APPLICATION OF A DISTRIBUTED-ROUTING RAINFALL-RUNOFF MODEL TO FLOOD-FREQUENCY ESTIMATION IN SOMERSET COUNTY, NEW JERSEY

By James L. Fulton

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 89-4210

Prepared in cooperation with the

SOMERSET COUNTY BOARD OF CHOSEN FREEHOLDERS

West Trenton, New Jersey
1990

U.S. DEPARTMENT OF THE INTERIOR

MANUEL LUJAN, JR., Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

District Chief
U.S. Geological Survey
Mountain View Office Park
810 Bear Tavern Road
Suite 206
West Trenton, NJ 08628

Copies of this report may be
purchased from:

U.S. Geological Survey
Books and Open-File Reports Section
Box 25425
Denver, CO 80225

CONTENTS

CONTENTS--Continued

ILLUSTRATIONS

ILLUSTRATIONS--Continued

TABLES

TABLES--Continued

CONVERSION FACTORS

Multiply inch-pound unit	by	To obtain metric unit
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
square foot (ft^2)	0.09294	square meter (m^2)
square mile (mi^2)	2.590	square kilometer (km^2)
inch per hour (in/hr)	25.4	millimeters per hour (mm/hr)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
cubic foot per second per square mile [$(ft^3/s)/mi^2$]	0.01093	cubic meter per second per square kilometer [$(m^3/s)/km^2$]
million gallons (Mgal)	3,785	cubic meter (m^3)

Sea level: In this report "sea level" refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)--a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

APPLICATION OF A DISTRIBUTED-ROUTING RAINFALL-RUNOFF MODEL TO FLOOD-FREQUENCY ESTIMATION IN SOMERSET COUNTY, NEW JERSEY

By James L. Fulton

ABSTRACT

The U.S. Geological Survey Distributed-Routing Rainfall-Runoff Model was calibrated for eight basins in Somerset County, New Jersey. The drainage areas of the modeled basins range from 1.20 to 26.2 square miles. Impervious area in the basins ranges from 3.2 to 13 percent of the total basin area. Model verification efficiencies were 75 to 93 percent for storm volumes and 65 to 93 percent for peak flows.

An evaporation model was used to provide synthetic evaporation estimates for model calibration and long-term simulation of flood peaks. During model calibration, the use of synthetic evaporation data yielded a closer fit between observed and simulated storm volumes and peak flows than the use of actual pan-evaporation data.

Long-term 5-minute precipitation data for 448 storms that were collected during 1914-79 at Trenton, New Jersey, were compiled. The storms were selected on the basis of storm volume and intensity, and antecedent soil-moisture conditions. Daily precipitation data collected at Trenton, New Jersey, also were used. The long-term 5-minute and daily precipitation data were used with the synthetic evaporation data to synthesize a long-term flood record at each of the eight study basins.

The recording increment of the storm precipitation data used for model calibration (0.1 inch) was different from the recording increment of the storm precipitation data used for long-term simulation (0.05 inch). Comparison of flood discharges synthesized using 0.1-inch increment precipitation data and using 0.05-increment precipitation data indicated that the recording increment of the data used for long-term simulation had a statistically significant (95 percent confidence level) effect on the synthesized flows. The long-term storm precipitation data were adjusted to account for differences between the recording increment of the long-term precipitation data and the recording increment of the precipitation data used to calibrate the model.

Statistically significant (95 percent confidence level) differences were found between precipitation data collected at Trenton and precipitation data collected in and near Somerset County. Annual total precipitation was 15.3 percent higher in the Somerset County area than in Trenton, on average. Annual maximum storm volumes were 13.5 percent higher in Somerset County than in Trenton, on average. In order to avoid bias in simulated flood data due to differences in precipitation data between Trenton and Somerset County, the long-term Trenton precipitation data were adjusted by increasing daily-data values by 15.5 percent and by increasing storm-data values by 13.5 percent.

The 2- to 500-year-recurrence-interval synthetic flood flows were developed for each basin by fitting a log-Pearson Type III distribution to the long-term synthetic peak-flow series. Long-term flood records were available for three of the basins studied. Flood-frequency estimates for these basins were developed using a weighted combination of the synthetic flood-frequency estimates and flood-frequency estimates developed by fitting a log-Pearson Type III distribution to the long-term flood records. At the remaining five basins, annual-peak data collected during model calibration were added to the long-term synthetic flood series, and flood-frequency estimates were developed by fitting a log-Pearson Type III distribution to the combined data.

INTRODUCTION

In recent years, flooding of small streams has become a problem in Somerset County, New Jersey. The flooding problem is becoming more critical with continuing urbanization. The flood of August 28, 1971 (tropical storm Doria) caused considerable damage and prompted County officials to install flood alarms on two small streams.

Two needs were recognized by the County. First, additional data were needed for use in flood monitoring, and second, precipitation and flow data collected on a real-time basis were needed to anticipate floods and to issue timely warnings.

Accurate techniques for estimating flood-flow frequencies at ungaged basins also were needed. Although Stankowski (1974) developed equations that are based on regression analysis for estimating peak-flow frequencies in New Jersey, one of the implicit assumptions of these equations is that the State of New Jersey is geologically and hydrologically homogeneous. In fact, the State of New Jersey includes parts of four physiographic provinces--the Coastal Plain, the Piedmont, the New England, and the Valley and Ridge (fig. 1). The Coastal Plain is characterized by high-permeability soils and low runoff, whereas the Piedmont province has low permeability soils and relatively high runoff. The soils in the New England and the Valley and Ridge provinces have moderate permeabilities. Because Stankowski's (1974) equations were developed for the entire State, they provide reasonable flood-frequency estimates for stations in the New England and the Valley and Ridge provinces, which have soils with moderate permeabilities. Stankowski's equations, however, tend to overestimate flood frequencies in the Coastal Plain and underestimate flood frequencies in the Piedmont province. Because Somerset County lies predominantly in the Piedmont province (fig. 1), the flood frequencies in the County tend to be underestimated by Stankowski's (1974) equations.

The equations developed by Stankowski (1974) were derived using data mostly from moderate- and large-sized (5 to 800 mi^2 (square miles)) basins. Of the 51 stations that were available for developing the 100-year flood-peak discharges, only four were at basins smaller than 5 mi^2. A set of equations was needed that would be representative of small (<5 mi^2) basins and that would provide improved flood-frequency estimates for Somerset County.

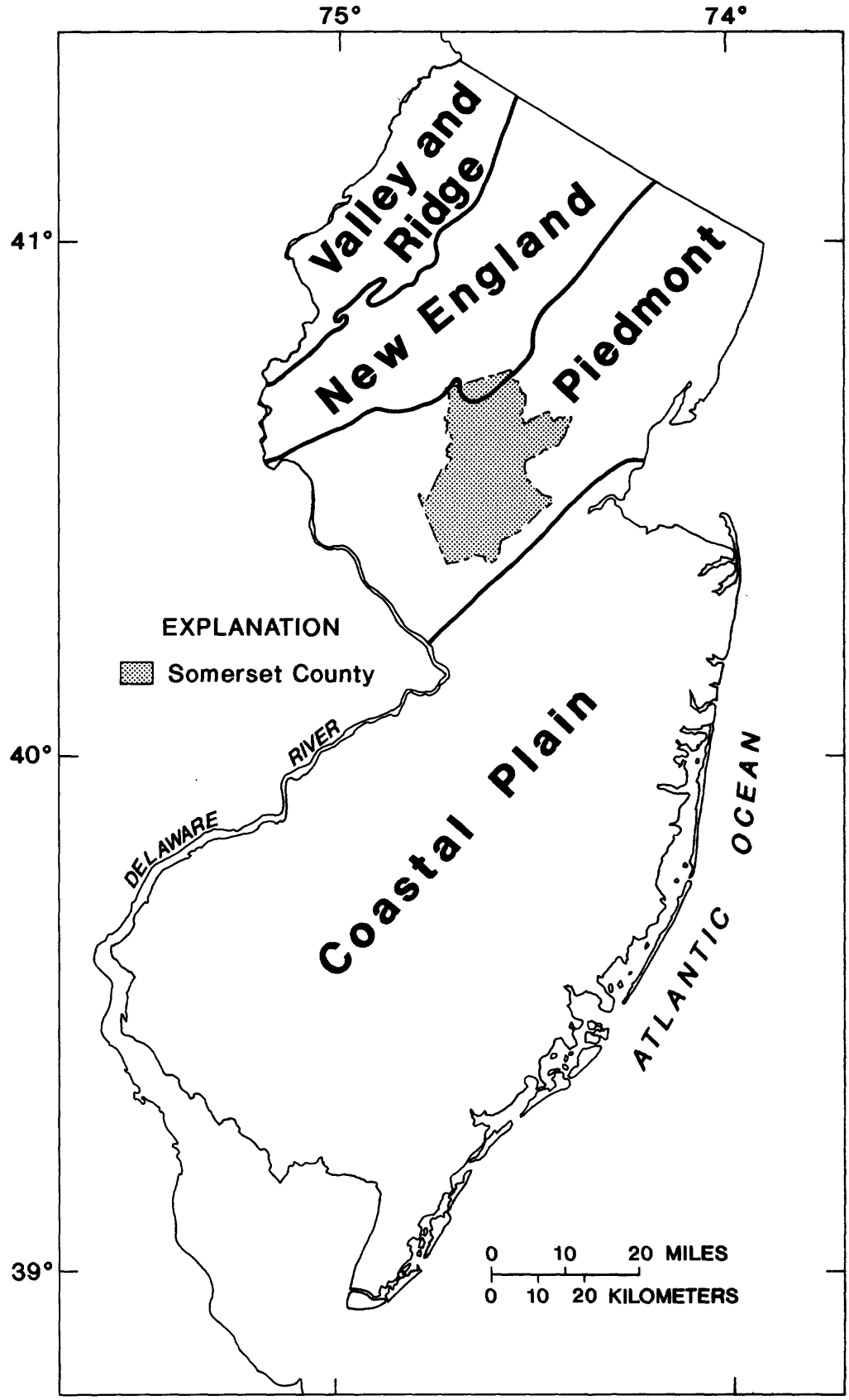


Figure 1.--Physiographic provinces in New Jersey and location of Somerset County. (Modified from Fennemann, 1946)

The U.S. Geological Survey, in cooperation with the Board of Chosen Freeholders of Somerset County, New Jersey, began hydrologic investigations throughout the County to meet these needs. In 1978, the present study was initiated with the following goals:

1. Establish a network of precipitation and streamflow gages in eight study basins. Equip the network with telemetry to provide real-time data to the County for flood-monitoring activities and to provide data for calibrating a rainfall-runoff model.

2. Develop equations for improved estimates of flood-frequency for Somerset County:
   a. Select a rainfall-runoff model for the study and calibrate it for each of the eight study basins, using the precipitation and streamflow data collected.
   b. Extend the peak-flow data base using synthetic flood data generated by the calibrated rainfall-runoff model from long-term precipitation data.
   c. Estimate peak-flow frequencies from the extended flood data at each of the basins studied, by use of the log-Pearson Type III distribution and the guidelines recommended by the Interagency Advisory Committee on Water Data (1982).
   d. Develop equations to estimate peak-flow frequencies for ungaged sites in Somerset County by using regression analysis of estimated peak-flow frequencies for the study basins and flood-frequency estimates from other long-term stream gages in the region.

The gage network was documented by Campbell (1987) and the development of equations for estimating flood-peaks is documented by R.D. Schopp (U.S. Geological Survey, written commun., 1988).

Purpose and Scope

This report documents (1) the selection and calibration of a distributed-routing rainfall-runoff model on the basis of 5 years of data collected during 1980-84, (2) the simulation of the long-term flood records for 1914-79, (3) the extension of the observed records, and (4) the estimation of peak-flow frequencies at each of the eight study basins in Somerset County, New Jersey.

Acknowledgments

The author is grateful to Mr. Thomas Harris and Mr. Michael Amarosa of the Office of the Somerset County Engineer for providing topographic maps of the study basins and for their advice and consultation throughout the project.

SELECTION OF THE RAINFALL-RUNOFF MODEL

Two rainfall-runoff models were considered for use in this investigation--the U.S. Geological Survey Rainfall-Runoff Model (Dawdy and others, 1972) and the Distributed-Routing Rainfall-Runoff Model (DR3M) (Dawdy and others, 1978). DR3M is a combination of the U.S. Geological Survey Rainfall-Runoff Model and the routing component of the Massachusetts

Institute of Technology catchment model (Leclerc and Schaake, 1973). In DR3M the unit-hydrograph routing scheme used in the U.S. Geological Survey Rainfall-Runoff model was replaced with a distributed-routing scheme that represents the drainage structure of a basin as a collection of overland-flow, channel, and reservoir segments.

DR3M was selected for two reasons:

1. DR3M uses a detailed representation of the drainage structure. A detailed representation was needed because the soils in the basins are thin and have a relatively low permeability. As a result, the basins respond quickly to many storms.
2. DR3M can simulate both inbank and overbank flow. Separate channel parameters for inbank and overbank flow can be specified, and the storage of water in the flood plain during floods can be simulated. This capability is important in a model that is used to simulate flood flows. Linear unit-hydrograph routing models do not have this capability.

DEPENDENCE OF MODELS ON INPUT DATA

Mathematical models that require the input of data are highly dependent on the data used. Data provide approximations of variables represented in a model. For example, precipitation and evaporation data provide approximations of precipitation and evaporation in watershed models. The way that data approximate the variables they represent can have a substantial effect on the response of the model. When simulating hydrologic processes in a basin, one needs to consider not only the way in which the interrelation of hydrologic variables is represented by the hydrologic model, but also the way in which the hydrologic variables are represented by the data used. A hydrologic model may compensate for deficiencies in the way that data represent hydrologic variables, just as a component within a hydrologic model may compensate for deficiencies in another component.

Troutman (1982, 1983) analyzed the effects of errors in rainfall-runoff input data on model calibration and simulation. He considered errors caused by using point estimates of precipitation to estimate basin-wide precipitation. Troutman showed that bias is introduced into the simulated runoff even when input errors are random and unbiased. He pointed out that, when the model is calibrated, this bias is reduced through the selection of model parameters. He showed that if model input data used for simulation have an error structure that is different from the error structure of the input data used to calibrate the model, then the simulated output is biased. This is true even if the data used for model simulation have less error than the data used for model calibration. Troutman concluded that, to minimize the bias in simulated runoff, input data that are used for model calibration should have the same error structure as the inputs used for simulation.

Because of the dependence of models on input data, every effort was made to consider the properties of input data throughout the simulation process in this study. In particular, model calibration and simulation were done so that the properties of the data used for long-term simulation were as similar as possible to the properties of the data used to calibrate the basin models.

MODEL DESCRIPTION

DR3M consists of two major components--a rainfall-excess and a routing component. The rainfall-excess component simulates direct runoff from pervious and impervious surfaces. DR3M provides an automatic calibration scheme to facilitate the estimation of the parameters of this component. The routing component of the model is used to transform the time series of rainfall excess into flood discharge. The routing component is calibrated manually as a separate step after calibration of the rainfall-excess component. The derivation and workings of these components are described in more detail by Alley and Smith (1982).

Rainfall-Excess Component

The rainfall-excess component of the model is used to compute direct runoff during storms. The parameters of this component (table 1) are lumped over the entire basin. Two sets of parameters can be used optionally; however, the estimation of a second set of parameters requires more data than usually are available. The rainfall-excess component consists of two subcomponents--a daily soil-moisture accounting model and a storm-infiltration and runoff model. The soil-moisture accounting model is used to simulate infiltration of precipitation and evaporation on nonstorm days. The storm-infiltration and runoff model is used to compute infiltration and runoff during storms.

Impervious-Area Runoff

In DR3M, two types of impervious surfaces are considered. Effective-impervious surfaces are impervious surfaces that drain directly to the channel-drainage system. Rainfall on effective-impervious surfaces is assumed to contribute to direct runoff after satisfying a small impervious-retention storage. The remaining, noneffective-impervious surfaces are assumed to drain to pervious surfaces. Precipitation falling on a noneffective-impervious surface is assumed to run off onto surrounding pervious surfaces. This runoff is assumed to be uniformly distributed over the pervious surface, and is added to the precipitation falling there.

Pervious-Area Runoff

The rainfall excess from pervious surfaces is computed as a function of the rainfall rate and the point-potential infiltration of the pervious areas. In DR3M, the point-potential infiltration, FR, is computed as a function of two soil-moisture storages, SMS and BMS:

$$FR = KSAT \left[1 + \frac{PSP \left[RGF - (RGF-1)(BMS/BMSN) \right]}{SMS} \right], \qquad (1)$$

where

FR = point-potential infiltration, in inches per hour;
BMS = antecedent base-moisture storage, in inches;
SMS = saturated-moisture storage, in inches; and
KSAT, PSP, RGF, and BMSN are parameters of the rainfall-excess component. (See table 1 for a description of these parameters. Table 2 shows typical ranges for these parameters.)

Table 1.--Distributed-Routing Rainfall-Runoff Model parameters for soil-moisture accounting and infiltration

[From Alley and Smith, 1982]

Soil-Moisture-Accounting Parameters
EVC--A pan coefficient for converting measured pan evaporation to potential evapotranspiration.
RR--The proportion of daily rainfall that infiltrates into the soil for the period of simulation excluding unit days.
BMSN--Available soil water at field capacity, in inches.
Infiltration Parameters
KSAT--The effective saturated value of hydraulic conductivity, in inches per hour.
RGF--Ratio of suction at the wetting front for soil moisture at wilting point to that at field capacity.
PSP--Suction at wetting front for soil moisture at field capacity, in inches.

Table 2.--Ranges of values typically used for the soil-moisture-accounting and infiltration parameters

Parameter[1]	Range of values[2]	
EVC	0.50 - 1.0	
RR	.70 - .95	
BMSN	1.0 - 10	inches
KSAT	.015 - .2	inches per hour
RGF	2.0 - 30.	
PSP	0.5 - 10.	inches

[1] Definitions of these parameters are given in table 1.

[2] Ranges given in this table are based on parameter estimates found in Hauth (1974), Colson and Hudson (1976), Wibben (1976), Curtis (1977), Doyle and Miller (1980), Livingston (1981), Sloto (1982), Inman (1983), Wandle (1983), and Franklin (1984). All of these studies used either the U.S. Geological Survey Rainfall-Runoff Model or the Distributed-Routing Rainfall-Runoff model; both models use the same soil-moisture-accounting and infiltration model. In some of the reports cited, parameter values were reported which fell outside the ranges shown. The ranges shown are meant to show the ranges most commonly observed.

Actual infiltration potential differs from point to point in the basin and FR represents a nominal infiltration value. In DR3M, the distribution of infiltration potential as a function of area is assumed to range from a value of zero to the nominal value provided by equation (1). On the basis of this assumption, rainfall excess, QR, is computed from the following relations:

$$QR = \frac{SR^2}{2FR}, \qquad SR \leq FR \tag{2a}$$

$$QR = SR - \frac{FR}{2}, \qquad SR > FR \tag{2b}$$

where

QR = rate of rainfall excess, in inches per hour,
SR = rainfall rate or intensity, in inches per hour, and
FR = nominal infiltration potential, in inches per hour.

Infiltration is the difference between the rainfall rate, SR, and the rainfall excess, QR. Equations (2a) and (2b) indicate that rainfall excess is highly dependent on both infiltration potential and rainfall intensity. To simulate rainfall excess accurately from pervious surfaces, it is essential to (1) simulate accurately the soil-moisture conditions that govern point-potential infiltration, and (2) measure the rainfall intensity accurately.

Routing Component

The routing component of the model is used to transform rainfall excess into streamflow. The component uses a network of overland-flow, channel, and reservoir segments to represent the drainage system of a basin.

Overland-Flow Segments

Overland-flow segments are used to simulate the overland flow of surface runoff to the channel system. Kinematic routing is used to route rainfall excess, the output from the rainfall-excess component of the model, through the overland-flow segments. Overland-flow segments are depicted as shallow, rectangular channels in which the ratio of the depth of flow to the width of the channel is small.

Either a laminar or a turbulent flow regime can be simulated. Only one flow regime, however, can be used for a segment, and the modeler needs to specify the type of flow regime as a model parameter. The flow regime can be determined on the basis of the maximum flows computed by the model during model calibration. If the Reynolds numbers are less than 3,000, then the flows usually are assumed to be laminar. However, this conclusion is based on the assumption that the complex topography of an overland-flow area can be approximated by a plane. When a detailed representation is applied to an urban environment, then this assumption might be reasonably accurate. In natural or suburban basins, however, flow might not be uniform over an overland-flow area, but might be concentrated in small channels and rivulets. In this case, the flow regime is likely to be turbulent, even if the computed flows indicate laminar flow.

For each overland-flow segment, the overland-flow length, slope, roughness coefficient, and percent of the area that is impervious needs to be specified. Because each overland-flow segment drains to a channel, the model determines the area of each overland-flow segment by multiplying the overland-flow length by the length of the channel that drains it.

Channel Segments

Channel segments are used to simulate the flow in the channel system. Kinematic routing is used to route the flow through the channel segments. Flow, Q, in a channel is assumed to be related to the channel-flow cross-section area, A, by:

$$Q = \alpha A^m, \tag{3}$$

where

Q = channel outflow, in cubic feet per second (ft^3/s);
A = channel cross section area, in square feet (ft^2);
m = flow-routing exponent; and
α = flow-routing coefficient.

The flow routing coefficient and constant are channel parameters. The parameter α is a function of channel cross-section geometry, slope, and roughness. The parameter m is a function of the cross-section geometry. The model computes α and m if provided with a description of cross-section geometry, slope, and roughness for triangular or circular channels. The model also permits the input of values of α and m. If this option is used, an additional set of parameters can be specified for simulating overbank flow. In this case, one of two equations is used, depending on whether the computed flow exceeds the bankfull flow, Q_c:

$$Q = \alpha_i A^{m_i}, \quad Q \leq Q_c \tag{4a}$$

$$Q = \alpha_o A^{m_o}, \quad Q > Q_c \tag{4b}$$

where

Q = segment flow, in ft^3/s,
A = segment-cross-section area, in ft^2,
Q_c = is the cutoff top-of-bank flow, in ft^3/s,
m_i = the exponent for inbank flow,
α_i = the coefficient for inbank flow,
m_o = the exponent for overbank flow, and
α_o = the coefficient for overbank flow.

The explicit specification of α and m also can be used to specify channel-routing parameters for channels with other than circular or triangular cross sections.

For channels with circular cross sections, such as pipe segments, flow cannot exceed the nonpressurized-flow capacity of the pipe. If the inflow to a pipe segment exceeds that capacity, the water is stored behind the pipe segment. When the inflow becomes less than the nonpressurized flow capacity, the stored water is released. The rate of release is the

difference between the nonpressurized-flow capacity of the pipe and the rate of inflow to the pipe. If this situation arises, the model may give incorrect results. If storage behind the pipe is insufficient, then flow bypasses the pipe and reaches the basin outlet sooner than the model predicts. If storage is sufficient to hold the stored water, but only with a significant head of water over the pipe, then the actual flow in the pipe is pressurized, and the pipe flow computed by the model is too low. If either of these descriptions applies, the pipe segment can be replaced with a modified-Puls routing (Soil Conservation Service, 1972) reservoir with a storage-outflow relation determined through a hydraulic analysis of the pipe segment, and the topography upstream of the pipe. This procedure is described in more detail in a later section.

Reservoir Segments

Reservoir segments are used to account for the storage of water in detention basins, or behind pipes and culverts. Either of two methods may be used to model reservoir segments. One method, which uses a linear-storage representation of the reservoir, assumes that the outflow from a reservoir is related linearly to the reservoir storage by:

$$S = K\,O, \tag{5}$$

where

S = reservoir storage; in cubic feet per second times hours (ft^3/s-hours);
K = reservoir storage parameter, in hours; and
O = reservoir outflow, in ft^3/s.

The other method uses the modified-Puls routing method, which requires that the relation between reservoir storage and outflow be determined. It can provide a more detailed representation of flow through a reservoir than the linear-reservoir method, because it can accommodate a nonlinear relation between storage and outflow.

Direct rainfall on the surface of reservoirs is not accounted for by DR3M. If the surface area of a reservoir is significant, then rainfall on the reservoir surface can be modeled through the definition of a completely impervious overland-flow segment and a channel segment that drains to the reservoir. The overland-flow and channel segments are defined so that the product of the overland-flow and channel lengths is equal to the reservoir area, and so the water flows through the segments quickly.

Similarly, the model does not account for the lateral inflow of water to reservoirs from adjacent overland-flow segments. At reservoirs for which this inflow is significant, a channel segment can be added upstream of the reservoir that receives the lateral inflow. This channel segment should have the same length as the length of the reservoir boundary that receives the lateral inflow. The channel should have a large flow parameter, α, so that flow passes through it quickly.

Simplified Basin Representations

A basin representation can be simplified by collecting similar segments into groups, and then defining a single aggregate segment for each group.

This segment has parameters that are averages of the parameters in the group. The single aggregate-group segment is substituted for each of the individual group segments in the drainage network. The outflow from the aggregate-group segment is computed only once, and then is used wherever any of the individual group segment outflows is needed. Because the model needs to route flow through a segment only once in each group, the computational demands are reduced. Although some error is introduced by averaging segment properties, the drainage structure of the basin is preserved. Basins having complex drainage structures can require more than the 99 segments permitted by the model. Simplified basin representations can reduce the number of segments to less than 99.

Estimation of Base Flow

DR3M simulates only the surface-runoff component of flood flows. During floods, this usually is the most significant component, particularly for small and urban basins. However, several of the basins had significant base-flow components in some of their flood flows. A simple daily base-flow model was used to estimate base flow during the long-term simulation period.

The base-flow model was a soil-moisture-accounting model having a single storage variable. Water was added to soil-moisture storage through infiltration. Infiltration was computed as the product of daily precipitation and a model parameter, RR_{BF}. The storage was depleted by evaporation and base flow. Evaporation was computed as the product of the estimated daily pan evaporation and a model parameter, EVC_{BF}. Base flow, in inches, was assumed to be a linear function of the soil-moisture storage:

$$S_t = S_{t-1} + RR_{BF}P_{t-1} - Q_{t-1} - EVC_{BF}E_{t-1}, \quad (6)$$

$$Q_t = K_S S_t, \quad (7)$$

$$BF_t = 26.9\, Q_t A, \quad (8)$$

where

S_t = soil-moisture storage on day t, in inches;

P_t = precipitation on day t, in inches;

E_t = estimated pan evaporation on day t, in inches;

Q_t = soil-moisture depletion due to base flow on day t, in inches;

BF_t = base flow on day t, in ft^3/s, and

the model parameters are:

K_S = a base-flow recession constant;

RR_{BF} = infiltration coefficient for base flow;

EVC_{BF} = the evaporation coefficient for base flow;

A = the drainage area of the basin, in mi^2.

If the value of soil moisture, S_t, computed from equation (6), was less than zero, it was set to zero.

During calibration of DR3M the model separates observed streamflows into direct-runoff and base-flow components. The base-flow component is a constant for each storm. The base-flow values, reported by DR3M for the calibration data set, were used as a basis for calibrating the base-flow model.

Data Requirements

Four types of time-series data are required to calibrate DR3M: short-time- (usually 5- or 15-minute) interval precipitation, short-time-interval discharge, daily precipitation, and daily evaporation. Short-time-interval precipitation and runoff data are required when storms are being simulated. Storm-precipitation data from up to three rain gages can be input to the model. Observed storm-discharge data are compared to simulated storm-discharge data, as a basis for model calibration. Between storms, daily precipitation and evaporation data are used for soil-moisture accounting.

DATA COLLECTION AND PROCESSING

Site Selection

The following criteria were used as guidelines to select the basins studied:

1. basins are upstream of areas with serious flooding problems and the data collected can be used to provide early warning of flooding at downstream locations;
2. the land use in the basins is representative of variations in land use throughout the County;
3. the basins are hydrologically and hydraulically stable;
4. the stage-discharge relation is definable at the gage site;
5. basins are distributed over the County; and
6. the drainage areas of the basins are less than 10 mi^2.

The final site selection was done with the cooperation of the Somerset County Engineer's Office. Some of the basins selected did not meet all of the criteria listed above, but were selected for flood-monitoring purposes. The eight basins selected, the station numbers of the corresponding gaging stations, and the drainage areas of the basins are listed in table 3. The locations and outlines of the basins are shown in figure 2. A detailed description of the study area and a summary of the data collected are given by Campbell (1987).

Calibration Data

A stream-gaging station, a primary rain gage, and an auxiliary rain gage were installed in each of the study basins. The primary rain gage in each basin was located at or near the stream-gaging station. Records from the auxiliary rain gage were used in each basin to assess spatial variability, and for redundancy. The records from the auxiliary rain gages, which were located in the upper third of each basin, also were used to replace missing or inaccurate precipitation data from the primary rain gage. The station numbers and locations of the auxiliary rain gages are listed in table 4.

Table 3.--Streamflow-gaging stations

Station name	Station number	Drainage area (mi^2)	Location
Holland Brook at Readington, N.J.	01398107	9.00	Lat 40°33'30", long 74°43'50", Somerset County, on right bank 15 ft downstream from bridge on Old York Road.
North Branch Raritan River near Far Hills, N.J.	01398500	26.2	Lat 40°42'30", long 74°38'11", Somerset County, on left bank 75 ft upstream from Ravine Lake Dam.
Peters Brook near Raritan, N.J.	01400300	4.19	Lat 40°35'37", long 74°37'51", Somerset County, on left bank 12 ft upstream from bridge on Garretson Road.
Pike Run at Belle Mead, N.J.	01401650	5.36	Lat 40°28'05", long 74°38'57", Somerset County, on right bank 20 ft upstream from Township Line Road.
Royce Brook tributary near Belle Mead, N.J.	01402600	1.20	Lat 40°29'56", long 74°39'05", Somerset County, 25 ft upstream from bridge on State Highway 514 (Amwell Road).
West Branch Middle Brook near Martinsville, N.J.	01403150	1.99	Lat 40°36'44", long 74°35'28", Somerset County, 150 ft upstream from bridge on Crim Road.
Green Brook at Seeley Mills, N.J.	01403400	6.23	Lat 40°39'58", long 74°24'15", Somerset County, at Seeley Mills, 250 ft downstream from Blue Brook.
East Branch Stony Brook at Best Lake, at Watchung, N.J.	01403535	1.57	Lat 40°38'25", long 74°26'52", Somerset County, 700 ft upstream from dam on Best Lake in Watchung.

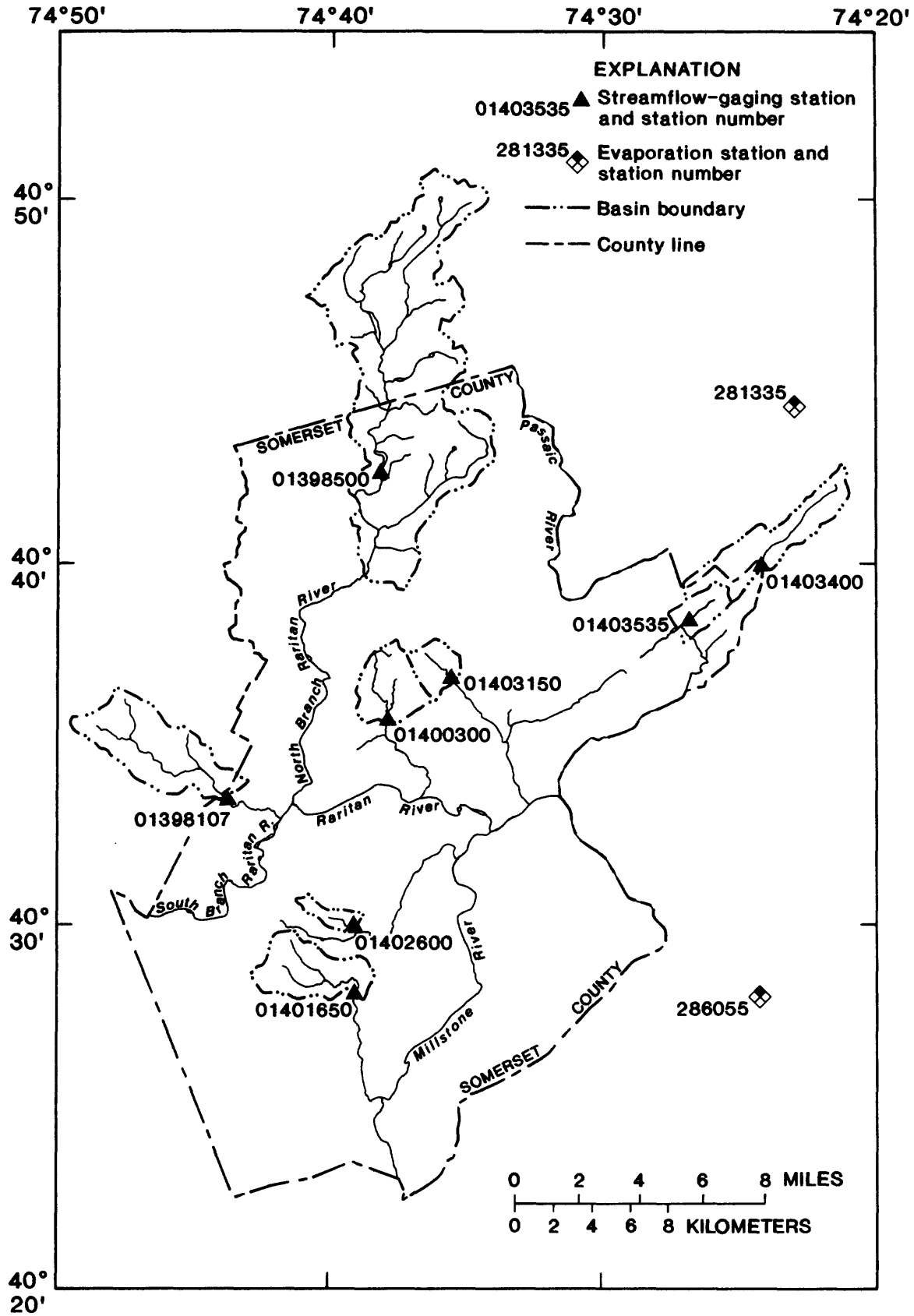


Figure 2.--Locations of modeled basins.

Table 4.--Location of auxiliary rain gages in Somerset County

[Elevations are approximate and expressed in feet above sea level]

Station name	Auxiliary rain gage station number	Location
Holland Brook at Readington, N.J.	403521074453501	Lat 40°35'21", long 74°45'35", on property of Readington School, in the fenced area for the sewage-treatment plant, 1.9 miles west of the town of Readington on Dreahook Road. Elevation, 240 feet, from topographic maps.
North Branch Raritan River near Far Hills, N.J.	404615074370001	Lat 40°46'15", long 74°37'00", on property of Saint John's School, about 100 ft north of the east end of the school, 0.8 miles east of the town of Mendham on State Route 24. Elevation, 620 feet, from topographic maps.
Peters Brook near Raritan, N.J.	403545074384201	Lat 40°35'45", long 74°38'42", on property of Transcontinental Gas pipeline company, about 300 feet west of Country Club Road, 1.1 miles north of the intersection of Country Club and Garretson Roads. Elevation, 160 feet, from topographic maps.
Pike Run at Belle Mead, N.J.	402826074414101	Lat 40°28'26", long 74°41'41", on property of Somerset County Park Commission, about 750 feet west of East Mountain Road, 0.7 miles north of the intersection of East Mountain and Blawenburg-Belle Mead Roads. Elevation, 140 feet, from topographic maps.
Royce Brook tributary near Belle Mead, N.J.	403025074401701	Lat 40°30'25", long 74°40'17", on property of gas pipeline and behind the second house east of Beekman Avenue on New Amwell Road, 2.3 miles west of the intersection of New Amwell Road and U.S. Route 206. Elevation, 120 feet, from topographic maps.

Table 4.--Location of auxiliary rain gages in Somerset County--Continued

[Elevations are approximate and expressed in feet above sea level]

Station name	Auxiliary rain gage station number	Location
West Branch Middle Brook near Martinsville, N.J.	403718074361701	Lat 40°37'18", long 74°36'17", on property of Bridgewater Township Park Commission about 300 feet north of Mount Vernon Road, 0.4 miles west from the intersection of Mount Vernon and Washington Valley Roads. Elevation, 270 feet, from topographic maps.
Green Brook at Seeley Mills, N.J.	404059074223301	Lat 40°40'59", long 74°33'33", on property of Morris County Park Commission about 700 feet west of the Trailside Museum, 1.6 miles west from the intersection of Central Avenue and U.S. Route 22. Elevation, 400 feet, from topographic maps.
East Branch Stony Brook at Best Lake, at Watchung, N.J.	403854074255401	Lat 40°38'54", long 74°25'54", on the roof of Bayberry School, about 600 feet south of Valley Road, 1.2 miles northeast of the intersection of Valley and Hillcrest Roads. Elevation, 320 feet, from topographic maps.

Stage and precipitation data were recorded every 5 minutes at seven of the eight basins. Five-minute streamflow data were computed from the 5-minute stage data. One of the basins, North Branch Raritan River near Far Hills, New Jersey, (station 01398500) had a drainage area of 26.2 mi^2, which was much larger than the drainage areas of the remaining seven basins. At this basin, a 15-minute recording interval for stage and streamflow was used.

The initial emphasis of this project was to establish a gaging network, with telemetry, to provide real-time data for flood monitoring. As a result, the gaging network was designed before the rainfall-runoff model was selected and without consideration of the error structure of the long-term precipitation data. Design of the rain-gage network focused on time-sampling errors and errors arising from the spatial variability of precipitation.

Although runoff simulation is influenced by the recording increment of the rain-gage data, the rain gages were selected mainly because they could be equipped with telemetry for flood-monitoring activities. Weighing-type rain gages with 0.1-inch recording increments were used to measure precipitation.

The computation of excessive precipitation in DR3M is highly dependent on rainfall intensity. For accurate runoff simulation, rainfall intensity must be measured accurately. If a 0.1-inch recording increment is used, a light rain falling over a long period is recorded as if it had fallen in 5 minutes. For example, if rain falls at a rate of 0.01 inch every 5 minutes for 50 minutes, 0.0 inches of rain is recorded for a period of 45 minutes followed by a 5-minute interval for which 0.1 inch of precipitation is recorded. The excess precipitation is computed on the basis of a single increment with an intensity of 0.02 inch per minute instead of the actual 10 increments with an intensity of 0.002 inch per minute.

Because of the recording increment used, the calibration rainfall data overestimate rainfall intensities while underestimating the number of rainfall increments, introducing errors into the model simulation. When the model was calibrated, these errors were overcome partially by automatically adjusting model parameters. Therefore, the calibrated infiltration parameters are not directly comparable to values obtained in other studies that used smaller precipitation increments.

DR3M permits the input of storm data from three rain gages. Data from only one rain gage were available for long-term simulation; therefore, data from only one rain gage were used for model calibration in order to make the error structures of the calibration data and the long-term data more similar.

Total daily precipitation was computed from 5-minute precipitation data for each rain gage.

Evaporation Data

DR3M requires the input of daily pan-evaporation data, used for simulating daily soil-moisture conditions. Daily pan evaporation and maximum and minimum air-temperature data at two National Oceanic and

Atmospheric Administration (NOAA) stations were obtained for the period of record from the National Climatic Data Center (NCDC). Initially, the evaporation data used for model calibration at a given basin were the data observed at the nearest of these two evaporation stations. The NOAA cooperative station names, station numbers, locations, and periods of record are given in table 5; the station locations are shown in figure 2.

Table 5.--National Oceanic and Atmospheric Administration evaporation stations

Station name	Station number	Latitude	Longitude	Period of record
Canoe Brook	281335	404500	0742100	1954 - 1984
New Brunswick	286055	402900	0742600	1968 - 1984

The period of record for the precipitation data used in the long-term simulation was 1914 through 1979. However, the measured evaporation data (table 5) does not span the entire period of record. Evaporation in this part of the County is rarely measured in winter because of low evaporation rates and freezing conditions. Often, a model is used to extend and fill in missing evaporation data with synthetic evaporation data.

Several models that could be used to extend and fill the pan-evaporation records were investigated. These models included a sinusoidal model, which estimated pan evaporation using a sine function with a period of one year; a model based on the average of historical pan-evaporation data for 122 three-day periods throughout the year; and a model that used mean-daily air temperature and possible hours of daylight to estimate pan evaporation (Hamon, 1961).

A split-sampling approach was used to evaluate the models. The observed evaporation records were split into two parts; the latter part of each record was used to fit each of the models that was being considered. The fitted models then were used to synthesize evaporation data for early periods of the observed evaporation records. The similarity between the observed and synthetic records for the early observed record periods provided the basis for comparing models. Because the simple log-sinusoidal model gave the best results, this model was fit again using the entire evaporation record for both evaporation stations.

Figure 3 shows a comparison between the observed and synthetic evaporation data for 1980 at Canoe Brook. The observed evaporation record was much more variable than the synthetic record. Because it might not be appropriate to use the less variable synthetic evaporation data for long-term simulations and the more variable observed record to calibrate the rainfall-runoff model, DR3M was calibrated using the synthetic evaporation data. The evaporation record at Canoe Brook is substantially longer than the record at New Brunswick (table 5); therefore, this record was used to derive a single evaporation model. The synthetic evaporation data are listed in table 6.

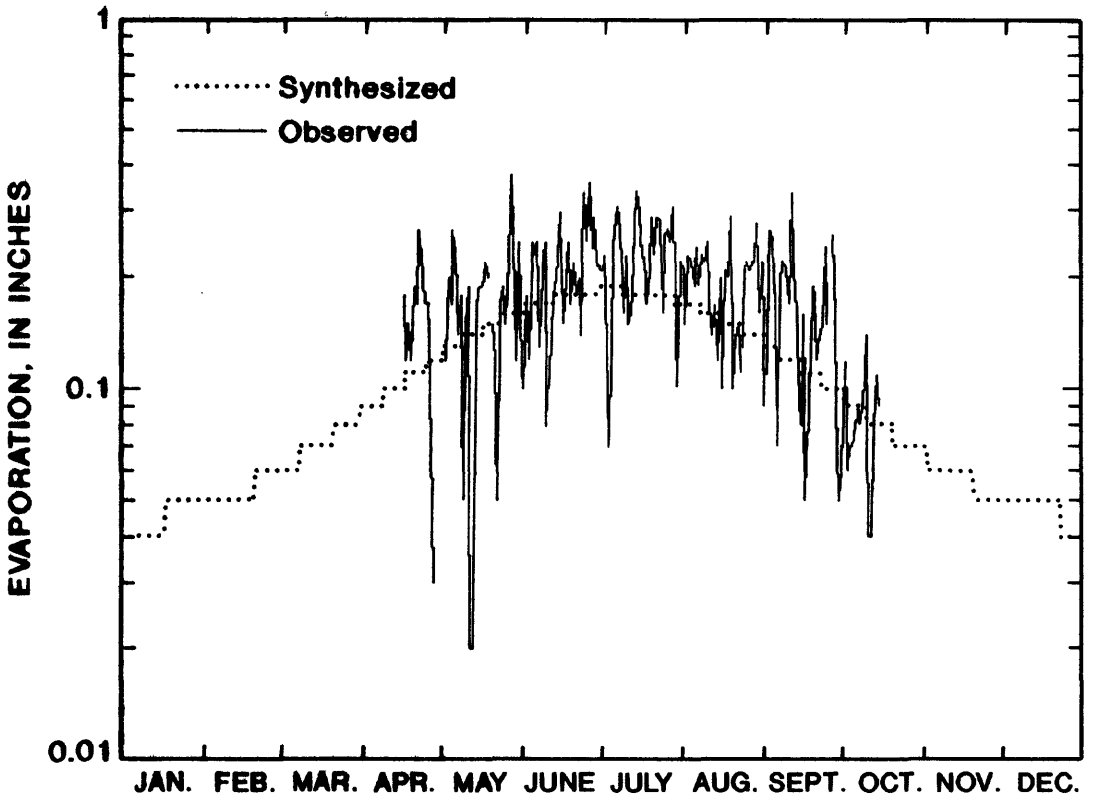


Figure 3.--Observed and synthetic evaporation data at Canoe Brook, 1980.

Table 6.--Synthetic pan-evaporation data for the evaporation station at Canoe Brook

[Evaporation is given in inches]

DAY	JAN	FEB	MAR	APR	MAY	JUNE	JULY	AUG	SEP	OCT	NOV	DEC
1	0.04	0.05	0.06	0.09	0.12	0.16	0.19	0.17	0.13	0.10	0.07	0.05
2	.04	.05	.06	.09	.13	.17	.19	.17	.13	.09	.07	.05
3	.04	.05	.06	.09	.13	.17	.19	.17	.13	.09	.06	.05
4	.04	.05	.06	.09	.13	.17	.19	.17	.13	.09	.06	.05
5	.04	.05	.06	.09	.13	.17	.19	.17	.13	.09	.06	.05
6	.04	.05	.06	.09	.13	.17	.19	.17	.13	.09	.06	.05
7	.04	.05	.06	.09	.13	.17	.19	.17	.13	.09	.06	.05
8	.04	.05	.07	.09	.13	.17	.19	.16	.12	.09	.06	.05
9	.04	.05	.07	.10	.14	.17	.19	.16	.12	.09	.06	.05
10	.04	.05	.07	.10	.14	.17	.19	.16	.12	.09	.06	.05
11	.04	.05	.07	.10	.14	.17	.18	.16	.12	.08	.06	.05
12	.04	.05	.07	.10	.14	.18	.18	.16	.12	.08	.06	.05
13	.04	.05	.07	.10	.14	.18	.18	.16	.12	.08	.06	.05
14	.04	.05	.07	.10	.14	.18	.18	.16	.12	.08	.06	.05
15	.04	.05	.07	.10	.14	.18	.18	.16	.11	.08	.06	.05
16	.04	.05	.07	.10	.14	.18	.18	.16	.11	.08	.06	.05
17	.05	.05	.07	.11	.15	.18	.18	.15	.11	.08	.06	.05
18	.05	.05	.07	.11	.15	.18	.18	.15	.11	.08	.06	.05
19	.05	.05	.07	.11	.15	.18	.18	.15	.11	.08	.06	.05
20	.05	.06	.07	.11	.15	.18	.18	.15	.11	.08	.05	.05
21	.05	.06	.08	.11	.15	.18	.18	.15	.11	.07	.05	.05
22	.05	.06	.08	.11	.15	.18	.18	.15	.11	.07	.05	.05
23	.05	.06	.08	.11	.15	.18	.18	.15	.10	.07	.05	.05
24	.05	.06	.08	.11	.16	.18	.18	.14	.10	.07	.05	.04
25	.05	.06	.08	.12	.16	.18	.18	.14	.10	.07	.05	.04
26	.05	.06	.08	.12	.16	.18	.18	.14	.10	.07	.05	.04
27	.05	.06	.08	.12	.16	.18	.18	.14	.10	.07	.05	.04
28	.05	.06	.08	.12	.16	.18	.18	.14	.10	.07	.05	.04
29	.05	.06	.08	.12	.16	.18	.17	.14	.10	.07	.05	.04
30	.05		.08	.12	.16	.19	.17	.14	.10	.07	.05	.04
31	.05		.09		.16		.17	.14		.07		.04
Total	1.39	1.55	2.22	3.12	4.48	5.29	5.65	4.80	3.43	2.48	1.71	1.47

Annual total = 37.59 inches.

MODEL CALIBRATION

Because DR3M is a nonlinear model, the linear-regression-based techniques for model evaluation cannot be used. Therefore, to evaluate model performance, split-sampling techniques usually are employed. In this approach, the data are separated into two parts, a calibration and a verification data set. The model first is calibrated using the calibration data set; then the model performance is evaluated using the verification data set. Model performance should be similar for both periods. Because of errors in the calibration data and approximations made in the model, rainfall-runoff models often may fit the calibration data well, but the model may be unable to simulate floods under varying hydrologic conditions. The model's performance during verification provides an independent estimate of its ability to simulate conditions different from those that prevailed during calibration.

Model verification provides an independent estimate of a model's predictive ability only if the verification data are independent of the calibration data. This implies that the calibration and verification data should be from different periods. If flood flows were independent of antecedent soil-moisture conditions, then individual floods might be assumed to be independent of one another, and calibration and verification data could be chosen randomly from a given period. If flood flows depend significantly on antecedent soil-moisture conditions, as they do in the study area, then individual flood flows depend on the contributions of other floods to soil-moisture conditions and on the effects of between-storm precipitation and evaporation. If the latter is true, and if calibration and verification storms are drawn from the same period, then it is unlikely that the verification results will give an independent estimate of the model's predictive ability. For this reason, the storms used for calibration were chosen from a different period than the storms used for verification.

In this study, 4 to 5 years of data were collected at each site. The storm data were split chronologically into two sets. The first set was used to calibrate the model, whereas the second set was used for model verification. The periods used for model calibration and verification, and various summary precipitation statistics for each site, are shown in table 7.

Calibration of DR3M consisted of four steps:

1. The basin drainage structure and values of measurable parameters were determined. The parameters included the number and configuration of channel and overland-flow segments, roughness values and slopes, channel geometry, and impervious cover area. These parameters have values that can be estimated from maps, aerial photographs, and field surveys.

2. The rainfall-excess component of the model was calibrated. The parameters of the rainfall-excess component were varied automatically to obtain a close fit between observed storm runoff and simulated rainfall excess.

Table 7.--Summary of calibration and verification data

Station number	Period	Period starting date	Period ending date	Number of storms	Precipitation statistics		
					Low	Median	High
01398500	calibration	12-24-1979	6-29-1982	27	0.7	1.4	3.9
	verification	8-25-1982	7-27-1984	28	.7	1.6	4.2
01398107	calibration	10-01-1979	6-29-1982	28	.7	1.1	3.1
	verification	9-27-1982	9-23-1984	29	.7	1.3	4.7
01400300	calibration	10-01-1979	4-27-1982	29	.7	1.2	3.2
	verification	6-16-1982	7-07-1984	30	.7	1.3	4.5
01403150	calibration	12-24-1979	4-26-1982	29	.7	1.3	3.3
	verification	4-27-1982	7-07-1984	28	.7	1.3	3.4
01403535	calibration	4-14-1981	4-10-1983	17	.8	1.6	3.5
	verification	4-16-1983	6-24-1984	18	.8	1.5	3.8
01403400	calibration	12-24-1979	11-12-1982	29	.8	1.5	3.3
	verification	11-28-1982	7-27-1984	30	.7	1.7	5.3
01402600	calibration	3-21-1980	4-26-1982	29	.7	1.3	3.2
	verification	4-27-1982	7-21-1984	30	.7	1.4	4.8
01401650	calibration	9-17-1980	11-28-1982	26	.7	1.2	3.3
	verification	3-18-1983	7-27-1984	26	1.0	1.7	5.1

3. The routing component of the model was calibrated. The routing parameters were adjusted to improve the fit between simulated and observed flood peaks.

4. Model calibration was verified through comparison of goodness-of-fit statistics computed for the calibration and verification periods using the calibrated model.

Identification of Basin Drainage Structure and Parameters

Determination of Drainage Structure

The basin drainage areas were determined from U.S. Geological Survey 7.5-minute National topographic quadrangle maps (7.5′ maps). After the basin drainage areas were determined, the channels and contributing overland-flow areas were identified using the 7.5′ maps. Two-foot contour maps (Maps Incorporated, 1964; Michael S. Kachorsky Associates, 1972) were provided by Somerset County, Bridgewater Township, and the Borough of Watchung. The 2-foot contour maps (2-ft maps) were used to identify channels and delineate overland-flow areas for basins with drainage areas less than 2 mi^2, and to determine the direction of overland flow where the 7.5′ maps did not provide adequate detail.

Simplified segmentation procedures were used whenever feasible. In this approach (Alley and Smith, 1982, attachment E), overland-flow or channel segments that have similar properties can be represented by a single segment that is used more than once. Although this approach makes the computation of overland-flow lengths and the estimation of headwater-channel lengths more complex, the reduction in routing computations and the simplification of the drainage structure justify its use.

Overland-Flow-Segment Parameters

Overland-flow areas were measured using a planimeter from the 7.5′ maps for the six largest basins. The overland-flow areas of the two smaller basins were measured from 2-ft maps. Overland-flow lengths were determined by dividing the overland-flow area of each overland-flow segment by the sum of the lengths of the channels that receive flow from the overland-flow area.

In addition to overland-flow length, the slope and roughness of each overland-flow segment must be specified. The overland-flow regime must be specified as laminar or turbulent. The slopes of the overland-flow segments were estimated from the 7.5′ maps. Two-ft maps were used for basins with drainage areas less than 2 mi^2, and when the 7.5′ maps did not provide adequate detail. Estimates of the roughness coefficients for overland-flow segments were based on the types of surfaces and vegetation, and on values provided by Alley and Smith (1982, p. 25). Initially, overland-flow regimes were assigned as laminar or turbulent based on segment slopes.

The final parameter that was specified for each overland-flow segment was the percent of the segment area covered by effective-impervious surfaces. Based on field observations, 50 percent of the impervious surfaces initially were assumed to be effective. The total impervious areas were estimated from topographic maps and aerial photographs. Roads were

divided into three categories by width. In each category, the lengths of roads were measured and the areas computed as the product of the lengths and the widths. Based on field inspection, the average area of houses was assumed to be 2,000 square feet (ft^2). The houses were counted and multiplied by this amount in each area. The areas of commercial buildings, apartments, parking areas, lakes, and reservoirs were measured using a planimeter or were estimated from aerial photographs. The estimate of impervious area in each overland-flow segment represented the total of all of these areas. The impervious-area percentage for each of the eight basins is given in table 8.

Table 8.--Impervious area, in percent of basin area

Station number	Impervious area (Percentage of total)
01398107	3.2
01398500	10.
01400300	13.
01401650	10.
01402600	11.
01403150	4.9
01403400	9.0
01403535	12.

Channel-Segment Parameters

The lengths of the main channels that receive flow from upstream channels were measured directly from topographic maps. Initially, lengths of the headwater channels that receive input solely from overland-flow segments were estimated from topographic maps. Headwater-channel lengths are uncertain and may change with changing flow regimes. These channel lengths were adjusted to provide reasonable computed overland-flow lengths.

In addition to a channel's length and position in the drainage network, estimates of channel slope, roughness, and cross-section geometry are required. From these, DR3M computes channel parameters α and m that satisfy equation (3). The formulas that are used by DR3M to compute α and m are given by Alley and Smith (1982, p. 27-31). DR3M computes α and m only for channels with triangular or circular cross sections; for these types of cross-section geometry it is possible to derive α and m analytically from Manning's equation.

Most of the channels in the modeled basins had cross sections that were nearly trapezoidal in shape. DR3M permits the user to input values of α and m directly. It is impossible to derive estimates of α and m for trapezoidal channels from Manning's equation that satisfy equation (3) exactly over a range of flows. However, it is possible to derive estimates of α and m for which equation (3) is satisfied approximately. α and m can be estimated by deriving a relation between the logarithms of flow and flow cross-section

area through hydraulic analysis of slope, roughness, and cross-section geometry, and by fitting a straight line to the derived values. The slope of this line provides the estimate of the m; the antilogarithm of the intercept of the line provides the estimate of α.

If α and m are provided by the user, then DR3M permits the use of two sets of these parameters--one for inbank and one for overbank flow. This feature was utilized in this study, and the parameters α and m were computed for both inbank and overbank flow. The relation between the logarithms of flow and flow-cross-section area was developed through hydraulic analysis of the slope, roughness, and cross-section geometry. Two straight lines were fit to this relation--one to the part of the relation governing inbank flows and one to the part of the relation governing overbank flows. The slopes of these lines provide estimates of the inbank and overbank flow parameters m_i and m_o; the antilogarithms of the intercepts provide estimates of the parameters α_i and α_o. The flow at the point at which the two lines intersect provides the cutoff flow, Q_c. At this point, the model switches from one set of parameters to another.

Channel slopes were estimated from topographic maps. Channel-roughness and cross-section-geometry estimates were obtained from field surveys and observations at six of the basins. At stations 01401650 and 01402600, cross-section and roughness data were available from a previous study (Federal Emergency Management Agency, 1980). Field surveys and observations were used to verify and supplement the data from these two stations.

Simplification of Drainage Representation at Stations 01398107 and 01398500

Stations 01398107 and 01398500 have drainage networks that are so complex that they required more than the 99 segments permitted by DR3M. For these basins, it was necessary to utilize the simplification strategy discussed earlier to reduce the number of segments.

The first step in simplifying the channel network for these basins was to assemble a list of the overland-flow segments and their parameters. For comparison, slopes and roughnesses were combined into a single parameter, SR, given by:

$$SR = \text{slope}^{1/2}\text{roughness}^{-1}. \quad (9)$$

This parameter represents the joint contribution of slope and roughness to the routing parameter, α, for both overland-flow and channel segments. All overland-flow segments were assumed to have turbulent flow regimes. Therefore, the flow regime was not used as a basis for comparing overland-flow areas. Three parameters were used for comparing overland-flow areas: SR, overland-flow length, and the percentage of effective-impervious area.

The overland-flow segments were divided into groups using a combination of cluster and graphical analyses. The total effective-impervious area was 10 percent or less of the total drainage area for both basins; also, the impervious areas were fairly evenly distributed. Therefore, the grouping of the overland-flow areas was based mostly on analysis of the slope-roughness parameter, SR, and the overland-flow length.

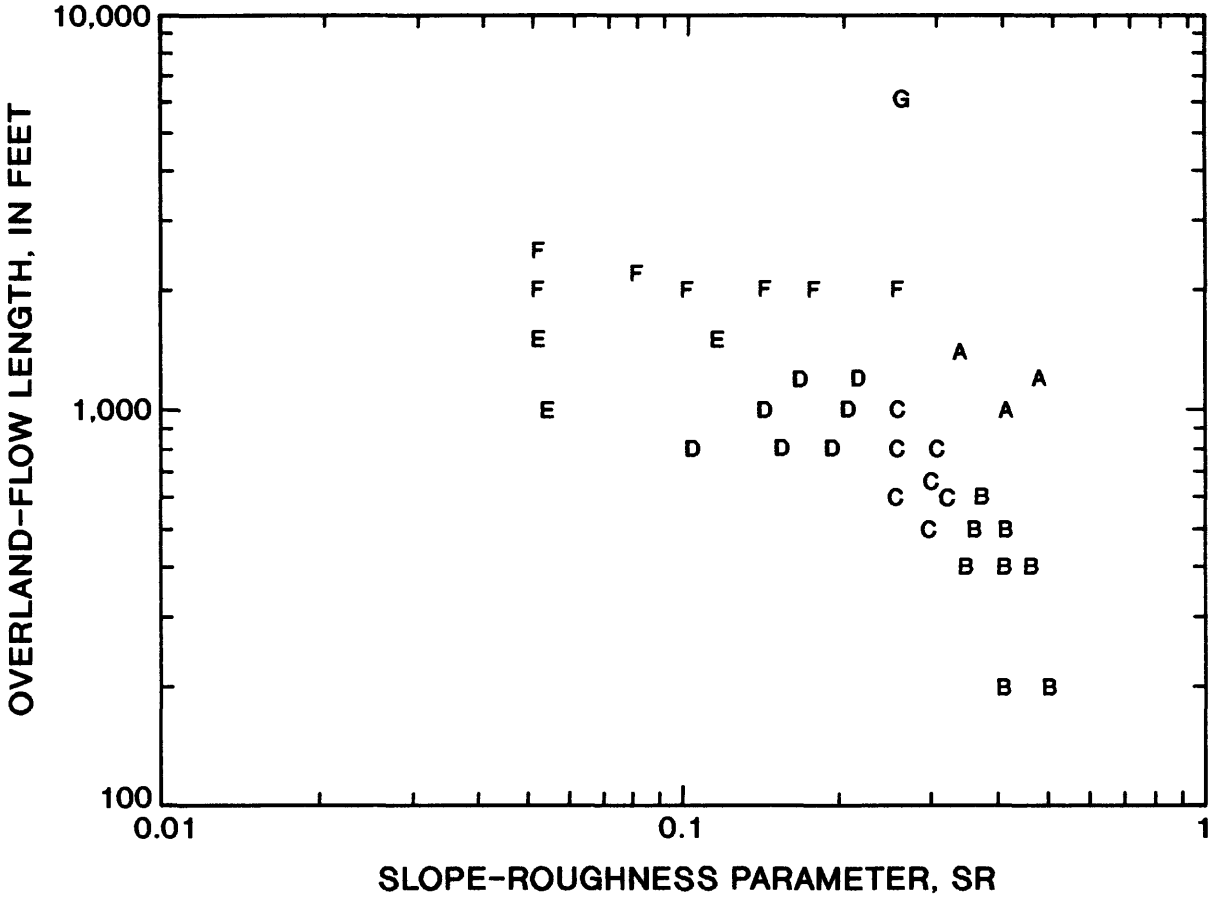
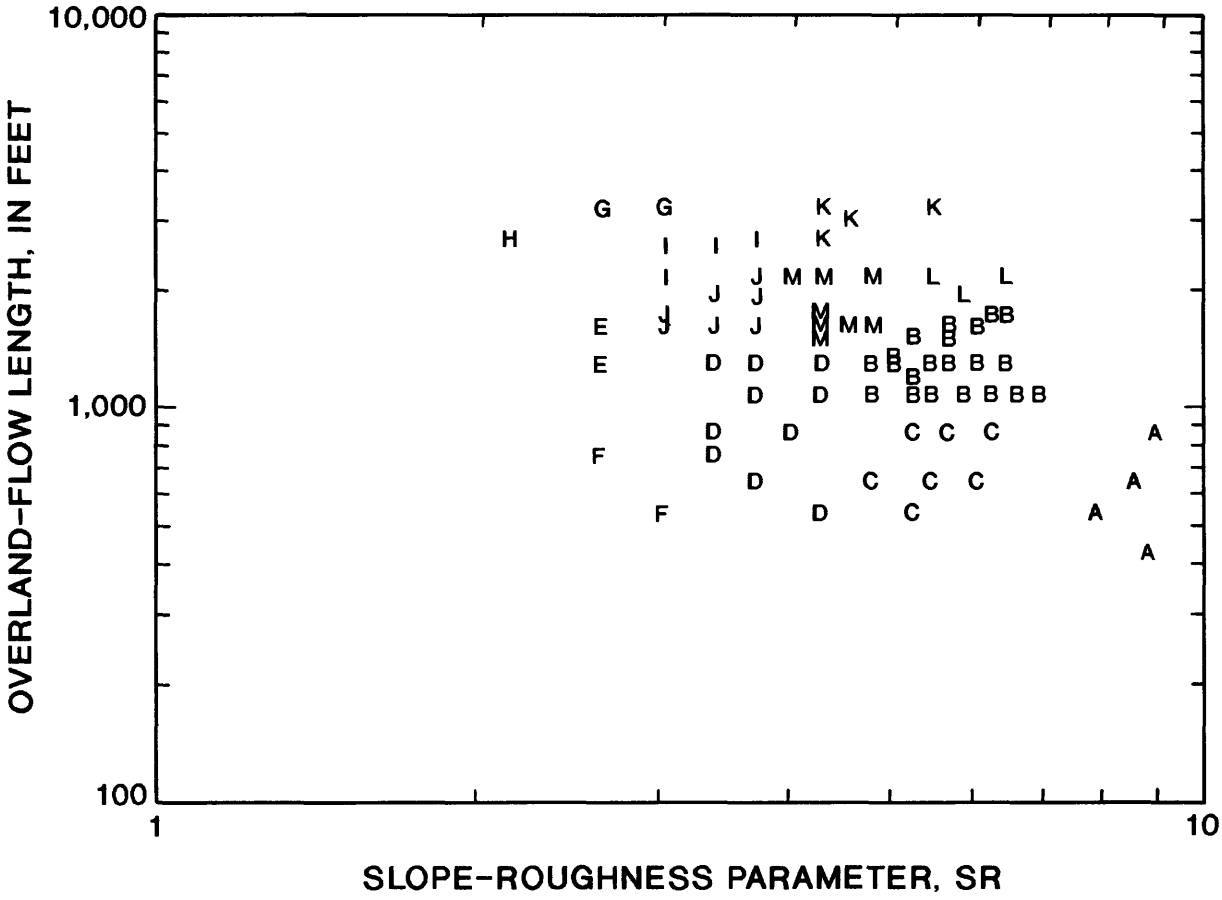


Figure 4.--Relation between the values of the slope-roughness parameter and the overland-flow lengths for overland-flow segments for station 01398107.



EXPLANATION

A - Overland-flow group A
B - Overland-flow group B
C - Overland-flow group C
D - Overland-flow group D
E - Overland-flow group E
F - Overland-flow group F
G - Overland-flow group G
H - Overland-flow group H
I - Overland-flow group I
J - Overland-flow group J
K - Overland-flow group K
L - Overland-flow group L
M - Overland-flow group M

Figure 5.--Relation between the values of the slope-roughness parameter and the overland-flow lengths for overland-flow segments for station 01398500.

Figures 4 and 5 show the distribution of the slope-roughness parameters and overland-flow lengths at stations 01398107 and 01398500. The overland-flow-segment groupings also are shown in these figures. The overland-flow lengths and percent effective-impervious areas were averaged arithmetically; the slope-roughness coefficients were averaged geometrically. Each overland-flow group was assigned a turbulent roughness coefficient of 0.1. The slopes for each overland-flow group were computed from:

$$\text{slope} = (0.1 \text{ SR})^2 \qquad (10)$$

The overland-flow segments in each group were replaced with the overland-flow segment derived for the group. The use of groups markedly reduced the number of required segments.

The headwater-channel segments, channel segments that received only lateral inflow, were classified according to the overland-flow groups that contributed lateral inflow to them. For headwater-channel segments that received lateral inflow from the same overland-flow groups, segments with similar channel lengths and slope-roughness coefficients were assembled into headwater-channel groups. For each of these groups, a group channel segment was derived that had parameters with values equal to the averages of the parameters for the segments in the group. The channel lengths were averaged arithmetically; the routing parameters were averaged geometrically. Channel parameters α and m were specified explicitly for the headwater-channel group segments. Each channel segment in a headwater-channel group was replaced by the appropriate group segment for that group. Because of the variety of combinations of overland-flow groups and the range of channel lengths and routing parameters, the reduction in the number of headwater-channel segments was minor (table 9).

Calibration of the Rainfall-Excess Component Parameters

Selection of Storms for Use in Calibration

The first step in the calibration of the rainfall-excess component was storm selection. All storms with at least 0.7 inch of precipitation were identified. This criterion was used to limit the data to storms for which direct runoff could be expected to provide the largest component of flow. Storms were defined as periods of precipitation having a maximum of less than one day between precipitation increments. The storm data were reviewed manually and were edited as needed. Because DR3M does not simulate snow accumulation or melt, storms that occurred during periods of below-freezing air temperature or for which there was evidence of snow accumulation or melt were not included in the data set.

Many storms had low runoff coefficients (the ratio of runoff to rainfall volume). In many cases, they were as low as 2 to 5 percent. The low runoff coefficients initially were assumed to be the result of erroneous rainfall or streamflow data. However, comparison of data from all the gages revealed that low runoff coefficients were common to all the basins. Because the storms in question followed prolonged dry periods, they were difficult to fit using soil-moisture and infiltration parameters that were within limits commonly used (table 2). Therefore, it was necessary to use values of the

Table 9.--Reduction in the number of segments due to model simplification at stations 01398107 and 01398500

	01398107		01398500	
	Before simplification	After simplification	Before simplification	After simplification
Overland-flow segments	76	8	133	14
Headwater-channel segments	20	20	35	23
Non-headwater-channel segments	18	18	32	32
Reservoir segments	0	0	2	2
Total segments	114	46	202	71

parameter RGF that were larger than those commonly used. The parameter RGF reflects the difference in infiltration potential between wet and dry conditions.

To assess the necessity of calibrations to the storms with low runoff coefficients, an attempt was made to calibrate the model while excluding data from these storms from the calibration data. When the low-runoff storms were excluded, the resulting parameter set caused the model to overestimate the runoff for these storms excessively. Although the observed peaks from the low-runoff storms were almost negligible, the simulated peaks, in some cases, equaled or exceeded the largest observed peaks for the years in which the storms occurred. Therefore, it was necessary to include the low-runoff storms in model calibration.

Objective Function Used for Automatic Calibration

Automatic calibration requires the choice of an objective function. This function, which is expressed as a function of the model parameters, reflects the goodness of fit between simulated model output and observed output (storm volumes) of the system. The model parameters are adjusted to achieve a minimum in the objective function. The objective function should be chosen to reflect the error structure of the system output being simulated. The chosen objective function should tend to give more weight to observations with low measurement errors.

The objective function that commonly is used in DR3M is a LLS (log-least squares) objective function. The LLS objective function consists of the sum of the squared differences between the logarithms of the observed and simulated storm volumes:

$$LLS = \sum_{i=1}^{n} [\ \log Q_{s,i} - \log Q_{o,i}\ ]^2, \quad (11)$$

where

$Q_{o,i}$ = observed storm volume for storm i, in inches,
$Q_{s,i}$ = simulated rainfall excess for storm i, in inches, and
n = number of storms.

An assumption implied by this objective function is that the standard deviations of the errors in the storm volumes are proportional to the magnitude of the storm volumes. This means that the errors are much higher for storms with high volumes than for storms with low volumes. An equivalent assumption implied by this objective function is that the log-errors have a constant variance, and that log-errors for large storms have the same magnitude as log-errors for small storms. The log-least squares objective function places a greater emphasis on small storms than on large storms, because of the assumption that the errors for the small storms are much less than the errors for large storms.

During the initial calibration the model tended to fit large storm volumes poorly. In some cases fit was extremely poor because the calibration procedure placed a great emphasis on small storms, including storms that had extremely small runoff coefficients as well as other storms for which the runoff was uncertain.

Although the error structure of the fitted model confirmed the assumptions implied by the LLS objective function, it did so because of the unreasonably poor fit of the large storm volumes.

Several alternative objective functions were considered for calibrating the model for storm volumes. These included an OLS (ordinary least squares) objective function, as well as three objective functions based on the Box-Cox transformation (BCT) (Box and Cox, 1964).

The OLS objective function is computed as the sum of the squared differences between the observed and predicted storm volumes:

$$OLS = \sum_{i=1}^{n} [\, Q_{s,i} - Q_{o,i} \,]^2, \quad (12)$$

where

$Q_{o,i}$ = observed storm volume for storm i, in inches,
$Q_{s,i}$ = simulated rainfall excess for storm i, in inches, and
n = number of storms.

This objective function implies an assumption that the errors have a constant variance, regardless of the volume magnitude. In addition, it implies that the log-errors are larger for smaller storms. This objective function gives equal emphasis to all storms, regardless of volume, and places greater emphasis on large storms than does a LLS objective function.

The objective functions based on the BCT have the advantage that they can adapt to the error structure of the model being calibrated. These objective functions require the estimation of a transformation parameter, λ, which reflects the variation of the magnitude of error variance as a function of output magnitude. A drawback in the use of objective functions based on the BCT with DR3M is that an additional parameter, the transformation parameter, must be estimated in addition to the model parameters. DR3M permits a maximum of 60 storms to be used for model calibration. With 60 or fewer storms, data may only be sufficient to estimate the model parameters. The addition of another parameter, even though it may improve the power of the objective function, can decrease the reliability of the parameter estimates obtained. This problem is particularly acute if the data are split into calibration and verification data sets. In this case 30 or fewer storms may be available.

The various objective functions mentioned above were tested during initial calibration runs. The only objective function that produced consistently satisfactory results was the OLS objective function. On the basis of these tests, the OLS objective function was used for the remainder of the study.

Goodness-of-Fit Statistic Used to Evaluate Model Performance

Various goodness-of-fit statistics were computed to evaluate the performance of the calibrated model. The most useful of these was the efficiency statistic (Nash and Sutcliffe, 1970; Loague and Freeze, 1985):

$$\text{Efficiency} = 100 \times \frac{\sum_{i=1}^{n} [\, Q_{o,i} - \bar{Q}_o \,]^2 - \sum_{i=1}^{n} [\, Q_{s,i} - Q_{o,i} \,]^2}{\sum_{i=1}^{n} [\, Q_{o,i} - \bar{Q}_o \,]^2}, \quad (13)$$

where

$Q_{o,i}$ = observed storm volume or peak for storm i,

$\bar{Q}_o$ = mean of observed storm volumes or peaks,

$Q_{s,i}$ = simulated storm volume or peak for storm i, and

n = number of storms.

An efficiency value of 100 indicates a perfect fit between predicted and observed values. An efficiency value of zero indicates that the model is no better than the observed mean at predicting observed values.

The efficiency is equivalent to the OLS objective function in the sense that, for a given calibration data set, a model parameter set that yields a lower value of the OLS objective function than some other parameter set will always yield a lower efficiency value. The efficiency can be rewritten as

$$\text{Efficiency} = 100 - 100 \times \left[\frac{\text{OLS}}{\sum_{i=1}^{n} [\, Q_{o,i} - \bar{Q}_o \,]^2} \right] \quad (14)$$

where

$Q_{o,i}$ = observed storm volume or peak for storm i,

$\bar{Q}_o$ = mean of observed storm volumes or peaks,

OLS = the OLS objective function, and

n = number of storms.

The denominator is constant for a given data set and basin; the efficiency is completely specified by the OLS objective function value.

The advantage in using the efficiency is that differences of scale between basins and calibration sets are removed through division by the observed sum of squares, facilitating the comparison of efficiencies between periods and between basins.

Estimation of Effective-Impervious Areas

The size of the effective-impervious areas usually is the first parameter in the rainfall-excess component to be estimated. An initial estimate of the size of impervious areas can be obtained from topographic

maps by measuring the lengths of roads, the number of houses and large buildings, the areas of parking lots, and the areas of lakes and reservoirs. A part of the total impervious area is assigned as effective. The percentage of impervious area that is effective may be estimated through field inspection of impervious surfaces. If the percentage of the total area that is impervious is small, then an approximate estimate of the effective-impervious area is adequate, because the model is insensitive to the value used. Fifty percent of the impervious area in each basin initially was assumed to be effective.

If storms during which the runoff is predominantly from effective impervious surfaces can be isolated, then the initial estimate of effective-impervious area can be adjusted to fit these storms. In some situations it may be difficult to identify storms with predominantly effective-impervious-area runoff. This is likely to be the case when either the extent of effective-impervious area is small or when the pervious areas are covered by low-permeability soils. In this situation, the sensitivity of simulated flows to the size of effective-impervious area is low and accurate estimation of this parameter is not critical. Theoretically, if there are no losses in the channel network, then the effective-impervious area, as a fraction of the basin area, should not exceed the minimum accurately observed runoff coefficient.

An attempt was made to identify those storms for which runoff was predominantly from effective-impervious surfaces and to use them to calibrate the effective-impervious area. These storms were identified by:

1. choosing the approximately one-fourth or one-fifth of the total storms that had the lowest runoff volumes;

2. choosing the approximately one-fourth or one-fifth of the total storms that had the lowest ratio of runoff to rainfall volume; or

3. calibrating the parameters that govern pervious-area runoff and then choosing those storms for which the simulated pervious-area runoff is zero or nearly zero.

The model permits the calibration of the effective-impervious area for a basin through the model parameter EAC. The estimates of the effective-impervious area for each subarea are multiplied by the value of EAC for each model run. Parameter EAC can be calibrated along with the soil-moisture-accounting and infiltration parameters.

Calibration of EAC was attempted using the storms selected with each of the above criteria separately and in combination. Calibration of EAC and the pervious-area runoff parameters together also was attempted. Estimates of EAC varied widely depending on the storms used and the values of the pervious-area runoff parameters.

Impervious areas in the basins generally are small (table 8), and the pervious surfaces are covered by low-permeability soils; therefore, computed runoff was less sensitive to effective-impervious area than to the pervious-area runoff parameters, even during dry periods. Also, computed runoffs generally were insensitive to changes in the estimate of effective-impervious area.

Because consistent estimates of EAC could not be obtained through automatic calibration, initial estimates of effective-impervious area were used.

Calibration of the Soil-Moisture-Accounting and Infiltration Parameters

After estimation of the effective-impervious area, the pervious-area runoff parameters were calibrated. The pervious-area runoff parameters are listed in table 1. Figures 6 through 9 show the response of the OLS objective function to variations in parameter values, for various pairs of pervious-area runoff parameters, for basin 01398500. As can be seen from these figures, some of the pervious-area runoff parameters interact strongly with one another. There is a three-way interaction among parameters PSP, KSAT, and RGF, as would be expected from equation (1). There is also a strong interaction between parameters EVC and RR.

Alley and Smith (1982, p. 17) note the strong interaction between EVC and RR, and between KSAT and PSP. Because independent estimates can be obtained for EVC and RR, they recommend that these parameters be left out of the optimization. They also argue that, because KSAT and PSP are highly interactive, they should not be calibrated simultaneously.

In this study, synthetic evaporation data rather than pan-evaporation data were used. In addition, the means of the pan-evaporation records at the two nearest pan-evaporation stations differed significantly, as can be seen from table 10. Therefore, it was impossible to estimate parameter EVC, which is used by the model to convert pan-evaporation data to potential-evaporation estimates, reliably in advance of model calibration.

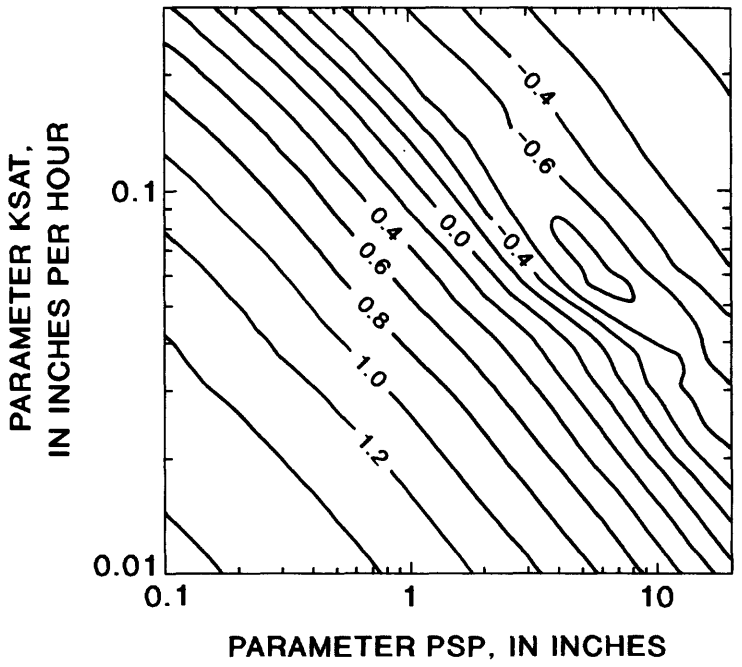
Table 10.--Mean evaporation at the Canoe Brook and New Brunswick evaporation stations computed for days when both stations were recording in period 1968-83

[Values are in inches per day]

Station	Geometric mean	Arithmetic mean
Canoe Brook	0.14	0.16
New Brunswick	.17	.21

In initial calibration runs, EVC and RR were calibrated simultaneously. The optimized parameter values were greater than their theoretical maximum values. As shown in figure 9, EVC and RR interact in such a way that large values of EVC are compensated for by large values of RR, and vice versa. Because optimizing EVC and RR led to unreasonable values for both parameters, one of the parameters was excluded from the optimization. Because of the uncertainty in estimating the value of EVC it was left in the optimization, and the value of RR was fixed.

The parameter RR is essentially equal to one minus the runoff coefficient for rainfall on nonstorm days. Actual runoff coefficients for

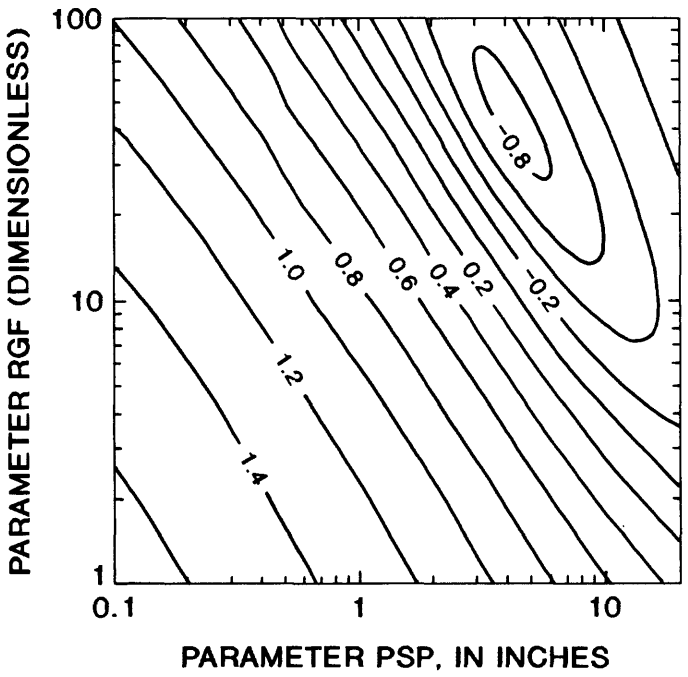


EXPLANATION

—— 0.8 —— Line of equal logarithm (base 10) of ordinary least-squares objective function.

Figure 6.--Response of base 10 log of ordinary-least-squares objective function for storm volumes at station 01398500 to variation in parameters PSP and KSAT.

(PSP is the suction at the wetting front for soil moisture at field capacity, in inches. KSAT represents the effective saturated value of hydraulic conductivity, in inches per hour.)

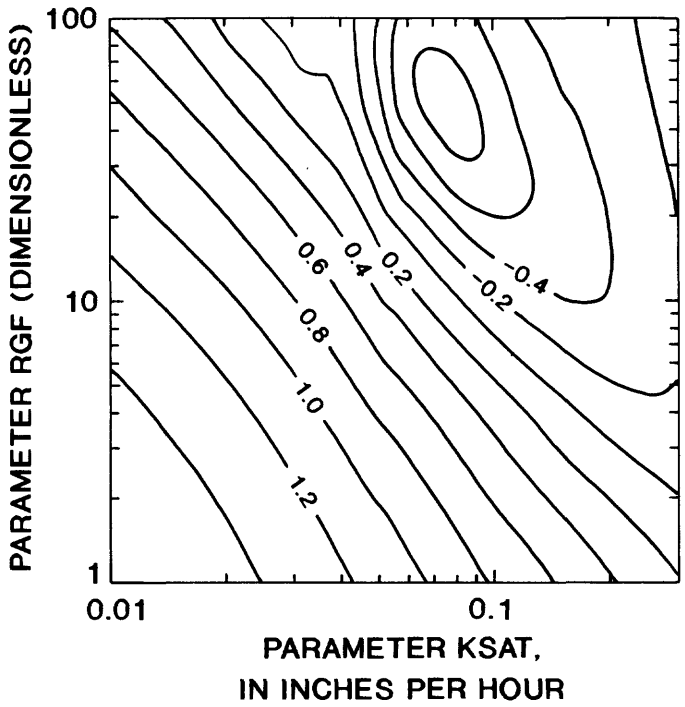


EXPLANATION

—0.8— Line of equal logarithm (base 10) of ordinary least-squares objective function.

Figure 7.--Response of base 10 log of ordinary-least-squares objective function for storm volumes at station 01398500 to variation in parameters PSP and RGF.

(PSP is the suction at the wetting front for soil moisture at field capacity, in inches. RGF is the ratio of suction at the wetting front for soil moisture at the wilting point to that at field capacity.)

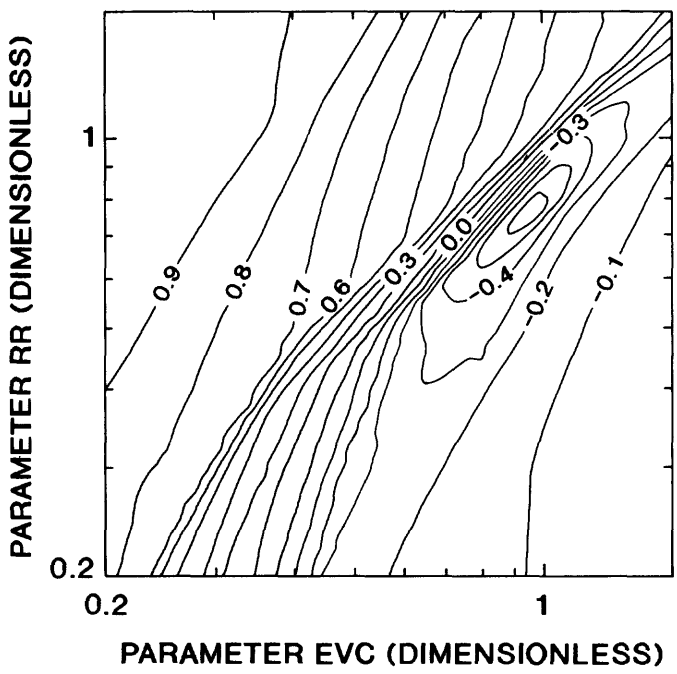


EXPLANATION

—0.8— Line of equal logarithm (base 10) of ordinary least-squares objective function.

Figure 8.--Response of base 10 log of ordinary-least-squares objective function for storm volumes at station 01398500 to variation in parameters KSAT and RGF.

(KSAT represents the effective saturated value of hydraulic conductivity, in inches per hour. RGF is the ratio of suction at the wetting front for soil moisture at the wilting point to that at field capacity.)



EXPLANATION

—0.8— Line of equal logarithm (base 10) of ordinary least-squares objective function.

Figure 9.--Response of base 10 log of ordinary-least-squares objective function for storm volumes at station 01398500 to variation in parameters EVC and RR.

(EVC is a pan coefficient for converting measured pan evaporation to potential evapotranspiration. RR is the proportion of daily rainfall that infiltrates into the soil for the period of simulation excluding unit days.)

nonstorm days are difficult to determine, because the streamflows on nonstorm days contain large percentages of subsurface flow. For this reason, the estimates of RR were based on runoff coefficients computed for the storms used for model calibration. At each basin, a linear-regression model was used to estimate the runoff coefficient from the total storm precipitation. The regression relationship was significant at the 0.05 probability level for only three of the basins--Holland Brook (station 01398107), Royce Brook (station 01402600), and Pike Run (station 01401650). At these basins, the regression model was used to calculate the predicted runoff coefficient for a 0.5-inch storm. This runoff coefficient was used to compute the estimate of RR. At the remaining basins, the median runoff coefficient for each basin was used to compute the estimate of RR.

Alley and Smith (1982, p. 17) recommend that only KSAT or PSP be optimized because of the high degree of interaction between the two (fig. 6). Because estimates of KSAT can be developed based on directly observable parameters such as soil type, KSAT is often excluded from the optimization.

In this study, parameters PSP and KSAT were calibrated simultaneously for two reasons. First, although parameters PSP and KSAT interact strongly with each other, they also interact strongly with parameter RGF (figs. 7 and 8). In fact, there is a three-way interaction among these variables, and it is unclear which variable or variables should be excluded from the calibration to eliminate the interaction.

Second, although KSAT corresponds to a physical parameter--the soil hydraulic conductivity--it is unclear that estimated values of soil hydraulic conductivity provide meaningful estimates of KSAT in DR3M. The soil hydraulic conductivity usually is estimated based on the type of soil in a basin. For example, ranges of KSAT for different soil types are given by Chow (1964, p. 12-26). Ideally, the value of KSAT that is used in a model such as DR3M compensates to some degree for approximations or other simplifications in the model and is probably different from the actual hydraulic conductivity of the soil.

Although many of the parameters in a model correspond to physical properties of a basin, they are not equivalent to them. For parameter KSAT, the ideal value used in DR3M probably is closer to double the average hydraulic conductivity of the soil than to the average value. This is because the average value of infiltration potential, FR, is used by DR3M to represent the maximum infiltration potential in the basin, which is equivalent to using a value of FR/2 to estimate average potential infiltration. To compensate for the use of FR/2 to represent average potential infiltration, the value of KSAT used in DR3M ideally should be about twice as large as the value of the hydraulic conductivity, so that the average infiltration potential computed by DR3M is equal to the point infiltration potential computed using the average hydraulic conductivity in the basin. This may explain why the values that define ranges of KSAT for various soil types, recommended by Alley and Smith (1982, p. 19), are about twice as large as those recommended by Chow (1964, p. 12-26). Because of the uncertainty associated with using soil hydraulic conductivity to estimate the model parameter KSAT, it was decided to calibrate the model for KSAT as well as PSP and RGF.

In this study, PSP, KSAT, and RGF were calibrated simultaneously. In most cases, the calibrated values of PSP and KSAT were within the limits recommended by Alley and Smith (1982, p. 19). In almost every case, the calibrated value of RGF was larger than the recommended maximum value of 20. A survey of similar modeling studies revealed that larger values of RGF have been used (table 2). Because of the uncertainties inherent in model calibration, the calibrated value of RGF was constrained to be less than 30. For most of the basins, constraining the value of RGF proved to be equivalent to fixing the value of RGF at the constraining value.

The final soil-moisture-accounting and infiltration-parameter estimates are listed in table 11.

Calibration of the Routing Component

Calibration of the routing component of the model consisted of three steps:

1. determine the number of subsegments used for finite-difference computations;

2. determine the flow regime for overland-flow segments; and

3. vary the α adjustment factor, ALPADJ, to achieve a close fit between simulated and observed hydrographs.

The order of these steps was varied often, and steps often were repeated in an iterative fashion.

Determination of Flow Regimes for Overland-Flow Segments

An analysis was made to determine the flow regime for overland-flow segments. To determine a range of flow values for the outflow from each segment, initial calibration runs were performed. The maximum Reynolds number for each overland-flow segment and storm was computed. The range of these Reynolds numbers for a given overland-flow segment was used to determine whether the flow regime was laminar or turbulent. If the analysis indicated a flow regime different from that originally assumed, then the appropriate changes were made in the model control information and the simulation was rerun. The new Reynolds numbers were checked to ensure that the results of the analysis still supported the given flow regime.

The analysis of the outflow Reynolds numbers indicated the predominance of laminar-flow regimes for the majority of overland-flow segments. However, the assumption of laminar flow based on computed Reynolds numbers caused the model to underestimate flows for small storms, while fitting large storms. In a nonurban setting, the complex topography that may exist in a basin may be poorly approximated by the idealized flow planes used by the model. It is likely that flow occurs in small rivulets and channels, and that the flow regime is turbulent. Converting the overland-flow regimes assumed by the model from laminar to turbulent significantly improved the model's performance for small storms, while large storms were unaffected.

Table 11.--Calibrated values of the Distributed-Routing Rainfall-Runoff Model soil-moisture accounting and infiltration parameters

[in. = inches, in/hr = inches per hour]

Station number	Parameters[1] PSP (in.)	KSAT (in/hr)	RGF	BMSN (in.)	EVC	RR
01398500	3.39	0.0648	29.9	5.99	1.03	0.87
01398107	3.88	.0475	23.9	3.83	.691	.87
01403400	3.89	.0499	30.0	5.58	1.03	.84
01403150	2.59	.0530	30.0	4.41	.551	.65
01400300	1.89	.0540	30.0	4.08	.725	.71
01401650	4.44	.0347	29.9	5.58	.934	.92
01402600	1.73	.0614	27.6	3.07	.792	.79
01403535	9.08	.0640	27.2	3.42	.816	.85

[1] Definitions of these parameters are given in table 1.

Estimation of Pressurized-Flow Parameters at Peter's Brook

The Peters Brook basin, station 01400300, is bisected by Interstate Highway 287. Several channels pass through culverts under this road. During model calibration, simulated flow in these culverts was insufficient to exceed their nonpressurized-flow capacities. During long-term simulation, however, the nonpressurized capacity was exceeded in three culvert segments. The volume of water stored was computed for each storm and segment. Table 12 summarizes the volumes stored behind the culvert segments for this initial simulation.

The culvert segment behind which the greatest amount of water was stored was segment PP05 (table 12). The maximum amount of water stored behind this segment was 92 million gallons (Mgal). An analysis of the topography behind the culvert revealed that at least 240 Mgal could be stored behind the culvert with up to 10 ft of pressure head. Because significant pressure head could develop for some storms, a representation of the segment was used that accounts for both nonpressurized and pressurized flow.

The first step in the analysis of pressurized flow in segment PP05 was to develop a relation between the pressurized-flow rate in the segment and the depth of water above the top of the culvert. Next, a relation was determined between the volume of water stored behind and in the culvert and the depth of water above the top of the culvert, based on the analysis of the topography above the culvert. By combining these relations, the relation between the volume of water in and behind the culvert, and the flow of water through the culvert was established. A modified Puls-routing (Soil Conservation Service, 1972) reservoir segment was defined using this relation.

Pipe segment PP05 was replaced by a simple channel segment placed downstream of the reservoir segment. The channel segment had routing parameters α and m that were computed for the original pipe segment. When flows are less than the nonpressurized flow capacity of the culvert, the flow is controlled by the channel segment. The ratio of volume to flow is so low for this range of flow that the reservoir segment has little or no effect. When flows exceed the culvert capacity, the flow is controlled by the reservoir segment. The storage in the reservoir segment increases rapidly for flows higher than the nonpressurized flow capacity of the culvert. This reduces the flows passed to the channel segment.

The simulated capacity of segment PP03 was exceeded during only one storm, but by a significant amount. An analysis of the topography behind this segment revealed that storage was insufficient to hold the simulated volume. In fact, any significant amount of water backed up behind the culvert would flow over the upstream channel banks toward a parallel segment, before enough pressure could develop in segment PP03 to increase its flow beyond the nonpressurized-flow capacity. DR3M does not provide a mechanism for changing flow paths in a network with changing flow conditions. The net effect of water bypassing segment PP03 is that the water will reach the basin outlet sooner than predicted by the model. To account for this effect, segment PP03 was replaced by a simple channel

Table 12.--Summary of culvert storages for selected storms in the Peters Brook basin

[ft^3/s = cubic feet per second; Mgal = million gallons]

Segment	Pipe capacity (ft^3/s)	Date	Mean storage (Mgal)	Maximum storage (Mgal)	Minutes capacity exceeded
PP02	509	55-08-07	0.40	0.57	15
PP03	603	55-08-07	3.2	5.9	35
PP05	365	19-07-18	.32	.54	20
		21-08-06	30.	51.	130
		26-08-12	.86	1.5	25
		32-09-02	.84	1.6	35
		32-11-01	1.0	1.9	35
		34-09-16	.19	.32	15
		39-06-29	18.	32.	100
		40-09-01	4.6	8.1	60
		55-08-07	38.	62.	120
		61-07-13	18.	32.	100
		65-07-17	.57	.94	25
		67-07-21	5.40	9.4	60
		71-08-26	24.	46.	150
		75-07-20	46.	92.	340

segment placed downstream of a modified Puls-routing (Soil Conservation service, 1972) reservoir segment. The channel segment had routing parameters α and m that were computed for the original pipe segment. The reservoir segment was identical to the segment used for culvert PP05.

A small amount of water was stored behind segment PP02 for only one storm (table 12). This segment was left unchanged.

Model Verification

A split-sample procedure was used to assess the reliability of parameter estimates. Four to five years of data were collected at each site. The storms were split chronologically into two sets of events. The first set was used to calibrate the model, whereas the second set was used for model verification. The storm-volume and peak-flow efficiencies for the calibration and verification periods are summarized in table 13. Plots of observed versus simulated volumes and efficiencies are shown in figures 10 through 25.

In general, the performance of volume simulation was poorer during the verification period than during the calibration period. The performance of peak simulation was similar during both periods. Because the model was used to synthesize long-term peaks, verification of the model for volume simulation is not critical. Because the model performed similarly in the verification and calibration periods for peaks, the model is adequate for simulation of peak data.

Figures 26 to 31 show typical simulated and observed hydrographs for large storms at a large basin (North Branch Raritan River, station 01398500), a medium-sized basin (Pike Run, station 01401650), and a small basin (Stoney Brook, station 01403535).

Sensitivity of Calibration Performance to Use of Synthetic Evaporation Data

To evaluate the effects of using synthetic evaporation estimates rather than pan-evaporation data, the rainfall-excess component of the model was calibrated using pan-evaporation data rather than synthetic pan-evaporation estimates. The calibration and verification efficiencies for volumes obtained by using pan-evaporation data and by using synthetic evaporation estimates are shown in table 14.

The efficiencies were similar whether synthetic evaporation estimates or pan-evaporation data were used. In most cases, model performance was better when synthetic evaporation data were used.

Calibration of the Base-Flow Model

The base-flow model parameters were estimated by the use of a Rosenbrock direct search procedure (Rosenbrock, 1960) to minimize the mean-square difference between computed and observed base-flow values, with respect to K_S, RR_{BF}, and EVC_{BF}. The parameter values obtained and the base-flow calibration efficiencies are summarized in table 15.

Table 13.--Calibration and verification efficiencies of storm-volume and peak-flow estimates

[Values are given in percent]

Station number	Storm-volume efficiencies		Peak-flow efficiencies	
	Calibration	Verification	Calibration	Verification
01398107	92	89	93	91
01398500	93	88	86	86
01400300	86	87	77	65
01401650	98	93	92	93
01402600	91	87	74	81
01403150	85	75	89	72
01403400	90	86	77	77
01403535	93	92	83	65

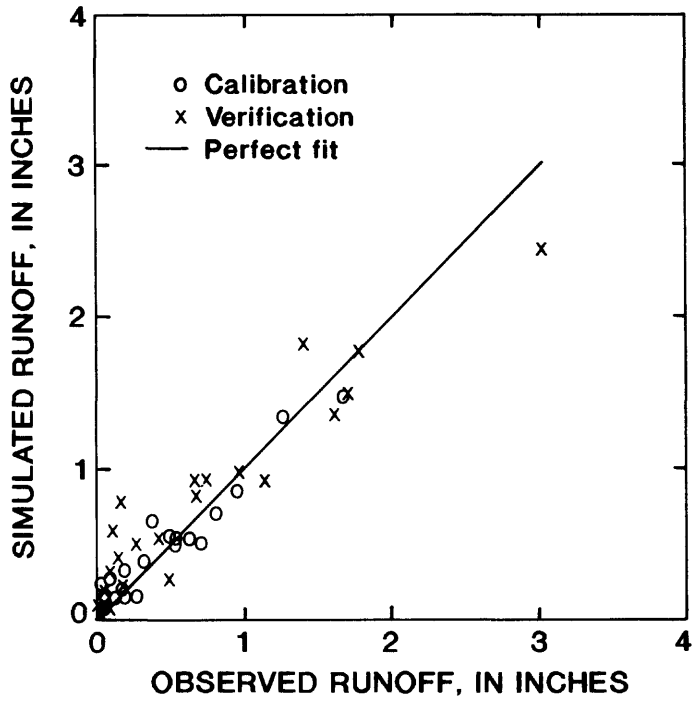


Figure 10.--Relation of simulated to observed runoff volumes at station 01398107.

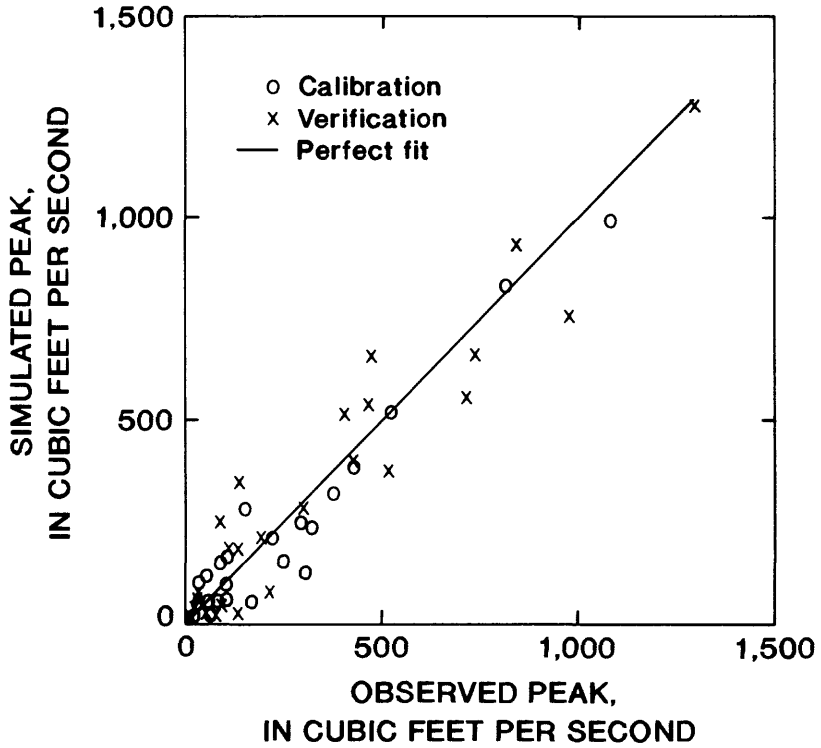


Figure 11.--Relation of simulated to observed peak flows at station 01398107.

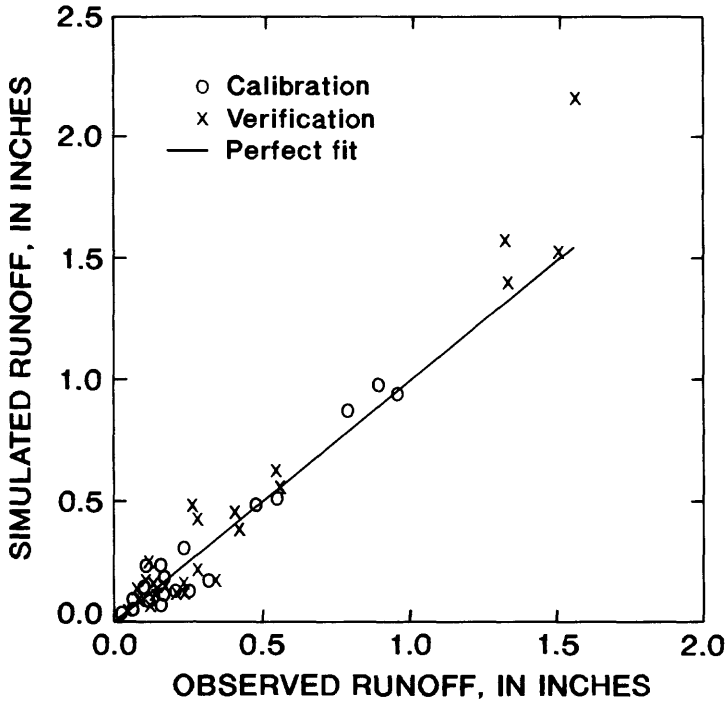


Figure 12.--Relation of simulated to observed runoff volumes at station 01398500.

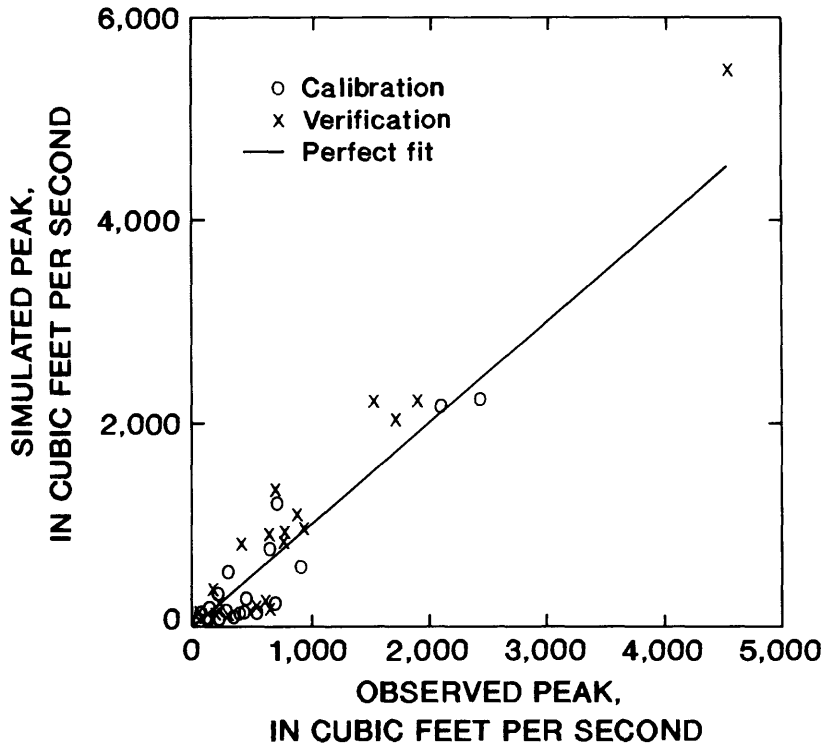


Figure 13.--Relation of simulated to observed peak flows at station 01398500.

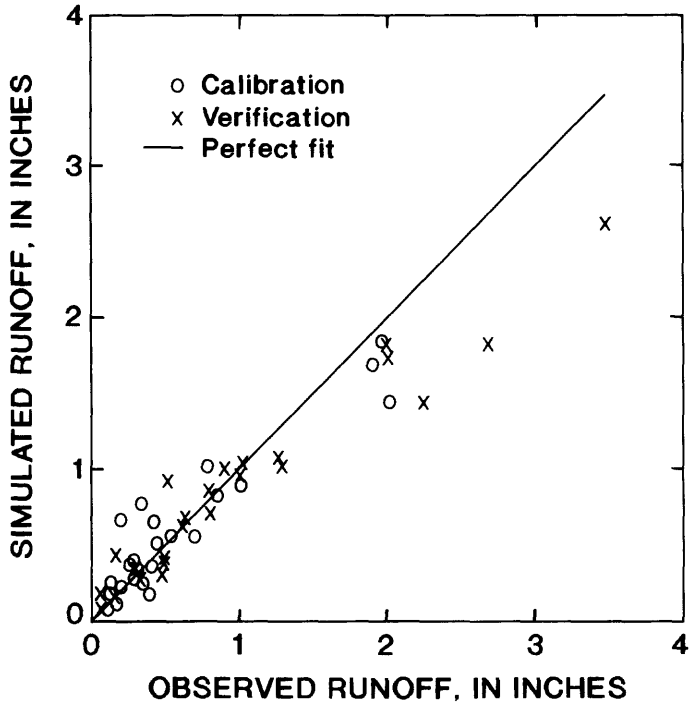


Figure 14.--Relation of simulated to observed runoff volumes at station 01400300.

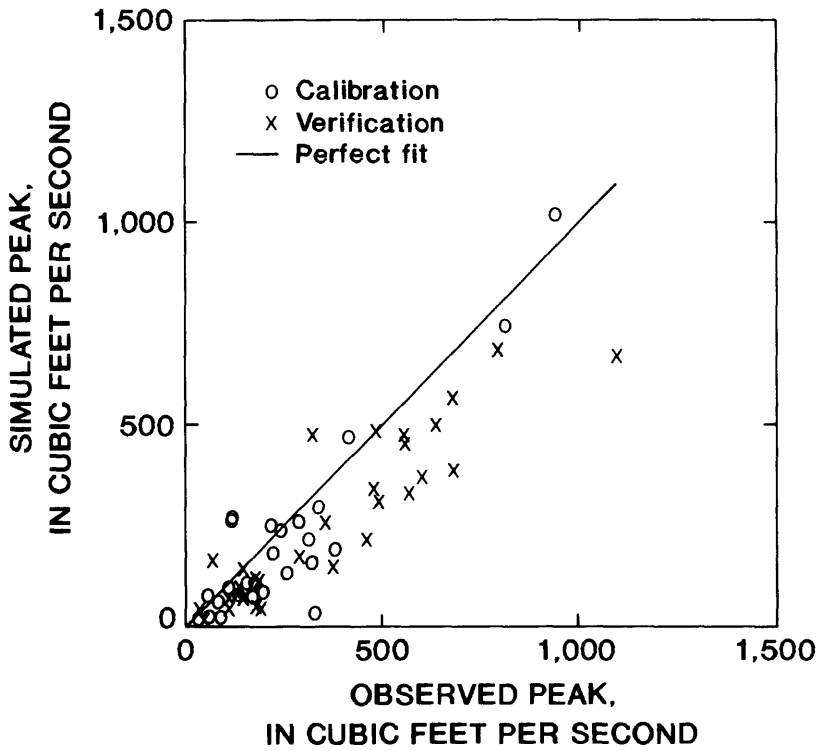


Figure 15.--Relation of simulated to observed peak flows at station 01400300.

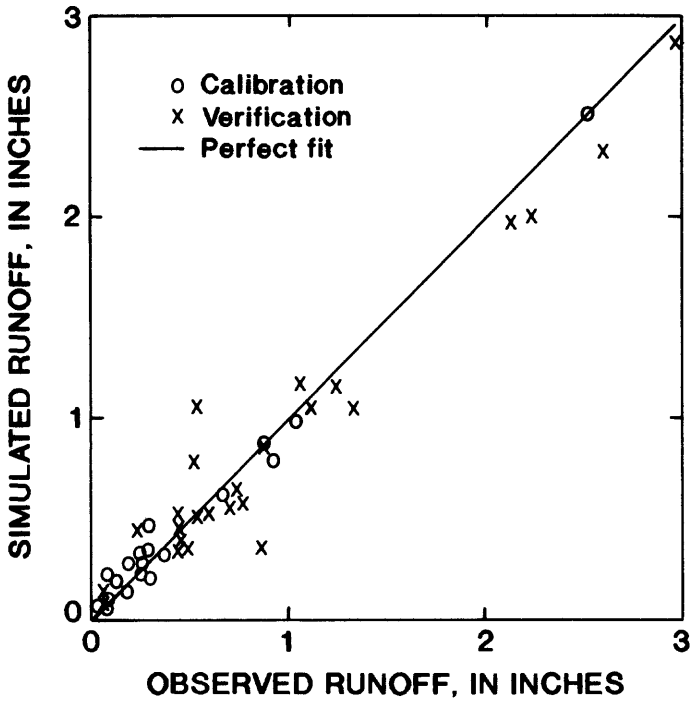


Figure 16.--Relation of simulated to observed runoff volumes at station 01401650.

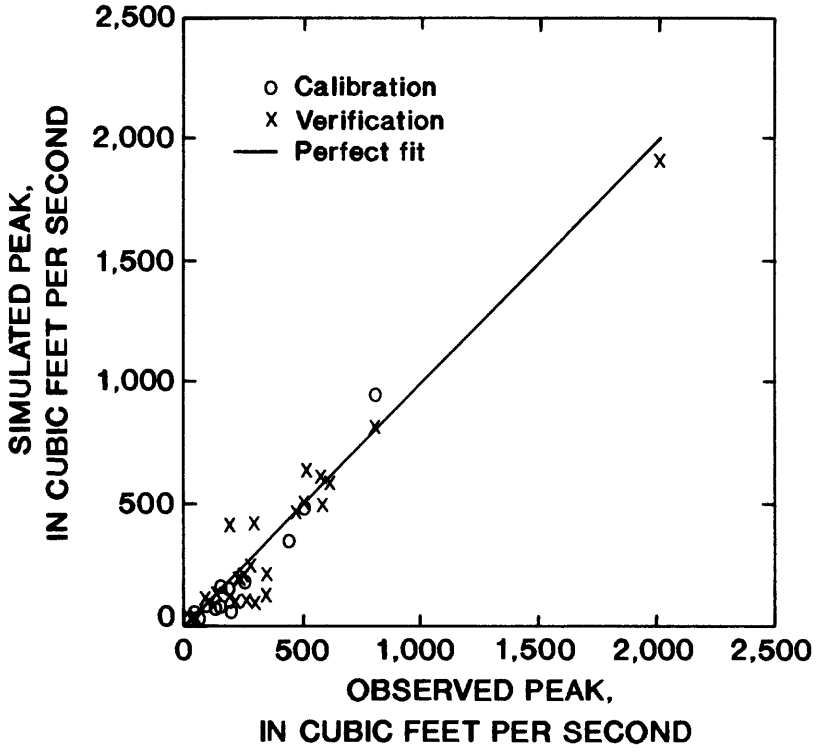


Figure 17.--Relation of simulated to observed peak flows at station 01401650.

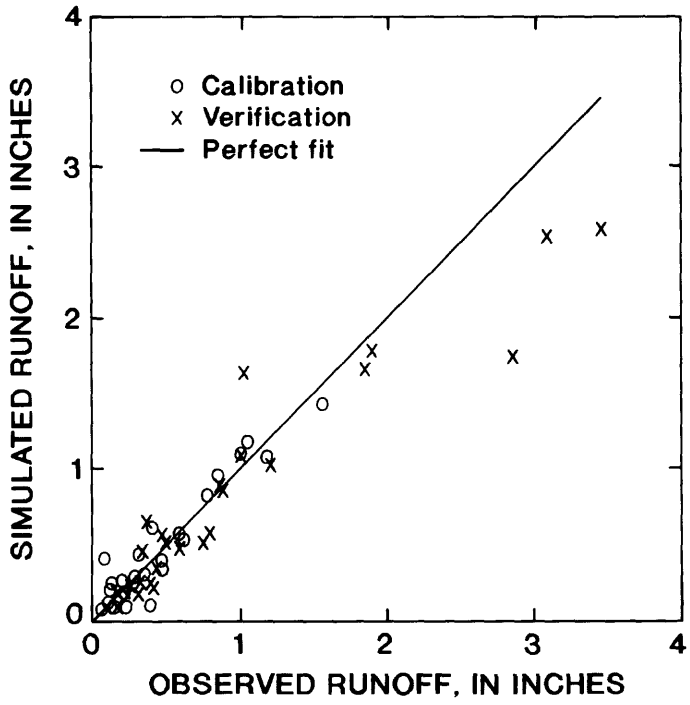


Figure 18.--Relation of simulated to observed runoff volumes at station 01402600.

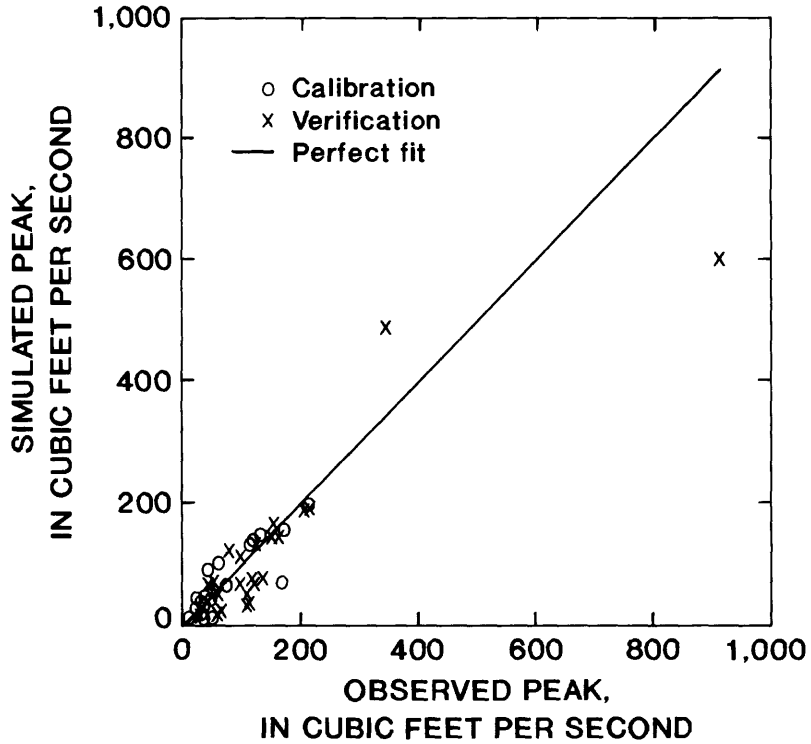


Figure 19.--Relation of simulated to observed peak flows at station 01402600.

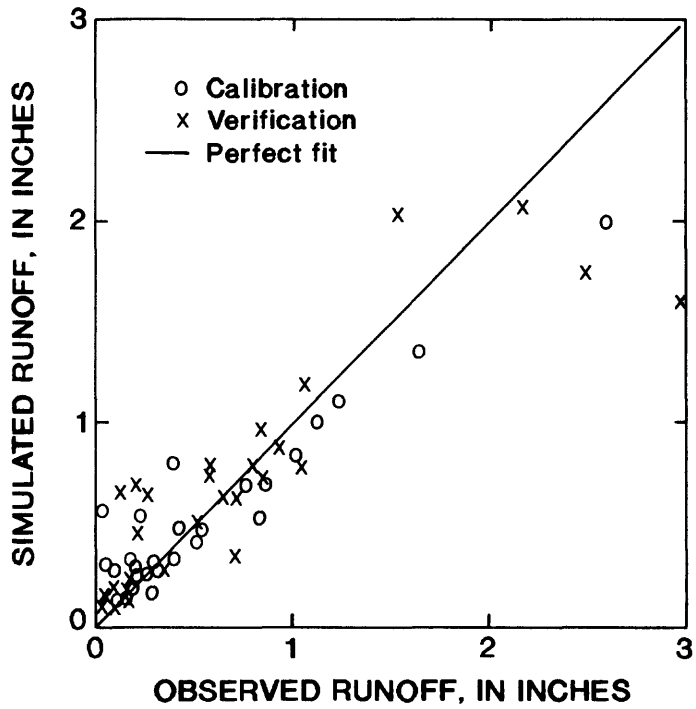


Figure 20.--Relation of simulated to observed runoff volumes at station 01403150.

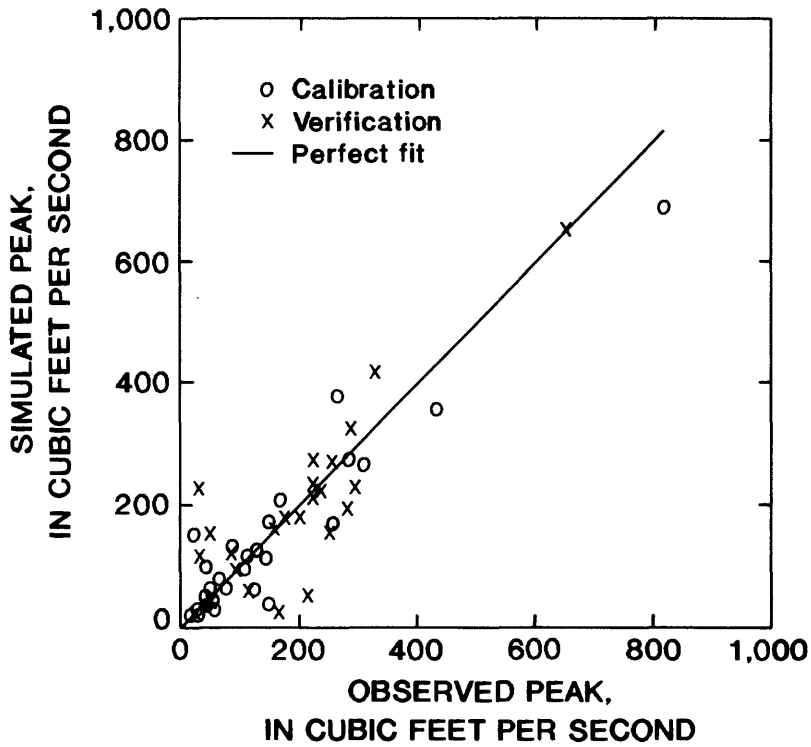


Figure 21.--Relation of simulated to observed peak flows at station 01403150.

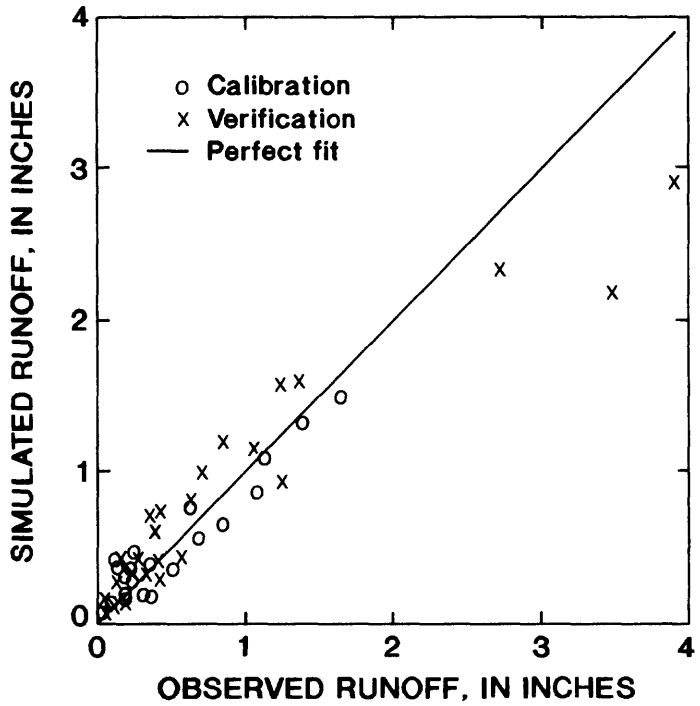


Figure 22.--Relation of simulated to observed runoff volumes at station 01403400.

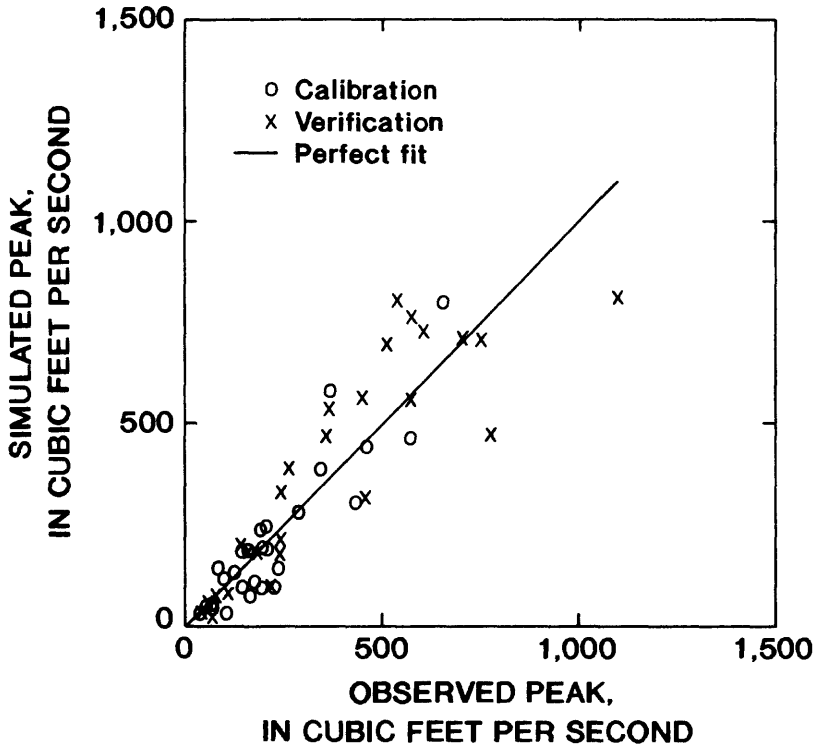


Figure 23.--Relation of simulated to observed peak flows at station 01403400.

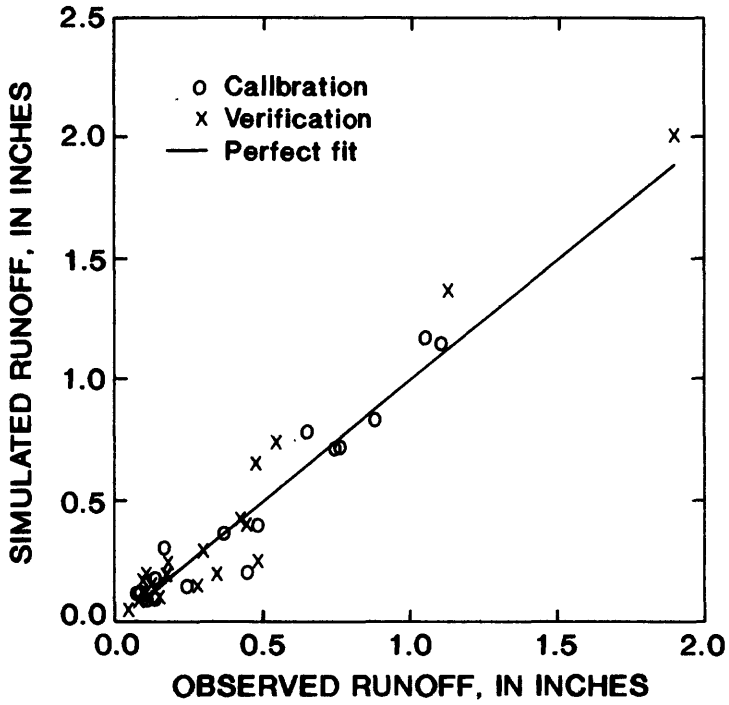


Figure 24.--Relation of simulated to observed runoff volumes at station 01403535.

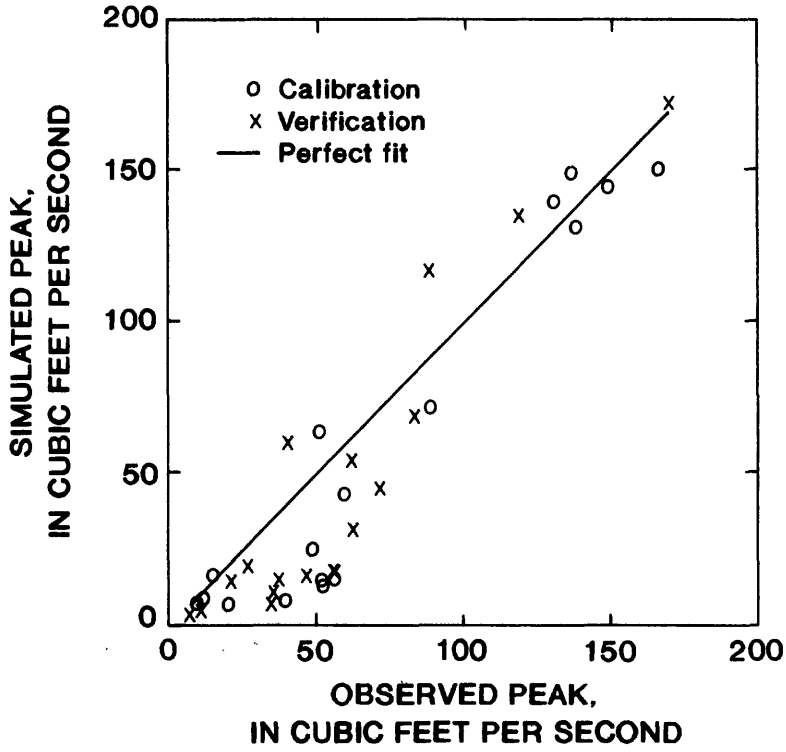


Figure 25.--Relation of simulated to observed peak flows at station 01403535.

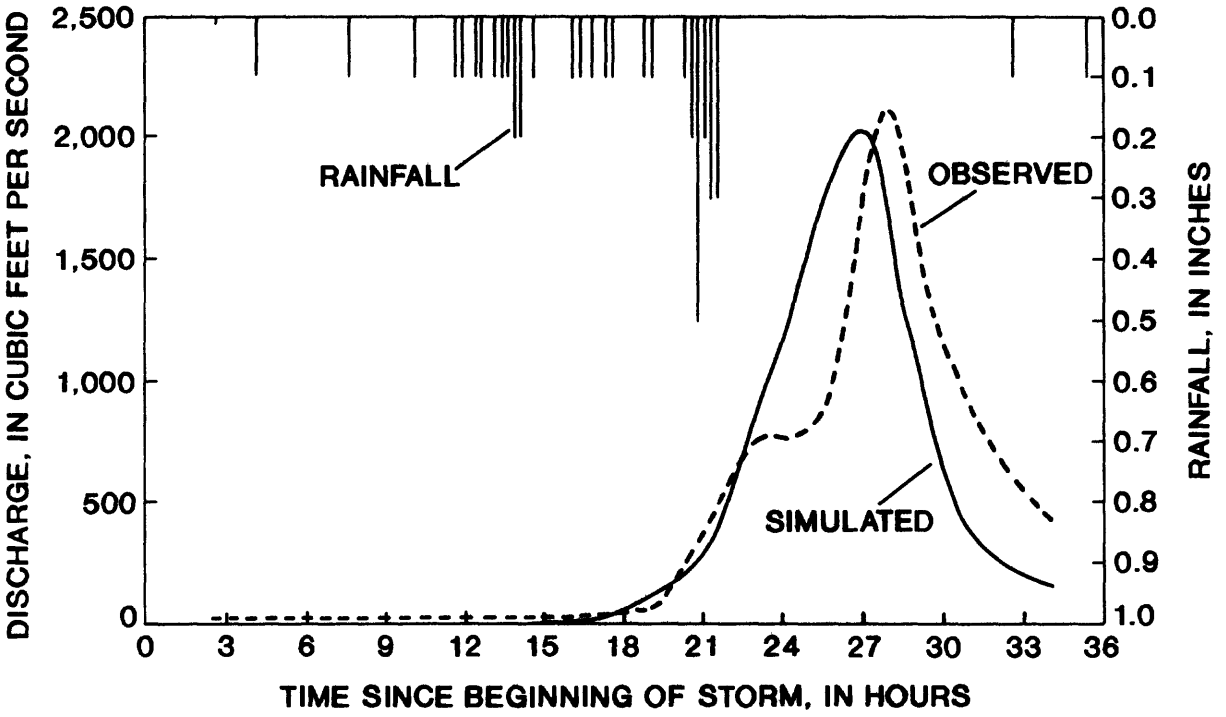


Figure 26.--Simulated and observed discharge for the May 11, 1981, storm at station 01398500.

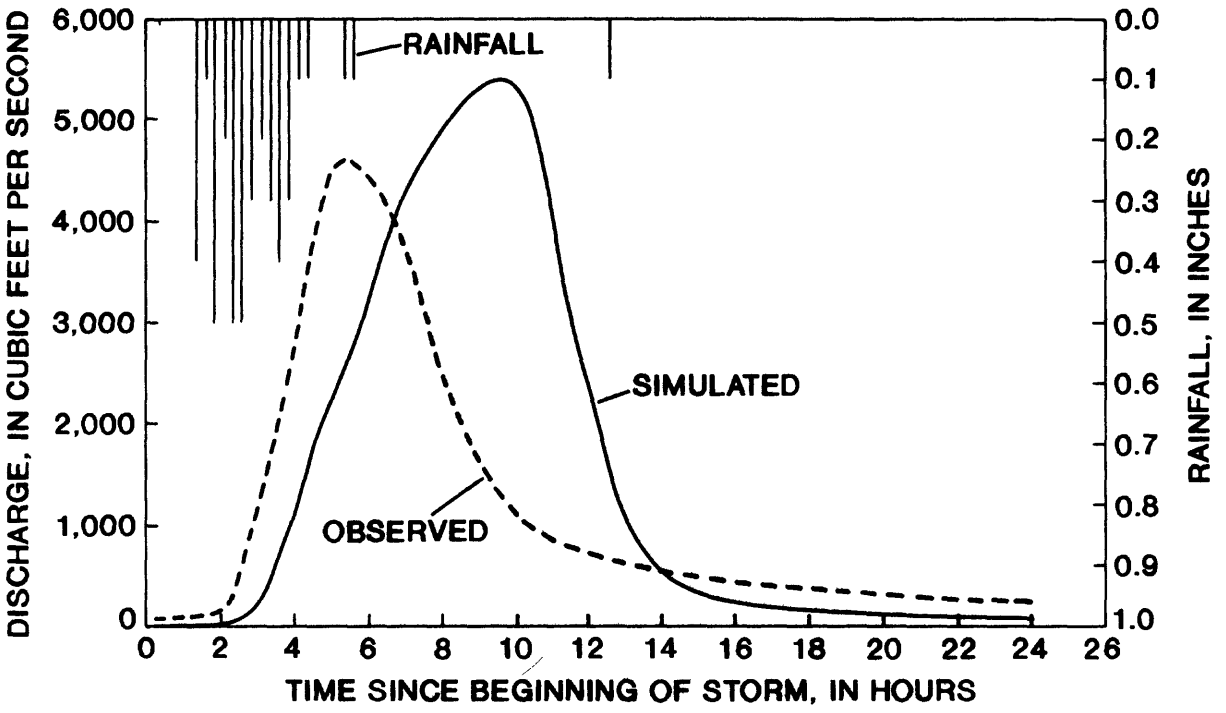


Figure 27.--Simulated and observed discharge for the July 7, 1984, storm at station 01398500.

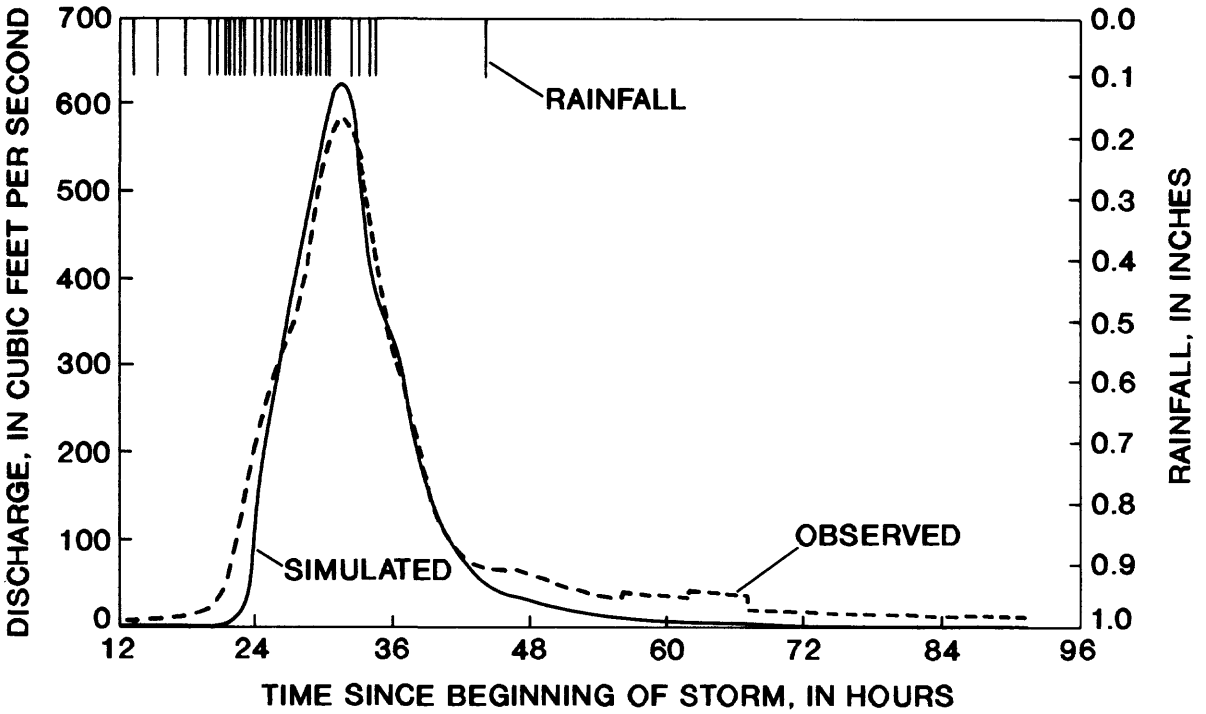


Figure 28.--Simulated and observed discharge for the April 4, 1984, storm at station 01401650.

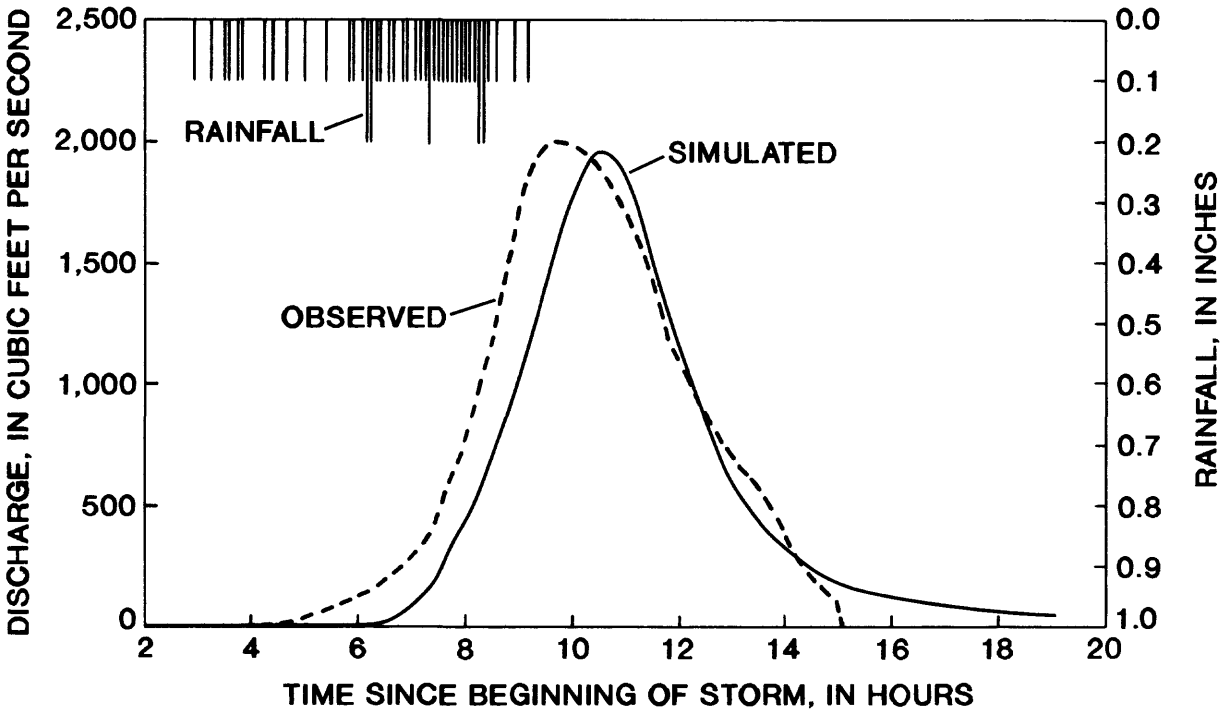


Figure 29.--Simulated and observed discharge for the July 7, 1984, storm at station 01401650.

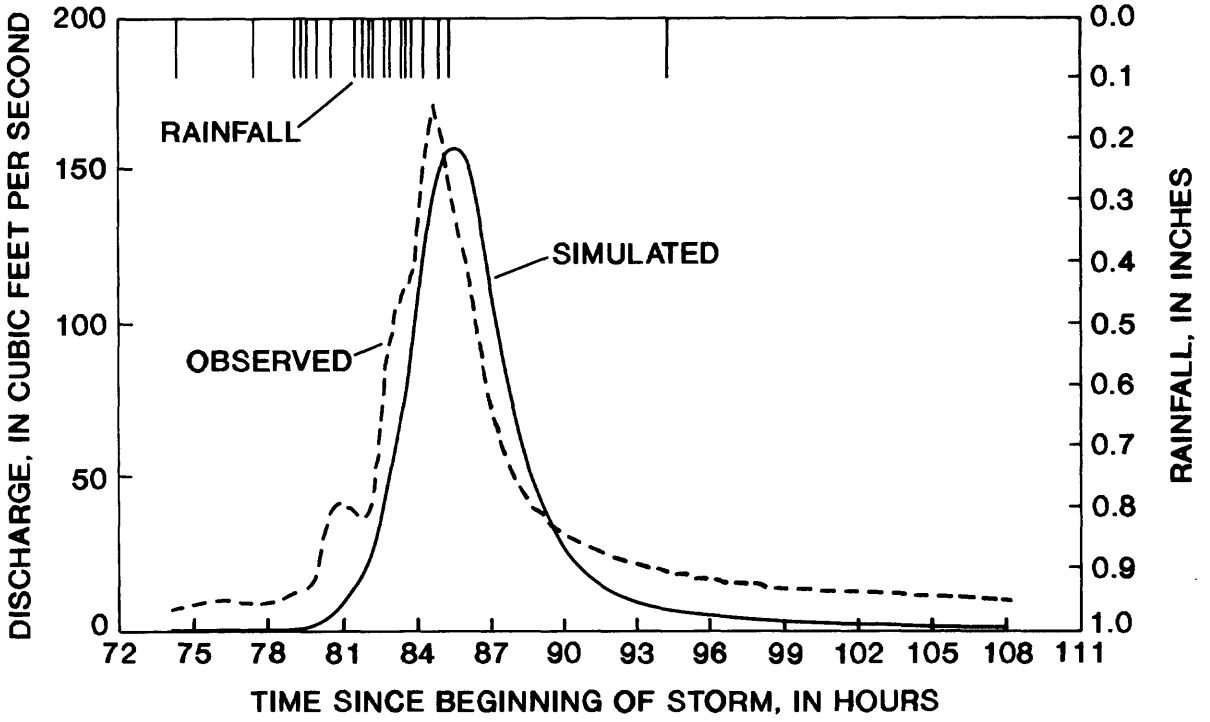


Figure 30.--Simulated and observed discharge for the April 10, 1983, storm at station 01403535.

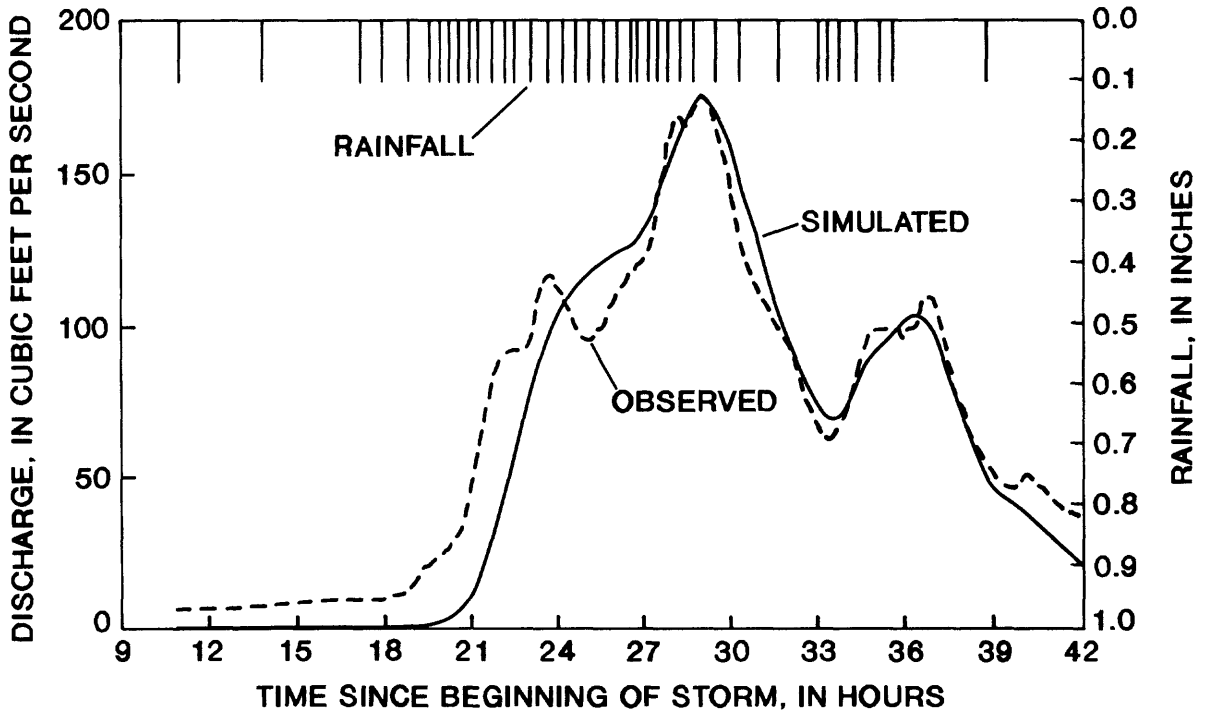


Figure 31.--Simulated and observed discharge for the April 4, 1984, storm at station 01403535.

Table 14.--Calibration and verification efficiencies for storm-runoff volumes using synthetic evaporation estimates and pan-evaporation data

[Values are given in percent]

Station number	Synthetic evaporation		Pan evaporation	
	Calibration	Verification	Calibration	Verification
01398107	92	89	94	79
01398500	94	89	91	42
01400300	87	87	85	86
01401650	98	93	96	92
01402600	91	87	91	91
01403150	85	75	83	75
01403400	90	86	84	89
01403535	93	92	93	86

Table 15.--Base-flow model parameter values and calibration efficiencies

[mi^2, square miles; K_S, base-flow recession constant; RR_{BF}, infiltration coefficient for base flow; EVC_{BF}, evaporation coefficient for base flow]

Station number	Drainage area (mi^2)	K_S	RR_{BF}	EVC_B	Calibration efficiency	Mean ratio of observed base flow to peak flow
01398107	9.00	0.11	0.53	0.44	57	0.048
01398500	26.2	.054	.80	.39	58	.10
01400300	4.19	.093	.29	.42	52	.0041
01401650	5.36	.11	.49	.55	72	.012
01402600	1.20	.11	.48	.61	45	.0075
01403150	1.99	.14	.27	.0015	35	.011
01403400	6.23	.10	.43	.32	47	.026
01403535	1.57	.064	.41	.22	51	.024

In general, the base-flow model efficiencies are low. However, because base flow is a small fraction of peak flow, as shown in table 15, accurate simulation of base flow is not critical to simulation of peak flow.

LONG-TERM SIMULATION AND FLOOD-FREQUENCY ANALYSIS

After DR3M and the base-flow model were calibrated for each of the eight basins, the models were used to derive long-term flood histories and frequencies. This required four steps:

1. The long-term data were assembled and prepared for input to the models.

2. DR3M was used to compute flood-runoff peaks from the long-term data.

3. The calibrated base-flow model was used to synthesize long-term base-flow estimates. To estimate long-term flood-peak flows, the synthesized long-term base flows were added to the long-term runoff peaks.

4. Flood-flow frequency estimates were developed using the synthetic flood peaks in combination with observed flood data at each basin.

Selection and Processing of Long-Term Simulation Data

Three types of data are required for long-term simulation with DR3M: daily precipitation, daily pan-evaporation, and short-time-interval precipitation data for storms. Long-term daily precipitation for Trenton, New Jersey, NOAA station number 288883, were obtained from the WATSTORE daily value files (Hutchinson, 1975). Long-term daily pan-evaporation data were derived using a sinusoidal function, as described earlier in this report.

Long-term 5-minute precipitation data at Trenton, New Jersey, NOAA station number 288883, were obtained from the WATSTORE unit values files (Hutchinson, 1975). The data set consists of 197 storms chosen from water years 1914 through 1977 (Bailey and others, 1989).

Most of the storms in the data set occurred during summer. Analysis of peak-flow data in Somerset County indicated that a large proportion of the actual annual peaks occur during spring and winter. The spring and winter peaks resulted from storms that had lower precipitation volumes and intensities than storms during the summer on average, but that occurred when antecedent soil moisture was high.

Additional 5-minute precipitation data at Trenton were obtained from the National Climatic Data Center in Nashville, North Carolina, to provide data from the spring and winter and to extend the period of record through 1979. The choice of these storms was based on consideration of estimated antecedent soil-moisture and precipitation volumes. The final data set consisted of 448 storms, spanning the water years 1914 through 1979.

The location of the Trenton rain gage relative to Somerset County is shown in figure 32. Also shown in this figure are the locations of several

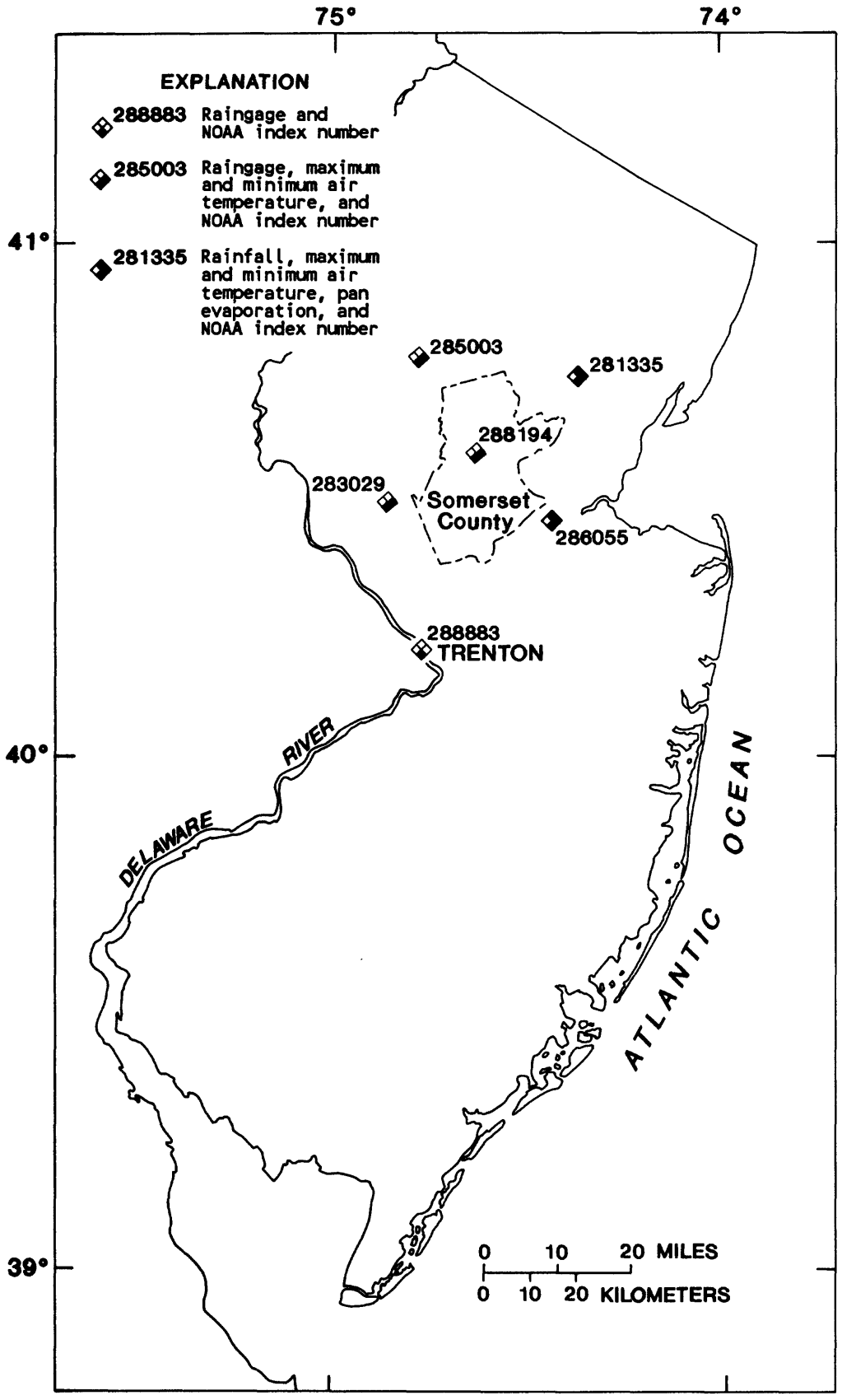


Figure 32.--Location of Trenton rain gage and five National Oceanic Atmospheric Administration rain gages in and near Somerset County, New Jersey.

NOAA rain gages. The data from these gages were used to evaluate differences in precipitation between the Trenton and Somerset County areas.

Adjustment of Long-Term Precipitation Data to Account for Areal Difference

Long-term daily precipitation data at five sites in and near Somerset County were obtained from the National Climatic Data Center in Asheville, North Carolina (table 16). The locations of the sites relative to the Trenton rain gage and to Somerset County are shown in figure 32. The period of record is from 1927 through 1983.

The long-term daily precipitation data at Trenton were compared to the data at four of these sites--Canoe Brook, Flemington, Somerville, and Long Valley--to investigate differences between Trenton and Somerset County precipitation. Data from New Brunswick were not used because the period of record at this site was much shorter than at the other sites.

Two statistics were computed each year for each station for the period of record: log annual maximum storm volume and log annual total precipitation. The annual maximum storm volume is the largest storm volume in each year. It is representative of precipitation amounts for the largest storms and was used to investigate bias in storm precipitation. The annual total precipitation is representative of rainfall volumes in general, and was used to investigate bias in daily precipitation amounts. The statistics were computed on a water-year basis. If there were less than 360 days of record in a year, then the statistics for that year were not computed. The statistics for the four stations in and near Somerset County were averaged for each year for which all of the stations had data (32 years). The averaged statistics for these stations were compared to the statistics for Trenton for the same years.

The annual total precipitation and annual maximum storm-volume data at the Trenton rain gage and the averaged data for stations in and near Somerset County are plotted in figures 33 and 34. The values of both statistics averaged for stations in and near Somerset County exceed the values computed for the Trenton rain gage in almost every year.

Paired T-tests were performed on the log statistics to assess the significance of the differences observed in the annual total precipitation and in the annual maximum storm volumes. The results of these T-tests are summarized in table 17. The values of both statistics were significantly greater in Somerset County than at Trenton. The difference was larger for log annual total precipitation than for log annual maximum storm volumes.

The long-term data from Trenton were adjusted to correct for the differences in precipitation between Trenton and Somerset County; the daily precipitation values were increased by 15 percent, and the storm-precipitation values were increased by 13.5 percent.

Aggregation of Long-Term 5-Minute Precipitation Data

The recording interval for the precipitation data used to calibrate DR3M is 0.1 inch, whereas the recording increment for the long-term precipitation data is 0.05 inch. The calibrated model parameters partially compensate for rainfall-intensity errors introduced by the 0.1-inch recording increment. If

Table 16.--Precipitation stations in or near Somerset County

Station	NOAA[1] station number	Latitude	Longitude	Elevation (feet above sea level)	Period of record
Canoe Brook	281335	40°45′	74°21′	180	1932 - 1983
New Brunswick	286055	40°28′	74°26′	86	1969 - 1983
Flemington	283029	40°30′	74°52′	180	1927 - 1983
Somerville	288194	40°36′	74°38′	160	1932 - 1983
Long Valley	285003	40°47[1]	74°47[1]	550	1933 - 1983

[1] National Oceanic and Atmospheric Administration

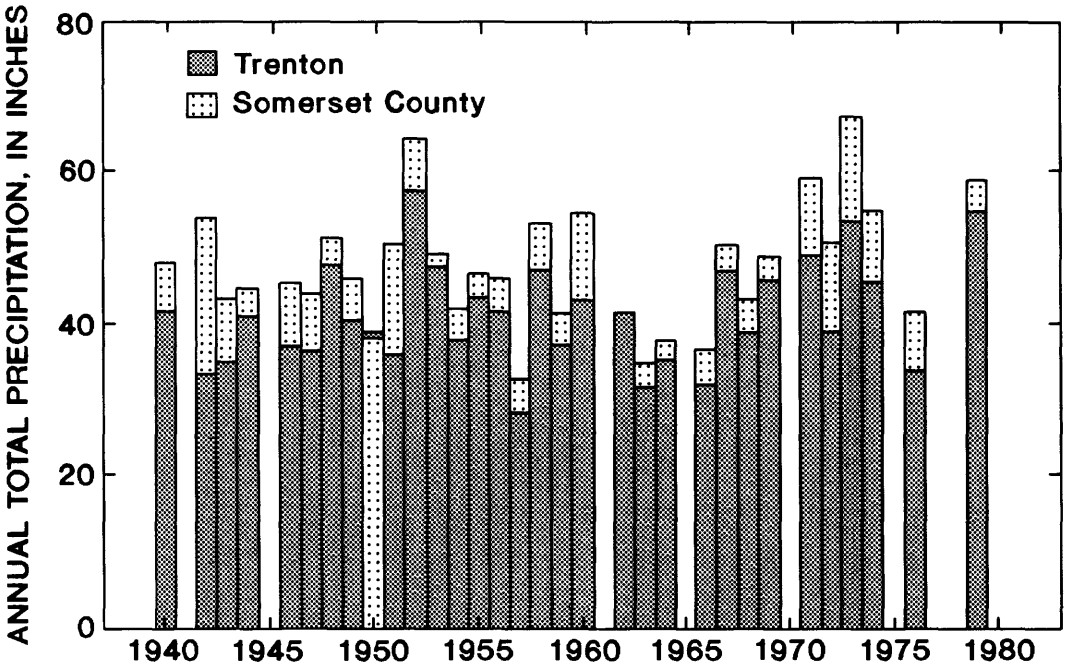


Figure 33.--Annual precipitation at Trenton and average annual precipitation for stations in and near Somerset County, New Jersey. (Only years with at least 360 days of record at all stations are shown.)

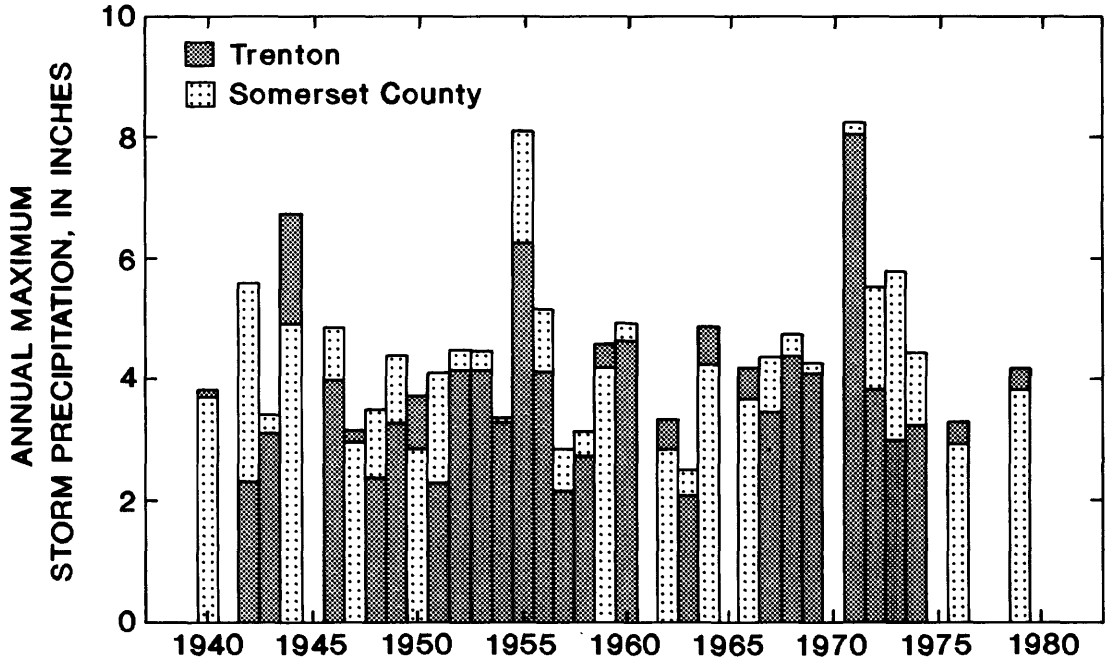


Figure 34.--Annual maximum storm precipitation at Trenton and average annual maximum storm precipitation for stations in Somerset County, New Jersey. (Only years with at least 360 days of record at all stations are shown.)

Table 17.--Results of T-test for differences in base 10 log precipitation between Trenton rain-gage data and mean data for rain-gage sites in the Somerset County area

[<, less than]

	Mean		Mean difference	Percent difference[1]	T statistic	Tail probability
	Trenton	Somerset County				
Log annual total precipitation	1.609	1.671	.062	15.3	8.06	<0.001
Log annual maximum storm volume	.565	.620	.055	13.5	2.72	.011

[1] The percent difference is the antilog of the mean difference. It is an estimate of the average difference between precipitation at Somerset County and Trenton, as a percentage of the precipitation at Trenton.

the 0.05-inch data is used for long-term simulation, then the results will be biased, because the model will compensate for errors that are not present. To avoid this bias, a 0.1-inch increment long-term precipitation-data record was computed by aggregating the 0.05-inch long-term precipitation-data record, after adjusting the record for bias.

An analysis was made to evaluate the impact of the precipitation-recording increment on the simulated flood peaks. DR3M was used to simulate long-term flood-peak flows, using both 0.05- and 0.1-inch increment precipitation data. Table 18 lists the logarithmic mean, standard deviation, and skew of annual floods generated by using 0.1- and 0.05-inch long-term precipitation data. A paired T-test was performed for each of these statistics to evaluate the statistical significance of any differences observed because of the choice of the precipitation-data recording interval. The results of these T-tests are summarized in table 19.

The estimated mean annual floods were lower when 0.05-inch-increment precipitation data were used than when 0.1-inch-increment precipitation data were used. The magnitude of the differences between the estimates of the annual mean ranged from 3 to 10 percent, after conversion from log units. The paired T-test indicated that these differences were significant at the 95-percent probability level. The mean annual flood estimates obtained by using 0.05-inch-increment precipitation data are lower than the estimates obtained by using 0.1-inch-increment precipitation data because the precipitation intensities are lower, infiltration is higher, and less direct runoff is produced, on average, when 0.05-inch-increment precipitation data are used.

The estimated standard deviation of annual floods was higher when 0.05-inch-increment precipitation data were used than when 0.1-inch-increment precipitation data were used. The magnitudes of the differences in this case ranged from 2 to 7 percent, after conversion from log units. The paired T-test performed indicated that these differences were significant at the 95-percent probability level.

Table 20 lists long-term flood-frequency estimates for 2-, 10-, and 100-year recurrence intervals (Q_2, Q_{10}, and Q_{100}), estimated by using 0.1-inch-increment and 0.05-inch-increment long-term precipitation data. A paired T-test was performed to evaluate the statistical significance of any differences observed in flood-frequency estimates obtained by using 0.05-inch- and 0.1-inch-increment precipitation data (table 21).

At the 2-year recurrence interval (Q_2), which corresponds to the median flood, the estimates derived by using 0.05-inch-increment precipitation data were 3 to 10 percent lower than those derived using 0.1-inch-increment precipitation data. The paired T-test indicated that these differences were significant at the 95-percent probability level. This result is similar to the result for the mean log flood.

As the recurrence interval increases, the difference between estimates derived from 0.05- and 0.1-inch-increment data decreases, until some recurrence interval is reached at which the estimates derived by using 0.05-inch-increment precipitation exceeds the estimate derived by using 0.1-inch-increment precipitation data. At the 100-year recurrence interval, flood-flow estimates tend to be higher when derived with 0.05-inch-increment

Table 18.--Statistics for long-term annual, logarithmically (base 10) transformed floods simulated by using precipitation data with 0.1-inch and 0.05-inch recording increments

Station number	0.1-inch increments			0.05-inch increments		
	Mean	Standard deviation	Skew	Mean	Standard deviation	Skew
01398107	2.97	0.21	-0.51	2.93	0.24	-0.77
01398500	3.21	.33	-.72	3.20	.35	-.70
01400300	2.95	.25	-.92	2.93	.26	-.88
01401650	2.87	.30	-.73	2.84	.32	-.64
01402600	2.48	.26	-.62	2.45	.28	-.70
01403150	2.84	.25	-.53	2.80	.28	-.70
01403400	2.85	.30	-.46	2.82	.32	-.47
01403535	2.31	.36	-.75	2.25	.38	-.62

Table 19.--Results of two-tailed paired T-test analysis for differences between simulated flood statistics derived by using 0.1-inch and 0.05-inch precipitation data

[s.d. is standard deviation; <, less than]

Statistic	Mean		Mean difference	Percent difference	T statistic	Tail probability
	0.1-inch	0.05-inch				
Log mean	2.809	2.777	-0.032	-7.069	-6.92	<0.001
Log s.d.	.282	.303	.021	4.930	10.76	<.001
Log skew	-.655	-.685	-.030	-6.782	-.64	.544

Table 20.--Long-term annual estimated return flows estimated by using precipitation data with 0.1-inch and 0.05-inch recording increments

Station number	0.1-inch increments			0.05-inch increments		
	Q_2	Q_{10}	Q_{100}	Q_2	Q_{10}	Q_{100}
01398107	940	1,690	2,640	880	1,660	2,660
01398500	1,700	4,250	8,190	1,640	4,280	8,530
01400300	936	1,770	2,770	880	1,760	2,860
01401650	767	1,690	3,100	710	1,700	3,220
01402600	306	630	1,060	290	630	1,090
01403150	698	1,420	2,420	660	1,410	2,440
01403400	730	1,710	3,270	680	1,680	3,340
01403535	213	550	1,120	180	530	1,170

Table 21.--Results of two-tailed paired T-test analysis for differences between simulated log return flows with 2-, 10-, and 100-year recurrence intervals derived by using 0.1-inch and 0.05-inch precipitation data

Statistic	Mean		Mean difference	Percent difference	T statistic	Tail probability
	0.1-inch	0.05-inch				
Q_2	2.823	2.792	-0.031	-6.886	-5.71	0.001
Q_{10}	3.158	3.154	-.004	-.848	-1.59	.157
Q_{100}	3.405	3.417	.012	2.792	5.41	.001

precipitation data than when derived with 0.1-inch-increment precipitation data. The differences, though small, were statistically significant at the 95-percent probability level.

The tendency for return-flood estimates at high recurrence intervals to be higher when 0.05-inch-increment precipitation data are used than when 0.1-inch-increment precipitation data are used is surprising considering that less intense precipitation usually yields less runoff. The explanation for this tendency is that, although the average flood is smaller when 0.05-inch-increment precipitation data are used, the variance of the annual flood flows is greater, thereby increasing the values of the estimates of the high-recurrence-interval floods.

These analyses indicate that the choice of recording increment for the precipitation data used for long-term simulation has a significant impact on the flood data synthesized by DR3M and on the flood frequencies estimated from the synthetic flood data. If the recording increment of the data used for long-term simulation differs from the recording increment of the data used to calibrate the model, then biased results can be expected. Therefore, 0.1-inch-increment, long-term precipitation data, obtained by aggregating 0.05-inch-increment data, were used to synthesize the flood peaks and to develop flood-frequency estimates for this study after adjusting for areal bias.

Preservation of Variance in Simulated Data

Simulated flood peak flows often have been found to have less variance than observed flood peak flows. Kirby (1975) explained this loss of variance as an inherent byproduct of model simulation that can affect any simulation data. Because of the sensitivity of flood-frequency estimates to the variance of flood peak flows, a bias in the variance of simulated flood data could significantly affect the accuracy of the resulting flood-frequency estimates.

Analyses were performed to assess whether there was a bias in variance of flood peak flows simulated in this study. For each of the basins studied, F-tests were used to determine whether variances of simulated and observed peak flows were different from one another during the model-verification period. Tests were performed on untransformed, as opposed to log-transformed, data to give more emphasis to the larger storms, which are indicative of maximum annual floods. No significant differences (at the 5-percent probability level) were found, although the variances of the simulated data were higher on average than the variances of the observed data.

Three of the basins, North Branch Raritan River (station 01398500), Royce Brook tributary (station 01402600), and Green Brook (station 01403400) had long observed-flood records that overlapped the simulated records. At each of these stations, the variances of the log-transformed annual maximum flood peak flows for the years for which the simulated and observed records overlapped were computed. The annual peaks were log-transformed because the flood-frequency estimates were to be based on the log-transformed peak flows. F-tests were used to determine whether the observed and simulated annual peak flow variances were different from one another during the

periods of overlapping flood records. No significant diffferences (at the 5-percent probability level) were found.

Because no evidence of bias in simulated flood peak flow variances was found, no adjustments were applied to the simulated flood peaks to correct for loss of variance.

Flood-Frequency Analysis

At each basin, a series of synthetic flood peaks was developed by adding the long-term runoff peaks computed by DR3M to base-flow estimates obtained from the base-flow model. A log-Pearson Type III frequency curve was fitted to each series of synthetic flood peaks in accordance with Interagency Committee on Water Data (1982) recommendations. The flood-frequency estimates for the 2-, 5-, 10-, 25-, 50-, 100-, and 500-year recurrence intervals (Q_2, Q_5, Q_{10}, Q_{25}, Q_{50}, Q_{100}, and Q_{500}) are summarized in table 22.

At three of the basins, North Branch Raritan River (station 01398500), Royce Brook (station 01402600), and Green Brook (station 01403400), the observed peak-flow record was sufficiently long to estimate a log-Pearson Type III frequency curve based on observed data. At these sites, a weighted frequency curve was developed by weighting and combining the flood-frequency estimates based on the synthetic flood series with flood-frequency estimates based on observed data. The weighting equation used is:

$$Q_{weighted,T} = \frac{N_{obs}\, Q_{obs,T} + N_{sim}\, Q_{sim,T}}{N_{obs} + N_{sim}}, \quad (15)$$

where

$Q_{weighted,T}$ - the weighted estimate of the flood with a recurrence interval of T years, in cubic feet per second,

N_{obs} - the length of the observed flood series, in years,

$Q_{obs,T}$ - the estimate, based on observed flood data, of the flood with a recurrence interval of T years, in cubic feet per second,

N_{sim} - the length of the simulated flood series, in years, and

$Q_{sim,T}$ - the estimate, based on synthetic flood data, of the flood with a recurrence interval of T years, in cubic feet per second.

The observed peak data for stations 01398500, 01402600, and 01403400 included historical peaks. For these stations, the historical period of record was used for the N_{obs} in equation 15. The weighted flood-frequency estimates are given in table 23.

The remaining five basins had from 4 to 5 years of observed peak data. These peaks were contained in the data for water years 1980 through 1984 used to calibrate DR3M. This period does not overlap with the period 1914-79, during which the long-term data were collected. Because the observed peaks occurred during a period that was separate from the long-term simulation period, it is reasonable to assume that the observed and simulated flood data are independent. For this reason, the simulated and observed peak-flow series for these stations were concatenated. Log-Pearson Type III flood-frequency estimates were developed based on the combined simulated and

Table 22.--Flood-frequency estimates from long-term synthesis

Station number	Drainage area (square miles)	Discharge (Q_2)	(Q_5)	(Q_{10})	(Q_{25})	(Q_{50})	(Q_{100})	(Q_{500})
		(cubic feet per second)						
01398107	9.00	940	1,390	1,690	2,070	2,360	2,640	3,290
01398500	26.2	1,700	3,150	4,250	5,770	6,960	8,190	11,200
01400300	4.19	936	1,440	1,770	2,180	2,480	2,770	3,420
01401650	5.36	767	1,300	1,690	2,230	2,660	3,100	4,210
01402600	1.20	306	494	626	796	925	1,060	1,360
01403150	1.99	698	1,120	1,420	1,820	2,120	2,420	3,160
01403400	6.23	730	1,290	1,710	2,300	2,770	3,270	4,530
01403535	1.57	213	403	553	765	936	1,120	1,580

Table 23.--Weighted flood-frequency estimates at stations 01398500, 1402600, and 01403400

Station number	Type of estimate	Length of record (years)	Discharge (Q_2)	(Q_5)	(Q_{10})	(Q_{25})	(Q_{50})	(Q_{100})	(Q_{500})
			(cubic feet per second)						
01398500	Synthetic	66	1,700	3,150	4,250	5,770	6,960	8,190	11,200
	Observed	66*	1,350	2,320	3,140	4,420	5,560	6,880	10,800
	Weighted		1,530	2,740	3,700	5,090	6,260	7,530	11,000
01402600	Synthetic	66	306	494	626	796	925	1,060	1,360
	Observed	13	250	420	583	863	1,140	1,490	2,690
	Weighted		297	482	619	807	961	1,130	1,580
01403400	Synthetic	66	730	1,290	1,710	2,300	2,770	3,270	4,530
	Observed	88*	848	1,590	2,220	3,160	3,980	4,890	7,460
	Weighted		797	1,460	2,000	2,790	3,460	4,200	6,200

* The observed records for these stations include historic peaks. The lengths of record shown and used to compute weighted estimates are for historic periods.

observed data at these six basins. The observed data for station 01403535 included one historical peak from 1973. An adjustment was made to the flood-frequency estimates for this station to account for the historic peak (Interagency Committee on Water Data, 1982). The flood-frequency estimates are summarized in table 24.

This study was undertaken in part because the flood-frequency estimates for Somerset County obtained by using the regional equations given by Stankowski (1974) are too low. Figure 35 compares the flood-frequency estimates derived in this study with estimates computed by using Stankowski's (1974) equations. As expected, the peaks estimated in this study are significantly higher than those estimated by using the equations given by Stankowski (1974).

Length of Record for Use in Regional Flood-Frequency Analysis

The weighted flood-frequency estimates from table 23 and the flood-frequency estimates from table 24 can be used in a regional flood-frequency analysis. However, in a regional analysis, the length of record associated with each estimate must be known, so that the estimate may be properly weighted relative to other estimates used in the analysis. The length of record at each station can be computed by adding the lengths of record of the observed and simulated flood data. However, the length of record of the simulated data should be adjusted to account for the fact that all eight simulated flood records are based on a single precipitation record and therefore are highly correlated. This adjustment can be accomplished by assuming that the information contained in the precipitation record is distributed evenly among the eight flood-frequency estimates, and by dividing the length of record of the simulated data by eight. Therefore, the adjusted total length of record associated with the flood-frequency estimates is computed as 8.25 years (66 years of precipitation data divided by 8) plus the length of record of the observed data at each station. The observed and total lengths of record are summarized for all eight stations in table 25.

SUMMARY

Rainfall-runoff data were collected at eight basins in Somerset County, New Jersey. The drainage areas of the basins ranged from 1.20 to 26.2 mi^2. Impervious area in the basins ranged from 3.2 to 13 percent of the total basin area.

The U.S. Geological Survey Distributed-Routing Rainfall-Runoff Model was calibrated for each of the eight basins. A split-sample analysis was used to verify the model. The data were split chronologically into two parts. The first part of the data set was used to calibrate the model; the second part was used to verify the calibrated model. Verification efficiency ranged from 75 to 93 percent for storm volumes, and from 65 to 93 percent for peak flows.

An evaporation model was used to provide evaporation estimates for model calibration and long-term simulation. During model calibration, the use of synthetic evaporation data, developed using the model, yielded improved model performance than did the use of actual pan-evaporation data.

Table 24.--Flood-frequency estimates from synthetic and observed peaks

Station number	Drainage area (square miles)	Discharge (Q_2)	(Q_5)	(Q_{10})	(Q_{25})	(Q_{50})	(Q_{100})	(Q_{500})
		(cubic feet per second)						
01398107	9.00	948	1,380	1,670	2,020	2,280	2,530	3,110
01400300	4.19	936	1,420	1,730	2,110	2,380	2,640	3,220
01401650	5.36	762	1,290	1,670	2,200	2,610	3,040	4,090
01403150	1.99	686	1,090	1,380	1,760	2,050	2,350	3,060
01403535	1.57	210	397	543	751	919	1,100	1,550

Table 25.--Observed and adjusted total lengths of record

Station number	Observed length of record	Adjusted total length of record
01398107	6	14
01398500	63	71
01400300	5	13
01401650	5	13
01402600	13	21
01403150	5	13
01403400	26*	34
01403535	5	13

* Adjusted from 19 years on the basis of historic data.

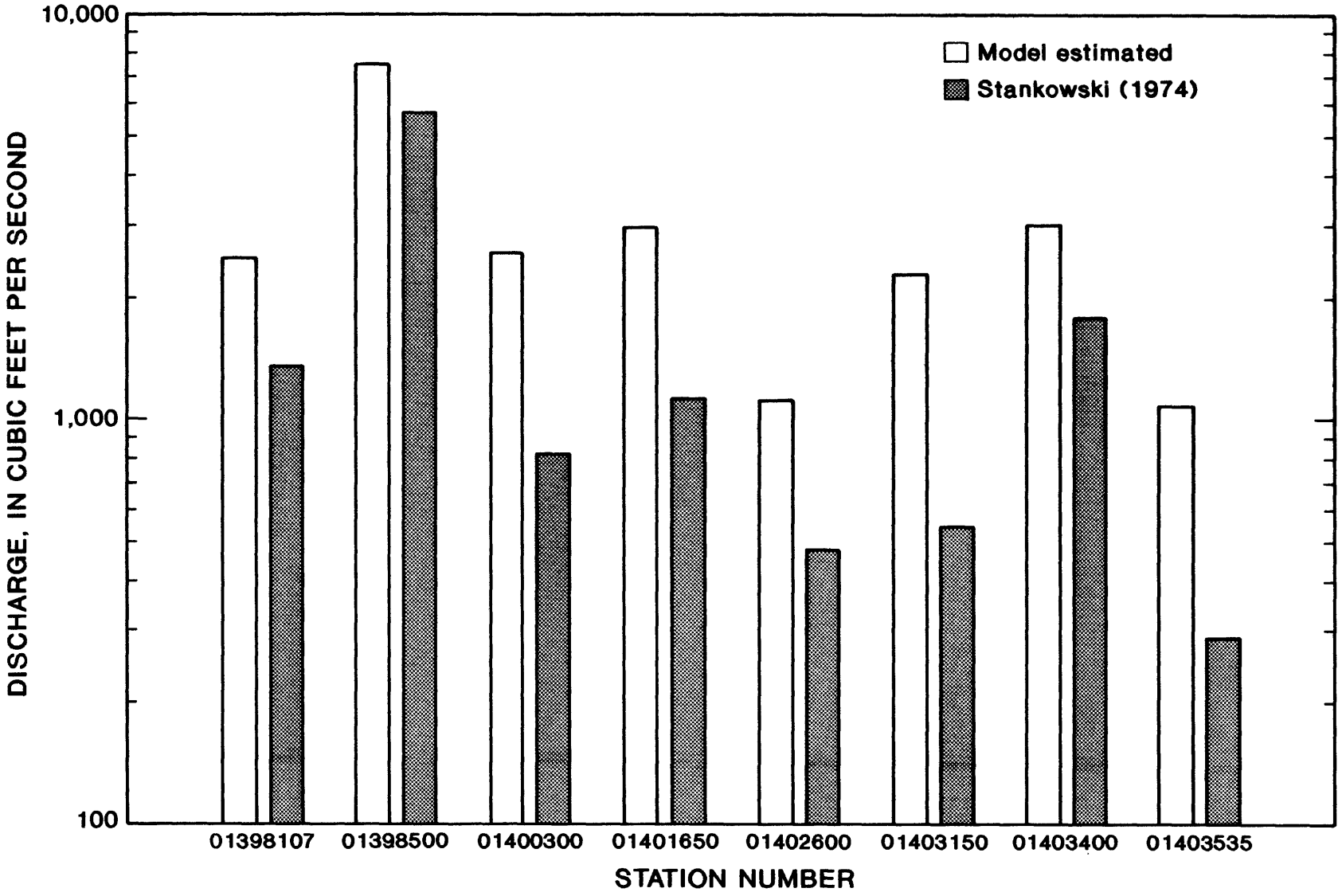


Figure 35.--One hundred year peak-flow estimates and estimates computed using equations given by Stankowski (1974).

Long-term 5-minute storm-precipitation data at Trenton, New Jersey, were assembled for selected storms. A total of 448 storms was selected from data collected during 1914-79. The storms were selected on the basis of storm volume and intensity, and antecedent soil-moisture conditions.

Statistically significant differences (95 percent confidence level) were found between precipitation data at Trenton and data from in and near Somerset County. Annual total precipitation was found to be 15.3 percent lower in Trenton than in Somerset County, on average. Annual maximum storm volumes were 13.5 percent lower in Trenton than in Somerset County, on average. The long-term daily and 5-minute data were adjusted to account for these differences.

The recording increment of the storm-precipitation data used for model calibration (0.1 inch) was different from the recording increment used for long-term simulation (0.05 inch). Analyses indicated that the recording increment of the data used for long-term simulation had a small, but statistically significant, effect on the synthesized flow values. The 2-year recurrence-interval flood peaks were 7 percent higher when 0.1-inch-increment data were used, on average. The 100-year recurrence-interval flood peaks were 3 percent lower when 0.1-inch-increment data were used. The long-term 5-minute data were aggregated to provide a 0.1-inch recording increment to account for differences between the recording increment of the long-term data and the precipitation data used to calibrate the model.

The model was used in combination with the long-term precipitation data and synthetic evaporation estimates to synthesize long-term flood series at each of the eight study basins. Because the model only simulates flood flow from surface runoff, a base-flow model was developed for synthesizing long-term base flows. The synthesized long-term base flows were added to the synthesized runoff flows. The 2- to 500-year floods were estimated at five of the basins by analyzing the combined synthetic and observed records. Flood-frequency estimates were made at the remaining three basins by using a weighted combination of frequency estimates based on the synthesized record and estimates based on observed records.

SELECTED REFERENCES

Alley, W.M., and Smith, P.E., 1982, Distributed-routing rainfall-runoff model version II: U.S. Geological Survey Open-File Report 82-344, [NSTL Station, Mississippi], 201 p.

Bailey, J.F., Thomas, W.O., Wetzel, K.L., and Ross, T.J., 1989, Estimation of flood-frequency characteristics and the effects of urbanization for streams in the Philadelphia, Pennsylvania area, U.S. Geological Survey Water-Resources Investigations Report 87-4194, 72 p.

Box, G.E.P., and Cox, O.R., 1964, The analysis of transformations: Journal of the Royal Statistical Society, Series B., v. 26, no. 2, p. 211-252.

Campbell, J.B., 1987, Rainfall-runoff data for Somerset County, New Jersey: U.S. Geological Survey Open-File Report 87-384, 161 p.

Chow, V.T., 1964, Handbook of applied hydrology: New York, McGraw-Hill, Inc., p. 12-26.

Curtis, G.W., 1977, Frequency analysis of Illinois floods using observed and synthetic streamflow records: U.S. Geological Survey Water-Resources Investigations Report 77-104, 32 p.

Colson, B.E., and Hudson, J.W., 1976, Flood frequency of Mississippi streams: U.S. Geological Survey Report, prepared for the Mississippi State Highway Department, in cooperation with the Federal Highway Administration, Department of Transportation, 34 p.

Dawdy, D.R., Lichty, R.W., and Bergmann, J.M., 1972, A rainfall-runoff simulation model for estimation of flood peaks for small drainage basins: U.S. Geological Survey Professional Paper 506-B, 28 p.

Dawdy, D.R., Schaake, J.C., and Alley, W.M., 1978, User's guide for Distributed-Routing Rainfall-Runoff Model: U.S. Geological Survey Water-Resources Investigations 78-90, 146 p.

Doyle, H.W., and Miller, J.E., 1980, Calibration of a Distributed-Routing Rainfall-Runoff Model at four urban sites near Miami, Florida: U.S. Geological Survey Water-Resources Investigations 80-1, 87 p.

Federal Emergency Management Agency, Federal Insurance Administration, 1980, Flood Insurance Study, Township of Hillsborough, Somerset County, New Jersey: Washington, D.C.

Fennemann, N.M., 1946, Physical divisions of the United States: U.S. Geological Survey Special Map, 1 sheet, scale 1:7,000,000.

Franklin, M.A., 1984, Magnitude and frequency of flood volumes for urban watersheds in Leon County, Florida: U.S. Geological Survey Water-Resources Investigations Report 84-4233, 20 p.

Hamon, W.R., 1961, Estimating potential evapotranspiration: Proceedings of the American Society of Civil Engineers, Journal of the Hydraulic Division, v. 87, no. HY3, p. 107-120.

Hauth, L.D., 1974, Model synthesis in frequency analysis of Missouri floods: U.S. Geological Survey Circular 708, 16 p.

Hutchinson, N.E., 1975, WATSTORE user's guide, volume 1: U.S. Geological Survey Open-File Report 75-426.

Inman, E.J., 1983, Flood-frequency relations for urban streams in metropolitan Atlanta, Georgia: U.S. Geological Survey Water-Resources Investigations Report 82-4203, 38 p.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency: Hydrology Subcommittee Bulletin 17B.

Kirby, William, 1975, Model smooth effect diminishes simulated flood peak variance, EOS, Transactions of the American Geophysical Union, v. 56, no. 6, p. 361.

Leclerc, Guy, and Schaake, J.C., Jr., 1973, Methodology for assessing the potential impact of urban development on urban runoff and the relative efficiency of runoff control alternatives: Ralph M. Parsons Laboratory Rept. no. 167, Massachusetts Institute of Technology, 257 p.

Livingston, R.K., 1981, Rainfall-runoff modeling and preliminary regional flood characteristics of small rural streams in the Arkansas River basin in Colorado: U.S. Geological Survey, Water-Resources Investigations 80-112, 43 p.

Loague, K.M., and Freeze, R.A., 1985, A comparison of rainfall-runoff modeling techniques on small upland catchments: Water Resources Research, v. 21, no. 2, p. 229-248.

Maps, Inc., 1964, Topographic maps of Bridgewater Township at a scale of 1"=100', Contour interval of two feet: April 27, 1964, Baltimore, Maryland.

Michael S. Kachorsky Associates, 1972, Topographic Maps, Scale 1:1200, Contour interval 2 feet: Wachung, New Jersey.

Nash, J.E., and Sutcliffe, J.V., 1970, River flow forecasting through conceptual models, 1, A discussion of principles: Journal of Hydrology, v. 10, p. 282-290.

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

Sloto, R.A., 1982, A stormwater management model for the West Branch Brandywine Creek, Chester County, Pennsylvania: U.S. Geological Survey Water-Resources Investigations 81-73, 34 p.

Soil Conservation Service, 1972, National engineering handbook, sec. 4, Hydrology, chap. 17, Flood routing: U.S. Department of Agriculture, p. 17-1 to 17-93.

Stankowski, S.J., 1974, Magnitude and frequency of floods in New Jersey, with effects of urbanization: New Jersey Department of Environmental Protection, Division of Water Resources, Special Report 38., 45 p.

Troutman, B.M., 1982, An analysis of input errors in precipitation-runoff models using regression with errors in independent variables: Water Resources Research, v. 18, no. 4, p. 947-964.

Troutman, B.M., 1983, Runoff prediction errors and bias in parameter estimation induced by spatial variability of precipitation: Water Resources Research, v. 19, no. 3, p. 791-810.

Wandle, S.W., 1983, Estimating peak discharges of small, rural streams in Massachusetts: U.S. Geological Survey Water-Supply Paper 2214, 26 p.

Wibben, H.C., 1976, Application of the U.S. Geological Survey Rainfall-Runoff simulation model to improve flood-frequency estimates on small Tennessee streams: U.S. Geological Survey Water-Resources Investigations 76-120, 53 p.