WATER RESOURCES ENGINEERING – ANALYSIS OF UNCERTAINTY

LILIANA TOPLICEANU

University of Bacau

Abstract: Water resources engineering design and analysis deal with the occurrence of water in various parts of a hydro-system and its effects on environmental, ecological, and socio-economical settings. Due to the extreme complex nature of the physical, chemical, biological, and socio-economical processes involved, tremendous efforts have been devoted by researchers attempting to have a better understanding of the processes. The models presented are a useful tool to assess the system performance under various scenarios based on which efficient designs or effective management schemes can be formulated.

Keywords: uncertainty analysis**,** hydro-systems, model, management scheme

1. INTRODUCTION

In water resources engineering most models are structural which take the forms of mathematical equations, tables, graphs, or computer programs. Despite numerous research efforts made to further our understanding of various processes in hydro-systems, there is still much more that are beyond our firm grasp. In general, uncertainty due to inherent randomness of physical processes cannot be eliminated. On the other hand, uncertainties such as those associated with lack of complete knowledge about the process, models, parameters, data, and etc. could be reduced through research, data collection, and careful manufacturing.

Uncertainties involved in water resources engineering, can be divided into four basic categories: hydrologic, hydraulic, structural, and economic. More specifically, in water resources engineering analyses and designs uncertainties could arise from the various sources including natural uncertainties, model uncertainties, parameter uncertainties, data uncertainties, and operational uncertainties. Natural uncertainty is associated with the inherent randomness of natural processes such as the occurrence of precipitation and flood events.

The occurrence of hydrological events often displays variations in time and in space. Their occurrences and intensities could not be predicted precisely in advance, due to the fact that a model is only an abstraction of the reality, which generally involves certain degrees of simplifications and idealizations. Model uncertainty reflects the inability of a model or design technique to represent precisely the system's true physical behavior.

Data uncertainties include: (1) measurement errors, (2) inconsistency of data, (3) data handling and transcription errors, and (4) inadequate representation of data sample due to time and space limitations. Operational uncertainties include those associated with construction, manufacture, deterioration, maintenance, and human. The magnitude of this type of uncertainty is largely dependent on the workmanship and quality control during the construction and manufacturing. Progressive deterioration due to lack of proper maintenance could result in changes in resistance coefficients and structural capacity reduction.

2. IMPLICATIONS OF UNCERTAINTY ANALYSIS

In water resources engineering design and analysis, the decisions on the layout, capacity, and operation of the system largely depend on the system response under some anticipated design conditions. When some of the components in a hydro-system are subject to uncertainty, the system responses under the design conditions cannot be assessed with certainty. Therefore, the conventional deterministic design practice is inappropriate because it is unable to account for possible variation of system responses. An engineer has to consider various criteria including, but not limited to, cost of the system, probability of failure, and consequence of failure so that a proper design can be made for the system.

In this type of analyses the design quantity and system output are functions of several system parameters and not all of them can be quantified with absolute accuracy. The task of uncertainty analysis is to determine the uncertainty features of the system outputs as a function of uncertainties in the system model itself and the stochastic variables involved. It provides a formal and systematic framework to quantify the uncertainty associated with the system output. Furthermore, it offers the designer useful insights regarding the contribution of each stochastic variable to the overall uncertainty of the system 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 of the system outputs.

2.1 Measures of uncertainty

Several expressions have been used to describe the degree of uncertainty of a parameter, a function, a model, or a system. In general, the uncertainty associated with the latter three is a result of combined effect of the uncertainties of the contributing parameters. 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 such a probability function cannot be derived or found precisely.

Another measure of the uncertainty of a quantity is to express it in terms of a reliability domain such as the confidence interval. A confidence interval is a numerical interval that would capture the quantity subject to uncertainty with a specified probabilistic confidence. Nevertheless, the use of confidence intervals has a few drawbacks: (1) the parameter population may not be normally distributed as assumed in the conventional procedures and this problem is particularly important when the sample size is small; (2) no means is available to directly combine the confidence intervals of individual contributing random components to give the overall confidence interval of the system.

A useful alternative to quantify the level of uncertainty is to use the statistical moments associated with a quantity subject to uncertainty. In particular, the variance and standard deviation which measure the dispersion of a stochastic variable are commonly used.

3. UNCERTAINTY ANALYSIS TECHNIQUES

Several techniques can be applied to conduct uncertainty analysis of water resources engineering problems. Each technique has different levels of mathematical complexity and data requirements. Broadly speaking, those techniques can be classified into two categories: analytical approaches and approximated approaches. The selection of an appropriate technique to be used depends on the nature of the problem including availability of information, resources constraints, model complexity, and type and accuracy of results de sired.

3.1Analytical Techniques

We describe several analytical methods that allow an analytical derivation of the exact PDF and/or statistical moments of a model as a function of several stochastic variables. Although the analytical techniques are rather restrictive in practical applications due to the complexity of most models, they are, nevertheless, powerful tools for deriving complete information about a stochastic process, including its distribution, in some situations. The analytical techniques described herein are straightforward. However, the success of implementing these procedures largely depends on the functional relation, forms of the PDFs involved, and analyst's mathematical skill.

Derived Distribution Technique - This derived distribution method is also known as the transformation of variables technique. Example applications of this technique can be found in modeling the distribution of pollutant decay process and rainfall-runoff modeling.

Fourier Transform Technique - The Fourier transform of the PDF of a stochastic variable X results in the so-called the characteristic function. The characteristic function of a stochastic variable always exists and two distribution functions are identical if and only if the corresponding characteristic functions are identical. Therefore, given a characteristic function of a stochastic variable, its PDF can be uniquely determined through the inverse Fourier transform. Also, the statistical moment of the stochastic variable X can be obtained by using the characteristic function. Fourier transform is particularly useful when stochastic variables are independent and linearly related. In such cases, the convolution property of the Fourier transform can be applied to derive the characteristic function of the resulting stochastic variable.

Laplace and Exponential Transform Techniques – The Laplace and exponential transforms of the PDF of a stochastic variable lead to the moment generating function. Similar the characteristic function, statistical moments of a stochastic variable X can be derived from its moment, generating function. There are two deficiencies associated with the moment generating functions: (1) the moment generating function of a stochastic variable may not always exist, and (2) the correspondence between a PDF and moment generating function may not necessarily be unique. However, the existence and unique conditions are generally satisfied in most situations. Fourier and exponential transforms are frequently used in uncertainty analysis of a model that involves exponentiation of stochastic variables. Examples of their applications can be found in probabilistic cash flow analysis and probabilistic modeling of pollutant decay.

Mellin Transform Technique - When the functional relation of a model satisfies the product form and the stochastic variables are independent and non-negative, the exact moments for model output of any order can be derived analytically by the Mellin transform. The Mellin transform is particularly attractive in uncertainty analysis of hydrologic and hydraulic problems because many models and the involved parameters satisfy the above two conditions. Similar to the convolution property of the Laplace and Fourier transforms, the Mellin transform of the convolution of the PDFs associated with independent stochastic variables in a product form is simply equal to the product of the Mellin transforms of individual PDFs. Applications of the Mellin transform can be found in economic benefit-cost analysis, and hydrology and hydraulics. One caution about the use of the Mellin transform is that under some combinations of distribution and functional form, the resulting transform may not be defined. This could occur especially when quotients or variables with negative exponents are involved.

Estimations of Probabilities and Quantiles Using Moments - Although it is generally difficult to analytically derive the PDF from the results of the integral transform techniques described above and the approximation techniques in the next section, it is, however, rather straightforward to obtain or estimate the statistical moments of the stochastic variable one is interested in. Based on the computed statistical moments, one is able to estimate the distribution and quantile of the stochastic variable. One possibility is to base on the asymptotic expansion about the norm al distribution for calculating the values of CDF and quantile, and the other is to base on the maxim um entropy concept.

3.2 Approximation Techniques

Most of the models or design procedures used in water resources engineering are nonlinear and highly complex. This basically prohibits any attempt to derive the probability distribution or the statistical moments of model output analytically. As a practical alternative, engineers frequently resort to methods that yield approximations to the statistical properties of uncertain model output. In this section, several methods that are useful for uncertainty analysis are briefly described.

First-order variance estimation (FOVE) method - The method, also called the variance propagation method, estimates uncertainty features associated with a model output based on the statistical properties of model's stochastic variables. The basic idea of the method is to approximate a model by the first-order Taylor series expansion. Commonly, the FOVE method takes the expansion point at the means of the stochastic variables. Consider a hydraulic or hydrologic design quantity W which is related to N stochastic variables

$$X = (X_1, X_2, X_3, ...X_N) \text{ as } W = g(X_1, X_2, ...X_N) \quad (1)$$

The mean of W, by the FOVE method, can be estimated as

$$E[W] = g((X_1, X_2, ...X_N) \quad (2)$$

in which *i* is the mean of the i-th stochastic variable. When all stochastic variables are independent, the variance of the design quantity W can be approximated as

$$Var[W] = s_1^2\ 1^2 + s_2^2\ 2^2 + \cdots\cdots + s_N^2\ N^2 \quad (3)$$

in which s_i is the first-order sensitivity coefficient of the *i* stochastic variable and s_i represents the corresponding standard deviation. From the above equation, the ratio $s_i^2\ i^2 / Var[W]$ indicates the proportion of overall uncertainty in the design quantity contributed by the uncertainty associated with the stochastic variable X_i.

In general, E[g(**X**)] is a linear function of **X**. Improvement of the accuracy can be made by incorporating higher-order terms in the Taylor expansion. The method can be expanded to include the second-order term to improve estimation of the mean to account for the presence of model nonlinearity and correlation between stochastic variables. The method does not require knowledge of the PDF of stochastic variables which simplifies the analysis. However, this advantage is also the disadvantage of the method because it is insensitive to the distributions of stochastic variables on the uncertainty analysis.

The FOVE method is simple and straightforward. The computational effort associated with the method largely depends on the ways how the sensitivity coefficients are calculated. For simple analytical functions the computation of derivatives are trivial tasks. However, for functions that are complex and/or implicit in the form of computer programs, or charts/ figures, the task of computing the derivatives could become cumbersome or difficult. In such cases probabilistic point estimation techniques can be viable alternatives. There are many applications of the FOVE method in the literature. Example applications of the method can be found in open channel flow, groundwater flow, water quality modeling, benefit-cost analysis, gravel pit migration analysis, storm sewer design, culverts, and bridges.

Probabilistic Point Estimation (PE) Methods - Unlike the FOVE methods, probabilistic PE methods quantify the model uncertainty by performing model evaluations without computing the model sensitivity. The methods generally are simpler and more flexible especially when a model is either complex or non-analytical in the forms of tables, figure, or computer programs. Several types of PE methods have been developed and applied to uncertainty analysis and each has its advantages and disadvantages. It has been shown that the FOVE method is a special case of the probabilistic PE methods when the uncertainties of stochastic variables are small.

Rosenblueth in 1975 developed a method for handling stochastic variables that are symmetric and the method is later extended to treat no n-sym metric stochastic variables in 1981. The basic idea of Rosenblueth's PE method is to approximate the original PDF or PMF of the stochastic variable by assuming that the entire probability mass is concentrated at two points. The four unknowns, namely, the locations of the two points and the corresponding probability masses, are determined in such a manner that the first three moments of the original stochastic ariable are preserved. For problems involving N stochastic variables, the two points for each variable are computed and permutated to produce a total of 2^N possible points of evaluation in the parameter space based on which the statistical moments of the model outputs are computed.

The potential drawback of Rosenblueth's PE method is its practical application due to explosive nature of the computation requirement. For moderate or large N, the number of required model evaluations could be too numerous to be implemented practically, even on the computer. Example applications of Rosenblueth’s PE method for uncertainty analysis can be found in groundwater flow model, dissolved oxygen deficit model, and bridge pier scouring model.

To circum vent the shortcoming in computation, was developed an alternative PE method that reduces the 2^N model evaluations required by Rosenblueth's method down to 2^N. This method utilizes the first two moments (that is, the mean and covariance) of the involved stochastic variables. The method is appropriate for treating stochastic variables that are normal. The theoretical basis of PE method is built on the orthogonal transformation using eigenvalue-eigenvector decomposition which maps correlated stochastic variables from their original space to a new domain in which they become uncorrelated. Hence, the analysis is greatly simplified. PE method has been applied to uncertainty analysis of a gravel pit migration model, regional equations for unit hydrograph parameters, groundwater flow models, and parameter estimation of a distributed hydrodynamic model.

Recently, was proposed a computationally practical PE method that allows incorporation of the first four moments of correlated stochastic variables. Among the probabilistic algorithms presented PE method is the most attractive from the computational viewpoint. However, the method cannot incorporate additional distribution al information of the stochastic variables other than the first two moments. Such distributional information could have important effects on the results of uncertainty analysis. To incorporate the information about the marginal distributions of involved stochastic variables, a transformation between non-normal parameter space and a multivariate standard normal space has been incorporated into the method. The resulting preserves the computational efficiency of PE method while extends its capability to hand le multivariate non-normal stochastic variables.

Monte-Carlo Simulation **-** Simulation is a process of replicating the real world based on a set of assumptions and conceived models of reality. Because the purpose of a simulation model is to duplicate reality, it is a useful tool for evaluating the effect of different designs on system performance. The Monte Carlo procedure is a numerical simulation to reproduce stochastic variables preserving the specified distributional properties.

Several books have been written for generating univariate random numbers. A number of computer programs are available in the public domain. The challenge of Monte Carlo simulation lies in generating multivariate random varieties. Compared with univarietie random, variated generators algorithms for multivariate random variates are much more restricted to a few joint distributions such as multivariate normal, multivariate lognormal, multivariate gamma, and few others. If the multivariate stochastic variables involved are correlated with a mixture of marginal distributions, the joint PDF is difficult to formulate.

Rather than preserving the full multivariate features, practical multivariate Monte Carlo simulation procedures for problems involving mixtures of non-normal stochastic variables have been developed to preserve the marginal distributions and correlation of involved stochastic variables.

In uncertainty analysis, the implementation of brutal force type of simulation is straightforward but can be very computationally intensive. Furthermore, because the Monte Carlo simulation is a sampling procedure, the results obtained inevitably involve sampling errors which decrease as the sample size increases. Increasing sample size for achieving higher precision generally means an increase in computer time for generating random varieties and data processing. Therefore, the issue lies on using the minimum possible computation to gain the maximum possible accuracy for the quantity under estimation. For this, various variance reduction techniques have been developed.

Applications of Monte Carlo simulation in water resources engineering are abundant. Examples can be found in groundwater, benefit-cost analysis, water quality model, pier-scouring prediction, and open channel.

Resembling Techniques - The Monte Carlo simulations are conducted under the condition that the probability distribution and the associated population parameters are known for the stochastic variables involved in the system. The observed data are not directly utilized in the simulation. Unlike the Monte Carlo simulation approach, resembling techniques reproduce random data exclusively on the basis of observed ones. The two resembling techniques that are frequently used are jackknife method and bootstrap method.

CONCLUSIONS

In many water resource engineering problems, uncertainties in data and in theory, including design and analysis procedures, warrant a probabilistic treatment of the problems. The failure associated with a hydraulic structure is the result of the combined effect from inherent randomness of extern al load and various uncertainties involved in the analysis, design, construction, and operation al procedures described previously. Failure of an engineering system occurs when the load (external forces or demands) on the system exceeds the resistance (strength, capacity, or supply) of the system. In hydraulic and hydrologic analyses, the resistance and load are frequently functions of a number of stochastic variables. Without considering the time-dependence of the load and resistance, static reliability model is generally applied to evaluate the system performance subject to a single worst load event. However, a hydraulic structure is expected to serve its designed function over an expected period of time. In such circumstances, time-dependent of models are used to incorporate the effects of service duration, random ness of occurrence of loads, and possible change of resistance characteristics over time.

REFERENCES

[1] Bevington, P. R. and Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences*. WCB/Mc Graw-Hill Boston USA, 1992.
[2] Cîrstoiu, D. I. *Sisteme expert*, Editura All, Bucuresti, 1994.
[4] Cullen, A.C., H.C. Frey, H.C. *Probabilistic Techniques in Exposure Assessment*, Plenum Publishing Corp., New York, USA, 1999.
[8] Dou, C. H., Woldt, W., Bogardi, I. *Steady state groundwater flow simulation with imprecise parameters*. Water Resources Research., 1995.
[7] Evans, G. W., Karwowski, W., Wilhelm, M. R. *Applications of Fuzzy Set Methodologies* in Industrial Engineering,. Elsevier, New York, 1986.
[6] Evans, G. W., Karwowski, W. *An Introduction to Fuzzy Set Methodologies for Industrial and Systems Engineering*, New York, 1992.
[5] Fishman, G. S. *Monte Carlo : Concepts, Algorithms, and Applications*. Springer Verlag, New York, 1996.
[3] Mandel, J. *The Statistical Analysis of Experimental Data*, Dover Publications New York, USA, 1984.
[9] Negoita, C.V. *Expert Systems and Fuzzy Systems*, Benjamin Cummings Publishing Company Inc., 1985.
[10] Yen, B.C., Tung, Y.K., *Reliability and Uncertainty Analyses in Hydraulic Designs*, the American Society of Civil Engineers, New York, NY, 1993.