

APPROXIMATING RAINFALL-RUNOFF MODELLING RESPONSE USING A STOCHASTIC INTEGRAL EQUATION

T. V. HROMADKA II

Department of Mathematics, California State University, Fullerton, CA 92634, USA

AND

R. J. WHITLEY

Department of Mathematics, University of California, Irvine, CA 92717, USA

ABSTRACT

Rainfall-runoff modelling uncertainty can be analysed by the use of a stochastic integral formulation. The stochastic integral equation can be based on the rainfall-runoff model input of model rainfall or model rainfall excess. Similarly, the stochastic integral equation can be based on the rainfall-runoff model output of the modelled runoff hydrograph. The residual between actual measured runoff data and modelled runoff (from the rainfall-runoff model) is analysed here by the use of a stochastic integral equation. This approach is used to develop a set of convolution integral transfer function realizations that represent the chosen rainfall-runoff modelling error. The resulting stochastic integral component is a distribution of possible residual outcomes that may be directly added to the rainfall-runoff model's deterministic outcome, to develop a distribution of probable runoff hydrograph realizations from the chosen rainfall-runoff model.

KEY WORDS uncertainty; rainfall-runoff models; hydrology; modelling; stochastic; stochastic integrals

INTRODUCTION

The state of the art in current rainfall-runoff models is to use computers to approximately solve the various complex mathematical partial differential equations (PDEs) that describe the hydrological cycle as distributed over the catchment (or watershed) and to approximately solve the flood flow timing PDE involved in conduit flow routing (i.e. time varying flow effects in streams, channels, pipes or other structures). Empirical equations are used to describe the modelling components of evaporation, plant transpiration, infiltration of moisture into the soil, percolation of soil moisture into deeper soils and ponding of water, among other effects. The hydraulic effects of flood flow routing in streams and conduits are described by the non-linear PDEs known as the Navier-Stokes equations, but are approximated by simplified algorithms such as the kinematic wave, Muskingum, convex, diffusion or other flow routing techniques (see Hromadka *et al.*, 1987b). The main thrust in computer modelling of the rainfall-runoff process is to subdivide the catchment into smaller subcatchments (or subareas) that are 'linked' together by the hydraulic flow routing models used to represent storm flow in streams and channels. Each subarea is assumed to have a representative rainfall-runoff response, described by a set of hydrological parameters and equations. The subarea runoff, which is assumed to depend only on the rainfall history (and subarea hydrological cycle characteristics), is generally assumed to concentrate at a 'nodal point'. The assemblage of all these links and nodes forms the catchment 'link-node' rainfall-runoff model.

The main objectives of this work is presented in four parts, as follows:

1. Development of a generalized stochastic integral equation representation of rainfall-runoff. Although over 100 link-node modelling techniques are currently reported, almost all these rainfall-runoff

Let each of the 40 plots' runoff be equated to $q_{41}^i(t)$ by use of the random variables $[\lambda_j]$ and $[\theta_j]$ for each R_j where for each storm event, i

$$q_j^i(t) = \lambda_j^i q_{41}^i(t - \theta_j^i) \tag{2}$$

where it is assumed in Equation (2) that the variations between a $q_j^i(t)$ and $q_{41}^i(t)$ are only first order with respect to magnitude and timing. The frequency distributions of $[\lambda_j]$ and $[\theta_j]$ are developed from a large collection of values determined from Equation (2). It is noted that the several random variables may be all mutually dependent.

The stream gauge runoff, $Q_g^i(t)$ can be written using Equation (2) as

$$Q_g^i(t) = \sum_{j=1}^{40} \lambda_j^i q_{41}^i(t - \tau_j - \theta_j^i) \tag{3}$$

where the set of values $\{\lambda_j^i\}$ and $\{\theta_j^i\}$ are samples of the corresponding random variables.

Our analysis now turns to the important problem of prediction. Assuming a hypothetical storm event to occur at the study site, resulting in the rainfall $P_g^D(t)$ and the plot R_{41} runoff, $q_{41}^D(t)$, what would be the estimate of runoff at the stream gauge? Because we are in a prediction mode, the values for each λ_j^D and θ_j^D are unknown for $j = 1, 2, \dots, 40$, which are samples of mutually dependent random variables distributed as $[\lambda_j]$ and $[\theta_j]$, respectively. Then our estimate for runoff at the stream gauge is the stochastic process $[Q_g^D(t)]$ where

$$[Q_g^D(t)] = \sum_{j=1}^{40} [\lambda_j] q_{41}^D(t - \tau_j - [\theta_j]) \tag{4}$$

In Equation (4), it is understood that the various distributions $\{[\lambda_j]; [\theta_j]; j = 1, 2, \dots, 40; [\theta_j], j = 1, 2, \dots, 40\}$ may be all mutually dependent. Also, the 'measured' $q_{41}^D(t)$ is used to develop $[Q_g^D(t)]$ to simplify the presentation; the $F[P_g^D(t)]$ could also have been used.

Stochastic integral equations in rainfall-runoff modelling

The work of Hjelmfelt and Burwell (1984) is recast into an idealized situation where our study catchment, R , can be subdivided into m equally sized small subareas, $R_j, j = 1, 2, \dots, m$, with each subarea being nearly identical in its rainfall-runoff properties. Additionally, at the rain gauge site another such small subarea, R_{m+1} , is specified and monitored so that for each storm event, i , the rainfall and runoff from that subarea are both measured (assume the rain gauge is placed in the centre of the subarea) — that is, for each storm event, i , we obtain the data $P_g^i(t)$ and $e_g^i(t)$, which are the measured rainfall and effective rainfall data, respectively, from the rain gauge site (see Figure 1). We assume that all subareas satisfy the cited Hjelmfelt and Burwell (1984) similarity criteria.

The effective rainfall distribution over subarea j , for storm i , is noted by $e_j^i(t)$. Assuming that for storm i there are characteristic travel times for translation channel routing, the runoff hydrograph at the stream gauge, $Q_g^i(t)$, equates to the m -subarea contributions by

$$Q_g^i(t) = Q_m^i(t) = \sum_{j=1}^m q_j^i(t - \tau_j^i) \tag{5}$$

where τ_j^i is the sum of the characteristic travel times for all channel links which connect subarea j to the catchment R stream gauge; and $Q_m^i(t)$ is the m subarea rainfall-runoff model estimate of runoff for storm event i . We now expand on the elements used in Equation (5).

Subarea effective rainfall, $e_j^i(t)$

Subarea j effective rainfall, $e_j^i(t)$, is unknown because there is neither a stream gauge nor rain gauge in R_j . Assuming that $e_j^i(t)$ can be written as a linear combination of translates of the available data $e_g^i(t)$ gives

$$e_j^i(t) = \sum_{k=1}^{n_j^i} \lambda_{jk}^i e_g^i(t - \theta_{jk}^i) \tag{6}$$

By a change of variables

$$q_j^i(t) = \int_{s=0}^t e_j^i(t-s) \sum_{k=1}^{n_j^i} \lambda_{jk}^i \phi_j^i(s - \theta_{jk}^i) ds \quad (9)$$

Combining Equations (5) and (9) gives the $Q_m^i(t)$ estimate for the runoff hydrograph at the stream gauge

$$Q_m^i(t) = \sum_{j=1}^m \int_{s=0}^t e_g^i(t-s) \sum_{k=1}^{n_j^i} \lambda_{jk}^i \phi_j^i(s - \theta_{jk}^i - \tau_j^i) ds \quad (10)$$

$$= \int_{s=0}^t e_j^i(t-s) \sum_{j=1}^m \sum_{k=1}^{n_j^i} \lambda_{jk}^i \phi_j^i(s - \theta_{jk}^i - \tau_j^i) ds \quad (11)$$

Stochastic integral equation formulation

Equation (11) can be written as a stochastic integral equation

$$Q_m^i(t) = Q_i^i(t) = \int_{s=0}^t e_g^i(t-s) \eta^i(s) ds \quad (12a)$$

where from Equation (11)

$$\eta^i(s) = \sum_{j=1}^m \sum_{k=1}^{n_j^i} \lambda_{jk}^i \phi_j^i(s - \theta_{jk}^i - \tau_j^i) \quad (12b)$$

In Equation (12), $\eta^i(s)$ is a transfer function, for storm i , for the entire catchment. Consequently, given a set of storm effective rainfalls, $\{e_g^i(t)\}$, there is an associated set of realizations, $\{\eta^i(s)\}$, which not only represent the several unknown variations in hydraulic response in R [represented in Equation (12b) by the parameters $\phi^i(s)$ and τ_j^i], but also the several variations in the effective rainfall distribution (i.e. the hydrological response) over R [represented in Equation (12b) by the parameters λ_{jk}^i , θ_{jk}^i , n_j^i]. Because all of these uncertainties and variations cannot be evaluated without a supply of rainfall-runoff data for each subarea and channel hydraulic link used in $Q_m^i(t)$, the modelling output of $Q_i^i(t)$ must be, in a predictive mode, considered a stochastic process. Given a design (or predicted) effective rainfall distribution at the rain gauge site of $e_g^D(t)$, then the model output is a stochastic process, $[Q_i^D(t)]$, where

$$[Q_i^D(t)] = \int_{s=0}^t e_g^D(t-s) [\eta(s)] ds \quad (13)$$

where $[\eta(s)]$ is the stochastic process with realizations developed from Equation (12). In Equation (13), the brackets are notation for a random or stochastic process. In Equation (13), $[\eta(s)]$ is the distribution of transfer functions developed by inserting the mutually dependent distributions of $[\lambda_{jk}]$, $[\theta_{jk}]$, $[\phi_{jk}(s)]$, $[\tau_j]$, $[\eta_j]$, into Equation (12b). The last result is important because even though the individual distributions used in Equation (12b) cannot be evaluated (due to the lack of flow data), the effect of the several interdependent random processes are properly represented by the distribution of transfer functions, $[\eta(s)]$, used in Equation (13).

Effects of channel routing

In this section, the development leading to the rainfall-runoff models of Equations (11) and (12) is extended to include the effects of unsteady flow routing due to channel storage effects. Channel routing effects are generally considered to be important, and this has fuelled the proliferation of rainfall-runoff models. Let $I_1(t)$ be the inflow hydrograph to a channel flow routing link (number 1) and $O_1(t)$ the outflow

For the above linear approximations for storm i , Equations (6), (9) and (18) can be combined to give the final form for our rainfall-runoff model

$$Q_m^i(t) = \sum_{j=1}^m \sum_{(l)_j} a_{(l)_j}^i \int_{s=0}^t e_g^i(t-s) \sum_{k=1}^{n_j^i} \lambda_{jk}^i \phi_j^i(s - \theta_{jk}^i - \alpha_{(l)_j}^i) ds \tag{20}$$

Because the measured effective rainfall distribution, $e_g^i(t)$, is independent of the model, Equation (20) is rewritten as

$$Q_m^i(t) = \int_{s=0}^t e_g^i(t-s) \sum_{j=1}^m \sum_{(l)_j} a_{(l)_j}^i \sum_{k=1}^{n_j^i} \lambda_{jk}^i \phi_j^i(s - \theta_{jk}^i - \alpha_{(l)_j}^i) ds \tag{21}$$

where all parameters are evaluated on a storm by storm basis.

We now consider an important extension of Equation (21). Suppose a simple storm classification system is defined where the effective rainfall distribution measured at the rain gauge, $e_g^i(t)$, can be classified as being in one of three categories: (1) severe; (2) moderate; or (3) minor. Thus, if $e_g^i(t)$ is a class 1 storm, we would expect all channel links to be flowing close to capacity due to high runoffs throughout the catchment. All routing parameters are defined as class 1 parameters and

$$Q_m^1(t) = \int_{s=0}^t e_g^1(t-s) \sum_{j=1}^m \sum_{(l)_j} a_{(l)_j}^1 \sum_{k=1}^{n_j^1} \lambda_{jk}^1 \phi_j^1(s - \theta_{jk}^1 - \alpha_{(l)_j}^1) ds \tag{22}$$

where the subarea transfer functions are similarly defined as being class 1 types. (It is noted that the use of superscript 1 indicates values dependent on storm class, and not $i = 1$.)

Suppose, in prediction, we are interested in the probable runoff at the stream gauge for a hypothetical storm event that is considered to be in storm class 1. Then the estimate of runoff is similar to the results of Equation (22), except that now we have a distribution of outcomes, represented by the stochastic process

$$[Q_m^D(t)] = \int_{s=0}^t e_g^D(t-s) \sum_{j=1}^m \sum_{(l)_j} a_{(l)_j}^1 \sum_{k=1}^{n_j^1} [\lambda_{jk}^1] \phi_j^1(s - [\theta_{jk}^1] - \alpha_{(l)_j}^1) ds \tag{23}$$

where $[\lambda_{jk}^1]$ and $[\theta_{jk}^1]$ are distributions of (possibly mutually dependent) random variables, which are now assumed to have a different probability distribution depending on the storm class.

Stochastic integral representation

The rainfall-runoff model of Equation (23) can be written as a set of stochastic integral equations which provide a variation in prediction due to the storm class system

$$[Q_m^D(t)]_\beta = \int_{s=0}^t e_g^D(t-s) [\eta(s)]_\beta ds \tag{24}$$

where $[\eta(s)]_\beta$ is the stochastic process of catchment transfer functions, associated with storm class β , when $e_g^D(t)$ is in storm class β ; and where $[\eta(s)]_\beta$ equates to the totality

$$[\eta(s)]_\beta = \sum_{j=1}^m \sum_{(l)_j} a_{(l)_j}^\beta \sum_{k=1}^{n_j^\beta} [\lambda_{jk}^\beta] \phi_j^\beta(s - [\theta_{jk}^\beta] - \alpha_{(l)_j}^\beta) ds \tag{25}$$

From Equation (24), the effects of the uncertainty in the effective rainfall over R , and the randomness in the unsteady flow channel routing parameters, are all properly integrated into the $[\eta(s)]_\beta$ realizations, which reflect the combined mutually dependent distributions of all the considered hydrological and hydraulic effects. This last result is important because almost all rainfall-runoff models in use today can be written

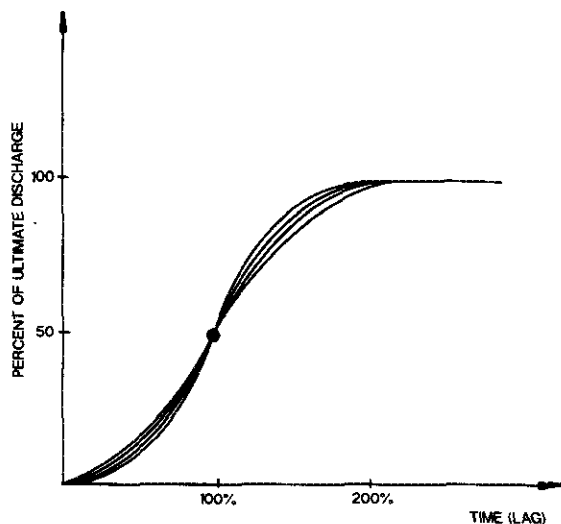


Figure 3. S graphs, $S_0^i(s)$, for storm class $\langle \xi_0 \rangle$

Based on the above normalizations, each summation graph, $M_0^i(s)$, is identified by the parameter set $P_0^i \equiv \{\text{lag}^i, U_0^i, Y^i\}$. Consequently, each realization, $\eta_0^i(s)$, is identified by the vector, P_0^i , for $i = 1, 2, \dots, n_0$, where again n_0 is the number of elements in storm class $\langle \xi_0 \rangle$.

The components of the parameter sets can be considered as random variables which are all mutually dependent. The marginal distributions are developed by plotting the frequency distributions of each component in the parameter set (see Figure 5).

The relative frequency estimate associated with vector, P_0^i , is given by the probability of $Pr(P_0^i)$ where

$$Pr(P_0^i) = Pr(\text{lag}^i, U_0^i, Y^i) \tag{28}$$

It is noted that the $Pr(P_0^i) \neq Pr(\text{lag}^i)Pr(U_0^i)Pr(Y^i)$ as the parameters are not mutually independent. Similarly, the distribution of realizations $[\eta(s)]_\beta$ in Equation (25) is not determined by letting the various parameters $[\lambda_{jk}^\beta]$ and $[\theta_{jk}^\beta]$ vary independently. The $[\eta(s)]_\beta$ frequency distribution properly provides the inherent variability in the rainfall-runoff model.

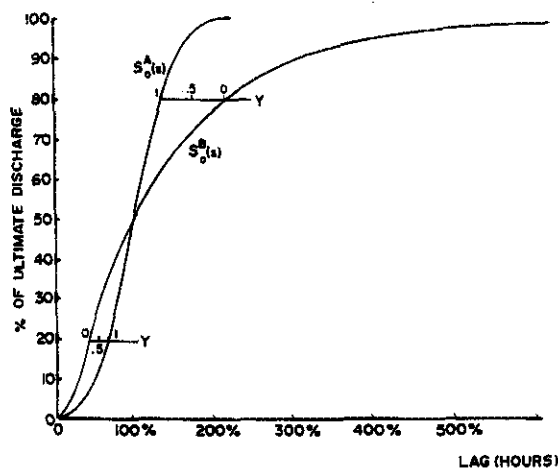


Figure 4. Definition of S graph parameter Y using $S_0^A(s)$ and $S_0^B(s)$

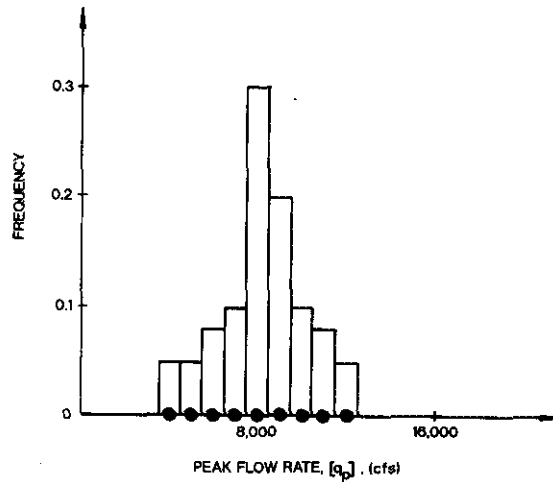


Figure 6. Frequency distribution for the estimate of peak flow-rate $[q_p]$

distributions shown in Figure 5, according to the mutually dependent probability of occurrence given by Equation (28). By scanning the entire set of $[\eta(s)]_0$ realizations developed for storm class $\langle \xi_0 \rangle$, the q_p discrete frequency distribution is constructed.

ESTIMATING UNCERTAINTY IN RAINFALL-RUNOFF MODELLING ESTIMATES

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

$$M : P_g^i(t) \rightarrow M^i(t) \tag{32}$$

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.

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

$$Q_g^i(t) = M^i(t) + E_m^i(t) + E_d^i(t) + E_r^i(t) = M^i(t) + E^i(t) \tag{33}$$

where $E_m^i(t)$ is the modelling 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 'inexplicable' error, such as due to the unknown variation of effective rainfall (i.e. rainfall less losses; rainfall excess) over the catchment, among other factors; and where

$$E^i(t) = E_m^i(t) + E_d^i(t) + E_r^i(t) \tag{34}$$

Because $E^i(t)$ depends on the model M used in Equation (32), then Equations (33) and (34) are combined as

$$Q_g^i(t) = M^i(t) + E_M^i(t) \tag{35}$$

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

Consequently, our final model structure can be used to study the effect on the runoff prediction (at the stream gauge) from arbitrary model M .

Stochastic integral equations and uncertainty estimates

The distributed parameter rainfall-runoff model of Equation (41) provides a useful approximation of almost any rainfall-runoff model in use today. A stochastic integral equation that is equivalent to Equation (41) for each event in the subject storm class is

$$[Q_M^D(t)] = \int_{s=0}^t F^D(t-s)[\eta(s)]_\beta ds \tag{43}$$

where now $[\eta(s)]_\beta$ is the stochastic process representing the various random variations defined on a storm class basis. (It is recalled that on a storm class basis, the hydraulic parameters of $a_{(l)}$, and $\alpha_{(l)}$, and the $\phi_j(s)$, may be assumed to not vary.) In prediction, the expected runoff estimate for storm events that are elements of the subject storm class is

$$E[Q_M^D(t)] = \int_{s=0}^t F^D(t-s)E[\eta(s)]_\beta ds \tag{44}$$

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 modelling approach in use today.

Then the model M structure of Equation (41), when unbiased, is given from Equation (44), by

$$M^D(t) = E[(Q_M^D(t))] \tag{45}$$

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

$$[E_M^D(t)] = [Q_M^D(t)] - E[Q_M^D(t)] \tag{46}$$

where all equations are defined on the storm class basis used in the previous equations.

APPLICATION OF THE STOCHASTIC INTEGRAL EQUATION

In our application problem, the model input functional $F : P_g^i(t) \rightarrow F^i(t)$ is specified as simply a yield type relationship (e.g. Hromadka and Whitley, 1988)

$$F : P_g^i(t) \rightarrow \lambda P_g^i(t) \tag{47}$$

where λ is a constant runoff coefficient typically estimated as a ratio of rainfall divided by runoff, for storm events in a specified class. The corresponding stochastic integral equation used to relate the rainfall-runoff data is

$$Q_g^i(t) = \lambda \int_{s=0}^t P_g^i(t-s)\eta^i(s) ds \tag{48}$$

In this application, storm classes are defined according to the 85 centile value of rainfall intensity in excess of one-half of the maximum five-minute mean intensity, z , and also according to the total rainfall mass which occurs within three days before 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, 16 stream gauges and 24 rain gauges were studied for catchment characteristics. Because of the sparcity of rainfall-runoff data, several catchments are considered to regionalize the statistical results. All storms considered are assumed to be elements of the same storm class considered important for flood control, which was found to be the case based on the severity of each storm event.

For each storm event and catchment, the rainfall-runoff data are used to directly develop the set of realizations, $\{\eta^i(s)\}$. 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. Note that

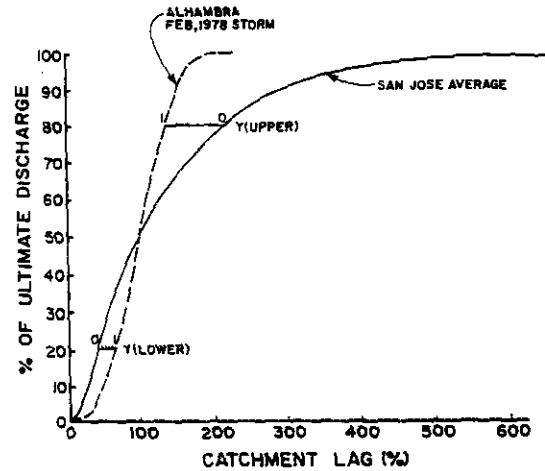


Figure 9. Normalized summation graph shape factor, Y , definition

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 Figure 9 and averaging the upper and lower reading for Y . The regionalized marginal distribution for the parameter Y is shown in Figure 10. 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 normal distribution as fit to the sample data. 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 on the model M defined by Equations (47) and (48), a severe storm of 1 March 1983 (which as not used in the development of $[\eta^i(s)]$) is analysed for the Alhambra Wash stream gauge. The outcomes of $[Q_M^D(t)]$ are plotted along with the recorded stream gauge data in Figure 11. From the figure, the uncertainty in the model prediction of $[Q_M^D(t)]$ is significant and should be included when analysing an operator A on the runoff predictions.

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

$$M^D(t) = E[Q_M^D(t)] = \lambda \int_{s=0}^t P_g^D(t-s) E[\eta(s)] ds \tag{49}$$

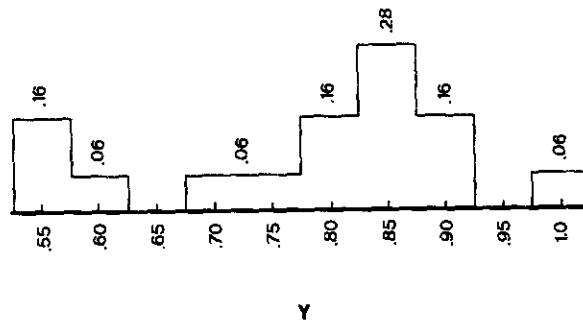


Figure 10. Distribution of Y , regionalized from data from several catchments

distribution of stochastic outcomes (of modelled runoff hydrographs) to the standard catchment unit hydrograph, given on a storm class basis. The work effort involved in developing a stochastic integral equation formulation of a rainfall-runoff model is essentially the same as in using standard unit hydrograph techniques, except that several equally likely computational runs are made (one for each transfer function) instead of just the single expected value unit hydrograph transfer function.

REFERENCES

- Becker, A. and Kundzewicz, Z. W. 1987. 'Nonlinear flood routing with multi-linear models', *Wat. Resour. Res.*, **23**, 1043-1048.
- Doyle, W. H., Shearman, J. O., Stiltner, G. J. and Krug, W. R. 1983. 'A digital model for streamflow routing by convolution methods', *USGS Wat. Res. Invest. Rep.* 83-4160.
- Hjelmfelt, A. and Burwell, R. 1984. 'Spatial variability of runoff', *J. Irrig. Drainage Engin.*, **110**.
- Hromadka II, T. V. and McCuen, R. H. 1986a. *The Orange County Hydrology Manual*. OCEMA, Orange County.
- Hromadka II, T. V. and McCuen, R. H. 1986b. *The San Bernardino County Hydrology Manual*. San Bernardino County.
- Hromadka II, T. V. and Whitley, R. J. 1988. 'The design storm concept in flood control design and planning', *Stochastic Hydrol. Hydr.* **2**, 213-239.
- Hromadka II, T. V., McCuen, R. H., and Yen, C. C. 1987a. *Computational Hydrology in Flood Control Design and Planning*. Light-house.
- Hromadka II, T. V., Yen, C. C., and Pinder, G. F. 1987b. *The Best Approximation Method — An Introduction*. Lecture Notes Series, No. 27 Springer Verlag, Berlin.
- Huff, F. A. 1970. 'Spatial distribution of rainfall rates', *Wat. Resour. Res.*, **6**, 254-260.
- Mockus, V. J. 1972. *National Engineering Handbook, Section 4, Hydrology*. US Department of Agriculture, SCS, Washington.
- Nash, J. and Sutcliffe, J. 1970. 'River flow forecasting through conceptual models part 1 — a discussion of principles', *J. Hydrol.*, **10**, 282-290.
- Tsokos, C. P. and Padgett, W. J. 1974. *Random Equation with Applications to Life, Sciences and Engineering*. Mathematics in Science and Engineering, Vol. 108. Academic Press, New York.