Stochastic Hydrology and Hydraulics 11: 173-192

Regionalization of unit hydrograph Parameters: 2. Uncertainty analysis

K.-C. Yeh and **J.-C. Yang**
Department of Civil Engineering, National Chiao-Tung University, Hsinchu, Taiwan, Republic of China

Y.-K. Tung
Wyoming Water Resources Center & Statistics Department, The University of Wyoming, Laramie, WY 82071, USA

Abstract: Hydrologic model parameters obtained from regional regression equations are subject to uncertainty. Consequently, hydrologic model outputs based on the stochastic parameters are random. This paper presents a systematic analysis of uncertainty associated with the two parameters, N and K, in Nash's IUH model from different regional regression equations. The uncertainty features associated with N and K are further incorporated to assess the uncertainty of the resulting IUH. Numerical results indicate that uncertainty of N and K from the regional regression equations are too significant to be ignored.

Keywords: Uncertainty analysis, unit hydrograph, regression analysis, probabilistic point estimation methods.

1 Uncertainties in regional regression equations

In hydrologic and hydraulic analysis and design, system responses are often described by empirical equations such as equation (1) in Part 1 of the accompanying paper (Tung et al., 1995). Due to the fact that model parameters $\boldsymbol{\theta}$'s are estimated from limited amount of data, they are subject to sampling errors. Furthermore, due to lack of perfect fit between the observed system responses and modeled responses, the empirical models are also subject to model uncertainty represented by the error term $\in$. Detail discussions on the presence of uncertainties in an empirical equation are given by Yeh and Tung (1993) and Tung (1994).

Suppose that parameters $\boldsymbol{\theta}$'s and $\in$ in equation (1) are estimated involving uncertainty and they are treated as random variables. Furthermore, assume that their first two moments (including their covariance) are quantified. Then, by the first-order approximation (Mays and Tung, 1992; Tung and Yen, 1993), the mean and variance of hydrologic response y for a given basin characteristics $\mathbf{x}_0$ can be estimated, respectively, by

$$E(y|\mathbf{x}_{\boldsymbol{\theta}}) \approx g(\mu_\theta|\mathbf{x}_{\boldsymbol{\theta}}) + \mu_\in \quad (32)$$

$$\mathrm{Var}(y_0|x_0) \;\approx\; \nabla g(\mu_\theta|x_0)^t\; \mathbf{C}(\boldsymbol{\theta})\nabla g(\mu_\theta|x_{\boldsymbol{\theta}} + \;\sigma^2_\epsilon \tag{33}$$

in which $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\mu}_\epsilon$ are vectors of means for the stochastic model parameters $\boldsymbol{\theta}$'s and the model error term, respectively; $\nabla g(\boldsymbol{\mu}_\theta|x_0)$ is a gradient vector measuring the sensitivity of model response to the unit change in model parameter, that is, $\nabla g(\boldsymbol{\mu}_\theta|x_0) = [\partial g/\partial\theta_q]$, evaluated at $\boldsymbol{\mu}_\theta$; $\mathbf{C}(\boldsymbol{\theta})$ is the covariance matrix of stochastic model parameters; $\boldsymbol{\sigma}^2_\epsilon$ = variance of the model error term.

When equation (1) can be expressed as equation (18a), equation (32) becomes equation (4a) and equation (33) reduces to

$$\mathrm{Var}(y_0|x_0) \approx\; \nabla g(\boldsymbol{\mu}_b|x_0)^t\; \mathbf{C}(\mathbf{b})\nabla g(\boldsymbol{\mu}_b|x_0) \;+\; \boldsymbol{\sigma}^2_\epsilon$$

$$=\; x_0^t\; \mathbf{C}(\mathbf{b})x_0 \;+\; \boldsymbol{\sigma}^2_\epsilon \tag{34a}$$

$$=\; \boldsymbol{\sigma}^2_\epsilon[1 + x_0^t(\mathbf{X}^t\mathbf{X})^{-1}x_0] \tag{34b}$$

which is identical to equation (4b) in Part 1.

It should be pointed out that, in the above quantification of uncertainty associated with the dependent variable of a regional regression equation, only the parameter and model uncertainties are considered while the independent variables are considered free of errors. This treatment is consistent with the classical regression analysis by which only the dependent variable is considered as random variable. If the uncertainties associated with the independent variables are to be considered, special treatments in regression analysis are required. Alternatively, a practical way is to apply methods such as first-order variance estimation method or probabilistic point estimation procedures to regional regression models by which model inputs, parameters, and model error term are all treated as random variables. Presently, this study limits its scope of quantifying uncertainties associated with hydrologic regional equations in the classical sense and demonstrates how the information provided by regression analysis can be directly utilized.

Note that the derived regional regression equations in Part 1 provide estimations of means, standard deviations, and correlation coefficient of N and K in the log-space, namely, $\mu_{\ln N}$, $\boldsymbol{\sigma}_{\ln N}$, $\mu_{\ln K}$, $\boldsymbol{\sigma}_{\ln K}$ and $\boldsymbol{\rho}_{\ln N, \ln K}$. Under the normality condition for ln(N) and ln(K) as shown in Part 1, the statistical moments of N and K in the original space can be analytically derived. Without making any parametric assumptions, the mean and variance of N and K in the original space can be approximated by methods such as first-order variance estimation technique (Tung and Yen, 1993).

Using the analytical approach assuming bivariate log-normal distribution for N and K, the means and variances of N and K can be computed by the following general formulas as

$$\mu_x \;=\; \exp\left[\mu_{\ln X} \;+\; \frac{\sigma^2_{\ln X}}{2}\right] \tag{35}$$

$$\sigma^2_x \;=\; \mu^2_x\left[\exp(\sigma^2_{\ln X} \;-\; 1)\right] \tag{36}$$

in which X can be N or K in the original space and ln(X) be ln(N) or ln(K). The correlation coefficient of N and K in the original space can be computed as

$$\boldsymbol{\rho}_{N,K} \;=\; \frac{\mathbf{c}^{\rho_{\ln N, \ln K}\sigma_{\ln N}\sigma_{\ln K}} \;-\; 1}{\Omega_N\Omega_K} \tag{37}$$

in which $\Omega_N = \sigma_N/\mu_N$ and $\Omega_K = \sigma_K/\mu_K$ are the coefficients of variation of N and K, respectively.

2 Assessment of uncertainties associated with regionalized Nash's IUH parameters

2.1 Regionalized Nash's IUH parameters from univariate regression (UVR)

For a selected watershed with known basin characteristics, $\mathbf{x}_0$ = (area, L_{ca}, length, slope), the means of predicted ln(N) and ln(K) can be estimated based on the regression coefficients in Table 4. The variances associated with the predicted ln(N) and ln(K) can be estimated by either equations (34a-b). By equation (34a) the covariance matrix of estimated regression coefficients for each regional regression equation is used. More specifically, the variances of ln(N) and ln(K) associated with a given basin characteristic $\mathbf{x}_0$, respectively, are

$$\sigma^2_{\ln(N_0)} = \mathrm{Var}\left[\ln(N_0)|\mathbf{x_0}\right] = \mathbf{x}^t_{0,\ln N}\mathbf{C}(\mathbf{b_{lnN}})\mathbf{x_{0,lnK}} + \sigma^2_{\in_{\ln N}} \tag{38}$$

$$\sigma^2_{\ln(K_0)} = \mathrm{Var}\left[\ln(K_0)|\mathbf{x_0}\right] = \mathbf{x}^t_{0,\ln K}\mathbf{C}(\mathbf{b_{lnK}})\mathbf{x_{0,lnK}} + \sigma^2_{\in_{\ln K}} \tag{39}$$

in which $\mathbf{x}_{0,\ln N}$ and $\mathbf{x}_{0,\ln K}$ are, respectively, the vectors of basin characteristics in the derived regional equations for predicting ln(N) and ln(K); $\mathbf{C}(\mathbf{b}_{\ln X})$ is the covariance matrix associated with $\mathbf{b}_{\ln X}$); and $\sigma_\in$ is the standard error of estimate for the corresponding regional equation. The elements in $\mathbf{x}_{0,\ln X}$ should correspond to the derived regional equation. The regression coefficients $\mathbf{b}_{\ln N}$, $\mathbf{b}_{\ln K}$, the model error $\sigma_\in$, and the covariance matrices for $\mathbf{b}_{\ln N}$ and $\mathbf{b}_{\ln K}$ from the UVR are provided by most statistical packages.

Using equation (34b), the matrix $(\mathbf{X}^t\mathbf{X})^{-1}$ associated with each derived regional equation has to be calculated and the variances corresponding to the predicted ln(N) and ln(K) can be calculated as

$$\sigma^2_{\ln(N_0),UVR} = \sigma^2_{\in_{\ln N},UVR}\left[1 + \mathbf{x}^t_{0,\ln N}\left(\mathbf{X}^t_{\ln N}\mathbf{X}_{\ln N}\right)^{-1}\mathbf{x}_{0,\ln N}\right] \tag{40}$$

$$\sigma^2_{\ln(K_0),UVR} = \sigma^2_{\in_{\ln K},UVR}\left[1 + \mathbf{x}^t_{0,\ln K}\left(\mathbf{X}^t_{\ln K}\mathbf{X}_{\ln K}\right)^{-1}\mathbf{x}_{0,\ln K}\right] \tag{41}$$

Using the conventional univariate regression, UVR0, the correlation between the residuals of ln(N) and ln(K) is not considered. As an approximation, the sample correlation between ln(N) and ln(K) can be used to account for the correlation between the predicted model parameters. Based on Table 3(b) in Part 1, a value of $\rho_{\ln N,\ln K}$ = -0.753 can be used and the covariance between the predicted model parameters is obtained as

$$\mathrm{Cov}\left[\ln(N_0), \ln(K_0)\right] = -0.753\ \sigma_{\in_{\ln(N_0)},UVR}\sigma_{\in_{\ln(K_0)},UVR}\ . \tag{42}$$

Based on the derived statistical moments of ln(N) and ln(K) from the regional equations, the means and variances of N and K can be computed by equations (35)-(37).

2.2 Regionalized Nash's IUH parameters from multivariate regression (MVR)

By the multivariate regression framework, the computations of means and variances of ln(N) and ln(K) are similar to those described for the UVR. Unlike the univariate case, the correlation between the residuals of ln(N) and ln(K) is used in equation (13) for computing the covariance of predicted model parameters. Furthermore, because the multivariate regression model contains the same independent variables for different dependent variables, the terms $(\mathbf{X}^t\mathbf{X})^{-1}$ in equations (40) and (41) are identical.

More specifically, for multivariate regression *with* intercept (MVRW), the vector of basin characteristics involved in the derived regional equations for both ln(N) and ln(K), referring to Table 4 in Part 1, is $\mathbf{x}^1_{0,\mathrm{MVRW}}$ = [1, ln(Area), ln(L_{ca}), ln(Slope), $\ln^2$(Area), $\ln^2$(Slope)]. The variances associated with ln(N) and ln(K) can be made as

$$\sigma^2_{\ln(N_0),\mathrm{MVRW}} = \sigma^2_{\epsilon_{\ln N},\mathrm{MVRW}}\left[1+\mathbf{x}^t_{0,\mathrm{MVRW}}\left(\mathbf{X}^t\mathrm{X}\right)^{-1}_{\mathrm{MVRW}}\mathbf{x}_{0,\mathrm{MVRW}}\right] \qquad (43)$$

$$\sigma^2_{\ln(K_0),\mathrm{MVRW}} = \sigma^2_{\epsilon_{\ln K},\mathrm{MVRW}}\left[1+\mathbf{x}^t_{0,\mathrm{MVRW}}\left(\mathbf{X}^t\mathrm{X}\right)^{-1}_{\mathrm{MVRW}}\mathbf{x}_{0,\mathrm{MVRW}}\right] \qquad (44)$$

The covariance between the predicted ln(N) and ln(K), according to Table 4, can be estimated by

$$\mathrm{Cov}\left[\ln(N_0),\ln(K_0)\right] = -0.77796\ \sigma_{\epsilon_{\ln(N_0)},\mathrm{MVRW}}\sigma_{\epsilon_{\ln(K_0)},\mathrm{MVRW}} \qquad (45)$$

Similarly, for multivariate regression *without* intercept (MVRWO), the vector of basin characteristics used in the derived regional equations for both ln(N) and ln(K) is $\mathbf{x}^t_{0,\mathrm{MVRWO}}$ = [ln(Area), ln(L_{ca}), ln(Slope), $\ln^2$(Area), $\ln^2$(Slope)]. The computation of the variances associated with ln(N) and ln(K) and their covariance can be made in the same way as equations (43)-(45) using the matrix $\mathbf{X}^t\mathbf{X})^{-1}$, the standard errors, and correlation coefficient provided by the statistical package.

2.3 Regionalized Nash's IUH parameters from seemingly unrelated regression (SUR)

According to Table 4 in Part 1, the basin characteristics used in the derived regional equations for IUH parameters by the SUR are

$$\mathbf{x}^t_{0,\ln N,\mathrm{SUR}} = [1,\ln(\mathrm{Area}),\ln(L_{ca}),\ln(\mathrm{Slope}),\ln^2(\mathrm{Area}),\ln^2(\mathrm{Slope})]$$

$$\mathbf{x}^t_{0,\ln K,\mathrm{SUR}} = [\ln(\mathrm{Area}),\ln(L_{ca}),\ln(\mathrm{Slope}),\ln^2(\mathrm{Area}),\ln^2(\mathrm{Slope})]\ .$$

Based on equation (14b), the covariance matrix of predicted ln(N) and ln(K) can be estimated as

$$\mathbf{C}(\ln N_0,\ln K_0)_{\mathrm{SUR}} = \begin{bmatrix} \sigma^2_{\ln N_0,\mathrm{SUR}} & \mathrm{Cov}(\ln N_{0,\mathrm{SUR}},\ln K_{0,\mathrm{SUR}}) \\ \mathrm{Cov}(\ln N_{0,\mathrm{SUR}},\ln K_{0,\mathrm{SUR}}) & \sigma^2_{\ln K_0,\mathrm{SUR}} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{x}^t_{\ln N,\mathrm{SUR}}\mathbf{C}(\mathbf{b}_{\ln N,\mathrm{SUR}},\mathbf{b}_{\ln N,\mathrm{SUR}}) & \mathbf{x}_{\ln N,\mathrm{SUR}}\mathbf{x}^t_{\ln N,\mathrm{SUR}}\mathbf{C}(\mathbf{b}_{\ln N,\mathrm{SUR}},\mathbf{b}_{\ln K,\mathrm{SUR}})\mathbf{x}_{\ln K,\mathrm{SUR}} \\ \mathbf{x}^t_{\ln K,\mathrm{SUR}}\mathbf{C}(\mathbf{b}_{\ln K,\mathrm{SUR}},\mathbf{b}_{\ln N,\mathrm{SUR}}) & \mathbf{x}_{\ln K,\mathrm{SUR}}\mathbf{x}^t_{\ln K,\mathrm{SUR}}\mathbf{C}(\mathbf{b}_{\ln N,\mathrm{SUR}},\mathbf{b}_{\ln K,\mathrm{SUR}})\mathbf{x}_{\ln K,\mathrm{SUR}} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2_{\epsilon_{\ln N,\mathrm{SUR}}} & \rho_{\epsilon_{\ln N,\mathrm{SUR}},\epsilon_{\ln K,\mathrm{SUR}}}\sigma_{\epsilon_{\ln N,\mathrm{SUR}}}\sigma_{\epsilon_{\ln K,\mathrm{SUR}}} \\ \rho_{\epsilon_{\ln N,\mathrm{SUR}},\epsilon_{\ln K,\mathrm{SUR}}}\sigma_{\epsilon_{\ln N,\mathrm{SUR}}}\sigma_{\epsilon_{\ln K,\mathrm{SUR}}} & \sigma^2_{\epsilon_{\ln K,\mathrm{SUR}}} \end{bmatrix} \qquad (46)$$

in which $\mathbf{C}(\mathbf{b}_{\ln N,SUR}, \mathbf{b}_{\ln N,SUR})$, is a 6×6 covariance matrix of estimated regression coefficients for ln(N), $\mathbf{C}(\mathbf{b}_{\ln N,SUR}, \mathbf{b}_{\ln K,SUR})$ is a 6×5 covariance matrix between the estimated regression coefficients for ln(N) and ln(K), $\mathbf{C}(\mathbf{b}_{\ln K,SUR}, \mathbf{b}_{\ln K,SUR})$ is a 5×5 covariance matrix of estimated regression coefficients for ln(K), $\boldsymbol{\sigma}_{\ln N,SUR}$ and $\boldsymbol{\sigma}_{\ln K,SUR}$ are the standard errors of estimate associated with the regional equations for ln(N) and ln(K), respectively, and $\boldsymbol{\rho}_{\epsilon\ln N,\epsilon\ln K}$ is the correlation coefficient between the residuals of ln(N) and ln(K).

3 Assessment of uncertainty in a regionalized unit hydrograph

From equations (15) and (16), the Nash's IUH ordinates, U(t), and those of the corresponding DUH ordinates, U_m, are functions of model parameters N and K. However, the values of model parameters, N and K, estimated from the regional regression equations for a given watershed should only be regarded as nominal and are inherently subject to uncertainties. These uncertainties, through equations (15) or (16), will be transmitted to the resulting UH ordinates. Hence, the derived IUH and DUH for a watershed involve uncertainty. The presence of uncertainty in UH can be incorporated in reliability analysis of hydraulic structures as shown by Yeh et al. (1993) and Zhao et al. (1995).

Since the same stochastic model parameters, namely, N and K, are used to determine the UH by equation (15) or (16), the UH ordinates at different times are not independent but correlated. Due to the nonlinear relationship between the UH ordinates and the model parameters, analytical derivations of the joint probability density function for the UH ordinates is practically impossible. As a practical alternative, this study focuses on estimating the moments of UH ordinates and the correlation among them.

3.1 Methods of uncertainty analysis

For a given watershed, several methods, based on the statistical information of N and K, can be applied to estimate the statistical moments of Nash's IUH ordinates and those of the corresponding DUH (Tung and Yen, 1993). Although the first-order variance method is frequently used in uncertainty analysis, it is less attractive in this study for the following reasons: (1) great degree of nonlinearity of Nash's IUH model; (2) rather large variances associated with the regionalized model parameter estimators; and (3) requirement of computing derivatives of u(t) which is a cumbersome exercise.

For the above reason, two probabilistic point-estimate (PE) methods, namely, Rosenblueth's method and Harr's method were applied. The PE methods evaluate uncertainty of a model output subject to stochastic model parameters by computing the model's responses at specified points in the parameter space. Therefore, probabilistic PE methods do not require computations of derivatives of U(t) with respect to N and K. By a probabilistic PE method, proper candidate points in the parameter space for function evaluations are selected to preserve the probabilistic characteristics of stochastic parameters. The general consideration is to preserve moments of stochastic model parameters. However, the required input moments vary among probabilistic PE algorithms. It has been shown by Karmeshu and Lara-Rosano (1987) that the first-order variance method is a special case of probabilistic PE methods when the uncertainty of stochastic parameters are small.

3.1.1 Rosenblueth's Point Estimation Method - In 1975, Rosenblueth proposed a probabilistic PE algorithm to deal with uncertainty analysis of a model involving symmetric stochastic parameters. It was later extended to handle non-symmetric variables (Rosenblueth 1981). The fundamental principle of the method is to determine the points for model evaluation in the parameter space in such a manner that the first three statistical moments of each individual stochastic parameter are preserved. The locations of the two points, x_+ and x_-, and the corresponding probability masses, p_+ and p_-, are obtained by solving

$$p_+ + p_- = 1$$

$$p_+z_+ - p_-z_- = \mu_z = 0$$

$$p_+z_+^2 - p_-z_-^2 = \sigma_z^2 = 1$$

$$p_+z_+^3 - p_-z_-^3 = \gamma \tag{47}$$

in which $z_- = |x_- - \mu|/\boldsymbol{\sigma}, z_+ = x_+ - \mu|/\boldsymbol{\sigma}$, and γ is the skew coefficient of the random variable X. The four unknowns in equation (47) can be obtained as

$$z_+ = \frac{\gamma}{2} + \sqrt{1 + \left(\frac{\gamma}{2}\right)^2}$$

$$z_- = z_+ - \gamma$$

$$p_+ = \frac{z_-}{z_+ + z_-}$$

$$p_- = 1 - p_+ \tag{48}$$

Once the two points in the standardized space are obtained, the corresponding points in the original space can be respectively determined from

$$x_- = \mu - z_-\boldsymbol{\sigma}$$

$$x_+ = \mu + z_+\boldsymbol{\sigma} \tag{49}$$

For a multivariate model involving n stochastically correlated parameters, $W = g(\mathbf{X}) = g(X_1, X_2, ..., X_n)$, the two-point representations for each of the variable are computed and permuted to form the 2^n points in the parameter space for model evaluations. Hence, the moments about the origin of model output, W, can be estimated as

$$\mu'_{W,m} = E(W^m) \approx \sum_{\delta_1=+,-} \cdots \cdots \sum_{\delta_n=+,-} p_{(\delta_1,\cdots,\delta_n)} w^m_{(\delta_1,\cdots,\delta_n)} \tag{50}$$

in which subscript, $\boldsymbol{\delta}_i$, I=1~n, is a sign indicator and can only be + or - representing the model parameter X_i having the value of x_{i+} or x_{i-}, respectively, as the two points locations obtained in the univariate case; $W_{(\delta^1,\cdots,\delta^n)}$ is the corresponding model output value evaluated at the selected point. If the stochastic model parameters are independent, the weighing factor, $p_{(\delta^1,\ldots,\delta^n)}$, is simply the product of the marginal probability masses from the univariate case. When correlation among the stochastic model parameters exists, a correction term is added to the product as

$$p_{(\delta_1,\cdots,\delta_n)} = \prod_{i=1}^{n-1} p_{i},\delta_1 + \sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n} \delta_i\delta_j a_{ij}\right) \tag{51}$$

in which p_{i+} and p_{i-} represent the probability masses at point locations x_{i+} and x_{i-}, respectively; a_{ij} is determined as

$$a_{ij} = \frac{\rho_{ij}/2^n}{\sqrt{\prod_{i=1}^{n}\left[1+\left(\frac{\gamma_i}{2}\right)^2\right]}} \tag{52}$$

where $\boldsymbol{\rho}_{ij}$ is the correlation coefficient between stochastic model parameters X_i and X_j. The m-th order central moment of W, $\mu_{W,m}$, can be obtained from the non-central moments, $\mu'_{W,m}$, by

$$\mu_{W,m} = \sum_{i=0}^{m}(-1)^i C_i^m(\mu_W)^i\mu'_{W,m-i} \tag{53}$$

where C_i^m=m!/[I!(m-I)!], a binomial coefficient and μ_W=E(W).

To yield each point estimate for the model output, the model W(**X**) must be evaluated once. From equation (50), one realizes that 2^n model evaluations are required for a model involving n stochastic model parameters. As the number of stochastic model parameters increasing, Rosenblueth's algorithm becomes computational less attractive.

3.1.2 Harr's Point Estimation Method - To circumvent the potential computational disadvantage of Rosenblueth's algorithm, an alternative probabilistic PE method was proposed by Harr (1989). Harr's algorithm ignores the skewness of the variables and, therefore, is theoretically appropriate for treating stochastic variables having normal distributions.

By Harr's algorithm, a hypersphere with radius $\sqrt{n}$ centered at the origin in the standardized parameter space is constructed for a model involving n correlated stochastic model parameters. The points for model evaluation are selected at the intersections of the hypersphere and the eigenvectors of the correlation matrix of the stochastic model parameters resulting in two intersections on each eigenvector. For problems involving n stochastic model parameters, the number of total points selected for model evaluations is 2n. Thus, the amount of computations using Harr's algorithm is significantly less than Rosenblueth's algorithm as n increases. To calculate the statistical moments of model output, the point estimates are weighted by the eigenvalues associated with the correlation matrix. Harr's probabilistic PE algorithm is summarized as the following:

1. The correlation matrix of the stochastic model parameters, $\mathbf{R}_X$, is decomposed as

$$\mathbf{R_X} = \mathbf{VLV}^t \tag{54}$$

where $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n]$ is the eigenvector matrix with $\mathbf{v}_1,\ \mathbf{v}_2, \cdots, \mathbf{v}_n$ being the column vectors of the eigenvectors; $\mathbf{L} = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ is the corresponding diagonal eigenvalue matrix with $\lambda_1, \lambda_2, \cdots, \lambda_n$ being the eigenvalues.

2. Along each eigenvector in the standardized parameter space, the coordinates of two intersection points on the hypersphere with a radius of $\sqrt{n}$, centered at the origin, is determined as

$$\mathbf{z}_{i\pm} = \pm\sqrt{n}\ \mathbf{v}_j,\ i = 1, 2, \cdots, n \tag{55}$$

where $\mathbf{z}_{i+}$ and $\mathbf{z}_{i-}$ is the column vectors containing the coordinates of two intersection points along the I-th eigenvector, $\mathbf{v}_i$, in the standardized parameter space. Then, the coordinates of the 2n intersection points in the original parameter space are obtained as

$$\mathbf{x}_{i\pm} = \boldsymbol{\mu} \pm \mathbf{D}^{1/2}\mathbf{z}_{i\pm},\ i = 1, 2, \cdots, n \tag{56}$$

where $\mathbf{x}$ is the column vector containing the coordinates of the intersection points in the original parameter space; $\boldsymbol{\mu}$ is the vector of the expected values of stochastic model parameters; and $\mathbf{D}^{1/2}$ is the diagonal matrix containing the standard deviations of the stochastic model parameters.

3. Based on the two points selected along each eigenvector, compute the corresponding model output values $w_{i\pm} = g(\mathbf{x}_{i\pm})$ for I=1 to n.

4. The m-th order moment about the origin for the model output can be computed as

$$\mu'_{W,m} = E(W^m) = \frac{\sum_{i=1}^{m} \lambda_i \overline{w_i^m}}{\sum_{i=1}^{n} \lambda_i} = \frac{\sum_{i=1}^{n} \lambda_i \overline{w_i^m}}{n} \tag{57}$$

where

$$\overline{w_i^m} = \frac{w_{i+}^m + w_{i-}^m}{2} \tag{58}$$

Then, the central moments can be obtained by equation (53).

3.2 Assessment of uncertainties of regionalized unit hydrograph

The stochastic model parameters involved in the determination of UH are N and K whose statistical characteristics can be obtained from the regional regression equations as illustrated in Section 3.1. Therefore, referring to the above two PE methods, n=2 and the model $W(\mathbf{X})$ is the IUH, U(t|N,K), by equation (15) or the DUH, U_m(N,K), by equation (16).

For a given watershed with known basin characteristics, depending on the regression procedure employed, the means and variances of N and K, and their correlation coefficient $\boldsymbol{\rho}$(N,K) can be obtained from the regional regression equations. Then, from the two PE algorithms, four points in the N-K parameter space are chosen for computing the IUH and/or DUH ordinates. According to equation (50) for Rosenblueth's approach or equation (57) for Harr's algorithm, the means and variances of U(t) and U_m can be computed.

To compute the correlation coefficients among ordinates in a IUH, the probabilistic PE methods can be applied to compute the expectation of the following model

$$W = g(N,K) = U(t)U(t') = \frac{1}{K^2\Gamma^2(N)}\left(\frac{tt'}{K^2}\right)^{N-1} c^{-\frac{t+t'}{K}}, t \neq t' \tag{59}$$

in which t and t′ represent different time instances. Once the expectation of U(t)U(t′) is obtained, the covariance and correlation coefficient of IUH ordinates at different times can be computed, respectively, as

$$\mathbf{Cov}[U(t),U(t')] = E[U(t)U(t')] - E[U(t)]EU(t')] \tag{60}$$

and

$$\boldsymbol{\rho}[U(t),U(t')] = \frac{\mathbf{Cov}[U(t),U(t')]}{\boldsymbol{\sigma}_{U(t)}\boldsymbol{\sigma}_{U(t')}} \tag{61}$$

in which E[U(t)] and $\boldsymbol{\sigma}_{U(t)}$ are the mean and standard deviation of U(t), respectively.

Similarly, to compute the correlation coefficient among ordinates of a DUH of a specified duration, the expectation of $U_m U_{m'}$ for $m \neq m'$, is calculated by the PE methods with U_m computed by equation (16). Then, the covariance and correlation coefficient of U_m and $U_{m'}$ can be obtained by equations (60) and (61), with U(t) and U(t′) replaced by U_m and $U_{m'}$, respectively.

4 Application

To illustrate the application of the two probabilistic PE methods to assess the uncertainty features of a IUH and DUH from a regional regression equations, a watershed was selected for which the uncertainties associated with its regional IUH and DUH were obtained. In this application, the results of uncertainty analysis affected by the three regional regression analysis procedures and the two PE methods were examined. The watershed selected in this application is the portion of Tan-Shui River upstream of Jei-Shou Bridge. The watershed characteristics are listed in Table 1 in Part 1.

Statistical moments of ln(N) and ln(K) from the various regional regression equations are shown in Table 7. The UVR0 and UVR1 have the identical values for the means and standard deviations of ln(N) and ln(K) except that UVR1 considers the sample correlation between the two parameters in the log-space. The statistical moments of N and K in their original space can be obtained by equations (35)-(37) and are tabulated in Table 7.

It is interesting to note that the values of N and K of the four points selected by the two probabilistic PE methods for uncertainty assessment of IUH and DUH as shown in Table 8. When N and K are treated as bivariate normal variables as in UVR1, MVRW, MVRWO, and SUR, the four points selected by Harr's and Rosenblueth's methods for uncertainty analysis are identical. Under the UVR0 in which N and K are considered as independent, the four points selected by the two PE methods are different with Rosenblueth's method giving its points in a narrower range. Under the analytical assumptions of lognormal distributions for N and K, the selected points by the two PE methods are different because primarily Rosenblueth's method considers

Table 7. Statistical moments of N and K for Tan-Shui watershed at Jei-Shou Bridge from different regional equations.

	$\mu_{\ln N}$	$\sigma_{\ln N}$	$\mu_{\ln K}$	$\sigma_{\ln K}$	$\rho_{\ln N,\ln K}$	μ_N	σ_N	μ_K	σ_K	$\rho_{N,K}$
UVRO	1.072	0.4322	0.5792	0.6035	0.0000	3.206	1.453	2.141	1.419	0.0000
UVR1	1.072	0.4322	0.5792	0.6035	-0.702	3.206	1.453	2.141	1.419	-0.557
MVRW	1.072	0.4322	0.5502	0.6327	-0.780	3.206	1.453	2.141	1.419	-0.604
MVRWO	0.875	0.4144	0.7081	0.5954	-0.782	2.613	1.131	2.424	1.581	-0 .622
SUR	0.988	0.4184	0.7081	0.5954	-0.756	2.931	1.282	2.424	1 .581	-0.602

Table 8. Points used in N and K parameter space by two PE methods for assessing IUH uncertainties.

	Point #1 (N,K)	Point #2 (N,K)	Point #3 (N,K)	Point #4 (N,K)
UVRO/H	(3.206, 4.148)	(3.206, 0.134)	(5.260, 2.141)	(1.151, 2.141)
UVRO/R	(2.465, 1.607)	(6.056, 1.607)	(2.465, 5.911)	(6.056, 5.911)
UVR1/H	(4.658, 0.722)	(1.753, 3.560)	(4.658, 3.560)	(1.753, 0.722)
UVR1/R	(2.465, 1.607)	(6.056, 1.607)	(2.465, 5.911)	(6.056, 5.911)
MVRW/H	(4.658, 0.632)	(1.753, 3.603)	(4.658, 3.603)	(1.753, 0.632)
MVRW/R	(2.465, 1.588)	(6.056, 1.588)	(2.465, 6.287)	(6.056, 6.287)
MVRWO/H	(3.744, 0.843)	(1.482, 4.005)	(3.744, 4.005)	(1.482, 0.843)
MVRWO/R	(2.019, 1.820)	(4.767, 1.820)	(2.019, 6.560)	(4.767, 6.560)
SUR/H	(4.213, 0.843)	(1.649, 4.005)	(4.213, 4.005)	(1.649, 0.843)
SUR/R	(2.262, 1.820)	(5.389, 1.820)	(2.262, 6.560)	(5.389, 6.560)

Note: H = Harr's PE method is used for estimating moments of IUH
R = Rosenblueth's PE method is used for estimating moments of IUH

the skewness of N and K whereas Harfs algorithm does not. Although Rosenblueth's method selects the four identical points for UVR0 and UVR1, the weighting functions in equation (51) for computing the moments are different. Therefore, the resulting mean and standard deviation of IUH and DUH will not be the same.

In the following discussions, illustrations are made using the IUH because the results for the DUH are similar. The means and standard deviations of IUH ordinates by the two PE methods with N and K obtained from different regression procedures are shown in Figures 2 and 5. Based on these figures, comparisons were made to evaluate the differences in the resulting IUHs by different methods under various considerations.

4.1 Comparison of UH's with parameters from different regional regression equations

Figures 2(a) and 2(b), respectively, are the means and standard deviations of UH obtained by Harr's PE method with N and K estimated from various regional regression equations assuming that N and K have a bivariate log-normal distribution. As can be seen, the mean IUH and the associated standard deviation obtained from the conventional univariate regression (UVR0), assuming independence between N and K, are quite different from those UHs considering their correlation. By Harr's method, consideration of correlation between N and K results in higher mean peak discharge for the IUI-I. Among the three regional regression equations that consider $\rho_{N,K}$, multivariate regression with intercept (MVRW) yields the highest peak discharge, followed by multivariate regression without intercept (MVRWO). Regional equations for N and K developed by the seemingly unrelated regression (SUR) results in a lower peak discharge for this test case. As for the associated standard deviation, the MVRWO, overall speaking, yields the lowest standard deviation than the other three regression equations considering correlation in N and K. The standard deviation of the IUH ordinates computed by Harr's algorithm shows a bimodal nature with N and K estimated from UVR1, MVRWO, and SUR during the earlier part of the IUH.

Figures 3(a)-(b) are similar to those in Figures 2(a)-(b) in that Rosenblueth's method were used to calculate the mean and-standard deviation of resulting IUHs. From Figure 3(a) one observes that the shape of mean IUHs for all regional regression equations are similar. The mean and standard deviation of the resulting IUHs at peak discharge by Rosenblueth's algorithm is about 1-1.5 cms/mm smaller than those obtained by Harr's algorithm. The UVR0, in contrast to Figure 2(a), results in the highest peak than all other regional equations which account for correlation in N and K. The relative position of IUHs among the four regional equations considering $\rho_{N,K}$ is about the same as those by Harr's method. There also exists greater similarity in temporal variation in standard deviation among different regional equations by Rosenblueth's method as shown in Figure 3(b).

4.2 Comparison between two probabilistic point estimate methods

For this particular test case, a glance over Figures 2-3 reveals that, when $\rho_{N,K}$ is considered in regional regression equations, Harr's method yields mean IUHs with higher peak discharge than those by Rosenblueth's. A more detailed comparison between

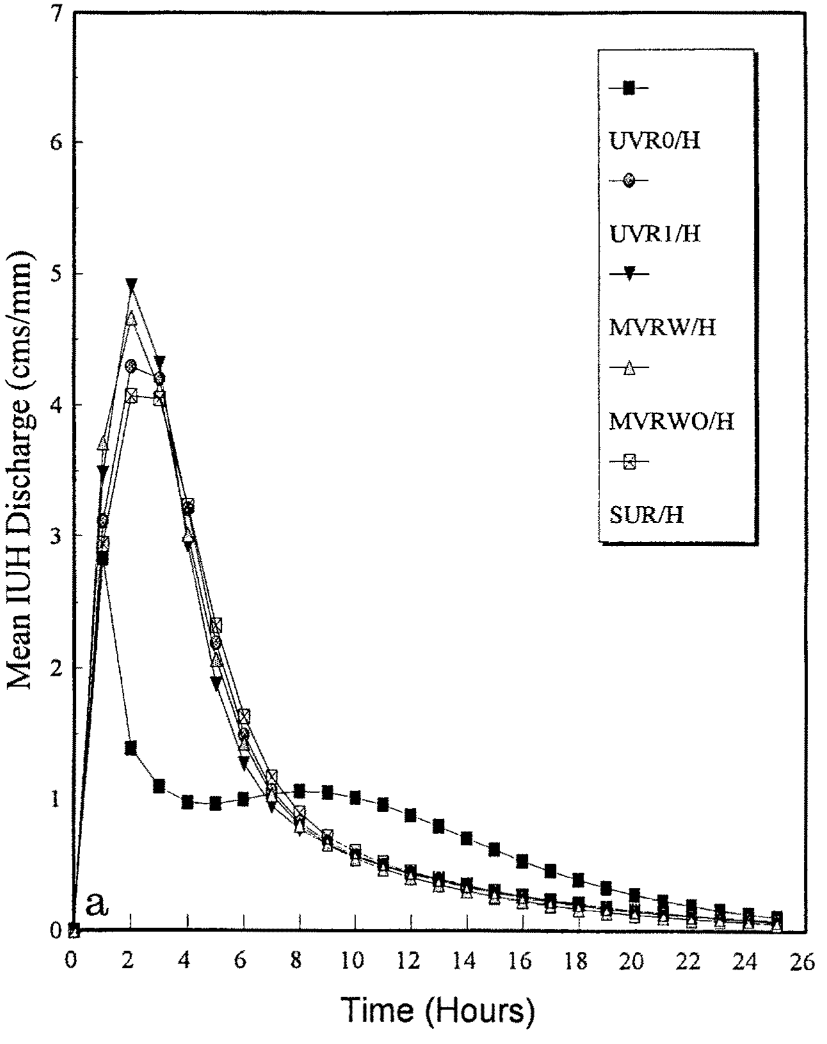

Figure 2a. Comparison of mean IUHs by Harr's method with N and K estimated from various regression procedures.

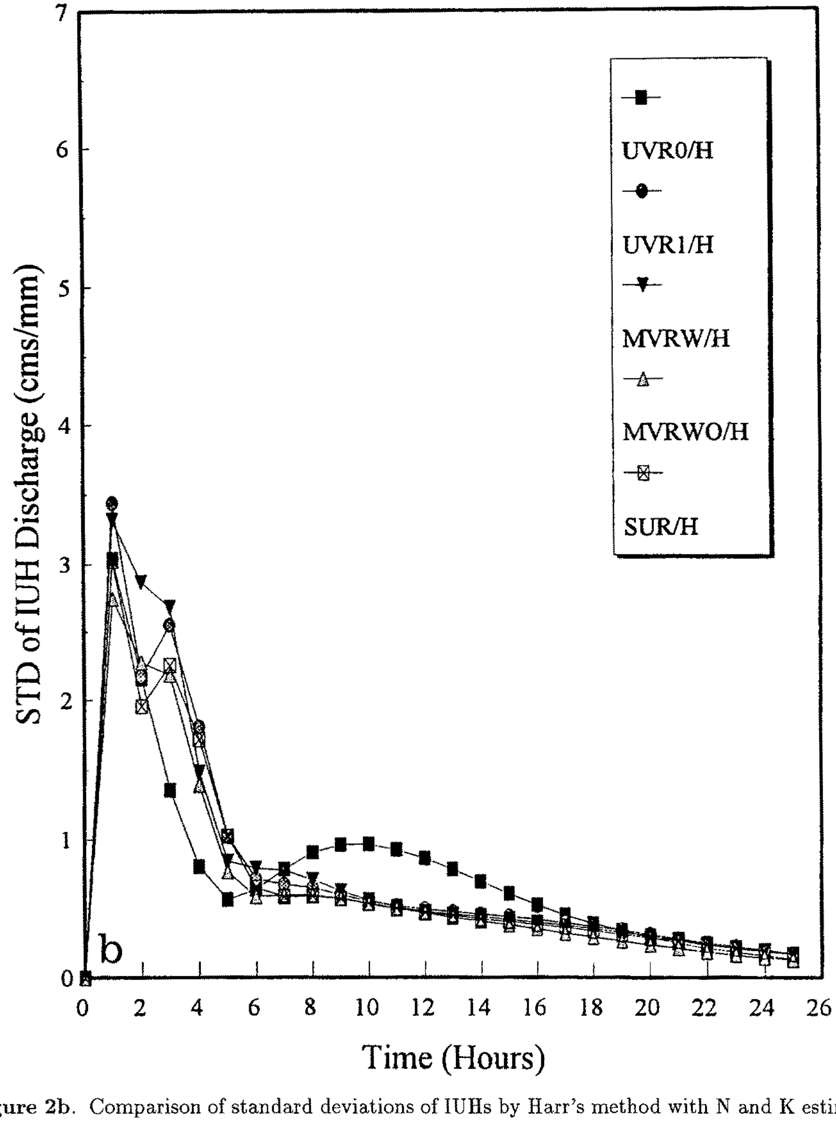

Figure 2b. Comparison of standard deviations of IUHs by Harr's method with N and K estimated from various regression procedures.

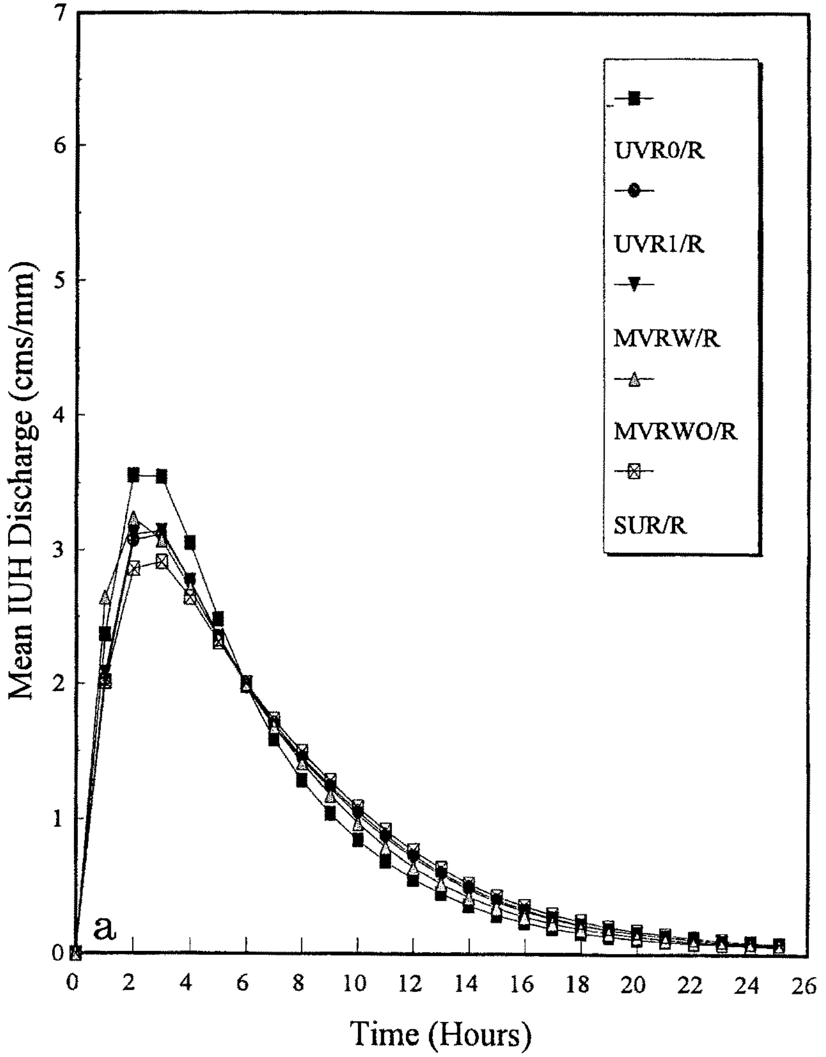

Figure 3a. Comparison of mean IUHs by Rosenblueth method with N and K estimated from various regression procedures.

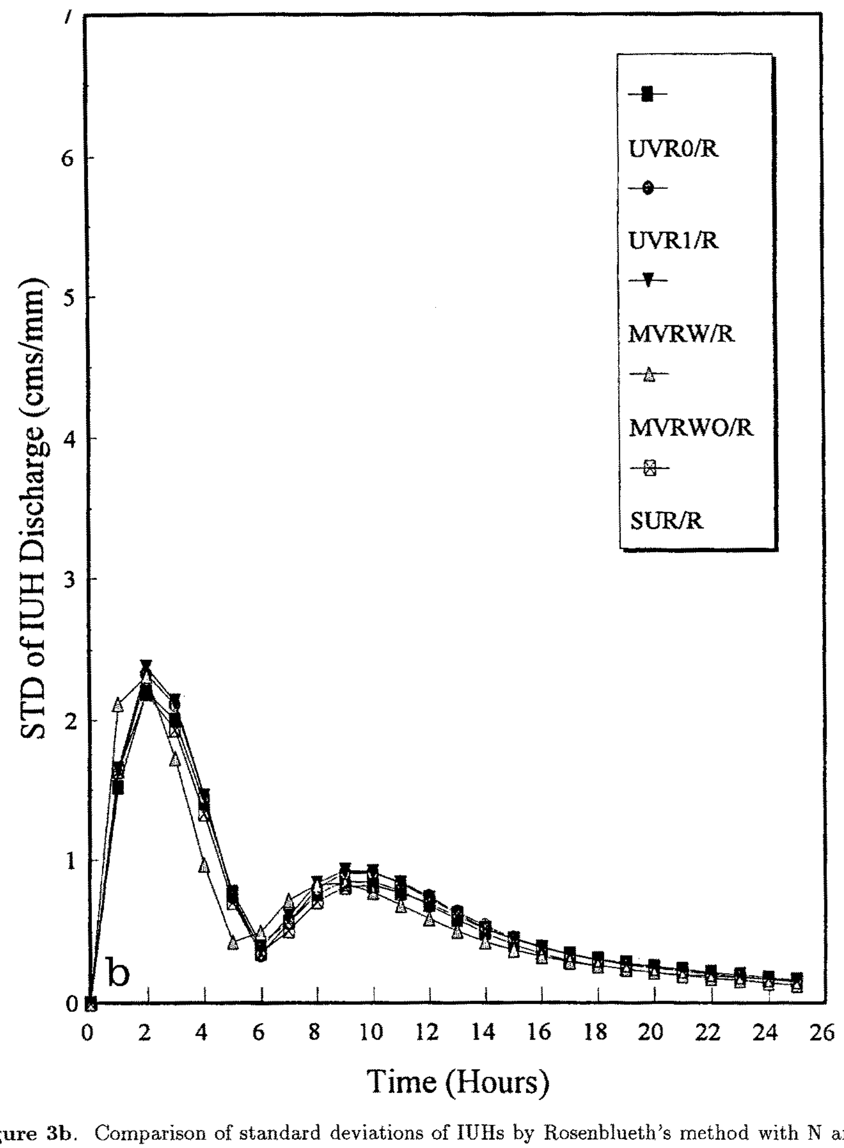

Figure 3b. Comparison of standard deviations of IUHs by Rosenblueth's method with N and K estimated from various regression procedures.

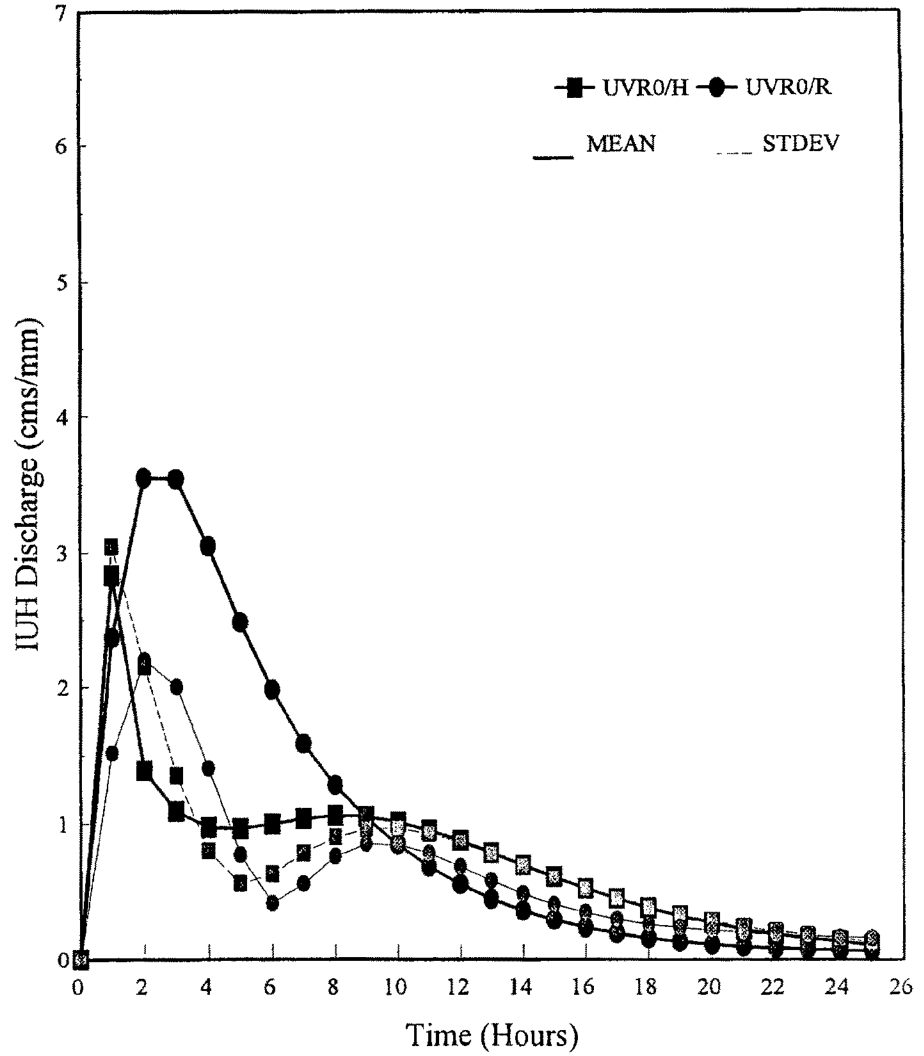

Figure 4. Comparison of means and standard deviations of IUHs by different PE methods with N and K estimated by UVRO.

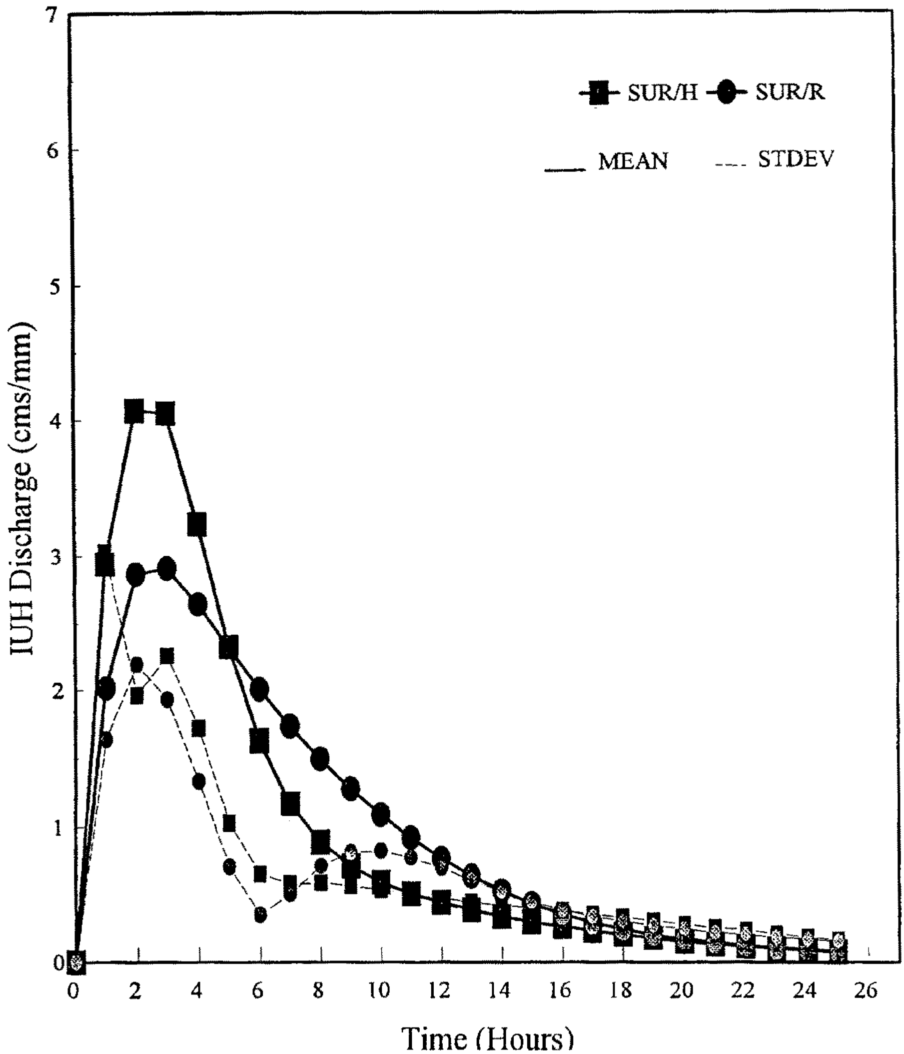

Figure 5. Comparison of means and standard deviations of IUHs by different PE methods with N and K estimated by SUR.

the two PE methods by the regional equations are shown in Figures 4-5. Except for the UVR0, Harr's method results in higher mean peak discharge and larger standard deviation for regional equations considering $\rho_{N,K}$.

The consistency in lower mean peak and standard deviation for the IUH by Rosenblueth's method is attributed to the underlying difference in point selection for model evaluation between the two PE methods. Recall that Harr's method preserves only the first two moments of N and K implying that they are bivariate normal random variables. On the other hand, Rosenblueth's method preserves the first three moments of model parameters. Under the assumption that N and K are bivariate log-normal random variables, Rosenblueth's method could account for the positive-skewed nature of N and K marginally. The nonlinear relationship between the UH ordinates and the model parameters, along with the presence of correlation, makes the explanations of such behavior difficult. Based on the recent study by Chang et al. (1995), the ability to incorporate the skew coefficients of involved random variables by Rosenblueth's method will improve the accuracy of uncertainty analysis.

To improve the competitiveness of Harr's method in this particular application, the selection of points for model evaluations can be made in the log-transformed space in which both ln(N) and ln(K) are bivariate normal random variables. Under a bivariate standardized normal parameter space, both Harr's and Rosenblueth's methods will select four identical points for model evaluations and the skew coefficient of lognormal random variables in the original space can be implicitly accounted for. Due to different weights are used in computing the statistical moments of model response (see equations (50) and (57)), the resulting uncertainty features of model response computed by the two probabilistic PE methods will be different. However, the numerical experiences have indicated that the differences in estimated statistical moments between the two PE methods, when they are placed on the same common ground, will be negligible (Chang, 1994).

5 Summary and conclusions

Due to the existence of uncertainties in regional equations, the parameters N and K in Nash's model cannot be predicted with absolute certainty. Consequently, the IUH and DUH derived based on the uncertain N and K are also subject to uncertainty. From the regional regression study, relevant statistical information can be readily incorporated into the assessment of uncertainty associated with the IUH and DUH of a watershed under consideration. To quantify the uncertainty features associated with the IUH and DUH by Nash's model, two probabilistic point estimate methods were applied and their relative performance compared.

Results from the uncertainty analysis indicate that the effect of uncertainty in N and K from the regional equations on the uncertainty of IUH and DUH is significant and cannot be ignored. The correlations among the dependent variables in the regional regression equations may have profound influence on the results of uncertainty analysis and should be considered accordingly. This effect is clearly observed in this study which shows that, around the peak, the mean and the standard deviation of IUH obtained from the conventional univariate regression procedure is significantly higher. The consideration of correlation coefficient between N and K, which is negative, leads to lower mean peak discharge and standard deviation. Therefore, using an IUH without considering the correlation between N and K could lead to a conservative design of a hydraulic structure or over-estimation of failure probability of an existing structure.

The paper demonstrates the applicability of two probabilistic point estimation methods to uncertainty analysis of hydrologic models. Rosenblueth's method has the theoretical advantage over Harr's for its ability to account for the skew coefficients of involved random variables, which often leads to improving the accuracy of uncertainty analysis. On the other hand, Harr's method is more computationally viable to handle problems involving large number of random variables. To enhance the accuracy of Harr's method, one possible way is to transform the non-normal random variables to their equivalent normal spaces (Liu and Der Kiureghian, 1986; Chang et al., 1994) from which the selection of points for model evaluations is made.

As a last note, the uncertainty analysis conducted herein considers only the parameter and model uncertainties associated with regional regression equations. The main purpose is to illustrate how the information from a regression analysis can be readily incorporated and used to quantify uncertainty features of a dependent variable. In case one wishes to further consider the uncertainty of independent variables, then every terms in a regression model are random variables and the classical results such as equation (34b) is no longer appropriate. In this circumstance, one practical approach is to apply an appropriate probabilistic point estimation method or other techniques to assess the uncertainty of the dependent variable.

Acknowledgments

The study was funded by the Agricultural Council of the Executive Yuan, Taiwan, Republic of China. The authors are grateful to Mr. Wen-Zhang Hu of the Agricultural Council for his support and encouragement. In addition, comments and criticisms by the two reviewers are greatly appreciated.

References

Chang, C.-H. 1994: Incorporating information of non-normal distribution in uncertainty analysis of hydrosystems. Ph.D. Dissertation, Department of Civil Engineering, National Chiao-Tung University, Hsinchu, Taiwan, Republic of China. 166pp

Chang, C.-H.; Tung, Y.-K.; Yang, J.-C. 1994: Mote carlo simulation for correlated variables with marginal distributions. Journal of Hydraulic Engineering, ASCE 120(2), 313-331

Chang, C.-H.; Tung, Y.-K.; Yang, J.-C. 1995: Evaluating performance of probabilistic point estimates methods. Journal of Mathematical Modelling 19(2), 95-105

Harr, M.E. 1989: Probabilistic estimates for multivariate analyses. Applied Mathematical Modelling 13, 313-318

Karmeshu; Lara-Rosano, F. 1987: Modelling data uncertainty in growth forecasts. Applied Mathematical Modelling 11, 62-68

Liu, P.-L.; Der Kiureghian, A. 1986: Multivariate distribution models with prescribed marginals and covariances. Probabilistic Engineering Mechanics 1(2), 105-112

Mays, L.W.; Tung, Y.K. 1992: Hydrosystems Engineering and Management, McGraw-Hill Book Company, New York, N.Y.

Rosenblueth, E. 1975: Point estimates for probability moments. Proceedings, National Academy of Science 72(10), 3812-3814

Rosenblueth, E. 1981: Two-point estimates in probabilities. Applied Mathematical Modelling 5, 329-335

Tung, Y.K.; Yen, B.C. 1993: Some recent progress in uncertainty analysis for hydraulic design. In: Reliability and Uncertainty Analyses in Hydraulic Design, edited by B.C. Yen and Y.K. Tung. 1734. ASCE

Tung, Y.K. 1994: Probabilistic hydraulic design: a next step to experimental hydraulics. d. of Hydraulic Research, IAHR 32(3), 323-336

Tung, Y.K.; Yeh, K.C.; Yang, J.C. 1995: Regionalization of unit hydrograph parameters: 1. Comparison of regional regression methods. (Accompanyingpaper under review)

Yeh, K.C.; Tung, Y.K., 1993: Uncertainty and sensitivity of a pit migration model. d. of Hydraulic Engr., ASCE 119(2), 262-281

Yeh, K.C.; Tung, Y.K.; Yang, J.C.; Zhao, B. 1993: Uncertainty analysis of hydrologic models and its implications on reliability of hydraulic structures (2). Agricultural Council, Executive Yuan, Taiwan

Zhao, B.; Tung, Y.K.; Yeh, K.C.; Yang, J..C. 1995: Reliability analysis of hydraulic structures considering unit hydrograph uncertainty. Journal of Stochastic Hydrology and Hydraulics (Accepted).