Stochastic Evaluation of Rainfall-Runoff Prediction Performance

By T. V. Hromadka II,[1] Member, ASCE

Abstract: Given a set of realizations of error data (i.e., the difference between model runoff estimates and stream gauge data) from rainfall-runoff hydrologic models, it is possible to generate a set of error transfer function realizations that, when convoluted with a suitable kernel function such as the hydrologic model output, equate to the original error data. In turn, these error transfer function realizations may be used to generate synthetic error data that is convolved from a separate design storm modeled runoff and the generated error transfer function realizations. The synthetic error data set is then added to the design storm modeled runoff to produce a set of equally likely outcomes for the model prediction. The set of equally likely outcomes is statistically analyzed to provide, for instance, a confidence interval for the possible outcomes of the design storm model. A four-section algorithm is presented that performs each of these tasks.

INTRODUCTION

The classification of rainfall-runoff hydrological models is well known in the literature. However, a uniform procedure to evaluate rainfall-runoff model performance, in prediction mode, has not been generally adopted. In the present paper, an algorithm is presented that provides a means for evaluating the relative performance between rainfall-runoff model structures as applied in prediction mode. Before beginning the mathematical development, a preview is presented.

Let X be a random variable for which we wish to estimate future outcomes. Various estimates of X are available, namely, outputs from model types Y_1, Y_2, and so forth. We wish to compare the models Y_k as to prediction performance.

To begin, we choose one particular model Y_k, say model Y_1. Next, for m historic events, we have m ordered pairs of data $\{(x^i, y_1^i); i = 1, 2, \ldots, m\}$, where x^i is the outcome of X for event i, and y_1^i is the outcome (model prediction) of Y for event i. The error for event i, for model 1, is e_1^i, where $e_1^i = x^i - y_1^i$. For m events, there are m error values (or realizations) for model 1, given by the set $\{e_1^i; i = 1, 2, \ldots, m\}$, or simply noted as the distribution of errors $[e]_1$ for model Y_1. The error realizations employed to characterize the distribution of modeling error, for a particular model, will be used as a sampling from a stochastic process.

Suppose we wish to predict the outcome of X, for a future event, d, noted as X_d. Then, an estimate of X_d is given, for model type 1, by the distribution of values of Y_1, for future event d, noted as $[Y]_1^d$ where

$$[Y]_1^d = y_1^d + [e]_1 \tag{1}$$

where $[Y]_1^d$ = set of m estimates for X_d, from model 1, for event d; y_1^d = predicted outcome for event d from model Y_1; and $[e_1]$ = set of historic error values, from model Y_1, given estimate y_1^d.

A similar approach is used to produce a set of equally likely outcomes for a rainfall-runoff model structure. Eq. (1) can be extended to the evaluation of storm runoff criterion variables (e.g., peak flow rate, peak three-hour runoff volume, sediment transport mean flow velocity for flow rates above a threshold, and so forth). The algorithm, which is shown in Fig. 1, uses results obtained by applying the theory of stochastic integral equations (Tsokos and Padgett 1974). Each rainfall-runoff model is used to simulate, in prediction mode, four separate storm events. That is, each rainfall-runoff model is first calibrated to a set of storm data, and then the calibrated models are applied to a set of four storms that were not part of the original calibration set of storms (i.e., "validation" test runs). For the case study (Abbott 1978), a nearly full urbanized small catchment in northern California was studied, where negligible runoff storage effects existed, and where good quality rainfall and runoff data are available.

To develop error realizations, for each particular model in prediction mode, the graphical output of each model (resulting from the validation test runs) is compared to observed stream gauge data at discrete time intervals and the difference between the two realizations is recorded as modeling error, in prediction mode, for that particular model, for the particular storm event. Inherent in this approach is the assumption that the obtained error realizations are representative of a simple random sampling of the underlying stochastic process. Additionally, the error realizations are dependent upon the runoff model used; hence, a comparison in the variance of modeling estimates can be made between model types, similar to the evaluation of statistical estimators according to their respective variances. Furthermore, if sufficient data sets are available, one can partition the data according to storm size, for example, in order to investigate modeling trends with respect to storm magnitude. The error realizations are then equated to a convolution of the model output with a new transfer function called the error transfer function. By solving the inverse problem, the error transfer function realization is evolved on a two-hour time interval basis for each model structure studied.

Bootstrapping directly from the error series assumes that the error intervals are independent of the class chosen, and not necessarily unique to any particular model prediction interval within the chosen class. The prior approach for generating transfer functions is one of four techniques mathematically developed in Hromadka and Whitley (1989). The approach assumes that there exists a relationship between the error series and the subject model predictions, and that this relationship can be represented by the stochastic integral equation (SIE), where each realization is equally likely. This chosen SIE also includes the assumptions that (1) the model error at time t is related to events prior or equal to time t and is not influenced by events following time t; and (2) the transfer function realizations are all equally likely on a class basis, where the total set of transfer functions may be partitioned into classes, such as according to model prediction magnitude, rainfall magnitude, or other criteria. In the current work, it is assumed that

[1]Sr. Managing Engr., Failure Analysis Associates, 4590 MacArthur Blvd., Ste. 400, Newport Beach, CA 92660-2027; Prof. of Math. and Envir. Studies, Dept. of Math., California State Univ., Fullerton, CA 92634-9480.

Note. Discussion open until March 1, 1998. 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 November 20, 1996. This paper is part of the *Journal of Hydrologic Engineering*, Vol. 2, No. 4, October, 1997. ©ASCE, ISSN 1084-0699/97/0004-0188–0196/$4.00 + $.50 per page. Paper No. 11432.

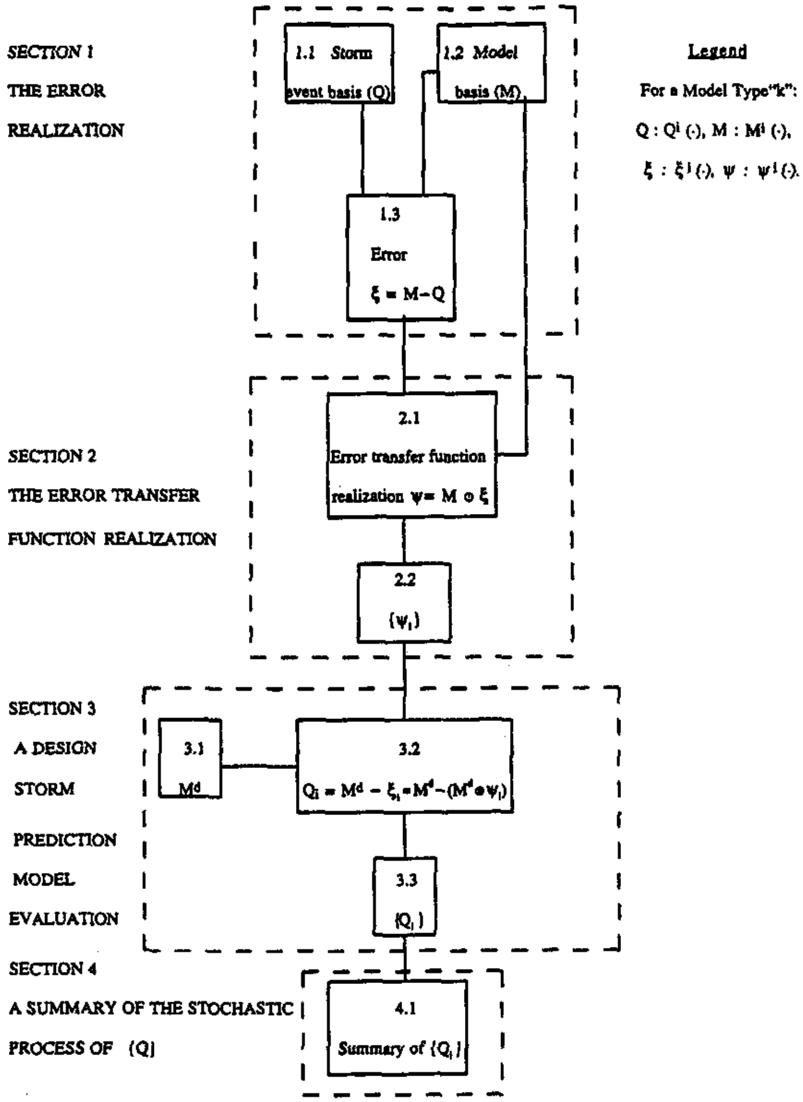


FIG. 1. Algorithm for Statistically Evaluating Rainfall-Runoff Model Prediction Performance

all of the available runoff data and model results are from the same class, and that the associated errors of model estimates are as likely to occur for one particular hydrograph as another. If the data were partitioned, then subsets of the error realizations would be deemed of the same class, and a different stochastic distribution would result (analogous to partitioning height measurements according to classes of male or female). This SIE technique attempts to explain how very similar intervals of measured storm rainfall have corresponding modeled runoff hydrograph intervals that are also very similar, yet the corresponding measured runoff hydrograph intervals differ significantly.

Two other techniques (suggested by this paper's reviewers) for generating a larger population of transfer functions, using a bootstrapping method, are

1. Assume errors are related to model predictions. Approximate this relationship using the SIE and synthesize transfer function realizations. Assuming these transfer function realizations are independent, bootstrap from them.
2. Assume errors are unrelated to model predictions, and bootstrap from the error series directly to develop a set of realizations.

A major goal in rainfall-runoff model studies is to predict the runoff hydrograph corresponding to some critical hypothetical storm event. This critical hypothetical storm event, called the design storm, can be evaluated by convoluting the resulting design storm model output (for a chosen model type) with the set of error transfer functions that correspond to the chosen model. This generates a new set of error realizations that are subtracted with the design storm model output to create a set of equally likely outcomes, which may then be statistically evaluated, analogous to the discrete case of (1).

This approach will be demonstrated in an example application to a series of actual data sets (Abbott 1978; Hromadka et al. 1987) to determine the error realizations in each of six hydrologic models, for a set of measured rainfall-runoff storm events in an urban catchment. The six models evaluated were (1) Continuous Flood Hydrographs (HEC-1); (2) Storm Water Management Model (SWMM); (3) Massachusetts Institute of Technology Catchment Model (MITCAT); (4) Storage Treatment Overflow Runoff Model (STORM); (5) Hydrocomp Simulation Program (HSP); and (6) Streamflow Synthesis and Reservoir Regulation (SSARR). [Descriptions of each of these models are contained in the rainfall-runoff model study prepared by Abbott (1978)].

The algorithm used to obtain the set of equally likely design storm model outcomes is comprised of four sections. The Error Realization Section describes how the error realizations are obtained from the model output (i.e., in prediction mode), and the observed stream gauge information. The Error Transfer Function Realizations section describes how numerical approximation to the Volterra integral is used to derive the error transfer function realizations. The Design Storm Prediction Model Evaluation section describes how a design storm model is analyzed using the set of error transfer function realizations obtained in the Error Transfer Function Realizations section. The result is in the section that covers the statistical analysis, giving a set of equally likely runoff hydrograph outcomes, on an assumed storm class basis.

MATHEMATICAL MODEL DEVELOPMENT

Error Realization

To develop a simple random sample of error realizations, measured runoff data are subtracted from model produced runoff hydrographs (from validation tests). In the present paper, due to the small sample size of only four validation tests, an assumption of independence of model error on a two-hour time interval basis is introduced in order to build a synthetically larger sample size. The choice of the two-hour window is based on the assumption that the final one-hour time interval of runoff only depends on effects during the final and prior one-hour time interval, and is due to a T_c of about one hour.

Fig. 1 shows how the observed stream gauge data (1.1) and the model output (1.2), for a given storm event, are subtracted to form the error realization (1.3) for the particular model under investigation. The error realization is denoted by $\xi_k^i(\cdot)$, where $i = 1, 2, 3, \ldots, n$ is a particular two-hour interval; $k = 1, 2, 3, \ldots, 6$ is the model under investigation, and $(\cdot)$ is time. As shown in step 1.3, the error realization is obtained by subtracting the observed stream gauge data $Q_k(\cdot)$ at a particular time for the cumulative storm from the corresponding model output $M_k(\cdot)$ at that time, resulting in $M_k(\cdot) - Q_k(\cdot)$. Observed low flows measured by the gauge were not included in order to preserve the assumption of mutually independent realizations on a class basis, and so model base flow estimates are not included in the analysis.

For use in this algorithm, the four validation storm events for each of the six models are combined in time sequence to form one large combined storm event realization, $M_k(\cdot)$, where k is the model type. The extended time frame is then divided into two-hour segments. Each two-hour segment is separated into 24 five-minutes intervals. The motivation for using two-hour segments is to synthetically develop a larger population of modeling error realizations; then a bootstrapping procedure is used to simulate random sampling. The choice of two-hour segments is based upon the hydrologic response time (i.e., time of concentration, T_c) of the study catchment being approximately one hour, and hence the last one-hour runoff hydrograph interval within the two-hour segment may be mutually independent for this analysis.

Error Transfer Function Realizations

The second section of the algorithm depicts how an error transfer function realization $\psi_k^i(\cdot)$ is synthesized by using the inverse of the convolution procedure (Hromadka and Whitley 1989). This concept is similar to evolving a unit hydrograph given rainfall excess (rainfall less losses) and a runoff hydrograph, such as can be accomplished using HEC-1. The convolution procedure is based on the stochastic integral equation

$$\xi_k^i(t) = \int_{s=0}^{t} M(t - s)\psi_k^i(s)\, ds \tag{2}$$

where s and t span the two-hour time intervals; and $M(\cdot)$ = two-hour duration segment of hydrologic model output under investigation. Although (2) defines the transfer function as determined from modeling estimates, an underlying assumption is that the resulting synthesized transfer functions are equally likely realizations from a stochastic process. It is noted that the synthesized transfer function is uniquely determined and has both positive and negative values (Hromadka and Whitley 1989), as long as $M(\tau)$ and $\xi_k^i(\tau)$ begin to have nonzero values simultaneously. For example, Tsokos and Padgett (1974) develop in detail the use of stochastic integral equations to model the uncertainties inherent in biological processes as a Volterra integral representation of data obtained by instrumentation and other measurements. In our case, the uncertainty in runoff estimates is modeled as a Volterra integral representation of rainfall-runoff modeling estimates. The Volterra integral may be approximated by use of the discrete convolution method. The convolution method for the algorithm is

$$\xi_k^i(\cdot) = M_k^i(\cdot) \otimes \psi_k^i(\cdot) \tag{3}$$

where $\otimes$ = notation for convolution process of (2). In (3), the error realization $\xi_k^i(\cdot)$ is known from the supplied data, and the hydrologic model output $M_k^i(\cdot)$ is known. As before, $i = 1, 2, 3, \ldots, n$. The underpinning of (3) is to represent a particular rainfall-runoff model's history of performance (i.e., modeling error) as a stochastic process. To describe the stochastic process, a discrete set of realizations are used as a simple random sample, which is then bootstrapped to evolve a synthetic population. Obviously, should a significant data supply exist, those data should be used instead of synthetic data. However, it is recalled that only the errors resulting from validation tests should be used in (1)–(3).

To solve for the error transfer function $\psi_k^i(\cdot)$, it will be necessary to rewrite (3). For notational purposes, we will use the figure $\odot$ to denote the inverse of the convolution procedure, so that

$$\psi_k^i(\cdot) = M_k^i(\cdot) \odot \xi_k^i(\cdot) \tag{4}$$

where $\xi_k^i(\cdot)$ = modeling error realization developed in (3). By numerically integrating (4) with respect to a two-hour duration and five-minute unit intervals, (4) may be solved using linear algebra by restating $M_k^i(\cdot)$ as a square (24 × 24) matrix and $\xi_k^i(\cdot)$ as a (24 × 1) column vector. The situation where there is no solution to the inverse problems occurring when the first model value input is zero is handled within the computer program that was created. The resultant error transfer function realization $\psi_k^i(\cdot)$ will be in the form of a (24 × 1) column vector where $i = 1, 2, 3, \ldots, n$. For $i = 1$, the first two-hour duration segment of the combined $M_k^i(\cdot)$ is analyzed, resulting in the error transfer function realization $\psi_k^i(\cdot)$. For $i = 2$, the time interval is shifted to encompass the interval starting at $t = 5$ and ending at t = two hours plus five minutes. Again, the extracted error $\xi_k^2(\cdot)$ is used to determine the error transfer function realization $\psi_k^2(\cdot)$. For $i = 3$, the time base is again shifted to span the time interval starting at $t = 10$ minutes and ending at t = two hours plus 10 minutes. The process of incrementing the time base by five minutes and extracting the error transfer function realization for the included two-hour segment is repeated for the length of the combined storm duration, $M_{k1}(\cdot)$. The result will be set $\{\psi_k^i\}$ of synthetic, equally likely, error transfer function realizations.

To expand the discrete set of transfer functions into a larger sampling domain, bootstrapping (Efron and Tribshirani 1993) is used. The bootstrap method consists of randomly selecting a transfer function from the set of transfer functions, with replacement. Selection with replacement means that each time a transfer function is selected, the selection is made from the entire set, regardless of how many times each function has been selected previously; i.e., each function has the same probability of being selected. Naturally, bootstrapping sampling is a weak substitute for actual data. However, in the present paper bootstrapping provides a synthetic population for use in illustrating the statistical analysis.

Design Storm Prediction Model Evaluation

To develop a distribution of runoff hydrograph possibilities, given the history of a particular rainfall-runoff model's performance, for a hypothetical design storm event, the stochastic integral equation is applied to the model output. This section of the algorithm incorporates the set of error transfer function realizations developed in the previous section. These transfer function realizations will be used to evaluate a rainfall-runoff model prediction, called the design storm model M_k^d (see step 3.1).

By rearranging the formulation shown in step 1.3

$$[Q(\cdot)]_k^d = M_k^d(\cdot) - [\xi(\cdot)]_k^d \tag{5}$$

As in step 3.2, $[Q]_k^d$ = distribution of runoff hydrographs, for model type k, for design storm event d; $[\xi(\cdot)]_k^d$ = distribution of error realizations, for model type k, in prediction mode; and $M_k^d(\cdot)$ = model type k runoff hydrograph, for design storm event, d. For evaluation purposes, the error distribution for the design storm model will be estimated by the stochastic process, in discrete form

$$[\xi(\cdot)]_k^d = M_k^d(\cdot) \otimes \{\psi_k^i(\cdot)\} \tag{6}$$

where $\{\psi_k^i(\cdot)\}$ = set of error transfer function realizations obtained in the previous section. Eq. (6) generates an ensemble of error realizations $[\xi(\cdot)]_k^d$, which are based on the error transfer function realizations developed for model type k discussed in the two previous sections, and here are convoluted with the design storm model output, $M_k^d(\cdot)$. For the design storm model, this step constitutes the transition from a single storm observation to a discrete statistical sample space capable of statistical analysis, described in the following section.

Eq. (5) may then be rewritten in terms of a stochastic process as

$$[Q(\cdot)]_k^d = M_k^d(\cdot) - [M_k^d(\cdot) \otimes \{\psi_k^i(\cdot)\}] \tag{7}$$

for $i = 1, 2, 3, \ldots, n$, where $[Q(\cdot)]_k^d$ = stochastic process of equally likely outcomes for design storm model d, and, where as before, $k = 1, 2, 3, \ldots, 6$.

Statistical Analysis of Stochastic Process $[Q(\cdot)]_k^d$

The output of the algorithm, $[Q(\cdot)]_k^d$, consists of a discrete distribution of equally likely outcomes for the particular design storm prediction under consideration. These stochastic results can be analyzed using traditional statistical methods. Since we assume each outcome is equally likely, statistical methods may be used to obtain useful information from the results of the algorithm, such as

1. The maximum value of the model output (i.e., peak flow rate)
2. The expected value of the peak flow rate
3. The median value of the peak flow rate estimates
4. The variance of the peak flow rate estimates (the variance is particularly interesting, since it calls attention to the distribution, or spread, of the outcomes of each model under consideration
5. Confidence intervals, which would establish an interval that includes the true value of the peak flow rate with a predetermined degree of certainty
   Total runoff volume

Other statistics may be similarly analyzed, such as runoff volume and depth-duration characteristics, among others. The algorithm may be used to generate statistical data for each of the k models under consideration. Since the data are based on the error realizations obtained in the Error Realization section, the output of the algorithm will reflect the actual discrepancy between the observed stream gauge data and the model under study, when applied in prediction mode. (Of course, an additional increase in variance will occur in transposing the model sets to watersheds that are not part of the rainfall-runoff data set.)

By using the algorithm, it is possible to rank the six models under evaluation with respect to a statistical measure, such as the mean, for each $[Q(\cdot)]_k^d$.

APPLICATION

As described in the previous development, the actual rainfall-runoff model output is used in the stochastic integral equation in developing the stochastic process of runoff hydrographs in prediction mode. It is recalled that error transfer function is defined by a set of equally likely realizations synthesized by equating a particular rainfall-runoff model output, in prediction mode (i.e., in this case, from validation tests), to the modeling error. That is, once the rainfall-runoff model is calibrated, the model is tested using known rainfall-runoff data that were not part of the calibration data set, resulting in realizations of modeling error for the given model in prediction mode. This approach describes the modeling error trends more appropriate to those applied in practice, rather than modeling error in matching its own calibration data.

To compare the six previously cited rainfall-runoff models in relative performance, three hypothetical design storm runoff hydrographs are defined, labeled as A, B, and C. Storms A, B, and C correspond to about the 100-, 5-, and 25-year design storm events, respectively. For output A, as an example, it is assumed that all six rainfall-runoff models have produced an output realization (runoff hydrograph) that is then used in each of the respective stochastic integral representations to develop the distribution of runoff predictions.

As described before, six rainfall-runoff hydrological models, in prediction mode, were each compared to observed stream

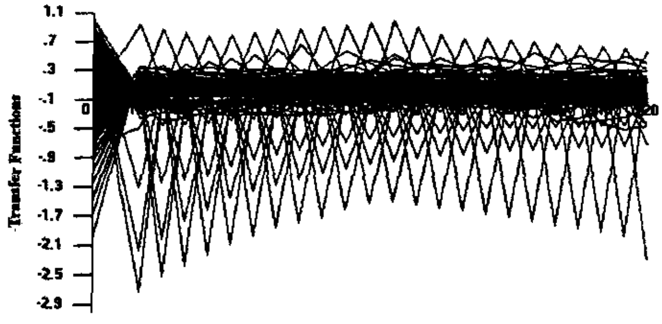


FIG. 2(a). Stochastic Process of Error Transfer Functions—Rainfall-Runoff Model 1

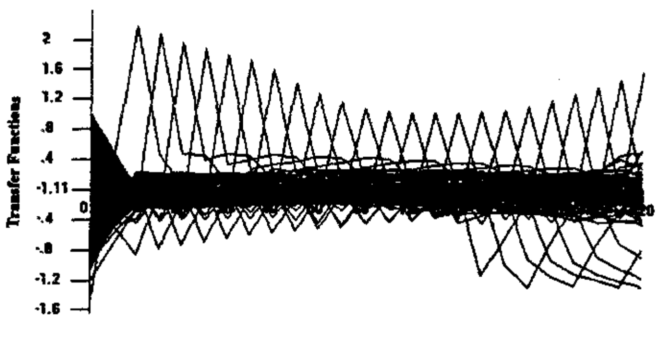


FIG. 2(b). Stochastic Process of Error Transfer Functions—Rainfall-Runoff Model 2

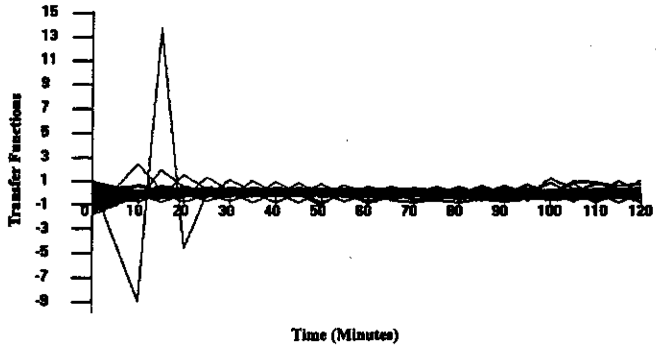


FIG. 2(*c*). Stochastic Process of Error Transfer Functions—Rainfall-Runoff Model 3

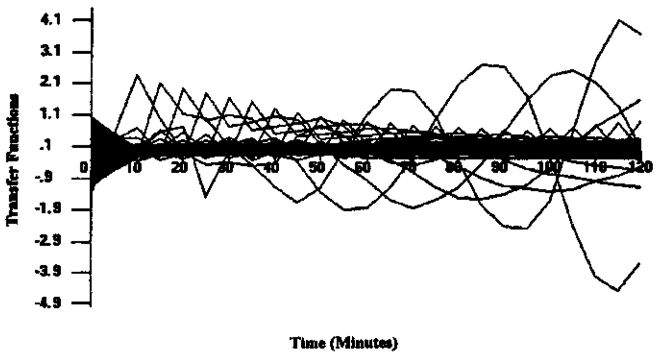


FIG. 2(*d*). Stochastic Process of Error Transfer Functions—Rainfall-Runoff Model 4

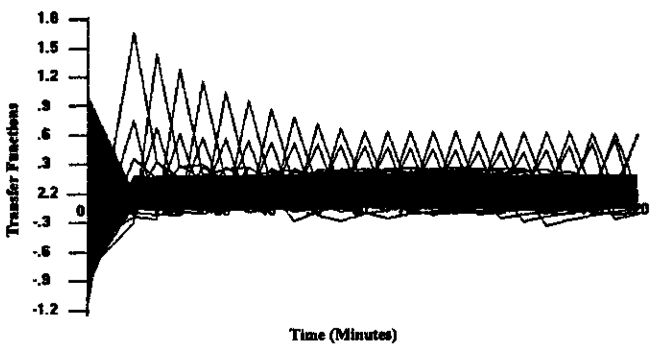


FIG. 2(*e*). Stochastic Process of Error Transfer Functions—Rainfall-Runoff Model 5

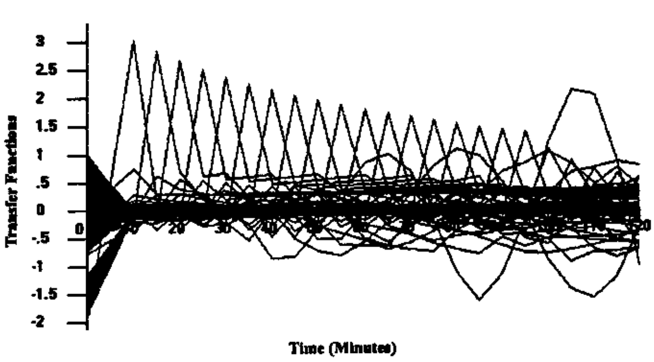


FIG. 2(*f*). Stochastic Process of Error Transfer Functions—Rainfall-Runoff Model 6

gauge data (from the model validation data set) at discrete time intervals, and a set of error realizations were obtained for each of the six models. A set of error transfer function realizations $\{\psi_k^i\}$ were obtained for each of the six models, $k = 1, 2, \ldots, 6$. The results may be seen in Figs. 2(a)–2(f), where the error transfer function realizations $\{\psi_k^i\}$ are plotted for the chosen two-hour time durations.

Each of the six hydrological models was used in conjunction with the three design storms to generate a set of probable design storm runoff hydrographs, as shown in Figs. 3(a)–3(f), 4(a)–4(f), and 5(a)–5(f). In each of the six figures (for each case of A, B, and C), the discrete statistical sample space is indicated by the dark plots, and the actual design storm prediction (i.e., A, B, or C) is shown by the white line within the shaded area.

Tables 1–3 give the actual values for the peak flow rate and then give the mean and standard deviation for the peak flow rate of each of the six rainfall-runoff hydrological models as applied to each of the three design storm runoff hydrographs. Further, the tables provide the actual total runoff volume, for the set of stochastic realizations of the runoff hydrograph, produced from (7), and then provide the mean and standard deviation for the total runoff volume for each of the six rainfall-

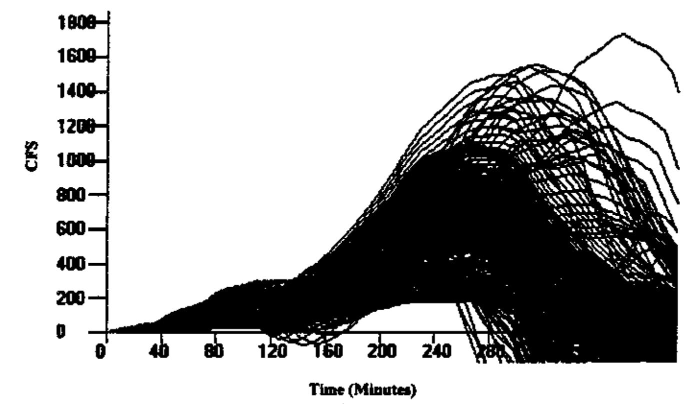


FIG. 3(*a*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph A—Rainfall-Runoff Model 1

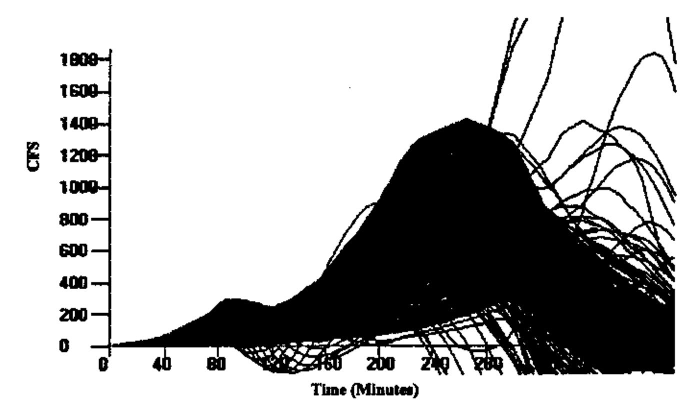


FIG. 3(*b*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph A—Rainfall-Runoff Model 2

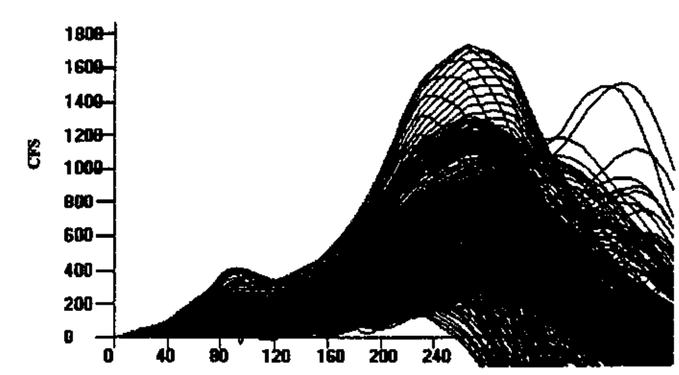


FIG. 3(*c*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph A—Rainfall-Runoff Model 3

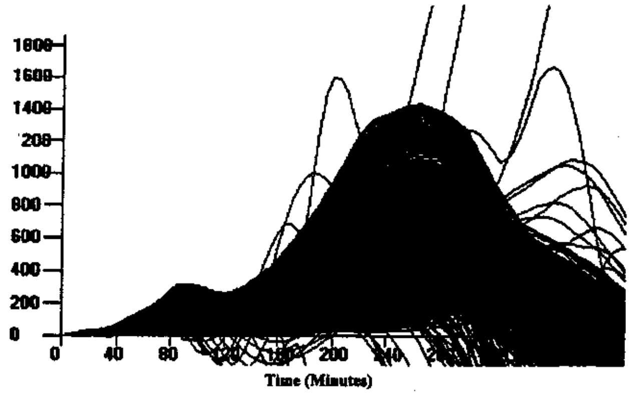


FIG. 3(*d*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph A—Rainfall-Runoff Model 4

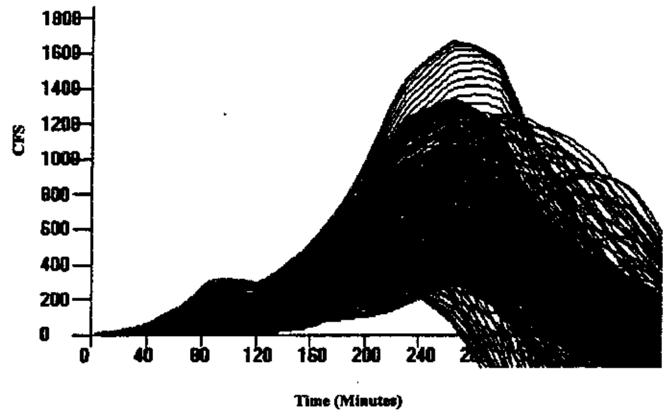


FIG. 3(*e*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph A—Rainfall-Runoff Model 5

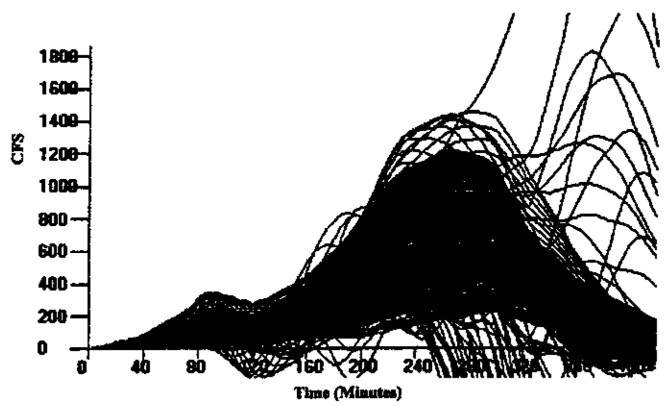


FIG. 3(*f*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph A—Rainfall-Runoff Model 6

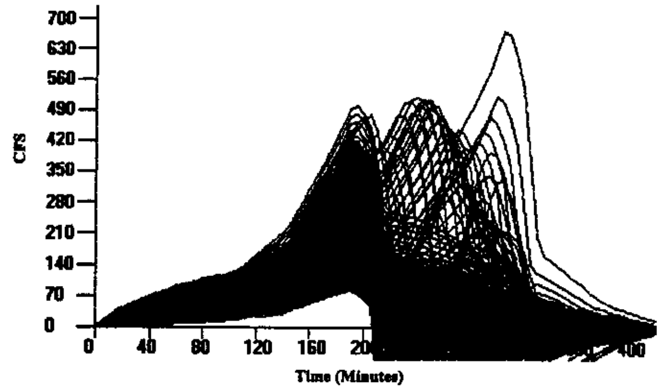


FIG. 4(*a*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph B (Shown in White)—Rainfall-Runoff Model 1

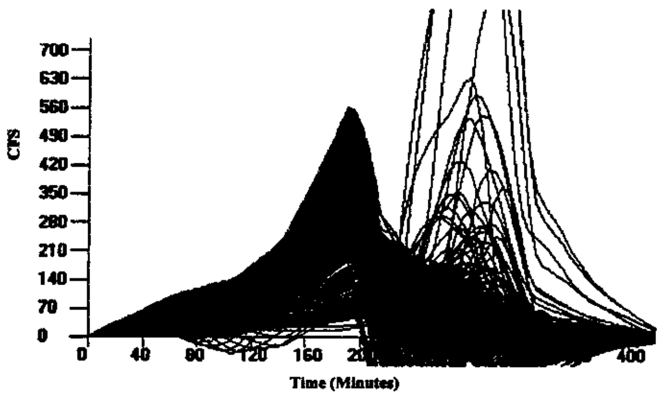


FIG. 4(*b*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph B (Shown in White)—Rainfall-Runoff Model 2

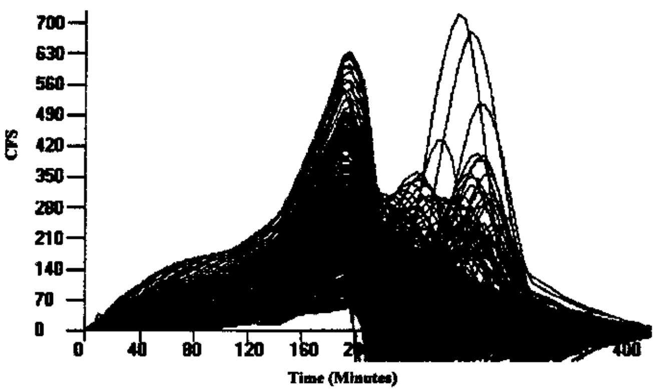


FIG. 4(*c*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph B (Shown in White)—Rainfall-Runoff Model 3

runoff hydrological models as applied to each of the three design storm runoff hydrographs.

In examining Figs. 3–5, it is seen that for all six rainfall-runoff models, there is a considerable dispersion in prediction results. Generally, the vast majority of probable prediction hydrographs have a shape that emulates the original rainfall-runoff model output. However, from the figures, there are occurrences of significantly delayed peak flows that reflect similar occurrences observed from the historic performance data of these models in the tests performed by Abbott (1978). That is, realizations shown in the figures reflect the history of rainfall-runoff model performance as noted in validation tests.

In the present case, the stochastic process shown in Figs. 2–5 reflects the use of bootstrapping in order to populate the low-population discrete distribution of the original data. If validation results were to be preserved by the engineering community in a common database, for all rainfall-runoff models, then bootstrapping may not be needed, as sufficient performance data may exist.

It is noted that the results reflect the variance in modeling estimates as developed from validation results. In practice, one generally does not have a calibrated rainfall-runoff model on a site-by-site basis, but may have a regionally calibrated model. This occurrence will result in less certainty in modeling results, and there will be a corresponding increase in the variance in prediction results addressed herein. To study this latter effect, the approach described earlier can be readily extended to develop error realizations from site uncalibrated rainfall-runoff models versus stream gauge data, and the corresponding stochastic integral equation formulation is developed analogously.

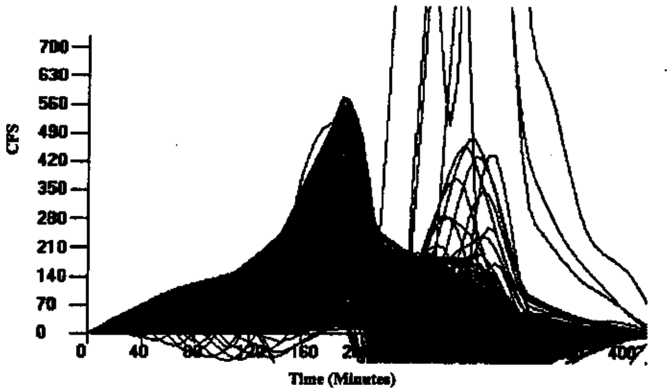


FIG. 4(*d*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph B (Shown in White)—Rainfall-Runoff Model 4

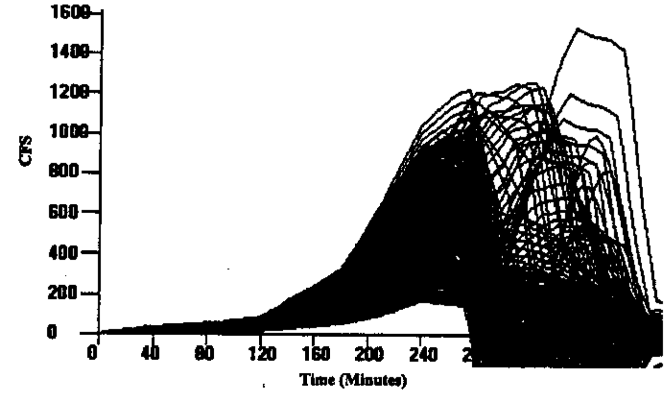


FIG. 5(*a*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph C (Shown in White)—Rainfall-Runoff Model 1

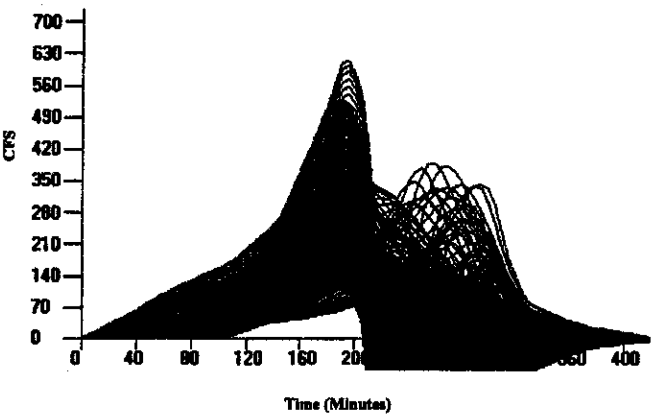


FIG. 4(*e*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph B (Shown in White)—Rainfall-Runoff Model 5

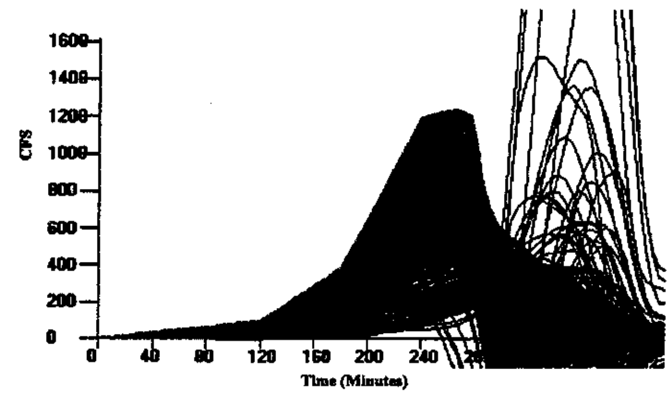


FIG. 5(*b*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph C (Shown in White)—Rainfall-Runoff Model 2

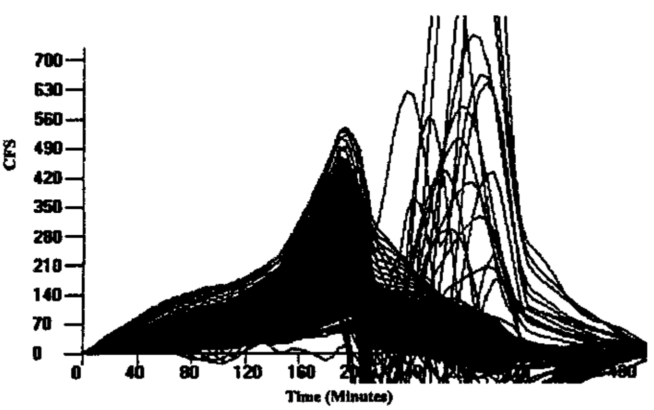


FIG. 4(*f*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph B (Shown in White)—Rainfall-Runoff Model 6

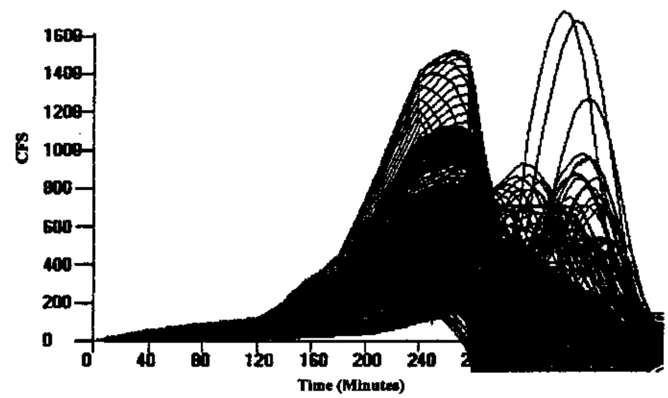


FIG. 5(*c*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph C (Shown in White)—Rainfall-Runoff Model 3

The objective evaluation based on error realizations may be used as a comparative tool to rank a set of hydrologic models under consideration. For example, reference to Tables 1–3 shows that, for the three design storm runoff hydrograph test cases, ranking the six hydrologic models with respect to how accurately their expected value estimates predict the peak flow rate [for the considered test cases of Abbott (1978)] may be possible. However, due to the limited data available to this analysis, the six models tested are assigned a random number for the rankings given in the tables.

FUTURE RESEARCH NEEDS

Several topics remain for future research, including (but by no means limited to) the sensitivity of model error variance estimates due to selection of storm class for partitioning purposes; extension of the SIE method to cases involving significant runoff storage effects (which may negate the fundamental assumption of error interval mutual independence on a class basis); alternative error analysis, such as using the storm rainfall rather than the model output [e.g., Hromadka and Whitley (1989)] or assuming independence of intervals of the error transfer function realizations; among many other topics. Other areas of research include evaluating other techniques for handling small sample sizes; comparison of statistical results using a single long-duration validation event versus several short-duration events; determining when the sample size is sufficient for statistical analysis; evaluating methods to determine storm classes (i.e., split-sampling decision techniques); sensitivity of SIE variance estimates to choice of sampling techniques and

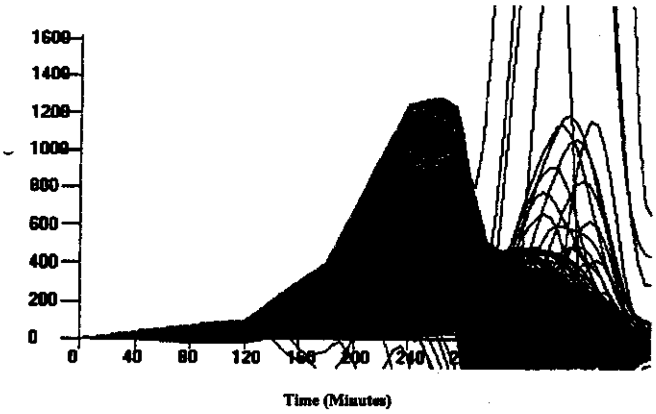
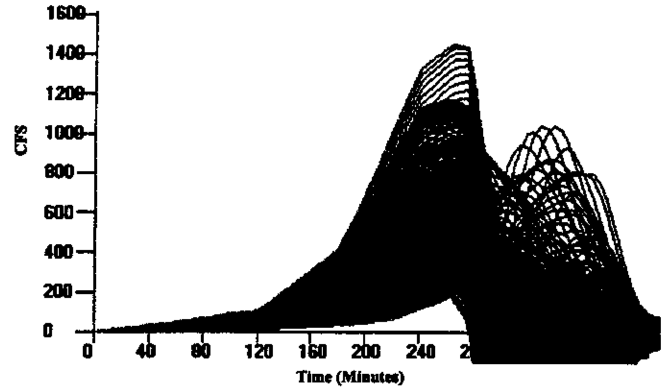


FIG. 5(*d*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph C (Shown in White)—Rainfall-Runoff Model 4



'. 5(*e*). Stochastic Process of Runoff Hydrographs for De-*n* Storm Runoff Hydrograph C (Shown in White)—Rainfall-Runoff Model 5

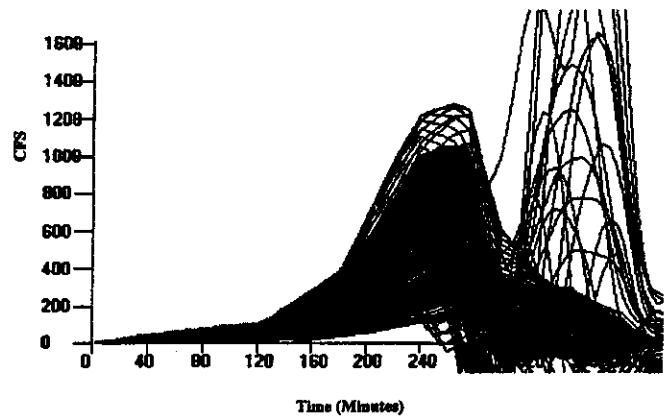


FIG. 5(*f*). Stochastic Process of Runoff Hydrographs for Design Storm Runoff Hydrograph C (Shown in White—Rainfall-Runoff Model 6

to choice of sample size and classification; and analysis of the mutually independent assumption used in the bootstrapping technique employed in the present paper.

The ability to include model performance with storm runoff estimates is a feasible alternative to using a single point estimate or a single runoff hydrograph estimate for design studies. h capabilities may aid in deciding which model structure -y be "best" for a particular application, and for risk analysis of a system. By creating a central database of hydrologic model validation studies, including the data needed for the SIE analysis, access to and input to the database are possible via the Internet. Eventually, a significant database may be built that can replace the need to synthetically populate a distribution of error intervals (and eliminate the bias created by bootstrapping or sampling from a limited sample).

TABLE 1. Statistical Summary of Peak Flow Rate Values, for Design Storm Runoff Hydrograph A

Actual value (1)	Model number (2)	Mean (3)	Standard deviation (4)
	(*a*) Peak Q; m^3/s (cfs)		
21 (748)	1	20 (719)	8 (293)
21 (748)	2	24 (854)	6 (337)
21 (748)	3	19 (686)	9 (312)
21 (748)	4	19 (679)	13 (454)
21 (748)	5	21 (726)	9 (313)
21 (748)	6	22 (787)	10 (354)
	(*b*) Volume; 1,000 m^3 (AF)		
207 (168)	1	197 (160)	89 (72)
207 (168)	2	234 (190)	90 (73)
207 (168)	3	190 (154)	91 (74)
207 (168)	4	187 (152)	113 (92)
207 (168)	5	203 (165)	94 (76)
207 (168)	6	221 (179)	96 (78)

TABLE 2. Statistical Summary of Peak Flow Rate Values, for Design Storm Runoff Hydrograph B

Actual value (1)	Model number (2)	Mean (3)	Standard deviation (4)
	(*a*) Peak Q; m^3/s (cfs)		
8 (292)	1	8 (284)	3 (105)
8 (292)	2	12 (334)	4 (136)
8 (292)	3	7 (264)	3 (119)
8 (292)	4	7 (262)	6 (209)
8 (292)	5	8 (274)	3 (115)
8 (292)	6	9 (305)	4 (158)
	(*b*) Volume; 1,000 m^3 (AF)		
46 (37)	1	46 (37)	20 (16)
46 (37)	2	53 (43)	21 (17)
46 (37)	3	44 (36)	20 (16)
46 (37)	4	43 (35)	27 (22)
46 (37)	5	47 (38)	20 (16)
46 (37)	6	52 (42)	22 (18)

TABLE 3. Statistical Summary of Peak Flow Rate Values, for Design Storm Runoff Hydrograph C

Actual value (1)	Model number (2)	Mean (3)	Standard deviation (4)
	(*a*) Peak Q; m^3/s (cfs)		
18 (648)	1	18 (640)	7 (247)
18 (648)	2	21 (757)	9 (318)
18 (648)	3	17 (613)	8 (280)
18 (648)	4	17 (604)	14 (484)
18 (648)	5	7 (614)	8 (267)
18 (648)	6	20 (714)	11 (381)
	(*b*) Volume; 1,000 m^3 (AF)		
117 (95)	1	121 (98)	49 (40)
117 (95)	2	139 (113)	52 (42)
117 (95)	3	116 (94)	52 (42)
117 (95)	4	112 (91)	68 (55)
117 (95)	5	122 (99)	52 (42)
117 (95)	6	133 (108)	57 (46)

CONCLUSIONS

The subsections of the section on mathematical model development show how a stochastic process has been developed that objectively evaluates rainfall-runoff hydrologic model performance in prediction mode. The process relies solely on data that the hydrologic models themselves produce in validation tests where calibrated rainfall-runoff models are tested against rainfall-runoff data sets not included in the calibration data sets. The runoff hydrographs from the rainfall-runoff hydrologic model, in prediction mode, are compared to the corresponding stream gauge data obtained during the same storm event. The deviation between the two realizations is recorded as an error realization (first subsection).

A convolution procedure (based on the Volterra integral) is used to obtain a set of error transfer function realizations using the storm model output and the error realizations (second subsection). The algorithm may now be used to statistically evaluate a design storm model runoff hydrograph prediction.

The algorithm is used to generate a synthetic set of error transfer function realizations, but here the error realization is obtained by the convolution of the design storm model output and the previously generated error transfer functions. This produces an ensemble of design storm error realizations (third subsection). The error realizations are then subtracted from the design storm model result to obtain a set of equally likely outcomes for the design storm model. This statistical information is based solely on the error realizations obtained from the models themselves, in prediction mode.

The work of Tsokos and Padgett (1974) demonstrates the predictive ability available by use of a stochastic integral equation such as (2). However, further research is needed to demonstrate the predictive ability of (2) in rainfall-runoff modeling, among many other topics. Perhaps model users could make available their validation/calibration data sets of stream gauge data and model results via the Internet. In time, a considerable data set "history" of model performance would be assembled and subsequently could be analyzed using a stochastic integral equation formulation such as presented herein. Further research is also needed regarding issues of "filtering" the error distribution to avoid so-called "negative runoff." Obviously, the presented results could have been filtered so as to present a more tractable appearance; however, by using a nonfiltered set of results, a more accurate representation of the modeling results is presented.

APPENDIX. REFERENCES

Abbott, J. (1978). "Testing of several runoff models on an urban watershed." *Tech. Paper no. 59*, U.S. Army Corps of Engineers, Hydro. Engrg. Ctr., Vicksburg, Miss.

Efron, B., and Tribshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman Hall, Ltd., London.

Hromadka, T. V. II, McCuen, R. H., and Yen, C. C. (1987). *Computational hydrology in flood control design and planning*. Lighthouse Publications.

Hromadka, T. V. II, and Whitley, R. J. (1989). *Stochastic equations and rainfall-runoff models*. Springer-Verlag New York, Inc., New York, N.Y.

Tsokos, C. P., and Padgett, W. J. (1974). *Random integral equations with applications to life science and engineering*. Academic Press, Inc., San Diego.