CONFIDENCE INTERVALS FOR FLOOD CONTROL DESIGN

T. V. Hromadka II, M.ASCE*

Abstract

The line drawn and labeled as the flood frequency curve is seldom identified as to what confidence is associated to the plot. For example while flood control designs are typically based on the flood frequency curve Q_{100} estimate, seldom is it considered that, on the average, and with the other factors and parameters being correct, this Q_{100} estimate has only a 50 percent chance of being greater than the true (unknown) Q_{100} but equivalently has a 50 percent chance of being less than the true Q_{100}.

Consequently, if the goal is to provide protection against Q_{100}, and this level of protection is adopted as the local policy statement for all design purposes, and there is liability should Q_{100} flooding occur, then confidence intervals should be incorporated into the flood control policy statement.

INTRODUCTION

In this paper, the unit hydrograph method (UH) is used to develop estimates of runoff modeling error in the frequently occurring cases where the uncertainty in the rainfall distribution over the catchment dominates all other sources of modeling uncertainty. Indeed, the uncertainty in the precipitation distribution appears to be a limiting factor in the successful development, calibration, and application of all surface runoff hydrologic models (e.g., Loague and Freeze, 1985; Beard and Chang, 1979; Schilling and Fuchs, 1986; Garen and Burges, 1981; Nash and Sutcliffe, 1970; Troutman, 1982).

Schilling and Fuchs (1986) write "that the spatial resolution of rain data input is of paramount importance to the accuracy of the simulated hydrograph" due to "the high spatial variability of storms" and "the amplification of rainfall sampling errors by the nonlinear transformation" of rainfall into runoff. They recommend that a model should employ a simplified surface flow model if there are many subbasins; a simple runoff coefficient loss rate; and a diffusion (zero inertia) or storage channel routing technique.

In their study, Schilling and Fuchs (1986) reduced the rainfall data set resolution from a grid of 81 gages to a single catchment-centered gage in an 1,800 acre catchment. They noted that variations in runoff volumes and peak flows "are well above 100 percent over the entire range of storms implying that the spatial resolution of rainfall has a dominant influence on the reliability of computed runoff." It is also noted that "errors in the rainfall input are amplified by the rainfall-runoff transformation so that "a rainfall depth error of 30 percent results in a volume error of 60 percent and a peak flow error of 80 percent." They also write that "it is inappropriate to use a sophisticated runoff model to achieve a desired level of modeling accuracy if the spatial resolution of rain input is low" (in their study, the raingage densities considered for the 1,800-acre catchment are 81, 9, and a single centered gage).

Similarly, Beard and Chang (1979) write that in their study of 14 urban catchments, complex models such as continuous simulation typically have 20 to 40 parameters and

---

* Director of Water Resources Engineering, Williamson and Schmid, Irvine, California and Research Associate, Princeton University, New Jersey

functions that must be derived from recorded rainfall-runoff data. "Inasmuch as rainfall data are for scattered point locations and storm rainfall is highly variable in time and space, available data are generally inadequate for reliably calibrating the various interrelated functions of these complex models."

In the extensive study by Loague and Freeze, (1985), three event-based rainfall-runoff models (a regression model, a unit hydrograph model, and a kinematic wave quasi-physically based model) were used on three data sets of 269 events from three small upland catchments. In that paper, the term "quasi-physically based", or QPB, is used for the kinematic wave model. The three catchments were 25 acres, 2.8 square miles, and 35 acres in size, and were extensively monitored with rain gage, stream gage, neutron probe, and soil parameter site testing. For example, the 25 acre site contained 35 neutron probe access sites, 26 soil parameter sites (all equally spaced), an on-site rain gage, and a stream gage. The QPB model utilized 22 overland flow planes and four channel segments. In comparative tests between the three modeling approaches to measured rainfall-runoff data it was concluded that all models performed poorly and that the QPB performance was only slightly improved by calibration of its most sensitive parameter, hydraulic conductivity. They write that the "conclusion one is forced to draw...is that the QPB model does not represent reality very well; in other words, there is considerable model error present. We suspect this is the case with most, if not all conceptual models currently in use." Additionally, "the fact that simpler, less data intensive models provided as good or better predictions than a QPB is food for thought."

Based on the literature, the main difficulty in the use, calibration, and development, of complex models appears to be the lack of precise rainfall data and the high model sensitivity to (and magnification of) rainfall measurements errors. Nash and Sutcliffe(1970) write that "As there is little point in applying exact laws to approximate boundary conditions, this, and the limited ranges of the variables encountered, suggest the use of simplified empirical relations."

CATCHMENT AND DATA DESCRIPTION

Let R be a free draining catchment with negligible detention effects. R is discretized into m subareas, R_j, each draining to a nodal point which is drained by a channel system. The m-subarea link node model resulting by combining the subarea runoffs for storm i, adding runoff hydrographs at nodal points, and routing through the channel system, is denoted as $Q_m^i(t)$. It is assumed that there is only a single rain gage and stream gage available for data analysis. The rain gage site is monitored for the 'true' effective rainfall distribution, $e_g^i(t)$. The motivation in using a measured $e_g^i(t)$ at the rain gage site is to avoid the necessity of using a multiparameter submodel to approximate $e_g^i(t)$; rather we assume that an accurate value of $e_g^i(t)$ is available, even though this data is measured at the rain gage site which may be located outside of the catchment. The stream gage data represents the entire catchment, R, and is denoted by $Q_g^i(t)$ for storm event i.

LINEAR EFFECTIVE RAINFALLS FOR SUBAREAS

The effective rainfall distribution (rainfall less losses) in R_j is given by $e_j^i(t)$ for storm i where $e_j^i(t)$ is assumed to be linear in $e_g^i(t)$ by:

$$e_j^i(t) = \sum \lambda_{jk}^i \, e_g^i(t - \theta_{jk}^i), \; j = 1,2,\cdots,m \tag{1}$$

where λ_{jk}^i and θ_{jk}^i are coefficients and timing offsets, respectively, for storm i and subarea R_j. In Eq. (1), the variations in the effective rainfall distribution over R due to magnitude and timing are accounted for by the λ_{jk}^i and θ_{jk}^i, respectively. As an alternative to Eq. (1), the $e_g^i(t)$ may be defined as a set of unit effective rainfalls, each unit associated with its own proportion factor; however for simplicity, the use of the entire $e_g^i(t)$ function will be carried forward in the model development.

SUBAREA RUNOFF

The storm i subarea runoff from R_j, $q_j^i(t)$, is given by the linear convolution integral:

$$q_j^i(t) = \int_{s=0}^{t} e_j^i(t - s)\, \phi_j^i(s)\, ds \tag{2}$$

where $\phi_j^i(s)$ is the subarea unit hydrograph (UH) for storm i such that Eq. (2) applies. Combining Eqs. (1) and (2) gives

$$q_j^i(t) = \int_{s=0}^{t} \sum e_g^i(t - \theta_{jk}^i - s)\, \lambda_{jk}^i\, \phi_j^i(s)\, ds \tag{3}$$

Rearranging variables,

$$q_j^i(t) = \int_{s=0}^{t} e_g^i(t - s) \sum \lambda_{jk}^i\, \phi_j^i(s - \theta_{jk}^i)\, ds \tag{4}$$

where throughout this paper, the argument of the arbitrary frunction $F(s - Z)$ is notation that $F(s - Z) = 0$ for $s < Z$.

LINEAR ROUTING

Let $I_1(t)$ be the inflow hydrograph to a channel flow routing link (number 1), and $Q_1(t)$ the outflow hydrograph. A linear routing model of the unsteady flow routing process is given by

$$O_1(t) = \sum_{k_1=1}^{n_1} a_{k_1}\, I_1(t - \alpha_{k_1}) \tag{5}$$

where the a_{k_1} are coefficients which sum to unity; and the α_{k_1} are timing offsets. Again, $I_1(t - \alpha_{k_1}) = 0$ for $t < \alpha_{k_1}$. Given stream gage data for $I_1(t)$ and $O_1(t)$, the best fit values for the a_{k_1} and α_{k_1} can be determined.

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

$$\begin{aligned} O_2(t) &= \sum_{k_2=1}^{n_2} a_{k_2}\, I_2(t - \alpha_{k_2}) \\ &= \sum_{k_2=1}^{n_2} a_{k_2} \sum_{k_1=1}^{n_1} a_{k_1}\, I_1(t - \alpha_{k_1} - \alpha_{k_2}) \end{aligned} \tag{6}$$

For L links, each with their own respective stream gage routing data, the above linear routing technique results in the outflow hydrograph for link number L, $Q_L(t)$, being given by

$$O_L(t) = \sum_{k_L=1}^{n_L} a_{k_L} \sum_{k_{L-1}=1}^{n_{L-1}} a_{k_{L-1}} \cdots \sum_{k_2=1}^{n_2} a_{k_2} \sum_{k_1=1}^{n_1} a_{k_1}\, I_1(t - \alpha_{k_1} - \alpha_{k_2} - \cdots - \alpha_{k_{L-1}} - \alpha_{k_L}) \tag{7}$$

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

$$O_L(t) = \sum_{<k>} a_{<k>} I_1(t - \alpha_{<k>}) \tag{8}$$

For subarea R_j, the runoff hydrograph for storm i, $q_j^i(t)$, flows through L_j links before arriving at the stream gage and contributing to the total measured runoff hydrograph, $Q_g^i(t)$. All of the constants $a^i_{<k>}$ and $\alpha^i_{<k>}$ are available on a storm by storm 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,

$$Q_m^i(t) = \sum_{j=1}^{m} \sum_{<k>_j} a^i_{<k>_j} q_j^i(t - \alpha^i_{<k>}) \tag{9}$$

where each vector $<k>_j$ is associated to a R_j, and all data is defined for storm i. It is noted that in all cases,

$$\sum_{<k>_j} a^i_{<k>_j} = 1 \tag{10}$$

LINK-NODE MODEL, $Q_m^i(t)$, AND MODEL REDUCTION

For the above linear approximations for storm i, Eqs. (1), (4), and (9) can be combined to give the final form for the m subarea link-node model, $Q_m^i(t)$.

$$Q_m^i(t) = \sum_{j=1}^{m} \sum_{<k>_j} a^i_{<k>_j} \int_{s=0}^{t} e_g^i(t - s) \sum \lambda^i_{jk} \phi_j^i(s - \theta^i_{jk} - \alpha^i_{<k>_j}) \, ds \tag{11}$$

Because the measured effective rainfall distribution, $e_g^i(t)$, is independent of the several indices, Eq. (11) is rewritten in the form

$$Q_m^i(t) = \int_{s=0}^{t} e_g^i(t - s) \sum_{j=1}^{m} \sum_{<k>} a^i_{<k>_j} \sum \lambda^i_{jk} \phi_j^i(s - \theta^i_{jk} - \alpha^i_{<k>_j}) \, ds \tag{12}$$

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

Equation (12) described a model which represents the total catchment runoff response based on variable subarea UH's, $\phi_j^i(s)$; variable effective rainfall distributions on a subarea-by-subarea basis with differences in magnitude (λ^i_{jk}), timing (θ^i_{jk}), and pattern shape (linearly assumption); and channel flow routing translation and storage effects (parameters $a^i_{<k>_j}$ and $\alpha^i_{<k>_j}$). All parameters employed in Eq. (12) must be evaluated by runoff data where stream gages are supplied to measure runoff from each subarea, R_j, and stream gages are located upstream and downstream of each channel reach (link) used in the model.

The m-subarea model of Eq. (12) is directly reudced to the simple single area UH model (no discretization of R into subareas) given by $Q_1^i(t)$ where

$$Q_1^i(t) = \int_{s=0}^{t} e_g^i(t - s) \, \eta^i(s) \, ds \tag{13}$$

where $\eta^i(s)$ is the correlation distribution between the data pair $\{Q_g{}^i(t), e_g{}^i(t)\}$, for storm event i.

STORM CLASSIFICATION SYSTEM

To proceed with the analysis, the full domain of effective rainfall distributions measured at the rain gage site are categorized into storm classes, $\langle\xi_x\rangle$. Because the storm classifications are based upon effective rainfalls, the measured precipitations, $P_g{}^i(t)$, may vary considerably yet produce similar effective rainfall distributions. That is, any two elements of a class $\langle\xi_x\rangle$ would result in nearly identical effective rainfall distributions at the rain gage site, and hence one would "expect" nearly identical runoff hydrographs recorded at the stream gage. Typically, however, the resulting runoff hydrographs differ and, therefore, the randomness of the effective rainfall distribution over the catchment, R, results in variations in the modeling "best-fit" parameters (i.e., in $Q_1{}^i(t)$, the $\eta^i(s)$ variations) in correlating the available rainfall-runoff data.

More precisely, any element of a specific storm class $\langle\xi_o\rangle$ has the effective rainfall distribution, $e_g{}^o(t)$. However, there are several runoffs associated to the single $e_g{}^o(t)$, and are noted by $Q_g^{oi}(t)$. In correlating $\{Q_g^{oi}(t), e_g{}^o(t)\}$, a different $\eta^i(s)$ results due to the variations in the measured $Q_g^{oi}(t)$ with respect to the single known input at the rain gage site, $e_g{}^o(t)$.

In the predictive mode, where one is given an assumed (or design) effective rainfall distribution, $e_g{}^D(t)$, to apply at the rain gage site, the storm class of which $e_g{}^D(t)$ is an element of is identified, $\langle\xi_D\rangle$, and the predictive output for the input, $e_g{}^D(t)$, must necessarily be the random variable or distribution,

$$[Q_1{}^D(t)] = \int_{s=0}^{t} e_g{}^D(t - s)\ [\eta(s)]_D\ ds \tag{14}$$

where $[\eta(s)]_D$ is the distribution of $\eta^i(s)$ distributions associated to storm class $[\xi_D]$.

THE VARIANCE OF HYDROLOGIC MODEL OUTPUT

Consider the $Q_1{}^i(t)$ model structure in correlating the single rain gage and stream gage. For storm class $\langle\xi_o\rangle$, there is an associated distribution of correlation distributions, $[\eta(s)]_o$. Then in the predictive mode, the predicted hydrologic model output is the distribution $[Q_1{}^o(t)]$ where

$$[Q_1{}^o(t)] = \int_{s=0}^{t} e_g{}^o(t-s)\ [\eta(s)]_o\ ds \tag{15}$$

For storm time z, the distribution of flow rate values is by $[Q_1{}^o(z)]$, where

$$[Q_1{}^o(z)] = \int_{s=0}^{t} e_g{}^o(z-s)\ [\eta(s)]_o\ ds \tag{16}$$

Let t_p be the storm time where the peak flow rate, Q_p, occurs for storm class $\langle\xi_o\rangle$. Noting that t_p is a function of $[\eta(s)]_o$, then the distribution of $[Q_p]_o$ is given by

$$[Q_p]_o = \int_{s=0}^{t_p} e_g{}^o(t_p - s)\ [\eta(s)]_o\ ds \tag{17}$$

Let D be a single time duration. Of interest is the maximum volume of runoff during duration, D, for storm class $<\xi_0>$. Then the distribution of this estimate is given by

$$[\max \int_D Q_1{}^o(t)dt] = \max \int_D \int_{s=0}^{t} e_g{}^o(t - s)[\eta(s)]_o \, ds \tag{18}$$

Let A be an operator which represents a hydrologic process algorithm (e.g., detention basin, etc.). Then the output of the operator for storm class $<\xi_0>$ is the distribution

$$[A]_o = A \left(\int_{s=0}^{t} e_g{}^o(t - s) \, [\eta(s)]_o \, ds \right) \tag{19}$$

The expected value of the hydrologic process A for storm class $<\xi_o>$ is

$$E[A]_o = \sum_{[\eta(s)]_o} A\left(\int_{s=0}^{t} e_g{}^o(t - s) \, \eta(s) \, ds \right) P(\eta(s)) \tag{20}$$

where $P(\eta(s))$ is the frequency of occurrance for distribution $\eta(s)$ in $[\eta(s)]_o$. The variance of predictions of hydrologic process A for storm class $<\xi_o>$ is (for A ( ) being a mapping into the real number line; i.e., giving a single number result),

$$\text{var}[A]_o = \sum_{[\eta(s)]_o} \left\{ A(\int_{s=0}^{t} e_g{}^o(t-s) \, \eta(s) \, ds) - E[A]_o \right\}^2 P(\eta(s)) \tag{21}$$

CONCLUSIONS

A lower bound for estimating the distribution of uncertainty in surface runoff modeling output is advanced. The bound is based on a linear unit hydrograph approach, which utilizes an arbitrary number of catchment subdivisions into subareas, a linear routing technique for channel flow effects, a variable effective rainfall distribution over the catchment, and calibration parameter distributions developed in correlating rainfall-runoff data by the model. Because all hydrologic parameters (e.g., subarea unit hydrographs, channel routing parameters, effective rainfall distribution factors) vary on a storm basis, the unit hydrograph methodology is a reasonable approximation for assessing uncertainty in hydrologic modeling estimates. The uncertainty bound developed reflects the dominating influence of the unknown rainfall distribution over the catchment and is expressed as a distribution function which can be reduced only by supplying additional rainfall-runoff data. It is recommended that this uncertainty distribution be included in flood control design studies in order to incorporate prescribed levels of confidence in flood protection facilities.

APPENDIX I - REFERENCES

1. Beard, L. and Chang, S., Urbanization Impact on Stream Flow, ASCE, Journal of the Hydraulics Division, June, 1979.
2. Doyle, Jr., W. H., Sherman, J. O., Stiltner, G. J. and Krug, W. R., A Digital Model for Streamflow Routing by Convolution Methods, U.S. Geological Survey, Water-Resources Investigation Report 83-4160, 1983.
3. Loague, K. and Freeze, R., A Comparison of Rainfall-Runoff Modeling Techniques on Small Upland Catchments, Water Resources Research, Vol. 21, No. 2, Feb., 1985.
4. Nash, J. and Sutcliffe, J., River Flow Forecasting Through Conceptual Models, Part I - A Discussion of Principles, Journal of Hydrology, Vol. 10, 1970.
5. Schilling, W. and Fuchs, L., Errors in Stormwater Modeling - A Quantitative Assessment, ASCE Journal of Hydraulic Engineering, Vol. 112, No. 2, Feb., 1986.
6. Troutman, B., An Analysis of Input in Precipitation - Runoff Models Using Regression with Errors in the Independent Variables, Water Resources Research, Vol. 18, No. 4, Aug., 1982.