

UNCERTAINTY MODELING OF MANY CORRELATED
AND SKEWED RANDOM VARIABLES

A Thesis
Submitted to the Faculty

of

Purdue University

by

Geethanjali Panchalingam

In Partial Fulfillment of the
Requirements for the degree

of

Doctor of Philosophy

May 1992

To my Mother and in the Memory of my Father

ACKNOWLEDGEMENTS

The author would like to express her deep appreciation, gratitude, and indebtedness to her advisor, Professor Milton E. Harr, for his continuous encouragement, support, and generosity in sharing his time and knowledge.

The author also wants to thank Professor Jean-Lou Chameau for his valuable constructive criticism.

Appreciative acknowledgement is made to Professors Darrell Leap and George McCabe for their contributions to the advisory committee.

The financial support of Sandia National Laboratories and the Department of Geotechnical Engineering through research and teaching assistantships is gratefully acknowledged.

The author would like to thank the faculty members and staff for their assistance during her course of study at Purdue University. The author is very grateful to all of her friends and colleagues at Purdue for making this long journey exciting and fulfilling.

Most of all, the author would like to express grateful thanks to her parents and sister for their support, encouragement, and affection.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vii
LIST OF FIGURES	x
ABSTRACT	xii
CHAPTER 1 INTRODUCTION	1.
CHAPTER 2 REVIEW OF THE LITERATURE	3
2.1 Point Estimates of Statistical Moments	3
2.2 Rosenblueth's Point Estimate Method	5
2.2.1 Two-point Representation	5
2.2.1.1 Functions of a Single Random Variable	5
2.2.1.2 Functions of Many Random Variables	7
2.2.2 Three-point Estimates	8
2.2.3 Point Representation for a Function of Two Correlated and Skewed Random Variables	8
2.3 Lind's Alternative Point Estimate Method	10
2.4 Harr's Alternative Point Estimate Method	12
2.5 Athanasiou-Grivas and Stiefel Three-point Representation	14
CHAPTER 3 ACCOUNTING FOR SKEWNESS	20
3.1 Model I of the New Point Distribution Methodology	21
3.1.1 Discrete Approximation for a Function of Two Random Variables	21
3.1.2 Discrete Approximation for a Function of Three Random Variables	25
3.1.3 Generalization of Model I Point Distribution Approximation	30

	Page
3.2 Model II of the New Point Distribution Methodology	34
3.2.1 Discrete Approximation for a Function of Two Random Variables	34
3.2.2 Discrete Approximation for a Function of Three Random Variables	39
3.2.3 Generalization of Model II Point Distribution Approximation	41
3.2.4 Alternative Solutions for Model II Applicability	44
 CHAPTER 4 NUMERICAL EXAMPLES OF THE DEVELOPED METHODODOLOGY	 46
4.1 Numerical Examples	46
4.2 Function of Three Random Variables	46
4.2.1 Point Estimate Results for Skewed and Correlated Variables	46
4.2.1.1 Results Using Model I	47
4.2.1.2 Results Using Model II	50
4.2.2 Point Estimate Results for Independent and Skewed Variables	55
4.2.3 Point Estimate Results for Independent and Lognormal Variables	57
 CHAPTER 5 APPLICATION TO LLUVIA	 62
5.1 Flow Code LLUVIA	62
5.1.1 The Model	63
5.1.1.1 Mathematical Model	63
5.1.1.2 Parameter Models	63
5.1.2 Boundary Conditions	64
5.1.3 Solution Procedure	64
5.1.4 Parameters Required	65
5.2 Specific Problems Solved	65
5.2.1 Example I - Single Layer Problem of Silty Clay	66
5.2.1.1 Example I - Model I	67
5.2.1.2 Example I - Model II	71
5.2.1.3 Example I - Other Models	72
5.2.1.4 Example I - Two Variables Study	74
5.2.2 Example II - Two Layer Problem of Silt and Silty Clay	75
5.2.2.1 Example II - Model I	79
5.2.2.2 Example II - Model II	80
5.2.2.3 Example II - Other Models	80

	Page
5.2.3 Example III - Four Layer Problem of Silt, Clay, and Silty Clay	85.
CHAPTER 6 DISCUSSION OF THE NEW POINT ESTIMATE MODELS	89
6.1 Applicability of Models I and II	89
6.1.1 Condition of Uniqueness	89
6.1.2 Requirement of Positive Locations	90
6.2 Effect of Skewness Coefficient in Point Estimate Methodologies	91
6.2.1 Results for Skewed and Correlated Variables	91
6.2.2 Results for Independent and Skewed Variables	92
6.2.3 Results for Independent Lognormal Variables	93
6.3 Application of New Models to flow code LLUVIA	94
6.3.1 Single Layer of Silty Clay	94
6.3.2 Two Layer of Silt and Silty Clay	96
6.3.3 Four Layer of Silt, Clay, and Silty Clay	97
CHAPTER 7 SUMMARY AND CONCLUSIONS	99
CHAPTER 8 SUGGESTIONS FOR FUTURE RESEARCH	102
LIST OF REFERENCES	103
APPENDICES	
Appendix A: Details of the Methodology	110
Appendix B: Models I and II Point Distribution Approximation	111
Appendix C: Probabilistic Calculations	118
Appendix D: Computer Program	123
VITA	144

LIST OF TABLES

Table	Page
2.1 Lind's Point Representation Coordinates - Two Variables.	12
3.1 Model I Point Estimate Locations for Three Random Variables.	27
3.2 Model I Number of Point Representation Unknowns.	31
3.3 Model II Point Estimate Locations for Two Random Variables.	34
3.4 Model II Point Estimate Locations for Three Random Variables.	39
4.1 Statistical Data of Input Random Variables.	47
4.2 Model I Point Estimate Locations for Skewed and Correlated Random Variables.	48
4.3 Model I Results for Point Estimates of Function $y = x_1 x_2/x_3$ - Skewed and Correlated Random Variables.	49
4.4 Model II Point Estimate Locations for Skewed and Correlated Random Variables.	54
4.5 Model II Results for Point Estimates of Function $y = x_1 x_2/x_3$ - Skewed and Correlated Random Variables.	54
4.6 Point Estimates for Statistical Moments of Function $y = x_1 x_2/x_3$ - Skewed and Correlated Random Variables.	55
4.7 Point Estimates for Statistical Moments of Function $y = x_1 x_2/x_3$ - Independent and Skewed Random Variables.	56
4.8 Results of the Monte Carlo Simulation Technique for Function $y = x_1 x_2/x_3$ - Beta Variables.	57

Table	Page
4.9 Statistical Parameters of Input Random Variables with Coefficient of Variation 60 % - Independent Lognormal Variates.	58
4.10 Results for the Expected value and Standard deviation of Function $y = x_1 x_2/x_3$ - Independent Lognormal Variates.	59
5.1 Representative Values of the Random Variables for Silty Clay.	66
5.2 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example I.	73
5.3 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example I (Two Variables Study).	75
5.4 Representative Values of the Random Variables for Silt.	79
5.5 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example II.	84
5.6 Representative Values of the Random Variables for Clay.	85
5.7 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example III.	88
Appendix	
Table	
C.1 Sets of Values using Model I - Example I.	118
C.2 Ground Water Travel Times Computed using Sets of Values given in Table C.1.	119
C.3 Sets of Values using Model II - Example I.	120
C.4 Possible Sets of Values using Model II - Example I.	120
C.5 Groundwater Travel Time Computed using Sets of Values given in Table C.4.	121
C.6 Sets of Values using Rosenblueth (1975) Methodology - Example I.	121
C.7 Sets of Values using Lind (1983) Methodology - Example I.	122

Appendix Table	Page
C.8 Sets of Values using Harr (1989) Methodology - Example I.	122

LIST OF FIGURES

Figure	Page
2.1 Two-point Representation.	6
2.2 Point Representation for a Function of Two Correlated and Skewed Random Variables.	9
2.3 Lind's Point Representation - Two Variables.	11
2.4 Three-point Representation.	15
2.5 Multiple-point Representation, Athanasίου-Grivas and Stiefel (1983).	17
2.6 Alternate Multiple-point Representation.	19
3.1 Model I Point Distribution Approximation - Two Variables.	22
3.2 Model I Point Distribution Approximation - Three Variables.	26
3.3 Model II Point Distribution Approximation - Two Variables.	35
3.4 Model II Point Distribution Approximation - Three Variables.	40
4.1 Comparison Between Point Estimate Method Solutions and Lognormal Solution for Expected Value of Function $y = x_1 x_2/x_3$ - Independent Lognormal Variates.	60
4.2 Comparison Between Point Estimate Method Solutions and Lognormal Solution for Standard deviation of Function $y = x_1 x_2/x_3$ - Independent Lognormal Variates.	61

Figure	Page
5.1 Steady-State Pressure Head Profile when Flux rate = 0.1 mm/year, Example I - Model I.	68
5.2 Steady-State Degree of Saturation Profile when Flux rate = 0.1 mm/year, Example I - Model I.	69
5.3 Steady-State Hydraulic Conductivity Profile when Flux rate = 0.1 mm/year, Example I - Model I.	70
5.4 Steady-State Pressure Head Profile when Flux rate = 0.1 mm/year, Example I (Two variables study) - Model I.	76
5.5 Steady-State Degree of Saturation Profile when Flux rate = 0.1 mm/year, Example I (Two variables study) - Model I.	77
5.6 Steady-State Hydraulic Conductivity Profile when Flux rate=0.1 mm/year Example I (Two variables study) - Model I.	78
5.7 Steady-State Pressure Head Profile when Flux rate = 1.0 mm/year, Example II - Model I.	81
5.8 Steady-State Degree of Saturation Profile when Flux rate = 1.0 mm/year, Example II - Model I.	82
5.9 Steady-State Hydraulic Conductivity Profile when Flux rate=1.0 mm/year Example II - Model I.	83
5.10 One-Dimensional Flow domain of Example III.	86

ABSTRACT

Panchalingam, Geethanjali. Ph.D., Purdue University, May 1992. Uncertainty Modeling of Many Correlated and Skewed Random Variables. Major Professor: Milton E. Harr.

This thesis presents a point estimate methodology directly suited to uncertainty modeling for systems with many correlated and skewed random variables. The methodology proposes two models. The models provide a simple but powerful procedure to approximate the joint probability density function of many correlated and skewed random variables. It is assumed that the only available information is the set of expected values, standard deviations, skewness coefficients, and correlation coefficients of the random variables.

Examples and step-by-step procedures are provided to illustrate the models. Comparisons are made with the results of Rosenblueth (1975), Lind (1983), and Harr (1989) point estimate methodologies, the Monte Carlo simulation technique, and also with the exact solution for a special case. Application is also made to the flow code LLUVIA for estimating the groundwater travel time in an unsaturated medium.

CHAPTER 1

INTRODUCTION

One of the major problems in modeling civil engineering systems today is the choice of values to use as system input variables. Add to this that for most situations, decisions have to be made under conditions of uncertainty. Values of input parameters are not single-valued and, consequently, the reliability of civil engineering designs must be framed in probabilistic terms. Available information is a significant factor in deciding the method of reliability analysis. In selecting the method of analysis, good engineering design necessitates that all reasonably possible scenarios be accounted for; but, on the other hand, no more information should be assumed than is available and reliable.

Today there is a constantly increasing need to solve problems, such as problems of resource management or environmental protection, for which there is little information. The complexity of the system functions and inadequate statistical data have given rise to a need for reliable procedures, which are capable of accounting for uncertainties resulting from inadequate information. Assumptions and approximations must be introduced in a "least-biased" manner.

Several probabilistic methods have been developed in the past to obtain measures of the reliability of engineering systems. These range from exact analytical procedures to approximate methodologies that can be accommodated by relatively simple algebraic computations. While each method has its own set of assumptions and advocates, the adoption of a reliable method of analysis depends on the amount of information available

concerning the uncertainty of the system variables, the nature of the system functions, and the desired level of accuracy.

In practice, it is not uncommon to find systems which have many skewed and correlated random variables. The combination of very many variables and of highly complex system functions have given rise to point-estimate methods (Rosenblueth (1975), Lind (1983), Harr (1989)). However, current point-estimate methods are restricted to systems of symmetrical and correlated random variables. The present work seeks to provide a methodology that will account for correlated and skewed random variables in uncertainty analyses. It is assumed that the only information available is the set of the first three statistical moments: expected values, standard deviations, and skewness coefficients; and the correlation structure of the random variables.

CHAPTER 2

REVIEW OF THE LITERATURE

2.1 Point Estimates of Statistical Moments

The fundamental concept of the point estimate method is the approximation of the probability density functions of random variables at selected points of the space defined by the random variables. The location of the points are determined so as to satisfy the requirement that the discrete approximations have the same statistical moments (up to a specified order) as the probability density functions.

Rosenblueth (1975) first introduced the point estimate method by requiring the equivalency of the first three statistical moments: mean, standard deviation, and skewness of a random variable and its point-wise approximation. The procedure was then generalized to functions of several random variables. In concept, Rosenblueth point distributions are at the corners of a hypercube defined in the variable space (in reality for non-equal variables the figure is a "hyperprism"). The number of point estimates of Rosenblueth's method is 2^N where N is the number of random variables. As an illustration, for a function of three random variables ($N=3$), Rosenblueth point distributions are at the corners of a 3-dimensional cube, and hence, the number of point estimates is $2^3=8$, which is equivalent to the number of corners of a 3-dimensional cube.

- Higher-order approximations were also derived by Rosenblueth (1975) for functions of a single random variable whose first four statistical moments: mean, standard deviation, skewness, and kurtosis were approximated as a gaussian random

variable.

In 1981, Rosenblueth derived a point distribution approximation for a function of two random variables by satisfying the mean, standard deviation, skewness, and correlation coefficient of these two random variables. However, this procedure was not generalized to functions of several skewed and correlated random variables.

Lind (1983) developed a methodology called the "face-center" point distribution. This methodology consists of $2N$ equal probability masses or weighting functions located "near the center of each face of the hypercube whose corners are the points of Rosenblueth's distribution." Lind's methodology approximates the first two moments: means and covariances of the random variables.

Harr (1989) developed an alternative point estimate methodology in which the estimates are obtained in the eigenspace of the correlation matrix. This methodology also requires only $2N$ point estimates.

Athanasiou-Grivas and Stiefel (1983) proposed a three-point representation for a beta variate. The representation approximates the first four statistical moments: mean, standard deviation, skewness, and kurtosis of a beta variate. An m -point representation was also suggested by Athanasiou-Grivas and Stiefel for a function of a single random variable.

In a private communication, Harr suggested an m -point representation to approximate the probability density function of a single random variable. This m -point representation involves the approximation of a definite integral as the sum of a number of terms using an algorithm for numerical integration.

2.2 Rosenblueth's Point Estimate Method

2.2.1 Two-point Representation

Rosenblueth (1975) developed a simple two-point representation to approximate a probability density function of a random variable, $f(x)$, up to its third-order statistical moment. He also provided a procedure to consider a joint probability function of many random variables, $f(x_1, x_2, \dots, x_N)$, using a second-order approximation. The generalization of the two-point procedure assumed the skewness of the random variables to be zero.

2.2.1.1 Functions of a Single Random Variable

In Figure 2.1 is shown schematically the two-point representation of $f(x)$ as x_+ and x_- , and the corresponding mass densities or weighting functions as p_+ and p_- , respectively. Rosenblueth solved for the four unknowns (i.e., x_+ , x_- , p_+ , p_-) to meet the following four conditions:

$$p_+ + p_- = 1 \quad (2.1)$$

$$p_+ x_+ + p_- x_- = \bar{x} \quad (2.2)$$

$$p_+ (x_+ - \bar{x})^2 + p_- (x_- - \bar{x})^2 = \sigma_x^2 \quad (2.3)$$

$$p_+ (x_+ - \bar{x})^3 + p_- (x_- - \bar{x})^3 = \beta(1) \sigma_x^3 \quad (2.4)$$

where x is the random variable with mean \bar{x} , standard deviation σ_x , and skewness coefficient $\beta(1)$. From equations (2.1) through (2.4), it is seen that for the case of a two-point approximation of $f(x)$, the determination of the four unknown quantities requires knowledge of the statistical moments of x up to the third order. The solution of the above system of simultaneous equations is

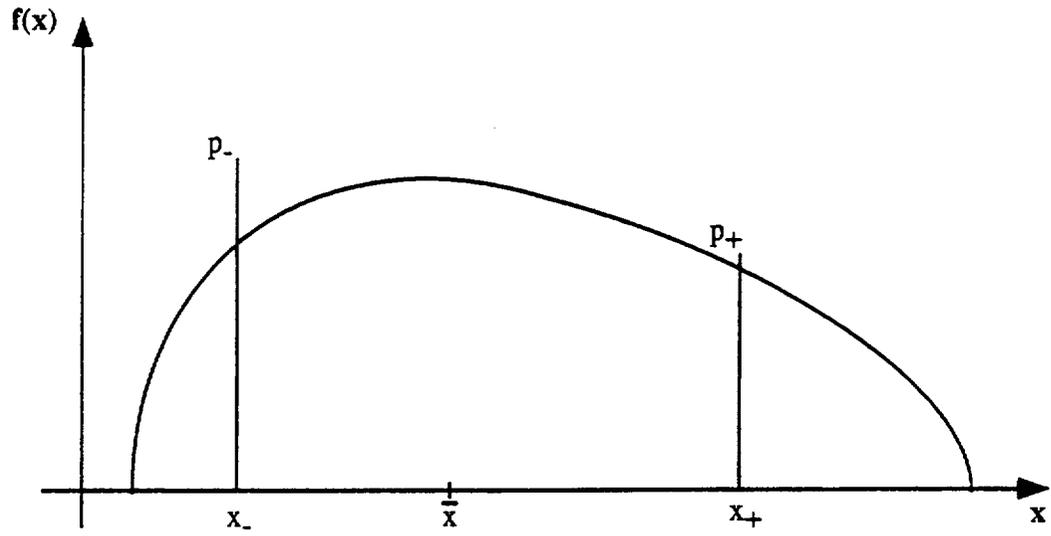


Figure 2.1 Two-point Representation.

$$p_{\pm} = \frac{1}{2} \left\{ 1 \pm \sqrt{1 - \frac{1}{1 + \left[\frac{\beta(1)}{2} \right]^2}} \right\} \quad (2.5)$$

$$p_{-} = 1 - p_{+} \quad (2.6)$$

$$x_{+} = \bar{x} + \sigma_x \sqrt{\frac{p_{-}}{p_{+}}} \quad (2.7)$$

$$x_{-} = \bar{x} - \sigma_x \sqrt{\frac{p_{+}}{p_{-}}} \quad (2.8)$$

In equation (2.5) the sign preceding the radical is that of $-\beta(1)$.

In the special case where the probability density function, $f(x)$, is symmetrical, ($\beta(1)=0$), and equations (2.5) through (2.8) take the simple form:

$$p_{+} = p_{-} = \frac{1}{2} \quad (2.9)$$

$$x_{+} = \bar{x} + \sigma_x \quad (2.10)$$

$$x_{-} = \bar{x} - \sigma_x \quad (2.11)$$

2.2.1.2 Functions of Many Random Variables

For a joint probability density function of N random variables, $f(x_1, x_2, \dots, x_N)$, Rosenblueth's distribution is composed of 2^N probability masses or weighting functions, $p_{\pm, \pm, \dots, \pm}$, where all the possible permutations of N '+'s and '-'s are considered. The correlations between pairs of the random variables are accounted for in the weighting functions. These weighting functions are concentrated at the corners of a multi-dimensional hyperprism of N -dimensional space defined by the random variables with

coordinates $x^i \equiv (\bar{x}_1 \pm \sigma_{x_1}, \bar{x}_2 \pm \sigma_{x_2}, \dots, \bar{x}_N \pm \sigma_{x_N})$, $i \equiv 1, 2, \dots, 2^N$. The magnitude of the weighting functions at the coordinates x^i are given by:

$$p_{\pm, \pm, \dots, \pm} = 2^{-N} \left\{ 1 + (-1)^{d_1} (-1)^{d_2} \rho_{12} + (-1)^{d_1} (-1)^{d_3} \rho_{13} + \dots + (-1)^{d_{N-1}} (-1)^{d_N} \rho_{N-1N} \right\}$$

where ρ_{ij} is the correlation coefficient between x_i and x_j ; and d_i is the i^{th} binary digit of the weighting function $p_{\pm, \pm, \dots, \pm}$, $d_i \in \{0, 1\}$. The binary digit d_i takes the value according to the i^{th} sign value of the weighting function $p_{\pm, \pm, \dots, \pm}$, $\{d_i = 0, \text{ if the sign is positive (+); } d_i = 1, \text{ if the sign is negative (-)}\}$.

2.2.2 Three-point Estimates

Rosenblueth (1975) derived a three-point representation to approximate the first four moments of a gaussian random variable. For the three-point estimate, the weighting functions p_+ , p_0 , and p_- are taken to be respectively at x_+ , \bar{x} , and x_- . The solution is obtained following the procedure above for the two-point representation (see section 2.2.1.1 for a single random variable). As the gaussian distribution is symmetrical and the kurtosis coefficient of the gaussian distribution is equal to three, the weighting functions of the three-point representation and the point estimate locations are as follows: $p_+ = p_- = 1/6$, $p_0 = 2/3$, $x_+ = \bar{x} + \sqrt{3}\sigma_x$, and $x_- = \bar{x} - \sqrt{3}\sigma_x$.

2.2.3 Point Representation for a Function of Two Correlated and Skewed Random Variables

Rosenblueth (1981) presented a point representation for a function of two correlated and skewed random variables. The representation is as shown in Figure 2.2. The coordinates of the rectangle, as shown in Figure 2.2, are determined so as to satisfy the first three moments: expected value (\bar{x}_i), standard deviation (σ_{x_i}), and skewness ($\beta_{x_i}(1)$) of the random variables x_1 and x_2 . The weighting functions at the corners of the rectangle are defined to satisfy the specified correlation coefficient ρ_{12} between x_1 and x_2 .

The point representation coordinates as given in Figure 2.2 are as follows: $x_{1+} =$

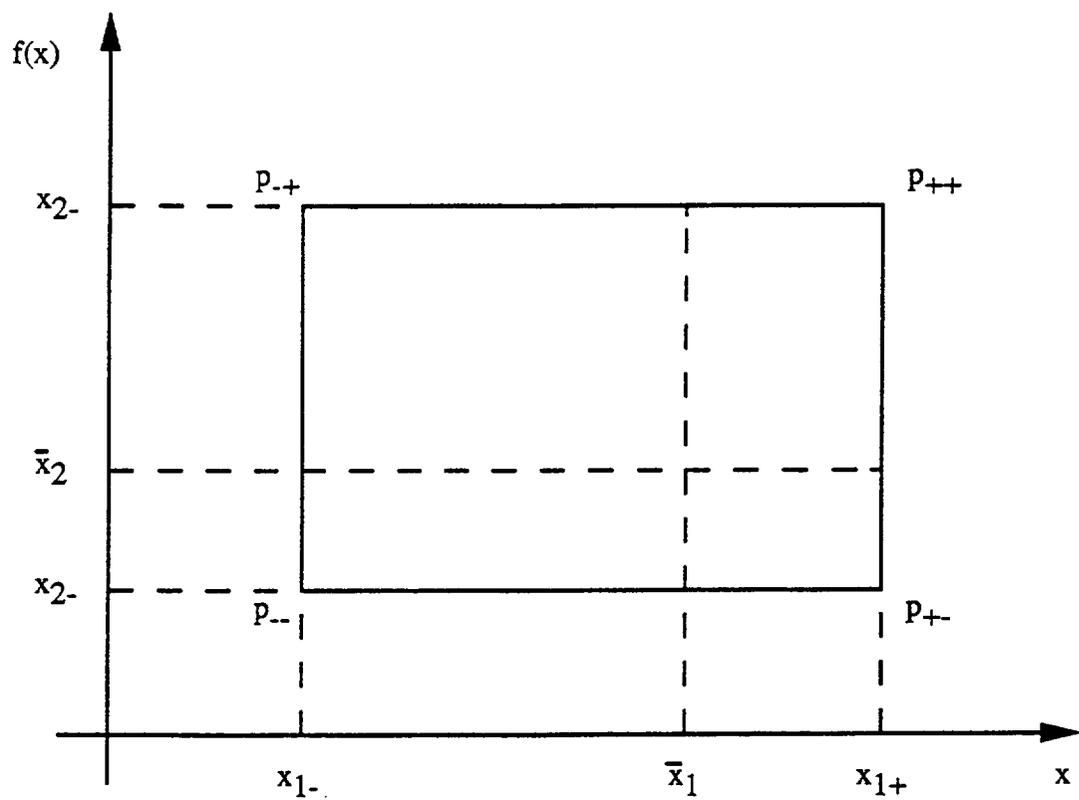


Figure 2.2 Point Representation for a Function of Two Correlated and Skewed Random Variables.

$\bar{x}_1 + (p_{1-}/p_{1+})^{0.5} \sigma_{x1}$, $x_{1-} = \bar{x}_1 - (p_{1+}/p_{1-})^{0.5} \sigma_{x1}$, $x_{2+} = \bar{x}_2 + (p_{2-}/p_{2+})^{0.5} \sigma_{x2}$, and $x_{2-} = \bar{x}_2 - (p_{2+}/p_{2-})^{0.5} \sigma_{x2}$; and the weighting functions are defined as follows: $p_{++} = p_{1+}p_{2+} + a$, $p_{+-} = p_{1+}p_{2-} - a$, $p_{-+} = p_{1-}p_{2+} - a$, and $p_{--} = p_{1-}p_{2-} + a$, where p_{i+} and p_{i-} ($i=1,2$) are as defined in the equations (2.5) and (2.6); and $a = \rho_{12} / \{ [4 + \beta_{x1}(1)^2] [4 + \beta_{x2}(1)^2] \}^{0.5}$.

2.3 Lind's Alternative Point Estimate Method

Lind (1983) introduced an alternative point estimate method with $2N$ point estimates. Lind's methodology is referred to as a "face-center" point representation. The representation is a simple point-symmetric distribution about the expected values of the random variables. The distribution approximates the first two statistical moments, mean and covariances, of the random variables. As with the point-symmetric distribution, the variables are assumed to be symmetrical. The distribution consists of $2N$ equal probability masses or weighting functions located at $x_i = \bar{x} \pm z_i$, $i = 1, 2, \dots, N$, with :

$$\begin{aligned} \bar{x} &= [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N] & (2.12) \\ z_1 &= [z_{11}, z_{12}, z_{13}, \dots, z_{1N}] \\ z_2 &= [0, z_{22}, z_{23}, \dots, z_{2N}] \\ z_3 &= [0, 0, z_{33}, \dots, z_{3N}] \\ &\vdots \\ &\vdots \\ z_N &= [0, \dots, 0, z_{NN}] \end{aligned}$$

As an illustration, Lind's point distribution for a function of two random variables is located at the points shown in Figure 2.3. The coordinates of these points are given in Table 2.1.

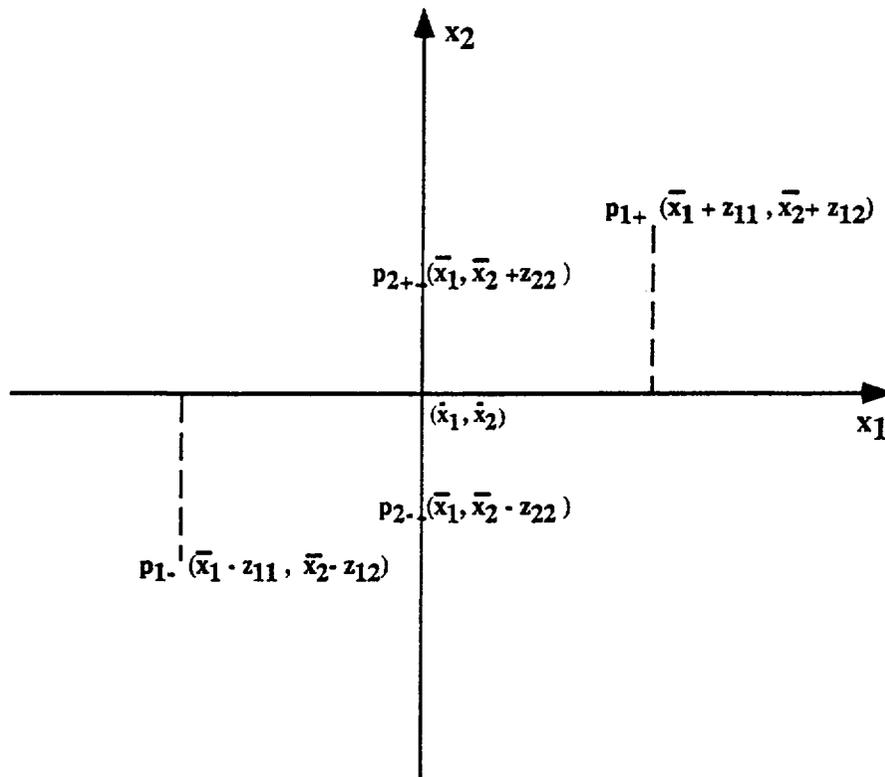


Figure 2.3 Lind's Point Representation - Two Variables.

Table 2.1 Lind's Point Representation Coordinates - Two Variables.

variable x_1	variable x_2	weighting function
$\bar{x}_1 + z_{11}$	$\bar{x}_2 + z_{12}$	0.25
$\bar{x}_1 - z_{11}$	$\bar{x}_2 - z_{12}$	0.25
\bar{x}_1	$\bar{x}_2 + z_{22}$	0.25
\bar{x}_1	$\bar{x}_2 - z_{22}$	0.25

The location of the weighting functions can be considered as N point-distribution pairs, $[(\bar{x} + z_i), (\bar{x} - z_i)]$, $i=1,2, \dots, N$. Each pair, $[(\bar{x} + z_i), (\bar{x} - z_i)]$, represents the point-representation coordinates which are placed symmetrically about the mean value of the random variables. The coordinates x_i of Lind's point-representation are determined sequentially. An algorithm is given in Appendix A.1.

In Lind's methodology it is assumed that the weighting functions are equally distributed among the N point-distribution pairs. Also, the coordinates of the point distribution depend upon the sequence of the variables in the \bar{x} vector as given in equation (2.12). As a result, if there are N random variables, it is possible to choose $N!$ sequences for the \bar{x} matrix, and consequently, $N!$ point distribution models are possible. For example, in the case of a function of two random variables (x_1 and x_2), it is possible to select $2! = 2$ sequences (i.e., sequence $[x_1, x_2]$ or $[x_2, x_1]$) to solve for the point distribution coordinates. For 3 random variables, 6 sequences are possible.

2.4 Harr's Alternative Point Estimate Method

Harr's (1989) alternative point estimate methodology stems from "principal component analysis". His methodology starts from the correlation matrix of the random variables. The use of a correlation matrix rather than a covariance matrix negates the problem of the arbitrariness of units of variables by standardizing the variables by the linear transformation:

$$ds_i = \frac{x_i - \bar{x}_i}{\sigma_{x_i}} \quad (2.13)$$

where x_i is the random variable with mean \bar{x}_i and standard deviation σ_{x_i} , and ds_i is the standardized random variable with mean as zero and standard deviation of unity. As in principal component analysis, the standardized random variables ds_i are then transformed into new random variables $[u_i, i=1,2, \dots, N]$ in the eigenspace of the correlation matrix. The variables u_i are linear combinations of the standardized variables ds_i (see Appendix A.2).

In Harr's methodology, the point estimates are first obtained in the direction of the eigenvectors of the correlation matrix. The first two statistical moments of the new random variables u_i are known in the eigenspace - the expected value (first-order statistical moment) of the new random variables is zero and the eigenvalues ($\lambda_i, i=1,2, \dots, N$) of the correlation matrix represent the variance (second-order statistical moment) of the new random variables. In the eigenspace, the variables u_i are independent of each other; hence, the correlations vanish in these directions. The independence among variables makes the $2N$ point-representation easily obtainable. The point-representations are then given by two points along each eigenvector which are placed symmetrically about the expected value \bar{u}_i of the new random variables u_i ($i=1,2, \dots, N$). In the eigensystem, the correlation matrix is represented by a hypersphere of radius \sqrt{N} centered at the expected values \bar{u}_i of the new random variables. The hypersphere follows from the weighting function along each eigenvector direction being a function of the eigenvalue λ_i in that direction. The weighting function is given by λ_i/N , where N is the number of random variables.

The coordinates of the point representation in the eigensystem are then transformed back into the space of the original random variables x_i . The point representations satisfy the first two statistical moments of the random variables and the

correlation structure. The skewness of the random variables are taken to be zero.

2.5 Athanasiou-Grivas and Stiefel Three-point Representation

Athanasiou-Grivas and Stiefel (1983) proposed a three-point representation to account for the first four statistical moments of a random variable's probability density function. This representation is shown schematically for a single random variable in Figure 2.4; the three discrete values for x are denoted as x_+ , \bar{x} , and x_- , and the corresponding mass densities as p_+ , p_0 , and p_- , respectively. The location of the point estimate associated with the probability p_0 is taken to coincide with the mean value \bar{x} .

Using the equations (2.14) through (2.18), Athanasiou-Grivas and Stiefel solved for the five unknowns, x_+ , x_- , p_+ , p_0 and p_- , of the three-point representation. Equation (2.14) satisfies the requirement that the sum of the probabilities associated with the discrete values of x (i.e., x_+ , \bar{x} , and x_-) must be equal to unity. The remaining four equations are obtained by equating the point estimates to the mean value and the second, third, and fourth central moments of the variable x .

$$p_+ + p_0 + p_- = 1 \quad (2.14)$$

$$p_+ x_+ + p_0 \bar{x} + p_- x_- = \bar{x} \quad (2.15)$$

$$p_+ (x_+ - \bar{x})^2 + p_- (x_- - \bar{x})^2 = \sigma_x^2 \quad (2.16)$$

$$p_+ (x_+ - \bar{x})^3 + p_- (x_- - \bar{x})^3 = \beta(1) \sigma_x^3 \quad (2.17)$$

$$p_+ (x_+ - \bar{x})^4 + p_- (x_- - \bar{x})^4 = \beta(2) \sigma_x^4 \quad (2.18)$$

where x is the random variable with mean \bar{x} , standard deviation σ_x , skewness coefficient $\beta(1)$, and kurtosis coefficient $\beta(2)$.

The solution of the system of equations (2.14) through (2.18) given by Athanasiou-Grivas and Stiefel is as follows:

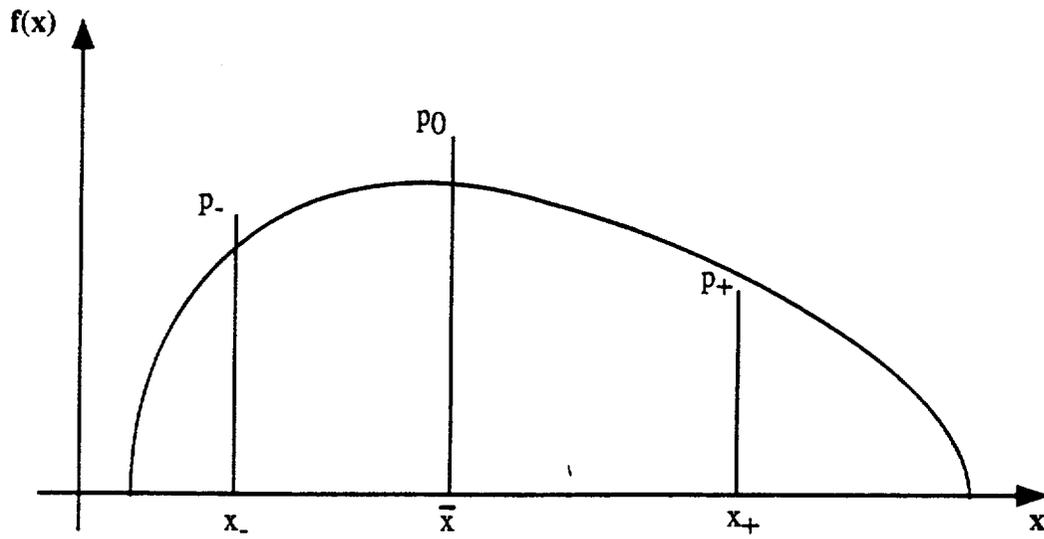


Figure 2.4 Three-point Representation.

$$p_- = \frac{2}{\beta(2) - \beta(1)^2} \left\{ 4 + \frac{\beta(1)^2}{\beta(2) - \beta(1)^2} \pm \sqrt{\frac{4\beta(1)^2}{\beta(2) - \beta(1)^2} + \left[\frac{\beta(1)^2}{\beta(2) - \beta(1)^2} \right]^2} \right\}^{-1} \quad (2.19)$$

$$p_+ = \frac{1}{\beta(2) - \beta(1)^2} - p_- \quad (2.20)$$

$$p_o = 1 - \frac{1}{\beta(2) - \beta(1)^2} \quad (2.21)$$

$$x_- = \bar{x} - \frac{p_+}{p_-} (x_+ - \bar{x}) \quad (2.22)$$

$$x_+ = \bar{x} + \sigma_x \left\{ p_+ \left[1 + \frac{p_+}{p_-} \right] \right\}^{-\frac{1}{2}} \quad (2.23)$$

In equation (2.19), the sign preceding the radical is that of $-\beta(1)$.

Athanasiou-Grivas and Stiefel (1983) also suggested an m-point representation for a random variable's probability density function. This representation is shown schematically in Figure 2.5. It specifies discrete values of x , located at multiple standard deviations away from the mean. The order of the statistical moments required in this case for the determination of weighting functions $p_i = p(x_i)$ at x_i , $i=1,2,\dots,m$, is $m-1$. Reliable higher order statistical moments are seldom attainable; nevertheless, this type of multiple point representation may be used for certain engineering applications.

In a private communication, Harr suggested an alternative m-point representation, called a revised point estimate method. It involves the approximation of a definite integral as the sum of a number of terms of the form as follows:

$$E[y(x)^q] = \int_a^b y(x)^q f(x) dx \approx \sum_{i=1}^m y_i^q w_i f_i \quad (2.24)$$

where $y(x)$ is the given function; $f(x)$ is the probability density function of a random

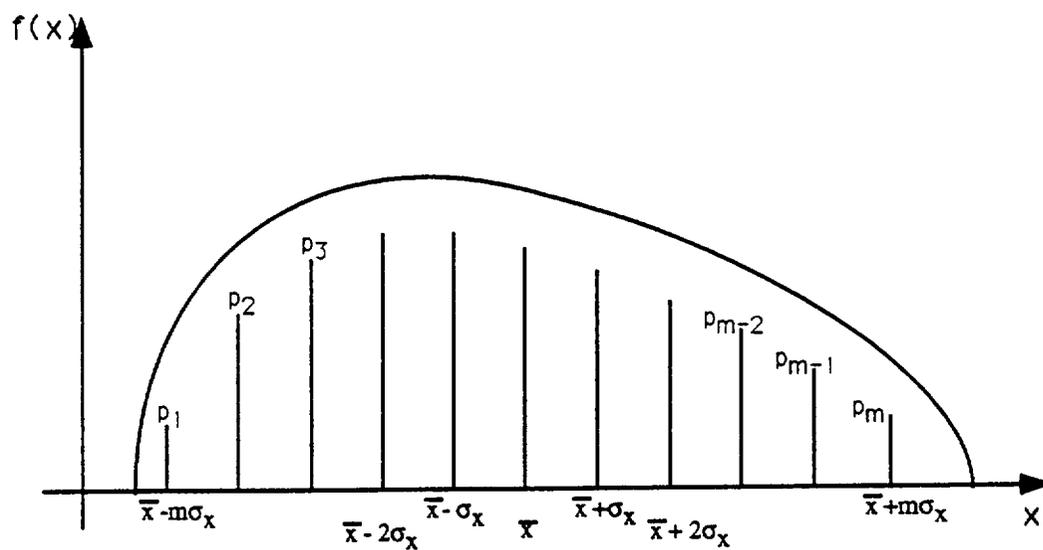


Figure 2.5 Multiple-point Representation,
Athanasίου-Grivas and Stiefel (1983).

variable x , defined over the interval (a,b) ; y_i 's are the values of the function defined at x_i 's (i.e., $y_i = y(x_i)$, $i=1,2, \dots, m$); w_i 's are the values corresponding to an algorithm for approximating a definite integral; and f_i 's are the ordinates of the probability density function at the points of subdivision as shown in Figure 2.6.

Among the many algorithms for numerical integration used in engineering applications, the most frequent are the trapezoidal rule and Simpson's rule. The probability density function, $f(x)$, is selected based on the amount of information available (Table 2.5.1, Harr 1987).

Athanasiou-Grivas and Stiefel (1983) also discussed the applicability of three-point and multiple-point estimate methods for functions of many random variables. The generalization of point-estimate methods for functions of many random variables involves a large number of non-linear simultaneous equations: these are necessary to solve the point representations of the random variables' joint probability density function. The unknowns of the simultaneous equations are the coordinates of the point representation and the weighting functions of the joint probability density function. The coordinates for functions of N -random variables are defined in N -dimensional space, and the weighting functions act at these coordinates. The number of simultaneous equations required to solve these unknowns of the point representations increase very rapidly with the number of variables. Difficulties multiply even more when the variables are correlated. Seldom are sufficient data available to obtain the statistical moments of the joint probability density function.

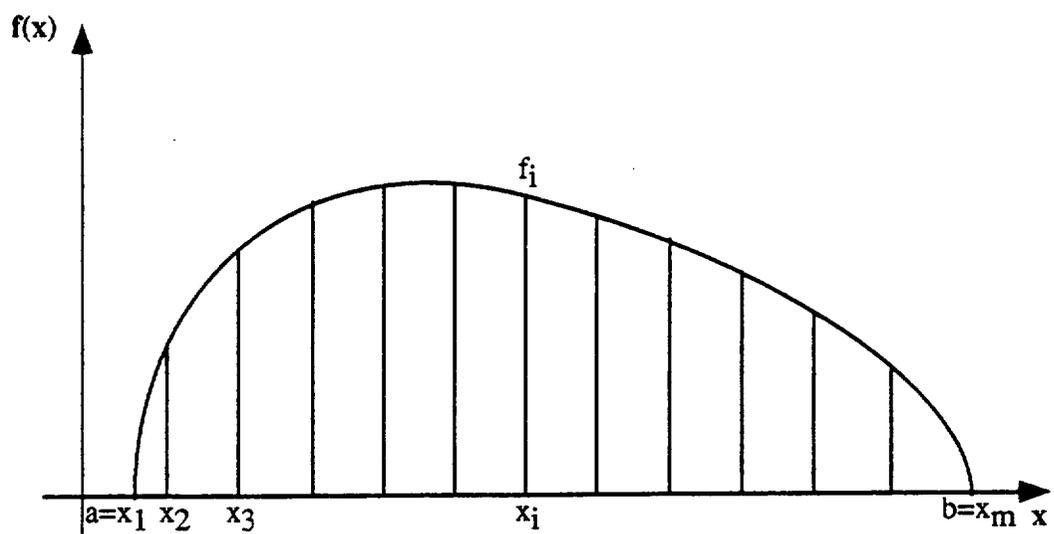


Figure 2.6 Alternate Multiple-point Representation.

CHAPTER 3

ACCOUNTING FOR SKEWNESS

A new point distribution methodology is presented in this chapter that incorporates both correlated and skewed random variables into uncertainty analyses. The methodology takes into consideration the first three statistical moments: expected value, standard deviation, and skewness coefficient of the random variables, and the correlation structure of the variables. As was noted previously in Chapter 2, prior point distribution methodologies were predicated for functions of unskewed random variables.

Two point distribution models (I and II) will be proposed to incorporate skewed and correlated random variables into uncertainty analyses. Model I is an extension of the point distribution pattern used by Rosenblueth (1975) and it requires 2^N point estimates. In model I, the point estimate locations are first determined so as to incorporate the skewness coefficient of the random variables and the required correlation structure of the variables are incorporated in the respective weighting functions associated with the point estimate locations. Model II uses an extension of the point distribution pattern proposed by Lind (1983); consequently, the second model requires only $2N$ point estimates. The algorithm of the second model defines both the point estimate locations and the respective weighting functions as functions of skewness coefficients and the required correlation structure of the random variables.

The two models will be first illustrated for functions of two and three random variables. Then a general solution of these models will be proposed.

3.1 Model I of the New Point Distribution Methodology

3.1.1 Discrete Approximation for a Function of Two Random Variables¹

Model I of the new point distribution methodology will be illustrated using Figure 3.1 for a function $y=y(x_1, x_2)$ of two random variables x_1 and x_2 . The statistical moments: expected value, standard deviation, skewness coefficient, and the correlation structure of the random variables are assumed to be known. As shown in Figure 3.1, the point estimates are placed unsymmetrically about the expected value of the random variables. The unknowns to be determined are (a) the locations of the point estimates, x_{1-} , x_{1+} , x_{2-} , and x_{2+} (or x_{11} , x_{12} , x_{21} , and x_{22}) and (b) the weighting functions associated with the discrete locations of x_1 and x_2 [i.e., (x_{1-}, x_{2-}) , (x_{1+}, x_{2-}) , (x_{1-}, x_{2+}) , and (x_{1+}, x_{2+})], denoted as p_{--} , p_{+-} , p_{-+} , and p_{++} (or p_{11} , p_{12} , p_{21} , and p_{22}), respectively.

Equations can be obtained as functions of the unknowns, x_{1-} , x_{1+} , x_{2-} , x_{2+} , p_{--} , p_{+-} , p_{-+} , and p_{++} , by specifying these as the means, standard deviations, skewness coefficients, and correlation coefficients of the random variables. In these equations, p_{i+} and p_{i-} , where $i=1,2$, represent the sum of the weighting functions associated with the point estimate locations x_{i+} and x_{i-} , respectively. For instance, as shown in Figure 3.1, p_{1-} and p_{1+} represent the weighting functions associated with the point estimate location x_{1-} , and hence, the weighting function p_{1-} will be defined as the sum of the weighting functions p_{--} and p_{-+} . The corresponding relationships of all the weighting functions are

$$p_{1+} = p_{+-} + p_{++} \quad (3.1)$$

$$p_{1-} = p_{--} + p_{-+} \quad (3.2)$$

$$p_{2+} = p_{-+} + p_{++} \quad (3.3)$$

$$p_{2-} = p_{--} + p_{+-} \quad (3.4)$$

¹ This model was first proposed by Professors Harr and Chameau of Purdue University

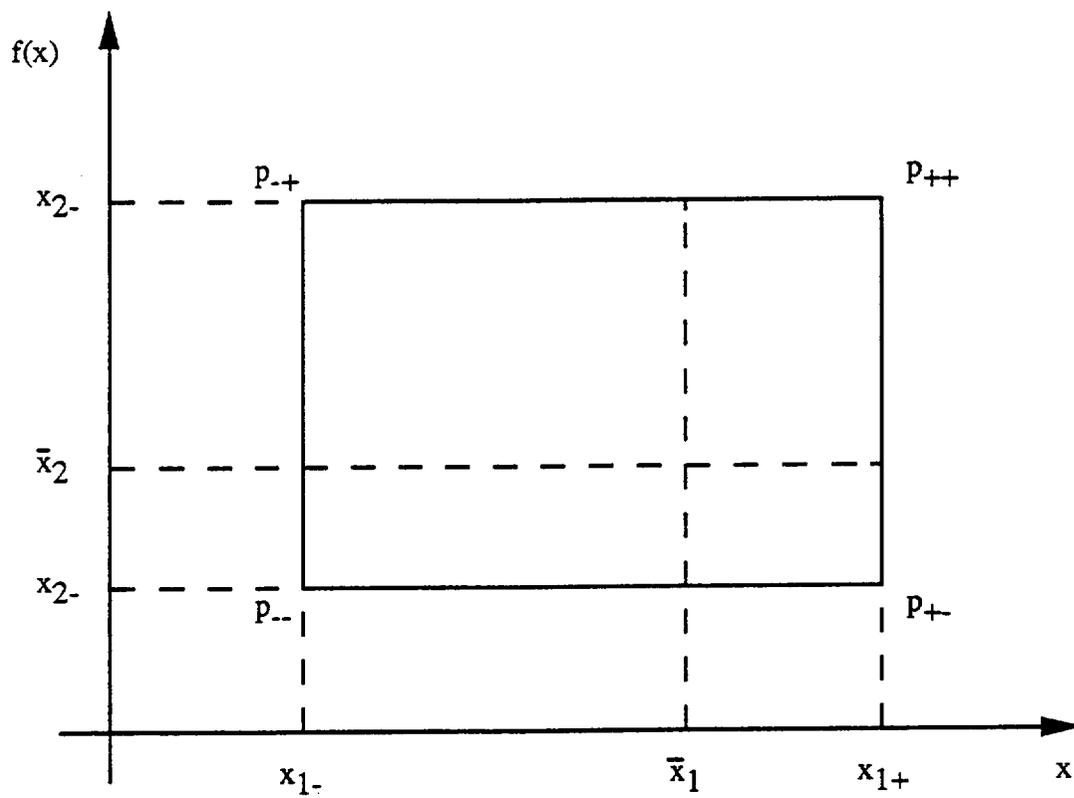


Figure 3.1 Model I Point Distribution
Approximation - Two Variables.

Using the definitions given in equations (3.1-3.4) and given the mean, standard deviation, skewness coefficient and correlation coefficient of the random variables, the resulting equations are

$$p_{--} + p_{-+} + p_{+-} + p_{++} = 1.0 \quad (3.5)$$

$$(x_{1-} - \bar{x}_1)[p_{--}(x_{2-} - \bar{x}_2) + p_{-+}(x_{2+} - \bar{x}_2)] \\ + (x_{1+} - \bar{x}_1)[p_{+-}(x_{2-} - \bar{x}_2) + p_{++}(x_{2+} - \bar{x}_2)] = \rho_{12} \sigma_{x_1} \sigma_{x_2} \quad (3.6)$$

$$p_{i+} x_{i+} + p_{i-} x_{i-} = \bar{x}_i \quad (3.7)$$

$$p_{i+} (x_{i+} - \bar{x}_i)^2 + p_{i-} (x_{i-} - \bar{x}_i)^2 = \sigma_{x_i}^2 \quad (3.8)$$

$$p_{i+} (x_{i+} - \bar{x}_i)^3 + p_{i-} (x_{i-} - \bar{x}_i)^3 = \beta_{x_i}(1) \sigma_{x_i}^3 \quad (3.9)$$

where x_i ($i=1,2$) is a random variable with mean \bar{x}_i , standard deviation σ_{x_i} , and skewness coefficient $\beta_{x_i}(1)$; ρ_{12} is the correlation coefficient between x_1 and x_2 .

Comparing equations (3.7) through (3.9) with the conditions for a single random variable (i.e., equations (2.2) through (2.4) of Chapter 2), it can be seen that the coordinates of the discrete approximation for functions with two random variables satisfy similar conditions as for a single random variable, and thus, the solution for the locations of the point estimates (i.e., x_{1-} , x_{1+} , x_{2-} , and x_{2+}) has the same form as for a single random variable. The solution is given as:

$$p_{i+} = \frac{1}{2} \left\{ 1 \pm \sqrt{1 - \frac{1}{1 + \left[\frac{\beta_{xi}(1)}{2} \right]^2}} \right\} \quad (3.10)$$

$$p_{i-} = 1 - p_{i+} \quad (3.11)$$

$$x_{i+} = \bar{x}_i + \sigma_{xi} \sqrt{\frac{p_{i-}}{p_{i+}}} \quad (3.12)$$

$$x_{i-} = \bar{x}_i - \sigma_{xi} \sqrt{\frac{p_{i+}}{p_{i-}}} \quad (3.13)$$

where $i=1,2$, and in equation (3.10) the sign preceding the radical is that of $-\beta_{xi}(1)$.

The above results show that the solution for p_{i+} and p_{i-} (sum of the weighting functions) for two random variables is of the same form as for a function with a single random variable. Having obtained the solution for p_{i+} and p_{i-} , the weighting functions at the point estimate locations, p_{--} , p_{+-} , p_{-+} , and p_{++} , are then determined. The weighting functions at the point estimate locations must satisfy the implied dependance among the variables. They are conveniently obtained by first setting the variables to be independent and then determining the solution for these weighting functions using the following relations:

$$p_{--} = p_{1-} p_{2-} + a_{12} \quad (3.14)$$

$$p_{+-} = p_{1-} p_{2-} - a_{12} \quad (3.15)$$

$$p_{-+} = p_{1-} p_{2+} - a_{12} \quad (3.16)$$

$$p_{++} = p_{1+} p_{2+} + a_{12} \quad (3.17)$$

where the first term of each weighting function represents the solution for the special case for a function with independent random variables. The sign preceding the factor a_{12}

is determined by the product of the subscripts (+’s and -’s) of the respective weighting functions, $p_{\pm\pm}$. For example, the plus sign preceding the factor a_{12} in equation (3.14) is of the product of a minus *times* a minus.

The solution obtained for the factor a_{12} that satisfies equation (3.6) is

$$a_{12} = \rho_{12} [p_{1+} p_{2+} p_{1-} p_{2-}]^{\frac{1}{2}} \quad (3.18)$$

When the random variables are not skewed ($\beta_1=0$), the solution reduces to the simpler form:

$$p_{1+} = p_{1-} = \frac{1}{2} \quad (3.19)$$

$$x_{i+} = \bar{x}_i + \sigma_{xi} \quad (3.20)$$

$$x_{i-} = \bar{x}_i - \sigma_{xi} \quad (3.21)$$

$$a_{12} = \frac{\rho_{12}}{4} \quad (3.22)$$

where $i=1,2$. The above result is precisely that given by Rosenblueth for symmetrically distributed random variables.

3.1.2 Discrete Approximation for a Function of Three Random Variables

Figure 3.2 illustrates model I of the new proposed point distribution methodology for a function $y=y(x_1, x_2, x_3)$ of three random variables. As explained for a function of two random variables, the point estimates are placed unsymmetrically about the mean value of the random variables. In the present case, the point estimates are placed in a three-dimensional space defined by the random variables x_1, x_2 , and x_3 . The unknowns to be determined are the locations of the point estimates (x_{i+} and x_{i-} , $i=1,2,3$) and the respective weighting functions as given in Table 3.1.

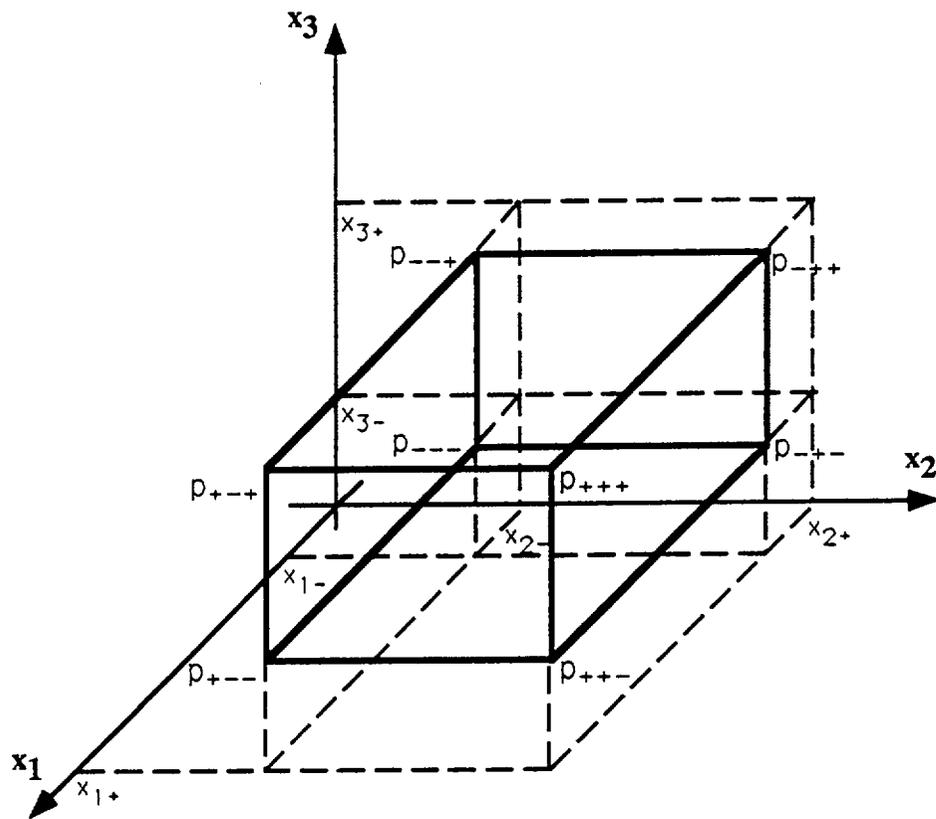


Figure 3.2 Model I Point Distribution
Approximation - Three Variables.

Table 3.1 Model I Point Estimate Locations for Three Random Variables.

Point estimate locations			weighting function
variable x_1	variable x_2	variable x_3	
x_{1+}	x_{2+}	x_{3+}	p_{+++}
x_{1-}	x_{2+}	x_{3+}	p_{-++}
x_{1+}	x_{2-}	x_{3+}	p_{+--}
x_{1-}	x_{2-}	x_{3+}	p_{---}
x_{1+}	x_{2+}	x_{3-}	p_{++-}
x_{1-}	x_{2+}	x_{3-}	p_{+-}
x_{1+}	x_{2-}	x_{3-}	p_{+--}
x_{1-}	x_{2-}	x_{3-}	p_{---}

The methodology adopted in model I is that the point estimate locations (x_{i+} and x_{i-}) are first determined by specifying the statistical moments: means, standard deviations, and skewness coefficients of the random variables as functions of the point estimate location unknowns x_{i+} and x_{i-} . The weighting functions at these determined locations are then computed by satisfying the required correlation structure of the variables.

Equations (B.1) through (B.8) in Appendix B.1 represent the model I point representation approximation for a function of three random variables, where the variables, x_i , $i=1,2,3$, are considered to be skewed and correlated. From equations (B.5-B.8), it can be seen that the coordinates of the discrete approximation (i.e., x_{i+} , x_{i-} , $i=1,2,3$) satisfy the same form of equations as for a single random variable. Thus, equations (3.10) through (3.13) represent the solution to the location of the point estimates x_{i+} and x_{i-} ; and the weighting functions p_{i+} and p_{i-} , where $i=1,2,3$. As given in equations (B.9-B.11), p_{i+} and p_{i-} represent the sum of the weighting functions associated with the point estimate locations x_{i+} and x_{i-} , respectively.

Next, the weighting functions associated with the point estimate locations (coordinates) as given in Table 3.1 are determined. The total number of unknown weighting functions is eight. These weighting functions have to satisfy the correlation

structure of the variables as given in equations (B.1 - B.4) and the conditions given in equations (B.9 - B.11). Hence, the total number of available conditions is seven. Therefore, with the point distribution pattern as shown in Figure 3.2, the total number of available conditions is one less than the eight unknown weighting functions. As a result of this, a unique solution will not be provided for the unknown weighting functions. The problem of non-uniqueness with the less number of available conditions is addressed below.

The first alternative solution to reconcile unknowns and equations for three random variables ignores a weighting function associated with one of the point estimate location and considers only the weighting functions associated with seven remaining point estimate locations, instead of all the eight locations given in Table 3.1. For example, neglecting the point estimate location associated with p_{--} (see Table 3.1), the solution obtained for the remaining seven weighting functions is as follows:

$$p_{+++} = p_{2+}p_{3+} - p_{1-}p_{3+} + p_{1-}p_{2-} + a_{12} + a_{13} + a_{23} \quad (3.23)$$

$$p_{--+} = p_{1-}p_{3+} - p_{1-}p_{2-} - a_{12} - a_{13} \quad (3.24)$$

$$p_{+-+} = p_{2-}p_{3+} - p_{1-}p_{2-} - a_{12} - a_{23} \quad (3.25)$$

$$p_{--+} = p_{1-}p_{2-} + a_{12} \quad (3.26)$$

$$p_{++-} = p_{2+}p_{3-} - p_{1-}p_{3-} - a_{13} - a_{23} \quad (3.27)$$

$$p_{-+-} = p_{1-}p_{3-} + a_{13} \quad (3.28)$$

$$p_{+--} = p_{2-}p_{3-} + a_{23} \quad (3.29)$$

where $a_{12} = \rho_{12}[p_{1+}p_{2+}p_{1-}p_{2-}]^{1/2}$, $a_{13} = \rho_{13}[p_{1+}p_{3+}p_{1-}p_{3-}]^{1/2}$, and $a_{23} = \rho_{23}[p_{2+}p_{3+}p_{2-}p_{3-}]^{1/2}$; and, p_{i+} and p_{i-} ($i=1,2,3$) are defined in equations (3.10-3.11).

A second alternative solution to achieve or balance the unknowns proposes that all the eight point estimate locations given in Table 3.1 (corners of Figure 3.2) be

considered. As noted, the solution provided will not be unique; however, the proposed solution does satisfy the necessary approximations to incorporate the skewed and correlated random variables. The proposed solution for the weighting functions is as given below:

$$p_{+++} = p_{1+}p_{2+}p_{3+} + a_{12}p_{3+} + a_{13}p_{2+} + a_{23}p_{1+} \quad (3.30)$$

$$p_{--+} = p_{1-}p_{2+}p_{3+} - a_{12}p_{3+} - a_{13}p_{2+} + a_{23}p_{1-} \quad (3.31)$$

$$p_{+-+} = p_{1+}p_{2-}p_{3+} - a_{12}p_{3+} + a_{13}p_{2-} - a_{23}p_{1+} \quad (3.32)$$

$$p_{---} = p_{1-}p_{2-}p_{3+} + a_{12}p_{3+} - a_{13}p_{2-} - a_{23}p_{1-} \quad (3.33)$$

$$p_{+--} = p_{1+}p_{2-}p_{3-} + a_{12}p_{3-} - a_{13}p_{2+} - a_{23}p_{1+} \quad (3.34)$$

$$p_{-+-} = p_{1-}p_{2+}p_{3-} - a_{12}p_{3-} + a_{13}p_{2+} - a_{23}p_{1-} \quad (3.35)$$

$$p_{+--} = p_{1+}p_{2-}p_{3-} - a_{12}p_{3-} - a_{13}p_{2-} + a_{23}p_{1+} \quad (3.36)$$

$$p_{---} = p_{1-}p_{2-}p_{3-} + a_{12}p_{3-} + a_{13}p_{2-} + a_{23}p_{1-} \quad (3.37)$$

where $a_{12} = \rho_{12}[p_{1+}p_{2+}p_{1-}p_{2-}]^{1/2}$, $a_{13} = \rho_{13}[p_{1+}p_{3+}p_{1-}p_{3-}]^{1/2}$, and $a_{23} = \rho_{23}[p_{2+}p_{3+}p_{2-}p_{3-}]^{1/2}$.

The above results for the weighting functions show that the number of point estimate terms required are all the permutations of three '+'s and '-'s (i.e., $2^3=8$). In the proposed solution given in equations (3.30-3.37), the factors a_{12} , a_{13} , and a_{23} are functions of the correlations between the variables, where the subscripts of these factors are those of the correlation coefficients by which these factors were defined. In addition, p_{i+} and p_{i-} ($i=1,2,3$) are defined in terms of the skewness coefficient of the variables (equations 3.10 and 3.11). The first terms of all the weighting functions (p_{1+2+3+} , p_{1-2+3+} , etc.) given in equations (3.30 - 3.37) represent the solution for the special case where the variables are independent ($a_{12} = a_{13} = a_{23} = 0$). The sign preceding the factors a_{12} , a_{13} , or a_{23} in equations (3.30 - 3.37) is determined as the product of the subscripts ('+'s and '-'s) of the respective weighting functions (p_{+++} , p_{--+} , etc.). The position of '+'s and '-'s in the weighting functions is dictated by the subscripts of these factors. For example,

in equation (3.31), the sign preceding the factor a_{12} is the product of - times + = - (i.e., the first two subscripts in $p_{.++}$); for the factor a_{13} , the sign is - times + = - (i.e., the first and third subscripts in $p_{.++}$). When the random variables are symmetrically distributed, the above solution again reduces to Rosenblueth's (1975) solution.

3.1.3 Generalization of Model I Point Distribution Approximation

Model I of the new point distribution methodology will be generalized from the point distribution patterns for functions of two and three random variables (see sections 3.1.1 and 3.1.2). For functions of N correlated and skewed random variables, the point-distribution will be defined in the N -dimensional variables space. It is again assumed that the statistical moments: expected value, standard deviation, skewness coefficient, and correlation structure of the random variables are known.

The point distribution pattern of model I requires $2N+2^N$ unknowns, where N is the number of input random variables of a given function; the $2N$ term represents the unknown coordinates of the point distribution pattern: the term 2^N represents the unknown weighting functions of the respective point estimate locations. With the given information on statistical moments: expected value, standard deviation, skewness, and correlation structure of the variables, the total number of available conditions to solve for the point-representation unknowns are $1+3N+N(N-1)/2$, where N is the number of random variables; the term 1 represents the condition obtained by the sum of the weighting functions of the point-representation, which is similar to the sum of the discrete probabilities is always unity, the term $3N$ represents the number of conditions obtained by specifying the first three statistical moments: expected value, standard deviation, and skewness of the variables, and the term $N(N-1)/2$ represents the number of conditions obtained by specifying the given covariance terms of the variables. Table 3.2 compares the number of unknowns of the point representation and the number of available conditions to solve for these unknowns.

Table 3.2 Model I Number of Point Representation Unknowns.

Number of Variables	Number of Unknowns			Number of Conditions
	unknown point estimate locations	unknown weighting functions	total number of unknowns	
2	4	4	8	8
3	6	8	14	13
4	8	16	24	19
5	10	32	42	26

As the number of unknowns exceeds the number of available conditions, the solution provided by model I point distribution approximation will not be unique for $N > 2$. A solution is proposed below which satisfies the specified number conditions given in Table 3.2. Also, the solution provided follows a pattern that can be used directly in computer applications.

The proposed generalized solution of model I first solves for the unknown point estimate locations (x_{i+} and x_{i-} , $i=1,2, \dots, N$) by specifying the first three statistical moments: expected value, standard deviation, and skewness coefficient of the random variables (for more details see equations (B.5-B.8) for three random variables). The solution to these point estimate locations is defined in N -dimensional space, with the coordinates given by:

$$[\bar{x}_1 + (-1)^{d_1} F_1 \sigma_{x_1}, \bar{x}_2 + (-1)^{d_2} F_2 \sigma_{x_2}, \dots, \bar{x}_N + (-1)^{d_N} F_N \sigma_{x_N}] \quad (3.38)$$

where x_i ($i=1,2, \dots, N$) is a random variable with mean \bar{x}_i , standard deviation σ_{x_i} , and skewness coefficient $\beta_{x_i}(1)$; $d_i \in \{0,1\}$; and $\{F_i = (p_{i+}/p_{i-})^{0.5}$ if $d_i=1$, $F_i = (p_{i-}/p_{i+})^{0.5}$ if $d_i=0\}$; where p_{i+} and p_{i-} are functions of the skewness coefficient of the random variables given by the following equations:

$$p_{i\pm} = \frac{1}{2} \left\{ 1 \pm \sqrt{1 - \frac{1}{1 + \left[\frac{\beta_{xi}(1)}{2}\right]^2}} \right\} \quad (3.39)$$

$$p_{i-} = 1 - p_{i+} \quad (3.40)$$

where $i=1,2, \dots, N$ and in equation (3.39) the sign preceding the radical is that of $-\beta_{xi}(1)$.

The point distribution coordinates of model I as given in equation (3.38) form a "hyper-prism" in the N -dimensional space. According to this equation, the F_i 's are unity for Rosenblueth's (1975) original distribution: his solution is symmetrical about the expected value of the random variables.

Having determined the point estimate locations, the point distribution unknowns to be defined are the respective weighting functions associated with the point estimate locations. As noted previously, a unique solution is not possible (see Table 3.2). Therefore, a solution is proposed that does incorporate skewed and correlated random variables. The proposed generalized solution provides 2^N point estimate locations. In addition, it considers all the possible permutations of N '+'s and '-'s. The proposed generalized solution for these respective weighting functions is

$$p_{\pm\pm, \dots, \pm} = p_{1\pm} p_{2\pm}, \dots, p_{N\pm} + (-1)^{d_1} (-1)^{d_2} a_{12} p_{12} \\ + (-1)^{d_1} (-1)^{d_3} a_{13} p_{13} + \dots + (-1)^{d_{N-1}} (-1)^{d_N} a_{N-1N} p_{N-1N} \quad (3.41)$$

where $p_{i\pm}$ ($i=1,2, \dots, N$) is defined in equations (3.39) and (3.40) as functions of the skewness coefficient of the variables; $a_{lm} = \rho_{lm} (p_{1+} p_{1-} p_{m+} p_{m-})^{0.5}$, $p_{lm} = p_{1\pm} p_{2\pm} \dots p_{k\pm} \dots p_{N\pm}$ ($k \neq l$ and $k \neq m$) where $l=1,2, \dots, N$, $m=1+1, 1+2, \dots, N$ and ρ_{lm} is the correlation coefficient between the random variables x_l and x_m ; and $d_i \in \{0,1\}$.

The product $[(-1)^{d_1} \times (-1)^{d_m}]$ in equation (3.41) determines the sign preceding the factor a_{lm} , which is also the product of the subscripts '+'s and '-'s of the respective weighting function, $p_{\pm\pm, \dots, \pm}$. The position of '+'s and '-'s in the weighting functions is dictated by the subscripts of the factor a_{lm} , which are 1 and m. The total number of terms of the factor a_{lm} is $N(N-1)/2$. When the skewness of the random variables is zero, the generalized of the weighting functions reduces to Rosenblueth's (1975) solution.

As an example, consider a function with four random variables ($N=4$). Then the total number of weighting functions is equal to 16 (2^4). The magnitudes of these weighting functions are given by the equation (3.41) in which all possible combinations of the binary digits d_1, d_2, d_3 , and d_4 are considered. For instance, the weighting function p_{+++} has binary digits ($d_i, i=1,2,3, \& 4$) of the form (1,0,0,0) and thus the magnitude of p_{+++} is given by:

$$\begin{aligned} p_{+++} &= p_{1-}p_{2+}p_{3+}p_{4+} + (-1)^1(-1)^0 a_{12}p_{3+}p_{4+} + (-1)^1(-1)^0 a_{13}p_{2+}p_{4+} + (-1)^1(-1)^0 a_{14}p_{2+}p_{3+} \\ &\quad + (-1)^0(-1)^0 a_{23}p_{1-}p_{4+} + (-1)^0(-1)^0 a_{24}p_{1-}p_{3+} + (-1)^0(-1)^0 a_{34}p_{1-}p_{2+} \\ &= p_{1-}p_{2+}p_{3+}p_{4+} - a_{12}p_{3+}p_{4+} - a_{13}p_{2+}p_{4+} - a_{14}p_{2+}p_{3+} + a_{23}p_{1-}p_{4+} + a_{24}p_{1-}p_{3+} + a_{34}p_{1-}p_{2+} \end{aligned}$$

where p_{i+} and p_{i-} ($i=1,2,3, \& 4$) are defined in equations (3.39) and (3.40);

$a_{12} = \rho_{12} (p_{1+}p_{1-}p_{2+}p_{2-})^{0.5}$, $a_{13} = \rho_{13} (p_{1+}p_{1-}p_{3+}p_{3-})^{0.5}$, $a_{14} = \rho_{14} (p_{1+}p_{1-}p_{4+}p_{4-})^{0.5}$,
 $a_{23} = \rho_{23} (p_{2+}p_{2-}p_{3+}p_{3-})^{0.5}$, and $a_{24} = \rho_{24} (p_{2+}p_{2-}p_{4+}p_{4-})^{0.5}$; $p_{12} = p_{3+}p_{4+}$, $p_{13} = p_{2+}p_{4+}$,
 $p_{14} = p_{2+}p_{3+}$, $p_{23} = p_{1-}p_{4+}$, $p_{24} = p_{1-}p_{3+}$, and $p_{34} = p_{1-}p_{2+}$. The weighting functions have coordinates given by the equation (3.38). Hence, for $N=4$ weighting function p_{+++} is located at:

$$\begin{aligned} x_{+++} &= (\bar{x}_1 + (-1)^1 F_1 \sigma_{x1}, \bar{x}_2 + (-1)^0 F_2 \sigma_{x2}, \bar{x}_3 + (-1)^0 F_3 \sigma_{x3}, \bar{x}_4 + (-1)^0 F_4 \sigma_{x4}) \\ &= (\bar{x}_1 - F_1 \sigma_{x1}, \bar{x}_2 + F_2 \sigma_{x2}, \bar{x}_3 + F_3 \sigma_{x3}, \bar{x}_4 + F_4 \sigma_{x4}) \end{aligned}$$

where $F_1 = (p_{1+}/p_{1-})^{0.5}$, $F_2 = (p_{2+}/p_{2-})^{0.5}$, $F_3 = (p_{3+}/p_{3-})^{0.5}$, and $F_4 = (p_{4+}/p_{4-})^{0.5}$.

Following the sequence of operations as given above, the resulting distributions of weighting functions for $N=4$ are given in Appendix B.2.

A computer program was developed to compute the point representation approximations of model I. The program listing is given in Appendix D.

3.2 Model II of the New Point Distribution Methodology

3.2.1 Discrete Approximation for a Function of Two Random Variables

The point-distribution pattern for model II is as shown in Figure 3.3 for a function $f(x_1, x_2)$ of two random variables x_1 and x_2 . To account for the skewness of the random variables, as shown in Figure 3.3, the point estimates are placed unsymmetrically about the expected value of the random variables (\bar{x}_1, \bar{x}_2). The unknowns are the locations of the point estimates and the respective weighting functions associated with the discrete locations of the point estimates. The coordinates of the point-representations are given in Table 3.3.

Table 3.3 Model II Point Estimate Locations for Two Random Variables.

Point estimate locations				Weighting functions
x_1	x_2	$x_1 - \bar{x}_1$	$x_2 - \bar{x}_2$	
$\bar{x}_1 + z_{11+}$	$\bar{x}_2 + z_{12}$	z_{11+}	z_{12}	p_{1+}
$\bar{x}_1 - z_{11-}$	$\bar{x}_2 - z_{12}$	$-z_{11-}$	$-z_{12}$	p_{1-}
\bar{x}_1	$\bar{x}_2 + z_{22+}$	0.	z_{22+}	p_{2+}
\bar{x}_1	$\bar{x}_2 - z_{22-}$	0.	$-z_{22-}$	p_{2-}

Following the methodology, equations are obtained as functions of the point representation unknowns ($z_{11+}, z_{11-}, z_{22+}, z_{22-}, z_{12}, p_{1+}, p_{1-}, p_{2+},$ and p_{2-}) by specifying the statistical moments: mean, standard deviation, skewness coefficient, and correlation coefficients of the random variables. The weighting functions are assumed to be equally

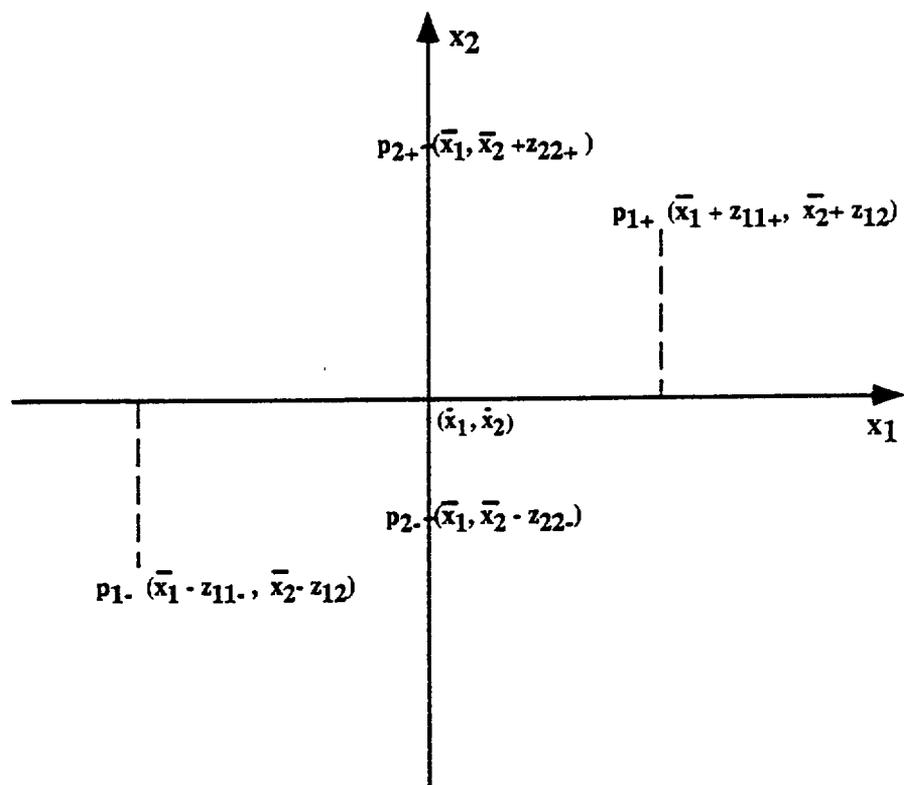


Figure 3.3 Model II Point Distribution Approximation - Two Variables.

distributed among the point distribution pairs. For example, the point distribution pairs for the point estimate locations given in Table 3.3 are $[p_{1+}, p_{1-}]$ and $[p_{2+}, p_{2-}]$, therefore,

$$p_{i+} + p_{i-} = \frac{1}{2} \quad (3.42)$$

where $i=1,2$.

Using the above assumptions and with the given information concerning the random variables' statistical moments, the equations given below are obtained.

For the variable x_1 :

$$E[(x_1 - \bar{x}_1)] = p_{1+} z_{11+} - p_{1-} z_{11-} = 0.0 \quad (3.43)$$

$$E[(x_1 - \bar{x}_1)^2] = p_{1+} z_{11+}^2 + p_{1-} z_{11-}^2 = \sigma_{x1}^2 \quad (3.44)$$

$$E[(x_1 - \bar{x}_1)^3] = p_{1+} z_{11+}^3 - p_{1-} z_{11-}^3 = \beta_{x1}(1) \sigma_{x1}^3 \quad (3.45)$$

For the variable x_2 :

$$E[(x_2 - \bar{x}_2)] = (p_{1+} - p_{1-}) z_{12} + p_{2+} z_{22+} - p_{2-} z_{22-} = 0.0 \quad (3.46)$$

$$E[(x_2 - \bar{x}_2)^2] = (p_{1+} + p_{1-}) z_{12}^2 + p_{2+} z_{22+}^2 + p_{2-} z_{22-}^2 = \sigma_{x2}^2 \quad (3.47)$$

$$E[(x_2 - \bar{x}_2)^3] = (p_{1+} - p_{1-}) z_{12}^3 + p_{2+} z_{22+}^3 - p_{2-} z_{22-}^3 = \beta_{x2}(1) \sigma_{x2}^3 \quad (3.48)$$

The covariance term of the variables x_1 and x_2 is

$$E[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)] = (p_{1+} z_{11+} + p_{1-} z_{11-}) z_{12} = \rho_{12} \sigma_{x1} \sigma_{x2} \quad (3.49)$$

where x_i ($i=1,2$) is a random variable with mean \bar{x}_i , standard deviation σ_{xi} , and skewness coefficient $\beta_{xi}(1)$: ρ_{12} is the correlation coefficient between x_1 and x_2 .

A special sequence is recommended to solve the above system of equations.

First, the unknowns z_{11+} , z_{11-} , p_{1+} , and p_{1-} are solved, and then the unknown z_{12} ; finally, the unknowns z_{22+} , z_{22-} , p_{2+} and p_{2-} are obtained. The sequence is to solve the point estimate locations associated with the weighting function p_1 (i.e., p_{1+} and p_{1-}) and then that of p_2 (p_{2+} and p_{2-}). The unknown locations associated with the weighting functions are simply the values in columns 3 and 4 of Table 3.3. Given below is the sequence adopted in solving the point estimate locations in vectorial form:

$$z_1 = [z_{11}, z_{12}]$$

$$z_2 = [0, z_{22}]$$

where $z_{ii} = \{z_{ii+}, z_{ii-}, p_{i+}, \text{ and } p_{i-}\}$, $i=1,2$.

Following the above sequence, the point representation unknowns are found. Equations (3.42-3.45) provide the four unknown quantities, p_{1+} , p_{1-} , z_{11+} , and z_{11-} , and have the following form:

$$p_{1+} = \frac{1}{4} \left\{ 1 \pm \sqrt{\frac{\beta_{xl}(1)^2}{8 + \beta_{xl}(1)^2}} \right\} \quad (3.50)$$

$$p_{1-} = \frac{1}{2} - p_{1+} \quad (3.51)$$

$$z_{11+} = \sqrt{2} \sqrt{\frac{p_{1-}}{p_{1+}}} \sigma_{xl} \quad (3.52)$$

$$z_{11-} = \sqrt{2} \sqrt{\frac{p_{1+}}{p_{1-}}} \sigma_{xl} \quad (3.53)$$

where the sign preceding the radical in equation (3.50) is that of $-\beta_{xl}(1)$.

The solution to the equation (3.49) provides the following expression for z_{12} :

$$z_{12} = \frac{\rho_{12} \sigma_{xl} \sigma_{12}}{p_{1+} z_{11+} + p_{1-} z_{11-}} \quad (3.54)$$

where p_{1+} , p_{1-} , z_{11+} , and z_{11-} are defined in equations (3.50-3.53).

Knowing the solution to p_{1+} , p_{1-} , z_{11+} , z_{11-} , and z_{12} , equations (3.46-3.48) are rearranged as follows:

$$p_{2+}z_{22+} - p_{2-}z_{22-} = -(p_{1+} - p_{1-})z_{12} = A \quad (3.55)$$

$$p_{2+}z_{22+}^2 + p_{2-}z_{22-}^2 = \sigma_{x2}^2 - (p_{1+} + p_{1-})z_{12}^2 = B \quad (3.56)$$

$$p_{2+}z_{22+}^3 - p_{2-}z_{22-}^3 = \beta_{x2}(1)\sigma_{x2}^3 - (p_{1+} - p_{1-})z_{12}^3 = C \quad (3.57)$$

where $p_{2+} + p_{2-} = 1/2$.

From equations (3.55-3.57), a quadratic equation is obtained as:

$$\left(\frac{B}{2} - A^2\right)z_{22}^2 - \left(\frac{C}{2} - AB\right)z_{22} + (AC - B^2) = 0.0 \quad (3.58)$$

where A, B, and C are defined in equations (3.55-3.57), respectively. This quadratic equation provides z_{22+} and z_{22-} , where the solution is given as $z_{22+} = z_{22}$ and $z_{22-} = (-z_{22})$. Finally the weighting functions p_{2-} and p_{2+} are determined from the expressions:

$$p_{2-} = \frac{\frac{1}{2}z_{22+} - A}{z_{22+} + z_{22-}} \quad (3.59)$$

$$p_{2+} = \frac{1}{2} - p_{2-} \quad (3.60)$$

There are some limitations in using model II. One limitation is in situations where the roots of the quadratic equation (3.58) are imaginary. The other is when the locations of the point distributions are negative. Alternatives are discussed in sub section 3.2.4.

3.2.2 Discrete Approximation for a Function of Three random variables

Figure 3.4 presents the point distribution scheme of model II for a function of three random variables. The point estimates are placed unsymmetrically about the expected value of the random variables ($\bar{x}_1, \bar{x}_2, \bar{x}_3$). The procedure adopted in placing the point estimates is the same as for a function of two random variables. Table 3.4 presents the location of the coordinates.

Table 3.4 Model II Point Estimate Locations for Three Random Variables.

Point estimate locations						Weighting Functions
x_1	x_2	x_3	$x_1 - \bar{x}_1$	$x_2 - \bar{x}_2$	$x_3 - \bar{x}_3$	
$\bar{x}_1 + z_{11+}$	$\bar{x}_2 + z_{12}$	$\bar{x}_3 + z_{13}$	z_{11+}	z_{12}	z_{13}	p_{1+}
$\bar{x}_1 - z_{11-}$	$\bar{x}_2 - z_{12}$	$\bar{x}_3 - z_{13}$	$-z_{11-}$	$-z_{12}$	$-z_{13}$	p_{1-}
\bar{x}_1	$\bar{x}_2 + z_{22+}$	$\bar{x}_3 + z_{23}$	0.	z_{22+}	z_{23}	p_{2+}
\bar{x}_1	$\bar{x}_2 - z_{22-}$	$\bar{x}_3 - z_{23}$	0.	$-z_{22-}$	$-z_{23}$	p_{2-}
\bar{x}_1	\bar{x}_2	$\bar{x}_3 + z_{33+}$	0.	0.	z_{33+}	p_{3+}
\bar{x}_1	\bar{x}_2	$\bar{x}_3 - z_{33-}$	0.	0.	$-z_{33-}$	p_{3-}

Given the statistical moments: expected value, standard deviation, skewness coefficient, and correlation coefficient of the random variables, $x_1, x_2,$ and x_3 ; equations are obtained as functions of the point representation unknowns (see Appendix B.3). These equations are based on the assumption that the weighting functions are equally distributed among the point distribution pairs, i.e., $p_{i+} + p_{i-} = 1/3$, where $i=1,2,3$. In solving these equations, a special sequence is followed such that the unknowns can be evaluated. Thus, the sequential order in which the unknowns are solved is given in a vector as follows:

$$z_1 = [z_{11}, z_{12}, z_{13}]$$

$$z_2 = [0, z_{22}, z_{23}]$$

$$z_3 = [0, 0, z_{33}]$$

where $z_{ii} = \{z_{ii+}, z_{ii-}, p_{i+}, p_{i-}\}$, $i=1, 2,$ and 3 .

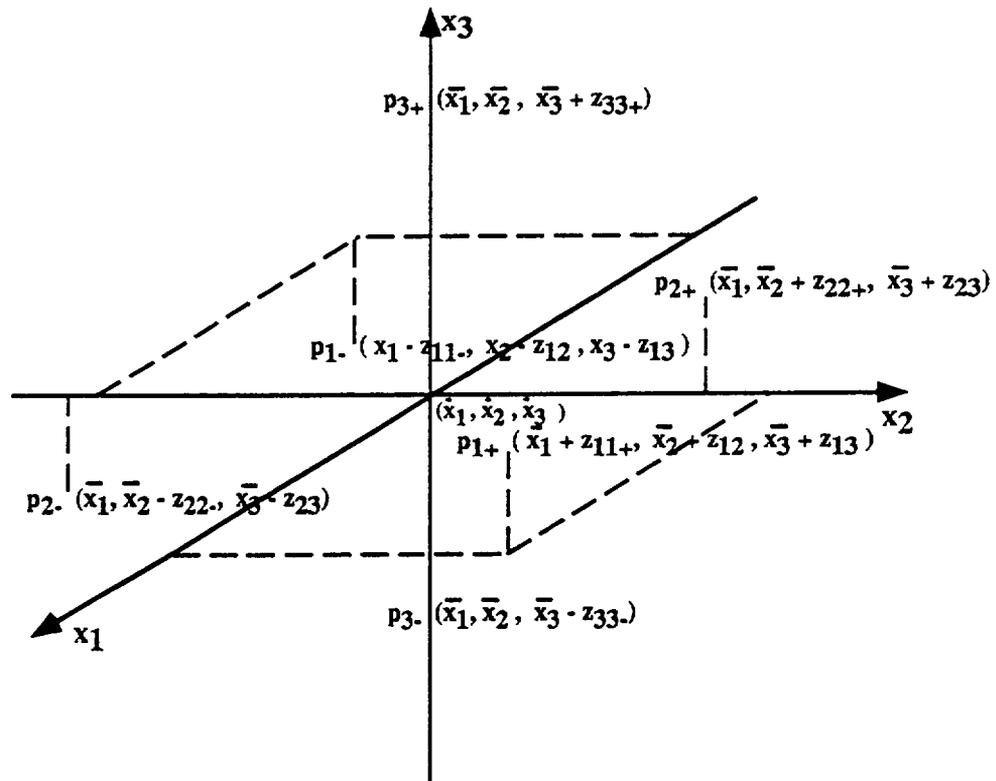


Figure 3.4 Model II Point Distribution Approximation - Three Variables.

The vectors z_1 , z_2 , and z_3 are associated with the weighting functions p_1 , p_2 , and p_3 , respectively. The above sequence follows the same pattern as explained for a function of two random variables.

The solutions for the unknowns are given in Appendix B.3 (equations B.25 - B.39). Equations (B.25 - B.28) represent the solution for p_{1+} , p_{1-} , z_{11+} , and z_{11-} . Knowing these solutions, z_{12} and z_{13} are then determined using equation B.29. Then z_{22+} , z_{22-} , p_{2+} , and p_{2-} are obtained. To solve these unknowns, equations (B.16 - B.18) are rearranged as (B.30-B.32), and a quadratic equation (B.33) is obtained as given in Appendix B.3. The solution to this quadratic equation provides the unknowns z_{22+} and z_{22-} . The weighting functions p_{2-} and p_{2+} are then determined using equations (B.34) and (B.35), respectively. Next, the unknown z_{23} is solved using equation (B.36). Finally, equations (B.19 - B.21) are rearranged to solve for the unknowns z_{33+} , z_{33-} , p_{3+} , and p_{3-} . A quadratic equation (B.37) is obtained, which yields the solution for z_{33+} and z_{33-} . Lastly, the solution to the weighting functions p_{3-} and p_{3+} are determined from equations (B.38) and (B.39), respectively.

The conditions of applicability of model II are discussed in subsection 3.2.4.

3.2.3 Generalization of Model II Point Distribution Approximation

The generalized coordinates of the point representation for a function of N random variables for model II can be given as N -point estimate pairs, located at $\bar{x} + z_{i+}$ and $\bar{x} - z_{i-}$ ($i=1,2, \dots, N$), with:

$$\begin{aligned}\bar{x} &= [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N] \\ z_{i+} &= [z_{i1}, z_{i2}, \dots, z_{i(i-1)}, z_{ii+}, z_{i(i+1)}, \dots, z_{in}] \\ z_{i-} &= [z_{i1}, z_{i2}, \dots, z_{i(i-1)}, z_{ii-}, z_{i(i+1)}, \dots, z_{in}]\end{aligned}\tag{3.61}$$

where \bar{x}_i is the expected value of the random variable x_i ($i = 1,2, \dots, N$) and $\{z_{ij} = 0$ for $j = 1,2, \dots, i-1\}$.

The locations $\bar{x} + z_{i+}$ and $\bar{x} - z_{i-}$ are associated with the weighting functions p_{i+} and p_{i-} , respectively. In providing the discrete approximation, the statistical moments: expected value, standard deviation, skewness coefficient, and correlation coefficient of the random variables are assumed to be known. For a function of N random variables, the number of point representation unknowns will be $2N + [N(N+3)/2]$, where the term $2N$ represents the number of weighting functions p_{i+} and p_{i-} ($i=1,2, \dots, N$), and the term $[N(N+3)/2]$ represents the number of unknown locations of the point representation, which are given in the vector forms as z_{i+} and z_{i-} (equation (3.61)), $i=1,2, \dots, N$.

The solution is based on the assumption that the weighting functions are equally distributed among the point-distribution pairs, that is,

$$p_{i+} + p_{i-} = \frac{1}{N} \quad (3.62)$$

where $i=1,2, \dots, N$.

The point representation unknowns are solved by specifying the statistical moments of the random variables as functions of these unknowns. By defining these statistical moments, the equations obtained are

$$E[(x_i - \bar{x}_i)] = p_{i+} z_{i+} - p_{i-} z_{i-} = 0.0 - \sum_{k=1}^{i-1} (p_{k+} - p_{k-}) z_{ki} = A_i \quad (3.63)$$

$$E[(x_i - \bar{x}_i)^2] = p_{i+} z_{i+}^2 + p_{i-} z_{i-}^2 = \sigma_{xi}^2 - \sum_{k=1}^{i-1} (p_{k+} + p_{k-}) z_{ki}^2 = B_i \quad (3.64)$$

$$E[(x_i - \bar{x}_i)^3] = p_{i+} z_{i+}^3 - p_{i-} z_{i-}^3 = \beta_{xi}(1) \sigma_{xi}^3 - \sum_{k=1}^{i-1} (p_{k+} - p_{k-}) z_{ki}^3 = C_i \quad (3.65)$$

where the random variable x_i ($i=1,2, \dots, N$) is with mean \bar{x}_i , standard deviation σ_{xi} , and skewness coefficient $\beta_{xi}(1)$.

Similarly, defining the covariance term of the random variables x_i and x_j (ρ_{ij}):

$$E[(x_i - \bar{x}_i)(x_j - \bar{x}_j)] = (p_{i+} z_{ii+} + p_{i-} z_{ii-}) z_{ij} = \rho_{ij} \sigma_{xi} \sigma_{xj} - \sum_{k=1}^{i-1} (p_{k+} + p_{k-}) z_{ki} z_{kj} \quad (3.66)$$

where p_{i+} , p_{i-} , z_{ii+} , z_{ii-} and z_{ij} ($i=1,2, \dots, N-1$ and $j=i+1, i+2, \dots, N$) are the unknowns of the point representation.

Equations (3.62-3.66) represent the required conditions to solve for the point representation unknowns of model II. In solving these unknowns, a special sequence is followed. The sequential order is given in vectorial forms z_{i+} and z_{i-} (see equation (3.61)). Equations (3.62 - 3.65) are rearranged to solve for the unknowns z_{ii+} , z_{ii-} , p_{i-} , and p_{i+} ($i=1,2, \dots, N$), and then the unknowns z_{ij} ($i=1,2, \dots, N-1$ and $j=i+1, i+2, \dots, N$) are computed using equation (3.66). The corresponding expressions are given below.

The quadratic equation which solves for the unknowns z_{ii+} and z_{ii-} is as follows:

$$\left(\frac{B_i}{N} - A_i^2\right) z_{ii+}^2 - \left(\frac{C_i}{N} - A_i B_i\right) z_{ii+} + (A_i C_i - B_i^2) = 0.0 \quad (3.67)$$

where the solution z_{ii} is defined as $z_{ii+} = z_{ii}$ and $z_{ii-} = (-z_{ii})$.

The expressions for the unknowns p_{i-} and p_{i+} are

$$p_{i-} = \frac{\frac{1}{N} z_{ii+} - A_i}{z_{ii+} + z_{ii-}} \quad (3.68)$$

$$p_{i+} = \frac{1}{N} - p_{i-} \quad (3.69)$$

The unknowns z_{ij} are obtained from the expression

$$z_{ij} = \frac{\rho_{ij} \sigma_{xi} \sigma_{xj} - \sum_{k=1}^{i-1} (p_{k+} + p_{k-}) z_{ki} z_{kj}}{p_{i+} z_{ii+} + p_{i-} z_{ii-}} \quad (3.70)$$

where $i=1,2, \dots, N-1$ and $j=i+1, i+2, \dots, N$. An algorithm was designed and is incorporated into a computer program to solve for the unknowns. The algorithm is analogous to that for solving functions of two and three random variables. When the random variables are symmetrically distributed, the generalized solution of model II reduces to Lind's (1983) solution.

3.2.4 Alternative Solutions for Model II Applicability

The applicability of model II is contingent on the roots of the quadratic equation (3.67) being real; and, of course in practical applications, by the condition that the coordinates of the point representation are greater than zero (because the variables are defined to be positive). In cases where these conditions are not met, two alternative procedures will be suggested. One alternative is to change the variables' sequence when solving for the point-estimates, and the other alternative is to redistribute the weighting functions. The suggested alternatives are discussed below:

- 1) The discrete approximation for model II is provided by following the extended 'face-centered' point representation (see Tables 3.3 and 3.4). By virtue of this representation, the coordinates of the discrete approximation depend on the sequence of the variables of the \bar{x} vector as given in equation (3.61). For example, if there are N random variables, it is possible to choose $N!$ sequences for the \bar{x} vector. Recall, there is no uniqueness. Therefore, one alternative is to use a trial and error procedure to produce a sequence of variables among all the possible sequences of $N!$ such that the required conditions for practical use are satisfied. But, it is not appropriate to consider all the possible $N!$ sequences as it increases tremendously with N . An example will be given in the next chapter.

2) Formally, the second alternative distributes the weighting functions equally among the point-distribution pairs (see equation 3.62). This assumption is consistent with the principle of maximum entropy, which says that if no additional information is available about the distribution, the least biased is the uniform distribution. However, if this alternative does not satisfy the required conditions, another alternative is to redistribute the weighting functions according to the available information. Again a trial and error procedure must be involved because of the sequential nature of the algorithm.

The alternatives suggested here are also applicable to Lind's (1983) methodology.

CHAPTER 4

NUMERICAL EXAMPLES OF THE DEVELOPED METHODOLOGY

4.1 Numerical Examples

Illustrations of the newly developed point estimate methodology is presented in this chapter using some numerical examples with detailed step-by-step procedures. The results of the new point estimate methodology are also compared with Rosenblueth (1975), Lind (1983), and Harr (1989) point estimate methodologies, and with the Monte Carlo simulation technique. A sensitivity analysis is also performed to assess the influence of the skewed and correlated random variables. In the next chapter, the methodology will be applied to the flow code LLUVIA.

4.2 Function of Three Random Variables

The first example will consider the function $y = x_1 x_2 / x_3$ of three random variables. Three cases will be investigated. In the first case, variables will be skewed and correlated. In the second case they will be independent and skewed. In the third case, the variables will be assumed to be statistically independent log-normal variates.

4.2.1. Point Estimate Results for Skewed and Correlated Variables

The statistical properties of the variables x_1 , x_2 , and x_3 are given in Table 4.1: the correlation structure of these variables are $\rho_{12}=0.8$, $\rho_{13}=0.7$, and $\rho_{23}=0.6$. Using this information, the expected value and standard deviation of the function $y = (x_1 x_2 / x_3)$ will be obtained using the two models of the new point estimate method. Estimates will also

be obtained using Rosenblueth's (1975), Lind's (1983), and Harr's (1989) methods.

Table 4.1 Statistical Data of Input Random Variables.

variable	expected value	coefficient of variation	skewness
x_1	20.0	20 %	-1.5
x_2	5.0	40 %	1.0
x_3	15.0	30 %	-1.0

4.2.1.1 Results Using Model I

The procedure to be followed in model I of the new point estimate methodology was given in section 3.1.3. The point representation coordinates of model I are computed using equations (3.38), (3.39), and (3.40) for which there are 2^3 coordinates. For example, using equations (3.39) and (3.40) for the random variable x_1 , the weighting functions p_{1+} and p_{1-} are given by:

$$p_{1+} = \frac{1}{2} \left\{ 1 + \sqrt{1 - \frac{1}{1 + \left[\frac{-1.5}{2} \right]^2}} \right\} = 0.8 \quad (4.1)$$

$$p_{1-} = 1 - 0.8 = 0.2 \quad (4.2)$$

Using equation (3.38), the factor F_1 is defined as $F_1 = (0.8/0.2)^{0.5}$ if $d_1=1$ and $F_1 = (0.2/0.8)^{0.5}$ if $d_1=0$. Hence, the point estimate locations x_{1+} and x_{1-} are then given by:

$$x_{1+} = 20.0 + (-1)^0 \left(\frac{0.2}{0.8} \right)^{0.5} (0.2) (20.0) = 22.0 \quad (4.3)$$

$$x_{1-} = 20.0 + (-1)^1 \left(\frac{0.8}{0.2} \right)^{0.5} (0.2) (20.0) = 12.0 \quad (4.4)$$

Similarly, for the random variables x_2 and x_3 , the point estimate locations are $x_{2+} = 8.236$, $x_{2-} = 3.764$, $x_{3+} = 17.781$, and $x_{3-} = 7.719$. Model I requires all the possible permutations

of the point estimate locations x_{i+} and x_{i-} , $i=1,2,3$. Table 4.2 gives all the eight permutations of x_{i+} and x_{i-} .

Table 4.2 Model I Point Estimate Locations for Skewed and Correlated Random Variables.

sign	variable x_1	variable x_2	variable x_3
---	12.000	3.764	7.719
+--	22.000	3.764	7.719
-+-	12.000	8.236	7.719
++-	22.000	8.236	7.719
--+	12.000	3.764	17.781
+ - +	22.000	3.764	17.781
- + +	12.000	8.236	17.781
+++	22.000	8.236	17.781

The next step in the procedure is the calculation of the weighting functions associated with the point estimate locations as given in Table 4.2. This is accomplished by using the generalized solution for the weighting functions given in equation (3.41). According to equation (3.41), the weighting functions are defined in terms of p_{i+} , p_{i-} , and the correlation structure of the variables. For example, considering the first sign value (---) as given in column one of Table 4.2, the corresponding weighting function for this sign value is given by:

$$p_{---} = p_{1-}p_{2-}p_{3-} + (-1)^1(-1)^1 a_{12} p_{3-} + (-1)^1(-1)^1 a_{13} p_{2-} + (-1)^1(-1)^1 a_{23} p_{1-}. \quad (4.5)$$

where $a_{12} = \rho_{12}(p_{1+}p_{1-}p_{2+}p_{2-})^{0.5}$, $a_{13} = \rho_{13}(p_{1+}p_{1-}p_{3+}p_{3-})^{0.5}$, and $a_{23} = \rho_{23}(p_{2+}p_{2-}p_{3+}p_{3-})^{0.5}$. The factors p_{im} in equation (3.41) are defined as $p_{12} = p_{3-}$, $p_{13} = p_{2-}$, and $p_{23} = p_{1-}$.

All the weighting functions p_{i+} and p_{i-} ($i=1,2,3$) in equation (4.5) are computed using equations (3.39) and (3.40). Therefore, $p_{1+} = 0.8$, $p_{1-} = 0.2$, $p_{2+} = 0.2764$, $p_{2-} = 0.7236$, $p_{3+} = 0.7236$, and $p_{3-} = 0.2764$. Substituting these weighting functions in equation (4.5), the factors a_{im} are defined as $a_{12} = 0.8 (0.8 \ 0.2 \ 0.2764 \ 0.7236)^{0.5} =$

0.1431, $a_{13} = 0.7 (0.8 \ 0.2 \ 0.7236 \ 0.2764)^{0.5} = 0.1252$, and $a_{23} = 0.6 (0.2764 \ 0.7236 \ 0.7236 \ 0.2764)^{0.5} = 0.1200$. Hence, the weighting function p_{---} of equation (4.5) reduces to: $p_{---} = 0.2 \ 0.7236 \ 0.2764 + 0.1431 \ 0.2764 + 0.1252 \ 0.7236 + 0.1200 \ 0.2 = 0.1942$.

The corresponding estimate of the function y for the sign value (---) is $y = (12.000)(3.764)/7.719 = 5.852$. Continuing this procedure provides the magnitudes of all the weighting functions for the signs listed in column one of Table 4.2 and the corresponding point estimates of the function $y = (x_1 \ x_2/x_3)$. The results are given in Table 4.3.

Table 4.3 Model I Results for Point Estimates of Function $y = x_1 \ x_2/x_3$ - Skewed and Correlated Random Variables.

sign	weighting function P_{ijk}	function y point estimates Y_{ijk}	$P_{ijk} Y_{ijk}$	$P_{ijk} Y_{ijk}^2$
---	0.1942	5.852	1.137	6.651
+- -	0.1258	10.728	1.350	14.478
-+ -	-.0137	12.804	-.175	-2.246
++ -	-.0299	23.474	-.702	-16.476
-- +	0.0937	2.540	0.238	0.605
+ - +	0.3099	4.657	1.443	6.721
- + +	-.0742	5.558	-.412	-2.292
+++	0.3942	10.190	4.017	40.932

With the point estimates of the function y as given in Table 4.3, the estimates for the statistical moments of the function y are then obtained using the relationship (for more details see section 4.8, Harr (1987)):

$$E[y^M] = \sum P_{ijk} Y_{ijk}^M \quad (4.6)$$

where the indices i , j , and k represent the signs listed in column one of Table 4.3; p_{ijk}

and y_{ijk} are given in Table 4.3.

The expected value and standard deviation of the function y are obtained using equation (4.6). The expected value of y , $E[y]$, results from $M=1$. The standard deviation of function y (s_y) follows from the relationship $s_y^2 = E[y^2] - (E[y])^2$. $E[y]$ is the sum of the terms in column four of Table 4.3, $E[y] = 6.894$; and $E[y^2]$ is the sum of the terms in column five of Table 4.3, $E[y^2] = 48.356$. Using the relationship $s_y^2 = 48.356 - (6.894)^2 = 0.829$, this gives $s_y = 0.911$ and the coefficient of variation $V = (s_y/E[y]) 100 = (0.911/6.894) 100 = 13.21 \%$. These results form the first row of Table 4.6.

4.2.1.2 Results Using Model II

The procedure followed in model II of the new point estimate methodology was given in sections 3.2.2 and 3.2.3. The point estimate locations and the corresponding weighting functions were given in Table 3.4. The necessary relations are equations (B.25-B.39) of Appendix B.3. To provide a balance between unknowns and equations (the equations of Appendix B.3), the assumption is made that the weighting functions are equally distributed among the point-distribution pairs, i.e., $p_{i+} + p_{i-} = 1/3$ (as was discussed in section 3.2.2).

The point estimates for model II were obtained with the sequence of the variables taken as $[x_3, x_1, x_2]$. That is, in providing the point distribution approximation, the statistical data of the variable x_3 is considered first and then the variables x_1 and x_2 , respectively. This particular sequence meets the required conditions of applicability (positive point estimate locations and non imaginary roots) discussed in section 3.2.4.

Using equations (B.25-B.28), the unknowns p_{1+} , p_{1-} , z_{11+} , and z_{11-} of the variable x_3 are first obtained. In solving for these unknowns, the first three statistical moments of the variable x_3 are substituted in the respective equations.

$$p_{1+} = \frac{1}{6} \left\{ 1 + \sqrt{\frac{(-1.0)^2}{12 + (-1.0)^2}} \right\} = 0.2129 \quad (4.7)$$

$$p_{1-} = \frac{1}{3} - p_{1+} = 0.3333 - 0.2129 = 0.1204 \quad (4.8)$$

$$z_{11+} = \sqrt{3} \sqrt{\frac{0.1204}{0.2129}} 0.3 (15.0) = 5.863 \quad (4.9)$$

$$z_{11-} = \sqrt{3} \sqrt{\frac{0.2129}{0.1204}} 0.3 (15.0) = 10.362 \quad (4.10)$$

The unknowns z_{12} and z_{13} are then solved using equation (B.29). For the selected sequence $[x_3, x_1, x_2]$, the point representation unknowns z_{12} and z_{13} , respectively, satisfy the covariance terms between x_3 and x_1 (ρ_{31}) and between x_3 and x_2 (ρ_{32}). Therefore, in equation (B.29), corresponding correlation coefficient terms are substituted as follows:

$$z_{12} = \frac{0.7 \cdot 0.3 (15.0) \cdot 0.2 (20.0)}{0.2129 (5.863) + 0.1204 (10.362)} = 5.048 \quad (4.11)$$

$$z_{13} = \frac{0.6 \cdot 0.3 (15.0) \cdot 0.4 (5.0)}{0.2129 (5.863) + 0.1204 (10.362)} = 2.163 \quad (4.12)$$

Therefore, the coordinates of the first pair of the point estimate locations are $(15.0 + 5.863, 20.0 + 5.048, 5 + 2.163) = (20.863, 25.048, 7.163)$ and $(15.0 - 10.362, 20.0 - 5.048, 5.0 - 2.163) = (4.638, 14.952, 2.837)$. The corresponding weighting functions are 0.2129 and 0.1204, respectively. These coordinates are entered in the first two rows of Table 4.4.

To solve for the unknowns z_{22+} , z_{22-} , p_{2+} , and p_{2-} , the factors A_2 , B_2 , and C_2 are computed using equations (B.30-B.32). In equations (B.30-B.32), the first three statistical moments of the variable x_1 are substituted into these equations. Because, the variable x_1 is the second variable in the sequence $[x_3, x_1, x_2]$.

$$A_2 = -(0.2129 - 0.1204) 5.048 = -0.4669$$

$$B_2 = (0.2(20.0))^2 - (0.2129 + 0.1204) 5.048^2 = 7.506$$

$$C_2 = (-1.5)(0.2(20.0))^3 - (0.2129 - 0.1204) 5.048^3 = -107.899$$

The quadratic equation (B.33) is then given by:

$$\begin{aligned} & \left(\frac{7.506}{3} - (-0.467)^2 \right) z_{22}^2 - \left(\frac{-107.899}{3} - (-0.467)(7.506) \right) z_{22} \\ & + ((-0.407)(-107.899) - 7.506^2) = 0.0 \end{aligned} \quad (4.13)$$

and the solution is $z_{22} = z_{22+} = 0.183$ and $z_{22} = (-z_{22-}) = -14.392$.

The respective weighting functions p_{2-} and p_{2+} are as follows:

$$p_{2-} = \frac{\frac{1}{3} 0.183 - (-0.467)}{0.183 + 14.392} = 0.0362 \quad (4.14)$$

$$p_{2+} = \frac{1}{3} - 0.0362 = 0.2971 \quad (4.15)$$

The unknown z_{23} is then determined by using equation (B.36). In solving the unknown z_{23} , the covariance term between x_1 and x_2 (ρ_{12}) is considered as x_1 and x_2 are the second and third variables in the sequence $[x_3, x_1, x_2]$.

$$z_{23} = \frac{0.8 \cdot 0.2(20.0) \cdot 0.4(5.0) - (0.2129 + 0.1204) 5.048 \cdot 2.163}{0.2971(0.183) + 0.0362(14.392)} = 4.799 \quad (4.16)$$

Hence, the coordinates of the second pair of the point estimate locations are $(15.0, 20.0 + 0.183, 5.0 + 4.799) = (15.0, 20.183, 9.799)$ and $(15.0, 20.0 - 14.392, 5.0 - 4.799) = (15.0, 5.608, 0.201)$. The respective weighting functions are 0.2971 and 0.0362. The coordinates are entered in the third and fourth rows of Table 4.4.

Finally, the unknowns z_{33+} , z_{33-} , p_{3+} , and p_{3-} are solved to satisfy the first three statistical moments of the third variable in the selected array, x_2 . Using the equation

(B.37), the quadratic equation which solves for the unknowns z_{33+} and z_{33-} is as follows:

$$\left(\frac{-5.236}{3} - (-1.452)^2 \right) z_{33}^2 - \left(\frac{-21.772}{3} - (-1.452)(-5.236) \right) z_{33} + ((-1.452)(-21.772) - (-5.236)^2) = 0.0 \quad (4.17)$$

where the factors A_3 , B_3 , and C_3 in equation (B.37) are enumerated as follows:

$$A_3 = 0.0 - (0.2129 - 0.1204) 2.163 - (0.2971 - 0.0362) 4.799 = -1.452$$

$$B_3 = (0.4 (5.0))^2 - (2.163)^2/3 - (4.799)^2/3 = -5.236$$

$$C_3 = 1.0 (0.4 (5.0))^3 - (0.2129 - 0.1204) 2.163^3 - (0.2971 - 0.0362) 4.799^3 = -21.772$$

The solution of the quadratic equation yields $z_{33} = z_{33+} = 4.120$ and $z_{33} = (-z_{33-}) = -0.265$. The weighting functions p_{3-} and p_{3+} are given by:

$$p_{3-} = \frac{\frac{1}{3} 4.120 - (-1.452)}{4.120 + 0.265} = 0.6444 \quad (4.18)$$

$$p_{3+} = \frac{1}{3} - 0.6444 = -0.3110 \quad (4.19)$$

Hence, the coordinates of the third pair of the point estimate locations are defined as follows $(15.0, 20.0, 5.0 + 4.120) = (15.0, 20.0, 9.120)$ and $(15.0, 20.0, 5.0 - 0.265) = (15.0, 20.0, 4.735)$. The respective weighting functions are -0.3110 and 0.6444 . The coordinates of the third pair are entered in the fifth and sixth rows of Table 4.4.

The above was presented to explain the details of model II solution procedure. Actual results were obtained by the use of a computer program given in Appendix D.

Table 4.4 Model II Point Estimate Locations for Skewed and Correlated Random Variables.

variable x_3	variable x_1	variable x_2	weighting functions
20.863	25.048	7.163	0.2129
4.638	14.952	2.837	0.1204
15.000	20.183	9.799	0.2971
15.000	5.608	0.201	0.0362
15.000	20.000	9.120	-.3110
15.000	20.000	4.735	0.6444

The point estimates of the function $y = (x_1 x_2/x_3)$ can now be computed using the results given in Table 4.4. For example, considering the first row of the point estimate location given in Table 4.4, the estimate of the function y is given by $y = (25.048) (7.163) / (20.863) = 8.600$. Continuing this procedure, the estimates of all the point estimate locations are given in Table 4.5.

Table 4.5 Model II Results for Point Estimates of Function $y = x_1 x_2/x_3$ - Skewed and Correlated Random Variables.

weighting function p	function y point estimates	$p y$	$p y^2$
0.2129	8.600	1.831	15.746
0.1204	9.146	1.101	10.071
0.2971	13.184	3.917	51.641
0.0362	0.075	0.003	0.0002
-.3110	12.161	-3.782	-45.994
0.6444	6.314	4.069	25.690

The expected value and standard deviation of the function y are again obtained using equation (4.6). The procedure followed beyond this point in determining these

statistical moments is exactly the same procedure as was explained for model I. Following this procedure, the expected value of the function y is given by the sum of the six point estimates in the third column of Table 4.5, $E[y] = 7.139$. The sum of the fourth column of Table 4.5 produces $E[y^2] = 57.155$, and hence, the standard deviation s_y of the function y is $s_y = (57.155 - (7.139)^2)^{0.5} = 2.490$. The coefficient of variation is $V = (2.490/7.139) 100 = 34.88 \%$. These results form the second row of Table 4.6.

For comparison purposes, the estimates for the statistical moments of the function $y=(x_1 x_2/x_3)$ are also obtained using Rosenblueth (1975), Lind (1983) and Harr (1989) point estimate methodologies. The results of all the point estimate methodologies are given in Table 4.6.

Table 4.6 Point Estimates for Statistical Moments of Function $y =x_1 x_2/x_3$ - Skewed and Correlated Random Variables.

methodology	expected value $E[y]$	standard deviation s_y	coefficient of variation V
model I of new method	6.894	0.911	13.21 %
model II of new method	7.139	2.490	34.88 %
Rosenblueth (1975)	6.960	3.253	46.73 %
Lind (1983)	6.957	2.903	41.73 %
Harr (1989)	7.004	3.487	49.79 %

4.2.2 Point Estimate Results for Independent and Skewed Variables

The same statistical properties of the variables will be used as given in Table 4.1, except that there will not be any correlation structure among the variables. The results of all the methodologies are tabulated in Table 4.7.

Table 4.7 Point Estimates for Statistical Moments of Function $y = x_1 x_2/x_3$
- Independent and Skewed Random Variables.

methodology	expected value $E[y]$	standard deviation s_y	coefficient of variation V
model I of new method	7.650	5.005	65.42 %
model II of new method	8.062	5.867	72.77 %
Rosenblueth (1975)	7.326	4.112	56.12 %
Lind (1983), Harr (1989)	7.489	4.213	56.25 %

A Monte Carlo simulations of the above cases was conducted using beta distributions (Harr, 1987) for the variates. The assumed beta distributions were selected to satisfy the given statistical moments about the variables x_1 , x_2 , and x_3 as given in Table 4.1 and also the constraint $\beta^2(1)+1 \leq \beta(2) \leq 1.5\beta^2(1)+3$, where $\beta(1)$ is the skewness coefficient, and $\beta(2)$ is the kurtosis coefficient of the beta distribution. The later constraint is imposed to meet the criteria for a possible beta distribution (as given in Figure 2.2.2 of Harr, 1987).

The lower bounds of the beta variables x_1 , x_2 , and x_3 for the Monte Carlo simulation were taken to be 0, 2, and 3, respectively, along with the statistical properties as given in Table 4.1. In selecting the bounds for the negatively skewed random variables x_1 and x_3 , the lower bounds were defined such that the upper bounds of these variables were one standard deviation from the mean. For the positively skewed random variable x_2 , the lowest possible bound was selected, which meets the constraint on $\beta(1)$ and $\beta(2)$ as stated above along with the statistical properties given in Table 4.1. The Monte Carlo simulation technique was then applied to the function y by varying the number of simulations from 100 to 10,000. The results obtained for the statistical moments of the function $y = (x_1 x_2/x_3)$ using the Monte Carlo simulation are tabulated in Table 4.8.

Table 4.8 Results of the Monte Carlo Simulation Technique for Function $y = x_1 x_2/x_3$ - Beta Variables.

Number of simulations	Expected value of function y $E[y]$	Standard deviation of function y s_y	coefficient of variation V
100	8.049	5.311	65.98 %
1000	8.278	6.406	77.39 %
5000	7.952	6.551	82.38 %
10000	7.882	6.166	78.23 %

4.2.3 Point Estimate Results for Independent Lognormal Variables

This case requires more information than is known about the variables in the previous cases (see section -4.2.1 and 4.2.2). That is, conceptually, all statistical moments of the random variables are completely known. Also, the exact solution can be obtained for the statistical moments for lognormal variates given as the function $y = x_1 x_2/x_3$.

The expected values of the input random variables x_i ($i=1,2,3$) are taken to be the same as given in Table 4.1. The coefficients of variation of the input random variables are assumed to range from 10% to 90%. For illustration purposes, the coefficient of variation of the random variables x_1 , x_2 , and x_3 will be taken to be the same. As the variables are taken to be lognormal variates, the skewness of the variables is a function of the coefficient of variation of the variables. The relevant relationships are given by Hahn and Shapiro (1967).

As an example, considering the case in which the coefficient of variation of the input random variables is 60%, the computed values for the statistical parameters using the relationship by Hahn and Shapiro (1967) are as given in Table 4.9. For instance, considering the variable x_1 , the variance of the lognormal variate $\ln x_1$ is given by $\ln(1+0.6^2)=0.308$, the expected value of the log normal variate is given by $\ln(20.0)$ -

$0.5(0.308) = 2.842$, and the skewness of the variable x_1 is given by $(e^{0.308}-1)^{0.5} (e^{0.308}+2) = 2.016$. As the coefficients of variation of all the input random variables ($x_i, i=1,2,3$) are assumed to be the same, the variance of the lognormal variates and the skewness of the random variables x_i are the same as given in Table 4.9.

Table 4.9 Statistical Parameters of Input Random Variables with Coefficient of Variation 60 % - Independent Lognormal Variates.

variable	expected value	expected value of lognormal variate	variance of lognormal variate	skewness
x_1	20	2.842	0.308	2.016
x_2	5	1.456	0.308	2.016
x_3	15	2.554	0.308	2.016

With the first three statistical moments: expected value, standard deviation (coefficient of variation), and skewness coefficient of the independent lognormal variables x_i ($i=1,2,3$) (for example, the expected value and skewness coefficient are respectively given in first and fourth columns of Table 4.9, and the coefficients of variation of the variables as 60 %), the estimates for the statistical moments: expected value and standard deviation of the function $y = x_1 x_2/x_3$ are computed using the two models of the new point estimate methodology. The results of Rosenblueth's (1975), Lind's (1983), and Harr's (1989) point estimate methodologies were computed (see also Figures 4.1 and 4.2). Furthermore, the exact solution (lognormal solution) for the expected value and standard deviation of the function $y = x_1 x_2/x_3$ as independent lognormal variates were also obtained. The results are tabulated in Table 4.10 and shown in Figures 4.1 and 4.2.

Table 4.10 Results for the Expected value and Standard deviation of Function $y = x_1 x_2/x_3$ - Independent Lognormal Variates.

coefficient of variation (%)	skewness	expected value of function y			standard deviation of function y		
		model I of new method	model II of new method	lognormal solution	model I of new method	model II of new method	lognormal solution
10	0.301	6.732	6.733	6.733	1.160	1.159	1.172
20	0.608	6.913	6.933	6.933	2.354	2.339	2.450
30	0.927	7.172	7.262	7.267	3.611	3.556	3.947
40	1.264	7.459	7.707	7.733	4.951	4.811	5.792
50	1.625	7.733	8.235	8.333	6.384	6.082	8.136
60	2.016	7.964	8.791	9.067	7.908	7.319	11.161
70	2.443	8.138	9.301	9.933	9.519	8.461	15.091
80	2.912	8.253	9.694	10.933	11.210	9.462	20.193
90	3.429	8.315	9.927	12.067	12.975	10.315	26.792

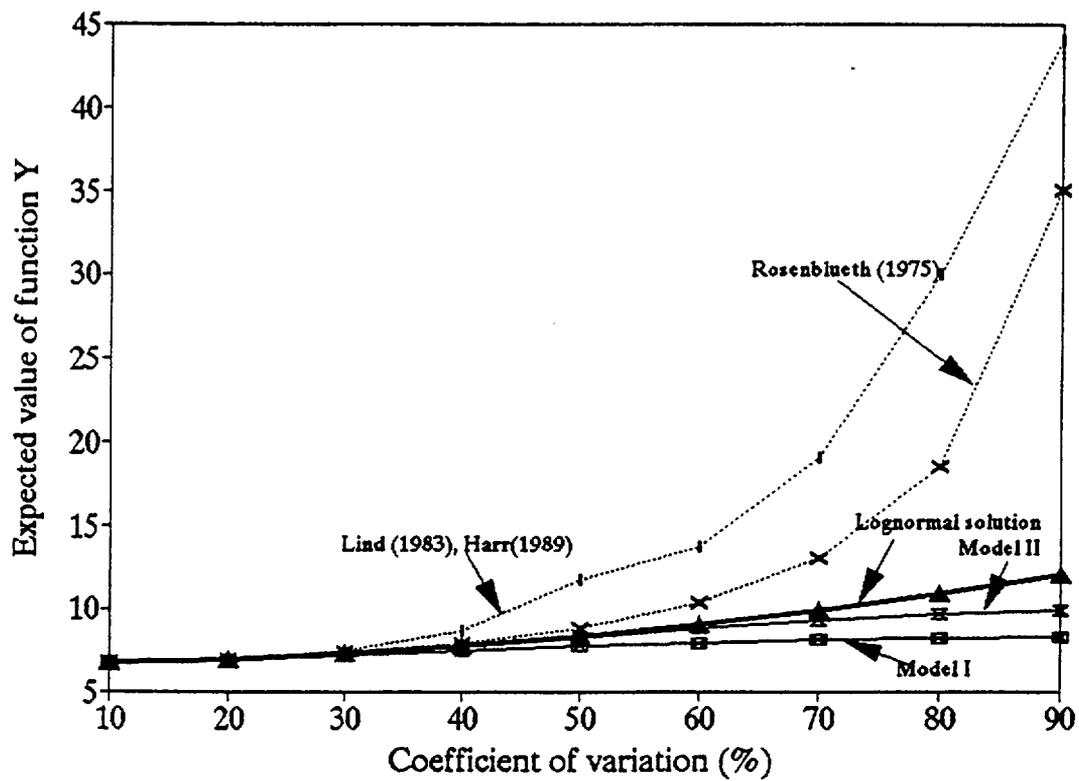


Figure 4.1 Comparison Between Point Estimate Method Solutions and Lognormal Solution for Expected Value of Function $y = x_1 x_2 / x_3$ - Independent Lognormal Variates.

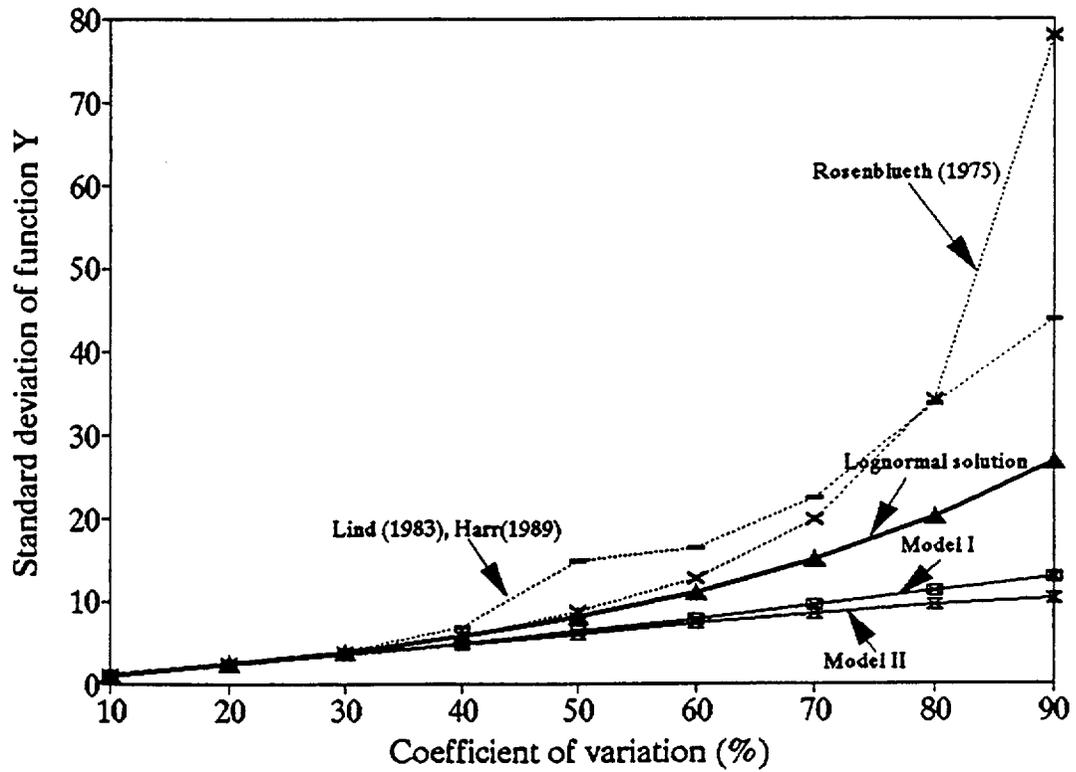


Figure 4.2 Comparison Between Point Estimate Method Solutions and Lognormal Solution for Standard deviation of Function $y = x_1 x_2/x_3$ - Independent Lognormal Variates.

CHAPTER 5

APPLICATION TO LLUVIA

LLUVIA is a flow code developed at Sandia National Laboratories to estimate the groundwater travel time computation for one-dimensional, steady-state flow problems through multiple layers of saturated or partially saturated media. The problem of special interest in the present chapter is the variability of the travel time caused by the uncertainties in the unsaturated zone parameters. The analysis of uncertainties in the prediction of travel time is supported by Carsel and Parrish (1988); Panian (1987); Rawls et al. (1982); and Kaplan and Yarrington (Report SAND88-2247C, Sandia National Laboratories).

The developed procedures described in Chapter 3 and illustrated in Chapter 4 were used to address the uncertainties of the input parameters. Three cases will be investigated. Representative statistical parameters are those reported by Carsel and Parrish (1988). The results of the new models will be compared to Rosenblueth (1975), Lind (1983), and Harr (1989) methods.

5.1 Flow Code LLUVIA

The flow problem will be treated as a steady-state, one-dimensional, unsaturated flow in a layered, fractured medium. The system is assumed to be isothermal, non-deforming, and the fluid is assumed to be of constant density. Only a single fluid phase - the liquid phase - will be considered. The flow in the system is assumed to be Darcian.

5.1.1 The Model

5.1.1.1 Mathematical Model

The mathematical model used in the flow code LLUVIA to describe the unsaturated flow in a fractured medium is the composite matrix/fracture model of Klavetter and Peters (1986, 1988). The model representation treats the media material as a continuum when solving for the pressure field. The major assumption in this development is that the pressure gradient in the matrix and the fractures are identical in the direction perpendicular to flow. The flow media may or may not contain fractures.

The flow in the system is assumed to Darcian; therefore:

$$v = -K(\psi) \nabla (\psi + z) \quad (5.1)$$

where v is the Darcian velocity (infiltration rate), ψ is the pressure head (p/γ , where γ is the unit weight of the liquid), $K(\psi)$ is the effective hydraulic conductivity for the prevailing degree of saturation, and z is the elevation head, above a chosen datum.

5.1.1.2 Parameter Models

The characterizing fluid movement in an unsaturated medium requires knowledge of the unsaturated parameters hydraulic conductivity $K(\psi)$ and the degree of saturation $S(\psi)$ as functions of the pressure head ψ . The recommended functional form for the matrix and fracture saturation is van Genuchten (1978) model:

$$S(\psi) = (S_s - S_r) \left\{ \frac{1}{1 + |\alpha \psi|^N} \right\}^{(1 - \frac{1}{N})} + S_r \quad (5.2)$$

where ψ is the pressure head; $S(\psi)$ is the degree of saturation for the prevailing pressure head ψ ; S_s is the maximum saturation; S_r is the residual saturation; and α and N are the curve fitting parameters of van Genuchten's model respectively describing the air entry parameter and the slope of the desaturation curve.

The recommended functional form for the matrix and fracture hydraulic conductivity is van Genuchten/Mualem (1976) model:

$$K(\psi) = K_s [1 + |\alpha \psi|^N]^{-\frac{\lambda}{2}} \left\{ 1 - \left[\frac{|\alpha \psi|^N}{1 + |\alpha \psi|^N} \right]^\lambda \right\}^2 \quad (5.3)$$

where $K(\psi)$ is the hydraulic conductivity for the pressure head ψ , K_s is the saturated hydraulic conductivity, α and N are the saturation-curve fitting parameters of van Genuchten's (1978) model, and $\lambda = 1 - 1/N$.

5.1.2 Boundary Conditions

Two boundary conditions must be stipulated as the flow is assumed to be one-dimensional (equation 5.1). The upper boundary is the constant infiltration flux boundary ($q=q_0$), and the lower boundary is taken as a water table boundary with the pressure head $\psi=0.0$. Hence,

$$\begin{aligned} q &= q_0 \text{ at } z = z_t \\ \psi &= 0.0 \text{ at } z = 0. \end{aligned}$$

5.1.3 Solution Procedure

Equation (5.1) is the governing flow equation to be solved for the pressure field. For partially saturated conditions, the hydraulic conductivity ($K(\psi)$) is a strong function of pressure head (ψ) and hence, equation (5.1) is extremely nonlinear. The solver in LLUVIA is based on Hindmarsh's code (1981). The solver is designed to provide an efficient numerical solution procedure for the one-dimensional, steady-state flow equation (5.1).

The pressure field obtained by solving equation (5.1) is subsequently used to compute the hydraulic conductivity, matrix saturation, and flow velocities in matrix and fractures (if present). The flow velocities are calculated following the equations from Peters et al. (1986). Using these flow velocities, minimum groundwater travel time is

computed from a given elevation to the bottom of the flow domain.

5.1.4 Parameters Required

The program separates the one-dimensional domain into a number of layers. Layers are numbered consecutively from the bottom to the top of the domain. For the groundwater travel time computation, the program requires the quantification of the infiltration rate, boundary pressure head at the bottom of the domain, and the material characterization for each layer. The material properties are the porosity (ϕ), saturated hydraulic conductivity (K_s), residual saturation (S_r), and the curve fitting parameters of van Genuchten's model (the air entry parameter (α) and van Genuchten model desaturation parameter (N)). If the layer is fractured, matrix and fracture properties are defined separately for the LLUVIA program. If there are no fractures in the material, the area fraction of fractures (referred to as a fracture porosity) is set to zero.

5.2 Specific Problems Solved

The developed methodology described in Chapter 3 was incorporated in to the flow code LLUVIA for the probabilistic estimate of the groundwater travel time in an unsaturated medium. The material input parameters: saturated hydraulic conductivity (K_s), residual saturation (S_r), and van Genuchten model parameters describing air entry parameter (α) and the slope of the desaturation curve (N) are taken to be the random variables. The layers will be assumed to be nonfractured. Therefore, within each layer (unit) of nonfractured medium there will be four random variables: K_s , S_r , α , and N . The representative values for the probabilistic analyses are selected in accordance with the data reported by Carsel and Parrish (1988). These results show that these four random variables within each unit are correlated and skewed.

Infiltration flux rates of 0.1 mm/year and 1 mm/year are considered in the illustrative examples presented in this chapter (Report SAND88-0942, Sandia National Laboratories).

5.2.1 Example I - Single Layer Problem of Silty Clay

The first application of the newly developed methodology to LLUVIA considers a silty clay layer of 10 m thickness. The groundwater travel time is taken to be the time for a drop of water to travel from the top of the layer to the groundwater table which is at the bottom of the layer. According to the reported data by Carsel and Parrish, silty clay has the highest coefficient of variation (453.3 %) for the saturated hydraulic conductivity (K_s). The fitted beta distribution (Harr, 1987) for silty clay saturated hydraulic conductivity (K_s) has a skewness of 6.26. In fitting a beta distribution for K_s , the reported values by Carsel and Parrish (1988) were used for the expected value, coefficient of variation, and the range of limits. Similarly, the representative values for the other three random variables S_r , α , and N of silty clay were selected as given by Carsel and Parrish, and the skewness of these random variables were determined by fitting a beta distribution within the given limits of variation. The reported data of Carsel and Parrish also give positive pair-wise correlation coefficients among the four random variables K_s , S_r , α , and N of silty clay. Table 5.1 gives the desired expected values, coefficients of variation, skewness coefficients, and correlation coefficients of the random variables.

Table 5.1 Representative Values of the Random Variables for Silty Clay.

Random variable	K_s (cm/hr)	S_r	α (1/m)	N
expected value	0.02	0.07	0.5	1.09
coefficient of variation (%)	453.3	33.5	113.6	5.0
skewness	6.26	0.00	2.10	0.78

The correlation matrix for the silty clay layer is

$$\rho = \begin{bmatrix} K_g & S_r & \alpha & N \\ 1.0 & 0.949 & 0.974 & 0.908 \\ 0.949 & 1.0 & 0.964 & 0.794 \\ 0.974 & 0.964 & 1.0 & 0.889 \\ 0.908 & 0.794 & 0.889 & 1.0 \end{bmatrix}$$

Having chosen the expected values, coefficient of variations, skewness coefficients, and correlation coefficients of the random variables, the point distribution approximation of the random variables were computed using the newly developed methodology. Models I and II, as outlined in Chapter 3, were considered. A computer program was developed to solve for the point distribution approximation of these two models.

5.2.1.1 Example I - Model I

For the first example, model I gives sixteen realizations for the LLUVIA run. These are the required number of runs for a system of four random variables ($2^N=2^4=16$). These values are tabulated in Table C.1.

For each set, infiltration flux rates of 0.1 mm/year and 1.0 mm/year were specified, and the groundwater travel time outputs of LLUVIA were obtained. Each output of the groundwater travel time was then weighted by the corresponding weighting functions of each sets of values as given in Table C.1. The evaluation of the statistical parameters, expected value and coefficient of variation of the groundwater travel time, are summarized in Appendix C.1 and the results are given in Table 5.2. Estimates of the expected pressure head, degree of saturation, and hydraulic conductivity profiles in the silty clay layer were also obtained (see Figures 5.1, 5.2, and 5.3).

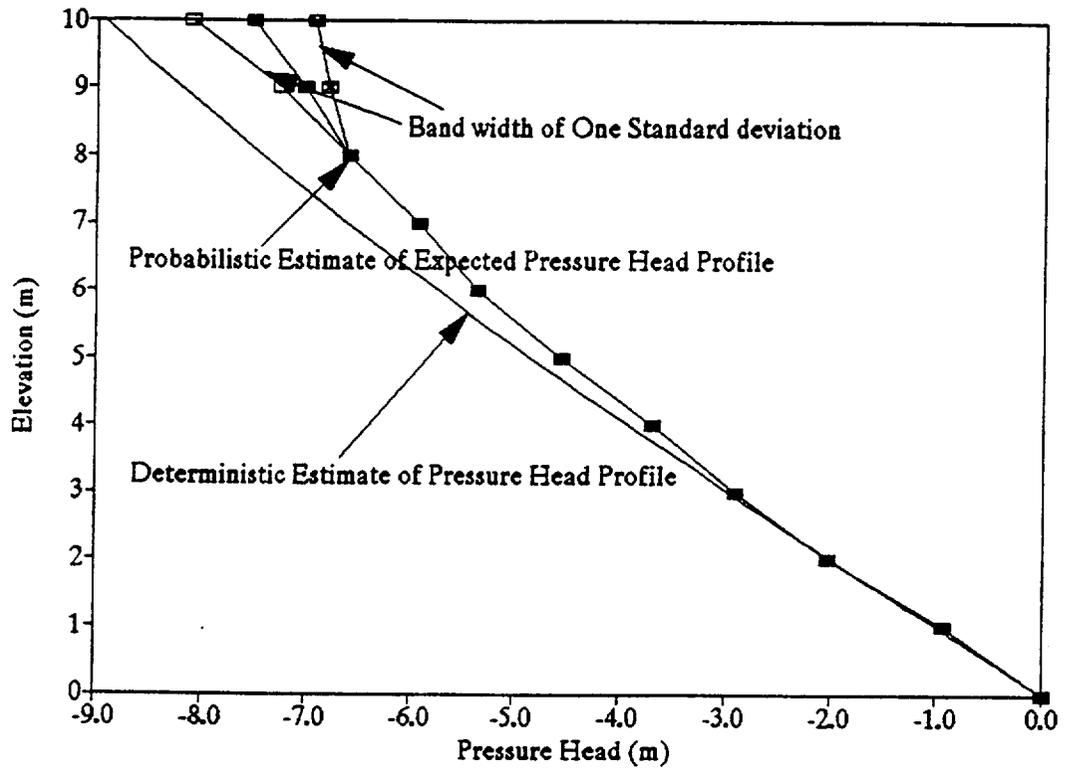


Figure 5.1 Steady-State Pressure Head Profile when Flux rate = 0.1 mm/year, Example I - Model I.

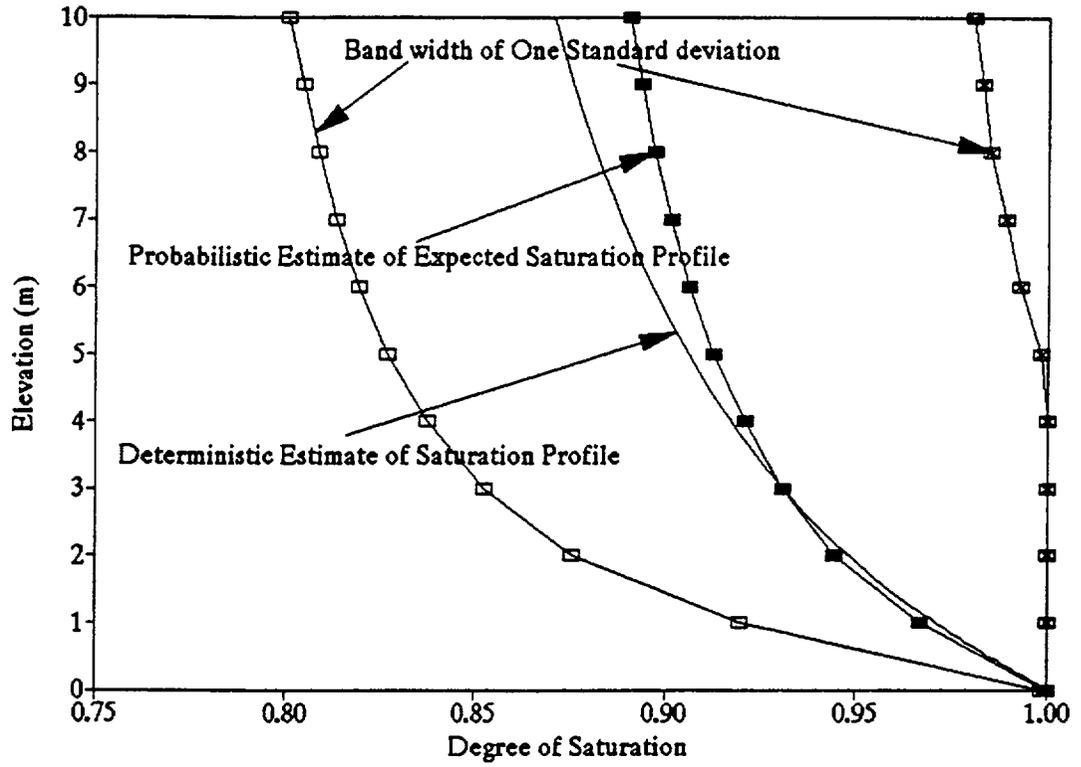


Figure 5.2 Steady-State Degree of Saturation Profile when Flux rate = 0.1 mm/year, Example I - Model I.

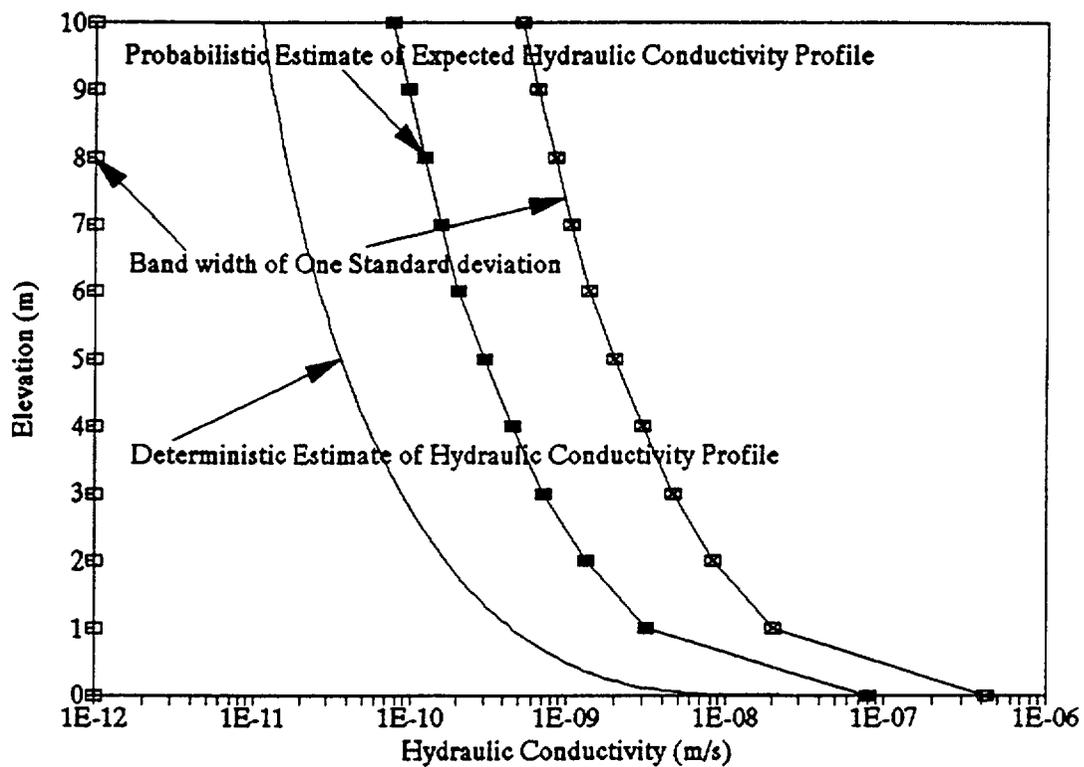


Figure 5.3 Steady-State Hydraulic Conductivity Profile when Flux rate = 0.1 mm/year, Example I - Model I.

5.2.1.2 Example I - Model II

Model II was applied with the same statistical properties for the random variables as given in Table 5.1. With the assumption of equally distributed weighting functions among the point estimate pairs ($p_{i+} + p_{i-} = 1/4 = 0.25$, as stated in equation (3.62)), model II failed to provide a point distribution approximation with all of the eight point estimate locations being positive. For example, model II requires eight point estimate locations in the case of four random variables, i.e., 2 times 4 = 8. Among the eight sets of values (see Table C.3), only four sets of values are possible (see Table C.4). That is, only four sets of point estimate locations satisfy the necessary condition of positive values.

As the original assumption of equally distributed weighting functions failed to provide the positive point estimate locations for all the point estimate pairs, alternative two suggested in section 3.2.4 with unequally distributed weighting functions among the point estimate pairs (i.e., $p_{i+} + p_{i-} \neq 1/4$) was considered. As given in equation (3.62), the weighting function of a point estimate pair is given by the sum of the weighting functions p_{i+} and p_{i-} . In the present example with four random variables there will be four pairs of $p_{i+} + p_{i-}$, where $i=1,2,3,4$. Therefore, the redistribution of the weighting functions were considered among these four pairs. In redistributing the weighting functions, the only condition adopted is that the sum of all the weighting functions pairs is equal to unity. As a result of this, there is no uniqueness. As stated before in section 3.2.4, a possible redistribution to meet the necessary condition of positive point estimates was a trial and error procedure.

In considering a possible redistribution of the weighting function, for the given statistical values (Table 5.1), the random variable K_1 (skewness 6.26 and coefficient of variation of 453.3%) had the greatest influence in producing negative point estimate locations. Therefore, the weighting function for the random variable K_1 was varied over the range of zero and one. With the statistical properties given in Table 5.1, the random variable K_1 failed to provide positive point estimates locations. Therefore, the alternative

approach suggested with unequally distributed weighting functions failed.

As an alternate solution for this case, an appropriate point distribution approximation was then obtained by redistributing the weighting functions among the positive point estimate locations. The resulting point distribution approximations are given in Table C.4. The four possible sets of values were then used to estimate the groundwater travel time (see Table C.5). Model II results for the probabilistic estimates of the groundwater travel time are given in Table 5.2 for the infiltration flux rates of 0.1 mm/year and 1.0 mm/year.

As another alternative to model II, the four random variables K_s , θ_r , α , and N were treated as independent skewed random variables, and the probabilistic estimates for the groundwater travel time were obtained. In treating the variables as independent skewed random variables, the representative statistical properties of the random variables were again taken to be as given in Table 5.1 except there was no correlation structure. Again, a point distribution approximation was obtained by redistributing the weighting functions equally as explained before and only six sets of values were possible among the eight sets of values. The results are given in Table 5.2.

5.2.1.3 Example I - Other Models

Results of Rosenblueth (1975), Lind (1983), and Harr (1989) point estimate methodologies were also obtained for Example I. In applying these point estimate methodologies, the statistical values as given in Table 5.1 were used except the skewness coefficients of the random variables were zero. In obtaining the probabilistic estimates of the groundwater travel time using these methodologies, the possible sets of values considered are tabulated in Tables C.6, C.7, and C.8. Possible sets of values implies that the locations are positive. Therefore, in applying Rosenblueth's methodology, only four sets of values were positive (possible) among the sixteen sets of values (2^4). Similarly, in applying Lind's methodology, only seven sets of values were possible among the eight sets of values ($2 \text{ times } 4 = 8$). In Harr's methodology also only four

sets of values were possible among the eight sets (2 times 4=8). In the later methodologies, the weighting functions for possible locations were redistributed as shown in the respective tabular values in Tables C.6, C.7, and C.8.

All the point estimate methodologies results for the probabilistic estimates of the groundwater travel time are tabulated in Table 5.2 for the infiltration flux rates of 0.1 mm/year and 1.0 mm/year. The single valued deterministic solution was also obtained for the groundwater travel time in the silty clay layer.

Table 5.2 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example I.

Point estimate methodology	Flux rate = 0.1 mm/year		Flux rate = 1.0 mm/year	
	expected value (years)	coefficient of variation	expected value (years)	coefficient of variation
Model I	30,682.62	10.94 %	3,150.31	8.88 %
Model II with redistributed weighting functions	20,669.08	≈ 0 %	2,210.58	≈ 0 %
Model II with uncorrelated variables	29,032.70	5.97 %	3,000.22	5.65 %
Rosenblueth (1975)	26,170.47	≈ 0 %	2,657.94	≈ 0 %
Lind (1983)	29,357.65	≈ 0 %	2,986.36	≈ 0 %
Harr (1989)	26,583.76	6.32 %	2,696.68	5.92 %
Deterministic Solution	30,447.50		3,096.53	

5.2.1.4 Example I - Two Variables Study

A study was performed to explain the effect of the skewness coefficient of the random variables in predicting the probabilistic estimate of the groundwater travel time. For this study, representative statistical values for the random variables K_1 and α were varied; the values for the other random variables (S_r and N) and the correlation structure were kept the same as before. The random variable K_1 was taken to have an expected value as before (0.02 cm/hour), coefficient of variation of 120 %, and skewness coefficient of 2.16. For the random variable α , an expected value as before (0.5 1/m), coefficient of variation of 80 %, and skewness coefficient of 1.51 were taken. The skewness coefficient of the random variables K_1 and α was computed by fitting a beta distribution within the range of 0.0 to 0.1 cm/hour and 0.0 to 15.0 1/m, respectively. An infiltration rate of 0.1 mm/year was used as the input.

The results are given in Table 5.3. As mentioned previously, model I requires sixteen ($2^4=16$) runs of LLUVIA for four random variables. The redistribution of the weighting functions was necessary for model II locations. With the statistical values considered in this study, model II gives only four sets of values with the positive point estimate locations, instead of all the possible eight point estimate locations involved in this methodology. An alternative for model II, independent skewed random variables were also considered. The results of the point estimate methodologies by Rosenblueth (1975), Lind (1983), and Harr (1989) were also obtained and tabulated in Table 5.3.

Table 5.3 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example I (Two Variables Study).

Point estimate methodology	Infiltration flux rate = 0.1 mm/year	
	expected value (years)	coefficient of variation
Model I	30,425.23	11.24 %
Model II with redistributed weighting functions	19,931.48	≈ 0 %
Model II with uncorrelated variables	29,238.01	5.62 %
Rosenblueth (1975)	26,773.52	≈ 0 %
Lind (1983)	29,450.37	≈ 0 %
Harr (1989)	27,073.78	5.72 %

As model I gave the largest coefficient of variation for the groundwater travel time, probabilistic estimates of the pressure head, degree of saturation, and hydraulic conductivity profiles in the silty clay layer were also obtained using this model (see Figures 5.4, 5.5, and 5.6).

5.2.2 Example II - Two Layer Problem of Silt and Silty Clay

A second example of the newly developed methodology considers two layers: silt and silty clay. In this example the layer thickness is also 10 m and the bottom layer is taken to be a silt layer of 5 m. According to the published data by Carsel and Parrish (1988), silt has the highest capillary tension force (see Table 3, Carsel and Parrish, 1988). Therefore, the second application is intended to illustrate the effect of the silt layer in the silty clay layer profile. An infiltration rate of 1 mm/year was selected to maintain the steady state (Report SAND88-0942, Sandia National Laboratories). Each layer is assumed to be independent.

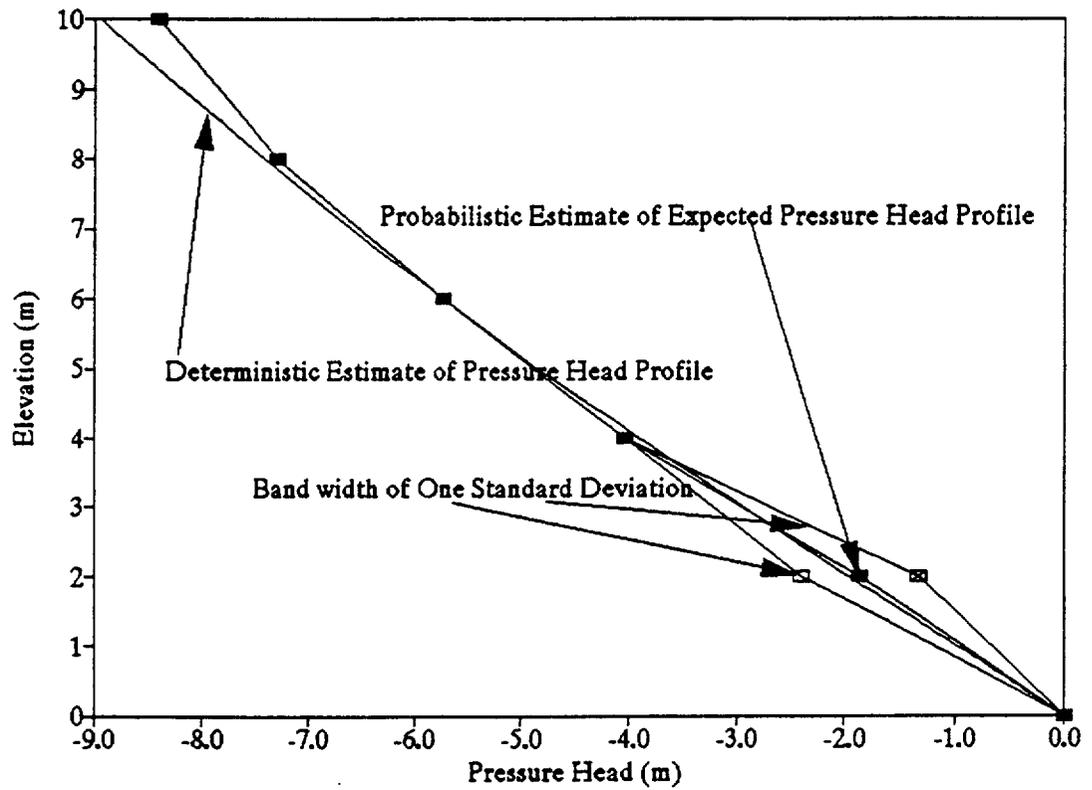


Figure 5.4 Steady-State Pressure Head Profile when Flux rate = 0.1 mm/year, Example I (Two variables study) - Model I.

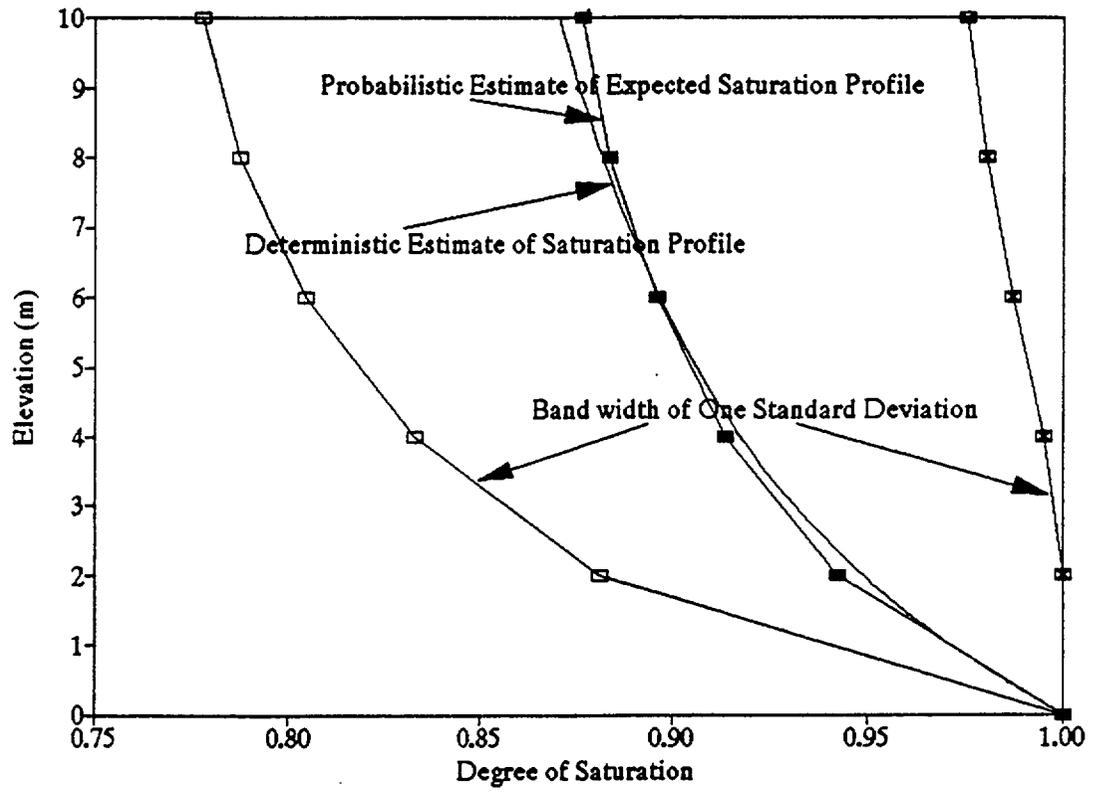


Figure 5.5 Steady-State Degree of Saturation Profile when Flux rate = 0.1 mm/year, Example I (Two variables study) - Model I.

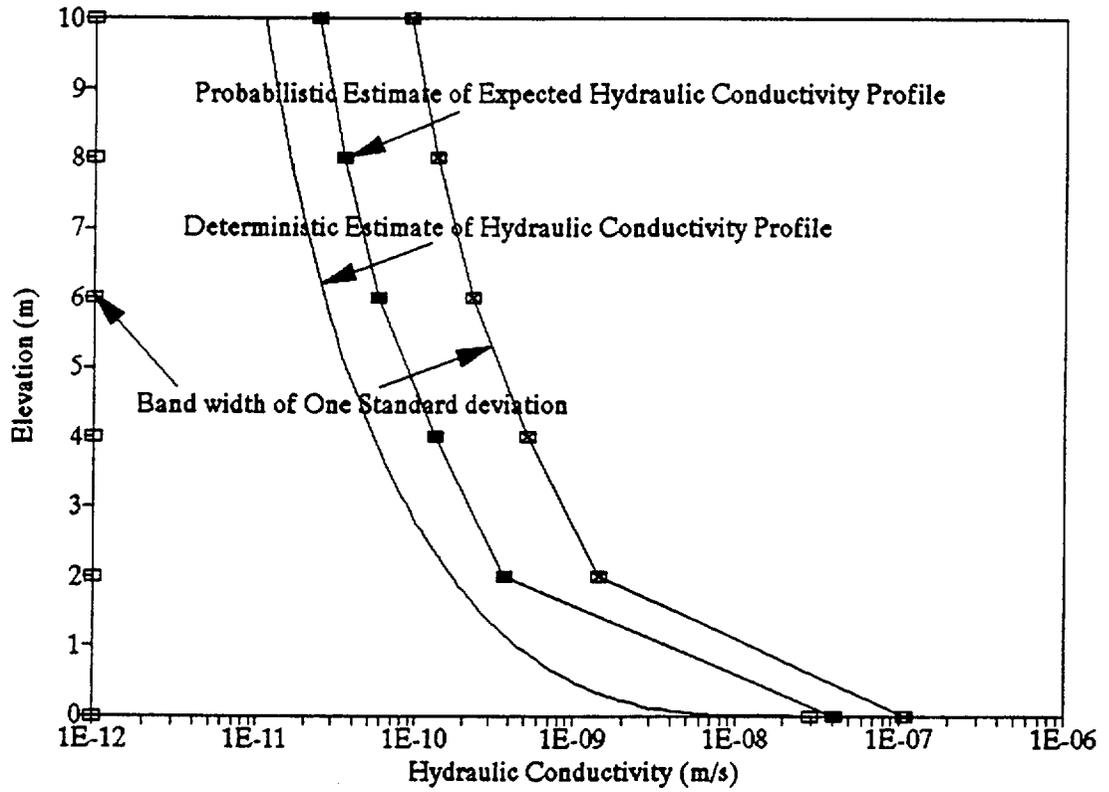


Figure 5.6 Steady-State Hydraulic Conductivity Profile when Flux rate=0.1 mm/year
Example I (Two variables study) - Model I.

The representative values for the silty clay layer are again those in Table 5.1. For the silt layer the values taken from Carsel and Parrish (1988) are given in Table 5.4. The skewness coefficient of the silt layer random variables K_s , S_r , α , and N were calculated as before by fitting a beta distribution with the available data of expected value, coefficient of variation, and limits of variation given by Carsel and Parrish. The global correlation matrix is given in Appendix C.2.

Table 5.4 Representative Values of the Random Variables for Silt.

Random variable	K_s (cm/hr)	S_r	α (1/m)	N
expected value	0.25	0.034	1.6	1.37
coefficient of variation (%)	129.9	21.6	64.7	8.5
skewness	1.79	0.17	0.97	0.27

The correlation matrix for the silt layer is

$$\rho = \begin{bmatrix} K_s & \theta_r & \alpha & N \\ 1.0 & -0.204 & 0.984 & 0.466 \\ -0.204 & 1.0 & -0.20 & -0.61 \\ 0.984 & -0.20 & 1.0 & 0.551 \\ 0.466 & -0.61 & 0.551 & 1.0 \end{bmatrix}$$

5.2.2.1 Example II - Model I

Using the statistical values as given in Tables 5.1 and 5.3 and the global correlation matrix, model I was applied to the LLUVIA code. The total number of runs required by model I is 256 ($2^8 = 256$). For each set of values, a LLUVIA input file was

programmed and the corresponding groundwater travel time was computed. The expected value and coefficient of variation of the groundwater travel time were computed by weighting the travel time output by the corresponding weighting functions. The results are given in Table 5.5.

Probabilistic estimates of the pressure head, degree of saturation, and hydraulic conductivity profiles of the two layer problem were also obtained using model I (see Figures 5.7, 5.8, and 5.9).

5.2.2.2 Example II - Model II

With the information given for the two layer problem, probabilistic estimates for the groundwater travel time were also obtained using model II. For two layers, model II requires sixteen sets of values ($2 \text{ times } 8 = 16$, where 8 is the number of random variables). However, for the statistical parameters given in the two layer problem, only ten sets of values satisfy the requirement of positive point estimate locations. These possible ten point locations were obtained with the assumption of equally distributed weighting functions among the point distribution pairs (see section 3.2.3). The estimates for the ten sets are given in Table 5.5. The alternate solution with a redistribution of the weighting functions which specifies the condition of positive point estimate locations for all the sixteen point estimate locations of the given two layer problem again failed. Consequently, probabilistic estimates using model II was undertaken for the condition of independent skewed random variables. With independent condition also, a possible point distribution approximation was obtained with only twelve sets of values, instead of all the sixteen point locations of model II.

5.2.2.3 Example II - Other Models

Probabilistic estimates of the groundwater travel time in the two layer problem were also obtained using the point estimate methodologies by Rosenblueth (1975), Lind (1983), and Harr (1989). In applying these methodologies to the two layer problem,

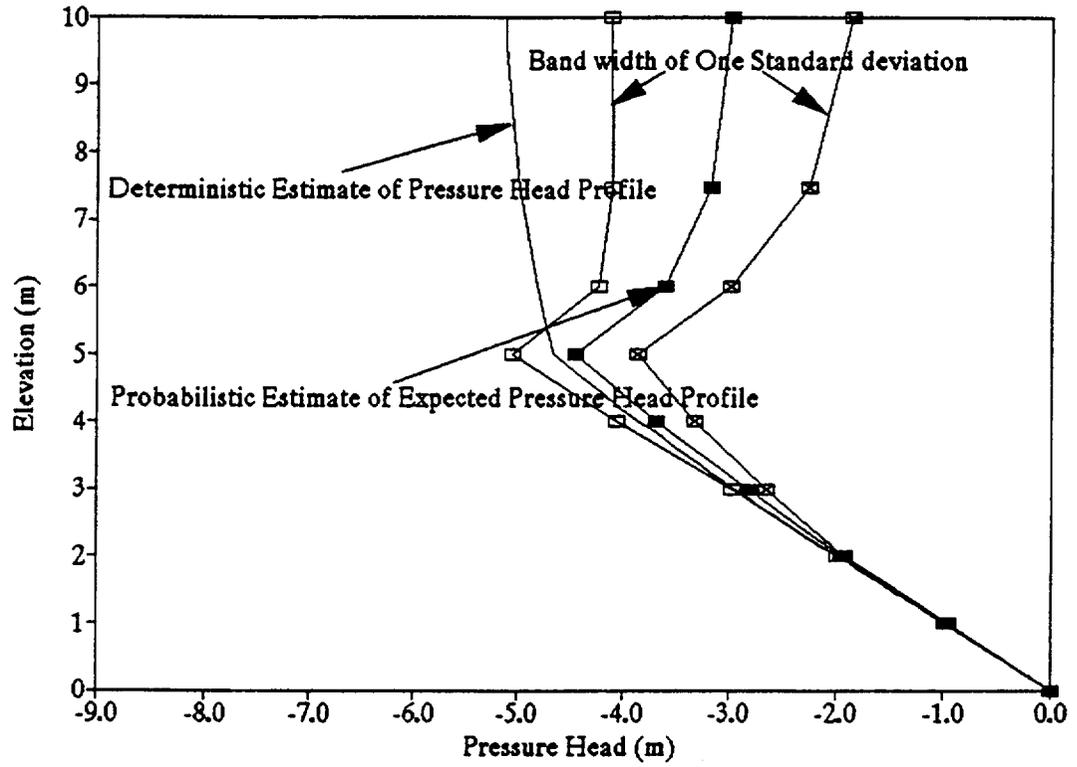


Figure 5.7 Steady-State Pressure Head Profile when Flux rate = 1.0 mm/year, Example II - Model I.

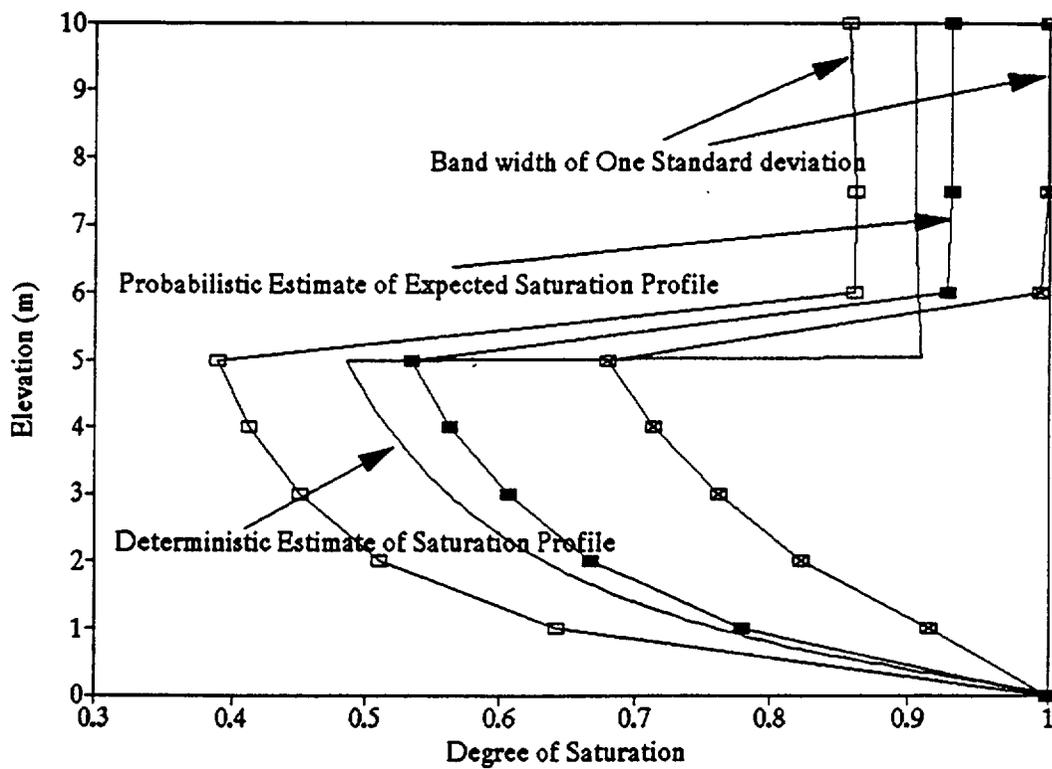


Figure 5.8 Steady-State Degree of Saturation Profile when Flux rate = 1.0 mm/year, Example II - Model I.

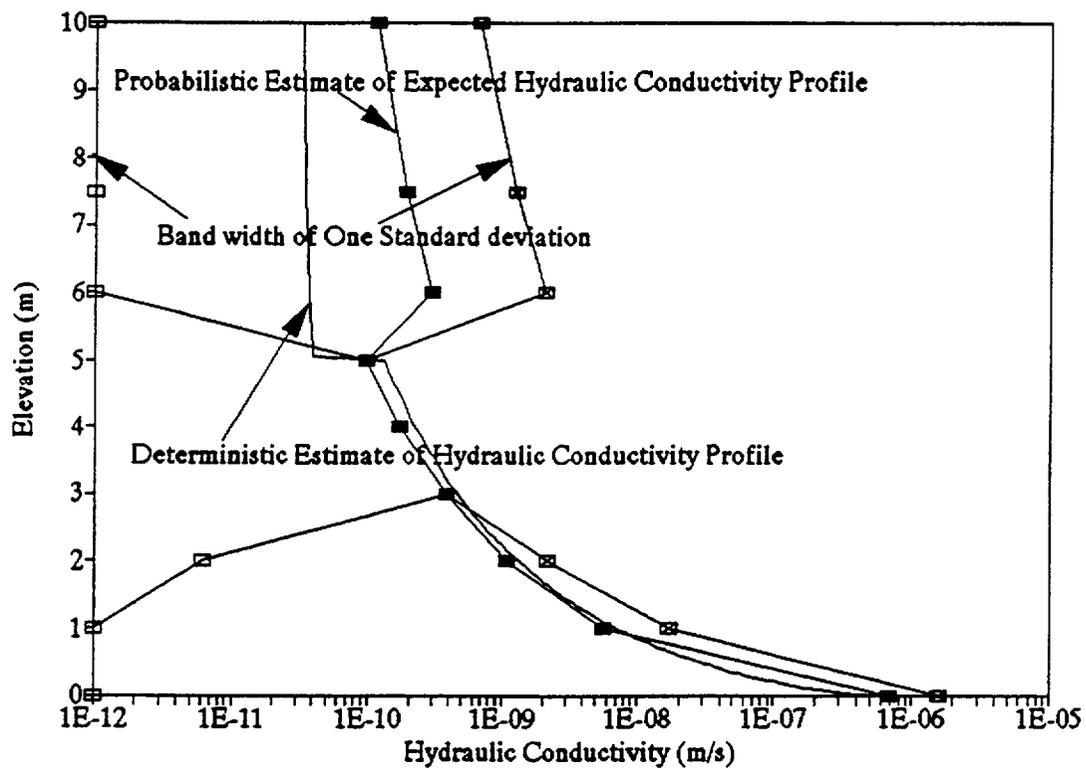


Figure 5.9 Steady-State Hydraulic Conductivity Profile when Flux rate=1.0 mm/year
Example II - Model I.

Rosenblueth's, Lind's, and Harr's methodologies respectively provide 256 (2^8), 16 (2×8), and 16 (2×8) sets of values. Among these sets of values, the possible sets of values were again selected for the condition of positive point estimate locations. With this condition, Rosenblueth's, Lind's, and Harr's methodologies respectively gave 32, 14, and 8 possible sets of values for the LLUVIA run. With these selected sets of values, the probabilistic estimates of the groundwater travel time were computed. The results of all the methodologies are given in Table 5.5.

For example II, the deterministic solution of the groundwater travel time is 2896.65 years for the influx rate of 1.0 mm/year.

Table 5.5 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example II.

Point estimate methodology	Infiltration flux rate 1.0 mm/year	
	expected value (years)	coefficient of variation
Model I	3,008.44	11.43 %
Model II with redistributed weighting functions	2,389.11	22.41 %
Model II with uncorrelated variables	2,903.33	10.22 %
Rosenblueth (1975)	2,572.29	6.76 %
Lind (1983)	2,831.26	10.48 %
Harr (1989)	2,620.62	9.72 %
Deterministic solution	2,896.65	

5.2.3 Example III - Four Layer Problem of Silt, Clay, and Silty Clay

As another example, a four layer problem was considered with a clay layer of 0.1 m thickness placed in the middle of the silt layer (bottom layer) presented in the second example. The total thickness of the flow domain is 10 m (see Figure 5.10). In particular, this example is interesting because the published data by Carsel and Parrish (1988) gives very low air entry value parameter (α) for the clay layer. Therefore, this example will explain the effect of having a clay seam in the silt layer. An infiltration flux rate of 1.0 mm/year (Report SAND88-0942, Sandia National Laboratories) was selected and the groundwater travel time is computed from the top of the silty clay layer to the water table at the bottom of the silt layer.

The representative values of silty clay and silt layers are those taken previously. The clay layer representative values taken from Carsel and Parrish (1988) are given in Table 5.6. The skewness coefficients of the random variables for the clay layer were again determined by fitting a beta distribution.

Table 5.6 Representative Values of the Random Variables for Clay.

Random variable	K_r (cm/hr)	S_r	α (1/m)	N
expected value	0.20	0.068	0.8	1.09
coefficient of variation (%)	210.3	49.9	160.3	7.9
skewness	3.40	0.14	2.64	0.31

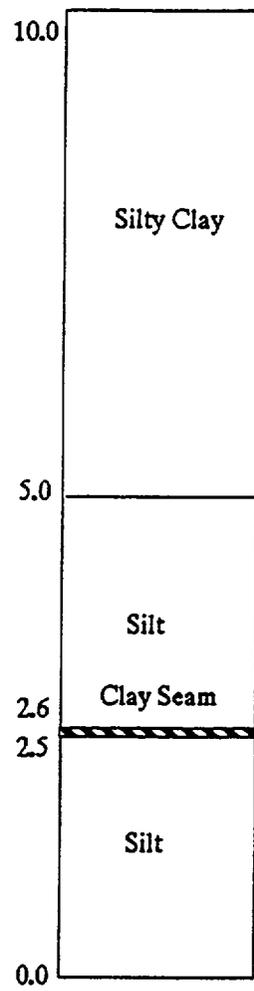


Figure 5.10 One-Dimensional Flow Domain of Example III.

The correlation matrix for the clay layer is

$$\rho = \begin{bmatrix} K_s & \theta_r & \alpha & N \\ 1.0 & 0.972 & 0.948 & 0.908 \\ 0.972 & 1.0 & 0.890 & 0.819 \\ 0.948 & 0.890 & 1.0 & 0.910 \\ 0.908 & 0.819 & 0.910 & 1.0 \end{bmatrix}$$

In the probabilistic analysis the properties between layers are independent: the total number of random variables is sixteen. Model I requires $2^{16} = 65,536$ number of runs of LLUVIA which are beyond reason; for the same system, model II requires only $2 \text{ times } 16 = 32$ number of runs. Consequently, only model II was employed.

Using model II, 18 sets of values are possible among the 32 point estimate locations. The probabilistic estimates of the groundwater travel time were then obtained with these possible sets of values. As in the previous examples, model II point distribution approximation was possible with equally redistributed weighting functions. Probabilistic estimates were also obtained using the point estimate methodologies by Lind (1983) and Harr (1989). Both of these methodologies provide 32 point estimate locations for the problem considered in example III. Among these 32 point estimates, 27 sets of values were possible in Lind's methodology, and in Harr's methodology 11 sets of values were possible. The results are given in Table 5.7. The deterministic solution of the groundwater travel time for example III is 2903.45 years.

Table 5.7 Probabilistic Estimates of the Groundwater Travel Time using Point Estimate Methodologies - Example III.

Point estimate methodology	Infiltration flux rate 1.0 mm/year	
	expected value (years)	coefficient of variation
Model II with redistributed weighting functions	2,800.17	9.26 %
Lind (1983)	2,868.71	8.14 %
Harr (1989)	2,720.67	9.30 %
Deterministic solution	2,903.45	

CHAPTER 6

DISCUSSION OF THE NEW POINT ESTIMATE MODELS

The preceding chapters presented and demonstrated the new point estimate models I and II for uncertainty analyses of systems with correlated and skewed random variables. Comparisons were made with the results of Rosenblueth (1975), Lind (1983), and Harr (1989) point estimate methodologies, the Monte Carlo simulation technique, and also with the exact solution for a special case. Application was also made to the flow code LLUVIA. In the present chapter various aspects of the models will be discussed.

6.1 Applicability of Models I and II

The purpose of the present study was to provide a methodology capable of accommodating functions of many correlated and skewed random variables into uncertainty analyses. Given were information concerning the expected values, standard deviations, skewness coefficients, and correlation coefficients of the random variables. Models I and II were devised to incorporate this information into uncertainty analyses. Existing methods are incapable of accommodating either the correlation structure or skewness or both.

6.1.1 Condition of Uniqueness

It was shown (see sections 3.1.3 and 3.2.3) that uniqueness of a solution cannot be obtained given only the knowledge of first three statistical moments and the correlation structure of the N random variables for $N > 2$. To overcome this deficiency,

the developed models introduced assumptions using information in the "least-biased" sense. Hence, in model I, the point estimates and the corresponding weighting functions were taken to be concentrated at the corners of a multi-dimensional hyperprism of N -dimensional space defined by the random variables. For model II, the weighting functions are taken to be equally distributed among the point estimate pairs (equation (3.62)).

The lack of uniqueness of a solution for $N > 2$ is also the case in Rosenblueth's (1975), Lind's (1983), and Harr's (1989) point estimate methods. In the Monte Carlo simulation technique complete probability distribution functions must be assumed "a priori." The necessity of doing so is often overlooked in the literature wherein the Monte Carlo simulations are thought to be the standard for comparisons. The necessary assumption of complete probability density functions is the major difference between the Monte Carlo simulation techniques and point estimate methodologies. Conceptually, if a probability distribution function is invoked all orders of statistical moments can be computed. Hence, the requirement of uniqueness could also be met by point estimate methodologies by assuming a probability distribution.

6.1.2 Requirement of Positive Locations

If variables are defined to be positive by their physical attributes then this requirement can be treated as an additional piece of information. This led to the assertion that negative point estimate locations should be disregarded. Examples presented in this study showed that model I point distribution approximation given in equation (3.38) meets the necessary condition of positive locations for highly skewed, scattered, and correlated random variables. For model II, alternative solutions had to be sought to meet the necessary requirement for variables having high coefficients of variation (i.e., coefficients of variation greater than 100%). Two alternatives for model II were suggested in section 3.2.4. As stated in the alternatives, a possible solution was first considered with equally distributed weighting functions. When this failed, redistribution of the weighting functions was considered among the selected positive

locations in order to satisfy the condition that the sum of the weighting functions must be equal to unity. Motivating the first alternative was the principle of maximum entropy which states that the uniform distribution is the least based solution for a system with no additional information (Table 2.5.1, Harr 1987).

6.2 Effect of Skewness Coefficient in Point Estimate Methodologies

The models were first used to demonstrate the effect of incorporating the skewness coefficients of random variables (see Chapter 4) for the function $y = x_1 x_2 / x_3$. Three cases were considered: 1) correlated and skewed, 2) independent and skewed, and 3) independent lognormal input variables.

6.2.1 Results for Skewed and Correlated Variables

Results given in Table 4.6 illustrated the effect of incorporating correlated and skewed random variables in the point estimate methodologies; specifically, in estimating the coefficients of variation for the function. Models I and II gave lower coefficients of variation than Rosenblueth (1975), Lind (1983), or Harr (1989) methodologies. Differences were also found between the two models. Recognizing that the developed models are not based on similar assumptions, differences in estimates should be expected. However, for the example considered with positively correlated and skewed variables, the difference was very high. Therefore, the purpose of the next case was to compare the difference when the variables are not correlated.

Comparisons for similar sets of assumptions were investigated. Model I and the point distribution pattern used by Rosenblueth (1975) are similar in this respect. Model II and the point distribution pattern used by Lind (1983) are similar. Hence, the comparison of similar models was thought to show the influence of the skewness coefficients. As expected, the results (see Table 4.6) showed the influence of skewness, specifically in estimating the coefficient of variation of the function. It was found that when accounting for skewness, lower estimates were obtained for the coefficient of

variation than when it is taken to be zero. Note that this result was for positively correlated variables.

6.2.2 Results for Independent and Skewed Variables

For independent and skewed random variables (see section 4.2.2), models I and II gave comparable results but higher coefficients of variation for the function than Rosenblueth (1975), Lind (1983), or Harr (1989) point estimate methodologies. As might be expected, skewness coefficients of the random variables influenced the estimates of the coefficient of variation.

Comparative studies were also made using the Monte Carlo simulation technique. However, the significance of such comparisons should be carefully noted. Monte Carlo simulations do not consider the correlation structure among the variables. In addition, they require complete " *a priori*;" knowledge of the entire distribution of all the variables. Recall that the hypothesis of the present thesis specifies only knowledge of the first three statistical moments and the correlation structure of the variables. Complete probability density functions of the random variables x_1 , x_2 , and x_3 must be defined to perform Monte Carlo simulations. Consequently, comparison between the results of the Monte Carlo simulation technique and the point estimate methodologies should be considered in a qualitative sense only.

As expected, the results given in Table 4.8 (section 4.2.2) show that the number of simulations performed in the Monte Carlo simulation technique dramatically affect the estimate of the statistical moments of the function. It is more noticeable when computing the coefficient of variation of the function. To provide a Monte Carlo simulation with a 95 % confidence level, two thousand times the 10,000 number of trials performed in Table 4.8 would have to be performed!

For the Monte Carlo simulation results given in Table 4.8, the coefficient of variation of the function should lie between 70 % and 80 %. When comparing the

results of the Monte Carlo simulation approach and the point estimate methodologies approximations for independent and skewed variables (see Table 4.7), it was shown that the results of the new models and the Monte Carlo simulation results are within range of each other. On the other hand, models I and II approximations for case one (see Table 4.6), where the variables are correlated and skewed, are not in line with the Monte Carlo results. Recall that the correlation structure of the variables is not taken into account in the Monte Carlo simulation approach.

6.2.3 Results for Independent Lognormal Variables

Figures 4.1 and 4.2 compared the point estimate results for independent variates (section 4.2.3) with the exact solution for lognormal variates. The third case considered is a special situation. Recall, complete lognormal probability density functions were assumed. On the whole, when comparing the results of the point estimate methodologies with the lognormal solution, the new models under-predict and Rosenblueth (1975), Lind (1983), and Harr (1989) point estimate methodologies over-predict. The effect of incorporating the skewness coefficients in the new models is clearly noticeable when the skewness coefficient of the random variables is greater than unity or when the coefficient of variation of the random variables is greater than 40 % (Figure 4.1). As shown in Figures 4.1 and 4.2, the results given by Lind (1983) and Harr (1989) methodologies are the same; because when the variables are statistically independent, the point distribution models for both methodologies are the same.

According to the example presented, the difference between models I and II estimates increases as the skewness (which is a function of coefficient of variation in this example) of the random variables increases. However, when compared to the difference with the results presented for correlated and skewed random variables, it is small. From the three cases of input variables considered, the following conclusions can be drawn: 1) models I and II indicated the influence of skewness in obtaining the probabilistic estimates for both independent and correlated variables 2) the observed influence was a combined function of both the skewness and the correlation structure, and 3) as expected,

models I and II showed differences in the approximation, but the difference was obvious for correlated and skewed variables. Specifically, this example is for positively correlated variables.

6.3 Application of New Models to flow code LLUVIA

Both models were applied to LLUVIA to estimate the steady-state groundwater travel time and to quantify the effect of uncertainty of the material parameters. The two models, as employed in the present work, considered the uncertainty associated with a system of correlated and skewed random variables. The required information is the first three statistical moments: expected values, standard deviations, and skewness coefficients and the correlation structure of the random variables.

As available information is a vital factor in engineering designs, the two models have a special significance by providing the approximations with a minimum of practical data available. Four random variables within each layer scale LLUVIA: the saturated hydraulic conductivity (K_s), residual saturation (S_r), and van Genuchten model parameters describing air entry parameter (α) and the slope of the desaturation curve (N). Applications used the data reported by Carsel and Parrish (1988). According to these data, the four random variables are correlated and skewed. Very low infiltration flux rates were considered in these examples to maintain the steady-state flow (Report SAND88-0942, Sandia National Laboratories).

6.3.1 Single Layer of Silty Clay

The first application of the models to LLUVIA (see section 5.2.1) was the determination of the travel time in a silty clay layer of 10 m thickness. The random variables were both positively skewed (except S_r) and positively correlated. Among the four random variables, K_s was highly skewed and scattered (see Table 5.1). Infiltration flux rates of 0.1 mm/year and 1.0 mm/year were considered.

Table 5.2 gives the estimates of all the point estimate methods. As expected, these estimates are seen to be dependent upon the methods. For example, model I gave the highest estimate for the expected groundwater travel time and coefficient of variation. In view of this, the estimates indicated the influence of accounting for both the skewness and the correlation structure. The predicted coefficient of variation for the steady-state travel time is low when compared to the amount of uncertainty involved in the input parameters K_s and α (see Table 5.1).

In using model II, with equally distributed weighting functions, available information was used in the "least-biased" manner. However, model II estimates were found not to be satisfactory. That is, computed estimates gave negative variances for both flux rates. This is believed to be due to the very high skewness and scatter of the variable K_s , which had the greatest influence in computing model II point estimates locations. Negative variances were also obtained in Rosenblueth (1975) and Lind (1983) methodologies (see Table 5.2). As an alternative in using model II, uncorrelated variables were considered (third row entry of Table 5.2). Comparison of the results with model I indicated the influence of the correlation structure as anticipated. In this case the correlation structure gave higher estimates than when it was neglected. This was shown for both flux rates (0.1 mm/year and 1.0 mm/year).

Comparison of model I estimates with Harr's methodology (sixth row of Table 5.2) clearly show the influence of skewness in estimating the travel time. For the case presented, the influence of positive skewness gave higher estimates for the expected travel time and coefficient of variation.

The results in Table 5.3 demonstrated the influence of saturated hydraulic conductivity (K_s) and air entry parameter (α), which were taken to have less skewness and lower coefficients of variation than the values assumed for the results given in Table 5.2 (see section 5.2.1.4). Comparison of the results given in Tables 5.2 and 5.3 illustrate the influence of both the skewness and correlation structure. For instance, the

comparison of model I results shows a decrease in the travel time by having lower values of skewness and coefficient of variation. This finding agrees with the observation that when the skewness is taken to be zero, lower estimates were obtained for the travel time. Note that these observations were for positively skewed variables. But the comparison of model I estimates for coefficient of variation illustrated that the estimates were greatly biased when the variables were taken to be symmetrical.

Another example of illustrating the influence of the correlation structure is that when it was neglected (i.e., model II (uncorrelated) estimates) with less skewness and lower coefficient of variation an increase in the travel time was observed which was not the case with model I estimates. Model I showed a decrease.

Comparison of the deterministic travel time for both flux rates agreed well with model I results than the other methods (see Table 5.2). However, model I estimates have the added advantage in assessing the uncertainty. According to the profiles shown in Figures 5.1, 5.2, and 5.3, the steady state hydraulic conductivity profile displayed a coefficient of variation in the range 400 - 600 %, which is compatible with the uncertainty involved in the input parameters K_s and α . The coefficients of variation in the pressure head and degree of saturation profile were in the range 3 - 10 %, which is very small when compared to K_s and α . The characteristics of the two variables study profiles also were found to be the same (see Figures 5.4, 5.5, and 5.6).

6.3.2 Two Layers of Silt and Silty Clay

This example was intended to illustrate the effect of adding a silt layer into the silty clay layer profile of the previous example (see section 5.2.2). The effect observed was a decrease in the expected ground travel time in the range of 70 - 140 years. The coefficient of variation of the travel time was increased by 2 - 5 %. Of special note is that the coefficient of variation of the steady-state groundwater travel time is very low when compared to the uncertainty in the input parameters K_s and α (see Tables 5.1 and 5.4).

The results given in Table 5.5 were for positively and negatively correlated random variables but the variables were only positively skewed. The results given support the conclusion as drawn before that the skewness and correlation structure have an important influence on estimates. As noted before for correlated random variables (section 6.2.1), results given by models I and II approximations were not the same. According to the results presented, model I approximations accommodated positively skewed and positively and negatively correlated variables by giving higher estimate for the expected travel time and lower coefficient of variation than model II estimates. The difference between the two models estimate for coefficient of variation was almost one hundred percent whereas in the travel time 20 % difference was observed.

The deterministic groundwater travel time is respectively at one third and one standard deviations from models I and II estimates. As in the previous problem, the scatter in the hydraulic conductivity profile (coefficient of variation 100 - 600 %, see Figure 5.9) was compatible with the uncertainty involved in the input parameters. Also, as before, the pressure head and degree of saturation profiles (Figures 5.8 and 5.9) showed very low scatter (coefficient of variation 5 - 30 %).

6.3.3 Four Layers of Silt, Clay, and Silty Clay

This example was intended to illustrate the effect of a more complex system: a clay seam with the silt layer of the previous example (see section 5.2.3). In this example, model II was within the practical range with respect to the number of simulations required in the analysis. As in the previous case, the variables were positively and negatively correlated but only positively skewed. According to model II (redistributed weighting functions), the effect of having a clay seam in the silt layer increased the expected travel time by 400 years and the coefficient of variation was decreased by 15 % (see Table 5.7). Lind's and Harr's methodologies respectively showed an increase of 40 and 100 years, whereas in the deterministic solution the increase was only 7 years. Lind's and Harr's methodologies showed very small differences in the coefficients of variation of the groundwater travel time. Therefore, the present models are clearly

illustrating the influence of accounting for the skewed and correlated random variables. However, the influence seems to be problem specific. That is when only the results given in Table 5.7 are compared, the effect of incorporating the skewness is small. The estimates of the three methodologies are comparable.

As in the other two examples, the coefficient of variation of the travel time is low when compared to the amount of uncertainty involved in the input parameters K_i and α (see Tables 5.1, 5.4, and 5.6). The deterministic travel time is within the bandwidth of 0.4 standard deviation of model II (redistributed weighting functions) estimate.

CHAPTER 7

SUMMARY AND CONCLUSIONS

Two point estimate models have been formulated that are directly suited to evaluating the uncertainty for systems with many correlated and skewed random variables. It was assumed that information is available concerning the first three statistical moments of the N random variables: expected values, standard deviations, and skewness coefficients and their correlation structure. On the basis of the developed procedures, it was shown that:

1. The proposed models will provide point estimates that compute the statistical moments of functions of correlated and skewed random variables.
2. For a function of N random variables, a maximum of 2^N point estimates are required for model I and $2N$ for model II. A computer program was developed in this study incorporating the algorithms of the two models.
3. Uniqueness of solution cannot be obtained given only knowledge of first three statistical moments and the correlation structure for N random variables with $N > 2$.
4. Additional assumptions were introduced in these models to render the models determinate. In model I, the point estimates and the corresponding weighting functions were assumed to be concentrated at the corners of a multi-dimensional hyperprism of N -dimensional space defined by the random variables. In model II, the weighting functions were assumed to be equally distributed among the point estimate pairs.
5. Alternative procedures were provided if the original assumptions of these

models failed to meet the necessary condition of providing positive point estimates locations.

6. Step-by-step procedures were presented to lead the reader through the developed procedures.

7. Examples were presented that illustrated the use of the models and compared results with those obtained using the point estimate procedures of Rosenblueth (1975), Lind (1983), and Harr (1989); as well as the Monte Carlo simulations.

7.1. The results of these examples showed the effect of incorporating the skewness coefficient of the random variables into the point estimate methodologies. The influence of skewness coefficient was very apparent when estimating the coefficient of variation.

7.2. Comparison made with exact solution (lognormal solution) showed that the new models obtained better approximations for highly skewed random variables than Rosenblueth (1975), Lind (1983), or Harr (1989) methodologies.

7.3. The Monte Carlo simulation results agreed with the new models results for independent and skewed random variables. But for correlated and skewed random variables, as expected, the results were not comparable.

8. The models were applied to the flow code LLUVIA to estimate the steady-state groundwater travel time in an unsaturated medium. Four random variables within each layer scale LLUVIA: saturated hydraulic conductivity, residual saturation, and van Genuchten model parameters describing air entry parameter and the slope of the desaturation curve. Three cases were presented using the parameters reported by Carsel and Parrish (1988). The saturated hydraulic conductivity and van Genuchten model air entry parameter are highly skewed and exhibit high coefficients of variation: all the four random variables were correlated. It was demonstrated that:

8.1 The coefficient of variation of the steady-state groundwater travel time is low (10 - 30 %) when compared to the amount of uncertainty

involved in the input parameters, specifically in the saturated hydraulic conductivity and in the air entry parameter.

8.2 The hydraulic conductivity profiles showed coefficients of variation within the range of 100 to 600 %. The pressure head and degree of saturation profiles showed coefficients of variation within the range of 5 to 30 %. For the examples considered, the deterministic groundwater travel time was within one standard deviation band width of models I and II estimates.

8.3 Model I gave higher estimates for the expected groundwater travel time than model II. On the other hand, model II gave higher estimates for the coefficient of variation of the groundwater travel time.

9. In sum, knowing only the expected values, standard deviations, skewness coefficients, and correlation coefficients of the random variables, models I and II provide methodologies that address the problem of incorporating the correlated and skewed random variables into uncertainty analyses.

10. Examples presented showed that the generalized point distribution algorithm provided by model I (equation (3.38)) meets the necessary condition of positive point estimates locations for highly skewed and scattered random variables.

11. For a system of very many random variables, model II is recommended as the methodology requires only a maximum of $2N$ point estimates.

CHAPTER 8

SUGGESTIONS FOR FUTURE RESEARCH

1. The examples presented in this dissertation clearly establish the importance of accommodating both the skewness and correlation. Therefore, the success associated with the application of the models suggest that they be applied to other problems.
2. The present study presented two models to incorporate correlated and skewed random variables. According to the examples presented, model I generalized algorithm met the necessary condition of positive locations. However, in model II the assumption of equally distributed weighting functions failed to meet this necessary condition in the cases highly skewed and scattered random variables. Therefore, a study should be under taken to examine other distribution of weighting functions.
3. Alternate algorithms for the models can be considered. Recall that uniqueness of a solution cannot be obtained. Therefore, the differences in estimates associated with the algorithms can be studied in more detail.
4. Alternate point distribution models can be developed to incorporate correlated beta variates. An attempt was made in the present study to expand Harr's methodology. However, complications arose. The transference from the eigenspace back to the solution space for beta variables required more information than was available. However, it appears that there might be a way to connect the eigenspace and the solution space. Successful accomplishment of this would provide approximations of higher order.

LIST OF REFERENCES

LIST OF REFERENCES

- Agterberg, F.P. (1974). *Geomathematics*. Elsevier Scientific Publishing Co., New York, N.Y.
- Alonso, E.E. (1976). Risk analysis of slopes and its application to slopes in Canadian sensitive clays. *Geotechnique*, 26(3), 453-472.
- Anderson, J., and Shapiro, A.M. (1983). Stochastic analysis of one-dimensional steady state unsaturated flow: A comparison of Monte Carlo and perturbation methods. *Water Resources Research*, 19, 121-133.
- Ang, A. H-S., and Tang, W.H. (1975). *Probability concepts in engineering planning and design* (Vol 1). John Wiley and Sons, Inc., New York.
- Athanasίου-Grivas, D., and Stiefel, U.C. (1983). Reliability assessment based on point estimates of probability moments. *Int. Conference on Mathematical Modelling*, Zurich, 289-294.
- Ayyub, B.M., and Haldar, A. (1984). Practical structural reliability techniques. *Journal of Structural Engineering*, 110(8), 1707-1724.
- Bakr, A.A., Gelhar, L.W., Gutjahr, A.L., and MacMillan, J.R. (1978). Stochastic analysis of spatial variability in subsurface flows 1. Comparison of one- and three-dimensional flows. *Water Resources Research*, 14(2), 263-271.
- Bear, J. (1972). *Dynamics of fluids in porous media*. American Elsevier.
- Bear, J. (1979). *Hydraulics of groundwater*. McGraw-Hill Book Co., New York.
- Benjamin, J.R., and Cornell, C.A. (1970). *Probability, statistics, and decision for civil engineers*. McGraw-Hill Book Co., New York.
- Brooks, R.H., and Corey, A.T. (1964). Hydraulic properties of porous media. *Colorado State Univ. Hydrol.* 3, 27.

- Brooks, R.H., and Corey, A.T. (1966). Properties of porous media affecting fluid flow. *Journal of the Irrigation and Drainage Division*, ASCE, 92(IR2), 61-88.
- Burdine, N.T. (1953). Relative permeability calculations from pore size distribution data. *Petroleum Transactions*, AIME, 198, T.P.3519, 71-78.
- Campbell, G.S. (1974). A simple method for determining unsaturated conductivity from moisture retention data. *Soil Science*, 117, 311-314.
- Carrigan, C.R., Bixler, N.E., Hopkins, P.L., and Eaton, R.R., (in review) *COVE2 Bench-marking calculations using NORIA* (SAND88-0942). Sandia National Laboratories, Albuquerque, NM.
- Carsel, R.F., and Parrish, R.S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5), 755-769.
- Carsel, R.F., Parrish, R.S., Jones, R.L., Hansen, J.L., and Lamb, R.L. (1988). Characterizing the uncertainty of pesticide movement in agriculture soils. *Journal of Contaminant Hydrology*, 2, 111-124.
- Chameau, J. L, and Clough, G.W. (1983). Probabilistic pore pressure analysis for seismic loading. *Journal of Geotechnical Engineering*, ASCE, 109(4), 507-524.
- Chowdhury, R.N., and Athanasiou-Grivas, D. (1982). Probabilistic model of progressive failure of slopes. *Journal of Geotechnical Engineering Division*, ASCE, 108(6), 803-819.
- Cornell, C.A. (1971). First-order uncertainty analysis of soils deformation and stability. *Proc. First International Conference on Applications of Statistics and probability in Soil and Structural Engineering*, Hong Kong, 129-144.
- Corotis, R.B., Azzouz, A.S., and Krizek, R.J. (1975). Statistical evaluation of soil index properties and constrained Modulus. *Proc. Second International Conference on Applications of Statistics and probability in Soil and Structural Engineering*, Aachen, 2, 273-293.
- Cox, D.C., and Baybutt, P. (1981). Methods for uncertainty analysis: A comparative survey. *Risk Analysis*, 1(4), 251-258.
- Devore, J. (1982). *Probability and statistics for engineering and the sciences*. Brooks/Cole Publishing Company.
- Ditlevsen, O.D. (1981). *Uncertainty modelling*. McGraw Hill Book Co., New York.

- El-Kadi, A.I., (1987). Variability of infiltration under uncertainty in unsaturated zone parameters. *Journal of Contaminant Hydrology*, 90(1), 61-80.
- Freeze, R.A. (1975). A stochastic-conceptual analysis of one-dimensional groundwater flow in non-uniform homogenous media. *Water Resources Research*, 11(5), 725-741.
- Freeze, R.A. (1979). A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media. *Water Resources Research*, 11(5), 725-741.
- Freeze, R.A. and Cherry, J.A. (1979). *Groundwater*. Prentice-Hall, Inc., New Jersey.
- Gelhar, L.W., (1986). Stochastic subsurface hydrology from theory to applications. *Water Resources Research*, 22(9), 135S-145S.
- Gelhar, L.W., Bakr, A.A., and MacMillan, J.R. (1977). Comments on: A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogenous media by R. Allen Freeze. *Water Resources Research*, 13(2), 477-479.
- Green, R.E., and Corey, J.C. (1971). Calculation of hydraulic conductivity: A further evaluation of some predictive methods. *Soil Science Society of America Proc.*, 35, 3-8.
- Green, W.H. and Ampt, G.A. (1911). Studies in soil physics. I. The flow of air and water through soils. *J. Agric. Science*, 4, 1-24.
- Hahn, G.J. and Shapiro, S.S. (1967). *Statistical models in engineering*. Wiley, New York.
- Harr, M.E. (1962). *Groundwater and seepage*. McGraw-Hill Book Co., New York.
- Harr, M.E. (1977). *Mechanics of particular media*. McGraw-Hill Book Co., New York.
- Harr, M.E. (1987). *Reliability based design in civil engineering*. McGraw-Hill Book Co., New York.
- Harr, M.E. (1988). Information: An engineering material. *Proc. Conference on Reliability-Based Design in Civil Engineering*, Lausanne, Switzerland.
- Harr, M.E. (1989). Probabilistic estimates for multivariate analyses. *Applied Mathematical Modelling*, 13, 313-318.

- Haverkamp, R., Vaculin, M., Touma, J., Wierenga, P., and Vachaud, G. (1977). A comparison of numerical solution models for one-dimensional infiltration. *Soil Science Society of America Journal*, 41(2), 285-294.
- Hillel, D. (1982). *Introduction to soil physics*. Academic Press, New York, N.Y.
- Hindmarsh, A.C. (1981). *ODE solvers for use with the method of lines* (UCRL-85293 (Rev. 1)), Lawrence Livermore National Laboratories, Livermore, CA.
- Holtz, R.D., and Krizek, R.J. (1971). Statistical evaluation of soil test data. *Proc. First International Conference on Applications of Statistics and probability in Soil and Structural Engineering*, Hong-Kong, 1, 230-253.
- Hopkins, P., and Eaton, R. (in review). *LLUVIA - A one-dimensional, steady-state program for flow through partially saturated porous media* (SAND88-0558). Sandia National Laboratories, Albuquerque, NM.
- Iman, R.L., and Conover, W.J. (1980). Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. *Communications in Statistics*, A9(17), 1749-1842.
- Iman, R.L., and Conover, W.J. (1982). A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics*, B11(3), 311-334.
- Iman, R.L., and Davenport, J.M. (1982). Rank correlation plots for use with correlated input variables. *Communications in Statistics*, 11, 335-360.
- Iman, R.L., and Helton, J.C. (1988). An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis*, 8, 71-90.
- Kaplan, P.G. and Yarrington, L. (in review). *Modeling the Uncertainties in the Parameter Values of a layered, Variably Saturated Column of Volcanic Tuff*, (SAND88-2247C). Sandia National Laboratories, Albuquerque, NM.
- Klavetter, E.M., and Peters, R.R. (1986). *Estimation of hydraulic properties of an unsaturated fractured rock mass* (SAND84-2642). Sandia National Laboratories, Albuquerque, NM.
- Klavetter, E.M., and Peters, R.R. (1988). A continuum model for water movement in an unsaturated fractured rock mass. *Water Resources research*, 24(3), 416-430.
- Lind, N.C. (1983). Modelling of uncertainty in discrete dynamical systems. *Applied Mathematical Modelling*, 7, 146-152.

- Lumb, P. (1966). The variability of natural soils. *Canadian Geotechnical Journal*, 3, 74-79.
- Lumb, P. (1974). Application of statistics in soil mechanics. Lee (Ed.), *Soil mechanics - New Horizons*, Butterworth, London.
- McGuffey, V., A-Grivas, D., Iori, J., and Kyfor, Z. (1982). Conventional and probabilistic embankment design. *Journal of Geotechnical Engineering*, ASCE, 108, 1246-1254.
- McKay, M.D., Conover, W.J., and Beckman, R.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21, 239-245.
- Mood, A. M., and Grayball, F. A., (1963). *Introduction to the theory of statistics*. McGraw-Hill Book Co., Inc., New York, N.Y.
- Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, 12(3), 513-522.
- Musket, M. (1937). *The flow of homogenous fluids through porous media*. McGraw Hill Pub. Co., New York, NY.
- Neuman, S.P. (1982). Statistical characterization of aquifer heterogeneities: An overview (Recent trends in hydrology). *Geological Society of America, Special Paper*, 189, 81-102.
- Nguyen, V.U., and Chowdhury, R.N. (1985). Simulation for risk analysis with correlated variables. *Geotechnique*, 35(1), 47-58.
- Nielsen, D.R., van Genuchten, M.Th., and Biggar, J.W., (1986). Water flow and solute transport processes in the unsaturated zone. *Water Resources Research*, 22(9), 89S-108S.
- Ordaz, M. (1987). On the use of probability concentrations. *Structural Safety*, 5, 317-318.
- Panián, T.F. (1987). *Unsaturated flow properties data catalog, Volume II* (Publication #45601). Water Resources Center, Desert research Institute, University of Nevada System.
- Persaud, N., Giraldez, J.V., and Chang, A.C. (1985). Monte-Carlo simulation of non interacting solute transport in a spatially heterogenous soil. *Soil Science Society of America Journal*, 49, 562-568.

- Peters, R.R., Klavetter, E.A., Hall, I.J., Blair, S.C., Heller, P.R., and Gee, G.W., (1984). *Fracture and matrix hydrologic characteristics of tuffaceous materials from Yucca Mountain, Nye County, Nevada* (SAND84-1471). Sandia National Laboratories, Albuquerque, NM.
- Peters, R.R., Gauthier, J.H., and Dudley, A.L. (1986). *The effect of percolation rate on water-travel time in deep, partially saturated zones* (SAND85-0854). Sandia National Laboratories, Albuquerque, NM.
- Rampello, S., Calabresi, G., and Gallo, C. (1987). Geotechnical characteristics of two large plains. *Proc. 9th Southeast Asian Geotechnical Conference*, Bangkok, 5, 73-84.
- Rawls, W.J. and Brakensiek, D.L. (1985). Prediction of soil water properties for hydrologic modelling. *Proc. Symposium on Watershed Management*, ASCE, New York, 293-299.
- Rawls, W.J., Brakensiek, D.L., and Saxton, K.E. (1982). Estimation of soil water properties. *Trans. ASAE*, 25(5), 1316-1320.
- Rosenblueth, E. (1975). Point estimates for probability moments. *Proc. National Academy of Science, USA*, 72(10), 3812-3814.
- Rosenblueth, E. (1981). Two-point estimates in probabilities. *Applied Mathematical Modelling*, 5, 329-335.
- Ross, B. (1984). A conceptual model of deep unsaturated zones with negligible recharge. *Water Resources Research*, 20(11), 1627-1629.
- Shampine, L.F., and Watts, H.A. (1980). *DEPAC-Design of a user oriented package of ODE solvers* (SAND79-2374). Sandia National Laboratories, Albuquerque, NM.
- Smith, L., and Freeze, R. (1979). Stochastic analysis of steady state groundwater flow in a bounded domain 1. one-dimension simulations. *Water Resources Research*, 15(3), 521-528.
- Smith, R.E., and Hebbert, R.H.B. (1979). A Monte Carlo analysis of the hydrologic effects of spatial variability of infiltration. *Water Resources Research*, 15(2), 419-429.
- Vachaud, G., and Thony, J.L. (1971). Hysteresis during infiltration and redistribution in a soil column at different initial water contents. *Water Resources Research*, 7, 111-127.

- Van Genuchten, R. (1978). Calculating the unsaturated hydraulic conductivity with a new closed form analytical model. *Water Resources Bulletin*, Princeton University Press, Princeton, NJ.
- Van Genuchten, M.Th. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Proc.*, 44(5), 892-898.
- Van Genuchten, M.Th. (1982). A comparison of numerical solutions of the one-dimensional unsaturated-saturated flow and mass transport equations. *Water Resources Research*, 5, 47-55.
- Vanmarcke, E.H. (1977). Probabilistic modeling of soil profile. *Journal of Geotechnical Engineering*, ASCE, 103(11), 1227-1246.
- Vanmarcke, E.H. (1977). Reliability of earth slopes. *Journal of Geotechnical Engineering*, ASCE, 103(11), 1247-1264.
- Whitman, R.V. (1984). Evaluating calculated risk in geotechnical engineering. *Journal of Geotechnical Engineering*, ASCE, 110(2), 145-189.
- Wolff, T.F., and Harr, M.E. (1987). slope design for earth dams. *Int. Conference on Applications of Statistics and probability in Soil and Structural Engineering*, Vancouver, 2, 725-732.
- Wu, T.H., Vyas, S.K., and Chang, N.Y. (1973). Probabilistic analysis of seepage. *Journal of Soil Mechanics and Foundations Division*, 99(4), 323-340.
- Yeh, T.C.J., Gelhar, L.W., and Gutjahr, A.L. (1985). Stochastic analysis of unsaturated flow in heterogeneous soils 1. Statistically isotropic media. *Water Resources Research*, 21(4), 447-456.

APPENDICES

Appendix A: Details of the Methodology

A.1 Lind's Point-representation Algorithm

The vectors z_i , $i=1,2, \dots,N$, of Lind's point-representation coordinates (as given in equation (2.12)) are determined as follows:

$$z_{ii} = \left\{ N\sigma_{ii}^2 - \sum_{k=1}^{i-1} z_{ki}^2 \right\}^{1/2} \quad i=1,2, \dots, N$$

$$z_{ij} = \frac{\left\{ N\sigma_{ij} - \sum_{k=1}^{i-1} z_{ki} z_{kj} \right\}}{z_{ii}} \quad \begin{array}{l} i=1,2, \dots, N-1 \\ \text{and } j=i+1, i+2, \dots N \end{array} \quad (\text{A.1})$$

where z_{ii} and z_{ij} are the point distribution coordinates of the vector z_i ; σ_{ii} and σ_{ij} are respectively the variance, covariance terms of the random variables; and N is the number of random variables.

A.2 The principal component analysis involved in Harr's methodology

The principal component analysis involved in transforming the standardized random variables ds_i (as given in equation (2.13)) into new random variables u_i in the eigenspace of the correlation matrix is as follows:

$$u_i = \alpha_{1i} ds_1 + \alpha_{2i} ds_2 + \dots + \alpha_{ji} ds_j + \dots + \alpha_{Ni} ds_N \quad (\text{A.2})$$

where $i=1,2,\dots,N$, $j=1,2,\dots,N$, and α_{ji} is the eigenvector component of the correlation matrix. The eigenvalues - eigenvectors of the correlation matrix is given by:

eigenvalue

$$[\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_N]$$

eigenvector

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1N} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2N} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \alpha_{N1} & \alpha_{N2} & \dots & \alpha_{NN} \end{bmatrix}$$

Appendix B: Models I and II Point Distribution Approximation

B.1 Model I Discrete Approximation for a Function of Three Random Variables

The equations of the discrete approximation to a joint probability density function $f(x_1, x_2, x_3)$ with three random variables are given below:

$$p_{+++} + p_{+..} + p_{-..} + p_{...} + p_{+..} + p_{-..} + p_{...} + p_{...} = 1 \quad (\text{B.1})$$

$$(x_{1-} - \bar{x}_1)[(p_{---} + p_{--+})(x_{2-} - \bar{x}_2) + (p_{--+} + p_{+..})(x_{2+} - \bar{x}_2)] \\ + (x_{1+} - \bar{x}_1)[(p_{+..} + p_{+..})(x_{2-} - \bar{x}_2) + (p_{+..} + p_{+..})(x_{2+} - \bar{x}_2)] = \rho_{12} \sigma_{x1} \sigma_{x2} \quad (\text{B.2})$$

$$(x_{1-} - \bar{x}_1)[(p_{---} + p_{--+})(x_{3-} - \bar{x}_3) + (p_{--+} + p_{+..})(x_{3+} - \bar{x}_3)] \\ + (x_{1+} - \bar{x}_1)[(p_{+..} + p_{+..})(x_{3-} - \bar{x}_3) + (p_{+..} + p_{+..})(x_{3+} - \bar{x}_3)] = \rho_{13} \sigma_{x1} \sigma_{x3} \quad (\text{B.3})$$

$$(x_{2-} - \bar{x}_2)[(p_{---} + p_{+..})(x_{3-} - \bar{x}_3) + (p_{--+} + p_{+..})(x_{3+} - \bar{x}_3)] + \\ + (x_{2+} - \bar{x}_2)[(p_{+..} + p_{+..})(x_{3-} - \bar{x}_3) + (p_{+..} + p_{+..})(x_{3+} - \bar{x}_3)] = \rho_{23} \sigma_{x2} \sigma_{x3} \quad (\text{B.4})$$

$$p_{i+} + p_{i-} = 1 \quad (\text{B.5})$$

$$p_{i+}(x_{i+} - \bar{x}_i) + p_{i-}(x_{i-} - \bar{x}_i) = 0 \quad (\text{B.6})$$

$$p_{i+}(x_{i+} - \bar{x}_i)^2 + p_{i-}(x_{i-} - \bar{x}_i)^2 = \sigma_{xi}^2 \quad (\text{B.7})$$

$$p_{i+}(x_{i+} - \bar{x}_i)^3 + p_{i-}(x_{i-} - \bar{x}_i)^3 = \beta_{xi}(1) \sigma_{xi}^3 \quad (\text{B.8})$$

where x_i , $i = 1, 2, 3$, is a random variable with mean \bar{x}_i , standard deviation σ_{xi} , and skewness coefficient $\beta_{xi}(1)$; and ρ_{ij} , $i=1, 2$, $j=2, 3$, is the correlation coefficient between x_i and x_j .

In defining the equations (B.5-B.8), p_{i-} and p_{i+} ($=1, 2, 3$) represent the sum of the weighting functions associated with the point estimate locations x_{i-} and x_{i+} ($i=1, 2, 3$), respectively. The corresponding relationships are given in equations (B.9-B.11).

For the point estimate locations x_i , ($i=1,2,3$) :

$$p_{1-} = p_{---} + p_{-+-} + p_{-+-} + p_{-+-} \quad (\text{B.9})$$

$$p_{2-} = p_{---} + p_{+--} + p_{+--} + p_{+--} \quad (\text{B.10})$$

$$p_{3-} = p_{---} + p_{-+-} + p_{-+-} + p_{-+-} \quad (\text{B.11})$$

where p_{1-} , p_{2-} , and p_{3-} are associated with the locations x_1 , x_2 , and x_3 , respectively. The similar expressions for p_{i+} ($i=1,2,3$) can be obtained from the relationships (B.9 - B.11) and (B.5).

B.2 Model I Discrete Approximation for a Function of Four Random Variables

The magnitudes of all the weighting functions are given in a matrix form as follows:

$$[P] = [C] [D] \quad (\text{B.12})$$

where the matrices $[P]$, $[D]$, and $[C]$ are given below:

$$[P]^T = \begin{bmatrix} p_{++++} & p_{+---} & p_{-+++} & p_{----} & p_{+---} & p_{-+++} & p_{-+++} & p_{-+++} \\ p_{+---} & p_{+---} \end{bmatrix}_{1 \times 16}$$

$$[D]^T = \begin{bmatrix} 1.0 & p_{12} & p_{13} & p_{14} & p_{23} & p_{24} & p_{34} \end{bmatrix}_{1 \times 7}$$

$$[C] = \begin{bmatrix} p_{1+}p_{2+}p_{3+}p_{4+} & a_{12} & a_{13} & a_{14} & a_{23} & a_{24} & a_{34} \\ p_{1-}p_{2+}p_{3+}p_{4+} & -a_{12} & -a_{13} & -a_{14} & a_{23} & a_{24} & a_{34} \\ p_{1+}p_{2-}p_{3+}p_{4+} & -a_{12} & a_{13} & a_{14} & -a_{23} & -a_{24} & a_{34} \\ p_{1-}p_{2-}p_{3+}p_{4+} & a_{12} & -a_{13} & -a_{14} & -a_{23} & -a_{24} & a_{34} \\ p_{1+}p_{2+}p_{3-}p_{4+} & a_{12} & -a_{13} & a_{14} & -a_{23} & a_{24} & -a_{34} \\ p_{1-}p_{2+}p_{3-}p_{4+} & -a_{12} & a_{13} & -a_{14} & -a_{23} & a_{24} & -a_{34} \\ p_{1+}p_{2-}p_{3-}p_{4+} & -a_{12} & -a_{13} & a_{14} & a_{23} & -a_{24} & -a_{34} \\ p_{1-}p_{2-}p_{3-}p_{4+} & a_{12} & a_{13} & -a_{14} & a_{23} & -a_{24} & -a_{34} \\ p_{1+}p_{2+}p_{3+}p_{4-} & a_{12} & a_{13} & -a_{14} & a_{23} & -a_{24} & -a_{34} \\ p_{1-}p_{2+}p_{3+}p_{4-} & -a_{12} & -a_{13} & a_{14} & a_{23} & -a_{24} & -a_{34} \\ p_{1+}p_{2-}p_{3+}p_{4-} & -a_{12} & a_{13} & -a_{14} & -a_{23} & a_{24} & -a_{34} \\ p_{1-}p_{2-}p_{3+}p_{4-} & a_{12} & -a_{13} & a_{14} & -a_{23} & a_{24} & -a_{34} \\ p_{1+}p_{2+}p_{3-}p_{4-} & a_{12} & -a_{13} & -a_{14} & -a_{23} & -a_{24} & a_{34} \\ p_{1-}p_{2+}p_{3-}p_{4-} & -a_{12} & a_{13} & a_{14} & -a_{23} & -a_{24} & a_{34} \\ p_{1+}p_{2-}p_{3-}p_{4-} & -a_{12} & -a_{13} & -a_{14} & a_{23} & a_{24} & a_{34} \\ p_{1-}p_{2-}p_{3-}p_{4-} & a_{12} & a_{13} & a_{14} & a_{23} & a_{24} & a_{34} \end{bmatrix} \quad 16 \times 7$$

where p_{lm} and a_{lm} ($l=1, 2, 3$ & $m=1+1, 1+2, \dots, 4$) are defined in equation (3.41).

B.3 Model II Discrete Approximation for a Function of Three Random Variables.

For the variable x_1 , the expressions obtained by specifying the mean \bar{x}_1 , standard deviation σ_{x1} , and skewness coefficient $\beta_{x1}(1)$, respectively, are as follows:

$$E[(x_1 - \bar{x}_1)] = p_{1+}z_{11+} - p_{1-}z_{11-} = 0.0 \quad (\text{B.13})$$

$$E[(x_1 - \bar{x}_1)^2] = p_{1+}z_{11+}^2 + p_{1-}z_{11-}^2 = \sigma_{x1}^2 \quad (\text{B.14})$$

$$E[(x_1 - \bar{x}_1)^3] = p_{1+}z_{11+}^3 - p_{1-}z_{11-}^3 = \beta_{x1}(1)\sigma_{x1}^3 \quad (\text{B.15})$$

Similarly, the expressions are obtained for the variables x_2 and x_3 as follows:

For the variable x_2

$$E[(x_2 - \bar{x}_2)] = (p_{1+} - p_{1-})z_{12} + p_{2+}z_{22+} - p_{2-}z_{22-} = 0.0 \quad (\text{B.16})$$

$$E[(x_2 - \bar{x}_2)^2] = (p_{1+} + p_{1-})z_{12}^2 + p_{2+}z_{22+}^2 + p_{2-}z_{22-}^2 = \sigma_{x_2}^2 \quad (\text{B.17})$$

$$E[(x_2 - \bar{x}_2)^3] = (p_{1+} - p_{1-})z_{12}^3 + p_{2+}z_{22+}^3 - p_{2-}z_{22-}^3 = \beta_{x_2}(1)\sigma_{x_2}^3 \quad (\text{B.18})$$

For the variable x_3

$$E[(x_3 - \bar{x}_3)] = (p_{1+} - p_{1-})z_{13} + (p_{2+} - p_{2-})z_{23} + p_{3+}z_{33+} - p_{3-}z_{33-} = 0.0 \quad (\text{B.19})$$

$$E[(x_3 - \bar{x}_3)^2] = (p_{1+} + p_{1-})z_{13}^2 + (p_{2+} + p_{2-})z_{23}^2 + p_{3+}z_{33+}^2 + p_{3-}z_{33-}^2 = \sigma_{x_3}^2 \quad (\text{B.20})$$

$$E[(x_3 - \bar{x}_3)^3] = (p_{1+} - p_{1-})z_{13}^3 + (p_{2+} - p_{2-})z_{23}^3 + p_{3+}z_{33+}^3 - p_{3-}z_{33-}^3 = \beta_{x_3}(1)\sigma_{x_3}^3 \quad (\text{B.21})$$

The covariance terms of the variables are given below:

For the variables x_1 and x_2

$$E[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)] = (p_{1+}z_{11+} + p_{1-}z_{11-})z_{12} = \rho_{12}\sigma_{x_1}\sigma_{x_2} \quad (\text{B.22})$$

For the variables x_1 and x_3

$$E[(x_1 - \bar{x}_1)(x_3 - \bar{x}_3)] = (p_{1+}z_{11+} + p_{1-}z_{11-})z_{13} = \rho_{13}\sigma_{x_1}\sigma_{x_3} \quad (\text{B.23})$$

For the variables x_2 and x_3

$$E[(x_2 - \bar{x}_2)(x_3 - \bar{x}_3)] = (p_{1+} + p_{1-})z_{12}z_{13} + (p_{2+}z_{22+} + p_{2-}z_{22-})z_{23} = \rho_{23}\sigma_{x_2}\sigma_{x_3} \quad (\text{B.24})$$

In the equations (B.13 - B.24), x_i ($i=1,2,3$) is a random variable with mean \bar{x}_i , standard deviation σ_{x_i} , and skewness coefficient $\beta_{x_i}(1)$. The correlation coefficient between the random variables x_i and x_j ($i=1,2$ & $j=2,3$) is ρ_{ij} , and according to the assumption of equally distributed weighting functions as mentioned in section 3.2.2, $p_{i+} + p_{i-} = 1/3$ ($i=1,2,3$).

The solution of the equations (B.13 - B.15) give the unknowns p_{1+} , p_{1-} , z_{11+} , and z_{11-} as follows:

$$p_{1+} = \frac{1}{6} \left\{ 1 \pm \sqrt{\frac{\beta_{x1}(1)^2}{12 + \beta_{x1}(1)^2}} \right\} \quad (\text{B.25})$$

$$p_{1-} = \frac{1}{3} - p_{1+} \quad (\text{B.26})$$

$$z_{11+} = \sqrt{3} \sqrt{\frac{p_{1-}}{p_{1+}}} \sigma_{x1} \quad (\text{B.27})$$

$$z_{11-} = \sqrt{3} \sqrt{\frac{p_{1+}}{p_{1-}}} \sigma_{x1} \quad (\text{B.28})$$

where the sign preceding the radical in equation (B.25) is that of $-\beta_{x1}(1)$.

According to the equations (B.22 and B.23), the unknowns z_{12} and z_{13} are given by:

$$z_{1j} = \frac{\rho_{1j} \sigma_{x1} \sigma_{xj}}{p_{1+} z_{11+} + p_{1-} z_{11-}} \quad (\text{B.29})$$

where $j = 2, 3$; and the unknowns p_{1+} , p_{1-} , z_{11+} , and z_{11-} are defined in the equations (B.25 - B.28).

Knowing the solution to p_{1+} , p_{1-} , z_{11+} , z_{11-} and z_{12} , equations (B.16 - B.18) are rearranged as follows:

$$p_{2+} z_{22+} - p_{2-} z_{22-} = -(p_{1+} - p_{1-}) z_{12} = A_2 \quad (\text{B.30})$$

$$p_{2+} z_{22+}^2 + p_{2-} z_{22-}^2 = \sigma_{x2}^2 - (p_{1+} + p_{1-}) z_{12}^2 = B_2 \quad (\text{B.31})$$

$$p_{2+} z_{22+}^3 - p_{2-} z_{22-}^3 = \beta_{x2}(1) \sigma_{x2}^3 - (p_{1+} - p_{1-}) z_{12}^3 = C_2 \quad (\text{B.32})$$

where $p_{2+} + p_{2-} = 1/3$.

From the equations (B.30-B.32), the quadratic equation which solves for z_{22+} and z_{22-} is obtained as given below:

$$\left(\frac{B_2}{3} - A_2^2\right)z_{22}^2 - \left(\frac{C_2}{3} - A_2 B_2\right)z_{22} + (A_2 C_2 - B_2^2) = 0.0 \quad (\text{B.33})$$

where the solution z_{22} is defined as $z_{22} = z_{22+}$ and $z_{22} = (-z_{22-})$.

The weighting functions p_{2-} and p_{2+} are determined from the expressions given below:

$$p_{2-} = \frac{\frac{1}{3}z_{22+} - A_2}{z_{22+} + z_{22-}} \quad (\text{B.34})$$

$$p_{2+} = \frac{1}{3} - p_{2-} \quad (\text{B.35})$$

From the equation (B.24), the unknown z_{23} is given by:

$$z_{23} = \frac{\rho_{23} \sigma_{x2} \sigma_{x3} - (p_{1+} + p_{1-})z_{12}z_{13}}{p_{2+}z_{22+} + p_{2-}z_{22-}} \quad (\text{B.36})$$

By rearranging the equations (B.19 - B.21), the quadratic equation which solves for z_{33+} and z_{33-} is obtained as given below:

$$\left(\frac{B_3}{3} - A_3^2\right)z_{33}^2 - \left(\frac{C_3}{3} - A_3 B_3\right)z_{33} + (A_3 C_3 - B_3^2) = 0.0 \quad (\text{B.37})$$

where A_3 , B_3 , and C_3 are defined as follows:

$$A_3 = 0.0 - (p_{1+} - p_{1-})z_{13} - (p_{2+} - p_{2-})z_{23}$$

$$B_3 = \sigma_{x3}^2 - \frac{z_{13}^2}{3} - \frac{z_{23}^2}{3}$$

$$C_3 = \beta_{x3}(1)\sigma_{x3}^3 - (p_{1+} - p_{1-})z_{13}^3 - (p_{2+} - p_{2-})z_{23}^3$$

From the equations (B.19 - B.20), the weighting functions p_{3-} and p_{3+} are given by:

$$p_{3-} = \frac{\frac{1}{3}z_{33+} - A_3}{z_{33+} + z_{33-}} \quad (\text{B.38})$$

$$p_{3+} = \frac{1}{3} - p_{3-} \quad (\text{B.39})$$

Appendix C: Probabilistic Calculations

C.1 Point Estimates Using Model I - Example I

Table C.1 Sets of Values using Model I - Example I.

K_s (cm/hour)	S_r	α (1/m)	N	weighting function p
0.0059	0.0465	0.2728	1.0527	0.7101
0.6014	0.0465	0.2728	1.0527	-0.0724
0.0059	0.0935	0.2728	1.0527	0.0929
0.6014	0.0935	0.2728	1.0527	-0.0006
0.0059	0.0465	1.9200	1.0527	-0.1157
0.6014	0.0465	1.9200	1.0527	0.0035
0.0059	0.0935	1.9200	1.0527	0.0419
0.6014	0.0935	1.9200	1.0527	0.0213
0.0059	0.0465	0.2728	1.1696	-0.0392
0.6014	0.0465	0.2728	1.1696	-0.0012
0.0059	0.0935	0.2728	1.1696	0.1290
0.6014	0.0935	0.2728	1.1696	0.0436
0.0059	0.0465	1.9200	1.1696	0.0051
0.6014	0.0465	1.9200	1.1696	0.0098
0.0059	0.0935	1.9200	1.1696	0.1521
0.6014	0.0935	1.9200	1.1696	0.0199

Table C.1 lists the sixteen different values obtained using model I of the new point estimate methodology for the single layer problem (example I - containing four random variables K_s , S_r , α , and N). As an aid in determining the various values of the random variables and the corresponding weighting functions, the reader is referred to sections 3.1.3 and 4.2.1.1. Table C.2 lists the ground water travel time computed using these sets of values for influx rates of 0.1 mm/year and 1.0 mm/year.

Table C.2 Ground Water Travel Times Computed using Sets of Values given in Table C.1.

weighting function p	ground water travel time T (years)	
	flux rate = 0.1 mm/year	flux rate = 1.0 mm/year
0.7101	33121.63	3364.94
-0.0724	32995.16	3300.78
0.0929	31488.99	3199.08
-0.0006	31368.75	3138.08
-0.1157	32226.11	3349.76
0.0035	30765.44	3115.30
0.0419	30637.61	3184.64
0.0213	29248.94	2961.74
-0.0392	30564.54	3099.64
-0.0012	30506.58	3051.18
0.1290	29057.95	2946.85
0.0436	29002.85	2900.78
0.0051	25971.49	2892.70
0.0098	24228.10	2451.98
0.1521	24691.30	2750.11
0.0199	23033.85	2331.12

The expected value and the standard deviation of the travel time was computed using equation 4.6 and as explained in section 4.2.1.1.

Considering the travel times for the flux rate = 0.1 mm/year, from equation (4.6) $E[T] = \sum p T = 30,682.62$ years and $E[T^2] = \sum p T^2 = 9.53 \cdot 10^8$. This gives, standard deviation $s_T = (E[T^2] - (E[T])^2)^{0.5} = 3357.47$ years and coefficient of variation $V = (3357.47/30,682.62) \cdot 100 = 10.94 \%$. Similarly, for the flux rate = 1.0 mm/year $E[T] = 3150.31$ years and coefficient of variation $V = 8.88 \%$.

C.2 Point Estimates Using Model II - Example I

Table C.3 lists the eight sets of values obtained using model II. The procedure followed in determining these values was explained in sections 3.2.3 and 4.2.1.2.

Table C.3 Sets of Values using Model II - Example I.

K_s (cm/hour)	S_r	α (1/m)	N	weighting function p
0.6403	0.1528	2.5543	1.2738	0.0197
-0.0330	-0.0128	-1.5543	0.9063	0.2303
0.02	0.0827	4.3203	1.4490	0.3798
0.02	-0.0271	-3.3203	0.7310	-0.1298
0.02	0.07	5.5813	1.9853	-0.1940
0.02	0.07	-0.6897	0.1947	0.4440
0.02	0.07	0.5	1.5058	0.4398
0.02	0.07	0.5	-0.1966	-0.1898

Table C.4 lists the possible four sets of values among the sets of values given in Table C.3 for the LUVIA run.

Table C.4 Possible Sets of Values using Model II - Example I.

K_s (cm/hour)	S_r	α (1/m)	N	weighting function	redistributed weighting function
0.6403	0.1528	2.5543	1.2738	0.0197	0.0305
0.02	0.0827	4.3203	1.4490	0.3798	0.5886
0.02	0.07	5.5813	1.9853	-0.1940	-0.3007
0.02	0.07	0.5	1.5058	0.4398	0.6816

The redistributed weighting functions given in column six of Table C.4 are obtained by dividing each weighting functions given in column five by the sum of the weighting functions of column five. Following this, the sum of the weighting functions of column five is given by 0.6453, therefore, the first row of redistributed weighting function is given by $0.0197/0.6453 = 0.0305$. Continuing this procedure, the redistributed weighting functions given in column six are obtained.

Table C.5 gives the ground water travel time computed using LLUVIA for the sets of values given in Table C.4

Table C.5 Groundwater Travel Time Computed using Sets of Values given in Table C.4.

weighting function p	ground water travel time T (years)	
	flux rate = 0.1 mm/year	flux rate = 1.0 mm/year
0.0305	16508.88	1694.98
0.5886	12918.43	1647.70
-0.3007	6256.55	939.96
0.6816	21190.42	2159.19

Table C.6 Sets of Values using Rosenblueth (1975) Methodology - Example I.

K_r (cm/hour)	S_r	α (1/m)	N	weighting function	redistributed weighting function
0.1107	0.0465	1.0680	1.0355	-0.0589	-0.1193
0.1107	0.0935	1.0680	1.0355	0.0810	0.1641
0.1107	0.0465	1.0680	1.1445	0.0665	0.1348
0.1107	0.07	1.0680	1.1445	0.4049	0.8204

Table C.7 Sets of Values using Lind (1983) Methodology - Example I.

K_s (cm/hour)	S_r	α (1/m)	N	weighting function	redistributed weighting function
0.2013	0.1146	1.6065	1.1890	0.125	0.1429
0.02	0.0848	0.6429	1.0666	0.125	0.1429
0.02	0.0552	0.3570	1.1134	0.125	0.1429
0.02	0.07	0.7140	1.1083	0.125	0.1429
0.02	0.07	0.2860	1.0717	0.125	0.1429
0.02	0.07	0.5	1.1247	0.125	0.1429
0.02	0.07	0.5	1.0553	0.125	0.1429

Table C.8 Sets of Values using Harr (1989) Methodology - Example I.

K_s (cm/hour)	S_r	α (1/m)	N	weighting function	redistributed weighting function
0.1129	0.0933	1.0814	1.1422	0.4677	0.8930
0.0142	0.0429	0.3204	1.1772	0.0269	0.0514
0.0258	0.0971	0.6796	1.0027	0.0269	0.0514
0.0646	0.0403	1.2530	1.0556	0.0022	0.0042

C.2 Global Correlation Matrix used for the Two Layer Problem - Example II

	<i>silt</i>				<i>siltyclay</i>			
	K_s	θ_r	α	N	K_s	θ_r	α	N
	1.0	-0.204	0.984	0.466	0.0	0.0	0.0	0.0
	-0.204	1.0	-0.20	-0.61	0.0	0.0	0.0	0.0
	0.984	-0.20	1.0	0.551	0.0	0.0	0.0	0.0
$\rho =$	0.466	-0.61	0.551	1.0	0.0	0.0	0.0	0.0
	0.0	0.0	0.0	0.0	1.0	0.949	0.974	0.908
	0.0	0.0	0.0	0.0	0.949	1.0	0.964	0.794
	0.0	0.0	0.0	0.0	0.974	0.964	1.0	0.889
	0.0	0.0	0.0	0.0	0.908	0.794	0.889	1.0

8x8

Appendix D: Computer Program

D.1 Point Estimates Program for LLUVIA

```

c      *****
c      This program calculates the point distribution
c      approximation of Model I, Model II,
c      Rosenblueth (1975), Lind (1983), and Harr (1989)
c      methodologies for LLUVIA run
c
c      *****
c
c      corr - global correlation matrix
c      xm - mean
c      xs - standard deviation
c      cs - skewness
c      P - weighting function
c      npg - number of variables
c
c      common /cor/corr(50,50)
c      common /inp/xm(40),xs(40),cs(40)
c      common /lay/ES(100),ESL(100),ESF(20)
c      common /seq/IS(30)
c      integer rowmax, colmax, row, col
c
c      input file : 'input'
c
c      open(unit=8,file='input',status='unknown')
c
c      output files: rose-Rosenblueth (1975),
c      lin-Lind(1983), alt-Harr(1989)
c      skro-Model I, slin - Model II
c
c      open(unit=10,file='rose',status='unknown')
c      open(unit=11,file='lin',status='unknown')
c      open(unit=12,file='alt',status='unknown')
c      open(unit=13,file='slin',status='unknown')
c      open(unit=14,file='skro',status='unknown')
c
c      SELECTING THE INPUT MODE FOR THE GLOBAL CORRELATION
c      MATRIX FOR THE LLUVIA RUN
c
c      - input subroutine called - 'dacor'
c      each sub layer correlation matrix file - 'entry'
c      profile correlation matrix file - 'global'
c
c      open(unit=15,file='global',status='unknown')
c      open(unit=16,file='entry',status='unknown')

```

```

c      print *, '.....select the input mode for.....'
      print *, '.....the correlation matrix.....'
      print *, '          "0" = SCREEN'
      print *, '          "1" = FILE(INPUT)'
      print *, '.....'
      read *,isc
      print *, ' '
      if (isc.eq.0) then
        print *, 'The sequence to be followed is from the bottom'
        print *, '          layer to the top layer'
c      print *, ' and'
c      print *, ' the matrix order is ks, theta res, par1, and par2'
        print *, ' '
        call dacor(npg)
      else
        print *, '  Enter the size of the Correlation matrix :'
        read *,npg
        rowmax = npg
        colmax = npg
c
c      Reading the Correlation matrix :'
c
      do 10 row=1, rowmax
        read (15,*)(corr(row, col), col=1, colmax)
10     continue
        end if
c
c      SETTING THE INPUT MODE FOR THE PARAMETERS FOR THE
C      CALCULATION OF POINT ESTIMATES
c
      print *, '.....Select the INPUT mode.....'
      print *, '.....for the input parameters.....'
      print *, ' '
      print *, '          "0" = SCREEN'
      print *, '          "1" = FILE(input)'
      print *, ' '
      read *,inp
      if (inp.eq.0) then
        print *, 'Enter the mean, standard deviation, & skewness'
        print *, 'The sequence is from the bottom layer to the top'
        print *, 'and enter matrix, and fracture properties for '
        print *, 'each layer'
        print *, 'Enter xm, xs, cs'
        do 30 i=1,npg
          print *,i
          read *,xm(i),xs(i),cs(i)
          write(8,*)xm(i),xs(i),cs(i)
30     continue
        else
          do 35 i=1,npg
            read(8,*)xm(i),xs(i),cs(i)

```

```

35     continue
      end if
c
      print *, ' '
      print *, '      If the columns has to be rearranged'
      print *, '      for the final output'
      print *, '      Enter ===== > 1'
      read *, ire
      if(ire.eq.1)then
      print *, 'Enter the correct sequence of the columns:'
      print *, ' '
      read *,(IS(iq), iq=1,npg)
      print *, ' The sequence is'
22     do 22 iq=1,npg
      print *,iq,'col',IS(iq)
      else
23     do 23 ip =1,npg
      IS(ip)=ip
      end if
c
c     subroutines of the methods: rosen-Rosenblueth(1975),
c     skros -Model I, lind-Lind(1983), sklind-Model II,
c     altpem-Harr(1989).
c
      call rosen(npg)
call skros(npg)
      call lind(npg)
      call sklind(npg)
      call altpem(npg)
      endfile (unit=10)
      endfile (unit=11)
      endfile (unit=12)
      endfile (unit=13)
      endfile (unit=14)
      close (unit=15)
      close (unit=16)
      print *, ' If the lluvia.inp file to be created'
      print *, '      Enter ===== > 1'
      print *, ' Else Enter ===== > 0'
      read *, isw
c     print *, ' This program can only deal with all layers'
c     print *, '      matrix or matrix and fractured'
      if(isw.eq.1)then
      rewind (unit=10)
      rewind (unit=11)
      rewind (unit=12)
      rewind (unit=13)
      rewind (unit=14)
      call read(npg)
      end if
      stop
      end

```

```

c *****
c subroutine skros calculates model I point distribution
c approximation
c
c *****
c NB - number of variables
c corr - global correlation matrix
c xm - mean
c xs - standard deviation
c cs - skewness
c P - weighting function
c si - sign matrix
c   pp - p-plus weighting function
c   pm - p-minus weighting function
c   fi - independent component weighting function
c
c subroutine skros(NB)
c common /cor/corr(50,50)
c common /inp/xm(40),xs(40),cs(40)
c common /loca/x(40)
c dimension P(300), pp(40), pm(40), fi(400)
c common /signal/si(400,20)
c common /seq/IS(30)
c
c   subroutine SIGN: assign sign for the variates
c
c   print *, ' Model I approximation'
c   print *, ' .....'
c   print *, ' '
c   call SIGN(NB)
c   N = 2**NB
c
c   calculation of p-plus(pp) and p-minus(pm)
c
c   do 5 kp= 1,NB
c   pf= 1. + (cs(kp)/2. )**2.
c   pf1= 1.0-(1.0/pf)
c   pf2=pf1**0.5
c   print *, 'cs, pf = ', pf2,cs(kp)
c
c   sign value assigned according to skewness
c
c   if(cs(kp).lt.0.0.or.cs(kp).eq.0.0)then
c   pp(kp) = 0.5*(1. + pf2)
c   else
c   pp(kp) = 0.5*(1.-pf2)
c   end if
c   pm(kp)= 1.0-pp(kp)
c   print *, 'pm, pp = ', pm(kp),pp(kp)
c 5   continue
c

```

```

c    calculation of independent component weighting
c    function
c
do 100 kc = 1,N
    fi(kc) = 1.0
    do 110 jc = 1,NB
        if(si(kc,jc).lt.0.0)then
            fi(kc) = fi(kc)*pm(jc)
        else
            fi(kc) = fi(kc)*pp(jc)
        end if
110    continue
c    print *, 'fi = ', fi(kc)
100    continue
c
c    Calculation of weighting functions
c    P-lm (see equation 3.38 given in Thesis)
c
do 30 k = 1,N
    fa = 0.0
    do 35 jk = 1,NB-1
        do 40 kk = jk + 1,NB
c
c            component corresponding to variable x -l
c
            if(si(k,jk).lt.0.0)then
                pf2 = fi(k)/pm(jk)
            else
                pf2 = fi(k)/pp(jk)
            end if
c
c            component corresponding to variable x -m
c
            if(si(k,kk).lt.0.0)then
                pf2 = pf2/pm(kk)
            else
                pf2 = pf2/pp(kk)
            end if
c            print *, 'pf2 = ', pf2
c
c    calculation of correlation component weighting
c    function
c
pf1 = (pp(jk)*pm(jk)*pp(kk)*pm(kk))**0.5
c    print *, 'pf1 = ', pf1
c
c    calculation of sum of all the correlation component
c    weighting functions
c
fa = fa + si(k,jk)*si(k,kk)*corr(jk,kk)*pf1*pf2
40    continue
35    continue

```

```

c
c   calculation of weighting functions ratio
c   according to given skewness
c
do 45 jkk = 1,NB
if(si(k,jkk).lt.0.0)then
fx = (pp(jkk)/pm(jkk))**0.5
else
fx = (pm(jkk)/pp(jkk))**0.5
end if
c   print *, ' fx = ',fx
c
c   calculation of point estimates coordinates
c
x(jkk) = xm(jkk) + si(k,jkk)*xs(jkk)*fx
45 continue
c
c   weighting function:summation of independent and correlation
c   component weighting function
c
P(k) = fi(k) + fa
c   print *,P(k)
c   print *,(x(kl),kl=1,NB),P(k)
c
c   output coordinates
c
write(14,99)k,(x(IS(kl)),kl=1,NB),P(k)
99 format(i5,9f14.6)
30 continue
return
end

c   *****
c   Subroutine rosen calculates Rosenblueth (1975)
c   point approximation
c   *****
c   NB - number of variables
c   corr - global correlation matrix
c   xm - mean
c   xs - standard deviation
c   cs - skewness
c   P - weighting function
c   si - sign matrix
c
subroutine rosen(NB)
common /cor/corr(50,50)
common /inp/xm(40),xs(40),cs(40)
common /loca/x(40)
dimension P(300)
common /signal/si(400,20)
common /seq/IS(30)

```

```

c
c   SIGN: read the sign of the variates
c
c   print *, 'Rosenblueth Point Estimate Method'
c   print *, '.....'
c   print *, ' '
c   call SIGN(NB)
c   N = 2**NB
c
c   calculation of correlation component
c   weighting functions - fa
c
c   do 30 k=1,N
c   fa = 0.0
c   do 35 jk=1,NB-1
c   do 40 kk = jk+1,NB
c   fa = fa + si(k,jk)*si(k,kk)*corr(jk,kk)
40  continue
35  continue
c
c   calculation of point locations
c
c   do 45 jkk = 1,NB
c   x(jkk) = xm(jkk) + si(k,jkk)*xs(jkk)
45  continue
c
c   calculation of weighting function
c
c   P(k) = (1. + fa)/real(N)
c   print *,(x(kl),kl=1,NB),P(k)
c
c   output of estimates
c
c   write(10,99)k,(x(IS(kl)),kl=1,NB),P(k)
99  format(i5,9f14.6)
30  continue
c   return
c   end
c
c   *****
c   subroutine sklind calculates model II point distribution
c   approximation
c   *****
c   subroutine sklind(NB)
c
c   NB - number of variables
c   corr - global correlation matrix
c   xm - mean
c   xs - standard deviation
c   cs - skewness
c   Pm - P-minus weighting function

```

```

c      Pp - P-plus weighting function
c      Zm - Z-minus coordinate
c      Zp - Z-plus coordinate
c
common /cor/corr(50,50)
common /inp/xm(40),xs(40),cs(40)
common /loca/x(40)
common /seq/IS(30)
dimension Pm(40),Pp(40),Zm(20,20),Zp(20,20)
dimension sig(2),Z(20,20),P(2)
data sig(1),sig(2)/1.,-1./
c
rad = real(NB)**0.5
print *, 'Model II approximation'
print *, '.....'
print *, ' '
c
c      with the assumption of equally distributed
c      weighting functions
c
do 30 i=1,NB
do 45 j = i,NB
if(j.eq.i)then
if(j.eq.1)then
fra = (cs(i)**2.)*0.5/((4.*real(NB) + cs(i)**2.)*0.5)
c
c      jn=1 equals P plus is calculated
c
if(cs(i).gt.0.)then
Ppl = (1. - fra)/(2.*real(NB))
else
Ppl = (1. + fra)/(2.*real(NB))
end if
Pmin = (1./real(NB)) - Ppl
Zpl = rad*((Pmin/Ppl)**0.5)*xs(j)
Zmi = rad*((Ppl/Pmin)**0.5)*xs(j)
Zm(i,i) = Zmi
Zp(i,i) = Zpl
Pm(i) = Pmin
Pp(i) = Ppl
c      print *, 'i      Zm      Zp      Pm      Pp'
c      print *, i, Zm(i,i), Zp(i,i), Pm(i), Pp(i)
go to 45
end if
cm1 = 0.
cm2 = 0.
cm3 = 0.
do 50 k=1,i-1
Pcm13 = (Pp(k) - Pm(k))
Pcm2 = (Pp(k) + Pm(k))
cm1 = cm1 + Pcm13*Z(k,i)
cm2 = cm2 + Pcm2 * (Z(k,i)**2.)

```

```

cm3 = cm3 + Pcm13 * (Z(k,i)**3.)
50  continue
cm1 = 0. - (cm1)
cm2 = xs(i)**2. - cm2
cm3 = cs(i)*(xs(i)**3.) - cm3

c
c  cm1 = A, cm2 = B, cm3 = C
c
c  print *, 'j      A      B      C'
c  print *, j, cm1, cm2, cm3
ac1 = (cm3/real(NB) - cm1*cm2)
ac2 = 4. *(cm1*cm3 - (cm2**2.))
ac3 = (cm2/real(NB)) - (cm1**2.)
ac4 = (ac1**2. - ac2*ac3)**0.5
c  print *, ac3, ac1, ac2/4
c  print *, 'delta'
delta = (ac1**2. - ac2*ac3)
if(delta.lt.0.0)then
print *, j, ' delta = ', delta
print *, ' '
print *, 'Try Different Sequence'
print *, ' '
print *, ' '
go to 120
end if
Zp1 = (ac1 - ac4)/(2.*ac3)
Zp2 = (ac1 + ac4)/(2.*ac3)

c
den = (Zp1**2./real(NB)) - cm2
c  print *, 'den = ', den
diff = 0.
do 67 kp = 1, i-1
diff = Pp(kp) - Pm(kp) + diff
c  print *, diff, Pp(kp), Pm(kp), kp
67  continue
if(diff.lt.0.0)then
diff = -1.*diff
end if
c  if(cs(i).eq.0.0.and.cs(i-1).eq.0.0)then
if(diff.lt.0.0001.and.cs(i).eq.0.0)then
Zm1 = Zp1
Zm2 = Zp2
else
Zm1 = (cm3 - cm1*(Zp1**2.))/((Zp1**2./real(NB))
c - cm2)
Zm2 = (cm3 - cm1*(Zp2**2.))/((Zp2**2./real(NB))
c - cm2)
end if
Pm1 = ((Zp1/real(NB)) - cm1)/(Zm1 + Zp1)
Pp1 = (1./real(NB)) - Pm1
Pm2 = ((Zp2/real(NB)) - cm1)/(Zm2 + Zp2)
Pp2 = (1./real(NB)) - Pm2

```

```

c
c  check
c
diff1 = Zm1 - (-Zp2)
diff2 = Zp1 - (-Zm2)
if(diff1.lt.0.0)then
diff1 = -1.*diff1
end if
if(diff2.lt.0.0)then
diff2 = -1.*diff2
end if
if(diff1.lt.0.01.and.diff2.lt.0.01)then
Zp(i,i) = Zp1
Zm(i,i) = Zm1
Pm(i) = Pm1
Pp(i) = Pp1
c  print *,Zm(i,i),Zp(i,i),Pm(i),Pp(i),i
go to 45
else
print *, 'Check quadratic equation '
print *, 'diff1 = ',diff1, 'diff2 = ',diff2
print *, 'Pm1 = ',Pm1, 'Pp1 = ',Pp1
print *, 'Zm1 = ',Zm1, 'Zp1 = ',Zp1
print *, 'Pm2 = ',Pm2, 'Pp2 = ',Pp2
print *, 'Zm2 = ',Zm2, 'Zp2 = ',Zp2
chb1 = (Pp1*(Zp1**3.)) - (Pm1*(Zm1**3.))
chb2 = (Pp2*(Zp2**3.)) - (Pm2*(Zm2**3.))
print *, 'chb1,chb2 (C) = ',chb1,chb2
go to 120
end if
end if
sum2=0.
do 51 k=1,i-1
sum2 = sum2 + Z(k,i)*Z(k,j)
51  continue
cof = Pm(i)*Zm(i,i) + Pp(i)*Zp(i,i)
Z(i,j) =(xs(i)*xs(j)*corr(i,j) - (sum2/real(NB)))/cof
c  print *, 'Z(' ,i , ' , ' ,j , ') = ',Z(i,j),sum2,corr(i,j)
c  print *,i,j, 'Zcr = ',Z(i,j),cof,sum2,xs(i),xs(j)
45  continue
c
c  calculation of pairs coordinates and weighting functions
c
do 70 kl = 1,2
do 60 kj = 1,NB
if(kj.eq.i)then
if(kl.eq.1)then
x(kj) = xm(kj) + Zp(kj,kj)
P(kl) = Pp(i)
else
x(kj) = xm(kj) - Zm(kj,kj)
P(kl) = Pm(i)

```

```

end if
go to 60
end if
x(kj) = xm(kj) + sig(kl)*Z(i,kj)
c print *,x(kj),sig(kl),Z(i,kj)
60 continue
print *,(x(kn),kn=1,NB),P(kl)
write(13,99)i,(x(IS(kn)),kn=1,NB),P(kl)
99 format(i5,9f14.6)
70 continue
30 continue
120 return
end

c *****
c generate the input sign matrix for model I and
c Rosenblueth (1975)
c *****
c NB - number of variables
c j=No. of runs (2**NB)
c i=No. of variables
c
subroutine SIGN(NB)
common /signal/si(400,20)
dimension s(20,400)
N = 2.**NB
do 10 i =1,NB
do 20 j =1,N
rq = real(j/(2.**i))
iq = (j/(2.**i))
re = rq - iq
if (re.eq.0.0) re = 1.
s(i,j) = 1.
if (re.eq.0.5.or.re.lt.0.5) s(i,j)=-1.
20 continue
10 continue
do 30 i=1,N
do 29 j=1,NB
si(i,j) = s(j,i)
29 continue
30 continue
return
end

```

```

c      *****
c      subroutine lind calculates the point distribution
c      approximation of Lind (1983)
c
c      *****
c      NB - number of variables
c      corr - global correlation matrix
c      xm - mean
c      xs - standard deviation
c      cs - skewness
c      P - weighting function
c      si - sign matrix
c      subroutine lind(NB)
c      common /cor/corr(50,50)
c      common /inp/xm(40),xs(40),cs(40)
c      common /loca/x(40)
c      common /seq/IS(30)
c      dimension P(300)
c      dimension sig(2),Z(20,20)
c      data sig(1),sig(2)/1.,-1./
c
c      calculation of weighting functions
c      P(i)- weighting function
c
c      print *,' Lind Point Estimate Method '
c      print *,' ..... '
c      print *,' '
c      do 30 i=1,NB
c
c      assumption of equally distributed weighting function
c
c      P(i)= 1./(2.*real(NB))
c      do 47 jn=1,2
c      do 45 j = i,NB
c      if(j.eq.i)then
c      sum = 0.
c      do 50 k=1,i-1
50      sum = sum + Z(k,i)**2.
c      continue
c      Z(i,j) = (real(NB)*(xs(i)**2.) - sum)**.5
c      go to 45
c      end if
c      sum 2=0.
c      do 51 k=1,i-1
c      sum 2 = sum 2 + Z(k,i)*Z(k,j)
51      continue
c      Z(i,j) =(real(NB)*xs(i)*xs(j)*corr(i,j) - sum2)/Z(i,i)
45      continue
c
c      calculation of coordinates
c

```

```

do 48 jl=1,NB
x(jl) = xm(jl) + sig(jn)*Z(i,jl)
48 continue
c
c output
c
print *,(x(j),j=1,NB),P(i)
write(11,99)i,(x(IS(j)),j=1,NB),P(i)
99 format(i5,9f14.6)
47 continue
30 continue
return
end

c *****
c subroutine altpem calculates the point distribution
c approximation of Harr (1989) methodology
c *****
c
c subroutine altpem(NB)
c
c NB - number of variables
c corr - global correlation matrix
c xm - mean
c xs - standard deviation
c cs - skewness
c P - weighting function
c evalue - eigenvalue
c evecr - eigenvector
c
c common /cor/corr(50,50)
c common /inp/xm(40),xs(40),cs(40)
c common /loca/x(40)
c common /seq/IS(30)
c dimension P(300)
c dimension sig(2)
c common /eig/evalue(50), evecr(50,50)
c data sig(1),sig(2)/1.,-1./

c
c print *, ' Alternate Point Estimate Method '
c print *, ' ..... '
c call eigen(NB)

c
c calculation of weighting functions
c
c rad = real(NB)**0.5
c do 30 k=1,NB
c P(k)= evalue(k)/(2.*real(NB))
c do 47 jn=1,2
c do 45 jkk = 1,NB
c radc = rad *evecr(jkk,k)

```

```

c
c   calculation of point locations
c
x(jkk) = xm(jkk) + sig(jn)*radc*xs(jkk)
45 continue
print *,(x(jkc),jkc=1,NB),P(k)
write(12,99)k,(x(IS(jkc)),jkc=1,NB),P(k)
99 format(i5,9f14.6)
47 continue
30 continue
return
end

c
c   This subroutine calculates the Eigen values and Eigen vectors
c   of a Correlation matrix in a descending order using the
c   subroutines - JACOBI & EIGSRT
c   Parameter Rowmax and Colmax depend on the size of the
c   correlation matrix
c   EVALUE and EVECTOR are printed in a certain pattern
c   Input file - Correlation Matrix
c   Name the Output file
c
subroutine eigen(ipara)
common /cor/corr(50,50)
common /eig/evalue(50), evector(50,50)
integer rowmax, colmax, row, col
rowmax = ipara
colmax = ipara

call jacobi (corr, rowmax, colmax, evalue, evector, nrotn)
call eigsrt (evalue, evector, rowmax, colmax)

c
do 60 col=1, colmax
  print *, evalue(col)
  print *,(evector(row,col),row = 1,rowmax)
60 continue
c
return
end

subroutine jacobi (a, n, np, d, v, nrot)
parameter (nmax=100)
dimension a(50, 50), d(50), v(50, 50), b(nmax), z(nmax)

do 12 ip=1, n
  do 11 iq=1, n
    v(ip, iq)=0.
11 continue
    v(ip, ip)=1.
12 continue

```

```

do 13 ip=1, n
  b(ip)=a(ip, ip)
  d(ip)=b(ip)
  z(ip)=0.
13  continue

nrot=0.

do 24 i=1, 50
  sm=0.
  do 15 ip=1, n-1
    do 14 iq=ip+1, n
      sm=sm+abs(a(ip,iq))
14    continue
15  continue
  if (sm .eq. 0.) return
  if (i .lt. 4) then
    tresh=0.2*sm/n**2
  else
    tresh=0.
  endif

  do 22 ip=1, n-1
    do 21 iq=ip+1, n
      g=100. *abs(a(ip,iq))
      if ((i.gt.4) .and. (abs(d(ip))+g .eq. abs(d(ip)))
&      .and. (abs(d(iq))+g .eq. abs(d(iq)))) then
        a(ip,iq) =0.
      else if (abs(a(ip,iq)).gt.tresh) then
        h=d(iq)-d(ip)
        if (abs(h)+g .eq. abs(h)) then
          t=a(ip, iq)/h
        else
          theta= 0.5*h/a(ip, iq)
          t= 1./(abs(theta)+sqrt(1.+theta**2))
          if (theta .lt. 0.) t=-t
        endif
        c=1./sqrt(1+t**2)
        s=t*c
        tau=s/(1.+c)
        h=t*a(ip, iq)
        z(ip)=z(ip)-h
        z(iq)=z(iq)+h
        d(ip)=d(ip)-h
        d(iq)=d(iq)+h
        a(ip,iq)=0.
        do 16 j=1, ip-1
          g=a(j,ip)
          h=a(j,iq)
          a(j,ip)=g-s*(h+g*tau)
          a(j,iq)=h+s*(g-h*tau)
16      continue

```

```

do 17 j=ip+1, iq-1
  g=a(ip,j)
  h=a(j,iq)
  a(ip,j)=g-s*(h+g*tau)
  a(j,iq)=h+s*(g-h*tau)
17  continue
do 18 j=iq+1, n
  g=a(ip,j)
  h=a(iq,j)
  a(ip,j)=g-s*(h+g*tau)
  a(iq,j)=h+s*(g-h*tau)
18  continue
do 19 j=1, n
  g=v(j,ip)
  h=v(j,iq)
  v(j,ip)=g-s*(h+g*tau)
  v(j,iq)=h+s*(g-h*tau)
19  continue
nrot=nrot+1
endif
21  continue
22  continue
do 23 ip=1, n
  b(ip)=b(ip)+z(ip)
  d(ip)=b(ip)
  z(ip)=0.
23  continue
24  continue

```

pause '50 iterations should never happen'

```

return
end

```

```

subroutine eigsrt (d, v, n, np)
dimension d(50), v(50, 50)
do 13 i=1, n-1
  k=i
  p=d(i)
  do 11 j=i+1, n
    if (d(j).ge.p) then
      k=j
      p=d(j)
    endif
11  continue
  if (k.ne.i) then
    d(k)=d(i)
    d(i)=p
    do 12 j=1, n
      p=v(j,i)
      v(j,i)=v(j,k)

```

```

12         v(j,k)=p
           continue
           endif
13 continue
           return
           end

c *****
c subroutine dacor(na)
c *****
c
c This program forms the correlation matrix from
c the partitioned sub layer correlation matrices
c
c The sub layer correlation matrices are entered as
c row entries of the matrix C and then the global
c matrix correlation is formed
c
c Ex. For a 3 x 3 correlation matrix
c Enter the Lower Triangular Matrix only with the
c diagonal elements and it is stored as
c C(.,1),C(.,2),..... C(.,6)
c . - Layer number
c
c To run the LLUVIA it is assumed that each sub layer
c is fractured and the matrix and fracture properties
c are uncorrelated. Therefore each matrix and fracture
c properties can be treated as sub layers to form
c the GLOBAL CORRELATION MATRIX
c
c common /cor/corr(50,50)
c dimension C(10,20)
c integer rowmax,colmax
c
c Reading the Sub Layers Correlation Matrices
c
c open(unit=15,file='global',status='unknown')
c open(unit=16,file='entry',status='unknown')
c print *, ' Input the Correlation Matrix '
c print *, '
c print *, ' Example :-'
c print *, ' -----'
c print *, ' LAYER 1'
c print *, ' matrix(sub layer 1) Fracture(sub layer 2)'
c print *, '
c print *, ' -----'
c print *, ' LAYER 2'
c print *, ' matrix(sub layer 3) Fracture(sub layer 4)'
c print *, '
c print *, ' -----'
c print *, ' total number of sub layers = 4'
c print *, '

```

```

print *, ' '
print *, 'Enter the total number of sub layers == >'
read *, nl
print *, ' '
print *, 'Enter the number of properties for each'
print *, 'sub layer == >'
read *, np

c
c   Calculating the number of elements in each sub
c   correlation matrix
c
do 2 i = 0, np-1
npe = npe + np - i
2 continue

c
print *, 'Entries per each sub layer = ', npe
print *, ' '

c
print *, ' If the "entry" file is already created'
print *, '       Enter "0"'
print *, '       Else Enter "1"'
read *, ien
print *, ' '
do 6 i = 1, nl
print *, '       Layer = ', i
do 5 j = 1, npe
if (ien.eq.0) then
read(16, *)C(i,j)
print *, 'C( ', i, j, ') = ', C(i,j)
go to 5
else
print *, 'C( ', i, j, ') = '
read *, C(i,j)
write(16, *)C(i,j)
end if
5 continue
6 continue

c
c   Initializing the Assembled Correlation Matrix
c
na = nl * np
do 7 i = 1, na
do 7 j = 1, na
corr(i,j) = 0.0
7 continue

c
c   Assembling the Correlation Matrix
c
n = 1
k = 1
do 10 i = 1, na, np
ie = k * np

```

```

do 20 kk = i, ie
do 30 j = kk,ie
  corr(kk,j) = C(k,n)
  n = n+1
  corr(j,kk) = corr(kk,j)
30 continue
20 continue
  n = 1
  k = k + 1
10 continue
c
c   Writing it to an output file
c
  rowmax = na
  colmax = na
  do 40 i = 1,na
  write(15,*)(corr(i,j),j=1,na)
40 continue
  return
  end

subroutine read(npg)
c *****
c   subroutine read creates file for the LLUVIA
c   run
c   *****
  common /lay/ES(100),ESL(100),ESF(20)
  open(unit=15,file='rose.in')
  open(unit=16,file='lin.in')
  open(unit=17,file='alt.in')
  open(unit=18,file='slin.in')
  open(unit=19,file='skro.in')
  print *, 'Enter the number of Layers ='
  read *,nl
  print *, ' '
c   print *, 'enter the number of properties'
c   print *, 'per layer'
c   print *, '           If the Layer is Fractured'
c   print *, '           Enter =====> 1'
c   read *,ifr
c   if(ifr.eq.1)then
c   np=8
c   else
c   np=4
c   end if
c   np = npg/nl
c   N=nl*np
c
c   first reading the Rosenblueth point estimates
c
  NN = 2**N

```

```

c
c   assigning the unit numbers 10 & 15 and km=1
c
      km = 1
      ku = 10
      kr = 15
c
110  do 10 ii=1,NN
      read(ku,*)k,(ES(i),i=1,N),P
      write(kr,*)ii,P
      j=1
      do 15 ij=1,np
        do 20 if=ij,N,np
          ESL(j)=ES(if)
c
c   Hydraulic conductivity model four parameters for
c   matrix and fractures and the unit is changed from
c   cm/hour to m/sec
c
          if(ij.eq.1.or.is.eq.5)then
            ESL(j)=ESL(j)/(100.*3600.)
          end if
          j=j+1
20   continue
c
c   integer j is updated therefore ESL(1)..ESL(nl)
c   keep changing
c
          j=1
          write(kr,*)(ESL(kl),',',kl=1,nl-1),ESL(nl)
15   continue
          if(np.eq.4)then
            do 23 if=1,np
              write(kr,*)(ESF(kl),',',kl=1,nl-1),ESF(nl)
23   continue
            end if
10   continue
            if(km.eq.1)then
              km=2
              ku = 11
              kr = 16
              NN = 2*N
              go to 110
            else if(km.eq.2)then
              km=3
              kr=17
              ku = 12
              NN = 2*N
              go to 110
            else if(km.eq.3)then
              km=4
              kr = 18

```

```
ku = 13
NN = 2*N
go to 110
else if(km.eq.4)then
km=5
kr = 19
ku = 14
NN = 2**N
go to 110
end if
return
end
```