

**UNCERTAINTY AND SENSITIVITY
ANALYSES OF PIT-MIGRATION MODEL**

Keh-Chia Yeh

Yeou-Koung Tung

Journal Article

1993

WWRC-93-05

In

Journal of Hydraulic Engineering

**Keh-Chia Yeh
Department of Civil Engineering
National Chiao Tung University
Hsinchu, Taiwan**

**Yeou-Koung Tung
Wyoming Water Resources Center
University of Wyoming
Laramie, Wyoming**

UNCERTAINTY AND SENSITIVITY ANALYSES OF PIT-MIGRATION MODEL

By Keh-Chia Yeh¹ and Yeou-Koung Tung,² Associate Member, ASCE

ABSTRACT: Sand and gravel mining from a river bed results in irregular pits on the river bed. The migration of the pits might potentially threaten the safety of downstream bridge piers and other in-stream hydraulic structures. Models were developed to simulate the movement of pits for predicting the effect of migrating pits on in-stream structures. In view of random characteristics inherently residing in hydraulic and hydrologic processes, it is essential for an engineer to assess the overall uncertainty features of a hydraulic model output subjected to its stochastic input parameters. As an illustration, this paper analyzes the uncertainties of a pit migration model recently proposed by Lee et al. (1990) using three methods including the first-order variance estimation method, point estimation technique, and Latin hypercube sampling. Comparisons of merits and limitations of these methods are also made.

INTRODUCTION

Sand and gravel mined from river beds are major sources of construction materials in Taiwan. Mining activities result in irregular pits on the river bed. As observed by Lee (H. Y. Lee, personal communication, 1991) in his laboratory experiments, the migrating pit disappears or diffuses rapidly in a very short distance under high flow conditions. On the other hand, the pit might travel a long distance downstream and maintain its distinctiveness during low- and medium-flow conditions. Therefore, the migration of pits during low- and medium-flow seasons might impose potential threats to the safety of downstream bridge piers and other in-stream hydraulic structures.

Pit migration and morphology are the result of a very complicated non-linear interaction among the flow, topography of erodible bed and banks, and sediment transport. To better understand the pit-migration process, Lee et al. (1990) conducted a series of experiments from which a set of empirical formulas were developed for predicting the progressive change of the geometry of a pit as it migrates downstream. Models such as this can be applied to evaluate the safety of in-stream structures in relation to migrating pits.

In hydraulic/hydrologic modelings, analyses, and designs, several factors contribute to different uncertainties (Yen et al. 1986), such as:

1. Uncertainties associated with the inherent randomness of natural processes
2. Model uncertainty reflecting the inability of the simulation model or design technique to represent the system's true physical behavior

¹Assoc. Prof., Dept. of Civ. Engrg., Nat. Chiao Tung Univ., Hsinchu, Taiwan, 30050, R.O.C.

²Assoc. Prof., Wyoming Water Res. Ctr. and Stat. Dept., Univ. of Wyoming, Laramie, WY 82071.

Note. Discussion open until July 1, 1993. 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 October 21, 1991. This paper is part of the *Journal of Hydraulic Engineering*, Vol. 119, No. 2, February, 1993. ©ASCE, ISSN 0733-9429/93/0002-0262/\$1.00 + \$.15 per page. Paper No. 2841.

3. Model-parameter uncertainties resulting from inability to quantify accurately the model input parameters

4. Data uncertainties including (a) measurement errors; (b) inconsistency and nonhomogeneity of data; and (c) data handling and transcription errors

5. Operational uncertainties including those associated with construction, manufacture, deterioration and maintenance, and other human factors that are not accounted for in the modeling or design procedures.

All of these uncertainties may contribute to the stochasticity of model input parameters, which, in turn, result in model output uncertainty. The purpose of uncertainty analysis is to determine how the stochastic input parameters affect model outputs. The analysis provides the modeler with insight about the contribution of each stochastic input parameter to the overall uncertainty of the model output. Such information is useful for identifying the important input parameters to which more attention should be given if the overall uncertainty of model output is to be reduced.

Another important aspect of model evaluation is the sensitivity analysis. Sensitivity analysis is concerned with how inputs influence the output and output variability. A local perspective for sensitivity analysis is concerned with output variability in the neighborhood of a point—often the nominal value—in the input space. Sensitivity analysis from the global perspective, on the other hand, is concerned with the variability of output over the entire input space. Although local sensitivity measures can provide, in principle, a more detailed description of the importance of input parameters than the global measures, the use of local measures, in practice, is often limited by the computational effort required to evaluate them, especially when the number of variables is large. Global sensitivity measures require less computational effort and also can be used to rank the relative importance of inputs. However, lack of resolution can limit its usefulness, especially when the effect of an input on the model output is drastically different in different parts of the parameter space.

In view of safety implications of migrating pits on in-stream hydraulic structures, the main objective of this study is to assess the uncertainty, in terms of statistical characteristics, of a hydraulic model in describing migration of a pit. For purposes of demonstration a simple pit-migration model developed by Lee et al. (1990) was adopted. Parameter uncertainties in the model are the main concerns. Methods for performing uncertainty analysis vary in sophistication. In principal, it would be ideal to derive the exact probability density function (PDF) of the model output as a function of the PDFs of the stochastic input variables. However, most of the models used in hydraulic and hydrologic analyses are highly complex. This usually prohibits attempts to analytically derive the PDF of the model output. Alternative methods therefore are useful in estimating the statistical properties of the model output. The methods considered herein are the first-order variance estimation (FOVE) method, the point estimate (PE) method, and Latin hypercube sampling (LHS). The relative performance of the three techniques in analyzing the uncertainty of the pit migration model is compared.

PIT-MIGRATION MODEL

The experiments conducted by Lee et al. (1990) were limited to pits of rectangular shape and uniform sand material. Additional conditions under

which the experiments were performed were: (1) The Froude number was less than 0.85 to guarantee a subcritical flow; (2) the particle size of the bed material, d_s , was larger than 0.6 mm to avoid the occurrence of ripples; and (3) the shear velocity of the flow was smaller than $3.4\sqrt{d_s}$ to prevent the occurrence of sediment in suspension. Based on the experimental data a set of empirical equations were developed using multiple regression analysis. Although equations were developed for predicting the geometry of a rectangular pit as it migrates downstream, this study is only concerned with the maximum depth after a pit travels a specified distance downstream. The model for predicting the maximum pit depth consists of five equations whose functional relations were obtained through a trial-and-error procedure. To facilitate discussion of the uncertainty analysis to be presented later the pit migration model given below retains regression coefficients and model error terms in the form of notations. The values of regression coefficients estimated by the least-squares method are given in column 2 of Table 1.

The model output of interest, i.e., the maximum pit depth H_d , is estimated by

TABLE 1. Statistical Properties of Stochastic Input Parameters in Pit-Migration Model

Input variable (1)	Mean (2)	Standard deviation (3)	Coefficient of variation (4)	Distribution (5)
n	0.0250	0.002041	0.0816	triangular
S (m/m)	0.0001	0.00001732	0.1732	uniform
γ_s	2.650	0.05774	0.0218	uniform
d_s (m)	0.001	0.0001155	0.1155	uniform
a_1	-0.3480	0.02547	0.0732	normal
a_2	0.4102	0.05781	0.1409	normal
a_3	-0.00352	0.00105	0.2983	normal
a_4	0.4354	0.02913	0.0669	normal
a_5	-0.4464	0.053	0.1187	normal
a_6	0.00013	0.000023	0.1769	normal
b_0	0.5692	0.0924	0.1623	normal
b_1	-1.187	0.1411	0.1189	normal
b_2	0.1715	0.0818	0.4770	normal
c_0	2.166	0.436	0.2013	normal
c_1	0.7795	0.3334	0.4277	normal
c_2	-0.2865	0.2216	0.7735	normal
c_3	0.6008	0.1544	0.2570	normal
d_0	-0.3797	0.0473	0.1246	normal
d_1	-0.2028	0.0208	0.0540	normal
d_2	-0.2975	0.047	0.1580	normal
r_1	-0.00653	0.000946	0.1449	normal
r_2	0.04611	0.002624	0.0569	normal
r_3	0.00289	0.00045	0.1557	normal
e_1	0.000	0.001112	—	normal
e_2	0.000	0.04131	—	normal
e_3	0.000	0.09423	—	normal
e_4	0.000	0.09442	—	normal
e_5	0.000	0.000454	—	normal

$$\ln \left(\frac{H_d}{H_{\max}} \right) = d_0 + d_1 \left(\frac{t - T_f}{T_f} \right) + d_2 \left(\frac{u_*}{u_{*c}} \right) + e_4 \dots \dots \dots (1)$$

where H_{\max} = the maximum pit depth at the end of a convection period by which time the upstream boundary of the pit reaches the original downstream pit boundary; T_f = the time span of the convection period; t = total travel time of the pit including the convection period and the diffusion period during which the pit loses its identity downstream; u_* and u_{*c} = the shear velocity and critical shear velocity, respectively; d_0 , d_1 , and d_2 = regression coefficients; and e_4 = the model error term. Based on the experimental data, (1) has a coefficient of determination $R^2 = 0.752$. The associated standard error $s_4 = 0.09442$.

The maximum pit depth at the end of the convection period, H_{\max} , in (1) is estimated by

$$\ln \left(\frac{H_{\max}}{H} \right) = b_0 + b_1 \ln \left(\frac{u_*}{u_{*c}} \right) + b_2 \ln \left(\frac{H}{y} \right) + e_2 \dots \dots \dots (2)$$

where H = the initial pit depth; y = the flow depth; b_0 , b_1 , and b_2 = regression coefficients; and e_2 = the error term associated with estimating $\ln(H_{\max}/H)$. Eq. (2) has an $R^2 = 0.937$ and standard error $s_2 = 0.04131$.

The total time [i.e., t in (1)] required for a pit to travel a specified distance D is

$$t = \frac{D}{u_{bd}} + \frac{L}{u_{bc}} \dots \dots \dots (3)$$

where L = the initial pit length; and u_{bd} and u_{bc} = the moving velocities of the pit during the diffusion period and convection period, respectively. The moving velocity u_{bd} is estimated as

$$\frac{u_{bd}}{u_{*c}} = a_1 \left(\frac{u_*}{u_{*c}} - 1 \right) + a_2 \left(\frac{H_{\max}}{H} \right) + a_3 \left(\frac{L_e}{H} \right) + a_4 \left(\frac{u_*}{u_{*c}} - 1 \right)^2 + a_5 \left(\frac{H_{\max}}{H} \right)^2 + a_6 \left(\frac{L_e}{H} \right)^2 + e_1 \dots \dots \dots (4)$$

where a_1 , a_2 , a_3 , a_4 , a_5 , and a_6 = regression coefficients; L_e = the effective pit length at the end of the convection period; and e_1 = the model error term. Eq. (4) has an $R^2 = 0.9989$ with a standard error of estimate $s_1 = 0.001112$. The term L_e/H in (4) is estimated by

$$\ln \left(\frac{L_e}{H} \right) = c_0 + c_1 \ln \left(\frac{u_*}{u_{*c}} \right) + c_2 \ln \left(\frac{H}{y} \right) + c_3 \ln \left(\frac{L}{y} \right) + e_3 \dots \dots \dots (5)$$

where c_0 , c_1 , c_2 , and c_3 = regression coefficients; and e_3 = the model error term. The R^2 and standard error of estimate s_3 for (5) are 0.806 and 0.09423, respectively.

The pit moving velocity during the convection period u_{bc} is estimated by

$$\frac{u_{bc}}{u_{*c}} = r_1 \ln\left(\frac{L}{y} + 1\right) + r_2 \left[\ln\left(\frac{u_*}{u_{*c}}\right)\right]^2 + r_3 \left[\ln\left(\frac{L}{y} + 1\right)\right]^2 + e_5 \quad (6)$$

where r_1 , r_2 , and r_3 = regression coefficients; and e_5 = the model error term. Eq. (6) has an $R^2 = 0.9954$ and standard error of estimate $s_5 = 0.0004542$.

To estimate the maximum pit depth H_d after a certain distance of travel, the use of the five aforementioned regression equations requires knowledge of the shear velocity u_* and the critical shear velocity u_{*c} . The shear velocity u_* , under the wide-channel condition, can be computed by

$$u_* = \sqrt{g_y S_f} \quad (7)$$

where g_y = the gravitational acceleration constant; S_f = the friction slope; and y = the flow depth, which can be computed by Manning's formula

$$y = \left(\frac{Qn}{BS^{1/2}}\right)^{3/5} \quad (8)$$

where n = the Manning's roughness; Q = the flow rate; and B = the channel width. The critical shear velocity u_{*c} can be determined by the Shields' diagram, which is a function of the specific weight of sand material (γ_s), representative particle size (d_s), and other flow characteristics.

UNCERTAINTIES IN PIT-MIGRATION MODEL

Using the pit-migration model [(1)–(6)] to estimate the maximum pit depth (H_d) as the pit travels downstream requires specifying the pit dimensions (H and L), hydraulic conditions (Q , B , n , S_f), bed material characteristics (γ_s , d_s), and regression coefficients (a , b , c , d , r) in the models. These can be regarded as the model input parameters that affect the estimated maximum pit depth; all are subject to uncertainty.

Hydraulic parameters such as Manning's roughness n and friction slope S_f cannot be assessed with absolute certainty. Bed material characteristics, such as specific weight γ_s and representative grain size d_s , generally vary spatially. Also, estimated values for regression coefficients given in column 2 of Table 1 cannot be treated as the true ones because they are estimated from a limited amount of experimental data. Sample errors exist in the estimated regression coefficients. Furthermore, the pit-migration model developed on the basis of experimental data can only be considered as an approximation of the underlying physical process. Note that the model errors associated with the regression equations (i.e., e_1 , e_2 , e_3 , e_4 , e_5) account only in part of the total model errors with respect to the real-world pit-migration process because some factors such as nonuniform sediment size and irregular-shaped pits are not considered. In the pit-migration model considered herein, a total of 28 input parameters listed in column 1 of Table 1 are considered random and subject to uncertainty.

The uncertainties associated with input parameters can be assessed subjectively based on personal experience and judgment, or they can be quantified statistically on the basis of measurements and proper statistical theories. For example, Manning's roughness n is a conceptual parameter that is not physically measurable. The values of Manning's roughness used in most hydraulic computations are determined on the basis of personal judgment

in comparing field channel conditions and hydraulic reference books. A good example illustrating the uncertainty of Manning's roughness coefficient are nominal values and ranges of Manning's roughness under a variety of channel conditions listed by Chow (1959). Further, probability distributions for those conceptual parameters can only be assumed.

In contrast, some parameters in a hydraulic model can be physically measured. In such cases, statistical tools can be applied to analyze the measured data and to quantify the associated uncertainty represented by the mean, standard deviation, or even the probability distribution. Examples of such parameters are representative sediment sizes, specific weights, and hydraulic geometry. Other types of model parameters are empirical coefficients whose statistical properties can be inferred via appropriate statistical theories. An example of this is the assessment of parameter and model uncertainties associated with the regression model, such as the pit-migration model presented previously. Based on the experimental data from Lee et al. (1990), the mean, standard deviation and coefficient of variation of the regression coefficients in the pit-migration model [(1)–(6)] are listed in columns 2–4 of Table 1, respectively. The standard deviations associated with the error terms of the regression equations are the standard errors that represent partial model uncertainties. Although the actual model uncertainty is larger than those indicated by regression standard errors, the assessment of such total model error is difficult because of the absence of a true model of the process. Without having evidence to adjust for its value, this study treats regression standard errors as model errors. The correlation matrices between regression coefficients within individual regression equations are shown in Tables 2–6. Correlation relationships are important for describing the linear relationship among stochastic model input parameters that directly affect the degree of uncertainty in the model output.

METHODS OF UNCERTAINTY ANALYSIS

First-Order Variance Estimation (FOVE) Method

This method estimates uncertainty of the model output as a function of the variances of stochastic input parameters. It uses Taylor's series expansion to estimate the local uncertainty of the model output at a selected expansion point. Consider that the output Y of a hydraulic or hydrologic model can be expressed as a function of stochastic input parameters X_s as

TABLE 2. Correlation Matrix of Regression Coefficients in Eq. (1)

Input variable (1)	d_0 (2)	d_1 (3)	d_2 (4)
d_0	1.0000	-0.2750	-0.9912
d_1	-0.2750	1.0000	0.1695
d_2	-0.9912	0.1695	1.0000

TABLE 3. Correlation Matrix of Regression Coefficients in Eq. (2)

Input variable (1)	b_0 (2)	b_1 (3)	b_2 (4)
b_0	1.0000	-0.9059	0.9291
b_1	-0.9059	1.0000	-0.7199
b_2	0.9291	-0.7199	1.0000

TABLE 4. Correlation Matrix of Regression Coefficients in Eq. (4)

Input variable (1)	a_1 (2)	a_2 (3)	a_3 (4)	a_4 (5)	a_5 (6)	a_6 (7)
a_1	1.0000	-0.6732	0.4441	-0.9657	0.6718	-0.4629
a_2	-0.6732	1.0000	-0.9576	0.5219	-0.9953	0.9634
a_3	0.4441	-0.9576	1.0000	-0.2877	0.9453	-0.9916
a_4	-0.9657	0.5219	-0.2877	1.0000	-0.5118	0.2889
a_5	0.6718	-0.9953	0.9453	-0.5118	1.0000	-0.9623
a_6	-0.4629	0.9634	-0.9916	0.2889	-0.9623	1.0000

TABLE 5. Correlation Matrix of Regression Coefficients in Eq. (5)

Input variable (1)	c_0 (2)	c_1 (3)	c_2 (4)	c_3 (5)
c_0	1.0000	-0.6503	0.8503	-0.8752
c_1	-0.6503	1.0000	-0.7255	0.2595
c_2	0.8503	-0.7255	1.0000	-0.5391
c_3	-0.8752	0.2595	-0.5391	1.0000

TABLE 6. Correlation Matrix of Regression Coefficients in Eq. (6)

Input variable (1)	r_1 (2)	r_2 (3)	r_3 (4)
r_1	1.0000	-0.3224	-0.9790
r_2	-0.3224	1.0000	0.1405
r_3	-0.9790	0.1405	1.0000

$$Y = g(\mathbf{X}') = g(X_1, X_2, \dots, X_n) \dots \dots \dots (9)$$

Where \mathbf{X} = an n -dimensional column vector of stochastic input parameters; the superscript t = the matrix or vector transpose; and $g(\)$ represents a functional relationship for the model. In the context of the present study, $g(\mathbf{X}')$ is the pit-migration model consisting of (1)-(8), the model output Y is the maximum pit depth H_d , and the stochastic input vector \mathbf{X} consists of elements indicated in column 1 of Table 1. The FOVE method considers the first-order Taylor series expansion term of (9):

$$Y \cong g(\mathbf{x}'_0) + \sum_{i=1}^n s_{i0}(X_i - x_{i0}) = g(\mathbf{x}'_0) + \mathbf{s}'_0(\mathbf{X} - \mathbf{x}_0) \dots \dots \dots (10)$$

where \mathbf{s}_0 = n -dimensional column vector of sensitivity coefficients with elements $s_{i0} = (\partial g / \partial X_i)_{\mathbf{x}_0}$ being the sensitivity coefficient of the model output Y with respect to the i th input parameter X_i at the expansion point \mathbf{x}_0 . Applying the expectation and variance operators to (10) with $\mathbf{x}_0 = \boldsymbol{\mu}$, the mean and variance of the model output Y can be estimated as

$$E(Y) = \mu_Y \cong g(\boldsymbol{\mu}') \dots \dots \dots (11)$$

and

$$\text{var}(Y) = \sigma_Y^2 \cong \mathbf{s}'\boldsymbol{\Omega}\mathbf{s} \dots \dots \dots (12)$$

where μ_Y and σ_Y = the mean and standard deviation of the model output, respectively; $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ = the vector of means and the covariance matrix of stochastic input parameters, respectively; and \mathbf{s} = the sensitivity coefficient vector evaluated at $\mathbf{x}_0 = \boldsymbol{\mu}$. If all stochastic input parameters are independent, the variance of the model output Y reduces to

$$\text{var}(Y) = \sigma_Y^2 \cong \sum_{i=1}^n s_i^2 \sigma_i^2 \dots \dots \dots (13)$$

As can be seen from (12) and (13), the uncertainty of the model output, $\text{var}(Y)$, depends not only on the uncertainty of individual stochastic input parameters as measured by σ_i^2 , but also on the associated sensitivity coefficients s_i . According to (13), the contribution of each stochastic input parameter, C_i to the overall uncertainty of the model output can be computed as

$$C_i = \frac{s_i^2 \sigma_i^2}{\sigma_Y^2} \quad i = 1, 2, \dots, n \dots \dots \dots (14)$$

The relative importance of stochastic input parameters, in terms of their contribution to the overall uncertainty of the model output, then can be assessed through examining the relative magnitude of C_i . In the case that some of the stochastic input parameters are correlated, the positive or negative contribution of such correlation to the overall uncertainty of the model output also can be evaluated.

Point Estimation (PE) Methods

The 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 original PDF of a random variable by discrete probability masses concentrated at two points in such a way that the first three moments of the original PDF are preserved.

Consider the model represented by (9) having n stochastic input parameters. By Rosenblueth's PE method, 2^n model evaluations (runs) are required to estimate the statistical moments of the model output. For large computerized hydraulic models involving many stochastic input parameters, Rosenblueth's PE method is computationally impractical. For the present analysis, the required model run by Rosenblueth's method would be $2^{28} = 268,435,456$. To avoid this difficulty, Harr (1989) proposed a modification that reduces the required model runs from 2^n to $2n$, and greatly enhances practical applicability of the method.

By Harr's modification, the correlation matrix \mathbf{C} of n stochastic input parameters, which is real and symmetric, is decomposed

$$\mathbf{C} = \mathbf{V}\mathbf{L}\mathbf{V}' \dots \dots \dots (15)$$

where \mathbf{V} = eigenvector matrix = $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$, in which $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ = column vectors of eigenvectors; and $\mathbf{L} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, a diagonal matrix with $\lambda_1, \lambda_2, \dots, \lambda_n$ being the corresponding eigenvalues. The correlation matrix can also be represented geometrically by a hypersphere of radius \sqrt{n} centered at the expected values of stochastic input parameters $\mu_1, \mu_2, \dots, \mu_n$ in the standardized coordinate system. Each

eigenvector, passing through the origin of the hypersphere, intersects the sphere surface at two points. Then, these $2n$ intersection points (for n input variables) are used to estimate the statistical moments of model output. Harr's (1989) modified PE method can be summarized by the following:

- Step 1. Decompose the correlation matrix C of the input variables into an eigenvector matrix V and corresponding diagonal eigenvalue matrix L , as (15)
- Step 2. Generate coordinates of the $2n$ intersecting points using

$$X_{i\pm} = \mu \pm \sqrt{n} \begin{bmatrix} \sigma_1 & \dots & \dots \\ \dots & 0 & \dots \\ \dots & \dots & \dots \\ 0 & \dots & \dots \\ \dots & \dots & \sigma_n \end{bmatrix} v_i \quad i = 1, 2, \dots, n \quad \dots \dots \dots (16)$$

- Step 3. Compute $Y_{i\pm} = g(X_{i\pm})$, and $Y_{i\pm}^2 = g^2(X_{i\pm})$ for $i = 1, 2, \dots, n$.
- Step 4. Compute the averaged model outputs for $i = 1, 2, \dots, n$, i.e.

$$\bar{Y}_i = \frac{Y_{i+} + Y_{i-}}{2} \text{ and } \bar{Y}_i^2 = \frac{Y_{i+}^2 + Y_{i-}^2}{2}$$

- Step 5. Compute the mean and variance of the model output Y :

$$E(Y) = \frac{\sum_{i=1}^n \bar{Y}_i \lambda_i}{\sum_{i=1}^n \lambda_i} = \frac{\sum_{i=1}^n \bar{Y}_i \lambda_i}{n} \quad \dots \dots \dots (17)$$

$$E(Y^2) = \frac{\sum_{i=1}^n \bar{Y}_i^2 \lambda_i}{n} \quad \dots \dots \dots (18)$$

$$\text{var}(Y) = E(Y^2) - E^2(Y) \quad \dots \dots \dots (19)$$

Latin Hypercubic Sampling (LHS)

The essence of Latin hypercubic sampling is to select, in a stratified manner, random samples for each stochastic input parameter over its range such that the overall uncertainty of the model output can be reasonably described by finite samples. Consider K sets of n random input parameters to be generated from which the corresponding model outputs are computed. By LHS, the plausible range of each of the stochastic input parameters is divided into K equal probability intervals from which a random sample for the input parameter is taken. More specifically, consider a stochastic input parameter X_i over the interval (l_i, u_i) following a specified PDF, $f_i(x_i)$. The range (l_i, u_i) is partitioned into K intervals,

$$l_i = a_{i0} < a_{i1} < \dots < a_{i,K-1} < a_{iK} = u_i \quad \dots \dots \dots (20)$$

where $P(a_{i,k} \leq X_i < a_{i,k+1}) = 1/K$ for all $k = 0, 1, 2, \dots, K - 1$. LHS involves randomly selecting a value in each of the intervals to form the K samples for the i th stochastic input parameter X_i . A random permutation

can be applied to randomize the sequence to make the generated sequence of K samples random. The process can be repeated for all stochastic input parameters resulting in K input data sets.

Using LHS, the usual estimators of the mean and distribution function of the model output are unbiased (McKay 1988). Moreover, when the mode $g(X)$ is monotonic in each of the X_i , the variance of the estimators are no more than, and often much less than, the variances based on inputs generated by a simple random sampling procedure, such as by Monte Carlo simulation. This variance reduction property of LHS implies that fewer samples or computer runs may be necessary to obtain a degree of precision comparable to that obtained from a simple random sampling of input parameters.

The aforementioned description of LHS assumes that the n model input parameters are uncorrelated. In the case that some of the stochastic input parameters are correlated, the joint PDF of the inputs are required. Comparing with the variety of the random number generation for univariate variable, multivariate random number generation is much more restricted to a few joint distributions such as multivariate normal, multivariate log-normal, and multivariate gamma (Johnson 1987). For more detailed discussions, readers are referred to McKay (1988).

APPLICATION

The conditions under which the analysis of the pit migration model are considered are: flow rate $Q = 200$ cms, channel width $B = 100$ m, initial pit length $L = 40$ m, initial pit depth $H = 3$ m, and traveling distance $D = 500$ m. The example adopted herein meets the required conditions under which Lee et al.'s (1990) experiments were conducted.

The distribution types and corresponding statistical properties of the 28 stochastic input parameters used in the three methods are listed in Table 1. Those distribution types are judged from the data, if available, or subjectively assumed. Normal distributions used for regression coefficients and model errors are based on the validity of the normality condition through the regression analysis. The correlation coefficient matrices of regression coefficients in the pit migration model are listed in Tables 2-6.

Due to the complexity of the model, numerical derivatives were adopted for sensitivity coefficients using the central difference scheme instead of analytical derivatives. In the numerical differencing, the mean values of the stochastic input parameters were used as nominal and the increment of X_i was set, after several trials for numerical convergence, equal to $0.01|\mu_i|$ if $\mu_i \neq 0$ and equalled to 0.001 if $\mu_i = 0$. By Harr's PE method, the correlation coefficient matrix $C_{28 \times 28}$ was decomposed into eigenvalues and eigenvectors using the computer software MATLAB (Matrix Laboratory) ("MATLAB" 1989).

In the framework of LHS the regression coefficients for a given equation were considered to have multivariate normal distributions. Regression coefficients between different regression equations were considered uncorrelated. Input parameters associated with channel bed characteristics such as n, S_f, d_s, γ_s were assumed not correlated with any other parameters. Sixty Latin hypercubic sample sets ($K = 60$) for the 28 stochastic input parameters were generated based on their respective statistical properties. According to McKay (1988), a sample size around $2n = 2(28) = 56$ would be sufficient in the LHS procedure for the purpose of uncertainty and sensitivity analyses. After Latin hypercubic samples were generated, the sample statistics such

as mean, standard deviation, and correlation structure were computed to check the compliance to the population statistical characteristics of the input parameters.

RESULTS AND ANALYSIS

FOVE Method

Two cases were considered in uncertainty analysis by the FOVE method. The first case assumed that all 28 stochastic input parameters are uncorrelated. The estimated mean and standard deviation of the model output H_d under this assumption are 1.304 m and 0.340 m, respectively. The sensitivity coefficients s_i and the relative contribution of each stochastic input parameter C_i to the overall uncertainty of H_d are listed in Table 7. It can be seen from Table 7 that some of the variables, such as a_6 and r_3 , have large sensitivity coefficients but have small contributions to the overall uncertainty of the model output H_d . On the other hand, some input param-

TABLE 7. Sensitivity Coefficients and Relative Contribution of Each Input Parameter to Uncertainty of Model Output H_d (Uncorrelated Condition)

Input parameter (1)	Sensitivity coefficient ($\partial H_d / \partial X_i$) (2)	% contribution (C_i) (3)	Rank of importance (4)
n	-36.0476	4.69	8
S_f	-6234.13	10.10	5
γ_s	0.6509	1.22	—
d_s	1073.94	13.33	3
a_1	0.1366	0.01	—
a_2	0.0971	0.03	—
a_3	10.5558	0.11	—
a_4	0.1588	0.02	—
a_5	0.0802	0.02	—
a_6	948.67	0.41	—
b_0	1.2719	11.98	4
b_1	0.9811	16.60	1 ^a
b_2	0.1691	0.17	—
c_0	0.2078	7.11	6
c_1	0.1603	2.47	—
c_2	0.0276	0.03	—
c_3	0.5659	6.62	7
d_0	1.3037	3.29	10
d_1	0.5465	0.11	—
d_2	0.1967	0.07	—
r_1	-9.7603	0.07	—
r_2	-2.0842	0.03	—
r_3	-27.20	0.13	—
e_1	0.1175	0.57	—
e_2	1.2719	2.39	—
e_3	0.2078	0.33	—
e_4	1.3037	13.68	2
e_5	-3.5025	4.41	9

^aRank 1 represents the most important parameter in the sense of percentage contribution to overall output uncertainty.

eters, such as b_1 and e_4 , having relatively small values of sensitivity coefficients, contribute significantly to the uncertainty of H_d . Therefore, the relative contribution of each input parameter to the overall uncertainty of a model output depends not only on the respective sensitivity coefficient but also on its own variance. On the basis of percentage contribution, the 10 most important input parameters to the overall uncertainty of the pit-migration model are identified in Table 7.

In the second case, correlation among input parameters was considered. The computed standard deviation of the model output was reduced from 0.340 m (under the uncorrelated condition) to 0.257 m, a variance reduction of 43%. Consequently, correlation among the stochastic input parameters should be included in the uncertainty analysis so as not to overestimate the variance of H_d . The important correlated stochastic input parameters contributing to the variance of H_d are listed in Table 8. From Table 8, the relative contribution of the variance of H_d resulting from the correlated pair (b_0, b_1) is the most important among all other pairs. Their effect on the total variance of H_d has about the same order of magnitude as the top five individually important parameters listed in Table 7, which include parameters b_0 and b_1 . This result suggests that two correlated and individually important input parameters can constitute an important pair affecting the overall uncertainty of model output. Two highly correlated parameters may not necessarily form an important pair if only one parameter is individually important and the other is insignificant in contributing to the overall uncertainty.

Harr's PE Method

In the pit-migration model considered, there are $2(n) = 2(28) = 56$ intersecting points on the hypersphere along the 28 principal axes defined by the eigenvectors. One of these 56 points resulted in zero pit depth downstream because the associated friction slope S_f is very small and the corresponding shear velocity is less than the critical shear velocity for the incipient motion of sediment. The estimated mean and standard deviation of the model output H_d were 1.595 m and 0.399 m, respectively. The uncertainty of H_d in terms of standard deviation, 0.399 m, obtained by the Harr's PE method is larger than the value (0.257 m) obtained from the FOVE method.

TABLE 8. Summary of Important Correlated Terms to Uncertainty of H_d

Correlated input and parameters (X_i and X_j) (1)	$s_i s_j \rho_{ij} \sigma_i \sigma_j$ (2)	$s_i s_j \rho_{ij} \sigma_i \sigma_j / 0.1154$ (3)	Rank of importance ^a (4)
b_0 and b_1	-0.0148	-0.1278	1
b_0 and b_2	0.0015	0.0131	4
b_1 and b_2	-0.0014	-0.0119	5
c_0 and c_1	-0.0031	-0.0273	3
c_0 and c_2	0.0005	0.0041	8
c_0 and c_3	-0.0069	-0.0600	2
c_1 and c_3	0.0012	0.0105	6
d_0 and d_1	-0.0007	-0.0058	7

^aIn terms of percentage contribution to the overall output uncertainty.

Note that the eigenvalues associated with the correlation matrix of the stochastic input parameters are the variances of standardized stochastic parameters in the transformed space via the eigenvalue-eigenvector orthogonal transform (Ang and Tang 1984). The relative contribution of each transformed parameter to the overall model output uncertainty can be computed separately by the ratio of the corresponding eigenvalue to the sum of all eigenvalues, which is equal to n . However, since the parameters in the transformed space are a linear combination of the parameters in the original space, inverting the process to provide the relative contribution of each original stochastic input parameter to the overall model output uncertainty is difficult. At the present stage, the different PE versions are not applicable in assessing sensitivity of input parameters on model output.

LHS

The histogram based on the 60 values of maximum pit depth is shown in Fig. 1. The mean and standard deviation of the maximum pit depth are 1.326 m and 0.273 m, respectively. The distribution of the model output is more or less symmetric with respect to its mean. On the basis of 60 LHS sets for the 28 stochastic input parameters and the corresponding maximum pit depths, an analysis to identify the importance and sensitivity of input parameters in determining the model output and its uncertainty can be made.

A useful index that measures the importance of input parameters is the simple correlation coefficient (CC), computed by

$$r_{iy} = \frac{\sum_{k=1}^K (x_{ki} - \bar{x}_i)(y_k - \bar{y})}{\sqrt{\left[\sum_{k=1}^K (x_{ki} - \bar{x}_i)^2 \right] \left[\sum_{k=1}^K (y_k - \bar{y})^2 \right]}} \quad (21)$$

where r_{iy} = the simple CC of input X_i and model output Y ; x_{ki} and y_k =

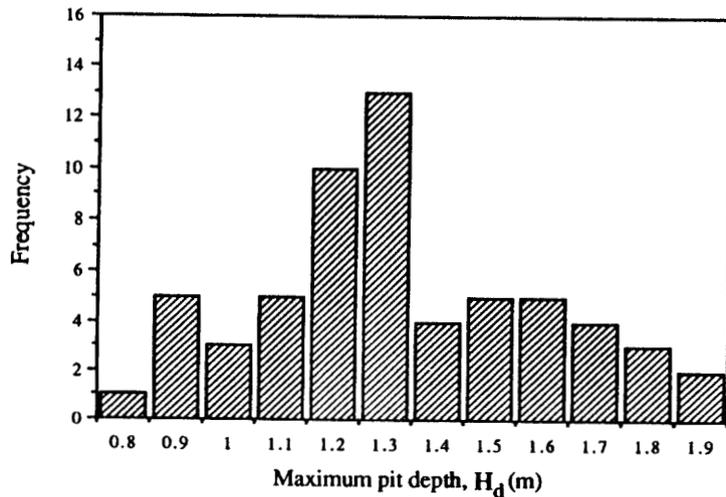


FIG. 1. Histogram of Maximum Pit Depth Based on 60 LHS Sets

the values of input X_i and the corresponding model output, respectively, for the k th LH sample; \bar{x}_i and \bar{y} = the sample means of input parameter X_i and model output based on the K LHS sets, respectively. The squared value of the correlation coefficient represents the percentage of variation (or uncertainty) in model output Y that can be explained by the input parameter under consideration. A given input parameter with a high correlation coefficient indicates that a large percentage of the model output variability (uncertainty) can be accounted for by that input parameter. In this sense, a stochastic input parameter with a high correlation coefficient indicates its importance in contributing to model output uncertainty.

Note that the simple CC computed by (21) does not take into account the presence of correlation between input parameters in the process of determining model output. A better indicator is the partial correlation coefficient (PCC) which accounts for linear dependency among the input parameters. The PCC between the model output Y and input parameter X_i can be determined as

$$c_{iy} = \frac{-c^{iy}}{\sqrt{c^{ii}c^{yy}}} \quad (22)$$

with c^{iy} , c^{ii} , and c^{yy} being the elements of the simple CC matrix inverse

$$\begin{bmatrix} C(X, X) & C(X, Y) \\ C(X, Y) & 1 \end{bmatrix}^{-1} = \begin{bmatrix} c^{11} & c^{12} & \dots & c^{1n} & | & c^{1y} \\ c^{21} & c^{22} & \dots & c^{2n} & | & c^{2y} \\ \dots & \dots & \dots & \dots & | & \dots \\ c^{n1} & c^{n2} & \dots & c^{nn} & | & c^{ny} \\ \dots & \dots & \dots & \dots & | & \dots \\ c^{y1} & c^{y2} & \dots & c^{yn} & | & c^{yy} \end{bmatrix} \quad (23)$$

where $C(X, X)$ is an $n \times n$ correlation matrix for the stochastic input parameters; and $C(X, Y)$ is an $n \times 1$ column vector of correlation coefficients between the n stochastic input parameters and the model output Y .

Although PCC is an improvement over the simple CC, it is like the simple CC in that it only measures the linear association between two random variables. To measure the possible nonlinear but monotonic relation between the two random variables, a rank correlation coefficient (RCC) and partial RCC (PRCC) can be used. The RCC and PRCC are computed in a manner similar to (22) and (23) by replacing the values of random variables under consideration by their respective integer ranks. After the rank transform the two random variables will have an exact linear relationship if they have a monotonic relation. Two random variables having low CC or PCC, but high RCC or PRCC, can be related strongly in a nonlinear fashion.

The simple CC, PCC, RCC, and PRCC of the maximum pit depth with all 28 stochastic input parameters are tabulated in Table 9. Judging from the values of the four correlation coefficients, the input parameters defining the channel characteristics (n , S_f , and d_s) and error terms associated with the regression equations for H_{max} [(1)] and u_{bc} [(6)] are significantly more important than the remaining input parameters. Among the 19 regression coefficients, only b_1 in (2) can be considered an important input parameter. Referring to Table 1, these important input parameters do not necessarily all have very high uncertainty individually. The relative importance of each input parameter also depends on the functional form in which the parameter is used. Comparing the values of PCC and PRCC, practically all input parameters have about the same values. This indicates that the use of a

TABLE 9. Simple CC, PCC, RCC, and PRCC of Maximum Pit Depth and 28 Input Parameters in Pit-Migration Model

Input variable (1)	CC (2)	PCC (3)	RCC (4)	PRCC (5)
<i>n</i>	-0.531	-0.822 [4]	-0.512	-0.629 [5]
<i>S_f</i>	-0.270	-0.888 [3]	-0.305	-0.840 [3]
<i>γ_s</i>	0.162	0.432 [8]	0.168	0.248 [10]
<i>d_s</i>	0.468	0.947 [1]	0.477	0.900 [1]
<i>a₁</i>	-0.028	-0.165	-0.098	-0.303 [8]
<i>a₂</i>	0.064	-0.143	0.133	-0.235
<i>a₃</i>	-0.057	-0.142	-0.092	-0.084
<i>a₄</i>	0.017	-0.183	0.061	-0.197
<i>a₅</i>	-0.079	-0.128	-0.149	-0.073
<i>a₆</i>	0.072	-0.119	0.116	0.083
<i>b₀</i>	-0.044	0.337 [9]	-0.012	0.199
<i>b₁</i>	0.215	0.613 [6]	0.181	0.460 [6]
<i>b₂</i>	0.051	0.086	0.067	0.110
<i>c₀</i>	-0.210	-0.002	-0.194	-0.268 [9]
<i>c₁</i>	0.211	0.087	0.217	0.151
<i>c₂</i>	-0.151	0.127	-0.155	0.420 [7]
<i>c₃</i>	0.184	0.079	0.194	-0.108
<i>d₀</i>	0.173	0.139	0.191	-0.001
<i>d₁</i>	-0.100	0.071	-0.120	-0.041
<i>d₂</i>	0.220	0.069	0.215	0.212
<i>r₁</i>	0.092	-0.235	0.012	0.197
<i>r₂</i>	-0.018	-0.245 [10]	-0.034	0.120
<i>r₃</i>	-0.093	-0.233	-0.010	0.252
<i>e₁</i>	0.306	0.039	0.232	-0.112
<i>e₂</i>	0.365	0.614 [5]	0.344	0.636 [4]
<i>e₃</i>	-0.040	0.445 [7]	-0.051	0.035
<i>e₄</i>	0.497	0.929 [2]	0.481	0.857 [2]
<i>e₅</i>	0.088	0.073	0.057	-0.077

Note: Number in the bracket represents the rank of importance.

nonlinear relation between model output (the maximum pit depth) and an individual input parameter does not improve the association.

In addition to the use of PCC and PRCC, the importance of input parameters can be identified by regression analysis. In regression analysis, the model outputs computed from the *K* input sets generated by the LHS technique are related to the *n* stochastic input parameters, in the simplest case, as

$$y_k = \beta_0 + \sum_{i=1}^n x_{ki}\beta_i, \quad k = 1, 2, \dots, K \quad (24)$$

Then, the least-squares estimates of β_s indicate the values of the sensitivity coefficient associated with each input parameter over the entire range of input space considered in generating LHS sets. In most situations, inputs and outputs may have different units and ranges, so it is advisable to standardize inputs and to centralize the output (the dependent variable). Then, the regression coefficient β^* s in the following model are sought

$$y_k^* = \beta_0^* + \sum_{i=1}^n z_{ki}\beta_i^*, \quad k = 1, 2, \dots, K \quad (25)$$

where $z_{ki} = (x_{ki} - \bar{x}_i)/\sigma_i$ and $y_k^* = y_k - \bar{y}$. Since β^* s have the same units as the output, they represent the change in output due to a change in input by one standard deviation. The values of *t*-statistics (called *t*-ratio) associated with β^* s indicate the relative predictive quality, or significance, of the stochastic input parameters.

Whereas correlation coefficients indicate the strength of the association between inputs and output, regression coefficients represent the intensity of the relation. Results of regression analysis, based on (25), are given in Table 10. The coefficient of determination *R*² associated with the regression model is 98.1% with a standard error = 0.05160 m. The important inputs identified on the basis of PCC and PRCC (shown in Table 9) have exceptionally high values for the *t*-ratio. This indicates that input parameters *n*, *S_f*, *d_s*, *b₁*, *e₂*, and *e₄* have very good predictive quality. From the viewpoint

TABLE 10. Results of Regressing Maximum Pit Depth on Input Parameters of Pit-Migration Model

Input variable (1)	β_s (2)	Standard deviation (3)	<i>t</i> -ratio (4)	Rank (5)
<i>n</i>	-0.074876	0.009274	-8.07	4
<i>S_f</i>	-0.096947	0.008682	-11.17	3
<i>γ_s</i>	0.027275	0.009302	2.93	7
<i>d_s</i>	0.143266	0.008654	16.56	1
<i>a₁</i>	-0.06020	0.09450	-0.64	—
<i>a₂</i> [*]	—	—	—	—
<i>a₃</i>	0.00349	0.09419	0.04	—
<i>a₄</i>	-0.05994	0.07086	-0.85	—
<i>a₅</i>	0.0496	0.1217	0.41	—
<i>a₆</i>	0.0519	0.1588	0.33	—
<i>b₀</i>	0.13106	0.07075	1.85	9
<i>b₁</i>	0.15130	0.03531	4.29	5
<i>b₂</i>	0.03259	0.04314	0.76	—
<i>c₀</i>	0.0234	0.1743	0.13	—
<i>c₁</i>	0.03292	0.04986	0.66	—
<i>c₂</i>	0.03011	0.05060	0.60	—
<i>c₃</i>	0.0722	0.1160	0.62	—
<i>d₀</i>	0.0930	0.1086	0.86	—
<i>d₁</i>	0.0537	0.1167	0.46	—
<i>d₂</i>	0.00804	0.03375	0.24	—
<i>r₁</i>	-0.1585	0.1166	-1.36	—
<i>r₂</i>	-0.02927	0.02001	-1.46	10
<i>r₃</i>	-0.1548	0.1157	-1.34	—
<i>e₁</i>	0.000700	0.009201	0.08	—
<i>e₂</i>	0.04490	0.01048	4.28	6
<i>e₃</i>	0.026252	0.009666	2.72	8
<i>e₄</i>	0.124358	0.008558	14.53	2
<i>e₅</i>	0.007214	0.009781	0.74	—

^{*}Deleted by the computer package due to its high correlation with other parameters.

of model parsimony, input parameters with less significance can be discarded from the regression model without jeopardizing model's predictive quality. In this sense, those input parameters that have little statistical significance from a regression study can be treated as constants in the uncertainty analysis of model output. The selection of input parameters can be done subjectively in an iterative manner (e.g., retain those with t -ratios greater than 1.0) or through the use of formal statistical procedures (e.g., stepwise regression).

Based on Table 10, the stochastic input parameters with little statistical significance are deleted from the full regression model. The results of the reduced regression model containing 21 of the original 28 input parameters are shown in the first four columns of Table 11. The corresponding R^2 is 98.0% and the standard error is 0.04753 m. The reduced regression model maintains practically the same level of R^2 , whereas the standard error is improved by 8% due to an increase in degrees of freedom. The PCCs of 21 input parameters in the reduced regression model are also listed in column 5 of Table 11. As can be seen, an input parameter with large value of t -ratio generally is associated with a high value of PCC. Many of the input parameters deleted from the full regression model are mainly due to high correlation with other variables. For example, only a_4 and a_5 in (4) for u_{bd} are retained in the reduced regression model. This is because other coefficients $a_1, a_2, a_3,$ and a_6 are redundant in the sense that they are highly correlated with a_4 and a_5 . The reduced regression model then identifies the important input parameters affecting the sensitivity and uncertainty of pit-migration model output in a global sense.

As mentioned previously, the regression coefficients provide a measure

TABLE 11. Results of Regressing H_d on Reduced Input Parameters of Pit-Migration Model and PCC between H_d and Input Parameters

Input variable (1)	β_s (2)	Standard deviation (3)	t -ratio (4)	PCC (5)
n	-0.074802	0.007972	-9.38	-0.8324
S_f	-0.096084	0.007214	-13.32	-0.9054
γ_s	0.027306	0.008219	3.32	0.4696
d_s	0.144657	0.007454	19.41	0.9519
a_4	-0.023149	0.008362	-2.77	-0.4052
a_5	-0.017435	0.008015	-2.18	-0.3289
b_0	0.13549	0.04985	2.72	0.3990
b_1	0.15208	0.02633	5.78	0.6790
b_2	0.02570	0.03081	0.83	0.1323
c_1	0.02779	0.01125	2.47	0.3678
c_2	0.03492	0.01288	2.71	0.3982
c_3	0.054153	0.009572	5.66	0.6713
d_1	-0.045797	0.007293	-6.28	-0.7090
d_2	0.034954	0.007839	4.46	0.5810
r_1	-0.12497	0.09892	-1.26	-0.1982
r_2	-0.02324	0.01732	-1.34	-0.2100
r_3	-0.12084	0.09709	-1.24	-0.1954
e_2	0.044209	0.008052	5.49	0.6602
e_3	0.027787	0.007795	3.56	0.4957
e_4	0.125115	0.007437	16.82	0.9374
e_5	0.006134	0.007084	0.87	0.1373

of global sensitivity of the maximum pit depth to the input parameters over the entire input domain. Referring to Table 10, the regression coefficients associated with those significant input parameters are generally, but not always, larger than those insignificant ones. Insignificant input parameters with large regression coefficients might indicate fluctuation of local sensitivity of opposite signs in different regions of the input domain, resulting an domain-averaged sensitivity not statistically significantly from zero. This lack of resolution is a limitation that could potentially reduce the effectiveness of LHS for sensitivity analysis.

Comparison of Three Uncertainty Analysis Methods

Comparing Tables 7, 9, and 10, the ranking of importance of parameters can vary with method. The ranking based on t -ratio (Table 10) is practically identical to that using PCC (column 3 of Table 9). PCC and PRCC (Table 9) yield basically the same ranking for the six parameters having the highest correlation coefficients with the model output. Note that the ranking of input parameters by the FOVE method is quite different from that by regression analysis. The main discrepancy is attributed to the difference in domain in which the two methods operate. That is, the FOVE method gives the indication of importance measured in the vicinity of a selected point in the entire parameter space, whereas measures such as PCC, PRCC, and t -ratio indicate the importance of a model parameter in a global sense over the entire parameter space. Discrepancy in the signs of sensitivity coefficients in the FOVE method and those of PCC, PRCC, and t -ratio for some parameters also indicates the local and global effects.

The means and standard deviations of the model output H_d computed by the three uncertainty analysis methods are summarized in Table 12. Values for the mean and standard deviation of H_d obtained by Harr's PE method are larger than those obtained by the other two methods. Numerical examples given in Harr (1989) show that Harr's PE method has the tendency to overestimate the mean and standard deviation of the model output over those obtained by Rosenblueth's PE method. This tendency becomes more pronounced as the degree of nonlinearity of the model increases. The FOVE method, when correlation among parameters is considered, yields close but slightly lower values than those obtained by the LHS technique. Although the three techniques yield approximations to the true statistical moments of the model output, the LHS procedure has better theoretical support for a nonlinear model, which is the case for the pit migration model.

For all three uncertainty analysis techniques considered, computational effort is largely determined by the number of computer runs of the model under consideration. In this aspect, the three techniques are comparable. Using the central differencing scheme, the FOVE method requires $2n + 1$

TABLE 12. Comparison of Mean and Standard Deviation of H_d among Three Methods

Statistics (1)	FOVE		Harr's PE (4)	LHS (5)
	Uncorrelated (2)	Correlated (3)		
Mean	1.304	1.304	1.595	1.326
Standard deviation	0.340	0.257	0.399	0.273

= $2(28) + 1 = 57$ computer runs of the pit-migration model to obtain the sensitivity coefficients. Of course, the number of the computer runs can be reduced by about half if one adopts the forward or backward differencing scheme. Harr's PE method needs $2n = 2(29) = 56$ computer runs. The LHS procedure, for purpose of subsequent sensitivity and uncertainty analyses, requires at least $n + 1$ computer runs based on the sample sets determined by the technique. McKay (1988) pointed out that $2n$ LHS sets are generally sufficient. In case the execution time for each model run is lengthy, subsets of the generated LHS sets can be selected for model evaluation.

With about the same amount of computational effort, the FOVE method and LHS procedure can produce more information than Harr's PE method (including the Rosenblueth original version). In addition to estimates of the mean and standard deviation of the model output, the FOVE method provides information concerning local sensitivity and percentage contribution of individual stochastic input parameters to the overall model uncertainty. Similarly, from LHS, the importance of stochastic input parameters with regard to the model sensitivity and uncertainty in a global sense can be identified. Although Harr's PE method allows assessments of relative contribution to overall model output uncertainty in the transformed space, the conversion back to the original parameter space is difficult.

Improvement in the accuracy of the FOVE method could be made by incorporating the correlation relationship of stochastic input parameters in estimating the mean model output. This results in a second-order approximation. The price to pay for the improvement is an increase of computer model runs to estimate the second-order partial derivatives by numerical differencing. Improvement in accuracy of LHS can be made through generating more LHS sets, hence increase computer model runs.

The FOVE method is simple, effective, and straightforward once the sensitivity coefficients, variances, and covariance matrix of the stochastic input parameters are quantified. The FOVE method is particularly accurate if the relations between model output and inputs are linear or close to linear. The merit of the FOVE method is that important stochastic input parameters can be identified easily. In addition, any changes in uncertainties of the input parameters can easily be incorporated to update the uncertainty of the model output. Note that calculations of sensitivity coefficients might be cumbersome and time-consuming for a large model in practice. Numerical differencing must be applied for models whose analytical derivatives are not obtainable. One intrinsic drawback of the FOVE method is that the method is insensitive to the distribution of stochastic input parameters. In fact, Harr's PE method is like the FOVE method in that only the first two moments of a random variable, not its distribution type, are used in the computation. Harr's method is an approximation to Rosenblueth's PE method, which is capable of considering the distribution type of the stochastic input parameters. On the other hand, the PE methods provide a general framework allowing the estimation of statistical moments of any order without significantly increasing the computational burden, as would the FOVE method when high-order moments are sought.

SUMMARY AND CONCLUSIONS

Sand and gravel mining from a river bed leaves pits of a variety of sizes and shapes in the river bed that can migrate downstream, threatening the safety of bridge piers and other in-stream hydraulic structures. A pit-migration model has been developed by Lee et al. (1990) based on a series of

laboratory experiments. Within the range of the experimental data, the model provides predictions of the moving speed, shape, and maximum depth of the pit as it travels downstream. The model can be utilized to regulate mining operations and to evaluate the impact of pits on in-stream structures. Like other hydraulic/hydrologic models, the pit-migration model involves parameters that are subject to uncertainty. Due to implications of using the model for predictions, an examination of the model behavior and its uncertainty is of value.

In this study, three uncertainty analysis techniques are applied to assess the uncertainty associated with the maximum pit depth predicted by the pit-migration model. In addition, features of the three uncertainty analysis techniques and their performance are compared. Numerical investigation indicates that the PE method proposed by Harr (1989) yields higher values in mean and standard deviation than the other methods, i.e., the FOVE method and LHS. This implies that the use of Harr's PE method would result in a conservative prediction for the maximum pit depth. Computationally, the three uncertainty analysis procedures are comparable. However, the FOVE method and LHS yield more information with regard to the relative importance of stochastic input parameters in model sensitivity and uncertainty. Each method has its own advantages and weaknesses. The selection of a method for analysis depends on study objectives.

This paper demonstrates capabilities of three uncertainty analysis techniques, which can be extended to analyze the behavior of more complex hydraulic and hydrologic models. Also, the paper illustrates that in the course of analyzing model uncertainty a lot of information (e.g., the individual contribution of input parameters to the uncertainty of the model output) with regard to the behavior of model inputs and output can be derived. This is not generally possible through a purely deterministic analysis.

ACKNOWLEDGMENTS

This research was financially supported by Water Resources Planning Commission, MOEA, Republic of China. The writers wish to express their gratitude to H. Y. Lee of the National Taiwan University for providing the experimental data of pit migration. Advice and discussions given by J. C. Yang during the course of study are also appreciated. Finally, the writers are thankful to the reviewers whose constructive criticisms really make us think hard about what we present in the paper.

APPENDIX I. 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.
- Chow, V. T. (1959). *Open-channel hydraulics*. McGraw Hill Book Co., New York, N.Y.
- Harr, M. E. (1989). "Probabilistic estimates for multivariate analysis." *Appl. Math. Modelling*, 13(5), 313-318.
- Johnson, M. E. (1987). *Multivariate statistical simulation*. John Wiley and Sons, Inc., New York, N.Y.
- Lee, H. Y., Young, D. L., and Huang, L. H. (1990). "A study of the effects of sand and gravel mining operations on the cross-river structures and channel stability of the Tansui river." *Tech. Report No. 119*, Hydr. Res. Lab., Nat. Taiwan Univ., Taiwan (in Chinese).

- "MATLAB for 80386-based MS-DOS personal computers." (1989). *User's guide*, The Math Works Inc., South Natick, Mass.
- McKay, M. D. (1988). "Sensitivity and uncertainty analysis using a statistical sample of input values." *Uncertainty Analysis*, Y. Ronen, ed., CRC Press, Inc., Boca Raton, Fla., 145-186.
- Rosenblueth, E. (1975). "Point estimates for probability moments." *Proc. Nat. Academy of Science*, 72(10), 3812-3814.
- Rosenblueth, E. (1981). "Two-point estimates in probabilities." *Appl. Math. Modeling*, 5(5), 329-335.
- Yen, B. C., Chen, S. T., and Melching, C. S. (1986). "First-order reliability analysis." *Stochastic and risk analysis in hydraulic engineering*, B. C. Yen, ed., Water Res. Publ., Littleton, Colo., 1-36.

APPENDIX II. NOTATION

The following symbols are used in this paper:

- a_1, \dots, a_6 = regression coefficients in (4);
 B = channel width;
 b_0, b_1, b_2 = regression coefficients in (2);
 C = correlation matrix of stochastic input parameters X ;
 C_i = contribution of X to overall uncertainty of Y (under independent condition);
 c_1, \dots, c_3 = regression coefficients in (5);
 c^{ij}, c^{ii}, c^{jj} = elements of simple correlation coefficient matrix inverse;
 D = specified distance of pit migration;
 d_0, d_1, d_2 = regression coefficients in (1);
 d_p = representative particle size;
 e_1, \dots, e_5 = model error terms;
 g_r = gravitational acceleration;
 H = initial pit depth;
 H_d = maximum pit depth after traveling some distance downstream;
 H_{max} = maximum pit depth at end of convection period;
 K = number of LHS sets;
 L = diagonal matrix with eigenvalues as its elements;
 L = initial pit length;
 L_e = effective length at end of the convection period;
 n = Manning's roughness;
 Q = flow discharge;
 R^2 = coefficient of determination;
 r_1, r_2, r_3 = regression coefficients in (6);
 S_f = friction slope;
 s_i = sensitivity coefficient of model output with respect to X_i ;
 s_1, \dots, s_5 = standard error of estimates;
 T_f = time span for convection period;
 t = total travel time of pit migration;
 u_* , u_{*c} = shear velocity and critical shear velocity, respectively;
 u_{bc} , u_{bd} = pit migrating velocities during convection period and diffusion period, respectively;
 V = eigenvector matrix;
 $\text{var}(Y)$ = variance of model output Y ;

- X = n -dimensional column vector of stochastic input parameters;
 x_0 = Taylor series expansion point;
 x_{ki}, y_k = k th LHS for input parameter X_i and corresponding model output, respectively;
 \bar{x}_i, \bar{y} = sample means of input parameter X_i and model output based on LHS sets, respectively;
 Y = model output;
 y = flow depth;
 y_k^* = $y_k - \bar{y}$;
 z_{ki} = standardized form of stochastic input parameter X_i in LHS sets;
 β_i^*, β_i = linear regression coefficients;
 γ_s = specific weight of sediment;
 λ = eigenvalue;
 μ = vector of means of stochastic input parameters;
 μ_i, σ_i = mean and standard deviation of input parameter X_i , respectively;
 μ_Y, σ_Y = mean and standard deviation of model output, respectively;
 ρ_{ij} = correlation coefficient between input parameters X_i and X_j ; and
 Ω = covariance matrix of X s.