frac{R}{r}\right)\ \right]^2 \left(Var\left[\ \frac{1}{K_s}\ \right]\right) . \tag{5.20}$$

Evaluation of Eq. (5.20) requires the estimation of the variance of $1/K_s$ on the right hand side of the equation. Use of theorem 5.1 will be made for the estimation. When $g(K_s) = 1/K_s$, we can write,

$$g^{(1)}(K_s) = -\frac{1}{K_s^2} . \tag{5.21}$$

which gives

$$g^{(1)}(\mu_{K_s}) = \left(-\frac{1}{\{E[K_s]\}^2} \right) . \tag{5.22}$$

Theorem 5.1, therefore, gives

$$Var\left[\frac{1}{K_s}\right] \approx \left\{ -\frac{1}{\{E[K_s]\}^2} \right\}^2 Var[K_s] = \frac{Var[K_s]}{\{E[K_s]\}^4} . \tag{5.23}$$

When Eq. (5.23) is used in Eq. (5.20), we get the relationship for *Var[h]* which is

$$Var[h] = \left[\frac{Q}{2\pi H_0} \ln\left(\frac{R}{r}\right) \right]^2 \left(\frac{Var[K_s]}{\{E[K_s]\}^4} \right) \tag{5.24}$$

This is the expression used to compute the "*analytical*" variance of *h*. Thus computed variance can then be compared with the "*estimated*" variance from the MC simulation procedure.

5.4 Monte Carlo Simulation

In order to perform the MC simulation, an estimate of the sample size necessary to achieve a predetermined error criterion is required. This will be done by following

the theory developed in section 4.2. Specifically, the sample size required to achieve a (1-δ)x100% reliability with an error less than ε is estimated from

$$N \geq - \ln \left(\frac{\delta}{2}\right) \frac{\left[(h_{max} - h_{min})/\varepsilon \right]^2}{2} \tag{5.25}$$

where h_{max} and h_{min} are the maximum and the minimum phreatic surface heights. The range can be written as a constant multiple of the standard deviation of the piezometric head. By using the analytical variance the sample size can therefore be estimated. Based upon the results of the MC simulation and using the correlation between the saturated hydraulic conductivity and the phreatic surface (this correlation is computed from the results of the MC simulation), an unbiased estimator of the mean phreatic surface can be estimated from the theory proposed in section 4.3 as

$$h_c(n) = \bar{h}(n) - \beta^* \left[\bar{K}_s(n) - \mu_{K_s}\right] \tag{5.26}$$

where the subscript *c* represents the unbiased estimate and the overbars represent the average from the traditional MC simulation. The optimal value of the weight, β^*, is computed using the results of the MC simulation using the equation (see section 4.3)

$$\beta^* = \frac{C_{h,K_s}}{\sigma^2_{K_s}} . \tag{5.27}$$

A confidence interval can then be placed around the unbiased estimator obtained in Eq. (5.26) using the relationship [*Hogg and Craig*, 1978]

$$h_c(n) \pm z\, S(n) / \sqrt{(n-1)} \tag{5.28}$$

where $S(n)$ is the standard deviation from n simulation runs. The expected head estimate obtained from the MC simulation will be compared with the head values obtained by using the arithmetic, geometric, and harmonic means as the effective conductivity. The application is presented in chapter 6.

6. Applications

6.1 Introduction

In this chapter various problems taken from the literature are analyzed and performance improvements with the present approach are reported. The problems are: (1) one-dimensional, steady state, saturated flow problem by *Freeze* [1975]; (2) one-dimensional, steady state, unsaturated flow problem by *Bear* [1972]; (3) two-dimensional, steady state, unsaturated flow problem by *Yeh* [1987]; (4) one-dimensional, transient, unsaturated flow problem; and (5) a radial well flow problem by *Bear* [1979]. The principal objective of these applications is to verify the effective medium relationship and the MC approach developed in this thesis. Both the methods perform quite well.

6.2 One-Dimensional, Steady state, Saturated Flow

The first problem is a one-dimensional, steady state, saturated flow problem solved by *Freeze* [1975]. It consists of a 100 cm long column with fixed head boundaries at both ends. The hydraulic head at one end (x=0) is 0 cm and that at the other end (x=100 cm) is 100 cm [Figure 6.1]. Saturated hydraulic conductivities are

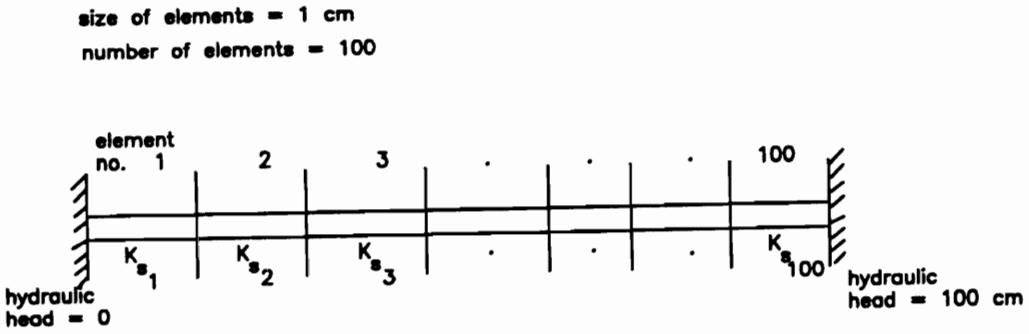

Figure 6.1 One-Dimensional, Steady State, Saturated Flow Problem

generated from a log-normal distribution with mean 1.0 and standard deviation 0.5.

The flow equation describing one-dimensional, steady state, saturated flow through a heterogeneous porous medium is given by

$$\frac{\partial}{\partial x}\left[K_s(x)\ \frac{\partial \phi}{\partial x}\right] = 0 \tag{6.1}$$

in which ϕ is the hydraulic head and $K_s(x)$ is the saturated hydraulic conductivity at x. The boundary conditions are the fixed head boundaries at the ends of the flow domain. Mathematically, these are expressed as

$$\phi(0) = \phi(x=0) = \phi_0 \tag{6.2}$$

and

$$\phi(L) = \phi(x=L) = \phi_L \tag{6.3}$$

where L is the length of the flow domain. If the flow domain is divided into m blocks of equal size and if K_{si} ($i = 1, 2, ..., m$) is the saturated hydraulic conductivity of the i^{th} block, then the hydraulic head for each block can be computed by using the following set of equations [*Freeze*, 1975]:

$$\phi_1 = \frac{K_{s_m}}{K_{s_1}} ax + \phi_0 \quad , \qquad i = 1 \tag{6.4}$$

$$\phi_i = \frac{K_{s_m}}{K_{s_i}} ax + \frac{aL}{m} \sum_{j=2}^{i} (j-1) \left(\frac{K_{s_m}}{K_{s_j}} \right) \left(\frac{K_{s_j}}{K_{s_{j-1}}} - 1 \right) + \phi_0 \quad , \quad 1 < i < m \tag{6.5}$$

and

$$\phi_m = ax - aL + \phi_L \quad , \qquad i=m \tag{6.6}$$

where a is given by the following equations

$$a = \frac{\phi_L - \phi_0}{\frac{L}{m} \left[(m-1) \left(\frac{K_{s_m}}{K_{s_{m-1}}} - 1 \right) + m \right]} \quad , \qquad \text{for} \qquad m=1,2 \tag{6.7}$$

and when $m \geq 3$,

$$a = \frac{(\phi_L - \phi_0)}{\frac{L}{m} \left[\left\{ \sum_{j=2}^{m-1} (j-1) \left(\frac{K_{s_m}}{K_{s_j}} \right) \left(\frac{K_{s_j}}{K_{s_{j-1}}} - 1 \right) \right\} + (m-1) \left(\frac{K_{s_m}}{K_{s_{m-1}}} - 1 \right) + m \right]} \quad . \tag{6.8}$$

In Eqs. (6.7) and (6.8), i is computed as

$$i = \frac{mx}{L} + 1 \quad . \tag{6.9}$$

The right hand side of Eq. (6.9) is truncated to the lower integer.

Generation of saturated hydraulic conductivities from a log-normal distribution requires the mean and the variance of the log hydraulic conductivity. The random log saturated hydraulic conductivities, Y's ($Y = lnK_s$), are then generated using

$$Y = \mu_Y + R_N \sigma_Y \tag{6.10}$$

where R_N is the standard normal random deviate generated; μ_Y and σ_Y are the mean and the standard deviations of log saturated hydraulic conductivity, respectively. However, in general, only the mean and the variance of the saturated hydraulic conductivity and the parameters without the log transformation are available. For the log transformed data, μ_Y and σ_Y have to be computed. The relationships between these two sets of parameters are [*Aitchison and Brown*, 1957]

$$\mu_{K_s} = \exp\left[\mu_Y + \frac{\sigma_Y^2}{2}\right] \tag{6.11}$$

and

$$\sigma_{K_s}^2 = e^{\left(2\mu_Y + \sigma_Y^2\right)}\left[e^{\sigma_Y^2} - 1\right] \tag{6.12}$$

In Eqs. (6.11) and (6.12), μ_{Ks} and σ_{Ks} are the mean and the standard deviation of the saturated hydraulic conductivity without the log transformation. Solution of Eqs. (6.11) and (6.12) results in μ_Y and σ_Y which are used in Eq. (6.10) along with the standard normal deviate to generate random Y. Exponential transformation of thus

generated Y value

$$K_s = e^{Y} \tag{6.13}$$

results in the necessary saturated hydraulic conductivity value. This is done for all blocks within the flow domain. This is schematically presented in Figure 4.1. The simulation is performed first by dividing the flow domain into m segments. Each MC simulation run, therefore, consists of generating m random saturated hydraulic conductivity values (one for each block), assigning one of these values to each block, and solving Eqs. (6.4) through (6.9) for hydraulic heads. This is repeated a pre-specified number of times and the statistics, namely the first and the second moments, are estimated.

MC simulation runs were performed initially for N=25. Based upon the simulation results of the 25 runs, it was observed that the maximum range of hydraulic heads occurred at the central node (node no. 51). Therefore, it was decided to update the sample size estimate using the maximum and the minimum hydraulic head values at this node. From the results of 25 simulation runs, the maximum and the minimum hydraulic head values, ϕ_{max} and ϕ_{min}, at node no. 51 were 53.7 cm and 46.7 cm, respectively. Using these values of ϕ_{max} and ϕ_{min} in Eq. (4.40) with a confidence coefficient of 0.95 (δ=0.05) and an absolute error of 1 cm (ε=1 cm), we get the revised sample size of 91. In other words, a total of 91 simulation runs are necessary if one were to rely upon the simulation results within the criterion specified

above. Therefore, an additional 66 (=91-25) realizations were made, each realization consisting of the generation of saturated hydraulic conductivity values and solving for the hydraulic heads. Analyses of hydraulic heads from the 91 simulation runs resulted in the maximum hydraulic head value of 53.9 cm and the minimum hydraulic head value of 46.7 cm. Use of these ϕ_{max} and ϕ_{min} values in the sample size estimate Eq. (4.40) resulted in the new sample size of 96. Hence, an additional 5 (= 96-91) runs were made and the range of hydraulic head was unaffected. A plot of the mean hydraulic heads, at different spatial locations within the flow domain, based upon the 96 simulation runs and a second plot showing the standard deviation of hydraulic heads at different spatial locations are shown in Figures 6.2 and 6.3, respectively. The results obtained by *Freeze* [1975] have also been superimposed in these figures. It is observed, from the plots, that the mean hydraulic heads and the standard deviations are quite close to the results of *Freeze* [1975]. It is important to note that the results obtained here was with 96 simulation runs, which is only a fraction (19.2%) of the total number of simulations performed by Freeze [500 runs]. Also, the mean and the standard deviation of hydraulic head distributions agree quite well with the results of *Freeze* [1975]. Therefore, no additional information is gained by performing a very large number of simulation runs.

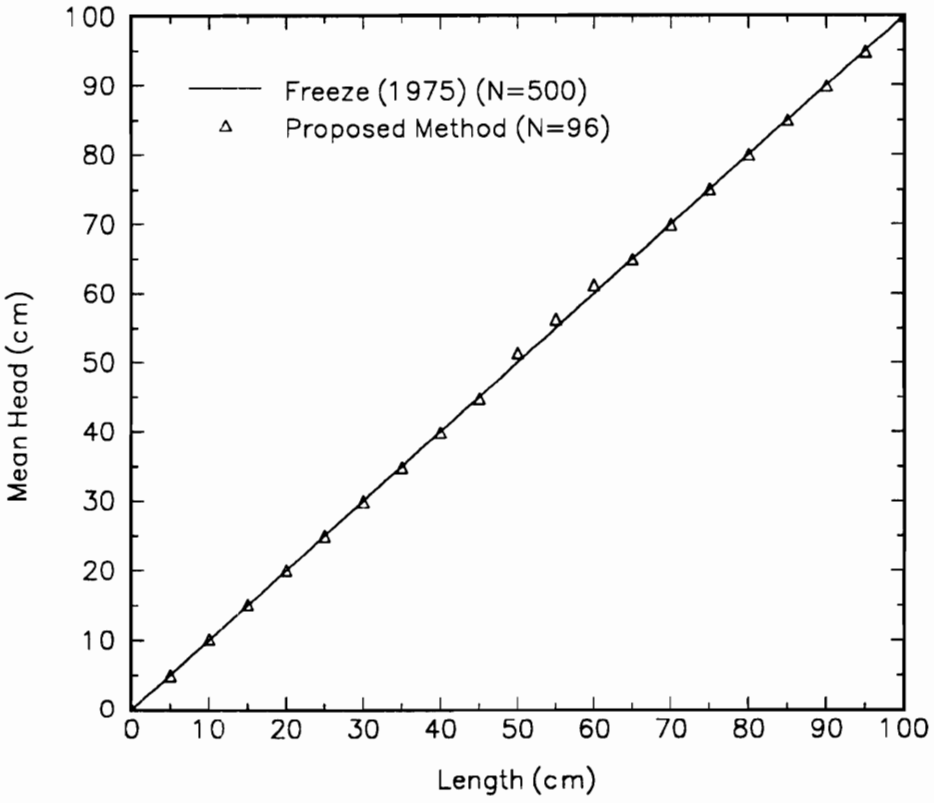

Figure 6.2 Mean Hydraulic Head - One-Dimensional, Steady State, Saturated Flow

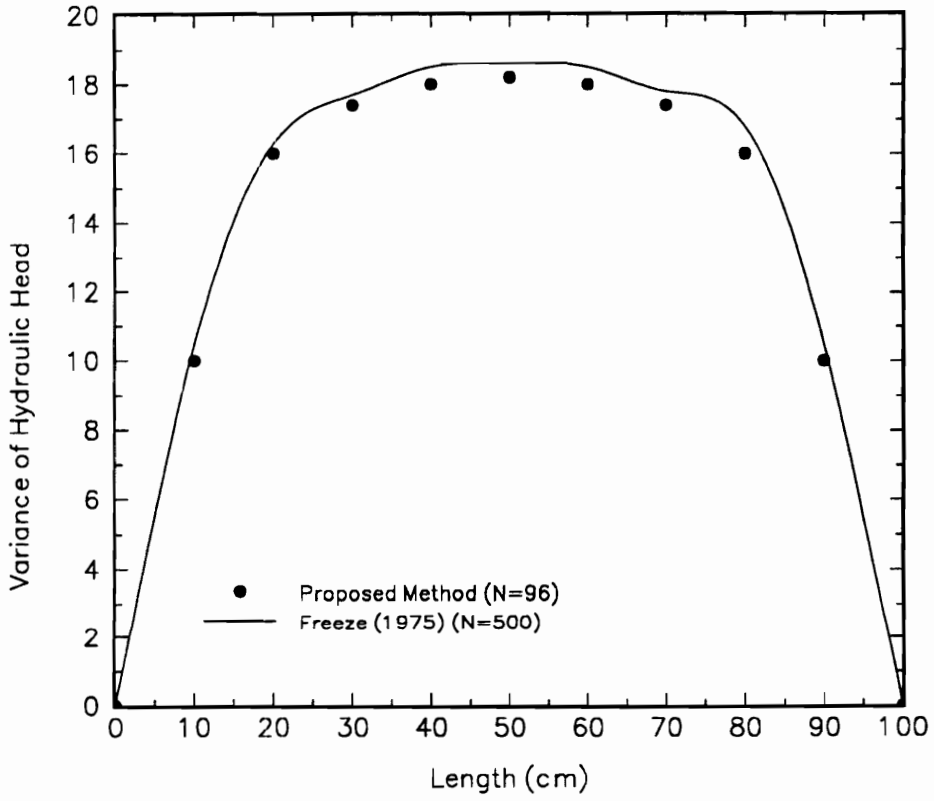

Figure 6.3 Hydraulic Head Variations: 1-D, Steady State, Saturated, Flow

6.3 One-Dimensional, Steady State, Infiltration

The next problem is a 200-cm thick soil column shown in Figure 6.4. A constant flux rate of 0.10 cm/day is to be maintained at the upper boundary whereas the lower boundary is a stationary water table. The mean and the standard deviation of the saturated hydraulic conductivity of the soil are 36.6 cm/day and 1.09 cm/day, respectively. Use of these values in Eqs. (6.11) and (6.12) results in μ_Y and σ_Y values of 3.0 and 1.09, respectively. The Gardner parameter, α, is 0.07 cm^{-1}. The column is divided into 200, 1-cm thick, layers.

Log-saturated hydraulic conductivities are generated from a normal distribution generator using the parameters given above. Eq. (6.13) is then used to transform the generated log-conductivity values to saturated conductivities. Two hundred saturated hydraulic conductivity values, one for each element, are to be generated for each realization of MC simulation. For each realization, the pressure head values are computed using the equation [see *Bear*, 1972]

$$\psi = \frac{1}{\alpha} \ln \left\{ \exp \left[-\alpha \, (z-\psi_0) \right] + \frac{q}{K_s} \left(e^{-\alpha z} - 1 \right) \right\} \tag{6.14}$$

with ψ_0, the prescribed head at the lower boundary q, the flux at the ground surface, and z is the elevation above the datum (taken positive upwards). Eq. (6.14) is valid only if

$$0 < \left\{ \exp\left[-\alpha\ (z-\psi_0)\right] + \frac{q}{K_s}\left(e^{-\alpha z} - 1\right) \right\} \leq 1\ . \qquad (6.15)$$

An initial sample size of 100 was chosen. The maximum and minimum pressure head values at each node were recorded to revise the sample size required for MC simulation. The maximum range in pressure heads was observed at node number 159, the extreme values being -82.406 cm and -86.191 cm. Using these pressure head values as ψ_{max} and ψ_{min} in Eq. (4.40) along with $\delta = 0.05$ and $\varepsilon = 1/3$ cm, the number of simulations required was found to be 238. Therefore, an additional 138 (=238-100) realizations were performed and the sample size estimate was updated based upon the results of 238 realizations. The revised sample size was estimated to be 243 (maximum and minimum pressure head values were -82.401 cm and -86.226 cm, respectively. The maximum range, once again, was found to occur at node 159). This necessitated an additional 5 simulation runs, which were performed starting with the generation of log-saturated hydraulic conductivities, exponential transformation of these values using Eq. (6.13) and pressure head computation using Eq. (6.14) with the assigned saturated hydraulic conductivity values. The results of 243 simulation runs indicated that there was no need for additional simulations.

The mean pressure head variation as a function of depth, computed based upon 243 simulation runs, is shown in Figure 6.5. Standard deviations computed at each

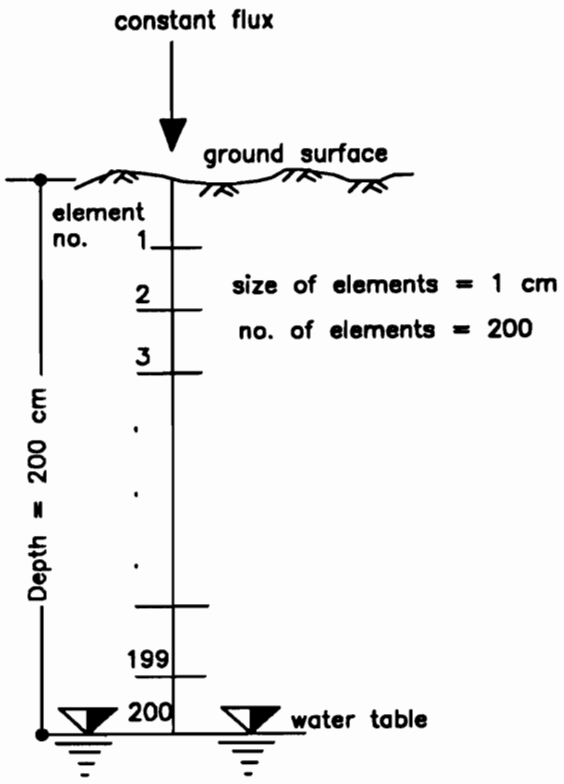

Figure 6.4 One-Dimensional, Steady State, Infiltration Problem

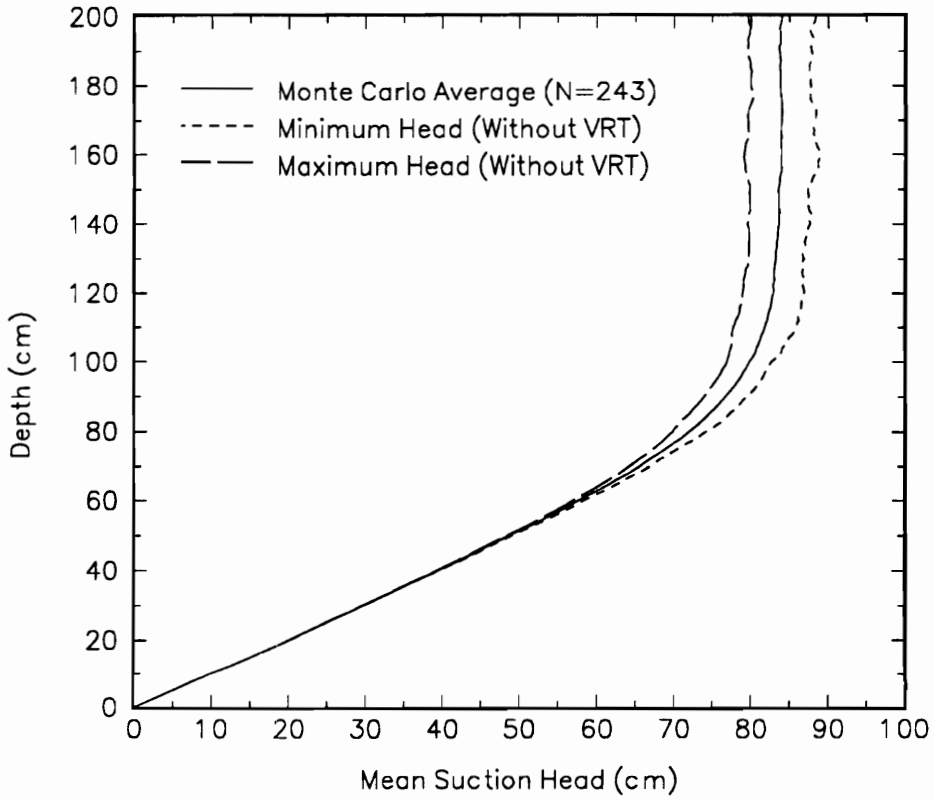

Figure 6.5 Mean Suction Head Distribution with 95% Confidence Interval, Without Variance Reduction

node are used in Eq. (4.63) to compute the confidence interval. The z value used for the computation is 1.96 corresponding to a confidence level of 0.95 [see *Beyer*, 1984].

Mean pressure head distribution is also computed using the proposed effective conductivity relationship, Eq. (3.31). The mean and the standard deviation of log-hydraulic conductivity are used in Eq. (3.38) to compute K_s^* and is 32.14 cm/day. This value of K_s^* when used in Eq. (3.39) results in the K_{eff} relationship as

$$K_{eff} = 32.14\ e^{-0.07H} \tag{6.16}$$

Therefore, when Eq. (6.14) is solved with K_s replaced by K_s^*, the result is the mean head distribution, $H(z)$. This mean head distribution is shown in Figure 6.6 along with the mean pressure head distribution obtained from the 243 MC simulation runs and that obtained by using Eq. (3.31), which does not involve any approximation. It is seen that the results of the mean pressure heads based upon both the effective conductivity relationships [Eqs. (3.31) and (3.39)] agree with the results of the MC simulation. Therefore, it can be concluded that the field scale Gardner's effective conductivity approximation is a reasonable first approximation.

Binley et al. [1989] have suggested the use of geometric mean as the effective conductivity. The flow problem was also solved by using the geometric mean. The mean pressure head distribution obtained for this case along with the mean pressure head distribution from MC simulation [N = 243] and the effective conductivity relationship [Eq. (6.16)] are shown in Figure 6.7. It is seen from the figure that the

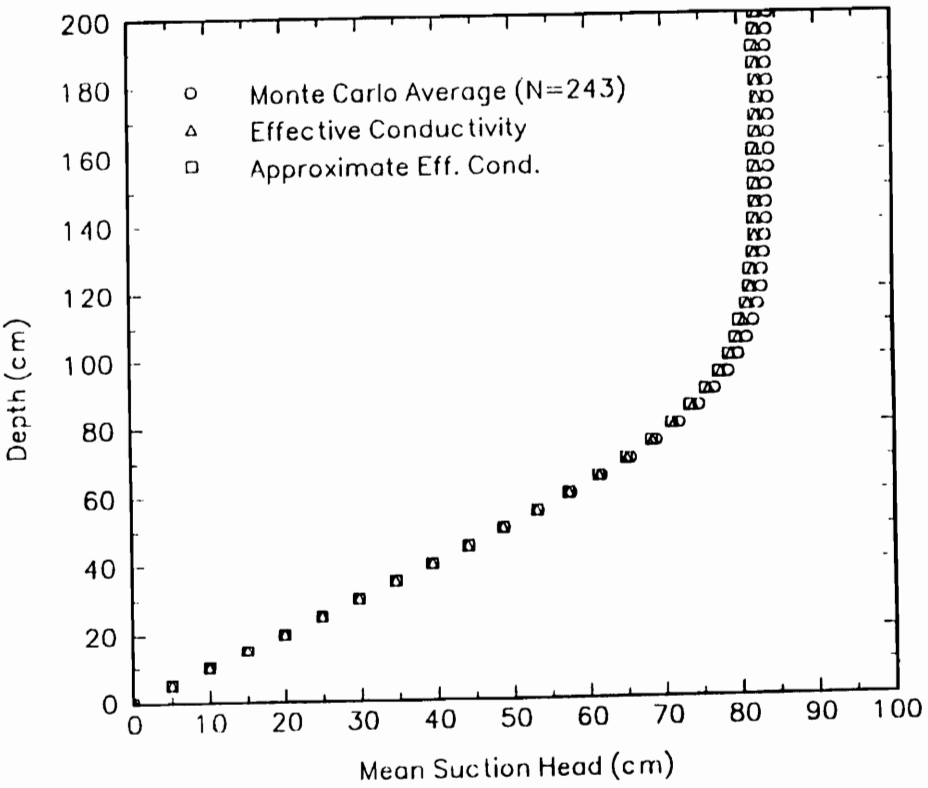

Figure 6.6 Comparison of Mean Head Distributions
One-Dimensional, Steady State, Unsaturated Infiltration

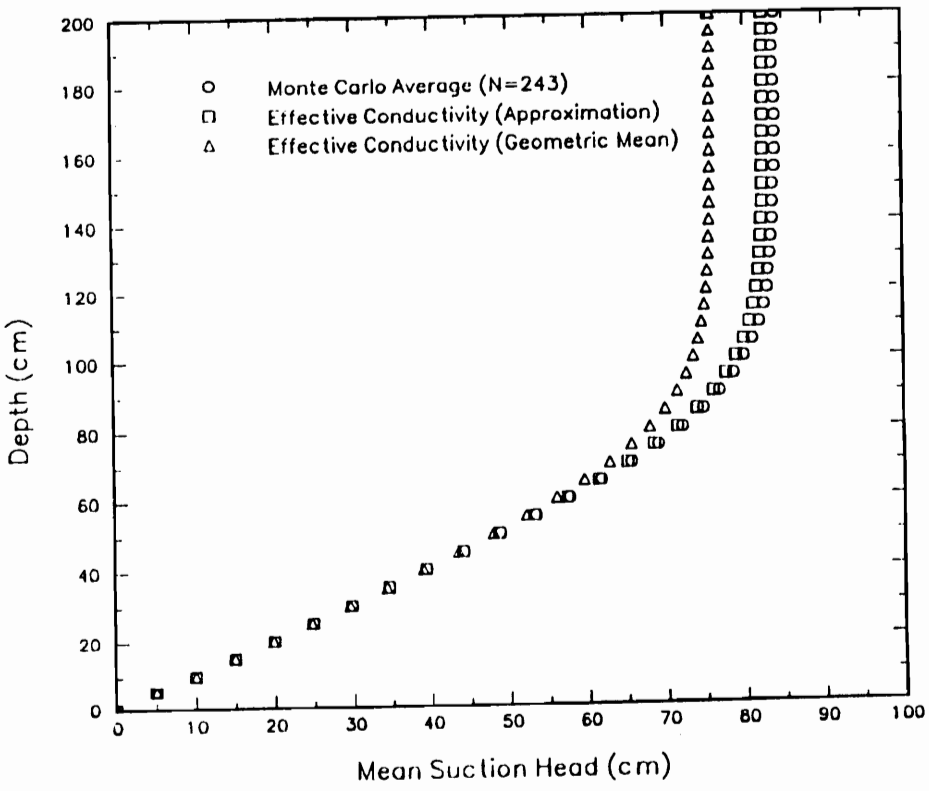

Figure 6.7 Comparison Between Proposed and Traditional Methods
One-Dimensional, Steady State, Unsaturated Flow

geometric mean results differ considerably from the other two methods.

To check the advantage of the proposed variance reduction technique (VRT), the weight β^* was computed using Eq. (4.60). These β^*'s were then used in Eq. (4.61) to obtain the unbiased estimator of the mean pressure head. This was repeated for each node and the result is plotted in Figure 6.8.

A comparison is also made between sample size estimates obtained by Chebyshev inequality and Hoeffding inequality. As already shown above, the Hoeffding inequality resulted in a sample size of 243 for a confidence coefficient of 0.95. To determine the sample size required to satisfy Chebyshev inequality, Eq. (4.46) is used along with the sample variance of 2.054 cm/day. Eq. (4.46) along with a confidence coefficient of 0.95, and an absolute error of 1/3 cm, gives

$$N_C \geq \frac{2.054^2}{0.05 \left\{\frac{1}{3}\right\}^2} = 760 \ , \tag{6.17}$$

which is significantly larger than the sample size actually used.

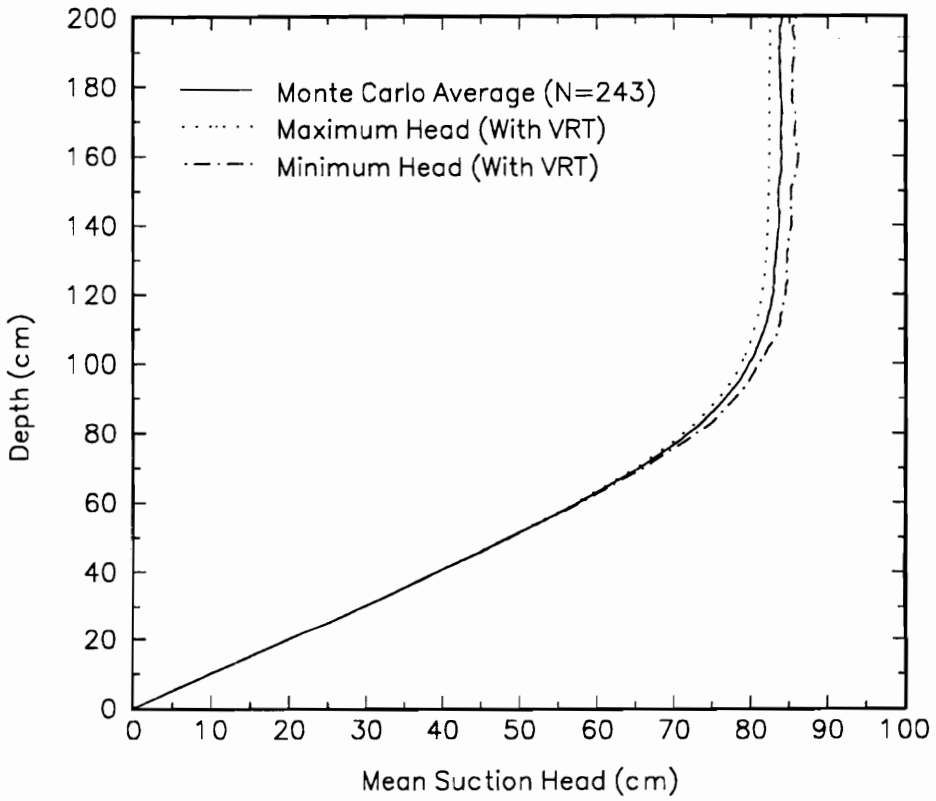

Figure 6.8 Mean Suction Head Distribution with 95% Confidence Interval, With Variance Reduction

6.4 Two-Dimensional, Steady State, Infiltration

Next problem is a two-dimensional, steady state, unsaturated flow. The flow domain consists of a two-dimensional vertical face (Figure 6.9). The lower boundary is maintained at a constant suction head of 150 cm whereas the central node on the upper boundary is maintained at a constant pressure head of zero. The rest of the upper face and the two vertical faces are no flux boundaries. The flow domain is 150 cm by 150 cm in size. The statistics of the saturated hydraulic conductivity is the same as in the one-dimensional steady state infiltration case. The flow domain is discretized into 100 (10 in each direction) elements and 121 nodes (11 in each direction).

Saturated hydraulic conductivity values were generated and assigned to each block [see Figure 4.1] and the boundary value problem was solved using FEMWATER [*Yeh*, 1990]. The procedure was repeated 100 times. Average suction head distribution from 100 simulations were computed at each node (total number of nodes = 121). The maximum range of suction head was found to be 6.281 cm which was used in Eq. (4.39) along with a confidence coefficient of 0.95 and an absolute error of 1 cm to yield the sample size of 100. Therefore, it was decided that no additional simulations were needed.

The suction head contour within the flow domain is shown in Figure 6.10. To check the validity of the approximate effective conductivity relationship, the problem

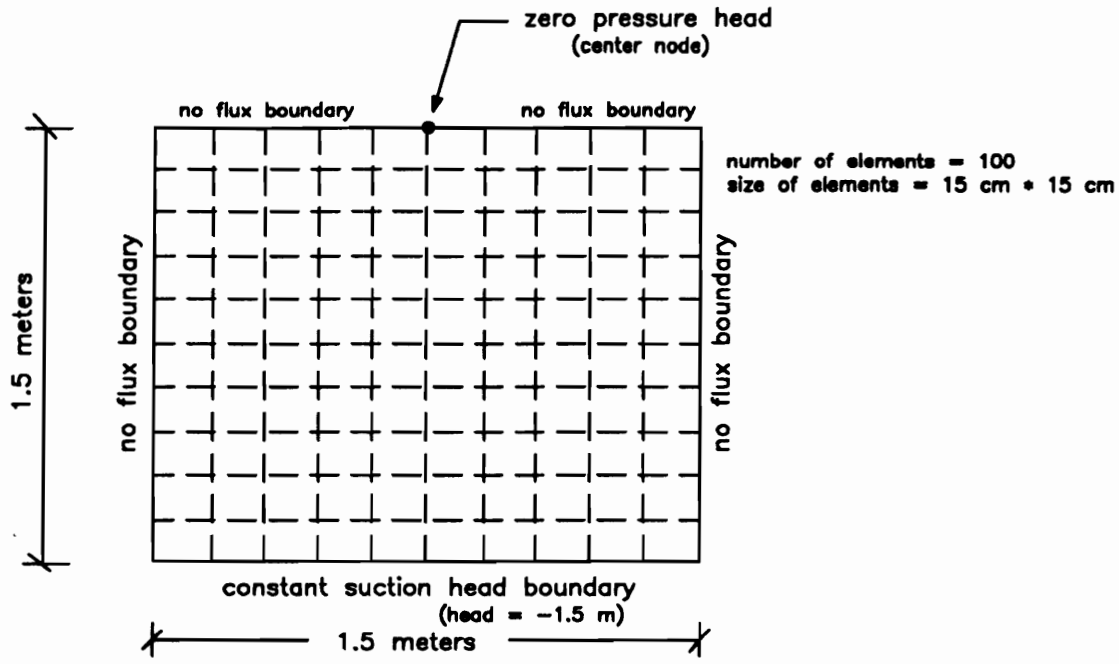

Figure 6.9 Two-Dimensional, Steady State, Infiltration Problem

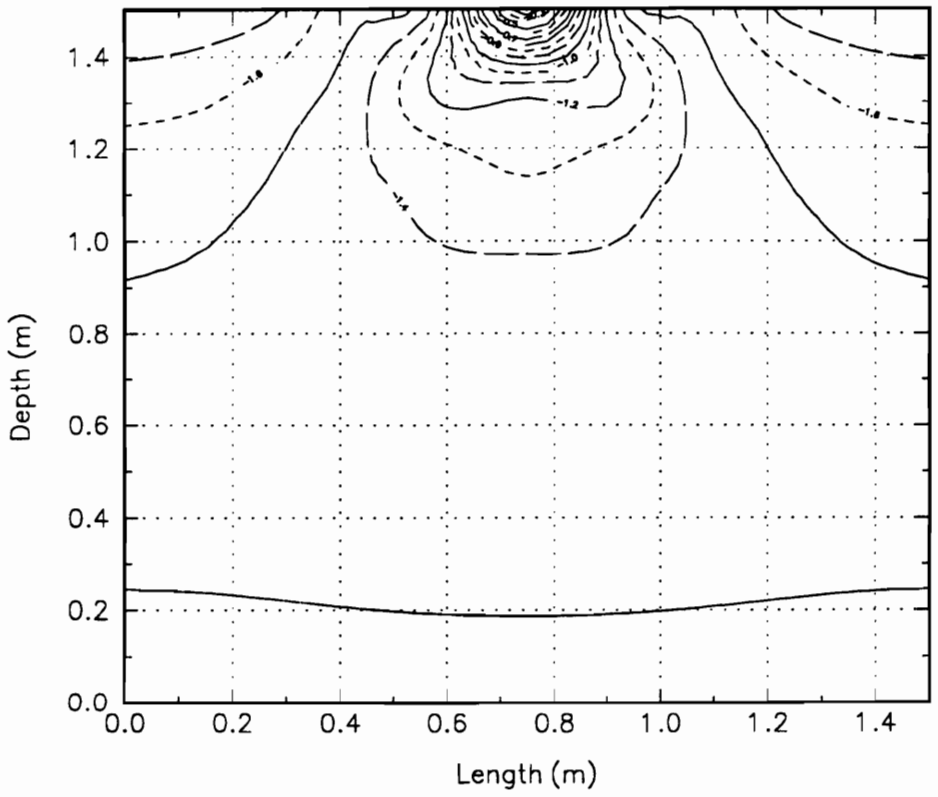

Figure 6.10 Average Suction Head Distribution, MC Simulation
Two-Dimensional, Steady State, Unsaturated Flow

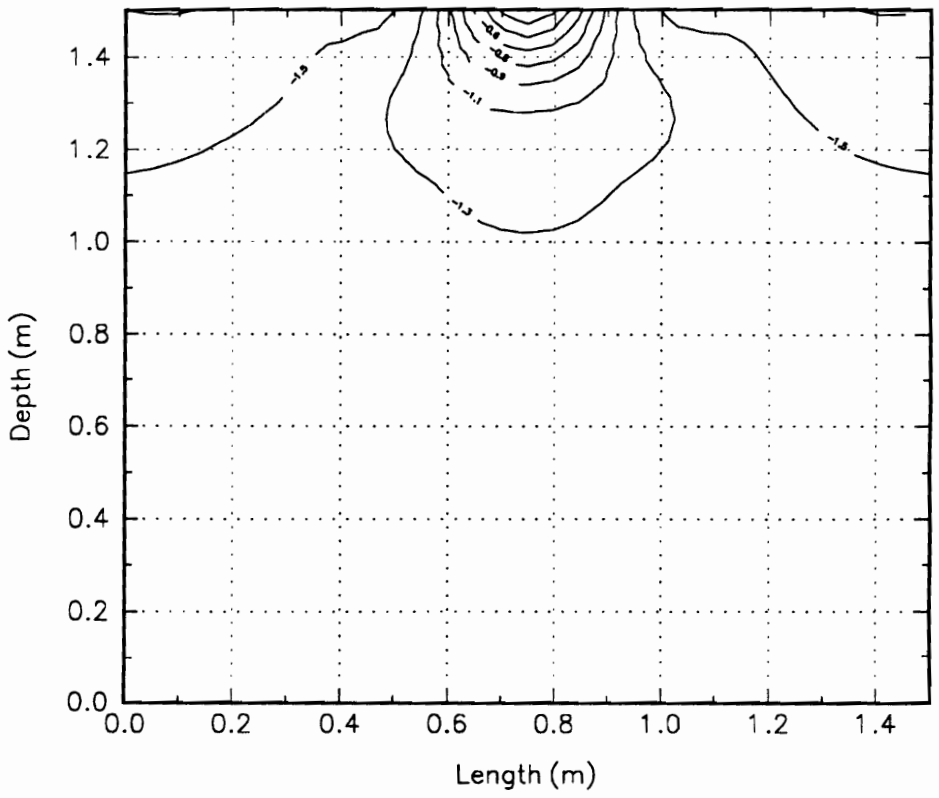

Figure 6.11 Mean Suction Head Distribution, Effective Medium Approximation, Two-Dimensional, Steady State, Unsaturated Flow

was solved with K_s^* instead of K_s in the local Gardner's equation. K_s^* was computed using Eq. (3.38) and the same parameters as in the one-dimensional example problem [section 6.3] were used. It was further assumed that $K_{sx} = K_{sz}$. As in the one-dimensional, steady-state, infiltration problem, K_s^* was computed as 32.14 cm/day and α was taken as 0.07 cm^{-1}. The mean head variations obtained by using Eq. (6.16) in place of local Gardner's equation is shown in Figure 6.11. It is seen that mean head distribution resulting from the effective conductivity is close to the mean head distribution obtained from the MC simulation.

6.5 One-Dimensional, Transient, Unsaturated Flow

The fourth problem is the one-dimensional, transient problem. Two problems with identical boundary conditions are solved, the only difference being the size of the domain. The first domain consists of a 100 cm deep soil column (Figure 6.12) and the second domain is larger with 200 cm depth. The ground surface is maintained at zero flux whereas the lower boundary is maintained at unit hydraulic gradient. Pressure heads at all points within the flow domain are assumed to be known at the beginning of simulation.

Saturated hydraulic conductivity values were generated from a lognormal generator. The boundary value problem consisting of Eqs. (2.10) through (2.13) was solved repeatedly. Simulations were performed for a total length of 50 days. A sample size of 50 was used.

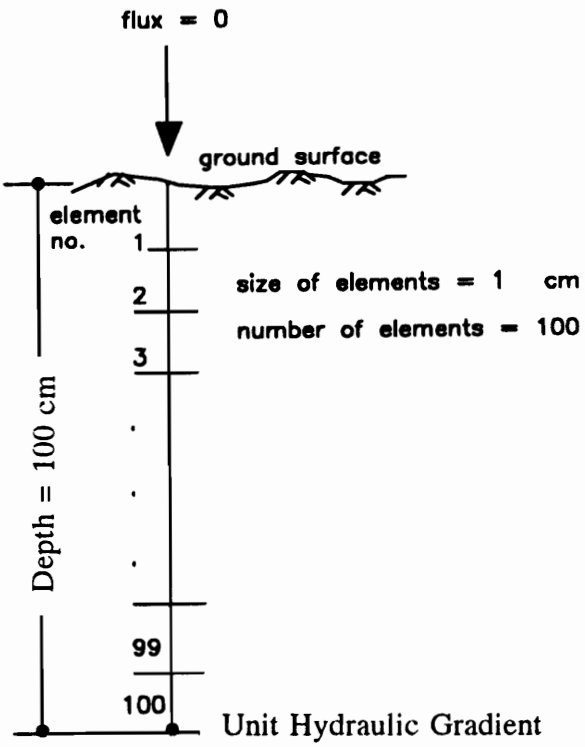

Figure 6.12 One-Dimensional, Transient, Unsaturated Flow Problem

The spatial distribution of mean pressure heads, based on 50 MC simulation runs, are shown in Figure 6.13. The figure shows the mean head distributions (variation over depth) at times 5, 25, and 50 days for both the smaller and the larger domains. There is good agreement between the mean pressure heads, except at the vicinity of the lower boundary of the smaller domain. The disagreement near the lower boundary of the smaller domain is attributed to the condition imposed at the lower boundary of the smaller domain. Figures 6.14 and 6.15 show comparison of variances for the two domains. It is observed that, in general, variances of pressure heads increase with depth. As can be seen from these figures, there is reasonable agreement between the results of the two domains within the top 70 cm. Effects of the boundary are seen through the disagreement at larger depths for the two domains. Furthermore, there seem to be no apparent variance behavior with time.

These problems were also solved with the Gardner effective conductivity, Eq. (3.35), proposed in this thesis. The mean pressure head field obtained by using the EMA for the two domains are shown in Figure 6.16. To facilitate the one-on-one comparison with the results of MC simulations, the mean pressure heads obtained by the effective medium approximation and MC simulations for the 100 cm deep column are shown on Figure 6.17. Similar results for the 200 cm deep column are shown on Figure 6.18.

Figures 6.17 and 6.18 show that the EMA performs well when compared with the MC simulation results. The slight discrepancies in mean pressure heads may be attributed to the approximation made in arriving at Eq. (3.35) from Eq. (3.29). It is

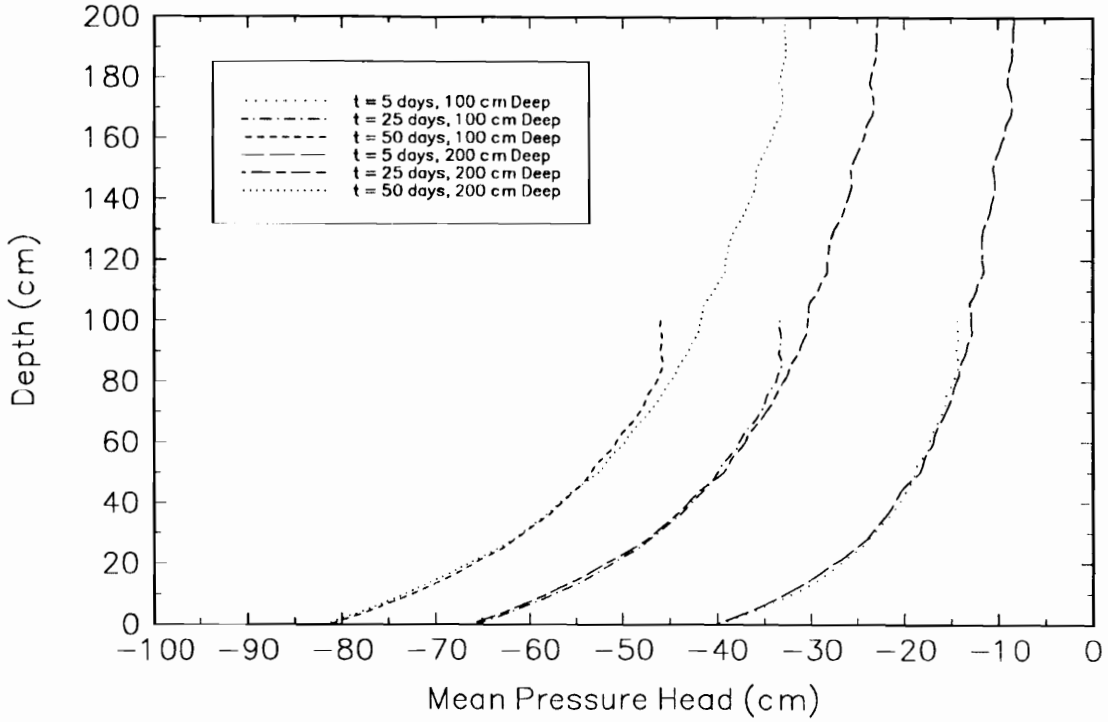

Figure 6.13 Mean Pressure Head Distribution
Comparison Between Different Sized Domains, Monte Carlo Simulation

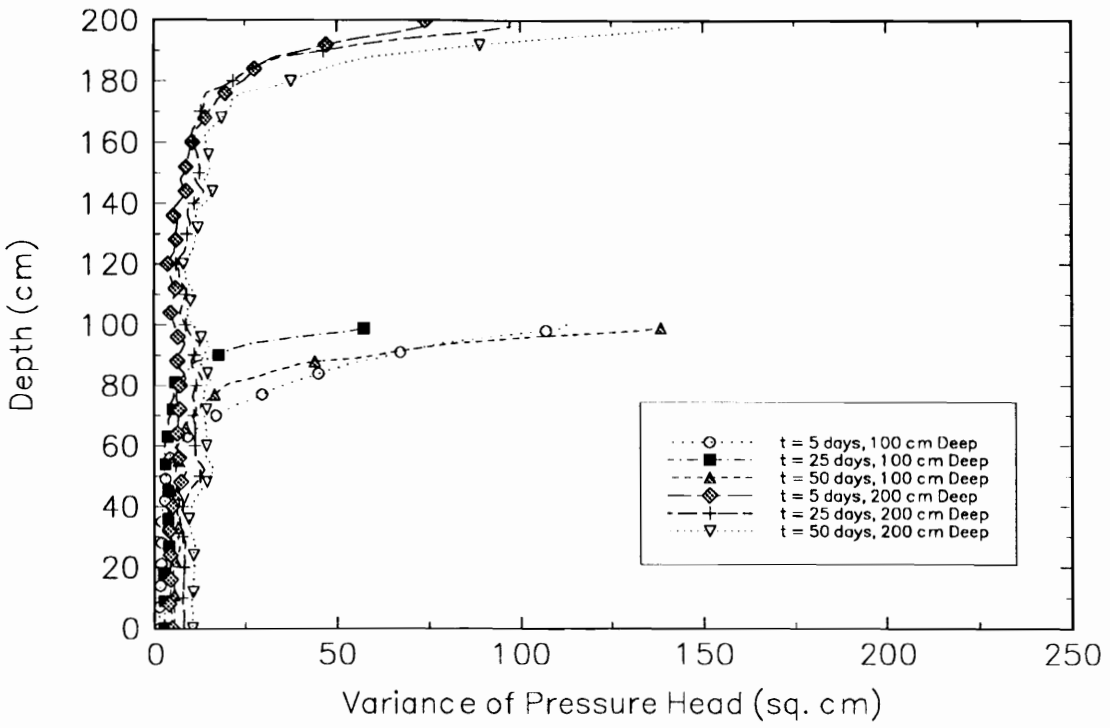

Figure 6.14 Comparison Between Variances of Pressure Heads Different Sized Domains, Monte Carlo Simulation

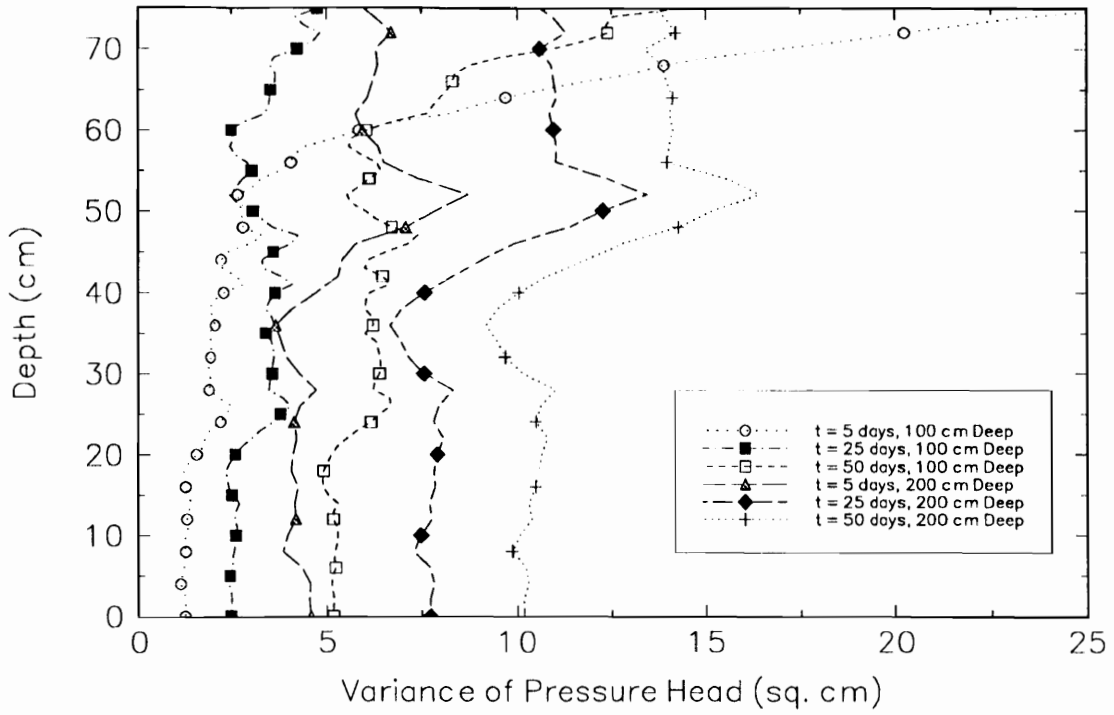

Figure 6.15 Comparison Between Variances of Pressure Heads
Upper 70 cm, Different Sized Domains, Monte Carlo Simulation

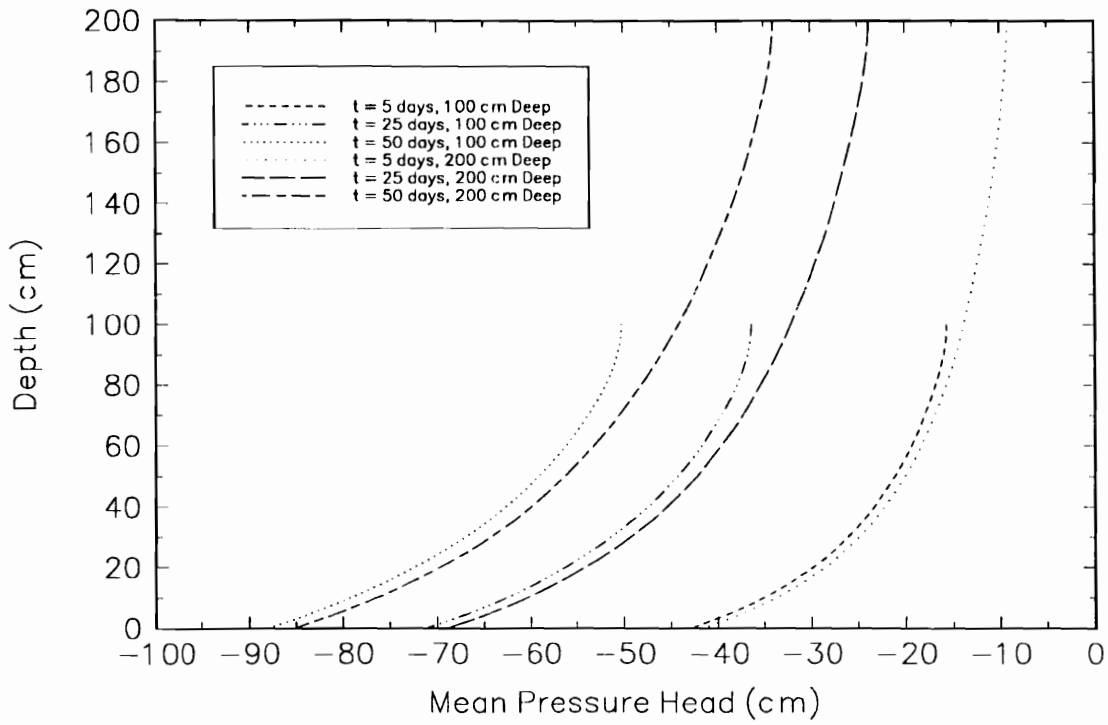

Figure 6.16 Mean Pressure Head Distribution
Comparison Between Different Sized Domains, Effective Medium Approximation

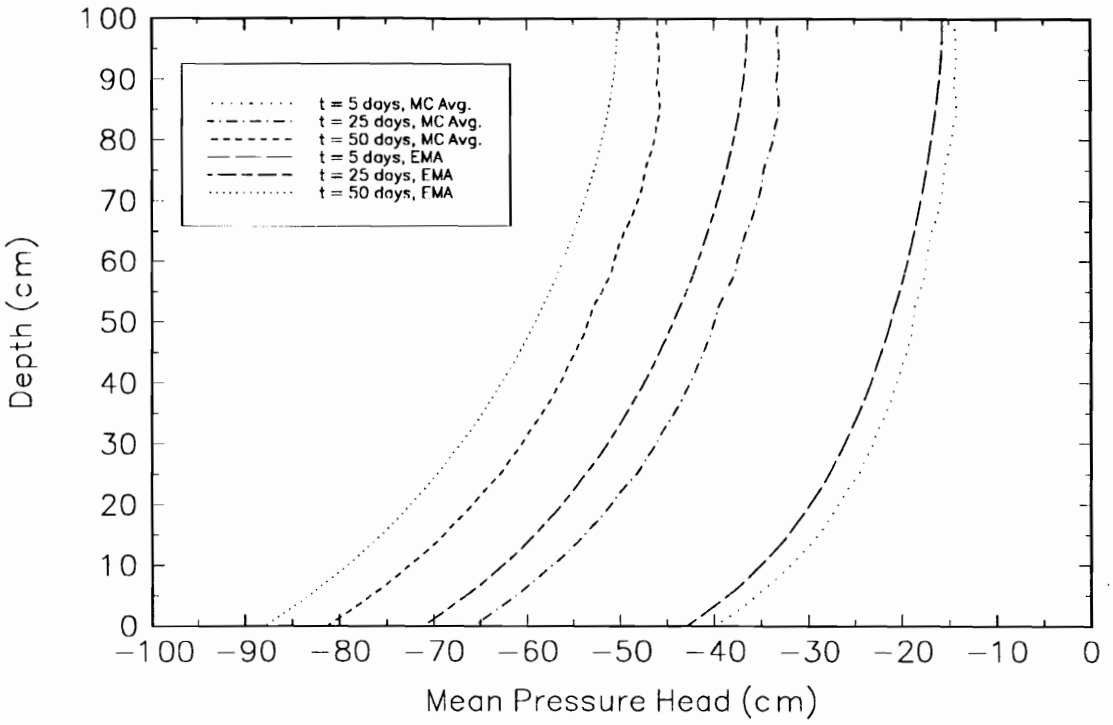

Figure 6.17 Mean Pressure Head Distribution
Comparison Between EMA & MC Simulation, 100 cm Deep Column

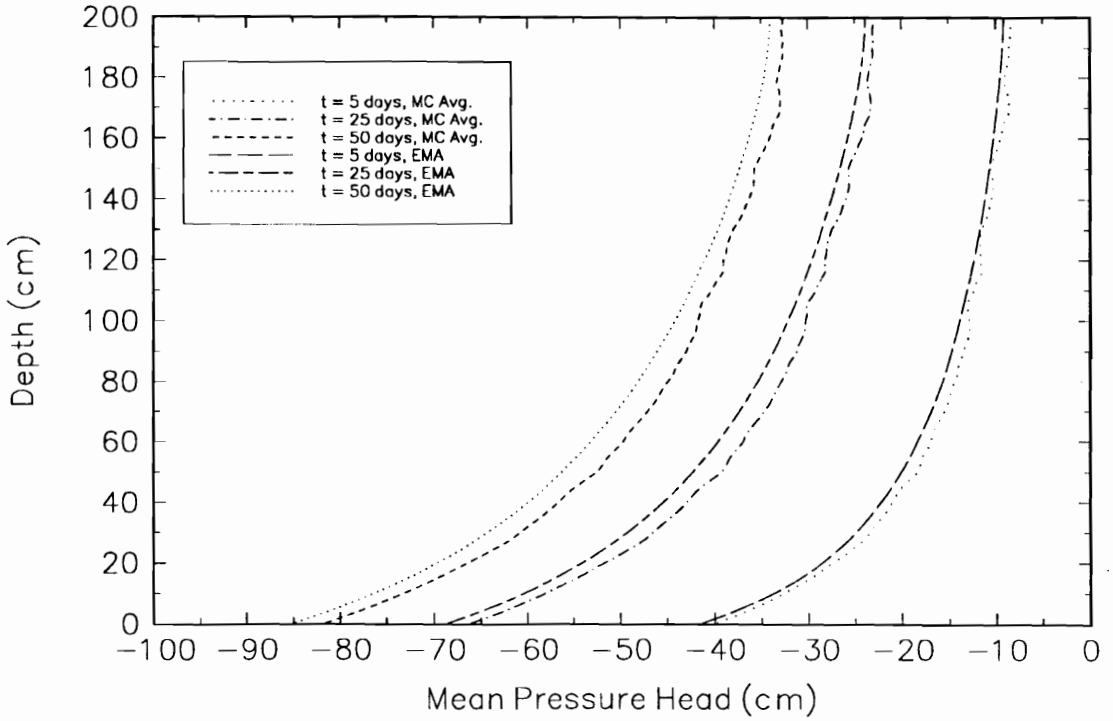

Figure 6.18 Mean Pressure Head Distribution
Comparison Between EMA & MC Simulation, 200 cm Deep Column

noted that the assumption of equality between the variance of log unsaturated and log saturated hydraulic conductivities was assumed in deriving Eq. (3.35). Nevertheless, these variances are within one standard deviation.

6.6 Well Flow Problem

The last problem consists of a pumping well in an unconfined aquifer. The height of the phreatic surface, h, located at a radial distance r as given in chapter 5 (Eq. (5.6)) is

$$h \approx H_0 - \left[\frac{Q}{2\pi H_0} \ln\left(\frac{R}{r}\right)\right] \frac{1}{K_s} \tag{6.18}$$

where H_0 is the initial height of the phreatic surface, R is the radius of influence, Q is the pumping rate, and K_s is the saturated hydraulic conductivity of the aquifer medium. Variance of h is then computed as (see Chapter 5 for derivation)

$$Var[h] = \left[\frac{Q}{2\pi H_0} \ln\left(\frac{R}{r}\right)\right]^2 \left(\frac{Var[K_s]}{(E[K_s])^4}\right). \tag{6.19}$$

For the example application, a constant pumping rate of 100 m^3/hr, radii of influence of 100-m and 500-m, and an initial phreatic surface height of 50 m are used. Then using the conductivity value for sand as reported by *Carsel and Parrish* [1988] which are

$$E[K_s] = 0.297 \text{ m/hr}$$

$$\text{Var}[K_s] = (0.156)^2 \text{ (m/hr)}^2$$

and assuming log-normally distributed K_s, the mean and the variance of lnK_s from Eqs. (6.11) and (6.12), are

$$\text{E}[lnK_s] = -1.3358$$

$$\text{Var}[ln\ K_s] = (0.4936)^2.$$

The range can be expressed by $p\sigma_h$ [*Crow et al.*, 1960]. By taking the range to be $5\sigma_h$ at a radial distance of 15 m from the well, the required sample size to achieve a confidence coefficient of 0.95 and an absolute error of 0.5-m is 200. The well flow problem is solved for the phreatic surface, for each realization of the hydraulic conductivity. The maximum range of phreatic surface levels from the MC simulation at r=15-m was found to be 5.441-m. Using this value as the range, $(h_{max}-h_{min})$ in Eq. (4.62) with an absolute error criteria of 0.5-m and a confidence coefficient of 0.95, we get the new sample size to be 220. This being close to the initial number of simulations (200), it was decided to analyze the results based upon the results of 200 simulations. Covariances between the phreatic surface heights and saturated hydraulic conductivity values were estimated using Eq. (4.49) and β^* values were computed using Eq. (5.27). The correlation coefficient required in Eq. (4.54) was computed using Eq. (4.49). The unbiased estimator, h_c, of the phreatic levels and the reduced variance were computed using Eqs. (5.26) and (4.54), respectively.

Standard deviations are computed at different radial distances from the pumping

well along with the mean head values. Next, the arithmetic (K_A), geometric (K_G), and harmonic (K_H) means of the generated hydraulic conductivity values are used, as effective conductivities, to compute the phreatic levels. These results are summarized in Tables 6.1 through 6.4. Table 6.2 also gives the estimates obtained from the variance reducing procedure. It is evident from Tables 6.1 and 6.2 that variance reduction was achieved without changing the mean values.

Traditionally, for steady state saturated flow analysis one of arithmetic, geometric, and harmonic means is taken as the effective conductivity [*Dagan*, 1989]. It is crucial to note that the computation of these means do need a sample size. Also, which of these means is the most appropriate is not known apriori. The present analysis overcomes all these drawbacks. The phreatic surface level computed from the present procedure and that from using the harmonic mean are quite close while the estimates from the arithmetic and the geometric means differ slightly. The performance of the analytical variance of the piezometric heads is good. It can be employed to estimate the confidence limits directly if a reliable average piezometric head is available, say using the harmonic mean. The advantage of using the harmonic mean is that only the random conductivities need be generated and there is no need to solve the well flow problem for each realization.

Table 6.1

Comparison of Phreatic Surfaces
(N = 200; R = 100 m)

Distance From the Well (m)	MC Average (m)	Estimate Using VRT (m)	Estimate* Using K_A (m)	Estimate* Using K_G (m)	Estimate* Using K_H (m)
0.2	41.180	41.210	42.798	42.072	41.260
0.5	42.641	42.664	43.934	43.332	42.661
1.0	43.697	43.717	44.775	44.262	43.691
2.0	44.720	44.736	45.599	45.172	44.698
4.0	45.712	45.725	46.410	46.064	45.682
7.0	46.493	46.503	47.053	46.772	46.462
10.0	46.982	46.991	47.459	47.218	46.952
15.0	47.530	47.538	47.916	47.720	47.503
25.0	48.210	48.216	48.486	48.344	48.188
40.0	48.826	48.829	49.005	48.912	48.810
60.0	49.350	49.351	49.448	49.396	49.340
80.0	49.717	49.718	49.759	49.737	49.713
100.0	50.000	50.000	50.000	50.000	50.000

* K_A - Arithmetic Mean of hydraulic conductivity
K_G - Geometric Mean of hydraulic conductivity
K_H - Harmonic Mean of hydraulic conductivity

Table 6.2

Comparison of Standard Deviations
(N = 200; R = 100 m)

Distance From the Well (m)	Analytical Standard Deviation (m)	Standard Deviation From MC Simulation (m)	Std. Dev. Using VRT (m)
0.2	3.498	4.352	2.402
0.5	2.983	3.510	1.888
1.0	2.592	2.942	1.557
2.0	2.202	2.419	1.262
4.0	1.812	1.932	0.996
7.0	1.497	1.560	0.797
10.0	1.296	1.332	0.677
15.0	1.068	1.081	0.546
25.0	0.780	0.776	0.389
40.0	0.516	0.505	0.252
60.0	0.288	0.278	0.138
80.0	0.126	0.120	0.059
100.0	0.000	0.000	0.000

Table 6.3

Comparison of Phreatic Surfaces
(N = 200; R = 500 m)

Distance From the Well (m)	Monte Carlo Average (m)	Estimate* Using K_A (m)	Estimate* Using K_G (m)	Estimate* Using K_H (m)
0.2	38.364	40.726	39.762	38.675
0.5	40.016	41.919	41.093	40.167
1.0	41.180	42.798	42.072	41.260
2.0	42.292	43.661	43.029	42.324
4.0	43.361	44.506	43.965	43.362
7.0	44.198	45.177	44.706	44.183
10.0	44.720	45.599	45.172	44.698
15.0	45.303	46.075	45.696	45.276
25.0	46.025	46.667	46.348	45.995
40.0	46.676	47.206	46.940	46.646
60.0	47.229	47.665	47.444	47.200
80.0	47.617	47.989	47.799	47.590
100.0	47.915	48.238	48.072	47.890
200.0	48.826	49.005	48.912	48.810
300.0	49.350	49.448	49.396	49.340
400.0	49.717	49.759	49.737	49.713
500.0	50.000	50.000	50.000	50.000

* K_A - Arithmetic Mean of hydraulic conductivity
K_G - Geometric Mean of hydraulic conductivity
K_H - Harmonic Mean of hydraulic conductivity

Table 6.4

Comparison of Standard Deviations
(N = 200; R = 500 m)

Distance From the Well (m)	Analytical Std. Dev. (m)	MC Standard Deviation (m)
0.2	4.404	6.296
0.5	3.889	5.086
1.0	3.498	4.352
2.0	3.108	3.704
4.0	2.718	3.120
7.0	2.403	2.683
10.0	2.202	2.419
15.0	1.974	2.130
25.0	1.686	1.781
40.0	1.422	1.474
60.0	1.194	1.219
80.0	1.032	1.042
100.0	0.906	0.908
200.0	0.516	0.505
300.0	0.288	0.278
400.0	0.126	0.120
500.0	0.000	0.000

7. Summary

In this study two approaches, namely an effective medium approximation and a refined Monte Carlo simulation procedure for solving the stochastic groundwater flow problem are presented. In the effective medium approach, an effective hydraulic conductivity which adheres to the same conventional definition of the ratio of (average) flow to (average) hydraulic gradient is employed. The method permits one to solve the stochastic groundwater flow problem in a single run to generate the expected pressure head field. A reduced form of the effective conductivity expression attains the form of the traditional Gardner equation. This field scale Gardner equation has the saturated hydraulic conductivity replaced by its effective counterpart and the capillary tension replaced by the mean pressure head. The major advantage of the relationship is that the same schemes used to solve the groundwater boundary value problem (with Richards' equation, Gardner's equation, initial, and boundary conditions) can be used to solve the mean flow equation. Example applications verify this proposed effective conductivity.

The refined MC simulation procedure presents an analytical method to estimate

the sample size. The key advantage of this estimate is that it uses the output variable, namely the pressure head behavior along with an error control and the desired confidence level. Use of the proposed approach in the one-dimensional, steady state, saturated flow has shown rather dramatic improvement in accurately duplicating the results of *Freeze* [1975] with significantly smaller number of simulation runs.

The variance reducing estimator which is different from the simple average for pressure head is derived. This estimator takes advantage of the correlation between the saturated conductivity and the pressure head distribution to reduce the variance and is unbiased. This reduced variance results in a smaller width of uncertainty about the predicted pressure head. For predictive purposes, the method not only yields the expected pressure head but also establishes a range to cover the variations at a given confidence level.

Recommendations for future work are:

(1) to apply the proposed method to real field problems and validate against field observed data.

(2) to derive a relationship for EMA relaxing the assumption of equality of variances of log unsaturated and log saturated hydraulic conductivities. This will facilitate in using the EMA relationship over a wide range of saturation.

8. Bibliography

Aitchison, J., and J. A. C. Brown, *The Lognormal Distribution*, Cambridge University Press, 1957.

Anderson, M. P., and W. W. Woessner, *Applied Groundwater Modeling: Simulation of Flow and Advective Transport*, Academic Press, Inc., New York, 1992.

Bear, J., *Hydraulics of Groundwater*, McGraw-Hill, New York, 1979.

Bear, J., *Dynamics of Fluids in Porous Media*, McGraw-Hill, New York, 1972.

Bennion, D. W., and J. C. Griffiths, A Stochastic Model for Predicting Variations in Reservoir Rock Properties, *Trans. AIME*, 237(2):9-16, 1968.

Beyer, W. H., *CRC Standard Mathematical Tables*, 27th Edition, CRC Press, Inc, Boca Raton, Florida, 1984.

Bickel, P. J., and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day, Inc., San Francisco, 1977.

Binley, A., K. Beven, and J. Elgy, A Physically Based Model of Heterogeneous Hillslopes, 2, Effective Hydraulic Conductivities, *Water Resour. Res.*, 25(6):1227-1233, 1989.

Bouwer, H., *Groundwater Hydrology*, McGraw-Hill, New York, 1978.

Bresler, E., Soil Spatial Variability, *In: J. Hanks and J. T. Ritchie (eds.), Modeling Plant and Soil Systems - Agronomy Monograph no. 31*, 1991.

Brooks, R. H., and A. T. Corey, Hydraulic Properties of Porous Media, *Hydrology Paper No. 3*, Colorado State University, Fort Collins, CO, 1964.

Carsel, R. F., and R. S. Parrish, Developing Joint Probability Distributions of Soil Water Retention Characteristics, *Water Resour. Res.*, 24(5):755-769, 1988.

Crow, E. L., F. A. Davis, and M. W. Maxfield, Statistics Manual, Dover Publications, Inc., New York, 1960.

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

Dagan, G., and E. Bresler, Unsaturated Flow in Spatially Variable Fields, 1, Derivation of Models of Infiltration and Redistribution, *Water Resour. Res.*, 19(2):413-412, 1983.

DeGroot, M. H., *Probability and Statistics*, Addison-Wesley Publishing Co., Inc., 1975.

de Marsily, G., *Quantitative Hydrogeology: Groundwater Hydrology for Engineers*, Academic Press, 1986.

de Marsily, G., *Spatial Variability of Properties in Porous Media: A Stochastic Approach*, Kluwer Publ. Co., 1984.

Dikow, E., Stochastic Analysis of Groundwater Flow in a Bounded Domain by Spectral Methods, *Transp. Porous Media*, 3:173-184, 1988.

El-Kadi, A. L., Modeling Variability in Groundwater Flow, International Ground Water Modeling Center, *Institute Paper no. 31*, Publication no. GWMI 84-10, 1984.

Fetter, C. W., *Applied Hydrogeology*, Second Edition, Macmillan, 1988.

Freeze, R. A., A Stochastic-Conceptual Analysis of One-Dimensional Groundwater Flow in Nonuniform Homogeneous Media, *Water Resour. Res.*, 11(5):725-741, 1975.

Freeze, R. A., and J. A. Cherry, *Groundwater*, Prentice-Hall, Inc., Englewoods Cliffs, New Jersey, 1979.

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.

Gelhar, L. W., Stochastic Subsurface Hydrology: From Theory to Applications, *Water Resour. Res.*, 22(9):135S-145S, 1986.

Gómez-Hernández, J. J., and S. M. Gorelick, Effective Groundwater Model Paramter Values: Influence of Spatial Variability of Hydraulic Conductivity, Leakance, and Recharge, *Water Resour. Res.*, 25(3):405-419, 1989.

Hoeffding, W., Probability Inequalities for Sums of Bounded Random Variables, *J. Am. Statist. Assn.*, 58:13-30, 1963.

Hoel, P. G., S. C. Port, and C. J. Stone, *Introduction to Stochastic Processes*, Houghton Mifflin Co., Boston, 1972.

Hogg, R. V., and A. T. Craig, *Introduction to Mathematical Statistics*, Macmillan Publishing Co., New York, 1978.

Konikow, L. F., and J. D. Bredehoeft, Ground-water Models Cannot be Validated, *Adv. Water Resour.*, 15:75-83, 1992.

Kotz, S., N. L. Johnson, and B. R. Campbell, *Encyclopedia of Statististical Sciences*, Wiley & Sons, New York, 1989.

Law, J., A Statistical Approach to the Interstitial Heterogeneity of Sand Reservoirs, *Trans. AIME*, 155:202-222, 1944.

Law, A. M., and W. D. Kelton, *Simulation Modeling and Analysis*, Second Edition, McGraw-Hill, New York, 1991.

Lumley, J. L., and H. A. Panofsky, *The Structure of Atmospheric Turbulence*, John Wiley & Sons, New York, 1964.

Ma, F., M. S. Wei, and W. H. Mills, Correlation Structuring and the Statistical Analysis of Steady-State Groundwater Flow, *SIAM J. Stat. Comput.*, 8(5):848-867, 1987.

Mantoglou, A., and L. W. Gelhar, Stochastic Modeling of Large-Scale Transient Unsaturated Flow Systems, *Water Resour. Res.*, 23(1):37-46, 1987a.

Mantoglou, A., and L. W. Gelhar, Capillary Tension Head Variance, Mean Soil Moisture Content, and Effective Specific Soil Moisture Capacity of Transient Unsaturated Flow in Stratified Soils, *Water Resour. Res.*, 23(1):47-56, 1987b.

Mantoglou, A., and L. W. Gelhar, Effective Hydraulic Conductivities of Transient Unsaturated Flow in Stratified Soils, *Water Resour. Res.*, 23(1):57-67, 1987c.

Mishra, S., J. L. Zhu, and J. C. Parker, How Effective are Effective Medium Properties?, ModelCARE 90: Calibration and Reliability in Groundwater Modelling, *IAHS Publ. no. 195*, 1990.

Nielsen, D. R., J. W. Biggar, and K. T. Erh, Spatial Variability of Field Measured Soil-Water Properties, *Hilgardia*, 42:215-259, 1973.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1965.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, Second Edition, McGraw-Hill, New York, 1984.

Russo, D., Upscaling of Hydraulic Conductivity in Partially Saturated Heterogeneous Formation, *Water Resour. Res.*, 28(2):397-409, 1992.

Shanmugan, K. S., and A. M. Breipohl, *Random Signals: Detection, Estimation and Data Analysis*, John Wiley & Sons, New York, 1987.

Smith, L., and R. A. Freeze, Stochastic Analysis of Steady State Groundwater Flow in a Bounded Domain, Part 1, One-Dimensional Simulations, *Water Resour. Res.*, 1979a.

Smith, L., and R. A. Freeze, Stochastic Analysis of Steady State Groundwater Flow in a Bounded Domain, Part 2, Two-Dimensional Simulations, *Water Resour. Res.*, 1979b.

Ünlü, K., D. R. Nielsen, and J. W. Biggar, Stochastic Analysis of Unsaturated Flow: One-Dimensional Monte Carlo Simulations and Comparisons With Spectral Perturbation Analysis and Field Observations, *Water Resour. Res.*, 26(9):2207-

2218, 1990.

Warren, J. E., and H. S. Price, Flow in Heterogeneous Porous Media, *Soc. Pet. Eng. J.*, 1:153-169, 1961.

Watt, J. P., G. F. Davies, and R. J. O'Connell, The Elastic Properties of Composite Materials, *Rev. Geophys. and Space Phys.*, 14(4):541-563, 1976

Wilson, G. V., J. M. Alfonsi, and P. M. Jardine, Spatial Variability of Saturated Hydraulic Conductivity of the Subsoil of Two Forested Watersheds, *Soil Sci. Soc. Am. J.*, 53(3):679-685, 1989.

Yeh, G. T., 3DFEMWATER: A Three-Dimensional Finite Element Model of WATER Flow Through Saturated-Unsaturated Media, Revised and Updated Version, *Personal Communications*, 1990.

Yeh, G. T., 3DFEMWATER: A Three-Dimensional Finite Element Model of WATER Flow Through Saturated-Unsaturated Media, *Publication No. 2904*, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 1987.

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(4):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(4):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(4):465-471, 1985c.

Yeh, T.-C. J., Stochastic Analysis of Effects of Spatial Variability on Unsaturated Flow, *Ph.D. Dissertation*, New Mexico Inst. of Min. and Tech., Soccoro, 1982.

Zhu, J. L., Effective Properties For Flow in Heterogeneous Porous Media, *Ph.D. Dissertation*, Virginia Polytechnic Institute and State University, Blacksburg, VA., 1990.

Appendix

This appendix presents definitions of stochastic terms used in this thesis.

Random variable: Assigns a real number for a member of the sample space made up of all realizations of a random experiment. The value of a random variable is not known with certainty prior to the outcome.

Stochastic process: A stochastic process $\{K(\boldsymbol{x}); \boldsymbol{x} \in \boldsymbol{X}\}$ is a set of random variables indexed by parameter space $\boldsymbol{X}$, i.e., for each fixed value $\boldsymbol{x} \in \boldsymbol{X}$, $K(\boldsymbol{x})$ is a random variable. If $\boldsymbol{x}$ has more than one component, $K(\boldsymbol{x})$ is called a random field.

Stationary process: The joint distribution of $K(x_1)K(x_2)$, ... for any choice of x_1, x_2, ... from X depend only on the distance between these random variables and not on the choice of x_1, x_2, ... For example $K(x_1)$, $K(x_2)$, $K(x_3)$, ... have all identical univariate probability distribution functions; $\{K(x_2), K(x_5)\}$, $\{K(x_7), K(x_{10})\}$, $\{K(x_1), K(x_4)\}$, ... all have the same bivariate distribution; $\{K(x_5), K(x_{10}), K(x_{12}), K(x_{15})\}$, $\{K(x_1), K(x_6), K(x_8), K(x_{11})\}$, ... have the same joint distribution. This aspect is guaranteed if $\{K(x_1),$

$K(x_2)$, ... } and {$K(x_1+h)$, $K(x_2+h)$, ... } have the same joint distribution function for all x_1, x_2, ... $\in X$ and for all h. This definition also yields the important characteristic that any joint statistics depend only on the distances between the random variables, e.g., $E[x_1x_7x_9]$ is the same as $E[x_{11}x_{17}x_{19}]$ and so on for $h=10$.

Weak stationarity: A stochastic process is said to be weakly stationary (also called second order stationary or stationary in the wide sense) if its expected value is a constant and its covariance function only depends on distance and not on specific points of reference. Note this characteristic is implied in the stationary process definition for any choice of two random variables.

Ensemble: The set of all possible realizations.

Ergodic hypothesis: Ergodicity implies that the ensemble average taken across all realizations of a stochastic process is the same as that of a time average of a single realization of the same stochastic process. Clearly for this to be true the mean value must not depend on at which time the ensemble average is taken or on which realization is used to form the time average, i.e., $E[X_t]=\mu=\bar{X}(\omega)$, the time average. For the process $X(t)$ to be distribution-ergodic, set $y(t)=1$ if $X(t)\leq x$ and $y(t)=0$ if $X(t)\rangle x$. Then $E[y(t)]=P[X(t)\leq x]=F(x)$ and therefore if $y(t)$ is ergodic (in the mean), $X(t)$ is ergodic in distribution [see *Papoulis*, 1984 for complete discussion].

Precision: If $T(\varsigma)$ is an estimation of the location, x, dependent parameter $K(x)$, then the mean squared error of this estimator is given by,

$$\text{MSE}\,[x,T] = E\,[T(\varsigma) - K(x)]^2 = Var\,[T(\varsigma)] + b^2(x,T)$$

where $b(x,T) = \text{E}[T(\varsigma)]-K(x)$ is known as the bias of the estimator. The variance in the above equation measures how tightly the distribution of T clusters about its origin. This aspect of T is called its precision [*Bickel and Doksum*, 1977]. For a normal distribution, this precision can be defined as the reciprocal of variance [*DeGroot*, 1975].

Vita

Surendra Prakash Shrestha was born in Nepal on August 1, 1957. He completed his High school from Laboratory school in Kathmandu. He then went to Calicut University in India to receive his Bachelor of Science degree in Civil Engineering. He obtained his Master of Science degree in Water Resources Engineering from University of Hawaii at Manoa and entered Virginia Polytechnic Institute and State University to pursue his study towards the degree of Doctor of Philosophy.