ESTIMATION OF FLOOD-FREQUENCY CHARACTERISTICS OF SMALL URBAN STREAMS IN NORTH CAROLINA

By Jeanne C. Robbins *and* Benjamin F. Pope

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 96-4084

Prepared in cooperation with the

NORTH CAROLINA DEPARTMENT OF TRANSPORTATION,
and the CITIES of ASHEVILLE, FAYETTEVILLE,
and RALEIGH, NORTH CAROLINA

Raleigh, North Carolina

1996

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary

U.S. GEOLOGICAL SURVEY
Gordon P. Eaton, Director

For addtional information write to:

District Chief
U.S. Geological Survey
3916 Sunset Ridge Road
Raleigh, NC 27607

Copies of this report can be purchased from:

U.S. Geological Survey
Branch of Information Services
Box 25286
Federal Center
Denver, CO 80225-0286

CONTENTS

ILLUSTRATIONS

TABLES

CONVERSION FACTORS AND VERTICAL DATUM

Multiply	By	To obtain
	Length	
inch (in.)	25.4	millimeter
foot (ft)	0.3048	meter
mile (mi)	1.609	kilometer
	Area	
acre	4.047	square meter
	0.4047	hectare
square mile (mi^2)	2.590	square kilometer
	Flow	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second

Sea level: In this report "sea level" refers to the National Geodetic Vertical Datum of 1929 (NGVD)--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.

ESTIMATION OF FLOOD-FREQUENCY CHARACTERISTICS OF SMALL URBAN STREAMS IN NORTH CAROLINA

By **Jeanne C. Robbins** *and* **Benjamin F. Pope**

ABSTRACT

A statewide study was conducted to develop methods for estimating the magnitude and frequency of floods of small urban streams in North Carolina. This type of information is critical in the design of bridges, culverts and water-control structures, establishment of flood-insurance rates and flood-plain regulation, and for other uses by urban planners and engineers.

Concurrent records of rainfall and runoff data collected in small urban basins were used to calibrate rainfall-runoff models. Historic rainfall records were used with the calibrated models to synthesize a long-term record of annual peak discharges. The synthesized record of annual peak discharges were used in a statistical analysis to determine flood-frequency distributions. These frequency distributions were used with distributions from previous investigations to develop a database for 32 small urban basins in the Blue Ridge-Piedmont, Sand Hills, and Coastal Plain hydrologic areas. The study basins ranged in size from 0.04 to 41.0 square miles. Data describing the size and shape of the basin, level of urban development, and climate and rural flood characteristics also were included in the database.

Estimation equations were developed by relating flood-frequency characteristics to basin characteristics in a generalized least-squares regression analysis. The most significant basin characteristics are drainage area, impervious area, and rural flood discharge. The model error and prediction errors for the estimating equations were less than those for the national flood-frequency equations previously reported. Resulting equations, which have prediction errors generally less than 40 percent, can be used to estimate flood-peak discharges for 2-, 5-, 10-, 25-, 50-, and 100-year recurrence intervals for small urban basins across the State assuming negligible, sustainable, in-channel detention or basin storage.

INTRODUCTION

Information on the magnitude and frequency of floods is critical in the design of bridges, culverts, and some water-control structures, establishment of flood-insurance rates and flood-plain regulation, and for other uses by urban planners and engineers. Because urbanization can significantly affect the magnitude and frequency of floods, existing rural flood-frequency relations are not applicable to urban streams. Recognizing the need for urban flood-frequency relations, the U.S. Geological Survey (USGS), in cooperation with the North Carolina Department of Transportation and the Cities of Asheville, Fayetteville, and Raleigh, conducted an investigation of flood-frequency characteristics of small urban streams in North Carolina.

Traditionally, flood-frequency characteristics are determined by analyzing long periods of homogeneous flood record. However, long periods of homogeneous record are hard to obtain in urban basins where the hydrologic character of the basin is subject to change. Therefore, calibrated digital models frequently are used to synthesize long-term flood records for flood-frequency analysis. This method allows a relatively short period of homogeneous basin conditions to be fixed, as a calibrated rainfall-runoff model, and applied to a long period of climatological conditions resulting in a long term, homogenous, synthetic record of peak flows that reflects a specific set of basin conditions. This long, synthetic record can then be statistically analyzed to develop an estimate of the magnitude of peak flows of selected recurrence intervals. In order to develop estimates of urban flood-frequency, other investigators also used rainfall-runoff models to extend streamflow records (Putnam, 1972; Gunter and others, 1987), or selected homogeneous periods of streamflow record from longer records at urban basins in stable urban areas (Sauer and others, 1983). This study incorporates urban flood-frequency data from these studies with the urban flood-frequency data developed from the data-collection component of this study in order to increase the amount and extend the geographic coverage of data available for analysis. Flood-frequency characteristics and basin characteristics are used in a regression analysis to determine equations for estimating flood characteristics at ungaged sites.

Purpose and Scope

This report describes the development and application of regression equations for estimating the magnitude and frequency of floods in ungaged, urban basins in North Carolina ranging in size from 0.04 to 41.0 square miles (mi^2). Regression equations include flood discharges having 2-, 5-, 10-, 25-, 50-, and 100-year recurrence intervals and are based on a database of 32 sites in 11 North Carolina cities. Data were collected by the USGS specifically for this study at 17 of these sites, in Asheville, Fayetteville, and Raleigh, from 1986 through 1993. Data for the other 15 sites, in Charlotte, Goldsboro, Greenville, Lenoir, Morganton, Wilmington, Wilson, and Winston-Salem were compiled from several previous flood-frequency investigations. These data were collected by USGS personnel for various periods between 1962 and 1991.

The following sections of the report describe the collection of streamflow and rainfall data at the urban basins selected for this study and the development of the historic climatological data used for record extension. This study used the USGS rainfall-runoff model (RRM) (Dawdy and others, 1972). The RRM, its calibration, and its use to generate synthetic, long-term records or peak flows; the development of urban flood-frequency characteristics from statistical analysis of those records and from the analyses of previous investigators; the compilation of urban basin characteristics; and the development of the regression equations also are described.

Previous Flood-Frequency Studies

Previous investigations of flood frequency in North Carolina have focused on small rural basins or urban basins in selected physiographic regions of the State. Putnam (1972) related basin lagtime to the length, slope, and percentage of impervious area in an urban basin and used it to develop flood-frequency relations for small urban streams in the Piedmont Province of North Carolina (fig. 1). Gunter and others (1987) developed regional relations for estimating the magnitude and frequency of flood discharges for rural basins throughout North Carolina. Although Gunter and others (1987) also showed that the nationwide flood-frequency relations developed by Sauer and others (1983) were applicable to the Coastal Plain Province of North Carolina, they were unable to show applicability of the relations to other regions of the State.

Acknowledgments

This study was coordinated by Mr. Robert R. Mason, Jr., of the USGS in cooperation with the North Carolina Department of Transportation, and the Cities of Asheville, Fayetteville, and Raleigh. The authors gratefully acknowledge the assistance of Mr. Archie Hankins of the North Carolina Department of Transportation, Mr. James Ewing and Mr. Tom Tarrant of the City of Asheville, Mr. Robert Bennett and Mr. Gerald Croll of the City of Fayetteville, and Mr. Jimmie Beckom and Mr. Danny Bowden of the City of Raleigh.

DATA COLLECTION AND ANALYSIS

For this study, concurrent records of rainfall and runoff data were collected in 24 small urban basins in the Blue Ridge-Piedmont and Sand Hills hydrologic areas of North Carolina from 1986 through 1993. These data were used to calibrate rainfall-runoff models for 17 basins. The RRM is a lumped parameter model that assumes no significant storage in the basin and uniform rainfall throughout the basin (Dawdy and others, 1972). Calibration was achieved through adjustment of model parameters representing soil moisture accounting, infiltration and surface runoff routing. Historic rainfall and evaporation records from six sites were then applied to each calibrated model to synthesize a long record of peak discharge. Annual flood-peak discharges were determined from the synthesized record. Synthesized annual flood peaks were used in a statistical analysis to determine discharges for selected recurrence intervals. A database, including flood characteristics for each modeled basin and flood characteristics for 15 additional basins from previous flood-frequency studies, was developed for use in the regression analysis. The database also included basin characteristics describing the size and shape of the basin, level of urban development, climate, and rural flood characteristics. A generalized least-squares regression analysis was used to develop equations relating flood frequency to basin characteristics.

Hydrologic Data

Streamflow and rainfall data were collected at 24 study basins, representing small urban basins in the Blue Ridge (5 sites), Piedmont, (12 sites) and Sand Hills (7 sites) regions of North Carolina (fig. 1; table 1). The study basins were selected to represent a broad range of land use and development. Basins were chosen for which no significant change in land use or basin development was anticipated during the data-collection period. Where possible, peak-flow data sites were selected to meet criteria required for application of indirect methods for flow measurement (Benson and Dalrymple, 1967).

Basins were avoided in which flow was regulated or in which significant temporary storage of flood flows occurred; however, two such basins in the Piedmont region were included to satisfy cooperative interests. Further investigation at five other basins revealed that these basins did not meet the selected model assumptions. Rainfall-runoff models were not calibrated and flood-frequency characteristics were not developed for these seven basins, although they are listed in table 1.

Typical instrumentation at a study basin consisted of a float counterweight inside a stilling well to sense stage and a float counterweight inside a standpipe raingage to measure rainfall accumulation. Stage and rainfall were recorded using either a mechanical analog-to-digital recorder or an electronic data-logger. Recording intervals or the unit-value interval for the analog-to-digital recorders were either 5 or 15 minutes, depending on the size and anticipated response time of the basin. The electronic data-loggers allowed 1-minute recording intervals, which were used at a number of the smaller basins. Additionally, there were at least two crest-stage indicators at each site, upstream and downstream from the gage. Current-meter discharge measurements were made at the sites during the study period and the sites were surveyed to obtain data for the development of stage-discharge relations.

The stage and rainfall data collected at the study sites were loaded into the USGS automated data-processing system (ADAPS). Discharge data were computed from the recorded stage data using a stage-discharge relationship, or rating. Ratings were typically developed using indirect methods for flow measurement (Benson and Dalrymple, 1967) and checked with available current-meter discharge measurements; however, a few sites were rated based on direct current-meter discharge measurements only. One site (site 14, table 1) in the Piedmont was discontinued because a satisfactory rating could not be developed.

Unit values of stage and rainfall data were plotted versus time after each site visit. The hydrographs were reviewed to identify timing errors, hung floats, plugged intakes, and other recorder or sensor malfunctions. Data from crest stage indicators were used as a check on recorded peak stages.

Rainfall records for selected storms were compared to records from nearby study basins for consistency. Unit data for selected storms were retrieved from ADAPS and formatted for loading into the rainfall-runoff model. Daily rainfall record was computed for each site; however, at a number of the Piedmont and Sand Hills sites, the electronic data-loggers did not function properly following their initial installation which caused small amounts of rainfall to be recorded on days for which there was no actual rainfall. Comparisons between nearby National Weather Service (NWS) daily rainfall record and study-site rainfall record indicated that this was significant only during dry periods; actual rainfall amounts were correctly recorded and unit rainfall data for selected storms was not affected. Because of this error, nearby NWS daily rainfall record was used to calibrate the models for most Piedmont and Sand Hills sites.

In addition to unit rainfall, unit discharge, and daily rainfall data, daily evaporation data also are required for calibration of the model. Pan evaporation data for the calibration period (1986-92) were obtained from sites at Chapel Hill, North Carolina (National Weather Service), and Franklin, North Carolina (Tennessee Valley Authority) (fig. 1; table 1).

Long-term historic records of short-interval (5- and 15-minute) storm rainfall, daily rainfall, and evaporation data were applied to the calibrated models to synthesize long-term peak-flow data. Short-interval historic rainfall data used for peak-flow synthesis were obtained from NWS rainfall hyetographs of 2-5 storms of greatest intensity per year from 1954 to 1984 for Raleigh, and from 1940 to 1984 for Asheville (fig. 1; table 1). Daily rainfall data for Raleigh and Asheville

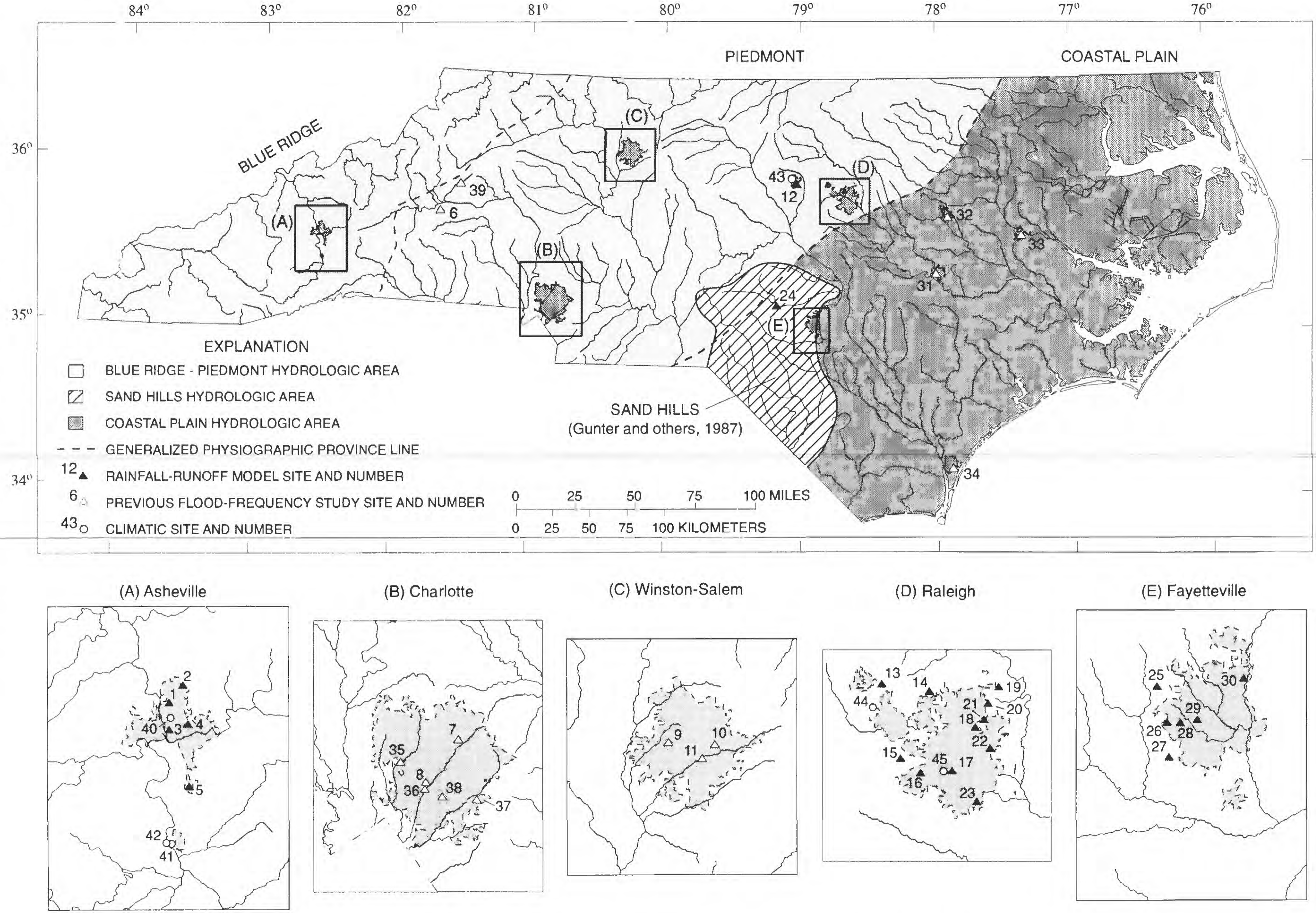

Figure 1. Locations of urban rainfall-runoff model sites, previous flood-frequency study sites, and climatic sites, North Carolina.

Table 1. Data-collection site descriptions in North Carolina

Site number (fig. 1)	Station name	USGS station number	Latitude	Longitude	Type of data	Period of record used
1	Reed Creek above Barnard Avenue at Asheville	03451510	35°36'52"	82°33'41"	Discharge, precipitation	7/18/86 - 12/31/88
[a]2	Spooks Branch near Woodfin	0345153800	35°38'17"	82°32'24"	Discharge, precipitation	6/18/87 - 1/1/90
3	Nasty Branch at Asheville	0345112600	35°34'44"	82°33'35"	Discharge, precipitation	7/3/86 - 12/31/88
4	Ross Creek at Beaucatcher Road at Asheville	0345092550	35°35'15"	82°31'49"	Discharge, precipitation	6/18/86 - 12/31/89
5	Dingle Creek near Skyland	03448068	35°30'22"	82°31'30"	Discharge, precipitation	3/23/88 - 1/1/90
[b]6	Hunting Creek at Morganton	02139610	35°44'17"	81°40'45"	Discharge, precipitation	1967 - 70
[b]7	Briar Creek tributary 7 at Shamrock Drive, Charlotte	02146436	35°14'07"	80°47'26"	Discharge, precipitation	1966 - 70
[b]8	Little Hope Creek at Seneca Place, Charlotte	02146470	35°09'53"	80°51'12"	Discharge, precipitation	12/1/82 - 10/30/91
[b]9	Silas Creek tributary at Pine Valley Road, Winston-Salem	02115765	36°06'19"	80°17'52"	Discharge, precipitation	1968 - 70
[b]10	Brushy Fork tributary no. 2 at U.S. Highway 311, Winston-Salem	02115839	36°06'10"	80°13'21"	Discharge, precipitation	1968 - 70
[b]11	Tar Branch at Walnut Street, Winston-Salem	02115843	36°05'02"	80°14'34"	Discharge, precipitation	1967 - 70
[a]12	Little Creek tributary near Chapel Hill	0209736050	35°55'02"	79°01'57"	Discharge, precipitation	2/4/87 - 9/30/93
13	Sycamore Creek near Lynn Crossroads	0208725600	35°54'03"	78°45'56"	Discharge, precipitation	9/3/87 - 1/14/91
[a]14	Haire Snipe tributary near Leesville	0208726835	35°53'28"	78°41'25"	Discharge, precipitation	7/24/87 - 2/7/91
15	Richlands Creek near Westover	0208726100	35°48'13"	78°44'07"	Discharge, precipitation	7/22/87 - 9/30/93
16	Bushy Branch tributary below Schaub Drive at Raleigh	0208734221 0208734220	35°47'04"	78°42'14"	Discharge, precipitation	6/8/88 - 9/30/93
17	Pigeon House Creek at Cameron Village at Raleigh	0208732534	35°47'14"	78°39'17"	Discharge, precipitation	8/19/87 - 9/30/93
18	Big Branch tributary at Wingate Drive at Millbrook	0208730025	35°50'38"	78°37'01"	Discharge, precipitation	9/22/87 - 2/7/91
19	Perry Creek tributary at Neuse	0208721290	35°53'47"	78°34'46"	Discharge, precipitation	10/1/86 - 4/26/89
[a]20	Perry Creek at SR 2012 near Millbrook	0208721055	35°52'30"	78°35'48"	Discharge, precipitation	3/1/86 - 9/30/89
21	Marsh Creek at SR 2030 at Millbrook	0208732810	35°51'13"	78°36'12"	Discharge, precipitation	2/18/86 - 9/30/89
[a]22	Marsh Creek near New Hope	0208732885	35°48'59"	78°35'37"	Discharge, precipitation	1/14/86 - 9/30/93
23	Walnut Creek tributary at Evers Street at Raleigh	0208735550	35°44'49"	78°36'54"	Discharge, precipitation	9/14/87 -11/15/90
[a]24	Flat Creek near Inverness	02102908	35°10'54"	79°10'40"	Discharge, precipitation	5/2/91 - 9/24/92

Table 1. Data-collection site descriptions in North Carolina--Continued

Site number (fig. 1)	Station name	USGS station number	Latitude	Longitude	Type of data	Period of record used
25	Jack Fords Creek at Fayetteville	0210434115	35°05'35"	78°57'57"	Discharge, precipitation	8/24/89 - 11/2/92
26	Buckhead Creek at Skibo	0210438680	35°03'34"	78°57'17"	Discharge, precipitation	6/29/89 - 9/30/93
27	Buckhead Creek near Owens	02104387	35°01'37"	78°57'08"	Discharge, precipitation	10/76 - 2/80 and 8/18/89 - 11/2/92
28	Branson Creek near Fayetteville	0210397520	35°03'31"	78°56'23"	Discharge, precipitation	6/29/89 - 9/30/93 8/30/89 - 1/28/93
[a]29	Hybart Creek tributary at Fayetteville	0210397475	35°03'41"	78°55'13"	Discharge, precipitation	8/17/89 - 11/3/92 8/7/89 - 9/29/92
30	Cape Fear River tributary near Fayetteville	0210367030	35°06'01"	78°52'03"	Discharge, precipitation	8/17/89 - 10/7/92
[c]31	Big Ditch at Retha Street at Goldsboro	02088682	35°22'16"	78°00'15"	Discharge, precipitation	1/20/80 - 9/30/84
[c]32	Hominy Swamp at Phillips Street at Wilson	02090512	35°42'39"	77°55'00"	Discharge, precipitation	8/1/78 - 9/30/85
[c]33	Greenmill Run at Arlington Boulevard, Greenville	02084070	35°35'57"	77°22'17"	Discharge, precipitation	3/1/80 - 9/30/85
[c]34	Hewletts Creek at SR 1102 near Wilmington	02093229	34°11'28"	77°53'32"	Discharge, precipitation	10/18/76 - 9/30/90
[d]35	Irwin Creek near Charlotte	02146300	35°11'50"	80°54'18"	Discharge	1962 - 1977
[d]36	Little Sugar Creek near Charlotte	02146500	35°09'13"	80°51'18"	Discharge	1962 - 1977
[d]37	McAlpine Creek at Sardis Road near Charlotte	02146600	35°08'14"	80°45'05"	Discharge	1962 - 1977
[d]38	McMullen Creek at Sharon View Road near Charlotte	02146700	35°08'27"	80°49'13"	Discharge	1964 - 1977
[d]39	Lower Creek at Mulberry Street at Lenoir	02141150	35°54'20"	81°31'59"	Discharge	1967 - 1977
[e]40	Asheville - Downtown	0301	35°35'40"	82°33'28"	Historic daily and unit precipitation	1/1/40 - 12/31/47
[e]41	Asheville Airport	0300	35°25'48"	82°33'00"	Historic daily precipitation	1/1/48 - 12/31/93
[f]42	Fletcher	766	35°25'51"	82°33'31"	Historic daily evaporation	8/1/72 - 9/30/80
[e]43	Chapel Hill 2 W	1677	35°56'49"	79°04'02"	Historic daily evaporation	1/1/48 - 12/31/93
[e]44	Raleigh - Durham Airport	7069	35°52'12"	78°46'48"	Historic daily and unit precipitation	1/1/48 - 12/31/93
[e]45	Raleigh - N.C. State University	7079	35°47'11"	78°40'05"	Historic daily precipitation	1/1/40 - 12/31/93

[a] Data not used in this study.
[b] From Putnam (1972).
[c] From Gunter and others (1987).
[d] From Sauer and others (1983).
[e] National Weather Service data provided by M.F. Brown, Southeast Regional Climate Center, South Carolina Water Resources Commission, written commun., 1994.
[f] Tennessee Valley Authority data provided by L.W. Hamberger, Tennessee Valley Authority, written commun., 1991

rain gages for the period also were obtained from the National Weather Service. Daily pan-evaporation data used in peak-flow synthesis were obtained for sites at Chapel Hill, North Carolina (National Weather Service), and Franklin, North Carolina (Tennessee Valley Authority) (fig. 1; table 1). Pan-evaporation data for periods of missing record were estimated as the average of the data for each day of the years available.

Model Description

The RRM used in this study originally was developed and documented by Dawdy and others (1972) as a means for predicting flood volumes and peak-discharge rates of surface runoff from small drainage basins. Restricting the model to small basins allows for (1) neglecting the contribution of ground water to the total hydrograph, (2) using a single gage to represent basin-wide rainfall, and (3) representing pertinent physical characteristics of the basin by basin-average parameters. Hence, RRM is classified as a lumped-parameter model that assumes a small homogeneous basin with no storage and uniformly distributed rainfall. The model uses 11 model parameters defined in table 2 to describe three significant hydrologic phenomena— soil-moisture retention, rainfall infiltration, and surface-runoff routing. A number of modifications have been made to the model since its original version, including the development of an automated model-parameter optimization routine (E.J. Inman, U.S. Geological Survey, written commun., 1994).

The soil-moisture accounting component evaluates antecedent moisture conditions affecting infiltration. This is a daily accounting scheme, between storm events, that incorporates four model parameters — EVC, RR, BMSM, and DRN (defined in table 2)— and daily values of rainfall and evaporation. This component simulates the redistribution of soil moisture in the soil column and evapotranspiration from the soil. Infiltration is modeled based on a modified form of the infiltration equation developed by Philip (1954). The runoff volume or excess rainfall for each storm event is determined by using unit value rainfall data, output from the soil-moisture accounting component, and four model parameters— PSP, KSAT, RGF, and EIA. The surface-runoff routing component applies a temporal distribution of the resulting runoff volume by use of a unit hydrograph (Clark, 1945). Model parameters KSW, TC, and TP/TC are applied to determine the shape of the resulting hydrograph.

The calibrated model is then applied to each basin using a long-term short-interval historic rainfall record to generate synthetic peak streamflows. These peaks are then analyzed in accordance with procedures outlined by the U.S. Water Resources Council (1981) to determine the t-year peak flow values, where t represents the recurrence interval associated with the flood peak. The model has been used in similar applications of flood-frequency analysis by Inman (1988 and 1995), Bohman (1992), and Sherwood (1994). Specific model algorithms, their theoretical basis, limitations, and applications are described further in Dawdy and others

Table 2. Description of model parameters
[----, a dimensionless parameter]

Parameter	Unit of measure	Definition
		Antecedent soil-moisture accounting component
EVC	----	Coefficient to convert pan evaporation to potential evapotranspiration.
RR	----	Proportion of daily rainfall that infiltrates the soil.
BMSM	inch	Soil-moisture storage volume at field capacity.
DRN	inches/hour	The constant rate of drainage for redistribution of soil moisture.
		Infiltration component
PSP	inch	Minimum value of the combined action of capillary suction and soil-moisture differential.
KSAT	inches/hour	Minimum saturated hydraulic conductivity used to determine soil-infiltration rates.
RGF	----	Ratio of combined action of suction and potential at wilting point to that at field capacity.
EIA	----	Effective ratio of impervious area to total basin area; a measure of the impervious area that is directly connected to the channel drainage system.
		Surface-runoff routing component
KSW	hour	Linear reservoir routing coefficient.
TC	minute	Duration of the triangular translation hydrograph (time of concentration).
TP/TC	----	Ratio of time to peak to time of concentration.

(1972), Carrigan (1973), Boning (1974), Carrigan and others (1977), and by E.J. Inman (U.S. Geological Survey, written commun., 1994).

Model Calibration and Synthesis

A rainfall-runoff model was successfully calibrated at 17 small urban basins in the State. Calibrated models were then supplied long-term historic rainfall records to generate a synthetic peak-flow record. The synthetic peak-flow record was then analyzed in accordance with recommendations from the U.S. Water Resources Council (1981) to determine the 2-, 5-, 10-, 25-, 50-, and 100-year peak-flow values.

Model calibration is defined as the process of determining a set of model parameter values that enable the model to best duplicate observed runoff data when furnished the corresponding observed rainfall data. For each study basin, rainfall-runoff events were reviewed, and 30 to 40 high-flow events with good quality data were selected for use in model calibrations.

Once the rainfall-runoff events were selected, calibration ensued by use of an automated optimization routine. The routine performs a systematic trial and error search, based on a method devised by Rosenbrock (1960), for an optimal set of model parameters to minimize five objective functions. The objective functions measure error between simulated and observed values of peaks and volumes for each rainfall-runoff event (E.J. Inman, U.S. Geological Survey, written commun., 1994). Initial model runs are made with estimated values for each model parameter and a specified range and increment for the automated optimization routine to vary each parameter. DRN and TP/TC were relatively insensitive parameters and were fixed at 1.0 and 0.5, respectively, as reported by Inman (1988) and Bohman (1992). The value of EVC, pan-evaporation coefficient, was fixed at 0.75 for Asheville sites (Blue Ridge), 0.72 for Raleigh sites (Piedmont), and 0.76 for Fayetteville sites (Sand Hills) (Kohler and others, 1959). RR, a soil-moisture accounting parameter insensitive at extreme conditions, also was fixed for each region based on average precipitation and runoff in each hydrologic region. RR values of 0.60, 0.67, and 0.66 were applied to sites in Asheville, Raleigh, and Fayetteville, respectively. Initial values of other soil-moisture and infiltration parameters were set based on data from previous studies in physiographically similar basins (Inman, 1988; Bohman, 1992). Surface runoff-routing parameters were estimated from study basin hydrographs and EIA, defined as the percentage of impervious area that is directly connected to the channel drainage system, and likely the most sensitive model parameter was initially estimated within 25 percent of the total impervious area. Some manual adjustment of model parameters was initially performed before application of the automated routine to ensure that a physically realistic set of model parameters was developed.

The calibration process continued until close agreement was achieved between simulated and observed values. Reasonable calibrations could not be obtained for five of the original gaged study basins, one in the Blue Ridge and two each in the Piedmont and Sand Hills regions. In the Blue Ridge, site 2 (fig. 1; table 1) did not have enough large, uniform rainfall-runoff events that met the model assumptions to develop a calibration. In the Piedmont, site 12 (fig. 1; table 1) was poorly rated and did not have an adequate daily value rainfall record or a suitable NWS alternate nearby, and at site 14 (fig.1; table 1) a rating relating stage to discharge could not be established. In the Sand Hills hydrologic region, site 24 (fig. 1; table 1) was dropped because of its basically rural nature, and site 29 (fig. 1; table 1) contained too much temporary storage, in the form of culverts and storm drains, in relation to total runoff of the basin.

Quality of each calibrated model was judged on the development of a reasonable set of model parameter values, the slope of the regression line between simulated and observed peak runoff and volume values, and the standard error of regression (table 3). A regression line slope criteria of 0.95 to 1.05 was established to evaluate model fit. Peak slopes were always within the established slope criteria. Five volume slopes fell below the 0.95 criteria (table 3), only two of which were less than 0.90, indicating that volumes may have been underestimated as a result of small amounts of storage in these basins, perhaps invalidating RRM model assumptions. However, these effects did not appear to be significant, so the sites were retained in the database.

Model fit was further evaluated based on the calculated standard error for the regression between simulated and observed values. Standard errors of regression for all sites were generally less than 50 percent with 10 of the 17 modeled sites recording

Table 3. Calibrated model parameters and statistics

Site number	PSP	KSAT	RGF	BMSM	EVC	RR	KSW	TC	EIA	Volume slope	Volume error (percent)	Peak slope	Peak error (percent)
1	1.88	0.31	24.7	6.37	0.75	0.60	0.51	45.7	0.08	0.98	34.5	1.04	26.1
3	4.08	.07	18.9	14.9	.75	.60	.21	10.1	.23	.97	35.8	1.00	40.2
4	1.86	.05	15.2	5.51	.75	.60	.62	27.5	.04	.91	54.5	1.03	63.4
5	1.59	.17	22.5	10.2	.75	.60	1.02	62.0	.15	.95	48.6	.95	38.9
13	.43	.14	30.0	2.52	.72	.67	1.39	149	.05	.86	61.0	1.00	56.9
15	1.69	.22	23.2	2.00	.72	.67	.50	30.0	.10	.95	37.0	1.02	36.0
16	2.00	.10	13.4	15.8	.72	.67	.17	18.9	.09	.96	69.0	.98	74.0
17	3.27	.35	5.65	14.9	.72	.67	.17	8.97	.52	.89	34.0	1.04	25.0
18	2.88	.15	5.09	13.9	.72	.67	.11	17.4	.18	.96	37.5	.98	40.7
19	1.69	.05	25.7	8.89	.72	.67	1.83	56.0	.01	.96	34.8	1.00	42.6
21	1.59	.01	19.8	5.59	.72	.67	.91	54.7	.22	.97	30.9	.98	37.4
23	1.07	.23	27.8	3.58	.72	.67	.40	19.4	.01	.93	33.2	1.04	35.2
25	2.17	.05	21.3	4.64	.76	.66	.89	65.4	.07	.97	35.0	1.03	42.5
26	.11	.40	20.0	16.0	.76	.66	.93	22.5	.32	1.01	21.8	.97	21.3
27	.82	.04	26.5	8.42	.76	.66	2.72	158	.09	.96	29.7	.95	38.7
28	.27	.70	7.50	8.10	.76	.66	1.13	53.8	.20	1.04	29.9	.96	26.4
30	3.74	.29	9.44	11.4	.76	.66	.16	11.2	.01	.93	32.5	1.00	29.7

standard error for peaks of less than 40 percent and 13 of the modeled sites recording standard error for volumes less than 40 percent. These values may be slightly higher than those reported by Inman (1988) and Bohman (1992), yet may result from some nonhomogeneity in the basins during the data-collection period or procedural differences in event-selection criteria. The calibrated parameters which represent the set of values that result in the best overall fit and associated regression statistics are presented in table 3. These statistics of model development for each modeled site in the study indicate good unbiased agreement between the observed and simulated data.

Comparative plots of the simulated and observed peak flows and depth of runoff volumes also were reviewed for bias. Plots for site 18, Big Branch tributary at Wingate Drive at Millbrook, N.C., site 28 Branson Creek near Fayetteville, N.C., and site 3, Nasty Branch at Asheville, N.C., show typical results (table 1; fig. 2) in each hydrologic area. The figure indicates that the points are generally equally distributed about the line of perfect fit.

Estimates of peak discharges for selected recurrence intervals were developed for each of the 17 modeled basins. Long-term record of peak flows was generated by applying the available historic unit rainfall, daily rainfall, and daily evaporation data to the set of calibrated model parameters determined for each basin.

For Asheville sites, historic climatic data from Asheville and Fletcher (table 1) were used; for Raleigh sites, historic climatic data from Raleigh and Chapel Hill (table 1) were used. For Fayetteville sites, however, the closest available historic climatic data were from Raleigh and Chapel Hill (table 1). Daily and unit rainfall amounts were multiplied by 1.02, the ratio of the mean annual rainfall at Fayetteville to the mean annual rainfall at Raleigh, to account for geographic variation in precipitation amounts.

The maximum annual peak flows were then selected from each record and that annual series of peak flows were log-transformed and fitted to a Pearson-Type III distribution, in accordance with recommendations of the U.S. Water Resources Council

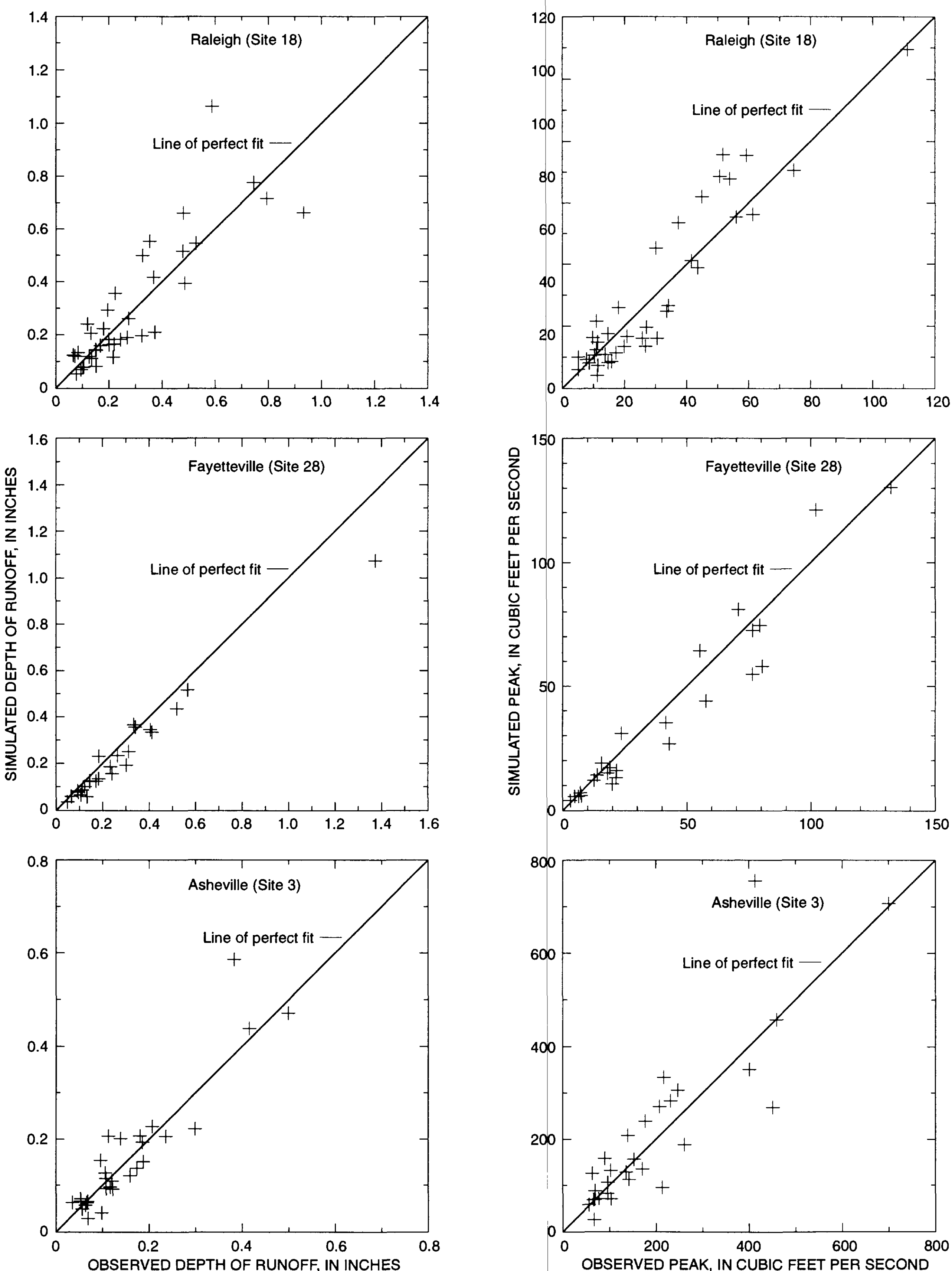

Figure 2. Observed and simulated depth of runoff and peak flows for three selected sites, one each in Asheville, Fayetteville, and Raleigh.

(1981). Using that distribution, flows corresponding to the 2-, 5-, 10-, 25-, 50-, and 100-year recurrence intervals were computed for each modeled basin. Skew coefficients as computed for each basin were used rather than regional skew because the regional skew data provided in U.S. Water Resources Council (1981) is for rural conditions only.

Previous estimates of the 2-, 5-, 10-, 25-, 50-, and 100-year peak discharges in small urban basins in North Carolina were compiled. Putnam (1972) and Gunter and others (1987) used rainfall-runoff modeling and historic rainfall record to generate synthetic peak-flow records at six sites in the Piedmont and four sites in the Coastal Plain, respectively. Sauer and others (1983) selected five Piedmont sites that had 11 to 16 years of homogeneous record and developed estimates of peak discharges from those records. Although all three investigators used log-transformed Pearson-Type III distribution to compute flows for selected recurrence intervals, only Gunter and others (1987) and Sauer and others (1983) did so according to guidelines of the U.S. Water Resources Council (1981), because Putnam's (1972) work was done before those recommendations were established. Putnam only developed estimates from synthetic record for the 25-, 50- and 100-year peak discharges. Consequently, there are 32 sites having the 25-, 50-, and 100-year peak discharges and 26 sites having the 2-, 5-, and 10-year peak discharges (table 4).

Urban Basin Characteristics

The urban basins for which the pooled flood-frequency data are derived constitute a significant aerial coverage of urban areas in North Carolina and include basins representing a wide range of urban development and hydrologic and climatic conditions. Seven hydrologically relevant basin characteristics that serve as measures of basin size and shape and as indices of urban development and channel improvement, climate, and rural or background flood characteristics were developed for use in this study (table 5).

Basin characteristics include basin size and shape. Contributing drainage area (DA), measured in square miles, was determined from topographic maps and included diversions across basin divides. Channel length (L) was measured in miles from the gaging station upstream along the channel to the basin divide, and channel slope (S) was measured in feet per mile, computed as the difference in elevation between the 10- and 85-percent points along the stream channel divided by the length between those two points.

Measures of urban development and channel improvement include percent impervious area (IA) and basin development factor (BDF). IA is a dimensionless value determined by overlaying a grid on basin maps, delineating the impervious areas such as roads, parking lots, and rooftops, and determining the percentage of grid cells that constitute areas impervious to infiltration of rainfall. BDF is a dimensionless value which represents a measure of the prevalence of drainage aspects of (1) channel improvements, (2) channel linings, (3) storm drains or sewers, and (4) curb and gutter streets. As described by Sauer and others (1983), the value of BDF is computed for each third of the drainage basin by evaluating prevalence of drainage aspects in each subarea. A value of zero is assigned if each of the aforementioned drainage improvements constitutes less than 50 percent of the drainage system; conversely, a value of one is assigned if drainage improvements constitute more than 50 percent. These values are summed for each drainage aspect in each subarea resulting in a total BDF value which ranges from 0 to 12.

Measures of climate include 2-year, 2-hour rainfall amount ($RI_{2,2}$) (Hershfield, 1961). The rural or background flood characteristics are provided as the rural flood-frequency values for 2-, 5-, 10-, 25-, 50-, and 100-year floods (RQ2, RQ5, RQ10, RQ25, RQ50, and RQ100). These measures of rural flood characteristics were derived from regression equations developed by Gunter and others (1987) for the Blue-Ridge Piedmont, Sand Hills, and Coastal Plain, and Sand Hills hydrologic regions of North Carolina. Contributing drainage area (DA), in square miles, is the only explanatory variable used in each regression equation (table 6).

Estimates of these characteristics were developed for each basin (table 6). The basins range in size from 0.04 to 41.0 mi^2 with impervious areas ranging from 2 to 54.6 percent of the basin. Basin development factors ranged from 2 to 11; however, the site with the greatest impervious area was not the site with the largest basin development factor.

Table 4. Flood-frequency data for selected recurrence intervals in urban basins

[ft^3/s, cubic feet per second]

Site number	Recurrence interval years (ft^3/s)					
	2	5	10	25	50	100
1	248	525	775	1,170	1,520	1,930
3	628	1,230	1,750	2,440	3,000	3,580
4	745	1,360	1,810	2,410	2,880	3,350
5	142	270	374	525	650	786
6				2,200	2,600	3,000
7				600	660	700
8				1,950	2,100	2,200
9				1,020	1,100	1,200
10				1,250	1,350	1,450
11				1,450	1,550	1,700
13	226	379	487	626	730	834
15	194	316	412	552	670	800
16	99.1	168	215	272	312	351
17	209	299	353	415	457	495
18	56.5	91.4	114	140	159	176
19	68.6	135	186	259	316	377
21	434	622	749	911	1,030	1,150
23	132	234	318	444	553	675
25	147	223	276	346	401	457
26	328	471	550	631	682	726
27	365	566	700	869	994	1,120
28	150	235	286	343	379	412
30	6.3	13.5	19.9	29.7	38.3	48.0
31	497	737	910	1,140	1,330	1,520
32	357	549	695	900	1,070	1,250
33	281	511	685	925	1,120	1,310
34	247	503	725	1,070	1,360	1,700
35	3,200	4,650	5,650	6,960	7,970	8,990
36	4,360	5,950	7,000	8,330	9,330	10,300
37	2,700	3,880	4,700	5,760	6,560	7,390
38	925	1,260	1,470	1,750	1,950	2,160
39	1,390	2,090	2,630	3,410	4,070	4,790

Putnam (1972)

Gunter and others (1987)

Sauer and others (1983)

Table 5. Selected basin characteristics

[DA, contributing drainage area; L, channel length; S, channel slope; IA, impervious area; BDF, basin development factor; RQ, rural equivalent peak discharge for 2-, 5-, 10-, 25-, 50-, and 100-year floods; $RI_{2,2}$, 2-year, 2-hour rainfall amount; mi^2, square mile; mi, mile; ft/mi, foot per mile; ft^3/s, cubic foot per second; in., inch]

Site number (fig. 1)	DA (mi^2)	L (mi)	S (ft/mi)	IA (percent)	BDF	RQ2 (ft^3/s)	RQ5 (ft^3/s)	RQ10 (ft^3/s)	RQ25 (ft^3/s)	RQ50 (ft^3/s)	RQ100 ft^3/s)	$RI_{2,2}$ (in.)
1	2.13	2.27	147	17.5	5	243	411	552	766	950	1,169	1.9
3	1.36	1.98	90	31.4	11	178	305	410	571	710	876	1.9
4	2.46	2.91	156	15.0	4	268	453	608	842	1,043	1,283	1.9
5	1.04	1.67	352	14.3	4	148	254	343	479	596	737	1.9
6	8.26	6.56	28	3	6	619	1,020	1,360	1,862	2,292	2,795	1.9
7	.52	1.07	70	20	9	91.7	160	216	304	380	472	1.9
8	2.72	2.66	41	15	9	288	484	650	899	1,113	1,368	1.9
9	.89	1.62	88	12	5	133	229	309	433	539	667	1.9
10	.55	1.10	143	37	9	95.3	166	224	316	394	490	1.9
11	.59	1.27	156	28	7	100	174	235	331	412	512	1.9
13	2.75	2.56	36	4.7	3	290	488	654	906	1,121	1,378	2.1
15	.98	1.06	64	10.4	3	142	244	330	461	573	710	2.1
16	.19	.60	127	34.2	6	45.7	81.5	111	157	200	247	2.1
17	.27	.61	162	54.6	10	58.2	103	140	198	248	310	2.1
18	.08	.55	121	41.7	9	25.1	45.6	62.3	89.3	113	142	2.1
19	1.09	1.98	89	3.85	3	153	262	354	494	614	760	2.1
21	1.27	1.93	80	32.4	8	170	291	392	546	679	838	2.1
23	.66	1.42	60	10.3	7	108	187	253	356	443	550	2.1
25	.64	.89	27	19.4	4	21.4	35.1	46.3	61.7	74.9	89.6	2.2
26	.82	1.00	20	48	11	25.7	42.1	55.0	74.3	90.4	108	2.2
27	2.74	3.60	14	26.2	6	62.2	102	136	184	225	272	2.2
28	.64	1.81	30	27.0	8	21.4	35.1	46.3	61.7	74.9	89.6	2.2
30	.04	.11	375	12	4	2.81	4.54	5.95	7.69	9.18	10.8	2.2
31	2.17	3.26	11	30	4	113	233	347	547	730	960	2.3
32	7.92	4.92	11	11	5	257	496	715	1,089	1,420	1,826	2.1
33	9.10	4.85	9	2	2	280	538	773	1,172	1,525	1,957	2.3
34	1.98	1.82	15	6	3	107	221	330	521	696	917	2.6
35	30.5	11.2	13.7	20	9	1,530	2,450	3,240	4,380	5,360	6,470	1.9
36	41.0	11.0	13.1	22.0	9	1,870	2,990	3,950	5,320	6,490	7,830	1.9
37	38.3	8.72	12.20	10.0	7	1,790	2,850	3,770	5,090	6,210	7,490	1.9
38	6.98	5.20	20.90	12.0	9	551	912	1,220	1,670	2,050	2,510	1.9
39	31.80	9.40	17.7	13.0	4	1,570	2,520	3,330	4,500	5,500	6,650	1.9

Table 6. North Carolina rural flood-frequency equations

[DA, contributing drainage area of basin, measured in square miles]

Rural flood-recurrence interval	Hydrologic region		
	Blue-Ridge Piedmont	Sand Hills	Coastal Plain
2	$144\ DA^{0.691}$	$29.7\ DA^{0.733}$	$69.4\ DA^{0.632}$
5	$248\ DA^{.670}$	$48.8\ DA^{.738}$	$149\ DA^{.582}$
10	$334\ DA^{.665}$	$64.4\ DA^{.740}$	$225\ DA^{.559}$
25	$467\ DA^{.655}$	$86.2\ DA^{.751}$	$362\ DA^{.532}$
50	$581\ DA^{.650}$	$105\ DA^{.757}$	$490\ DA^{.514}$
100	$719\ DA^{.643}$	$126\ DA^{.763}$	$653\ DA^{.497}$

ESTIMATION OF FLOOD-FREQUENCY CHARACTERISTICS

Generalized least-squares regression techniques were used to develop the final estimating equations (table 7) that relate peak discharges for 2-, 5-, 10-, 25-, 50-, and 100-year floods (table 4) to basin characteristics at urban basins in North Carolina (table 5). The response or dependent variables for the regression analysis were estimates of peak discharges for selected recurrence-interval floods. Explanatory or independent variables were the set of seven basin characteristics identified as potential predictors of peak flow. Regression analyses were performed on the logarithms of the response and explanatory variables in order to linearize the relations between peak flows and basin characteristics. Regression equations were tested for both parameter and geographic bias and sensitivity to errors in the independent variables, and the equations can be used to estimate flood-frequency characteristics for small, ungaged, urban North Carolina streams with insignificant surface and embankment storage.

Regression Analysis

Initial or exploratory, ordinary least-squares regression was used to determine the best model for all combinations of one through seven explanatory variables using adjusted coefficient of determination as the criterion. This exploratory regression resulted in two different three-variable regression models. The first, most significant model, included drainage area, relative impervious area, and rural equivalent flood discharge. The second, slightly less significant model consisted of drainage area, basin development factor, and rural equivalent flood discharge. Although ordinary least-squares regression was used to preliminarily determine the most significant explanatory variables, the nature of the peak discharge data used as the response variable is such that ordinary least-squares regression is inappropriate for the regional regression analysis.

Ordinary least-squares is an appropriate and efficient regression model for use when estimates of response variables are independent of each other (no correlation between pairs of sites) and when the reliability and variability of flow estimates used as response variables are approximately equal. Most of the peak discharges used in this regression were developed from synthetic records of annual peaks generated by a rainfall-runoff model; such records are highly correlated with other sites in the same city because common rainfall records were used to generate data for sites within the same city. The synthetic records at gages in Fayetteville and Raleigh were generated using the same rainfall record. In this case, not only are records from the same city highly correlated with each other, but records from two different cities also have a high degree of correlation. Peak discharges at the five sites taken from Sauer and others (1983) are based on actual records. These records have varying degrees of correlation among themselves and with the other records used.

Because of the correlations between the annual peak records, estimates of peak discharges developed for use in the regional regressions cannot be assumed to be independent; also, because of the different sources of record, the assumption of equal reliability and variability of the estimates is violated. For this reason, regression equations were developed using generalized least-squares regression techniques.

Table 7. North Carolina urban flood-frequency equations

[DA, contributing drainage area of basin, in square miles; IA, relative impervious area, in percent; RQ, rural equivalent peak discharge, in cubic feet per second]

Urban flood-recurrence interval (years)	Urban flood-frequency equation	Model error (percent)	Prediction error (percent)
2	$7.87\ DA^{0.539}\ IA^{0.686}\ RQ2^{0.290}$	32.8	40.4
5	$16.3\ DA^{.489}\ IA^{.572}\ RQ5^{.286}$	31.7	38.5
10	$22.7\ DA^{.463}\ IA^{.515}\ RQ10^{.289}$	31.9	38.3
25	$28.5\ DA^{.390}\ IA^{.436}\ RQ25^{.338}$	34.2	38.7
50	$37.4\ DA^{.391}\ IA^{.396}\ RQ50^{.325}$	33.3	37.8
100	$48.0\ DA^{.392}\ IA^{.358}\ RQ100^{.312}$	33.0	37.8

Generalized least-squares regression (GLS), as described by Stedinger and Tasker (1985), is a regression technique that takes into account differences in the variability and reliability of, as well as the correlation between, dependent variables. These factors are accounted for in GLS by assigning different weights to each observation of the dependent variable used in the regression based on its contribution to the variance of the sample-flow statistic used as the dependent variable. Ordinary least-squares regression analysis assumes equal record length and variability for all sites, and no cross-correlation between flow records, so each sample flow statistic is assumed to have equal variance and is assigned equal weight in the regression.

Use of generalized least-squares regression to model the relations between the estimated peak discharges and basin characteristics of the sites in this study required estimates of the standard deviations, effective record length, and cross-correlation coefficients of the series of annual peaks. A generalized least-squares regression of sample standard deviations against drainage area and impervious area was used to estimate the standard deviations of annual peaks.

Effective record length is used to compare the reliability of estimates of peak discharge generated from synthetic record to estimates generated from actual record. Lichty and Liscum (1978) developed estimates of effective record length for sites where synthetic record was used. Those estimates were used in this study for all sites except the five sites from Sauer and others (1983), for which actual record length were used.

Average cross-correlations among sites from the same city or source were estimated using sample cross correlations. These correlations were estimated to be 0.88 for urban sites in Raleigh and Fayetteville, where the same historic rainfall record was used, and 0.96 for urban sites in Asheville. Correlations for sites taken from Putnam (1972) and Gunter and others (1987) were estimated to be 0.90 in both cases. Cross correlations for the sites taken from Sauer and others (1983) were found to be much less than those correlations among sites with synthetic records; an average value of 0.40 was used for all of these sites. Average cross correlations between sites from different cities or sources also were estimated using sample cross correlations. A value of 0.88 was used for the cross correlation between Raleigh and Fayetteville sites, because the same historic rainfall record was used to generate synthetic record at sites in both cities. Cross correlation between other cities or sources of record was estimated to be 0.30, except between sites taken from Gunter and others (1987) and sites taken either from Putnam (1972) or those from Asheville, and between sites from Sauer and others (1983) and any other source. Those correlations were assumed to be zero.

Uncertainty in a prediction of a flow value at an ungaged site using the regression models can be measured by the standard error of the prediction. This standard error is computed as the square root of the mean square error of the prediction, MSE_p, which is made up of two parts—the mean square error as a result of having an imperfect model, γ^2, and the sampling mean square error as a result of estimating model parameters from a sample, $MSE_{s,i}$. The mean square model error, γ^2, is assumed to be constant for all sites, but $MSE_{s,i}$ depends on the values of the explanatory variables (DA, IA, and RQ) at the site indexed by i. Therefore, the standard error of a prediction at site i, given by

$$S_{p,i} = (\gamma^2 + MSE_{s,i})^{1/2} \qquad (1)$$

varies from site to site. By assuming that the values of the basin characteristics of the gaging stations in the regression are a representative sample of all sites in the area, the average prediction accuracy for the regression model can be appraised by the average standard error of prediction, given by

$$S_p = \left(\gamma^2 + \frac{1}{n}\sum_{i=1}^{n} MSE_{s,i}\right)^{1/2}. \qquad (2)$$

The standard error of the model can be transformed from log (base 10) units to percent by the formula:

$$SE_{model}(\text{in percent}) = 100[e^{(5.302*\gamma^2)}-1]^{0.5}. \qquad (3)$$

Similarly, the average standard error of prediction can be transformed from log units to percent by substituting S_p^2 for γ^2 in the formula above. These standard errors are shown for each equation in table 7. The procedure for computing $S_{p,i}$ for an ungaged site is given in the Appendix.

Generalized least-squares regression was used to evaluate two 3-variable models that had been suggested by preliminary ordinary least-squares regression. The final regression model relates peak discharge to drainage area, relative impervious area, and rural equivalent flood discharge for each recurrence interval (table 7). A regression model using BDF was tested using generalized least-squares regression, and BDF was not found to be significant at the 1-percent level. Model standard error and average prediction standard errors were generally less than 40 percent for all recurrence-interval regression equations. These errors are less than the standard errors of estimate and prediction errors reported by Sauer and others (1983) for both the 3-parameter and 7-parameter national urban equations.

Regression Bias and Sensitivity

The regression relations were tested for parameter and geographic bias. Parameter bias was tested by plotting the residuals (differences between the equation estimates and the observed values, as defined in table 4) against the independent variables, drainage area and impervious area, for each equation. Rural equivalent discharge was not used because it is also a function of drainage area. Visual inspection of these plots showed no tendency to overestimate or underestimate within a given range of parameter values and were found to be free of parameter bias. Geographic bias was tested by plotting equation estimates against observed values for each of the six sources of data--three sets from modeled cities and three sets from previous investigations. These six plots for each regression equation also were visually inspected to determine if there was a tendency for overestimation or underestimation. Only the plots for Raleigh and Fayetteville indicated a tendency to overpredict at recurrence intervals greater than 25 and 50 years, respectively. This may result from the use of the same historic rainfall record for Raleigh and Fayetteville in developing the flood-frequency data applied to the regression database. However, the percentage differences are on the order of the model prediction errors, with only two sites in Raleigh and one site in Fayetteville recording greater than 50-percent error.

The sensitivity of the equations to errors in the independent variables, DA and IA, was evaluated by varying each variable by selected percentages and comparing the resulting peak-flow estimates. The sensitivity of the 2-year and 100-year estimated peak discharge to changes in drainage area and impervious area for a site in the Blue Ridge-Piedmont hydrologic area is presented in table 8. The table indicates that a 50-percent underestimation of drainage area will result in a 33.7-percent underestimation of the resulting 100-year peak discharge. Similarly a 50-percent

Table 8. Sensitivity of the 2-year and 100-year computed urban peak discharges to errors in independent variables in the Blue Ridge-Piedmont hydrologic area

[DA, contributing drainage area; IA, impervious area; mi^2, square miles]

Percent change in independent variable	Percent change in computed peak discharge: 2-year, DA (mi^2)	2-year, IA (percent of overall basin area)	100-year, DA (mi^2)	100-year, IA (percent of overall basin area)
-50	-40.1	-37.8	-33.7	-22.0
-40	-31.5	-29.6	-26.1	-16.7
-30	-23.2	-21.7	-19.1	-12.0
-20	-15.2	-14.2	-12.4	-7.7
-10	-7.5	-7.0	- 6.1	-3.7
10	7.3	6.8	5.8	3.5
20	14.4	13.3	11.4	6.7
30	21.4	19.7	16.8	9.8
40	28.2	26.0	22.1	12.8
50	35.0	32.1	27.2	15.6

overestimation of the percentage of impervious area will result in a 32.1-percent overestimation of the resulting 2-year peak discharge. Each set of equations is more sensitive to drainage area than impervious area because drainage area also is applied in computing the rural equivalent discharge used in the regression equations. Furthermore, the 2-year equation is more sensitive to changes in independent variables than the 100-year equation because the exponents are greater. Sensitivity of the regression equations for the 5-, 10-, 25-, and 50-year equations are expected to behave similarly. This analysis is specific to the Blue Ridge-Piedmont area; however, the sensitivity to errors in the independent variables for the Coastal Plain and Sand Hills hydrologic areas are expected to be comparable.

Use of Regression Equations

Within applicable ranges of basin characteristics, and in the absence of sustainable in-channel detention or basin storage, flood-frequency relations can be used to estimate flood-frequency characteristics of ungaged, urban streams in North Carolina.

To apply the relations, estimates of the input variables are first needed. Flood-peak discharges for the 2-, 5-, 10-, 25-, 50- and 100-year recurrence intervals can be estimated by first determining the drainage area and impervious area for the site of interest. The site must then be located within one of the three hydrologic areas to determine the rural equivalent discharge for the specified recurrence interval (table 6). The corresponding urban equation (table 7), which essentially adjusts the rural equivalent discharge, can then be applied. An example for the Blue Ridge-Piedmont hydrologic area is listed below.

Applying Gunter and other's (1987) rural flood-frequency equation for the 10-year peak discharge (table 6) to a site in the Blue Ridge-Piedmont hydrologic area with a drainage area of 2.4 mi^2 yields a value of 598 cubic feet per second (ft^3/s). The rural equivalent discharge is then adjusted for urbanization by applying the urban flood-frequency relation which incorporates the relative impervious area of 18.6 percent at the site. The resulting 10-year urban peak discharge from the estimating equations (table 7) is 974 ft^3/s— more than a 60-percent increase in predicted peak discharge for the rural area.

Users are cautioned against using the equations outside of the range of values of the independent variables used to develop the equations. The relative impervious area ranged from a minimum of 2 percent to a maximum of 54.6 percent and total drainage area ranged from 0.04 to 41.0 mi^2. In basins with relative impervious areas less than 10 percent, the computed urban peak discharge may be less than the computed rural peak discharge. It is left to the discretion of each user, based on hydrologic judgment and knowledge of the area, to decide which computed peak discharge to use. Furthermore, the equations are only applicable in basins with insignificant surface and embankment storage.

SUMMARY

Information on the magnitude and frequency of floods is critical in the design of bridges and culverts, establishment of flood-insurance rates and flood-plain regulation, and for other uses by urban planners and engineers. Recognizing the need for urban flood-frequency relations, the U.S. Geological Survey, in cooperation with the North Carolina Department of Transportation and the Cities of Asheville, Fayetteville, and Raleigh conducted an investigation of flood-frequency characteristics of small urban streams in North Carolina. The investigation included the collection of rainfall and runoff data and determination of basin characteristics in small urban basins, the calibration and application of rainfall-runoff models to synthesize peak-flow records, the analysis of the peak-flow record to generate flood-frequency characteristics, and the development of regression equations relating urban basin characteristics to flood-frequency characteristics. This report documents the development and application of regression equations quantifying the 2-, 5-, 10-, 25-, 50-, and 100-year flood flows in small ungaged urban basins in North Carolina, ranging in size from about 0.04 to 41 mi^2.

Concurrent records of rainfall and runoff data were collected in small urban basins in the Blue Ridge-Piedmont and Sand Hills hydrologic areas of North Carolina from 1986 through 1993. These data were used to calibrate rainfall-runoff models for each basin. Long-term record of peak flows were then generated by applying the historic rainfall and evaporation data to each set of calibrated model parameters. The maximum annual peak flows were then selected from each record and fitted to a Pearson-Type III distribution. Using this distribution, flows corresponding to the 2-, 5-, 10-, 25-, 50-, and 100-year recurrence intervals were computed for each modeled basin. Estimates of the 2-, 5-, 10-, 25-, 50-, and 100-year peak discharges also were compiled from previous studies to expand the database.

A database of these flood characteristics was coupled with seven hydrologically relevant basin characteristics describing the size and shape of each basin, level of urban development, and climate and rural flood characteristics for use in developing equations relating flood frequency to basin characteristics. Generalized least-squares regression techniques were used to develop the estimating equations that relate peak discharges for 2-, 5-, 10-, 25-, 50-, and 100-year floods to drainage area, relative impervious area, and rural equivalent flood discharge. Model error or regression error and prediction errors were generally less than 40 percent for all recurrence-interval regression equations. These errors are less than the regression errors and prediction errors previously reported for the 3-parameter and 7-parameter national urban equations.

The regression relations were tested for both parameter and geographic bias. Visual inspection of plots of residuals against independent variables showed no tendency to overestimate or underestimate within a given range of parameter values and were free of parameter bias. Geographic bias was tested by plotting equation estimates against observed values for each of the six sources of data--three sets from modeled cities and three sets from previous investigations. Only the plots for Raleigh and Fayetteville indicated a tendency to overpredict at recurrence intervals greater than 25 and 50 years, respectively. However, the percentage differences are on the order of the model prediction errors, with only two sites in Raleigh and one site in Fayetteville recording greater than 50-percent error.

The sensitivity of the equations to errors in the independent variables was evaluated by systematically varying each of the independent variables by selected percentages and comparing the resulting peak-flow estimates. Each set of equations is more sensitive to drainage area than impervious area because drainage area is also applied in computing the rural equivalent discharge used in the regression equations. Furthermore, the 2-year equation is more sensitive to changes in independent variables than the 100-year because the exponents of the independent variables are greater.

Within applicable ranges of basin characteristics and in the absence of sustainable in-channel detention or basin storage, flood-frequency relations can be used to estimate flood-frequency characteristics of small, ungaged, urban streams in North Carolina. The user is cautioned against using the equations outside of the range of values of the independent variables used to develop the equations. The relative impervious area ranged from a minimum of 2 percent to a maximum of 54.6 percent and total drainage area ranged from 0.04 to 41.0 mi^2.

REFERENCES

Benson, M.A., and Dalrymple, Tate, 1967, General field and office procedures for indirect discharge measurements: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A1, 30 p.

Bohman, L.R., 1992, Determination of flood hydrographs for streams in South Carolina: Volume 2. Estimation of peak-discharge frequency, runoff volumes, and flood hydrographs for urban watersheds: U.S. Geological Survey Water-Resources Investigations Report 92-4040, 79 p.

Boning, C.W., 1974, User's guide for a U.S. Geological Survey rainfall-runoff model: U.S. Geological Survey Open-File Report 74-33, 252 p.

Carrigan, P.H., Jr., 1973, Calibration of U.S. Geological Survey rainfall-runoff model for peak flow synthesis - natural basins: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A7, 28 p.

Carrigan, P.H., Jr., Dempster, G.R., Jr., and Bower, D.E., 1977, User's guide for U.S. Geological Survey rainfall-runoff models - revision of Open-File Report 77-33: U.S. Geological Survey Open-File Report 77-884, 260 p.

Clark, C.O., 1945, Storage and the unit hydrograph: American Society of Civil Engineers Transactions, v. 110, p. 1419-1488.

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.

Gunter, H.C., Mason, R.R., and Stamey, T.C., 1987, Magnitude and frequency of floods in rural and urban basins of North Carolina: U.S. Geological Survey Water-Resources Investigation Report 87-4096, 52 p.

Hershfield, D.M., 1961, Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years: U.S. Department of Commerce, National Weather Bureau, Technical Paper no. 40, 113 p.

Inman, E.J., 1988, Flood-frequency relations for urban streams in Georgia: U.S. Geological Survey Water-Resources Investigations Report 88-4085, 36 p.

_____ 1995, Flood-frequency relations for urban streams in Georgia--1994 update: U.S. Geological Survey Water-Resources Investigations Report 95-4017, 27 p.

Kohler, M.A., Nordenson, T.J., and Baker, D.R., 1959, Evaporation Maps for the United States: U.S. Department of Commerce, National Weather Bureau, Technical Paper no. 37, 13 p.

Lichty, R.W., and Liscum, Fred, 1978, A rainfall-runoff modeling procedure for improving estimates of T-year annual floods for small drainage basins: U.S. Geological Survey Water-Resources Investigations 78-7, 44 p.

Philip, J.R., 1954, An infiltration equation with physical significance in Proceedings: Soil Science Society of America, v. 77, p. 153-157.

Putnam, A.L., 1972, Effects of urban development on floods in the Piedmont Province of North Carolina: U.S. Geological Survey Open-File Report, 87 p.

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

SAS Institute, Inc., 1989, SAS user's guide: Statistics, 583 p.

Sauer, V.B., Thomas, W.O., Jr., Stricker, V.A., and Wilson, K.V., 1983, Flood characteristics of urban watersheds in the United States: U. S. Geological Survey Water-Supply Paper 2207, 63 p.

Sherwood, J.M., 1994, Estimation of peak-frequency relations, flood hydrographs, and volume-duration-frequency relations of ungaged small urban streams in Ohio: U.S. Geological Survey Water-Supply Paper 2432, 42 p.

Stedinger, J.R., and Tasker, G.D., 1985, Regional hydrologic analysis 1--ordinary, weighted, and generalized least squares compared: American Geophysical Union, Water-Resources Research, v. 21, no. 9, p. 1421-1432.

Tasker, G.D., and Stedinger, J.R., 1989, An operational GLS model for hydrologic regression: Journal of Hydrology, v. 111, p. 361-375.

U.S. Water Resources Council, 1981, Guidelines for determining flood flow frequency: U.S. Water Resources Council Bulletin 17B, 183 p.

GLOSSARY

The following are definitions for selected acronyms and terms used in this report; however, they are not necessarily the only definitions for these acronyms and terms.

BDF Basin-development factor. A measure of basin development that takes into account channel improvements, impervious channel linings, storm sewers, and curb and gutter streets. It is measured on a scale from 0 (little or no development) to 12 (fully developed).

DA Contributing drainage area, in square miles. The area of a basin that contributes drainage to a specified location on a stream, measured in a horizontal plane. It is usually computed by planimeter, digitizer, or grid method from U.S. Geological Survey 7.5-minute topographic quadrangle maps.

IA Impervious area. The percentage of drainage area covered by impervious surfaces, such as streets, parking lots, and buildings.

L Channel length, in miles, from the basin outlet to the basin divide.

$RI_{2,2}$ 2-year, 2-hour rainfall amount, in inches, reported in Hershfield (1961).

RQ2 Recurrence interval of a 2-year flood, in cubic feet per second; RQ5 is for a 5-year flood, and so forth up to and including RQ100 for a 100-year flood.

RRM USGS rainfall-runoff model. A lumped parameter model that assumes no significant storage in a basin and uniform rainfall throughout the basin.

S Channel slope, in feet per mile, measured from points 10 percent and 85 percent of the main channel length from the basin outlet.

SER Average standard error of regression, in percent. A measure of the error associated with estimating a streamflow characteristic of a site used in the regression analysis.

T Recurrence interval, in years. An average interval of time in which a given hydrologic event will be equaled or exceeded once.

Peak flow The maximum discharge, in cubic feet per second, associated with an observed or estimated flood hydrograph.

Urban basin A basin for which the basin development factor (BDF) is generally greater than 3.

APPENDIX

The value of the mean square error (MSE_s) at a specific site can be estimated as follows: Denote the column vector of n logarithms of observed peak-discharge characteristics at n sites in a region by Y. For example,

$$Y = \begin{bmatrix} \log Q_{50,\,1} \\ \log Q_{50,\,2} \\ "" \\ "" \\ \log Q_{50,\,n} \end{bmatrix},$$

in which, $Q_{50,i,}$ represents the observed 50-year peak at the ith gaging station in the region. Further, let X represent a (n by p) matrix of p-1 basin characteristics augmented by a column of ones at n gaging stations and B represent a column vector of p regression coefficients.

For example,

$$X= \begin{bmatrix} 1 & \log\left(DA_1\right) & \log\left(IA_1\right) & \log\left(RQ50_1\right) \\ 1 & \log\left(DA_2\right) & \log\left(IA_2\right) & \log\left(RQ50_2\right) \\ - & - & - & - \\ 1 & \log\left(DA_n\right) & \log\left(IA_n\right) & \log\left(RQ50_n\right) \end{bmatrix} \text{ and } B = \begin{bmatrix} a \\ b_1 \\ b_2 \\ b_4 \end{bmatrix}.$$

The linear model can be written as

$$Y=XB$$

The mean square sampling error, $MSE_{s,0}$, for an ungaged site with basin characteristics given by the row vector $\boldsymbol{x}_0$=[1 log (DA_0) log (IA_0) log ($RQ50_0$)], for example, is calculated as

$$MSE_{s,0} = x_0 \{X^T \Lambda^{-1} X\}^{-1} x_0{}^T$$

in which Λ is the (n by n) convariance matrix associated with Y. The diagonal elements of Λ are model error variance, γ^2, plus the time-sampling error for each site i, (i=1,2,3,...n) which is estimated as a function of a regional estimate of the standard deviation of annual peaks at site i, the recurrence interval of the dependent variable and the number of years of record at site i. The off-diagonal elements of Λ are the sample covariance of the estimated t-year peaks at sites i and j. These off-diagonal elements are estimated as a function of a regional estimate of the standard deviation of annual peaks at sites i and j, the recurrence interval of the dependent variable and the number of concurrent years of record at sites i and j (Tasker and Stedinger, 1989). The (p by p) matrix $\{X^T \Lambda^{-1} X\}^{-1}$ for each equation is given in Appendix table 1. The mean square error of a prediction, in log (base 10) units, at specific ungaged sites can be estimated as

$$MSE_{p,0}= (\gamma^2 + MSE_{s,0}).$$

The standard error of a prediction, $SE_{prediction,}$ in percent, can be calculated as

$$SE_{prediction}=100\{e^{5.302*(MSEp,0)}-1\}^{0.5}.$$

Appendix table 1. Matrix $\{X^T \Lambda^{-1} X\}^{-1}$ for the equations in table 7 (p. 14)

[These matrices can be used to compute the standard error of prediction and prediction intervals as explained in the text. Numbers are given in scientific notation— for example, 0.43958E-01 = 0.43958*10^{-1} = 0.043958.]

2-year recurrence interval (γ^2 = .19252E-01)

0.90869E-01	0.13294E-01	-0.18773E-01	-0.28369E-01
0.13294E-01	0.11280E-01	0.17836E-02	-0.96925E-02
-0.18773E-01	0.17836E-02	0.10257E-01	0.15357E-02
-0.28369E-01	-0.96925E-02	0.15357E-02	0.13137E-01

5-year recurrence interval (γ^2 = .18115E-01)

0.93973E-01	0.13749E-01	-0.17563E-01	-0.28227E-01
0.13749E-01	0.98769E-02	0.16057E-02	-0.85002E-02
-0.17563E-01	0.16057E-02	0.95948E-02	0.14128E-02
-0.28227E-01	-0.85002E-02	0.14128E-02	0.11775E-01

10-year recurrence interval (γ^2 = .18320E-01)

0.98930E-01	0.14865E-01	-0.17340E-01	-0.29275E-01
0.14865E-01	0.96007E-02	0.15068E-02	-0.83182E-02
-0.17340E-01	0.15068E-02	0.95444E-02	0.13891E-02
-0.29275E-01	-0.83182E-02	0.13891E-02	0.11591E-01

25-year recurrence interval (γ^2 = .20901E-01)

0.10049	0.17014E-01	-0.14689E-01	-0.30413E-01
0.17014E-01	0.85300E-02	0.15477E-02	-0.80817E-02
-0.14689E-01	0.15477E-02	0.85430E-02	0.10392E-02
-0.30413E-01	-0.80817E-02	0.10392E-02	0.11441E-01

50-year recurrence interval (γ^2 = .19799E-01)

0.10181	0.16168E-01	-0.14678E-01	-0.29671E-01
0.16168E-01	0.80657E-02	0.15708E-02	-0.74821E-02
-0.14678E-01	0.15708E-02	0.83678E-02	0.10437E-02
-0.29671E-01	-0.74821E-02	0.10437E-02	0.10722E-01

100-year recurrence interval (γ^2 = .19479E-01)

0.10706	0.15683E-01	-0.15487E-01	-0.29922E-01
0.15683E-01	0.79234E-02	0.16867E-02	-0.71533E-02
-0.15487E-01	0.16867E-02	0.86145E-02	0.11040E-02
-0.29922E-01	-0.71533E-02	0.11040E-02	0.10401E-01