A SIMPLE APPROACH TO RELIABILITY, RISK, AND UNCERTAINTY ANALYSIS OF HYDROLOGIC, HYDRAULIC, AND ENVIRONMENTAL ENGINEERING SYSTEMS

BY

ADITYA TYAGI

Bachelor of Civil Engineering
University of Roorkee
Roorkee, UP, India
1987

Master of Civil Engineering
University of Roorkee
Roorkee, UP, India
1989

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2000

A SIMPLE APPROACH TO RELIABILITY, RISK, AND UNCERTAINTY ANALYSIS OF HYDROLOGIC, HYDRAULIC, AND ENVIRONMENTAL ENGINEERING SYSTEMS

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

PREFACE

Application of uncertainty and reliability analysis is an essential part of many problems related to modeling and decision making in the area of environmental engineering and water resources. Computation efficiency, understandability, and easier application have made the first order approximation (FOA) method a favored tool for uncertainty analysis. In many instances doubtful situations may arise where the accuracy of FOA estimates becomes questionable. Presently, no clear-cut guidelines specifying where FOA should be used are available.

The objective of this dissertation was to investigate the important factors affecting the exactness of FOA estimates and develop a simple correction procedure useful for practicing engineers to correct the FOA estimates for the mean and the variance of a model output. To carryout reliability and risk analysis, knowledge of distribution for a model output is very important. Therefore the other objective of this thesis was to develop a simple approach for calculating the higher-order moments of a model output from which an appropriate distribution can be chosen.

Methods to correct FOA estimates for the mean and variance of a model output were developed. Further, a generic expectation function approach was developed to determine higher-order moments of a model output correctly.

I sincerely thank my doctoral advisory committee - Drs. Charles T. Haan (Chair), Billy J. Barfield, Glenn O. Brown, and John N. Veenstra - for their excellent guidance and support in the completion of this research.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my major advisor, Regents Professor and Sarkey's Distinguished Professor Dr. C. T. Haan for his intelligent supervision, constructive guidance, inspiration, encouragement, and friendship throughout the completion of this work. My sincere appreciation extends to my other committee members Dr. Billy J. Barfield, Dr. Glenn O. Brown, and Dr. John N. Veenstra, whose guidance, assistance, encouragement, and friendship are also invaluable. For all of my committee members, I owe a debt of gratitude.

Moreover, I wish to express my sincere appreciation to Dr. Ellen Stevens, Dr. Jeanne Hill, and Dr. Patricia Haan for their suggestions and assistance for this study. I am very grateful to all of my friends and staff members of Biosystems and Agricultural Engineering Department who have helped me directly or indirectly in attaining this goal.

I would also like to give my deep appreciation to my mother and wife for supporting my decision to embark on this adventure and their continuous strong moral support, and encouragement throughout this study.

Finally, I would like to dedicate this work to the memory of my father, Deo Dutt Shastri, who always encouraged his son toward attaining higher goals and made his son a dedicated and disciplined hard worker.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

BACKGROUND, LITERATURE REVIEW, AND OBJECTIVES

Background

The design and analysis of hydrologic, hydraulic, and environmental projects are subject to uncertainty because of the inherent uncertainty in natural systems, a lack of understanding of the causes and effects in various physical, chemical, and biological processes occurring in natural systems, and insufficient data. As a result of these uncertainties, the performance reliability of a project is uncertain. A reliable assessment of the performance of any water resources project requires an assessment of the validity of predicted loads (such as discharges and pollutant loads) and capacities (ability to perform under a given load without any harm). Typically the loads are assessed using models having a number of parameters which can be determined with varying degrees of certainty. These parameters are best represented as random variables. Consequently, the model response, being a function of random variables, is best represented as a random variable. For reliable design and analysis of a project, it becomes necessary to address the uncertain nature of model outputs. Reliability, risk, and uncertainty analysis are therefore becoming increasingly important in modeling and designing water resources infrastructure and decision support systems. In some cases, uncertainty analysis is mandatory, particularly when critical decisions involve potentially high levels of risk.

Many problems are best approached using probabilistic and reliability methods. Examples include determining the probability of a structural failure or the life expectancy of a hydraulic structure under uncertainty. The prediction and evaluation of pollution of surface and/or subsurface environments and decisions regarding remedial actions often rely on probabilistic approaches. Quantification of the underlying uncertainty is central to each of these problems.

Probabilistic and reliability analyses are based on knowledge of the underlying parameter uncertainties. Two methods, Monte Carlo Simulation (MCS) and First-Order Approximation (FOA) are generally used to assess the uncertainty associated with model outputs. A MCS requires several thousand repetitive runs of a model and is therefore computationally demanding. With the advent of high-speed computers, the computational challenges to probabilistic analysis have been largely removed but philosophical and conceptual aspects remain. MCS has other limitations. Rules for determining the number of simulations required for convergence are not available (Melching, 1995). Often, the information about distribution function(s) of input variable(s) required to conduct a MCS is not available and can not be obtained. Therefore, it requires judgement on the part of the modeler to create theoretical input sample distributions that are representative of the parameter populations.

FOA is an approximate method that gives estimates of means and variances only. It has several advantages over MCS. FOA is more computationally efficient and provides a measure of a model's sensitivity to each input random variable, thus providing a better understanding of the processes being modeled. The assumptions typically cited for FOA to yield good results (Melching, 1995) are: (1) linearity in functional relationships, (2) small coefficients of variation of the most sensitive uncertain variables, and (3) normal distributions for the uncertain variables. Moore and Clarke (1981) expressed that FOA assumptions are "rarely likely to be justifiable with models containing nine or twelve parameters." Despite its several conceptual shortcomings, FOA has been used quite successfully in a wide variety of fields. The exactness of the estimates is influenced in part by the degree of nonlinearity in the functional relationship and parameter uncertainty. Due to unavailability of clear-cut guidelines as to when FOA should be applied, FOA has been misused in many instances. Conclusions based on such applications may be highly misleading, and in any design and/or decision-making process may have serious consequences.

Various researchers have suggested a number of criteria for FOA. Garen and Burges (1981) found satisfactory results with FOA when the CV of the input parameters was less than or equal to 0.25. Cornell (1972) and Burges (1979) suggested that FOA is applicable to moderately nonlinear systems when the CV is less than or equal to 0.2. Gardner et al. (1981) found that the validity of the linear approximation deteriorates rapidly when the CV was more than 0.3. The best agreement between MCS and FOA estimates occurs when MCS output distributions are symmetric (Scavia et al., 1981a).

Several researchers detected significant nonlinearity effects while comparing variance estimates from FOA and MCS. To overcome the problem of nonlinearity, several predictors were proposed (Beale, 1960; Bates and Watts, 1980; Bates, 1988; Stevens, 1993). Bates (1988) and Stevens (1993) indicated that the predictors for nonlinearity developed so far work well only in specific applications and that no well-accepted, generalized nonlinearity measure is available.

To date, the only widely used criterion for the validity of FOA variance is to restrict the parameter CV to less than 0.2. This is a very restrictive assumption in water resource systems modeling, where there is often a great uncertainty in the parameters (Johnson, 1996). It can be shown that FOA has performed well in some situations when the parameter uncertainty is higher than 0.2. Smith and Charbeneau (1990) suggested that FOA can be used if the difference between function gradients at the mean and one standard deviation away from the mean are less than some acceptable range (5-10%), however, this method also has limitations

Literature Review

Water resources and environmental engineering systems deal with the extremely complex nature of the physical, chemical, biological, and socio-economical processes. While designing and/or analyzing a given system, most often a mathematical model describing the interrelationships and interactions among its component processes is used. Despite a tremendous research effort to evolve a better understanding of various processes, a number of uncertainties still exist due to lack of perfect knowledge concerning the phenomena and processes involved. Therefore, most models used in designing and analyzing engineering systems involve a number of uncertainties.

In water resources and environmental engineering, the decisions on the layout, capacity, and operation of a system largely depend on the system response under some anticipated design conditions. If the response of any of the components in a system is considered uncertain, the response of the system under the design conditions must also be considered uncertain. The presence of uncertainties makes the conventional deterministic design practice inappropriate due to its inability to account for possible variations of system responses. The issues involved in the design and analysis of water resources and environmental engineering systems under uncertainty are multi-dimensional. Therefore, quantification of system uncertainties is imperative in order to design and/or operate a project successfully. A systematic quantitative uncertainty analysis provides insight into the level of confidence warranted in model estimates and in understanding judgements associated with the modeling process. It may also play an illuminating role in identifying how robust the conclusions about model results are and help target data gathering efforts.

Uncertainty refers to lack of knowledge about specific factors, parameters, or models. There are a number of distinct sources of uncertainty in the analysis and design of engineering systems. In general, in the field of water resources and environmental engineering, uncertainties can be classified under the general headings (Yen et al., 1986; Beck, 1987; Melching and Anmangandla, 1992; Melching, 1995) of natural uncertainties, model uncertainties, parameter uncertainties, and data uncertainties.

Natural uncertainty is associated with the inherent randomness of natural processes such as the occurrence of precipitation, flood events, and change in climatological conditions. According to Beck (1987) uncertainty resulting from natural variability includes environmental variability due to system disturbances, aggregation uncertainty due to spatial heterogeneity, and genetic variability which could be indistinguishable from the errors of parameter estimation.

The structure of mathematical models employed to represent a phenomenon of interest is often a key source of uncertainty. Models are only an abstraction of a real-world system. The problem boundary encompassed by a model may be incorrect or incomplete. Significant approximations and idealizations are often an inherent part of the assumptions upon which a model is built. Competing models may be available based on different scientific or technical assumptions. Model uncertainty reflects the inability of the model to precisely represent the true physical behavior of a system. Model uncertainty includes uncertainty due to necessary simplification of real-world processes, misspecification of the model structure, model misuse, or use of inappropriate surrogate variables.

Parameter uncertainty is a result of the inability to quantify the input parameters of a model accurately due to measurement errors, sampling errors, systematic errors, etc. Most of the models used in hydrologic, hydraulic, and environmental engineering involve several physical or empirical parameters that cannot be quantified accurately. Parameter uncertainty could also be caused by changes in the operational conditions of a system, inherent variability of inputs and parameters in time and in space, and insufficiency in the quantity or quality of data.

Data uncertainties include measurement errors, measurement limitations, inconsistency and non-homogeneity of data, and lack of data due to time and money constraints

In this thesis, the impact of input parameter uncertainties on the probabilistic and reliability analyses of hydrologic, hydraulic, and environmental engineering systems is studied. Parameter uncertainties are the variation in a parameter due to an inability to precisely quantify that parameter. Parameter uncertainty may be partially quantified by the coefficient of variation (CV). The CV of a parameter is the ratio of the standard deviation to the mean and offers a normalized measure useful and convenient for comparison.

Uncertainty Analysis

The main objective of uncertainty analysis is to assess the statistical properties of model outputs as a function of stochastic input parameters. In water resources engineering projects, design quantity and model outputs are functions of several parameters, not all of which can be quantified with absolute accuracy. The task of uncertainty analysis is to determine the uncertainty features of the model outputs as a function of uncertainties in the model itself and in the stochastic parameters involved. It provides a formal and systematic framework to quantify the uncertainty associated with the model outputs. Furthermore, it offers the designer useful insights regarding the contribution of each stochastic parameter to the overall uncertainty of the model outputs. Such knowledge is essential to identify the important parameters to which more attention should be given to have a better assessment of their values and, accordingly, to reduce the overall uncertainty in model output. Quantitative characterization of uncertainty provides an estimate of the degree of confidence that can be placed on the analysis and findings.

As an example, water quality models are formulated to describe both observed conditions and predict planning scenarios that may be substantially different from observed conditions. Planning and management activities such as checking basin wide water quality for regulatory compliance, waste load allocation, etc., require the assessment of hydrologic, hydraulic, and water quality conditions beyond the range of observed data. These inadequacies regarding model parameters or inputs force water quality modelers to characterize the impacts of parameter uncertainties quantitatively so that appropriate decisions regarding water pollution abatement programs can be made. The most complete and ideal description of uncertainty is the probability density function

(PDF) of the quantity subject to uncertainty. However, in most practical problems a probability function is very difficult, if not impossible, to derive precisely. In most situations, the main objective of uncertainty analysis is to evaluate the first and second moments of a model output in terms of input random variables.

Reliability and Risk Analysis

Reliability and risk analysis is a technique for identifying, characterizing, quantifying, and evaluating the probability of a pre-identified hazard. It is widely used by private and government agencies to support regulatory and resource allocation decisions. In most hydrologic, hydraulic, and environmental engineering problems, empirically developed or theoretically derived mathematical models are used to evaluate a system's performance. These models involve several uncertain parameters that are difficult to accurately quantify. An accurate reliability assessment of such models would help the designer build more reliable systems and aid the operator in making better maintenance and scheduling decisions.

The reliability of a system can be most realistically measured in terms of probability. The failure of a system can be considered as an event in which the demand or loading, L, on the system exceeds the capacity or resistance, R, of the system so that the system fails to perform satisfactorily for its intended use. The objective of reliability analysis is to ensure that the probability of the event ($R < L$) throughout the specified useful life is acceptably small. The risk, P_f, defined as the probability of failure, can be expressed as (Ang and Tang, 1984; Yen et al., 1986)

$$P_f = P(L > R) \tag{1-1}$$

where P denotes the probability function. Equation (1-1) can be rewritten in terms of the performance function Z as

$$P_f = P(Z < 0) \tag{1-2}$$

where Z is defined alternatively as

$$Z = R - L \tag{1-3}$$

$$Z = \frac{R}{L} - 1 \tag{1-4}$$

$$Z = \ln\left(\frac{R}{L}\right) \tag{1-5}$$

The reliability, $\Re$, of the system can be written as

$$\Re = P(Z > 0) = 1 - P_f \tag{1-6}$$

In general, from (1-1), the risk can be expressed as

$$P_f = \int_a^b \int_c^l p_{R,L}(r,l)\,dr\,dl \tag{1-7}$$

where $p_{R,L}(r,l)$ is the joint probability density function of R and L; c is the lower bound of R; and a and b are the lower and upper bounds of L respectively. The resistance, R, and load, L, are random variables given as

$$R = g_1(\underline{U}) \tag{1-8}$$

$$L = g_2(\underline{V}) \tag{1-9}$$

where, $\underline{U}$ is the vector representing input parameters of the model representing R; and $\underline{V}$ is the vector representing input parameters of the model representing L. In some problems $\bar{L}$ may be a deterministic quantity representing a hydrologic/hydraulic/environmental target level such as peak discharge, volume, contaminant concentration in soil, water, and air, minimum dissolved oxygen in a stream, critical cancer risk, etc. Alternatively, by using the performance variable Z defined in (1-3), (1-4), and (1-5), the risk can be written as

$$P_f = P(Z < 0) = \int_{-\infty}^{0} p_Z(z)dZ \tag{1-10}$$

where $p_Z(z)$ is the probability density function of Z. The probability distribution of Z is unknown, or difficult to obtain. In most cases the exact distribution of Z may not be required as any of several distributions can be used to make a decision if correct information about the moments of $p_Z(z)$ is available.

Uncertainty, Risk, and Reliability Analysis Methods

Ideally, a probability distribution function should be obtained to do a complete assessment of the uncertainty, risk, and reliability analysis of a given system. This requires determination of the joint probability distribution function for all the significant sources of uncertainty affecting the output of the system. However, the determination of probability distributions for the basic variables is quite difficult and involves several assumptions. Further, the multivariate combination and integration of the input variable distributions is a daunting task. In real life problems, the aggregation of uncertainties in the basic variables of a model into measures of overall model-output uncertainty/reliability are done in an approximate manner. Several methods that have been used in water resources and environmental engineering have been discussed.

First Order Approximation Method

The first order approximation (FOA) method can be used to estimate the amount of uncertainty, or scatter, of a dependent variable due to uncertainty about the independent variables included in a functional relationship. Benjamin and Cornell (1970) have described first order approximation (FOA) technique in detail.

To present the general methodology of first order approximation, consider a output random variable, Y, which is a function of n random variables. Mathematically, Y can be expressed as

$$Y = g(\underline{X}) \tag{1-11}$$

where $\underline{X} = (X_1, X_2, ..., X_n)$, a vector containing n random variables. In FOA, a Taylor series expansion of the model output is truncated after the first-order term

$$Y = g(\underline{X}_e) + \sum_{i=1}^{n} (X_i - X_{ie})\left(\frac{\partial g}{\partial X_i}\right)_{X_e} \tag{1-12}$$

where $\underline{X}_e = (X_{1e}, X_{2e}, ..., X_{ne})$, a vector representing the expansion points. In FOA applications to water resources and environmental engineering, the expansion point is commonly the mean value of the basic variables. Thus, the expected value and variance of Y are

$$E[Y] \approx g(\underline{\overline{X}}) \tag{1-13}$$

$$Var(Y) = \sigma_Y^2 \approx \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\partial g}{\partial X_i}\right)_{\overline{X}_i}\left(\frac{\partial g}{\partial X_j}\right)_{\overline{X}_j} E\left[\left(X_i - \overline{X}_i\right)\left(X_i - \overline{X}_j\right)\right] \tag{1-14}$$

where σ_Y is the standard deviation of Y; $\overline{X} = \left(\overline{X}_1, \overline{X}_2, ...\overline{X}_n\right)$, a vector of mean values of the input basic variables. If the basic variables are statistically independent, the expression for $Var(Y)$ becomes

$$Var(Y) = \sigma_Y^2 \approx \sum_{i=1}^{n}\left[\left(\frac{\partial g}{\partial X_i}\right)_{\overline{X}_i} \sigma_{X_i}\right]^2 \tag{1-15}$$

To estimate the reliability of the system, $\mathfrak{R}$, it is typically assumed that Z is normally distributed. Using $p_Z(z)$ to be a normal distribution with its parameters $E[Z]$ and σ_Z determined by FOA, (1-2) and (1-6) are used to determine the risk and reliability of a given system.

An alternative method to define a system reliability is the reliability index, β, which is defined as the reciprocal of the coefficient of variation of Z, given as

$$\beta = \frac{E[Z]}{\sigma_Z} \tag{1-16}$$

The great advantage of FOA is its simplicity, requiring knowledge of only the first two statistical moments of the basic variables and simple sensitivity calculations about selected central values. FOA is an approximate method that may suffice for many applications (Ku, 1966), but the method does have several theoretical and/or conceptual shortcomings (Melching, 1992a; Cheng, 1982). The main weakness of the FOA method is that it is assumed that a single linearization of the system performance function at the central values of the basic variables is representative of the statistical properties of system performance over the complete range of basic input variables. The accuracy of the estimates is influenced in part by the degree of nonlinearity in the functional relationship, and the importance of higher-order terms which are truncated in the Taylor series expansion (Burn and McBean, 1985). In applying FOA in risk and reliability analyses, it is generally assumed that the performance function is normally distributed, which is seldom true. Any attempt to characterize the tails of the actual distribution based on an assumption of normality is likely to result in an inexact answer (Burn and McBean, 1985).

Despite its shortcomings FOA has been used very widely in hydrologic, hydraulic, and environmental engineering. Examples of FOA application in hydrologic and hydraulic engineering design include Tang and Yen (1972), Yen and Tang (1976), Burges (1979), Yen et al. (1980), Tung and Mays (1980, 1981), Lee and Mays (1986), and Cesare (1991). Application examples related to hydrologic modeling include Garen and Burges (1981), Townley (1984), Townley and Wilson (1985), Melching (1992a, 1992b), Kuczera (1988), Bates and Townely (1988), Jones (1989), Lei and Schilling (1993). Examples in groundwater contamination modeling include Loague and Green (1988), Loague et al. (1989,1990), Smith and Charbeneau (1990), and Loague (1991). Examples of applying FOA in water quality and ecological modeling include Burges and Lettenmaier (1975), Argentesi and Olivi (1976), Lettennmaier and Richey (1979), Reckhow (1979a, 1979b), Scavia (1980), Dettinger and Wilson (1981), Beck (1981a, 1981b), Scavia et al. (1981a, 1981b), Devary and Doctor (1982), Chadderton et al. (1982), Walker (1982), Van Straten (1983), Burn and McBean (1985), Tung and Hathhorn (1989), Melching and Anmangandla (1992), Melching and Yoon (1996), and Zhang and Haan (1996).

Response Surface Methods

The response surface (SR) method is very similar to the FOA method. While the FOA method deals directly with the performance function, the RS approach involves approximating the original, complicated system performance function with a simpler, more computationally tractable system model. This approximation typically takes the form of a first or second order polynomial

$$Y = g(\underline{X}) \approx G(\underline{X}) \approx a_0 + a_1 X_1 + \ldots\ldots + a_n X_n + a_{n+1} X_1^2 + \ldots\ldots + + a_{2n} X_n^2 + + a_{2n+1} X_1 X_2 + \ldots \quad (1\text{-}17)$$

where $G(\underline{X})$ is the approximate function representing the original function $g(\underline{X})$. Determination of the constants is accomplished through a linear regression about some nominal value, typically the mean. Given the new performance function, the analysis proceeds in exactly the same manner as the FOA method. This method has not been used much in the area of water resources and environmental engineering.

Monte Carlo Simulation

In Monte Carlo Simulation (MCS), probability distributions are assumed for the uncertain input variables for the system being studied. Random values of each of the uncertain variables are generated according to their respective probability distributions and the model describing the system is executed. By repeating the random generation of variable values and model execution steps many times, the statistics and an empirical probability distribution of the model output can be determined. The accuracy of the statistics and probability distribution obtained from MCS is a function of the number of simulations performed and the adequacy of the assumed parameter distributions.

MCS is an art (Burges and Lettenmair, 1975). It requires judgement on the part of the modeler to create theoretical input sample distributions that are representative of the populations and to estimate the number of trials needed to generate the input and output density functions. There is no strictly defined answer to either of these questions.

A key problem in applying the MCS method is estimating the necessary sample size. One empirical test to determine the adequacy of the sample size consists of iterating the sample program with increasingly greater sample sizes and estimating the convergence rate of the sample mean value towards the population mean (Burges and Lettenmair, 1975). The error in the estimation of the population mean is inversely proportional to the square root of the number of trials. To improve the estimate by a factor of two, the sample size must increase by a factor of four. If the sample size is n, the standard deviation of the mean is $1/\sqrt{n}$ times the standard deviation of the population. This indicates that the sample size must be large (Siddall, 1983). As the sample size increases, the precision of the empirical percentile estimates of a model output improves (Modarres, 1993). However, Martz (1983) noted that the rate of convergence to the true distribution decreases as the size of sample increases.

The requirement of generating very large samples is a serious problem with MCS (Siddall, 1983). The method often entails sample sizes that are in the range of 5,000 to 20,000 members. Generally, the number of required samples increases with the variances and the coefficient of skewness of the input distributions (Burges and Lettenmair, 1975).

MCS has been used to analyze uncertainty, risk, and reliability of many water resources and environmental engineering systems. Many of these applications of MCS were to provide a check of less computationally intensive methods. Examples in hydrology includes Freeze (1975), Smith and Freeze (1979), Smith and Hebbert (1979), Gardner et al (1980), Smith and Schwartz (1980), Clifton and Neuman (1982), Takasao and Takara (1989), Warwick and Wilson (1990), Goldman et al (1990), Binley et al (1991), Krajewski et al. (1991), Beven and Binley (1992), Harlin and Kung (1992), etc. In the area of environmental engineering MCS has been used very extensively in water quality modeling studies. Some examples are O'Neill (1971), Burges and Lettenmaier (1975), Tiwari and Hobbie (1976), Gardner et al. (1980, 1981) O'Neill et al. (1980), Hornberger (1980), Montgomery et al. (1980), Smith and Schwartz (1981), Fedra et al. (1981), Scavia et al. (1981b), Walker (1982), Gardner and O'Neill (1983), Malone et al. (1984), Van De Kramer (1983), Black and Freyburg (1987), and Batchelor et al.(1998).

Another simulation technique similar to MCS is the Latin hypercube sampling (LHS) in which stratified sampling approach is used. In LHS the probability distribution of each basic variable are subdivided into non-overlapping intervals (say m) each with equal probability (1/m). Random values of the basic variables are simulated such that each range is sampled only once. The order of the selection of the ranges is randomized and the model is executed m times with the random combination of basic variables from each range for each basic variable. The output statistics and distributions may then be approximated from the sample of m output values. McKay et al. (1979) has shown that the stratified sampling procedure of LHS converges more quickly than an equidistribution sampling employed in MCS. Examples of LHS application in water resources engineering are Yeh and Tang (1993) and Chang et al. (1992). The main shortcoming with this stratification scheme is that it is one-dimensional and does not provide good uniformity properties on a k-dimensional unit hypercube (Diwekar and Kalagnanam, 1997). Except reducing computation effort to some extent, LHS has the same problems that are associated with MCS.

Second Order Approximation Method

In the second order approximation (SOA) method, a Taylor series expansion of a model is truncated after the second-order term. Consider a model represented by (1-11), the second order Taylor series expansion of Y is given as

$$Y = g(\underline{X}_e) + \sum_{i=1}^{n} (X_i - X_{ie})\left(\frac{\partial g}{\partial X_i}\right)_{X_e} + \frac{1}{2}\sum_{i=1}^{n} (X_i - X_{ie})^2\left(\frac{\partial^2 g}{\partial X_i^2}\right)_{X_e} \tag{1-18}$$

In SOA, the expansion point is commonly the mean value of the basic variables. Considering that all input variables are statistically independent and taking expectation of (1-18), the expected value Y is given as

$$E[Y] \approx g(\overline{X}) + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial^2 g}{\partial X_i^2}\right) Var(X_i) \tag{1-19}$$

The variance of Y is given as

$$Var(Y) = \sigma_Y^2 \approx \sum_{i=1}^{n}\left(\frac{\partial g}{\partial X_i}\right)^2_{\overline{X}_i} Var(X_i) - \frac{1}{4}\sum_{i=1}^{n}\left(\frac{\partial^2 g}{\partial X_i^2}\right)^2_{\overline{X}_i} Var^2(X_i) + \sum_{i=1}^{n}\left(\frac{\partial g}{\partial X_i}\right)_{\overline{X}_i}\left(\frac{\partial^2 g}{\partial X_i^2}\right)^2_{\overline{X}_i} E\left[(X_i - \overline{X}_i)^3\right]$$
$$+ \frac{1}{4}\sum_{i=1}^{n}\left(\frac{\partial^2 g}{\partial X_i^2}\right)^2_{\overline{X}_i} E\left[(X_i - \overline{X}_i)^4\right] + \sum_{i=1}^{n-1}\sum_{j=1+1}^{n}\left(\frac{\partial^2 g}{\partial X_i^2}\right)\left(\frac{\partial^2 g}{\partial X_j^2}\right) Var(X_i) Var(X_j)$$

(1-20)

Bates and Townley (1988) and Tung and Hathhorn (1989) used SOA only for evaluating the mean of the model output. They preferred FOA to estimate variance of the model output due to involvement of complicated calculations in approximating the model output variance based on SOA.

<u>First Order Reliability Method</u>

The first order reliability (FORM) method is characterized by the iterative, linear approximation to the performance function. Fundamentally, this method can be considered as an extension to the FOA method and is also known as advanced first order approximation (AFOA) method, which was developed to address technical difficulties of FOA. One of the major problems with the FOA technique was the lack of invariance of the solution relative to the formulation of the performance function. Simple algebraic changes in the problem formulation can lead to significant changes in assessing the propagation of uncertainty. Hasofer and Lind (1974) presented a methodology, which specifically addressed this issue by requiring expansion about a unique point in the feasible solution space. It should be mentioned that Fruedenthal (1956) also proposed a method suggesting similar restrictions on the expansion point.

Hasofer and Lind (1974) proposed taking the Taylor series expansion at a likely point on the failure surface of the performance function. Rackwitz (1976) implemented the ideas of Hasofer and Lind. The failure surface is defined by the equation $Z = 0$. The perpendicular drawn on the failure surface from the origin cuts the failure surface at a point called the failure point. The distance of the failure point from the origin is a measure of reliability. The expected value and variance of Z can be obtained by first solving $Z = 0$ to find the failure point $\underline{X}^*$ and then expanding Z about $\underline{X}^*$ using a Taylor series expansion as

$$E[Z] \approx \sum_{i=1}^{n} \left(\frac{\partial Z}{\partial X_i} \right)_{X_i^*} \left(\overline{X}_i - X_i^* \right) \tag{1-21}$$

$$Var(Z) = \sigma_Z^2 \approx \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial Z}{\partial X_i} \right)_{X_i^*} \left(\frac{\partial Z}{\partial X_j} \right)_{X_j^*} E\left[\left(X_i - X_i^* \right)\left(X_i - X_j^* \right) \right] \tag{1-22}$$

where σ_Z is the standard deviation of Z. For the case of statistically independent basic variables *Var(Z)* is rewritten as

$$Var(Z) = \sigma_Z^2 \approx \sum_{i=1}^{n} \left(\frac{\partial Z}{\partial X_i} \right)_{X_i^*}^2 \sigma_{X_i}^2 \tag{1-23}$$

Now, (1-2), (1-6), and (1-16) can be used to determine P_f, $\Re$, and β respectively. For models having a linear failure surface and all the basic variables normally distributed, the estimates of P_f, $\Re$ are exact.

For most modeling problems, it is very unlikely that all basic input variables will be normally distributed. Rackwitz (1976) proposed a transformation technique in which the values of the CDF and PDF of the non-normal distributions are the same as those of the equivalent normal distributions at the failure point $\underline{X}^*$. Consider an input random variable X_i for which PDF and CDF are given as $p_{Xi}(x_i)$ and $P_{Xi}(x_i)$ respectively. Equating the cumulative probabilities at the failure point

$$\Phi\left(\frac{x_i^* - \mu_{X_i}^N}{\sigma_{X_i}^N}\right) = P_{X_i}(x_i^*) \tag{1-24}$$

where $\mu_{X_i}^N$ and $\sigma_{X_i}^N$ are the mean value and standard deviation of the equivalent normal distribution for X_i; $P_X(x_i^*)$ is the original CDF of X_i; and $\Phi(.)$ is the CDF of the standard normal distribution. Using (1-24), the mean of the equivalent normal distribution can be written as

$$\mu_{X_i}^N = x_i^* - \sigma_{X_i}^N \Phi^{-1}\left[P_{X_i}(x_i^*)\right] \tag{1-25}$$

Now equating the corresponding PDF ordinates at x_i^*

$$\frac{1}{\sigma_{X_i}^N}\phi\left(\frac{x_i^* - \mu_{X_i}^N}{\sigma_{X_i}^N}\right) = p_{Xi}(x_i^*) \tag{1-26}$$

where $\phi(.)$ is the PDF of the standard normal distribution. Based on (1-26), the standard deviation of the equivalent normal distribution can be written as

$$\sigma_{X_i}^N = \frac{\phi\{\Phi^{-1}[P_{X_i}(x_i^*)]\}}{p_{X_i}(x_i^*)} \tag{1-27}$$

The key to FORM is the determination of the failure point for the Taylor series expansion. Shinozuka (1983) has shown that for FORM the reliability index, β, is the shortest distance in the standardized space between the system mean state and the failure surface. Thus, if the failure point is determined correctly, it represents the most likely combination of input variable values which produce the critical target level. The determination of β requires application of a constrained nonlinear optimization such as the generalized reduced-gradient algorithm used by Cheng (1982), a Lagrange multiplier approach used by Shinozuka (1983), and an iterative optimization method suggested by Rackwitz (1976).

FORM has been used quite successfully in a wide variety of fields for reliability and risk analyses. For example, Melching et al. (1990), Melching et al. (1991), Cesare (1991), and Melching and Anmangandla (1992) used it in hydrologic and hydraulic design; Sitar et al., (1987), Cawlfield and Wu, (1993), Hamed et al. (1995, 1996a, 1996b) applied it to ground water contamination modeling; Hamed and Bedient (1997), Hamed (1997, 1999), Mishra (1998) used it in probabilistic human health risk assessments. FORM can also be used to carry out uncertainty analysis by repeating the procedure of calculating the linearization point to match the pre-specified output value whose exceedance probability is sought. Some examples of using FORM in water quality uncertainty analyses are Melching and Anmangandla (1992).

FORM is quite accurate because it is able to overcome model non-linearity problems, and no additional assumption about the distribution type of the performance function is required. It is still an approximation method because the performance function is approximated by a linear function at the design point, and accuracy problems may arise when the performance function is strongly nonlinear (Cawlfield and Wu, 1993; Zhao and Ono, 1999). Another disadvantage of the FORM is that determination of the lineraziation point is generally not easy, depending upon the nature and complexity of the system for which the reliability, risk, or uncertainty analysis is being studied (Melching and Anmangandla 1992). Further, the magnitude of acceptable convergence may affect the accuracy of the reliability estimates. In some cases, the magnitude of the convergence error may not be reduced after a certain level.

Second Order Reliability Methods

The second order reliability method (SORM) has been used extensively in structural reliability analyses. It has been established as an attempt to improve the accuracy of FORM. SORM is obtained by approximating the limit state surface function at the design point by a second order surface, and the failure probability is given as the probability content outside the second order surface. There are two kinds of second order reliability approximations: curvature-fitting SORM (Breitung 1984; Tvedt 1983, 1988, 1990) and point-fitting SORM (Kiureghian et al. 1987, 1991; Zhao and Ono 1999). Both methods involve complex numerical algorithms and extensive computational efforts.

Hamed et al. (1995) and Hamed (1997) compared risk assessments due to groundwater contamination based on FORM and SORM and reported that their results were in good agreement when the limit-state surface at the design point in the standard normal space is nearly flat. On the other hand, when the limit state function contains highly nonlinear terms, or when the input random variables have an accentuated non-normal character, SORM tends to produce more accurate results than FORM. But computational requirements of SORM are much higher than FORM.

Point Estimation Methods

The point estimation (PE) method was originally proposed by Rosenblueth (1975) to deal with symmetric, correlated, stochastic input parameters. The method was later extended to the case involving asymmetric random variables (Rosenblueth, 1981). The idea is to approximate the given PDF of an input random variable by discrete probability masses concentrated at two points in such a way that its first three moments are preserved.

Consider the model represented by (1-11) having *n* stochastic input parameters. Rosenblueth (1975, 1981) demonstrated that the r^{th}-order moment of output random variable *Y* about the origin could be approximated via a point-probability estimate of the first-order Taylor series expansion. This method requires 2^n model evaluations to estimate a single statistical moment of the model output. For a large model with a large number of parameters, Rosenblueth's PE method is computationally impractical. Further, a reliability analysis requires knowledge of higher order moments in order to approximate the distribution of the output random variable. This makes the method even more computationally extensive. Thus, while Rosenblueth's method is quite efficient for problems with a small number of uncertain basic variables, its computational requirements are similar to those of MCS for a model having a large number of parameters. For example, a model having between 10 and 15 parameters will require 1024 to 32768 model evaluations (Melching, 1995). Examples of applying the

Rosenblueth's method to watershed hydrology include Rogers et al. (1985), Binely et al. (1991), and Melching (1992b).

Harr (1989) modified the Rosenblueth's method to reduce its computational requirements from 2^n to 2n for an n-parameter model by using the first two moments of the random variables. This method does not provide the flexibility to incorporate known higher order moments of input random variables. Chang et al. (1995) showed that the estimated uncertainty feature of model output could be inaccurate if the skewness of a random variable is not accounted for. Yeh and Tung (1993) and Chang et al. (1992) are some of the examples of applying Harr's method in hydraulic engineering.

Transform Methods

Tung (1990) used the Mellin transform to calculate the higher-order moments of a model output. The application of the Mellin transform is not only cumbersome, but also it can not be universally applied. As pointed out by Tung, the Mellin transform may not be analytic under certain combinations of distribution and functional forms. In particular, problems may arise when a functional relationship consists of input variable(s) with negative exponent(s). When component functions of a given model have other forms than power functions, it can not be applied. Further, no formulation was suggested to obtain the moments of a model output having non-standard normally distributed input variable(s).

Review Summary

Based on literature survey it can be said that FOA and MCS are the two most commonly used methods employed for uncertainty analyses of water resources and environmental engineering systems. Both methods have some limitations. The MCS is computationally intensive with the number of simulations required for convergence not well defined (Melching, 1995). In most engineering problems, the true probability distributions of the input variables are seldom known. Theoretical distributions for the input variables are assumed to conduct the MCS. FOA is very computationally efficient but provides approximate model output estimates for the mean and variance only. The quality of these estimates is influenced by the coefficient of variation of input variables and non-linearity in the model (Burn and McBean, 1985; Tung, 1990). Many researchers (Burges, 1979; Dettinger and Wilson, 1981) concluded that FOA should be applied in cases where nonlinearity effects are not significant and uncertainties in input variables are not too large.

In many studies MCS estimates have been used to check the accuracy of FOA estimates. However, in reality, the MCS method is also an approximate method (Bates and Townley, 1988), the quality of which is affected by appropriateness of the chosen distribution function(s) for the input variable(s) and the number of simulations employed in the analysis. The inference drawn from the comparative analyses of Burges (1979), Garen and Burges (1981), Walker (1982), and Malone et al. (1984) indicates that for all practical purposes both methods produced identical results. Scavia et al. (1981b), Gardner et al. (1981), Gardner and O'Neill (1983), Smith and Charbeneau (1990) however, revealed contrary results. These studies suggest both significant and subtle differences in variance estimates from the two approaches. Thus, it appears that despite the fact that FOA is one of the few relatively tractable techniques available to evaluate the effect of parameter uncertainty, doubts about its validity have limited its wide application. There is a state of ambiguity whether FOA should be used for a given problem because no clear guidance about suitability of FOA is available (Zhang and Haan, 1996).

On the other hand, in reliability and risk analyses of water resources and environmental engineering system, three methods namely FOA, MCS, and FORM are used most frequently. Often, failures of engineering systems occur at extreme values rather than near the mean values of the input variables. Extremes are most likely associated with probability distributions having large variances and skewnesses (Yen et al., 1986). Since FOA uses expansion about the mean values of the input variables, attempts to characterize the tails of the output distribution are likely to result in inexact estimates (Burn and McBean, 1985). In addition to the problems due to nonlinearity in the functional form, FOA has some additional problems when employed for risk and/or reliability analysis of engineering systems. FOA does not provide the form of the distribution for the performance function required to carry out the risk/reliability analysis. A normal distribution is generally assumed when confidence limits on the output, risk, and reliability of the system are determined. Furthermore, using FOA, it is not possible to incorporate information about the forms of input variable distributions, even if they are known (Yen et al., 1986).

Using FORM the flaws of FOA due to model non-linearity can be removed by linearizing the functional relationship at the point on the limit-state surface nearest to the origin, rather than at the mean point. Calculation of the linearization point requires determination of the nearest point on the limit-state surface. FORM is quite accurate when the performance function is not strongly non-linear. The disadvantage of the FORM method is that it is quite complicated. This is because it requires transformation of non-normal distributions and determination of the failure point using a non-linear constrained optimization, which is generally not easy depending upon the nature and complexity of the system. Further, the magnitude of convergence error may affect the accuracy of the reliability estimates and in some case it may not be possible to reduce the convergence error below a certain level.

SORM and PE methods have not been used much in the area of water resources and environmental engineering. Hamed et al. (1995), Hamed (1997) observed that reliability estimates based on FORM and SORM were in good agreement when the limit-state surface at the design point in the standard normal space is nearly flat. In cases where a performance function is strongly non-linear, SORM reliability estimates are better than that of the FORM. But computational requirements and a complicated calculation process make it typically unsuitable for practicing engineers. As far as PE methods are concerned, both the methods give approximate statistical moments of a model output. While, Rosenblueth's method preserves the first three moments of the original PDF, Harr's method is able to preserve only the first two moments. In many problems where input variables do not have zero skewness and coefficient of kurtosis equal to 3, it is obvious that both methods will give inaccurate moments of the performance function and therefore inaccurate uncertainty and reliability estimates. The other drawback of PE methods is that they do not provide the distribution type for the performance function.

Objectives

The first objective of this thesis was to determine the impacts of model nonlineraity, magnitude of input parameter uncertainty, and distribution form of the input parameters on the uncertainty, risk, and reliability analysis of hydrologic, hydraulic, and environmental engineering systems. To be useful for practicing engineers, research emphasis was focussed towards development of simple, accurate, and generic methods.

Computation efficiency, understandability, and easier application have made the first order approximation (FOA) method a favored tool for uncertainty analysis. Due to several theoretical and/or conceptual drawbacks, in many instances specific situations may arise where the accuracy of FOA estimates becomes questionable. The second objective of this thesis was to develop a correction procedure to correct the FOA estimates for model nonlinearity, parameter uncertainty, and parameter distribution types. The developed method could be used to judge the suitability of FOA in ambiguous situations as well as to determine the exact values of mean and variance of a model output.

Literature review indicates that in many cases the true form of the output distribution is not required. A very good estimate of system reliability can be obtained if higher-order moments of model output are known correctly. Therefore, the third objective of this work was to develop a simple and generalized technique to determine higher-order moments of a model output as a function of the means, the CVs, and the distribution types for input random variables.

References

Ang, A. H. S., and Tang, W. H. (1984). Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk, and Reliability, John Wiley & Sons, Inc., New York, N.Y.

Argentesi, F. and Olivi, L. (1976). Statistical sensitivity analysis of a simulation model for the biomass-nutrient dynamics in aquatic ecosystems, Proceedings, 4th Summer Computer Simulation Conference, Simulation Counc., La Jolla, Calif., 389-393.

Batchelor, B., Valdes, J., and Araganth (1998). Stochastic risk assessment of sites contaminated by hazardous wastes, Journal of Environmental Engineering, ASCE, 124(4), 380-388.

Bates, B. C. (1988). Nonlinear, discrete flood event models, 2. Assessment of statistical nonlinearity, Journal of Hydrology, 99, 77-89.

Bates, B. C., and Townley, L.R. (1988). Nonlinear discrete flood event models, 3. Analysis of prediction uncertainty, Journal of Hydrology, 99, 91-101.

Bates, D. M., and Watts, D.G. (1980). Relative curvature measures of nonlinearity. Journal of the Royal Statistical Society, B, 42(1), 1-25.

Beale, E. M. L. (1960). Confidence regions in nonlinear estimation, Journal of the Royal Statistical Society, B, 22(1), 41-88.

Beck, M. B. (1981a). Operational estimation and prediction of nitrification dynamics in an activated sludge process, Water Research, 15, 1313-1330.

Beck, M. B. (1981b). Hard or soft environmental systems? Ecological Modeling, 11, 237-251.

Beck, M. B. (1987). Water quality modeling: A review of the analysis of uncertainty, Water Resources Research, 23(5), 1393-1441.

Benjamin, J. R., and Cornell, C. A. (1970). Probability, Statistics, and Decisions for Civil Engineers, McGraw-Hill, New York, N.Y.

Beven, K. J. and Binley, A. M. (1992). The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279-298.

Binley, A. M., Beven, K. J., Calver, A., and Watts, L. G. (1991). Changing responses in hydrology: Assessing the uncertainty in physical based model predictions, Water Resources Research, 27(6) 1253-1261.

Black, T. C. and Freyburg, D. L. (1987). Stochastic modeling of vertically averaged concentration uncertainty in a perfectly stratified aquifer, Water Resources Research, 23, 997-1004.

Breitung, K. (1984). Asymptotic approximation for multi-normal integrals. Journal of Engineering Mechanics, ASCE, 110(3), 357-366.

Burges, S. J. (1979). Analysis of uncertainty in flood plain mapping. Water Resources Bulletin, 15(1), 227-243.

Burges, S. J., and Lettenmaier, D. P. (1975). Probability methods in stream quality management, Water Resources Bulletin, 11(1), 115 –130.

Burn, D. H. and McBean, E. A. (1985). Optimization modeling of water quality in an uncertain environment, Water Resources Research, 21(7), 934-940.

Cawlfield, J. D., and Wu, M. C. (1993). Probabilistic sensitivity analysis for one dimentional reactive transport in porous media, Water Resources Research, 29(3), 661-672.

Cesare, M. A. (1991). First-order analysis of open-channel flow, Journal of Hydraulic Engineering, ASCE, 117(2), 242-247.

Chadderton, R. A., Miller, A. C., and McDonnell, A. J. (1982). Uncertainty analysis of dissolved oxygen model, Journal of the Environmental Engineering Division, ASCE, 108(5), 1003-1013.

Chang, C. H., Tung, Y. K., and Yang, J. C. (1995). Evaluation of probabilistic point estimate methods, Applied Mathematical Modeling, 19(2), 95-105.

Chang, J. H., Tung, Y. K., Yang, J. C., and Yeh, K. C. (1992). Uncertainty analysis of a computerized sediment transport model, Stochastic Hydraulics' 92, Proceedings Sixth International Association for Hydraulic Research Symposium on Stochastic Hydrology, Tapai, Taiwan, J. T. Kuo and G. F. Lin, eds., Water Resources Publications, Littleton, Colorado, 123-130.

Cheng, S. T. (1982). Overtopping Risk Evaluation for an Existing Dam, Ph.D. Thesis, Department of Civil Engineering, University of Illinois at Urbana-Champaign.

Clifton, P. M. and Neuman, S. P. (1982). Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in southern Arizona, Water Resources Research, 18(4), 1215-1213.

Cornell, C. A. (1972). First-order analysis of model and parameter uncertainty, Proc. Int. Symposium on Uncertainties in Hydrologic and Water Resources, Systems, Univ. of Arizona, vol. III, 1245-1276.

Dettinger, M. D., and Wilson, J. L. (1981). First order analysis of uncertainty in numerical models of groundwater flow, Part 1, Mathematical development, Water Resources Research 17(1), 149-161.

Devary, J. L., and Doctor, P. G. (1982). Pore velocity estimation uncertainties, Water Resources Research, 18(4), 1157-1164.

Diwekar, U. M., and Kalagnanam, J. R. (1997). Efficient sampling technique for optimization under uncertainty, AIChE Journal, 43(2), 440-447.

Fedra, K., van Starten, G., and Beck, M. B. (1981). Uncertainty and arbitrariness in ecosystems modeling: A lake modeling example, Ecological Modeling, 13, 87-110.

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.

Fruedenthal, A. M. (1956). Safety and probability of structural failure, Transactions of the ASCE, 121, 1337-1397.

Gardner R. H., Huff, D. D., O'Neill, R. V., Mankin, J. B., Carney, J. H., and Jones, J. (1980). Application of error analysis to a marsh hydrology model, Water Resources Research, 16(4), 659-664.

Gardner, R. H. and O'Neill, R. V. (1983). Parameter uncertainty and model predictions: A review of Monte Carlo results, Uncertainty and forecasting of water quality, M. B. Beck and G. van Straten, eds., Springer Verlag, New York, 245-257.

Gardner, R. H., Mankin, J. B., and Shugart, H. H., Jr. (1976). The COMEX computer code documentation and description, Oak Ridge National Laboratory, Oak Ridge, Tenn., EDFB/IBP-76/4.

Gardner, R. H., O'Neill, R. V., Carney, J. H., and Mankin, J. B. (1981). A comparison of sensitivity analysis and error analysis based on a stream ecosystem model, Ecological Modeling, 12, 173-190.

Garen, D. C. and Burges, S. J., (1981). Approximate error bounds for simulated hydrographs, Journal of Hydraulic Engineering Division, ASCE, 107(11), 1519-1534.

Goldman, D. M., Marino, M. A., and Feldman, A. D. (1990). Runoff prediction uncertainty for ungaged watersheds, Journal of Irrigation and Drainage Engineering., ASCE, 116(6), 752-768.

Hamed, M. M. (1997). First-order reliability analysis of public health risk assessment, Risk Analysis, 17(2), 177-185.

Hamed, M. M. (1999). Probabilistic sensitivity analysis of public health risk assessment from contaminated soil, Journal of Soil Contamination, 8(3), 285-306.

Hamed, M. M., Bedient, P. B, and Dawson, C. N. (1996a). Probabilistic modeling of aquifer heterogeneity using reliability methods, Advances in Water Resources, 19(5), 277-295.

Hamed, M. M., Bedient, P. B., and Conte, J. P.(1996b). Numerical stochastic analysis of groundwater contaminant transport and plume containment, Journal of Contaminant Hydrology, 24(1), 1-24.

Hamed, M. M., Conte, J. P., and Bedient, P. B. (1995). Probabilistic screening tool for groundwater contamination assessment, Journal of Environmental Engineering., ASCE, 121(11), 767-775.

Hamed, M. M. and Bedient, P. B. (1997). On the performance of computational methods for the assessment of risk from groundwater contamination, Ground Water, 35(4), 638-646.

Harlin, J. and Kung, C. S. (1992). Parameter uncertainty and simulation of design flood in Sweden, Journal of Hydrology, 137, 209-230.

Harr, M. E. (1989). Probability estimates for multivariate analyses. Applied Mathematical Modeling, 13, 313-318.

Hasofer, A. M., and Lind, N. C. (1974). Exact and invariant second-moment code format. Journal of Engineering Mechanics Division, ASCE, 100, 111-121.

Hornberger, G. M. (1980). Uncertainty in dissolved oxygen prediction due to variability in algal photosynthesis, Water Research, 14, 355-361.

Johnson, P. A. (1996). Uncertainty of hydraulic parameters, Journal of Hydraulic Engineering, ASCE, 122(2), 112-114.

Jones, L. (1989). Some results comparing Monte Carlo simulation and first-order Taylor series approximation for steady groundwater flow, Journal of Stochastic Hydrological Hydraulics, 3(3). 179-190.

Kiureghian, D. A., Lin, H. Z., and Hwang, S. J. (1987). Second-order reliability analysis. Structural safety, 14, 267-276.

Kiureghian, D. A., Stefano, D. M. (1991). Efficient algorithm for second-order reliability analysis, Journal of Engineering Mechanics, ASCE, 117(12), 1208-1225.

Krajewski, W. F., Lakshmi, V., Georgakakos, K. P., and Jain, S. C. (1991). A Monte Carlo study of rainfall sampling effects on distributed catchment model, Water Resour. Res., 27(1), 119-128.

Ku, H. H. (1966). Notes on the use of propagation of error formulas, Journal of Research of the National Bureau of Standards, Section C, Engineering and Instrumentation, 70C(4), 263-273.

Kuczera, G. (1988). On the validity of first-order prediction limits for conceptual hydrologic models, Journal of Hydrology, 103, 229-247.

Lee, H. L. and Mays, L. W. (1986). Hydraulic uncertainty in flood levee capacity. Journal of the Hydraulic Engineering Division, ASCE, 112(10), 928-934.

Lei, J. and Schilling, W. (1993). Propagation of model uncertainty, Proceedings Sixth International Conference on Urban Storm Drainage, Niagara Falls, Canada, eds. J. Marsalek and H., Torno, Seapoint Publishing, Victoria, British Columbia, Canada, 465-470.

Lettennmaier, D. P. and Richey, J. E. (1979). Use of first order analysis in estimating mass balance errors and planning sampling activities, Theoretical systems ecology: Advances and case studies, eds E. Halfon, Academic Press, Orlando, 80-109.

Loague, K. (1991). The impact of land use on estimates of pesticide leaching potential: Assessments and uncertainties, Journal of Contaminant Hydrology, 8,157-175.

Loague, K. M. and Green, R. E. (1988). Impact of data-related uncertainties in a pesticide leaching assessment, Methods for Groundwater studies - Proceedings of a National Workshop, Agricultural Research Service, University of Nebraska-Lincon, NE., 99-119.

Loague, K. M., Green, R. E., Giambelluca, T. W., Liang, T.C., and Yost, R. S. (1990). Impact of uncertainty in soil, climate, and chemical information in a pesticide leaching assessment, Journal of Contaminant Hydrology, *5*, 171-194.

Loague, K. M., Yost, R. S., Green, R. E., and Liang, T. C. (1989). Uncertainty in a pesticide leaching assessment for Hawaii, Journal of Contaminant Hydrology, 4,139-161.

Malone, R. F., Saidi, H., and Wegener, K. (1984). Predictive accuracy determination applied to a linear model of phosphorus loading resulting from urban runoff, Applied Math. Modeling, 8, 81-88.

Martz, H. F. (1983). A comparison of methods for uncertainty analysis of nuclear plant safety system fault tree models, U.S. Nuclear regulatory Commission and Los Alamos National Laboratory, NUREG/CR-3263, Los Alamos, NM.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21(2), 239-245.

Melching, C. S. (1992a). An improved first order reliability approach for assessing uncertainties in hydrologic modeling, Journal of Hydrology, 132, 157-177.

Melching, C. S. (1992b). A comparison of models for estimating variance of water resources model predictions, Proceedings sixth international association for hydraulic research symposium on stochastic hydraulics, *Taipei, Taiwan,* J. T. Kuo and G. F. Lin, eds., Water Resources Publications, Littleton, Colorado, 663-670.

Melching, C. S. (1995). Reliability estimation. Computer models of watershed hydrology, V. P. Singh, ed., Water Resources Publication, Littleton, Colorado, 69 –118.

Melching, C. S. and Yen, B. C., and Wenzel, H. G. Jr. (1991). Output reliability as guide for selection of rainfall-runoff models, Journal of Water Resources Planning and Management, ASCE, 117(3), 383-398.

Melching, C. S. and Yoon, C. G. (1996). Key sources of uncertainty in QUAL2E model of Passaic River. J. Water Resources Planning and Management, ASCE, 122(2), 105-114.

Melching, C. S., and Anmangandla, S. (1992). Improved first order uncertainty method for water quality modeling, Journal of Environmental Engineering, ASCE, 118(5), 791-805.

Melching, C. S., Yen, B. C., and Wenzel, H. G., Jr. (1990). A reliability estimation in modeling watershed runoff with uncertainties, Water Resources Research, 26(10), 2275-2286.

Mishra, S. (1998). A methodology for probabilistic risk-based performance assessments using the first-order reliability method, Proceeding Risk-Based Performance Assessment and Decision-Making Topical Meeting, *American Nuclear Society*, Pasco, WA, 1-6.

Modarres, M. (1993). Reliability and risk analysis, Marcel Dekker, Inc., New York.

Montgomery, R. H., Lee, U. D., and Reckhow, K. H.(1980). A comparison of uncertainty analysis techniques: First-order analysis vs. Monte Carlo simulation, IAGLR, Great Lakes, Mich.

Moore, R. J., and Clarke, R. T. (1981). A distribution function approach to rainfall runoff modeling, Water Resources Research, 17(5), 1367-1382.

O'Neill, R. V. (1971). Error analysis of ecological models, Proceedings of the third national symposium in radioecology, U.S. Atomic Energy Commission, Oak Ridge Laboratory, Oak Ridge, Tenn., CONF-710501.

O'Neill, R. V., Gardner, R. H., and Mankin, J. B. (1980). Analysis of parameter error in a non-linear model, Ecological Modeling, 8, 297-311.

Rackwitz, R. (1976). Practical probabilistic approach to design, Comite European du Beton, Paris, Bulletin No. 112.

Reckhow, K. H. (1979a). The use of a simple model and uncertainty analysis in lake management, Water Resources Bulletin, 15, 601-611.

Reckhow, K. H. (1979b). Uncertainty analysis applied to Vollenweider's phosphorus loading criterion, Journal of Water Pollution Control Federation, 51(8), 2123-2128.

Rogers, C. C. M., Beven, K. J., Morris, E. M., and Andreson, M. G. (1985). Sensitivity analysis, calibration and predictive uncertainty of the Institute of Hydrology Distributed Model, Journal of Hydrology, 81, 179-191.

Rosenblueth, E. (1975). Point estimates for probability moments, Proc. National Academy of Sciences U.S.A., 72(10), 3812-3814.

Rosenblueth, E. (1981). Point estimates in probabilities, Applied Mathematical Modeling, 72(10), 3812-3814.

Scavia, D. (1980). Uncertainty Analysis of a Lake Eutrophication Model, Ph.D. dissertation, Univ. of Mich., Ann Arbor.

Scavia, D., Canale, R. P., Powers, W. F., and Moody, J. L. (1981b). Variance estimates for a dynamic eutrophication model of Saginaw Bay, Lake Huron, Water Resources Research, 17(4), 1115-1124.

Scavia, D., Powers, W. F., Canale, R. P., and Moody, J. L. (1981a). Comparison of first-order error analysis and Monte Carlo simulation in time-dependent lake eutrophication models, Water Resources Research, 17(4), 1051-1059.

Shinozuka, M. (1983). Basic analysis of structural safety. Journal of Structural Engineering, ASCE, 109(3), 721-740.

Siddal, J. N. (1983). Probabilistic analysis. Probabilistic Engineering Design, Marcel Dekker, Inc., New York, N.Y., 145-236.

Sitar, N., Cawlfield, J. D., and Kiureghian, A. D. (1987). First-order reliability approach to stochastic analysis of subsurface flow and contaminant transport, Water Resources Research, 23(5), 794-804.

Smith, L. and Freeze, R. A. (1979). Stochastic analysis of steady state groundwater flow in a bounded domain, 1, One-dimensional simulations, Water Resources Research, 15(3), 521-528.

Smith, L. and Schwartz, W. F. (1980). Mass transport Part 1, Stochastic analysis of macrodispersion, Water Resources Research, 16 (2), 303-313.

Smith, L. and Schwartz, W. F. (1981). Mass transport Part 2, Analysis of uncertainty in prediction. Water Resources Research, 17, 351-369.

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.

Smith, V. J., and Charbeneau, R. J. (1990). Probabilistic soil contamination exposure assessment procedures, Journal of Environmental Engineering, ASCE, 116(6), pp.1143-1163.

Stevens, E. (1993). A Method to Predict the Accuracy of a First Order Approximation of Model Output Variance, Ph. D. Thesis, Department of Biosystems and Agricultural Engineering, Oklahoma State University, Stillwater, OK.

Takasao, T. and Takara, K. (1989). Evaluation of rainfall-runoff models from the stochastic viewpoint, Journal of Hydrology, 102, 381-406.

Tang, W. H., and Yen, B. C. (1972). Hydrologic and hydraulic design under uncertainties. Proceedings International Symposium on Uncertainties in Hydrology and Water Resources Systems, University of Arizona, Tucson, Arizona, 2, 868-882.

Tiwari, J. L. and Hobbie, J. E. (1976). Random differential equations as models of ecosystems: Monte Carlo simulation approach, Math. Bioscience, 28, 25-44.

Townley, L. (1984). Second-order effects of uncertain transmissivities on predictions of piezometric heads, Finite Element in Water Resources, Proceedings of the Fifth International Conference, Springer-Verlag, New York.

Townley. L. R., and Wilson, J. L (1985). Computationally efficient algorithm for parameter estimation and uncertainty propagation in numerical models of groundwater flow, Water Resources Research, 21(12), 1851-1860.

Tung, Y. K. and Hathhorn, W. E. (1989). Assessment of probability distribution of dissolved oxygen deficit, Journal of Environmental Engineering, ASCE, 114(6), 1421-1435.

Tung, Y. K., and Mays, L.W.(1981). Risk Models for Flood Levee Design, Water Resources Research, 17(4), 833-841.

Tung, Y. K. (1990). Mellin transformation applied to uncertainty analysis in hydrology / hydraulics, Journal of Hydraulic Engineering, ASCE, 116(5), 659-674.

Tung, Y. K., and Mays, L. W. (1980). Optimal risk-based design of water resources engineering projects, Technical Report No. 171, Center for Research in Water Resources, Univ. of Texas, Austin, Austin, Tex.

Tvedt, L. (1983). Two second-order approximation to the failure probability, Veritas Rep. RDIV/20-004083, Det norske Veritas, Oslo.

Tvedt, L. (1988). Second order reliability by an exact integral. Proceedings 2nd IFIP Working Conference on Reliability and Optimization on Structural Systems, Eds. P, Thoft-Chistensen, Springer, New York, 377-384.

Tvedt, L. (1990). Distribution of quadratic forms in normal space - application to structural reliability, Journal of Engineering Mechanics, ASCE, 119(6), 1183-1197.

Van de Kramer, S. (1983). Monte Carlo simulation and first-order error analysis: Two possible methods to cope with uncertainties in water quality modeling, Working paper No. WP-83-9, International Institute for Applied Systems Analysis, Vienna, Austria.

Van Straten, G. (1983). Maximum likelihood estimation of parameters and uncertainty in phytoplankton models, Uncertainty Analysis and Forecasting of Water Quality, Eds. M. B. Beck and G. van Straten, 157-171, Springer Verlag, New York.

Walker, W. W. (1982). A sensivity analysis and error analysis framework for lake eutrophication modeling, Water Resources Bulletin, 18,53-60.

Warwick, J. J. and Wilson, J. S. (1990). Estimating uncertainty in stormwater runoff computations, Journal of Water Resources Planning and Management, ASCE, 116(2), 187-204.

Yeh, K. C., and Tung, Y. K. (1993). Uncertainty and sensitivity analyses of pit-migration model, Journal of Hydraulic Engineering, ASCE, 119 (2), 262-283.

Yen, B. C. and Tang, W. H. (1976). Risk-safety factor relation for storm sewer design, Journal of Environmental Engineering Division, ASCE, 102 (2), 509-516.

Yen, B. C., Cheng, S. T., and Melching, C. S. (1986) First order reliability analysis. Stochastic and risk analyses in hydraulic engineering, B. C. Yen eds., Water resources publications, Littleton, Colorado, 79-89.

Yen, B. C., Cheng, S. T., and Tang, W. H. (1980). Reliability of hydraulic design of culverts, Proceedings international conference on water resources development, International Association for Hydraulic Research Asian Pacific Division Second Congress, 2, Taipei, Taiwan, 991-1001.

Zhang, J., and Haan, C. T. (1996). Evaluation of uncertainty in estimated flow and phosphorus loads by FHANTM, Trans. Soil and Water Division, ASAE, 12(6), 663-669.

Zhao, Y. G. and Ono, T. (1999). New approximation for SORM: Part 2. Journal of Engineering Mechanics, ASCE, 125(1), 86-93.

CHAPTER II

UNCERTAINTY ANALYSIS USING CORRECTED FIRST ORDER APPROXIMATION METHOD

Abstract

Application of uncertainty and reliability analysis is an essential part of many problems related to modeling and decision making in the area of environmental engineering and water resources. Computation efficiency, understandability, and easier application have made the first order approximation (FOA) method a favored tool for uncertainty analysis. In many instances doubtful situations may arise where the accuracy of FOA estimates becomes questionable. Often FOA application is considered acceptable if the coefficient of variation (CV) of the uncertain parameter(s) is less than 0.2. This criterion is not correct in all the situations. Analytical as well as graphical relationships for relative error are developed and presented for a generic power function which can be used as a guide for judging the suitability of FOA for a specified acceptable error of estimation. Further, these analytical and graphical relationships enable one to correct FOA estimates for means and variances of model components to their true values. Using these corrected values of means and variances for model components one can determine the exact values of mean and variance of a model output. This technique is applicable when an output random variable is a function of several independent random variables in

multiplicative, additive, or in combined (combination of multiplicative and additive) forms. Three examples are given to demonstrate the application of the technique.

Introduction

Reliability modeling and other probabilistic techniques are becoming increasingly important tools in modeling water resources systems and decision making. Many problems in environmental engineering are best approached using probabilistic and reliability methods. Examples are determining the probability of failure of a structure or the life expectancy of a hydraulic structure under uncertainty. The prediction and evaluation of pollution of surface and/or subsurface environments and decisions regarding remedial actions often rely on probabilistic approaches. Quantification of the underlying uncertainty is central to each of these problems.

Uncertainty is present in the design of hydrologic, hydraulic, and environmental projects because of inherent variation, a lack of understanding of all the causes and effects in various processes (physical, chemical, and biological) occurring in the system, and insufficient data. As a result of these uncertainties, the performance reliability of a project may be severely affected. A reliable assessment of the performance of any water resources project requires an assessment of the validity of predicted loads (such as discharges and pollutant loads) and capacities (ability to perform under a given load without any harm). Typically the loads are assessed using various models, generally having a number of parameters which can be determined with varying degrees of accuracy. These parameters are best represented as random variables. Consequently, the model response, being a function of random variables, is also a random variable.

Probabilistic and reliability analyses are based on knowledge of the underlying parameter uncertainties. Parameter uncertainties are the variation in a parameter due to an inability to precisely quantify that parameter. Two methods, Monte Carlo simulation (MCS) and first-order approximation (FOA) are generally used to model the uncertainty associated with input parameters. A MCS requires several thousand repetitive runs of a model and is therefore computationally demanding. With the advent of high-speed computers, the computational challenges to probabilistic analysis have been removed but philosophical and conceptual aspects remain unsolved. MCS has other limitations too. Rules for determining the number of simulations required for convergence are not available (Melching, 1995). The method often entails sample sizes that are in the range of 5,000-20,000 members (Siddall, 1983). Often, the information about distribution function(s) of input variable(s) required to conduct MCS simulations is not available and can not be obtained due to the constraints of time and money. Therefore, it requires judgement on the part of the modeler to create theoretical input sample distributions that are representative of the parameter populations.

FOA is an approximate method that gives only estimates of means and variances. The exactness of the estimates is influenced in part by the degree of nonlinearity in the functional relationship and parameter uncertainty. It has several advantages over MCS. FOA is more computationally efficient and provides a measure of model's sensitivity to each input random variable thus providing a better understanding of the processes being modeled. The assumptions typically cited for FOA to yield good results (Melching, 1995) are: linearity in functional relationships, small coefficients of variation of the most sensitive uncertain variables, and normal distributions for the uncertain variables. Moore and Clarke (1981) expressed that FOA assumptions are "rarely likely to be justifiable with models containing nine or twelve parameters". Despite its several conceptual shortcomings, FOA has been used quite successfully in a wide variety of fields such as hydrologic design (Tang and Yen, 1972; Yen et al., 1980; Tung and Mays, 1980 and 1981), water quality modeling (Burges and Lettenmaier, 1975; Scavia et al., 1981; Chadderton et al.1982; Melching and Anmangandla, 1992), watershed modeling (Garen and Burges, 1981; Melching, 1992a, 1992b; Kuczera, 1988; Bates and Townley, 1988), subsurface flow and contaminant transport modeling (Sagar, 1978; Dettinger and Wilson, 1981; Devary and Doctor, 1982; Townley and Wilson, 1985), and probabilistic human health risk assessment (Batchelor et al., 1998).

In many studies MCS estimates have been used to check the accuracy of FOA estimates. However, in reality, the MCS method is also an approximate method (Bates and Townley, 1988), the quality of which is affected by appropriateness of the chosen distribution function(s) for the input variable(s) and the number of simulations employed in the analysis. The inference drawn from the comparative analyses of Burges (1979), Garen and Burges, (1981), Walker (1982), and Malone et al. (1984) indicates that for all practical purposes both methods produced identical results. Scavia et al. (1981), Gardner et al. (1981), Gardner and O'Neill (1983), Smith and Charbeneau (1990) however, revealed contrary results. These studies suggest both significant and subtle differences in variance estimates from the two approaches. Thus, it appears that despite the fact that FOA is one of the few relatively tractable techniques available to evaluate the effect of parameter uncertainty, doubts about its validity have limited its wide application. There is a state of ambiguity whether FOA should be used for a given problem because no clear guidance about suitability of FOA is available (Zhang and Haan, 1996).

Garen and Burges (1981) found satisfactory results with FOA when the CV of the input parameters was ≤ 0.25. Cornell (1972) and Burges (1979) suggested that FOA is applicable to moderately nonlinear systems when the CV ≤ 0.2. Gardner et al. (1981) found that the validity of the linear approximation deteriorates rapidly when the CV > 0.3. Best agreement between MCS and FOA estimates occurs when MCS output distributions are symmetric (Scavia et al., 1981).

Several researchers detected significant nonlinearity effects while comparing variance estimates from FOA and MCS. To overcome the problem of nonlinearity, several predictors were proposed (Beale, 1960; Bates and Watts, 1980; Bates, 1988; Stevens, 1993). Bates (1988) and Stevens (1993) indicated that the predictors for nonlinearity developed so far work well only in specific applications and that no well-accepted, generalized nonlinearity measure is available.

To date, the only widely used criterion for the validity of FOA variance is to restrict the parameter CV to less than 0.2. This is a very restrictive assumption in water resource systems modeling where there is often a great uncertainty in the parameters (Johnson, 1996). It can be shown that FOA has performed well in some situations when the parameter uncertainty is higher than 0.2. Smith and Charbeneau (1990) suggested that FOA can be used if the difference between function gradients at the mean and one standard deviation away from the mean are less than some acceptable range (5-10%), however, this method also has limitations.

The main objective of uncertainty analysis, in general, is to evaluate the first and second moments of a model output in terms of input random variables. This paper describes a procedure to correct FOA estimates of model components for nonlinearity, CV, and distribution type. Using these corrected estimates of means and variances for model components, one can determine exact values of first and second moments of model output. This procedure provides a deep insight and understanding of the conceptual aspect of uncertainty analysis and hence can be a very useful tool for judging the suitability of FOA in ambiguous situations. The developed procedure demonstrates its application in the uncertainty analysis of problems related to hydrology, hydraulics, and probabilistic human health risk assessment.

Allowable Ranges of Input Parameter's CVs

A consistent measure often used in describing the amount of variation in a population is its CV. In hydrology, hydraulics, and environmental engineering applications most of the quantities of interest are non-negative. Parameter uncertainty represented by the CV can assume a value falling in a specific allowable range depending upon the underlying distribution. Table 2 -1 gives the allowable ranges for some of the commonly used distributions considered in this study.

Functional Forms

In most hydrologic and hydraulic engineering problems, empirically developed or theoretically derived mathematical equations are used which involve several uncertain parameters that have a significant amount of uncertainty and varied distribution characteristics. Further, a mathematical equation may have different degrees of nonlinearity with respect to these uncertain parameters. A multitude of functional forms for $g(\underline{X})$ are possible. In this paper, multiplicative forms, additive forms, and their combined forms are considered.

A multiplicative type model is frequently encountered in hydrological studies (e.g., daily stream flow, peak runoff, annual floods, and annual, monthly, and daily rainfall, soil loss and sediment transport). In hydraulics many equations are of multiplicative type. Examples are flow over control structures such as weirs, spillways, overfalls, and sluices, channel control equations such as Manning's equation (Haan et al., 1994), and pipe flow resistance equations such as Hazen-Williams and Darcy-Weisbach equations (Mays, 1999). In environmental engineering, many equations predicting water quality and pollution (Krenkel, 1979; Novotny and Olem, 1994), and equations used in risk assessment are of multiplicative type (USEPA, 1989). Tung and Mays (1980), Lee and Mays (1986), and Tung (1990) are some examples of the multiplicative forms encountered in hydraulic/hydrologic systems. In this form, the output random variable Y is expressed as the multiplication of n power functions as shown in (2-1).

$$Y = C_0 X_1^{r_1} X_2^{r_2} \ldots\ldots\ldots\ldots X_n^{r_n} = C_0 \prod_{i=1}^{n} X_i^{r_i} \qquad (2\text{-}1)$$

where C_0 and r_i are constants and X_is are independent stochastic input random variables.

Another form of interest is the additive form obtained when two or more power functions are added. It is often encountered in reliability analysis of engineering systems (Hasofer and Lind, 1974; Ang and Tang, 1984; Melching, 1995). In reliability evaluation a performance function (also known as state function) is defined as a combination of demand (loading) and capacity (resistance) of the system where both loading and capacity are random variables. Examples of hydrologic systems include storm sewer design (Yen and Tang, 1976; Tung, 1990), and reliability of a compound channel under extreme events (Cesare, 1991). The general additive form is written as:

$$Y = C_1 X_1^{r_1} + C_2 X_2^{r_2} + \dots\dots\dots + C_n X_n^{r_n} = \sum_{i=1}^{n} C_i X_i^{r_i} \tag{2-2}$$

The other functional form is the combination of multiplicative and additive forms. This form is obtained when two or more multiplicative forms having common power function(s) are added. Examples are application of Manning's equation in a compound channel with same slope for each section (Cesare, 1991; Burges, 1979) and assessment of overall human health risk due to multiple pollutants through different pathways (Batchelor et al., 1998). The general form can be represented as:

$$Y = C_0 X_1^{r_1} X_2^{r_2} \dots\dots X_m^{r_m} \left(C_1 X_{m+1}^{r_{m+1}} + C_2 X_{m+2}^{r_{m+2}} + \dots\dots + C_n X_{m+n}^{r_{m+n}} \right) = C_0 \prod_{i=1}^{m} X_i^{r_i} \sum_{j=1}^{n} C_j X_j^{r_j} \tag{2-3}$$

Approximate Moments Using FOA

Benjamin and Cornell (1970) and Cornell (1972) have provided a detailed description of FOA. Consider a random variable *Y*, which can be expressed as a function of n random independent variables

$$Y = g(\underline{X}) \tag{2-4}$$

where $\underline{X} = (X_1, X_2 \dots\dots\dots\dots, X_n)$, is a vector containing n random independent variables X_i. Through the use of Taylor's expansion and its first order approximation, the mean of the model output can be approximated by

$$\hat{\mu}_Y = g(\overline{\underline{X}}) \tag{2-5}$$

The variance of the model output can be approximated as

$$\hat{\sigma}_Y^2 = \sum_{i=1}^{n} \left[\frac{\partial g(\underline{X})}{\partial X_i} \right]_{\underline{\bar{X}}}^2 \sigma_{X_i}^2 \tag{2-6}$$

where σ_{Xi} = standard deviation of X_i. Using (2-5) and (2-6) and a given mathematical form for Y, the approximate moments of Y can be determined. For the multiplicative form (2-1), the approximate mean of the model output, $\hat{\mu}_Y$, can be written as

$$\hat{\mu}_Y = C_0 \prod_{i=1}^{n} \mu_{X_i}^{r_i} \tag{2-7}$$

where μ_{Xi} = mean of X_i. Using (2-6), the approximate variance of the multiplicative form (2-1), $\hat{\sigma}_Y^2$, can be approximated as

$$\hat{\sigma}_Y^2 = C_0^2 \prod_{i=1}^{n} \mu_{X_i}^{2r_i} \sum_{i=1}^{n} r_i^2 CV_{X_i}^2 \tag{2-8}$$

where r_i is the exponent of ith power function; $CV_{X_i} = \dfrac{\sigma_{X_i}}{\mu_{X_i}}$ = coefficient of variation of X_i. Dividing (2-8) by the square of (2-7), the approximate coefficient of variation of Y, $\hat{C}V_Y$, can be evaluated as

$$\hat{C}V_Y = \left(\frac{C_0^2 \prod_{i=1}^{n} \mu_{X_i}^{2r_i} \sum_{i=1}^{n} r_i^2 CV_{X_i}^2}{C_0^2 \prod_{i=1}^{n} \mu_{X_i}^{2r_i}} \right)^{0.5} = \left(\sum_{i=1}^{n} r_i^2 CV_{X_i}^2 \right)^{0.5} \tag{2-9}$$

When Y is represented by the additive form, the approximate mean of Y is given as

$$\hat{\mu}_Y = \sum_{i=1}^{n} C_i \mu_{X_i}^{r_i} \tag{2-10}$$

Similarly, the variance of the additive model can be approximated by

$$\hat{\sigma}_Y^2 = \sum_{i=1}^{n} C_i^2 r_i^2 \mu_{X_i}^{2r_i} CV_{X_i}^2 \tag{2-11}$$

When Y is represented by the combined form, the mean and variance of Y are determined using (2-5) and (2-6).

Exact Moments

In this section, the properties of statistical expectation of a random variable are used to derive moments of various considered forms. When Y is represented by a multiplicative form, the first moment or mean of Y, μ_Y, can be written as

$$\mu_Y = E[Y] = C_0 \prod_{i=1}^{n} E\left[X_i^{r_i}\right] = C_0 \prod_{i=1}^{n} \mu_i \tag{2-12}$$

where $E[]$ is an expectation operator, and μ_i is the mean of the ith power function given as

$$\mu_i = E\left[X_i^{r_i}\right] . \tag{2-13}$$

Similarly, the second moment of model output about the origin can be written as

$$E\left[Y^2\right] = C_0^2 \prod_{i=1}^{n} E\left[\left(X_i^{r_i}\right)^2\right] = C_0^2 \prod_{i=1}^{n} \left(\mu_i^2 + \sigma_i^2\right). \tag{2-14}$$

where σ_i^2 is the variance of the i^{th} power function given as

$$\sigma_i^2 = Var(X_i^r) = E\left[X_i^{2r_i}\right] - \mu_i^2 \tag{2-15}$$

The variance of Y, σ_Y^2, can be expressed in terms of first and second moment (Haan, 1977) as

$$\sigma_Y^2 = E\left[Y^2\right] - \{E[Y]\}^2 \tag{2-16}$$

Substituting, (2-12) and (2-14) in to (2-16)

$$\sigma_Y^2 = C_0^2 \prod_{i=1}^{n} \left(\mu_i^2 + \sigma_i^2\right) - C_0^2 \prod_{i=1}^{n} \mu_i^2 \tag{2-17}$$

The coefficient of variation of Y, CV_Y can be written as

$$CV_Y^2 = \frac{\sigma_Y^2}{\mu_Y^2} = \frac{C_0^2 \prod_{i=1}^{n} (\mu_i^2 + \sigma_i^2) - C_0^2 \prod_{i=1}^{n} \mu_i^2}{C_0^2 \prod_{i=1}^{n} \mu_i^2} \tag{2-18a}$$

Simplifying (2-18a), CV_Y can be written as

$$CV_Y^2 = \sum_{i=1}^{n} CV_i^2 + \sum_{i,j=1,i \neq j}^{n} CV_i^2 CV_j^2 + \dots\dots\dots\dots\dots + CV_1^2 CV_2^2 \dots\dots\dots CV_n^2 \tag{2-18b}$$

where $CV_i = \frac{\sigma_i}{\mu_i}$ = coefficient of variation of i^{th} power function. Equation (2-18b) shows that the output uncertainty of a multiplicative model is governed by the most uncertain component functions. For the convenience of computation, (2-18b) can be shown to be equal to

$$CV_Y = \left[\prod_{i=1}^{n} (CV_i^2 + 1) - 1 \right]^{0.5} \tag{2-18c}$$

Equation (2-18c) can be used to determine the uncertainty in model output if CVs of component power functions are known correctly. Using CV_Y and a correct value of μ_Y, the exact variance of model output can be evaluated.

Using the additive form (2-2), the mean of Y, μ_Y, is given as

$$\mu_Y = C_1 E[X_1^{r_1}] + C_2 E[X_2^{r_2}] + \dots\dots\dots + C_n E[X_n^{r_n}] = \sum_{i=1}^{n} C_i \mu_i \tag{2-19}$$

Similarly, the variance of Y, σ_Y^2, can be written as

$$\sigma_Y^2 = C_1^2 Var[X_1^{r_1}] + C_2^2 Var[X_2^{r_2}] + \dots\dots\dots + C_n^2 Var[X_n^{r_n}] = \sum_{i=1}^{n} C_i^2 \sigma_i^2 \tag{2-20}$$

Equation (2-20) shows that magnitude of C_i is also equally important as uncertainty of a component function (X_i^r).

For evaluating the mean and variance of combined forms of Y such as (2-3), the mean and variance of the additive part must be determined first using (2-19) and (2-20). Next (2-12), (2-17), and (2-18c) are used to determine the mean, variance, and CV of Y by treating the combined form as a multiplicative form for which the additive part is assumed to be a multiplicative component with known mean and variance.

It is noted that the mean and variance of Y for both the multiplicative and additive forms are a function of the exact mean and variance of individual power functions, whereas, FOA estimates for the mean and variance of Y are a function of mean and variance of input parameters. In order to determine exact mean and variance of Y, it is necessary to know the correct mean and variance of the individual power functions. FOA can be used to approximate the mean and variance of individual power functions. Since FOA estimates are not exact, they need to be corrected before using them to determine moments of the overall model output. In the following section a technique is suggested to correct FOA estimates for the mean and variance of a power function.

Correcting FOA Mean and Variance Estimates

Consider a power function

$$Y = f(X) = cX^r \tag{2-21}$$

where r and c are constants. Using (2-5), the FOA estimate for the mean, $\hat{\mu}_Y$, is given as

$$\hat{\mu}_Y = c\mu_X^r \tag{2-22}$$

Using (2-6), the FOA estimate for the variance of Y, $\hat{\sigma}_Y^2$, is given as

$$\hat{\sigma}_Y^2 = c^2 r^2 \mu_X^{2(r-1)} \sigma_X^2 = c^2 r^2 \mu_X^{2r} CV_X^2 \qquad (2\text{-}23)$$

The estimates provided by (2-22) and (2-23) for μ_y and σ_y contain errors. The relative error, E, in the FOA estimate for a moment of any order can be computed as

$$\text{E} = \frac{\textit{Exact value} - \textit{FOA estimate}}{\textit{Exact value}} = 1 - \frac{\textit{FOA estimate}}{\textit{Exact value}} \qquad (2\text{-}24)$$

The relative error in FOA estimates for the mean and variance of a power function depends upon the CV of the input parameter, magnitude of exponent r, and type of distribution for the input parameter. Rewriting (2-24)

$$\textit{Exact value} = \frac{\textit{FOA estimate}}{(1 - \text{E})} \qquad (2\text{-}25)$$

FOA estimates for the mean and variance of a power function can be corrected if their corresponding relative errors are known.

Development of Relative Error Functions

The exact estimates of the mean and variance of a power function can be evaluated analytically. The mean of Y can be calculated from

$$\mu_Y = E[Y] = cE[X^r] = \int_{-\infty}^{\infty} f(X) p_X(x) dX = c\int_{-\infty}^{\infty} X^r p_X(x) dX \qquad (2\text{-}26)$$

where $p_X(x)$ is the probability density function of X. Similarly, the variance, σ_Y^2, can be determined from

$$\sigma_Y^2 = E[(Y-\mu_Y)^2] = E[Y^2] - \mu_Y^2 = c^2 E[X^{2r}] - \{cE[X^r]\}^2 = c^2 \int_{-\infty}^{\infty} X^{2r} p_X(x) dX - \mu_Y^2 \qquad (2\text{-}27)$$

Using (2-26) and (2-27) analytical estimates of μ_Y and σ_Y^2 can be determined for a given function *f(X)* and distribution $p_X(x)$. Equations (2-26) and (2-27) involves determination of $E[X^r]$ and $E[X^{2r}]$. If these two moments of Y about the origin can be determined for a distribution, its central moments μ_Y and σ_Y^2 can be fully characterized.

Using (2-26) and (2-27), exact values of the mean and variance are determined for a given power function $f_X(x)$ and probability density function $p_X(x)$. Substituting the FOA estimates and the exact values of the mean and variance in (2-24), the corresponding expressions for E are derived for commonly used distributions (Appendix I). The derived expressions of E for different distributions are presented here.

Uniform Distribution

The probability density function $p_X(x)$ for the continuous uniform distribution is

$$p_X(x) = \frac{1}{(\beta - \alpha)} \quad , \alpha \leq X \leq \beta \tag{2-28}$$

where α and β are the distribution parameters. Using the methods of moments, the estimates for α and β are given (Haan, 1977) as

$$\hat{\alpha} = \mu_X - \sqrt{3}\sigma_X = \mu_X\left(1 - \sqrt{3}CV_X\right) \tag{2-29}$$

$$\hat{\beta} = \mu_X + \sqrt{3}\sigma_X = \mu_X\left(1 + \sqrt{3}CV_X\right) \tag{2-30}$$

The expression for the relative error in the FOA predicted mean, $\mathrm{E}(\hat{\mu}_Y)$ is given as

$$\mathrm{E}(\hat{\mu}_Y) = 1 - \frac{2\sqrt{3}(r+1)CV_X}{\left[\left(1 + CV_X\sqrt{3}\right)^{(r+1)} - \left(1 - CV_X\sqrt{3}\right)^{(r+1)}\right]} \tag{2-31}$$

Figure 2-1 shows a plot $\mathrm{E}(\hat{\mu}_Y)$ vs. exponent r for CV_X ranging from 0.01 to 0.57. The relative error in FOA predicted variance, $\mathrm{E}(\hat{\sigma}_y^2)$, is expressed as

$$E(\hat{\sigma}_y^2) = 1 - \frac{12\,(2r+1)\,r^2(r+1)^2 CV_X^4}{\left\{2\sqrt{3}CV_X(r+1)^2\left[\left(1+CV_X\sqrt{3}\right)^{2r+1} - \left(1-CV_X\sqrt{3}\right)^{2r+1}\right] - (2r+1)\left[\left(1+CV_X\sqrt{3}\right)^{r+1} - \left(1-CV_X\sqrt{3}\right)^{r+1}\right]^2\right\}}$$

(2-32)

Figure 2-2 depicts a plot of $E(\hat{\sigma}_y^2)$ vs. r for CV_X values ranging from 0.01 to 0.57.

Symmetrical Triangular Distribution

The probability density function $p_X(x)$ for the triangular distribution is

$$p_X(x) = \frac{2}{(\beta-\alpha)}\frac{(X-\alpha)}{(\gamma-\alpha)} \quad \text{, when } \alpha \le X \le \gamma \tag{2-33a}$$

$$p_X(x) = \frac{2}{(\beta-\alpha)}\frac{(\beta-X)}{(\beta-\gamma)} \quad \text{, when } \gamma \le X \le \beta \tag{2-33b}$$

where α, β, γ are the minimum, maximum, and mode values of X. For a symmetric triangle, $\gamma = \mu_X$. The method of moments estimates for α and β are

$$\hat{\alpha} = \mu_X - \sqrt{6}\sigma_X = \mu_X\left(1-\sqrt{6}CV_X\right) \tag{2-34}$$

$$\hat{\beta} = \mu_X + \sqrt{6}\sigma_X = \mu_X\left(1+\sqrt{6}CV_X\right) \tag{2-35}$$

The expression for $E(\hat{\mu}_Y)$ is

$$E(\hat{\mu}_Y) = 1 - \frac{6(r+1)(r+2)CV_X^2}{\left[\left(1+CV_X\sqrt{6}\right)^{(r+2)} + \left(1-CV_X\sqrt{6}\right)^{(r+2)} - 2\right]} \tag{2-36}$$

Equation (2-36) has been represented graphically in Figure 2-3 for different values of r and CV_X ranging from 0.01 to 0.4. The $E(\hat{\sigma}_y^2)$ can be expressed as

$$E(\hat{\sigma}_y^2) = 1 - \frac{36\left(2r+1\right) r^2\left(r+1\right)^2\left(r+2\right)^2 CV_X^4}{\left\{3\left(r+1\right)\left(r+2\right)^2 CV_X^2\left[\left(1+CV_X\sqrt{6}\right)^{2r+2}+\left(1-CV_X\sqrt{6}\right)^{2r+2}-2\right]-\left(2r+1\right)\left[\left(1+CV_X\sqrt{6}\right)^{r+2}+\left(1-CV_X\sqrt{6}\right)^{r+2}-2\right]^2\right\}}$$

(2-37)

Figure 2-4 plots equation (2-37) for various r and CV_X values ranging from 0.01 to 0.4.

Lognormal Distribution

If X is lognormally distributed with mean μ_X and variance σ_X^2, its probability density function is given (Haan, 1977) as

$$p_X(x) = \frac{1}{\sigma_V X\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln X-\mu_V}{\sigma_V}\right)^2}, X > 0 \qquad (2\text{-}38)$$

where $V = ln(X)$ is normally distributed with parameters μ_V and σ_V^2. The parameters μ_V and σ^2_V are defined (Haan, 1977) as

$$\mu_V = \frac{1}{2}\ln\left[\frac{\mu_X^2}{CV_X^2+1}\right] \qquad (2\text{-}39)$$

$$\sigma_V^2 = \ln\left(CV_X^2+1\right) \qquad (2\text{-}40)$$

The expression for $E(\hat{\mu}_Y)$ is

$$E(\hat{\mu}_Y) = 1 - \left(1+CV_X^2\right)^{\frac{1}{2}r(1-r)} \qquad (2\text{-}41)$$

$E(\hat{\sigma}_y^2)$ can be rewritten as

$$E(\hat{\sigma}_y^2) = 1 - \frac{r^2 CV_X^2\left(CV_X^2+1\right)^r}{\left(CV_X^2+1\right)^{r^2}\left[\left(CV_X^2+1\right)^{r^2}-1\right]} \qquad (2\text{-}42)$$

$E(\hat{\mu}_Y)$ and $E(\hat{\sigma}_y^2)$ vs. r plots are shown in Figures 2-5 and 2-6 for CV_X values ranging from 0.01 to 1.0.

Gamma Distribution

The gamma distribution density function is given by

$$p_X(x)=\frac{\lambda^{\alpha}e^{-\lambda x}X^{(\alpha-1)}}{\Gamma(\alpha)},\ X,\ \alpha,\ \text{and}\ \lambda>0 \tag{2-43}$$

where α and λ are the distribution parameters. The method of moments estimates for α and λ are given (Haan, 1977) as

$$\hat{\lambda}=\frac{\mu_X}{\sigma_X^2} \tag{2-44}$$

$$\hat{\alpha}=\frac{\mu_X^2}{\sigma_X^2}=\frac{1}{CV_X^2} \tag{2-45}$$

The expression for $E(\hat{\mu}_Y)$ is

$$E(\hat{\mu}_Y)=1-\frac{CV_X^{-2r}\Gamma\left(CV_X^{-2}\right)}{\Gamma\left[CV_X^{-2}\left(1+rCV_X^2\right)\right]} \tag{2-46}$$

Figure 2-7 shows a plot $E(\hat{\mu}_Y)$ vs. r for CV_X values ranging from 0.01 to 1.0. $E(\hat{\sigma}_y^2)$ is expressed as

$$E(\hat{\sigma}_y^2)=1-\frac{r^2CV_X^{2(1-2r)}\left[\Gamma\left(CV_X^{-2}\right)\right]^2}{\Gamma\left[CV_X^{-2}\left(1+2rCV_X^2\right)\right]\Gamma\left(CV_X^{-2}\right)-\left\{\Gamma\left[CV_X^{-2}\left(1+rCV_X^2\right)\right]\right\}^2} \tag{2-47}$$

Figure 2-8 shows a plot $E(\hat{\sigma}_y^2)$ vs. r for CV_X values ranging from 0.01 to 1.0.

Exponential Distribution

The exponential distribution is a special case of the gamma distribution with α = 1. Substituting $\hat{\alpha}$ = 1 into (2-45), CV_X = 1. The expression for $E(\hat{\mu}_Y)$ is obtained by substituting CV_X = 1 into (2-46) as

$$\mathrm{E}(\hat{\mu}_Y) = 1 - \frac{1}{\Gamma(r+1)} \tag{2-48}$$

On substituting $CV_X = 1$ into (2-47), $\mathrm{E}(\hat{\sigma}_y^2)$ can be expressed as

$$\mathrm{E}(\hat{\sigma}_y^2) = 1 - \frac{r^2}{\left[\Gamma(2r+1) - \Gamma^2(r+1)\right]} \tag{2-49}$$

In Figure 2-9, $\mathrm{E}(\hat{\mu}_Y)$ and $\mathrm{E}(\hat{\sigma}_y^2)$ have been plotted with respect to exponent r.

Normal Distribution

The probability density function of normal distribution is

$$p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2} \tag{2-50}$$

where μ_x and σ_X^2 are the parameters of the normal distribution. When $CV_X < 1.0$, the general expression for $E[X^r]$ is

$$E\left[X^r\right] = \mu_X^r\left[1 + rCV_X E[Z] + \frac{r(r-1)}{2!} CV_X^2 E[Z^2] + ... + \frac{r(r-1)(r-2)....(r-n+1)}{n!} CV_X^n E[Z^n] + ...\right]$$

(2-51)

where r is any exponent (positive, negative, integer or fraction); n is the term number in the expansion; and Z is the standard normal variate. The term $E[Z^n]$ is defined (Benjamin and Cornell, 1970) as

$$E\left[Z^n\right] = \frac{2^{n/2}\Gamma[(n+1)/2]}{\sqrt{\pi}} = \frac{n!}{2^{n/2}(n/2)!} = (n-1)(n-3)........(3)(1), \text{ when } n \text{ is even} \tag{2-52a}$$

and

$$E\left[Z^n\right] = 0, \text{ when } n \text{ is odd.} \tag{2-52b}$$

Substituting (2-52a) and (2-52b) into (2-51)

$$E\left(X^r\right)=\mu_X^r\left[1+\frac{r(r-1)}{2}CV_X^2+\frac{r(r-1)(r-2)(r-3)}{2^{4/2}(4/2)!}CV_X^4\ldots\ldots+\frac{r(r-1)(r-2)\ldots..(r-n+1)}{2^{n/2}(n/2)!}CV_X^n+\ldots\right] \tag{2-53}$$

When r is a positive integer, the RHS of (2-53) is finite and terminates at $n = r + 1$. Consequently, (2-53) can be written as

$$E\left(X^r\right)=\mu_X^r\sum_{n=0}^{r/2}\binom{r}{2n}\frac{2n!}{2^n n!}CV_X^{2n}, \text{ when } r \text{ is even and;} \tag{2-54a}$$

$$E\left(X^r\right)=\mu_X^r\sum_{n=0}^{(r-1)/2}\binom{r}{2i}\frac{2n!}{2^n n!}CV_X^{2n}, \text{ when } r \text{ is odd.} \tag{2-54b}$$

For values of r other than a positive integer (2-53) does not converge. In order to determine $E[X^r]$, (2-53) needs to be truncated. When r is a positive fraction, $E[X^r]$ can be obtained using (2-54a) and (2-54b) with rounded value of r to its nearest whole number. In cases when r is negative, the truncating error depends upon the magnitudes of r and CV_X. Further, there exists a minimum truncating error for a given combination of r and CV_X, beyond which no improvement in $E[X^r]$ is possible. To evaluate the approximate value of $E[X^r]$, a trial and error procedure was used to determine the number of terms to be summed up to obtain $E[X^r]$ corresponding to the minimum truncating error for a given combination of r and CV_X.

After estimating $E[X^r]$ and $E[X^{2r}]$, (2-26) and (2-27) are used to determine μ_Y and σ_Y^2. Substituting μ_Y, σ_Y^2 and the FOA estimates $\hat{\mu}_Y$ and $\hat{\sigma}_Y^2$ into (2-24), the relative error in FOA predicted estimates of the mean and variance, $\mathrm{E}(\hat{\mu}_Y)$ and $\mathrm{E}(\hat{\sigma}_y^2)$, can be determined. Figures 2-10 and 2-11 show plots of $\mathrm{E}(\hat{\mu}_Y)$ and $\mathrm{E}(\hat{\sigma}_y^2)$ vs. r for various values of CV_X ranging from 0.02 to 0.33.

Salient Features of Relative Error Curves

When applying FOA, it is assumed that the functional relationship between the dependent and independent parameters can be approximated by a linear relationship. This assumption is often valid, but specific situations may arise when the function is very nonlinear (represented either by a very large or very small exponent of a power function). For this reason, applying FOA to models containing a power function with a large exponent is not common. These situations can be identified and dealt with by visualizing the relative error plots. The relative error plots also show where FOA estimates are nearly acceptable and where they are unacceptable and need to be corrected. Observing these error plots, the following salient points are noted:

1.The relative error is zero for a power function of different distributions at certain values of the exponent. These exponents are 0 and 1 as shown by $E(\hat{\mu}_Y)$ plots (Figures 2-1, 2-3, 2-5, 2-7, 2-10) for all the considered distributions. The exponent value of 0 represents a constant function and the exponent value of 1 corresponds to a linear function. In the same way, there are two exponent values for $E(\hat{\sigma}_y^2)$ where FOA estimates for the variance have no error. One of these exponents is 1 and the other changes with the distribution type and CV of the input parameter as shown in Table 2-2.

Table 2-2 shows that when the exponent of a power function lies within the tabulated range for each distribution, the FOA variance estimate will have almost no error and the power function will behave like a linear function as far as the variance prediction is concerned. These situations are depicted by $E(\hat{\sigma}_y^2)$ vs. r plots in Figures 2-2, 2-4, 2-6, 2-8, 2-11.

2. The variation of relative error in FOA predicted variance also changes with respect to CV, exponent, and type of distribution. When the exponent falls between 1 and 1.7 (an approximate value) for normal, uniform, and triangular distributed parameters, the FOA overestimates the actual variance. However, the overestimation is small as shown by the negative values of $E(\hat{\sigma}_y^2)$ in Figures 2-2, 2-4, and 2-11. When the exponent falls outside this range, the FOA underestimates the actual variance. When the parameter is lognormally distributed and the exponent falls between 0 and 1, the FOA may highly overestimate the actual variance depending upon the parameter CV value as shown negative values of $E(\hat{\sigma}_y^2)$ in Figure 2-6. When the exponent falls outside this range, FOA underestimates the actual variance. In the case of the gamma distribution, when the exponent falls between 0.3 and 1, the FOA overestimates the actual variance. For exponents outside this range the FOA underestimates the actual variance.

3. It is clear from the error plots that the type of distribution may affect the accuracy of the FOA predicted variance. For example, with a power function exponent between 0 and 1 the FOA overestimates the actual variance when a parameter is lognormally distributed and underestimates when a parameter has normal, uniform, or triangular distribution. Whereas, in case of gamma distributed parameter, the FOA overestimates the actual variance when exponent lies between 1.0 and 0.2 and underestimates elsewhere.

4. Even a very small exponent (very close to zero) may give a very high relative error in FOA predicted variance for some of the distributions (normal, uniform, and triangular).

5. Error plots of the normal distribution (Figures 2-10 and 2-11) show that significant errors occur in both the mean and variance of a power function predicted using FOA. This contradicts previous findings that FOA works well when input variables are normally distributed (e.g. Scavia et al., 1981; Johnson, 1996).

6. When a power function has its exponent in the vicinity of those tabulated in Table 2-2, the relative error is very small regardless of the CV values of the input variable. This contradicts previous findings that FOA works well only when CV ≤ 0.2.

Examples

Three examples demonstrating the use of the corrected first order uncertainty method involving multiplicative, additive, and combined form models are presented.

Example No. 1 (Uncertainty in water distribution)

Hydraulic modeling of a water distribution network is a critical component in the planning, design, maintenance, and operational control of water supply systems. Analysis of water distribution networks involves the determination of nodal heads and pipe flow rates. A basic relationship describing the dependence of discharge on head loss caused by friction between the flow of fluid and the pipe wall is used in the hydraulic design of a pipeline system (Mays, 1999). One of the most widely used head loss relationships is the Hazen-Williams equation, which is given as (Mays, 1999)

$$h_f = \frac{10.654 L Q^{1.852}}{C^{1.852} D^{4.87}} = 10.654 L Q^{1.852} C^{-1.852} D^{-4.87} \qquad (2\text{-}55)$$

where h_f is head loss (m), L is length of pipe (m), D is pipe diameter (m), Q is flow (m^3/sec), and C is the Hazen-Williams roughness coefficient which varies with pipe materials and age (Mays and Tung, 1992). h_f is uncertain due to uncertainty in Q, C, and D. L is assumed to be exact (1500 m). Table 2-3 gives the mean, CV, and assumed distribution (Mays and Tung, 1992) for the uncertain variables in Q, C, and D.

The FOA estimate for the mean, $\hat{\mu}_{h_f}$, is calculated using (2-7) as

$$\hat{\mu}_{h_f} = 10.654(1500)(0.915)^{1.852}(130)^{-1.852}(0.305)^{-4.87} = 535.29\text{ m.}$$

Using (2-9) the FOA estimate for the CV of h_f is

$$C\hat{V}_{h_f} = \left[(1.852)^2(0.1)^2 + (-1.852)^2(0.15)^2 + (-4.87)^2(0.05)^2\right]^{0.5} = 0.413$$

Multiplying the mean and coefficient of variation calculated above, the FOA estimate of the standard deviation, $\hat{\sigma}_{h_f}$ is 221.2 m. The FOA estimates for the mean and variance for the component power functions are calculated using (2-22) and (2-23) and listed in columns 2 and 5 of Table 2-4. To correct the FOA estimates for the means and variances of the component power functions, the relative error equations (2-41) and (2-42) developed for the lognormally distributed variables are used. The calculated relative errors have been listed in columns 3 and 6 of Table 2-4. The exact estimates of means and variances for the component power functions are determined using (2-25) and are shown in columns 4 and 7 of Table 2-4. Values of CV_i^2 in column 8 shows that all component functions are important for determining uncertainty in h_f, however, the contribution of $C^{-1.852}$ is the maximum.

Using (2-12) and corrected means of component power functions from column 4 of Table 2-4, the value of μ_{hf} = 592.96 m. To calculate CV_{hf}, (2-18c) and column 8 of

Table 2-4 are used to give $C\hat{V}_{h_f} = [(1+0.035)(1+0.079)(1+0.061)) - 1]^{0.5} = 0.43$. Using these values for μ_{hf} and CV_{hf}, the standard deviation, σ_{hf}, is 254.98 m.

The MCS technique is also used to estimate the mean and variance of h_f. Figures 2-12a and 2-12b show plots of μ_{hf} and σ_{hf} vs. number of simulations. It can be seen from the plots that there is quite a bit of fluctuation in standard deviation even after 20,000 simulations. The μ_{hf} and σ_{hf} values based on 20,000 simulations are 594.12 m and 256.16 m respectively.

Example No. 2 (Hydraulic uncertainty for flood levee capacity)

Manning's equation is the most commonly used resistance equation to find the flow in a section (Chow, 1959). It is expressed as

$$Q = \frac{1}{n} A R^{\frac{2}{3}} S^{\frac{1}{2}} \tag{2-56}$$

where Q is flow (m^3/sec), A is the cross sectional area of the channel (m^2), R is the hydraulic radius of the channel (m), S is the channel slope (m/m), and n is Manning's coefficient (SI units). Natural channels often have a main channel section and an overbank section. Most flow occurs in the main channel, however, during flood events overbank flows may occur. Considering a symmetric river-flood plain section, the overall flow capacity, Q, for the compound section can be expressed as

$$Q = \left(\frac{Y_c}{n_c} + 2\frac{Y_b}{n_b}\right) S^{\frac{1}{2}} = \left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right) S^{0.5} \tag{2-57}$$

where $Y_i = AR^{2/3}$ is called the section factor. Y_c and Y_b represent section factors for main channel and overbank sections respectively. Considering section factors to be deterministic (Y_c = 296.9 m$^{8/3}$ and Y_b = 0.6 m$^{8/3}$), there are three random variables in (2-57) n_c, n_b, and S. Substituting values of Y_c and Y_b, (2-57) is rewritten as

$$Q = \left(296.9 n_c^{-1} + 1.2 n_b^{-1}\right) S^{0.5} = \phi S^{0.5} \tag{2-58}$$

where ϕ is a dummy variable representing the additive form defined as $\phi = 269.9 n_c^{-1} + 1.2 n_b^{-1}$. In the literature (Tung, 1990; Cesare, 1991; Mays and Tung, 1992; Yeh and Tung, 1993), when applying FOA to Manning's equation a small CV is assumed for n. In this example greater CV values for both banks and the main channel have been assumed as reported by Johnson (1996). Table 2-5 gives the mean, CV, and distribution type (Johnson, 1996) for the uncertain variables n_c, n_b, and S in (2-58).

Using the data of Table 2-5, the FOA estimates for mean and standard deviation of Q are found to be 618.72 m^3/sec and 130.39 m^3/sec respectively. To determine exact values of mean and standard deviation of Q, first FOA estimates of component power functions are corrected as given in Table 2-6.

Using (2-19) and corrected means for the power functions from column 4 of Table 2-6, the exact mean of the additive form, μ_ϕ, is 9019.89 m^3/s. Similarly using (2-20) and corrected variances of the component power functions from column 7 of Table 2-6, σ_ϕ = 1586.19 m^3/s. The corresponding CV_ϕ is 0.176. Now, treating Q as a multiplicative form with ϕ and $S_0^{0.5}$ as its components with known means and CV values μ_Q = 632.99 using (2-12), CV_Q = 0.216 using (2-18c) are obtained. Multiplying values of μ_Q and CV_Q, σ_Q = 136.74 is obtained. In this example it can be noted that n_b is the most uncertain parameter but its contribution to the uncertainty of Q is negligible as the additive form ϕ is governed mainly by n_c because of its very large coefficient in comparison to that of n_b.

Figures 2-13a and 2-13b are plots of μ_Q and σ_Q obtained using different number of MCS simulations. There is still a sizable uncertainty in Q, as convergence is not reached after 20,000 simulations. The values of μ_Q and σ_Q corresponding to 20,000 MCS simulations are 634.61 m^3/sec and 137.08 m^3/sec respectively.

Example No. 3 (Probabilistic human health risk assessment)

Quantitative risk assessment has received increased attention because of the recognition of both the potential threat to human health from hazardous substances and the potential for releases into the environment. Recognizing the extent of the hazardous waste problem and role of risk assessment, the EPA has developed assessment procedures that are used for a variety of purposes. Some examples are designating substances as hazardous, establishing minimum quantities for reporting releases when they would present substantial danger, evaluating the relative dangers of various sites in order to establish priorities for response actions, developing, and selecting appropriate response actions at the contaminated sites. Risk assessment is also used to evaluate threats to public health posed by superfund sites (USEPA, 1989).

The risk assessment process used by the EPA is carried out in four steps (USEPA, 1989).

The first step is hazard identification during which contaminants of concern are selected based on their toxicity, mobility, spatial distribution and concentration. The second step is exposure assessment in which all possible pathways (e.g., inhalation, ingestion, dermal, etc.) are identified through which contaminants are exposed to the human body. In the third step, intake doses of the pre-identified contaminants absorbed through various exposure routes are estimated. The final step is the risk characterization in which the magnitude of the risk is calculated. Quantitative uncertainty analysis is necessary when screening level calculations indicate a potential problem, remediation may result in high costs, or it is necessary to establish the relative importance of contaminants and exposure pathways.

To demonstrate an application of the developed method in risk characterization, risk assessment due to ingestion of contaminated soils is considered. Ingestion of soils contaminated by high molecular weight contaminants such as polychlorinated biphenyl (PCBs) is a potential source of human exposure to toxicants. The following equation (USEPA, 1990) is used to estimate the probability of life-time cancer (R_c) due to ingestion of soil

$$R_c = \frac{C_s I_r C_f F_i E_f E_d}{B_w A_t} S_f = C_s I_r C_f F_i E_f E_d B_w^{-1} A_t^{-1} S_f \qquad (2\text{-}59)$$

where C_s is the chemical concentration in the soil (mg/kg), C_f is a conversion factor (10^{-6} kg/mg), I_r is the ingestion rate (mg soil/day), F_i is the fraction ingested from contaminated sources (non-dimensional), E_f is the exposure frequency (days/year), E_d is the exposure duration (years), B_w is the body weight (kg), A_t is the averaging time (period over which exposure is averaged in days), and S_f is the slope factor or cancer potency factor $(\text{mg/kg-day})^{-1}$.

There is always some uncertainty about each of these elements in risk estimation. A large number of references are available to describe the extent of uncertainty in each of these elements. Talcott (1992) has summarized the available information in detail. In this analysis the mean, CV, and distributions of the variables are taken from Batchelor et al. (1998) corresponding to the age group of 1-6 years. The distribution of F_i was assumed to be the lognormal instead of the beta distribution as reported by Batchelor et al. (1998). This data are given in Table 2-7.

In (2-59) there are two constants. One constant is C_f and the other is (1/365) to convert A_t in days. Combining these two a new constant $C_0 = 10^{-6}/365 = 2.74E-09$ is obtained. Using (2-7), the FOA estimate of $\hat{\mu}_{Rc}$ is 1.80E-05. Similarly, using (2-9), the FOA estimate for $\hat{CV}_{Rc}$ is 2.57. Multiplying these two, the estimate for $\hat{\sigma}_{Rc}$ is 4.62E-05. To determine the exact mean and variance of R_c, FOA estimates for means and variances of component functions are corrected as shown in Table 2-8.

Substituting values of C_0 and corrected estimates for mean of component power functions from column 4 of Table 2-8 in (2-12), the exact mean of R_c (μ_{Rc}) is 1.97E-05. Similarly, substituting values of CV_i^2 from column 8 of Table 2-8 in (2-18c), the correct CV_{Rc} is 6.95. Multiplying μ_{Rc} and CV_{Rc}, the standard deviation of R_c, σ_{Rc}, is 1.37E-04.

Using MCS, μ_{Rc} and σ_{Rc} are determined for different number of simulations. These plots are shown in Figures 2-14a and 2-14b. It is clear from Figure 2-14b that there is a significant amount of uncertainty in risk prediction even after 20,000 simulations. The estimates of μ_{Rc} and σ_{Rc} corresponding to 20,000 simulations are 1.87E-05 and 1.06E-04 respectively.

Obtained results, using FOA, MCS and corrected FOA methods for examples 1, 2 and 3 are compared in Table 2-9.

Conclusions

In this paper analytical relationships are developed to determine the relative errors in FOA estimates for the means and variances of power functions. Using these relative error functions, one can correct the FOA estimates for the means and variances of component power functions for nonlinearity, and distribution type to evaluate the exact mean and variance of model output. For ease in application relative error curves for commonly used distributions are presented graphically. These plots can be used to determine an approximate relative error for a given exponent of a power function and CV of its random variable. Three examples are presented which shows that this technique is not only easy to use but also provides more insight into the process by analyzing each component function of the model separately. Special cases are identified when applying FOA to a nonlinear power function for estimating its variance will give a negligible or no error. This method provides a procedure for incorporating the known information of the types of input variable distributions. This technique is applicable when an output random variable is a function of several mutually independent random variables in multiplicative, additive, or in combined forms.

References

Ang, A. H. S., and Tang, W. H. (1984). Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk, and Reliability. John Wiley & Sons, Inc., New York, N.Y.

Batchelor, B., Valdes, J., and Araganth (1998). Stochastic risk assessment of sites contaminated by hazardous wastes. Journal of the Environmental Engineering Division, ASCE, 124(4), 380-388.

Bates, B. C. (1988). Nonlinear, discrete flood event models, 2. Assessment of statistical nonlinearity. Journal of Hydrology, 99, 77-89.

Bates, B. C., and Townley, L. R. (1988). Nonlinear discrete flood event models, 3. Analysis of prediction uncertainty. Journal of Hydrology, 99, 91-101.

Bates, D. M., and Watts, D. G. (1980). Relative curvature measures of nonlinearity. Journal of the Royal Statistical Society, B, 42(2-1), 1-25.

Beale, E. M. L. (1960). Confidence regions in nonlinear estimation. Journal of the Royal Statistical Society, B, 22(2-1), 41-88.

Benjamin, J. R., and Cornell, C. A. (1970). Probability, Statistics, and Decisions for Civil Engineers. McGraw-Hill, New York, N.Y.

Burges, S. J. (1979). Analysis of uncertainty in flood plain mapping. Water Resources Bulletin, 15(2-1), 227-243.

Burges, S. J., and Lettenmaier, D. P. (1975). Probability methods in stream quality management. Water Resources Bulletin, 11(2-1), 115 –130.

Cesare, M. A. (1991). First-order Analysis of open-channel flow. Journal of the Hydraulic Engineering Division, ASCE, 117(2-2), 242-247.

Chadderton, R. A., Miller, A. C., and McDonnell, A. J. (1982). Uncertainty analysis of dissolved oxygen model. Journal of the Environmental Engineering Division, ASCE, 108(5), 1003-1013.

Chadderton, R. A., and Miller, A. C. (1980). Analysis of Uncertainty in Floodplain Mapping, Water Resources Bulletin, 16(4), 752-754.

Chow, V. T. (1959). Open Channel Hydraulics, McGraw-Hill, New York, N.Y.

Cornell, C. A. (1972). First-order analysis of model and parameter uncertainty. Proc. Int. Symposium on Uncertainties in Hydrologic and Water Resources Systems, Univ. of Arizona, vol. III, 1245-1276.

Dettinger, M. D., and Wilson, J. L. (1981). First order analysis of uncertainty in numerical models of groundwater flow. Part 1. Mathematical development. Water Resources Research, 17(2-1), 149-161.

Devary, J. L., and Doctor, P. G. (1982). Pore velocity estimation uncertainties. Water Resources Research, 18(4), 1157-1164.

Gardner, R. H. and O'Neill, R. V. (1983). Parameter uncertainty and model predictions: A review of Monte Carlo results. Uncertainty and Forecasting of Water Quality, M. B. Beck and G. van Straten, eds., Springer Verlag, New York, 245-257.

Gardner, R. H., O'Neill, R. V., Carney, J. H., and Mankin, J. B. (1981). A comparison of sensitivity analysis and error analysis based on a stream ecosystem model, Ecological Modeling, 12, 173-190.

Garen, D. C. and Burges, S. J., (1981). Approximate error bounds for simulated hydrographs. Journal of the Hydraulic Engineering Division, ASCE, 107(11), 1519-1534.

Haan, C. T. (1977). Statistical Methods in Hydrology, The Iowa State University Press, Ames, Iowa.

Haan, C. T., Barfield, B. J., and Hayes, J. C. (1994). Design Hydrology and Sedimentology for Small Catchments. Academic Press, San Diego.

Hasofer, A. M., and Lind, N. C. (1974). Exact and invariant second-moment code format. Journal of the Engineering Mechanics Division, ASCE, 100, 111-121.

Johnson, P. A. (1996). Uncertainty of hydraulic parameters. Journal of the Hydraulic Engineering Division, ASCE, 122(2-2), 112-114.

Krenkel, P. A., and Ruane, R. J. (1979). Chapter 18: Basic approach to water quality modeling. Modeling of Rivers, H. W. Shen, ed., Wiley, New York.

Kuczera, G. (1988). On the validity of first-order prediction limits for conceptual hydrologic models. Journal of Hydrology, 103, 229-247.

Lee, H. L. and Mays, L. W. (1986). Hydraulic uncertainty in flood levee capacity. Journal of the Hydraulic Engineering Division, ASCE, 112(10), 928-934.

Malone, R. F., Saidi, H., and Wegener, K. (1984). Predictive accuracy determination applied to a linear model of phosphorus loading resulting from urban runoff. Applied Math. Modeling, 8, 81-88.

Mays, L. W. (1999). Water Distribution Systems Handbook, McGraw-Hill, New York, N.Y.

Mays, L. W., and Tung, Y. K. (1992). Hydrosystems Engineering and Management, McGraw-Hill, Inc., New York, N.Y.

Melching, C. S. (1992a). An improved first order reliability approach for assessing uncertainties in hydrologic modeling. Journal of Hydrology, 132, 157-177.

Melching, C. S. (1992b). A comparison of models for estimating variance of water resources model predictions. Proceedings Sixth International Association for Hydraulic Research Symposium on Stochastic Hydraulics, Taipei, Taiwan, J. T. Kuo and G. F. Lin, eds., Water Resources Publications, Littleton, Colorado, 663-670.

Melching, C. S. (1995). Reliability estimation. Computer Models of Watershed Hydrology, V. P.Singh, ed., Water Resources Publication, Littleton, Colorado, 69 – 118.

Melching, C. S., and Anmangandla, S. (1992). Improved first order uncertainty method for water quality modeling. Journal of Environmental Engineering Division, ASCE, 118(5), 791-805.

Moore, R. J., and Clarke, R. T. (1981). A distribution function approach to rainfall runoff modeling. Water Resources Research, 17(5), 1367-1382.

Novotny, V. and Olem, H. (1994). Water Quality: Prevention, Identification, and Management of Diffuse Pollution, Van Nostrand Reinhold, New York.

Sagar, B. (1978). Galerkin finite element procedure for analyzing flow through random media. Water Resources Research, 14(6), 1035-1044.

Scavia, D., Powers, W. F., Canale, R. P., and Moody, J. L (1981). Comparison of first-order error analysis and Monte Carlo simulation in time-dependent lake eutrophication models. Water Resources Research, 17(4), 1051-1059.

Siddall, J. N. (1983). Probabilistic Analysis, Probabilistic Engineering Design, Marcel Dekker, Inc., New York, N.Y., 145-236.

Smith, V. J., and Charbeneau, R. J. (1990). Probabilistic soil contamination exposure assessment procedures. Journal of the Environmental Engineering Division, ASCE, 6(116), 1143-1163.

Stevens, E. (1993). A Method to Predict the Accuracy of a First Order Approximation of Model Output Variance,. Ph. D. Thesis, Department of Biosystems and Agricultural Engineering, Oklahoma State University, Stillwater, OK.

Talcott, F. W. (1992). How certain is that environmental risk estimate? Resources for the Future, 107, 10-15.

Tang, W. H., and Yen, B. C. (1972). Hydrologic and hydraulic design under uncertainties. Proceedings International Symposium on Uncertainties in Hydrology and Water Resources Systems, University of Arizona, Tucson, Arizona, 2, 868-882.

Townley. L. R., and Wilson, J. L (1985). Computationally efficient algorithm for parameter estimation and uncertainty propagation in numerical models of groundwater flow, Water Resources Research, 21(12), 1851-1860.

Tung, Y. K. (1990). Mellin transformation applied to uncertainty analysis in hydrology/ hydraulics. Journal of the Hydraulic Engineering Division, ASCE, 116(5), 659-674.

Tung, Y. K., and Mays, L. W. (1981). Risk Models for Flood Levee Design. Water Resources Research, 17(4), 833-841.

Tung, Y. K., and Mays, L. W. (1980). Optimal risk-based design of water resources engineering projects. Technical Report No. 171, Center for Research in Water Resources, Univ. of Texas, Austin, Austin, Tex.

USEPA (1990). Guidance on remedial actions for superfund sites with PCB contamination, Rep. No. EPA/540/G-90/007, U.S. Office of Emergency and Remedial Response, Washington, D.C.

USEPA (1989). Risk assessment guidance for superfund volume 1 human health evaluation manual (part A) interim final, Rep. No. EPA/540/1-89/002, US. Office of Emergency and Remedial Response, Washington, D. C.

Walker, W. W. (1982). A sensitivity analysis and error analysis framework for lake eutrophication modeling, Water Resources Bulletin, 18, 53-60.

Yeh, K. C., and Tung, Y. K. (1993). Uncertainty and sensitivity analyses of pit-migration model. Journal of the Hydraulic Engineering Division, ASCE, 119 (2), 262-283.

Yen, B. C. and Tang, W. H. (1976). Risk-safety factor relation for storm sewer design, Journal of the Environmental Engineering Division, ASCE, 102 (2), 509-516.

Yen, B. C., Cheng, S. T., and Tang, W. H. (1980). Reliability of hydraulic design of culverts. Proceedings International Conference on Water Resources Development, International Association for Hydraulic Research Asian Pacific Division Second Congress, 2, Taipei, Taiwan, 991-1001.

Zhang, J., and Haan, C. T. (1996). Evaluation of uncertainty in estimated flow and phosphorus loads by FHANTM. Trans. Soil and Water Division, ASAE, 12(6), 663-669.

Table 2-1: Allowable ranges of coefficient of variation

Distribution	CV range
Uniform	$0 \le CV \le 1/\sqrt{3}$
Triangular	$0 \le CV \le 1/\sqrt{6}$
Normal	$0 \le CV \le 1/3^{*}$
Exponential	CV = 1
Lognormal	No restriction
Gamma	No restriction

*Any value of CV is theoretically possible with the normal distribution; however, for CV > 1/3, the probability of a negative number from the distribution increases rapidly.

Table 2-2: Exponent corresponding to zero error in FOA estimated variance

Distribution	Variation in the CV of input parameter	Variation in the exponent corresponding to $E(\hat{\sigma}_Y^2) \approx 0$	Exponent point value
Uniform	0.01 to 0.57	1.7 to 1.8	1.751
Triangular	0.01 to 0.40	1.6 to 1.8	1.700
Normal	0.01 to 0.33	1.6 to 1.7	1.650
Exponential	1.00	0.279	0.279
Lognormal	0.01 to 1.00	-0.333 to –0.347	-0.340
Gamma	0.01 to 1.00	0.20 to 0.40	0.300

Table 2-3: Uncertain parameters of Hazen-Williams equation

Variable	Distribution	Mean	CV
Q (m^3/s)	Lognormal	0.915	0.10
D (m)	Lognormal	0.305	0.05
C (SI units)	Lognormal	130.0	0.15

Table 2-4: Computation of the exact mean and variance of head-loss using corrected FOA method

Power Function	Mean			Variance			CV_i^2
	FOA Estimate	Relative Error*	Corrected estimate	FOA Estimate	Relative Error*	Corrected estimate	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Q^{1.852}$	0.848	7.82E-3	0.855	0.0247	0.027	0.025	0.035
$C^{-1.852}$	1.22E-4	0.057	1.29E-4	1.14E-9	0.135	1.32E-9	0.079
$D^{-4.87}$	324.68	0.035	336.48	6250.54	0.095	6907.12	0.061

*Relative errors for mean and variance can also be determined using relative error plots, the values read from plots may be less accurate because of individual error. Figures 2-5 and 2-6 contain the plots for the relative errors for the mean and variance for lognormally distributed random variables.

Table 2-5: Uncertain parameters of Manning's equation

Variable	Distribution	Mean	CV
n_c	Uniform	0.034	0.17
n_b	Uniform	0.068	0.38
S	Lognormal	0.005	0.25

Table 2-6: Computation of the exact mean and variance of flood levee capacity using corrected FOA method

Power Function	Mean			Variance			CV^2
	FOA Estimate	Relative Error	Corrected estimate	FOA Estimate	Relative Error	Corrected estimate	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
n_c^{-1}	29.41	0.029	30.31	24.99	0.124	28.54	0.031
n_b^{-1}	14.71	0.166	17.64	31.22	0.556	70.31	0.226
$S^{0.5}$	0.071	-0.008	0.07	7.81×10^{-5}	-0.039	7.52×10^{-5}	0.015

Column 2 is calculated using Equation (2-22), column 3 from Equation (2-31) for uniform distribution and Equation (2-41) for lognormal distribution, column 4 from Equation (2-25). Similarly, column 5 is calculated from Equation (2-23), column 6 from Equations (2-32) and (2-42), and column 7 from Equation (2-25). Column 8 = column 7/(column 4)^2.
Note: Columns 3 and 6 can also be determined using relative error plots.

Table 2-7: Statistical data for human health risk assessment

Parameter	Symbol	Distribution	Parameter values	
			Mean	CV
Contaminant concentration (mg/kg)	C_s	Lognormal	155	0.39
Ingestion rate (mg/day)	I_r	Lognormal	100	1.26
Fraction ingested	F_i	Lognormal	0.909	0.03
Exposure frequency (days/yr.)	E_f	Exponential	17.4	1.0
Exposure duration (yr.)	E_d	Exponential	13.0	1.0
Body weight (kg)	B_w	Lognormal	15.6	0.23
Averaging time (yr.)	A_t	Normal	70.0	0.19
Slope factor (kg-day/mg)	S_f	Lognormal	2.25	1.66

Table 2-8: Computation of the exact moments of human health risk using corrected FOA method

Component power function (1)	Mean FOA estimate (2)	Mean Relative error (3)	Mean Corrected estimate (4)	Variance FOA estimate (5)	Variance Relative error (6)	Variance Corrected estimate (7)	CV_i^2 (8)
C_s	155.000	0.000	155.000	3600.00	0.000	3600.00	0.0150
I_r	100.000	0.000	100.000	15876.00	0.000	15876.00	1.5880
F_I	0.909	0.000	0.909	7.29E-4	0.000	7.29E-4	0.0010
E_f	17.400	0.000	17.400	302.76	0.000	302.76	1.0000
E_d	13.000	0.000	13.000	169.00	0.000	169.00	1.0000
B_w^{-1}	0.064	0.050	0.067	2.13E-4	0.098	2.36E-4	0.0052
A_t^{-1}	0.014	0.039	0.015	7.04E-6	0.306	1.01E-5	0.0460
S_f	2.250	0.000	2.250	13.99	0.000	13.99	2.7630

Table 2-9: Comparison between output results using FOA and corrected FOA methods

Example	Output	Mean FOA	Mean Exact	Mean MCS*	Standard deviation FOA	Standard deviation Exact	Standard deviation MCS*
1	h_f(m)	535.3	592.96	594.12	221.2	254.98	256.16
2	Q (m^3/s)	618.7	632.99	634.61	130.39	136.74	137.08
3	R_c	1.8E-5	1.97E-5	1.87E-5	4.62E-5	1.37E-4	1.06E-4

*Values based on 20,000 number of MCS simulations.

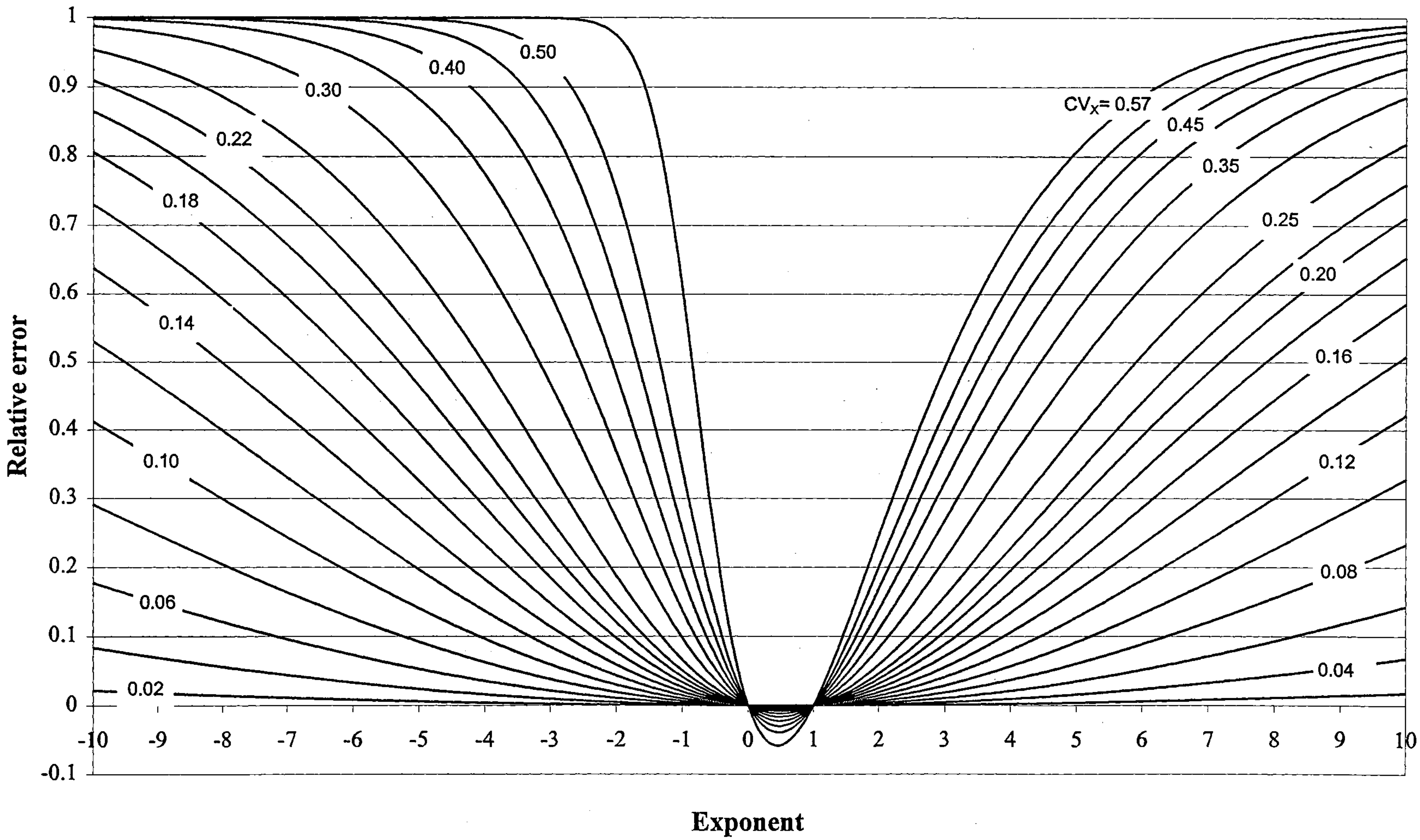

Fig. 2-1: Relative error in FOA predicted mean for CV_X ranging from 0.01 to 0.57, where X is uniformly distributed

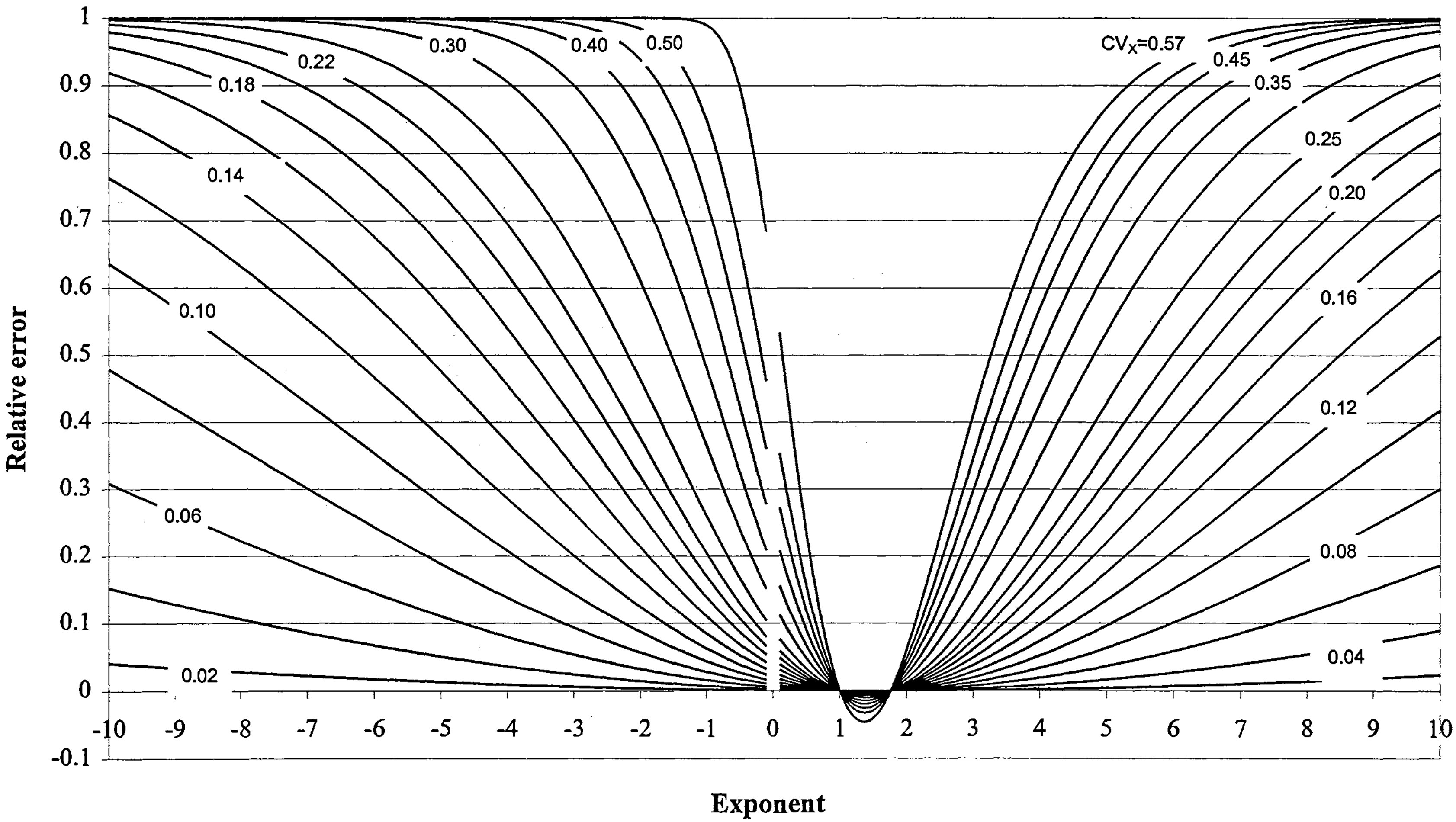

Fig. 2-2: Relative error in FOA predicted variance for CV_X ranging from 0.01 to 0.57, where X is uniformly distributed

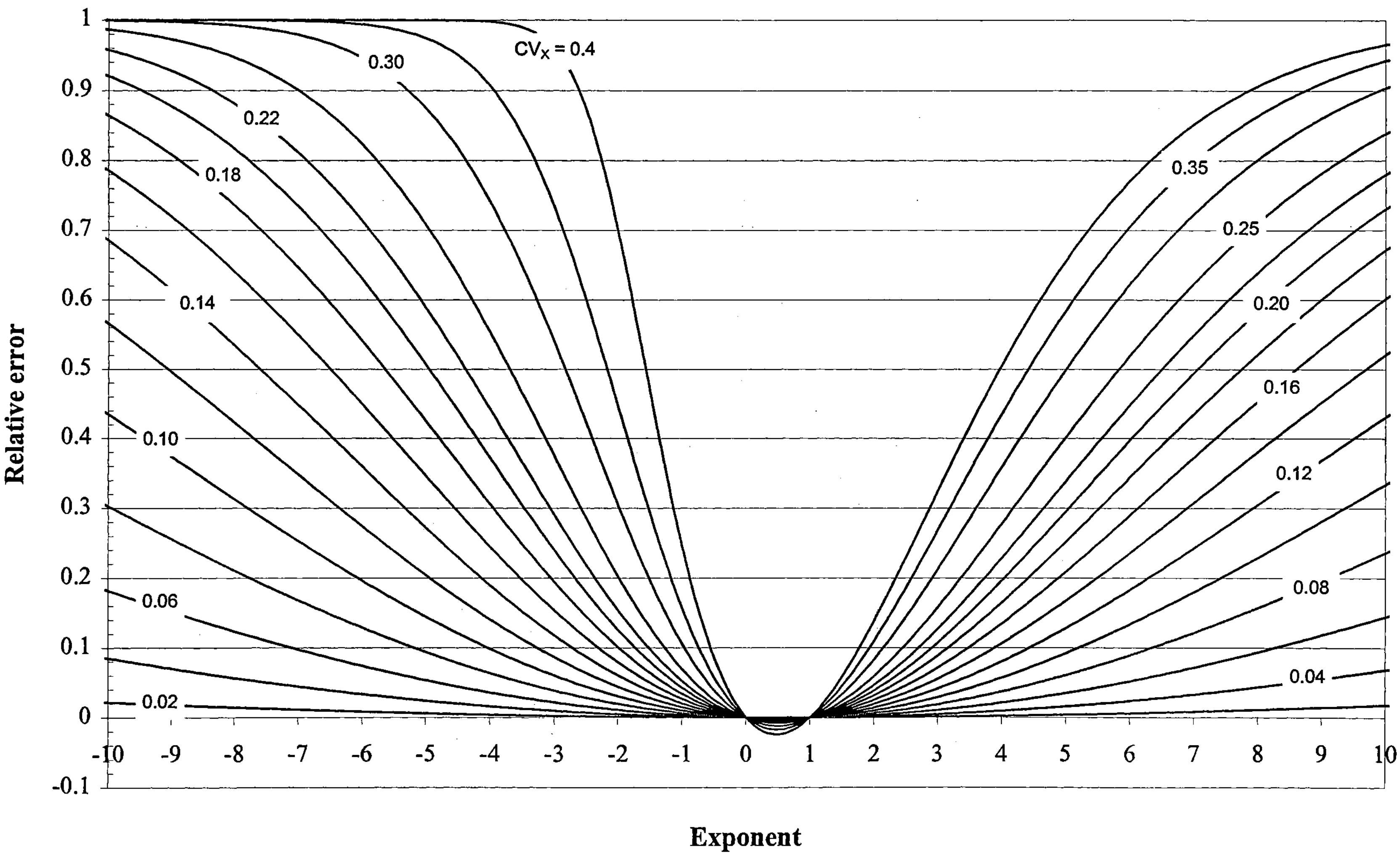

Fig.2-3: Relative error in FOA predicted mean for CV_X ranging from 0.01 to 0.4, where X is triangularly distributed

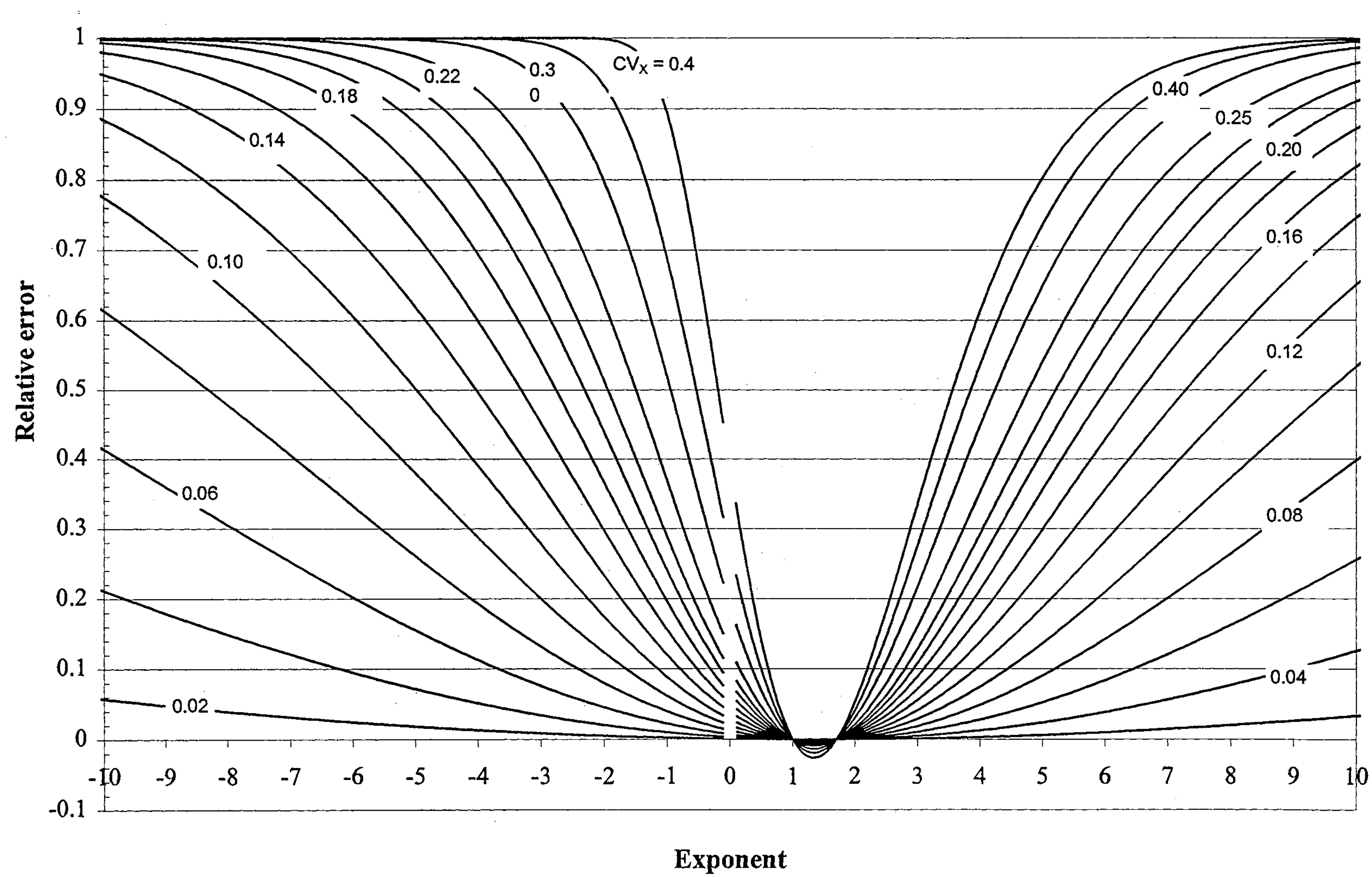

Fig. 2-4: Relative error in FOA predicted variance for CV_X ranging from 0.01 to 0.40, where X is triangularly distribution

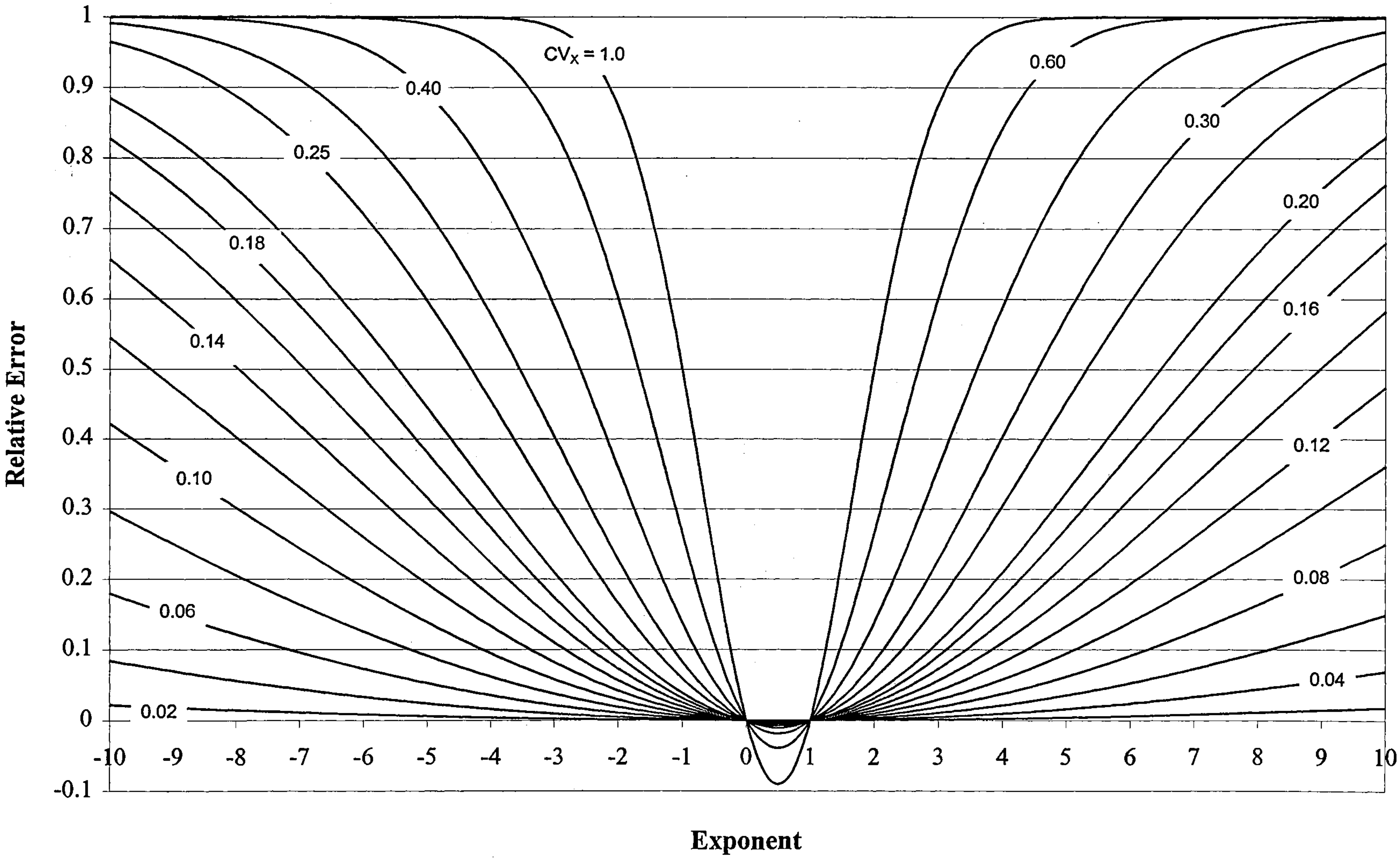

Fig.2-5: Relative error in FOA predicted mean for CV_X ranging from 0.01 to 1.0, where X is lognormally distributed

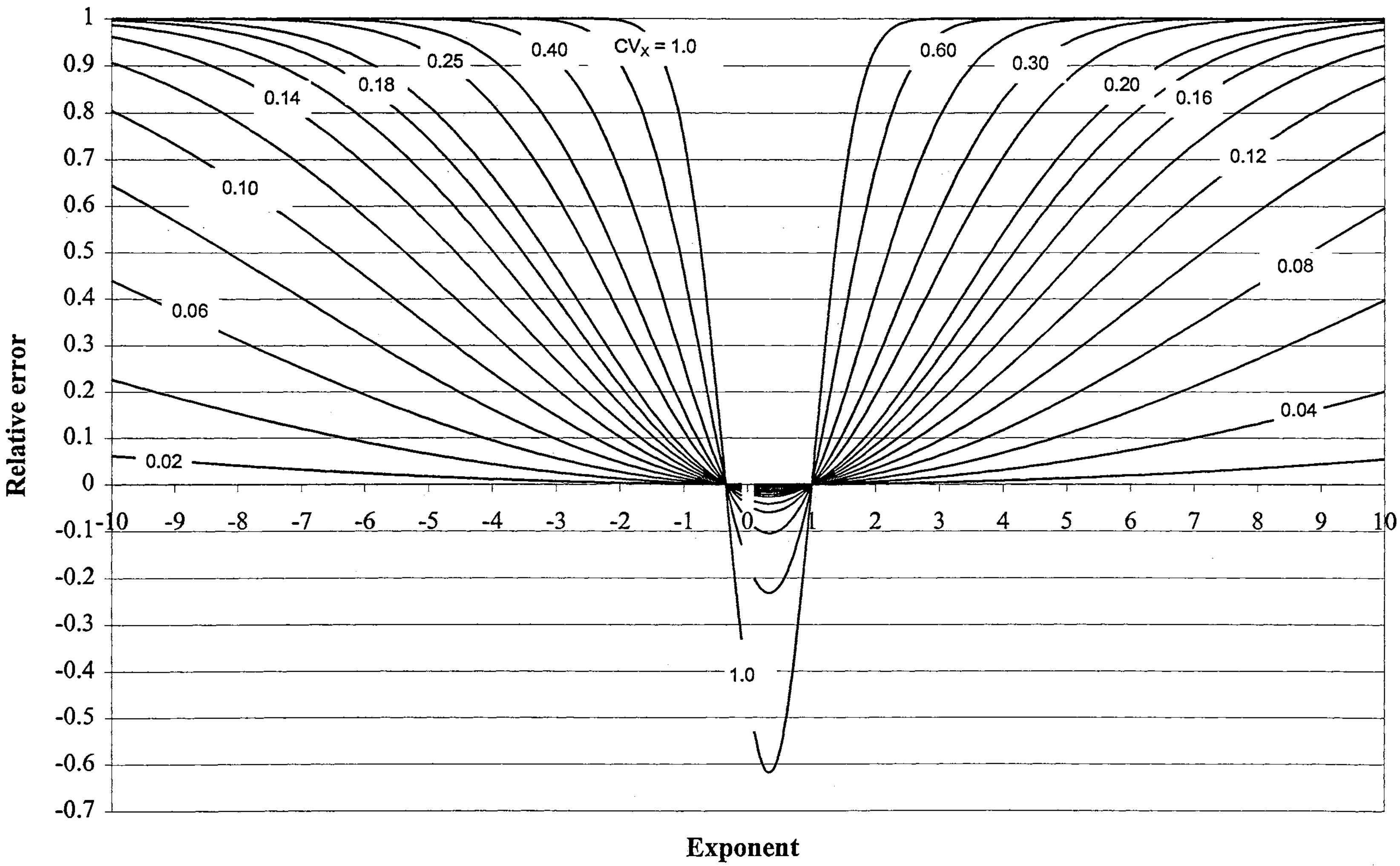

Fig.2-6: Relative error in FOA predicted variance for CV_X ranging from 0.01 to 1.0, where X is lognormally distributed

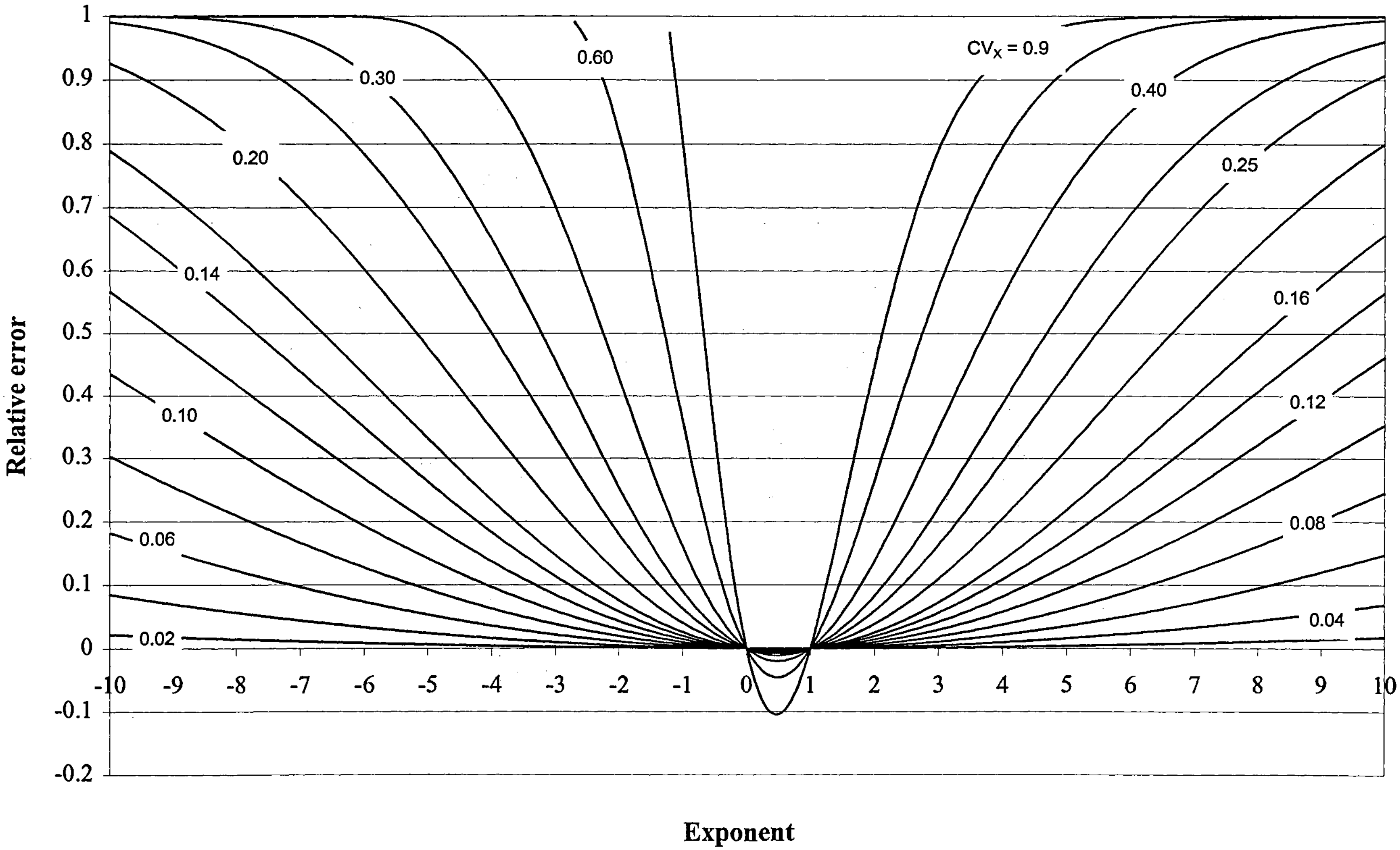

Fig.2-7: Relative error in FOA predicted mean for CV_X ranging from 0.01 to 0.90, where X is defined by the gamma distribution

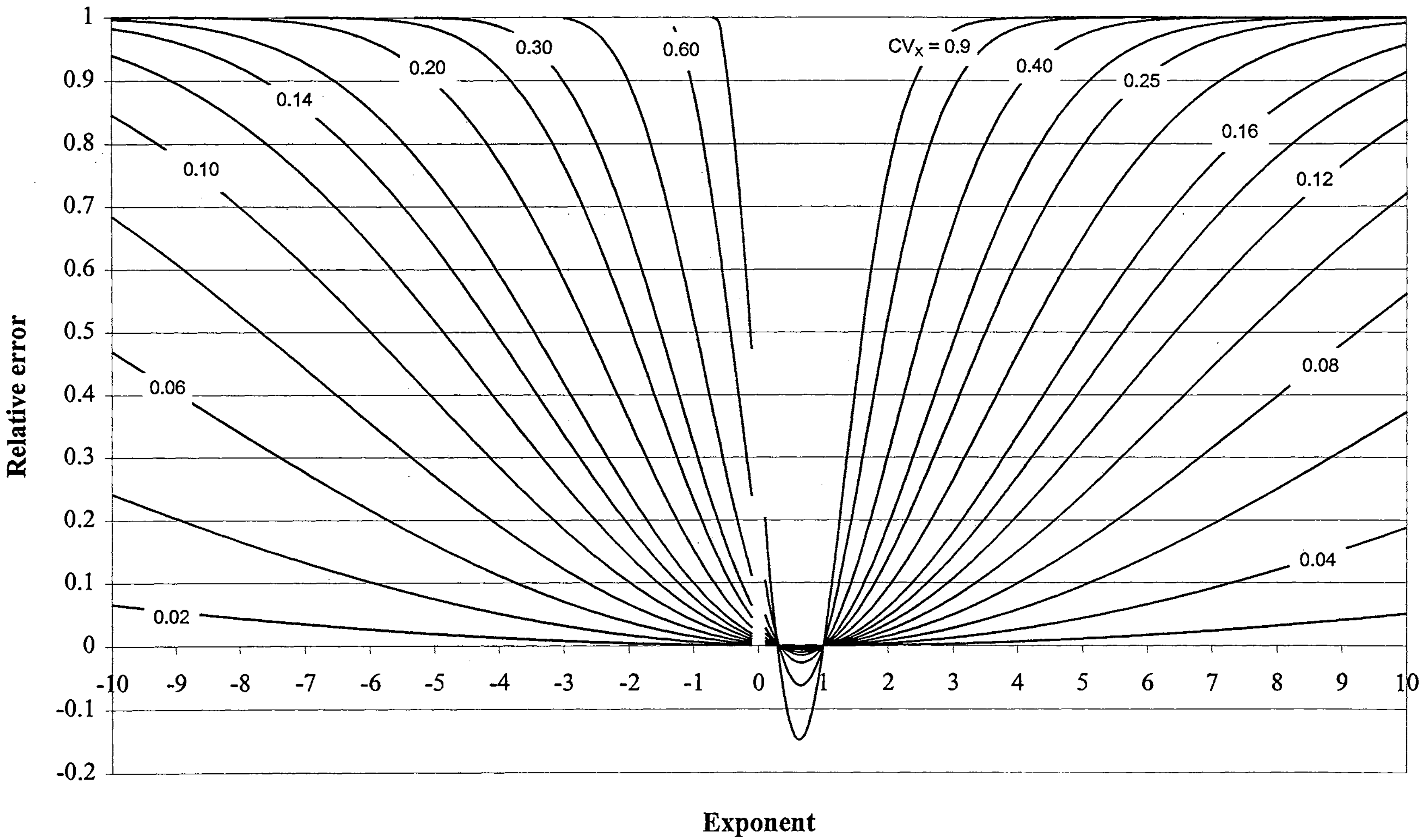

Fig. 2-8 : Relative error in FOA predicted variance for CV_X ranging from 0.01to 0.90, where X is defined by the gamma distribution

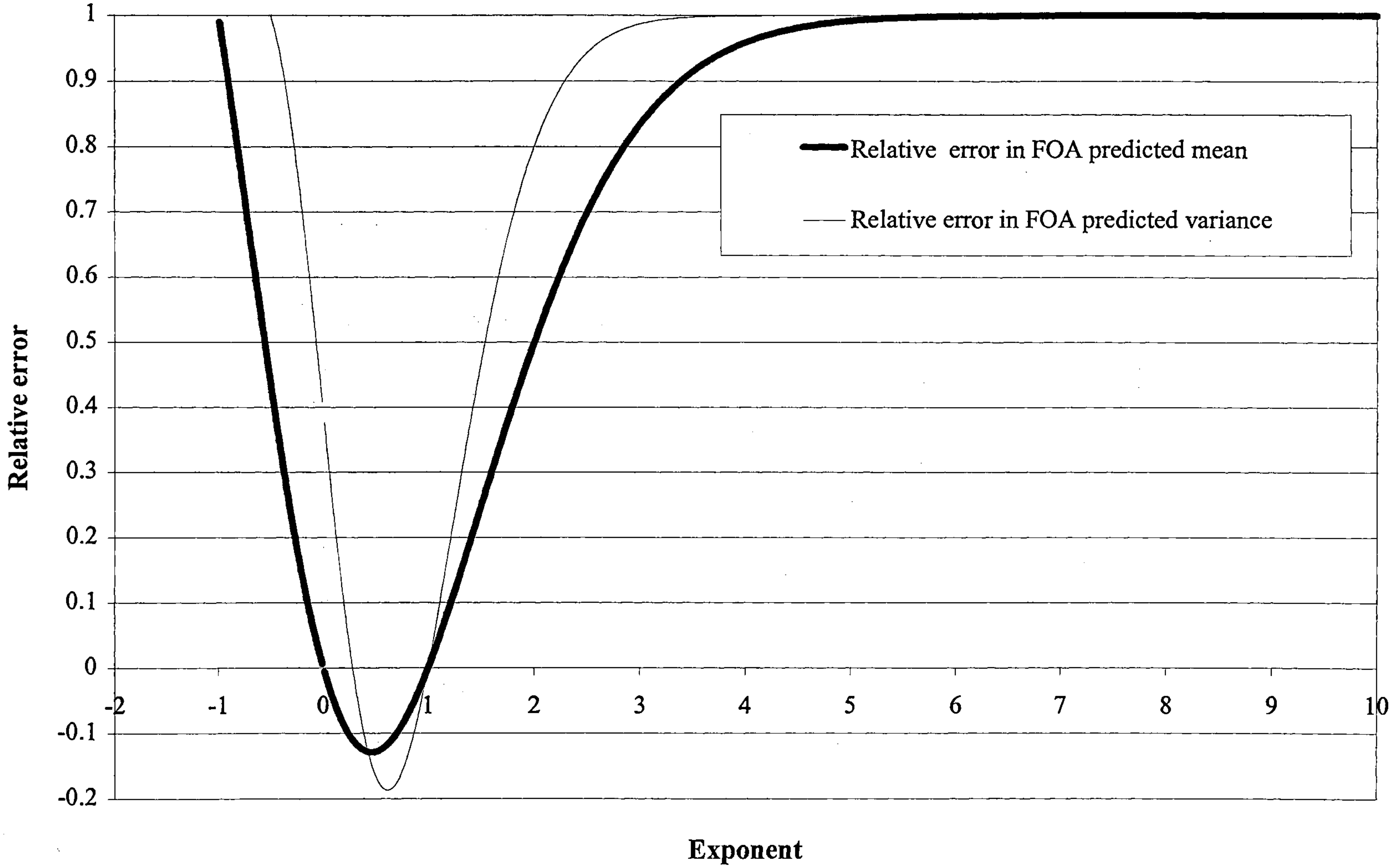

Fig. 2-9: Relative errors for FOA predicted mean and variance, where X is from the exponential distribution

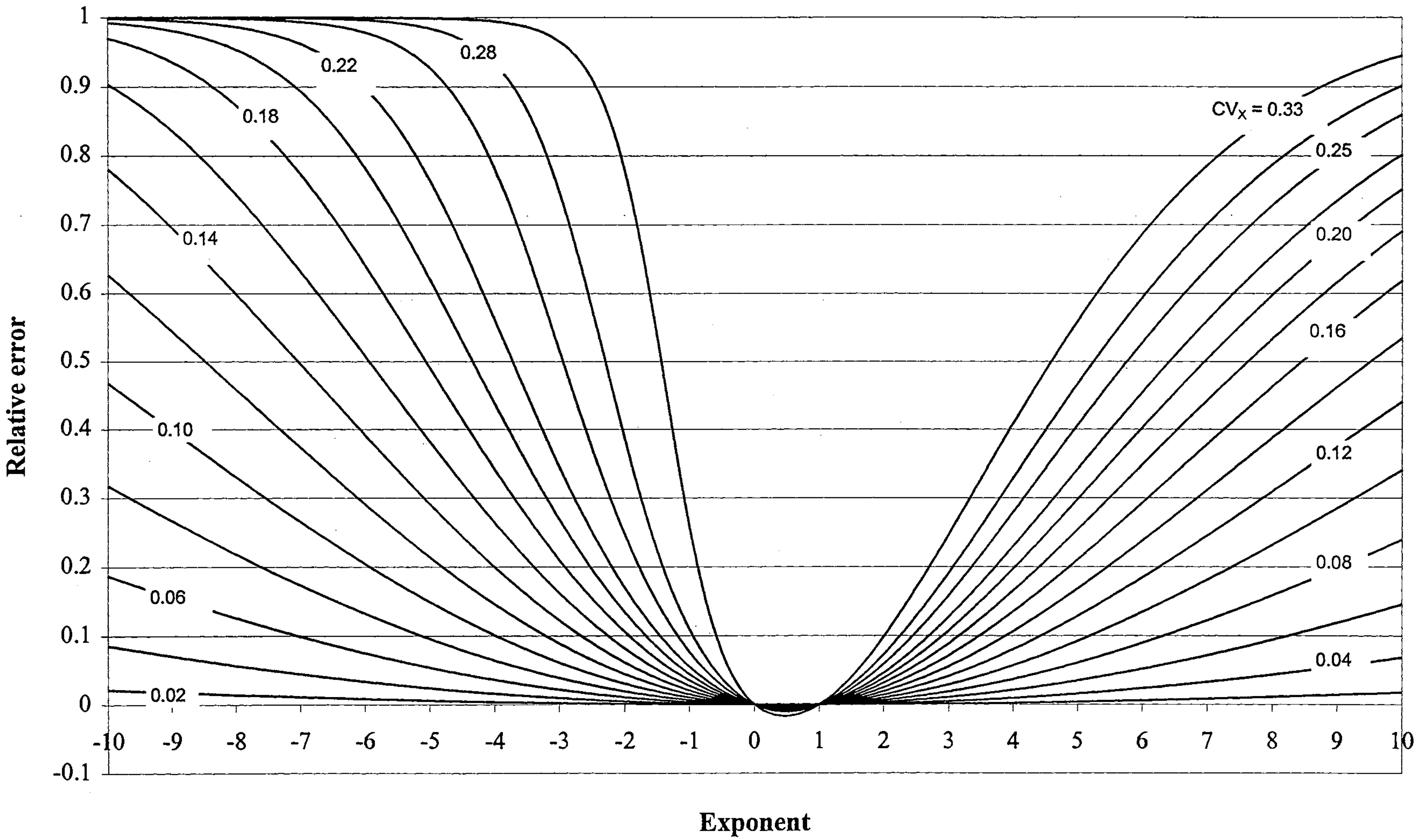

Fig. 2-10: Relative error in FOA predicted mean for CV_X ranging from 0.01 to 0.33, where X is normally distributed

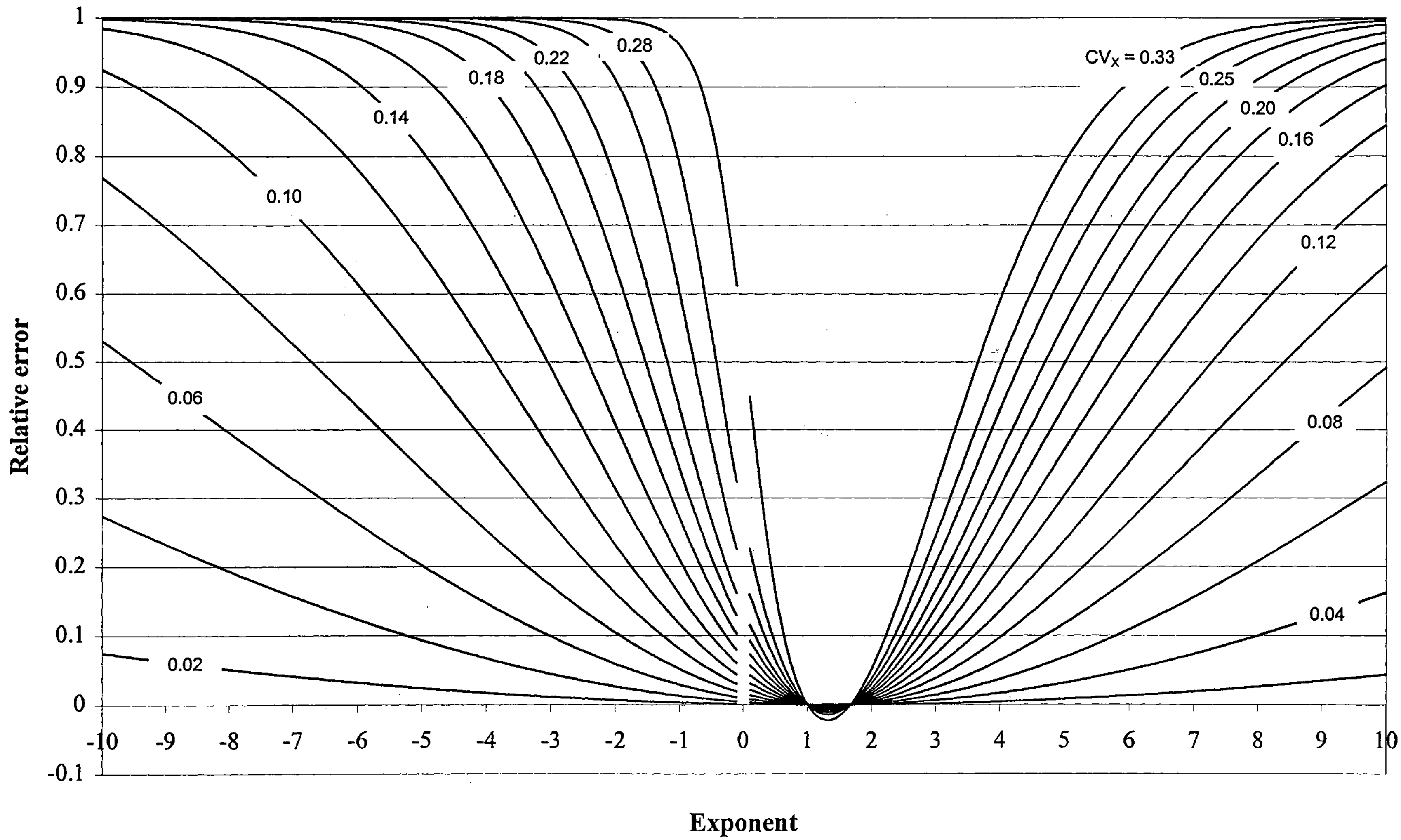

Fig. 2-11: Relative error in FOA predicted variance for CV_X ranging from 0.01 to 0.33, where X is normally distributed

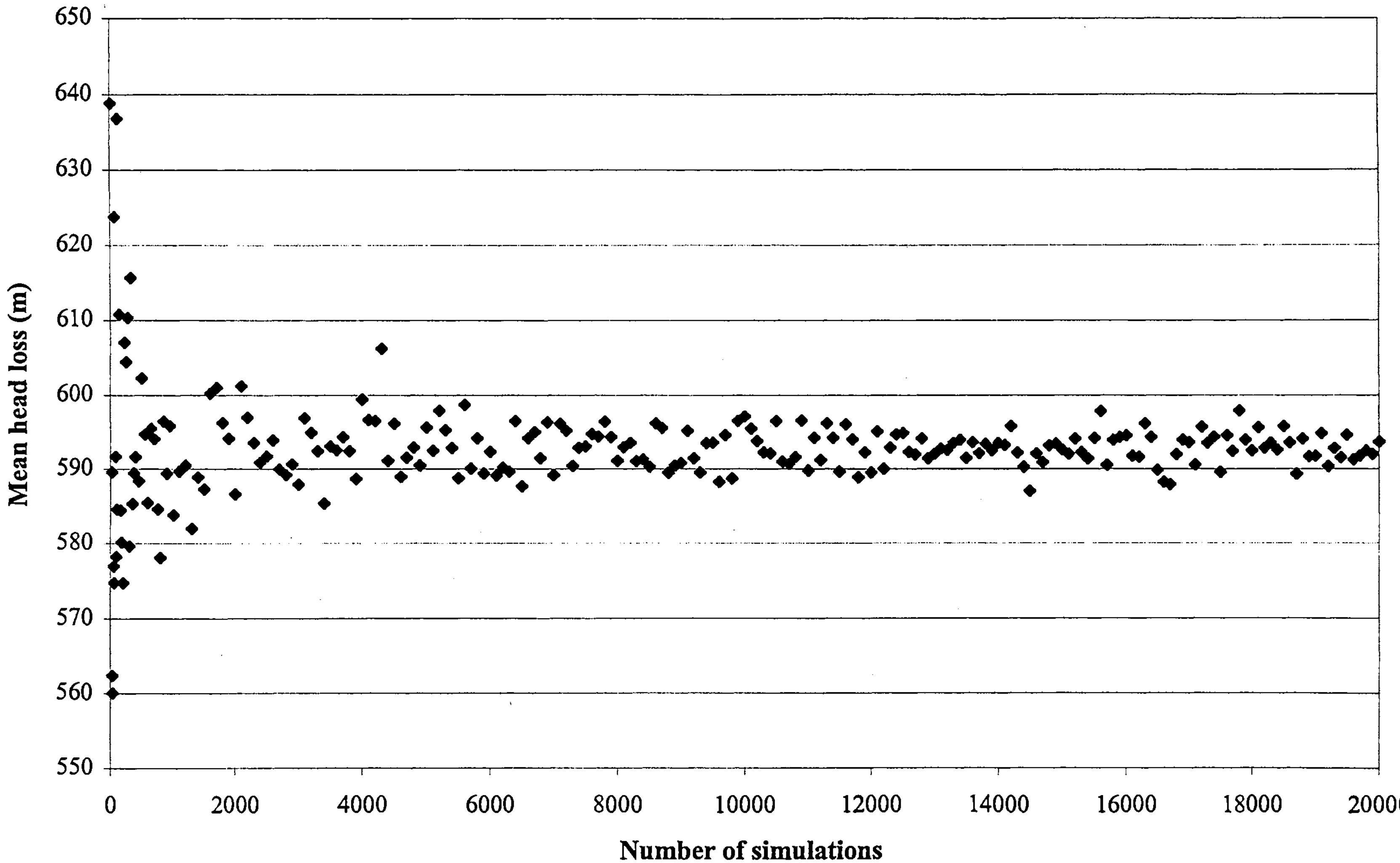

Fig. 2-12a: Variation of the mean with number of simulations for Example # 1

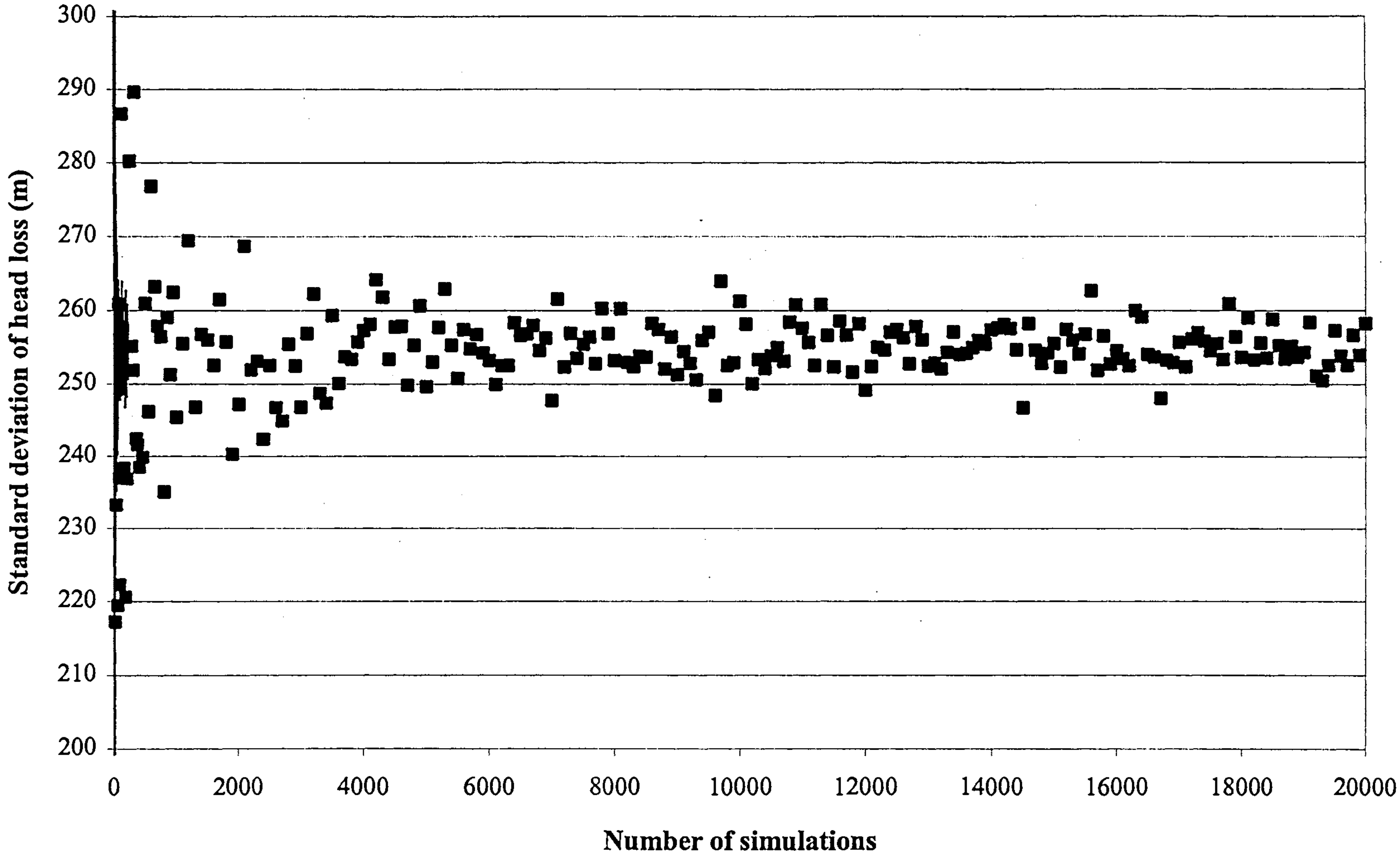

Fig. 2-12b: Variation of the standard deviation Vs number of simulations for example # 1

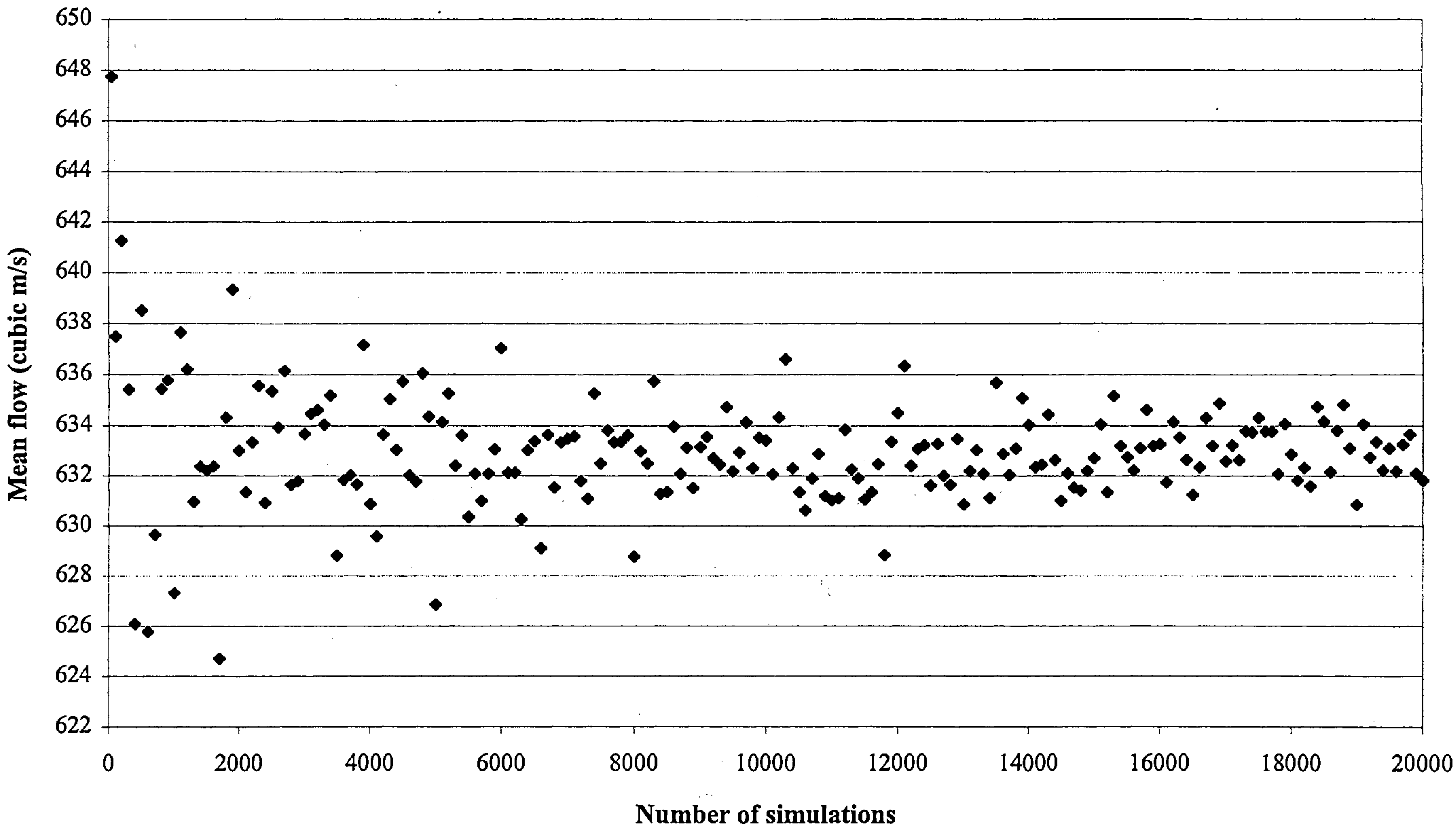

Fig. 2-13a :Variation of the mean with number of simulations for Example # 2

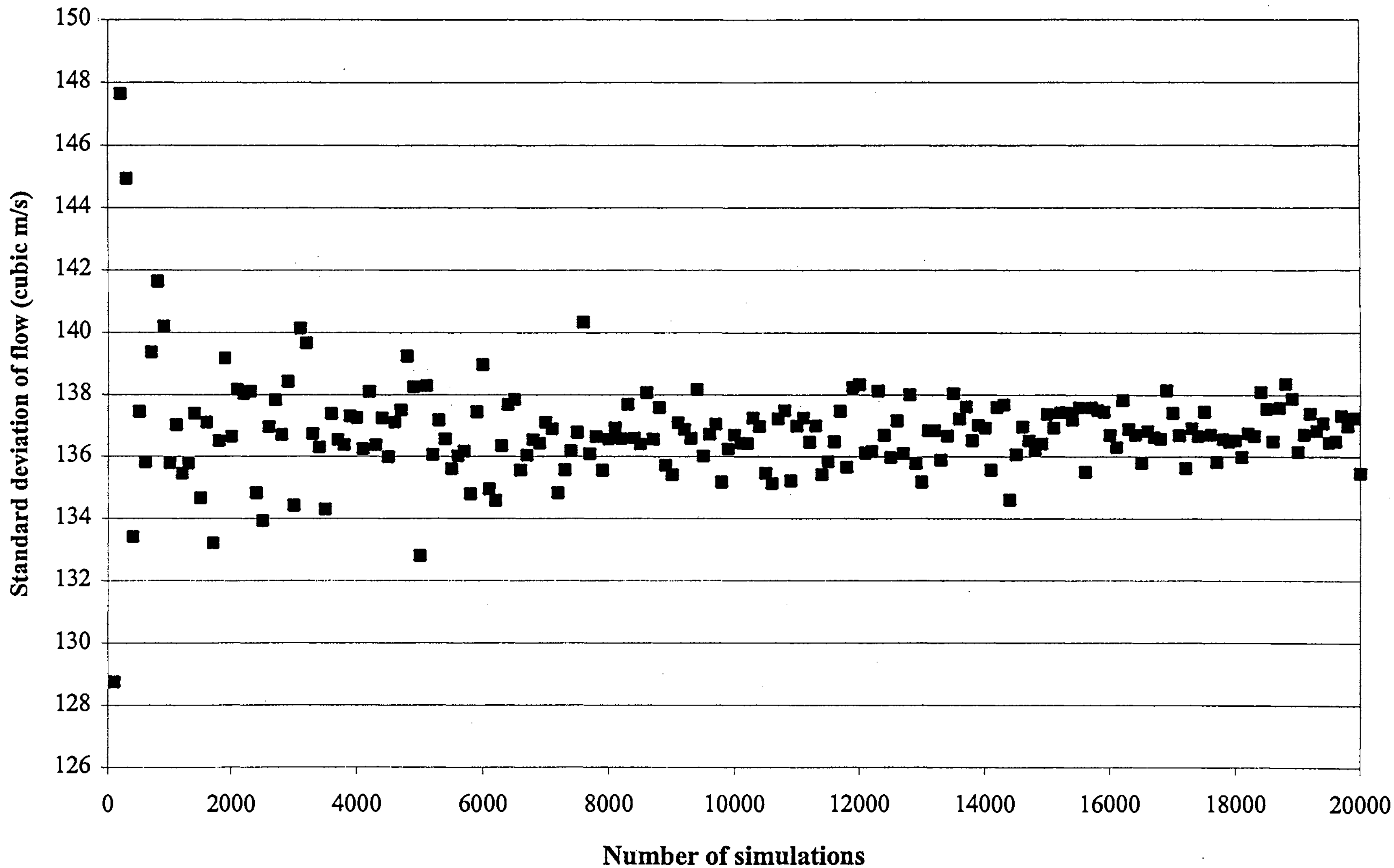

Fig. 2-13b : Variation of the standard deviation vs number of simulations for example # 2

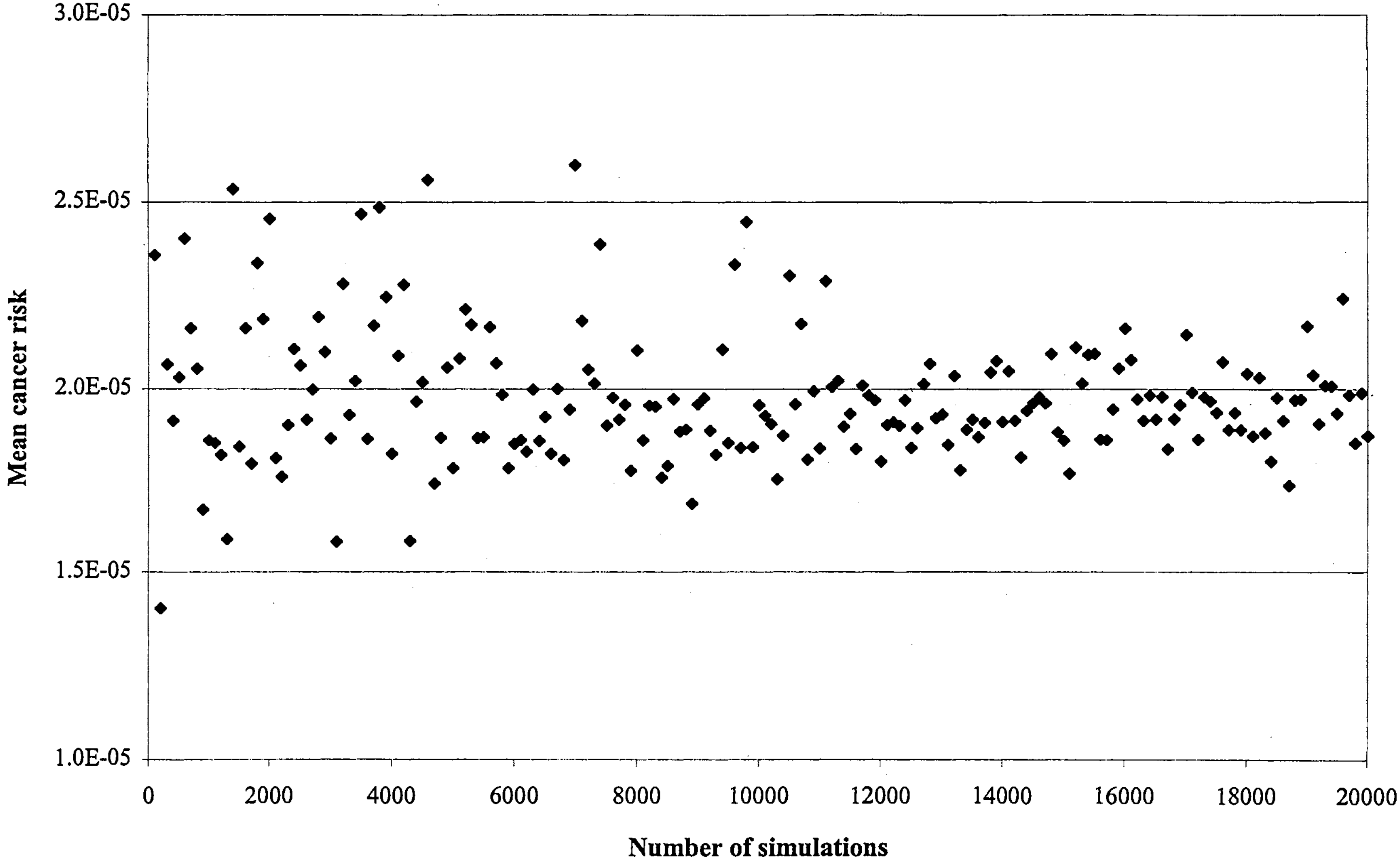

Fig. 2-14a : Plot of the mean value of risk (Example # 3)

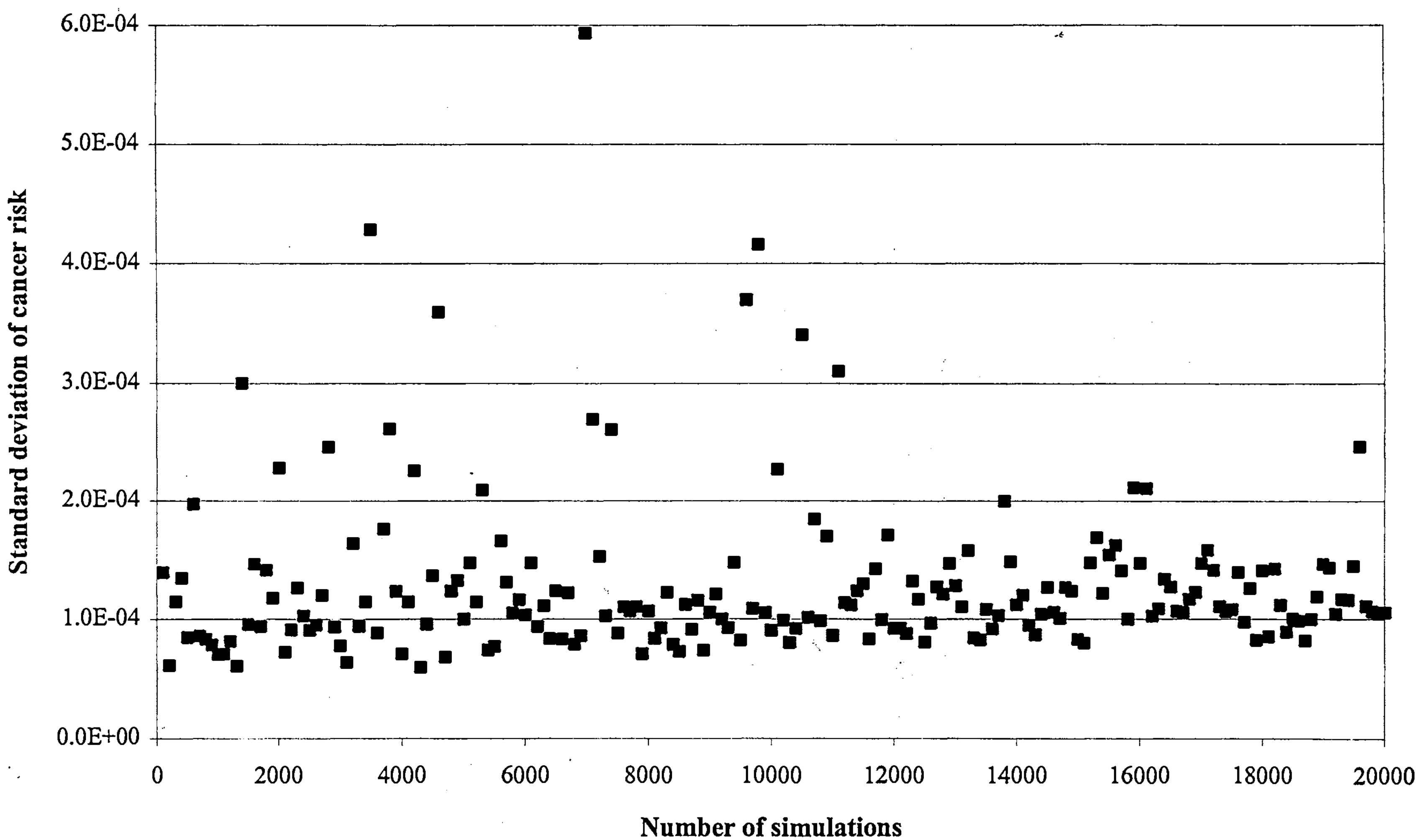

Fig. 2-14b : Plot of the standard deviation of risk (Example # 3)

CHAPTER III

RELIABILITY, RISK, AND UNCERTAINTY ANALYSIS USING GENERIC EXPECTATION FUNCTIONS

Abstract

In engineering design and analysis mathematical models are frequently employed for decision making which generally involve a number of uncertain parameters. Over the years, a number of techniques have been developed to quantify model output uncertainty contributed by uncertain input parameters. Typically the methods which are easy to apply may give inaccurate estimates of model output uncertainty. Other methods which reliably produce very accurate results are either difficult to apply or require intensive computational effort. This paper describes the development of generic expectation functions as a function of means and CVs of input random variables. The generic expectation functions are easy to develop and simple to apply to problems related to reliability, risk, and uncertainty analysis. Several expectation functions based on commonly used probability distributions have been developed. The developed expectation functions are general. Using them any order of moment can be estimated exactly. It is found that if exact moments of model output are available, one can find a good estimate of reliability, risk and uncertainty of a system without knowing its model

output distribution. Two practical examples are presented to demonstrate the application of generic expectation functions.

Introduction

In engineering design and analysis, very often models are employed. These models generally involve a number of uncertain parameters, which are determined with varying degrees of accuracy. These parameters are best represented as random variables. Consequently, model outputs on which engineering design and analysis are based are also uncertain and should be represented as random variables. As a result of uncertainty in model response, the performance of a project designed based on the model will be uncertain as well. To incorporate uncertainty in the decision-making and design process quantification of uncertainty is required. Many problems related to hydrology, hydraulics, and environmental engineering are best approached using uncertainty and reliability methods. Reliability and uncertainty analyses are becoming mandatory, particularly where critical decisions involving potentially high adverse consequences are made.

Two major types of uncertainties in the field of water resources (Tung and Mays, 1980) are model uncertainty and parameter uncertainty. Model uncertainties arise through simplifying assumptions used to derive simple mathematical relationships between the inputs and outputs in describing a complex process. Whereas, parameter uncertainty, represented by the coefficient of variation (CV), arises because of inherent natural variability, measurement limitations, and lack of sufficient data.

Two most commonly used methods for reliability, risk, and uncertainty analyses are the Monte Carlo simulation (MCS) and the first-order approximation (Benjamin and Cornell, 1970) known as FOA. Both methods have some limitations. The MCS is computationally intensive with the number of simulations required for convergence not well defined (Melching, 1995). In most engineering problems, the true probability distributions of the input variables are seldom known. Theoretical distributions for the input variables are assumed to conduct the MCS. The quality of MCS estimates is affected by appropriateness of the chosen distribution functions for the input variables and the number of simulations used in the analyses (Bates and Townley, 1988). FOA is very computationally efficient but provides approximate model output estimates for the mean and variance only. The quality of these estimates is influenced by the CVs of input variables and non-linearity in the model (Burn and McBean, 1985; Tung, 1990). Further, FOA does not provide the form of the output distribution. Based on the central limit theorem output is assumed to be normally distributed when confidence limits on the output, risk, and reliability of the system are determined.

In reliability and risk analyses, one is concerned with system failure. Often, failures of engineering projects occur at extreme values (rather than near the mean values) of the input variables. Extremes are most likely associated with probability distributions having large variance and skewness (Yen et al., 1986). FOA uses expansion about the mean values of the input variables indicating that any attempt to characterize the tails of the output distribution is likely to result in an inexact estimate (Burn and McBean, 1985). Furthermore, using FOA, it is not possible to incorporate the information about the forms of input variable distributions, if they were known (Yen et al., 1986).

Hasofer and Lind (1974) showed that flaws in FOA due to model non-linearity can be removed by linearizing the functional relationship at the point on the limit-state surface nearest to the origin, rather than at the mean point. Calculation of the linearization point requires determination of the nearest point on the limit-state surface. This method involves an assumption that all the input variables are normally distributed, giving a normally distributed output. However, in most real modeling problems, all the basic variables are not normally distributed. Rackwitz (1976) proposed a transformation procedure in which the values of non-normal distribution are the same as those of the equivalent normal distributions at the failure point. This method is known as advanced first-order second-moment method (AFOSM). The AFOSM is widely used in reliability and risk analyses (e.g., Melching et al., 1991; Sitar et al., 1987; Cawlfield and Wu, 1993; Mishra, 1998; Cesare, 1991). The AFOSM can also be used to carry out uncertainty analysis by repeating the procedure of calculating the linearization point to match the pre-specified output value whose exceedance probability is sought. Examples of using AFOSM in uncertainty analyses are Melching and Anmangandla (1992) in water quality modeling and Mishra (1998) in environmental probabilistic risk assessment. The AFOSM is very accurate because using it one is able to overcome model non-linearity problems and can utilize the available information about the input variable distributions, without having to make any additional assumptions. The disadvantage of the AFOSM is that determination of the lineraziation point is generally not easy depending upon the nature and complexity of the system for which the reliability, risk, or uncertainty analysis is being studied (Melching and Anmangandla 1992).

Rosenblueth (1975, 1981) proposed the point estimation (PE) method to evaluate uncertainty at specified points in the parameter space. To estimate the statistical moments of a model output, 2^n model evaluations are required for a model involving n uncertain parameters. As the number of stochastic parameters increases, the computation requirement of Rosenblueth's algorithm becomes similar to that of MCS method (Melching, 1995). An alternative computationally efficient PE method was proposed by Harr (1989) by utilizing the first two moments of the random variables. Chang et al. (1995) showed that the estimated uncertainty feature of model output could be inaccurate if the skewness of a random variable is not accounted for.

In many cases, the true form of the output distribution is not required. A very good estimate of system reliability can be obtained if moments of model output are known correctly. As far as the distribution of model output is concerned, several forms of distributions can be assumed. The knowledge of the higher-order moments of a model output helps in identifying the candidate distributions for the model output and provides more flexibility to include those distribution forms, which require higher order moments. Tung (1990) used the Mellin transform to calculate the higher-order moments of a model output. The application of the Mellin transform is not only cumbersome but also it can not be universally applied. As pointed out by Tung, the Mellin transform may not be analytic under certain combinations of distribution and functional forms. In particular, problems may arise when a functional relationship consists of input variable(s) with negative exponent(s). Further, no formulation was suggested to obtain the moments of a model output having non-standard normally distributed input variable(s).

This paper describes the development of generic expectation functions as a function of means and CVs of input random variables. These functions are easy to apply in any general application. Further, a procedure has been suggested to apply the developed expectation functions to reliability, risk, and uncertainty analyses. Two examples are presented to demonstrate the application of generic expectation functions.

Uncertainty and Reliability Analyses

In most hydrologic and hydraulic engineering problems, empirically developed or theoretically derived mathematical equations are used which involve several uncertain parameters that are difficult to quantify accurately. Further, a mathematical equation, $g(\underline{X})$, may have different degrees of nonlinearity with respect to its uncertain parameters represented as an array $\underline{X}$. The term nonlinearity is difficult to define and no well accepted definition is available. A multitude of functional forms for $g(\underline{X})$ is possible. In this paper a multiplicative form is considered.

A multiplicative type model is frequently encountered in hydrological studies (e.g., daily stream flow, peak runoff, annual floods, and annual, monthly, and daily rainfall, soil loss and sediment transport). In hydraulics many equations are of multiplicative type. Examples are flow over control structures such as weirs, spillways, overfalls, and sluices (Haan et al., 1994), channel control equations such as Manning's equation (Haan et al., 1994), pipe flow resistance equations such as Hazen-Williams and Darcy-Weisbach equations. In environmental engineering, many equations predicting water quality and pollution (Krenkel, 1979; Novotny and Olem, 1994), and risk (USEPA, 1989) are of multiplicative type. Tung and Mays (1980), Lee and Mays (1986), and Tung (1990) are some of the examples of uncertainty analysis of multiplicative forms encountered in hydraulic/hydrologic systems. In this form, the output random variable Y is expressed as the multiplication of n power functions.

$$Y = C_0 \prod_{i=1}^{i=n} X_i^{r_i} \qquad (3\text{-}1)$$

where C_0 and r_i are constants, and X_is are n independent stochastic input random variables. The k^{th} moment of Y about the origin, μ'_k, is defined as (Haan, 1977)

$$\mu'_k = E\left[Y^k\right] = C_0^k \prod_{i=1}^{n} E\left[X_i^{kr_i}\right] \tag{3-2}$$

where $E[\ ]$ is an expectation operator. The k^{th}-central moment of Y, μ_k, can be obtained using the following equation (Haan, 1977)

$$\mu_k = E\left[(Y-\mu_Y)^k\right] = \sum_{i=0}^{k} (-1)^i \binom{k}{i} \mu_Y^i \mu'_{k-i} \tag{3-3}$$

where, $\mu_Y = \mu'_1$= mean of Y. Substituting k =1 in (3-2), μ_Y is given as

$$\mu_Y = E[Y] = C_0 \prod_{i=1}^{n} E\left[X_i^{r_i}\right] \tag{3-4}$$

Substituting μ'_{k-i} from (3-2) and μ_Y from (3-4) in (3-3), μ_k can be expressed as

$$\mu_k = C_0^k \sum_{i=0}^{k} (-1)^i \binom{k}{i} \left\{\prod_{i=1}^{n} E\left[X_i^{r_i}\right]\right\}^i \prod_{i=1}^{n} E\left[X_i^{(k-i)r_i}\right] \tag{3-5}$$

Eqs. (3-2) and (3-5) show that moments of Y of any order k about the mean and the origin can be obtained if expectation of individual power functions is known.

In most situations distributional properties of a random variable are characterized in terms of its mean, variance, coefficient of skewness, and coefficient of kurtosis. The variance of Y, σ^2_Y, is defined as the second moment about the mean. Substituting $k = 2$ in (3-3), σ^2_Y is given as

$$\mu_2 = \sigma^2 = E\left[Y^2\right] - \mu_Y^2 \tag{3-6}$$

where μ_2 is the second moment of Y about the mean. The coefficient of skewness of Y, γ_Y, is defined as (Haan, 1977)

$$\gamma_Y = \frac{\mu_3}{\mu_2^{3/2}} \tag{3-7}$$

where μ_3 is the third moment of Y about the mean which can be obtained by substituting $k = 3$ in (3-3) as

$$\mu_3 = E\left[Y^3\right] - 3\mu_Y E\left[Y^2\right] + 2\mu_Y^3 \tag{3-8}$$

The kurtosis of Y, κ_Y, is defined as (Haan, 1977)

$$\kappa_Y = \frac{\mu_4}{\mu_2^2} \tag{3-9}$$

where μ_4 is the fourth moment of Y about the mean which can be obtained by substituting $k = 4$ in (3-3) as

$$\mu_4 = E\left[Y^4\right] - 4\mu E\left[Y^3\right] + 6\mu^2 E\left[Y^2\right] - 3\mu^4 \tag{3-10}$$

The reliability of a system can be more realistically measured in terms of probability. The failure of a system can be considered as an event that the demand or loading, L, on the system exceeds the capacity or resistance, R, of the system so that the system fails to perform satisfactorily for its intended use. The objective of reliability analysis is to ensure the probability of event ($R > L$) throughout the specified useful life is acceptably small. To study this event, a performance function, Z, is defined as (Ang and Tang, 1984; Mays and Tung, 1992; Tung, 1990)

$$Z = R - L \tag{3-11}$$

The risk is defined as the probability of failure of the system, which can be written as

$$P_f = P(Z < 0) = \int_{-\infty}^{0} p_Z(z)dZ \tag{3-12}$$

where P_f is the probability of failure, P is the probability operator, and $p_Z(z)$ is the probability density function of Z. The reliability of the system can be written as

Reliability $= P(Z > 0) = 1 - P_f$ (3-13)

The probability distribution of Z is unknown, or difficult to obtain. In most cases the exact distribution may not be required, as several distributions can be used to make a decision if correct information about its moments is available. Further, higher order moments are helpful in both identifying the candidate distributions for $p_Z(z)$ and using the distributions requiring higher order moments.

In most cases both R and L can be represented as a multiplicative form as (3-1). To characterize the failure event ($R<0$), it is necessary to define the random variable Z statistically, i.e., its various moments and distribution. The statistical moments of Z about the origin can be expressed in terms of moments of R and L as

$$E[Z] = E[R] - E[L] \quad (3\text{-}14)$$

$$E[Z^2] = E[R^2] - 2E[R]E[L] + E[L^2] \quad (3\text{-}15)$$

$$E[Z^3] = E[R^3] - 3E[R^2]E[L] + 3E[R]E[L^2] - E[L^3] \quad (3\text{-}16)$$

$$E[Z^4] = E[R^4] - 4E[R^3]E[L] + 6E[R^2]E[L^2] - 4E[R]E[L^3] + E[L^4] \quad (3\text{-}17)$$

As clear from (3-14) to (3-17), the moment of Z about the origin can be evaluated once moments of R and L are determined. Using these moments about the origin one can easily determine the central moment of Z. As mentioned earlier determining the true probability distribution of Z is difficult if not impossible. For calculating risk, several distributions can be selected based on the higher order moments. Using (3-12) and the selected distribution of Z, the risk can be estimated. Yen et al. (1986) have derived risk formulas for some selected probability distributions.

Development of Generic Expectation Functions

Consider a power function

$$Y = X^r \tag{3-18}$$

The k^{th} order moment of Y about the origin can be obtained as

$$\mu_k' = E[Y^k] = E[X^{kr}] = \int_{-\infty}^{\infty} X^{kr} p_X(x) dX \tag{3-19}$$

where $p_X(x)$ = probability density function of X.

Uniform Distribution

The probability density function for the continuous uniform distribution is

$$p_X(x) = \frac{1}{b-a}, \quad a \leq X \leq b \tag{3-20}$$

where a and b are the distribution parameters. The methods of moments estimates for a and b are given as (Haan, 1977)

$$\hat{a} = \mu_X - \sqrt{3}\sigma_X = \mu_X\left(1 - \sqrt{3}CV_X\right) \tag{3-21}$$

$$\hat{b} = \mu_X + \sqrt{3}\sigma_X = \mu_X\left(1 + \sqrt{3}CV_X\right) \tag{3-22}$$

where CV_X = coefficient of variation of X, defined as

$$CV_X = \frac{\sigma_X}{\mu_X} \tag{3-23}$$

Using Eqs. (3-19), (3-20), (3-21), (3-22) and (3-23) the $E[X^r]$ is given as

$$E[X^r] = \int_a^b \frac{1}{(b-a)} X^r dX = \frac{\mu_X^r}{2\sqrt{3}(r+1)CV_X}\left[\left(1 + CV_X\sqrt{3}\right)^{r+1} - \left(1 - CV_X\sqrt{3}\right)^{r+1}\right] \tag{3-24}$$

Triangular Distribution

The probability density function $p_X(x)$ for the triangular distribution is

$$p_X(x)=\frac{2}{b-a}\frac{(X-a)}{(c-a)},\ \text{when } a\le X\le c \tag{3-25a}$$

$$p_X(x)=\frac{2}{b-a}\frac{(b-X)}{(b-c)},\ \text{when } c\le X\le b \tag{3-25b}$$

where *a, b, c* are the minimum, maximum, and mode values of *X*. These parameters can be obtained by the following equation (Appendix I)

$$\bar{a}=\mu_X\left\{1+2\sqrt{2}CV_X\cos\left[\frac{2\pi n}{3}+\frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_x\right)\right]\right\} \tag{3-26}$$

where $\bar{a}$ = a vector containing *b, a,* and *c* which can be obtained by substituting n = 0, 1, and 2, respectively, in (3-26); and γ_X is the coefficient of skew of *X*. Using (3-19), (3-25a) and (3-25b) the $E[X^r]$ is given as

$$E\left[X^r\right]=\frac{2\left[(b-c)a^{r+2}+(c-a)b^{r+2}+(a-b)c^{r+2}\right]}{(r+1)(r+2)(b-c)(c-a)(b-a)} \tag{3-27}$$

For symmetrical triangle γ_X = 0 and the parameters *a, b,* and *c* can be obtained corresponding to n = 1, 0, and 2. The obtained *c* is the μ_X and the parameters *a* and *b* are the same as obtained using the methods of moments. The estimates of a and b are given as

$$\hat{a}=\mu_X\left(1-\sqrt{6}CV_X\right) \tag{3-28}$$

$$\hat{b}=\mu_X\left(1+\sqrt{6}CV_X\right) \tag{3-29}$$

Using (3-19), (3-25a), (3-25b), (3-28) and (3-29) the $E[X^r]$ is given as

$$E\left[X^r\right]=\frac{\mu_X^r}{6(r+1)(r+2)CV_X^2}\left[\left(1+CV_X\sqrt{6}\right)^{r+2}+\left(1-CV_X\sqrt{6}\right)^{r+2}-2\right] \tag{3-30}$$

<u>Lognormal Distribution</u>

If X is lognormally distributed with mean μ_X and variance σ^2_X, its probability density function is given (Haan, 1977) as

$$p_X(x)=\frac{1}{\sigma_V X\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln X-\mu_V}{\sigma_V}\right)^2}, X>0 \tag{3-31}$$

where $V = ln(X)$ is normally distributed with parameters μ_V and σ^2_V. The parameters μ_V and σ^2_V are defined (Haan, 1977) as

$$\mu_V=\frac{1}{2}\ln\left[\frac{\mu_X^2}{CV_X^2+1}\right] \tag{3-32}$$

$$\sigma_V^2=\ln\left(CV_X^2+1\right) \tag{3-33}$$

Substituting (3-31) in (3-19), the $E[X^r]$ is given as

$$E\left[X^r\right]=\frac{1}{\sigma_V\sqrt{2\pi}}\int_0^\infty X^{r-1}e^{-\frac{1}{2}\left(\frac{\ln X-\mu_V}{\sigma_V}\right)^2}dx \tag{3-34}$$

Assuming, $\frac{\ln(X)-\mu_V}{\sigma_V}=z$, the random variable X can be written as, $X=e^{(\mu_V+\sigma_V z)}$, (3-34) is rewritten as

$$E\left[X^r\right]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{r(\mu_V+\sigma_V z)}e^{-\frac{1}{2}z^2}dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{\frac{1}{2}\left(2\mu_V r+\sigma_V^2 r^2\right)}e^{-\frac{1}{2}(z-r\sigma_V)^2}dz \tag{3-35}$$

But, $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty e^{-\frac{1}{2}(z-r\sigma_V)^2}dz=1.0$. Therefore, (3-35) can be written as

$$E\left[X^r\right]=e^{\left(\mu_V r+\frac{1}{2}\sigma_V^2 r^2\right)} \tag{3-36}$$

Gamma Distribution

The gamma density function is given by

$$p_X(x) = \frac{\lambda^\alpha e^{-\lambda x} X^{(\alpha-1)}}{\Gamma(\alpha)}, \; X,\ \alpha,\ and\ \lambda > 0 \tag{3-37}$$

where α and λ are the distribution parameters. Using the method of moments, α and λ are expressed (Haan, 1977) as

$$\hat{\lambda} = \frac{\mu_X}{\sigma_X^2} \tag{3-38}$$

$$\hat{\alpha} = \frac{\mu_X^2}{\sigma_X^2} = \frac{1}{CV_X^2} \tag{3-39}$$

Substituting (3-37) in (3-19), the $E[X^r]$ is written as

$$E\left[X^r\right] = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \frac{e^{-\lambda X} X^{(\alpha+r-1)}}{\Gamma(\alpha)} dX = \frac{\Gamma(\alpha+r)}{\lambda^r \Gamma(\alpha)} \tag{3-40}$$

Replacing α and λ in (3-40) by their estimates given in (3-38) and (3-39), (3-40) is rewritten as

$$E\left[X^r\right] = \frac{CV_X^{2r} \mu_X^r \Gamma\left(CV_X^{-2} + r\right)}{\Gamma(CV_X^{-2})} \tag{3-41}$$

Exponential Distribution

The exponential distribution is a special case of the gamma distribution with α = 1 and $\lambda = 1/\mu_X$. Substituting these parameter values in (3-40), the $E[X^r]$ is given as

$$E\left[X^r\right] = \mu_X^r \Gamma(r+1) \tag{3-42}$$

Normal Distribution

The probability density function of normal distribution is

$$p_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2} \tag{3-43}$$

where μ_x and σ^2_x are the parameters of normal distribution. Assuming, $\frac{X-\mu_X}{\sigma_X} = z$, the random variable X can be written as, $X = (\mu_X + \sigma_X z)$, the $E[X^r]$ can be written as

$$E[X^r] = E[(\mu_X + \sigma_X z)^r] = \mu_X^r E[(1 + CV_X z)^r \tag{3-44}$$

For $CV_X < 1.0$, (3-44) can be expanded using Binomial Theorem as

$$E[X^r] = \mu_X^r E\left[1 + rCV_X z + \frac{r(r-1)}{2!} CV_X^2 z^2 + \ldots\ldots + \frac{r(r-1)(r-2)\ldots.(r-n+1)}{n!} CV_X^n z^n + \ldots\right] \tag{3-45}$$

Taking expectation of all the terms, (3-45) is written as

$$E[X^r] = \mu_X^r \left[1 + rCV_X E[z] + \frac{r(r-1)}{2!} CV_X^2 E[z^2] + \ldots\ldots + \frac{r(r-1)(r-2)\ldots.(r-n+1)}{n!} CV_X^n E[z^n] + \ldots\right] \tag{3-46}$$

where n is the term number plus 1 in the expansion. The $E[z^n]$ is given as

$$E[X^r] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^n e^{-\frac{1}{2}z^2} dz \tag{3-47}$$

The integral of (3-47) is

$$E[z^n] = \frac{2^{n/2}\Gamma[(n+1)/2]}{\sqrt{\pi}} = \frac{n!}{2^{n/2}(n/2)!} = (n-1)(n-3)\ldots\ldots.(3)(1), \text{ when n is even} \tag{3-48a}$$

$$E[z^n] = 0, \text{ when } n \text{ is odd.} \tag{3-48b}$$

Substituting (48b) in (3-46), the resulting equation is written as

$$E[X^r] = \mu_X^r \left[1 + \frac{r(r-1)}{2!} CV_X^2 E[z^2] + \ldots\ldots + \frac{r(r-1)(r-2)\ldots.(r-n+1)}{n!} CV_X^n E[z^n] + \ldots\right]. \tag{3-49}$$

When r is a positive integer, the RHS of (3-49) is finite and terminates when $n = r + 1$. Consequently, (3-49) can be written as

$$E\left[X^r\right]=\mu_X^r\sum_{n=0}^{r/2}\binom{r}{2n}CV_X^{2n}E\left[z^{2n}\right], \text{ when } r \text{ is even,} \tag{3-50a}$$

$$E\left[X^r\right]=\mu_X^r\sum_{n=0}^{(r-1)/2}\binom{r}{2n}CV_X^{2n}E\left[z^{2n}\right], \text{ when } r \text{ is odd.} \tag{3-50b}$$

When r is anything but a positive integer, the RHS of (3-49) does not converge. Equation (3-49) can be further simplified by substituting $E[z^n]$ from (48a) as

$$E\left[X^r\right]=\mu_X^r\left[1+\frac{r(r-1)}{2}CV_X^2+\ldots\ldots+\frac{r(r-1)(r-2)\ldots.(r-n+1)}{2^{n/2}(n/2)!}CV_X^n+\ldots\right]. \tag{3-51}$$

When r is a positive fraction, a very good estimate of $E[X^r]$ can be obtained using (3-50a) and (3-50b) with rounded value of r to its nearest whole number. In cases when r is negative, it is observed that the truncation error depends upon the magnitudes of r and CV_X. Further, there exists a minimum error of truncation for a given combination of r and CV_X, beyond which no improvement in $E[X^r]$ is possible. To evaluate approximate value of $E[X^r]$, a trial and error procedure was used to determine the number of terms to be summed up to give the minimum error in $E[X^r]$ for a given combination of r and CV_X. It is worth to note that when $CV_X < 0.1$, the truncation error is very small (< 0.1%) but as CV_X increases, the magnitude of this error increases rapidly.

Examples

To demonstrate the use of developed generic expectation function for a power function, two examples are presented.

Example No. 1 (Probabilistic human health risk assessment)

Quantitative risk assessment has received increased attention because of the recognition of both the potential threat to human health from hazardous substances and the potential for releases into the environment. Recognizing the extent of the hazardous waste problem and role of risk assessment, the EPA has developed assessment procedures that are used for a variety of purposes. Risk assessment is used for designating substances as hazardous and establishing minimum quantities for reporting releases when they would present substantial danger. In addition, risk assessment is used to evaluate the relative dangers of various sites in order to establish priorities for response actions and for developing, evaluating, and selecting appropriate response actions at the contaminated site. For example, risk assessment is used to evaluate threats to public health posed by a superfund site.

The risk assessment is carried out in four steps (USEPA, 1989). The first step is hazard identification in which chemicals of concern are selected based on their toxicity, mobility, spatial distribution and concentration. The second step is exposure assessment in which all possible exposure pathways (e.g., inhalation, ingestion, dermal, etc.) are identified. In the third step, intake doses of the pre-identified contaminants absorbed through the various exposure routes are estimated. The final step is the risk characterization, in which the magnitude of the risk is calculated. Quantitative uncertainty analysis is necessary when screening level calculations indicate a potential problem, remediation may result in high costs, or it is necessary to establish the relative importance of contaminants and exposure pathways.

To demonstrate an application of the developed method to risk characterization, risk assessment due to ingestion of contaminated soils is considered. Ingestion of soils contaminated by high molecular weight contaminants such as polychlorinated biphenyl (PCBs) is a potential source of human exposure to toxicants. The following equation

(USEPA, 1990) is used to estimate the probability of excess lifetime cancer, R_c, due to ingestion of contaminated soil

$$R_c = \frac{C_s I_r C_f F_i E_f E_d}{B_w A_t} S_f = C_s I_r C_f F_i E_f E_d B_w^{-1} A_t^{-1} S_f \tag{3-52}$$

where C_s = chemical concentration in the soil (mg/kg), C_f = a conversion factor (10^{-6} kg/mg), I_r = ingestion rate (mg soil/day), F_i = fraction ingested from contaminated sources (non-dimensional), E_f = exposure frequency (days/year), E_d = exposure duration (years), B_w = body weight (kg), A_t = averaging time (period over which exposure is averaged in days), and S_f = slope factor or cancer potency factor $(\text{mg/kg-day})^{-1}$.

There is always some uncertainty about each of these elements in risk estimation. A large number of references are available to describe the extent of uncertainty in each of the elements of (3-52). Talcott (1992) has summarized the available information in detail. Statistical properties of the variables in (3-52) are taken from Batchelor et al. (1998), and are applicable to individuals 1-6 years of age. The distribution of F_i was assumed to be the lognormal instead of the beta distribution as reported (Batchelor et al. 1998). This data is listed in Table 3-1.

In (3-52), there are two constants. One constant is C_f and the other is (1/365) to convert A_t to time in years. Combining these two, a new constant $C_0 = 10^{-6}/365$ = 2.74E-09 is obtained. Using generic expectation functions corresponding to distribution types of input variables and their means and *CV*s listed in Table 3-1, different orders of expectations of all the component power functions of (3-52) were obtained. These expectations have been listed in Table 3-2. Substituting these computed expectations into (3-2), moments of R_c about the origin were obtained. Using (3-3) or (3-5), different orders of central moments are computed. The computed moments of R_c about the origin and the mean have also been listed in Table 3-2.

The exact moments of R_c calculated in Table 3-2 can be used to characterize distribution of R_c. Based on first and second moments of R_c (1.972 E-05 and 1.885 E-08) the CV_{Rc} is calculated to be 6.96. Substituting values of second and third moments of R_c in (3-7), γ_{Rc} is obtained as 2.034 E+02. Using (3-9) and calculated values of second and fourth moments of R_c, the value of κ_{Rc} is determined as 9.192 E05. These characteristics of R_c provide a clear picture of its distribution. These characteristics and non-negative property of R_c indicate that R_c has lognormal distribution.

Example No. 2 (Risk analysis of storm sewer design)

For storm sewers, failure and potential property damage occurs when the peak runoff, Q_L, exceeds the storm sewer capacity, Q_C. Using rational method Q_L is expressed as:

$$Q_L = \lambda_L CIA \tag{2-53}$$

where λ_L = correction factor for model uncertainty, C = runoff coefficient; I = rainfall intensity; and A = drainage area. Using Manning's equation, Q_C is estimated using (Mays and Tung, 1992; Tung, 1990; Melching and Yen, 1986):

$$Q_C = \frac{0.463}{n} \lambda_m d^{\frac{8}{3}} S_0^{\frac{1}{2}} \tag{3-54}$$

where n = Manning's roughness; λ_m = model correction factor; d = pipe diameter; S_0 = pipe slope. In this example R is the designed capacity of sewer, Q_C, and L is the peak runoff Q_L. Using (3-11), the performance function can be defined as

$$Z = Q_C - Q_L \tag{3-55}$$

The statistical data of the variables included in (3-53) and (3-54) is taken from Mays and Tung (1992) and Tung (1990) and is presented in Table 3-3. Using (3-30) for variables having the symmetrical triangular distribution and (3-41) for variables having the gamma distribution and Table 3-3, the 1st, 2nd, 3rd, and 4th order expectations of all the component power functions of (3-53) and (3-54) were obtained as listed in Table 3-4. Using (3-2) and computed expectations of various component power functions, different orders of moments of Q_L and Q_C about the origin were calculated. Similarly, central moments of different orders were obtained for both Q_L and Q_C using (3-3). Using (3-14), (3-15), (3-16), and (3-17) and various orders of moments of Q_L and Q_C about the origin different orders of moments of Z about the origin were calculated. Substituting these moments about the origin in (3-3), various orders of central moments were obtained. Using (3-7) and (3-9), skewness and coefficient of kurtosis were calculated for Q_L, Q_C, and Z. All of these calculations are carried out in a tabular form as listed in Table 3-4.

Now exact moments and other distribution characteristics of Z are available. Using this information, several suitable probability distributions can be selected for Z, and risk corresponding to each of these assumed distributions can be calculated. For estimating the range, risk corresponding to the normal and uniform distributions can be estimated. The risk obtained assuming these two distributions may be regarded as extremes, since in reality the Z distribution of most cases probably falls between the normal and uniform distributions (Yen et al., 1986). Using the extremes and risk calculated assuming other distributions, an appropriate decision can be taken about the system risk.

Seeing the distribution characteristics of Z, the normal distribution may be a good choice as it has a negligible skew and kurtosis close to 3. The CV of Z is quite high indicating negative values of Z, which will be true when Q_L is more than Q_C. Using (3-12) the risk corresponding to normal distribution is

$$P_f = P(Z < 0) = P(z < \frac{0 - 12.10}{10.48}) = P(z < -1.15) = \Phi(-1.15) = 0.124 \tag{3-56}$$

where z is the standard normal variate defined as $z = \frac{Z - \mu_Z}{\sigma_Z}$, and $\Phi(z)$ is the standard normal cumulative distribution function. Assuming the triangular distribution for Z as defined in (3-25a) and (3-25b), the risk can be calculated from (Yen et al., 1986)

$$P_f = P(Z < 0) = \frac{a}{(b-a)(c-a)} \qquad \text{for } c > 0 \tag{3-57a}$$

$$P_f = P(Z < 0) = \frac{c-a}{b-c} - \frac{c+2ab}{(b-a)(b-c)} \qquad \text{for } c < 0 \tag{3-57b}$$

Using (3-26) and computed values of μ_z, CV_z, and γ_z from Table 3-4, the values of 38.31, -13.23, and 11.22 were obtained corresponding to n = 0, 1, and 2 respectively. Arranging these values in order of minimum, maximum, and mode, values of a, b, and c can be determined. Therefore, a = -13.23, b = 38.31, and c = 11.22. As $c > 0$, (3-57a) can be used to calculate P_f. Substituting the values of a, b, and c in (3-57a), P_f = 0.139 was obtained.

Assuming 3-parameter lognormal distribution for Z, $p_Z(z)$ is given as (Haan, 1977)

$$p_Z(Z) = \frac{1}{(Z-\varepsilon)^2 \sigma_y \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{\ln(Z-\varepsilon) - \mu_y}{\sigma_y}\right]^2\right\} \tag{3-58}$$

where $y = ln(Z)$, and ε is a location parameter. The relationships between ε, y, and Z are given as

$$\mu_Z = \varepsilon + \exp\left(\mu_y + \frac{1}{2}\sigma_y^2\right) \tag{3-59}$$

$$\sigma_Z^2 = \exp\left(2\mu_y + \sigma_y^2\right)\left[\exp\left(\sigma_y^2\right) - 1\right] \tag{3-60}$$

$$\gamma_Z = \left[\exp\left(\sigma_y^2\right) - 1\right]^{\frac{3}{2}} + 3\left[\exp\left(\sigma_y^2\right) - 1\right]^{\frac{1}{2}} \tag{3-61}$$

To solve, substitute $\gamma_Z = 0.055$, in (3-61) and find $\left[\exp\left(\sigma_y^2\right) - 1\right]^{\frac{1}{2}}$. This cubic equation has one real and two imaginary roots. The real root gives $\sigma_y = 0.0183$. Substituting values of σ_y and σ_Z in (3-60), $\mu_y = 6.35$ was obtained. Substituting, σ_y, μ_y, and μ_Z in (3-59), ε = -560.7 was determined. Using σ_y, μ_y, and ε, the standard normal variate corresponding to $Z = 0$ was found as $z = -1.14$. The corresponding risk is obtained as

$$P_f = P(Z < 0) = P(z < -1.14) = \Phi(-1.14) = 0.127 \tag{3-62}$$

Using the Edgeworth asymptotic expansion (Abramowitz and Stegun, 1972; Kendall et al., 1987; Tung, 1996), P_f can be obtained as

$$P_f \approx \Phi(\xi) - \phi(\xi)\left[\frac{\gamma_Z}{6}H_2(\xi) + \left(\frac{\kappa_Z - 3}{24}\right)H_3(\xi) + \left(\frac{\gamma_Z^2}{72}\right)H_5(\xi)\right] \tag{3-63}$$

where $\phi(\xi)$ is the standard normal probability density function; and $H_r(\xi)$ is r^{th}-order Hermite polynomial (Abramowitz and Stegun, 1972). Calculating various order of $H_r(\xi)$ and substituting values of $\Phi(\xi)$, $\phi(\xi)$, γ_Z, and κ_Z in (3-63) $P_f = 0.124$ was obtained. To use the Fisher-Cornish expansion, a correction has to be applied to the standard normal variate using the following formula (Fisher and Cornish, 1960; Kendall et al., 1987; Tung, 1996)

$$\zeta = \xi + \frac{\gamma_Z}{6} H_2(\xi) + \left(\frac{\kappa_Z - 3}{24}\right) H_3(\xi) - \left(\frac{\gamma_Z^2}{36}\right) [2H_3(\xi) + H_1(\xi)] \qquad (3\text{-}64)$$

Substituting various values in (3-64), $\zeta = -1.124$ was obtained. The corresponding risk, $P_f = \Phi(-1.124) = 0.125$ was obtained. Now, to see the upper bound of risk, Z was assumed to have a uniform distribution as given in (3-20). Using (3-21) and (3-22), a and b were calculated as –6.13, and 30.33 respectively. The P_f is determined from (Yen et al., 1986)

$$P_f = \frac{1}{2} - \frac{1}{\sqrt{12} CV_Z} = 0.168 \qquad (3\text{-}65)$$

In Table 3-5, the different risk estimates obtained assuming different distributions, have been listed along with their parameters. Comparing different risk estimates presented in Table 3-6, it can be seen that computed risk varies from 12% to 17%. The normal, 3-parameter lognormal, and the Edgeworth asymptotic expansion give more or less similar results, which are almost equal to the lower bound of the risk. The Fisher-Cornish asymptotic expansion and triangular distributions both give risk estimates falling in between the extreme bounds obtained using the normal and uniform distributions. Practically speaking, with the possible exception of the uniform distribution, all distributional assumptions yield the same risk

Conclusions

In this paper, a simple approach of developing generic expectation functions is described. Using several commonly used distributions, analytical expressions for expectation functions are derived. These expectation functions can be used to determine exact estimates of any order of model output moments. Further, a simple and practical approach of evaluating the probability of failure of a system is suggested using the triangular distribution for the model output. An analytical equation is derived that will give the parameters of the triangular distribution, given the mean, CV, and coefficient of skewness of the output random variable. After delineating the triangular distribution, risk, reliability of the system can be estimated by calculating the appropriate area of the triangular distribution.

References

Abramowitz, M., and Stegun, I. A. (1972). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th ed., Dover Publications, New York.

Ang, A. H. S., and Tang, W. H. (1984). Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk, and Reliability. John Wiley & Sons, Inc., New York, N.Y.

Batchelor, B., Valdes, J., and Araganth, V. (1998). Stochastic risk assessment of sites contaminated by hazardous wastes. Journal of the Environmental Engineering Division, ASCE, 124(4), 380-388.

Bates, B. C., and Townley, L. R. (1988). Nonlinear discrete flood event models, 3. Analysis of prediction uncertainty. Journal of Hydrology, 99, 91-101.

Benjamin, J. R., and Cornell, C. A. (1970). Probability, Statistics, and Decisions for Civil Engineers. McGraw-Hill, New York, N.Y.

Borofsky, S. (1950). Elementary Theory of Equations, The Macmillan Company, New York.

Burn, D. H. and McBean, E. A. (1985). Optimization modeling of water quality in an uncertain environment. Water Resources Research, 21(7), 934-940.

Cawlfield, J. D., and Wu, M. C. (1993). Probabilistic sensitivity analysis for one dimensional reactive transport in porous media. Water Resources Research, 29(3), 661-672.

Cesare, M. A. (1991). First-order analysis of open-channel flow. Journal of the Hydraulic Engineering Division, ASCE, 117(2), 242-247.

Chang, C. H., Tung, Y. K., and Yang, J. C.(1995). Evaluation of probabilistic point estimate methods. Applied Mathematical Modeling, 19(2), 95-105.

Fisher, R. A., and Cornish, E. A. (1960). The percentile points of distributions having known cumulants. Technometrics, 2(2), 209-225.

Haan, C. T. (1977). Statistical Methods in Hydrology, The Iowa State University Press, Ames, Iowa.

Haan, C. T., Barfield, B. J., and Hayes, J. C. (1994). Design Hydrology and Sedimentology for Small Catchments. Academic Press, San Diego.

Harr, M. E. (1989). Probability estimates for multivariate analyses. Applied Mathematical Modeling, 13, 313-318.

Hasofer, A. M., and Lind, N. C. (1974). Exact and invariant second-moment code format. Journal of the Engineering Mechanics Division, ASCE, 100, 111-121.

Kendall, M., Stuart, A., and Ord, J. K. (1987). Kendall's Advanced Theory of Statistics, Vol. 1: Distribution Theory, 5th Ed., Oxford University Press, New York, N.Y.

Krenkel, P. A., and Ruane, R. J. (1979). Chapter 18: Basic approach to water quality modeling. Modeling of Rivers, H. W. Shen, ed., Wiley, New York.

Lee, H. L. and Mays, L. W. (1986). Hydraulic uncertainty in flood levee capacity. Journal of the Hydraulic Engineering Division, ASCE, 112(10), 928-934.

Mays, L. W., and Tung, Y. K. (1992). Hydrosystems Engineering and Management, McGraw-Hill, Inc., New York, N.Y.

Melching, C. S. and Yen, B. C. (1986). Slope influence on storm sewer risk. Stochastic and Risk Analyses in Hydraulic Engineering, B. C. Yen eds., Water resources publications, Littleton, Colorado, 79-89.

Melching, C. S. (1995). Reliability estimation, Computer Models of Watershed Hydrology, V. P. Singh, ed., Water Resources Publication, Littleton, Colorado, 69 – 118.

Melching, C. S. and Yen, B. C., and Wenzel, H. G. Jr. (1991). Output reliability as guide for selection of rainfall-runoff models. Journal of Water Resources Planning and Management, ASCE, 117(3), 383-398.

Melching, C. S., and Anmangandla, S. (1992). Improved first order uncertainty method for water quality modeling. Journal of the Environmental Engineering Division, 118(5), 791-805.

Mishra, S. (1998). A methodology for probabilistic risk-based performance assessments using the first-order reliability method. Proc., Risk-Based Performance Assessment and Decision Making Topical Meeting, American Nuclear Society, Pasco, WA, 1-6.

Novotny, V. and Olem, H. (1994). Water Quality: Prevention, Identification, and Management of Diffuse Pollution, Van Nostrand Reinhold, New York.

Rackwitz, R. (1976). Practical probabilistic approach to design, Comite European du Beton, Paris, Bulletin No. 112.

Rosenblueth, E. (1975). Point estimates for probability moments, Proc. National Academy of Sciences U.S.A., 72(10), 3812-3814.

Rosenblueth, E. (1981). Point estimates in probabilities. Applied Mathematical Modeling, 72(10), 3812-3814.

Sitar, N., Cawlfield, J. D., and Kiureghian, A. D. (1987). First-order reliability approach to stochastic analysis of subsurface flow and contaminant transport. Water Resources Research, 23(5), 794-804.

Talcott, F. W. (1992). How certain is that environmental risk estimate? Resources for the Future, 107, 10-15.

Tung, Y. K. (1990). Mellin transformation applied to uncertainty analysis in hydrology / hydraulics, Journal of the Hydraulic Engineering Division, ASCE, 116(5), 659-674.

Tung, Y. K. (1996). Uncertainty analysis in water resources engineering. Stochastic Hydraulics'96, K. S. Tickle, I. C. Goulter, C. Xu, S. A. Wasimi, and F. Bouchart eds., A. A. Balkema, Rotterdam, the Netherlands., 29-46.

Tung, Y. K., and Mays, L. W. (1980). Optimal risk-based design of water resources engineering projects, Technical report No. 171, Center for Research in Water Resources, Univ. of Texas, Austin, Austin, Tex.

USEPA (1990). Guidance on remedial actions for superfund sites with PCB contamination, Rep. No. EPA/540/G-90/007, U.S. Office of Emergency and Remedial Response, Washington, D.C.

USEPA (1989). Risk assessment guidance for superfund volume 1 human health evaluation manual (part A) interim final, Rep. No. EPA/540/1-89/002, US. Office of Emergency and Remedial Response, Washington, D.C.

Yen, B. C. and Tang, W. H. (1976). Risk-safety factor relation for storm sewer design, Journal of the Environmental Engineering Division, ASCE, 102 (2), 509-516.

Yen, B. C., Cheng, S. T., and Melching, C. S. (1986) First order reliability analysis, Stochastic and risk analyses in hydraulic engineering, B. C. Yen eds., Water resources publications, Littleton, Colorado, 79-89.

Table 3-1: Statistical data for example No. 1

Parameter	Symbol	Distribution	Parameter values	
			Mean	CV
(1)	(2)	(3)	(4)	(5)
Contaminant concentration (mg/kg)	C_s	Lognormal	155	0.39
Ingestion rate (mg/day)	I_r	Lognormal	100	1.26
Fraction ingested	F_I	Lognormal	0.909	0.03
Exposure frequency (days/yr.)	E_f	Exponential	17.4	1.0
Exposure duration (yr)	E_d	Exponential	13.0	1.0
Body weight (kg)	B_w	Lognormal	15.6	0.23
Averaging time (yr.)	A_t	Normal	70.0	0.19
Slope factor (kg-day/mg)	S_f	Lognormal	2.25	1.66

Table 3-2: Computation of moments of R_c in Example No. 1

Expectation	Order of Expectation, k			
	1	2	3	4
$E[C_s^k]$	1.550E+02	2.768E+04	5.695E+06	1.350E+09
$E[I_r^k]$	1.000E+02	2.588E+04	1.733E+07	3.002E+10
$E[F_i^k]$	9.090E-01	8.270E-01	7.531E-01	6.864E-01
$E[E_f^k]$	1.740E+1	6.055E+02	3.161E+04	2.200E+06
$E[E_d^k]$	1.300E+01	3.380E+02	1.318E+04	6.855E+05
$E[B_w^{-k}]$	6.749E-02	4.796E-03	3.589E-04	2.827E-05
$E[A_t^{-k}]$	1.487E-02	2.318E-04	3.825E-06	6.794E-08
$E[S_f^k]$	2.250	1.901E+01	6.034E02	7.191E+04
$E[R_C^k]$	1.972E-05	1.924E-08	5.274E-10	3.266E-10
$E[(R_C-\mu_C)^k]$	0	1.885E-08	5.263E-10	2.034E+02

Statistics of R_c: μ_{Rc} =1.972E-05; σ_{Rc} = 1.373E-04; CV_{Rc}= 6.96;

γ_{Rc} = 2.034 E+02; and κ_{Rc} = 9.192E+05

Table 3-3: Statistical data for Example No. 2

Variable (1)	Mean (2)	Standard deviation (3)	Distribution (4)
λ_m	1.100	0.0891	Triangular
N	0.015	0.0553	Gamma
D (ft)	3.000	0.0410	Triangular
S_0 (ft / ft)	0.005	0.1640	Triangular
λ_L	1.000	0.1230	Triangular
C	0.825	0.0618	Triangular
I (in/hr)	4.000	0.1535	Triangular
A (acre)	10.00	0.0408	Triangular

1 ft = 0.305m; 1 in = 2.54 cm; 1 acre = 4047 m^2

Table 3-4: Calculation of expectations for storm sewer design

Expectation	Order of Expectation, k			
	1	2	3	4
$E[\lambda_m^k]$	1.100	1.220	1.363	1.534
$E[d^{8k/3}]$	1.879E+1	3.573E+2	6.873E+3	1.337E+5
$E[S_0^{0.5k}]$	7.047E-2	5.000E-3	3.571E-4	2.567E-5
$E[n^{-k}]$	6.687E+1	4.486E+3	3.018E+5	2.037E+7
$E[\lambda_L^k]$	1.000	1.015	1.045	1.091
$E[C^k]$	8.250E-1	6.832E-1	5.680E-1	4.739E-1
$E[I^k]$	4.000	1.637E+1	6.849E+1	2.923E+2
$E[A^k]$	1.000E+1	1.000E+2	1.005E+3	1.010E+4
$E[Q_C^k]$	4.510E+1	2.095E+3	1.002E+5	4.929E+6
$E[(Q_C - \mu_C)^k]$	0	6.127E+1	1.884E+2	4.929E+6
$E[Q_L^k]$	3.300E+1	1.138E+3	4.087E+4	1.527E+6
$E[(Q_L - \mu_L)^k]$	0	4.857E+1	1.252E+2	6.957E+3
$E[Z^k]$	1.210E+1	2.562E+2	5.819E+3	1.572E+5
$E[(Z - \mu_Z)^k]$	0	1.098E+2	6.315E+1	3.627E+4

Coefficient of variation of Q_C, Q_L, and Z are 0.17, 0.21, and 0.87 respectively.
Coefficient of skew for Q_C, Q_L, and Z are 0.39, 0.37, and 0.055 respectively.
Coefficient of kurtosis for Q_C, Q_L, and Z are 3.05, 2.95 and 3.01 respectively.

Table 3-5: Comparison of different risk estimates for storm sewer design

Distribution assumed (1)	Parameters (2)	Risk (3)
Normal	$\mu_z = 12.1$, $\sigma^2_z = 109.8$	0.124
Uniform	$a = -6.13$, $b = 30.33$	0.168
Triangular	$a = -13.2$, $b = 38.3$, $c = 11.2$	0.139
Three parameter lognormal	$\varepsilon = -560.69$, $\mu_y = 6.35$, $\sigma_y = 0.018$	0.123
Edgeworth asymptotic expansion of CDF	$\mu_z = 12.1$, $\sigma^2_z = 109.8$, $\gamma_z = 0.055$, $\kappa_z = 3.01$	0.124
Fisher-Cornish asymptotic expansion of quantile	$\mu_z = 12.1$, $\sigma^2_z = 109.8$, $\gamma_z = 0.055$, $\kappa_z = 3.01$	0.125

CHAPTER IV

UNCERTAINTY ANALYSIS OF EXPONENTIAL MODELS

Abstract

Exponential models are of general interest as they have variety of applications in science and engineering. In particular, first-order reaction kinetics, which produce exponential models, are the most commonly used kinetics in modeling, designing, and performance evaluation of environmental engineering systems. Application of uncertainty and reliability analysis is essential for many problems related to environmental engineering systems since they involve a number of uncertain input parameters. As the exponent of an exponential model increases, its non-linearity also increases, and thus, application of FOA becomes doubtful. This paper describes a procedure for correcting the FOA estimates for parameter uncertainty, distribution type, and model non-linearity, in order to determine true values for the first and second moments of a model output. When confidence limits on the output or system reliability are of concern, the output distribution is required. This paper also describes the development of generic expectations as a function of the mean and the CVs of input random variables. Generic expectation functions can be used to determine higher order moments of model output. This knowledge helps in identifying the candidate distributions for the model output and provides more flexibility to include those

distributions, which require higher order moments. Both techniques, the correction procedure and generic expectation function method, are easy to use in any general application. Three examples are presented to demonstrate the use of developed techniques.

Key words: Exponential functions, first-order kinetics, uncertainty and reliability analysis, first-order approximation method, exact estimates of parameter uncertainty.

Introduction

First-order kinetics models are of general interest as they describe many events in science and engineering. These include problems involving change in population, pollution, temperature, bank savings, drugs in the bloodstream, and radioactive materials. In environmental engineering, a number of kinetic models are used to model various physical, chemical, and biological processes occurring in both natural environments (such as streams, aquifers, and air) and artificially controlled environments (such as water or wastewater treatment units). Among the most widely used models are the first-order reaction kinetics models. In water quality modeling, the first-order models are used (Schnoor, 1997; Thomann and Mueller, 1987; Baughman and Lassiter, 1978) to represent constituent reactions, microorganism decay/growth, volatilization, sorption, and biodegradation rates. While modeling mobility, fate, and transport of hazardous waste, first-order degradation kinetics are assumed for simplicity and because of the unavailability of other practical mathematical expressions (LaGrega et al. 1994). In air pollution modeling, first-order reaction kinetics are used for microbial viability decay of air borne microorganisms (Lighthart and Frisch, 1976) and for overall chemical decay of air pollutants (APIDSS) as they travel from the source to the receptor.

There are many examples where exponential models are used as basic performance models in water and wastewater treatment systems. Some of the examples include aeration and disinfection kinetics in water treatment, chlorine decay in drinking water systems, trickling filters, CSTR, and plug flow activated sludge systems in wastewater treatment systems. Similarly, facultative pond and constructed wetland systems employed in natural treatment systems also rely on 1st-order removal kinetics. In non-point source pollution modeling, nutrient components of commonly used watershed models (EPIC, AGNPS, OPUS) assume first-order reaction kinetics for nutrient transformations, pesticide leaching, and decomposition of crop residues.

First-order models are also known as the exponential models in which the input random variables occur as an exponent. The exponent consists of two variables: the rate coefficient and time. In natural environments, both of them are characteristically random variables. In surface water quality modeling, not only stream flow and waste flow are inherently random (Loucks and Lynn, 1966), but there are a number of uncertainties associated with the various physical and biological processes occurring within the stream environment (Tung and Hathhorn, 1988). In the subsurface environment, the fate and transport modeling of an organic contaminant is dependent on uncertain flow dynamics through porous media having varied physical properties as well as interactions of a variety of physical, chemical, and biological processes which are yet to be clearly defined (Smith and Charbeneau, 1990). The chemical characteristics of the contaminants also impact the transport formulation. Furthermore, the transport of constituents in the unsaturated zone is also dependent upon variations in the rate of rainfall and infiltration.

In air pollution modeling, model parameters are uncertain due to variations in wind speed, turbulence, temperature, humidity, atmospheric stability, and the presence of any barrier which might entrap the particle. However, in highly controlled systems where the time parameters and other processes are regulated, the rate coefficient will always be associated with some uncertainty related to measurement errors.

Due to the presence of uncertainty in the exponent parameters, the output of an exponential model is considered to be a random variable. Two commonly used methods for analyzing parameter uncertainty are the Monte Carlo simulation (MCS) and the first-order approximation (FOA) (Benjamin and Cornell, 1970). Both methods have limitations. The MCS is computationally intensive because of the number of simulations required for convergence, which is not well defined (Melching, 1995). In most engineering problems, the true probability distributions of the input variables are seldom known and commonly used distributions are typically assumed. The quality of the MCS estimates is affected by the appropriateness of the chosen distribution functions for the input variables and the number of simulations used in the analyses (Bates and Townley, 1988). FOA is computationally efficient but provides approximate model output estimates for the mean and variance only. The quality of these estimates is influenced by the coefficient of variation (CV) of input variables and non-linearity in the model (Burn and McBean, 1985; Tung, 1990).

To date, the only widely used criterion to ensure the validity of FOA approximation is to restrict the parameter's coefficient of variation (CV) to less than 0.2 (Benjamin and Cornell, 1970; Burges, 1979; Dettinger and Wilson, 1981). Smith and Charbeneau (1990) suggest FOA can be used if the difference between function gradients at the mean and one standard deviation away from the mean are less than some acceptable percentage (5-10%). Both of these criteria have limitations. It has been observed that error in FOA estimates depends upon parameter CV, parameter distribution, and model non-linearity. The nonlinearity of exponential models depends upon the magnitude of the exponent. As the mean value of the exponent increases, the non-linearity of exponential models increases. For the same CV and model non-linearity, the error in FOA estimates varies with the type of parameter distribution. Therefore, any criteria judging the suitability of FOA must include these three elements.

The main objective of uncertainty analysis is to evaluate the first and second moments of a model output in terms of input random variables. Exponential models become significantly nonlinear when the magnitude of the exponent > 1, and thus, the validity of FOA application becomes questionable. This paper describes a procedure to correct FOA estimates for parameter uncertainty, distribution type, and model non-linearity in order to determine true values for first and second moments of model output. When confidence limits on the output or system reliability are of concern, the output distribution is required. This paper also describes the development of generic expectations as a function of the mean and CVs of input random variables. Generic expectation functions can be used to determine higher order moments of the output. This knowledge helps in identifying the candidate distributions for the model output and provides more flexibility to include those distributions that require higher order moments. Both techniques, the correction procedure and generic expectation function approach, are easy to use in any general application. Three practical examples related to volatilization of organic compounds from streams, pesticide leaching assessment, and decay of chlorine in water distribution systems are presented to demonstrate the use of the developed techniques.

First Order Approximation Method

Benjamin and Cornell (1970) and Cornell (1972) have provided detailed description of FOA. Mathematically a random variable Y which is a function of n random independent variables can be expressed as

$$Y = g(\underline{X}) \tag{4-1}$$

where $\underline{X} = (X_1, X_2 \ldots\ldots\ldots, X_n)$, a vector containing n random independent variables X_i. Through the use of Taylor's expansion and its first order approximation, the mean of the model output can be approximated by

$$\hat{\mu}_Y = g\left(E\left[\underline{\overline{X}}\right]\right) \tag{4-2}$$

where $\underline{\overline{X}}$ is the vector containing the mean values of all the random variables, $\hat{\mu}_Y$ is the FOA predicted mean for a model output. The variance of the model output can be approximated as

$$\hat{\sigma}_Y^2 = \sum_{i=1}^{n} \left[\frac{\partial g}{\partial X_i}\right]_{\overline{X}_i}^2 \sigma_{X_i}^2 \tag{4-3}$$

where $\sigma_{X_i}^2$ is the variance of input parameter X_i and $\hat{\sigma}_Y^2$ is the FOA predicted variance for the model output. Since, the FOA is an approximate method giving only estimates for the means and variances of a model output, there is always some error associated with estimates obtained using it.

Relative Error in FOA Estimates

Consider an exponential function

$$Y = f(x) = be^{cx} \tag{4-4}$$

where b and c are constants. Using (4-2), the FOA estimate for the mean of Y, $\hat{\mu}_Y$ is given as

$$\hat{\mu}_Y = be^{c\mu_x} \tag{4-5}$$

μ_x is the mean of the input variable x. Using (4-3), the FOA estimate for the variance of Y, $\hat{\sigma}_Y^2$ is given as

$$\hat{\sigma}_Y^2 = b^2c^2e^{2c\mu_x}\sigma_x^2 = b^2c^2e^{2c\mu_x}\mu_x^2CV_x^2 \tag{4-6}$$

where σ_x^2 is the variance of x; and CV_x is the coefficient of variation of x which is defined as

$$CV_x = \frac{\sigma_x}{\mu_x} \tag{4-7}$$

The estimates obtained from (4-5) and (4-6) for μ_Y and σ_Y^2 contain errors. The relative error, E, in FOA estimates is defined as

$$\mathrm{E} = \frac{Exact\ value - FOA\ estimate}{Exact\ value} = 1 - \frac{FOA\ estimate}{Exact\ value} \tag{4-8}$$

The exact value of the mean and variance of an exponential function and therefore the corresponding relative error in FOA estimates for the mean and variance depend upon CV, mean, and type of distribution of input parameter(s). Rewriting (4-8)

$$Exact\ value = \frac{FOA\ estimate}{(1 - \mathrm{E})} \tag{4-9}$$

Using (4-9), FOA estimates of the mean and variance of an exponential function can be corrected if E is known.

Generic Expectation Function

The generic expectation function is defined as the r^{th} moment of Y about the origin (μ_r'). Mathematically, it is defined as

$$E\left[Y^r\right] = \mu_r' = \int_{-\infty}^{\infty} [f(x)]^r \, p_X(x)dx = b^r \int_{-\infty}^{\infty} e^{rcx} p_X(x)dx \tag{4-10}$$

where $E[\ ]$ is an expectation operator, and $p_X(x)$ is the probability density function of X. The r^{th}-central moment of Y, μ_r, can be obtained using the following equation (Haan, 1977)

$$\mu_r = E\left[(Y - \mu_Y)^r\right] = \sum_{i=0}^{r} (-1)^i \binom{r}{i} \mu_Y^i \mu_{r-i}' \tag{4-11}$$

where, μ_Y is the mean of Y, which can be evaluated from (4-10) by substituting $r = 1$, as

$$\mu_y = E[Y] \tag{4-12}$$

In most situations, distributional properties of a random variable are characterized in terms of their mean, variance, coefficient of skewness, and coefficient of kurtosis. The variance of Y, σ^2_Y, is defined as the second moment about the mean. Substituting $r = 2$ in (4-11), σ^2_Y is given as

$$\mu_2 = \sigma_Y^2 = E\left[Y^2\right] - \mu_Y^2 \tag{4-13}$$

where μ_2 is the second moment of Y about the mean. The coefficient of skewness of Y, γ_Y, is defined as (Haan, 1977)

$$\gamma_Y = \frac{\mu_3}{\mu_2^{3/2}} \tag{4-14}$$

where μ_3 is the third moment of Y about the mean which can be obtained by substituting $r = 3$ in (4-11) as

$$\mu_3 = E[Y^3] - 3\mu_Y E[Y^2] + 2\mu_Y^3 \tag{4-15}$$

The kurtosis of Y, κ_Y, is defined as (Haan, 1977)

$$\kappa_Y = \frac{\mu_4}{\mu_2^2} \tag{4-16}$$

where μ_4 is the fourth moment of Y about the mean which can be obtained by substituting $r = 4$ in (4-11) as

$$\mu_4 = E[Y^4] - 4\mu E[Y^3] + 6\mu^2 E[Y^2] - 3\mu^4 \tag{4-17}$$

Development of Relative Error and Generic Expectation Functions

Uniform Distribution

The probability density function $p_X(x)$ for the continuous uniform distribution is

$$p_X(x) = \frac{1}{(\beta - \alpha)} \qquad , \alpha \le x \le \beta \tag{4-18}$$

where α and β are the distribution parameters. Using the methods of moments, the estimates for α and β are given (Haan, 1977) as

$$\alpha = \mu_x - \sqrt{3}\sigma_x = \mu_x\left(1 - \sqrt{3}CV_x\right) \tag{4-19}$$

$$\beta = \mu_x + \sqrt{3}\sigma_x = \mu_x\left(1 + \sqrt{3}CV_x\right) \tag{4-20}$$

Substituting $p_X(x)$ into (4-10), the $E[Y^r]$ is given as

$$E[Y^r] = \frac{b^r}{2\sqrt{3}rc\mu_x CV_x}\left[e^{rc\mu_x\left(1+CV_x\sqrt{3}\right)} - e^{rc\mu_x\left(1-CV_x\sqrt{3}\right)}\right] \tag{4-21}$$

Substituting $r = 1$ and 2 into (4-21), $E[Y]$ and $E[Y^2]$ are given as

$$E[Y] = \frac{b}{2\sqrt{3}c\mu_x CV_x}\left[e^{c\mu_x\left(1+CV_x\sqrt{3}\right)} - e^{c\mu_x\left(1-CV_x\sqrt{3}\right)}\right] \tag{4-22}$$

$$E\left[Y^2\right] = \frac{b^2}{4\sqrt{3}c\mu_x CV_x}\left[e^{2c\mu_x\left(1+CV_x\sqrt{3}\right)} - e^{2c\mu_x\left(1-CV_x\sqrt{3}\right)}\right] \tag{4-23}$$

Substituting (4-22) and (4-23) into (4-13), σ_y^2 is given as

$$\sigma_y^2 = \frac{b^2}{12c^2\mu_x^2 CV_x^2}\left[e^{2c\mu_x\left(1+CV_x\sqrt{3}\right)}\left(1+\sqrt{3}c\mu_x CV_x\right) + 2e^{2c\mu_x}\right] \tag{4-24}$$

Substituting FOA predicted mean of Y from (4-5) and true mean of Y from (4-22) into (4-8), the expression for relative error in FOA predicted mean, $\mathrm{E}(\hat{\mu}_y)$ is given as

$$\mathrm{E}(\hat{\mu}_y) = 1 - \frac{2\sqrt{3}c\mu_x CV_x e^{\sqrt{3}c\mu_x CV_x}}{\left(e^{2\sqrt{3}c\mu_x CV_x} - 1\right)} \tag{4-25}$$

Figure 4-1 shows a plot of $\mathrm{E}(\hat{\mu}_y)$ versus exponent mean for various CV values ranging form 0.01 to 0.57. Now substituting the FOA estimated variance (4-6) and correct variance (4-24) into (4-8), the relative error in FOA predicted variance, $\mathrm{E}(\hat{\sigma}_y^2)$ is given as

$$\mathrm{E}(\hat{\sigma}_y^2) = 1 - \frac{12c^4\mu_x^4 CV_x^4 e^{2\sqrt{3}c\mu_x CV_x}}{\left(e^{2\sqrt{3}c\mu_x CV_x} - 1\right)\left[\left(\sqrt{3}c\mu_x CV_x - 1\right)e^{2\sqrt{3}c\mu_x CV_x} + \sqrt{3}c\mu_x CV_x + 1\right]} \tag{4-26}$$

Figure 4-2 depicts a plot of $\mathrm{E}(\hat{\sigma}_y^2)$ versus exponent mean for various CV values ranging form 0.01 to 0.57.

Symmetrical Triangular Distribution

The probability density function $p_X(x)$ for the symmetrical triangular distribution is

$$p_X(x) = \frac{2}{(\beta-\alpha)}\frac{(x-\alpha)}{(\omega-\alpha)}, \text{ when } \alpha \le x \le \omega$$

$$p_X(x) = \frac{2}{(\beta-\alpha)}\frac{(\beta-x)}{(\beta-\omega)}, \text{ when } \omega \le x \le \beta \tag{4-27}$$

where α, β, and ω are the minimum, maximum, and mode values of X. These parameters can be obtained by the following equation (Appendix II)

$$\underline{\alpha} = \mu_X \left\{ 1 + 2\sqrt{2}CV_X \cos\left[\frac{2\pi n}{3} + \frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_x \right) \right] \right\} \tag{4-28}$$

where $\underline{\alpha}$ = a vector containing β, α, and ω, which can be obtained by substituting $n = 0$, 1, and 2, respectively, into (4-28); and γ_X is the coefficient of skew of X. Substituting (4-27) into (4-10) and integrating, the $E[Y^r]$ is given as

$$E\left[Y^r\right] = \frac{2b^r\left[(\alpha-\beta)\exp(rc\omega) + (\beta-\omega)\exp(rc\alpha) + (\omega-\alpha)\exp(rc\beta)\right]}{r^2c^2(\beta-\alpha)(\omega-\alpha)(\beta-\omega)} \tag{4-29}$$

For symmetrical distribution $\gamma_X = 0$ and the parameters α, β, and ω can be obtained corresponding to $n = 1$, 0, and 2. The estimates of α, β, and ω are given as

$$\hat{\alpha} = \mu_X\left(1 - \sqrt{6}CV_X\right) \tag{4-30}$$

$$\hat{\beta} = \mu_X\left(1 + \sqrt{6}CV_X\right) \tag{4-31}$$

$$\hat{\omega} = \mu_X \tag{4-32}$$

Substituting estimates of α, β, and ω into (4-29), the $E[X^r]$ for the symmetrical triangular distribution is given as

$$E\left[Y^r\right]=\frac{b}{6r^2c^2\mu_x^2CV_x^2}\left[e^{\frac{1}{2}rc\mu_x\left(1+CV_x\sqrt{6}\right)}-e^{\frac{1}{2}rc\mu_x\left(1-CV_x\sqrt{6}\right)}\right] \tag{4-33}$$

Substituting r =1 and 2 into (4-33), $E[Y]$ *and* $E[Y^2]$ are expressed as

$$E\left[Y\right]=\frac{b}{6c^2\mu_x^2CV_x^2}\left[e^{\frac{1}{2}c\mu_x\left(1+CV_x\sqrt{6}\right)}-e^{\frac{1}{2}c\mu_x\left(1-CV_x\sqrt{6}\right)}\right] \tag{4-34}$$

$$E\left[Y^2\right]=\frac{b^2}{24c^2\mu_x^2CV_x^2}\left[e^{c\mu_x\left(1+CV_x\sqrt{6}\right)}-e^{c\mu_x\left(1-CV_x\sqrt{6}\right)}\right] \tag{4-35}$$

Substituting $E[Y]$ and $E[Y^2]$ into (4-13), σ_y^2 is written as

$$\sigma_y^2=\frac{b^2\left\{\left(3c^2\sigma_x^2-2\right)\left[e^{2c\mu_x\left(1+CV_x\sqrt{6}\right)}+e^{2c\mu_x\left(1-CV_x\sqrt{6}\right)}\right]-12\left(c^2\sigma_x^2+2\right)e^{2c\mu_x}+8e^{c\mu_x\left(2-CV\sqrt{6}\right)}+8e^{c\mu_x\left(2+CV\sqrt{6}\right)}\right\}}{72c^4\sigma_x^4} \tag{4-36}$$

where $\sigma_x = \mu_x CV_x$. Substituting (4-5) and (4-34) into (4-8), the expression for relative error into FOA predicted mean, $\mathrm{E}(\hat{\mu}_y)$ is given as

$$\mathrm{E}(\hat{\mu}_y)=1-\frac{6c^2\mu_x^2CV_x^2e^{c\mu_xCV_x\sqrt{6}}}{\left(e^{c\mu_xCV_x\sqrt{6}}-1\right)^2} \tag{4-37}$$

Equation (4-37) has been represented graphically in Figure 4-3 for different values of exponent mean and exponent CV ranging form 0.01 to 0.4. Substituting (4-6) and (4-36) into (4-8), the relative error in FOA predicted variance $\mathrm{E}(\hat{\sigma}_y^2)$ is expressed as

$$\mathrm{E}(\hat{\sigma}_y^2)=1-\frac{72c^6\mu_x^6CV_x^6e^{2\sqrt{6}c\mu_xCV_x}}{\left(e^{\sqrt{6}c\mu_xCV_x}-1\right)^2\left[\left(3c^2\mu_x^2CV_x^2-2\right)\left(e^{2\sqrt{6}c\mu_xCV_x}+1\right)+2e^{\sqrt{6}c\mu_xCV_x}\left(3c^2\mu_x^2CV_x^2+2\right)\right]} \tag{4-38}$$

Figure 4-4 plots (4-38) for various exponents mean values and CV values ranging from 0.01 to 0.40.

Normal Distribution

If X is normally distributed, its probability density function is given (Haan, 1977) as

$$p_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x - \mu_X}{\sigma_X} \right)^2 \right] \tag{4-39}$$

where μ_X and σ_X^2 are the distribution parameters. Substituting (4-39) into (4-10), $E[Y^r]$ is given as

$$E[Y^r] = b^r \exp\left(rc\mu_x + \frac{1}{2} r^2 c^2 \mu_X^2 CV_X^2 \right) \tag{4-40}$$

Substituting $r = 1$ and 2 into (4-40), $E[Y]$ and $E[Y^2]$ are given as

$$E[Y] = b \exp\left(c\mu_x + \frac{1}{2} c^2 \mu_X^2 CV_X^2 \right) \tag{4-41}$$

$$E[Y^2] = b^2 \exp\left[2\left(c\mu_x + c^2 \mu_X^2 CV_X^2 \right) \right] \tag{4-42}$$

Substituting $E[Y]$ and $E[Y^2]$ into (4-13), σ_y^2 is written as

$$\sigma_y^2 = b^2 \exp\left(2c\mu_x + c^2 \sigma_x^2 \right) \left[\exp\left(c^2 \sigma_x^2 \right) - 1 \right] \tag{4-43}$$

Substituting (4-5) and (4-41) into (4-8), the relative error in FOA predicted mean $\mathrm{E}(\hat{\mu}_y)$ is expressed as

$$\mathrm{E}(\hat{\mu}_y) = 1 - \exp\left[-\frac{1}{2} c^2 \mu_x^2 CV_x^2 \right] \tag{4-44}$$

Figure 4-5 presents a plot of $\mathrm{E}(\hat{\mu}_y)$ versus exponent mean for various CV values ranging from 0.01 to 0.33. Substituting (4-6) and (4-43) into (4-8), the relative error in FOA predicted variance $\mathrm{E}(\hat{\sigma}_y^2)$ can be expressed as

$$\mathrm{E}(\hat{\sigma}_y^2) = 1 - \frac{c^2\sigma_x^2}{\exp(c^2\sigma_x^2)[\exp(c^2\sigma_x^2) - 1]} \tag{4-45}$$

In Figure 4-6, $\mathrm{E}(\hat{\sigma}_y^2)$ is plotted against the exponent mean for various CV values ranging from 0.01 to 0.33.

Gamma Distribution

The gamma density function is given by

$$p_X(x) = \frac{\lambda^{\alpha} e^{-\lambda x} x^{(\alpha-1)}}{\Gamma(\alpha)} \qquad x,\ \alpha, \text{ and } \lambda > 0 \tag{4-46}$$

where α and λ are the distribution parameters. Using method of moments α and λ are expressed (Haan, 1977) as

$$\hat{\lambda} = \frac{\mu_x}{\sigma_x^2} = \frac{1}{\mu_x CV_x^2} \tag{4-47}$$

$$\hat{\alpha} = \frac{\mu_x^2}{\sigma_x^2} = \frac{1}{CV_x^2} \tag{4-48}$$

Substituting (4-46) into (4-10) and integrating, $E[Y^r]$ is obtained as

$$E[Y^r] = b^r\left(1 - cr\mu_x CV_x^2\right)^{-\frac{1}{CV_x^2}} \tag{4-49}$$

Substituting $r = 1$ and 2 into (4-40), $E[Y]$ and $E[Y^2]$ are given as

$$E[Y] = b\left(1 - c\mu_x CV_x^2\right)^{-\frac{1}{CV_x^2}} \tag{4-50}$$

$$E[y^2] = b^2\left(1 - 2c\mu_x CV_x^2\right)^{-\frac{1}{CV_x^2}} \tag{4-51}$$

Substituting $E[Y]$ and $E[Y^2]$ into (4-13), σ_Y^2 is written as

$$\sigma_y^2 = b^2\left[\left(1-2c\mu_x CV_x^2\right)^{-\frac{1}{CV_x^2}} - \left(1-c\mu_x CV_x^2\right)^{-\frac{2}{CV_x^2}}\right] \tag{4-52}$$

Substituting (4-5) and (4-50) into (4-8), the relative error in FOA predicted mean $\mathrm{E}(\hat{\mu}_y)$ is expressed as

$$\mathrm{E}(\hat{\mu}_y) = 1-\left(1-c\mu_x CV_x^2\right)^{\frac{1}{CV_x^2}} \exp(c\mu_x) \tag{4-53}$$

Figure 4-7 shows a plot of $\mathrm{E}(\hat{\mu}_y)$ versus exponent mean for various CV values ranging from 0.01 to 1.0. Substituting (4-6) and (4-52) into (4-8), the relative error in FOA predicted variance $\mathrm{E}(\hat{\sigma}_y^2)$ can be written as

$$\mathrm{E}(\hat{\sigma}_y^2) = 1-\frac{c^2\mu_x^2 CV_x^2 \exp(2c\mu_x)}{\left(1-2c\mu_x CV_x^2\right)^{-\frac{1}{CV_x^2}} - \left(1-c\mu_x CV_x^2\right)^{-\frac{2}{CV_x^2}}} \tag{4-54}$$

Figure 4-8 shows a plot of $\mathrm{E}(\hat{\sigma}_y^2)$ versus exponent mean for various CV values ranging from 0.01 to 1.0.

Exponential Distribution

The exponential distribution is a special case of the gamma distribution with $\hat{\alpha}$ = 1. Substituting $\hat{\alpha}$ = 1 in (4-48), CV_x = 1 is obtained. Substituting CV_x = 1 into (4-49), $E[Y^r]$ is given as

$$E\left[Y^r\right] = \frac{b^r}{\left(1-cr\mu_x\right)} \tag{4-55}$$

Substituting CV_x = 1 in (4-53), the relative error in FOA predicted mean, $\mathrm{E}(\hat{\mu}_y)$ is given as

$$\mathrm{E}(\hat{\mu}_y) = 1-\left(1-c\mu_x\right)\exp(c\mu_x) \tag{4-56}$$

On substituting $CV_x = 1$ in (4-54), the relative error in FOA predicted variance $E(\hat{\sigma}_y^2)$ is expressed as

$$E(\hat{\sigma}_y^2) = 1 - \frac{c^2 \mu_x^2 \exp(2c\mu_x)}{(1-2c\mu_x)^{-1} - (1-c\mu_x)^{-1}} \qquad (4\text{-}57)$$

In Figure 4-9, $E(\hat{\mu}_y)$ and $E(\hat{\sigma}_y^2)$ have been plotted with exponent mean values.

Examples

Three examples are used to illustrate the application of the developed procedures. In problems where only the mean and variance are required, the correction technique can be used to correct the FOA estimates for the mean and variance of a model output. In reliability and risk problems where the distribution of a model output or performance function is required, the generic expectation function can be used to determine higher order moments. Based on these moments, one may be able to determine an output variable distribution whose higher order moments are in exact match or choose a distribution among the commonly used distributions based on finding the closest fit by comparing the computed moments. All three examples are simple so that results can be easily interpreted.

Example No. 1 (Volatilization of organic compounds from streams)

Various physical, chemical, and biological processes occurring in the stream environment determine the fate of organic compounds discharged into streams and rivers. Among these processes, one of the most important for many compounds is volatilization, which is the physical transport of the compound through the air-water interface into the air (Rathbun and Tai, 1982). The loss of chemical due to air-water exchange is governed by first-order reaction kinetics (Rathbun and Tai, 1982; Schwarzenbach et al., 1995). The concentration, C, at distant x from the input point is given as

$$C = C_0 \exp\left(-\frac{K_L \Delta t}{d}\right) \tag{4-58}$$

where C_0 is the concentration of chemical after mixing of the release with the stream water; Δt is the travel time of flow; d is the average depth of water; and K_L is the mass-transfer coefficient for the volatilization of concerned chemical from the stream.

There is always some uncertainty involved in each of the variables of (4-58). C_0 is uncertain because of the complete mixing assumption and measurement errors. Δt is uncertain because of spatial and temporal variations in velocity of flow. Average depth also varies from cross-section to cross-section. K_L is uncertain because it is a mathematical approximation of several complex processes and also because of the number of uncertainties that are associated with those various processes occurring within the stream environment. Rathbun and Tai (1982) developed regression equations for K_L for the volatilization of ethylene and propane in terms of hydraulic and geometric characteristics of streams. To obtain mass transfer coefficients for other organics, corrections for molecular diffusivity, molecular diameter, or molecular weight were suggested.

In this example (Rathbun and Tai, 1982), an accidental release of a wastewater containing carbon tetrachloride into a stream is considered. The problem is to determine the mean and variance of chemical concentration after a given travel time. The mean values of C_0, K_L, d, and Δt are taken from Rathbun and Tai (1982) as 100 mg/L, 1.43 m/day, 0.40 m, and 2.6 days respectively. Assuming, C_0, d, and Δt as constants, the impact of uncertainty in K_L on the mean and variance of C is studied using different CV_{K_L} values and distributions as shown in Table 4-1.

Based on the mean values of input parameters, the FOA estimates for the means are determined using (4-5) as given in column 2 of Table 4-1. Using relative error functions (4-25), (4-37), (4-44) and (4-53) for the uniform, triangular, normal and gamma distributions respectively, $E(\hat{\mu}_C)$ is determined corresponding to each distribution and different values of CV_{K_L} as listed in columns 3, 5, 7, and 9 in Table 4-1. $E(\hat{\mu}_C)$ can also be determined using relative error plots such as Figures 4-1, 4-3, 4-5, and 4-7 for the uniform, triangular, normal, and gamma distributions, respectively. Substituting $\hat{\mu}_C$ and $E(\hat{\mu}_C)$ in (4-9), correct mean values, μ_C, are obtained as listed in Table 4-1 for various assumed distributions. From Table 4-1, it can be noticed that while $\hat{\mu}_C$ remains constant regardless of CV_{K_L}, the μ_C values increase with increasing in CV_{K_L}, indicating impact of K_L uncertainty on mean of C. Furthermore, Table 4-1 also depicts the impact of distribution type used for K_L. It can be observed from Table 4-1 that at smaller CV_{K_L}, the μ_C values are almost similar. Their differences increase with an increase in CV_{K_L}. The differences of μ_C values based on the normal distributions are found to be the most pronounced.

In Table 4-2, FOA estimates for the variance of C using different distributions and CVs of K_L are calculated as given in column 2. Relative errors in these estimates are determined using (4-26) for the uniform distribution, (4-38) for the triangular distribution, (4-45) for the normal distribution, and (4-54) for the gamma distribution.

The calculated $E(\hat{\sigma}^2_{C_x})$ are given in columns 3, 5, 7, and 9 of Table 4-2 respectively. Substituting $\hat{\sigma}^2_{C_x}$ and $E(\hat{\sigma}^2_{C_x})$ into (4-9), correct values of $\sigma^2_{C_x}$ are determined as listed in columns 4,6, 8, and 10 of Table 4-2.

In Table 4-2, estimates of $E(\hat{\sigma}^2_{C_x})$ for various distributions indicate that there is significant error in FOA estimates even at small values of CV_{K_L}. The impact of distribution type is also clearly depicted by comparing $E(\hat{\sigma}^2_{C_x})$ values for different distributions. Their differences become more pronounced at higher CV_{K_L} values. It can be noticed that at $CV_{K_L} = 0.2$, variation in relative errors for different distributions is 90% to 100% indicating CV $\leq$ 0.2 criteria is not valid.

Example No. 2 (Uncertainty in a pesticide leaching assessment)

For the past decade, there have been concerns over the problem of nonpoint source pollution of groundwater with organic chemicals. Several screening indices to determine a pesticide's leaching potential were suggested by various researchers. These chemical fate indices are based on the relative travel time needed for the pesticides to migrate through the vadose zone and on the relative mass emission from the vadose zone (Rao et al., 1985). These indices, which are less data demanding than deterministic-conceptual models, have not been widely accepted due to concerns regarding their reliability (Loague and Green, 1988).

Consider the attenuation factor (*AF*) index proposed by Rao et al. (1985) to rank pesticides with respect to their potential to leach to groundwater. The primary processes controlling the rate of pesticide leaching considered in the *AF* index are advection, sorption, and transformation. Sorption is incorporated into *AF* by a retardation factor, *RF*, defined (Rao et al., 1985; Loague and Green, 1988) as

$$RF = 1 + \frac{\rho_b f_{oc} K_{oc}}{\theta_{FC}} + \frac{n_a K_H}{\theta_{FC}} \quad (4\text{-}59)$$

where ρ_b is the bulk density of soil (kg/m^3); f_{oc} is the organic carbon content in the soil (mass fraction); K_{oc} is the pesticide sorption coefficient (m^3/kg); θ_{FC} is the water content of soil at field capacity (volume fraction); n_a is the air-filled porosity of soil (fraction); and K_H is Henry's constant (dimensionless). Assuming first-order reaction kinetics for the pesticide degradation, *AF* is defined as the fraction of surface-applied pesticide that reaches the groundwater. Mathematically, *AF* is given (Rao et al.,1985; Loague and Green, 1988) as

$$AF = \exp(-kt) \quad (4\text{-}60)$$

where k is the first-order degradation rate coefficient (days^{-1}); and t is the total travel time required for a pesticide to travel from soil surface to the water table. The total travel time can be approximated (Rao et al.,1985; Loague and Green, 1988) as

$$t = \frac{d\ RF\ \theta_{FC}}{q} \quad (4\text{-}61)$$

where d is the depth of the water table (m); and q is the net annual ground water recharge rate (m/day).Substituting (4-59) into (4-61), t is rewritten as

$$t = \frac{d}{q}\left(\theta_{FC} + \rho_b f_{oc} K_{oc} + n_a K_H\right) \quad (4\text{-}62)$$

The pesticide half-life, $T_{1/2}$, is related to k as

$$k = \frac{0.693}{T_{1/2}} \quad (4\text{-}63)$$

$$AF = \exp\left[-\frac{0.693d}{q\ T_{1/2}} \left(\theta_{FC} + \rho_b f_{oc} K_{oc} + n_a K_H \right) \right] \tag{4-64}$$

The assumptions and limitations of the *AF* index are described by Rao et al. (1985) and Loague et al. (1989, 1990). The parameters of (4-64) include soil properties, hydrogeologic and climatic characteristics, and chemical coefficients. An extensive data set of the Pearl Harbor Basin, Hawaii, has been given by Loague (1991), Loague and Green (1988), and Loague et al. (1989, 1990). It is noticed that the coefficient of variation of the parameters in (4-64) ranges from 0.2 to 0.96, indicating a very large uncertainty. Therefore, it becomes imperative to characterize the impact of parameter uncertainties on the estimates of AF. The impact of data uncertainty in pesticide leaching assessments has been addressed by Loague et al. (1989, 1990), Loague and Green (1988, 1990), and Loague(1991) using FOA. Very large parameter CV values and model nonlinearity make the reliability of the FOA estimates for the mean and variance of AF questionable. Table 4-3 presents the statistical properties of the parameters of (4-64) corresponding to the inceptisols soil order in the Pearl Harbor Basin (Loague and Green, 1988; Loague et al., 1990). The distributions of θ_{FC} and f_{oc} are assumed to be lognormal (Labieniec et al., 1994). Recognizing the high CV values and non-negative constraint for the rest of the parameters, a lognormal distribution is assumed.

Considering application of *AF* for a single near-surface layer, *d* is assumed to be a constant value of 0.5 m. The pesticide selected for the leaching potential assessment was diuron, for which K_H was assumed to be zero (Loague and Green, 1988; Loague et al., 1990). Substituting $K_H = 0$, (4-64) is rewritten as

$$AF = \exp\left[-\frac{0.693d}{q\ T_{1/2}} \left(\theta_{FC} + \rho_b f_{oc} K_{oc} \right) \right] \tag{4-65}$$

Applying FOA on (4-65) with the statistical data given in Table 4-3, the FOA estimates for the mean ($\hat{\mu}_{AF}$) and standard deviation ($\hat{\sigma}_{AF}$) of *AF* are determined as 3.18E-21 and 2.26E-19 respectively. These values are similar to what were calculated by Loague and Green (1988), Loague et al. (1990), and Loague (1991).

To calculate the correct estimates of mean and variance of *AF*, the developed technique is used. For simplification, (4-65) is rewritten as

$$AF = \exp(c\tau) \tag{4-66a}$$

where $c = -1$; and τ is a random variable defined as

$$\tau = 0.693d\ q^{-1} T_{1/2}^{-1} \left(\theta_{FC} + \rho_b f_{oc} K_{oc} \right) \tag{4-66b}$$

To use the developed relative error equations or plots, the distribution of τ must be known. If higher order moments of τ are known, an appropriate distribution can be determined either by incorporating higher-order moments exactly using the method of entropy (Tung, 1996) or by choosing an approximate distribution form already available distributions based on the information about moments. The exact moments of τ can be evaluated using the generic expectation function approach (GEFA) for a power function as presented in Chapter III.

The generic expectation function for *X*, where *X* is lognormally distributed is given as

$$E\left[X^r\right] = \exp\left(\mu_Y r + \frac{1}{2} \sigma_Y^2 r^2 \right) \tag{4-67}$$

where $Y = ln(X)$ is normally distributed with parameters μ_Y and σ_Y^2. The parameters μ_Y and σ_Y^2 are defined (Haan, 1977) as

$$\mu_Y = \frac{1}{2}\ln\left[\frac{\mu_X^2}{1+CV_X^2}\right] \tag{4-68}$$

$$\sigma_Y^2 = \ln(1+CV_X^2) \tag{4-69}$$

Using (4-67) and data given in Table 4-3, different orders of moments of component power functions of (4-66b) about the origin are obtained. Multiplying moments of individual power functions $E[\rho_b^r]$, $E[f_{oc}^r]$, and $E[K_{oc}^r]$ for a required order, $E[(\rho_b f_{oc} K_{oc})^r]$ is determined for various values of r as listed in Table 4-4. Using $E[\theta_{Fc}^r]$ and $E[(\rho_b f_{oc} K_{oc})^r]$ and applying the linearity property of expectation, $E[(\theta_{Fc} + \rho_b f_{oc} K_{oc})^r]$ is determined for different values of r. Multiplying 0.693 with d = 0.5, 0.3465 is obtained. For a required value of r, $(0.3465)^r, E[q^{-r}]$, $E[t_{1/2}^{-r}]$ and $E[(\theta_{Fc} + \rho_b f_{oc} K_{oc})^r]$ are multiplied to obtain $E\left[\left\{\frac{0.693d}{qt_{1/2}}(\theta_{Fc} + \rho_b f_{oc} K_{oc})\right\}^r\right]$. Using these moments and (4-11), central moments of τ are obtained. Using obtained central moments with (4-12), (4-13), (4-14) and (4-16), mean, variance, coefficient of skewness, coefficient of kurtosis of τ are evaluated. The calculation is presented in Table 4-4.

Seeing the relative error plots corresponding to μ_τ =128.96 and CV_τ = 2.23, it is noticed that estimates of mean and variance have almost 100% relative error regardless of the distribution of τ.

Statistics indicate that τ can be assumed to be approximately lognormally distributed and can be verified by comparing coefficients of skewness and kurtosis. Assuming lognormal distribution for τ with parameters μ_τ =128.96 and CV_τ = 2.23, the coefficients of skewness and kurtosis are calculated to be 17.8 and 1800 respectively. This indicates that a lognormal distribution is reasonable. Using lognormal distribution for τ, the μ_{AF} = 3.29E-03 and σ_{AF} = 2.74E-02 are calculated using the Gauss-Laguerre quadrature method (Zwillinger, 1996).

To verify the above results, MCS was used to determine the mean and standard deviation of *AF*. To ensure the convergence, 100,000 simulations were used in the MCS. The values of μ_{AF} and σ_{AF} obtained using MCS are presented in Table 4-5 along with those obtained from FOA and GEFA. Table 4-5 indicates the effect of choice of uncertainty analysis method. The results of FOA are totally erroneous and may have serious consequences on the decision making. The results of MCS are comparable with that of GEFA. But, it can be noticed that using such a large number of simulations, the MCS has not converged to the exact estimates. Further, it is observed that the discrepancy in the coefficients of skewness and kurtosis is large. This may affect risk or reliability analysis of a project involving *AF*.

Example No. 3 (Uncertainty in residual chlorine in water distribution systems)

In drinking water distribution systems, it is current practice to maintain a desired level of residual chlorine concentration to provide protection against leaks, regrowth of microbial contamination, and other breakdowns. Most network modeling packages assume that chlorine decay follows first-order kinetics (Powell et al., 2000). The chlorine concentration, C, at any time t (mg/L) is given by the following equation:

$$C = C_0 \exp(-kt) \tag{4-70}$$

where C_0 is the initial chlorine concentration (mg/L); and $k =$ overall decay constant (L/h). In literature considerable variability was observed in the value of the first-order decay constant (Powell et al., 2000). There are a number of factors affecting chlorine decay in the water distribution system such as reactions both within the bulk fluid and with pipe material, organic matter, or presence of other chlorine demanding impurities. Whereas, the reactions within the bulk fluid are affected by water temperature and organic content of water, the reactions with pipe wall are related to the corrosiveness of the ferrous pipe materials and the perimeter. Some of this variability is likely due to changes in the concentration and chemical nature of the compounds that chlorine is reacting with. Other uncertain factors are changes in temperature, chlorine dose, and organic content of the water. Furthermore, the network's ratio of chlorine to organic reactants in the water varies significantly with time or space. Since the reactions do not exactly follow first order decay, the value of k changes with respect to time as well.

In order to provide safe drinking water, the impact of uncertainty in the first-order decay constant must be investigated. To obtain statistical characteristics of k, frequency distribution data of k was read from Powell et al. (2000). The mean and coefficient of variation of k were found to be 0.14 L/h and 0.63 respectively. Three distributions, namely exponential, lognormal, and gamma were fitted. Using 6 classes, the chi-square statistics, χ_c^2 (Haan, 1977), were calculated as 6.52, 57.71, and 5.0, respectively. Comparing χ_c^2 with $\chi_{0.90,3}^2 = 6.25$, it is concluded that the gamma distribution adequately describes the data at the 10% significance level. In Table 4-6, statistical characteristics of C at different times t ranging from 1 hr to 24 hr were computed, which give a thorough understanding of the uncertainty in C at different times. In Table 4-6, $\hat{\mu}_C$, is calculated from (4-5), $E(\hat{\mu}_C)$ is calculated from (4-53), and μ_C is determined by substituting $\hat{\mu}_C$ and $E(\hat{\mu}_C)$ in (4-9). Similarly, using (4-6), (4-54), and (4-9), $\hat{\sigma}_C^2$, $E(\hat{\sigma}_C^2)$, and σ_C^2, are calculated as listed in Table 4-6.

It can be noticed from Table 4-6 that there is substantial amount of error in FOA estimates of means and variances of residual chlorine concentration in water distribution systems. The FOA underestimates the mean throughout the 24 hours, whereas, it overestimates the variance during first 10 hours and underestimates afterwards. To have an idea about distribution type, distributional characteristics of C can be determined using GEFA as given in Table 4-7. Using (4-49), 1st, 2nd, 3rd, and 4th orders of moments of C about the origin are estimated as listed in columns 2, 3, 4, and 5 of Table 4-7. Using these moments and (4-11), various orders of central moments are estimated and used to determine distributional characteristics (mean, variance, CV, coefficient of skew, and coefficient of kurtosis) of C as listed in columns 6, 7, 8, 9, and 10 respectively.

From Table 4-7, it can be concluded that C can not be represented by a normal distribution. Therefore, any analysis based on FOA and normal distribution for C will be misleading. A typical requirement for minimum free chlorine residual is 0.2 mg/L (Anonymus, 2000). In order to estimate the probability of failing to meet this requirement, i.e., $P(C \leq 0.2$ mg/L), the distribution of C is required. Using the variable transformation technique (Haan, 1977), the exact distribution of C is obtained as

$$p_C(c) = \frac{\left(\frac{\lambda}{t}\right)^{\alpha}}{C_0 \Gamma(\alpha)} \left(\frac{C}{C_0}\right)^{\frac{\lambda}{t}-1} \left[\ln\left(\frac{C}{C_0}\right)\right]^{\alpha-1} \tag{4-71}$$

where α and λ are the parameters of gamma distribution fitted for k. Using this distribution, one can compute the desired probabilities. It is general practice to use the normal distribution with FOA estimated mean and variance. In Table 4-8, different probability estimates, namely normal distribution with FOA estimated mean and variance, normal distribution with corrected mean and variance, and the exact derived distribution of C, are compared.

Table 4-8 indicates that the probability of residual chlorine concentration ≤ 0.2 mg/L, is overestimated by FOA. Based on FOA results, decision may be made to increase the residual chlorine concentration. This increased chlorination may prompt formation of undesirable by-products such as trihalomethanes (THM's) and other halogenated hydrocarbons, which are toxic to human health. Further, using normal distribution with correct means and variances, the probabilities of residual chlorine concentration ≤ 0.2 mg/L, are underestimated. This may result of reducing the residual chlorine concentration in the water, which may be inadequate against microbial protection. Due to the conflicting objectives, chlorine disinfection needs an exact analysis of residual chlorine concentration to safeguard water consumers.

Conclusions

In this paper two approaches, the correction procedure for correcting the FOA estimates for parameter uncertainty, parameter distribution type, and model non-linearity and the generic expectation function approach for evaluating exact moments of model output, are described. Analytical and graphical relationships for relative error, using several commonly used distributions, are developed to correct the FOA estimates to obtain exact values of the mean and variance of model output. This technique is particularly useful for determining exact values of the mean and variance of model output. When the distribution of a model output is required, the generic expectation method can be used. Analytical expressions for generic expectation functions using several commonly used distributions are derived. These functions can be used to determine exact model output moments of any order. Knowledge of higher-order moments helps in identifying the appropriate distribution for the model output. Three practical examples are solved which show that the developed techniques are not only easy to use but also provide more understanding of the process being considered.

References

Anonymus (2000), Drinking Water Information : Spring 2000 Edition, City of Chandler, Public Works Department, Water Quality Division. *http://www.chandlerz.org/html/service/whitebook%20200.htm*

Bates, B. C., and Townley, L. R. (1988). Nonlinear discrete flood event models, 3. Analysis of prediction uncertainty. Journal of Hydrology, 99, 91-101.

Baughman, G. L. and Lassiter, R. R. (1978). Prediction of environmental pollutant concentration. Estimating the Hazard of Chemical Substances to Aquatic Life, John Cairns, Jr., K.L. Dickson, and A.W. Maki, eds., ASTM STP 657., 35-54.

Benjamin, J. R., and Cornell, C. A. (1970). Probability, Statistics, and Decisions for Civil Engineers. McGraw-Hill, New York, N.Y.

Burges, S. J. (1979). Analysis of uncertainty in flood plain mapping. Water Resources Bulletin, 15(1), 227-243.

Burn, D. H. and McBean, E. A. (1995). Optimization modeling of water quality in an uncertain environment. Water Resources Research, 21(7), 934-940.

Cornell, C. A. (1972). First-order analysis of model and parameter uncertainty. Proceedings of Int. Symposium on Uncertainties in Hydrologic and Water Resources Systems, University of Arizona, vol. III, 1245-1276.

Dettinger, M. D. and Wilson, J. L. (1981). First-order analysis of uncertainty in numerical models of groundwater flow. Part 1. Mathematical development. Water Resources Research, 17(1), 149-161.

Haan, C. T. (1977). Statistical Methods in Hydrology, The Iowa State University Press, Ames, Iowa.

Labieniec, P. A., Dzombak, D. A., and Siegrist, R. L. (1994). Risk variability from uniform soil remediation goals for PCBs. Journal of the Environmental Engineering Division, ASCE, 120(3), 495-512.

LaGrega, M. D., Buckingham, P. L., and Jeffrey, C. E. (1994). Hazardous Waste Management, McGraw-Hill, Inc., New York.

Lighthart, B. and Frisch, A. S. (1976). Estimation of viable airborne microbes downwind from a point source. Applied Environmental Microbiology, 31(5), 700-704.

Loague, K. (1991). The impact of land use on estimates of pesticide leaching potential: Assessments and uncertainties. Journal of Contaminant Hydrology, 8,157-175.

Loague, K. M. and Green, R. E. (1990). Uncertainty in areal estimates of pesticide leaching potential, Transactions of 14th International Congress of Soil Science, International Soil Science Society, 162-167.

Loague, K. M., Green, R. E.(1988). Impact of data-related uncertainties in a pesticide leaching assessment. Methods for Groundwater Studies - Proceedings of a National Workshop, Agricultural Research Service, University of Nebraska-Lincon, NE., 99-119.

Loague, K. M., Green, R. E., Giambelluca, T. W., Liang, T. C., and Yost, R. S.(1990). Impact of uncertainty in soil, climate, and chemical information in a pesticide leaching assessment. Journal of Contaminant Hydrology, 5, 171-194.

Loague, K. M., Yost, R. S., Green, R. E., and Liang, T.C. (1989). Uncertainty in a pesticide leaching assessment for Hawaii. Journal of Contaminant Hydrology, 4,139-161.

Loucks, D. P. and Lynn, W. R. (1966). Probabilistic models for predicting stream quality. Water Resources Research, 2(3), 593-605.

Melching, C. S. (1995). Reliability estimation. Computer Models of Watershed Hydrology, V. P. Singh, ed., Water Resources Publication, Littleton, Colorado, 69 – 118.

Powell, J. C., West, J. R., Hallam, N. B., Forster, C. F., and Simms, J. (2000). Performance of various kinetic models for chlorine decay. Journal of the Environmental Engineering Division, ASCE, 126(1), 13-19.

Rao, P. S. C, Hornby, A. G., and Jessup, R. E. (1985). Indices for ranking the potential for pesticide contamination of ground water. Soil Crop Science Society, Florida Proc., 44, 1-8.

Rathbun, R. E. and Tai, D. R. (1982). Volatilization of organic compounds from streams. Journal of the Environmental Engineering Division, ASCE, 108(5), 973-989.

Schnoor, J. L. (1997). Environmental Modeling, John Wiley and Sons, New York.

Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. (1995). Environmental Organic Chemistry, John Wiley & Sons, Inc., New York.

Smith, V. J. and Charbeneau, R. J. (1990). Probabilistic soil contamination exposure assessment procedures. Journal of the Environmental Engineering Division, ASCE, 116(6), 1143-1163.

Thomann, R. V. and Mueller, J. A. (1987). Principles of Surface Water Quality Modeling and Control, Harper and Row, Publishers, New York.

Tung, Y. K. (1990). Mellin transformation applied to uncertainty analysis in hydrology/hydraulics. Journal of the Hydraulic Engineering Division, ASCE, 116(5), 659-674.

Tung, Y. K. and Hathhorn, W. E. (1988). Assessment of probability distribution of dissolved oxygen deficit. Journal of the Environmental Engineering Division, ASCE, 114(6), 1421-1435.

Zwillinger, D. (1996). Standard Mathematical Tables and Formulae, 30th ed., CRC Press Inc., Boca Raton, Florida.

Table 4-1: Correcting FOA estimates for the mean chemical concentration

CV_{K_L}	FOA Estimate	Exact estimates for mean values of $C(\mu g/L)$							
		Uniform		Triangular		Normal		Gamma	
	$\hat{\mu}_C$	E($\hat{\mu}_C$)	μ_C	E($\hat{\mu}_C$)	μ_C	E($\hat{\mu}_C$)	μ_C	E($\hat{\mu}_C$)	μ_C
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0.05	9.19	0.10	10.21	0.10	10.21	0.10	10.21	0.10	10.21
0.10	9.19	0.33	13.71	0.34	13.92	0.35	14.14	0.33	13.72
0.15	9.19	0.56	20.89	0.59	22.41	0.62	24.18	0.57	21.37
0.20	9.19	0.74	35.35	0.78	41.77	0.82	51.06	0.75	36.76
0.25	9.19	0.86	65.64	0.89	83.54	0.93	131.28	0.86	65.64
0.30	9.19	0.92	114.87	0.95	183.80	0.98	459.5	0.92	114.88

Table 4-2: Correcting FOA estimates for the variance of chemical concentration

CV_{K_L}	FOA Estimate	Exact estimates for variance of $C\,(\mu g/L)^2$							
		Uniform		Triangular		Normal		Gamma	
	$\hat{\sigma}^2_{C_x}$	E($\hat{\sigma}^2_{C_x}$)	$\sigma^2_{C_x}$	E($\hat{\sigma}^2_{C_x}$)	$\sigma^2_{C_x}$	E($\hat{\sigma}^2_{C_x}$)	$\sigma^2_{C_x}$	E($\hat{\sigma}^2_{C_x}$)	$\sigma^2_{C_x}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0.05	1.82E1	0.16	2.16E1	0.22	2.33E1	0.28	2.53E1	0.24	2.39E1
0.10	7.29E1	0.48	1.40E2	0.61	1.85E2	0.73	2.75E2	0.64	2.03E2
0.15	1.62E2	0.75	6.48E2	0.86	1.17E3	0.95	3.53E3	0.88	1.33E3
0.20	2.92E2	0.90	2.83E3	0.96	7.48E3	1.00	8.21E4	0.96	7.83E3
0.25	4.56E2	0.96	1.23E4	0.99	5.03E4	1.00	4.12E6	0.99	3.96E4
0.30	6.56E2	0.99	5.43E4	1.00	3.54E5	1.00	4.79E8	1.00	1.66E5

Table 4-3: Statistical properties of the parameters used in pesticide leaching assessment

Parameter	Symbol	Parameter values		Distribution
		Mean	CV	
(1)	(2)	(3)	(4)	(5)
Soil water content (fraction)	θ_{FC}	0.41	0.22	Lognormal
Soil bulk density (kg/m^3)	ρ_b	688	0.35	Lognormal
Soil organic content (fraction)	f_{oc}	0.09	0.56	Lognormal
Pesticide sorption coefficient (m^3/kg)	K_{oc}	0.383	0.72	Lognormal
Net annual groundwater recharge (m/day)	Q	5.4E-4	0.96	Lognormal
Pesticide half-life (days)	$T_{1/2}$	328	0.65	Lognormal

Table 4-4: Computation of the exact moments of τ

Expectation	Order of expectation, r			
	1	2	3	4
$E\left[\theta_{Fc}^{r}\right]$	4.10E-1	1.76E-1	7.94E-2	3.75E-2
$E\left[\rho_b^{r}\right]$	6.88E+2	5.31E+5	4.58E+8	4.44E+11
$E\left[f_{oc}^{r}\right]$	9.00E-2	1.06E-2	1.63E-3	3.29E-4
$E\left[K_{oc}^{r}\right]$	3.83E-1	2.23E-1	1.98E-1	2.69E-1
$E\left[q^{-r}\right]$	3.57E+3	2.45E+7	3.25E+11	8.32E+15
$E\left[t_{1/2}^{-r}\right]$	4.32E-3	2.65E-5	2.30E-7	2.83E-9
$E\left[(\rho_b f_{oc} K_{oc})^r\right]$	2.37E+1	1.26E+3	1.49E+5	3.93E+7
$E\left[(\theta_{Fc} + \rho_b f_{oc} K_{oc})^r\right]$	2.41E+1	1.27E+3	1.50E+5	3.95E+7
$E\left[\left\{\frac{0.693d}{qt_{1/2}}(\theta_{Fc} + \rho_b f_{oc} K_{oc})\right\}^r\right]$	1.29E+2	9.96E+4	4.68E+8	1.34E+13
Central moments of τ	0	8.29E+4	4.34E+8	1.32E+13

Statistics of τ: mean = 128.96, standard deviation = 288.07, CV = 2.23, Coefficient of skew = 18.15, and coefficient of kurtosis = 1916.93

Table 4-5: Comparison of the means and the standard deviations of AF

Method	Statistics of attenuation factor, AF		
	μ_{AF}	σ_{AF}	CV_{AF}
(1)	(2)	(3)	(4)
FOA	3.18E-21	2.26E-19	7.11
GEFA	3.29E-03	2.74E-02	8.33
MCS	2.75E-03	2.42E-02	8.80

Table 4-6: Correcting FOA estimates using relative error functions

t (hr)	Mean (mg/L)			Variance $(mg/L)^2$		
	FOA estimate	Relative error	Corrected Estimate	FOA estimate	Relative error	Corrected estimate
	$\hat{\mu}_C$	$E(\hat{\mu}_C)$	μ_C	$\hat{\sigma}_C^2$	$E(\hat{\sigma}_C^2)$	σ_C^2
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	0.7824	0.0037	0.7854	0.0048	-0.1005	0.0043
2	0.6802	0.0144	0.6901	0.0144	-0.1783	0.0122
3	0.5913	0.0311	0.6103	0.0245	-0.2323	0.0199
4	0.5141	0.0529	0.5428	0.0329	-0.2628	0.0260
5	0.4469	0.0790	0.4853	0.0388	-0.2712	0.0306
6	0.3885	0.1087	0.4359	0.0423	-0.2596	0.0336
7	0.3378	0.1412	0.3933	0.0435	-0.2306	0.0353
8	0.2937	0.1758	0.3563	0.0429	-0.1872	0.0362
9	0.2553	0.2120	0.3240	0.0411	-0.1321	0.0363
10	0.2219	0.2492	0.2956	0.0383	-0.0685	0.0359
11	0.1929	0.2869	0.2706	0.0350	0.0012	0.0351
12	0.1677	0.3248	0.2484	0.0315	0.0745	0.0341
13	0.1458	0.3625	0.2287	0.0280	0.1493	0.0329
14	0.1268	0.3996	0.2111	0.0245	0.2238	0.0316
15	0.1102	0.4359	0.1954	0.0213	0.2968	0.0302
16	0.0958	0.4713	0.1812	0.0183	0.3669	0.0289
17	0.0833	0.5056	0.1685	0.0156	0.4335	0.0275
18	0.0724	0.5385	0.1569	0.0132	0.4959	0.0262
19	0.0630	0.5702	0.1465	0.0111	0.5538	0.0249
20	0.0547	0.6003	0.1369	0.0093	0.6070	0.0237
21	0.0476	0.6290	0.1283	0.0078	0.6554	0.0225
22	0.0414	0.6563	0.1203	0.0064	0.6992	0.0214
23	0.0360	0.6820	0.1131	0.0053	0.7385	0.0204
24	0.0313	0.7062	0.1064	0.0044	0.7736	0.0193

Table 4-7: Uncertainty in chlorine concentration after a given time *t*

t (hr)	Moments of about the origin				Distributional characteristics of *C*				
	$E[C]$	$E[C^2]$	$E[C^3]$	$E[C^4]$	μ_C	σ_C^2	CV_C	γ_C	κ_C
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	0.785	0.621	0.494	0.396	0.785	0.004	0.084	-0.934	3.986
2	0.690	0.489	0.353	0.260	0.690	0.012	0.160	-0.672	3.192
3	0.610	0.392	0.262	0.181	0.610	0.020	0.231	-0.452	2.724
4	0.543	0.321	0.201	0.132	0.543	0.026	0.297	-0.262	2.451
5	0.485	0.266	0.158	0.100	0.485	0.031	0.360	-0.092	2.307
6	0.436	0.224	0.127	0.078	0.436	0.034	0.420	0.062	2.252
7	0.393	0.190	0.104	0.062	0.393	0.035	0.478	0.204	2.264
8	0.356	0.163	0.086	0.050	0.356	0.036	0.534	0.336	2.327
9	0.324	0.141	0.072	0.041	0.324	0.036	0.588	0.461	2.431
10	0.296	0.123	0.062	0.034	0.296	0.036	0.641	0.580	2.570
11	0.271	0.108	0.053	0.029	0.271	0.035	0.692	0.693	2.740
12	0.248	0.096	0.046	0.025	0.248	0.034	0.743	0.803	2.936
13	0.229	0.085	0.040	0.021	0.229	0.033	0.793	0.909	3.157
14	0.211	0.076	0.035	0.019	0.211	0.032	0.842	1.012	3.401
15	0.195	0.068	0.031	0.016	0.195	0.030	0.890	1.113	3.666
16	0.181	0.062	0.028	0.014	0.181	0.029	0.938	1.211	3.952
17	0.168	0.056	0.025	0.013	0.168	0.028	0.985	1.308	4.257
18	0.157	0.051	0.022	0.011	0.157	0.026	1.032	1.404	4.582
19	0.146	0.046	0.020	0.010	0.146	0.025	1.078	1.497	4.925
20	0.137	0.042	0.018	0.009	0.137	0.024	1.125	1.590	5.287
21	0.128	0.039	0.016	0.008	0.128	0.023	1.170	1.682	5.668
22	0.120	0.036	0.015	0.008	0.120	0.021	1.216	1.773	6.067
23	0.113	0.033	0.014	0.007	0.113	0.020	1.262	1.863	6.485
24	0.106	0.031	0.013	0.006	0.106	0.019	1.307	1.953	6.920

Table 4-8: Comparison between different probability estimates

T (hr)	$P(C \leq 0.2)$		
	Normal distribution with FOA estimated parameters	Normal distribution with corrected parameters	True distribution
(1)	(2)	(3)	(4)
1	0	0	1.43E-10
2	3.14E-05	4.63E-06	4.95E-05
3	6.19E-03	1.80E-03	2.74E-03
4	0.04	0.02	0.02
5	0.11	0.05	0.05
6	0.18	0.10	0.11
7	0.25	0.15	0.17
8	0.33	0.21	0.24
9	0.39	0.26	0.31
10	0.46	0.31	0.37
11	0.52	0.35	0.43
12	0.57	0.40	0.48
13	0.63	0.44	0.53
14	0.68	0.47	0.57
15	0.73	0.51	0.61
16	0.78	0.54	0.64
17	0.82	0.58	0.67
18	0.87	0.60	0.69
19	0.90	0.63	0.71
20	0.93	0.66	0.73
21	0.96	0.68	0.74
22	0.98	0.71	0.76
23	0.99	0.73	0.76
24	0.99	0.75	0.77

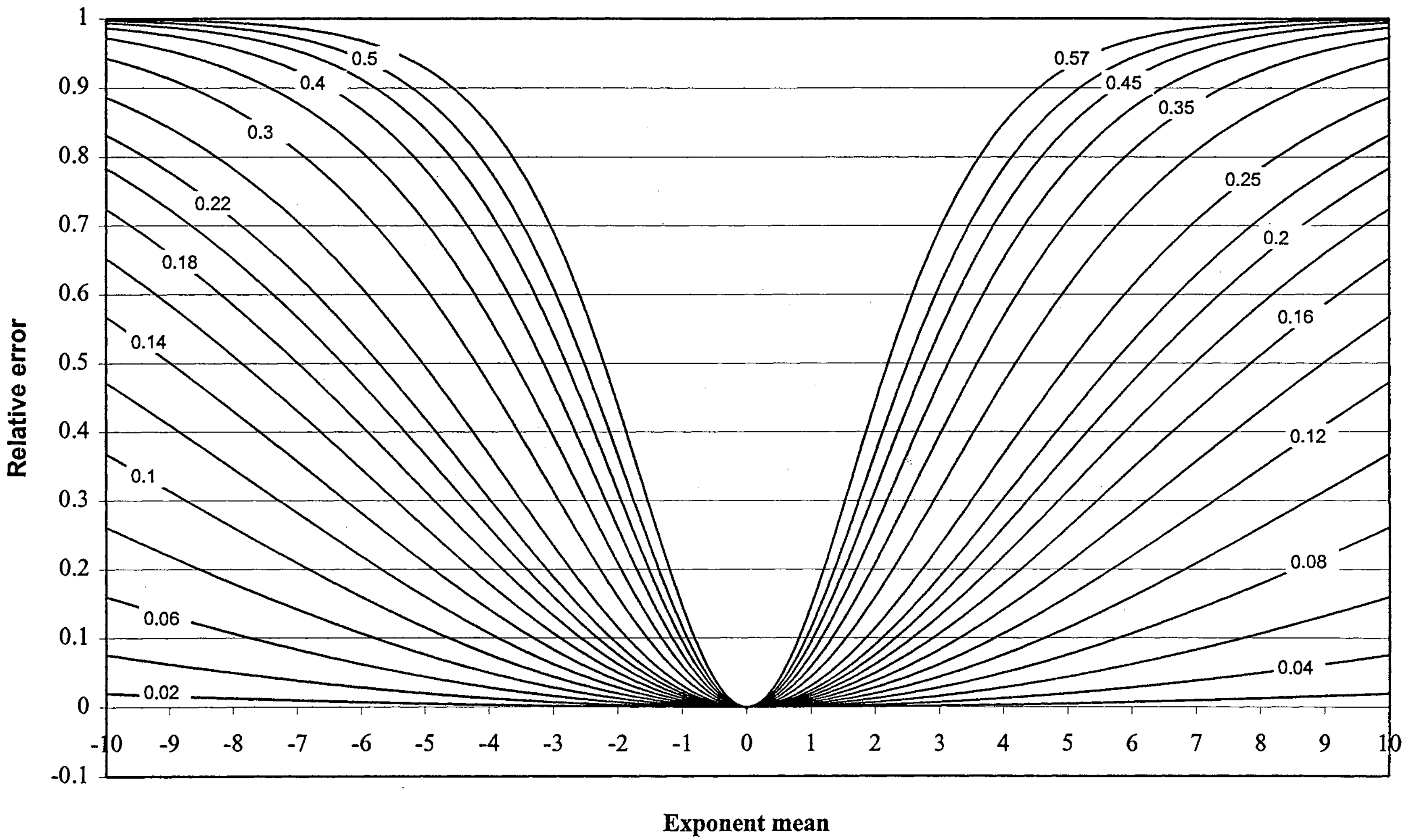

Fig. 4-1: Relative error in FOA predicted mean of an exponential function for CV_X ranging from 0.01to 0.57, where, exponent X is uniformly distributed

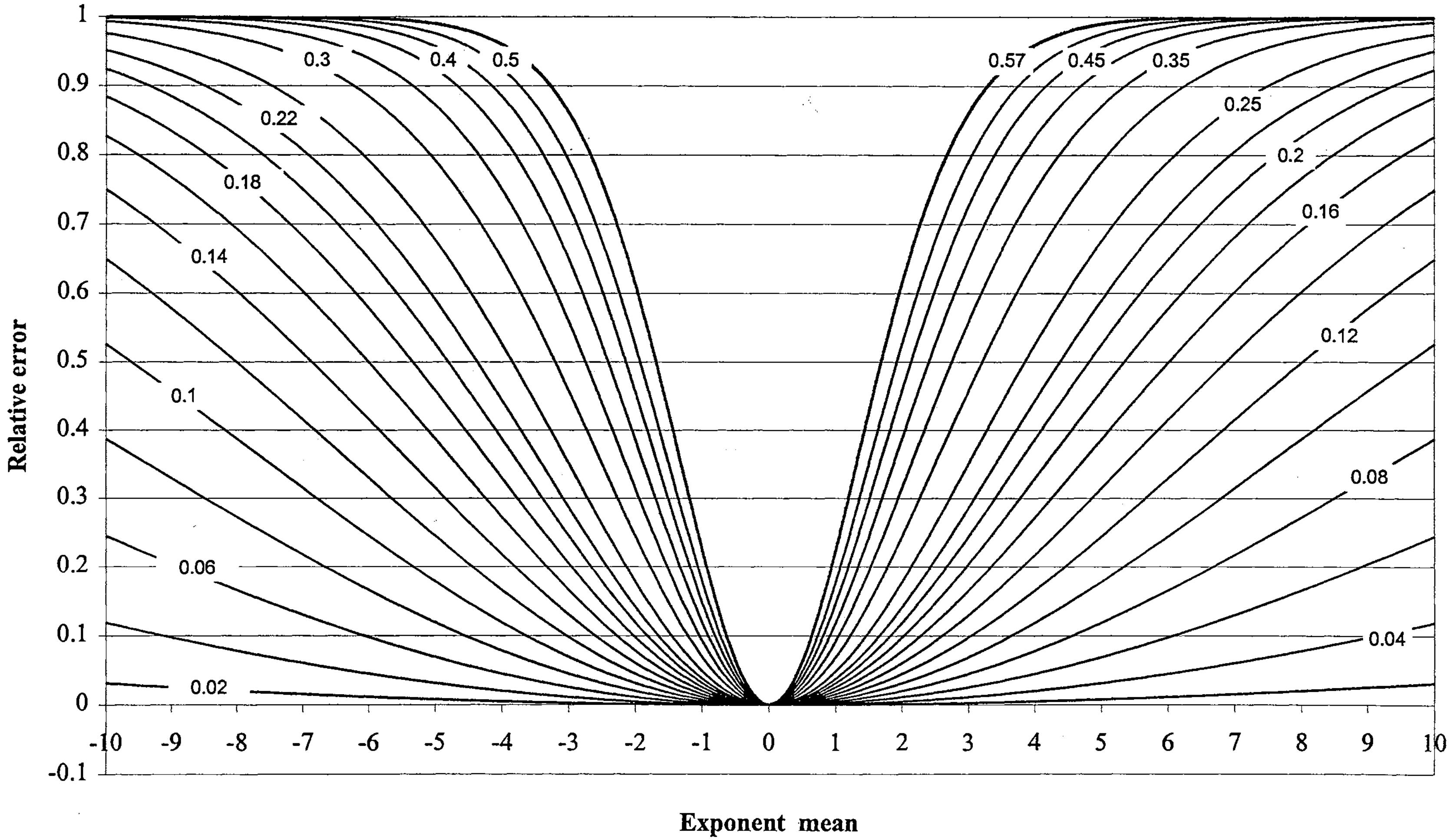

Fig. 4-2: Relative error in FOA predicted variance of an exponential function for CV_X ranging from 0.01to 0.57, where exponent X is unformly distributed

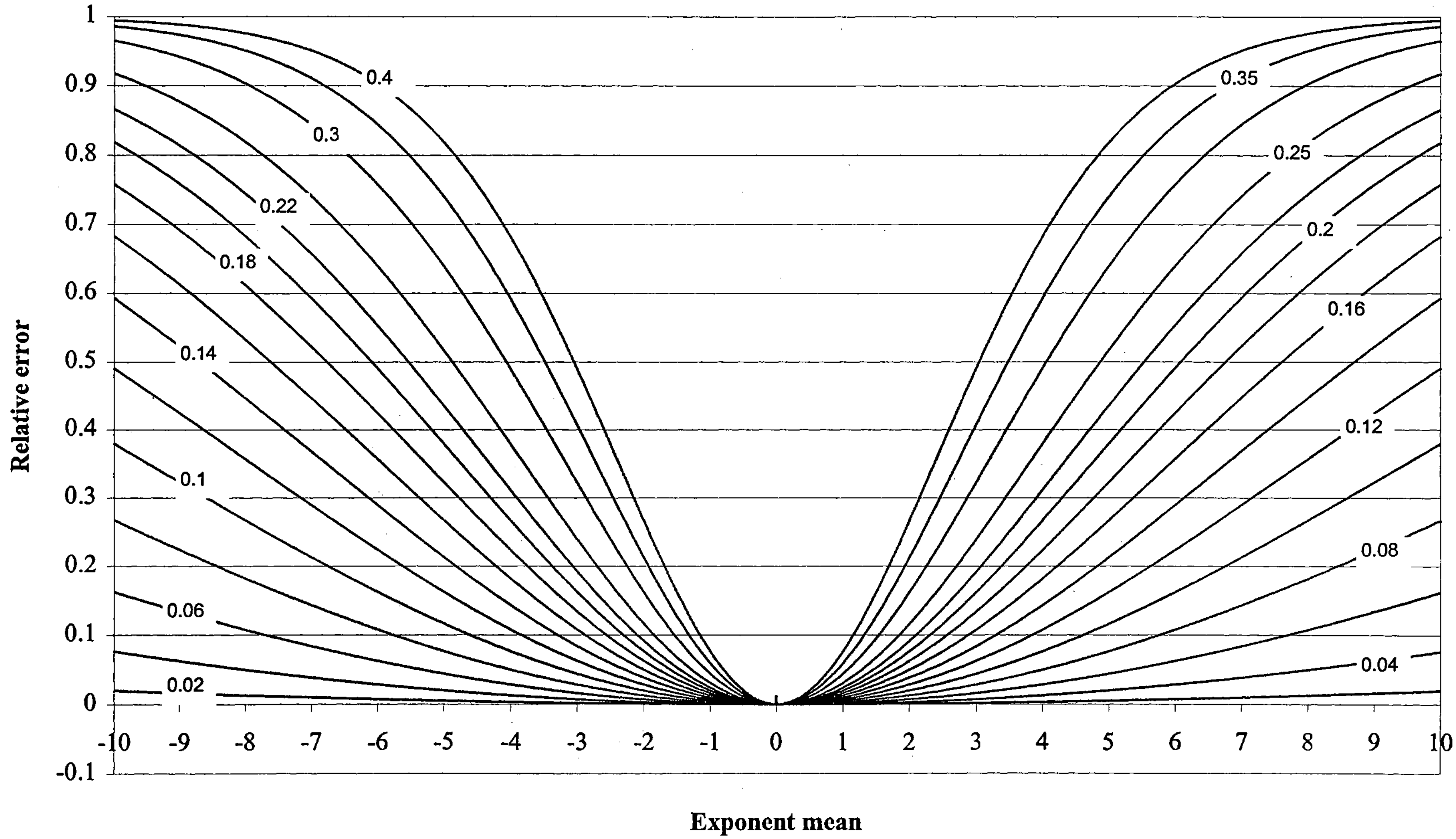

Fig. 4-3: Relative error in FOA predicted mean of an exponential function for CV_X ranging from 0.01to 0.40, where exponent X is triangularly distributed

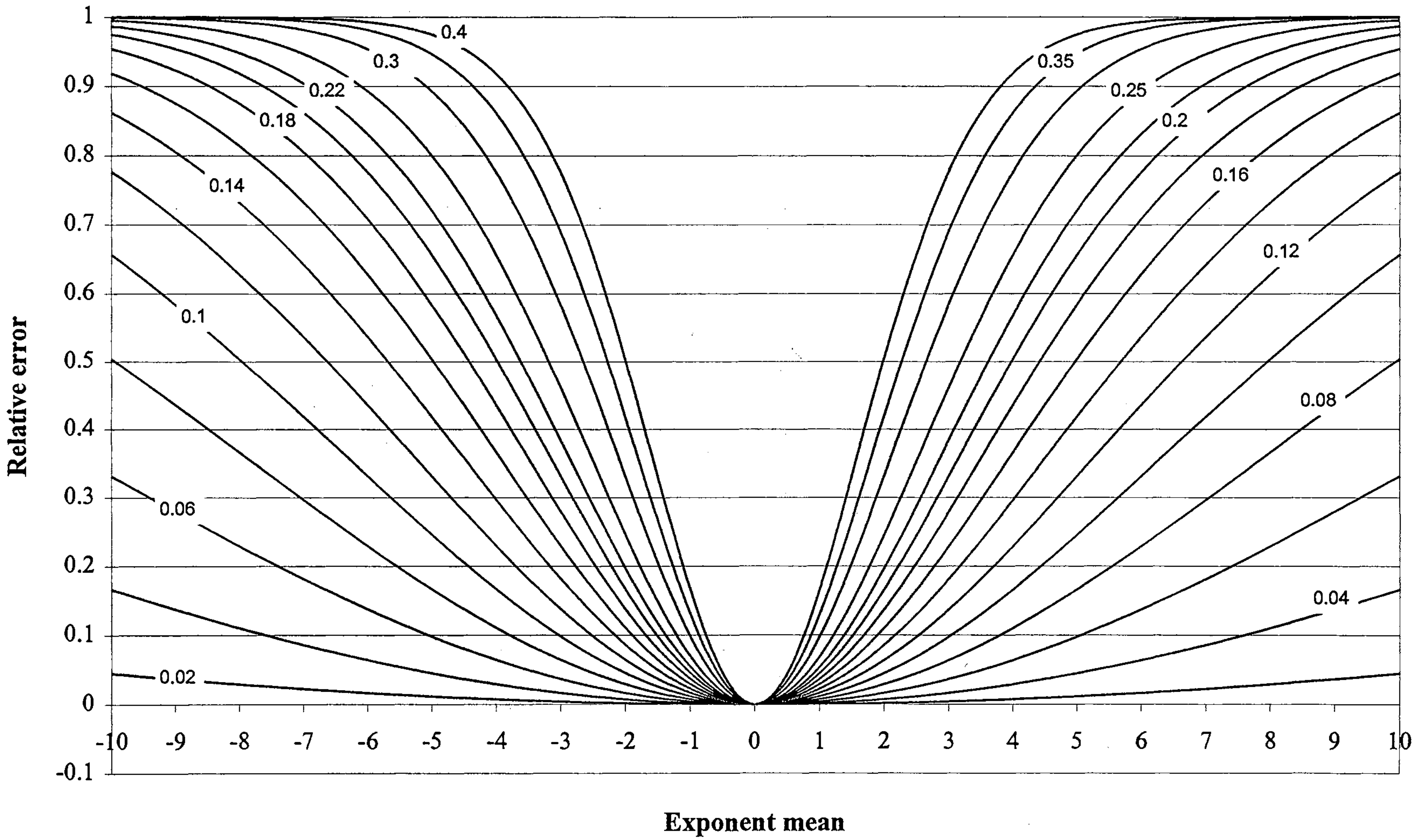

Fig. 4-4:Relative error in FOA predicted variance of an exponential function for CV_X ranging from 0.01to 0.40, where X exponent is triangularly distributed

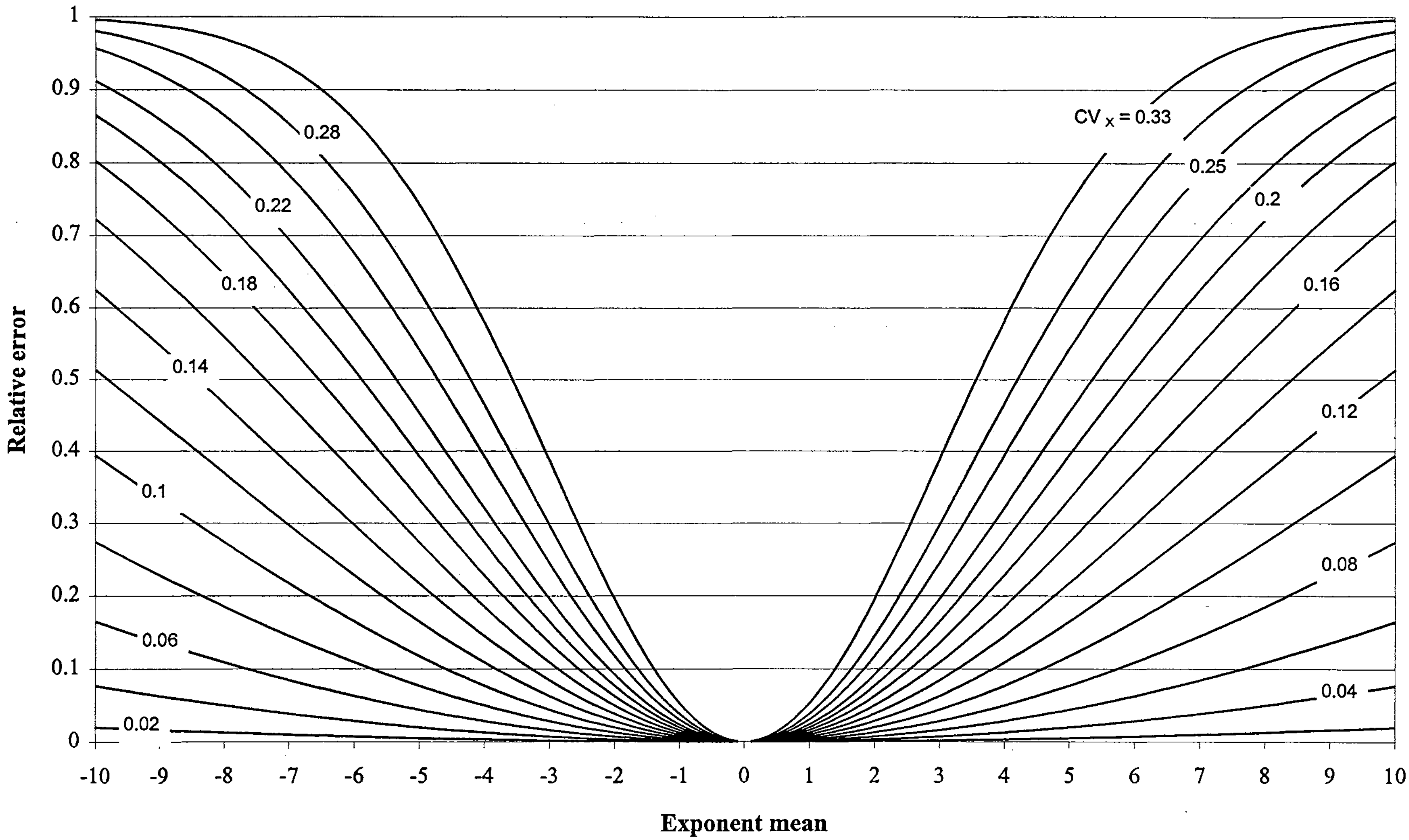

Fig.4-5: Relative error in FOA predicted mean of an exponential function for CV_X ranging from 0.01to 0.33, where exponent X is normally distributed

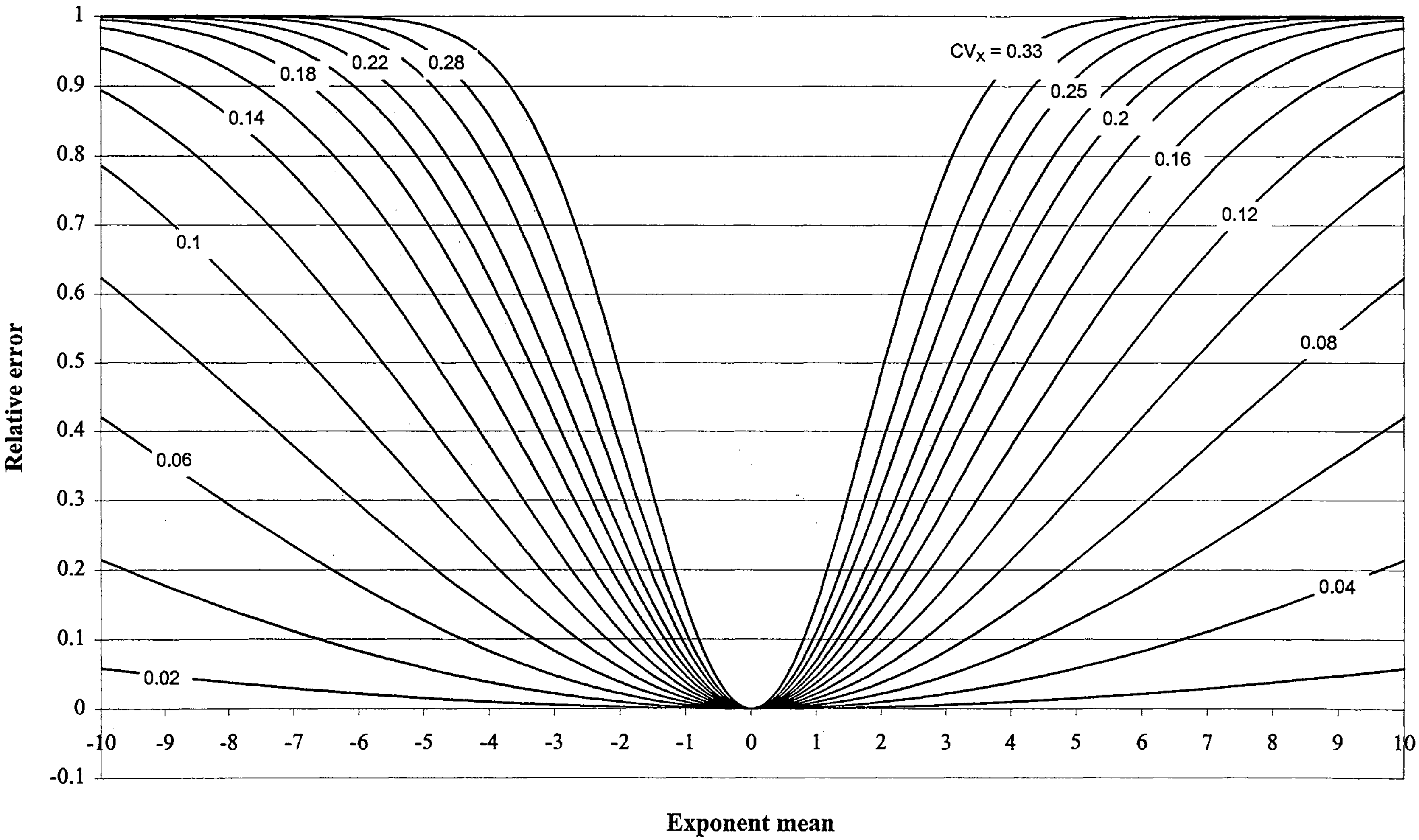

Fig. 4-6: Relative error in FOA predicted variance of an exponential function for CV_X ranging from 0.01to 0.33, where exponent X is normally distributed

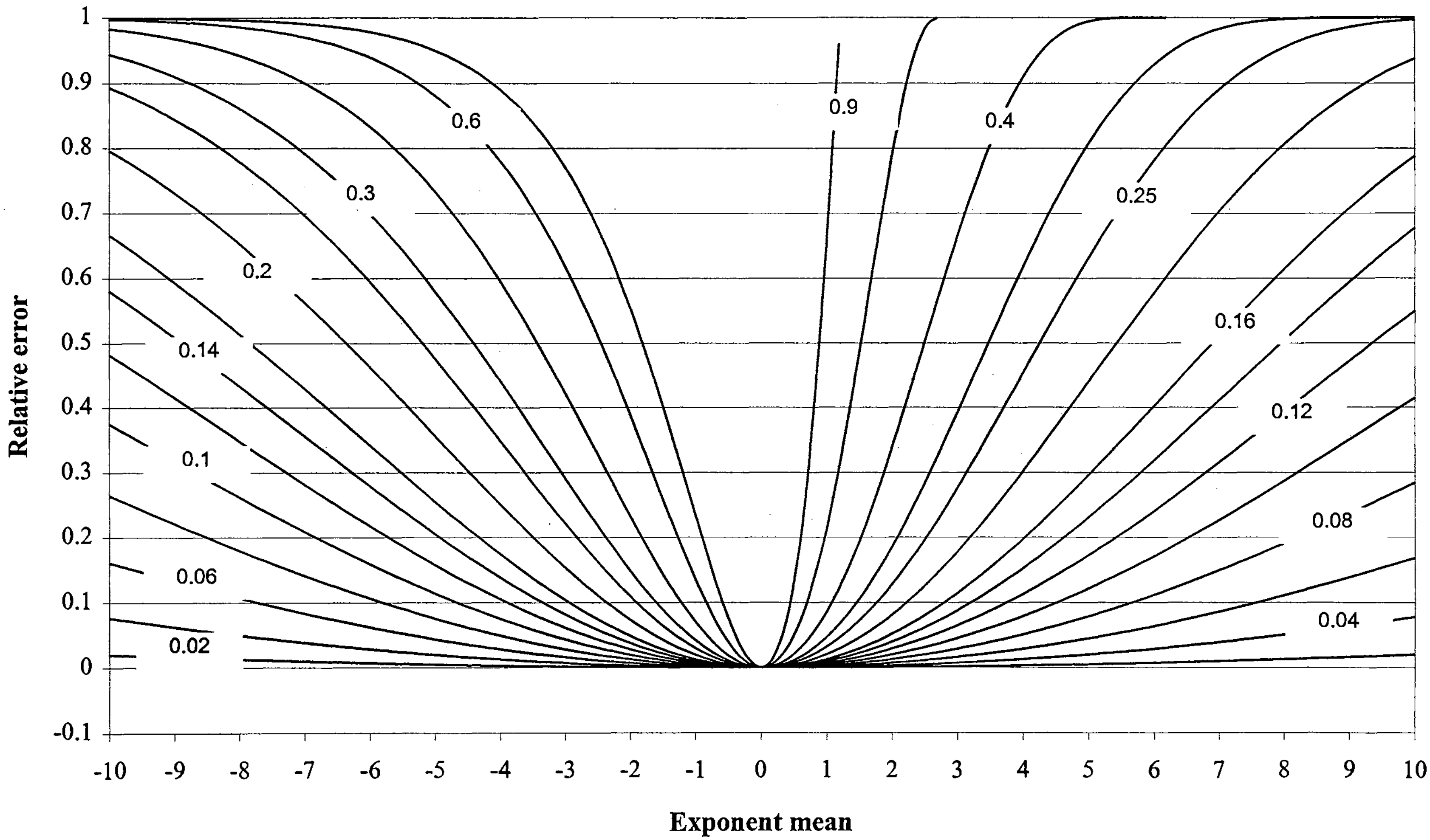

Fig. 4-7: Relative error in FOA predicted mean of an exponential function for CV_X ranging from 0.01to 0.90, where exponent X is from the gamma distribution

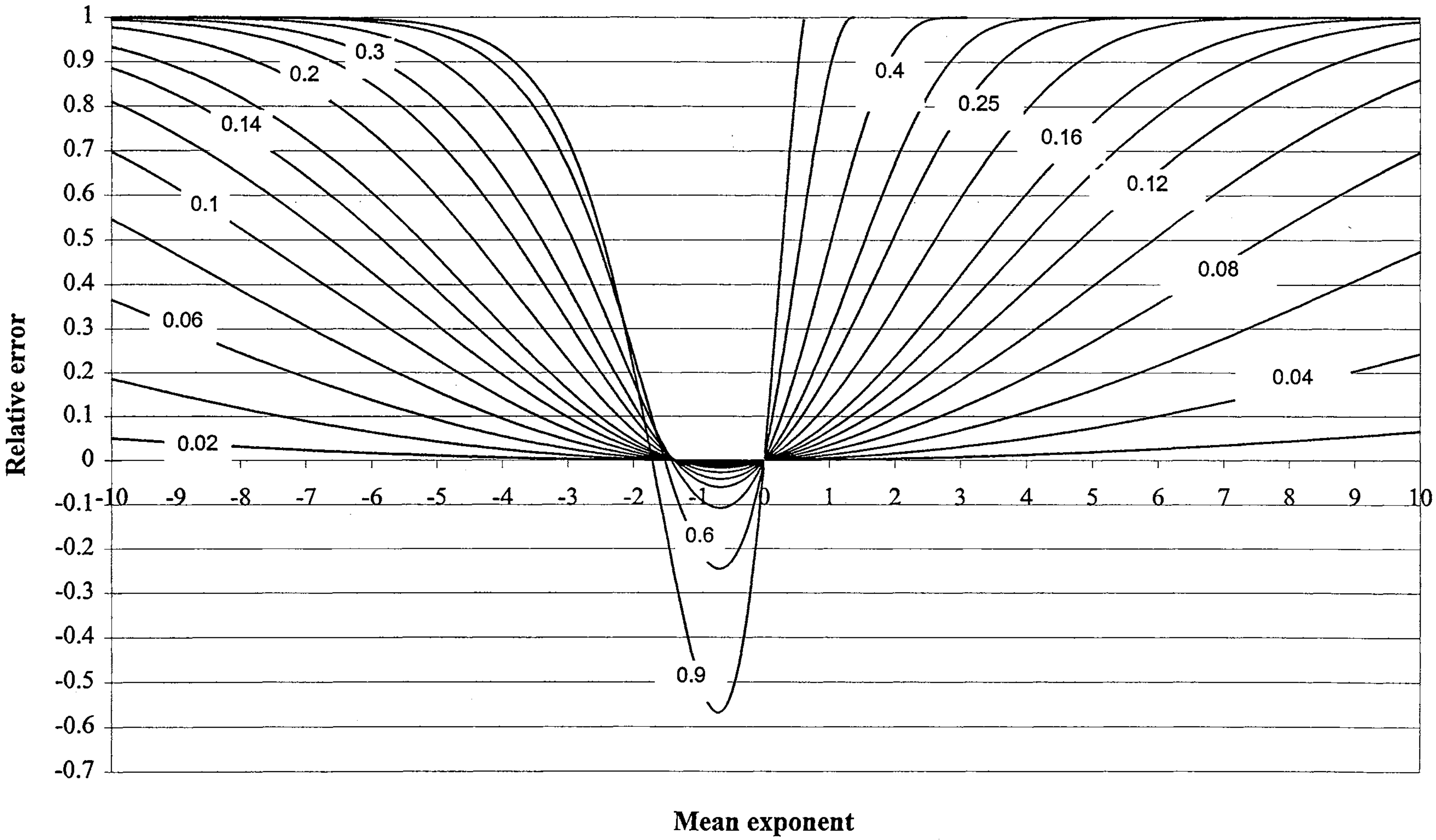

Fig. 4-8 : Relative error in FOA predicted variance of an exponential function for CV_X ranging 0.01to 0.90, where exponent X is from the gamma distribution

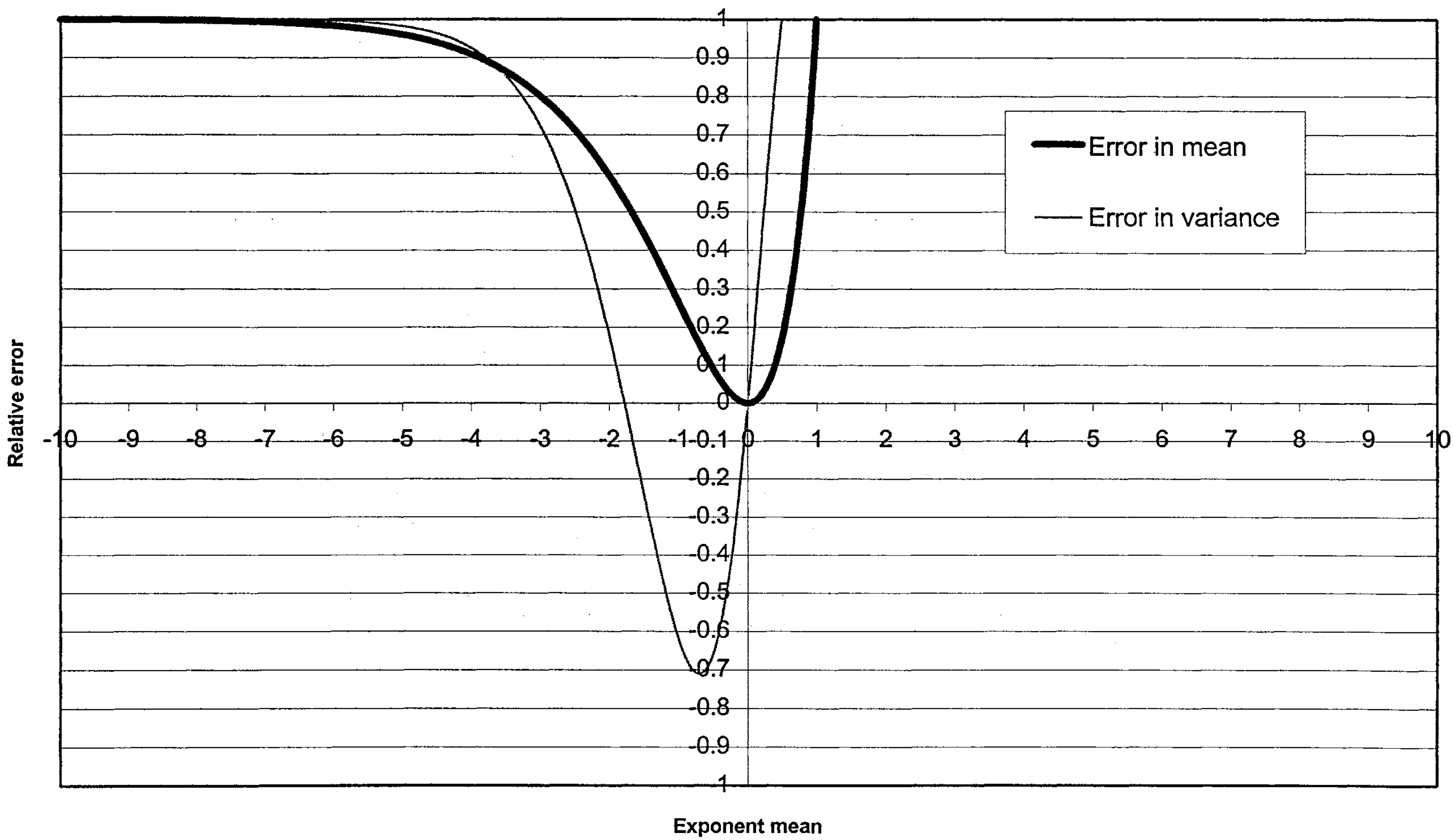

Fig. 4-9: Relative error in the FOA mean and variance of an exponential function, where exponent is from the exponential distribution

CHAPTER V

RELIABILITY ANALYSIS OF OPEN CHANNEL FLOW USING GENERIC EXPECTATION FUNCTIONS

Abstract

Traditionally, the location of a floodway boundary involves a solution of Manning's equation for the depth of flow for a given storm discharge corresponding to a design return period. In this approach, no uncertainty is considered in the parameters of Manning's equation or in the probability distribution of the observed maximum yearly flow. To incorporate uncertainties into the parameters of Manning's equation, researchers have used the first-order reliability method. This approach does not consider uncertainty in the probability distribution for maximum yearly flow even though the true distribution is not known. Using the generic expectation function approach, GEFA, the exact observed sample statistics can be incorporated in the determination of uncertainty in the depth of inundation. This method is easy in application because transformation of non-normal distributions and determination of the failure point using non-linear constrained optimization are not required. Furthermore, by applying GEFA, one can estimate the exact higher order moments of a performance function. Based on these moments, one can choose an appropriate distribution to improve understanding of the performance function in comparison to a lump parameter such as the reliability index. In this paper, a

comprehensive study is carried out using GEFA employing Manning's equation for a compound channel. The impact of parameter uncertainties on the depth of inundation is investigated. The reliability estimates obtained using GEFA are compared with those from the first-order reliability method, FORM, assuming different distributions for the design discharge.

Introduction

Flood plains are subject to periodic inundation that may result in loss of life and property, health and safety hazards, disruption of commerce and governmental services, extraordinary public expenditures for flood protection and relief. The boundary of a flood plain may vary according to the frequency of the flooding event, such as a 10-year, a 50-year, or a 100-year flood. Flood plain mapping is an inherently complicated process, full of uncertainties due to complexities in the hydrological/hydraulic models used, the availability and quality of data, and the subjectivity of human judgement in the process (Burges, 1979; Jones, 1980).

Traditionally, location of the floodway boundary involves a solution of Manning's equation for the depth of flow for a given storm discharge corresponding to a design return period. This approach accounts for uncertainty only in the peak flow and does not allow consideration of any uncertainty in the parameters of Manning's equation or in the probability distribution of the observed maximum yearly flow. To incorporate uncertainties in the parameters of Manning's equation, Cesare (1991) used the first-order reliability method (FORM). The use of FORM for reliability analysis of open channel flow has some drawbacks. First, it is implicitly assumed that the maximum yearly flow has been exactly described by a theoretical probability distribution function. In most situations, however, the underlying probability distribution for the peak annual flow is seldom known with certainty. It involves subjectivity on the part of analyst, particularly in situations where it is difficult to choose among several equally suited distributions. Any error in fitting the distribution may affect the reliability/risk estimates severely. Second, FORM involves the transformation of non-normal, random input variables to their equivalent normally distributed random variables and the determination of the lineraziation point using a nonlinear optimization technique. This is generally not an easy task depending upon the nature and complexity of the system. Third, the magnitude of acceptable convergence may affect the accuracy of the reliability estimates. In some cases, the magnitude of convergence error may not be reduced below a certain level.

Using the generic expectation function approach, GEFA, (Chapter III), the exact observed sample statistics can be incorporated in determining the uncertainty in the depth of inundation. This method is easy in application because transformation of non-normal distributions and determination of the failure point using non-linear constrained optimization are not required. By applying GEFA, one can estimate the exact higher order moments of a performance function. Based on these moments one can choose an appropriate distribution and improve understanding of the performance function in comparison to a lumped parameter approach such as the reliability index (Hasofer and Lind, 1974).

In this paper, GEFA is employed to determine the impacts of the magnitude of the CV and distribution types for the Manning's coefficients, and distribution types used to describe maximum yearly flow data on the risk for a given level of flooding along a channel. The reliability estimates obtained using the GEFA are compared with those from FORM, assuming different distributions for the design discharge.

Reliability Analysis of Open Channel Flow

Manning's equation is the most commonly used resistance equation to determine the flow capacity of a channel section corresponding to a given depth (Chow, 1959). In SI units, it is expressed as

$$Q_C = \frac{1}{n} A R^{\frac{2}{3}} S^{\frac{1}{2}} \quad (5\text{-}1)$$

where Q_C is the flow in m^3/sec; A is the cross sectional area of the channel in m^2; R is the hydraulic radius of the channel in m; S is the channel slope and n is the Manning's coefficient. Natural channels often have a main channel section and an overbank section. Most flow occurs in the main channel. However, during flood events overbank flows may occur. Considering a symmetric river-flood plain section as shown in Figure 5-1, the overall flow capacity, Q_C, for the compound section can be expressed as

$$Q_C = \left(\frac{Y_c}{n_c} + 2\frac{Y_b}{n_b} \right) S^{\frac{1}{2}} = \left(Y_c n_c^{-1} + 2Y_b n_b^{-1} \right) S^{0.5} \quad (5\text{-}2)$$

where $Y = AR^{2/3}$ is called the section factor. Y_c and Y_b represent section factors for main channel and overbank sections, respectively. Assuming the geometry of the channel to be deterministic, Q_C is a random variable due to uncertainties in the Manning's coefficients n_c and n_b and slope of the channel. It is assumed that n_c, n_b and S are independent to each other.

The annual peak flow, Q_L, is also a random variable, which is usually represented by a theoretical probability distribution by fitting observed annual peak flood flow data.

The exceedance probability (or return period) of a certain depth of flow (flood level) can be estimated by determining the probability of the event ($Q_C < Q_L$). The objective of reliability analysis is to determine the probability of exceedance of a certain flood level in a flood zone. To study this event, a performance function, Z, is defined as (Ang and Tang, 1984)

$$Z = Q_C - Q_L \tag{5-3}$$

Substituting expression for Q_C from (5-2) into (5-3) and incorporating uncertain factors to account for the empirical nature of Manning's equation and observational errors, Z can be written as

$$Z = C_1\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)S^{0.5} - C_2 Q_L \tag{5-4}$$

where C_1 and C_2 are the modeling factors included to account for modeling uncertainty in Manning's equation and observational uncertainty in the observed maximum flow data respectively. The probability, P_f, that a certain depth of flow is exceeded, can be estimated as

$$P_f = P(Z < 0) = \int_{-\infty}^{0} p_Z(z)dZ \tag{5-5}$$

where P is the probability operator; and $p_Z(z)$ is the probability density function of Z. The probability distribution of Z is unknown and generally difficult to obtain. In most cases the exact distribution may not be required, as several distributions can be used to make reliability analysis if correct information about its moments is available. Furthermore, higher order moments are helpful in both identifying the candidate distributions for $p_Z(z)$ and using the distributions requiring higher order moments.

Various orders of moments of Z can be obtained by applying the theory of statistical expectation if general expressions for evaluating the moments, including negative and fractional moments, of its input variables are known. The generic expectation functions can be used to evaluate any order of moment of an input random variable.

Generic Expectation Functions

Consider a power function

$$g = X^r \tag{5-6}$$

The k^{th} order moment of g about the origin can be obtained as

$$\mu_k' = E\left[g^k\right] = E\left[X^{kr}\right] = \int_{-\infty}^{\infty} X^{kr} p_X(x) dX \tag{5-7}$$

where $p_X(x)$ is the probability density function of X. Assuming different distribution types for $p_X(x)$, generic expressions for $E[X^r]$ have been derived analytically in chapter III as listed in Table 5-1.

Data Used

For this study, 31 years of maximum yearly flow data (Table 5-2) observed at the Beargrass Creek, Cannons Lane, Louisville, Kentucky, is considered. Table 5-3 presents the frequency analysis for this data, which has been discussed in detail by Haan et al. (1994). Figure 5-2 presents a comparison of the lognormal and the extreme value type-I distributions for this data. From Figure 5-2, it can be noticed that it is difficult to select which distribution best describe the observed data. Furthermore, no distribution is able to represent the observed data exactly. To study the influence of distribution types for Q_L on the risk analysis of open channel flow, the lognormal distribution, extreme value type-I distribution, and the actual observed data are considered to represent Q_L in the performance function given in (5-4).

Johnson (1996) has summarized the coefficient of variations and the distribution types of different uncertain hydraulic variables. To study the impact of variation in CV and distribution types generally used for the Manning's coefficients, their CVs are varied from 0.05 to 0.3 with different types of distributions. The channel slope, S, is assumed to be fixed with its CV and distribution type. Table 5-4 presents the statistics of uncertain variables in (5-4).

Using traditional frequency methods the discharges corresponding to the 5, 10, 25, 50, 100, and 200-year return periods are calculated using the extreme value type-I distribution. Corresponding to these discharges, depths of flow are determined to be 2.23 m, 2.56 m, 2.92 m, 3.19 m, 3.45 m, and 3.69 m respectively.

Distributional Characteristics of the Performance Function

Consider two independent random variables, X_1 and X_2. Performing expectation operation on $(X_1 + X_2)^k$ and $(X_1 - X_2)^k$, the following equations are obtained.

$$E\left[(X_1 + X_2)^k\right] = E\left[X_1^k\right] + \binom{k}{1} E\left[X_1^{k-1}\right] E\left[X_2\right] + \binom{k}{2} E\left[X_1^{k-2}\right] E\left[X_2^2\right] + \ldots\ldots\ldots + E\left[X_2^k\right] \quad (5\text{-}8)$$

$$E\left[(X_1 - X_2)^k\right] = E\left[X_1^k\right] - \binom{k}{1} E\left[X_1^{k-1}\right] E\left[X_2\right] + \binom{k}{2} E\left[X_1^{k-2}\right] E\left[X_2^2\right] + \ldots\ldots\ldots + (-1)^k E\left[X_2^k\right] \quad (5\text{-}9)$$

For the sake of convenience, rewrite (5-4) as

$$Z = R - L \quad (5\text{-}10)$$

where $R = C_1\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)S^{0.5}$; and $L = C_2 Q_L$. The statistical moments of Z about the origin can be obtained using (5-9) if moments of R and L about the origin are known. The k^{th}-order moment of R can be obtained as

$$E\left[R^k\right] = E\left[C_1^k\right]E\left[\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)^k\right]E\left[S^{0.5k}\right] \quad (5\text{-}11)$$

In (5-11) $E\left[C_1^k\right]$ and $E\left[S^{0.5k}\right]$ can be directly evaluated using a generic expectation functions for given distributions tabulated in Table 5-1. To evaluate $E\left[\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)^k\right]$, (5-8) can be used. As C_2 is taken to be deterministic and equal to unity, the moments of L are equal to the moments of Q_L.

Having determined moments of Z about the origin, the k^{th}-central moment of Z, μ_k, can be obtained using the following equation (Haan, 1977)

$$\mu_k = E\left[(Z-\mu)^k\right] = \sum_{i=0}^{k}(-1)^i\binom{k}{i}\mu^i\mu'_{k-i} \quad (5\text{-}12)$$

where, μ is the mean of Z; μ'_{k-i} is the $(k\text{-}i)^{th}$ order moment of Z about the origin. Using the second and third order moments of Z, the coefficient of skewness, γ_Z, is defined as (Haan, 1977)

$$\gamma_Z = \frac{\mu_3}{\mu_2^{3/2}} \quad (5\text{-}13)$$

where μ_2, μ_3 are the second and third order moments of Z about the mean. The kurtosis of Z, κ_Z, is defined as

$$\kappa_Z = \frac{\mu_4}{\mu_2^2} \quad (5\text{-}14)$$

where μ_4 is the fourth moment of Z about the mean.

To demonstrate the calculation procedure, a flow depth of 3.19 m is considered. The corresponding section factors, Y_c and Y_b, are determined to be 947.37 $m^{8/3}$ and 0.094 $m^{8/3}$ respectively. The distributions of n_c and n_b are assumed to be lognormal with coefficient of variation of 0.25. Table 5-5 presents the calculation procedure of the moments of Z about the origin as well as about the mean. Using these calculated central moments, distributional characteristics (mean, variance, skewness, and kurtosis) of Z are determined. This procedure is repeated for each depth for an assumed set of parameters i.e., distribution types and coefficient of variations of n_c and n_b, and distribution of Q_L. The obtained distributional characteristics of Z are summarized in Tables 5-6 to 5-10.

It can be observed from Tables 5-6 through 5-10 that the type of distribution, e.g., the extreme value type-I, the lognormal, and the actual observed data, assumed to represent peak annual flow in the performance function did not affect the mean and standard deviation of Z. However, the coefficient of skewness and the coefficient of kurtosis of Z are affected by the type of distribution of Q_L. An interesting relationship has been observed among the coefficient of skewness and coefficient of kurtosis of Z when Q_L is represented by the extreme value type-I, the lognormal, and the actual data. The coefficient of skewness of Z, when Q_L is represented by the actual data, matches very closely to the coefficient of skewness of Z, when Q_L is assumed to be lognormally distributed. Similarly, the coefficient of kurtosis of Z, when Q_L is represented by the actual data, matches very closely to the coefficient of kurtosis of Z, when Q_L is assumed to have the extreme value type-I distribution.

Determining Exceedance Probabilities

In order to evaluate P_f using (5-5), $p_Z(z)$ is required. As mentioned earlier, most cases do not require knowledge of the exact distribution of Z, as several distributions can be used based on available distributional characteristics of Z. The other most commonly used non-Gaussian distribution is the Edgeworth asymptotic expansion (Cieslikiewicz, 1990). For most practical applications, the truncated four term Edgeworth expansion (Abramowitz and Stegun, 1972; Kendall et al., 1987; Tung, 1996) has been used, which is given as

$$P_f \approx \Phi(\xi) - \phi(\xi)\left[\frac{\gamma_Z}{6}\left(\xi^2 - 1\right) + \left(\frac{\kappa_Z - 3}{24}\right)\left(\xi^3 - 3\xi\right) + \left(\frac{\gamma_Z^2}{72}\right)\left(\xi^5 - 10\xi^3 - 15\xi\right)\right] \qquad (5\text{-}15)$$

where $\Phi(\xi)$ is the standard normal cumulative density function; $\phi(\xi)$ is the standard normal probability density function; and ξ is the standard normal variate. The use of the Edgeworth expansion has some drawbacks, particularly when used in tail portions of the distribution as it may give negative values for the probability density function and cumulative density function. Obviously, this violates the definition of a probability density function. To ensure that P_f does not become negative, the Edgeworth expansion is used for the first three smaller flow depths.

Observing the distributional characteristics of Z listed in Tables 5-6 through 5-10, it is noticed that most of the time the distribution of Z is negatively skewed and leptokurtic ($\kappa_Z > 3$). Given these constraints, the Pearson type-III distribution is the best choice among the most commonly used distribution functions. The Pearson type-III density function may be expressed as (Matalas and Wallis, 1973)

$$p_Z(z) = \frac{1}{|\alpha|\Gamma(\beta)}\left(\frac{z-\lambda}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{z-\lambda}{\alpha}\right)\right] \tag{5-16}$$

where α, β, and λ are the parameters. If the distribution is positively skewed, α is positive, and $Z \geq \lambda$; otherwise α is negative, and $Z \leq \lambda$. The parameters α, β, and λ are related to the distributional characteristics of random variable Z as follows

$$\mu_Z = \lambda + \alpha\beta \tag{5-17}$$

$$\sigma_Z = |\alpha|\beta^{\frac{1}{2}} \tag{5-18}$$

$$\gamma_Z = 2\beta^{-\frac{1}{2}} \tag{5-19}$$

The Pearson type-III distribution is always leptokurtic as indicated by

$$\kappa_Z = 3\gamma_Z^2 + 3 \tag{5-20}$$

Substituting (5-16) in (5-5), P_f is given as

$$P_f = \frac{1}{|\alpha|\Gamma(\beta)}\int_{\lambda}^{0}\left(\frac{z-\lambda}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{z-\lambda}{\alpha}\right)\right]dz \tag{5-21}$$

Substituting $w = \left(\frac{z-\lambda}{\alpha}\right)$, (5-21) is rewritten as

$$P_f = \frac{1}{\Gamma(\beta)}\int_{0}^{\frac{z-\lambda}{\alpha}} w^{\beta-1} \exp(-w)dw \tag{5-22}$$

From (Abramowitz and Stegun, 1972), P_f can be calculated as

$$P_f = F_{\chi^2}\left(\chi^2 / \upsilon\right) \tag{5-23}$$

where $F_{\chi^2}\left(\chi^2 / \upsilon\right)$ is the chi-square distribution with parameter χ^2 and υ degree of freedom which is given as

$$\upsilon = 2\beta \tag{5-24}$$

The parameter χ^2 is given as

$$\chi^2 = 2w = 2\left(\frac{z-\lambda}{\alpha}\right) \tag{5-25}$$

The P_f corresponding to higher flow depths which may fall in the tails of $p_Z(z)$ are determined using the Pearson type-III distribution. The calculated values of P_f using the Edgeworth expansion and the Pearson type-III distribution have been tabulated in Tables 5-11 through 5-13 for different settings of input variables. In Tables 5-10 and 5-11, the P_f values obtained using the FORM are also tabulated along with those obtained using GEFA in order to facilitate a comparison between the two methods. It is necessary to point out that the P_f values corresponding to GEFA are only an approximation of the exact values because approximate distributions are used while calculating them. It can be observed from Tables 5-10 and 5-11 that P_f values using GEFA match very closely with those of the FORM in a number of cases. For the CV value of 0.05 and the gamma distribution for n_c and n_b, the P_f values could not be calculated. The P_f values presented in Table 5-13 correspond to the case where Q_L is represented by the actual observed data in (5-4). As FORM can not be used in this case, only GEFA P_f values are tabulated in Table 5-13.

Impact of Parameter Distributions and Variation in CV Values

Figure 5-3 presents a typical plot between CV of Manning's coefficients (n_c and n_b) and P_f values for different flow depths using GEFA and FORM. The distribution types for the Manning's coefficients and Q_L correspond to the lognormal and the extreme value type-I, distributions, respectively. Figure 5-3 indicates that both GEFA and FORM are in close match with each other. Figure 5-3 also shows that as the magnitude of CV of

Manning's coefficients increases, the magnitude of the probability of exceedance corresponding to a given depth increases.

To visualize the influence of distribution type of Manning's coefficients on risk, Figures 5-4 and 5-5 present plots between depth and exceedance probability using different distributions for Manning's coefficients and CV values of 0.1 and 0.20, respectively. Figure 5-4 shows that when the CV of Manning's coefficients is small, the distribution type does not make much difference on the risk estimates. However, as the CV values increase, the influence of the distribution types of Manning's coefficients becomes more discernable. Figure 5-4 shows that assuming a uniform distribution for Manning's coefficients give the highest estimates, and the assumption of lognormal distribution gives the lowest estimates of risk for a given flow depth. The risk estimates corresponding to the normal and the triangular distributions match exactly. The risk estimates obtained using the gamma distribution are found to be slightly higher than those of the lognormal distribution.

Figures 5-6 and 5-7 show depth versus exceedance probability plots corresponding to the extreme value type-I distribution, the lognormal distribution, and the actual observed data for the peak annual flow using two different CV values of 0.1 and 0.2 for the Manning's coefficients. The curves corresponding to the extreme value type-I and the lognormal distributions have been derived using FORM, whereas, the curves corresponding to the observed data have been generated using the GEFA. Both plots Figures 5-6 and 5-7 indicate that there is significant influence of the type of distribution of the peak annual flow. On comparing the impacts of CV magnitude of Manning's coefficients, distribution types of the Manning's coefficients, and the distribution types of

Q_L on the risk estimates, the impact of distribution type for the peak annual flow is found to be much higher than that of the rest.

Figures 5-6 and 5-7 also compare the depth vs. exceedance probability plots obtained by traditional method corresponding to both the lognormal and the extreme value type-I distributions. It is clear from these plots that this method underestimates the exceedance probabilities. Consideration of uncertainty in Manning's roughness coefficients improves the risk estimates but it still underestimates the exceedance probabilities due to the fact that neither of the distributions is able to fit the given data exactly.

Conclusions

In this paper, reliability analysis of open channel flow is carried out using the generic expectation function approach. This method is simple and general in application. Using GEFA exact distributional characteristics and moments of any order of a performance function can be obtained. The exactness of risk estimates using GEFA depends upon satisfying the distributional characteristics of the performance function by its assumed or derived distribution. By comparison, the FORM estimates show that in most cases a commonly used distribution can be employed to evaluate an approximate risk. Using GEFA and FORM impacts due to the variation in magnitudes of CV and distribution types for the Manning's coefficients and distribution types for the peak annual flow are studied. It is observed that an increase in the CV values of the Manning's coefficients increases the risk estimate. Whereas, distribution types of Manning's coefficients at smaller CV have a negligible impact on risk, this becomes more pronounced at higher CVs for the Manning' coefficients. The impact of distribution types for the peak annual flow has been found to be the most prominent at both smaller and higher CVs of the Manning's coefficients. The problems due to distribution fitting can be removed by incorporating the actual observed data in the performance function. FORM does not provide any flexibility to incorporate actual observed data into the performance function. GEFA can be used irrespective whether the peak annual flow is represented by a distribution or by the actual observed data.

References

Abramowitz, M., and Stegun, I. A. (1972). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th ed., Dover Publications, New York.

Ang, A. H. S., and Tang, W. H. (1984). Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk, and Reliability. John Wiley & Sons, Inc., New York, N.Y.

Burges, S. J.(1979), Analysis of uncertainty in flood plain mapping, Water Resources Bulletin, 15(1), 227-243.

Cesare, M. A. (1991). First-order analysis of open-channel flow, Journal of Hydraulic Engineering Division, ASCE, 117(2), 242-247.

Chow, V. T. (1959). Open Channel Hydraulics, McGraw-Hill, New York, N.Y.

Cieslikiewicz, W. (1990). Determination of the surface elevation probability distribution of wind waves using maximum entropy principle, Water Waves Kinematics, eds. Torum and O. T. Gudmestad, NATO ASI, Ser. E: Appl. Sci., 178, Kluwer Acad. Publ., Dordrecht-Boston-London, 345-348.

Haan, C. T. (1977). Statistical Methods in Hydrology, The Iowa State University Press, Ames, Iowa.

Haan, C. T., Barfield, B. J., and Hayes, J. C. (1994). Design Hydrology and Sedimentology for Small Catchments. Academic Press, San Diego.

Hasofer, A. M., and Lind, N. C. (1974). Exact and invariant second-moment code format, Journal of the Engineering Mechanics Division, ASCE, 100(1), 111-121.

Johnson, P. A. (1996). Uncertainty of hydraulic parameters. Journal of Hydraulic Engineering, ASCE, 122(2), 112-114.

Jones, D. E. Jr. (1980). Approximate method for quick flood plain mapping. Journal of Water Resources Planning and Management, 106(2): 103-22.

Kendall, M., Stuart, A., and Ord, J. K. (1987). Kendall's Advanced Theory of Statistics, Vol. 1: Distribution Theory. 5th Ed., Oxford Unversity Press, New York, N.Y.

Matalas, N. C. and Wallis, J. R. (1973). Eureka! It fits a Pearson type-III distribution, Water Resources Research, Vol. 9(2), 281-289.

Tung, Y. K. (1996). Uncertainty analysis in water resources engineering. Stochastic Hydraulics'96, K. S. Tickle, I. C. Goulter, C. Xu, S. A. Wasimi, and F. Bouchart eds., A. A. Balkema, Rotterdam, the Netherlands., 29-46.

Table 5-1: Generic Expectation Functions for Some Commonly Used Probability Density Functions

Name	Generic expectation function, $E[X^r]$
Uniform	$\frac{\mu_X^r}{2\sqrt{3}(r+1)CV_X}\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right]$
Symmetrical triangular	$\frac{\mu_X^r}{6(r+1)(r+2)CV_X^2}\left[\left(1+CV_X\sqrt{6}\right)^{r+2}+\left(1-CV_X\sqrt{6}\right)^{r+2}-2\right]$
Unsymmetrical triangular*	$\frac{2\left[(b-c)a^{r+2}+(c-a)b^{r+2}+(a-b)c^{r+2}\right]}{(r+1)(r+2)(b-c)(c-a)(b-a)}$
Lognormal	$\mu_X^r\left(1+CV_X^2\right)^{\frac{r(r-1)}{2}}$
Gamma	$\frac{CV_X^{2r}\mu_X^r\Gamma\left(CV_X^{-2}+r\right)}{\Gamma(CV_X^{-2})}$
Exponential	$\mu_X^r\Gamma(r+1)$
Normal	$\mu_X^r\sum_{n=0}^{r/2}\binom{r}{2n}\frac{(2n)!}{2^n n!}CV_X^{2n}$, when r is an even positive integer; $\mu_X^r\sum_{n=0}^{(r-1)/2}\binom{r}{2n}\frac{(2n)!}{2^n n!}CV_X^{2n}$, when r is an odd positive integer; and when r is anything but a positive integer, $\mu_X^r\left[1+\frac{r(r-1)}{2}CV_X^2+......+\frac{r(r-1)(r-2).....(r-n+1)}{2^{n/2}(n/2)!}CV_X^n+...\right]$.

* b, a, and c are the maximum, minimum, and mode values of X, which can be obtained by substituting n = 0, 1, and 2 respectively in the relationship (Appendix II) $a,b,c=\mu_X\left\{1+2\sqrt{2}CV_X\cos\left[\frac{2\pi n}{3}+\frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_x\right)\right]\right\}$.

Table 5-2: Annual peak flows in Middle Fork, Beargrass Creek, Cannons Lane, Louisville, Kentucky

Year	Q (m^3/sec)	Year	Q (m^3/sec)	Year	Q (m^3/sec)
1945	51.26	1956	30.02	1967	20.16
1946	22.40	1957	42.20	1968	41.06
1947	23.76	1958	25.04	1969	20.02
1948	49.56	1959	37.38	1970	147.27
1949	25.43	1960	93.46	1971	60.89
1950	60.04	1961	67.97	1972	33.14
1951	34.55	1962	27.64	1973	58.91
1952	36.53	1963	26.00	1974	35.40
1953	21.75	1964	111.02	1975	64.29
1954	44.46	1965	32.57		
1955	35.12	1966	24.75		

Table 5-3: Flood frequency analysis for Middle Fork, Beargrass Creek, Cannons Lane, Louisville, Kentucky

Return Period (years)	Flood flow (m^3/esc)	
	Lognormal distribution	Extreme value type-I
5	62.33	65.79
10	80.36	82.47
25	105.36	103.54
50	125.52	119.17
100	146.92	134.68
200	169.69	150.14

Table 5-4: Statistical data used for channel analysis

Variable (1)	Distribution (2)	Mean (3)	CV (5)
n_c	Uniform, Triangular, Normal, Lognormal, Gamma	0.08	0.05, 0.10, 0.15, 0.20, 0.25, 0.30
n_b	Uniform, Triangular, Normal, Lognormal, Gamma	0.11	0.05, 0.10, 0.15, 0.20, 0.25, 0.30
C1	Normal	1	0.10
C2	Normal	1	0
S	Lognormal	1.012E-4	0.25
Q	Extreme value type - I, Lognormal, Observed data	45.279	0.629

Table 5-5: Computation of moments using generic expectation functions

Expectation	Order of expectation, *k*			
	1	2	3	4
$Y_c^k E\left[n_c^{-k}\right]$	1.26E+04	1.68E+08	2.39E+12	3.61E+16
$Y_b^k E\left[n_b^{-k}\right]$	0.90	0.87	0.88	0.96
$E\left[\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)^k\right]$	1.26E+04	1.68E+08	2.39E+12	3.61E+16
$E\left[C_1^k\right]$	1.00	1.01	1.03	1.06
$E\left[S^{0.5k}\right]$	9.98E-03	1.01E-04	1.04E-06	1.09E-08
$E\left[\left\{C_1\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)S^{0.5}\right\}^k\right]$	125.63	1.72E+04	2.56E+06	4.16E+08
$E\left[Q_L^k\right]$	45.29	2.86E+03	2.30E+05	2.25E+07
$E\left[\left\{C_1\left(Y_c n_c^{-1} + 2Y_b n_b^{-1}\right)S^{0.5} - Q_L\right\}^k\right]$	80.34	8.68E+03	1.08E+06	1.54E+08
Central moments of *Z*	0	2.22E+03	2.20E+04	1.94E+07
Statistics of Z: mean = 80.34, standard deviation = 47.16, CV = 0.59, Coefficient of skew = 0.21, and coefficient of kurtosis = 3.93				

Table 5-6: Distributional characteristics of Z where Manning's coefficients are assumed to be for lognormally distributed with CV range of 0.05-0.30

Depth (m)	Extreme value type - I				Lognormal				Observed data			
	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z
CV = 0.05												
2.23	20.17	30.53	-0.91	4.83	20.16	30.52	-1.72	9.89	20.17	30.53	-1.69	4.54
2.56	36.75	31.63	-0.80	4.59	36.75	31.62	-1.53	8.99	36.75	31.63	-1.50	4.35
2.92	57.70	33.30	-0.65	4.31	57.70	33.29	-1.28	7.89	57.70	33.30	-1.26	4.11
3.19	73.25	34.71	-0.55	4.13	73.24	34.71	-1.10	7.15	73.25	34.71	-1.08	3.96
3.45	88.68	36.25	-0.45	3.97	88.67	36.25	-0.93	6.51	88.68	36.25	-0.92	3.82
3.69	104.05	37.89	-0.36	3.83	104.05	37.89	-0.78	5.97	104.05	37.89	-0.77	3.71
CV = 0.10												
2.23	20.66	31.10	-0.84	4.70	20.65	31.09	-1.61	9.41	20.66	31.10	-1.59	4.44
2.56	37.36	32.49	-0.71	4.45	37.36	32.48	-1.38	8.39	37.36	32.49	-1.36	4.22
2.92	58.47	34.58	-0.54	4.16	58.47	34.58	-1.10	7.23	58.47	34.58	-1.08	3.98
3.19	74.13	36.34	-0.42	3.98	74.13	36.34	-0.90	6.50	74.13	36.34	-0.89	3.84
3.45	89.68	38.23	-0.32	3.84	89.68	38.23	-0.73	5.90	89.68	38.23	-0.72	3.72
3.69	105.17	40.23	-0.22	3.73	105.16	40.23	-0.57	5.40	105.17	40.23	-0.56	3.63
CV = 0.15												
2.23	21.47	32.06	-0.74	4.53	21.47	32.06	-1.44	8.69	21.47	32.06	-1.42	4.29
2.56	38.39	33.93	-0.57	4.26	38.38	33.93	-1.16	7.57	38.39	33.93	-1.14	4.07
2.92	59.76	36.70	-0.37	3.99	59.75	36.70	-0.84	6.41	59.76	36.70	-0.83	3.85
3.19	75.61	39.00	-0.24	3.84	75.61	38.99	-0.63	5.75	75.61	39.00	-0.62	3.74
3.45	91.35	41.43	-0.13	3.74	91.35	41.43	-0.45	5.23	91.35	41.43	-0.44	3.66
3.69	107.03	43.99	-0.03	3.67	107.03	43.99	-0.30	4.84	107.03	43.99	-0.29	3.60
CV = 0.20												
2.23	22.61	33.44	-0.59	4.35	22.61	33.43	-1.21	7.86	22.61	33.44	-1.19	4.15
2.56	39.82	35.96	-0.39	4.09	39.81	35.95	-0.89	6.72	39.82	35.96	-0.87	3.95
2.92	61.56	39.63	-0.17	3.89	61.55	39.62	-0.54	5.67	61.56	39.63	-0.52	3.79
3.19	77.68	42.62	-0.02	3.81	77.68	42.62	-0.32	5.14	77.68	42.62	-0.31	3.73
3.45	93.69	45.76	0.09	3.76	93.68	45.76	-0.15	4.76	93.69	45.76	-0.14	3.70
3.69	109.64	49.02	0.19	3.74	109.63	49.02	0.00	4.50	109.64	49.02	0.00	3.70
CV = 0.25												
2.23	24.08	35.25	-0.42	4.21	24.08	35.25	-0.94	7.05	24.08	35.25	-0.93	4.04
2.56	41.66	38.58	-0.18	4.03	41.66	38.58	-0.58	6.01	41.66	38.58	-0.57	3.91
2.92	63.87	43.34	0.07	3.93	63.86	43.34	-0.22	5.18	63.87	43.34	-0.21	3.86
3.19	80.34	47.16	0.21	3.93	80.34	47.16	-0.01	4.82	80.34	47.16	0.00	3.88
3.45	96.70	51.13	0.32	3.95	96.69	51.13	0.15	4.59	96.70	51.13	0.16	3.91
3.69	112.99	55.22	0.42	3.98	112.99	55.21	0.28	4.45	112.99	55.22	0.28	3.95
CV = 0.30												
2.23	25.88	37.52	-0.21	4.17	25.87	37.52	-0.64	6.38	25.88	37.52	-0.63	4.04
2.56	43.91	41.81	0.05	4.11	43.91	41.81	-0.26	5.55	43.91	41.81	-0.25	4.03
2.92	66.69	47.83	0.31	4.16	66.69	47.83	0.10	5.00	66.69	47.83	0.10	4.11
3.19	83.59	52.60	0.45	4.23	83.59	52.60	0.29	4.81	83.59	52.60	0.29	4.20
3.45	100.37	57.51	0.56	4.32	100.37	57.50	0.43	4.72	100.37	57.51	0.44	4.29
3.69	117.09	62.51	0.64	4.39	117.08	62.50	0.54	4.68	117.09	62.51	0.55	4.38

Table 5-7: Distributional characteristics of Z where Manning's coefficients are assumed to be for normally distributed with CV range of 0.05-0.30

Depth (m)	Extreme value type - I				Lognormal				Observed data			
	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z
					CV = 0.05							
2.23	20.17	30.53	-0.91	4.83	20.16	30.52	-1.72	9.89	20.17	30.53	-1.69	4.54
2.56	36.75	31.63	-0.80	4.59	36.75	31.63	-1.52	8.99	36.75	31.63	-1.50	4.35
2.92	57.71	33.30	-0.65	4.31	57.70	33.30	-1.28	7.89	57.71	33.30	-1.26	4.11
3.19	73.25	34.72	-0.55	4.13	73.25	34.72	-1.10	7.15	73.25	34.72	-1.08	3.96
3.45	88.68	36.26	-0.45	3.97	88.68	36.26	-0.93	6.51	88.68	36.26	-0.92	3.82
3.69	104.05	37.90	-0.36	3.83	104.05	37.90	-0.78	5.96	104.05	37.90	-0.77	3.71
					CV = 0.10							
2.23	20.68	31.15	-0.83	4.70	20.67	31.14	-1.60	9.37	20.68	31.15	-1.57	4.44
2.56	37.39	32.56	-0.70	4.44	37.39	32.56	-1.36	8.36	37.39	32.56	-1.34	4.22
2.92	58.51	34.69	-0.52	4.16	58.50	34.69	-1.07	7.20	58.51	34.69	-1.06	3.99
3.19	74.17	36.48	-0.40	4.00	74.17	36.47	-0.87	6.48	74.17	36.48	-0.86	3.86
3.45	89.72	38.40	-0.29	3.86	89.72	38.39	-0.69	5.89	89.72	38.40	-0.68	3.75
3.69	105.22	40.43	-0.19	3.76	105.21	40.42	-0.54	5.41	105.22	40.43	-0.53	3.67
					CV = 0.15							
2.23	21.58	32.34	-0.68	4.54	21.58	32.34	-1.36	8.56	21.58	32.34	-1.34	4.31
2.56	38.53	34.35	-0.49	4.32	38.52	34.34	-1.06	7.48	38.53	34.35	-1.04	4.14
2.92	59.94	37.31	-0.26	4.14	59.93	37.30	-0.71	6.41	59.94	37.31	-0.69	4.01
3.19	75.82	39.75	-0.11	4.07	75.81	39.75	-0.48	5.83	75.82	39.75	-0.47	3.97
3.45	91.58	42.34	0.02	4.05	91.58	42.34	-0.28	5.42	91.58	42.34	-0.27	3.97
3.69	107.29	45.05	0.13	4.05	107.28	45.04	-0.12	5.12	107.29	45.05	-0.11	3.99
					CV = 0.20							
2.23	23.02	34.60	-0.33	4.70	23.02	34.60	-0.89	7.77	23.02	34.60	-0.87	4.53
2.56	40.33	37.65	-0.03	4.84	40.32	37.64	-0.46	7.03	40.33	37.65	-0.45	4.72
2.92	62.20	42.03	0.31	5.20	62.19	42.02	0.00	6.61	62.20	42.03	0.00	5.12
3.19	78.42	45.56	0.51	5.51	78.41	45.56	0.27	6.54	78.42	45.56	0.27	5.46
3.45	94.52	49.25	0.68	5.83	94.52	49.24	0.49	6.58	94.52	49.25	0.49	5.79
3.69	110.56	53.04	0.82	6.13	110.56	53.04	0.67	6.68	110.56	53.04	0.67	6.10
					CV = 0.25							
2.23	25.32	39.07	0.10	5.24	25.32	39.06	-0.29	7.13	25.32	39.07	-0.28	5.13
2.56	43.22	43.98	0.43	5.82	43.21	43.98	0.16	7.00	43.22	43.98	0.17	5.75
2.92	65.82	50.80	0.75	6.58	65.82	50.80	0.58	7.24	65.82	50.80	0.58	6.54
3.19	82.59	56.17	0.93	7.09	82.59	56.17	0.80	7.53	82.59	56.17	0.80	7.06
3.45	99.24	61.66	1.06	7.51	99.23	61.66	0.96	7.82	99.24	61.66	0.96	7.49
3.69	115.82	67.24	1.16	7.86	115.82	67.23	1.08	8.08	115.82	67.24	1.09	7.85
					CV = 0.30							
2.23	28.99	44.52	-0.38	5.02	28.99	44.52	-0.64	6.13	28.99	44.52	-0.64	4.95
2.56	47.81	51.48	-0.30	5.45	47.81	51.48	-0.47	6.08	47.81	51.48	-0.46	5.42
2.92	71.59	60.91	-0.25	5.94	71.59	60.90	-0.35	6.26	71.59	60.91	-0.34	5.92
3.19	89.23	68.19	-0.22	6.23	89.23	68.19	-0.30	6.43	89.23	68.19	-0.29	6.21
3.45	106.75	75.56	-0.21	6.45	106.74	75.56	-0.26	6.59	106.75	75.56	-0.26	6.44
3.69	124.19	82.98	-0.20	6.63	124.19	82.98	-0.24	6.72	124.19	82.98	-0.24	6.63

Table 5-8: Distributional characteristics of Z where Manning's coefficients are assumed to be for uniformly distributed with CV range of 0.05-0.30

Depth (m)	Extreme value type - I				Lognormal				Observed data			
	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z
CV = 0.05												
2.23	20.16	30.52	-1.72	9.89	20.17	30.53	-0.91	4.83	20.17	30.53	-1.69	4.54
2.56	36.75	31.62	-1.53	8.99	36.75	31.63	-0.80	4.59	36.75	31.63	-1.50	4.34
2.92	57.70	33.30	-1.28	7.89	57.70	33.30	-0.65	4.31	57.70	33.30	-1.26	4.11
3.19	73.25	34.71	-1.10	7.15	73.25	34.72	-0.55	4.12	73.25	34.72	-1.08	3.95
3.45	88.68	36.25	-0.93	6.51	88.68	36.25	-0.45	3.96	88.68	36.25	-0.92	3.82
3.69	104.05	37.89	-0.78	5.96	104.05	37.90	-0.36	3.83	104.05	37.90	-0.77	3.71
CV =0.10												
2.23	20.66	31.11	-1.61	9.39	20.67	31.12	-0.84	4.70	20.67	31.12	-1.58	4.43
2.56	37.37	32.51	-1.38	8.37	37.38	32.52	-0.71	4.43	37.38	32.52	-1.36	4.21
2.92	58.49	34.62	-1.10	7.20	58.49	34.62	-0.54	4.14	58.49	34.62	-1.08	3.97
3.19	74.15	36.39	-0.90	6.46	74.16	36.39	-0.42	3.96	74.16	36.39	-0.89	3.81
3.45	89.70	38.29	-0.73	5.85	89.70	38.29	-0.32	3.81	89.70	38.29	-0.71	3.69
3.69	105.19	40.30	-0.57	5.36	105.20	40.31	-0.22	3.69	105.20	40.31	-0.56	3.60
CV =0.15												
2.23	21.53	32.16	-1.42	8.61	21.53	32.16	-0.73	4.49	21.53	32.16	-1.40	4.26
2.56	38.46	34.07	-1.15	7.46	38.47	34.08	-0.56	4.20	38.47	34.08	-1.13	4.02
2.92	59.85	36.91	-0.82	6.28	59.86	36.91	-0.37	3.91	59.86	36.91	-0.81	3.77
3.19	75.72	39.26	-0.62	5.59	75.73	39.26	-0.23	3.74	75.73	39.26	-0.60	3.64
3.45	91.47	41.75	-0.44	5.06	91.48	41.75	-0.12	3.62	91.48	41.75	-0.43	3.54
3.69	107.17	44.36	-0.29	4.66	107.17	44.36	-0.02	3.52	107.17	44.36	-0.28	3.46
CV =0.20												
2.23	22.82	33.77	-1.17	7.63	22.82	33.77	-0.57	4.24	22.82	33.77	-1.15	4.05
2.56	40.07	36.44	-0.84	6.44	40.08	36.44	-0.37	3.95	40.08	36.44	-0.83	3.81
2.92	61.88	40.31	-0.50	5.35	61.88	40.32	-0.14	3.68	61.88	40.32	-0.49	3.59
3.19	78.05	43.46	-0.29	4.78	78.06	43.47	-0.01	3.55	78.06	43.47	-0.28	3.48
3.45	94.11	46.77	-0.12	4.38	94.11	46.77	0.10	3.46	94.11	46.77	-0.11	3.41
3.69	110.10	50.18	0.01	4.10	110.11	50.19	0.20	3.40	110.11	50.19	0.02	3.36
CV =0.25												
2.23	24.61	36.13	-0.85	6.57	24.62	36.13	-0.36	3.99	24.62	36.13	-0.83	3.84
2.56	42.32	39.84	-0.49	5.47	42.33	39.84	-0.13	3.73	42.33	39.84	-0.48	3.63
2.92	64.70	45.10	-0.14	4.60	64.70	45.10	0.11	3.53	64.70	45.10	-0.13	3.47
3.19	81.30	49.30	0.05	4.20	81.31	49.30	0.24	3.45	81.31	49.30	0.06	3.41
3.45	97.78	53.64	0.20	3.94	97.78	53.65	0.35	3.41	97.78	53.65	0.20	3.38
3.69	114.20	58.09	0.31	3.77	114.20	58.09	0.43	3.38	114.20	58.09	0.32	3.36
CV =0.30												
2.23	27.06	39.56	-0.46	5.59	27.07	39.56	-0.09	3.79	27.07	39.56	-0.45	3.69
2.56	45.39	44.66	-0.10	4.73	45.40	44.67	0.16	3.62	45.40	44.67	-0.09	3.56
2.92	68.56	51.73	0.23	4.14	68.56	51.74	0.39	3.53	68.56	51.74	0.23	3.50
3.19	85.74	57.28	0.39	3.92	85.74	57.29	0.51	3.51	85.74	57.29	0.40	3.48
3.45	102.79	62.96	0.51	3.78	102.80	62.96	0.61	3.50	102.80	62.96	0.52	3.49
3.69	119.79	68.71	0.60	3.71	119.79	68.71	0.67	3.51	119.79	68.71	0.61	3.50

Table 5-9: Distributional characteristics of Z where Manning's coefficients are assumed to be for triangularly distributed with CV range of 0.05-0.30

Depth (m)	Extreme value type - I				Lognormal				Observed data			
	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z
CV =0.05												
2.23	20.17	30.53	-0.91	4.83	20.16	30.52	-1.72	9.89	20.17	30.53	-1.69	4.54
2.56	36.75	31.63	-0.80	4.59	36.75	31.62	-1.53	8.99	36.75	31.63	-1.50	4.34
2.92	57.71	33.30	-0.65	4.31	57.70	33.30	-1.28	7.89	57.71	33.30	-1.26	4.11
3.19	73.25	34.72	-0.55	4.13	73.25	34.72	-1.10	7.15	73.25	34.72	-1.08	3.95
3.45	88.61	36.25	-0.45	3.97	88.68	36.25	-0.93	6.51	88.68	36.26	-0.92	3.82
3.69	104.05	37.90	-0.36	3.83	104.05	37.90	-0.78	5.96	104.05	37.90	-0.77	3.71
CV =0.10												
2.23	20.67	31.13	-0.84	4.70	20.67	31.12	-1.60	9.38	20.67	31.13	-1.58	4.43
2.56	37.38	32.54	-0.70	4.44	37.38	32.53	-1.37	8.36	37.38	32.54	-1.35	4.22
2.92	58.50	34.66	-0.53	4.15	58.50	34.65	-1.09	7.20	58.50	34.66	-1.07	3.98
3.19	74.16	36.43	-0.41	3.97	74.16	36.43	-0.89	6.47	74.16	36.43	-0.87	3.83
3.45	89.64	38.34	-0.30	3.83	89.71	38.34	-0.71	5.86	89.71	38.34	-0.70	3.72
3.69	105.21	40.36	-0.21	3.72	105.20	40.36	-0.56	5.38	105.21	40.36	-0.55	3.63
CV =0.15												
2.23	21.56	32.24	-0.71	4.50	21.55	32.24	-1.40	8.57	21.56	32.24	-1.38	4.27
2.56	38.49	34.20	-0.54	4.23	38.49	34.19	-1.11	7.45	38.49	34.20	-1.09	4.05
2.92	59.89	37.09	-0.33	3.97	59.89	37.08	-0.78	6.29	59.89	37.09	-0.76	3.84
3.19	75.77	39.47	-0.19	3.83	75.76	39.47	-0.56	5.64	75.77	39.47	-0.55	3.73
3.45	91.45	42.00	-0.07	3.73	91.52	42.00	-0.38	5.15	91.53	42.01	-0.37	3.65
3.69	107.23	44.66	0.04	3.67	107.22	44.66	-0.22	4.78	107.23	44.66	-0.21	3.61
CV =0.20												
2.23	22.90	34.06	-0.51	4.30	22.90	34.06	-1.09	7.56	22.90	34.06	-1.07	4.11
2.56	40.18	36.87	-0.28	4.07	40.18	36.86	-0.74	6.45	40.18	36.87	-0.72	3.94
2.92	62.01	40.92	-0.02	3.92	62.01	40.92	-0.36	5.49	62.01	40.92	-0.35	3.83
3.19	78.20	44.21	0.13	3.87	78.20	44.21	-0.13	5.02	78.20	44.21	-0.13	3.81
3.45	94.20	47.64	0.26	3.86	94.28	47.64	0.05	4.71	94.28	47.65	0.05	3.81
3.69	110.30	51.20	0.37	3.87	110.29	51.20	0.19	4.51	110.30	51.20	0.20	3.83
CV =0.25												
2.23	24.86	37.02	-0.18	4.25	24.85	37.01	-0.63	6.59	24.86	37.02	-0.62	4.12
2.56	42.63	41.10	0.11	4.21	42.63	41.09	-0.22	5.76	42.63	41.10	-0.21	4.13
2.92	65.09	46.84	0.40	4.30	65.08	46.84	0.18	5.21	65.09	46.84	0.18	4.25
3.19	81.74	51.41	0.57	4.40	81.74	51.40	0.40	5.03	81.74	51.41	0.40	4.36
3.45	98.20	56.11	0.69	4.51	98.28	56.11	0.56	4.95	98.28	56.11	0.57	4.48
3.69	114.76	60.91	0.79	4.61	114.75	60.91	0.69	4.93	114.76	60.91	0.69	4.59
CV =0.30												
2.23	27.71	42.05	0.37	4.91	27.71	42.05	0.06	6.31	27.71	42.05	0.07	4.83
2.56	46.21	48.11	0.71	5.31	46.21	48.11	0.51	6.13	46.21	48.11	0.51	5.26
2.92	69.58	56.39	1.02	5.81	69.58	56.39	0.89	6.24	69.58	56.39	0.89	5.78
3.19	86.92	62.84	1.18	6.12	86.91	62.83	1.09	6.40	86.92	62.84	1.09	6.10
3.45	104.05	69.38	1.30	6.37	104.12	69.38	1.23	6.56	104.13	69.39	1.23	6.36
3.69	121.27	76.00	1.38	6.58	121.27	76.00	1.33	6.71	121.27	76.00	1.33	6.57

Table 5-10: Distributional characteristics of Z where Manning's coefficients are assumed to be gamma distributed with CV range of 0.05-0.30

Depth (m)	Extreme value type - I				Lognormal				Observed data			
	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z	μ_Z	σ_Z	γ_Z	κ_Z
CV =0.05												
2.23	20.17	30.53	-0.91	4.83	20.16	30.52	-1.72	9.89	20.17	30.53	-1.69	4.54
2.56	36.75	31.63	-0.80	4.59	36.75	31.62	-1.53	8.99	36.75	31.63	-1.50	4.35
2.92	57.70	33.30	-0.65	4.31	57.70	33.29	-1.28	7.89	57.70	33.30	-1.26	4.11
3.19	73.25	34.72	-0.55	4.13	73.24	34.71	-1.10	7.15	73.25	34.72	-1.08	3.96
3.45	88.68	36.25	-0.45	3.97	88.67	36.25	-0.93	6.51	88.68	36.25	-0.92	3.82
3.69	104.05	37.90	-0.36	3.83	104.05	37.89	-0.78	5.97	104.05	37.90	-0.77	3.71
CV =0.10												
2.23	20.66	31.11	-0.84	4.70	20.66	31.11	-1.61	9.40	20.66	31.11	-1.58	4.44
2.56	37.37	32.51	-0.71	4.44	37.37	32.51	-1.38	8.38	37.37	32.51	-1.36	4.22
2.92	58.48	34.62	-0.54	4.16	58.48	34.61	-1.09	7.22	58.48	34.62	-1.07	3.99
3.19	74.15	36.38	-0.42	3.98	74.14	36.38	-0.89	6.49	74.15	36.38	-0.88	3.84
3.45	89.69	38.28	-0.31	3.84	89.69	38.28	-0.72	5.89	89.69	38.28	-0.71	3.73
3.69	105.18	40.29	-0.21	3.74	105.18	40.29	-0.56	5.40	105.18	40.29	-0.55	3.64
CV =0.15												
2.23	21.51	32.14	-0.72	4.53	21.50	32.14	-1.42	8.65	21.51	32.14	-1.40	4.29
2.56	38.43	34.05	-0.55	4.26	38.42	34.04	-1.14	7.53	38.43	34.05	-1.12	4.08
2.92	59.81	36.87	-0.35	4.01	59.81	36.87	-0.80	6.39	59.81	36.87	-0.79	3.87
3.19	75.67	39.21	-0.21	3.88	75.67	39.20	-0.59	5.74	75.67	39.21	-0.58	3.77
3.45	91.42	41.69	-0.09	3.79	91.42	41.68	-0.41	5.24	91.42	41.69	-0.40	3.71
3.69	107.11	44.29	0.01	3.73	107.10	44.28	-0.25	4.87	107.11	44.29	-0.24	3.67
CV =0.20												
2.23	22.72	33.70	-0.55	4.36	22.72	33.69	-1.15	7.77	22.72	33.70	-1.13	4.17
2.56	39.95	36.34	-0.33	4.15	39.95	36.33	-0.81	6.67	39.95	36.34	-0.80	4.01
2.92	61.73	40.17	-0.09	4.01	61.72	40.17	-0.44	5.70	61.73	40.17	-0.43	3.91
3.19	77.88	43.29	0.07	3.97	77.88	43.28	-0.22	5.22	77.88	43.29	-0.21	3.90
3.45	93.91	46.55	0.19	3.97	93.91	46.55	-0.04	4.91	93.91	46.55	-0.03	3.92
3.69	109.89	49.94	0.30	3.99	109.88	49.93	0.11	4.70	109.89	49.94	0.12	3.95
CV =0.25												
2.23	24.36	35.92	-0.30	4.33	24.35	35.92	-0.80	6.97	24.36	35.92	-0.78	4.18
2.56	42.00	39.54	-0.03	4.29	42.00	39.54	-0.41	6.09	42.00	39.54	-0.39	4.19
2.92	64.30	44.68	0.24	4.37	64.29	44.68	-0.01	5.48	64.30	44.68	-0.01	4.31
3.19	80.84	48.80	0.40	4.48	80.83	48.79	0.21	5.26	80.84	48.80	0.21	4.44
3.45	97.25	53.05	0.53	4.60	97.25	53.05	0.38	5.16	97.25	53.05	0.38	4.57
3.69	113.61	57.42	0.63	4.72	113.61	57.41	0.51	5.12	113.61	57.42	0.51	4.69
CV =0.30												
2.23	26.46	39.01	0.04	4.71	26.46	39.01	-0.35	6.61	26.46	39.01	-0.34	4.61
2.56	44.64	43.90	0.35	5.01	44.63	43.90	0.08	6.20	44.64	43.90	0.09	4.95
2.92	67.61	50.70	0.65	5.46	67.60	50.69	0.47	6.12	67.61	50.70	0.48	5.42
3.19	84.65	56.04	0.81	5.76	84.64	56.04	0.68	6.21	84.65	56.04	0.69	5.74
3.45	101.56	61.51	0.94	6.03	101.56	61.51	0.84	6.34	101.56	61.51	0.84	6.01
3.69	118.41	67.07	1.03	6.26	118.41	67.07	0.95	6.47	118.41	67.07	0.96	6.24

Table 5-11: Comparison of exceedance probability obtained using GEFA and FORM

Depth (m)	Normal		Triangular		Uniform		Gamma		Lgnormal	
	GEFA	FORM	GEFA	FORM	GEFA	FORM	GEFA	FORM	GEFA	FORM
					CV =0.05					
2.23	0.2119	0.2188	0.2119	0.2188	0.2119	0.2196	0.2119		0.2119	0.2182
2.56	0.1131	0.1175	0.1132	0.1175	0.1132	0.1184	0.1131		0.1131	0.1170
2.92	0.0555	0.0524	0.0555	0.0524	0.0555	0.0531	0.0555		0.0555	0.0521
3.19	0.0235	0.0214	0.0240	0.0214	0.0245	0.0219	0.0242		0.0246	0.0213
3.45	0.0117	0.0108	0.0125	0.0108	0.0127	0.0111	0.0124		0.0128	0.0107
3.69	0.0059	0.0055	0.0064	0.0055	0.0069	0.0057	0.0066		0.0069	0.0054
					CV = 0.10					
2.23	0.2121	0.2215	0.2122	0.2215	0.2123	0.2244	0.2121	0.2198	0.2121	0.2189
2.56	0.1148	0.1204	0.1150	0.1205	0.1151	0.1236	0.1148	0.1193	0.1148	0.1186
2.92	0.0568	0.0547	0.0569	0.0549	0.0569	0.0572	0.0568	0.0541	0.0568	0.0538
3.19	0.0296	0.0232	0.0295	0.0234	0.0251	0.0249	0.0297	0.0228	0.0269	0.0227
3.45	0.0142	0.0120	0.0152	0.0121	0.0139	0.0131	0.0142	0.0118	0.0154	0.0117
3.69	0.0069	0.0062	0.0066	0.0063	0.0073	0.0069	0.0071	0.0061	0.0072	0.0061
					CV = 0.15					
2.23	0.2120	0.2255	0.2130	0.2257	0.2132	0.2314	0.2125	0.2219	0.2125	0.2200
2.56	0.1159	0.1250	0.1178	0.1254	0.1184	0.1312	0.1172	0.1225	0.1173	0.1212
2.92	0.0575	0.0584	0.0587	0.0590	0.0592	0.0631	0.0585	0.0571	0.0586	0.0564
3.19	0.0306	0.0259	0.0309	0.0265	0.0330	0.0291	0.0310	0.0253	0.0293	0.0249
3.45	0.0153	0.0139	0.0156	0.0143	0.0170	0.0159	0.0156	0.0135	0.0161	0.0133
3.69	0.0063	0.0074	0.0082	0.0077	0.0082	0.0087	0.0078	0.0073	0.0078	0.0072
					CV = 0.20					
2.23	0.2080	0.2308	0.2147	0.2312	0.2152	0.2398	0.2129	0.2246	0.2130	0.2213
2.56	0.1079	0.1308	0.1215	0.1318	0.1236	0.1402	0.1193	0.1267	0.1201	0.1244
2.92	0.0462	0.0632	0.0603	0.0643	0.0628	0.0702	0.0594	0.0611	0.0605	0.0598
3.19	0.0293	0.0295	0.0366	0.0306	0.0364	0.0342	0.0353	0.0286	0.0347	0.0280
3.45	0.0067	0.0163	0.0178	0.0171	0.0204	0.0194	0.0207	0.0159	0.0178	0.0156
3.69	0.0008	0.0091	0.0087	0.0096	0.0110	0.0110	0.0121	0.0089	0.0095	0.0086
					CV = 0.25					
2.23	0.2113	0.2369	0.2184	0.2376	0.2193	0.2490	0.2129	0.2276	0.2136	0.2226
2.56	0.1096	0.1375	0.1262	0.1391	0.1312	0.1500	0.1201	0.1315	0.1229	0.1280
2.92	0.0407	0.0687	0.0604	0.0705	0.0686	0.0780	0.0579	0.0657	0.0621	0.0638
3.19	0.0353	0.0338	0.0477	0.0353	0.0428	0.0401	0.0424	0.0326	0.0399	0.0316
3.45	0.0063	0.0193	0.0190	0.0205	0.0271	0.0234	0.0168	0.0188	0.0201	0.0183
3.69	0.0000	0.0111	0.0062	0.0119	0.0123	0.0137	0.0088	0.0110	0.0101	0.0108
					CV = 0.30					
2.23	0.2048	0.2435	0.2272	0.2446	0.2266	0.2584	0.2109	0.2307	0.2142	0.2239
2.56	0.1243	0.1449	0.1336	0.1471	0.1427	0.1602	0.1167	0.1367	0.1255	0.1318
2.92	0.0771	0.0748	0.0593	0.0772	0.0786	0.0864	0.0505	0.0709	0.0634	0.0681
3.19	0.0953	0.0386	0.0433	0.0407	0.0510	0.0465	0.0320	0.0371	0.0492	0.0356
3.45	0.0789	0.0228	0.0022	0.0244	0.0375	0.0280	0.0098	0.0223	0.0280	0.0215
3.69	0.0673	0.0135	0.0553	0.0147	0.0198	0.0169	0.0015	0.0135	0.0121	0.0133

Table 5-12: Comparison of exceedance probability obtained using GEFA and FORM

Depth (m)	Normal		Triangular		Uniform		Gamma		Lognormal	
	GEFA	FORM	GEFA	FORM	GEFA	FORM	GEFA	FORM	GEFA	FORM
CV = 0.05										
2.23	0.1406	0.1912	0.1406	0.1912	0.1406	0.1920	0.1406		0.1406	0.1907
2.56	0.0689	0.1058	0.0689	0.1058	0.0689	0.1065	0.0689		0.0689	0.1055
2.92	0.0586	0.0518	0.0586	0.0518	0.0586	0.0524	0.0586		0.0586	0.0516
3.19	0.0261	0.0260	0.0262	0.0260	0.0264	0.0264	0.0263		0.0264	0.0259
3.45	0.0196	0.0156	0.0198	0.0156	0.0200	0.0159	0.0199		0.0201	0.0155
3.69	0.0115	0.0096	0.0117	0.0096	0.0119	0.0097	0.0118		0.0120	0.0095
CV = 0.10										
2.23	0.1425	0.1936	0.1426	0.1937	0.1428	0.1964	0.1425	0.1922	0.1425	0.1914
2.56	0.0732	0.1081	0.0734	0.1082	0.0735	0.1107	0.0732	0.1072	0.0732	0.1067
2.92	0.0593	0.0536	0.0593	0.0537	0.0593	0.0554	0.0593	0.0530	0.0593	0.0527
3.19	0.0314	0.0273	0.0365	0.0274	0.0258	0.0286	0.0243	0.0270	0.0267	0.0268
3.45	0.0179	0.0165	0.0162	0.0166	0.0208	0.0174	0.0183	0.0163	0.0218	0.1620
3.69	0.0096	0.0102	0.0090	0.0103	0.0104	0.0109	0.0113	0.0101	0.0112	0.0100
CV = 0.15										
2.23	0.1451	0.1974	0.1465	0.1976	0.1470	0.2029	0.1458	0.1942	0.1459	0.1926
2.56	0.0781	0.1117	0.0802	0.1121	0.0810	0.1168	0.0794	0.1097	0.0795	0.1086
2.92	0.0589	0.0563	0.0600	0.0567	0.0605	0.0599	0.0599	0.0551	0.0601	0.0545
3.19	0.0383	0.0293	0.0356	0.0297	0.0394	0.0319	0.0301	0.0287	0.0380	0.0283
3.45	0.0189	0.0180	0.0206	0.0182	0.0197	0.0198	0.0212	0.0176	0.0200	0.0173
3.69	0.0101	0.0113	0.0107	0.0115	0.0123	0.0125	0.0116	0.0110	0.0115	0.0108
CV = 0.20										
2.23	0.1435	0.2023	0.1530	0.2028	0.1542	0.2109	0.1505	0.1969	0.1506	0.1940
2.56	0.0731	0.1165	0.0885	0.1172	0.0910	0.1242	0.0858	0.1130	0.0865	0.1110
2.92	0.0455	0.0599	0.0597	0.0607	0.0624	0.0655	0.0594	0.0580	0.0606	0.0569
3.19	0.0365	0.0320	0.0393	0.0327	0.0410	0.0360	0.0386	0.0310	0.0393	0.0303
3.45	0.0145	0.0199	0.0239	0.0205	0.0249	0.0227	0.0224	0.0193	0.0230	0.0189
3.69	0.0022	0.0127	0.0112	0.0131	0.0141	0.0146	0.0117	0.0123	0.0127	0.0120
CV = 0.25										
2.23	0.1591	0.2081	0.1637	0.2090	0.1654	0.2198	0.1559	0.2000	0.1565	0.1956
2.56	0.0817	0.1220	0.0979	0.1233	0.1039	0.1326	0.0908	0.1169	0.0937	0.1139
2.92	0.0352	0.0642	0.0573	0.0655	0.0663	0.0718	0.0559	0.0614	0.0608	0.0597
3.19	0.0510	0.0353	0.0509	0.0364	0.0496	0.0406	0.0450	0.0338	0.0443	0.0328
3.45	0.0184	0.0223	0.0211	0.0232	0.0309	0.0260	0.0213	0.0215	0.0251	0.0208
3.69	0.0016	0.0144	0.0103	0.0151	0.0172	0.0170	0.0111	0.0139	0.0131	0.0135
CV = 0.30										
2.23	0.1801	0.2145	0.1819	0.2158	0.1818	0.2291	0.1605	0.2033	0.1634	0.1972
2.56	0.1130	0.1282	0.1102	0.1301	0.1208	0.1414	0.0913	0.1212	0.1004	0.1170
2.92	0.0739	0.0690	0.0536	0.0708	0.0749	0.0786	0.0466	0.6530	0.0608	0.0629
3.19	0.0937	0.0390	0.0546	0.0405	0.0652	0.0457	0.0410	0.0371	0.0469	0.0357
3.45	0.0835	0.0251	0.0053	0.0263	0.0354	0.0298	0.0181	0.0240	0.0275	0.0231
3.69	0.0697	0.0164	0.0553	0.0174	0.0267	0.0198	0.0055	0.0159	0.0180	0.0153

Table 5-13: Exceedance probability obtained using GEFA

Depth (m)	Normal	Triangular	Uniform	Gamma	Lognormal
CV = 0.05					
2.23	0.2598	0.2598	0.2598	0.2598	0.2598
2.56	0.1432	0.1433	0.1433	0.1433	0.1433
2.92	0.0586	0.0586	0.0586	0.0586	0.0585
3.19	0.0231	0.0233	0.0234	0.0234	0.0235
3.45	0.0168	0.0170	0.0172	0.0171	0.0172
3.69	0.0094	0.0095	0.0097	0.0096	0.0098
CV = 0.10					
2.23	0.2528	0.2531	0.2535	0.2533	0.2535
2.56	0.1410	0.1413	0.1416	0.1413	0.1414
2.92	0.0626	0.0625	0.0624	0.0624	0.0623
3.19	0.0268	0.0314	0.0356	0.0336	0.0368
3.45	0.0210	0.0195	0.0168	0.0220	0.0176
3.69	0.0093	0.0091	0.0107	0.0114	0.0115
CV = 0.15					
2.23	0.2403	0.2429	0.2443	0.2434	0.2444
2.56	0.1353	0.1381	0.1393	0.1379	0.1386
2.92	0.0661	0.0669	0.0670	0.0665	0.0664
3.19	0.0329	0.0354	0.0307	0.0313	0.0298
3.45	0.0203	0.0215	0.0206	0.0199	0.0215
3.69	0.0105	0.0113	0.0116	0.0116	0.0116
CV = 0.20					
2.23	0.2165	0.2307	0.2346	0.2316	0.2342
2.56	0.1158	0.1337	0.1376	0.1331	0.1352
2.92	0.0552	0.0688	0.0711	0.0681	0.0690
3.19	0.0336	0.0413	0.0432	0.0410	0.0374
3.45	0.0158	0.0239	0.0239	0.0224	0.0220
3.69	0.0035	0.0108	0.0141	0.0117	0.0127
CV = 0.25					
2.23	0.2041	0.2197	0.2271	0.2191	0.2246
2.56	0.1073	0.1294	0.1384	0.1263	0.1319
2.92	0.0437	0.0662	0.0752	0.0649	0.0696
3.19	0.0480	0.0486	0.0496	0.0462	0.0444
3.45	0.0175	0.0271	0.0301	0.0254	0.0269
3.69	0.0015	0.0095	0.0183	0.0094	0.0135
CV = 0.30					
2.23	0.2066	0.2158	0.2250	0.2062	0.2166
2.56	0.1271	0.1285	0.1439	0.1161	0.1293
2.92	0.0792	0.0601	0.0822	0.0543	0.0688
3.19	0.0900	0.0531	0.0540	0.0379	0.0479
3.45	0.0799	0.0667	0.0321	0.0494	0.0301
3.69	0.0670	0.0553	0.0251	0.0052	0.0162

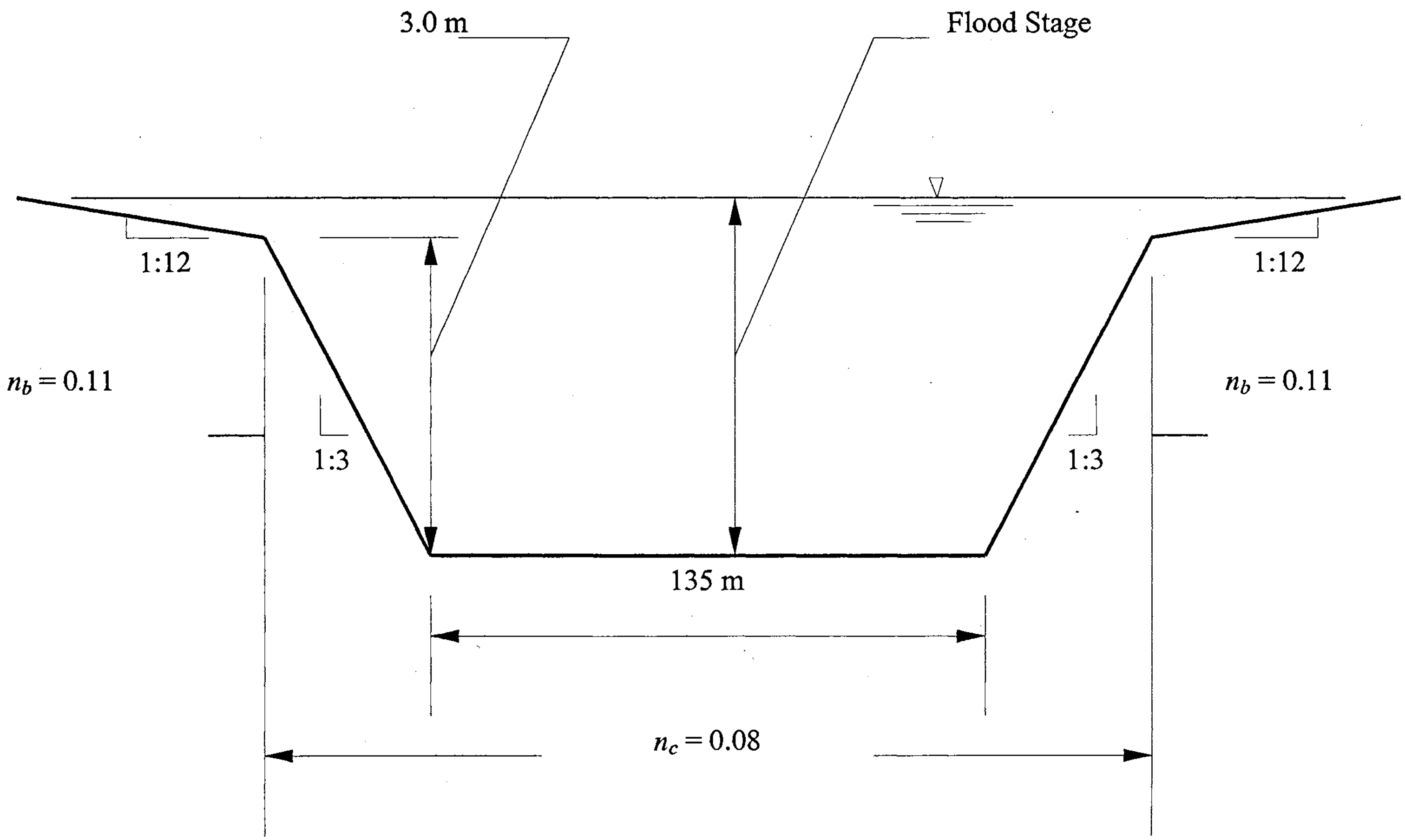

Fig. 5-1: Geometry of considered compound channel

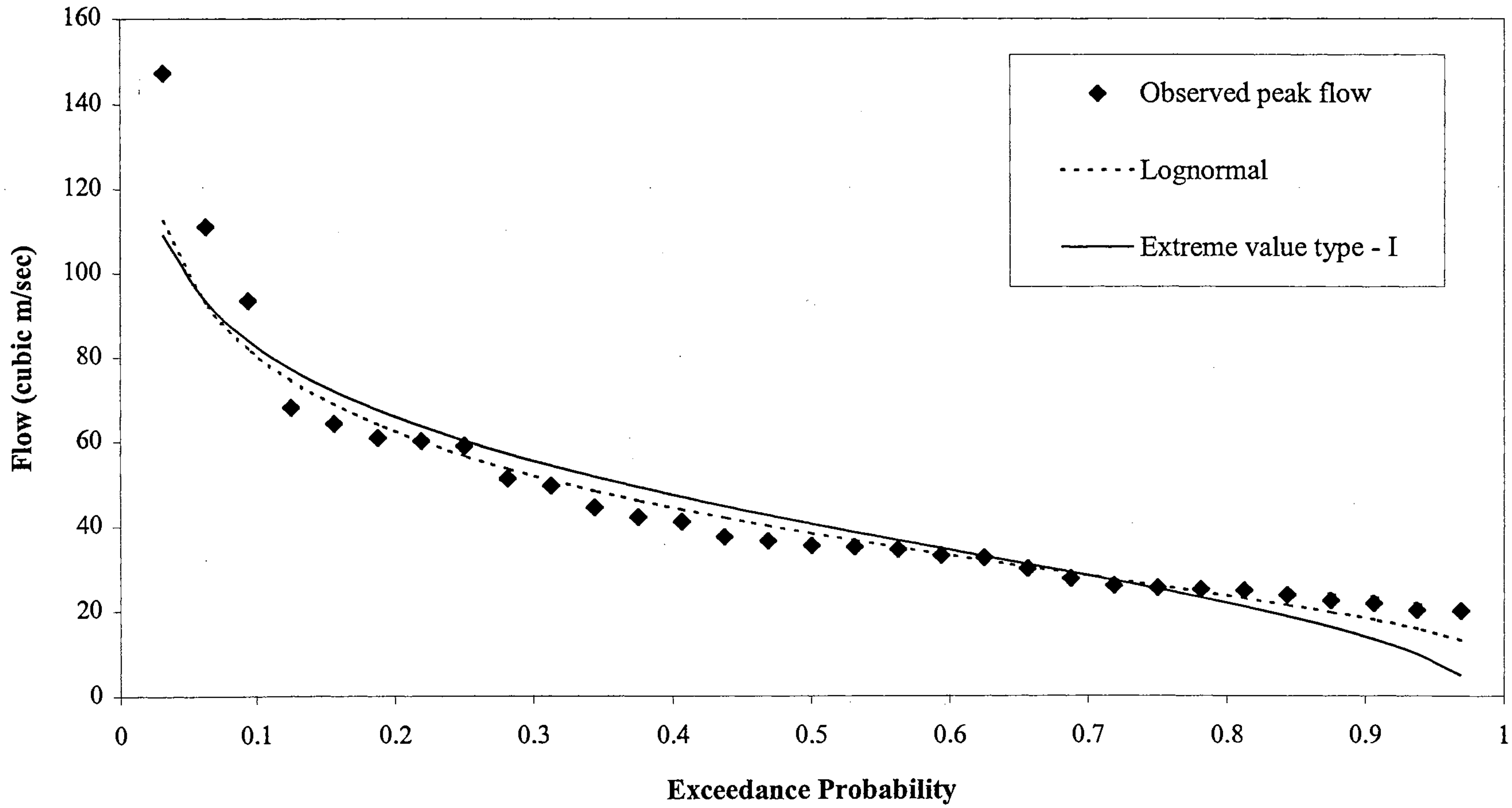

Figure 5-2: Comparison of the lognormal and the extreme value type-I distributions for the Beargrass Creek data

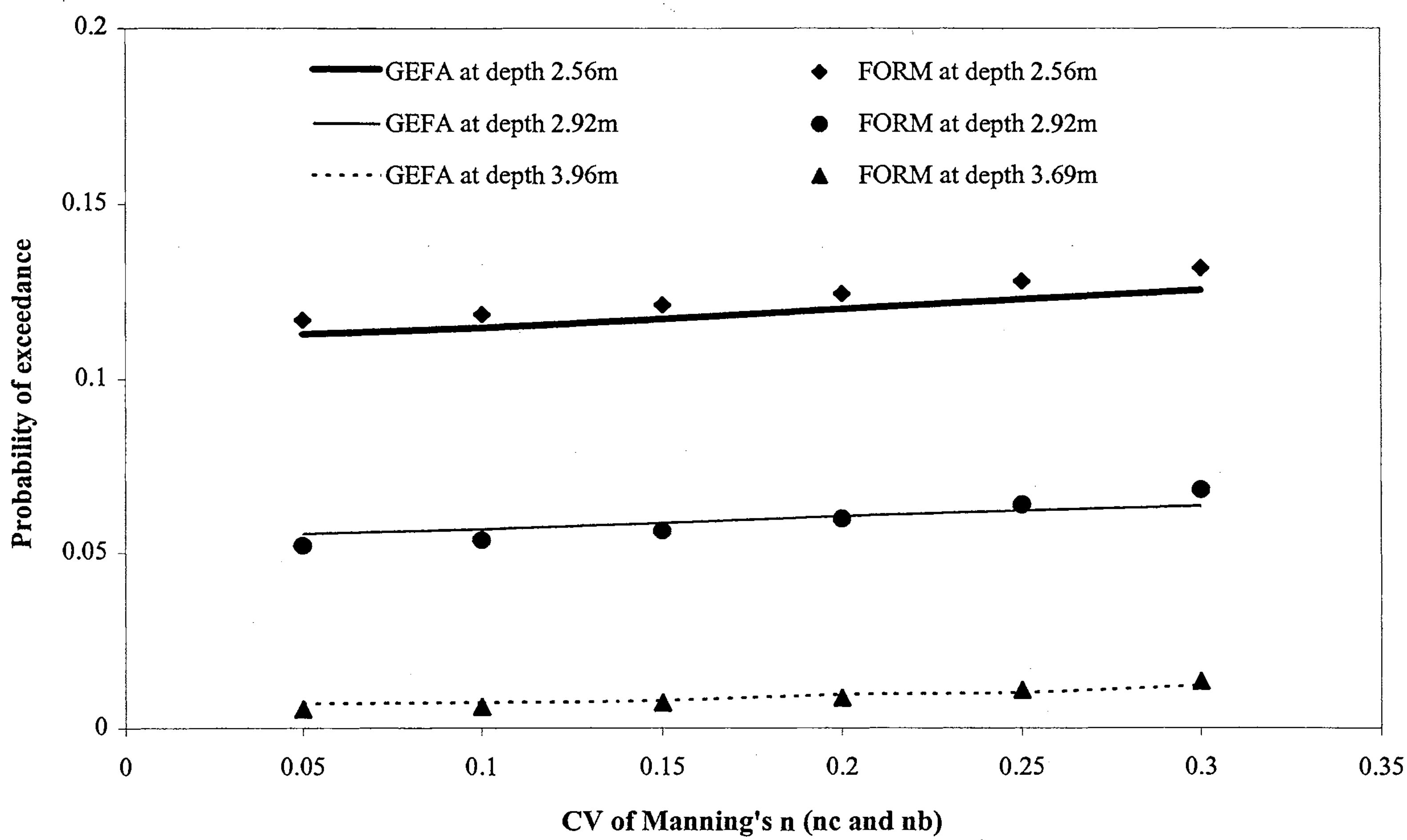

Fig. 5-3: Plot of probability of exceedance of a given depth using GEFA and FORM vs CV of Mannign's coefficients, where Manning' s coefficients and observed annual peak flow are assumed to have the lognarmal and the gamma distributions.

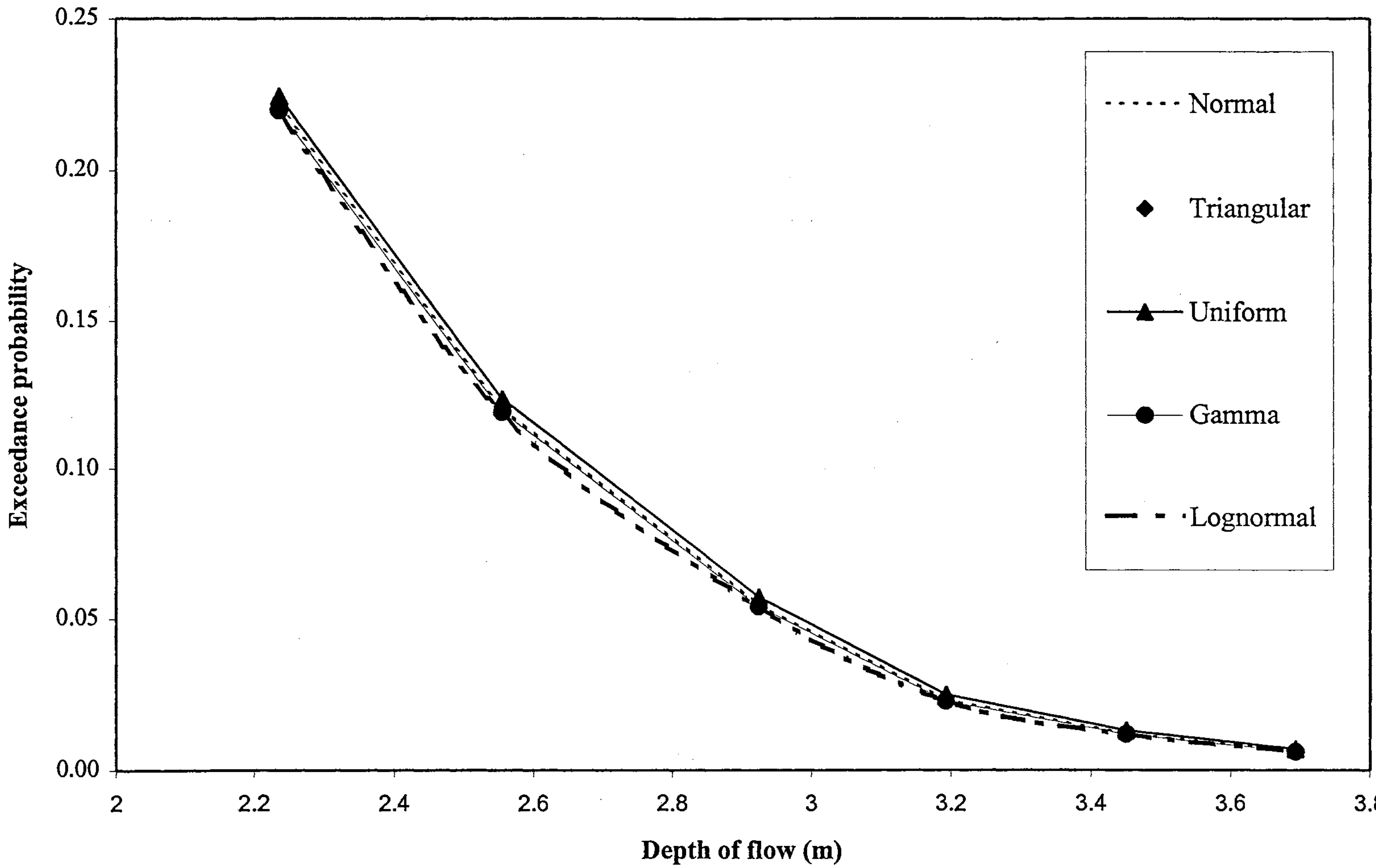

Fig. 5-4: Depth vs exceedance probability plot for a fixed CV of 0.10 and different distributions for Manning's coefficients

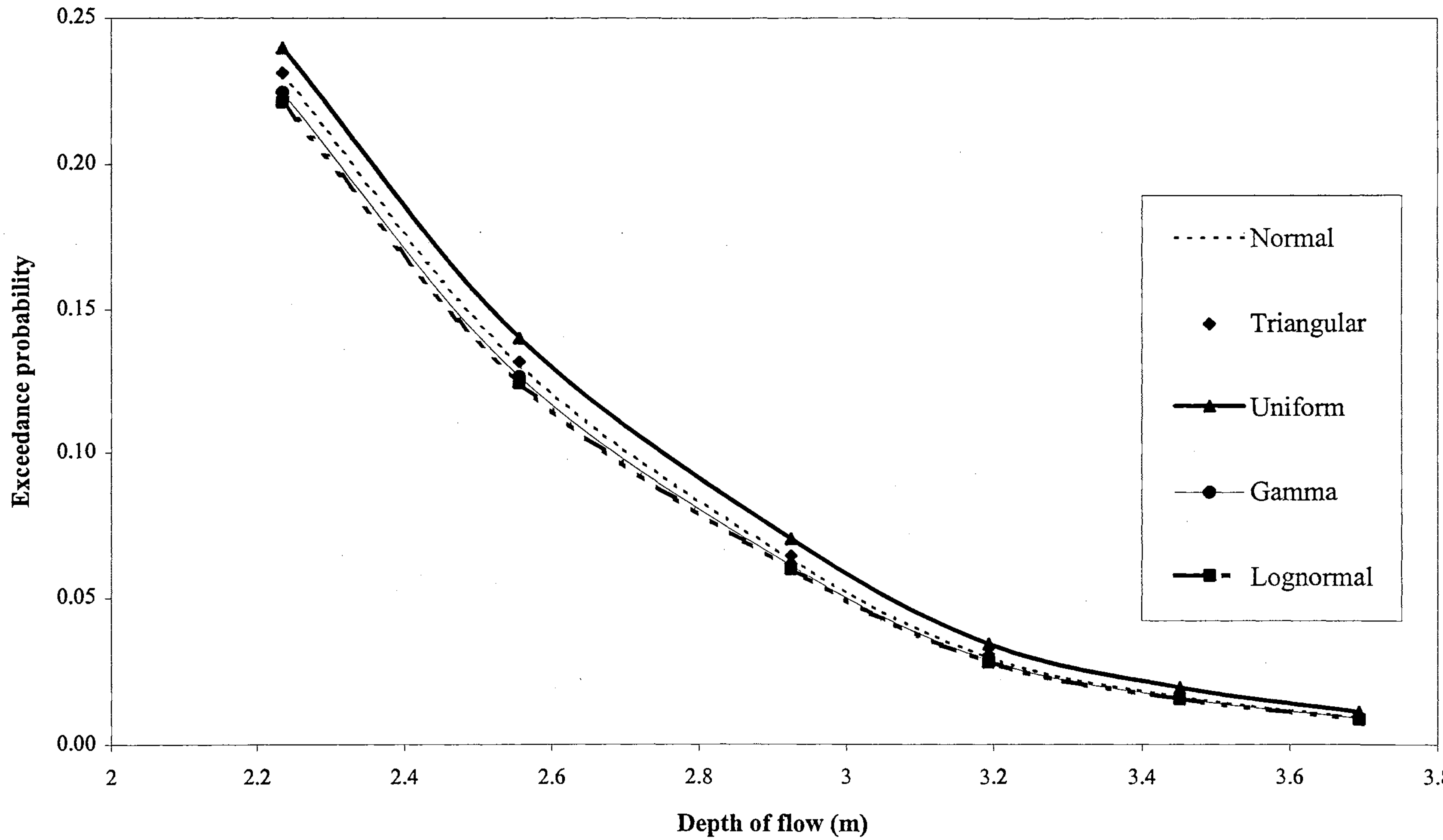

Fig.5-5: Depth vs exceedance probability plot for a fixed CV of 0.20 and different distributions for Manning's coefficients

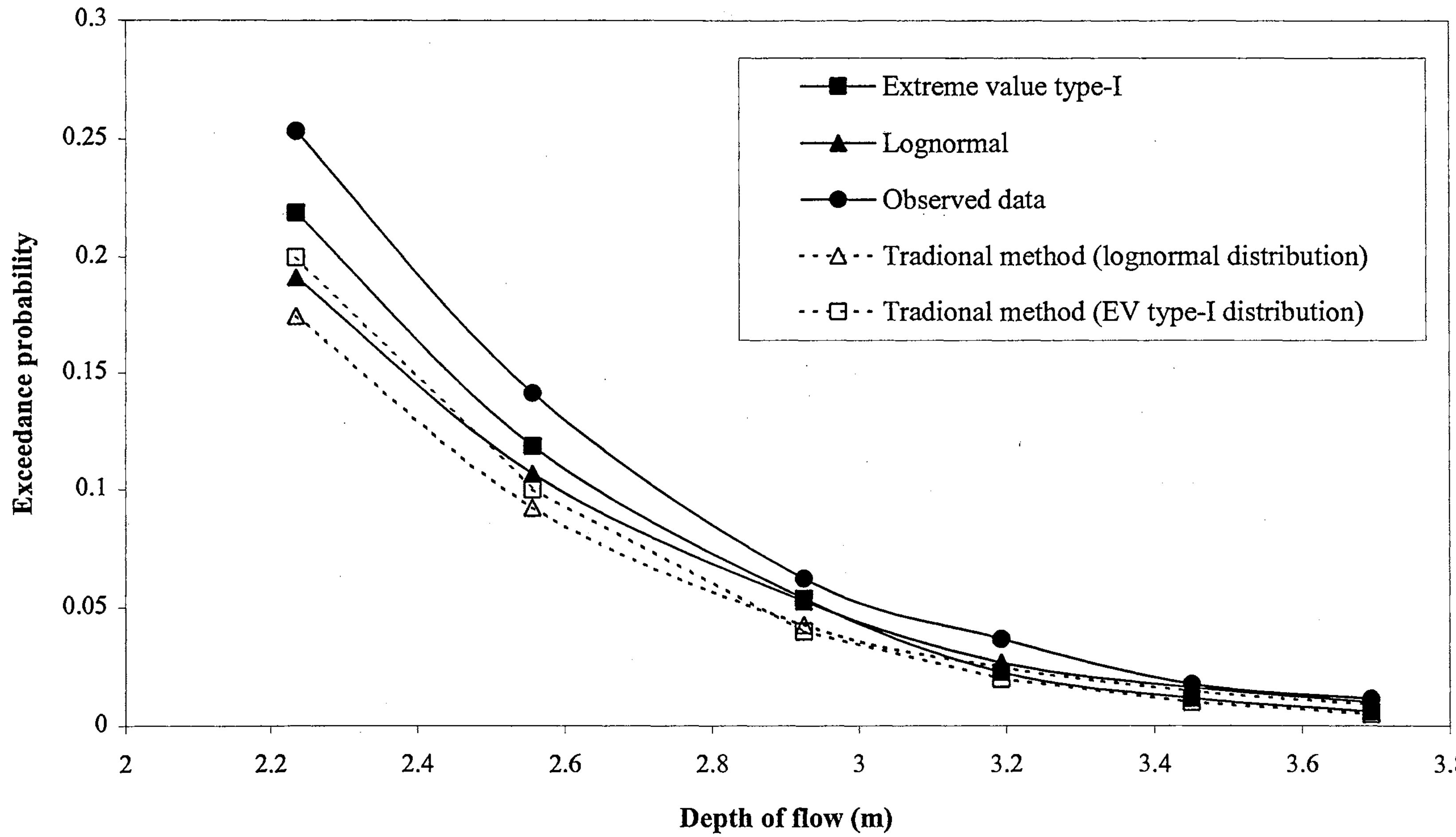

Fig.5- 6: Depth vs exceedance probability plot for a fixed CV of 0.10 for Manning's coefficients and different distributions for peak annual flow

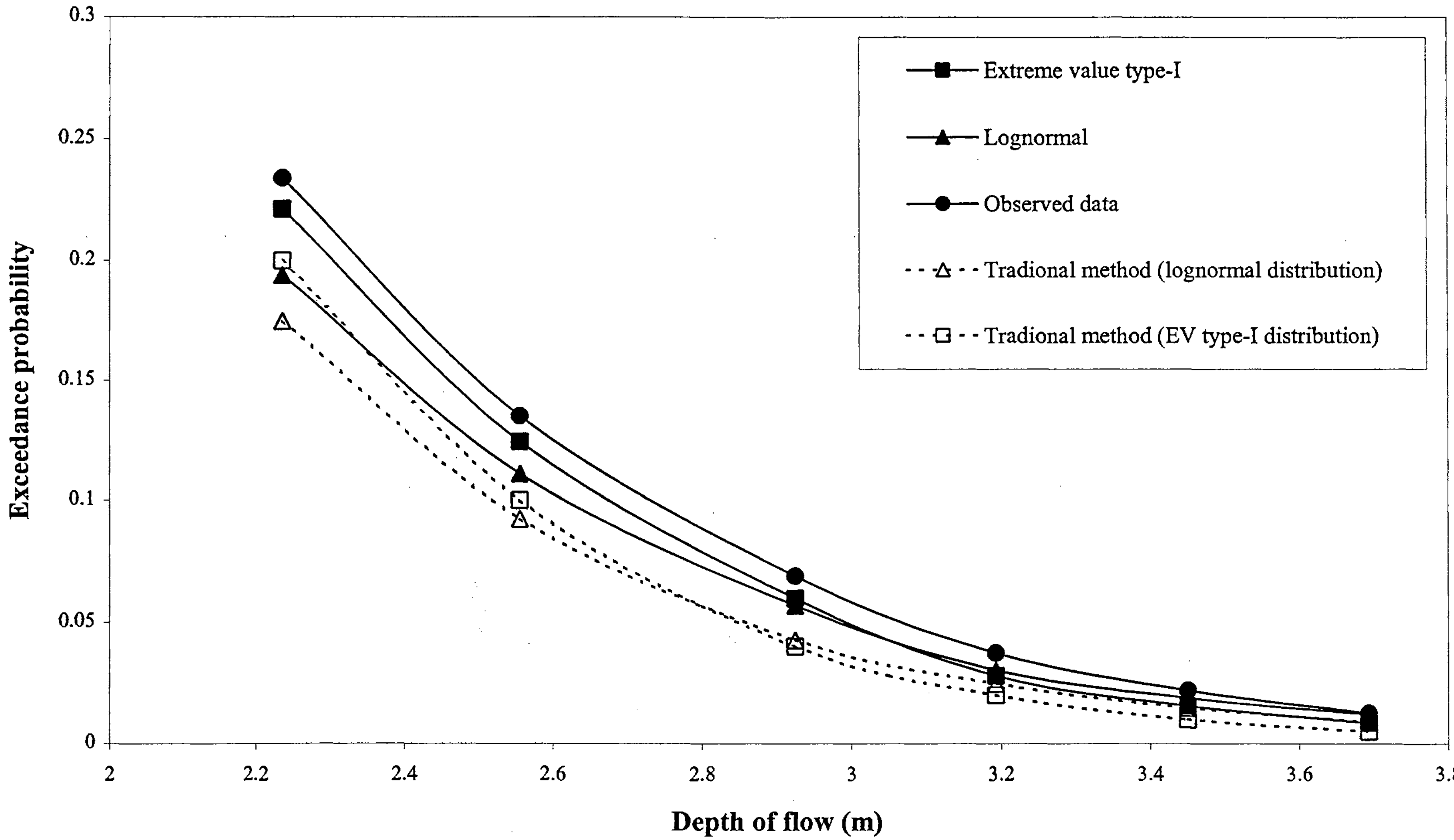

Fig.5-7: Depth vs exceedance probability plot for a fixec CV of 0.20 for Manning's coefficients and different distributions for peak annual flow

CHAPTER VI

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

The great advantage of FOA is its simplicity, requiring knowledge of only the first two statistical moments of the basic variables, and ease of application, requiring simple sensitivity calculations about the selected central values. FOA is an approximate method that may suffice for many applications, but the method does have several conceptual shortcomings. Despite these, FOA has been used quite successfully in a wide variety of fields. It is believed that the exactness of FOA estimates is influenced in part by the degree of nonlinearity in the functional relationship and by parameter uncertainty. To overcome nonlinearity problems, several nonlinearity predictors were proposed by various researchers, which work well only in specific situations. No generalized nonlinearity predictor has been developed so far. To ensure the validity of FOA application for output variance calculations, researchers suggested a number of criteria based on restricting input parameter CVs, but all these criteria have limitations. Presently, no clear-cut guidelines specifying where FOA should be used are available.

The objective of this dissertation was to investigate the important factors affecting the exactness of FOA estimates and develop a simple correction procedure useful for practicing engineers to correct the FOA estimates for the mean and the variance of a model output. The safe design of a hydraulic structure (spillway, channel, culvert, sewer, etc) must be in compliance with regulatory standards specified for the safety and public health (surface and ground water pollution, air pollution, and soil contaminants, etc) require reliability and risk analysis relative to system failure. FOA has been used frequently to carry out reliability and risk analysis of many water resources and environmental engineering systems. This approach has a drawback in that a normal distribution is typically assumed for the model output which is seldom true. Consequently, estimates of the risk/reliability of a model are severely affected, particularly when a probability calculation is sought in the tail portions of the distribution. Better identification of a suitable distribution for the model output is possible if knowledge of higher moments is available. Therefore the final objective of this thesis was to develop a simple approach for calculating the higher-order moments of a model output.

As multitudes of mathematical forms of models are available, it is very difficult to address all possible forms individually. However, most of the mathematical functions can be thought of as an integrated form of individual component functions such as a power function, an exponential function, etc. Their exponents can be used as a surrogate parameter to represent their nonlinearity. Therefore, nonlinearity of component functions can be accounted for.

To study the impacts of parameter CV and functional nonlinearity, a simple power function with positive integral exponents is considered. Table 6-1 presents the relative error expressions for FOA predicted variance, $E(\hat{\sigma}_Y^2)$, for the uniform, the symmetrical triangular, and the normal distributions. These error expressions indicate that exactness of FOA estimates depends not only on parameter uncertainty and functional nonlinearity but also on the distribution of the input parameter.

Table 6-1: Variance error expressions for a simple power function

Function	Relative error in FOA predicted variance, $E(\hat{\sigma}_Y^2)$		
	Uniform	Symmetrical Triangular	Normal
$Y = cX^2$	$\frac{CV^2}{CV^2+5}$	$\frac{7CV^2}{7CV^2+20}$	$\frac{CV^2}{CV^2+2}$
$Y = cX^3$	$\frac{CV^2(3CV^2+14)}{3CV^4+14CV^2+7}$	$\frac{3CV^2(2CV^2+7)}{6CV^4+21CV^2+7}$	$\frac{CV^2(5CV^2+12)}{5CV^4+12CV^2+3}$
$Y = cX^4$	$\frac{9CV^2(CV^4+15CV^2+15)}{9CV^6+135CV^4+135CV^2+25}$	$\frac{3CV^2(48CV^4+390CV^2+265)}{144CV^6+1170CV^4+795CV^2+100}$	$\frac{3CV^2(4CV^4+16CV^2+7)}{12CV^6+4CV^4+16CV^2+7}$

It is observed that the relative error in FOA estimates of means and variances of a model output for a given exponent and CV changes with the type of distribution. This shows that the type of distribution is also important when judging the exactness of FOA. The relative error in FOA estimates for the mean and variance of a power function depends upon the CV of the input parameter, magnitude of exponent *r*, and type of distribution for the input parameter. Knowledge of relative error corresponding to FOA estimates ($\hat{\mu}_Y$ and $\hat{\sigma}_Y^2$) can be used to correct them to obtain their exact values. The exact value of an FOA estimate can be obtained as

$$Exact\ value = \frac{FOA\ estimate}{1 - \mathrm{E}(.)} \tag{6-1}$$

where E(.) is the relative error corresponding to the FOA estimate. The exponent of a component power function in an integrated mathematical form may assume any value including a negative or positive and integer or fractional number. Therefore, generalized mathematical expressions are required for the component functions to apply this technique.

Analytical relationships for the relative error in FOA estimates of the means and the variances of component functions were developed for a generic power function and a generic exponential function using five common distributions. These analytical expressions can be used as a guide for judging the suitability of FOA by determining the relative errors in the most sensitive parameters. Further, when relative error is more than the acceptable error, these analytical relationships enable one to correct FOA estimates for means and variances of model components to their true values. Using these corrected values of means and variances for model components, one can determine the exact values of mean and variance of an overall model output. Tables 6-2 and 6-3 present the developed expressions for $E(\hat{\mu}_Y)$ and $E(\hat{\sigma}_Y^2)$ for a power function $(Y = cX^r)$.

Table 6-2: Generalized relative error in FOA predicted mean of a power function

Distribution	Relative error in FOA predicted variance, $E(\hat{\mu}_Y)$
Uniform	$1 - \frac{2\sqrt{3}(r+1)CV_X}{\left[\left(1+CV_X\sqrt{3}\right)^{(r+1)} - \left(1-CV_X\sqrt{3}\right)^{(r+1)}\right]}$
Symmetrical triangular	$1 - \frac{6(r+1)(r+2)CV_X^2}{\left[\left(1+CV_X\sqrt{6}\right)^{(r+2)} + \left(1-CV_X\sqrt{6}\right)^{(r+2)} - 2\right]}$
Lognormal	$1 - \left(1+CV_X^2\right)^{\frac{1}{2}r(1-r)}$
Gamma	$1 - \frac{CV_X^{-2r}\Gamma\left(CV_X^{-2}\right)}{\Gamma\left[CV_X^{-2}\left(1+rCV_X^2\right)\right]}$
Exponential	$1 - \frac{1}{\Gamma(r+1)}$

Table 6-3: Generalized relative error in FOA predicted variance of a power function

Distribution	Relative error in FOA predicted variance, $E\left(\hat{\sigma}_Y^2\right)$
Uniform	$1-\frac{12\,(2r+1)\,r^2(r+1)^2 CV_X^4}{\left\{2\sqrt{3}CV_X(r+1)^2\left[\left(1+CV_X\sqrt{3}\right)^{2r+1}-\left(1-CV_X\sqrt{3}\right)^{2r+1}\right]-(2r+1)\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right]^2\right\}}$
Symmetrical triangular	$1-\frac{36\,(2r+1)\,r^2(r+1)^2(r+2)^2 CV_X^6}{\left\{3(r+1)(r+2)^2 CV_X^2\left[\left(1+CV_X\sqrt{6}\right)^{2r+2}+\left(1-CV_X\sqrt{6}\right)^{2r+2}-2\right]-(2r+1)\left[\left(1+CV_X\sqrt{6}\right)^{r+2}+\left(1-C\right.\right.\right.}$
Lognormal	$1-\frac{r^2 CV_X^2\left(CV_X^2+1\right)^r}{\left(CV_X^2+1\right)^{r^2}\left[\left(CV_X^2+1\right)^{r^2}-1\right]}$
Gamma	$1-\frac{r^2 CV_X^{2(1-2r)}\left[\Gamma\left(CV_X^{-2}\right)\right]^2}{\Gamma\left[CV_X^{-2}\left(1+2rCV_X^2\right)\right]\Gamma\left(CV_X^{-2}\right)-\left\{\Gamma\left[CV_X^{-2}\left(1+rCV_X^2\right)\right]\right\}^2}$
Exponential	$1-\frac{r^2}{\left[\Gamma(2r+1)-\Gamma^2(r+1)\right]}$

To further simplify the correction procedure, these analytical relationships have been presented graphically. The relative error plots show where FOA estimates are acceptable and where they are unacceptable and need to be corrected. In specific situations, a given function may be very nonlinear (represented either by a very large or very small exponent of a power function). These situations can be identified and dealt with by using the relative error plots. There are several other features of error plots, which are discussed in the following section.

The relative error is zero for a power function at certain values of the exponents, which changes with the type of distribution used for the input random variable. These exponents are 0 and 1 as shown by $\mathrm{E}(\hat{\mu}_Y)$ plots (Figures 2-1, 2-3, 2-5, 2-7, 2-10) for all the considered distributions. In the same way, there are two exponent values for $\mathrm{E}\left(\hat{\sigma}_Y^2\right)$ where FOA estimates for the variance have no error. One of these exponents is 1 and the other changes with the distribution type and CV of the input parameter as shown in Table 2-2. The average values of these second exponents are 1.75, 1.70, 1.65, 0.28, -0.34 and 0.30 for the uniform, the triangular, the normal, the exponential, the lognormal, and the gamma distributions respectively.

Table 2-2 shows that when the exponent of a power function lies within the tabulated range for each distribution, the FOA variance estimate will have almost no error and the power function will behave like a linear function as far as the variance prediction is concerned. These situations are depicted by $E(\hat{\sigma}_Y^2)$vs. r plots in Figures 2-2, 2-4, 2-6, 2-8, and 2-11. In general, it can be concluded that application of FOA will provide good results when the exponent of a power function lies in the vicinity of these exponent values, and the error will be small regardless of the CV values of the input variable. This contradicts previous findings that FOA works well only when CV $\leq$ 0.2.

When the exponent of a power function falls between 1 and 1.7 for normal, uniform, and triangular distributed parameters, the FOA overestimates the actual variance. However, the overestimation is small as shown by the negative values of $E(\hat{\sigma}_Y^2)$ in Figures 2-2, 2-4, and 2-11. When the exponent falls outside this range, the FOA underestimates the actual variance. When the parameter is lognormally distributed and the exponent falls between -0.34 and 1, the FOA may highly overestimate the actual variance depending upon the parameter CV values as shown by negative values of $E(\hat{\sigma}_Y^2)$ in Figure 1.6. When the exponent falls outside this range, FOA underestimates the actual variance. In the case of the gamma distribution, when the exponent falls between 0.3 and 1, the FOA overestimates the actual variance. For exponents outside this range the FOA underestimates the actual variance. Relative error plots further endorse the conclusion that the parameter distribution type affects the accuracy of FOA predicted means and variances.

Even a very small exponent (very close to zero) may give a very high relative error in FOA predicted variance for some of the distributions (normal, uniform, and triangular). Error plots of the normal distribution (Figures 2-10 and 2-11) show that significant errors occur in both the mean and variance of a power function predicted using FOA.

Similarly Tables 6-4 and 6-5 present the developed expressions for $E(\hat{\mu}_Y)$ and $E(\hat{\sigma}_Y^2)$ for an exponential function $(Y = be^{cX})$.

Table 6-4: Generalized relative error in FOA predicted mean of a power function

Distribution	Relative error in FOA predicted variance, $E(\hat{\mu}_Y)$
Uniform	$1-\frac{2\sqrt{3}c\mu_x CV_x e^{\sqrt{3}c\mu_x CV_x}}{\left(e^{2\sqrt{3}c\mu_x CV_x}-1\right)}$
Symmetrical triangular	$1-\frac{6c^2\mu_x^2 CV_x^2 e^{c\mu_x CV_x\sqrt{6}}}{\left(e^{c\mu_x CV_x\sqrt{6}}-1\right)^2}$
Normal	$1-\exp\left[-\frac{1}{2}c^2\mu_x^2 CV_x^2\right]$
Gamma	$1-\left(1-c\mu_x CV_x^2\right)^{\frac{1}{CV_x^2}}\exp(c\mu_x)$
Exponential	$1-(1-c\mu_x)\exp(c\mu_x)$

Table 6-5: Generalized relative error in FOA predicted variance of a power function

Distribution	Relative error in FOA predicted variance, $E(\hat{\sigma}_Y^2)$
Uniform	$1-\frac{12c^4\mu_x^4CV_x^4e^{2\sqrt{3}c\mu_xCV_x}}{\left(e^{2\sqrt{3}c\mu_xCV_x}-1\right)\left[\left(\sqrt{3}c\mu_xCV_x-1\right)e^{2\sqrt{3}c\mu_xCV_x}+\sqrt{3}c\mu_xCV_x+1\right]}$
Symmetrical triangular	$1-\frac{72c^6\mu_x^6CV_x^6e^{2\sqrt{6}c\mu_xCV_x}}{\left(e^{\sqrt{6}c\mu_xCV_x}-1\right)^2\left[\left(3c^2\mu_x^2CV_x^2-2\right)\left(e^{2\sqrt{6}c\mu_xCV_x}+1\right)+2e^{\sqrt{6}c\mu_xCV_x}\left(3c^2\mu_x^2CV_x^2+2\right)\right]}$
Normal	$1-\frac{c^2\sigma_x^2}{\exp\left(c^2\sigma_x^2\right)\left[\exp\left(c^2\sigma_x^2\right)-1\right]}$
Gamma	$1-\frac{c^2\mu_x^2CV_x^2\exp(2c\mu_x)}{\left(1-2c\mu_xCV_x^2\right)^{-\frac{1}{CV_x^2}}-\left(1-c\mu_xCV_x^2\right)^{-\frac{2}{CV_x^2}}}$
Exponential	$1-\frac{c^2\mu_x^2\exp(2c\mu_x)}{(1-2c\mu_x)^{-1}-(1-c\mu_x)^{-1}}$

These mathematical relationships have been presented graphically in Figures 3.1 to 3.9. These relative error plots indicate that relative error in both the means and the variances of an exponential function is small when exponent mean value is small. However, an exception of this generalization has been observed in relative error plots of variance when input parameter has the gamma or the exponential distribution. In these cases, the relative error is small at two points of exponent mean values. For example, in case of the gamma distribution, one of such point is zero and the other varies from -1.33 to -1.71 with the CV of input parameter.

In order to determine the higher-order moments of a model output correctly, a simple approach of using generic expectation functions (GEFA) as a function of means and CVs of input random variables is proposed. GEFA are easy to develop and simple to apply to problems related to reliability, risk, and uncertainty analysis. Several expectation functions based on commonly used probability distributions have been developed (Tables 6-6 and 6-7) for a power and an exponential function.

Table 6-6: Generic expectation functions for some commonly used probability density functions

Name	Generic expectation function, $E[X^r]$
Uniform	$\dfrac{\mu_X^r}{2\sqrt{3}(r+1)CV_X}\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right]$
Symmetrical triangular	$\dfrac{\mu_X^r}{6(r+1)(r+2)CV_X^2}\left[\left(1+CV_X\sqrt{6}\right)^{r+2}+\left(1-CV_X\sqrt{6}\right)^{r+2}-2\right]$
Unsymmetrical triangular	$\dfrac{2\left[(\beta-\omega)\alpha^{r+2}+(\omega-\alpha)\beta^{r+2}+(\alpha-\beta)\omega^{r+2}\right]}{(r+1)(r+2)(\beta-\omega)(\omega-\alpha)(\beta-\alpha)}$
Lognormal	$\mu_X^r\left(1+CV_X^2\right)^{\frac{r(r-1)}{2}}$
Gamma	$\dfrac{CV_X^{2r}\mu_X^r\Gamma\left(CV_X^{-2}+r\right)}{\Gamma(CV_X^{-2})}$
Exponential	$\mu_X^r\Gamma(r+1)$
Normal	$\mu_X^r\sum_{n=0}^{r/2}\binom{r}{2n}\dfrac{(2n)!}{2^n\,n!}CV_X^{2n}$, when r is an even positive integer; $\mu_X^r\sum_{n=0}^{(r-1)/2}\binom{r}{2n}\dfrac{(2n)!}{2^n\,n!}CV_X^{2n}$, when r is an odd positive integer; and when r is anything but a positive integer, $\mu_X^r\left[1+\dfrac{r(r-1)}{2}CV_X^2+......+\dfrac{r(r-1)(r-2).....(r-n+1)}{2^{n/2}(n/2)!}CV_X^n+...\right]$.

Table 6-7: Generic expectation functions for some commonly used probability density functions

Name	Generic expectation function, $E[X^r]$
Uniform	$\frac{b^r}{2\sqrt{3}rc\mu_x CV_x}\left[e^{rc\mu_x(1+CV_x\sqrt{3})} - e^{rc\mu_x(1-CV_x\sqrt{3})}\right]$
Symmetrical triangular	$\frac{b}{6r^2c^2\mu_x^2 CV_x^2}\left[e^{\frac{1}{2}rc\mu_x(1+CV_x\sqrt{6})} - e^{\frac{1}{2}rc\mu_x(1-CV_x\sqrt{6})}\right]$
Unsymmetrical triangular	$\frac{2b^r[(\alpha-\beta)\exp(rc\omega)+(\beta-\omega)\exp(rc\alpha)+(\omega-\alpha)\exp(rc\beta)]}{r^2c^2(\beta-\alpha)(\omega-\alpha)(\beta-\omega)}$
Normal	$b^r \exp\left(rc\mu_x + \frac{1}{2}r^2c^2\mu_x^2 CV_x^2\right)$
Gamma	$b^r\left(1-cr\mu_x CV_x^2\right)^{-\frac{1}{CV_x^2}}$
Exponential	$\frac{b^r}{(1-cr\mu_x)}$

The parameters β, α, and ω of a unsymmetrical triangular distribution are the maximum, minimum, and mode values of X, which can be obtained by substituting n = 0, 1, and 2 respectively in the following relationship.

$$\alpha, \beta, \omega = \mu_X\left\{1 + 2\sqrt{2}CV_X \cos\left[\frac{2\pi n}{3} + \frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_x\right)\right]\right\} \quad (6\text{-}2)$$

The developed expectation functions are general. Using them, any order of moment can be estimated exactly. The knowledge of higher-order moments can be used to identify the most suitable distributions from the commonly used distributions as well as to determine an appropriate distribution by satisfying higher-order moments. Using such a distribution with its correct parameters, one can find a good estimate of reliability, risk and uncertainty of a system.

A comprehensive study on reliability analysis of open channel flow has been carried out using GEFA. It has been concluded that an increase in uncertainty of the Manning's roughness coefficient increases the estimated risk that a given depth of flow will be exceeded. The variation in types of distribution for the Manning's roughness coefficient has negligible effect at smaller CVs, however, the effect is more pronounced at higher CVs. The most significant impact was observed due to the type of distribution used to represent the peak annual flow in the performance function at both smaller and higher CVs of the Manning's coefficients. The problems due to distribution fitting can be removed by incorporating the actual observed data in the performance function. No other method provides any flexibility to incorporate actual observed data into the performance function. GEFA can be used regardless of whether the peak annual flow is represented by a distribution or by the actual observed data.

Several other practical examples related to uncertainty, risk, and reliability analysis of hydrologic, hydraulic, and environmental systems are presented which show that the developed techniques are easy to apply, simple to understand, efficient to compute, and give accurate results. The limitation of the proposed technique is that it can only be applied to functions that involve un-correlated input random variables.

Conclusions

It is investigated and mathematically proven that exactness of FOA estimates depends up on three factors parameter CV, parameter distribution type, and degree of nonlinearity in the functional relationship, as opposed to earlier belief that only two factors the degree of nonlinearity in the functional relationship and parameter uncertainty were relevant. However, in some specific situations all distribution types may give comparable results. Further it has been shown that FOA estimates are erroneous even when input parameters are normally distributed. This clears the misconception that FOA works well when input variables are normally distributed. Two approaches for reliability, risk, and uncertainty analysis are presented. The first is a correction procedure for correcting the FOA estimates for parameter uncertainty, parameter distribution type, and model non-linearity. The second is the use of generic expectation functions for evaluating exact moments of model output.

The FOA estimate correction technique is particularly useful for determining exact values of the first two moments of model output generally required for uncertainty analysis. Analytical relationships are developed to determine the relative errors in FOA estimates for the means and variances of power and exponential functions using several commonly used distributions. Using these relative error functions, one can correct the FOA estimates for the means and variances of component power functions for nonlinearity, and distribution type to evaluate the exact mean and variance of model output. For ease in application, analytical relative error expressions are presented graphically. These plots can be used to determine an approximate relative error for a given exponent of a power function and CV of its random variable. Another advantage of relative error plots is they present an overall idea about the suitability of FOA in a given situation without making any calculations. Special cases are identified when applying FOA to a nonlinear power function for estimating its variance will give a negligible or no error. This method provides a procedure for incorporating known information on the types of input variable distributions.

In reliability and risk analysis, information on the model output distribution is very important. Knowledge of higher-order moments helps to identify the appropriate distribution for the model output. A simple approach of developing generic expectation functions is described. Analytical expressions of generic expectation functions for generalized power and exponential functions were derived using several commonly used input parameter distributions. These expectation functions can be used to determine exact estimates of any order of model output moments.

Many hydrologic applications involve representing a random variable by a certain distribution, but it is seldom possible to fit given data exactly. Incorporating the actual observed data in the performance function could solve problems due to distribution fitting. No other method provides such flexibility to incorporate actual observed data into the performance function. GEFA can be used irrespective whether the peak annual flow is represented by a distribution or by the actual observed data.

Recommendations

As summarized and concluded in the preceding sections and presented in the previous chapters, the objectives of this study have been fully accomplished. Proposed and developed uncertainty analysis using corrected first order approximation method, and reliability and risk analysis using GEFA can be used as very efficient, easy, accurate tools to carryout reliability, risk, and uncertainty analysis of an engineering system. For further enhancement, future research should be focussed on the following areas.

1. While developing the two approaches, the correction procedure for correcting the FOA estimates for parameter uncertainty, parameter distribution type, and model non-linearity and the generic expectation function approach for evaluating the exact

moments of a model output, input parameters are assumed to be independent. This is a practical limitation of the developed method as there are many situations where input parameters are highly correlated. Thus, there is need to develop parallel expressions considering correlation among input variables.

2. For a power function of normally distributed input random variable with a negative exponent, the relative errors in FOA predicted estimates were obtained using a trial and error procedure. This procedure works well when CV and magnitude of exponent of input random variable are small but gives a significant error at higher CV and exponent magnitude. This need further research work of developing an analytical expression for determining the exact relative error in FOA estimated estimates.
3. For an exponential function of a lognormally distributed input random variable, the relative error expressions could not be derived either analytically or numerically. This also needs to be further investigated in order to develop relative error functions.
4. While applying GEFA to reliability analysis of open channel flow, it has been observed that in most cases a commonly used distribution can be employed to evaluate risk or reliability corresponding to a given depth of flow. But specific situations may arise where no available distribution can be used with confidence. The exactness of reliability estimates improves by incorporating the higher-order moments into determining the distribution of performance function. Thus, future research should be focussed to develop a simple procedure of deriving an output distribution involving higher-order moments exactly. It is recommended that output distribution should incorporate at least first four moments exactly.

5. Many engineering problems can be solved using relative error expressions and generic expectation functions for power and exponential functions. However, some specific fields use other component functional forms quite frequently. For example a complementary error function, $erfc(.)$, has appeared very frequently in analytical groundwater contaminant transport models. Thus, it is recommended that similar relative error functions should be developed for other frequently used functional forms.

APPENDIX I

DEVELOPMENT OF RELATIVE ERROR FUNCTIONS FOR A POWER FUNCTION

Uniform Distribution

The probability density function $p_X(x)$ for the continuous uniform distribution is

$$p_X(x)=\frac{1}{(\beta-\alpha)},\ \alpha\leq X\leq\beta \tag{I-1}$$

where α and β are the distribution parameters. Using the methods of moments, the estimates for α and β are given (Haan, 1977) as

$$\hat{\alpha}=\mu_X-\sqrt{3}\sigma_X=\mu_X\left(1-\sqrt{3}CV_X\right) \tag{I-2a}$$

$$\hat{\beta}=\mu_X+\sqrt{3}\sigma_X=\mu_X\left(1+\sqrt{3}CV_X\right) \tag{I-2b}$$

where CV_X is the coefficient of variation of X, defined as

$$CV_X=\frac{\sigma_X}{\mu_X} \tag{I-3}$$

Using equations (2-26), (I-1), (I-2a) and (I-2b) $E[X^r]$ is given as

$$E\left[X^r\right]=\int_{\alpha}^{\beta}\frac{1}{(\beta-\alpha)}X^r dX=\frac{\mu_X^r}{2\sqrt{3}(r+1)CV_X}\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right] \tag{I-4}$$

Similarly $E[X^{2r}]$ can also be written as

$$E\left[X^{2r}\right]=\int_{\alpha}^{\beta}\frac{1}{(\beta-\alpha)}X^{2r}dx=\frac{\mu_X^{2r}}{2\sqrt{3}(2r+1)CV_X}\left[\left(1+CV_X\sqrt{3}\right)^{2r+1}-\left(1-CV_X\sqrt{3}\right)^{2r+1}\right] \tag{I-5}$$

Substituting $E[X^r]$ in (2-26), μ_Y is given as

$$\mu_Y=c\frac{\mu_X^r}{2\sqrt{3}(r+1)CV_X}\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right] \tag{I-6}$$

Similarly substituting $E[X^r]$ from (I-4) and $E[X^{2r}]$ from (I-5) in (2-27), the expression for σ_Y^2 becomes

$$\sigma_Y^2=\frac{c^2\left\{2\sqrt{3}CV_X(r+1)^2\left[\left(1+CV_X\sqrt{3}\right)^{2r+1}-\left(1-CV_X\sqrt{3}\right)^{2r+1}\right]-(2r+1)\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right]^2\right\}\mu_X^{2r}}{12(r+1)^2(2r+1)CV_X^2} \tag{I-7}$$

Substituting (2-22) and (I-6) in (2-24), the expression for relative error in FOA predicted mean, $E(\hat{\mu}_Y)$ is given as

$$E(\hat{\mu}_Y)=1-\frac{2\sqrt{3}(r+1)CV_X}{\left[\left(1+CV_X\sqrt{3}\right)^{(r+1)}-\left(1-CV_X\sqrt{3}\right)^{(r+1)}\right]} \tag{I-8}$$

Now substituting (2-23) and (I-7) in (2-24), the relative error in FOA predicted variance $E(\hat{\sigma}_y^2)$ can be represented as

$$E(\hat{\sigma}_y^2)=1-\frac{12(2r+1)r^2(r+1)^2CV_X^4}{\left\{2\sqrt{3}CV_X(r+1)^2\left[\left(1+CV_X\sqrt{3}\right)^{2r+1}-\left(1-CV_X\sqrt{3}\right)^{2r+1}\right]-(2r+1)\left[\left(1+CV_X\sqrt{3}\right)^{r+1}-\left(1-CV_X\sqrt{3}\right)^{r+1}\right]^2\right\}} \tag{I-9}$$

Symmetrical Triangular Distribution

The probability density function $p_X(x)$ for the symmetrical triangular distribution is

$$p_X(x) = \frac{2}{(\beta-\alpha)}\frac{(X-\alpha)}{(\gamma-\alpha)} \quad \text{, when } \alpha \le X \le \gamma$$

$$p_X(x) = \frac{2}{(\beta-\alpha)}\frac{(\beta-X)}{(\beta-\gamma)} \quad \text{, when } \gamma \le X \le \beta \tag{I-10}$$

where α, β, γ are the minimum, maximum, and mode values of X. When it is symmetrical, $\gamma = \mu_X$. The methods of moments estimates for α and β are

$$\hat{\alpha} = \mu_X - \sqrt{6}\sigma_X = \mu_X\left(1-\sqrt{6}CV_X\right) \tag{I-11a}$$

$$\hat{\beta} = \mu_X + \sqrt{6}\sigma_X = \mu_X\left(1+\sqrt{6}CV_X\right) \tag{I-11b}$$

Using (2-26), (I-10), (I-11a) and (I-11b) $E[X^r]$ is given as

$$E\left[X^r\right] = \frac{\mu_X^r}{6(r+1)(r+2)CV_X^2}\left[\left(1+CV_X\sqrt{6}\right)^{r+2} + \left(1-CV_X\sqrt{6}\right)^{r+2} - 2\right] \tag{I-12}$$

Similarly $E[X^{2r}]$ can also be written as

$$E\left[X^{2r}\right] = \frac{\mu_X^{2r}}{12(r+1)(2r+1)CV_X^2}\left[\left(1+CV_X\sqrt{6}\right)^{2r+2} + \left(1-CV_X\sqrt{6}\right)^{2r+2} - 2\right] \tag{I-13}$$

Substituting $E[X^r]$ from (I-12) into (2-26), μ_Y is given as

$$\mu_Y = \frac{c\mu_X^r}{6(r+1)(r+2)CV_X^2}\left[\left(1+CV_X\sqrt{6}\right)^{r+2} + \left(1-CV_X\sqrt{6}\right)^{r+2} - 2\right] \tag{I-14}$$

Similarly substituting $E[X^r]$ from (I-12) and $E[X^{2r}]$ from (I-13) in (2-27), the expression for σ^2_Y is written as

$$\sigma_Y^2 = \frac{c^2\left\{3(r+1)(r+2)^2CV_X^2\left[\left(1+CV_X\sqrt{6}\right)^{2r+2} + \left(1-CV_X\sqrt{6}\right)^{2r+2} - 2\right] - (2r+1)\left[\left(1+CV_{X_x}\sqrt{6}\right)^{r+2} + \left(1-CV_X\sqrt{6}\right)^{r+2} - 2\right]^2\right\}\mu_X^{2r}}{36(2r+1)(r+1)^2(r+2)^2CV_X^4}$$

(I-15)

Substituting (2-22) and (I-14) in (2-24), the expression for relative error in FOA predicted mean, E($\hat{\mu}_Y$) is given as

$$E(\hat{\mu}_Y) = 1 - \frac{6(r+1)(r+2)CV_X^2}{\left[\left(1+CV_X\sqrt{6}\right)^{(r+2)} + \left(1-CV_X\sqrt{6}\right)^{(r+2)} - 2\right]} \tag{I-16}$$

Now substituting (2-23) and (I-15) in (2-24), the relative error in FOA predicted variance $E(\hat{\sigma}_y^2)$ can be expressed as

$$E(\hat{\sigma}_y^2) = 1 - \frac{36(2r+1)r^2(r+1)^2(r+2)^2 CV_X^6}{\left\{3(r+1)(r+2)^2 CV_X^2\left[\left(1+CV_X\sqrt{6}\right)^{2r+2} + \left(1-CV_X\sqrt{6}\right)^{2r+2} - 2\right] - (2r+1)\left[\left(1+CV_X\sqrt{6}\right)^{r+2} + \left(1-CV_X\sqrt{6}\right)^{r+2} - 2\right]^2\right\}} \tag{I-17}$$

Lognormal Distribution

If X is lognormally distributed with mean, μ_X, and variance, σ^2_X, its probability density function is given (Haan, 1977) as

$$p_X(x) = \frac{1}{\sigma_V X\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln X - \mu_V}{\sigma_V}\right)^2}, X > 0 \tag{I-18}$$

where $V = ln(X)$ is normally distributed with parameters μ_V and σ^2_V. The parameters μ_V and σ^2_V are defined (Haan, 1977) as

$$\mu_V = \frac{1}{2}\ln\left[\frac{\mu_X^2}{CV_X^2 + 1}\right] \tag{I-19}$$

$$\sigma_V^2 = \ln\left(CV_X^2 + 1\right) \tag{I-20}$$

Substituting (I-18) in (2-26), $E[X^r]$ is written as

$$E\left(X^r\right) = \frac{1}{\sigma_V\sqrt{2\pi}} \int_0^\infty X^{r-1} e^{-\frac{1}{2}\left(\frac{\ln X - \mu_V}{\sigma_V}\right)^2} dx \tag{I-21}$$

Assuming, $\frac{\ln(X)-\mu_V}{\sigma_V}=Z$, the random variable X can be written as, $X=e^{(\mu_V+\sigma_V Z)}$, (I-21)

is rewritten as

$$E(X^r)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{r(\mu_V+\sigma_V Z)}\exp^{-\frac{1}{2}z^2}\,dz=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{1}{2}(2\mu_V r+\sigma_V^2 r^2)}e^{-\frac{1}{2}(Z-r\sigma_V)^2}\,dZ \tag{I-22}$$

But

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(Z-r\sigma_V)^2}\,dZ=1$$

Equation (I-22) can be written as

$$E(X^r)=e^{\left(\mu_V r+\frac{1}{2}\sigma_V^2 r^2\right)} \tag{I-23}$$

Substituting $2r$ in place of r, $E[X^{2r}]$ can be written as

$$E(X^{2r})=e^{2(\mu_V r+\sigma_V^2 r^2)} \tag{I-24}$$

Substituting $r = 1$ in (I-23), μ_X can be written as

$$\mu_X=e^{\left(\mu_V+\frac{1}{2}\sigma_V^2\right)} \tag{I-25}$$

Substituting (I-25) in (22), the FOA estimate for mean μ_Y can be written as

$$\hat{\mu}_Y=c\exp\left(r\mu_V+\frac{1}{2}r\sigma_V^2\right) \tag{I-26}$$

The exact value of μ_Y can be obtained by substituting value of $E[X^r]$ from (I-23) in (2-26) as:

$$\mu_Y=c\exp\left(r\mu_V+\frac{1}{2}r^2\sigma_V^2\right) \tag{I-27}$$

Substituting (I-26) and (I-27) in (2-24), the relative error in FOA predicted mean $E(\hat{\mu}_Y)$ can be expressed as

$$E(\hat{\mu}_Y) = 1 - \frac{c\exp\left(r\mu_V + \frac{1}{2}r\sigma_V^2\right)}{c\exp\left(r\mu_V + \frac{1}{2}r^2\sigma_V^2\right)} = 1 - \exp\left[\frac{1}{2}r(1-r)\sigma_V^2\right] \quad \text{(I-28)}$$

Substituting (I-20) in (I-28), $E(\hat{\mu}_Y)$ can be rewritten as

$$E(\hat{\mu}_Y) = 1 - \left(1 + CV_X^2\right)^{\frac{1}{2}r(1-r)} \quad \text{(I-29)}$$

Substituting (I-25) in (23), the FOA estimate for variance σ^2_Y can be written as

$$\hat{\sigma}_Y^2 = c^2 r^2 \exp\left(2r\mu_V + r\sigma_V^2\right)CV_X^2 \quad \text{(I-30)}$$

Substituting $E[X^r]$ from (I-23) and $E[X^{2r}]$ from (I-24) in (2-27), the exact variance can be written as

$$\sigma_Y^2 = c^2 \exp\left(2r\mu_V + r^2\sigma_V^2\right)\left[\exp\left(r^2\sigma_V^2\right) - 1\right] \quad \text{(I-31)}$$

Substituting (I-30) and (I-31) in (2-24), the relative error in FOA predicted variance $E(\hat{\sigma}_y^2)$ can be expressed as

$$E(\hat{\sigma}_y^2) = 1 - \frac{c^2 r^2 \exp\left(2r\mu_V + r\sigma_V^2\right)CV_X^2}{c^2 \exp\left(2r\mu_V + r^2\sigma_V^2\right)\left[\exp\left(r^2\sigma_V^2\right) - 1\right]} = 1 - \frac{r^2 \exp\left(r\sigma_V^2\right)CV_X^2}{\exp\left(r^2\sigma_V^2\right)\left[\exp\left(r^2\sigma_V^2\right) - 1\right]} \quad \text{(I-32)}$$

Substituting (I-20) in (I-32), $E(\hat{\sigma}_y^2)$ can be rewritten as

$$E(\hat{\sigma}_y^2) = 1 - \frac{r^2 CV_X^2 \left(CV_X^2 + 1\right)^r}{\left(CV_X^2 + 1\right)^{r^2}\left[\left(CV_X^2 + 1\right)^{r^2} - 1\right]} \quad \text{(I-33)}$$

<u>Gamma Distribution</u>

The gamma density function is given by

$$p_X(x)=\frac{\lambda^{\alpha}e^{-\lambda x}X^{(\alpha-1)}}{\Gamma(\alpha)},X,\ \alpha,\ \text{and}\ \lambda>0 \tag{I-34}$$

where α and λ are the distribution parameters. Using method of moments α and λ are expressed (Haan, 1977) as

$$\hat{\lambda}=\frac{\mu_X}{\sigma_X^2} \tag{I-35}$$

$$\hat{\alpha}=\frac{\mu_X^2}{\sigma_X^2}=\frac{1}{CV_X^2} \tag{I-36}$$

Substituting (I-34) in (2-26), $E[X^r]$ is written as

$$E\left(X^r\right)=\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\int_0^{\infty}\frac{e^{-\lambda X}X^{(\alpha+r-1)}}{\Gamma(\alpha)}dX=\frac{\Gamma(\alpha+r)}{\lambda^r\Gamma(\alpha)} \tag{I-37}$$

Replacing *r* by *2r* in (I-37), $E[X^{2r}]$ is written as

$$E\left(X^{2r}\right)=\frac{\Gamma(\alpha+2r)}{\lambda^{2r}\Gamma(\alpha)} \tag{I-38}$$

Substituting $r = 1$ in (I-37), μ_X can be given as

$$\mu_X=\frac{\Gamma(\alpha+1)}{\lambda\Gamma(\alpha)}=\frac{\alpha}{\lambda} \tag{I-39}$$

Substituting (I-39) in (2-22), the FOA estimate for mean μ_Y can be given as

$$\hat{\mu}_Y=c\left(\frac{\alpha}{\lambda}\right)^r \tag{I-40}$$

The exact value of μ_Y can be obtained by substituting (I-37) in (2-26) as

$$\mu_Y=c\frac{\Gamma(\alpha+r)}{\lambda^r\Gamma(\alpha)} \tag{I-41}$$

Substituting (I-40) and (I-41) in (2-24), the relative error in FOA predicted mean $E(\hat{\mu}_Y)$ can written as

$$\mathrm{E}(\hat{\mu}_Y) = 1 - \frac{\alpha^r \Gamma(\alpha)}{\Gamma(\alpha + r)} \qquad \text{(I-42)}$$

Substituting α in terms of CV_X from (I-36), (I-42) is rewritten as

$$\mathrm{E}(\hat{\mu}_Y) = 1 - \frac{CV_X^{-2r}\Gamma\left(\frac{1}{CV_X^2}\right)}{\Gamma\left(\frac{1 + rCV_X^2}{CV_X^2}\right)} \qquad \text{(I-43)}$$

Substituting (I-39) in (2-23), the FOA estimate for variance σ^2_Y can be written as

$$\hat{\sigma}_Y^2 = \frac{c^2 r^2 \alpha^{2r} CV_X^2}{\lambda^{2r}} \qquad \text{(I-44)}$$

Substituting $E[X^r]$ from (I-37) and $E[X^{2r}]$ from (I-38) in (2-27), the exact variance can be written as

$$\sigma_Y^2 = \frac{c^2}{\lambda^{2r}}\left[\frac{\Gamma(\alpha)\Gamma(\alpha + 2r) - \Gamma^2(\alpha + r)}{\Gamma^2(\alpha)}\right] \qquad \text{(I-45)}$$

Substituting (I-44) and (I-45) in (2-24), the relative error in FOA predicted variance $\mathrm{E}(\hat{\sigma}_y^2)$ can be expressed as

$$\mathrm{E}(\hat{\sigma}_y^2) = 1 - \frac{r^2 \alpha^{2r} CV_X^2 \Gamma^2(\alpha)}{\Gamma(\alpha)\Gamma(\alpha + 2r) - \Gamma^2(\alpha + r)} \qquad \text{(I-46)}$$

Substituting (I-36) in (I-46), $\mathrm{E}(\hat{\sigma}_y^2)$ can be rewritten as

$$\mathrm{E}(\hat{\mu}_Y) = 1 - \frac{r^2 CV_X^{2(1-2r)}\left[\Gamma\left(\frac{1}{CV_X^2}\right)\right]^2}{\Gamma\left(\frac{1 + 2rCV_X^2}{CV_X^2}\right)\Gamma\left(\frac{1}{CV_X^2}\right) - \left[\Gamma\left(\frac{1 + rCV_X^2}{CV_X^2}\right)\right]^2} \qquad \text{(I-47)}$$

Exponential Distribution

Exponential distribution is a special case of the gamma distribution with $\alpha = 1$. Substituting $\alpha = 1$ in (I-42), the relative error in FOA predicted mean is given as

$$E(\hat{\mu}_Y) = 1 - \frac{1}{\Gamma(r+1)} \tag{I-48}$$

Substituting $\alpha = 1$ in (I-36), $CV_X = 1$. On substituting $CV_X = 1$ in (I-47), the relative error in FOA predicted variance $E(\hat{\sigma}_y^2)$ can be expressed as

$$E(\hat{\sigma}_y^2) = 1 - \frac{r^2}{\left[\Gamma(2r+1) - \Gamma^2(r+1)\right]} \tag{I-49}$$

Normal Distribution

The probability density function of normal distribution is

$$p_X(x) = \frac{1}{\sigma_X \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2} \tag{I-50}$$

where μ_x and σ^2_x are the parameters of normal distribution.

Substituting (I-50) in (2-26), the $E[X^r]$ is written as

$$E\left(X^r\right) = \frac{1}{\sigma_X \sqrt{2\pi}} \int_0^{\infty} X^r e^{-\frac{1}{2}\left(\frac{X-\mu_X}{\sigma_X}\right)^2} dX \tag{I-51}$$

Assuming, $\frac{X-\mu_X}{\sigma_X} = Z$, the random variable X can be written as, $X = \left(\mu_X + \sigma_X Z\right)$, (I-51) is rewritten as

$$E\left(X^r\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu_X + \sigma_X Z)^r e^{-\frac{1}{2}Z^2} dZ \tag{I-52}$$

Equation (I-52) is difficult to integrate. Its integral exists when x is represented by the standard normal distribution for which $\mu_X = 0$, $\sigma_X = 1$. The resulting equation is

$$E\left(Z^r\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} Z^r e^{-\frac{1}{2}Z^2}\, dZ \tag{I-53}$$

The integral of (I-53) is

$$E\left(Z^r\right)=\frac{2^{r/2}\,\Gamma[(r+1)/2]}{\sqrt{\pi}}=(r-1)(r-3).........(3)(1)\text{, when } r \text{ is even}$$

$$E\left(Z^r\right)=0\text{, when } r \text{ is odd} \tag{I-54}$$

Equation (I-54) can be used to compute $E[X^r]$.

$$E\left(X^r\right)=E\left[(\mu_X+\sigma_X Z)^r\right]=\mu_X^r E\left[(1+CV_X Z)^r\right] \tag{I-55}$$

When $CV_X < 1.0$, (I-55) can be expanded using Binomial Theorem as

$$E\left(X^r\right)=\mu_X^r E\left[1+rCV_XZ+\frac{r(r-1)}{2!}CV_X^2Z^2+......+\frac{r(r-1)(r-2)....(r-n+1)}{n!}CV_X^nZ^n+...\right] \tag{I-56}$$

Taking expectation of all terms, (I-56) is written as

$$E\left(X^r\right)=\mu_X^r\left[1+rCV_XE[Z]+\frac{r(r-1)}{2!}CV_X^2E[Z^2]+......+\frac{r(r-1)(r-2)....(r-n+1)}{n!}CV_X^nE[Z^n]+...\right] \tag{I-57}$$

From (I-54), $E[Z^r]$ =0 when r is odd. Therefore all the terms containing odd powers of Z will vanish from (I-57) and the resulting equation is written as

$$E\left(X^r\right)=\mu_X^r\left[1+\frac{r(r-1)}{2!}CV_X^2E[Z^2]+......+\frac{r(r-1)(r-2)....(r-n+1)}{n!}CV_X^nE[Z^n]+...\right] \tag{I-58}$$

When r is a positive integer, RHS of (I-58) is finite and terminates when $n = r + 1$. Consequently, (I-57) can be written as

$$E\left(X^r\right)=\mu_X^r\sum_{i=0}^{r/2}\binom{r}{2i}CV_X^{2i}E\left[Z^{2i}\right]\text{, when } r \text{ is even} \tag{I-59a}$$

$$E\left(X^r\right) = \mu_X^r \sum_{i=0}^{(r-1)/2} \binom{r}{2i} CV_X^{2i} E\left[Z^{2i}\right], \text{ when } r \text{ is odd} \quad \text{(I-59b)}$$

For values of r other than a positive integer (2-53) does not converge. In order to determine $E[X^r]$, (2-58) needs to be truncated. When r is a positive fraction, $E[X^r]$ can be obtained using (2-59a) and (2-59b) with rounded value of r to its nearest whole number. In cases when r is negative, the truncating error depends upon the magnitudes of r and CV_X. Further, there exists a minimum truncating error for a given combination of r and CV_X, beyond which no improvement in $E[X^r]$ is possible. To evaluate the approximate value of $E[X^r]$, a trial and error procedure was used to determine the number of terms to be summed up to obtain $E[X^r]$ corresponding to the minimum truncating error for a given combination of r and CV_X.

After estimating $E[X^r]$ and $E[X^{2r}]$, (2-26) and (2-27) are used to determine μ_Y and σ_Y^2. Substituting μ_Y, σ_Y^2 and the FOA estimates $\hat{\mu}_Y$ and $\hat{\sigma}_Y^2$ into (2-24), the relative error in FOA predicted estimates of the mean and variance, $E(\hat{\mu}_Y)$ and $E(\hat{\sigma}_y^2)$, can be determined. Figures 2-10 and 2-11 show plots of $E(\hat{\mu}_Y)$ and $E(\hat{\sigma}_y^2)$ vs. r for various values of CV_X ranging from 0.02 to 0.33.

APPENDIX II

CHARACTERIZATION OF A TRIANGULAR DISTRIBUTION

Consider a random variable $X \sim$ Triangular (a, b, c), where a, b, and c are the minimum, maximum , and mode values of a random variable X. The values of μ_x, CV_x, and γ_x can be represented in terms of a, b, and c as

$$\mu_X = \frac{1}{3}(a+b+c) \tag{II-1}$$

$$CV_X = \frac{1}{\sqrt{2}} \frac{\sqrt{a^2+b^2+c^2-(ab+bc+ca)}}{a+b+c} \tag{II-2}$$

$$\gamma_X = \frac{\sqrt{2}}{5} \frac{2(a^3+b^3+c^3)-3[ab(a+b)+bc(b+c)+ca(c+a)]+12abc}{[a^2+b^2+c^2-(ab+bc+ca)]^{3/2}} \tag{II-3}$$

If the values of μ_x, CV_x, and γ_x are known then a unique triangle can be delineated by determining its parameters a, b, and c. Simplifying (II-1), (II-2), and (II-3) and writing known quantities on the RHS as

$$a+b+c = 3\mu_X \tag{II-4}$$

$$ab+bc+ca = 3\mu_X^2(1-2CV_X^2) \tag{II-5}$$

$$abc = \mu_X^3(10\gamma_X CV_X^3 - 6CV_X^2 + 1) \tag{II-6}$$

Eliminating a and b, (II-4), (II-5), and (II-6) can be expressed in terms of c as

$$c^3 - 3\mu_X c^2 + 3\mu_X^2(1-2CV_X^2)\, c - \mu_X^3(10\gamma_X CV_X^3 - 6CV_X^2 + 1) = 0 \tag{II-7}$$

Equation (II-7) is a cubic equation which can be solved using Cardan's method (Borofsky, 1950). In order to solve (II-7), the quadratic term needs to be eliminated. Substituting $c = y + \mu_X$, (II-7) can be rewritten as

$$y^3 - 6\mu_X^2 CV_X^2 y - 10\gamma_X \mu_X^3 CV_X^3 = 0 \qquad \text{(II-8)}$$

For a real triangle, (II-8) does not involve any imaginary quantities. This can be enforced with the help of the triple angle formula

$$4\cos^3\theta - 3\cos\theta = \cos 3\theta \qquad \text{(II-9)}$$

Substituting $w = \cos\theta$, (II-9) is rewritten as

$$w^3 - \frac{3}{4}w - \frac{\cos 3\theta}{4} = 0 \qquad \text{(II-10)}$$

It is clear that $w = \cos\theta$ is a root of (II-10). Using this, (II-8) can be solved by analogy. Substituting $y = \lambda w$ in (II-8)

$$w^3 - \frac{6\mu_X^2 CV_X^2}{\lambda^2} w - \frac{10\gamma_X \mu_X^3 CV_X^3}{\lambda^3} = 0 \qquad \text{(II-11)}$$

Comparing the coefficients of w in (II-10) and (II-11)

$$\frac{6\mu_X^2 CV_X^2}{\lambda^2} = \frac{3}{4} \qquad \text{(II-12)}$$

Comparing the constant terms of (II-10) and (II-11)

$$\frac{10\gamma_X \mu_X^3 CV_X^3}{\lambda^3} = \frac{\cos 3\theta}{4} \qquad \text{(II-13)}$$

Simplifying (II-12) and retaining only positive value, λ is given as

$$\lambda = 2\sqrt{2}\mu_X CV_X \qquad \text{(II-14)}$$

On simplifying (II-13)

$$\cos 3\theta = \frac{5}{2\sqrt{2}}\gamma_X = \cos\left[\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_X\right)\right] \qquad \text{(II-15)}$$

Equation (II-15) is a trigonometric equation, which can be solved for θ as

$$\theta = \frac{2\pi n}{3} + \frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_X\right) \qquad \text{(II-16)}$$

Equation (II-16) gives a general value of θ, substituting $n = 0$, 1, and 2 particular values of θ can be obtained. Using (II-16), the general solution of (II-8) is

$$y = \lambda z = \lambda\cos\theta = 2\sqrt{2}\mu_X CV_X \cos\left[\frac{2\pi n}{3} + \frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_X\right)\right] \qquad \text{(II-17)}$$

Using (II-17), the solution of (II-7) is

$$c = y + \mu_X = \mu_X\left\{1 + 2\sqrt{2}CV_X \cos\left[\frac{2\pi n}{3} + \frac{1}{3}\cos^{-1}\left(\frac{5}{2\sqrt{2}}\gamma_X\right)\right]\right\} \qquad \text{(II-18)}$$

Equation (II-18) shows that maximum magnitude of coefficient of skewness for a triangular distribution is $\frac{2\sqrt{2}}{5}$.

VITA

Aditya Tyagi

Candidate for the Degree of

Doctor of Philosophy

Thesis: A SIMPLE APPROACH TO RELIABILITY, RISK, AND UNCERTAINTY ANALYSIS OF HYDROLOGIC, HYDRAULIC, AND ENVIRONMENTAL ENGINEERING SYSTEMS

Major Field: Biosystems Engineering

Biographical:

Education: Graduated from Sanatan Dharm Inter College, Muzaffarnagar, UP, India in 1981; received Bachelor of Engineering degree in Civil Engineering from University of Roorkee, Roorkee, India in May 1987; received Master of Civil Engineering degree with major in Environmental Engineering from University of Roorkee, Roorkee, India in May 1989; completed the requirements for the Doctor of Philosophy degree with major in Biosystems Engineering at Oklahoma State University in December 2000.

Experience: Employed by Punjab University, Chandigarh, India, as an assistant professor in civil engineering department, April 1989- January 1990; employed by National Institute of Hydrology, Roorkee, India as a scientist, February 1990-December 1997; Employed by Oklahoma State University, Department of Biosystems and Agricultural Engineering as a graduate research assistant, January 1998-December 2000.

Professional Memberships: American Geophysical Union, The Honor Society of Agricultural Engineers.