A comparison of techniques for evaluating hydrologic model uncertainty

T. V. Hromadka II

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

C. C. Yen and M. H. Seits

Williamson and Schmid, Irvine, CA, USA

Hydrologic models are composed of several components which are all parameter dependent. In the general setting, parameter values are selected based on regionalization of observed rainfall-runoff events, or upon calibration at local stream gauge data when available. Based on these data, a selected parameter set is then used for the hydrologic model. However, seldom are hydrologic model outputs examined as to the total variations in output due to the independent but coupled variations in parameter input values. In this paper, three of the more common techniques for evaluating model output distributions are compared as applied to a selected hydrologic model; i.e., an exhaustion techniques, Monte Carlo simulation method, and the more recently advanced Rosenblueth technique. It is concluded that, for the hydrologic model considered, the Monte Carlo technique provides more accuracy in comparison to Rosenblueth technique (for the same computational effort), but is less accurate than Exhaustion.

I. INTRODUCTION

Almost all policy statements regarding the design of flood control systems involve the use of a hydrologic mathematical model to estimate design flows for subsequent channel sizing and floodplain delineation. Typically, these models are hybridizations of other 'standardized' models which have been calibrated to regions dissimilar to the region where the model is intended for use. Or if the model has been regionalized by calibration to several local catchments, a prescribed set of rules are decided upon (e.g., a 'hydrology manual') as to how to select model parameter values for use in the subsequent estimation of runoff quantities. And finally, if the model has been calibrated to catchment rainfall-runoff data when available, an 'optimized' set of parameters are concluded which minimize, in some norm sense, the errors between model-produced and stream gauge measured runoff hydrographs.

In all of the above settings, the hydrologic model of the catchment is based upon a set of parameters which produce a single model output for decision making purposes. Oftentimes these model output values are compared to other models' single-output values and decisions are concluded as to which model is 'best'.

Because all hydrologic models are gross simplifications of the several hydrologic processes, and due to the sparse data (if any) available to compute the model parameters, each of the model parameters and submodels (e.g., loss functions, routing methods, etc.) have an associated distribution-frequency which relates parameter values to the probability of exceedance (or frequency of occurrence), or relates the probability of which submodel (e.g., which loss function to use, which flow routing submodel to use, etc.) best represents the actual ongoing hydrologic/hydraulic process. The use of a single set of parameter values or choices of submodel algorithms to be used in the global model is but a single vector in an infinity of possible vectors.

Consequently, the policy decisions need to include the uncertainty of the hydrologic model in the selection of a level for flood protection. Generally, policy statements reflect the control of a peak flow rate or \( Q \) in the design of flood control facilities. The policy may incorporate the accommodation of a time distribution of runoff volume by means of a design storm runoff hydrograph (for example), but ultimately a \( Q \) is developed for channel sizing, for the delineation of a floodplain, and for the planning of urbanization. In this paper, only the uncertainty in model estimated \( Q \) (i.e., peak flow rate) values is considered in order to simplify the presentation in the comparison of the uncertainty evaluation techniques.

The policy statement, then, is reflected by the calculation of a specific \( Q_m \) based upon using a single input vector (which represents parameter estimates, submodels to use, land use assumptions, etc.) into the model. Should another input vector be used, another \( Q \) estimate would result where typically \( Q \neq Q_m \). The question then becomes: what is the distribution of possible \( Q \) values as all possible input vectors are considered according to their respective probabilities of occurrence? To address this issue, a frequency distribution of the range of model estimated \( Q \) needs to be computed. But the enormity of this task has precluded a precise evaluation for most problems. For example, in a simple four parameter unit hydrograph model based
upon a timing parameter (lag), a unit hydrograph (S-graph equivalent), and a two parameter loss function \( Fm = \text{maximum loss rate or F-index technique, } \hat{Y} = \text{constant percentage loss rate fraction} \) where the parameter values can be presumably represented by a simple distribution-frequency histogram of 9-, 5-, 6- and 7-values respectively, 1350 test runs must be made with the model to exhaust all possible input vector definitions. Should the catchment model be discretized into, say, 4 subareas linked together by a single reach two-parameter channel routing submodule where both of the routing parameter distribution-frequency histograms have 5 values, then \( 8.3 \times 10^3 \) model runs are required. It is apparent that an alternative approach is needed in order to determine a model output \( Q \) distribution-frequency relationship.

The main objective of this paper is to present the results in the development of a distribution-frequency relationship for a hydrologic model output, the peak \( Q \), using three techniques: (i) Exhaustion; that is, using every possible vector to evaluate the model output \( Q \) values; (ii) Monte Carlo simulation method\(^{80}\); that is, randomly choosing input vectors to evaluate the Q distribution-frequency; and (iii) The Rosenbluth technique\(^{82}\) which utilizes an approximation to estimate the statistics of the \( Q \), i.e., the mean and variance of \( Q \). Details of the considered hydrologic model are presented in Section II.

An examination of the 'model' response function to parameter uncertainty is contained in Section III. Specifics of the Monte Carlo simulation and Rosenbluth techniques are contained in Section IV. Section IV also presents the computational results from the three considered approaches for comparison purposes.

II. HYDROLOGIC MODEL

Model selection

Of the over 100 models available, a design storm/unit hydrograph model (i.e., 'model') is selected for this particular application. Some of the reasons are as follows: (1) the design storm approach – the multiple discrete event and continuous simulation categories of models have not been clearly established to provide better predictions of flood flow frequency estimates for evaluating the impact of urbanization and for design of flood control systems than a calibrated design storm model\(^{3,4,57,67}\); (2) the unit hydrograph method – it has not been shown that alternate approaches (i.e., the kinematic wave modelling technique) provide a significantly better representation of watershed hydrologic response than a model based on unit hydrographs (locally calibrated or regionally calibrated) that represent free-draining catchments\(^{1,2,4,6,8,23,26,30,31,49,50,52,55,63,66,69,70,73,74,77}\); (3) model usage – this class of 'model' has been used extensively nationwide and has proved generally and relatively\(^{3,10,15,19,28,37,39,41,42,43,55}\): (4) parameter calibration – the 'model' used in this application is based upon a minimal number of parameters, giving higher accuracy in calibration of model parameters to rainfall-runoff data, and the design storm to local flood flow frequency tendencies\(^{3,4,12,18,20,22,23,29,32,36,38,44,45,46,58,59,62,66,70,79}\); (5) calibration effort – the 'model' does not require large data or time requirements for calibration\(^{3,4,12,18,20,29,36,38,66,70}\); (6) application effort – the 'model' does not require excessive computation for application\(^{3,6,70}\); (7) acceptability – the 'model' uses algorithms that are accepted in engineering practice\(^{3,37,66}\); (8) model flexibility for planning – data handling and computational submodels can be coupled to the 'model' (e.g., channel and basin routing) resulting in a highly flexible modelling capability\(^{3,66}\); (9) model certainty evaluation – the certainty of modelling results can be readily evaluated as a distribution of possible outcomes over the probabilistic distribution of parameter values\(^{34,39,60,66,75,76}\).

Runoff hydrograph model parameters

The design storm unit hydrograph model ('model') is based upon several parameters; namely, two loss rate parameters (a phi-index coupled with a fixed percentage), an S-graph, catchment lag, storm pattern (shape, location of peak rainfall, duration), depth area (or depth-area-duration) adjustment, and the return frequency of rainfall.

Loss function

The loss function, \( f(t) \), used in the 'model' is defined by

\[
 f(t) = \begin{cases} \hat{Y} I(t), & \text{for } \hat{Y} I(t) < Fm \\ Fm, & \text{otherwise} \end{cases}
\]

(1)

where \( \hat{Y} \) is the low loss fraction and \( Fm \) is a maximum loss rate defined by

\[
 Fm = \sum a_p F_p
\]

(2)

where \( a_p \) is the actual pervious area fraction with a corresponding maximum loss rate of \( F_p \); the infiltration rate for impervious area is set to zero; and \( I(t) \) is the design storm rainfall intensity at storm time \( t \).

The use of a constant percentage loss rate fraction \( \hat{Y} \) in equation (1) is reported in Scully and Bender\(^{57}\), Williams et al.\(^{77}\), and Schilling and Fuchs\(^{56}\). The use of a phi index (phi-index) method in effective rainfall calculations is also well-known.

The low loss rate fraction is estimated from the SCS loss rate equation (US Dep. of Agric., 1972) by

\[
 \hat{Y} = 1 - Y
\]

(3)

where \( Y \) is the catchment yield computed by

\[
 Y = \sum a_j Y_j
\]

(4)

In equation (4), \( Y_j \) is the yield corresponding to the catchment area fraction \( a_j \) and is estimated using the SCS curve number (CN) by

\[
 Y_j = \frac{(P_{24} - Ia)^2}{(P_{24} - Ia + S)P_{24}}
\]

(5)

where \( P_{24} = \text{the 24-hour T-year precipitation depth; } Ia = \text{the initial abstraction of } Ia = 0.2S; \text{ and } S = (1000/CN) - 10.\)

From the above relationships, the low loss fraction, \( \hat{Y} \), acts as a fixed loss rate percentage, whereas \( Fm \) serves as an upper bound to the possible values of \( f(t) = \hat{Y} I(t) \).

Values for \( Fm \) are based on the actual pervious area cover percentage (\( a_p \)) and a maximum loss rate for the pervious area, \( F_p \). Values for \( F_p \) are developed from rainfall-runoff calibration studies of several significant storm events for several watersheds within the region.
under study. Further discussions regarding the estimation of parameter values are contained in a subsequent section.

A distinct advantage afforded by the loss function of equation (1) over loss functions such as Green-Ampt or Horton is that the effect of the location of the peak rainfall intensities in the design storm pattern on the model peak flow rate \( Q(t) \) becomes negligible. That is, front-loaded, middle-loaded, and rear-loaded storm patterns all result in nearly equal peak flow estimates. Consequently, the shape (but not magnitude) of the design storm pattern is essentially eliminated from the list of parameters to be calibrated in the runoff hydrograph 'model' (although the time distribution of runoff volumes are affected by the location of the peak rainfalls in the storm pattern which is a consideration in detention basin design).

S-graph

The S-graph representation of the unit hydrograph (e.g., McCuen and Bondeld \( 18 \), Chow and Kulandalswamy \( 14 \), Mays and Coles \( 29 \)) can be used to develop unit hydrographs corresponding to various watershed lag estimates. The S-graph was developed by rainfall-runoff calibration studies of several storms for several watersheds. By averaging the S-graphs for each watershed studied, a representative S-graph is developed for each watershed. By comparing the representative S-graphs, regional S-graphs were derived to represent the average of watershed-averaged S-graphs.

Lag

Fundamental to any hydrologic model is a catchment timing parameter. For the 'model', watershed lag is defined as the time from the beginning of effective rainfall to that time corresponding to 50-percent of the S-graph ultimate discharge. To estimate catchment lag, it is assumed that lag is related to the catchment time of concentration \( T_c \) as calculated by a sum of normal depth flow calculated travel times; i.e., a mixed velocity method (e.g., Beard and Chang \( 4 \), McCuen et al. \( 45 \)). To correlate lag to \( T_c \) estimates, lag values measured from watershed calibrated S-graphs were plotted against \( T_c \) estimates. A least-squares best fit line gives the estimator

\[
\text{lag} = 0.80 T_c
\]  

(6)

Design storm pattern

A 24-hour duration design storm composed of nested 5-minute unit intervals (with each principal duration nested within the next longer duration) was adopted as part of the policy. The storm pattern provides equal return frequency rainfalls for any storm duration, i.e., the peak 5-minute, 30-minute, 1-hour, 3-hour, 6-hour, 12-hour, and 24-hour duration rainfalls are all of the selected \( T \)-year return frequency. Such a storm pattern construction is found in HEC Training Document No. 15 (1982) which uses a nested central-loaded design storm pattern.

Runoff hydrograph model

The 'model' produces a time distribution of runoff \( Q(t) \) given by the standard convolution integral representation of

\[
Q(t) = \int_0^t e(s)u(t-s) \, ds
\]  

(7)

where \( Q(t) \) is the catchment flow rate at the point of concentration; \( e(s) \) is the effective rainfall intensity; and \( u(t-s) \) is the unit hydrograph developed from the particular S-graph. In equation (7), \( e(s) \) represents the time distribution of the 24-hour duration design storm pattern modified according to the loss function definition of equation (1).

In the use of equation (7) for a particular watershed, an estimate of catchment lag is used to construct a unit hydrograph \( u(x) \). Then, based on the catchment area (depth-area adjustment) and loss rate characteristics, \( e(s) \) is determined. Because the peak flow rate \( Q = \max Q(t) \) shows a negligible variation due to a change in storm pattern shape (except for a severe front loaded, near-monotonically decreasing pattern or a rear loaded, near-monotonically increasing pattern); the model parameters that affect \( Q \) are loss rates \( (\gamma_1(t) \text{ and } F_p) \), S-graphs, lag estimates, depth-area adjustment curve set, and design storm rainfall return frequency. Note that \( F_m \) is not a calibration parameter as \( F_m = a_p F_p \), where \( a_p \) is the actual measured pervious area fraction.

Parameter calibration

Considerable rainfall-runoff calibration data has been prepared by the Corps of Engineers (COE) for use in their flood control design and planning studies. The watershed information available includes rainfall-runoff calibration results for three or more significant storms for watershed, which is used to develop optimized estimates for the S-graph, lag, and loss rate at the peak rainfall intensities. Although the COE used a more rational Horton type loss function which decreases with time, only the loss rate that occurred during the peak storm rainfalls was used in the calibration effort reported herein.

![Fig. 1: Location of drainage basins used in model parameter calibration effect](image-url)

Table 1. Watershed characteristics

<table>
<thead>
<tr>
<th>Watershed name</th>
<th>Area (m²)</th>
<th>Length (m)</th>
<th>Length of centroid (m)</th>
<th>Slope (°/m)</th>
<th>Percent impervious (%)</th>
<th>Tc (hrs)</th>
<th>Storm date</th>
<th>Peak Fₚ (inch/hr)</th>
<th>Lag (hrs)</th>
<th>Basin factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alhambra Wash¹</td>
<td>13.67</td>
<td>8.62</td>
<td>4.17</td>
<td>82.4</td>
<td>45</td>
<td>0.89</td>
<td>Feb 78</td>
<td>0.59, 0.24</td>
<td>0.62</td>
<td>0.015</td>
</tr>
<tr>
<td>Compton ²³</td>
<td>24.66</td>
<td>12.69</td>
<td>6.63</td>
<td>13.8</td>
<td>55</td>
<td>2.22</td>
<td>Mar 78</td>
<td>0.35, 0.29</td>
<td>0.36</td>
<td>0.015</td>
</tr>
<tr>
<td>Verdugo Wash¹</td>
<td>26.8</td>
<td>10.98</td>
<td>5.49</td>
<td>316.9</td>
<td>20</td>
<td>–</td>
<td>Feb 80</td>
<td>0.24</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Limekiln¹</td>
<td>10.3</td>
<td>7.77</td>
<td>3.41</td>
<td>295.7</td>
<td>25</td>
<td>–</td>
<td>Feb 78</td>
<td>0.36</td>
<td>0.94</td>
<td>0.015</td>
</tr>
<tr>
<td>San Jose²</td>
<td>83.4</td>
<td>23.00</td>
<td>8.5</td>
<td>60.0</td>
<td>18</td>
<td>–</td>
<td>Feb 80</td>
<td>0.44</td>
<td>0.64</td>
<td>0.016</td>
</tr>
<tr>
<td>Sepulveda²</td>
<td>152.0</td>
<td>19.0</td>
<td>9.0</td>
<td>143.0</td>
<td>24</td>
<td>–</td>
<td>Feb 78</td>
<td>0.27</td>
<td>0.73</td>
<td>0.026</td>
</tr>
<tr>
<td>Eaton Wash¹</td>
<td>11.02</td>
<td>8.14</td>
<td>3.41</td>
<td>90.9</td>
<td>40</td>
<td>1.05</td>
<td>Feb 80</td>
<td>0.39</td>
<td>1.66</td>
<td>0.020</td>
</tr>
<tr>
<td>Rubio Wash¹</td>
<td>12.20</td>
<td>9.47</td>
<td>5.11</td>
<td>125.7</td>
<td>40</td>
<td>0.08</td>
<td>Feb 80</td>
<td>0.22, 0.21</td>
<td>1.12</td>
<td>0.017</td>
</tr>
<tr>
<td>Arcadia Wash¹</td>
<td>7.70</td>
<td>5.87</td>
<td>3.03</td>
<td>156.7</td>
<td>45</td>
<td>0.60</td>
<td>Feb 78</td>
<td>0.32</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Compton ¹³</td>
<td>15.08</td>
<td>9.47</td>
<td>3.79</td>
<td>14.3</td>
<td>55</td>
<td>1.92</td>
<td>Feb 78</td>
<td>0.39</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Dominguez³</td>
<td>37.30</td>
<td>11.36</td>
<td>4.92</td>
<td>7.9</td>
<td>60</td>
<td>2.08</td>
<td>Feb 78</td>
<td>0.22</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Santa Ana Delhi³</td>
<td>17.6</td>
<td>8.71</td>
<td>4.17</td>
<td>16.0</td>
<td>40</td>
<td>1.73</td>
<td>Feb 78</td>
<td>0.39</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Westminster³</td>
<td>6.7</td>
<td>5.65</td>
<td>1.39</td>
<td>13</td>
<td>40</td>
<td>–</td>
<td>Feb 78</td>
<td>0.22</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>El Modena-Irvine³</td>
<td>11.9</td>
<td>6.34</td>
<td>2.69</td>
<td>52</td>
<td>40</td>
<td>0.78</td>
<td>Feb 78</td>
<td>0.39</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Garden Grove-Winterberg³</td>
<td>20.8</td>
<td>11.74</td>
<td>4.73</td>
<td>10.6</td>
<td>64</td>
<td>1.98</td>
<td>Feb 78</td>
<td>0.39</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>San Diego Creek³</td>
<td>36.8</td>
<td>14.2</td>
<td>8.52</td>
<td>95.0</td>
<td>20</td>
<td>1.39</td>
<td>Feb 78</td>
<td>0.39</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Notes:
1. Watershed Geometry based on review of quadrangle maps and LACFCD storm drain maps.
2. Watershed Geometry based on COE LACDA Study.
4. Area reduced 57% due to 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.

A total of 12 watersheds were considered in detail in this study. Seven of the watersheds are located in Los Angeles County while the other five catchments are in Orange County (Fig. 1). Several other local watersheds were also considered in light of previous COE studies that resulted in additional estimates of loss rates, S-graphs, and lag values. Table 1 provides an itemization of data obtained from the COE studies, and watershed data assumed for catchments considered hydrologically similar to the COE study catchments.

Peak loss rate, Fₚ

From Table 1, several peak rainfall loss rates are tabulated which include, when appropriate, two loss rates for double-peak storms. The range of values for all Fₚ estimates lie between 0.30 and 0.65 inch/hour with the highest value occurring in Verdugo Wash which has substantial open space in foothill areas. Except for Verdugo Wash, 0.20 ≤ Fₚ ≤ 0.60 which is a variation in values of the order noted for Alhambra Wash alone. Fig. 2 shows a histogram of Fₚ values for the several watersheds. It is evident from the figure that 88 percent of Fₚ values are between 0.20 and 0.45 inch/hour, with 77 percent of the values falling between 0.20 and 0.40 inch/hour. Consequently, a regional mean value of Fₚ equal to 0.30 inch/hour is proposed; this value contains nearly 80 percent of the Fₚ values, for all watersheds, for all storms, with 0.10 inch/hour.

Fig. 2. Distribution-frequency of pervious area loss function, Fₚ
S-graph

Each of the watersheds listed in Table 1 has S-graphs developed for each of the storms where peak loss rate values were developed. For example, Fig. 3 shows the several S-graphs developed for Alhambra Wash. By averaging the several S-graph ordinates (developed from rainfall-runoff data), an average S-graph was obtained. By combining the several watershed average S-graphs (Fig. 4) into a single plot, an average of averaged S-graphs is obtained. This regionalized S-graph (Urban S-graph in Fig. 4) can be proposed as a regionalized S-graph for the several watersheds.

In order to quantify the effects of variations in the S-graph due to variations in storms and in watersheds (i.e., for ungauged watersheds not included in the calibration data set), the scaling of Fig. 5 was used where the variable \( X \) signifies the average value of an arbitrary S-graph as a linear combination of the steepest and flattest S-graphs obtained. That is, all the S-graphs (all storms, all catchments) lie between the Feb. 1978 storm Alhambra S-graph \((X = 1)\) and the San Jose S-graph \((X = 0)\). To approximate a particular S-graph of the sample set,

\[
S(X) = XS_1 + (1 - X)S_2 \quad (8)
\]

where \( S(X) \) is the S-graph as a function of \( X \), and \( S_1 \) and \( S_2 \) are the Alhambra (Feb. 1978 storm) and San Jose S-graphs, respectively. Fig. 6 shows the population distribution of \( X \) where each watershed is weighted equally in the total distribution (i.e., each watershed is represented by an equal number of \( X \) entries). Table 2 lists the \( X \) values obtained from the Fig. 5 scalings of each catchment S-graph. In the table, an 'upper' and 'lower' \( X \)-value that corresponds to the \( X \)-coordinate at 80 percent and 20 percent of ultimate discharge values, respectively, is listed. An average of the upper and lower \( X \) values is used in the population distribution of Fig. 6.

Catchment lag

In Table 2, the Urban S-graph, which represents a regionalized S-graph for urbanized watersheds in valley type topography, has an associated \( X \) value of 0.85. It is noted, however, that when the Urban S-graph is compared to the standard SCS S-graph, a striking similarity is seen (Fig. 7). Because the new Urban S-graph is a near duplicate of the SCS S-graph, it was assumed that catchment lag (COE definition) is related to the catchment time of concentration, \( T_c \), as is generally assumed in the SCS approach.

Catchment \( T_c \) values are estimated by subdividing the watershed into subareas with the initial subarea less than 10 acres and a flowlength of less than 1,000 feet. Using a Kirpich formula, an initial subarea \( T_c \) is estimated, and a \( Q \) is calculated. By subsequent routing downstream of the peak flowrate \((Q)\) through the various conveyances (using normal depth flow velocities) and adding successive estimated subarea contributions, a catchment \( T_c \) is estimated as the sum of travel times analogous to a mixed velocity method.

In this study, lag values are developed directly from available COE calibration data, or by using 'basin factors' calibrated from neighboring catchments (see
### Table 2. Catchment S-graph X-values

<table>
<thead>
<tr>
<th>Watershed name</th>
<th>Storm</th>
<th>( X ) (Upper)</th>
<th>( X ) (Lower)</th>
<th>( X ) (Avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alhambra</td>
<td>Feb. 78</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Feb. 80</td>
<td>0.95</td>
<td>0.60</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Mar. 78</td>
<td>0.70</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td>Limekiln</td>
<td>Feb. 78</td>
<td>0.50</td>
<td>0.80</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Feb. 80</td>
<td>0.80</td>
<td>1.00</td>
<td>0.90 (2)</td>
</tr>
<tr>
<td>Supulveda</td>
<td>Avg.</td>
<td>0.90</td>
<td>0.80</td>
<td>0.85 (3)</td>
</tr>
<tr>
<td>Compton</td>
<td>Avg.</td>
<td>0.90</td>
<td>1.00</td>
<td>0.95 (3)</td>
</tr>
<tr>
<td>Westminster</td>
<td>Avg.</td>
<td>0.90</td>
<td>0.60</td>
<td>0.60 (3)</td>
</tr>
<tr>
<td>Santa Ana Delhi</td>
<td>Avg.</td>
<td>0.80</td>
<td>1.00</td>
<td>0.90 (3)</td>
</tr>
<tr>
<td>Urban</td>
<td>Avg.</td>
<td>0.90</td>
<td>0.80</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Note: In Table 2, the numbers in parenthesis indicate the weighting of the average \( X \) value. That is, due to only the average S-graph (previously derived by the COE) being available, it is weighted to be equally represented in the sample set with respect to the other catchments.

Fig. 7. Comparison of the standard SCS and the regionalized S-graphs

Fig. 1. The COE standard lag formula is:

\[
\text{lag (hours)} = 24\bar{n}\left(\frac{L_{\text{ref}}}{S^{0.5}}\right)^{0.38}
\]  

(9)

where \( L \) is the watershed length in miles; \( L_{\text{ref}} \) is the length to the centroid along the watercourse in miles; \( S \) is the slope in ft/mile; and \( \bar{n} \) is the basin factor.

Because Eaton Wash, Rubio Wash, Arcadia Wash and Alhambra Wash are all contiguous (see Fig. 1), have similar shape, slopes, development patterns, and drainage systems, the basin factor of \( \bar{n} = 0.015 \) developed for Alhambra Wash was also used for the other three neighboring watersheds. Then the lag was estimated using equation (9).

Compton Creek has two stream gauges, and the \( \bar{n} = 0.015 \) developed for Compton 2 was also used for the Compton 1 gauge. The Dominguez catchment, which is contiguous to Compton Creek, is also assumed to have a lag calculated from equation (9) using \( \bar{n} = 0.015 \).

The Santa Ana-Delhi and Westminster catchment systems of Orange County have lag values developed from prior COE calibration studies. Fig. 8 provides a summary of the local lag versus \( T_C \) data. A least-squares best fit results in

\[
\text{lag} = 0.72T_C
\]

(10)

McCuen et al.\(^{42}\) provide additional measured lag values and mixed velocity \( T_C \) estimates which, when lag is modified according to the COE definition, can be plotted with the local data such as shown in Fig. 9. A least-squares best fit results in:

\[
\text{lag} = 0.80T_C
\]

(11)

In comparison, McCuen\(^{41}\) gives standard SCS relationships between lag, \( T_C \), and time-to-peak which, when modified to the COE lag definition, results in:

\[
\text{lag} = 0.77T_C
\]

(12)

Adopting a lag of 0.80\( T_C \) as the estimator, the distribution of (lag/0.8\( T_C \)) values with respect to equation (11) is shown in Fig. 10.
of assuming lag equal to 0.8Tc, Fp equal to 0.30 inch/hour, and X equal to 0.85 (Urban S-graph). For different model parameters, the Q/Qm plots were all very close to Fig. 11 as a function of Tc. Therefore, Fig. 11 is taken to represent the overall Q/Qm distribution, which also reflects the inherent uncertainty in design Q values which is typically used in a flood control policy statement.

An important question arises as to whether or not the certainty of outcomes from the calibrated model can be reduced (i.e., the model made more certain) by introducing additional parameters. It is not clear in the current literature whether such a claim has validity. However, some pointed remarks can be taken from Klemes and Bulu14 who evaluate the 'limited confidence in confidence limits derived by operational stochastic hydrologic models'. They note that advocates of modelling 'sidestep the real problem of modelling – the problem of how well a model is likely to reflect the future events – and divert the user to a more tractable, though less useful, problem of how best to construct a model that will reproduce the past events'. In this fashion, 'by the time the prospective modeller has dug himself out of the heaps of technicalities, he either will have forgotten what the true purpose of modelling is or will have invested so much effort into the modelling game that he would prefer to avoid questions about its relevance'. Of special interest is their conclusion that 'Confidence bands derived by more sophisticated models are likely to be wider than those derived by simple models'. That is, 'the quality of the model increases with its simplicity'.

In a reply by Nash and Sutcliffe45 to comments by Fleming46, the simple model structure used by Nash and Sutcliffe45 is defined as to modelling completeness in comparison to the Stanford Watershed Model variant, HSP. Nash and Sutcliffe write that '...We believe that a simple model structure is not only desirable in itself but is essential if the parameter values of the component parts are to be determined reliably through an optimization procedure ... One must remember that the data always constitute a limited sample and the optimized values are 'statistics' derived from this sample and therefore subject to sampling variance. The more complex the model structure the greater is the difficulty in obtaining optimum parameter values with low sampling variance. This difficulty becomes particularly acute ... when two or more model components are similar in their operation ...'.

Should a comparison be made between a simple model, such as described herein, and an 'advanced' model, the results would not be conclusive. As Nash and Sutcliffe write, '...The more complex model would almost certainly provide a better fit, as a linear regression analysis on a large number of variables will almost invariably provide a better fit than one of those independent variables whose significance has been established'. Hence, the hydrologist must be careful to evaluate modelling results obtained from a verification test rather than obtained from a calibration data set. Nash and Sutcliffe also note the dominating importance of errors in rainfall and effective rainfall estimates in complex models such as HSP: 'We wonder, however, how the parameters expressing spatial variation of rainfall or infiltration capacity, can be optimized at all, let alone with stability or significance, in the typical case where the short-term rainfall data are based on a single recording rain gauge ...'.

III. PARAMETER UNCERTAINTY AND MODEL RESPONSE

Each of the 'model' parameters (lag, Fp, and S-graph) are assumed to have the probability distribution functions (pdf) shown in a discrete histogram form in Figs 2, 6, and 10 for Fp, S(X), and lag = 0.8Tc, respectively. For example, if the 'model' is applied at a gauged site, say Alhambra Wash, then the variability in the S-graph is not given by Fig. 6 for S(X), where 0.60 ≤ X ≤ 1, but for 0.75 ≤ X = 1 (see Table 2). Similarly, the estimate for lag is much more certain for Alhambra Wash than shown in Fig. 10. Consequently, the uncertainty in the 'model' output for a gauged site will typically show a significantly smaller range in possible outcomes than if the total range of parameter values of Figs 6 and 10 are assumed (as is done for the ungauged sites, or sites where an inadequate length of data exist for a constant level of watershed development).

To evaluate the 'model' uncertainty, a simulation that exhausts all combinations of parameter values shown in the several pdf distributions was prepared (i.e., the exhaustion approach). Because the lag/Tc plot is a function of Tc, several Tc values were assumed and lag values varied freely according to Figs 9 and 10. The resulting Q/Qm distribution is shown in Fig. 11 for the case of Tc equal to 1 hour and a watershed area of 1 square mile (hence, depth-area adjustments are not involved). In the figure, Q is a possible 'model' peak flow rate outcome, and Qm is the peak flow rate obtained from the 'model' policy

---

**Fig. 9.** Relationship between measured catchment lag and computed Tc

**Fig. 10.** Distribution-frequency of lag/(0.8Tc)

---

and the standard deviation is

\[ s = \left( \sum_{i=1}^{3} \sum_{j=1}^{2} \sum_{k=1}^{3} P_{1i}P_{2j}P_{3k}[F(X_{1i},X_{2j},X_{3k}) - \bar{Q}]^2 \right)^{1/2} \]  

(14)

where \( P_{1i} \) is the probability weighting assigned to \( X_1 \) at outcome \( i \). Equations (13) and (14) can be extended directly for any finite number of parameters and number of parameters-histogram values.

**Monte Carlo simulation method**

If probability density functions for each of the parameters can be obtained or estimated, the Monte Carlo simulation method can be used. Because the Monte Carlo simulation method involves randomly selected input vectors, the computed model output statistics (e.g., mean, variance) are themselves random variables and the estimates are also subject to statistical fluctuations. Thus any estimate will be a random variable and will have an associated error band. The larger the number of trials in the simulation, it is hoped the more precise will be the estimates for the statistics.

**Moments generation method**

Generally, the density functions are not available for most of the parameters in a hydrology model. However, oftentimes the mean and standard deviation of each parameter can be estimated from the limited information.

---

Exhaustion model

Simple discrete probability density functions can be derived using histograms for each random variable. For demonstration purposes, consider a hydrology model output \( Q \) based upon three parameters; i.e., \( Q = F(X_1, X_2, X_3) \), where the \( X_j \) are represented by three different discrete probability density functions (Fig. 12).

The mean of the subject model output, \( \bar{Q} \), is

\[ \bar{Q} = \sum_{i=1}^{3} \sum_{j=1}^{2} \sum_{k=1}^{3} P_{1i}P_{2j}P_{3k}F(X_{1i},X_{2j},X_{3k}) \]  

(13)
By using the previous example, the mean of the model output can be estimated from the moments generation method as

\[
E[Q] = E[F(\bar{X}_1, \bar{X}_2, \bar{X}_3)] + 1/2 \sum_{i=1}^{3} \frac{\partial^2 F(\bar{X}_1, \bar{X}_2, \bar{X}_3)}{\partial X_i^2} \text{Var}(X_i)
\]  

(15)

and the variance of the system performance can be estimated as

\[
\text{Var}[Q] = \sum_{i=1}^{3} \left( \frac{\partial F(\bar{X}_1, \bar{X}_2, \bar{X}_3)}{\partial X_i} \right)^2 \text{Var}(X_i)
\]

(16)
in which \(X_1, X_2, X_3\) are the estimated means and \(\text{Var}(X_i)\) is the estimated variance for each parameter, respectively.

Rosenblueth method

If the first and second partial derivatives in the above equations are not available, then the two-point estimate method\(^2\) can be considered. The estimated mean \(\bar{Q}\) and standard deviation(s) for the model output \(Q\) from the two-point estimate method are

\[
\begin{align*}
\bar{Q} &= 1/8F(\bar{X}_1 + s_1, \bar{X}_2 + s_2, \bar{X}_3 + s_3) \\
&+ 1/8F(\bar{X}_1 + s_1, \bar{X}_2 - s_2, \bar{X}_3 + s_3) \\
&+ 1/8F(\bar{X}_1 - s_1, \bar{X}_2 + s_2, \bar{X}_3 + s_3) \\
&+ 1/8F(\bar{X}_1 - s_1, \bar{X}_2 - s_2, \bar{X}_3 + s_3) \\
&+ 1/8F(\bar{X}_1 + s_1, \bar{X}_2 + s_2, \bar{X}_3 - s_3) \\
&+ 1/8F(\bar{X}_1 + s_1, \bar{X}_2 - s_2, \bar{X}_3 - s_3) \\
&+ 1/8F(\bar{X}_1 - s_1, \bar{X}_2 + s_2, \bar{X}_3 - s_3) \\
&+ 1/8F(\bar{X}_1 - s_1, \bar{X}_2 - s_2, \bar{X}_3 - s_3)
\end{align*}
\]

(17)

![Map of March 1, 1983 recorded storm rainfall distribution over Los Angeles, California](attachment:image_url)

**Fig. 13.** March 1, 1983 recorded storm rainfall distribution

<table>
<thead>
<tr>
<th>Statistical model</th>
<th>1st and 2nd Partial derivative</th>
<th>Probability density function</th>
<th>Iteration number</th>
<th>Mean and standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhaustion model</td>
<td>No</td>
<td>No</td>
<td>1! / (2π)^1/2</td>
<td>No</td>
</tr>
<tr>
<td>Moment generation model</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Rosenblueth method</td>
<td>No</td>
<td>No</td>
<td>2^n (2)</td>
<td>Yes</td>
</tr>
<tr>
<td>Monte Carlo simulation method</td>
<td>No</td>
<td>Yes</td>
<td>500 - 5000 (3)</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes: (1). \(N(p)^1\) denotes the discrete points on the discrete distribution function for parameter \(p_i\).
(2). \(n\) denotes the number of parameters.
(3). Number of iterations based on this study's results.
and
\[
\sigma = \left\{ \frac{1}{8} \left[ F(\hat{X}_1 + s_1, \hat{X}_2 + s_2, \hat{X}_3 + s_3) - Q \right]^2 + \frac{1}{8} \left[ F(\hat{X}_1 - s_1, \hat{X}_2 - s_2, \hat{X}_3 - s_3) - Q \right]^2 \right\}^{1/2}
\]

in which \(s_1, s_2, s_3\) are the standard deviations of \(X_1, X_2, X_3\), respectively.

**Discussion**

The exhaustion model is suitable for models involving few parameters with sparse discrete pdf's. When the discrete pdf's become dense, or the number of parameters becomes large, resulting in an infeasible number of model outcome runs for the exhaustion technique, then the Monte Carlo simulation or Rosenbluth techniques become more attractive. In most engineering problems, only the means and variances of the parameters need to be estimated; in which case, the moments generation method is preferred when the first and second partial derivatives can be evaluated or approximated. If the model output/parameters relationship is linear, then either the moments generation method or the two-point estimate method is suitable. For a non-linear system with parameters that have small coefficients of variation, the results from the moment generation method and the two-point estimate method are usually satisfactory. The exhaustion model and the Monte Carlo simulation method are suitable for non-linear systems and/or for parameters having large coefficients of variation.

Since different criteria and information are needed for each statistical model (Table 3), the selection of a statistical model should be conducted with care, and attention must be given to development of a realistic description of the underlying physical situation to serve as input to a statistical model.

**Comparison of model uncertainty evaluation techniques: peak flow estimates**

The estimation of the peak flow rates, \(Q\), is a basic problem in hydrology. An important source of uncertainty in this estimate is that caused by the uncertain estimation of model parameters. This uncertainty can have a significant effect on the flood design value, and its quantification is an important aspect of evaluating the risk involved in a chosen level of flood protection.

For simplicity, the hydrologic model of Section II is assumed to be a function of only the three representative parameters of soil loss (phi index), watershed time of concentration and the unit hydrograph (S-graph form) variable, (X). Figs 2, 6, and 10 show the assumed histograms for these three random variables. In this application, the model output considered is the runoff hydrograph peak flow rate, or \(Q\). Using the several parameter values from each histogram, an exhaustion study was performed as described in this section and the result is depicted in Fig. 11. For the second analysis using Monte Carlo technique, a uniform \([0, 1]\) distribution is used to represent each histogram shown in Figs 2, 6, and 10. The Monte Carlo simulation is used to develop random model outcomes by use of random input vectors containing randomly selected parameter values. Finally, the mean and standard deviation values for each histogram were used in the Rosenbluth method as the third model. Table 4 shows the results developed from a runoff hydrograph model of 1 square mile. Because only 3 parameters are used in this model setting, the Rosenbluth technique requires only 2 or 8 trials of the model, potentially affording a considerable cost savings over an exhaustion analysis.

As shown in Table 4, the Rosenbluth model gives the highest estimated mean value of the \(Q\) and also the largest standard deviation of the estimated \(Q\). This probably reflects the high coefficients of variation of watershed time of concentration (and similarly, lag) and soil loss function (phi index) in the above study.

For further comparison of the Monte Carlo and Rosenbluth models, the pdf's of watershed lag time and soil loss function were modified to have small coefficients of variation. Table 4 also includes the results of the three statistical approaches for this second study. In this case, the Rosenbluth model shows acceptable results compared to the exhaustion model and the Monte Carlo method.

It is noteworthy that in the actual field case study where the coefficients of variation of two model parameters are large, the Rosenbluth technique performs poorly, even when compared to a Monte Carlo analysis with a small sample set. Of course, the Monte Carlo technique estimates are random variables themselves, but in this study it was found that for over 90% of the time, the Monte Carlo technique resulted in better estimates of the mean and standard deviation than the Rosenbluth method for the same number of trials (i.e., sample size of 8). And when the model parameters are modified to have a small coefficient of variation, the Rosenbluth technique produced acceptable estimates but so did the Monte Carlo method for the same effort.

Because it is not clear beforehand whether a model uncertainty analysis based upon the Rosenbluth technique will result in adequate estimates of the mean and standard deviation, and because the Monte Carlo performed as good as or better than the Rosenbluth technique (as applied to this hydrologic model), the Monte Carlo technique may be preferable for use in other studies as well where information about the parameters' pdf's are known.

**Comparison of model uncertainty evaluation techniques: rainfall-runoff analysis**

Rainfall-runoff analysis may be used for generating
estimates of the peak flow rates, $Q$, needed for the design of flood control channels, and estimates for a time distribution of runoff volume for the design of detention basins.

The hydrologic model (described in Section II) was applied to a severe storm condition which occurred on March 1, 1983 in southern California. This storm resulted in various rainfall intensities ranging between 10-year and 200-year return frequencies, causing severe flooding damage throughout the region. The variation in rainfall is reflected by Fig. 13 which shows the rain gauge measured time-distributions in the Los Angeles area. As can be seen, the variation in rainfall is significant even though the storm was of a rare return frequency. Stream gauge data was also available at several catchments. Consequently, it is feasible to prepare a rainfall-runoff analysis using the previously described hydrologic model.

It is noted that this storm was not included in the data set used to determine the parameter calibration of the hydrologic model. It is also noted that in this study effort, the focus is upon the modelling of a particular storm event rather than the development of a probabilistic design storm runoff hydrograph for flood control purposes. Hence, the nested design pattern (see Section II) used for flood control design purposes is replaced, in this application, by the actual measured storm pattern recorded at the available rainfall gauges.

The subject model is based upon 5-minute unit intervals for both rainfall and runoff. Assuming each 5-minute unit interval of runoff to be a random variable, an uncertainty analysis can be prepared for the entire runoff hydrograph as a collection of 5-minute unit interval random variables. Such an analysis was prepared for two catchments using the parameter value histograms of Figs 2, 6, and 10, for both the Rosenblueth and Monte Carlo techniques. For comparison purposes, additional studies were also prepared using the Monte Carlo approach, but with successively larger sample sizes. A comparison of results in the estimates of Alhambra Wash and Compton (2) Creek of the several peak flow rate statistics are contained in Table 5. Figs 14 and 15 show the modelling outcomes developed from the two uncertainty analysis techniques considered.

As with the previous application, the Monte Carlo technique provides in general, a ‘better’ estimate for the mean and standard deviation than provided by the Rosenblueth technique. Even with the same computational effort (8 trials), the Monte Carlo technique outperformed (for this study) the Rosenblueth technique.

V. CONCLUSION

Advances in hydrologic modelling techniques typically involve the incorporation of higher complexity into the hydrology model by use of hydraulic submodels. With

Fig. 14a. Rosenblueth analysis of the March 1, 1983 storm for the Compton Creek watershed

Fig. 14b. Monte Carlo analysis of the March 1, 1983 storm for the Compton Creek watershed

<table>
<thead>
<tr>
<th>Watershed runoff hydrograph</th>
<th>Monte Carlo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$n=8$</td>
</tr>
<tr>
<td>5-minute interval no.</td>
<td>$\bar{Q}$</td>
</tr>
<tr>
<td>Alhambra Wash</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>2511</td>
</tr>
<tr>
<td>109</td>
<td>6486</td>
</tr>
<tr>
<td>152</td>
<td>2618</td>
</tr>
<tr>
<td>234</td>
<td>5602</td>
</tr>
<tr>
<td>271</td>
<td>1585</td>
</tr>
<tr>
<td>Compton (2) Creek</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>1828</td>
</tr>
<tr>
<td>105</td>
<td>3984</td>
</tr>
<tr>
<td>151</td>
<td>2570</td>
</tr>
<tr>
<td>233</td>
<td>946</td>
</tr>
<tr>
<td>291</td>
<td>762</td>
</tr>
</tbody>
</table>

Notes: $n=$ sample size. $\bar{Q}=$ mean value. $s=$ standard deviation.
over 150 models reported in the open literature, it is still not clear whether the general level of success afforded by the many types of complex models provide a marked improvement over that achieved by the more commonly used and simpler models such as the unit hydrograph method. Such a review indicates that it is still not clear, in general, whether as modelling complexity increases, modelling accuracy increases.

It is important to consider the uncertainty in a hydrologic model, because the estimates of the peak flow rates, Q, and estimates of runoff volume will have significant effect on the flood design values, e.g., the design of flood control channels, and the design of detention basins.

In order to prepare such an analysis, a technique to develop model outcome statistics is needed. Due to computational effort limitations, an exhaustion study which considers the total universe of parameter inputs is usually precluded. Two alternatives to an exhaustion study is a Monte Carlo emulation and the more recently advanced Rosenbluth technique. Although the Rosenbluth technique potentially affords a significant savings in computational effort over that usually needed with a Monte Carlo analysis, it was found that in this application the Monte Carlo technique was superior in accuracy, for even the same computational effort.