Monte Carlo Simulation for Correlated Variables with Marginal Distributions

By Che-Hao Chang,[1] Yeou-Koung Tung,[2] Associate Member, ASCE, and Jinn-Chuang Yang,[3] Associate Member, ASCE

Abstract: As computation speed increases, Monte Carlo simulation is becoming a viable tool for engineering design and analysis. However, restrictions are often imposed on multivariate cases in which the involved stochastic parameters are correlated. In multivariate Monte Carlo simulation, a joint probability distribution is required that can only be derived for some limited cases. This paper proposes a practical multivariate Monte Carlo simulation that preserves the marginal distributions of random variables and their correlation structure without requiring the complete joint distribution. For illustration, the procedure is applied to the reliability analysis of a bridge pier against scouring.

INTRODUCTION

As reliability related issues are becoming more critical in engineering design and analysis, proper assessment of stochastic behavior of an engineering system is essential. The true distribution for the system response subject to parameter uncertainty should be derived, if possible. However, due to the complexity of physical systems and mathematical functions, derivation of the exact solution for the random characteristics of the system response is difficult. In such cases, Monte Carlo simulation is a viable tool to provide numerical estimations of the stochastic features of the system response.

Monte Carlo simulation is like to repeatedly measuring the system response of interest under various system parameter sets generated from the known or assumed probabilistic laws. It offers a practical approach to reliability analysis because the stochastic behavior of the system response can be probabilistically duplicated.

Two major concerns in practical applications of Monte Carlo simulation are: (1) The requirement of tremendous computations for generating random variates; and (2) the presence of correlation among stochastic system parameters. In fact, the former concern is diminishing as the computing power increases. As for the second concern, it has been pointed out that neglecting correlation could have significant effect on the result of reliability analysis (Thoft-Christensen and Baker 1982). Therefore, a proper assessment of joint probability density function (PDF) for the correlated parameters is necessary in the generation of multivariate random variables. Compared with a variety of univariate random variate generators, generating multivariate random variates is much more restricted to a few joint distributions such as multivariate normal, multivariate lognormal (Parrish 1990),

[1]Grad. Student, Dept. of Civ. Engrg., Nat. Chiao-Tung Univ., Hsinchu, Taiwan, Republic of China.

[2]Prof., Wyoming Water Res. Ctr. and Statistics Dept., Univ. of Wyoming, Laramie, WY 82071.

[3]Prof., Dept. of Civ. Engrg., Nat. Chiao-Tung Univ., Hsinchu, Taiwan, Republic of China.

Note. Discussion open until August 1, 1994. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on May 6, 1993. This paper is part of the *Journal of Hydraulic Engineering*, Vol. 120, No. 3, March, 1994. ©ASCE, ISSN 0733-9429/94/0003-0313/$2.00 + $.25 per page. Paper No. 6129.

and multivariate gamma (Ronning 1977). If the random variables involved are correlated with a mixture of marginal distributions, the multivariate PDF is difficult to formulate.

In most practical engineering problems involving multivariate random variables, one can determine the marginal distribution of each individual random variable and the first two moments including covariances or correlations. To generate correlated random variables (not necessarily all normal), a methodology adopting a bivariate distribution model was suggested by Li and Hammond (1975). Based on Nataf bivariate distribution model, a set of semi-empirical formulas was derived by Der Kiureghian and Liu (1985) to reduce the necessary calculations. This set of formulas transforms the correlation coefficient of a pair of nonnormal random variables to its equivalent correlation coefficient in a bivariate standard normal space. Through this transformation, multivariate Monte Carlo simulation can be performed in a correlated standard normal space in which several efficient algorithms have been developed.

This paper considers multivariate Monte Carlo simulation for correlated random variables with known marginal PDFs. Brief descriptions are given to some basic concepts of Monte Carlo simulation and the multivariate distribution model from which the semi-empirical formulas were originated. Random variates generated by the proposed Monte Carlo simulation are examined to check whether the distributional properties of the original stochastic variables are preserved. For demonstration, an example applying the proposed Monte Carlo simulation procedure is applied to a reliability analysis of bridge pier scouring.

MONTE CARLO SIMULATION

Simulation is a process of replicating the real world based on a set of assumptions and conceived models of reality (Ang and Tang 1984). Because the purpose of a simulation model is to duplicate reality, it is an effective tool for evaluating the effect of different designs on system performance. Monte Carlo procedure is a numerical simulation to reproduce random variables preserving the specified distributional properties.

For a univariate problem, many algorithm have been developed to generate univariate random numbers of various distributions (Dagpunar 1988). These univariate algorithms often serve as the building blocks for multivariate Monte Carlo simulation problems. In a multivariate problem, the joint CDF for the random variables involved is required. If all random variables are statistically independent, multivariate generation can be accomplished by the appropriate univariate algorithms.

In most applications of multivariate Monte Carlo simulation, an assumption is often made of a multivariate normal distribution for the stochastic parameters involved. To generate multivariate standard normal random variates, several algorithms can be applied. One commonly used approach is the orthogonal transformation (Ang and Tang 1984). The algorithm decomposes the correlated normal random variables into uncorrelated ones, thus, each individual uncorrelated normal random variable can be dealt with separately. Then, several procedures such as the Box-Muller algorithm (Box and Muller 1958) can be applied to generate univariate normal random variates. Once the uncorrelated standard normal variates are produced, they are converted back to the correlated normal variates through the inverse orthogonal transform.

PROPOSED MULTIVARIATE MONTE CARLO SIMULATION PROCEDURE

In many practical engineering analyses, random variables are often statistically and physically dependent. Furthermore, distribution types for the random variables involved can be a mixture of different distributions. To properly replicate such systems, Monte Carlo simulation should be able to preserve the correlation relationship among the stochastic parameters and their distributions.

However, derivation of the joint CDF which describes the complete multivariate characteristics of random variables is generally difficult. This difficulty, in both theory and practice, increases with the number of random variables and the type of corresponding distributions. As a practical alternative, this section describes a procedure to generate multivariate random variates that preserves the marginal distributions and correlation structure of the random variables involved. In doing so, the difficulty of requiring a complete joint PDF in multivariate Monte Carlo simulation is circumvented.

Transformation Formulas

The proposed Monte Carlo simulation procedure employs a set of semi-empirical formulas developed by Der Kiureghian and Liu (1985). Instead of obtaining the joint distribution, these formulas transform the correlation coefficients of the original stochastic parameters to the equivalent correlation coefficients in standard normal space by using the following Nataf's bivariate distribution model:

$$\rho_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{x_i - \mu_i}{\sigma_i}\right)\left(\frac{x_j - \mu_j}{\sigma_j}\right)\phi_{ij}(y_i, y_j, \rho_{ij}^*)\, dy_i\, dy_j \quad \text{(1)}$$

in which Y_i and Y_j = the two correlated standard normal random variables having the equivalent marginal cumulative probabilities corresponding to a pair of stochastic parameters, X_i and X_j, in the original space; ρ_{ij}^* = the correlation coefficient between Y_i and Y_j; ρ_{ij} = the correlation coefficient between X_i and X_j; μ_i and σ_i are, respectively, the mean and standard deviation of X_i; ϕ_{ij} = the bivariate normal PDF of zero means, unit standard deviations, and correlation coefficient ρ_{ij}^*.

To avoid the required computation for solving ρ_{ij}^* in (1) when the correlation coefficient ρ_{ij} and the marginal distributions of X_i and X_j are given, Der Kiureghian and Liu (1985) developed a set of semi-empirical formulas as

$$\rho_{ij}^* = T_{ij} \cdot \rho_{ij} \quad \text{(2)}$$

in which T_{ij} = transformation factor depending on the marginal distributions and correlation of the two random variables considered. In the case that the pair of random variables under consideration are both normal, the transformation factor, T_{ij}, has a value of one. Given the marginal distributions and correlation for a pair of stochastic parameters, the formulas of Der Kiureghian and Liu (1985) compute the corresponding transformation factor, T_{ij}, to obtain the equivalent correlation ρ_{ij}^* as if the two stochastic parameters were bivariate normal random variables. After all pairs of stochastic parameters are treated, the correlation matrix in the correlated normal space, $\mathbf{R}_Y$, is obtained.

As aforementioned, the transformation factor is a function of the correlation coefficient between a pair of stochastic variables and their marginal

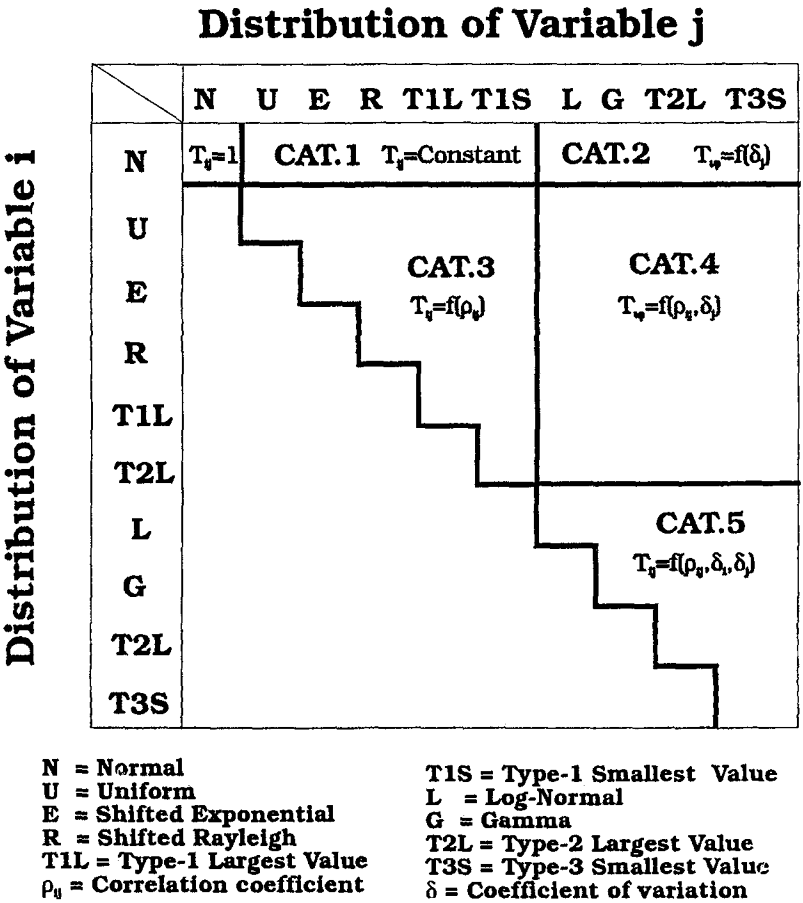

FIG. 1. Schematic Description of Categories of Transformation Factor T_{ij}

distributions. Therefore, for each combination of two distributions, there is a corresponding formula. A total of 54 formulas for 10 different distributions were developed and they were divided into five categories as shown in Fig. 1. The complete forms of these formulas can be referred to Der Kiureghian and Liu (1985) or Liu and Der Kiureghian (1986).

Two conditions are inherently considered in the bivariate distribution model of (1):

1. The mapping relationship given below is a one-to-one correspondence satisfying

$$Y_i = \Phi^{-1}[F_i(X_i)] \qquad (3)$$

in which Φ = the standard normal CDF. This condition preserves the probability content in both original and normal spaces.

2. The value of the correlation coefficient in the normal space lies between -1 and 1.

J. Hydraul. Eng. 1994.120:313-331.

Proposed Procedures

The proposed multivariate Monte Carlo simulation involves two basic steps:

1. Step 1. Transformation to standard normal space. Through (2), the operational domain is transformed to a standard normal space in which the transformed stochastic parameters are treated as if they were multivariate standard normal random variables with the correlation matrix $\mathbf{R}_Y$. As a result, multivariate normal random variates can be obtained by the orthogonal transform technique described previously.
2. Step 2. Inverse transformation. Once the standardized multivariate normal random variates are generated, then, according to (3), one can do the following inverse transformation

$$X_i = F_i^{-1}[\Phi(Y_i)] \quad \text{(4)}$$

to compute the values of multivariate random variates in the original space. Fig. 2 is a flow chart outlining the proposed procedure.

VERIFICATION OF PROPOSED SIMULATION PROCEDURE

The major concern with the proposed Monte Carlo simulation procedure is its ability to preserve the correlation structure and marginal distributions of stochastic parameters in the original space. The original correlation matrix is checked against the one generated from the proposed Monte Carlo simulation. Furthermore, consistency of the assigned distributions of stochastic parameters and the random variates generated from the proposed procedure is examined by Kolmogorov-Smirnov test. More specifically, this verification study examines respectively the means, standard deviations, and correlation coefficients of the generated random variates along with the significance probabilities in Kolmogorov-Smirnov test. The antithetic-variates technique was applied here to reduce the sampling variability of the relevant statistics and test index.

Experiments

In the verification experiments, three random variables with seven different mixtures of distributions, listed in Table 1, were used. Referring to Table 2, although different in sample sizes and numbers of simulation runs, a total of 5,000 random variates was kept constant for each individual random variable in each of the seven cases shown in Table 1. The purpose of doing it is to examine the effect of sample size on the simulation results. Representative results are shown in Tables 3, 4, 5, and 6. Furthermore, different correlation structures were used to examine the performance of the proposed simulation procedure. In all cases, the population (true) values of the correlation coefficients, means, and standard deviations used in the verification are given in the second column of Tables 3–8.

In each simulation run the performance criteria using percentage errors of relevant statistics are computed and they are:

1. Percentage of biasness (e_b):

$$e_b = \frac{1}{M}\sum_{i=1}^{M} e_i \quad \text{(5)}$$

J. Hydraul. Eng. 1994.120:313-331.

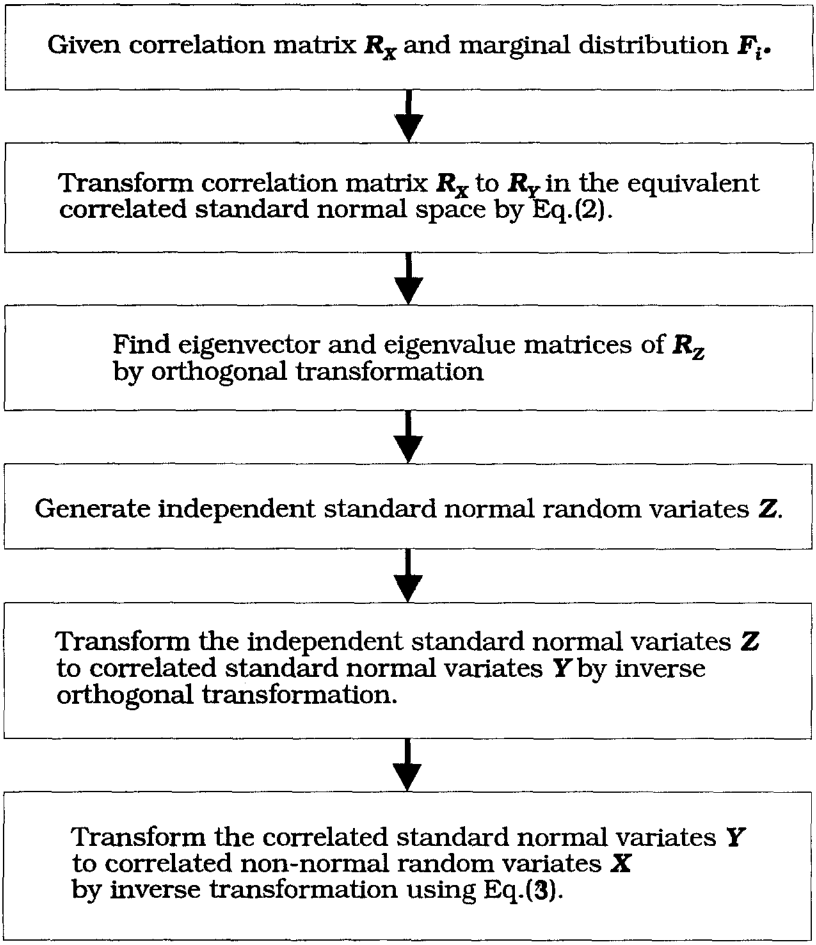

FIG. 2. Flow Chart of Proposed Monte Carlo Simulation Procedure

TABLE 1. Different Cases Considered in Verification Study

Case (1)	Variable 1 (2)	Variable 2 (3)	Variable 3 (4)
1	Normal	Normal	—
2	Uniform	Uniform	Uniform
3	Log-normal	Log-normal	Log-normal
4	Gamma	Gamma	Gamma
5	Normal	Uniform	Log-normal
6	Normal	Log-normal	Gamma
7	Log-normal	Gamma	Gumbel (max)

TABLE 2. Different Sample Sizes and Simulation Runs Used in Verification Study

Combination (1)	Sample size (2)	Number of simulation runs (3)
1	50	100
2	100	50
3	200	25
4	500	10

TABLE 3. Verification of Proposed Simulation Procedure Based on 50 Samples along with 100 Simulation Runs for Correlation Coefficients

Variable pair (1)	True value (2)	Error criteria (3)	Case 1 (4)	2 (5)	3 (6)	4 (7)	5 (8)	6 (9)	7 (10)
Variables 1 and 2	0.45	e_b (%)	0.26	−.46	−5.69	0.05	−1.19	−3.09	0.37
Variables 1 and 2	0.45	e_{mabs} (%)	13.42	12.94	15.26	13.73	14.13	13.44	13.62
Variables 1 and 2	0.45	e_{rms} (%)	16.18	15.89	19.41	16.95	17.55	17.09	16.86
Variables 1 and 3	0.30	e_b (%)	1.17	−3.81	−3.90	−7.67	−0.40	−0.12	−0.04
Variables 1 and 3	0.30	e_{mabs} (%)	26.18	25.50	26.03	29.49	22.51	22.20	29.05
Variables 1 and 3	0.30	e_{rms} (%)	32.25	31.35	32.83	36.00	27.98	27.65	35.44
Variables 2 and 3	0.42	e_b (%)	3.01	−4.22	−3.13	−2.06	−1.95	−2.42	−4.08
Variables 2 and 3	0.42	e_{mabs} (%)	15.86	16.98	16.63	15.41	14.49	17.28	16.41
Variables 2 and 3	0.42	e_{rms} (%)	19.67	21.43	21.23	20.08	18.10	21.08	21.34

where M = the number of simulation runs and e_i = the percentage error between the true and simulated values for a specified statistical parameter θ of interest. That is

$$e_i = \frac{\theta_t - \theta_{s_i}}{\theta_t} \quad (i = 1, 2, \ldots, M) \qquad (6)$$

with θ_t and θ_{si} being the true values and simulated values from the ith run, respectively. In the verification, the relevant statistics, θ, are the means, standard deviations, and correlation coefficients of the stochastic parameters involved.

2. Mean-absolute percentage error (e_{mabs}):

$$e_{mabs} = \frac{1}{M}\sum_{i=1}^{M} |e_i| \qquad (7)$$

3. Root-mean-squared percentage error (e_{rms}):

$$e_{rms} = \left(\frac{1}{M}\sum_{i=1}^{M} e_i^2\right)^{1/2} \qquad (8)$$

Antithetic-Variates Technique

Since the relevant statistics in verification are computed from the generated random variates, certain degrees of sampling error and variability exist. To reduce variability in the estimated statistics, the antithetic-variates

J. Hydraul. Eng. 1994.120:313-331.

TABLE 4. Verification of Proposed Simulation Procedure Based on 50 Samples along with 100 Simulation Runs for Correlation Coefficients for Marginal Statistical Properties and Distributions

Variable (1)	True value (2)	Error criteria (3)	Case 1 (4)	Case 2 (5)	Case 3 (6)	Case 4 (7)	Case 5 (8)	Case 6 (9)	Case 7 (10)
1 Mean	7.25	e_b (%)	0.12	0.15	0.47	−0.10	0.23	0.05	0.23
1 Mean	7.25	e_{mabs} (%)	3.21	3.28	2.82	3.28	3.00	2.78	2.72
1 Mean	7.25	e_{rms} (%)	3.93	4.21	3.51	4.05	3.86	3.48	3.53
1 STD	2.73	e_b (%)	−0.07	−0.48	−1.70	−0.92	−0.81	−0.80	−0.73
1 STD	2.73	e_{mabs} (%)	4.78	2.59	7.13	6.10	5.37	4.95	6.87
1 STD	2.73	e_{rms} (%)	5.84	3.27	8.79	7.56	6.53	6.06	8.60
1 K-S Pr	[Average (%)]		52.84	51.21	51.96	49.12	54.56	55.85	51.35
1 K-S Pr	[Number of rejections]		0	0	1	1	1	0	0
2 Mean	7.67	e_b (%)	0.23	−0.24	0.20	−0.07	0.38	0.03	0.25
2 Mean	7.67	e_{mabs} (%)	2.77	2.66	2.45	2.78	3.04	2.57	2.33
2 Mean	7.67	e_{rms} (%)	3.47	3.38	3.03	3.53	3.77	3.23	2.98
2 STD	2.60	e_b (%)	−0.52	−0.66	−2.55	−0.34	0.44	−1.79	−0.92
2 STD	2.60	e_{mabs} (%)	5.23	3.80	6.57	6.14	4.29	7.39	5.69
2 STD	2.60	e_{rms} (%)	6.35	4.79	8.24	7.67	5.40	9.30	7.30
2 K-S Pr	[Average (%)]		50.64	54.02	52.76	54.97	50.24	56.24	55.66
2 K-S Pr	[Number of rejections]		0	0	0	1	0	0	0
3 Mean	6.92	e_b (%)	−0.02	0.95	0.81	−0.19	−0.42	−0.13	0.51
3 Mean	6.92	e_{mabs} (%)	2.90	3.19	3.33	3.23	3.34	3.20	2.83
3 Mean	6.92	e_{rms} (%)	3.60	4.08	4.28	4.00	4.01	3.93	3.43
3 STD	2.82	e_b (%)	−0.20	0.01	0.50	−0.80	−3.09	−0.81	−0.16
3 STD	2.82	e_{mabs} (%)	4.53	2.46	8.57	6.01	7.46	6.17	7.93
3 STD	2.82	e_{rms} (%)	5.47	3.08	10.39	7.61	9.43	7.75	10.07
3 K-S Pr	[Average (%)]		54.56	51.25	50.77	51.71	52.19	51.76	54.71
3 K-S Pr	[Number of rejections]		0	1	1	0	0	0	0

TABLE 5. Verification of Proposed Simulation Procedure Based on 500 Samples along with 10 Simulation Runs for Correlation Coefficients

Variable pair (1)	True value (2)	Error criteria (3)	Case 1 (4)	Case 2 (5)	Case 3 (6)	Case 4 (7)	Case 5 (8)	Case 6 (9)	Case 7 (10)
Variables 1 and 2	0.45	e_b (%)	−0.54	−1.60	2.77	−1.34	−4.38	−0.04	1.75
Variables 1 and 2	0.45	e_{mabs} (%)	4.57	2.60	5.14	5.99	5.77	4.14	5.28
Variables 1 and 2	0.45	e_{rms} (%)	5.58	3.41	6.75	7.71	6.61	4.37	6.14
Variables 1 and 3	0.30	e_b (%)	2.69	−0.70	5.14	−1.41	−4.72	2.68	0.05
Variables 1 and 3	0.30	e_{mabs} (%)	6.47	6.08	8.44	10.73	8.43	5.45	6.31
Variables 1 and 3	0.30	e_{rms} (%)	7.91	8.24	11.14	13.00	10.08	6.66	7.58
Variables 2 and 3	0.42	e_b (%)	0.19	−1.60	0.47	−2.38	−2.32	−0.29	−0.11
Variables 2 and 3	0.42	e_{mabs} (%)	2.28	3.65	4.87	5.68	3.54	3.36	5.63
Variables 2 and 3	0.42	e_{rms} (%)	2.83	4.19	6.13	6.93	4.54	4.12	6.76

TABLE 6. Verification of Proposed Simulation Procedure Based on 500 Samples along with Simulation Runs for Marginal Statistical Properties and Distributions

Variable (1)	True value (2)	Error criteria (3)	Case 1 (4)	2 (5)	3 (6)	4 (7)	5 (8)	6 (9)	7 (10)
1 Mean	7.25	e_b (%)	0.41	−0.16	−0.35	0.17	0.57	0.17	0.45
1 Mean	7.25	e_{mabs} (%)	1.01	0.78	0.49	0.98	1.09	0.78	0.76
1 Mean	7.25	e_{rms} (%)	1.11	0.91	0.72	1.10	1.23	0.99	0.93
1 STD	2.73	e_b (%)	−0.06	−0.39	0.37	−0.47	−0.71	−0.20	1.40
1 STD	2.73	e_{mabs} (%)	1.06	0.84	1.26	2.74	1.10	1.59	2.82
1 STD	2.73	e_{rms} (%)	1.56	0.93	1.65	3.48	1.36	2.12	3.59
1 K-S Pr	[Average (%)]		43.03	44.93	56.62	42.92	43.51	41.63	54.53
1 K-S Pr	[Number of rejections]		0	0	0	0	0	1	0
2 Mean	7.67	e_b (%)	0.50	−0.19	0.17	−0.52	0.42	0.72	0.18
2 Mean	7.67	e_{mabs} (%)	0.81	0.65	0.70	1.11	0.79	0.86	0.83
2 Mean	7.67	e_{rms} (%)	1.11	0.76	0.88	1.36	0.85	1.09	0.95
2 STD	2.60	e_b (%)	−0.01	−0.44	1.48	1.18	−1.18	1.37	−0.61
2 STD	2.60	e_{mabs} (%)	1.38	1.23	1.99	1.64	1.29	2.55	2.24
2 STD	2.60	e_{rms} (%)	1.61	1.58	2.41	1.91	1.56	3.52	2.98
2 K-S Pr	[Average (%)]		49.95	58.47	59.74	45.27	51.93	62.46	56.02
2 K-S Pr	[Number of rejections]		0	0	0	0	0	0	0
3 Mean	6.92	e_b (%)	0.22	0.60	−0.09	−0.35	0.82	0.35	0.13
3 Mean	6.92	e_{mabs} (%)	0.86	1.11	0.52	1.16	1.19	1.00	0.78
3 Mean	6.92	e_{rms} (%)	1.14	1.50	0.65	1.44	1.61	1.47	0.91
3 STD	2.82	e_b (%)	0.32	−0.08	2.11	−1.51	1.68	−0.70	−0.39
3 STD	2.82	e_{mabs} (%)	1.95	0.47	3.13	2.33	4.34	1.89	1.24
3 STD	2.82	e_{rms} (%)	2.30	0.64	3.98	2.99	4.62	2.13	1.68
3 K-S Pr	[Average (%)]		47.21	49.84	55.74	50.27	47.15	42.77	67.58
3 K-S Pr	[Number of rejections]		0	0	0	0	0	0	0

TABLE 7. Verification of Proposed Simulation Procedure Based on 500 Samples with 10 Simulation Runs for Case of Strong and Negative Correlation for Correlation Coefficients

Variable pair (1)	True value (2)	Error criteria (3)	Case 1 (4)	2 (5)	3 (6)	4 (7)	5 (8)	6 (9)	7 (10)
Variables 1 and 2	0.80	e_b (%)	0.10	0.28	1.14	0.09	0.52	0.25	−0.48
Variables 1 and 2	0.80	e_{mabs} (%)	0.99	1.35	1.56	0.71	0.79	0.91	1.18
Variables 1 and 2	0.80	e_{rms} (%)	1.11	1.59	1.78	0.86	0.93	1.12	1.40
Variables 1 and 3	−0.30	e_b (%)	−0.60	−2.45	−0.54	−1.49	3.25	−3.06	2.68
Variables 1 and 3	−0.30	e_{mabs} (%)	8.24	7.15	5.97	5.59	4.66	7.19	7.53
Variables 1 and 3	−0.30	e_{rms} (%)	10.06	8.52	7.19	6.06	6.53	9.33	9.20
Variables 2 and 3	−0.42	e_b (%)	0.54	−0.31	−1.41	−1.66	1.20	−3.21	0.52
Variables 2 and 3	−0.42	e_{mabs} (%)	4.97	4.41	3.63	3.89	3.62	4.61	3.85
Variables 2 and 3	−0.42	e_{rms} (%)	5.88	5.06	4.84	5.30	4.11	5.87	5.06

TABLE 8. Verification of Proposed Simulation Procedure Based on 500 Samples with 10 Simulation Runs for Case of Strong and Negative Correlation for Marginal Statistical Properties and Distributions

Variable (1)	True value (2)	Error criteria (3)	Case 1 (4)	Case 2 (5)	Case 3 (6)	Case 4 (7)	Case 5 (8)	Case 6 (9)	Case 7 (10)
1 Mean	7.25	e_b (%)	−0.03	−0.56	0.14	−0.31	−0.27	0.69	−0.11
1 Mean	7.25	e_{mabs} (%)	0.78	1.14	0.68	1.24	0.66	1.01	1.00
1 Mean	7.25	e_{rms} (%)	0.88	1.35	0.82	1.57	0.84	1.10	1.13
1 STD	2.73	e_b (%)	−0.09	−0.19	0.26	−0.85	0.51	−0.10	−0.06
1 STD	2.73	e_{mabs} (%)	1.32	1.31	2.68	1.57	1.95	1.43	2.61
1 STD	2.73	e_{rms} (%)	1.65	1.45	3.36	2.11	2.27	1.74	3.46
1 K-S Pr	[Average (%)]		58.38	48.70	49.49	41.37	60.62	51.71	42.77
1 K-S Pr	[Number of rejections]		0	0	0	0	0	0	0
2 Mean	7.67	e_b (%)	−0.47	−0.27	0.45	0.01	0.09	0.73	0.03
2 Mean	7.67	e_{mabs} (%)	0.77	0.85	0.75	0.96	1.08	1.07	0.61
2 Mean	7.67	e_{rms} (%)	0.84	1.01	0.88	1.17	1.18	1.22	0.81
2 STD	2.60	e_b (%)	−0.22	0.81	1.11	0.56	−0.01	0.61	−0.59
2 STD	2.60	e_{mabs} (%)	1.78	1.04	2.13	1.29	1.07	2.88	2.48
2 STD	2.60	e_{rms} (%)	2.10	1.13	2.48	1.53	1.26	3.43	2.67
2 K-S Pr	[Average (%)]		47.81	59.23	51.26	46.00	48.21	49.51	62.69
2 K-S Pr	[Number of rejections]		0	0	0	0	0	0	0
3 Mean	6.92	e_b (%)	0.27	0.40	−0.08	−0.31	−0.14	−0.64	0.32
3 Mean	6.92	e_{mabs} (%)	0.88	1.01	0.57	1.06	1.36	1.31	1.18
3 Mean	6.92	e_{rms} (%)	0.95	1.16	0.74	1.21	1.50	1.61	1.56
3 STD	2.82	e_b (%)	0.16	−0.36	−0.94	−0.12	1.95	−0.41	0.67
3 STD	2.82	e_{mabs} (%)	2.04	0.81	1.98	1.42	2.80	3.02	1.84
3 STD	2.82	e_{rms} (%)	2.38	1.05	3.15	1.90	3.81	3.53	2.30
3 K-S Pr	[Average (%)]		58.12	48.36	59.49	41.37	45.84	44.07	45.92
3 K-S Pr	[Number of rejections]		0	0	0	0	0	1	0

technique (Ang and Tang 1984) is applied. By the antithetic-variate technique, the statistical properties of interest, θ, are computed by

$$\hat{\theta} = \frac{1}{2}(\hat{\theta}_1 + \hat{\theta}_2) \quad \text{(9)}$$

where $\hat{\theta}_1$ and $\hat{\theta}_2$ are the two unbiased estimators of θ. A simple antithetic-variates algorithm (Dagpunar 1988) to reduce the variance is to generate $\hat{\theta}_1$ and $\hat{\theta}_2$ based on random number sets U and 1-U, with U being the uniform random variable in (0, 1), which results a negative correlation between $\hat{\theta}_1$ and $\hat{\theta}_2$.

Kolmogorov-Smirnov Test

To examine whether the proposed Monte Carlo simulation procedure could generate random variates that preserve the known marginal distributions of the stochastic parameters involved, the generated random variates are examined by the Kolmogorov-Smirnov (KS) test (Press et al. 1989). The reason for choosing KS test is to avoid unnecessary binning of data that could cause loss of information. In the KS test, the null hypothesis is that the generated random variates have the same marginal distribution as the

J. Hydraul. Eng. 1994.120:313-331.

one used in simulation. The test statistic D is the maximum absolute difference between the distribution functions for the true stochastic variable under consideration and the generated random variates. Then, the significance level associated with the observed maximum absolute difference shows the likelihood that the random maximum absolute difference would exceed the observed value. A smaller probability means a stronger disparity between the null hypothesis and the data. A significance level of 0.05 is commonly used in the test below which the null hypothesis is rejected.

Results

To show the relative precision of the simulated results, the two extreme cases in Table 2 of 50 samples with 100 simulation runs and 500 samples with 10 simulation runs are presented in Tables 3, 4, 5, and 6, respectively. In all seven cases, the values of e_{mabs} and e_{rms} associated with the simulated statistical properties are large when the sample size is small and their values decrease monotonically as the sample size increases. However, the values of percentage bias, e_b, for all different sample sizes are practically zero. This shows that the proposed simulation procedure could generate statistics that have little or no systematic error. Based on this observation, the verification experiment adopts the larger sample size.

Table 5 shows the values of the different error criteria for the correlation coefficients computed from the proposed Monte Carlo simulation procedure based on 500 samples. Judging the values of e_{mabs} and e_{rms}, the proposed Monte Carlo simulation procedure is capable of generating random variates that closely preserve the original correlation structure of the stochastic parameters when the sample size is moderate or large. Furthermore, Table 6 lists the true and simulated statistical moments for each stochastic parameter. The small values of percentage errors associated with e_{mabs} and e_{rms} indicate that the simulated results well preserve the true moments.

To show the ability of the proposed simulation procedure to preserve the original marginal distributions, values of averaged significance probability of the KS test and the number of rejections during the 10 simulation runs are shown in Table 6. The only rejection occurred for variable 1 in case 6. The averaged significance probabilities are much greater than the commonly adopted rejection levels of 1–5%. This indicates that the generated samples by the proposed simulation algorithm follow the true distribution quite closely.

Condition adopted in Tables 5 and 6 for the experiment considers positive yet somewhat weak correlations among the random variables. To examine the performance of the proposed simulation procedure in a different condition, a stronger correlation coefficient 0.80 and two negative correlation coefficients, as listed in column 2 of Table 7, are used. The results shown in Tables 7 and 8, in general, do not reveal anything different from what have been observed in Tables 5 and 6.

According to the results obtained from the verification experiment, the proposed Monte Carlo simulation procedure exhibits its ability to preserve the marginal distribution and other relevant statistical properties of nonnormal correlated random variables. Therefore, the proposed procedure expands the applicability of the present multivariate simulation to accommodate correlated, nonnormal random variables.

APPLICATION

For demonstration, the proposed Monte Carlo simulation procedure was applied to the reliability analysis of bridge pier scour using a simple model developed by Johnson (1992). Different combinations of distributions of correlated stochastic parameters in the bridge pier scour model were used to examine the effect of distribution on the probability that the scour depth would exceed the designed pier depth.

Pier Scour Model

Johnson (1992) proposed an empirical pier scour model based on experimental data from various sources

$$D_s = 2.02\lambda y \left(\frac{b}{y}\right)^{0.98} \mathsf{F}^{0.21}\sigma^{-0.24} \quad \text{(10)}$$

in which D_s = the predicted scour depth; λ = the model correction factor; y = the flow depth; b = pier width; F = the Froude number; and σ = the sediment gradation. Because the model is empirical by nature, uncertainties exist in both model itself and the inputs/parameters involved (Yeh and Tung 1993). Consequently, the scour depth computed from (10) is subject to uncertainty and it is likely that a specified design pier depth could be exceeded resulting in potential threat to bridge safety.

Failure Probability Evaluation

The stochastic parameters considered in (10) are λ, y, F, and σ. The stochasticity of model correction factor, λ, represents the model uncertainty associated the pier scour model whereas the randomness of y, F, and σ results from model input uncertainties. Their means and coefficients of variation are listed in Table 9. According to Johnson (1992), all stochastic variables, except the model correction factor λ, are correlated random variables with the correlation matrix given in Table 10. The model correction factor λ is treated herein as an independent random variable. To examine the effect of distributions of stochastic parameters on the scour risk, three sets of distributional assumptions were used, which are: (1) All normal; (2)

TABLE 9. Means and Coefficients of Variation (CV) of Parameters used in the Pier Scour Model (from Johnson 1992)

Variables (1)	Mean (2)	CV (3)
λ	1.000	0.18
y	4.250	0.20
F	0.537	0.38
σ	4.000	0.20

TABLE 10. Correlation among Parameters used in the Pier Scour Model (from Johnson 1992)

Variables (1)	λ (2)	y (3)	F (4)	σ (4)
λ	1.00	0.00	0.00	0.00
y	0.00	1.00	−0.33	−0.79
F	0.00	−0.33	1.00	0.29
σ	0.00	−0.79	0.29	1.00

J. Hydraul. Eng. 1994.120:313-331.

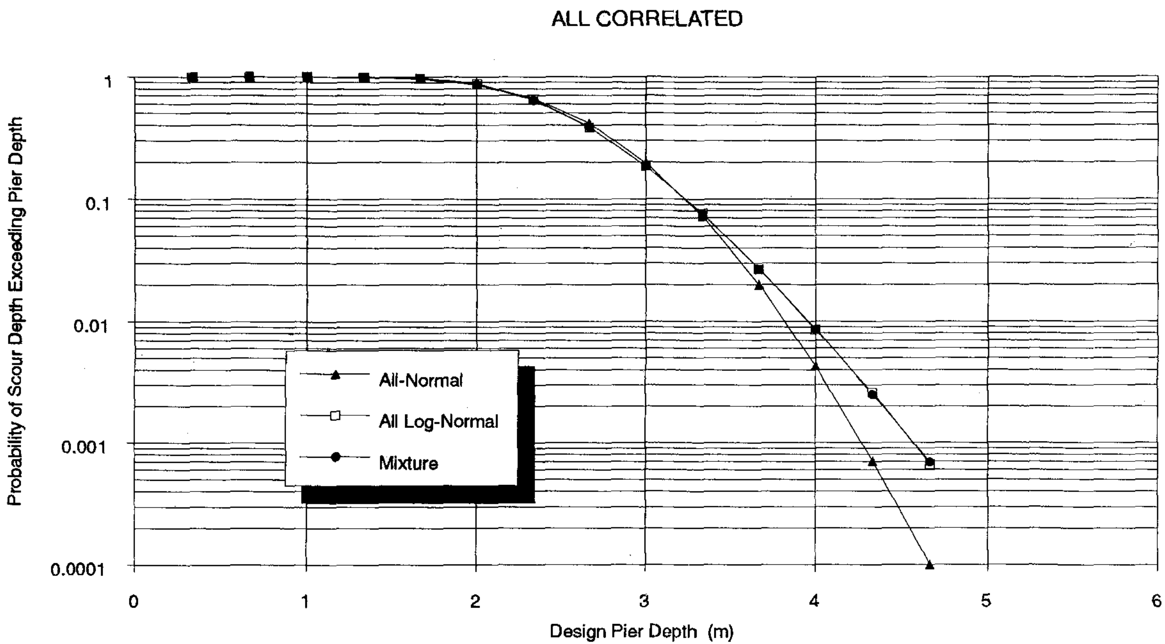

FIG. 3. Effect of Parameter Distribution on Failure Probability Curves

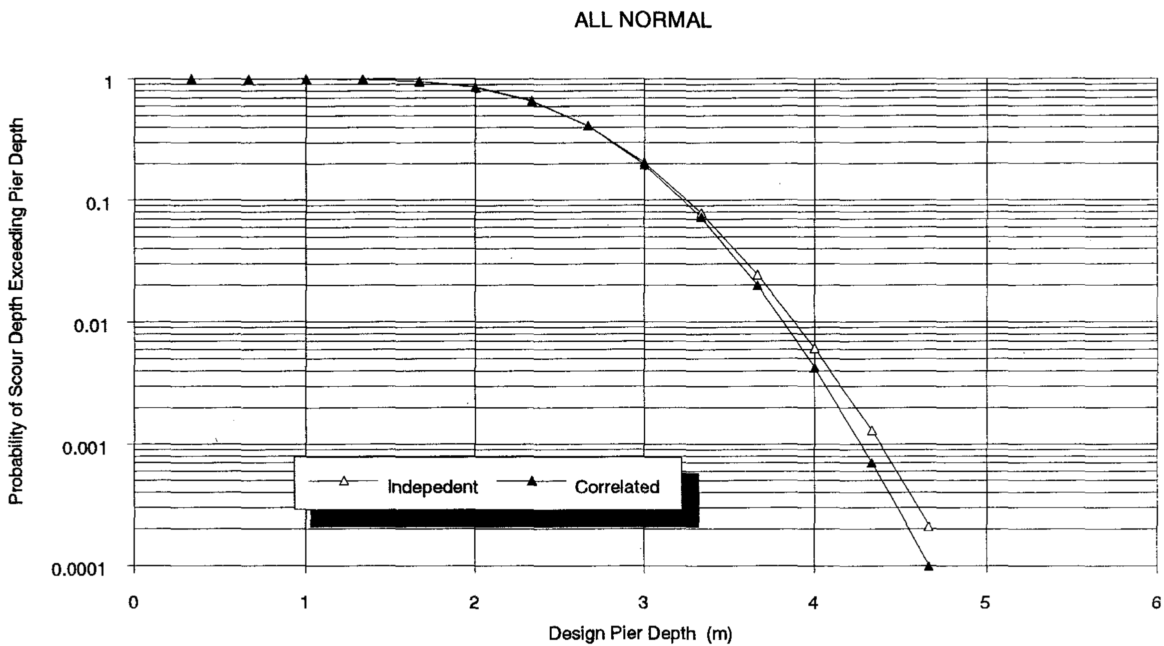

FIG. 4. Effect of Correlation on Failure Probability Curves Using All Normal Parameters

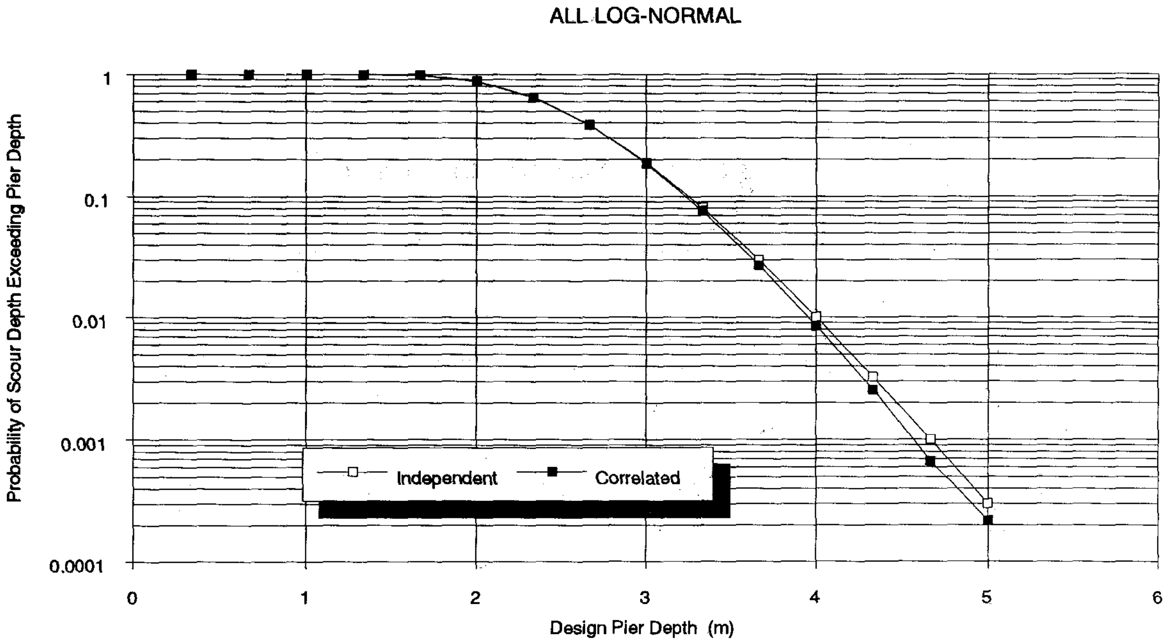

FIG. 5. Effect of Correlation of Failure Probability Curves Using All Log-Normal Parameters

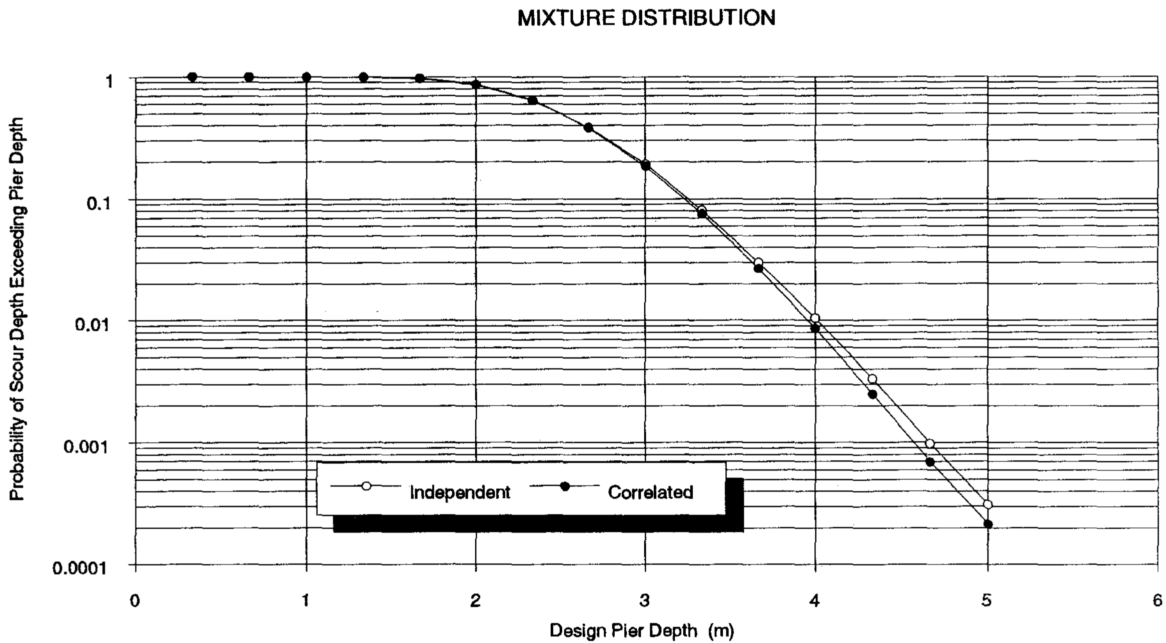

FIG. 6. Effect of Correlation on Failure Probability Curves Using Parameters with Mixture Distributions

all lognormal; and (3) mixture distributions that assign log-normal to λ and σ, gamma to y, and Weibull to F.

Based on the marginal distributions and correlations for the stochastic parameters, 100,000 samples were generated by the proposed Monte Carlo simulation procedure to calculate the possible realizations of scour depths. Therefore, the failure probability P_f for a certain designed pier depth, D_p, can be estimated as

$$P_f = Pr(D_s > D_p) = \frac{\text{number of } (D_s > D_p)}{100{,}000} \qquad (11)$$

Fig. 3 shows the effect of the probability distribution of stochastic parameters on the failure probability curve. The discrepancy between the failure probability curves enlarges as the design pier depth increases. This suggests that accurate assessments of the marginal distributions of the parameters involved is crucial for an accurate determination of failure probability when it is small. The cases with all log-normal distribution and mixture of distributions have longer tails than the case with all normal distribution. Using a multivariate normal distribution results in underestimating the potential risk. In many reliability analyses, especially for highly safe structures where failure probability is small, accuracy in the tail probability estimation is important (Tung and Mays 1980).

The effect of correlation among the stochastic parameters on the failure probability is shown in Figs. 4–6. As can be observed, the discrepancy among the failure probability curves grows as the design pier depth increases. It is interesting to note that without considering the correlation among stochastic parameters, the resulting failure probability is higher than that considering correlation. Therefore, in this example, considering stochastic parameters as independent random variables yields a conservative estimation of failure probability.

Although the failure probabilities shown in Figs. 4–6 for the pier scour example vary less than order of magnitude, this does not imply that incorporation of correlation information of stochastic variables is not essential in reliability analysis. In cases other than what is being considered, particularly where there are high negative correlations, the accounting for correlated variables may change the resulting failure probability by orders of magnitude (Thoft-Christensen and Baker 1982).

SUMMARY AND CONCLUSIONS

Monte Carlo simulation procedures are frequently applied in probabilistic analysis of engineering problems. The approach provides design engineers with essential information on system response under the stochastic environment and with valuable insight about the system behavior.

In spite of being an effective tool, multivariate Monte Carlo simulation is much restricted by the dimension and, perhaps, more by the distribution type of correlated stochastic parameters. This study proposes a procedure to generate multivariate, nonnormal, correlated random variates based on the specified marginal distributions and correlation coefficients. The procedure is based on the empirical equations derived by Liu and Der Kiureghian (1986) which, according to the marginal distribution types of the stochastic parameters, transform the correlation coefficients from the original parameter space to those of standard normal space. In doing so, many efficient algorithms for generating multivariate normal random variates can be applied. Therefore, the proposed procedure extends the applicability of multivariate Monte Carlo simulation.

To ensure the proposed procedure would preserve the statistical characteristics of the original stochastic parameters, the generated correlation matrix and distribution are examined by a numerical verification. The results from the numerical verification indicate that the proposed procedure for multivariate Monte Carlo simulation can preserve the marginal distributions and the corresponding correlation structure very satisfactorily.

For illustration, the proposed procedure is applied to assess the failure probability in a bridge pier scour problem. The results indicate that accurate estimation of reliability, especially in the tail part of the distribution, should account for relevant stochastic information including correlation and marginal distributions.

In practical engineering problems, probability distributions associated with the stochastic parameters are generally subject to uncertainty. This distributional model uncertainty could potentially have a significant effect on the results of reliability analysis, especially on the tail part of the distribution. When the distribution types of stochastic parameters are uncertain, distributions that are potential candidates should be applied. In the past, multivariate Monte Carlo simulation can only be applied to a few distributional cases. The proposed procedure allows examination of the effect of probability model uncertainty due to its ability to handle various nonnormal, correlated random variables.

ACKNOWLEDGMENTS

The writers wish to express their gratitude to Water Resources Planning Commission, MOEA, Republic of China for the financial support of this study. Comments from the reviewers are appreciated.

APPENDIX. 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 and Sons, Inc., New York, N.Y.

Box, G. E. P., and Muller, M. E. (1958). "A note on generation of random normal deviates." *Ann. Math. Stat.*, 29, 610–611.

Dagpunar, J. (1988). *Principles of random variates generation*. Oxford University Press, New York, N.Y.

Der Kiureghian, A., and Liu, P. L. (1985). "Structural reliability under incomplete probability information." *J. Engrg. Mech.*, ASCE, 112(1), 85–104.

Johnson, P. A. (1992). "Reliability-based pier scour engineering." *J. Hydr. Engrg.*, ASCE, 118(10), 1344–1358.

Li, S. T., and Hammond, J. L. (1975). "Generation of pseudo-random numbers with specified univariate distributions and covariance matrix." *IEEE Trans. on Systems, Man. and Cybernetics*, Sep., 557–561.

Liu, P. L., and Der Kiureghian, A. (1986). "Multivariate distribution models with prescribed marginals and covariances." *Probabilistic Engrg. Mech.*, 1(2), 105–112.

Parrish, R. S. (1990). "Generating random deviates from multivariate Pearson distributions." *Computational Statistics and Data Analysis*, 9, 283–296.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1989). *Numerical recipes*. Cambridge University Press, New York, N.Y.

Ronning, G. (1977). "A simple scheme for generating multivariate gamma distributions with nonnegative covariance matrix." *Technometrics*, 19(2), 179–183.

Thoft-Christensen, P., and Baker, M. J. (1982). *Structural reliability theory and its applications*. Springer-Verlag, New York, N.Y.
Tung, Y. K., and Mays, L. W. (1980). "Risk analysis for hydraulic design." *J. Hydraul. Div.*, ASCE, 106(5), 893–913.
Yeh, K. C., and Tung, Y. K. (1993). "Uncertainty and sensitivity of a pit migration model." *J. Hydr. Engrg.*, ASCE, 119(2), 262–281.