ESTIMATING UNCERTAINTY IN DESIGN STORM RAINFALL-RUNOFF MODELS USING A STOCHASTIC INTEGRAL EQUATION

T.V. Hromadka II
Director, Water Resources Engineering
Williamson and Schmid
17782 Sky Park Boulevard, Irvine, California
and
Associate Professor, Department of Mathematics
California State University, Fullerton, California

J.J. DeVries
Professor, Department of Civil Engineering
University of California, Davis, California

Abstract

Almost all rainfall-runoff models in use today involve the subdivision of the catchment into smaller areas, linked together by a system of channel links. These "link-node" hydrologic models represent the flow processes within the channel links by a translation (moving in time) and an attenuation (reduction of the maximum or peak flow rate) of the runoff (floodwater) hydrograph. The runoff in each subarea is based upon the available rainfall data, modified according to an assumed "loss rate" due to soil-infiltration, ponding, exaporation, and other effects. The net effect of all these approximations is to result in a vast spectrum of possible modeling structures. Using a stochastic integral equation, we can mathematically approximate many of these rainfall-runoff models with a generalized model that is more tractable to detailed analysis of the model structure. We can then proceed to evaluate rainfall-runoff model uncertainty in overall generality. The approach used in this paper is to isolate the uncertainty in runoff predictions from the expected value of the model runoff estimate, and then attempt to analyze the uncertainty as a separate form of information. In this way, the uncertainty may be analyzed as a stochastic process. Once the underlying distributions are identified, they can be normalized with respect to certain catchment characteristic variables, so that these distrtibutions can be rescaled for application at arbitrary study sites.

PAPER "Estimating Uncertainty in Design Storm Rainfall-Runoff Models Using a Stochastic Integral Equation" presented at American Geophysical Union (AGU) Fall 1989 Meeting, San Francisco, Session H31C, Paper No. 12.

INTRODUCTION

Due to the nondeterministic nature of the rainfall-runoff processes occurring over the catchment, the mathematical descriptions of these processes result in stochastic equations. Additionally, the so-called deterministic rainfall-runoff models used to describe the several physical processes contain parameters or coefficients which have well-defined physically-based meanings, but whose exact values are unknown. The boundary conditions of the problem itself are unknown (e.g., the temporal and spatial distribution of rainfall over the catchment for the storm event under study and also for all prior storm events) and often exhibit considerable variations with respect to the assumed boundary conditions, the measured rainfall at a single location (e.g., Nash and Sutcliffe, 1970; Huff, 1970). Thus the physically-based parameters and coefficients, and also the problem boundary conditions, are not the assumed values but are instead random variables and stochastic processes whose variations about the assumed values are governed by certain probability distributions.

In this paper, the uncertainty problem is addressed by providing a methodology which can be incorporated into almost all rainfall-runoff models. The methodology is based upon the standard theory of stochastic integral equations which has been successfully applied to several problems in the life sciences and chemical engineering (e.g., Tsokos and Padgett, 1974, provide a thorough development). The stochastic integral formulation is used to represent the total error between a record of measured rainfall-runoff data and the model estimates, and provides an answer to the question: "based upon the historic rainfall-runoff data record and the model's accuracy in estimating the measured runoff, what is the distribution of probable values of the subject criterion variable given a hypothetical rainfall event?" Using the analysis results (Hromadka, 1989), we now extend our findings in order to generalize the analysis to arbitrary rainfall-runoff model structures.

Almost all rainfall-runoff models in use today involve the subdivision of the catchment into smaller areas, linked together by a system of channel links. These "link-node" hydrologic models represent the flow processes within the channel links by a translation (moving in time) and an attenuation (reduction of the maximum or peak flow rate) of the inflow hydrograph. The runoff in each subarea is based upon the available rainfall data, modified according to an assumed "loss rate" due to soil-infiltration, ponding, evaporation, and other effects. The net effect of all these approximations is to result in a vast spectrum of possible modeling structures. Using the model structure presented in Hromadka, (1989), we can mathematically approximate many of these rainfall-runoff models with a single model structure, and can proceed to evaluate rainfall-runoff model uncertainty in overall generality. The approach used in this paper is to isolate the uncertainty in runoff predictions from the expected value of the model runoff estimate, and then attempt to analyze the uncertainty as a separate form of information. In this way, the uncertainty may be analyzed as a stochastic process. Once the underlying distributions are identified, they can be normalized with respect to certain catchment characteristic variables, so that these distributions can be rescaled for application at arbitrary study sites.

STOCHASTIC INTEGRAL EQUATION

Rainfall-Runoff Model Errors

Let M be a deterministic rainfall-runoff model which transforms rainfall data for some storm event, i, noted by $P_g^i(t)$, into an estimate of runoff, $M^i(t)$, by

$$M : P_g^i(t) \longrightarrow M^i(t) \tag{1}$$

where t is time. In our problem, rainfall data are obtained from a single rain gauge. The operator M may include loss rate and flow routing parameters, memory of prior storm event effects, and other factors. It is noted that precipitation data are now used in the current analysis rather than using a measured effective rainfall such as employed in Hromadka (1989).

Let $P_g^i(t)$ be the rainfall measured from storm event i, and $Q_g^i(t)$ be the runoff measured at the stream gauge. Various error (or uncertainty) terms are now defined such that for arbitrary storm event i,

$$Q_g^i(t) = M^i(t) + E_m^i(t) + E_d^i(t) + E_r^i(t) \tag{2}$$

where

$E_m^i(t)$ is the modeling error due to inaccurate approximations of the physical processes (spatially and temporally);

$E_d^i(t)$ is the error in data measurements of $P_g^i(t)$ and $Q_g^i(t)$ (which is assumed hereafter to be of negligible significance in the analysis);

$E_r^i(t)$ is the remaining "inexplainable" error, such as due to the unknown variation of effective rainfall (i.e., rainfall less losses; rainfall excess) over the catchment, among other factors.

Let $E^i(t)$ be redefined to equal the total error

$$E^i(t) = E_m^i(t) + E_d^i(t) + E_r^i(t) \tag{3}$$

where $E^i(t)$ is necessarily highly correlated to $E_r^i(t)$ due to the given assumptions. Because $E^i(t)$ depends on the model M used in Eq. (1), then Eqs. (2) and (3) are combined as

$$Q_g^i(t) = M^i(t) + E_M^i(t) \tag{4}$$

where $E_M^i(t)$ is a conditional notation for $E^i(t)$, given model type M.

The several terms in Eq. (4) are each a realization of a stochastic process. And for a future storm event D, the $E_M^D(t)$ is not known precisely, but rather is an unknown realization of a stochastic process distributed as $[E_M^D(t)]$ where

$$[Q_M^D(t)] = M^D(t) + [E_M^D(t)] \tag{5}$$

In Eq. (5), $[Q_M{}^D(t)]$ and $[E_M{}^D(t)]$ are the stochastic processes associated to the catchment runoff and total modeling error, respectively, associated with model M, for hypothetical storm event D. Hence in prediction, the model output of Eq. (5) is not a single outcome, but instead is a stochastic distribution of outcomes, distributed as $[Q_M{}^D(t)]$. Should A be some functional operator on the possible outcome (e.g., detention basin volume; peak flow rate; median flow velocity, etc.) of storm event D, then the possible value of A for storm event D, noted as $A_M{}^D$, is a random variable distributed as $[A_M{}^D]$, where

$$[A_M{}^D] = A[Q_M{}^D(t)] \tag{6}$$

Developing Distributions for Model Estimates

The distribution for $[E_M{}^D(t)]$ may be estimated by using the available sampling of realizations of the various stochastic processes:

$$\{E_M{}^i(t)\} = \{Q_g{}^i(t) - M^i(t)\}, \ i = 1,2,\cdots \tag{7}$$

Assuming elements in $\{E_M{}^i(t)\}$ to be dependent upon the "severity" of $Q_g{}^i(t)$, one may partition $\{E_M{}^i(t)\}$ into classes of storms such as mild, major, flood, or others, should ample rainfall-runoff data be available to develop significant distributions for the resulting subclasses. To simplify development purposes, $[E_M{}^D(t)]$ will be based on the entire set $\{E_M{}^i(t)\}$ with the underlying assumption that all storms are of "equivalent" error; storm classes will be used later.

The second assumption involved is to assume each $E_M{}^i(t)$ is strongly correlated to some function of precipitation, $F^i(t) = F(P_g{}^i(t))$, where F is an operator which includes parameters, memory of prior rainfall, and other factors. Assuming that $E_M{}^i(t_o)$ depends only on the values of $F^i(t)$ for time $t \leq t_o$, then $E_M{}^i(t)$ is expressed as a causal linear filter (for only mild conditions imposed on $F^i(t)$), given by the stochastic integral equation (see Tsokos and Padgett, 1974)

$$E_M{}^i(t_o) = \int_{s=0}^{t_o} F^i(t_o - s)\, h_M{}^i(s)\, ds \tag{8}$$

where $h_M{}^i(t)$ is the transfer function between $(E_M{}^i(t), F^i(t))$. Other convenient candidates to be used in Eq. (8), instead of $F^i(t)$, are the storm rainfall, $P_g{}^i(t)$, and the model estimates itself, $M^i(t)$.

Given a significant set of storm data, an underlying distribution $[h_M(t)]$ of the $\{h_M{}^i(t)\}$ may be identified, or the $\{h_M{}^i(t)\}$ may be used directly as in the case of having a discrete distribution of equally-likely realizations. Using $[h_M(t)]$ as notation for both cases of distributions stated above, the predicted response from M for future storm event D is estimated to be

$$[Q_M{}^D(t)] = M^D(t) + [E_M{}^D(t)] \tag{9}$$

Combining Eqs. (8) and (9),

$$[Q_M{}^D(t)] = M^D(t) + \int_{s=0}^{t} F^D(t - s)\ [h_M(s)]\ ds \tag{10}$$

and for the functional operator A, Eq. (6) is rewritten as

$$[A_M{}^D] = A[Q_M{}^D(t)] = A\left(M^D(t) + \int_{s=0}^{t} F^D(t - s)\ [h_M(s)]\ ds\right) \tag{11}$$

Confidence interval estimates for the chosen criterion variable can now be obtained from the frequency-distribution, $[A_M{}^D]$. It is noted that $[A_M{}^D]$ is necessarily a random variable distribution that depends on the model structure, M.

DEVELOPMENT OF TOTAL ERROR DISTRIBUTIONS

A translation unsteady flow routing rainfall-runoff model

The previous concepts are now utilized to directly develop the total error distributions, $[E_M(t)]$, for a set of three idealized catchment responses. Besides providing a set of applications, additional notation and concepts are introduced, leading to the introduction of storm classes.

Let F be a functional which operates on rainfall data, $P_g{}^i(t)$, to produce the realization, $F^i(t)$, for storm i by

$$F: P_g{}^i(t) \longrightarrow F^i(t) \tag{12}$$

The catchment R is subdivided into m homogeneous subareas, $R = \cup R_j$, (see Fig. 1; where, m = 9), such that in each R_j, the effective rainfall, $e_j{}^i(t)$, is assumed given by

$$e_j{}^i(t) = \lambda_j(1 + X_j{}^i)\, F^i(t) \tag{13}$$

where λ_j is a constant proportion factor; and where $X_j{}^i$ is a sample of a random variable, which is constant for storm event i. The parameter λ_j is defined for subarea R_j and represents the relative runoff response of R_j in comparison to $F^i(t)$, and is a constant for all storms, whereas $X_j{}^i$ is a sample of the random variable distributed as $[X_j]$, where the set of distributions, $\{[X_j];\ j = 1,2,\cdots,m\}$ may be mutually dependent.

The subarea runoff is

$$q_j{}^i(t) = \int_{s=0}^{t} e_j{}^i(t-s)\, \phi_j{}^i(s)\, ds = \int_{s=0}^{t} \lambda_j(1 + X_j{}^i)\, F^i(t-s)\, \phi_j{}^i(s)\, ds \tag{14}$$

At this stage of development, unsteady flow routing along channel links (see Fig. 1) is assumed to be pure translation. Thus, each channel link, L_k, has the constant translation time, T_k. Hence from Fig. 1, the total runoff response at the stream gauge for storm event i, $Q_g{}^i(t)$, is the sum of subarea runoffs, each translated by the sum of associated link travel times:

$$Q_g{}^i(t) = \sum_{j=1}^{9} q_j{}^i(t - \tau_j) \tag{15}$$

where $q_j{}^i(t - \tau_j)$ is defined to be zero for negative arguments; and τ_j is the sum of link travel times (e.g. from Fig. 1, $\tau_1 = T_1 + T_2 + T_3$; $\tau_6 = T_5 + T_6$; $\tau_9 = 0$).

For the above particular assumptions,

$$\begin{aligned} Q_g{}^i(t) &= \sum_{j=1}^{9} \int_{s=0}^{t} \lambda_j(1 + X_j{}^i)\, F^i(t - s)\, \phi_j{}^i(s - \tau_j)\, ds \\ &= \int_{s=0}^{t} F^i(t - s) \left[\sum_{j=1}^{9} \lambda_j(1 + X_j{}^i)\, \phi_j{}^i(s - \tau_j) \right] ds \end{aligned} \tag{16}$$

In a final form, the runoff response for the given simplification is

$$\begin{aligned} Q_g{}^i(t) &= \int_{s=0}^{t} F^i(t - s) \sum_{j=1}^{9} \lambda_j\, \phi_j{}^i(s - \tau_j)\, ds \\ &+ \int_{s=0}^{t} F^i(t - s) \sum_{j=1}^{9} \lambda_j X_j{}^i \phi_j{}^i(s - \tau_j)\, ds \end{aligned} \tag{17}$$

In the above equations, the samples $\{X_j{}^i\}$ are unknown to the modeler for any storm event i. From Eq. (17), the model structure, M, used in design practice is

$$M^i(t) = \int_{s=0}^{t} F^i(t - s) \sum_{j=1}^{9} \lambda_j\, \phi_j{}^i(s - \tau_j)\, ds \tag{18}$$

Then, $Q_g{}^i(t) = M^i(t) + E_M{}^i(t)$ where

$$E_M{}^i(t) = \int_{s=0}^{t} F^i(t - s)\, h_M{}^i(s)\, ds \tag{19}$$

where $h_M{}^i(s)$ follows directly from Eqs. (17) and (18).

Should the subarea UH all be assumed fixed, (i.e., $\phi_j{}^i(t) = \phi_j(t)$, for all i), as is assumed in practice, then the above equations can be further simplified as

$$M^i(t) = \int_{s=0}^{t} F^i(t-s)\,\Phi(s)\,ds \tag{20}$$

where $\Phi(s) = \sum_{j=1}^{9} \lambda_j\, \phi_j(s - \tau_j)$. Additionally, the distribution of the stochastic process $[h_M(t)]$ is readily determined for this simple example,

$$[h_M(t)] = \sum_{j=1}^{9} [X_j]\ \lambda_j\ \phi_j(t - \tau_j) \tag{21}$$

where $[h_M(t)]$ is directly equated to the 9 random variables, $\{X_j, j = 1,2,\cdots,9\}$. It is again noted that the random variables, X_j, may be all mutually dependent.

In prediction, the estimated runoff hydrograph is the distribution $[Q_M{}^D(t)]$ where $[Q_M{}^D(t)] = M^D(t) + [E_M{}^D(t)]$, and M refers to the model structure of Eqs. (18) or (20).

For this example problem, the stochastic integral formulation is

$$[Q_M{}^D(t)] = \int_{s=0}^{t} F^D(t-s)\ \Phi(s)\,ds + \int_{s=0}^{t} F^D(t-s)\ [h_M(s)]\,ds \tag{22}$$

where the error distribution, $[E_M{}^D(t)]$, is assumed to be correlated to the model input, $F^D(t)$, as provided in Eqs. (19) and (21).

Multilinear unsteady flow routing and storm classes

The above application is now extended to include the additional assumption that the channel link travel times are strongly correlated to some set of characteristic descriptions of the runoff hydrograph being routed, such as some weighted mean flow rate of the associated hydrograph. For example, the widely used Convex Routing technique (Mockus, 1972) often utilized the 85-percentile of all flows in excess of one-half of the peak flow rate as a statistic used to estimate the routing parameters. But by the previous development (i.e., definition of $e_j{}^i(t)$), all runoff hydrographs in the link-node channel system would be highly correlated to an equivalent weighting of the model input, $F^i(t)$. Hence, storm classes, $[\xi_z]$, of "equivalent" $F^i(t)$ realizations could be defined where all elements of $[\xi_z]$ have the same characteristic parameter set, $C(F^i(t))$, by

$$[\xi_z] = \{F^i(t) \mid C(F^i(t)) = z\} \tag{23}$$

And for all $F^i(t) \in [\xi_z]$, each respective channel link travel time is identical, that is $T_k = T_{k_z}$ for all $F^i(t) \in [\xi_z]$. In the above definition of storm class, z is a characteristic parameter set in vector form. (An example of such a characteristic parameter set is given in a subsequent section.)

This extension of the translation unsteady flow routing algorithm to a multilinear formulation (involving a set of link translation times) modifies the previous runoff equations (20) and (21) to be, in general

$$M^i(t) = \int_{s=0}^{t} F^i(t-s) \sum_j \lambda_j \, \phi_j(s-\tau_j^z)ds = \int_{s=0}^{t} F^i(t-s)\, \Phi_z(s)ds; \; F^i(t) \in [\xi_z] \quad (24)$$

where $\Phi_z(s) = \sum_j \lambda_j \, \phi_j(s-\tau_j^z)$, and

$$E_M{}^i(t) = \int_{s=0}^{t} F^i(t-s) \; h_{M_z}{}^i(s)\, ds; \; F^i(t) \in [\xi_z] \quad (25)$$

The structure of the new set of equations motivates an obvious extension of the definition of the subarea UH, the subarea λ_j proportion factor, and the subarea random variable distribution $[X_j]$, to all be also defined on the storm class basis of $[\xi_z]$. Thus, Eq. (24) is extended as

$$M^i(t) = \int_{s=0}^{t} F^i(t-s) \sum_j \lambda_j^z \, \phi_j^z(s-\tau_j^z)ds$$

$$= \int_{s=0}^{t} F^i(t-s) \; \Phi_z(s)\, ds; \; F^i(t) \in [\xi_z] \quad (26)$$

The stochastic process $[h_{M_z}(t)]$ is distributed as

$$[h_{M_z}(t)] = \sum_j [X_j^z] \;\lambda_j^z \; \phi_j^z(s-\tau_j^z); \; F^i(t) \in [\xi_z] \quad (27)$$

And in prediction,

$$[Q_M{}^D(t)] = M^D(t) + [E_M{}^D(t)]; \; F^D(t) \in [\xi_D] \quad (28)$$

where

$$[E_M{}^D(t)] = \int_{s=0}^{t} F^D(t-s)\,[h_M(s)]\,ds;\; F^D(t) \in [\xi_D] \tag{29}$$

A Multilinear Rainfall-Runoff Model

The previous two model derivations resulted in the development of the total error distribution, $[E_M(t)]$, for some particular model structures. In this section, the above results are generalized to include a wide range of possibilities.

As before, Let F be a functional defined on the assumed rainfall data, $F: P_g{}^i(t) \longrightarrow F^i(t)$. The catchment R is subdivided into m subareas, $\{R_j;\ j = 1,2,\cdots,m\}$ linked together by unsteady flow routing models. The link-node model drains freely to the single stream gauge where the data, $Q_g{}^i(t)$, is measured. The problem is to predict the runoff response at the stream gauge corresponding to a hypothetical storm event rainfall, $P_g{}^D(t)$.

Each subarea's effective rainfall, $e_j{}^i(t)$, is now defined to be the sum of proportions of $F^i(t)$ translates by

$$e_j{}^i(t) = \sum_k \lambda_{jk}(1 + X_{jk}{}^i)\, F^i(t - \theta_{jk}{}^i);\; F^i(t) \;\in [\xi_Z] \tag{30}$$

where $X_{jk}{}^i$ and $\theta_{jk}{}^i$ are samples of the random variables distributed as $[X_{jk}]$ and $[\theta_{jk}]$, respectively. In the above equation and all equations that follow, it is assumed that a storm class system is defined, $[\xi_Z]$, such that for $F^i(t) \in [\xi_Z]$, all parameters and probabilistic distributions are uniquely defined, and there is no loss in understanding by omitting the additional notation needed to indicate the storm class.

The subarea runoff is

$$q_j{}^i(t) = \int_{s=0}^{t} \sum_k \lambda_{jk}\,(1 + X_{jk}{}^i)\, F^i(t - \theta_{jk}{}^i - s)\;\; \phi_j(s)\, ds \tag{31}$$

or in a simpler form,

$$q_j{}^i(t) = \int_{s=0}^{t} F^i(t-s) \sum_k \lambda_{jk}\,(1 + X_{jk}{}^i)\;\; \phi_j(s - \theta_{jk}{}^i)ds \tag{32}$$

It is assumed that the unsteady flow channel routing effects are highly correlated to the magnitude of runoff in each channel link, which is additionally correlated to the magnitude of the model input realization, $F^i(t)$. On a storm class basis, each channel link is assumed to respond linearly in that (e.g., Doyle et al, 1983)

$$O_1{}^i(t) = \sum_{\ell} a_\ell \, I_1{}^i(t - \alpha_\ell) \tag{33}$$

where $O_1{}^i(t)$ and $I_1{}^i(t)$ are the outflow and inflow hydrographs for link 1, and storm event i; and $\{a_\ell\}$ and $\{\alpha_\ell\}$ are constants which are defined on a storm class basis which is also used for the model input, $F^i(t)$. Thus, the channel link flow routing algorithm is multilinear with routing parameters defined according to the storm class, $[\xi_z]$ (see Becker and Kundezewicz, 1987, for an analogy based on multilinear approximation of nonlinear routing).

Should the above outflow hydrograph, $O_1(t)$, now be routed through another link (number 2), then $I_2(t) = O_1(t)$ and from the above

$$\begin{aligned} O_2(t) &= \sum_{\ell_2=1}^{n_2} a_{\ell_2} I_2(t - \alpha_{\ell_2}) \\ &= \sum_{\ell_2=1}^{n_2} a_{\ell_2} \sum_{\ell_1=1}^{n_1} a_{\ell_1} I_1(t - \alpha_{\ell_1} - \alpha_{\ell_2}) \end{aligned} \tag{34}$$

For L links, each with their own respective stream gauge routing data, the above linear routing techniques result in the outflow hydrograph for link number L, $O_L(t)$, being given by

$$\begin{aligned} O_L(t) = \sum_{\ell_L=1}^{n_L} a_{\ell_L} \sum_{\ell_{L-1}=1}^{n_{L-1}} a_{\ell_{L-1}} \cdots \\ \cdots \sum_{\ell_2=1}^{n_2} a_{\ell_2} \sum_{\ell_1=1}^{n_1} a_{\ell_1} I_1\left(t - \alpha_{\ell_1} - \alpha_{\ell_2} - \cdots - \alpha_{\ell_{L-1}} - \alpha_{\ell_L}\right) \end{aligned} \tag{35}$$

Using an index notation, the above $O_L(t)$ is written as

$$O_L(t) = \sum_{<\ell>} a_{<\ell>} \; I_1(t - \alpha_{<\ell>}) \tag{36}$$

For subarea R_j, the runoff hydrograph for storm i, $q_j^i(t)$, flows through L_j links before arriving at the stream gauge and contributing to the total modeled runoff hydrograph, $M^i(t)$. All of the parameters $a^i_{<\ell>}$ and $\alpha^i_{<\ell>}$ are constants on a storm class basis. Consequently from the linearity of the routing technique, the m-subarea link node model is given by the sum of the m, $q_j^i(t)$ contributions,

$$M^i(t) = \sum_{j=1}^{m} \sum_{<\ell>_j} a^i_{<\ell>_j} q_j^i \left(t - \alpha^i_{<\ell>_j}\right) \tag{37}$$

Finally, the predicted runoff response for storm event D is the stochastic integral formulation

$$[Q_M^D(t)] = \int_{s=0}^{t} F^D(t-s) \left\{ \sum_{j=1}^{m} \sum_{<\ell>_j} a^i_{<\ell>_j} \sum_k \lambda_{jk}(1 + [X_{jk}]) \; \phi_j(s - [\theta_{jk}] - \alpha^i_{<\ell>_j}) \right\} ds; \; F^D(t) \in [\xi_D] \tag{38}$$

Given $F^i(t) \in [\xi_z]$, all subarea runoff parameters $\{\lambda_{jk}, \phi_j(t)\}$ and distributions $\{[X_{jk}], [\theta_{jk}]\}$ are uniquely defined for $j = 1,2,\cdots,m$; and all link routing parameters $\{a_\ell, \alpha_\ell\}$ are also uniquely defined. Then the entire link-node model is linear on a storm class basis and once more Eqs. (26)-(29) apply without modification.

The above multilinear rainfall-runoff model structure represents a highly detailed and distributed parameter model of the rainfall-runoff process which not only can be used to represent the catchment runoff response itself, but also can be used to approximate the response of other hydrologic modeling structures.

Consequently, our final model structure can be used to study the effect on the runoff prediction (at the stream gauge) from arbitrary model M, due to the randomness exhibited by the mutually dependent set of random variables, $\{X_{jk}, \theta_{jk}\}$. Hence for any operator, A, on the predicted runoff response of Eq. (38), the outcome of A for storm event $P_g{}^D(t)$ is the distribution $[A_M{}^D]$, where for all model parameters defined,

$$[A_M{}^D] = A[M^D(t)] = A(\{[X_{jk}], [\theta_{jk}]\}) \tag{39}$$

STOCHASTIC INTEGRAL EQUATIONS AND UNCERTAINTY ESTIMATES

The distributed parameter rainfall-runoff model of Eq. (38) provides a useful approximation of almost any rainfall-runoff model in use today. A stochastic integral equation that is equivalent to Eq. (38) is

$$[Q_M{}^D(t)] = \int_{s=0}^{t} F^D(t-s)\,[\eta(s)]ds;\; F^D(t) \in [\xi_D] \tag{40}$$

where now $[\eta(s)]$ is the distribution of the stochastic process representing the random variations from the set of mutually dependent random variables, $\{X_{jk}, \theta_{jk}\}$, defined on a storm class basis. (It is recalled that on a storm class basis, the hydraulic parameters of $a_{<\ell>j}$ and $\alpha_{<\ell>j}$, and the $\phi_j(s)$, do not vary.) In prediction, the expected runoff estimate for storm events that are elements of $[\xi_D]$ is

$$E[Q_M{}^D(t)] = \int_{s=0}^{t} F^D(t-s)\,E[\eta(s)]ds;\; F^D(t) \in [\xi_D] \tag{41}$$

which is a multilinear version of the well-known unit hydrograph method (e.g., Hromadka et al, 1987), which is perhaps the most widely used rainfall-runoff modeling approach in use today.

Then the model M structure of Eq. (38), when unbiased, is given from Eq. (41), by

$$M^D(t) = E[Q_M{}^D(t)] \tag{42}$$

The total error distribution (for the subject model M) can be developed by

$$[E_M{}^D(t)] = [Q_M{}^D(t)] - E[Q_M{}^D(t)] \tag{43}$$

where all equations are defined on the storm class basis used in the previous equations. Given sufficient rainfall-runoff data, the total error distribution can be approximately developed by use of Eq. (43). Should another rainfall-runoff model structure be used, then $E[Q_M{}^D(t)]$ is replaced by the alternative model, and another set of realizations of $[E_M{}^D(t)]$ is obtained from (43). Equation (43) is important in that given a specified model, the total error in model estimation is approximately given by a stochastic process. And similar to any sampling process, the modeling total error distribution becomes better defined as the sampling population increases. Through the equivalence between Eqs. (38) and (40), the uncertainty of the rainfall-runoff model of Eq. (38) can be evaluated by use of Eq. (40). That is, due to the limited data available, one cannot evaluate each of the random variables and processes utilized in Eq. (38), but one can evaluate the total model error, as developable from Eq. (43).

APPLICATION

In our application problem, the model input functional $F: P_g{}^i(t) \longrightarrow F^i(t)$ is specified as

$$F: P_g{}^i(t) \rightarrow \lambda P_g{}^i(t) \tag{44}$$

where λ is a constant runoff coefficient. The corresponding stochastic integral equation used to related rainfall-runoff data is

$$Q_g{}^i(t) = \lambda \int_{s=0}^{t} P_g{}^i(t - s)\, \eta^i(s)\, ds \tag{45}$$

In this application, storm classes are defined (z) according to the 85-percentile value of rainfall intensity in excess of one half of the maximum 5-minute mean intensity, and also according to the total rainfall mass which occurs within 3 days prior to the subject storm event. Storm classes are then assembled according to the characteristic z-value, at 0.5-inch increments.

For the study location of Southern California, Table 1 summarizes the study catchment characteristics. Table 2 lists the available rain gauge sites and the storm dates of events used in the rainfall-runoff data analysis. Because of the sparcity of rainfall-runoff data, several catchments are considered in order to regionalize the statistical results. All storms considered in Table 2 are assumed to be elements of the same storm class considered important for flood control. That is, it is hypothetisized that the variations in the various random variables and processes identified in Eq. (38), can be considered samples from distributions that apply for each of the considered storm events of Table 2.

For each storm event and catchment, the rainfall-runoff data is used to directly develop the $\{\eta^i(s)\}$ by use of Eq. (45). On a catchment basis, the several resulting $\eta^i(s)$ are pointwise averaged together to determine an estimate for $E[\eta(s)]$ for the prescribed storm class, for the considered catchment.

Summation (or distribution) graphs of the $\{\eta^i(s)\}$ indicate that normalizing could be performed by plotting mass along the y-axis from 0 to 100-percent of mass, and the x-axis as time with respect to the parameter "lag" where 100 percent of lag equals the time at 50 percent of total mass. Plots of normalized summation graphs of the $\eta^i(s)$ realizations for Alhambra Wash, for several storms, are shown in Fig. 2, and plots of summation graphs of the estimates of $E[\eta(s)]$ for the several catchments are shown in Fig. 3. From the data used in Fig. 2, the expected value (for the Alhambra Wash stream gauge) of the charactistic parameters lag and ultimate discharge, U, are obtained.

A comparison between Figs. 2 and 3 shows that the variation in the summation graphs of the $E[\eta(s)]$ among the several considered catchments is of a magnitude similar to the variation between the summation graphs of $\eta^i(s)$ for Alhambra Wash alone. Therefore in order to regionalize the total error distributions, and to increase the population of the random process sampling, the

TABLE 1. WATERSHED CHARACTERISTICS

Watershed Name	Area (mi²)	Length (mi)	Length of Centroid (mi)	Slope (ft/mi)	Percent Impervious (%)	Tc (Hrs)	Storm Date	Peak F_p (inch/hr)	Lag (hrs)	Basin factor
	Watershed Geometry						Calibration Results			
Alhambra Wash[1]	13.67	8.62	4.17	82.4	45	0.89	Feb.78 Mar.78 Feb.80	0.59,0.24 0.35,0.29 0.24	0.62	0.015
Compton2[1]	24.66	12.69	6.63	13.8	55	2.22	Feb.78 Mar.78 Feb.80	0.36 0.29 0.44	0.94	0.015
Verdugo Wash[1]	26.8	10.98	5.49	316.9	20	--	Feb.78	0.65	0.64	0.016
Limekiln[1]	10.3	7.77	3.41	295.7	25	--	Feb.78 Feb.80	0.27 0.27	0.73	0.026
San Jose[2]	83.4	23.00	8.5	60.0	18		Feb.78 Feb.80	0.20 0.39	1.66	0.020
Sepulveda[2]	152.0	19.0	9.0	143.0	24	--	Feb.78 Mar.78 Feb.80	0.22,0.21 0.32 0.42	1.12	0.017
Eaton Wash[1]	11.02[4] (57%)	8.14	3.41	90.9	40	1.05	---	---	---	0.015[7]
Rubio Wash[1]	12.20[5] (3%)	9.47	5.11	125.7	40	0.68	---	---	---	0.015[7]
Arcadia Wash[1]	7.70[6] (14%)	5.87	3.03	156.7	45	0.60	---	---	---	0.015[8]
Compton1[1]	15.08	9.47	3.79	14.3	55	1.92	---	---	---	0.015[8]
Dominguez[1]	37.30	11.36	4.92	7.9	60	2.08	---	---	---	0.015[8]
Santa Ana Delhi[3]	17.6	8.71	4.17	16.0	40	1.73	---	---	---	0.053[9] 0.040[10]
Westminster[3]	6.7	5.65	1.39	13	40		---	---	---	0.079[9] 0.040[10]
El Modena-Irvine[3]	11.9	6.34	2.69	52	40	0.78	---	---	---	0.028[9]
Garden Grove-Wintersberg[1]	20.8	11.74	4.73	10.6	64	1.98	---	---	---	---
San Diego Creek[1]	36.8	14.2	8.52	95.0	20	1.39	---	---	---	---

Notes 1: Watershed Geometry based on review of quadrangle maps and LACFCD strom drain maps.
2: Watershed Geometry based on COE LACDA Study.
3: Watershed Geometry based on COE Reconstitution Study for Santa Ana Delhi and Westminster Channels (June, 1983).
4: Area reduced 57% due to several debris basins and Eaton Wash Dam reservoir, and groundwater recharge ponds.
5: Area reduced 3% due to debris basin.
6: Area reduced 14% due to several debris basins.
7: 0.013 basin factor reported by COE (subarea characteristics, June, 1984).
8: 0.015 basin factor assumed due to similar watershed values of 0.015.
9: Average basin factor computed from reconstitution studies
10: COE recommended basin factor for flood flows.
11: COE = U.S. Army Corps of Engineers.
12: LACDA = Los Angeles County Drainage Area Study by COE.
13: LACFCD = Los Angeles County Flood Control District.

TABLE 2. PRECIPITATION GAUGES USED IN LOS ANGELES COUNTY FLOOD RECONSTITUTIONS

Stream Gauge Location	Storm Reconstitution	LACFCD Rain Gauge No.#
Alhambra Wash near Klingerman Street	Feb 78	191, 303, 1114B
	Mar 78	191, 303, 1114
	Feb 80	191B, 235, 280C, 1014
Compton Creek near Greenleaf Drive	Feb 78	116, 291
	Mar 78	116, 291
	Feb 80	116, 291, 716
Limekiln Creek above Aliso Creek	Feb 78	57A, 446
	Feb 80	259, 446
San Jose Creek Channel above Workman Mill Rd.	Feb 78	92, 1078, 1088X
	Feb 80	96CE, 347E, 1088
Sepulveda Dam (inflow)	Feb 78	57A, 292DE, 446, 735H
	Mar 78	57A, 435, 762
	Feb 80	292, 446, 735
Verdugo Wash at Estelle Ave.	Feb 78	280C, 373C, 498, 758

*No.	Station Name	Lat.	Long.	Elev.	Type
L057A	Camp Hi Hill (OPIDS)	34-15-18	118-05-41	4240	SR
L0092	Claremont-Pomona College	34-05-48	117-42-33	1185	SR
L0096CE	Puddingstone Dam	34-05-31	117-48-24	1030	SR
L0116	Inglewood Fire Station	33-47-53	118-21-22	153	SR
L0191(B)	Los Angeles-Alcazar	34-03-46	118-11-54	400	SR
L0235	Henninger Flats	43-11-38	118-05-17	2550	SR
L0259	Chatsworth-Twin Lakes	34-16-43	118-35-41	1275	SR
L0280C	Sacred Heart Academy	34-10-54	118-11-08	1600	R
L0291	Los Angeles-96th & Central	33-56-56	118-15-17	121	R
L0292(DE)	Encino Reservoir	34-08-56	118-30-57	1075	SR
L0303	Pasadena-Cal Tech	34-08-14	118-07-25	800	SR
L0347E	Baldwin Park-Exp. Station	34-05-56	117-57-40	384	SR
L0373C	Briggs Terrace	34-14-17	118-13-27	2200	SR
L0435	Monte Nido	34-04-41	118-41-35	600	SR
L0446	Aliso Canyon-Oat Canyon	34-18-53	118-33-25	2367	SR
L0498	Angeles Caest Hwy-Drk Cny Tr	34-15-21	118-11-45	2800	R
L0716	Los Angeles-Ducommun Street	34-03-09	118-14-13	306	SR
L0735(H)	Bell Canyon	34-11-40	118-39-23	895	R
L0758	Griffith Park-Lower Spr Cyn	34-08-02	118-17-27	600	R
L0762	Upperstone Canyon	34-07-27	118-27-15	943	R
L1014	Rio Hondo Spreading	33-59-57	118-06-04	170	SR
L1078	Covina-Griffith	34-04-10	117-50-47	975	SR
L1088(X)	LaHabra Hts-Mut Water Co	33-56-55	117-57-51	445	SR
L1114(B)	Whittier Narrows Dam	34-01-29	118-05-02	239	SR

S = Standard 8" raingauge (non-recording).
R = Recording raingauge.

variations among all the catchment $\eta^i(s)$ summation graph realizations are normalized and assembled together to form one regionalized distribution of summation graph realizations.

To describe the data, a "shape" scaling parameter, Y, is introduced by plotting each summation graph realization on Fig. 4 and averaging the upper and lower reading for Y. The regionalized marginal distribution for the parameter Y is shown in Fig. 5. With the normalization process, the variations in the timing parameter, lag^i, and the summation graph total mass (i.e., ultimate discharge, U^i), must be also accounted, and were assumed to be distributed according to the sampled frequency-distributions of Figs. 6 and 7. From these descriptor variables, each $\eta^i(s)$ is represented, in summation graph form, by the parameter values of $\{lag^i, U^i, Y^i\}$.

Based upon the model M defined by Eqs. (42)-(45), a severe storm of March 1, 1983 (which was not used in the development of $[\eta(s)]$) is analyzed for the Alhambra Wash stream gauge. The outcomes of $[Q_M{}^D(t)]$ are plotted along with the recorded stream gauge data in Fig. 8. From the figure, the uncertainty in the model prediction of $[Q_M{}^D(t)]$ is significant, and should be included when analyzing an operator A on the runoff predictions.

If the underlying distribution between the variables in $\{Y, lag, mass\}$ could be ascertained, the distribution of $[\eta(s)]$ could be identified (associated to F). One simple approach is to assume each of the variables to be independent, and to generate $[Q_M{}^D(t)]$ by probabilistic modeling.

DISCUSSION AND CONCLUSIONS

In the use of the above rainfall-runoff model, $M^D(t)$ is given by

$$M^D(t) = E[Q_M{}^D(t)] = \lambda \int_{s=0}^{t} P_g{}^D(t-s)\, E[\eta(s)]\, ds \tag{46}$$

The model uncertainty is then evaluated from Eq. (43) by

$$E_M{}^D(t) = \lambda \int_{s=0}^{t} (P_g{}^D(t-s))\, ([\eta(s)] - E[\eta(s)])\, ds \tag{47}$$

where from Eq. (29),

$$[h_M(s)] = [\eta(s)] - E[\eta(s)] \tag{48}$$

Hence it is seen that the distribution of $[\eta(s)]$ includes the effects of both the rainfall-runoff model itself and the associated uncertainty, where from Eqs. (38) and (40),

$$E[\eta(s)] = \sum_{j=1}^{m} \sum_{<\ell>_j} a^i{}_{<\ell>_j} \sum_k \lambda_{jk} (1 + E[X_{jk}]) \phi_j (s - E[\theta_{jk}] - \alpha^i{}_{<\ell>_j})$$

$$\text{for } F^D(t) \in [\xi_D] \tag{49}$$

The various stochastic distributions utilized are estimated from regional rainfall-runoff data and the chosen model structure. Because runoff data are available for the precise catchment point under study (i.e., we have a stream gauge), the various distributions represented by Figs. 2 through 7 can be rescaled to correspond to the selected study point (because from the stream gauge data being studied, we can estimate the expected value for lag and ultimate discharge). However, in order to utilize these distributions at ungauged points in the catchment, or at other catchments where there are no runoff data, a method of transferring these distributions is needed. That is, a method is needed for estimating the expected values for lag and ultimate discharge (or other description variables used) for the point under study. Given these estimates, the various distributions can be rescaled, and a distribution $[\eta(s)]$ can be estimated from the rainfall-runoff data pool.

In a subsequent paper, the above results will be generalized in order to facilitate use of the model uncertainty distribution (such as Eq. (48)) at other catchments, including catchments that have no runoff data to develop site specific information.

REFERENCES

1. Becker, A. and Kundezewicz, Z.W., Nonlinear Flood Routing with Multilinear Models, Water Resources Research, Vol. 23, No. 6, pp. 1043-1048, 1987.

2. Huff, F.A., Spatial Distribution of Rainfall Rates, Water Resources Research, Vol. 6, No. 1, pp. 254-260, 1970.

3. Hromadka II, T.V., McCuen, R.H., and Yen, C.C., Computational Hydrology in Flood Control Design and Planning, Lighthouse Publications, 1987.

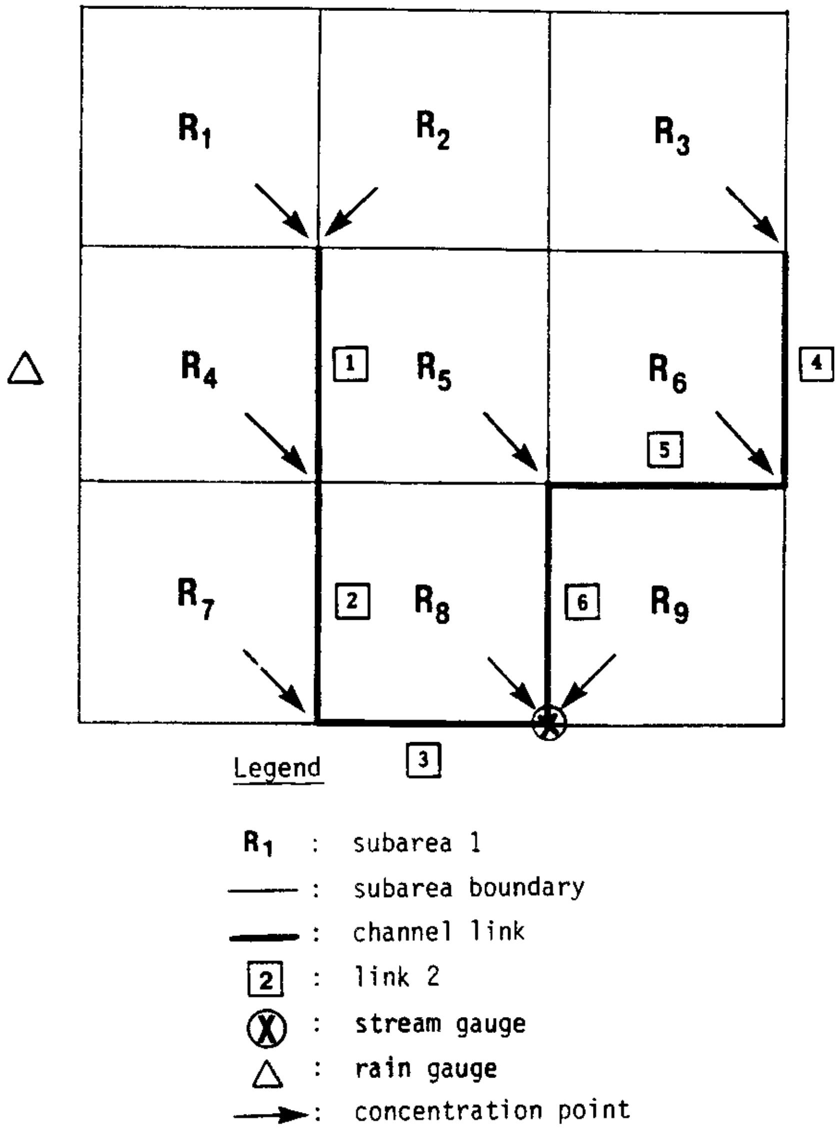
4. Hromadka II, T.V., Approximating Rainfall-Runoff Modeling Response Using a Stochastic Integral Equation, Communications in Statistics, Simulations & Computation, (1989), in-review.

5. Mockus, V.J., National Engineering Handbook, Section 4, Hydrology, U.S. Dept. of Agriculture, SCS, Washington, DC, 1972.

6. Nash, J. and Sutcliffe, J., River Flow Forecasting Through Conceptual Models Part 1 - A Discussion of Principles, Journal of Hydrology, Vol. 10, pp. 282-290, 1970.

7. Tsokos, C.P. and Padgett, W.J., Random Integral Equations with Applications to Life, Sciences and Engineering, Academic Press, Vol. 108, Mathematics in Science and Engineering, 1974.

8. Doyle, W.H., Shearman, J.O., Stiltner, G.J. and Krug, W.R., A Digital Model for Streamflow Routing by Convolution Methods, U.S.G.S. Water Res. Investigations Report 83-4160, 1983.



R1
R2
R3
R4
R5
R6
R7
R8
R9
1
2
3
4
5
6
Legend
R1 : subarea 1
: subarea boundary
: channel link
2 : link 2
X : stream gauge
: rain gauge
: concentration point

Figure 1.

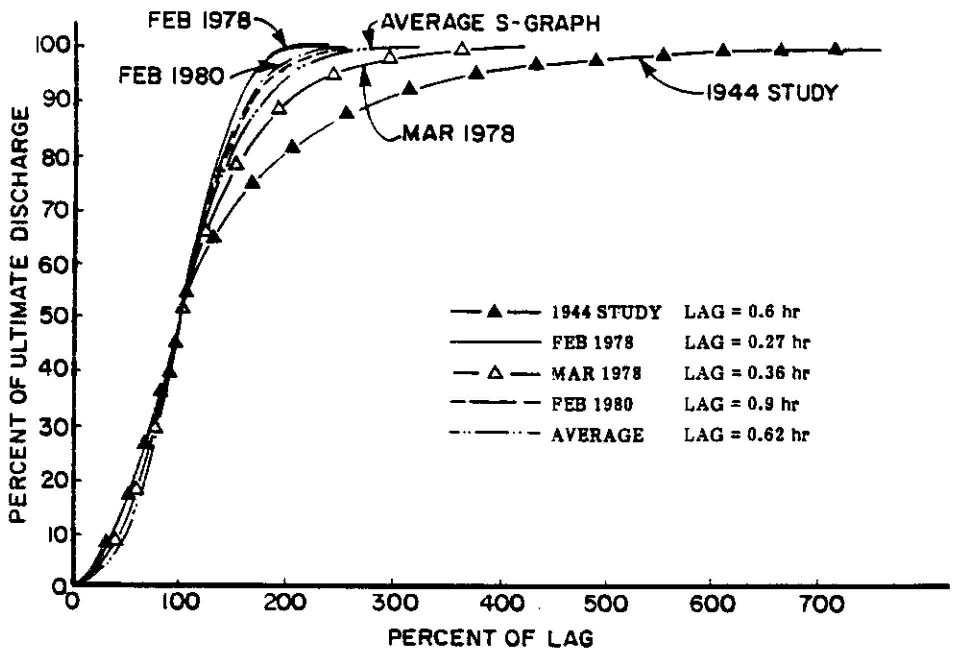


Figure 2.

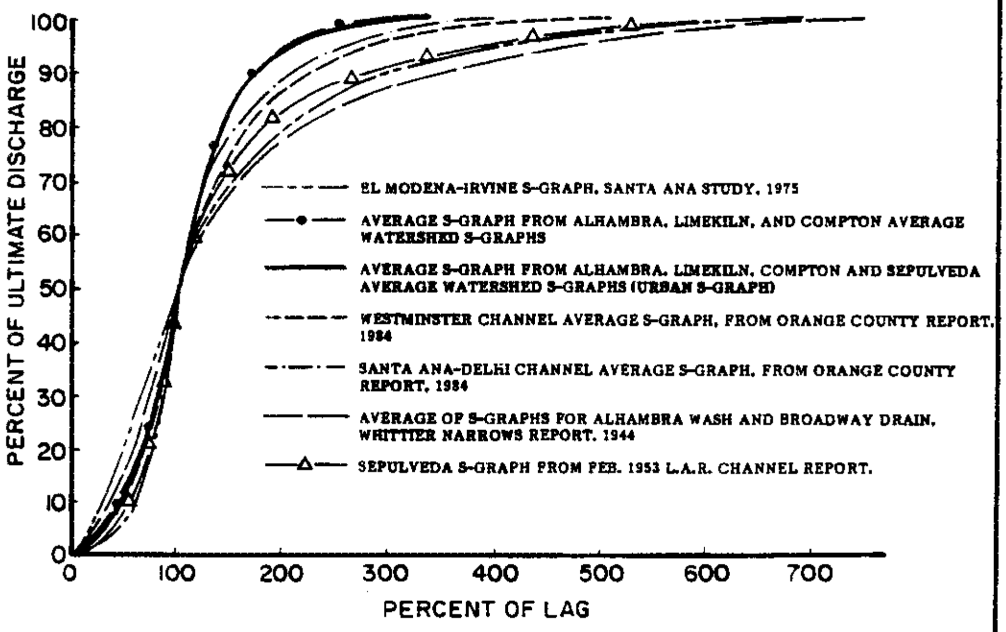


Figure 3.

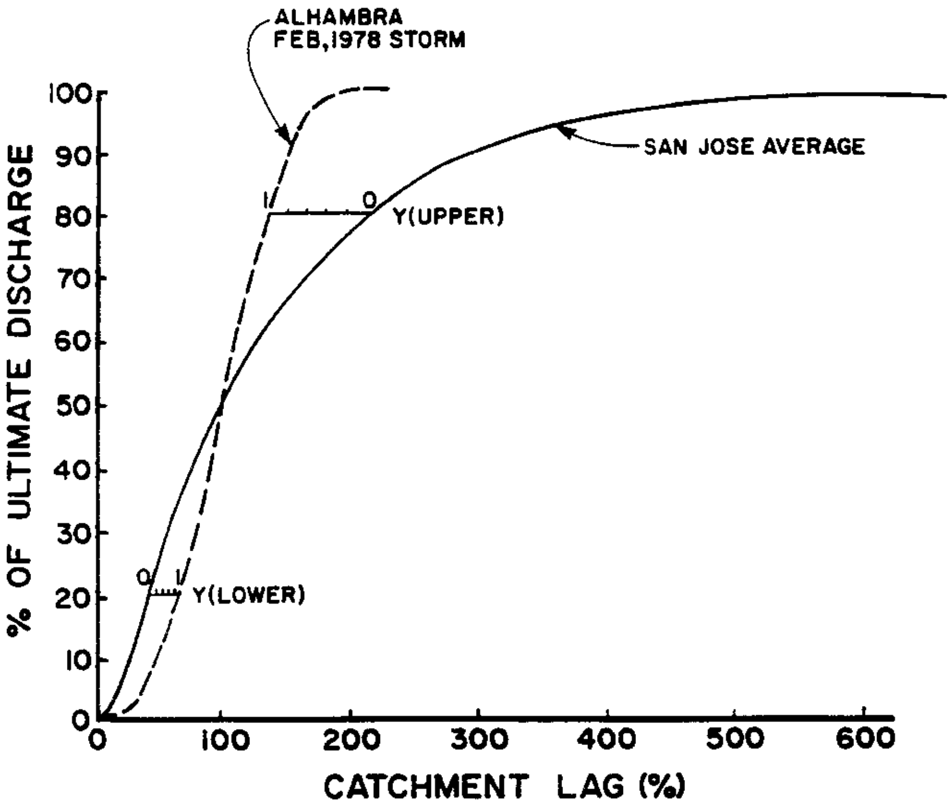


Figure 4.

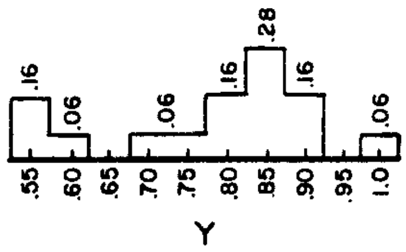


Figure 5.

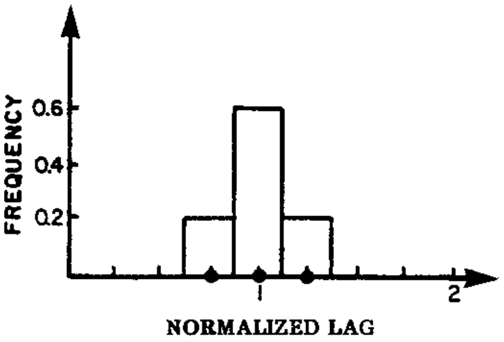


Figure 6.

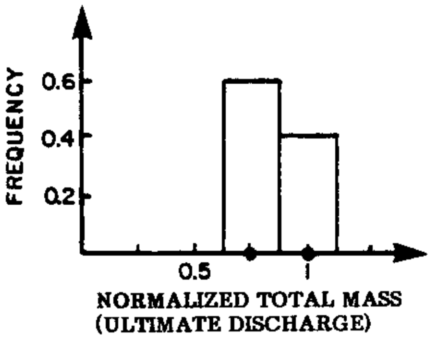


Figure 7.

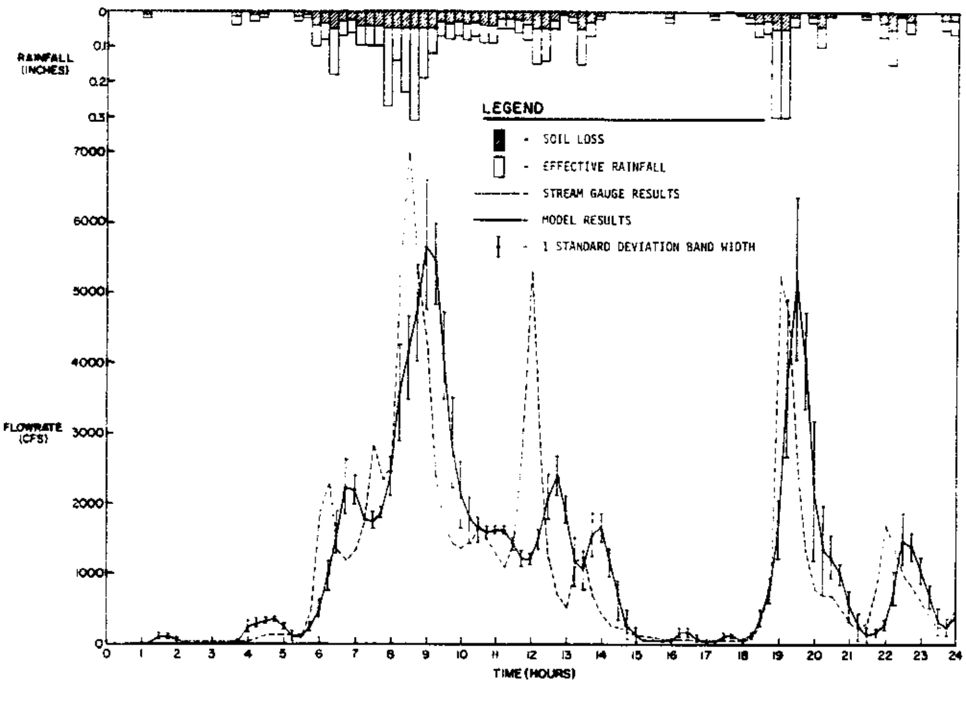


Figure 8.