A RAINFALL-RUNOFF SIMULATION MODEL FOR ESTIMATION OF FLOOD PEAKS FOR SMALL DRAINAGE BASINS--APROGRESS REPORT

BY

David R. Dawdy, Robert W. Lichty, and James M. Bergmann

U.S. Geological Survey
Open File Report

1970

A RAINFALL-RUNOFF SIMULATION MODEL FOR ESTIMATION OF FLOOD PEAKS FOR SMALL DRAINAGE BASINS--A PROGRESS REPORT

BY

David R. Dawdy, Robert W. Lichty, and James M. Bergmann

U.S.Geological Survey
Open File Report
1970

Contents

A Rainfall-Runoff Simulation Model for Estimation of Flood Peaks for Small Drainage Basins--A Progress Report

By David R. Dawdy, Robert W. Lichty, and James M. Bergmann

A parametric rainfall-runoff simulation model is used with point rainfall and daily potential evapotranspiration data to predict flood volume and peak rates of runoff for small drainage areas. The model is based on bulk-parameter approximations to the physical laws governing infiltration, soil moisture accretion and depletion, and surface streamflow. An objective fitting method is used for determining optimal best-fit sets of parameter values for the data available for use in predicting flood peaks for three case studies. Errors of prediction result from both errors of rainfall input and lack of model equivalence to the physical prototype. These two sources of error seem to be of the same order of magnitude for a model of the level of simplicity of that presented. Major gains in accuracy of simulation will require improvements in both data and model. The limit of accuracy of prediction of flood peaks by simulation with a bulk-parameter model using a single rain gage seems to be on the order of 25 percent.

INTRODUCTION

The development of the digital computer has added a new dimension to hydrology. Methods developed in a day when each step took hours with pen or pencil now take seconds with the computer. In addition, much more complex methods of analysis are feasible because of the speed of solution by the computer. The impact of the computer has been particularly great in the area of rainfall-runoff modeling. Surface water hydrology historically has been concerned with modeling, for flood routing and unit hydrograph analysis are mathematical modeling. Complete rainfall-runoff simulation models date back at least to the 1920's. However, the present burst of activity in hydrologic simulation is a direct result of widespread availability of the computer.

Computers have made rainfall-runoff simulation on a large scale economically feasible. Practicality, however, depends upon applicability and accuracy of the results of simulation. Simulation may be practical if one of the following applications is realized.

1. A rainfall record can be used to add to the information content of a streamflow record which has a shorter period of record than the rainfall record.

2. Model parameters for ungaged sites can be estimated on the basis of those derived for gaged sites, and information can be gained at the ungaged sites through the use of recorded or simulated rainfall and estimated parameters at the ungaged sites.

3. The effect of man-made changes on a basin can be related to changes in model parameters so that measured "before" conditions can be compared with simulated "after" conditions of sufficient accuracy for planning purposes.

In order for any one of these applications to be realized, some knowledge must be gained of the accuracy of prediction through the use of rainfall-runoff simulation models. Measures of accuracy must be presented to the user in understandable terms. Accuracy should be measured in terms of prediction rather than in terms of fitting. Accuracy of fitting indicates only how well the model can be used to reproduce a set of data by adjusting the model parameters. Accuracy of prediction indicates how well the model can reproduce a set of data which was not used to derive the parameter values. Therefore, prediction involves an independent test of accuracy of the model.

The Geological Survey Water Resources Division research program is developing rainfall-runoff simulation models. Emphasis has been on the utility of the models for practical field application to current programs of the Division, and has centered upon both development of models and testing of their accuracy of prediction. This report is a statement of progress to date on model development, with examples given of the results of prediction for several basins in differing hydrologic settings.

Hydrologic models have been developed in response to hydrologic needs. The introduction of computers has led to the development of more sophisticated models. The more sophisticated models should be more accurate in order to justify their existence. Their accuracy must be measured in terms of their ultimate use.

The models discussed in this report are parametric models, or models which try to simulate physical conditions by a deterministic mathematical description which includes, as much as possible, approximations to the physical laws which govern surface water hydrology. Wherever possible a physical interpretation is placed upon the parameters used in the models. A separate field of modeling not covered in this study is that of stochastic simulation, which includes models which describe the hydrologic record in statistical terms and use that statistical description to generate synthetic "equally likely" records. Each type of model has its advantages and disadvantages for application to meet a particular need.

The derivation of a set of optimum parameters which represent the hydrology of a basin must be based on data. A parametric model requires both streamflow and rainfall data and, perhaps, other hydrologic data. The data other than streamflow do contain streamflow information, and the added information should reduce the time length of streamflow data collection necessary to achieve a given level of accuracy of prediction.

Most studies concerning rainfall-runoff models assume a stationary time series, at least during a period of calibration. Thus the model parameter values can not change over time. An assumption often is made that if parameters do change that any such changes can be related to physical changes on the drainage basin, particularly to man-made changes.

Historical Development of Parametric Rainfall-Runoff Models

Parametric hydrology is that field of mathematical hydrology which attempts to synthesize a model of the land phase of the hydrologic cycle by approximating the physical laws governing the various components of the rainfall-runoff system. Infiltration, soil moisture storage, percolation to ground water, evapotranspiration and surface and subsurface flow routing are modeled by sets of equations which hopefully give a response equivalent to the response of the component modeled. The components and all necessary interrelations among components are described by means of parameters, some of which are empirical and some of which have a physical interpretation.

One of the earliest overall models of the hydrologic cycle was developed by Folse (1929). Development on that model began about 1916, and continued throughout the 1920's. During the 1930's, advances were made in the description of all components of the hydrologic cycle. Sherman (1932) introduced the theory of the unit hydrograph, which led to a flurry of developments culminating with Dooge's general linear theory of flood flow routing (1959). The Horton (1939) infiltration equation was an empirical attempt to describe unsaturated flow. Philip (1954) extended this by deriving an approximation based upon the Darcy equation for infiltration at a point. Theis (1935) showed the analogy of the Darcy equation for flow through saturated porous media to the heat flow equation. Many simplifications for specified boundary conditions followed, and became the basis for routing of ground water discharges, such as the equation of Kraijenhoff (1958) for instantaneous recharge in the two-dimensional case.

The digital computer made it possible to combine these many approximations into one overall approximation describing the operation of the land phase of the hydrologic cycle. Linsley was the first to take advantage of this possibility, and his efforts led to the development of the Stanford Watershed Model (Crawford and Linsley, 1966). Similar models have been developed at many universities and in government agencies, both in the USA and abroad.

The many models currently available or being developed must meet certain criteria in order to be useful in practical application. They must:

1. Require only input data which are generally available.

2. Be simple enough for the user to operate and to understand.

3. Provide the output desired at an acceptable level of accuracy for the application for which it is used.

The U. S. G. S. flood hydrograph simulation model follows directly from the historical developemtns described above, and is designed to meet the criteria outlined.

Transferability of Results of Modeling

Transferability of results requires that the parameters derived from simulation studies at measured sites be constant, or possess invariant relations with physical variables which can be measured in other basins. Time invariance is required or else any changes in time must be the result of measureable physical changes within the basin. Certain types of information may be transferable without the use of simulation. For instance, Benson (1962), as shown that raw data analysis leads to regionalization of flood frequency characteristics for a region. Simulation might aid such a study by extending the data base available for analysis. In addition, simulation is necessary if the time sequence of flows rather than just their frequency of occurrence is needed.

Parametric simulation is so structured as to contain parameters which are related to physical measures of the basin. Therefore, transferability is implicit in parametric simulation. No studies to date, however, have presented results leading to regionalization of the parameters of either stochastic or parametric simulation, although some thought has been given to the problem (Benson and Matalas, 1967; Matalas and Gilroy, 1968). The test of feasibility of any such regionalization of parameters depends upon the sensitivity of results of simulation to the accuracy of the parameter estimates. The fact that a parameter, whether statistical or physical, can be related to some characteristic of a basin is of no use if the standard error of estimate of the resulting relation is such that the simulation may be grossly in error. Therefore, transferability of parameters will be limited by the sensitivity of the modeling results to errors in parameter values.

Advantages and Disadvantages of Parametric Simulation

Rainfall-runoff models, in general, are lumped parameter models, although often the surface streamflow routing is accomplished by the use of a finite difference approximation to the drainage system. A lumped parameter model attempts to use a single parameter value to represent a physical measure which has spatial variability. The models are therefore at least one step removed from the reality of flow mechanics at each point in the watershed. Derived parameter values are at best average values for the basin, and are an index to, rather than a measure of, the underlying physical system. This approximation introduces a major source of error into a lumped parameter model, and limits the accuracy of prediction obtained by the use of the model.

The parameters in parametric simulation models should require a shorter period of record in order to be as well defined as those for either deterministic or stochastic black-box models. This has advantages when data collection must be commenced, and analysis postponed until sufficient data are available. Transferability should be easier for parametric models, although this is yet to be demonstrated. Parametric models require more types of data for each event modeled, both for system identification (fitting of parameters) and for simulation of synthetic records.

The emphasis of the models presented in this study is on flood hydrograph simulation for small drainage areas. Generally there is little or no data on small drainage areas. Therefore results must be obtained on the basis of short records. In addition, only a small precentage of smaller basins can be gaged. Therefore results must have transferability if the ungaged smaller basins are to be simulated. Concentration on the development of a parametric model thus seemed warranted.

Acknowledgements

This report represents several years of research by the authors. The project was initiated as the result of the efforts of R. W. Carter, Chief, Surface Water Branch, U. S. Geological Survey, and his helpful encouragement throughout was important. Terence O'Donnell, Imperial College, London, was a part of the research group during a years sabbatical with the Geological Survey and his continued interest has contributed to the development of ideas, particularly on the use of objective fitting methods. Jaime Amorocho, University of California at Davis, Chester Kisiel, University of Arizona, and Jacob Rubin, U. S. Geological Survey, Menlo Park, California, have all been helpful in trying to encourage the authors to say what they mean and mean what they say.

STRUCTURE OF THE MODEL

General Structure

The rainfall-runoff model described in this report deals with three components of the hydrologic cycle. They are antecedent moisture, infiltration, and surface runoff. The structure of the model is shown in figure 1.

Figure 1. -- near here

The attempt has been made to design a model which has a degree of equivalence to the physical system. Therefore, this model should be quite similar in structure to any other bulk parameter model for rainfall-runoff simulation. The antecedent moisture accounting component is a more sophisticated version of the antecedent precipitation index (API), which is designed to determine the starting infiltration rate for a storm. The infiltration component uses the Philip equation, which is supposedly a somewhat better approximation to the differential equation which describes unsaturated flow than is the classical Horton exponential decay infiltration equation. Surface routing is based on a linear approximation developed over 20 years ago (Clark, 1945).

The operation of the antecedent moisture accounting is designed to simulate moisture redistribution in the soil column and evapotranspiration from the soil. It contains four parameters: EVC, a pan coefficient converting measured pan evaporation to potential evapotranspiration, RR, a coefficient which determines the relative amounts of infiltration and surface runoff for periods with daily rainfall input, BMSM, a maximum effective amount of base moisture storage, and DRN, a coefficient controlling the rate of drainage of the infiltrated soil moisture. The input to this component is daily rainfall and daily pan evaporation. The output is the amount of base moisture storage (BMS) and of infiltrated surface moisture storage (SMS). BMS represents a uniform antecedent moisture content of the active soil column, and its range of values should simulate the range from wilting point to field capacity. SMS represents the moisture content of the surface layer which develops during infiltration.

The infiltration component is based on an approximation to the differential equation which describes unsaturated flow (Philip, 1954). The equation is based on a two-part accounting of the soil moisture, with a wetting layer overlying a layer of uniform moisture content determined by antecedent events. The parameters of this component are the capillary potential or soil suction at the wetting front for field capacity conditions (SWF), a parameter which varies the effective capillary potential over a range (RGF) as a function of BMS, and the saturated soil conductivity (KSAT). Imputs to this component are unit rainfall data and the values of BMS and SMS derived from previous times. The output is rainfall excess, which is the remaining rainfall after abstractions by infiltration.

Surface runoff routing is based on the Clark form of the instantaneous unit hydrograph. The single parameter is a linear reservoir routing coefficient (KSW), but in addition a time area curve is derived which distributes the excess rainfall into a translation hydrograph. The input to this component is the rainfall excess output computed from the infiltration component, and the output is the storm runoff hydrograph, Table 1 summarizes the eight model parameters.

Table 1. -- near here

The output from one component is the input to the next. Even a model as simple as this one has many interactions among the parameters. This is particularly true of the antecedent moisture accounting and the infiltration components. It is often possible that adjustments of a parameter in one component can be compensated for by an adjustment in a different component in another parameter. Over some error range, there may be many sets of parameter values which fit a given set of data equally well. Even though the parameters of the model are chosen so as to be analogous to physical parameters in a basin, the degree of non-uniqueness in the optimum set of derived parameter values may mask the relation of the values to their supposed physical prototype. Thus the conceptual physical equivalence of the model may be lost in the fitting process. This point will be covered more fully later.

Table 1. -- The eight model parameters
and their application in the modeling process.

Parameter Identifier	Units	Application
SWF	inches	Suction at the wetted front for soil moisture at field capacity
RGF	- -	Ratio of the suction at the wetted front for soil moisture at wilting point to that at field capacity
KSAT	inches per hour	The minimum (saturated) value of hydraulic conductivity used to determine infiltration soil rates
BMSM	inches	Soil moisture storage volume at field capacity
EVC	- -	Coefficient to convert pan evaporation to potential evapotranspiration values
DRN	inches per hour	A constant drainage rate for redistribution of soil moisture
RR	- -	Proportion of daily rainfall that infiltrates the soil
KSW	hours	Time characteristic for linear reservior routing

The Infiltration Component

Infiltration is the term used to describe the entry into the soil of water available at the soil surface. When rain falls on a soil it either infiltrates, goes into detention storage, or become surface runoff. The rate of infiltration into the soil is, of course, limited by the supply rate of rainfall. Darcy's law describes flow of a liquid in a homogeneous porous medium, and is the basic mathematical description of the infiltration process.

Many empirical equations have been used to spproximate the infiltration process. One of the more physically meaningful equations is that of Philip (1954, Green, 1911), which has been used as the basis for the infiltration component in the flood hydrograph synthesis program. The Philip equation assumes a two-part soil moisture distribution, as shown in figure 2.

Figure 2. -- near here

A soil column of initial moisture content, m_o, is infiltrated by water which wets a thickness of soil, x, to a uniform liquid content, $\bar{m}$. Both $\bar{m}$ and m_o are relative moisture contents of their respective soil columns, with $\bar{m}$ representing moisture content at field capacity.

The wetting front is at the depth, x, below the soil surface. The equation assumes that the velocity of flow throughout the wetted column and the soil suction at the wetting front each is constant. The capillary potential which exists in an unsaturated soil acts to move moisture from wetter to drier portions of a soil column.

With these assumptions, Darcy's law reduces to

$$\frac{Vx}{k_h} = P + x + H$$

or

$$V = k_h \left[1 + \frac{P + H}{x} \right] \tag{1}$$

Where V is the downward velocity of flow in the infiltrating column (units of L/T), k_h is the capillary conductivity (units of L/T) at soil moisture $\overline{m}$, P is capillary potential at the wetting front (units of L), and H is the depth of ponded water at the surface (units of L). The capillary potential usually is several orders of magnitude larger than the depth of ponded water, so that the H term may be ignored. Because

$$V = di/dt \tag{2}$$

and

$$X = i/(\overline{m}-m_o) \tag{3}$$

equation 1 becomes

$$\frac{di}{dt} = k_h \left[1 + \frac{P(\overline{m}-m_o)}{i} \right] \tag{4}$$

where i is the accumulated infiltration in the wetting column, and is denoted by the hacured area on figure 2. The mnemonic identifiers used to designate equation 4 in the computer program and in this paper are

$$FR = KSAT \left[1 + \frac{PS}{SMS} \right] \tag{5}$$

where

FR	=	di/dt	(units of L/T)
KSAT	=	k_h	(units of L/T)
PS	=	$P(\overline{m}-m_o) \big\rvert_{effective}$	(units of L)
SMS	=	i	(units of L)

The capillary potential at the wetting front is not a constant, but varies depending upon initial soil moisture condition. Colman and Bodman, (1944) state, in a paper used by Philip for some of the justification for his equation, that "of the changed conditions brought about by using moist rather than air-dry soils, the observed results indicate the particular importance of the lowered potential gradient at the wet front". However, there is no method given by Philip or Colman and Bodman for determining the variation of the potential. The flood hydrograph simulation program determines the effective value of PS as varying linearly between a value at wilting point and a value at field capacity. This requires two parameters. The first is the effective value of the product $P.(\bar{m}-m_o)$ at field capacity (SWF). The other is the ratio (RGF) of the product at wilting point to that at field capacity. The effective value of the product of capillary potential and soil moisture deficit is described by a linear relation to soil moisture deficit and is computed as

$$PS = SWF\left[RGF - (RGF-1)\frac{BMS}{BMSM}\right] \tag{6}$$

where BMS is the beginning soil moisture storage in the soil column and BMSM is the maximum moisture storage in the soil column at field capacity. Therefore, the ratio of BMS to BMSM is equal to the ratio of m_o to $\bar{m}$. This relationship is shown in figure 3.

Figure 3. -- near here

Equations 5 and 6 represent the approximation used for infiltration at a point. Equation 5 is a differential equation with a variable coefficient because the soil suction coefficient, PS, is a function of soil moisture, as shown in equation 6. Infiltration occurs over a basin at varying rates, however, the flood hydrograph synthesis program uses a scheme first presented by Crawford and Linsley (1966,p.210) in order to convert point potential infiltration to net infiltration over a basin. Letting SR represent the supply rate of rainfall for infiltration and QR represent the rate of generation of excess precipitation which does not infiltrate, the equations are

$$QR = SR^2/2FR \qquad SR < FR \tag{7a}$$

$$QR = SR-(FR/2) \qquad SR \geq FR \tag{7b}$$

The schematic representation of the relations is shown in figure 4. The

---

Figure 4.-- near here

---

relation maybe interpreted as describing the probability distribution of potential infiltration over the basin by a straight line, with net infiltration being the average over the basin. However, no claim is made that equation 7 actually is a representation of the probability distribution of potential infiltration. Certainly such a distribution would not be linear, as implied by the equation, nor would its shape be similar in time. Rather, equation 7 is an empirical tool which eliminates the absolute threshold value for infiltration. Thus, there is some runoff from any volume of rainfall, although for low intensity rains with dry soil conditions the runoff is quite small. The major justification for equation 7 is that it aids in the modeling of the runoff volumes for the smaller, low-intensity storms.

Equations 5, 6, and 7 together describe the infiltration component. The flow chart for the infiltration component is shown in figure 5.

Figure 5. -- near here

The Soil Moisture Accounting Component

The soil moisture accounting component in a rainfall-runoff simulation model determines the effect of antecedent conditions on the infiltration component. Although the moisture accounting system in this model was designed to represent the physical process to a large degree, the lack of full physical equivalence, in application, may result in a curve-fitting process so that the fitted parameter values have more apparent than real physical meaning. In addition, there is a necessary constraint that the soil moisture accounting component must be compatible with the infiltration component if a water budget is to be maintained throughout the system. These two facts should be kept in mind throughout the description of the soil moisture accounting component.

The soil moisture component in the flood hydrograph simulation program is based upon the Philip scheme described for the infiltration component. The total moisture in storage in the soil column is divided into two parts. The first is contained in a base moisture storage (BMS) at a soil moisture which may vary from field capacity to wilting point. The second is a surface moisture storage (SMS) at field capacity. Thus, the total infiltrated column is assumed to be at field capacity. This is based upon the results shown by Colman and Bodman in the paper mentioned earlier in the description of the infiltration component. The thin saturated soil layer at the surface which exists during active infiltration is ignored. A schematic diagram of the soil moisture accounting is shown in figure 2.

SMS depicts accumulated infiltration, and all infiltration during storm periods is added to SMS. BMS, on the other hand, is used to compute the relative soil moisture deficit. The unhachured area in figure 2 represents BMS. BMS and the ratio RGF are together used to compute PS, the effective value of the product of the capillary potential and the soil moisture deficit, which also is a part of the infiltration equation.

Evapotranspiration losses are assumed to occur at the potential rate. All evapotranspiration demand is met from SMS, if possible. When storage in SMS is zero, evapotranspiration then occurs from BMS.

Drainage occurs from SMS to BMS at a constant rate as long as storage exists in SMS. Storage in BMS has a maximum value (BMSM) which is equivalent to the field-capacity moisture storage of an active soil zone. Zero storage in BMS is assumed to correspond to wilting-point conditions in the active soil zone. When storage in BMS exceeds BMSM, the excess is spilled to deeper storage. The spills could be the basis for routing interflow and base flow components, if desired. However, these components of streamflow are not modeled in the flood hydrograph simulation program. If other components of flow comprise a significant portion of the flood peak, a routing of these spills would be necessary.

The Surface Routing Component

The excess precipitation generated in the flood hydrograph simulation program must be converted into a flood hydrograph by a routing scheme. The Clark flood-routing method (1945) is used to develop the basin unit hydrograph. The Clark method has two parts. First, the excess precipitation is converted into a translation hydrograph which represents the effect of varying travel times in the basin. The translation hydrograph for the basin is represented by a time-discharge histogram. The time-discharge histogram is developed from the distance-area histogram for the basin. In essence the derivation assumes that distance and travel time are directly proportional. Because of variation of both resistance to flow and of channel slope over the basin, the assumption of proportionality of distance and travel time does not necessarily hold. Therefore a comparison of the shapes for simulated and observed hydrograph for several flood events may be used to revise the time-area histogram to a more appropriate shape for a study basin.

The translation hydrograph must be routed through some element representing storage in the basin. For an instantaneously developed excess precipitation of one inch, this results in the instantaneous unit hydrograph. The Clark method assumes a linear time-invariant storage. Dooge (1959) presents an excellent discussion of unit hydrograph methods, and the place of the Clark method in the general theory. Figure 6 pictures the operation of the Clark method.

Figure 6. -- near here

SYSTEM IDENTIFICATION

The method of determining optimum parameter values is based on an optimization technique devised by Rosenbrock (1960) and referred to by Wilde (1964) as the method of rotating coordinates. It is a hill climbing procedure that does not require the evaluation of partial derivatives of the objective function with respect to the parameters. All parameters must be bounded for the method to be used. Thus, parameter values may be constrained to a range of "reasonable" values if desired. The utility of the procedure as related to system identification in the field hydrologic modeling has been discussed by Dawdy and O'Donnell (1965).

The method revises the parameter values and recomputes the objective function with the revised set. If the result is an improvement, the revised sets is accepted; if not, the method returns to the previous best set of parameters. The objective function, or U-function, throughout this study is based upon the sum of the squared deviations of the squared deviations of the logarithms of peak flows, storm volumes, or some combination of both. Thus, the fitting procedure develops a non-linear least-squares solution.

The logarithms of flows are used because streamflow errors generally are more nearly equal in percentage than in absolute terms. Thus, if a peak of 1000 cfs is estimated in error by an average of 100 cfs (10 percent), a peak of 5000 cfs will have a greater probability of an average error of 500 cfs (10 percent) than of 100 cfs (2 percent). The logarithmic transformation is meant to make the error of estimation more commensurable for the large and the small peaks. The sum of the squared errors is used as an objective function because of the mathematical property that it is a convex function, and because of its direct analogy to least squares fitting in standard linear statistical theory. More concerning this point will be discussed later in the report.

Rosenbrock's method of optimization proceeds by stages. During the first stage each parameter represents one axis until arbitrary end of stage criteria are satisfied. At the end of each stage a new set of orthogonal directions is computed, based on the experience of parameter movement during the preceding stage. The major feature of this procedure is that after the first stage, one axis is aligned in a direction reflecting the net parameter movement experienced during the previous stage.

To start the fitting process, the hydrologic model is assigned an initial set of parameter values and the resulting simulated flood hydrograph response is computed. The objective function is calculated and stored as a reference value, which is used to evaluate the results of subsequent trials. A step of arbitrary length is attempted in the first search direction. If the resulting value of the objective function is less than, or equal to, the reference value, the trial is registered as a success, and the appropriate step-size, e, is multiplied by $\partial > 1$. If a failure results, the step is not allowed and e is multiplied by $-\beta$, where $0 < \beta < 1$. An attempt is made in the next search direction, and the process continues until the end of stage criteria are met. At this point, a new search pattern is determined and another stage of optimization undertaken. Only a limited amount of information is output during optimization. The U-function value and associated parameter values are printed for each successful trial. In addition, a listing by flood event of the simulated hydrologic response and of observed data are output at the start of each stage.

It must be stressed that the concept of automatically determining optimum model parameters requires that the objective function be compatible with the intended use for which the fitting is undertaken. In order to give weight to both the volume and shape characteristics of the flood hydrograph, a weighted objective function (U3), which included both peak and volume error has been used. One component of the objective function used in optimization is the sum of squared log deviations between recorded and simulated flood peaks (U1). Another component, (U2), is the sum of squares log deviations between estimated and simulated surface runoff for each storm period.

Estimated surface runoff is calculated by a crude hydrograph-separation technique that integrates the volume of runoff under the flood hydrograph, from the start of the storm period through the period of rise, and for a duration of recession after the peak. The contribution from base flow is deducted and assumed to equal the volume derived from projecting the level of discharge at the start of the rise throughout the period of integration. Recorded flood peaks are similarly reduced by the antecedent discharge level to account for the contribution from base flow.

General Discussion of Response of the Model

The game of hydrologic simulation is based upon engineering approximations. Approximations introduce errors into simulation results. In order to properly utilize a model, some understanding must be gained concerning the magnitude of errors produced through use of the model.

Errors in data are reflected in errors in the fitted parameters in a simulation model. If perfect input data are routed through a perfect model, the output produced would agree perfectly with an error-free output record. If errors are introduced into the input or output record or both, the output will not be exactly reproduced even by a perfect model. If a fitting process is used, the parameters will deviate from their true values in order to minimize the deviations between the simulated and recorded traces as specified in the objective function. The "optimal" set of parameters will now be in error, and the value of the objective function after fitting will be less than its "true" value. This is so because the value has been so derived as to find the minimum value for the objective function.

This process is analogous to statistical least squares analysis. The fitted parameters deviate from their population values because of random errors in the data. The standard error of estimate is a measure of error in the data. The standard error of prediction, however, is somewhat greater than the standard error of estimate, for it includes both the measure of lack of fit of the data used to calibrate the model and the measure of error in the fitted parameters. These relationships are shown in Table 2.

Table 2. -- near here

Table 2. -- Qualitative comparison of errors involved in hydrologic modeling with analogous errors resulting from standard linear statistical analysis.

Source of error	Qualitative size of error variance	Statistical analog
Measured data	a	Measurement and sampling error variance
Differences between measured and simulated flows during the calibration period	a-b	Square of standard error of estimate
Differences between measured and simulated flows <u>outside</u> the calibration period	a+b	Square of standard error of prediction

If the assumptions of regression theory were valid, for a linear model with normally distributed and homoscedastic errors of the dependent variable, the standard error of prediction could be computed from the standard error of estimate, the deviations of the independent variables from their mean, and the error in the coefficients for the independent variables. These assumptions seldom hold, however, so that competent statisticians often resort to split sample testing. The assumptions also fail for hydrologic simulation and, in addition, the models are non-linear, so that there is no theory by which to compute the error of prediction. Therefore, split sample testing must be used in this case whenever possible.

Non-linearity of the hydrologic process precludes at this time any theoretical description of the mechanism by which errors in data are transferred to model parameters and then are combined with input errors in the test period to produce errors in the simulated streamflow. An empirical study for the response of the model is shown in table 3.

Table 3. -- near here

A recorded rainfall trace was assumed error free and routed through an optimized set of parameters for the Little Beaver Creek basin near Rolla, Missouri. The optimized parameter values were assumed to be correct values,to obtain a "true" streamflow trace. Then a random error with mean zero and standard deviation of ten percent was applied to all rainfall values. These "erroneous" rainfall values were then routed through the model with the "true" parameter values, and the resulting value of the objective function for the simulated streamflow trace computed. An optimization run then was made which adjusted the parameters to minimize that value. The "optimized" set of parameters is shown, along with the resulting value. The "true" rainfall trace was then routed through the new optimized parameters and the objective function evaluated. Assuming independence of the two souces of error, one in the input data and the other in the model parameters, the error of prediction should be approximately equal to the square root of the sum of the squares of the two separate estimates of error. To test this relation nine independent sets of random errors were applied to the rainfall values and routed through the model using the optimized parameter values. The average U value for these nine test runs is also shown in table 3.

Table 3. -- Results of an empirical study of the response of the model to input and output errors.

Parameter	Assumed True Values	Values Optimized to Rainfall errors of 10%	20%	Values Optimized to Streamflow errors of 5%	10%
SWF (in)	3.6	3.6	3.8	3.7	3.7
KSAT (in-hr)	.063	.063	.06	.063	.061
KSW (hr)	1.0	1.04	1.06	.98	.98
EVC	.56	.57	.58	.559	.56
BMSM (in)	4.0	4.02	3.98	4.04	4.04
RGF	12.0	11.9	11.94	12.12	12.21
RR	.8	.796	.8	.796	.796
DRN(in-hr)	.020	.018	.017	.020	.019
U, Pd*	--	.0150(12)Δ	.0538(23)	.00233(4.8)	.00915(9.6)
pd	--	.0097(9.9)	.0493(22)	.00170(4.1)	.00708(8.3)
pD	--	.0039(6.3)	.0152(12)	--	--
U, Test**	--	.0196(14)	.0890(30)	--	--

*P = True parameters, p = optimized parameters, D = correct data, d = erroneous data.

**Average of nine separate test runs.

ΔFirst value is the average of two thirds of the squares of differences of natural logarithms of the sample peaks plus one third of the squares for the sample storm volumes. The second value in parentheses converts the first value to an equivalent "percent standard error" by SE = antilog U, and averaging plus and and minus percentages.

For the case of 10 percent rainfall errors, the error introduced by data errors (Pd) is 0.0150, while that for parameter errors (pD) is 0.0039. The sum of these is 0.0189 which is to be compared with 0.0196 (U Test). The comparison of the error of prediction based upon the variances above is 13.8 percent (Pd+ pD) as compared to 14.3 (U Test).

Similar results are shown for input rainfall errors with a 20 percent standard error. As was the case for the 10 percent errors, the error in simulated output was magnified so that it is about 20 percent greater than the rainfall error (Pd is 23 percent as compared to the previous value of 12 percent). Once again Pd+ pD should combine to produce a value comparable to that for the test results, and 0.0538 + 0.0152 = 0.0690 is to be compared with 0.0890. The respective percentages are 26 and 30 for estimates of the error of prediction.

Errors in streamflow measurement are transferred to model parameters in the fitting process. An example of this is shown in Table 3. Errors of 5 and 10 percent were introduced into runoff estimates, and a set of best fit parameters derived. The rainfall and runoff errors are independent in this study, so that the square of the error of prediction for 10 percent runoff errors and 20 percent rainfall errors would be of the order of the sum of the two variance terms, 0.0890 for rainfall errors and 0.00915 for runoff errors, which yields a 32 percent error for the two combined, as compared to 30 percent for rainfall alone.

Two points are of interest in the above results. First, rainfall errors have a magnified effect on the simulated streamflow for basins similar to the one chosen for the study. This probably is true for most basins with drainage area less thanten square miles. Therefore, rainfall errors probably are the controlling factor determining accuracy of streamflow simulation. Second, the response of this quite non-linear hydrologic model is approximately linear for errors in rainfall of the order of magnitude assumed, which probably are of the order of magnitude generally encountered in a field case. One would expect that errors in rainfall input would result in greater proportional errors in predicted excess rainfall. This results from the fact that abstractions through infiltration, at least in this model, are not affected by rainfall intensity. However, amounts of excess rainfall for different time periods from different parts of the basin are combined in the translation routing, and the storage routing attenuates errors by averaging by means of the storage process. Therefore, the relative size of rainfall errors and of errors of estimated streamflow depends upon the extent to which the model of the routing process attenuates the magnification of errors produced in model estimates of excess rainfall.

The fact that the errors of streamflow estimates are approximately linearly related to errors of rainfall input data is particularly important. The linearity of errors indicates that there may be some hope for the derivation of a theory of errors for streamflow simulation. In addition, the linearity gives some post hoc justification for the non-linear least squares fitting technique used in the fitting process.

Case Studies Using the Simulation Model

The proof of the pudding is in the eating thereof. The empirical study described in the preceding section does give insight into the modeling process and, in particular, into the operation of the model. However, to better illustrate the utility of the model in field application, three case studies have been undertaken, They represent a range in location and hydrology. The basins are Santa Anita Creek near Pasadena, California, a semi-arid basin, Beetree Creek near Swannanoa, North Carolina, a humid basin, and Little Beaver Creek near Rolla, Missouri, with a hydrology typical of the interior United States. All three basins have pronounced relief.

The data available varied from basin to basin. In addition, the relative stage of development of the model led to emphasis of difference research areas. Sufficient rainfall data were available for Santa Anita Creek Basin so that a study could be made of the effect of bias of rainfall measurements and of the effect of time and space variability of rainfall on modeling results. Beetree Creek Basin was used to study the effects of split sample testing, and, in addition, was used to study the methodology of the use and limitations of the objective curve-fitting method. Each will be discussed separately, then a discussion of the overall results and of the problems encountered will be presented.

SANTA ANITA BASIN DATA

General Physiography

The Santa Anita Creek drainage basin is a 9.7 squre mile (25 square kilometers) area of the San Gabriel mountains in southern California. The rugged topography ranges in elevation from 1,500 to 5,700 feet (460 to 1,700 meters) above sea level with the mean about 3,600 feet (1,100 meters). Thin porous soils covering a highly fractured bedrock combine to give the basin high moisture retention and absorption properties. The southerly facing basin receives about three-fourths of its rainfall during cool winters. The climate and soils support a thin to dense growth of chaparral native to the area.

Precipitation

The precipitation measuring network on the Santa Anita Creek basin consists of 6 stations for a 14-year period ending with the 1962 water year. The six sites provide good areal (figure 7) and elevation (table 4) coverage

---

Figure 7. -- near here

Table 4. -- near here

---

of the basin annual precipitation. A double-mass analysis of the 14-year annual precipitation values show fair measurement consistency among the 6 stations. Three sites with continuous recorders (Stations 477, 60, and 338) provided the rainfall data required by the simulation model.

Table 4. -- Mean Annual Rainfall (1949-1962)

Station	Elevation feet	Mean Annual Rainfall inches	Deviation from Basin Mean[1] Percent
58 - Sturdavent	3,255	30.57	
60 - Hoegee	2,500	30.44	+ 3.2
63 - Big Santa Anita Dam	1,400	19.47	
338 - Mt. Wilson	5,710	25.21	- 13.5
432 - Fern Lodge	2,035	26.30	
477 - Spring Camp	4,670	28.89[2]	+ 0.6
Basin mean:			
Thiessen:			
All station		28.3	
Three-station		29.5	
Elevation-area		29.5	

[1]Basin mean used was 29.5 inches.

[2]Adjusted on the basis of double-mass analysis.

Two methods were used to determine the basin mean annual rainfall during the 14-year period. First the standard Thiessen method gave a mean of 28.3 inches (71.9 centimeters) when all 6 stations were considered and a mean of 29.5 inches when only the 3 recording sites were used. The second method, a numerical integration of the relationships for elevation-percent area obtained from topographic maps and elevation-annual rainfall defined by the 6 gage records (figure 8), gave an annual

---

Figure 8. -- near here

---

mean of 29.5 inches. 29.5 inches was chosen as the estimate of the mean annual and used to evaluate the relation of individual station rainfall to basin-mean rainfall.

The 24 storm periods selected for simulation had both complete records of the rainfall occurring at the 3 recording stations and a significant rise in stream discharge. The records for the storm periods were reduced to 15-minute volumes. Daily rainfall records were used between storms. The storm data were compiled from gage charts provided by the Los Angeles County Flood Control District.

Streamflow

The streamflow data used for fitting the model to the Santa Anita Creek Basin were those for the U. S. Geological Survey.gaging station near Pasadena, California. The site has been gaged since 1916. The mean flow has been 5.5 cfs (0.16 m^3/sec) or 7.7 inches (0.2 meters) per year. The maximum flow of about 5,200 cfs (147 m^3/sec) occurred in March 1938. The peak discharges during the storm periods selected for study ranged from 17 to 2,530 cfs (0.5 to 71.6 m^3/sec) (see Table 5).

Table 5. -- near here

Evaporation

Daily values of pan evaporation at Tanbark Flat were obtained from the Pacific Southwest Forest and Range Experiment Station, U. S. Department of Agriculture. The Tanbark Flat climatic station is located in the San Dimas Experimental Forest about 10 miles (16 kilometers) east of the Santa Anita Creek basin and is at an elevation of 2,800 feet (850 meters). The mean annual evaporation from a standard Weather Bureau pan is in excess of 60 inches (1.5 meters).

Table 5.--Simulated peak dischrages using fitted parameters

		Discharge in cubic feet per second[1]			
No.	Observed	Simulated with Adjusted Data (C)			
		Station			
		477	60	338	Mean[2]
1	2529	2506	2318	2722	2515
2*	1472	2846	1667	1198	1904
3*	338	917	296	346	520
4	194	224	221	225	223
5	342	532	405	454	464
6	45.8	44.4	53.4	63.8	53.9
7	30.6	37.2	38.5	35.2	37.0
8	108	184	146	148	159
9*	34.1	35.9	21.5	30.3	29.2
10*	50.1	51.9	14.5	27.3	31.2
11	660	601	518	472	547
12	150	121	149	130	133
13	156	166	190	129	162
14	111	113	99.6	96.6	103
15	361	283	417	318	339
16	332	238	228	253	240
17	837	779	696	512	662
18*	709	372	97.9	42.3	171
19	243	190	196	161	182
20	15.4	31.6	28.9	35.4	32.0
21*	58.4	18.3	36.2	14.3	22.9
22	55.0	28.1	29.7	21.9	26.6
23	91.2	54.1	45.9	69.3	56.4
24	1235	1238	1303	1362	1301
U1		2.07	1.76	2.73	1.92

*Events not included in fitting or the computation of U1.

[1] Observed and simulated discharges do not include base flow.

[2] Average of the simulated peaks for the three stations.

Data Screening

The amount of rainfall data available for the Santa Anita basin was sufficient to investigate the effects of variability of measured rainfall upon simulation results. The record of several storm events indicated a large spatial variation in total storm rainfall over the basin, as indicated by the deviation in percent of measured storm volume at each site from the weighted mean for each storm. Several storms also appeared to have a large spatial variation in total storm rainfall over the basin, as indicated by the deviation in percent of measured storm volume at each site from the weighted mean for each storm. Several storms also appeared to have a large spatial variation of rainfall intensities over the basin. On the basis of preliminary screening, six storms of the twenty-four available for analysis were not used in fitting the parameters. However, these peaks were simulated, and results are shown on the scatter diagrams. Only one of the excluded storms might have significantly changed the results. The records for that storm show extremely high intensities for very short periods of time; the 15-minute time interval used to define the rainfall records appears to be inadequate for accurate simulation for that storm. The purpose of screening is to eliminate storms with extreme errors in data input so as to minimize the effect of data errors on the fitting process.

Parameter Definition

Nine model parameter determinations were made as a series of three fittings for each of the three rainfall stations. The first in the series of three fittings was made by using the data as recorded at the stations (set A). These results are analogous to those for simulation studies for which a single recording rain gage is available in a basin, and for which there is no basis for adjusting the record to obtain a better estimate of mean basin rainfall.

The second series of parameter determinations was made for each station by adjusting the recorded storm volumes by a constant station factor (set B). These factors were computed so as to adjust the mean annual depth at the station site to 29.5 inches computed for basin mean annual rainfall as explained earlier. These results are analogous to those for simulation studies for which a recording gage is available in a basin, and supplementary data are available to determine an average annual rainfall on the basin and at the gaged site.

The third fit was made to the data with the storm volumes adjusted to a three-station Thiessen weighted mean for each event (set C) i.e., the mean basin volume was distributed in time in accordance with the rainfall intensity pattern for each individual station. These results are analogous to those for simulation studies for which a recording gage plus several non-recording rain gages are available in a basin. Thus, a weighted mean basin rainfall for each storm can be derived.

To summarize, the various rainfall intensities are adjusted as follows:

Set	Adjustment
A	$\hat{R}_{ij} = R_{ij}$
B	$\hat{R}_{ij} = a_{.j} \cdot R_{ij}$
C	$\hat{R}_{ij} = a_{ij} \cdot R_{ij}$

where R_{ij} is the measured intensity for period i at station j, $\hat{R}_{ij}$ is the adjusted intensity used in the given simulation set, $a_{.j}$ is an average adjustment which is the ratio of mean annual rainfall over the basin to the mean annual rainfall measured at station j, and a_{ij} is the ratio of average rainfall over the basin for storm i to that volume measured for storm i at station j.

The results of the nine fittings are given in table 6. In addition to parameter values the average squares of deviations between logarithms of

---

Table 6. -- near here

---

simulated and observed peaks, U1, is given for each set of parameters. Table 5 shows the values of the simulated peaks for data set C. Figure 9 shows typical scatter diagrams for data set A.

Figure 9. -- near here

Table 6. -- Fitted Parameter Values

Parameter	Input Series*	Station 477	60	338
SWF	A	21	22	11
	B	20	20	16
	C	20	18	17
RGF	A	7.5	6.1	4.4
	B	6.7	5.9	5.5
	C	5.6	6.5	6.1
KSAT	A	0.32	0.32	0.31
	B	0.31	0.32	0.25
	C	0.32	0.32	0.31
BMSM	A	4.1	3.4	2.1
	B	4.0	3.4	3.6
	C	3.5	3.5	3.5
EVC	A	0.52	0.73	0.52
	B	0.59	0.71	0.80
	C	0.74	0.72	0.74
DRN	A	0.049	0.058	0.030
	B	0.045	0.057	0.043
	C	0.056	0.059	0.057
RR	A	1.14	0.098	1.03
	B	1.08	1.00	0.90
	C	0.96	1.01	0.95
KSW	A	2.4	2.8	2.2
	B	2.5	2.7	2.3
	C	2.6	2.8	2.3
U1, fit criteria	A	.097(32)	.123(35)	.440(-)
	B	.100(32)	.122(35)	.438(-)
	C	.115(35)	.098(32)	.153(40)

* Input Series: A - data as recorded.
B - rainfall volumes adjusted by mean annual factors
C - rainfall volumes adjusted by mean storm factors

Figures in parentheses are average errors in percent.

Parameter Sensitivity

The sensitivity of the goodness of fit criterion to changes in parameter values is helpful in discussing parameter importance and simulation results. An expression of sensitivity of the error criterion to given parameters can be obtained by performing repeated simulations while incrementing the parameter, holding all other parameters to their fitted value, and observing the change in value of the fitting criterion. This gives no measure of interaction of the parameters, but is a simple measure of how critically the simulation results are dependent upon the individual parameters. The results of this procedure when applied to data set C for station 60 (C-60 on table 6) are shown graphically in figure 10. The figure is a plot of criterion value versus the percent

---

Figure 10. -- near here

---

change in parameter values. Applying this procedure to the other data sets produced similar relationships.

Analysis of Results

The results of the nine separate optimization runs, three for each of the three stations, are summarized in table 6. Shown are fitted parameter values and the resulting goodness of fit, U1, which is the peak simulation portion of the total fitting criterion used. The representativeness of the rainfall data is least for inputs A and most for inputs C. It would be expected that as the data for the three records became more similar to each other that the fitted parameter values would converge to common values for the three stations. As the data become more representative the accuracy of fitting should increase, and the U1 values decrease. The effects of the various components of error can be seen by comparison of variability of parameter values and goodness of fit between stations for a given input set.

Parameter Values

Prediction depends upon the fitted parameter values for the model as well as upon the data used for the prediction period. The more stable the estimated parameter values the better the possibility of relating the fitted values to measures of the basin. Thus, variability of fitted parameters for the nine optimization runs may give insight into the degree to which model parameters are influenced by data errors.

A wide range in fitted parameter values resulted when the data were used as recorded at the ~~time~~ three stations, input A. None of the three sets of parameters can be considered unlikely when viewed individually. Together the sets of values illustrate a possible range depending on the data representativeness. In a practical case available data may consist of only one record which must be used without knowledge of its degree of representativeness. The variability of fitted parameter values such as those for set A will affect the feasibility and accuracy of any regionalization of parameter values.

Input set B contains both time distribution and total volume errors but has been adjusted to reduce the gaging bias as a result of errors in the estimate of mean annual precipitation over the basin. The reduced range in parameter values except for KSAT and RR indicates a better estimation than was obtained in series A. The relative insensitivity shown for KSAT in figure 10 is for the independent effect of KSAT in the first term of the Philip infiltration equation. Accurate determination is not possible and may not be important. The range of values for RR between series A and B are about equal. Values of RR greater than 1.0 reflect curve fitting in the model, and partly result from the differences between rainfall measured at a point and average rainfall over the basin. No constraint was placed on the value of RR for these optimization runs.

Input set C has the same estimate for storm volume at all stations for each storm. The only variability is that introduced by the different time distributions within a given storm as recorded at the three stations. All parameters have relatively stable fitted values. The variation has been reduced to within $\pm$ 10 percent for all parameters, with only the infiltration parameters, PSP and RGF, and the routing parameter, KSW varying by more than 5 percent. The overall correspondence between stations 477 and 60 in series C is very close, especially when the direct interaction of PSP and RGF is considered.

Fitting Errors

The measure of goodness of fit, U1 in table 6, is the average of the squared deviations between logarithms of computed and simulated peaks, and is analogous to a variance or the square of a standard error.

There are several components of error in the simulation results. Assuming the errors are independent, the results of a simulation run if stated in terms of components of variance might be represented as

$$Q + M + R + V + T - C = U \tag{8}$$

where Q is the variance of error in the computation of discharge, and results from measurement error, from error in rating analysis, and from undefined rating changes. M is the variance due to approximations of the model, and results from the fact the physical laws are not exactly known and, where know, may be approximated for convenience or speed in computation. Both Q and M remain the same for all three sets of data. R is bias error resulting from the use of incorrect mean annual rainfall conditions for the basin. The purpose of the adjustments for data set B was to minimize this bias as much as possible for the given amount of data. This was accomplished, as stated, by using all the data to estimate mean basin rainfall, then adjusting each measured station mean to the estimate of the basin mean. V is error introduced by the fact that a point measurement of volume for a particular storm differs from the mean basin volume for that storm. The purpose of the adjustments to obtain input set C was to minimize this error component. This was accomplished by using all data available to estimate mean storm volume for each storm. T is error introduced by the fact that point measurements of time variability of intensity within a storm differ, and any point measurement differs from an "effective time distribution" which best represents average conditions over the basin for simulation purposes. Probably the only way to minimize the component V would be to use an input which varies over the basin. C is the curve-fitting error introduced into the model parameters by a fitting process. The parameter values are perturbed from a global "best" set of values in order to minimize the fitting criterion, U, so that C is negative in sign. For use of the model in prediction, the curve fitting adds to the error, as indicated in table 2.

The fitted error criteria of set A for all three stations are quite similar to those for set B, although set A rainfall values are not adjusted to mean basin conditions. The bias in the recorded rainfall at each station was compensated for by the curve-fitting ability of the model to adjust parameter values. On the basis of these data, bias in amount of recorded rainfall affects the resulting fitted parameter values rather than the accuracy of fit. As the result of a change in value of the fit criterion of less than 1 percent, the parameter values from station 338 have changed so much that the parameter values for set B have a maximum of 1.36 for the ratio of highest to lowest value, the ratio for parameter EVC. For Set A five parameters had ratios greater than 1.36, PSP, RGF, BMSM, EVC, and DRN. The fitted parameter value for station 338 is one of the extreme values for each of those parameters in both set A and set B. Thus, the errors seem to be transferred from the data to the parameters, as is particularly evident for station 338.

Input set C contains variability among the three inputs only in the time distribution of rainfall. The goodness of fit for this set ranged from 0.100 for station 60 to 0.152 for station 338. Converting the range of 0.052 to an average percentage error for the peaks yields an estimate of about a 23 percent error in peak discharge reproduction introduced by time variability alone. Therefore for a basin with this degree of variation in rainfall patterns and the relative smoothing action introduced by the model and, hopefully, by the hydrology, an average error of as much as 20 percent for simulated flood peaks can be introduced by the time distribution error alone. Considering only the two "better" or seemingly more representative gages, the difference in fitted U1 values is, 0.017 which gives an average percentage error of 13 percent introduced by time distribution error in a "good" record.

In set C the most representative gage, judged in terms of goodness of fit, was that closest to the center of the basin. The least representative was on the perimeter and at the highest elevation of the basin. Therefore, relative representativeness was about as expected.

Input set B contains both time distribution erros within a storm and storm volume errors. The records have been adjusted to minimize only the station bias in relation to basin mean annual rainfall. The results of Set B runs indicate that station 447 probably is the most representative station for predicting storm volumes, just as results of set C runs indicate that station 60 probably is the most representative for time distribution of rainfall within a storm.

An estimate of the volume error component for station 60 should be about the sum of the differences between the values of the objective functions for the B and C runs for the two stations. Thus, volume errors can introduce as much as 0.04 to U1, which is on the order of 20 percent errors. The compounding of the time distribution errors of station 477 and the storm volume errors of station 60 would give a U1 of 0.057, which leads to a possible combined rainfall data error component on the order of a 24 percent standard error.

Effect of Screened Data

All data used in fitting was screened for gross flyers or outliers. The fitted parameters will predict within the indicated range of accuracy for other data which contain the same range of errors as in the screened data. The screened data used for fitting contain the usual range of errors normally encountered. However, grossly inadequate or unrepresentative data will produce outliers well beyond the errors of the indicated prediction. If data are grossly in error, modeling results using that erroneous data should be expected to be in error also.

Discussion of Accuracy of Simulation for Santa Anita Basin

In general, accuracy of simulation of flood peaks for the 18 peaks used in the analysis was on the order of standard error of 32 to 35 percent. Errors introduced by rainfall variability over the basin were on the order of 24 percent. Assuming independence of data errors and model errors, other sources of error contributed about the same amount to the total error. This follows from the fact that, with independence of errors, variances should be additive. Therefore, the vaiance contributed by data error ($24^2 = 576$) plus that contributed by model approximations (M) is equal to the total variance (say $34^2 = 1156$). In order to reduce errors of simulation on this basin, the rainfall input must be refined by the use of information from more than one gage or by some means of using estimates of areal variability in the model other than by the assumption of uniform rainfall distribution as is assumed in the model.

Beetree Basin Data

General Physiography

Beetree Creek drains an area of 5.41 square miles (14 square kilometers) of rough terrain near Swannanoa, North Carolina, on the Western slope of the Great Craggy Mountains in the Blue Ridge province of the Appalachian Highlands (Fenneman, 1938). Land and channel slopes are steep, with elevations ranging from 2,700 feet (820 meters) at the stream-gaging station to 5,600 feet (1,700 meters) at the headwater drainage divide. The basin is approximately rectangular with a main channel length of about 3.2 miles (5.1 kilometers) and an average width of about 1.5 miles (2.4 kilometers). The index of channel slope, given by the ratio of fall over the reach of channel from .1 to .85 of main channel length, is 490 feet per mile (0.00928 feet per foot). The predominant soil is mapped as "stony rough land of Porters soil material" and described as a gray-brown podzolic type derived from granite, gneiss, and schist (Goldton, and others, 1952). Practically all the land supports native forest with small areas of pasture at lower elevations.

Streamflow

The streamflow data used for fitting the model to the Beetree Basin were those for the U. S. Geological Survey gaging station near Swannanoa, North Carolina. The site has been gaged since 1926. The mean discharge during the period 1926-60 was 10.4 cfs (0.29 m^3/sec) or 25 inches over the basin (0.64 meters). The maximum flow of 1,370 cfs (39 m^3/sec) occurred August 13, 1940. The peak discharges during the periods selected for study ranged from 82 to 1,370 cfs (2.3 to 39 m^3/sec), as shown on Table 7.

Table 7. -- near here

Evaporation

Daily values of pan evaporation were obtained from the Tennessee Valley Authority, which maintains a climatic station 4,000 feet (1,200 meters) downstream from the gaging station, at an elevation of 2,540 feet (770 meters). The evaporation record has been collected since 1935, and during the period 1935-1959 the average annual pan evaporation has been 39.9 inches (1.01 meters).

Table 7. -- Storm-Period Data

No.	Date	Storm rainfall (in.)	Peak discharge (cfs)	Surface runoff (in.)
		Sample A		
1.	Apr. 4, 5, 1936	2.08	220	.66
3.	Nov. 14, 15, 1938	2.29	82	.12
5.	Aug. 17, 18, 1939	2.49	236	.50
7.	Aug. 29, 30, 1940	7.36	1180	4.28
9.	Aug. 24, 25, 1941	1.22	94	.15
11.	Mar. 8, 9, 1942	1.27	151	.43
13.	Sept. 20, 21, 1944	1.42	115	.09
15.	Oct. 5, 6, 1945	2.22	117	.39
		Sample B		
2.	Oct. 15, 16, 1936	3.08	218	.62
4.	Jan. 29, 30, 1938	1.74	167	.43
6.	Aug. 11, 13, 1940	10.33	1370	4.42
8.	Dec. 27, 28, 1940	2.59	263	.59
10.	Feb. 16, 17, 1942	1.72	107	.26
12.	Dec. 29, 30, 1942	2.06	208	.74
14.	Mar. 26, 27, 1945	1.88	100	.28
16.	Feb. 10, 11, 1946	1.82	141	.41

Precipitation

The Tennessee Valley Authority has operated a recording rain gage since 1935 at Beetree Dam, located 4,000 feet (1,200 meters) downstream from the stream-gaging station. For the period 1935-1959 the mean annual precipitation was measured as 46.4 inches (1.18 meters) (Tennessee Valley Authority, 1961). In 1948 new equipment was installed for the recording gage, and problems of calibration caused the installation of a non-recording rain gaga beside the recording gage. In addition a recording gage has been maintained at various points in the upper area of the basin, as indicated on figure 11.

Figure 11. -- near here

Data for 40 flood events occurring during the period from April 1936, through October 1964, were assembled by personnel of the Geological Survey from published records and copies of original recording charts. Storm-period rainfall data were compiled on a 15-minute time basis. An analysis of annual rainfall data indicated that an inconsistency occurred in the Beetree Dam record in 1949. A review of the history of the rain gage showed that a change in instrumentation was made in July 1948, when the original installation of a Ferguson recording gage was replaced by a Universal recording gage. On the basis of this information 16 flood events occurring prior to July 1948 were selected for detailed study.

Parameter Definition

To facilitate a split-sample comparison of the results of simulation, the screened test sample of 16 storms was divided into 2 sets of 8 events each. In order to achieve an approximate balance in the range of magnitude of peak-discharge rates represented in each sample, the odd-numbered events were slected to comprise Sample A and the even-numbered events were assigned to Sample B. A summary of the storm-period data appear in Table 7.

Three separate types of optimization were performed on the pre-1948 flood events. First, Sample A was used for fitting and optimum model parameters were derived to predict the events of Sample B. In the second, Sample B was used for fitting to produce a set of optimum parameters which were used to predict Sample A. In the third, all 16 events were used to determine the best-fit parameters for the pre-1948 record.

In each optimization run a 5-week period of daily rainfall and pan evaporation was monitored, prior to the first storm event, to reduce the effect of arbitrarily initializing storage values for SMS and BMS (0 and BMSM respectively). A similar lead-in period was used for all basins, and for all results shown in this paper. In addition, initial optimization runs for all three types were started with the same set of initial parameter values, which were assigned on the basis of (a) assumptions about average soil characteristics, (b) an estimate of the ratio of potential evapotranspiration to pan evaporation, and (c) the recession and timing characteristic of observed flood hydrographs.

Results for the three optimization runs are shown in Table 8.

Table 8. -- near here

Both the optimum fitted parameter values and the fitted objective function values are shown. In addition, for each set of 8 peaks used for fitting, the remaining set of 8 peaks is used as a test sample and the accuracy of prediction is shown. An adjusted accuracy of prediction is also given in which that peak most in error is removed from the predicted set, in order to give an indication of the effect of extreme errors on the fitting criterion.

Table 8. -- Results of fitting of model parameters to data and of split sample testing for Beetree Creek near Swannanoa, North Carolina.

Parameter	Optimum fitted parameter value		
	Sample A	Sample B	All storms
SWF	3.36	4.26	3.62
KSAT	.101	.097	.095
KSW	4.97	6.24	5.67
EVC	.597	.541	.58
BMSM	1.60	1.67	1.87
RGF	14.0	8.15	14.0
RR	.78	.81	.75
DRN	.0050	.0051	.0048
U3A	.069 (27)	.191 (*)	.074 (27)
U3B	.132 (*)	.099 (32)	.107 (33)
U3 all	.101 (32)	.145 (*)	.090 (30)
U3 test-adj**	.079 (28)	.098 (32)	

Figures in parentheses are root mean square error presented as average percentage.

*Average error not computed.

**Peak most in error is removed from the predicted set.

Results and Conclusion

The response of the objective function during two optimization runs is shown in Figure 12. Figure 12 a shows the response with Sample A as the control used for estimation of parameters and the corresponding

---

Figure 12 and 12a. -- near here

---

response for the test-sample B used for independent prediction of flood peaks. Similarly, Figure 12b illustrates the results of optimization with Sample B as the control. In both cases the rate of improvement of the objective function for the control samples decreased markedly, with little progress achieved after about 30 trials, when a plateau of best

---

Figure 12b. -- near here

---

fit was encountered. Rapid improvement of the objective function during the early state of fitting, followed by an extended period of decreasing improvement, is a characteristic of the optimization procedure. Figure 12a shows that test-sample B is virtually unaffected by, and independent of, parameter adjustments made to improve the goodness-of-fit measured over control Sample A. However, Figure 12b shows that the response of the error criterion for test-sample A is strongly related to that of control-sample B during the early stage of optimization. Eventually, the response diverges, becoming progressively worse after a near-optimum solution has been achieved for the control sample.

The degradation of the error criterion measured over test-sample A (Figure 12b) can be attributed to the influence of episodes of low magnitude which produced highly variable, simulated flood runoff in response to small changes in the parameters associated with antecedent moisture accounting. However, the variable response of these events does not appear to bias the parameters generated from a control sample in which they are included. For example, the results of simulation for test-sample B using parameter values derived for sample A compare favorably with the results based on optimization. Furthermore, the results of simulation for test-sample A are similar to those based on optimization when the influence of those events are discounted. With the exclusion of event 9, for instance, the objective function for test-sample A would be reduced by about 50 percent and would compare favorably with a best-fit results of 0.069, illustrating the fact that an understanding of the distribution of error is important in evaluating the results of optimization.

The simulated response from the split-sample fitting and testing procedure is shown in figure 13. Figure 13a is a scatter diagram of observed versus simulated flood peaks based on optimization to Sample A. Similarly, figure 13b shows the observed versus simulated peaks based on optimization to Sample B. Figure 13c shows the scatter of fit when all 16 events were used in the optimization. The distribution of errors is related both to the approximations and simplifications inherent in the hydrologic model and to errors in storm rainfall, which is known to vary considerably throughout the area.

Figure 13.-- near here

The analysis of objective-function response to change in optimum parameter values offers a means of evaluating the significance of the optimum solution, and illustrates interaction between indivivual parameters and groups of parameters. However, the objective function will be importantly influenced by the nature of the events over which it is computed, and may not reflect the overall significance of model parameters. For example, figure 14 shows the response of the objective function at 5-percent increments

---

Figure 14. -- near here

---

from the optimum value of the parameter RR for both control-samples A and B (RR equal 0.78 and 0.81 respectively). The plots indicate that optimization provided best-fit solutions for both samples in the sense that the objective function would be degraded by either positive or negative incrementations. However, the objective function computed for Sample B is much less sensitive to the parameter RR than is that for Sample A. The sensitivity of RR for control Sample A results from the critical nature of antecedent soil moisture conditions in determining the peak of several of the smaller storms. For comparison, the sensitivity is shown for Sample A with event number 9 deleted. This has little effect for drier conditions (RR small) but brings control Samples A and B into relative agreement for wetter conditions (RR large). Apparently too high a value of RR causes event 9 to be over-estimated, and the optimum value (0.780) is a result of reducing this value sufficiently to estimate event 9 without reducing the accuracy of estimation of other events. It is to be noted that without event 9, a value of RR of 0.819 yields a lower error for the remaining eight events than does the overall optimum value of 0.78.

The final optimization to determine best-fit parameters for the pre-1948 flood events produced an objective function of 0.090. Results of the optimization procedure are shown in Table 8 for several different test runs. With the sample of 16 events the model produces a fit quite similar to that achieved for the smaller control samples. For example, the magnitude of errors in the optimum solution for all storm events was only 8 percent greater than the average of the objective functions for the control samples A and B.

Inspection of objective-function sensitivity for each of the three control samples indicated a consistent hierarchy of parameter influence. The parameters associated with the method of antecedent moisture accounting (RR, EVC and DRN) grossly controlled the objective function. The Philip infiltration parameters (SWF and KSAT) and the routing coefficient (KSW) were intermediate in importance. The range factor (RGF) and field capacity moisture storage (BMSM) had little influence on the objective function for the various control samples, and may be poorly identified.

A sufficient number of events is not the only requirement to obtain a meaningful identification of model parameters. Equally important is the need for a wide range in both antecedent and storm-period conditions. For example, if all flood events included in a control sample were associated with similar antecedent conditions then one or more parameters may exert little influence on the results of simulation and be poorly identified, while others may be "overdetermined." In addition, an interpretation of the hierarchy of parameter sensitivity must be tempered by an understanding of not only the limitations of the model and its lack of equivalence to the physical system, but also by consideration of the characteristics of the criterion used to express the sensitivity. The response of low-magnitude events to small changes in some parameter values prevents a straightforward assessment of model sensitivity and demonstrates the need for the development of alternative measures of sensitivity.

It cannot be overstressed that in the split sample testing for this station eight events were used to determine eight model parameters. This clearly places this study in the area of small sample theory. The relative consistency of results both in accuracy and derived parameter values is therefore quite encouraging. The various results of split sample testing indicate that the root mean square error of prediction is about 30 percent for these data, with, apparently, about one small storm being grossly in error for each test case.

LITTLE BEAVER BASIN

General Physiography

The Little Beaver Creek drainage basin is a 4.61 square miles (12 square kilometers) area of the Gasconade Hills in the Ozark Mountains just west of Rolla, Missouri. The range in elevation is from 790 feet (240 meters) at the U. S. Geological Survey gaging station to 1,180 feet (360 meters). The gently rolling hills are covered with a stony, porous soil. The southerly facing basin receives rainfall which is fairly evenly distributed throughout the year, but with somewhat greater amounts in the summer than in the winter months.

Precipitation

The U. S. Geological Survey maintains a recording rain gage, Rolla 3 W, near the center of the Basin (see Figure 15). This record was used

---

Figure 15. -- near here

---

for simulation throughout the period of record. In addition a rain gage is maintained at the School of Mines about a mile east of the east boundary of the basin. The average annual rainfall during 1948-64 was 36.7 inches (0.93 meters).

Data for 29 flood events covering the period 1948-1964 were reduced to 15 minute intensities. These storms were split into a control sample of 14 events during the period 1948-53 and a test period of 15 events during the period 1954-1964.

Streamflow

The streamflow data used for fitting the model to the Little Beaver Creek Basin were those for the U. S. Geological Survey gaging station near Rolla, Missouri. The site has been gaged since 1948. The mean discharge during the period 1948-64 was 3.77 cfs (0.11m^3/sec) or 11.1 inches (0.28 meters) over the basin. The maximum flow of 7,420 cfs (210 m^3/sec) occurred July 17, 1958. The annual peak discharges during the period of study varied from 524 cfs (15m^3/sec) to 7,420 cfs (210m^3/sec). The peaks selected for study ranged down to as low as 200 cfs (5.8m^3/sec).

Evaporation

Daily values of pan evaporation were obtained from the U. S. Weather Bureau, which maintains a pan evaporation station at Lakeside, Missouri, which is located about 45 miles west of the Little Beaver Basin and at an elevation of 595 feet (181 meters). The average pan evaporation during the period 1948-64 was 53 inches (1.35 meters).

PARAMETER DEFINITION

Three sets of model parameter determinations were made using the control period 1948-53. The results of these fittings plus two sets of starting parameters are shown in Table 9. The first derivation was of

---

Table 9.-- near here

---

set 2 from the starting set 1. The accuracy of fit of 0.065 gives a standard error of fit of about 25 percent. The value of RR of 0.98 seemed high, and was felt to be too much of a curve-fitting parameter. Therefore set 3 was derived by fixing the evaporation pan coefficient (EVC) at its optimum value and the daily rainfall infiltration coefficient (RR) at 0.8. A lower limit for RR should be 0.7, because the mean annual flow is about 30 percent of the mean annual rainfall. Therefore 0.8 to 0.85 is a reasonable value. The accuracy of fit for the parameters for set 3 is 0.075, or about 27 percent.

On the basis of hydrograph plots for the results of set 3, the routing component was recomputed. Both the time area histrogram and the surface routing coefficient (KSW) were revised, and KSW was included in the next optimization run. RR was held fixed at 0.85. The fit of set 5 is 0.055, which yields about a 23 percent accuracy. The test group of 15 floods during the period 1954-64 were then simulated with set 5 parameter values. The accuracy of fit for the test set was 0.073, which yields an estimate of 27 percent for a standard error of prediction.

Table 9.-- Results of fitting of model parameters to data for Little Beaver Creek near Rolla, Missouri, using the Rolla 3 W rain gage.

Parameter	Start 1	Optimum 2	Optimum 3	Start 4	Optimum 5	Test (No.5 parameters)
SWF	.2	.20	. 8	.2	.194	.194
KSAT	.1	.08	.07	.05	.047	.047
KSW	1.0	1.0*	1.0*	.85	.84	.84
EVC	.7	.56	.56*	.55	.52	.52
BMSM	2.0	2.8	2.3	3.0	2.4	2.4
RGF	10.0	9.4	9.3	10.0	11.7	11.7
RR	.8	.98	.8*	.85	.85*	.85
DRN	.1	.28	.28	.5	.47	.47
U1		0.065	0.075	0.061	0.055	0.073
Standard error (o/o)		25	27	25	23	27

* Parameter values held constant for the run indicated.

A separate fitting for the Little Beaver Basin was made to the School of Mines rain gage, which lies outside the basin. The results of the the fitting are shown in Table 10. A comparison of rainfall volumes for the two gages, and of the simulated volumes and peaks is shown in Table 11.

Table 10.-- near here

Table 11.-- near here

Table 10.-- Results of fitting of model parameters to data for Little Beaver Creek near Rolla, Missouri, using the School of Mines rain gage.

Parameter	Start 1	Optimum 2
SWF	0.20	0.11
KSAT	.05	.063
KSW	.85	.97
EVC	.55	.39
BMSM	3.0	2.2
RGF	10.0	8.0
RR	.85	.85*
DRN	.50	.60
U1 (13 events)	.21	.19
	46	44
U1 (9 events)	.121	.099
	35	31

* Parameter value held fixed for the run.

Table 11.--

Date	Measured		Rolla 3 W			School of Mines		
			Measured	Simulated		Measured	Simulated	
	Runoff inches	Peak* cfs	RF inches	Runoff inches	Peak* cfs	Rainfall inches	Runoff inches	Peak* cfs
6-17-48	0.12	376	1.17	0.13	351	0.76	0.21	545
6-2-49**	1.05	1228	2.59	.82	1328	.89	--	--
7-22-49	.33	1199	1.21	.47	1253	.85	.47	1247
10-11-49	2.76	3121	4.26	2.90	2848	6.05[a]	4.64	3589
10-20-49	.55	1142	1.33	.79	1639	.95	.54	1321
1-13-50	.64	1348	1.03	.38	990	1.10	.49	1124
4-10-50	.24	811	.85	.37	1053	.88	.52	1392
5-19-50	1.06	1575	1.85	.94	1446	1.77	.92	1145
5-26-50	.25	742	1.34	.43	1167	.48[a]	.15	406
6-9-50	1.78	4177	3.36	1.73	3683	2.01[a]	1.33	2461
6-22-51	.31	848	1.16	.38	979	1.08	.58	1338
6-30-51	1.64	2079	2.40	1.35	1514	2.67	1.69	2218
4-23-53	.73	2054	1.56	.74	1829	2.66[a]	2.19	5380
5-17-53	.15	416	.56	.10	301	.38	.11	308

* Peak rates are surface runoff rates only; base flow has been subtracted from measured rate.

** Not included in School of Mines optimization because measured storm runoff exceeded measured rainfall.

[a] School of Mines gage storm rainfall apparently grossly in error.

Two conclusions can be drawn from this second fitting. First, the School of Mines gage is not a very representative measure of rainfall on the basin, even though it is just outside the basin. The accuracy of fit is 44 percent. Five measured storm volumes are grossly different from those for Rolla 3 W. One was excluded from the fitting, but the other four influenced the fitting, and probably caused the higher value of KSW and reduced volumes of infiltration. The School of Mines gage does give some indication of the effect of variability of storm volume over the basin, however. For 9 of the 14 storms the Rolla 3 W gage simulation overestimated peaks when its measured storm volume exceeded that at the School of Mines. This held true for 7 of the 9 peaks above 1,000 cfs and 9 of the 12 above 500 cfs. Thus, although the School of Mines gage alone gives much less accurate results than those for Rolla 3 W, the two used together could give a better estimate for flood peak simulation.

Comparison of Derived Parameter Values

The model is based, at least in part, upon a simulation of the physical processes operating upon the basin modeled. The parameter values derived, therefore, should be related to the physical parameters involved. However, the model is a bulk-parameter model. That is, it models all the infiltration in the basin as if it were uniform over the basin. The parameter values derived are in some way optimal average values, and may be, at best, indices to the "true" parameters, or to their distribution over the basin.

If the model is to be used in regional studies, it can serve either of two roles. First, it can be used to extend a record in time. For that use, the most important consideration is the error of prediction. For the three basins for which results are presented in this study, a standard error of prediction of about 30 to 35 percent was achieved. This was found to be largely dependent upon accuracy of rainfall measurement. In particular, the use of a single rain gage to estimate rainfall variability over the basin seems to introduce an error of about 20 to 25 percent into the simulation. A decision must be made as to whether point rainfall data that produce errors of this magnitude add information to the record.

Second, the model can be used in regional studies by relating the derived parameter values to physical characteristics measurable in the basins which are simulated. The derived relations could then be used to estimate parameter values for ungaged sites. The accuracy of prediction in this case would be a function both of the errors in rainfall input and of the errors in predicted values for the model parameters. This accuracy of prediction would be compared to the accuracy of flood frequency methods presently in use.

The derived parameter values for the three basins used in this developmental study are shown in Table 12. All are reasonable values.

Table 12.-- near here

However, there are too few results to draw any general conclusions at this time. Each parameter will be discussed as to its relation among stations and the reasons for variability. RR is a measure of percentage of infiltration for daily rainfall amounts for periods not simulated in detail, either because rainfall amounts are too small or records are not accurate enough to use for detailed simulation. Also shown in Table 12 are values

one minus the ratio of measured runoff to measured rainfall for each basin during the study period. This sets a lower limit on RR, and in each case the fitted value exceeds this lower limit. Actually, the lower limit should be somewhat higher, because all base flow should be subtracted from the runoff to derive the limiting value. Bee Tree Creek Basin has the highest base flow, and thus the fitted value exceeding the limiting value by a relatively large amount is consistent.

Table 12.-- Summary of results of optimization for the three study basins.

	Santa Anita	Bee Tree	Little Beaver
SWF	20	3.6	4.1
KSAT	.32	.1	.05
KSW	2.7	5.7	.84
EVC	.73	.58	.52
BMSM	3.5	1.9	2.4
RGF	6	14	12
RR	1.0	.75	.85
DRN	.058	.0058	.022
L (miles)	4.7	3.2	3.25
S (ft/ft)	.12	.00929	.0124
$L/\sqrt{S}$	13.5	33.5	29.3
1-RO/RF	.74	.46	.7

KSAT, SWF, and RGF determine the infiltration equation during detailed strom simulation and therefore should be discussed together. SWF determines the soil suction characteristics for wet conditions, SWF times RGF determines it for dry conditions, and KSAT represents the soil saturated permeability, or minimum infiltration rate. The range of soil suction is from 4 to 50 inches (10 to 125 centimeters) for both Bee Tree Creek and Little Beaver Creek Basins, and is from 20 to 120 inches (50 to 300 centimeters) for Santa Anita Creek Basin. Comparable experimental ranges for a sandy loam are about 30 to 130 centimeters for Yolo sandy loam and 30 to 200 centimeters for Yolo silt loam (Coleman, 1944). Seemingly, the minimum infiltration rates are anomalous for the measured basins, in that 0.3 inches per hour seems to correspond to a sandy loam rate, whereas 0.05 to 0.10 seems to correspond to a rate for a silt loam (Musgrave, 1955). Some attempts should be made to relate the fitted values to ring infiltrameter or other data collected for study basins.

BMSM represents an effective maximum soil moisture retention, and the low values indicate shallow soils. Bee Tree Creek Basin appears to have the thinnest effective soil mantle and Santa Anita the least shallow. This agrees qualitatively with descriptions of the geology and soils. DRN represents the drainage rate from the saturated layer to the unsaturated layer. This parameter is critical for determining antecedent conditions for some storms, but has no effect on most storms. Therefore it is probably poorly defined for all cases. The derived values are considerably less than KSAT in each case, which is as expected, but nothing can be said as to the reasonableness of the values otherwise.

EVC should represent an effective average pan coefficient for the basin. However this meaning is compounded by the fact that for each basin a correction also must be made to adjust the pan evaporation to average basin conditions. For Little Beaver Creek Basin, the nearest pan evaporation record was 45 miles away, for Santa Anita Creek Basin it was 10 miles away. For Bee Tree Creek Basin the evaporation record was nearby, but at a lower elevation. All records are for U. S. Weather Bureau Class A pans, for which the pan coefficient should range from 0.6 to 0.8. EVC should be somewhat lower than these values, if an altitude correction is involved. There should be little or no altitude correction for Santa Anita Creek Basin, because the pan is at an elevation well above the lowest point in the basin. Both of the other records are for sites at elevations below the lowest point in the basin, and for Little Beaver Creek considerably lower. Therefore, the derived values seem to be of the right order of magnitude.

Both KSW, which is the hydrograph recession rate, and the translation hydrograph ordinates do not enter directly into the fitting process. They are derived from the measured hydrograph shapes. The Little Beaver Basin has an unusually rapid recession. Values of $L/\sqrt{S}$ are shown, where L is length of the main channel in miles and S is the slope of the basin in feet per foot for the reach from 10 percent to 85 percent of the distance from the discharge gaging station to the point on the ridge which represents the extension of the main channel (Benson, 1962). Although $L/\sqrt{S}$ values for Bee Tree and Little Beaver Creek Basins are quite similar, the values of KSW differ by a ratio of seven. Santa Anita Creek Basin is consistent with Bee Tree Creek Basin in this regard, in that both $L/\sqrt{S}$ and KSW are about half the values for Bee Tree. The reason for the anomalous value for Little Beaver is unknown, but it may be related to the drainage pattern. Both Santa Anita and Bee Tree Creek Basins are dendritic, whereas Little Beaver Creek Basin seems to be more palmate.

Sources of Error and Their Impact

The accuracy of fit for the three basins studied was similar. An accuracy of about a 30 percent standard error is obtainable. The detailed study for Santa Anita Creek Basin indicated that about a 20 percent standard error was attributable to rainfall sampling alone. If the rainfall errors are independent of other modeling errors, then

$$RE^2 + ME^2 = TE^2$$

where RE is the modeling error resulting from rainfall input error, ME is other modeling error, and TE is the total error of simulation. For the case above,

$$20^2 + ME^2 = 30^2 \; ; \; ME^2 = 500$$

According to Eagleson (1967), if one rain gage gives an error of 20 percent, then two properly placed rain gages would give an error of about 15 percent. The use of the information from two gages with the present model structure should thus result in an error of

$$TE^2 = 15^2 + 500 = 725$$

or a standard error of 27 percent rather than 30 percent.

The improvement of the structure of themodel can also lead to improvement of prediction. If the model error were cut in half, the resulting standard error would then be

$$TE^2 = 20^2 + 250 = 650$$

or a standard error of 25 percent. Thus for any major improvement in the accuracy of simulation, there must be simultaneous improvement in both the model and in the accuracy of rainfall input. Model improvements alone will increase the accuracy of prediction, but there will be a limiting accuracy which must be accepted if the constraint of a single rain gage is to be maintained.

The marginal gains in accuracy which should be expected from model improvement influence the strategy for judging model improvements. Changes should be accepted as improvements if they add to the simplicity of the model, if they aid in the regionalization of the parameter values, or if they gain accuracy. The search will continue for a better model, but an imperfect model must be accepted.

Conclusions

The development of the model demonstrates the feasibility of rainfall-runoff simulation. Such simulation is not new, so that such a demonstration of feasibility is not unexpected. However, the constraints placed upon the model developed were that a single rain gage be used for simulation on a basin. This led to the development of a bulk parameter model. Thus, model parameter values are indices of average conditions on the basin and only approximate real parameter values. Both errors of rainfall input and lack of model equivalence to the physical prototype limit the predictive ability of simulation. These two sources of error are of similar order of magnitude for the basins studied. Therefore major gains in accuracy will depend upon improvement in both. The limit of accuracy of prediction of flood peaks by simulation with a single rain gage seems to be on the order of about 25 percent, and this level of accuracy should be accepted as resulting from the imposed constraint.

References Cited

Benson, M. A., 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geol. Survey Water-Supply Paper 1580-B, 64 p.

———1964, Factors influencing the occurrence of floods in the Southwest: U.S. Geol. Survey Water-Supply Paper 1580-D, 72 p.

Benson, M. A., and Matalas, N. C., 1967, Synthetic hydrology based on regional statistical parameters: Water Resources Research, v. 3, no. 4, p. 931-945.

Clark, C. O., 1945, Storage and the unit hydrograph: Am. Soc. Civil Engineers Trans., v. 110, p. 1419-1488.

Coleman, E. A., and Bodman, G. B., 1944, Moisture and energy conditions during downward entry of water into moist and layered soils: Soil Sci. Soc. America Proc., v. 9, p. 3-11.

Crawford, N. H., and Linsley, R. K., 1966, Digital simulation in hydrology: Stanford Watershed Model IV, Tech. Rept. No. 39, Civil Eng. Dept., Stanford Univ., Stanford, Calif., 210 p.

Dawdy, David R., and O'Donnell, Terence, 1965, Mathematical models of Catchment behavior: Am. Soc. Civil Engineers Proc., v. 91, no. HY4, Paper 4410, p. 123-137.

Dooge, J. C. I., 1959, A general theory of the unit hydrograph: Jour. Geophys. Research, v. 14, no. 2, p. 241-256.

Eagleson, P. S., 1967, Optimum density of rainfall networks: Water Resources Research, v. 3, no. 4, p. 1021-1033.

Fenneman, N. M., 1938, Physiography of Eastern United States: New York, McGraw-Hill Book Co., Inc.

Folse, J. A., 1929, A new method of estimating streamflow: Carnegie Institution, Washington, D. C.

Goldton, E. F., and others, 1952, Soil survey of Buncombe County, North Carolina: U.S. Dept. Agr., S. 1942, no. 6, 122 p.

Green, W. H., and Ampt, G. A., 1911, Studies on soil physics: I Flow of air and water through soils: Jour. Agr. Research, v. 4, p. 1-24.

Horton, R. E., 1939, Approach toward a physical interpretation of infiltration capacity: Soil Sci. Soc. Am. Proc., v. 5, p. 399-417.

Kraijenhoff van de Leur, D. A., 1958, A study of non-steady groundwater flow with special reference to a reservoir coefficient: De Ingenieur, v. 70, no. 19, p. 87-94.

Matalas, N. C., and Gilroy, E. J., 1968, Some comments on regionalization in hydrologic studies: Water Resources Research, v. 4, no. 6, p. 1361-1369.

Musgrave, G. W., 1955, How much of the rain enters the soil? *in* Water U.S. Dept. Agr. Handbook, p. 151-159.

Philip, J. R., 1954, An infiltration equation with physical significance: Soil Sci. Soc. Am. Proc., v. 77, p. 153-157.

Rosenbrock, H. H., 1960, An automatic method of finding the greatest or least value of a function: Computer Jour., v. 3, p. 175-184.

Sherman, L. K., 1932, Streamflow from rainfall by the unit hydrograph method: Eng. News Record, v. 108, p. 501-505.

Tennessee Valley Authority, 1961, Evaporation in Tennessee River Basin: Division of Eater Control Planning, Hyd. Data Branch, rept. 0-243-310-A, 49 p.

Theis, C. V., 1935, The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground water storage: Am. Geophys. Union Trans., pt. 2, p. 519-524.

Wilde, Douglass J., 1964, Optimum seeking methods: Englewood Cliffs, Prentice-Hall, Inc.

ANTECEDENT MOISTURE ACCOUNTING COMPONENT	INFILTRATION COMPONENT	ROUTING COMPONENT
saturated-unsaturated soil moisture regimes	Philip infiltration equation	Clark instantaneous unit hydrograph
	$\frac{di}{dt} = K\left(1 + \frac{P \cdot (\bar{m} - m_o)}{i}\right)$	

PAR	VAR	PAR	VAR	PAR	VAR
EVC	BMS	PSP	BMS	KSW	SW
RR	SMS	KSAT	SMS	(time-area curve)	
BMSM		RGF			
DRN					

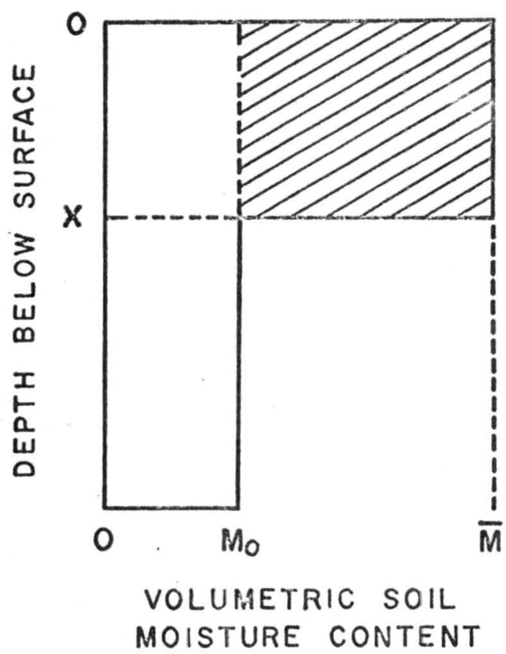
INPUT DATA

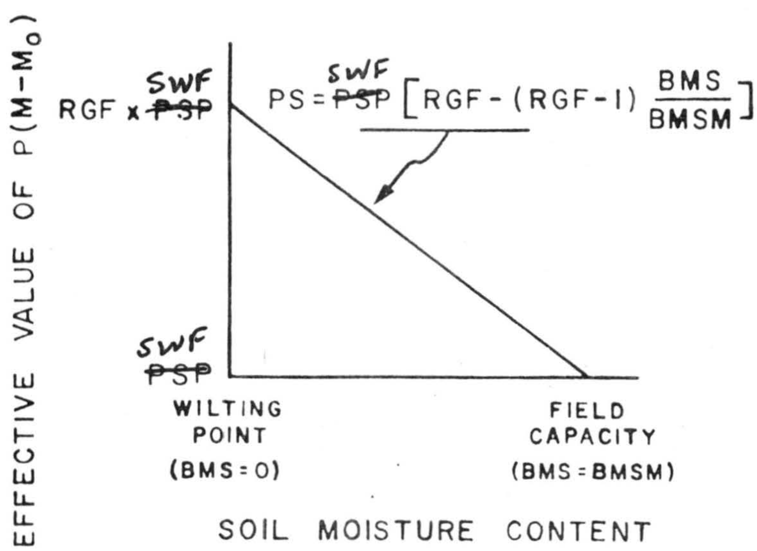
DAILY RAINFALL	UNIT RAINFALL	RAINFALL
DAILY PAN EVAP.	"BMS"	EXCESS
INITIAL CONDITION	"SMS"	

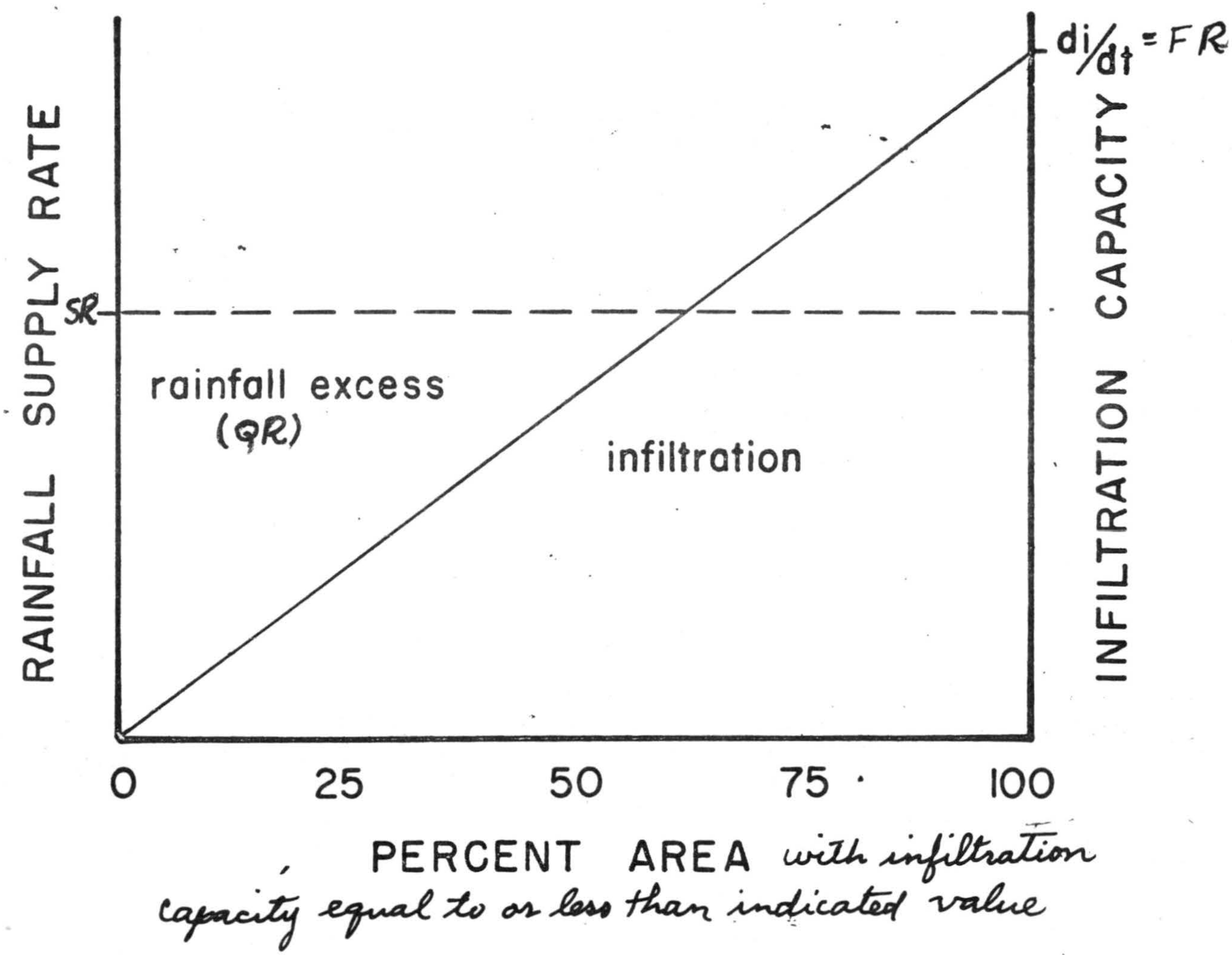
OUTPUT

BMS	RAINFALL	DISCHARGE
SMS	EXCESS	

1

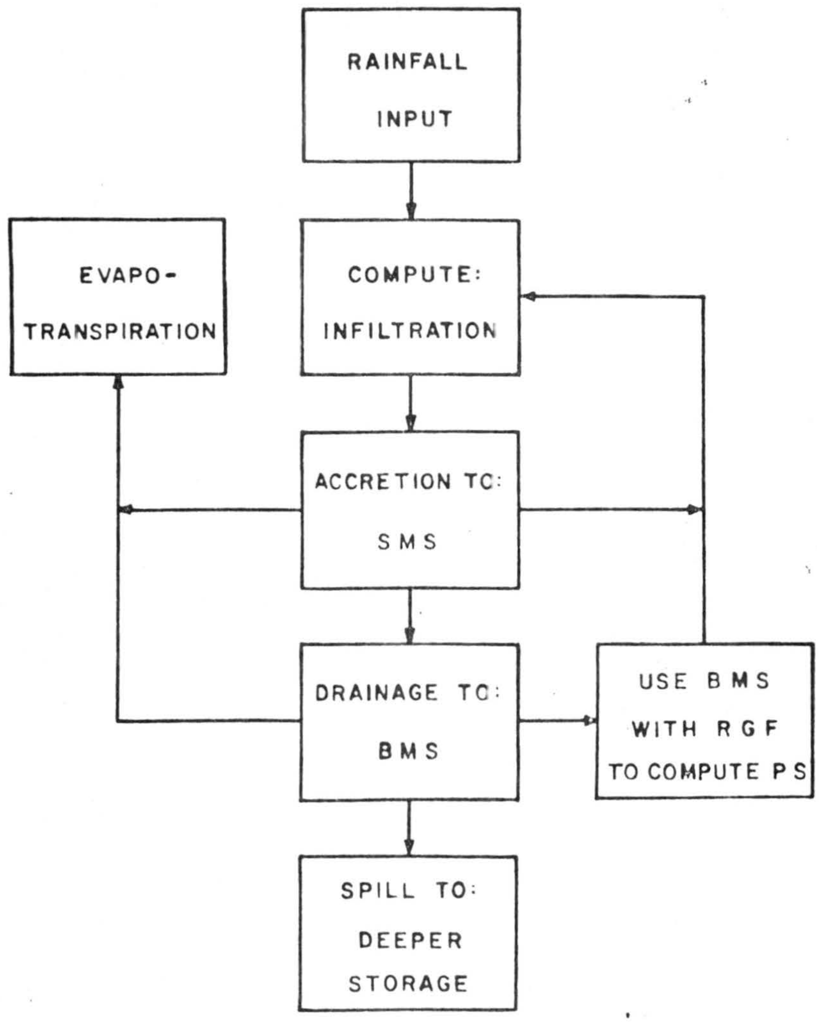

DEPTH BELOW SURFACE
O
X
O
Mo
M̄
VOLUMETRIC SOIL
MOISTURE CONTENT


EFFECTIVE VALUE OF P(M−M$_o$)
RGF x ~~PSP~~ SWF
PS = ~~PSP~~ SWF [RGF − (RGF − 1) BMS/BMSM]
~~PSP~~ SWF
WILTING POINT (BMS = 0)
FIELD CAPACITY (BMS = BMSM)
SOIL MOISTURE CONTENT


RAINFALL SUPPLY RATE
INFILTRATION CAPACITY
di/dt = FR
SR
rainfall excess
(QR)
infiltration
0
25
50
75
100
PERCENT AREA with infiltration capacity equal to or less than indicated value

91

4


RAINFALL INPUT
EVAPO-TRANSPIRATION
COMPUTE: INFILTRATION
ACCRETION TO: SMS
DRAINAGE TO: BMS
USE BMS WITH RGF TO COMPUTE PS
SPILL TO: DEEPER STORAGE

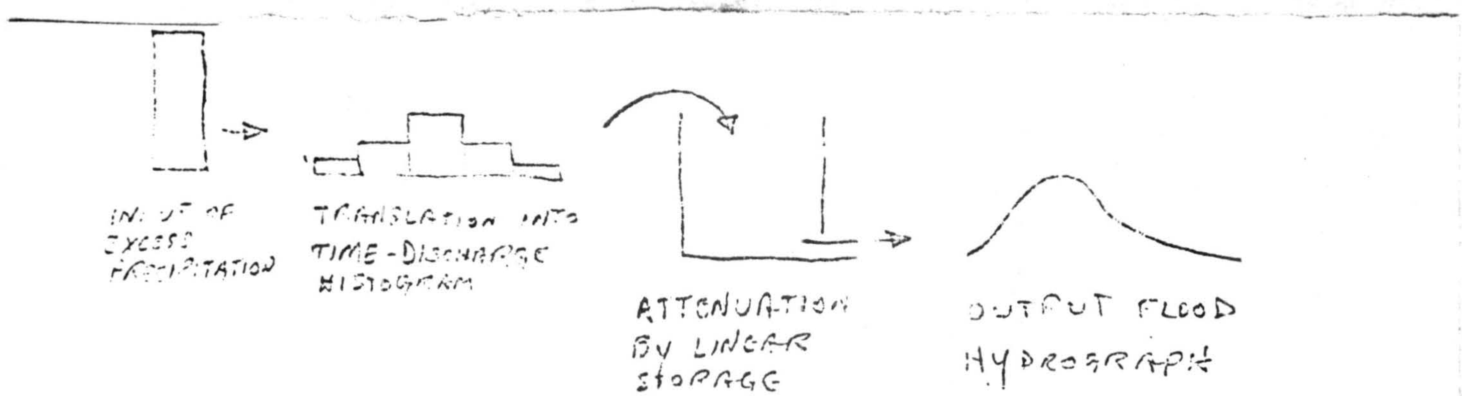
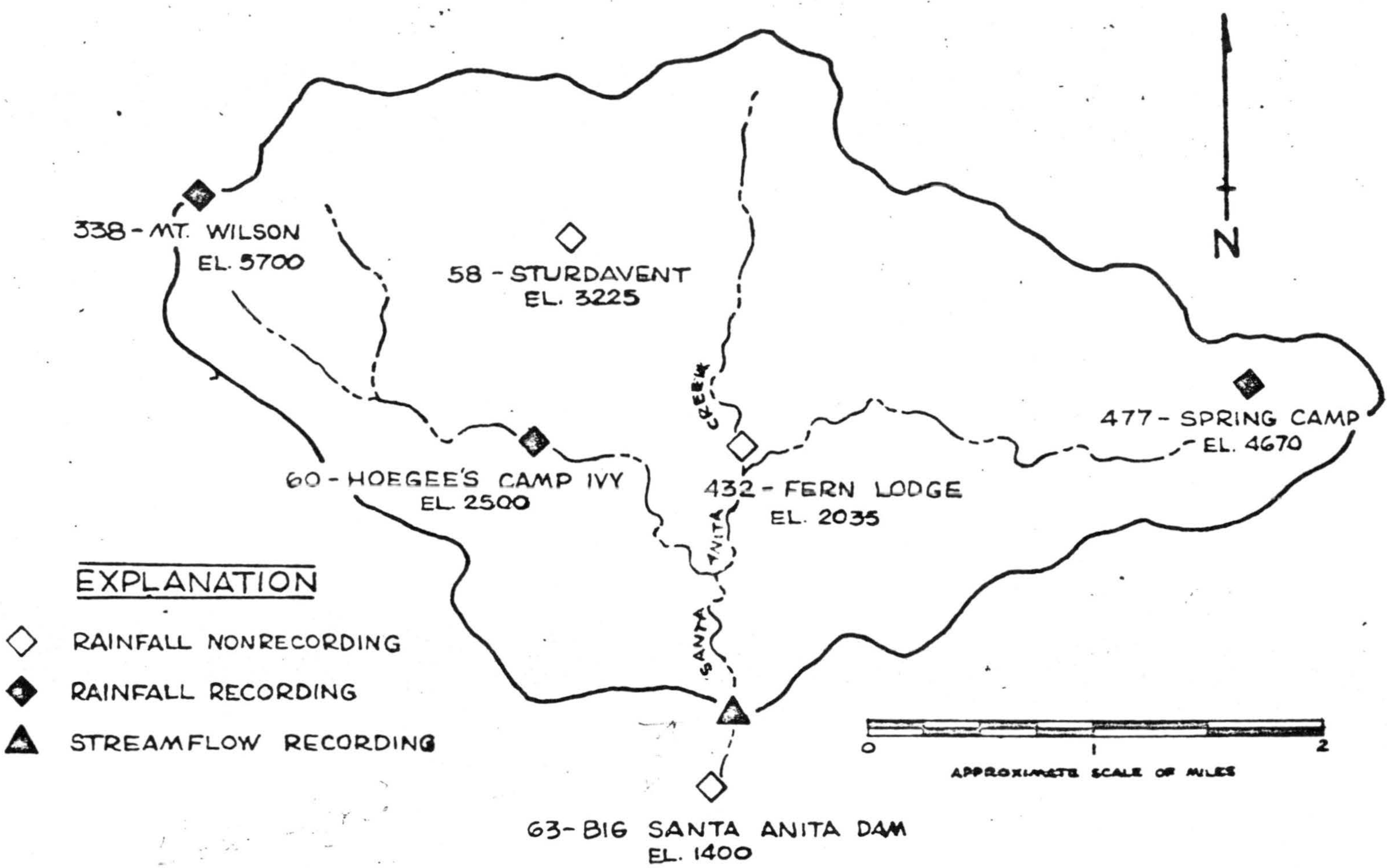
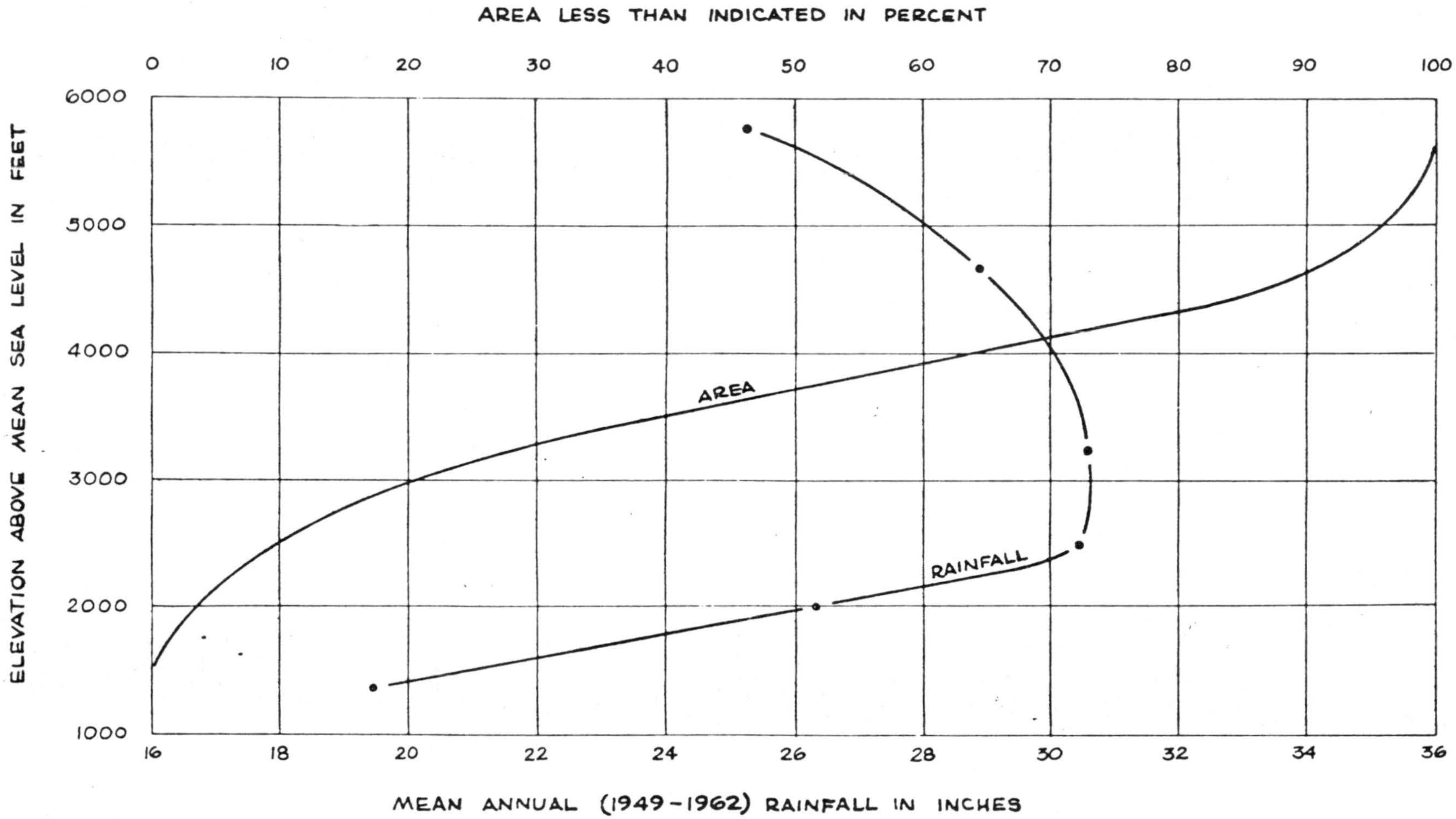


Figure 6. Schematic representation of the surface runoff routing component.

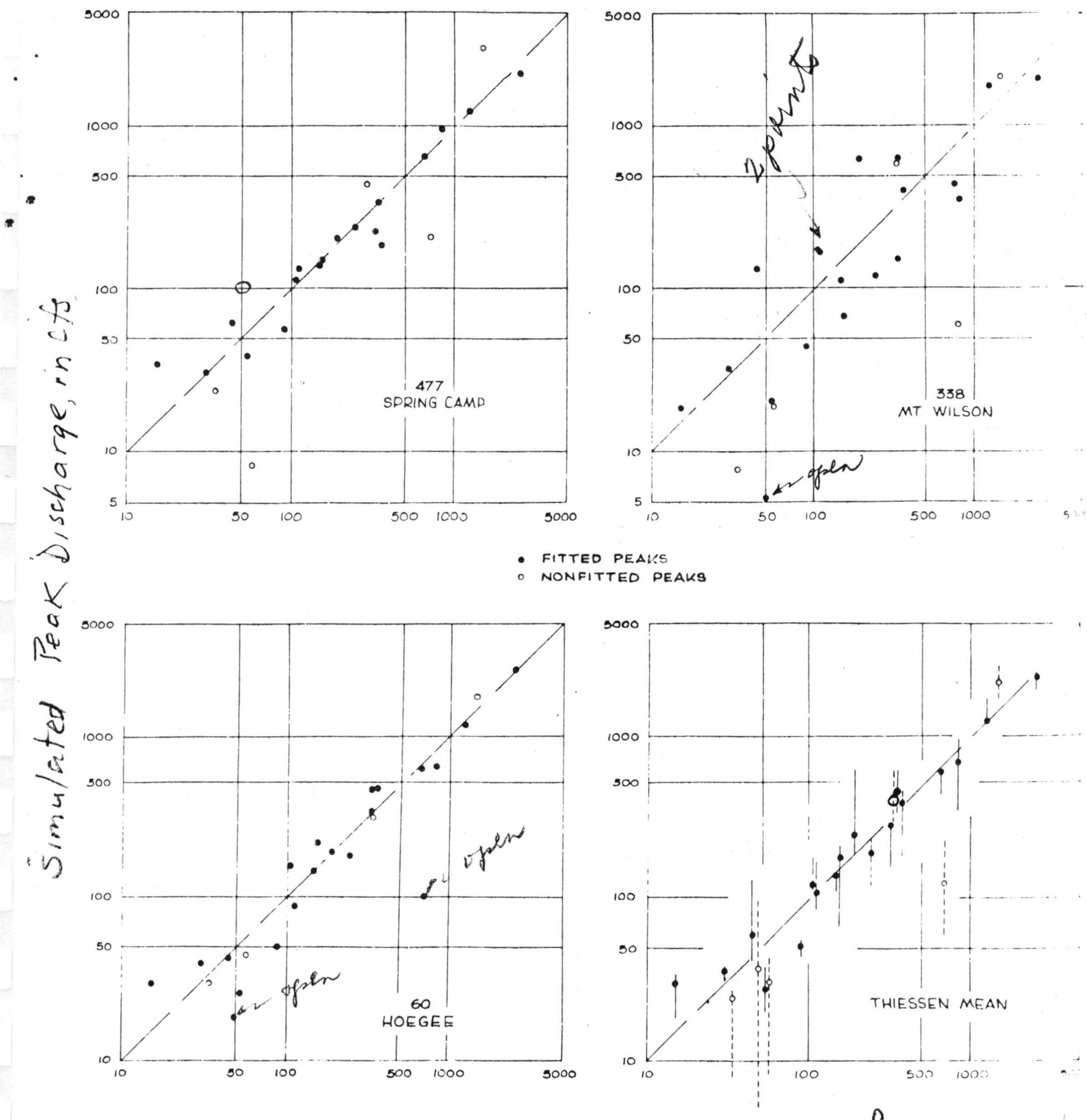


SANTA ANITA CREEK

94



ELEVATION-AREA AND ELEVATION-RAINFALL RELATIONSHIPS
FOR SANTA ANITA CREEK

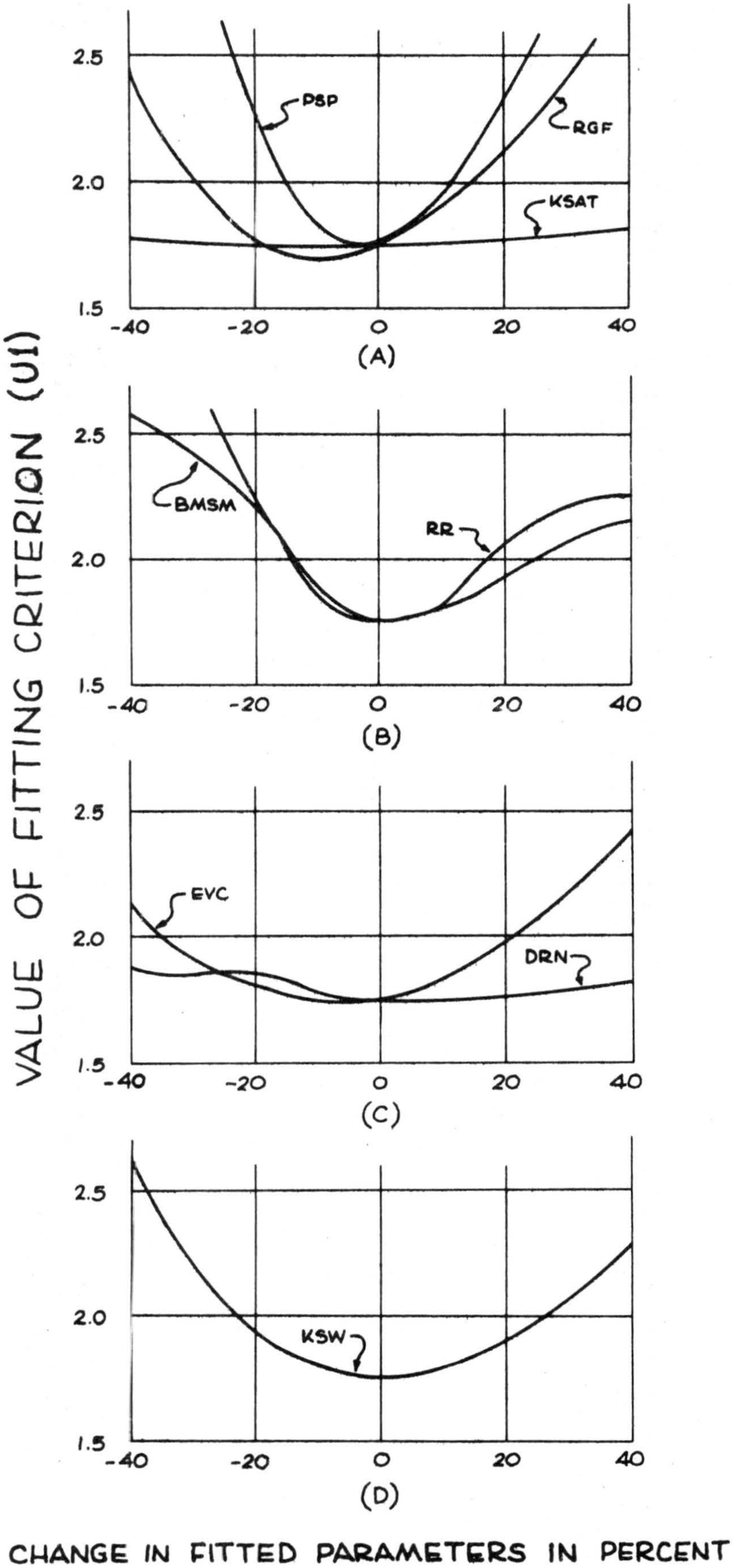


OBSERVED PEAK DISCHARGE, in cfs

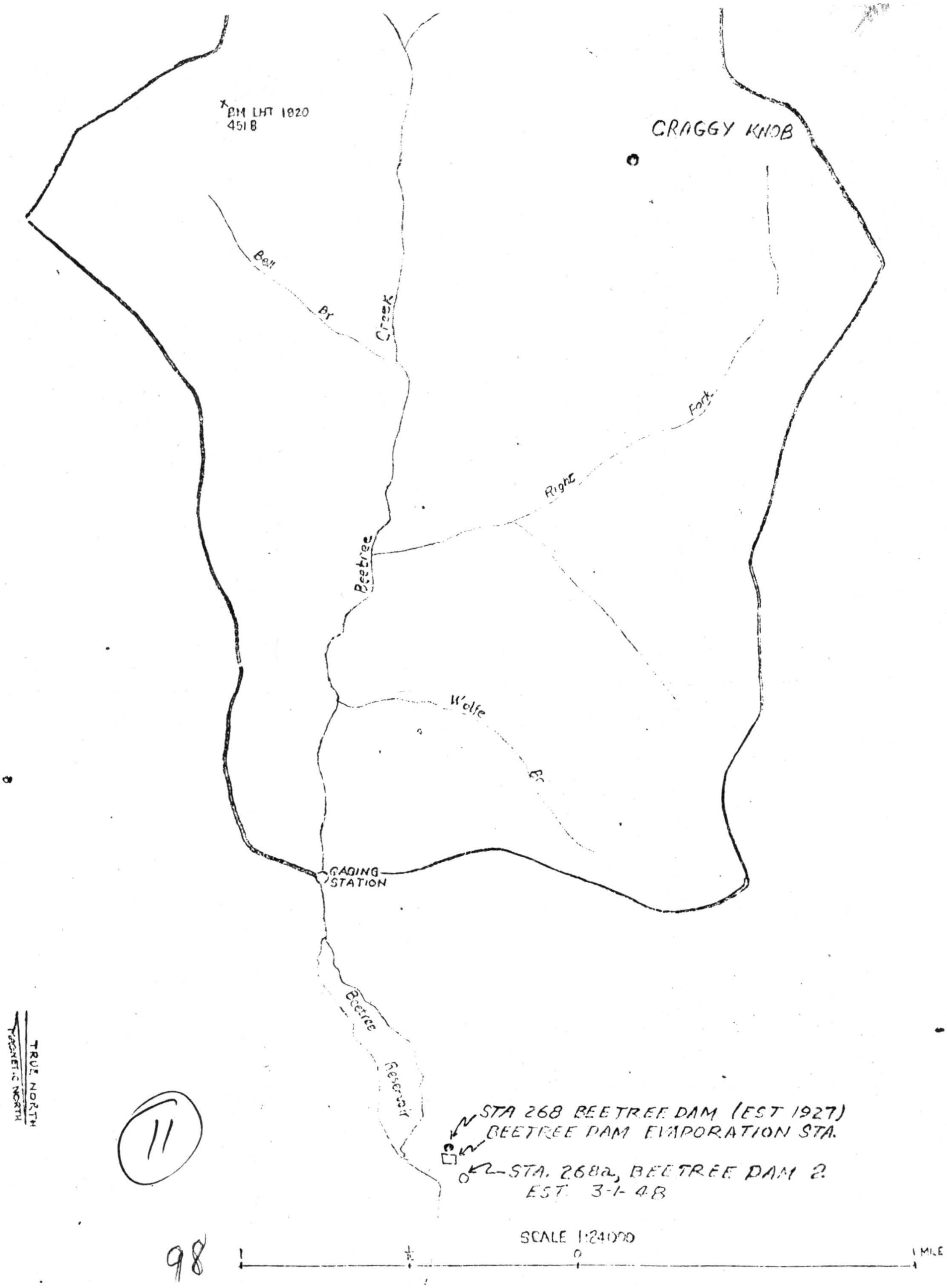
~~PEAK DISCHARGE SIMULATION (A)~~

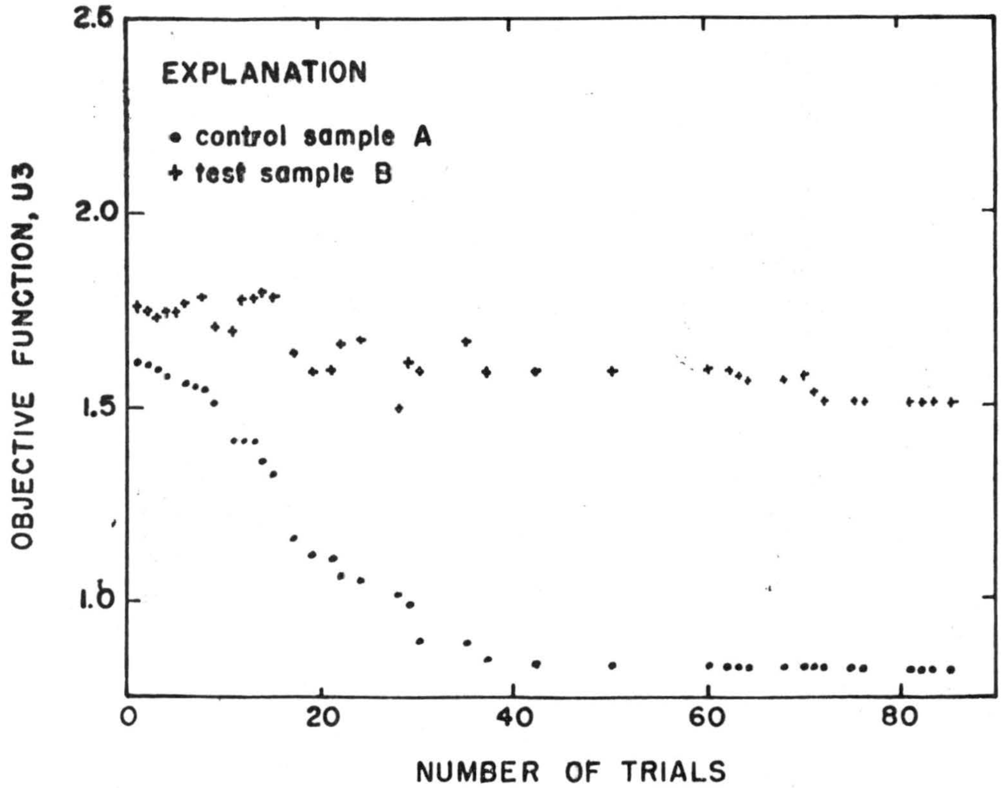
96

~~7~~

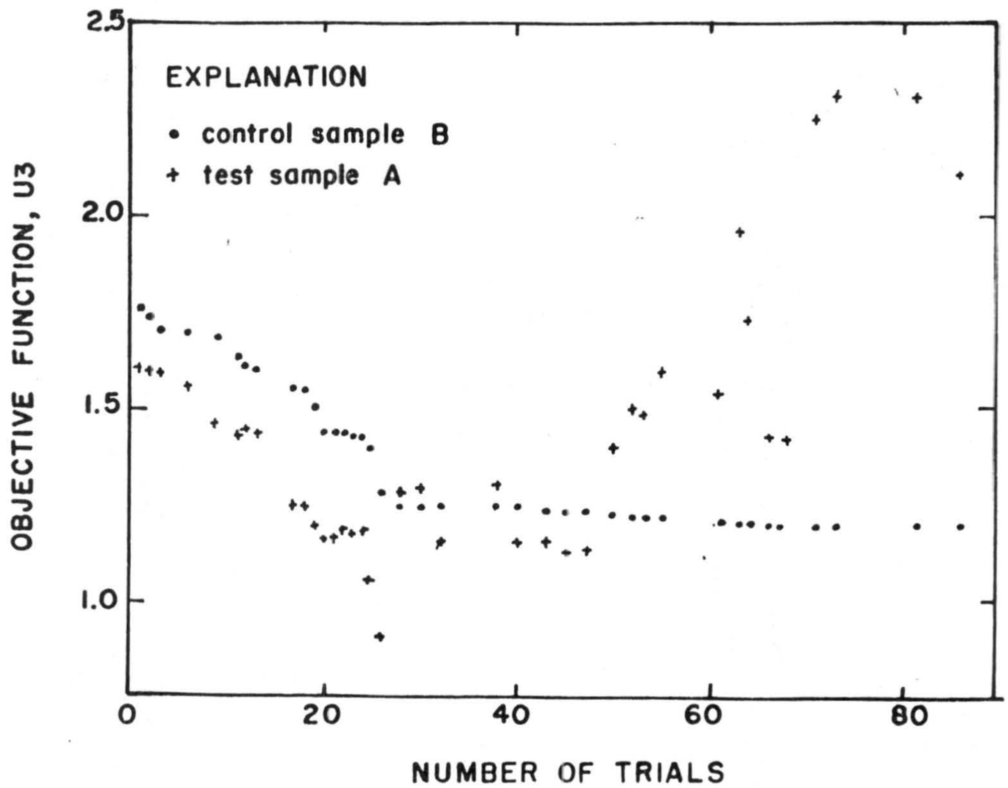


~~PARAMETER SENSITIVITY - INPUT SO(2)~~


BM LHT 1820
4518
CRAGGY KNOB
Bell Br
Creek
Right Fork
Beetree
Wolfe Br
GAGING STATION
Beetree Reservoir
TRUE NORTH
MAGNETIC NORTH
11
STA 268 BEETREE DAM (EST 1927)
BEETREE DAM EVAPORATION STA.
STA. 268a, BEETREE DAM 2
EST 3-1-48
SCALE 1:24000
1 ½ 0 1 MILE

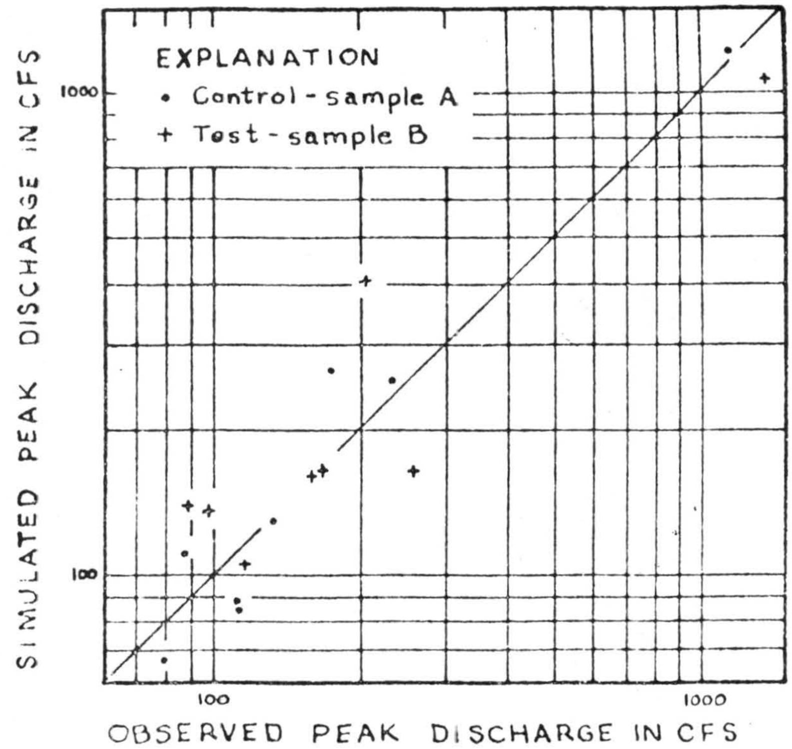


Response of objective function during optimization to control sample A



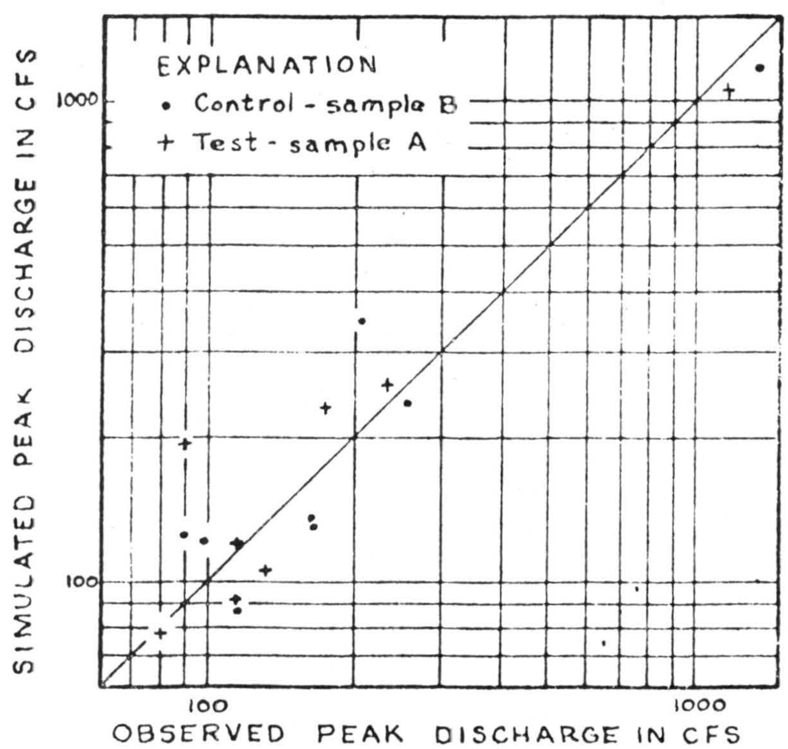
Response of objective function during optimization to control sample B

Figure 12.



13a

Fig. 3--Simulation results based on optimization to control-sample A



13b

Fig. 4--Simulation results based on optimization to control-sample B

13a - 13b 100

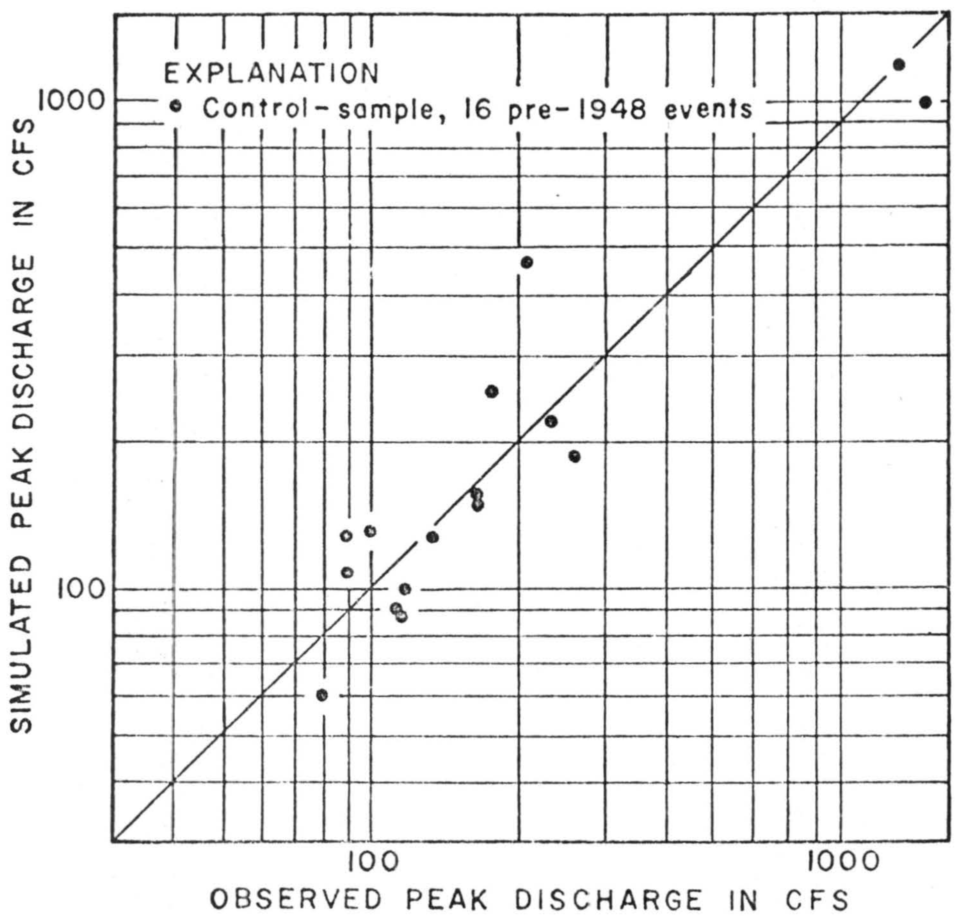


FIGURE 13C. SIMULATION RESULTS BASED ON OPTIMIZATION TO ALL PRE-1948 FLOOD EVENTS.

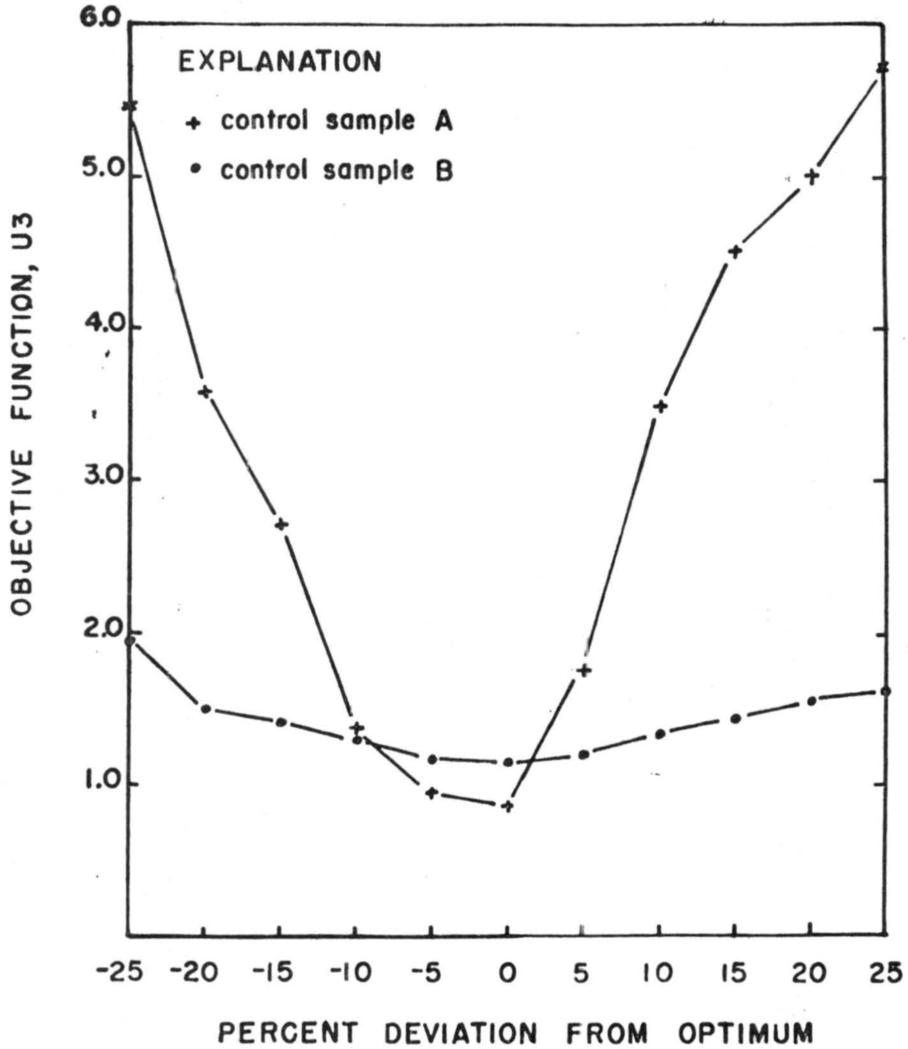


Figure 14.--Sensitivity of objective function to changes from optimum value of parameter RR

FIGURE 15. -- LITTLE BEAVER CREEK BASIN NEAR ROLLA, MISSOURI

102