REPRINTED FROM

HYDROLOGICAL SCIENCE AND TECHNOLOGY

VOLUME 3
NUMBER 1-2 1987

AMERICAN INSTITUTE OF HYDROLOGY

HYDROLOGICAL SCIENCE AND TECHNOLOGY
Volume 3, Number 1-2
American Institute of Hydrology
Pp. 37-45

Use of Subareas in Rainfall-Runoff Models, II: Reducing Modeling Uncertainty

T.V. HROMADKA II

Williamson and Schmid, Irvine, CA 92714 and
Dept. of Mathematics, California State Univ.
Fullerton, Ca 92634

ABSTRACT

The single rain gauge/stream gauge rainfall-runoff modeling problem is important to hydrologists, and yet a definitive analysis is still not available for the analysis of uncertainty. In this paper, an approximation of rainfall-runoff models is provided by use of a multi-linear model. This approximation can then be used for the development of uncertainty estimates in modeling output, and the need for additional data to reduce the variance in the estimates.

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 effective rainfall distribution (i.e., rainfall less losses; rainfall excess) over the catchment dominates all other sources of modeling uncertainty. Indeed, just 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 Change, 1979; Schilling and Fuchs, 1986; Garen and Burges, 1981; Nash and Sutcliffe, 1970; Troutman, 1982). The comparion paper (Hromadka, 1988) develops a multi-linear model of the rainfall-runoff modeling approach, which is useful in describing the mathematical underpinnings of other modeling structures. The multi-linear model provides a convenient and useful means to approximate modeling output uncertainty, and evaluate the need for additional rainfall-runoff data.

EFFECTIVE RAINFALL UNCERTAINTY AND THE DISTRIBUTIONS, $[\eta(s)]_x$

In Hromadka (1988), the Voltera integral is used to relate effective rainfall to runoff by means of the transfer function, $\eta^i(s)$, for storm event i. It was shown that $\eta^i(s)$ includes all the uncertainty in the effective rainfall distribution over the catchment R, as well as the uncertainty in the runoff and flow routing processes. That is, $\eta^i(s)$ must be an element of the random process

$[\eta(s)]_x$ where

$$\eta^i(s) = \sum_{j=1}^{m} \sum_{<k>_j} a^i_{<k>_j} \sum \lambda_{jk}{}^i \phi_j{}^i(s - \theta_{jk}{}^i - \alpha^i_{<k>_j}) \tag{1}$$

where the $a^i_{<k>_j}$ and $\alpha^i_{<k>_j}$ are sets of proportions and timing translates used in linear flow routing; $\lambda_{jk}{}^i$ and $\theta_{jk}{}^i$ are proportions and timing translates due to the variation in effective rainfall between the subarea R_j data and the known data, $e_g{}^i(t)$; $\phi_j{}^i(s)$ is the unit hydrograph; and j is the R_j subarea number, for an m-subarea link-node model; and Eq. (1) applies to storm event i for some storm class $<\xi_x>$. (The reader is referred to Hromadka, 1988, for further notation definition). For severe storms of flood control interest, one would be dealing with only a subset of the set of all storm classes. In a particular storm class, $<\xi_o>$, should it be assumed that the subarea runoff parameters and channel flow routing uncertainties are minor in comparison to the uncertainties in the effective rainfall distribution over R (e.g., Schilling and Fuchs, 1986; among others), then the collection of $\eta^i(s)$ may be written as

$$[\eta(s)]_o = \sum_{j=1}^{m} \sum_{<k>} \bar{a}_{<k>_j} \sum [\lambda_{jk}] \bar{\phi}_j \; (s - [\theta_{jk}] - \bar{\alpha}_{<k>_j}) \tag{2}$$

where the overbars are notation for mean values of the parameters for storm class $<\xi_o>$. But the mean values for the linear routing parameters are essentially the calibrated parameters corresponding to a class of hydrographs which accommodate a range of hydrograph magnitudes. And for a highly discretized catchment model, the use of a mean value UH for each subarea, $\phi_j(s)$, has only a minor influence in the total model results (Schilling and Fuchs, 1986). Equation (2) is useful in motivating the use of the probabilistic distribution concept in design and planning studies for all hydrologic models, based on just the magnitude of the uncertainties in the effective rainfall distribution over R. That is, although one may argue that a particular model is "physically based" and represents the "true" hydraulic response distributed throughout the catchment, the uncertainty in rainfall still remains and is not reduced to increasing hydraulic routing modeling complexity. Rather, the uncertainty in rainfall is reduced only by the use of additional rainfall-runoff data.

DISCRETIZATION ERROR

In the general case, the practitioner generally assigns the recorded precipitation from the single available rain gauge, $P_g{}^i(t)$, to occur simultaneously over each subarea, R_j. That is from Eq. (1), the $\theta_{jk}{}^i \equiv 0$ and the $\lambda_{jk}{}^i$ are set to constants $\hat{\lambda}_j$ which reflect only the variations in loss rate nonhomogeneity. Hence, the "true" $Q_m{}^i(t)$ model becomes the estimator $\hat{Q}_m{}^i(t)$ where

$$\hat{Q}_m{}^i(t) = \int_{s=0}^{t} \hat{e}_g{}^i(t-s) \sum_{j=1}^{m} \sum_{<k>} \hat{a}^i_{<k>_j} \sum \hat{\lambda}_j \, \hat{\phi}_j{}^i(s-\hat{\alpha}^i_{<k>_j}) \, ds \tag{3}$$

where hats are notation for estimates. These incorrect assumptions result in "discretization error". Indeed, an obvious example of discretization error is the case where a subarea R_j actually receives no rainfall, and yet one assumes that $P_g{}^i(t)$ occurs over R_j in the discretized model.

DISCRETIZATION CALIBRATION ERROR

A current trend among practitioners is to develop an m-subarea link-node model estimator $\hat{Q}_m{}^i(t)$ such as Eq. (3), and then "calibrate" the model parameters using the available (single) rain gauge and stream gauge data pair. Because subarea rainfall-runoff data are unavailable, necessarily it is assumed that the random variables associated to the subarea effective rainfalls are given by

$$\left.\begin{array}{l}[\theta_{jk}] = 0 \\ [\lambda_{jk}] = \hat{\lambda}_j\end{array}\right\} \text{(estimator, } Q_m{}^i(t)\text{, assumptions)} \qquad (4)$$

But these assumptions violate the previously stated premise that the uncertainty in the effective rainfall distribution over R has a major effect in hydrologic modeling accuracy. The impact in using Eq. (4) becomes apparent when calibrating the model to only storms of a single storm class, $<\xi_o>$.

Again, for all storms in $<\xi_o>$, the effective rainfall distributions are all nearly identical and are given by the single input, $e_g{}^o(t)$. But due to the variability in rainfall across the R_j, the associated runoff hydrographs, $Q_g{}^{oi}(t)$, differ even though $e_g{}^o(t)$ is the single model input.

It is recalled that in Eq. (3), the effective rainfall distribution is now the estimator, $\hat{e}_g{}^{oi}(t)$. That is, due to the several assumptions leading to Eq. (4) for the discretized model estimator, $\hat{Q}_m{}^i(t)$, the variations due to $[\lambda_{jk}]$ and $[\theta_{jk}]$ are transferred from the $[\eta(s)]$ distribution to the $\hat{e}_g{}^i(t)$ function. For storm class $<\xi_o>$, the estimator $Q_m{}^{oi}(t)$ can be written from Eqs. (2) and (3) as

$$\hat{Q}_m{}^{oi}(t) = \int_{s=0}^{t} \hat{e}_g{}^{oi}(t - s) \sum_{j=1}^{m} \sum_{<k>} \bar{a}_{<k>_j} \sum \hat{\lambda}_j \bar{\phi}_j(s - \bar{\alpha}_{<k>_j})\, ds \qquad (5)$$

where in Eq. (5), it is assumed that the variations in model output due to using mean values (overbar notation) are minor in comparison to the variations in model output due to $[\lambda_{jk}]$ and $[\theta_{jk}]$. That is, even though the rainfall distributions over the catchment, R, are variable with respect to the single input, $e_g{}^{oi}(t)$, the resulting subarea runoffs still fall within a single linear routing parameter class for each channel routing link, respectively. But then Eq. (5) is but another single area UH model:

$$\hat{Q}_m{}^{oi}(t) = \int_{s=0}^{t} \hat{e}_g{}^{oi}(t - s)\, \hat{\eta}_o(s)\, ds \qquad (6)$$

where $\hat{\eta}_o(s)$ is an estimated distribution which is "fixed" for all storms in a specified storm class $<\xi_o>$. In calibrating $\hat{Q}_m{}^{oi}(t)$, therefore, the work effort is focused towards finding the best fit effective rainfall distribution $\hat{e}_g{}^{oi}(t)$, which correlates the data pairs $\{Q_g{}^i(t), \hat{\eta}_o(s)\}$, for each storm i. That is, the "true" single $e_g{}^o(t)$ is modified to be $\hat{e}_g{}^{oi}(t)$ in order to correlate the $\{Q_g{}^{oi}(t), \hat{\eta}_o(s)\}$, for each storm i. This contrasts with finding the best fit $\eta^i(s)$ which correlates the pairs, $\{Q_g{}^{oi}(t), e_g{}^o(t)\}$. It is recalled that from Eqs. (2), (3), and (6), $\hat{\eta}_o(s)$ is a single outcome due to the assumptions of Eq. (4), and due to using a single storm class, $<\xi_o>$, which develops runoffs that fall within a single class of linear routing hydrographs.

The effective rainfall estimator, $\hat{e}_g{}^{oi}(t)$, used in Eqs. (5) and (6) is the correlation between the data pair $\{Q_g{}^{oi}(t), \hat{\eta}_o(s)\}$. Consequently, similar to the $\eta^i(s)$ outcomes, the $\hat{e}_g{}^{oi}(t)$ must have an infinite degrees of freedom in order to provide the needed correlation. However, hydrologic models prescribe a given model structure to the effective rainfall estimator which involves only a finite number of degrees of freedom, or parameters. This fixed model structure develops effective rainfalls, noted as $\tilde{e}_g{}^i(t)$, for storm event i.

Convoluting $\tilde{e}_g{}^i(t)$ with the $\hat{\eta}_o(s)$ estimated for storm class $<\xi_o>$ develops the general hydrologic model, $\tilde{Q}_g{}^i(t)$, for storm i. The model $\tilde{Q}_g{}^i(t)$ is the model that practitioners use. For storm class $<\xi_o>$, the correlation distribution is the fixed $\hat{\eta}_o(s)$, and the effective rainfall estimator is the single calibrated distribution $\tilde{e}_g{}^o(t)$. Thus, for storm class $<\xi_o>$, the "true" hydrologic model structure becomes the point estimate:

$$\tilde{Q}_m{}^o(t) = \int_{s=0}^{t} \tilde{e}_g{}^o(t - s)\ \hat{\eta}_o(s)\ ds \tag{7}$$

Because the effective rainfall submodel used in $\tilde{Q}_m{}^i(t)$ has a prescribed structure, it cannot match the best fit $\hat{e}_g{}^{oi}(t)$ for all storms and, consequently, modeling error is introduced into the parameters of the loss rate submodel, $\tilde{e}_g{}^o(t)$, when calibrated to storm class $<\xi_o>$.

An error which results due to use of Eq. (7) is that the estimator modeling distribution $[\hat{Q}_m(t)]$ for storm class $<\xi_o>$ will be imprecise due to the variation in derived loss rate parameters in $\tilde{e}_g{}^o(t)$ not achieving the true variation in $\hat{e}_g{}^{oi}(t)$ needed to correlate $\{Q_g{}^{oi}(t), \hat{\eta}_o(s)\}$ in Eq. (6).

HYDROLOGIC MODEL OUTPUT DISTRIBUTIONS

The previous development resulted in the identication of four modeling structures:

(i) $Q_m{}^i(t)$ -- this is the m-subarea link node model with channel links connecting the subareas. Stream gauge data is supplied for each subarea (or overland flowplane) and also along each channel link so that all modeling parameters and subarea effective rainfall factors are accurately determined for each storm event i. For storm class $<\xi_o>$, (measured at the single "available" rain gauge site), $Q_m{}^i(t)$ results in the distribution $[Q_m{}^o(t)]$.

(ii) $Q_1{}^i(t)$ -- this is a simple single area UH model. For only a single rain gauge and stream gauge, $Q_1{}^i(t)$ is equal to $Q_m{}^i(t)$ in predicting runoff at the stream gauge. For storm class $<\xi_o>$, $Q_1{}^i(t)$ becomes the distribution $[Q_1{}^o(t)]$ where $[Q_1{}^o(t)] = [Q_m{}^o(t)]$.

(iii) $\hat{Q}_m{}^i(t)$ -- should all the parameters in $Q_m{}^i(t)$ be estimated for a storm class, then $Q_m{}^i(t)$ is approximated by the estimator $\hat{Q}_m{}^i(t)$. However on a storm class basis, $\hat{Q}_m{}^i(t)$ reduces to another single area UH model of Eq. (6) where the realization, $\hat{\eta}_o(s)$, is fixed for storm class $<\xi_o>$. $\hat{Q}_m{}^i(t)$ equates to $Q_m{}^i(t)$ when the effective rainfall estimator, $\hat{e}_g{}^i(t)$, is given an infinite number of degrees of freedom.

(iv) $\tilde{Q}_m{}^i(t)$ -- because the effective rainfall estimates in an m-subarea link node model are of a prescribed structure, the estimates have a finite number of degrees of freedom. For storm class $<\xi_o>$, $\tilde{Q}_m{}^i(t)$ reduces to another single area UH model where the correlation distribution is identical to that used in $\hat{Q}_m{}^i(t)$. But the effective rainfall distribution in the single area UH representation is $\tilde{e}_g{}^i(t)$ where $\tilde{e}_g{}^i(t)$ is calibrated to best fit the distribution of $\hat{e}_g{}^i(t)$ distributions which are needed to correlate the data pairs, $\{Q_g{}^i(t), \hat{\eta}_o(s)$, in storm class $<\xi_o>$.

From the above modeling structures, the parameter calibration process can be interpretted. For storm class $<\xi_o>$, distributions are developed for $[Q_m{}^o(t)]$ and $[Q_1{}^o(t)]$. A distribution of $\hat{Q}_m{}^i(t)$, noted as $[\hat{Q}_m{}^o(t)]$, can be developed provided the effective rainfall estimator is given an infinite number of degrees of freedom. However, the "calibrated" model of $\tilde{Q}_m{}^i(t)$ develops only a single point estimate $\tilde{Q}_m{}^o(t)$ for storm class $<\xi_o>$.

For storm class $<\xi_o>$, the several modeling output distributions are as follows:

$$[Q_m{}^o(t)] = \int_{s=0}^{t} e_g{}^o(t-s) \sum_j \sum_{<k>_j} a^o{}_{<k>_j} \sum [\lambda_{jk}{}^o] \, \phi_j{}^o(s-[\theta_{jk}{}^o] - \alpha^o{}_{<k>_j}) \, ds \tag{8}$$

$$[Q_1{}^o(t)] = \int_{s=0}^{t} e_g{}^o(t-s) \, [\eta(s)]_o \, ds \tag{9}$$

$$[\hat{Q}_m{}^o(t)] = \int_{s=0}^{t} [\hat{e}_g{}^o(t-s] \, \hat{\eta}_o(s) \, ds \tag{10}$$

$$[\tilde{Q}_m{}^o(t)] = \tilde{Q}_m{}^o(t) = \int_{s=0}^{t} \tilde{e}_g{}^o(t-s) \, \hat{\eta}_o(s) \, ds \tag{11}$$

Again, $[Q_m{}^o(t)] = [Q_1{}^o(t)]$. $[\hat{Q}_m{}^o(t)] = [Q_m{}^o(t)]$ only when $\hat{e}_g{}^{oi}(t)$ is given an infinite number of degrees of freedom such as to correlate $Q_g{}^{oi}(t)$ to $\hat{\eta}_o(s)$ for each storm i. Finally, $\tilde{e}_g{}^o(t)$ is some weighted average of the distribution of $[\hat{e}_g{}^o(t)]$, usually, the expected value is used:

$$\tilde{e}_g{}^o(t) = E[\hat{e}_g{}^o(t)] \tag{12}$$

THE VARIANCE OF HYDROLOGIC MODEL OUTPUT

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

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

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

$$[Q_1{}^o(s)] = \int_{s=0}^{z} e_g{}^o(z-s) \, [\eta(s)]_o \, ds \tag{13}$$

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

$$[Q_p]_o = \int_{s=0}^{t_p} e_g{}^o(t_p - s) \, [\eta(s)]_o \, ds \tag{14}$$

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

$$[\max \int_{D} Q_1^{o}(t)dt] = \max \int_{D} \int_{s=0}^{t} e_g^{o}(t - s) \; [\eta(s)]_o \; ds \qquad (15)$$

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 $<\xi_o>$ is the distribution

$$[A]_o = A(\int_{s=0}^{t} e_g^{o}(t - s) \; [\eta(s)]_o \; ds) \qquad (16)$$

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

$$E[A]_o = \sum_{[\eta(s)]_o} A(\int_{s=0}^{t} e_g^{o}(t - s) \; \eta(s) \; ds \; P(\eta(s)) \qquad (17)$$

where $P(\eta(s))$ is the frequency of occurrence 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),

$$\text{var } [A]_o = \sum_{[\eta(s)]_o} \left\{ A(\int_{s=0}^{t} e_g^{o}(t-s) \; \eta(s) \; ds) - E[A]_o \right\}^2 P(\eta(s)) \qquad (18)$$

APPLICATION

Dominguez Wash is a fully developed 35 square-mile catchment located in Los Angeles, California. It has been essentially fully improved with a well-drained flood control system for nearly 50-years. Of concern is the design of a flood control detention basin at the stream gauge site.

The design objective is to build a flow-through type detention basin which provides a level of protection for a prescribed storm pattern and loss rate. The available rainfall data is a single rain gauge located off-site of the catchment.

In reviewing the rainfall data, no storms were found which precisely matched the design condition effective rainfall distribution, $e_g^D(t)$. Consequently, a storm class $<\xi_D>$ could not be developed.

The assumption that similar storm classes, $<\xi_x>$, have similar correlation distributions, $[\eta(s)]_x$, was then involved. By examining the available rainfall records and the runoff data from the Dominguez Wash stream gauge, only 5 storms were identified which were considered similar enough to $e_g^D(t)$ to have similar correlation distributions. More data would be needed to have statistical significance; however, this information is used for demonstration purposes.

The five correlation distributions, $\eta^i(s)$, are shown in mass-curve form in Fig. 1. Each $\eta^i(s)$ is assumed to have a probability of 0.20. The $\eta^i(s)$ of Fig. 1 were derived by a least-squares fit between estimated effective rainfall from the rain gauge and the stream gauge using the $Q_1^i(t)$ model structure.

For the prescribed design effective rainfall storm condition (rainfall less losses) given by $e_g^D(t)$ at the rain gauge, the hydrologic model estimate for runoff is given by the distribution $[Q_1^D(t)]$ of Eq. (1).

By routing each $Q_1^D(t)$ model, (using a different $\eta^i(s)$ for each trial), through the detention basin, a different demand on the basin volume is determined. Figures 2 and 3 show the resulting distribution of $Q_1^D(t)$ and the associated detention basin volume requirements, respectively. Also shown in Fig. 3 are confidence estimates from the modeling results.

CONCLUSIONS

The correlation of the effective rainfall to the runoff hydrograph from the catchment R will result in a different UH (for the single area model) for each storm event. However, the resulting collection of UH's reflect the dominating uncertainty in the variation in the magnitude, timing, and shape of the effective rainfall distribution over R. When the data base consists of only a single rain gauge and stream gauge these three uncertainties cannot be reduced by including additional complexitities into the hydrologic model (e.g., subareas linked by hydraulic routing submodels, additional soil-moisture accounting algorithms, etc.). *Only additional measured* rainfall-runoff data within the catchment R will reduce the uncertainty. Without this additional data, the uncertainty in the effective rainfall over R will remain and should be included in flood control design and planning studies by the development of confidence levels inthe modeling results.

REFERENCES

1. Beard, L. and Chang, S., 1974, Urbanization Impact of Stream Flow, ASCE Journal of the Hydraulics Division.

2. Garen, D. and Burges, S., 1981, Approximate Error Bounds for Simulated Hydrographs, ASCE Journal of the Hydraulics Division, Vol. 107, No. 11.

3. Hromadka II, T.V. and McCuen, R.H., 1986, Orange County Hydrology Manual, OCEMA, Santa Ana, California.

4. Loague, K. and Freeze, R., 1985, A Comparison of Rainfall-Runoff Modeling Techniques on Small Upland Catchments, Water Resources Research, Vol. 21, No. 2.

5. Nash, J. and Sutcliffe, J., 1970, River Flow Forecasting Through Conceptual Models, Part I - A Discussion of Principles, Journal of Hydrology, Vol. 10.

6. Schilling, W. and Fuchs, L., 1986, Errors in Stormwater Modeling - A Quantitative Assessment, ASCE Journal of Hydraulic Engineering, Vol. 112, No. 2.

7. Troutman, B., 1982, An Analysis of Input in Precipitation - Runoff Models Using Regression with Errors in Independent Variables, Water Resources Research, Vol. 18, No. 4.

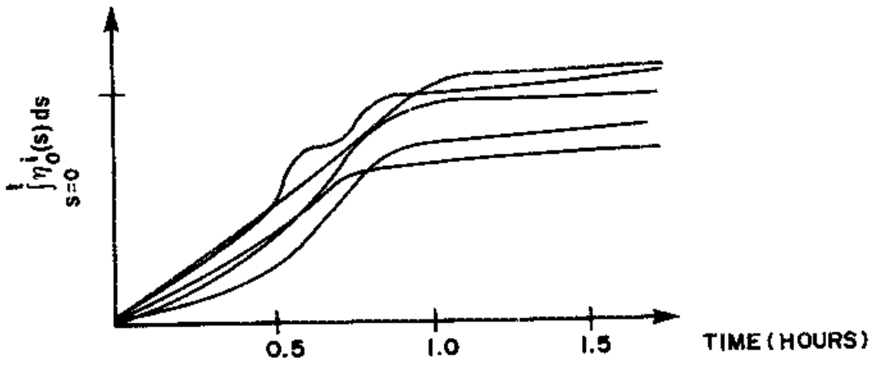

Fig. 1. Correlation Distributions $\eta_0{}^i(s)$, in Correlating $Q_g{}^i(t)$ and $e_g{}^i(t)$ for the Application Problem, Plotted in Summation (Distribution) Graph Form.

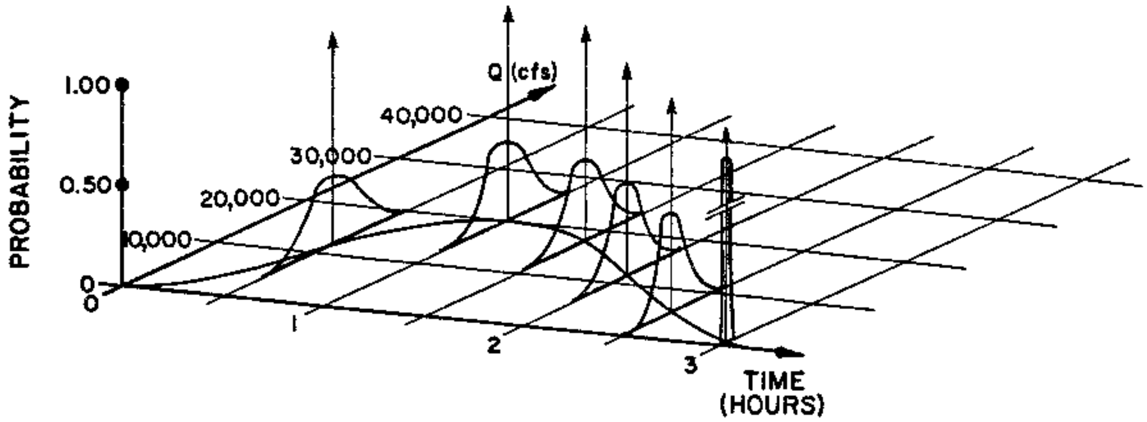

Fig. 2. The Hydrologic Model Distribution for a Predicted Response, $[Q_D(t)]$, from Input, $e_g{}^D(t)$. Heavy line is the Expected Distribution, $E[Q_D(t)]$.

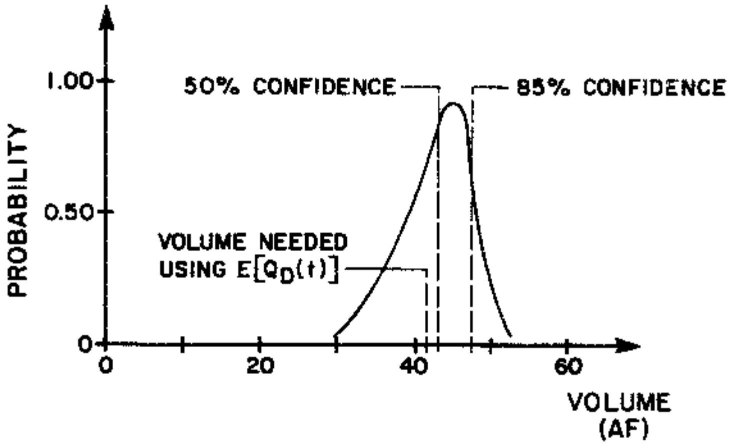

Fig. 3. Detention Basin Volume Requirements