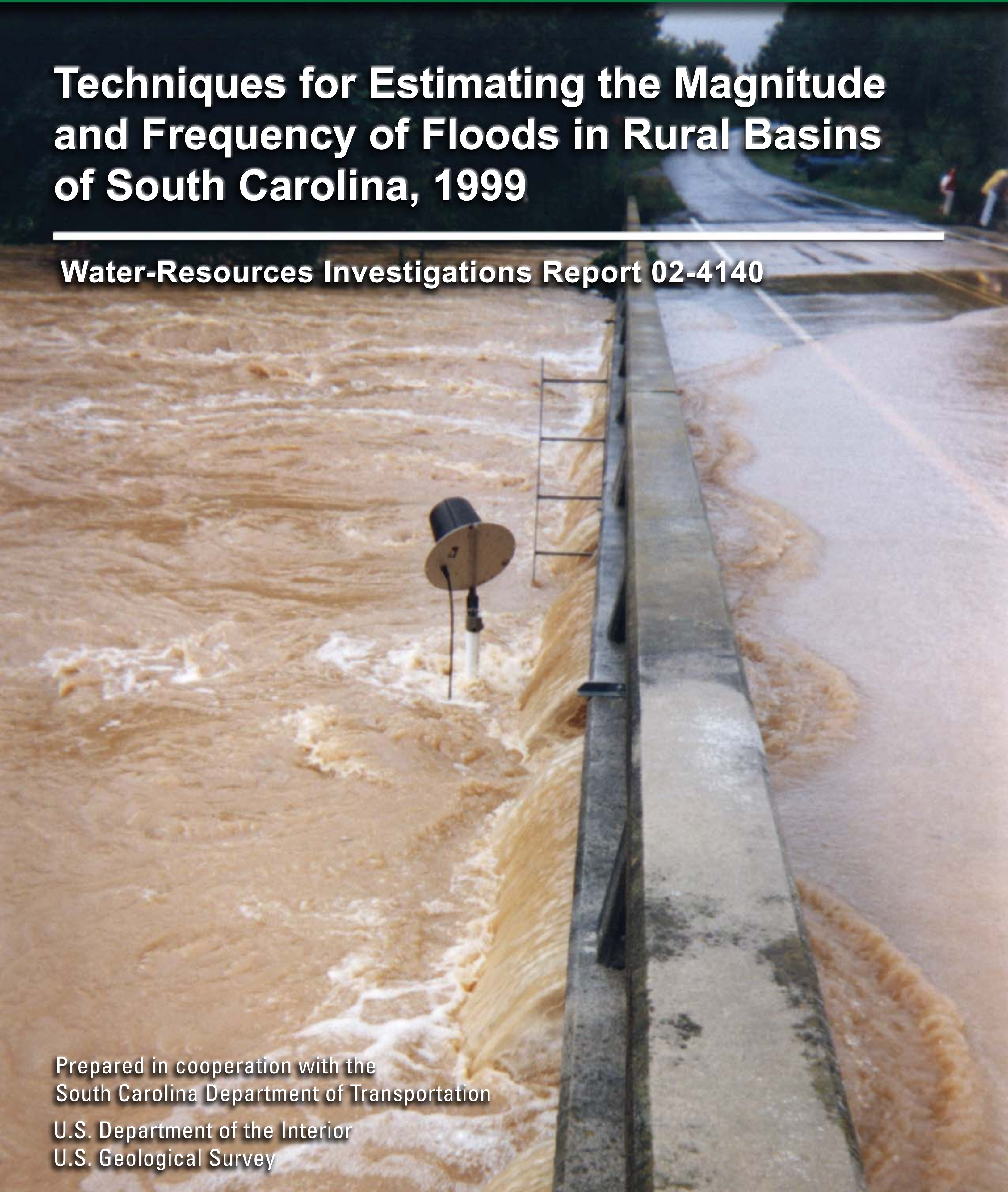
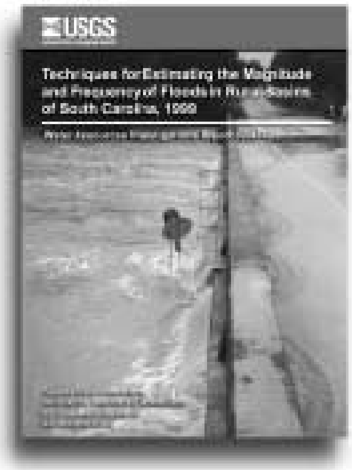
Techniques for Estimating the Magnitude and Frequency of Floods in Rural Basins of South Carolina, 1999

Water-Resources Investigations Report 02-4140



Prepared in cooperation with the
South Carolina Department of Transportation

U.S. Department of the Interior
U.S. Geological Survey



COVER PHOTOGRAPH: Station 02160390, Enoree River
near Woodruff, South Carolina. S.C. Highway 202 (August 27, 1995)

COVER PHOTOGRAPH: Station 02160390, Enoree River
near Woodruff, South Carolina. S.C. Highway 202 (September, 1993)

Cover Photographs by Michael Hall

Techniques for Estimating the Magnitude and Frequency of Floods in Rural Basins of South Carolina, 1999

By Toby D. Feaster and Gary D. Tasker

U.S. Geological Survey

Water-Resources Investigations Report 02-4140

U.S. DEPARTMENT OF THE INTERIOR
March 3, 1849

Columbia, South Carolina
2002

U.S. DEPARTMENT OF THE INTERIOR
GALE A. NORTON, Secretary

U.S. GEOLOGICAL SURVEY
Charles G. Groat, Director

For additional information
write to:

District Chief
U.S. Geological Survey
Suite 129
720 Gracern Road
Columbia, SC 29210-7651

Copies of this report can be
purchased from:

U.S. Geological Survey
Branch of Information Services
Box 25286
Denver, CO 80225-0286
888-ASK-USGS

Additional information about water resources in South Carolina is available on the World Wide Web at **http://sc.water.usgs.gov**

CONTENTS

Abstract 1
Introduction 2
    Purpose and Scope 2
    Previous Investigations 2
    Description of Study Area 2
    Acknowledgments 4
Peak-Flow Data 4
    Quality Assurance/Quality Control 5
Estimation of Flood Magnitude and Frequency at Gaging Stations 11
    Flood Frequency 11
    Skew Coefficient 12
    Synthesized Peak Flow 15
Estimation of Flood Magnitude and Frequency at Ungaged Sites 16
    Ordinary Least-Squares Analysis 16
    Generalized Least-Squares Analysis 18
    Accuracy of the Method 21
    Region-of-Influence Method 22
    Comparison of Methods 24
    Limitations 24
    Computer Software 24
Application of Methods 25
    Flood Frequency at Gaged Sites for Streams Draining One Physiographic Province 25
    Flood Frequency Near Gaged Sites on the Same Stream Draining One Physiographic Province 26
    Flood Frequency at Ungaged Sites on Streams Draining More Than One Physiographic Province 26
Flood Frequency at Gaged Sites on Streams Draining More Than One Physiographic Province 27
    Broad River 27
    Little Pee Dee River 29
    Lynches River 29
    Edisto River 29
    Salkehatchie River 29
Flood Frequency at Gaged Sites on Regulated Streams 30
    Pee Dee River 30
    Catawba River 30
    Wateree River 30
    Saluda River 31
    Congaree River 31
    Santee River 31
    Savannah River 31
    Pacolet River 32
Summary 32
Selected References 33

Appendix A: Water-year peak-flow and stage data at gaging stations in South Carolina through 1999 .... CD in pocket

Appendix B: Station, regional, and weighted flood estimates for selected recurrence intervals for rural streamflow stations in South Carolina used in the regionalization study .... CD in pocket

Appendix C: Generalized least-squares model description and assumptions .... CD in pocket

Appendix D: Water-year peak-flow and stage data at gaging stations in South Carolina on regulated streams and on streams draining more than one physiographic province through 1999 .... CD in pocket

FIGURES

1. Maps showing physiographic provinces and locations of streamflow gaging stations on rural streams in South Carolina and parts of North Carolina and Georgia 3
2–5. Graphs showing:
  2. Distribution of systematic peak-flow record lengths for rural streamflow stations used in the regionlization study 10
  3. Examples of distributions with zero skew, positive skew, and negative skew 12
  4. Relation between drainage area and 100-year recurrence-interval flow for stations on streams draining the Piedmont physiographic province in South Carolina, North Carolina, and Georgia....19
  5. Relation between drainage area and 100-year recurrence-interval flow for stations on streams draining the upper Coastal Plain physiographic province in South Carolina and North Carolina 20
6. Map showing locations of streamflow gaging stations on regulated streams and gaging stations on streams draining more than one physiographic province in South Carolina 27
7. Graph showing relation between drainage area and 100-year recurrence-interval flows for period of record and water years 1939-83 for Station 02153500, Broad River near Gaffney, S.C., Station 02156500, Broad River near Carlisle, S.C., and Station 02161500, Broad River at Richtex, S.C. 28

TABLES

1. Streamflow stations in South Carolina used to develop regional regression equations. 6
2. Streamflow stations in North Carolina and Georgia used to develop regional regression equations for South Carolina...7
3. Distribution of stations used in the regionalization analysis by State and physiographic province....10
4. Ordinary least-squares regression statistics for the generalized skew coefficients 13
5. Weighted generalized skew coefficients and associated statistics for rural South Carolina gaging stations 14
6. Sensitivity of South Carolina generalized skew coefficients (not weighted) to high and low values....15
7. Station and weighted flood-frequency data for the synthesized peak flows used in the regionalization analysis 16
8. South Carolina rural flood-frequency equations 21
9. Accuracy statistics for the regional regression equations, presented by physiographic province and recurrence interval....23
10. Ratio of PRESS (ROI) statistic to PRESS (RRE) statistic for each physiographic province 24

CONVERSION FACTORS, VERTICAL DATUM, AND ABBREVIATIONS

Multiply	By	To obtain
foot (ft)	0.3048	meter
mile (mi)	1.609	kilometer
acre	4,047	square meter
square mile (mi^2)	2.590	square kilometer
acre-foot (acre-ft)	1.233×10^3	cubic meter
cubic foot per second (ft^3/s)	0.02832	cubic meter per second

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. Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD83).

ACRONYMS

GLM	=	Generalized linear model	ROI	=	Region of influence
GLS	=	Generalized least squares	RRE	=	Regional regression equations
OLS	=	Ordinary least squares	R^2	=	Coefficient of determination
PRESS	=	Prediction error sum of squares	SCDOT	=	South Carolina Department of Transportation
QAQC	=	Quality assurance / quality control	USGS	=	U.S. Geological Survey

Techniques for Estimating the Magnitude and Frequency of Floods in Rural Basins of South Carolina, 1999

By Toby D. Feaster *and* Gary D. Tasker

ABSTRACT

Data from 167 streamflow-gaging stations in or near South Carolina with 10 or more years of record through September 30, 1999, were used to develop two methods for estimating the magnitude and frequency of floods in South Carolina for rural ungaged basins that are not significantly affected by regulation. Flood-frequency estimates for 54 gaged sites in South Carolina were computed by fitting the water-year peak flows for each site to a log-Pearson Type III distribution. As part of the computation of flood-frequency estimates for gaged sites, new values for generalized skew coefficients were developed. Flood-frequency analyses also were made for gaging stations that drain basins from more than one physiographic province. The U.S. Geological Survey, in cooperation with the South Carolina Department of Transportation, updated these data from previous flood-frequency reports to aid officials who are active in floodplain management as well as those who design bridges, culverts, and levees, or other structures near streams where flooding is likely to occur.

Regional regression analysis, using generalized least squares regression, was used to develop a set of predictive equations that can be used to estimate the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence-interval flows for rural ungaged basins in the Blue Ridge, Piedmont, upper Coastal Plain, and lower Coastal Plain physiographic provinces of South Carolina. The predictive equations are all functions of drainage area. Average errors of prediction for these regression equations ranged from -16 to 19 percent for the 2-year recurrence-interval flow in the upper Coastal Plain to -34 to 52 percent for the 500-year recurrence-interval flow in the lower Coastal Plain.

A region-of-influence method also was developed that interactively estimates recurrence- interval flows for rural ungaged basins in the Blue Ridge of South Carolina. The region-of-influence method uses regression techniques to develop a unique relation between flow and basin characteristics for an individual watershed. This, then, can be used to estimate flows at ungaged sites. Because the computations required for this method are somewhat complex, a computer application was developed that performs the computations and compares the predictive errors for this method. The computer application includes the option of using the region-of-influence method, or the generalized least squares regression equations from this report to compute estimated flows and errors of prediction specific to each ungaged site. From a comparison of predictive errors using the region-of-influence method with those computed using the regional regression method, the region-of-influence method performed systematically better only in the Blue Ridge and is, therefore, not recommended for use in the other physiographic provinces.

Peak-flow data for the South Carolina stations used in the regionalization study are provided in appendix A, which contains gaging station information, log-Pearson Type III statistics, information on stage-flow relations, and water-year peak stages and flows. For informational purposes, water-year peak-flow data for stations on regulated streams in South Carolina also are provided in appendix D. Other information pertaining to the regulated streams is provided in the text of the report.

INTRODUCTION

One of the most important factors in the design of bridges, highway embankments, culverts, levees, and other structures near streams is the magnitude and frequency of floods that are likely to occur during the life of the structure. Federal, State, regional, and local officials also need these data for effective floodplain management and to delineate areas susceptible to flooding.

In an effort to continue to improve the flood-frequency estimates for South Carolina, the U.S. Geological Survey (USGS), in cooperation with the South Carolina Department of Transportation (SCDOT), updated previous flood-frequency reports by incorporating additional data collected through the 1999 water year. A water year, which is the 12-month period from October 1 to September 30, is designated by the calendar year in which it ends. Thus, the 12-month period ending September 30, 1999, is called the “1999 water year.” Throughout this report, “peak flow” refers to the maximum peak for the water year.

Purpose and Scope

Two methods for predicting the magnitude and frequency of floods in South Carolina at ungaged, rural basins that are not significantly affected by regulation are presented with the estimated error of each method. The two methods are the regional regression method and the region-of-influence method (Tasker and Slade, 1994; Hodge and Tasker, 1995; Pope and Tasker, 1999). Flood-frequency estimates at streamflow-gaging stations, and methods for estimating the magnitude and frequency of floods at or near these streamflow-gaging stations also are provided. Peak-flow data were analyzed for 167 streamflow-gaging stations (54 in South Carolina, 65 in North Carolina, and 48 in Georgia) for streams that drain rural basins without significant regulation and have at least 10 years of peak-flow data through the 1999 water year (fig. 1; data for South Carolina stations shown in appendices A and B).

Previous Investigations

The earliest investigation of flood frequency of streams in South Carolina was made by Speer and Gamble (1964), who presented methods for estimating the magnitude of floods for selected recurrence intervals for streams in the South Atlantic slope basin. This area extends from the James River in Virginia to the Savannah River along the South Carolina-Georgia State line (Guimaraes and Bohman, 1991). Whetstone (1982) used multiple regression analyses to define the relation between flows and basin characteristics at recurrence intervals of 2, 5, 10, 25, 50, and 100 years, for unregulated, rural streams in South Carolina with drainage areas greater than 1.0 mi^2. Frequencies of peak flows were regionalized by Guimaraes and Bohman (1991) using generalized least squares regression methods to define the relation of magnitude and frequency of flows to various basin characteristics on ungaged, rural streams that were not significantly affected by regulation. Bohman (1992) described methods for determining flood-frequency relations for urban streams in South Carolina. This report updates and supersedes the previous flood-frequency report for rural streams in South Carolina by Guimaraes and Bohman (1991).

Description of Study Area

The study area includes the entire State of South Carolina, which is located on the South Atlantic slope adjacent to the Atlantic Ocean. The State has an area of 31,055 mi^2, and generally is divided into three major physiographic provinces: Blue Ridge, Piedmont, and Coastal Plain (Cooke, 1936). The Coastal Plain is further divided into the upper Coastal Plain and the lower Coastal Plain. The physiographic provinces and locations of data-collection sites in South Carolina and in adjacent areas in North Carolina and Georgia are shown on figure 1.

The Blue Ridge physiographic province in South Carolina is a mountainous region of steep terrain with some stream gradients greater than 250 ft per mile (Bloxham, 1979). Land-surface elevation ranges from

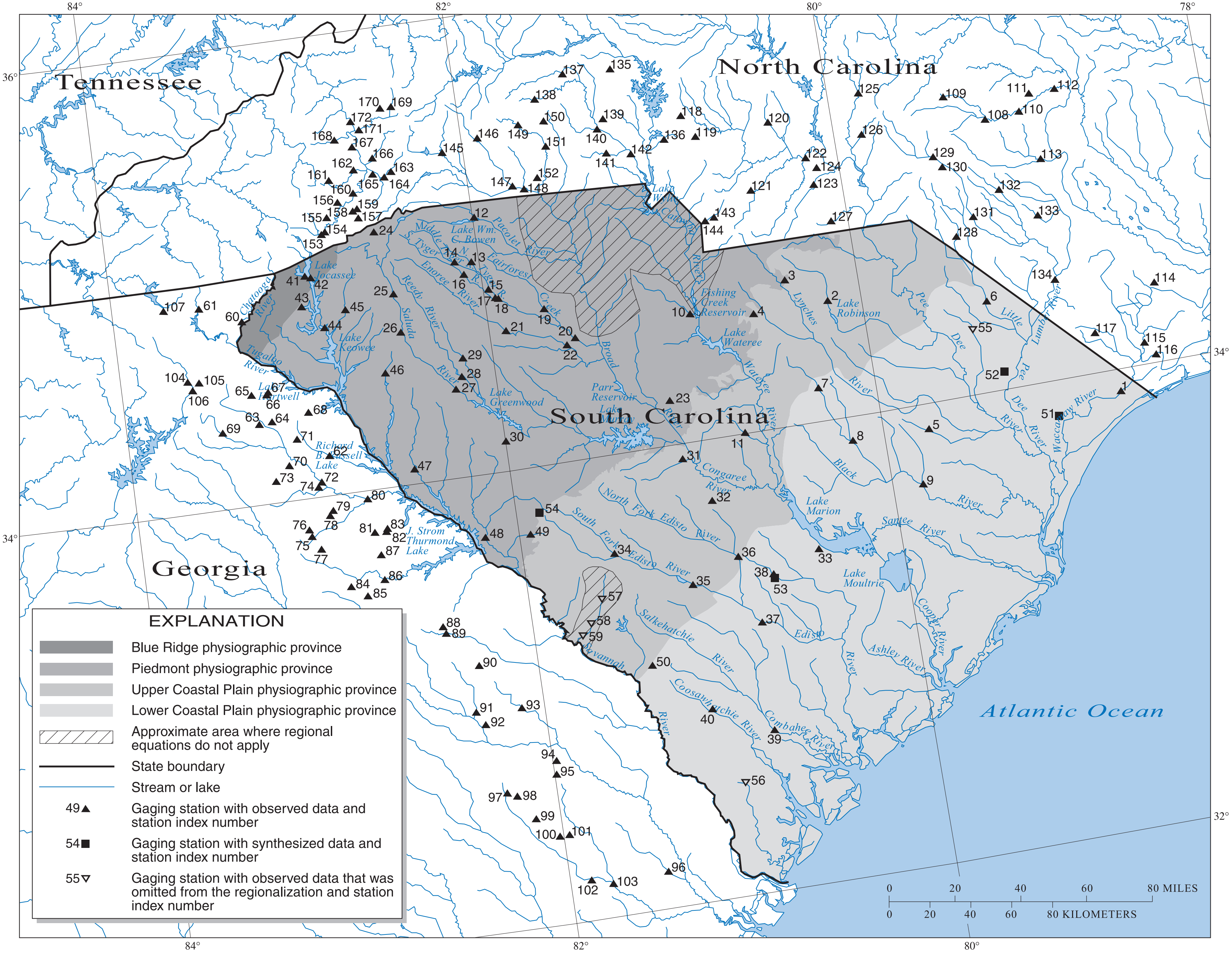


Figure 1. Physiographic provinces and locations of streamflow gaging stations on rural streams in South Carolina and parts of North Carolina and Georgia.

1,000 to more than 3,500 ft above sea level. Surface fractures in crystalline rock provide channels for runoff. Overlying the crystalline bedrock is a layer of weathered bedrock or saprolite. Although some rainfall infiltrates the saprolite layer, the steep-sided slopes and semipermeable soils in this region cause much of the rainfall to run off rapidly into stream channels (South Carolina Water Resources Commission, 1983).

Rolling hills, elongated ridges, and moderately deep to shallow valleys characterize the Piedmont physiographic province of South Carolina. Piedmont land-surface elevations range from about 1,000 ft above sea level at the Blue Ridge foothills to about 400 ft above sea level at the Fall Line, which is the name given to the boundary between the Piedmont and Coastal Plain physiographic provinces. In general, this boundary is characterized by a series of rapids or falls where the streams tumble off the more resistant rocks of the Piedmont into the deeper valleys worn in the softer sediments of the Coastal Plain (Cooke, 1936). The Piedmont is underlain by fractured crystalline rock consisting of intrusive granite, gneiss, schist, and metamorphosed volcanic rock. Most overlying soil is moderately to poorly permeable silty clay loams. Alluvial deposits of clay, silt, and sand are found along the valley floors (Bloxham, 1981).

Gradual slopes and rounded summits characterize the upper Coastal Plain physiographic province in South Carolina, although there are several areas of intensely irregular terrain. Some hilltop elevations exceed 700 ft above sea level near the Fall Line, but land-surface elevations commonly are less than 200 ft above sea level at the boundary of the lower Coastal Plain. Extensive swamps and very wide floodplains are common to the four large, through-flowing rivers (fig. 1) (Bloxham, 1976).

In the lower Coastal Plain physiographic province, the land surface slopes from elevations of about 200 ft above sea level near the boundary of the upper Coastal Plain to sea level at the coast. Topographic relief in this area is much less than that in other areas of the State, and small stream drainage patterns are characteristically more erratic in the seaward direction. Large parts of the lower Coastal Plain river systems are swamplands. The highly permeable soils in this region are similar to those of the upper Coastal Plain, which readily absorb rainfall and retard runoff to stream channels, causing streamflow to rise and fall gradually (Bloxham, 1981).

Acknowledgments

The authors gratefully acknowledge the assistance and support of William H. Hulbert formerly with the SCDOT. Mr. Hulbert worked cooperatively with the USGS on this and numerous other projects in South Carolina. The peak-flow data used in the analyses described in this report were collected throughout South Carolina and adjoining States at streamflow gages operated in cooperation with a variety of Federal, State, and local agencies. The authors also acknowledge the dedicated work of the USGS field office staff in collecting, processing, and storing the peak-flow data necessary for the completion of this report.

PEAK-FLOW DATA

The empirical basis for estimating specific recurrence-interval flows is the water-year peak flows collected at streamflow-gaging stations. The first streamflow data collected in the study area was for the Savannah River at Augusta, Ga. The U.S. Weather Bureau, now the National Weather Service, began collecting data at this site in 1884. By 1930, streamflow data were collected at a network of sites by the USGS, in cooperation with the South Carolina State Highway Department, now the SCDOT (Guimaraes and Bohman, 1991). By 1999, the USGS systematic data-collection program in South Carolina included 118 streamflow gaging stations and 53 crest-stage partial-record sites. These data are collected in cooperation with many Federal, State, and local agencies in South Carolina (Cooney and others, 1999). With World Wide Web Internet access, peak-flow data can be obtained at http://water.usgs.gov/sc/nwis/sw.

At continuous-record stations, the water-surface elevation, or stage, of the stream is recorded at fixed intervals, typically ranging from 5 to 60 minutes. At crest-stage partial-record stations, only the crest, or highest, stages that occur between site visits, usually 6 to 8 weeks, are recorded. Measurements of flow are determined throughout the range of recorded stages and a relation between stage and flow is developed for the gaged site. Using this stage-flow relation, or rating, flows for recorded stages are estimated. Because stream channels are dynamic, periodic streamflow measurements are made to verify that the hydraulic conditions at the site remain stable. If the measurements indicate conditions have changed, additional data are collected and used to make adjustments to the stage-streamflow relation. At some crest-stage sites, indirect-flow computation methods are used to develop a theoretical rating. This method has been used extensively to compute flows from small drainage areas (Bodhaine, 1968).

Data from 167 gaging stations on streams in South Carolina (54 stations) and adjacent areas of North Carolina (65 stations) and Georgia (48 stations) were used in the development of regionalized flood-frequency relations presented in this report (fig. 1, tables 1 and 2). The distribution of gaging stations used to develop these regression equations are presented by State and physiographic province in table 3. The distribution of systematic peak-flow record lengths used in the regional analyses is shown in figure 2. Appendix A of this report contains the following data for each of the South Carolina stations: station description, drainage area, type of data recorder, extreme gage heights and flows for the period of record, description of the stage-flow relation, and water-year peak stages and flows for the period of record. Flood-frequency data derived from the peak flows for South Carolina stations also are included in appendix A. Similar information can be obtained for the Georgia and North Carolina stations in reports by Stamey and Hess (1993) and Pope and others (2001), respectively. Those flood-frequency data, however, may be slightly different from the data used in this study because of additional data that may have been collected since those studies were completed and because different regionalized skew coefficients were used. Station, regional, and weighted flows for selected recurrence intervals for the South Carolina stations used in the regional regression are presented in appendix B.

Five stations for which streamflow data are available were not included in the regionalization analysis for various reasons (fig. 1). Stations 02197300, 02197310, and 02197315 (map index numbers 57, 58, and 59, respectively, fig. 1) in the Upper Three Runs basin of the upper Coastal Plain physiographic province were not included in the regionalization analysis because of the effects of large sand deposits in the upper end of the basin on rainfall runoff. Substantial amounts of rainfall runoff are stored in the sand deposits, diminishing the magnitude of peak flows for the Upper Three Runs basin compared to other basins in the upper Coastal Plain province. Therefore, the regionalized flood-frequency equations developed for the upper Coastal Plain province do not apply to this basin or to other basins physiographically similar to the Upper Three Runs basin (Guimaraes and Bohman, 1991). In addition, Lanier (1996), noted that Station 02197315 is affected by backwater from the Savannah River and, therefore, the data are not suitable for a flood-frequency analysis.

Catfish Canal near Conway, S.C. (map index number 55, station number 02131150) and Great Swamp near Ridgeland, S.C. (map index number 56, station number 02176875) were not included in the regional analysis because the flows are affected by channelization.

Quality Assurance/Quality Control

For this study, a comprehensive review was made of the peak-flow data. A minimum of 10 years of record was required for a station to be included in the analysis. The data at each station also were reviewed for homogeneity (time trends), which implies relatively constant watershed conditions during the period of record. The Kendall's tau statistic was chosen to assess the homogeneity of the record at each station. If it was determined that a station's record was not homogeneous, the station was excluded from the analysis. Additionally, the drainage basin for each station had to be substantially unaffected by regulation or urbanization. The peak-flow data for the stations that met these minimum criteria were then reviewed for quality assurance and quality control (QA/QC). Several computer programs were developed using commercial statistical software to automate the QA/QC reviews (C.L. Sanders, Jr., U.S. Geological Survey, written communs., December 1999).

The QA/QC computer program for reviewing continuous-record stations performs the following checks:

1. Computes a Kendall's tau to check for trends in the data over time;
2. Plots the peak streamflow by water year;
3. Plots peak stage against peak streamflow, which are then overlaid on the most recent rating to check for significant or abrupt changes;
4. Plots peak streamflow with the daily-value hydrographs by water year to compare the date of the peak with the date of the water-year maximum daily flow;

Table 1. Streamflow stations in South Carolina used to develop regional regression equations

Map index number (fig. 1)	Station number	Station name and location	Drainage area (in square miles)
		Blue Ridge	
24	02162350	Middle Saluda River near Cleveland	21.0
41	02184500	Whitewater River near Jocasse	47.3
42	02185000	Keowee River near Jocasse	148
43	02185200	Little River near Walhalla	72.0
44	02185500	Seneca River near Newry	455
		Piedmont	
3	02131309	Fork Creek near Jefferson	24.3
4	02131472	Hanging Rock Creek near Kershaw	23.7
10	02147500	Rocky Creek at Great Falls	194
12	02154500	North Pacolet River at Fingerville	116
13	02157000	North Tyger near Fairmont	44.4
14	02157500	Middle Tyger River at Lyman	68.3
15	02158000	North Tyger River near Moore	162
16	02158500	South Tyger River near Reidville	106
17	02159000	South Tyger River near Woodruff	174
18	02159500	Tyger River near Woodruff	351
19	02160000	Fairforest Creek near Union	183
20	02160105	Tyger River near Delta	759
21	02160500	Enoree River near Enoree	307
22	02160700	Enoree River at Whitmire	444
23	02162010	Cedar Creek near Blythewood	48.9
25	02162500	Saluda River near Greenville	295
26	02163000	Saluda River near Pelzer	405
27	02163500	Saluda River near Ware Shoals	581
28	02165000	Reedy River near Ware Shoals	236
29	02165200	South Rabon Creek near Gray Court	29.5
30	02166970	Ninety Six Creek near Ninety Six	17.4
45	02186000	Twelve Mile Creek near Liberty	106
46	02187900	Broadway Creek near Anderson	26.4
47	02192500	Little River near Mt. Carmel	217
54	02195660	Log Creek near Edgefield	1.26
48	02196000	Stevens Creek near Modoc	545
49	02196250	Horn Creek near Colliers	13.9
		Upper Coastal Plain	
2	02130900	Black Creek near McBee	108
6	02132500	Little Pee Dee River near Dillon	524
7	02135300	Scape Ore Swamp near Bishopville	96.0
11	02148300	Colonel Creek near Leesburg	40.2
31	02169550	Congaree Creek near Cayce	122
32	02169630	Big Beaver Creek near St. Matthews	10.1
34	02172500	South Fork Edisto near Montmorenci	198
35	02173000	South Fork Edisto near Denmark	720
36	02173500	North Fork Edisto at Orangeburg	683
37	02174000	Edisto River near Branchville	1,720

Table 1. Streamflow stations in South Carolina used to develop regional regression equations (Continued)

Map index number (fig. 1)	Station number	Station name and location	Drainage area (in square miles)
		Lower Coastal Plain	
1	02110500	Waccamaw River near Longs	1,110
51	02110700	Crabtree Swamp near Conway	14.0
5	02132100	Two Mile Branch near Lake City	18.4
52	02135050	Reedy Creek near Rains	10.4
8	02135500	Black River near Gable	401
9	02136000	Black River at Kingstree	1,252
33	02169960	Lake Marion Tributary near Vance	2.12
38	02174250	Cow Castle Creek near Bowman	23.4
53	02174300	Buck Branch at Bowman	11.9
39	02176000	Combahee River near Yemassee	1,100
40	02176500	Coosawhatchie River near Hampton	203
50	02197410	Miller Creek Tributary near Baldoc	7.82

Table 2. Streamflow stations in North Carolina and Georgia used to develop regional regression equations for South Carolina

Map index number (fig. 1)	Station number	Station name and location	Drainage area (in square miles)
		Blue Ridge	
60	02177000	Chattooga River near Clayton, Ga.	207
61	02178400	Tallulah River near Clayton, Ga.	56.5
104	02331000	Chattahoochee River near Leaf, Ga.	150
105	02331500	Soque River near Demorest, Ga.	156
106	02331600	Chattahoochee River near Corneila, Ga.	315
107	03545000	Hiwasse River at Presley, Ga.	45.5
153	03439000	French Broad River at Rosman, N.C.	67.9
154	03439500	French Broad River at Calvert, N.C.	103
155	03440000	Catheys Creek near Brevard, N.C.	11.7
156	03441000	Davidson River near Brevard, N.C.	40.4
157	03441440	Little River above High Falls near Cedar Mt., N.C.	26.8
158	03441500	Little River near Penrose, N.C.	41.4
159	03442000	Crab Creek near Penrose, N.C.	10.9
160	03443000	French Broad River at Blantyre, N.C.	296
161	03444500	South Fork Mills River at The Pink Beds, N.C.	9.99
162	03446000	Mills River near Mills River, N.C.	66.7
163	03446410	Laurel Branch near Edneyville, N.C.	0.57
164	03446500	Clear Creek near Hendersonville, N.C.	42.2
165	03447000	Mud Creek at Naples, N.C.	109
166	03447500	Cane Creek at Fletcher, N.C.	63.1
167	03448000	French Broad River at Bent Creek N.C.	676
168	03448500	Hominy Creek at Candler, N.C.	79.8
169	03449000	North Fork Swannanoa River near Black Mountain, N.C.	23.8
170	03450000	Beetree Creek near Swannanoa, N.C.	5.46

Table 2. Streamflow stations in North Carolina and Georgia used to develop regional regression equations for South Carolina (Continued)

Map index number (fig. 1)	Station number	Station name and location	Drainage area (in square miles)
171	03451000	Swannanoa River at Biltmore, N.C.	130
172	03451500	French Broad River at Asheville, N.C.	945
		Piedmont	
108	02102908	Flat Creek near Inverness, N.C.	7.63
118	02124060	North Prong Clarke Creek near Huntersville, N.C.	3.61
119	02124130	Mallard Creek near Charlotte, N.C.	20.7
120	02125000	Big Bear Creek near Richfield, N.C.	55.6
121	02125410	Chinkapin Creek near Monroe N.C.	8.50
122	02126000	Rocky River near Norwood, N.C.	1,372
123	02127000	Brown Creek near Polkton, N.C.	110
124	02127390	Palmetto Branch at Ansonville, N.C.	0.90
125	02128000	Little River near Star, N.C.	106
126	02128260	Cheek Creek near Pekin, N.C.	15.4
127	02129440	South Fork Jones Creek near Morven, N.C.	16.7
135	02142480	Hagan Creek near Catawba, N.C.	7.80
136	02142900	Long Creek near Paw Creek, N.C.	16.4
137	02143000	Henry Fork near Henry River, N.C.	83.2
138	02143040	Jacob Fork at Ramsey, N.C.	25.7
139	02143310	Lithia Inn Branch near Lincolnton, N.C.	1.00
140	02143500	Indian Creek near Laboratory, N.C.	69.2
141	02144000	Long Creek near Bessemer City, N.C.	31.8
142	02145000	South Fork Catawba River at Lowell, N.C.	628
143	02146890	East Fork Twelve Mile Creek near Waxhaw N.C.	41.8
144	02146900	Twelve Mile Creek near Waxhaw, N.C.	76.5
145	02149000	Cove Creek near Lake Lure, N.C.	79.0
146	02150420	Camp Creek near Rutherfordton, N.C.	13.0
147	02151000	Second Broad River at Cliffside, N.C.	220
148	02151500	Broad River near Boiling Springs, N.C.	875
149	02152100	First Broad River near Casar, N.C.	60.5
150	02152420	Big Knob Creek near Fallston, N.C.	16.4
151	02152500	First Broad River near Lawndale, N.C.	200
152	02152610	Sugar Branch near Boiling Springs N.C.	1.42
62	02188500	Beaverdam Creek at Dewy Rose, Ga.	35.8
63	02189020	Indian Creek near Carnesville, Ga.	7.63
64	02189030	Stephens Creek tributary at Carnesville, Ga.	0.39
65	02189600	Bear Creek near Mize, Ga.	3.62
66	02190100	Toms Creek near Eastanollee, Ga.	3.79
67	02190200	Toms Creek tributary near Avalon, Ga.	1.20
68	02190800	Double Branch at Bowersville, Ga.	0.50
69	02191200	Hudson River at Homer, Ga.	60.9
70	02191270	Scull Shoal Creek near Danielsville, Ga.	8.75
71	02191280	Mill Shoal Creek near Royston, Ga.	0.32
72	02191300	Broad River above Carlton, Ga.	760
73	02191600	Double Branch near Danielsville, Ga.	4.77
74	02191750	Fork Creek at Carlton, Ga.	13.8
75	02191890	Brooks Creek near Lexington, Ga.	12.3
76	02191910	Trouble Creek at Lexington, Ga.	2.70
77	02191930	Buffalo Creek near Lexington, Ga.	5.60

Table 2. Streamflow stations in North Carolina and Georgia used to develop regional regression equations for South Carolina (Continued)

Map index number (fig. 1)	Station number	Station name and location	Drainage area (in square miles)
78	02191960	Macks Creek near Lexington, Ga.	3.45
79	02191970	Little Macks Creek near Lexington, Ga.	1.73
80	02192000	Broad River near Bell, Ga.	1,430
81	02192300	Hog Fork Fishing Creek tributary near Tignall, Ga.	0.10
82	02192400	Anderson Mill Creek near Danburg, Ga.	5.49
83	02192420	Anderson Mill Creek tributary near Danburg, Ga.	0.92
84	02193300	Stephens Creek near Crawfordville, Ga.	6.30
85	02193400	Harden Creek near Sharon, Ga.	3.98
86	02193500	Little River near Washington, Ga.	291
87	02193600	Rocky Creek near Washington, Ga.	1.14
88	02197520	Brier Creek near Thomson, Ga.	55.0
89	02197550	Little Brier Creek near Thomson, Ga.	24.0
		Upper Coastal Plain	
109	02102910	Dunhams Creek tributary near Carthage, N.C.	2.20
110	02103000	Little River at Manchester, N.C.	348
111	02103390	South Prong Anderson Creek near Lillington, N.C.	7.57
112	02103500	Little River at Linden, N.C.	459
113	02104500	Rockfish Creek near Hope Mills, N.C.	292
128	02132230	Bridge Creek tributary at Johns N.C.	6.20
129	02133500	Drowning Creek near Hoffman, N.C.	183
130	02133590	Beaverdam Creek near Aberdeen, N.C.	4.66
131	02133624	Lumber River near Maxton, N.C.	365
132	02133960	Raft Swamp near Red Springs, N.C.	39.8
133	02134380	Tenmile Swamp near Lumberton N.C.	16.1
134	02134500	Lumber River at Boardman, N.C.	1,228
		Lower Coastal Plain	
114	02108960	Buckhead Branch near Bolton, N.C.	15.3
115	02109500	Waccamaw River at Freeland, N.C.	680
116	02109640	Wet Ash Swamp near Ash, N.C.	16.0
117	02110020	Mill Branch near Tabor City, N.C.	3.80
90	02197600	Brushy Creek near Wrens, Ga.	28.0
91	02200900	Big Creek near Louisville, Ga.	95.8
92	02200930	Spring Creek near Louisville, Ga.	14.2
93	02201350	Buckhead Creek near Waynesboro, Ga.	64.0
94	02201800	Richardson Creek near Millen, Ga.	43.0
95	02201830	Sculls Creek near Millen, Ga.	4.38
96	02202600	Black Creek near Blitchton, Ga.	232
97	02202800	Canoochee Creek near Swainsboro, Ga.	46.0
98	02202820	Reedy Creek near Twin City, Ga.	9.36
99	02202850	Reedy Branch near Metter, Ga.	3.41
100	02202900	Fifteenmile Creek near Metter, Ga.	147
101	02202910	Tenmile Creek tributary at Pulaski, Ga.	1.14
102	02203000	Canoochee River near Claxton, Ga.	555
103	02203280	Canoochee River near Daisy, Ga.	833

Table 3. Distribution of stations used in the regionalization analysis by State and physiographic province

Physiographic province	South Carolina	North Carolina	Georgia	Total
Blue Ridge	5	20	6	31
Piedmont	27	29	28	84
Upper Coastal Plain	10	12	0	22
Lower Coastal Plain	12	4	14	30
Totals	54	65	48	167

5. Computes the peak gage height using the peak streamflow and the current rating, then compares the computed value with the peak gage height listed in the peak-flow file;

6. Computes the peak streamflow from the peak gage height and the current rating, then compares the computed value with the peak streamflow listed in the peak-flow file;

7. Relates the peak streamflow with the water-year maximum daily-value streamflow by regression;

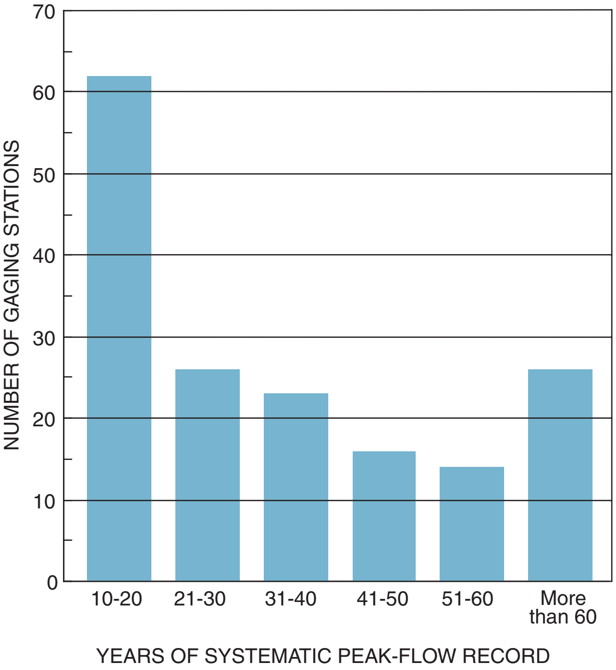


Figure 2. Distribution of systematic peak-flow record lengths for rural streamflow stations used in the regionalization study.

Wateree River dam being overtopped by high flows on July 25, 1997 (photo by Toby Feaster).

8. Plots the peak streamflow regression line from the relation previously described (7) with the upper 95 percent confidence limit for daily-value streamflow, and the 1:1 line. If the peak-flow plots above the 95 percent confidence limit of the regression or below the 1:1 line, an error in the data may have occurred;

9. Flags September 30 and October 1 peak streamflow to verify that the maximum value did not occur at midnight at the beginning or end of the water year, on the recession from or rise toward a higher peak in the adjoining water year.

10. Checks for changes in flow patterns that might be a result of regulation or changes in regulation by plotting the cumulated ratio of 90 to 50 percentile daily-value streamflow by water year, cumulated weekday water-year 50 percentile streamflow, cumulated water-year mean weekday streamflow, cumulated

water-year mean weekend streamflow, cumulated ratio of 10 to 50 percentile streamflow by climatic year (April 1 to March 31), and cumulated weekday climatic-year 50 percentile streamflow; and

11. Presents data used in plots and regressions in tables.

The QA/QC computer program for reviewing crest-stage, partial record sites performs the following checks:

1. Computes a Kendall's tau to check for trends in the data over time;
2. Plots the peak streamflow by water year;
3. Determines an index station (if one exists) by correlating concurrent water-year peak streamflow at the crest-stage station with water-year peak streamflow at all other stations statewide. The criterion for an index station to be chosen was that the index station peaks must be within ±8 days of the crest-stage station peaks, must have more concurrent peaks than half the number of peaks at the crest-stage station, and must have a coefficient of determination (R^2) greater than 0.60;
4. Relates concurrent peak flows at the index station and the station being reviewed and plots the 95 percent confidence limits of the regression. If the peak-flow plots above the 95 percent confidence limit of the regression or below the 1:1 line, an error in the data may have occurred;
5. Flags September 30 and October 1 peak streamflow to verify that the maximum value did not occur at midnight at the beginning or end of the water year, on the recession from or rise toward a higher peak in the adjoining water year;
6. Plots the peak gage height against the peak streamflow and current rating (if available) to check for anomalies;
7. Plots water-year hydrographs of peak streamflow at the crest-stage station and daily-value streamflow at the index station to check the estimated date of the crest-stage peak;
8. Presents data used in plotting and regressions in tables.

According to Rantz and others (1982), "... only as a last resort should the rating be extrapolated beyond a discharge value equal to twice the greatest measured discharge." Therefore, an additional check was made at the stations with established ratings by using the flow-measurement files. The peak flows were plotted against water year along with the maximum (at this point in time) measured flow and the value equal to two times that maximum measured flow. This plot was used to review peaks that may have been estimated from an excessive stage-flow rating extension but not updated when the rating was later defined by higher flow measurements.

ESTIMATION OF FLOOD MAGNITUDE AND FREQUENCY AT GAGING STATIONS

A frequency analysis of water-year peak-flow data at a gaging station on a stream provides an estimate of the flood magnitude and frequency at that specific site. The estimates are typically presented as a set of exceedance probabilities or, alternatively, recurrence intervals along with the associated flows. Exceedance probability is defined as the probability of exceeding a specified flow in a 1-year period and is expressed as decimal fractions less than 1.0 or as percentages less than 100. A flow with an exceedance probability of 0.01 has a 1 percent chance of being exceeded in any given year. Recurrence interval is defined as the number of years, on average, during which the specified flow is expected to be exceeded one time. A flow with a 100-year recurrence interval is one that, on average, will be exceeded once every 100 years. Recurrence interval and exceedance probability are the mathematical inverses of one another; therefore, a flow with an exceedance probability of 0.01 has a recurrence interval of 1/0.01 or 100 years. It is important to understand, however, that recurrence intervals, regardless of length, always refer to the average number of occurrences over a period of time. For example, a 100-year flood is one that might occur one time in a 100-year period, rather than exactly once every 100 years. Therefore, a flood with a 100-year recurrence interval can occur more frequently than once every 100 years and could even occur more than once in a given year. When interpreting the flood-frequency probabilities and recurrence intervals, it is helpful to remember that the analysis is based on water-year peaks; therefore, the results are relative to probability of exceedance in any given year.

Flood Frequency

Flood-frequency estimates at gaged sites can be computed by fitting the water-year peak flows to a known statistical distribution. For this study, flood-frequency estimates were computed by fitting the logarithms (base 10) of the water-year peak flows to a log-Pearson Type III distribution, following the guidelines and computational methods described in Bulletin 17B of the Hydrology Subcommittee of the Interagency Advisory Committee on Water Data (1982). The equation for fitting the log-Pearson Type III distribution to an observed series of water-year peak flows is as follows:

$$\log Q_T = \bar{X} + KS , \qquad (1)$$

where

Q_T is the T-year recurrence-interval flow, in cubic feet per second (ft^3/s);
$\bar{X}$ is the mean of the log-transformed water-year peak flows;
K is a factor dependent on recurrence interval and the skew coefficient of the log-transformed water-year peak flows; and
S is the standard deviation of the log-transformed water-year peak flows.

Values for *K* for a wide range of recurrence intervals and regionalized skew coefficients are published in appendix 3 of Bulletin 17B (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982).

A series of water-year peak flows at a station may include low or high outliers, which are data points that depart significantly from the range of the remaining data. The station record also may include information about peak flows that occurred outside of the period of regularly collected, or systematic, record. These peak flows are known as historic peaks and are often the peak flows known to have occurred during an extended period of time, longer than the period of collected record. Bulletin 17B (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982) provides guidelines for detecting and interpreting low and high outliers and historic data points and provides computational methods for making appropriate corrections to the distribution to account for their presence. In some cases, low or high outliers may be excluded from the record, so that the number of systematic peaks may not be equal to the number of years in the period of record.

Skew Coefficient

A skew coefficient measures the symmetry of the distribution of a set of peak flows about the median of the distribution. A peak-flow distribution with the mean equal to the median is said to have zero skew (fig. 3a). A positively skewed distribution has a mean that exceeds the median typically as a result of one or more extremely high peak flows (fig. 3b). A negatively skewed distribution has a mean that is less than the median, typically because of one or more extremely low peak flows (fig. 3c) (Pope and others, 2001).

The skew coefficient computed for a series of water-year peak flows at a single station can be weighted with a generalized, or regional, skew coefficient to obtain a better estimate of the station skew

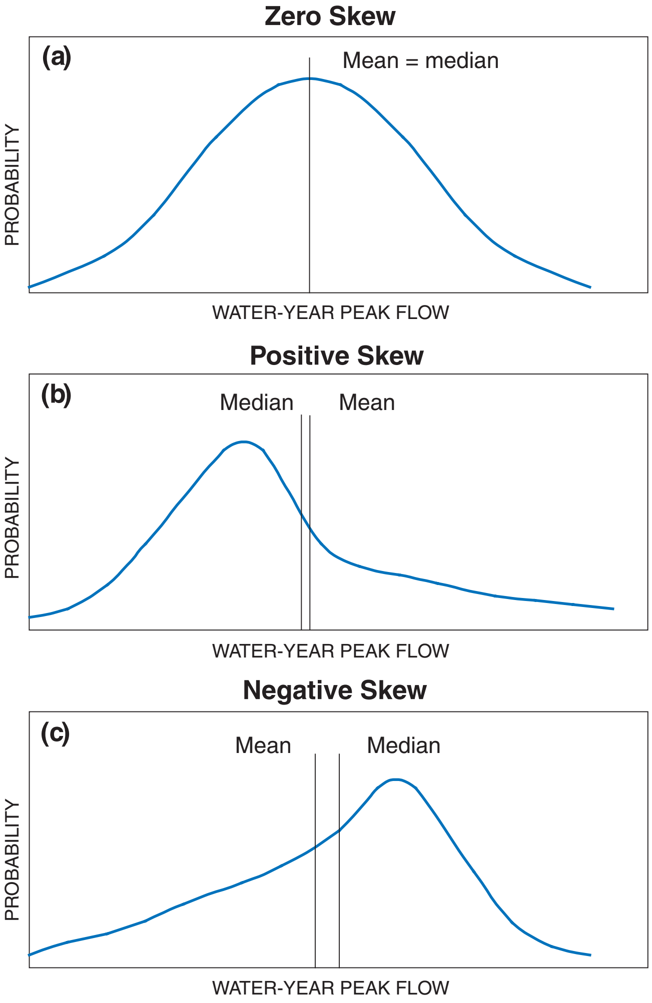


Figure 3. Examples of distributions with (a) zero skew, (b) positive skew, and (c) negative skew.

coefficient. A generalized skew coefficient can be obtained by combining skew estimates from nearby, similar sites. A nationwide skew study was conducted for the investigation documented in Bulletin 17B. Skew coefficients for long-term gaging stations throughout the Nation were computed and used to produce a map of isolines of generalized skew.

Bulletin 17B describes three methods for developing generalized skews using skew coefficients computed from long-term (25 or more years of record) gaging stations: (1) plot computed skew coefficients on a map and construct skew isolines, (2) use regression techniques to develop a skew prediction equation that would relate station skew coefficients to some set of basin characteristics, or (3) use the arithmetic mean of computed skew coefficients from long-term sites in the area. For this study, all three methods were attempted.

To develop the skew isoline map, skew coefficients from long-term gaging stations were plotted at the centroid of the watershed for each station. The map was then reviewed to determine if any geographic or topographic trends were visually apparent. No clearly definable patterns were found. Additionally, commercial contouring software was used to draw the skew isolines. Again, no definite geographic or topographic trends were revealed and the attempt to develop the skew isoline map was abandoned. It was noted, however, that skews in the Piedmont Province were predominately negative (24 of 35 stations). For the Blue Ridge, upper Coastal Plain, and lower Coastal Plain, the ratio of stations with negative skews to the total number of stations for that physiographic province were 10 of 20, 9 of 23, and 5 of 14, respectively.

Ordinary least squares regression was used to determine if a relation between the station skew and selected basin characteristics could be obtained. The regressions were performed on the complete data set and on the four physiographic provinces determined during the previous flood-frequency study (table 4).

Multiple regression analysis, using ordinary least squares regression, was used to determine the relation between the station skew coefficients and selected basin characteristics. Table 4 list the explanatory variables that were found to be statistically significant for each region and for the State as a whole. Although the variables shown were determined to be statistically significant, the R^2 values only ranged from 0.11 to 0.53, which indicates the fraction of the variance explained by the regression (Helshel and Hirsch, 1992). Consequently, it was concluded the relation between the station skews and the basin characteristics was weak and determining the generalized skew coefficients based on the regression analysis was not justified.

The third method for computing the generalized skew coefficients was to use the arithmetic mean of computed skew coefficients from long-term sites in the area. Before this was done, all stations with 25 or more years of record were evaluated to verify that all stations with significant regulation or that drained more than one physiographic province were removed from the data set. In addition, data at stations having low and high outliers in the frequency study as computed during the log-Pearson Type III analysis were not included in the skew computations. Descriptive statistics were computed for the different regions and the whole data set. Results from the statistics suggested that the skews were normally distributed. The SAS procedure Generalized Linear Model (GLM) was used to compare the

Table 4. Ordinary least-squares regression statistics for the generalized skew coefficients

[R^2, coefficient of determination; --, no data}

Region	Number of sites	Explanatory variables	R^2	Standard error	Mean square error
All stations	102	Longitude, precipitation times elevation	0.15	0.490	0.241
Blue Ridge	21	Latitude times latitude	0.38	0.380	0.145
Piedmont	47	Latitude	0.11	0.439	0.193
Upper Coastal Plain	16	No variables found	--	--	--
Lower Coastal Plain	18	Elevation, percent forest	0.53	0.253	0.064

station skews (SAS Institute, Inc., 1989). The results indicated that the skews in the Piedmont region had a larger difference from the other regions than would be expected by chance at a significance level of 0.05. No significant difference was observed among the three remaining regions. Consequently, the Blue Ridge, upper Coastal Plain and lower Coastal Plain skews were combined and a weighted skew was computed. The skews were weighted based on record length. Additionally, a weighted skew also was computed for the Piedmont (table 5).

Table 5. Weighted generalized skew coefficients and associated statistics for rural South Carolina gaging stations

	Blue Ridge, upper Coastal Plain, and lower Coastal Plain	**Piedmont**
Number of stations	54	35
Mean	0.082	-0.190
Maximum	0.811	0.524
Median	0.087	-0.177
Minimum	-0.547	-0.728
Mean square error	0.105	0.090
Root mean square error	0.324	0.299

To review the sensitivity of the mean skew to maximum and minimum skews, the mean skews were recomputed after excluding the maximum and minimum skews from each physiographic province and after excluding the two highest and two lowest skews. The results, which are provided in table 6, indicated that the weighted skews were not significantly affected by high and (or) low outliers.

The skew coefficients shown in table 5 were weighted based on years of record at the station. The skew coefficients shown in table 6 were not weighted. Although, it seems reasonable and appropriate to weight the skew coefficients by years of record, a comparison of the two tables suggest that, for the South Carolina generalized mean skew coefficients, the data were not significantly changed by weighting. The weighted skews from table 5, however, were used to compute the log-Pearson Type III frequency distributions for the regional regression analysis.

Bulletin 17B (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982) states that a mean square error of generalized skew of 0.302 should be used when the generalized skews are read from Plate I. As noted in table 5, the mean square errors of the generalized skews for South Carolina are much lower. The generalized skew map provided in Bulletin 17B was generated by averaging groups of 15 or more stations in areas covering four or more one-degree quadrangles of latitude and longitude. Consequently, the regions used for the current study in South Carolina do not compare geographically with the contours on the generalized skew map in Bulletin 17B and, therefore, a comparison of the means from the two studies does not necessarily provide substantial information. However, based on the results from the GLM test and on the improved mean square error when compared to the Bulletin 17B mean square error, the weighted skews listed in the table 5 were used to compute the log-Pearson Type III distributions.

Following Bulletin 17B guidelines, the assumption is made that the regionalized skew coefficient is an unbiased and independent estimate of the station skew. The mean square error of the weighted estimate, therefore, is minimized by weighting the station and regionalized skew in inverse proportion to their individual mean square errors. The following weighting equation was adopted from Tasker (1978), and is used in computing the weighted skew coefficients:

$$Gw = \frac{MSE_{\overline{G}}(G) + MSE_G(\overline{G})}{MSE_{\overline{G}} + MSE_G}, \qquad (2)$$

where

G_w is the weighted skew coefficient;
G is the station skew coefficient;
$\overline{G}$ is the regionalized skew coefficient;
$MSE_{\overline{G}}$ is the mean square error of regionalized skew coefficient; and
MSE_G is the mean square error of station skew coefficient.

The function MSE_G can be approximated with sufficient accuracy by the following equation (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982):

Table 6. Sensitivity of South Carolina generalized skew coefficients (not weighted) to high and low values

	Piedmont	Piedmont, with the maximum and minimum skews removed	Piedmont, with the two highest and two lowest skews removed	Blue Ridge, upper Coastal Plain, and lower Coastal Plain combined	Blue Ridge, upper Coastal Plain, and lower Coastal Plain combined, with the maximum and minimum skews removed	Blue Ridge, upper Coastal Plain, and lower Coastal Plain combined, with the two highest and two lowest skews removed
Number of stations	35	33	31	54	48	42
Mean	-0.168	-0.172	-0.174	0.095	0.096	0.091
Maximum	0.524	0.418	0.334	0.811	0.771	0.588
Median	-0.152	-0.152	-0.152	0.089	0.089	0.089
Minimum	-0.728	-0.717	-0.665	-0.547	-0.377	-0.296
Mean square error	0.103	0.086	0.070	0.104	0.074	0.054
Root mean square error	0.321	0.292	0.265	0.323	0.271	0.232

$$MSE_G \cong 10^{A - B(\log_{10}(N/10))}, \qquad (3)$$

where

$$A = \begin{cases} -0.33 + 0.08\,|G| & \text{if } |G| \le 0.90 \\ -0.52 + 0.30\,|G| & \text{if } |G| > 0.90; \text{ and} \end{cases}$$

$$B = \begin{cases} 0.94 - 0.26\,|G| & \text{if } |G| \le 1.50 \\ 0.55 & \text{if } |G| > 1.50; \end{cases}$$

where MSE_G is previously defined, |G| is the absolute value of the station skew coefficient (used as an estimate of population skew coefficient), and N is the record length, in years.

The USGS computer program PEAKFQ (W.O. Thomas, Jr., A.M. Lumb, K.M. Flynn, and W.H. Kirby, written commun., January 1998), was used to compute the relation between flood magnitude and probability of occurrence. PEAKFQ includes the features described in Bulletin 17B, but requires the user to exercise judgment when providing data on historic peaks, specifying screening levels for outliers, and interpreting the appropriateness of the resultant frequency curve to the observed data set. The weighted and station skew coefficients for the South Carolina stations used in the flood-frequency analyses are listed in appendix A.

Synthesized Peak Flow

Long-term flow records are not always available at a site; however, if concurrent rainfall and flow data are available, a rainfall-runoff model may be used to extend the records. Frequency data for four stations used in this report were from a study by Whetstone (1982) in which the USGS Rainfall-Runoff Model (Dawdy and others, 1972) was used to synthesize a series of water-year peak flows. The model utilizes 10 parameters to simulate the hydrologic processes of antecedent soil moisture, infiltration, and surface-runoff routing. Daily rainfall and evaporation data, and unit-value flow and rainfall data from several storm events collected over a period of 3 to 5 years were used to calibrate the model for each of the four basins. Long-term precipitation and evaporation data (more than 50 years of record) were used to synthesize a series of water-year peaks for use in subsequent frequency analyses (Guimaraes and Bohman, 1991). Synthesized flows for these stations are not listed in this report; however, flood-frequency data derived from the synthesized flows are listed in table 7 along with the weighted flows computed for each site using the appropriate regression equations for each physiographic province and individual station flood-frequency data.

Table 7. Station and weighted flood-frequency data for the synthesized peak flows used in the regionalization analysis

Map index number (fig. 1)	Station name and number	Drainage area (square miles)	Physiographic province	Type of data	Flows in cubic feet per second by recurrence interval (years)							
					2	5	10	25	50	100	200	500
51	Crabtree Swamp near Conway, S.C. (02110700)	14.0	Lower Coastal Plain	Station[1] Weighted[2]	206 214	388 411	544 579	784 835	994 1,060	1,230 1,310	1,500 1,590	1,910 2,020
52	Reedy Creek near Rains, S.C. (02135050)	10.4	Lower Coastal Plain	Station[1] Weighted[2]	242 242	478 472	680 667	991 967	1,260 1,230	1,570 1,530	1,920 1,860	2,440 2,370
53	Buck Branch at Bowman, S.C. (02174300)	11.9	Lower Coastal Plain	Station[1] Weighted[2]	210 215	438 446	639 646	955 956	1,240 1,230	1,550 1,540	1,920 1,890	2,460 2,430
54	Log Creek near Edgefield, S.C. (02195660)	1.26	Piedmont	Station[1] Weighted[2]	42.0 43.7	93.0 97.8	140 149	215 230	282 304	360 386	450 479	586 618

[1] The station data are computed by fitting the logarithms of water-year peak flows to a log-Pearson Type III distribution.

[2] The weighted data are computed by weighting the station data with flows computed using the regional flood-frequency relations. The weighted data should be used rather than the station data.

ESTIMATION OF FLOOD MAGNITUDE AND FREQUENCY AT UNGAGED SITES

For this study, two regional analyses were used to develop methods for estimating flows for ungaged rural basins in South Carolina. Traditional regional regression analysis using generalized least squares regression was used to define a set of predictive equations. These equations relate peak flows for the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence-interval flows to selected basin characteristics for ungaged, rural basins that are not significantly affected by regulation in each of four physiographic provinces of South Carolina (fig. 1). A second analysis was made using the region-of-influence method, which resulted in the development of a computer application to derive, for any given ungaged rural site in the Blue Ridge physiographic province, a unique predictive relation between the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence intervals and selected basin characteristics. Just as in the traditional regional regression, generalized least squares regression is used to develop this predictive relation. In the region-of-influence analysis, however, regression techniques are applied to only a selected subset of gaged sites, rather than the entire database of gaged sites.

Ordinary Least Squares Analysis

Ordinary least squares (OLS) regression techniques were used to select the explanatory variables that would define the final regression equations. In OLS regression, linear relations between the explanatory and response variables are necessary; consequently, variables must often be transformed. For example, the relation between drainage area and peak flow is typically not linear; however, the relation between the logarithms of drainage areas and the logarithms of peak flows often is linear. Homoscedasicity (a constant variance in the response variable over the range of the explanatory variables) about the regression line and normality of residuals also are requirements for OLS regression. Transformation of the flow data and other variables to logarithms enhances the homoscedasicity of the data about the regression line. Linearity, homoscedasicity, and normality of residuals were examined in residual plots.

Initially, eight explanatory variables for the gaged, unregulated streams in rural drainage basins in South Carolina and additional gaging stations in North Carolina and Georgia were included in this study. These eight explanatory variables were: *drainage area*, the area, in square miles, of a basin; *main-channel length*, the length of the main channel, in miles, from the gaging station to the basin divide; *main-channel slope*, the slope of the main channel, in feet per mile, between points that are 10 percent and 85 percent of the main-channel length from the gaging station; *elevation*, the mean basin elevation in feet above sea level; *forest cover*, forested area, in percent of contributing drainage area; *storage*, area of lakes, ponds, and swamps, in percent of contributing drainage area; *precipitation*, the water-year mean precipitation, in inches, at the centroid of the basin; and *runoff*, the water-year mean runoff, in inches, at the centroid of the basin. In addition, cross products, such as drainage area times length, computed using the eight explanatory variables also were included in the exploratory OLS regression analysis. The use of cross products allows the regression lines to converge or diverge, thereby decreasing or increasing the effect of one variable with the effect of another variable. For example, the effect of main-channel slope on flood frequency may decrease with increasing drainage area.

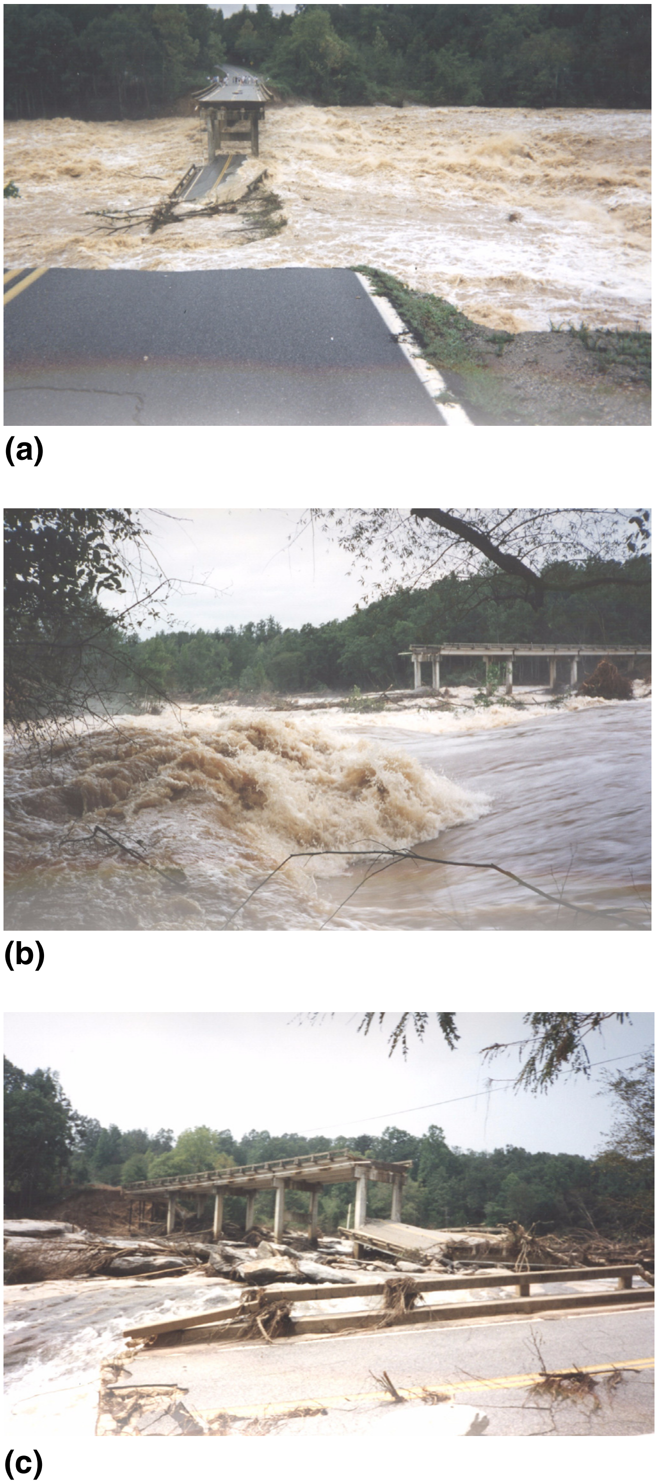
In the initial OLS regressions, the regionalization scheme used by Guimaraes and Bohman (1991), which divided the State into the Blue Ridge, Piedmont, upper Coastal Plain, and lower Coastal Plain, was found to be valid based on statistical significance of explanatory variables, comparisons of R^2, and standard errors of the OLS regression equations for the regions and the State as a whole. In addition, plots of the logarithms of drainage area and logarithms of the 100-year recurrence-interval flows were used to graphically compare the regions and to help determine the validity of the boundaries of those regions. Regression residuals (for the 100-year peak flows) also were plotted at the centroid of their respective drainage basins and inspected for geographical patterns of bias. No distinct patterns were apparent in the mapped residuals for the four regions.

The hydrologic model used in the regression analysis is of the form:

$$Q_T = aA^bB^cC^d\ldots, \qquad (4)$$

where

Q_T	is the flood magnitude having T-year recurrence interval;
A, B, C	are explanatory variables (basin characteristics); and
a, b, c, d	are regression coefficients.



Bridge over the Enoree River at S.C. Highway 418 collapsed on August 27, 1995, due to severe flooding caused by Tropical Depression Jerry:

(a) August 27, 1995; photo by Michael Hall
(b) August 28, 1995; photo by Toby Feaster
(c) September 1, 1995; photo by Toby Feaster

If the dependent and explanatory variables are logarithmically transformed, the hydrologic model has the following linear form:

$$\text{Log } Q_T = \log a + b\,(\log A) + c\,(\log B) + d\,(\log C) + \ldots, \qquad (5)$$

where the variables are previously defined. The logarithmic relation is the form that was used in this study.

Basins in the Piedmont province of South Carolina tend to be more elongated than basins in Georgia and North Carolina, and the recurrence-interval flows for the Piedmont stations in South Carolina tend to be lower than those in Georgia and North Carolina. This difference in the recurrence-interval flows could be attributed to the difference in basin shapes because an elongated basin would tend to have a lower peak with a more spread-out hydrograph than a similarly sized basin that was wider and shorter. Consequently, a “qualitative variable” for State was included in the regression model to account for these differences. This qualitative variable also was found to be significant in the upper Coastal Plain. Guimaraes and Bohman (1991) also used a qualitative variable to differentiate between flows in the Piedmont of South Carolina and flows in the Piedmont of Georgia and North Carolina.

An example of the linear form of the hydrologic model including qualitative variables is:

$$\log Q_T = \log a + b\,(\log A) + c\,(\log B) + d\,(\log C) + \ldots + eV, \qquad (6)$$

where V is a qualitative variable that is set to 0 if the variable is geographically in State “x” or 1 if geographically in State “y”. If the qualitative variable is determined to be significant by the regression analysis, the regression lines for the two States have the same slope, but different intercepts. The qualitative variable can be used to detect significant differences between States and to utilize data from both States where data are sparse.

Within each physiographic province, the 100-year recurrence-interval flow was regressed against the explanatory variables and a qualitative variable that denotes location by State. As previously mentioned, the qualitative variable was determined to be statistically significant by the regression analysis in the Piedmont and upper Coastal Plain (figs. 4 and 5) and indicated no statistically significant difference between the States for the Blue Ridge and lower Coastal Plain. For the Piedmont, the regressions indicated that Georgia and North Carolina data were not significantly different from each other, but were both significantly different from South Carolina (fig. 4). Note that on figure 4, the South Carolina data plots generally below the data points from Georgia and North Carolina. Although it is possible to use only South Carolina data in the regressions, the relation using the data from all three States with the qualitative variable was considered to be more definitive because of sparseness of the South Carolina data. Inclusion of the qualitative variable allows the data from each State to influence the slope of the overall relation, and at the same time permits a unique intercept value for the South Carolina relation.

As can be seen on figure 5, no Georgia stations were included in the South Carolina upper Coastal Plain regression analysis. The North Carolina Sandhills physiographic province is hydrologically similar to the South Carolina upper Coastal Plain. Georgia also has a Sandhill region that is narrowly banded near the South Carolina-Georgia border. However, there were no Georgia stations near the South Carolina border that were found to be hydrologically similar to the South Carolina upper Coastal Plain.

In addition to the qualitative variable described above, the explanatory variable determined to be the most significant in the OLS analyses for all regions was drainage area. In the Blue Ridge, other significant variables were chosen by the OLS regression model but did not reduce the standard error enough to warrant inclusion or were significant in only a few of the Q_T flows and then became statistically insignificant in the higher recurrence intervals (50 years and above). The explanatory variables determined in the OLS regression analyses were used in the generalized least squares regression procedures.

Generalized Least Squares Analysis

Generalized least squares (GLS) regression, as described by Stedinger and Tasker (1985), was used to compute the final coefficients and the measures of accuracy for the regression equations, using the USGS computer program GLSNET (G.D. Tasker, K.M. Flynn, A.M. Lumb, and W. O. Thomas, Jr., written commun., 1995). Stedinger and Tasker (1985) found that GLS regression equations are more accurate and provide a

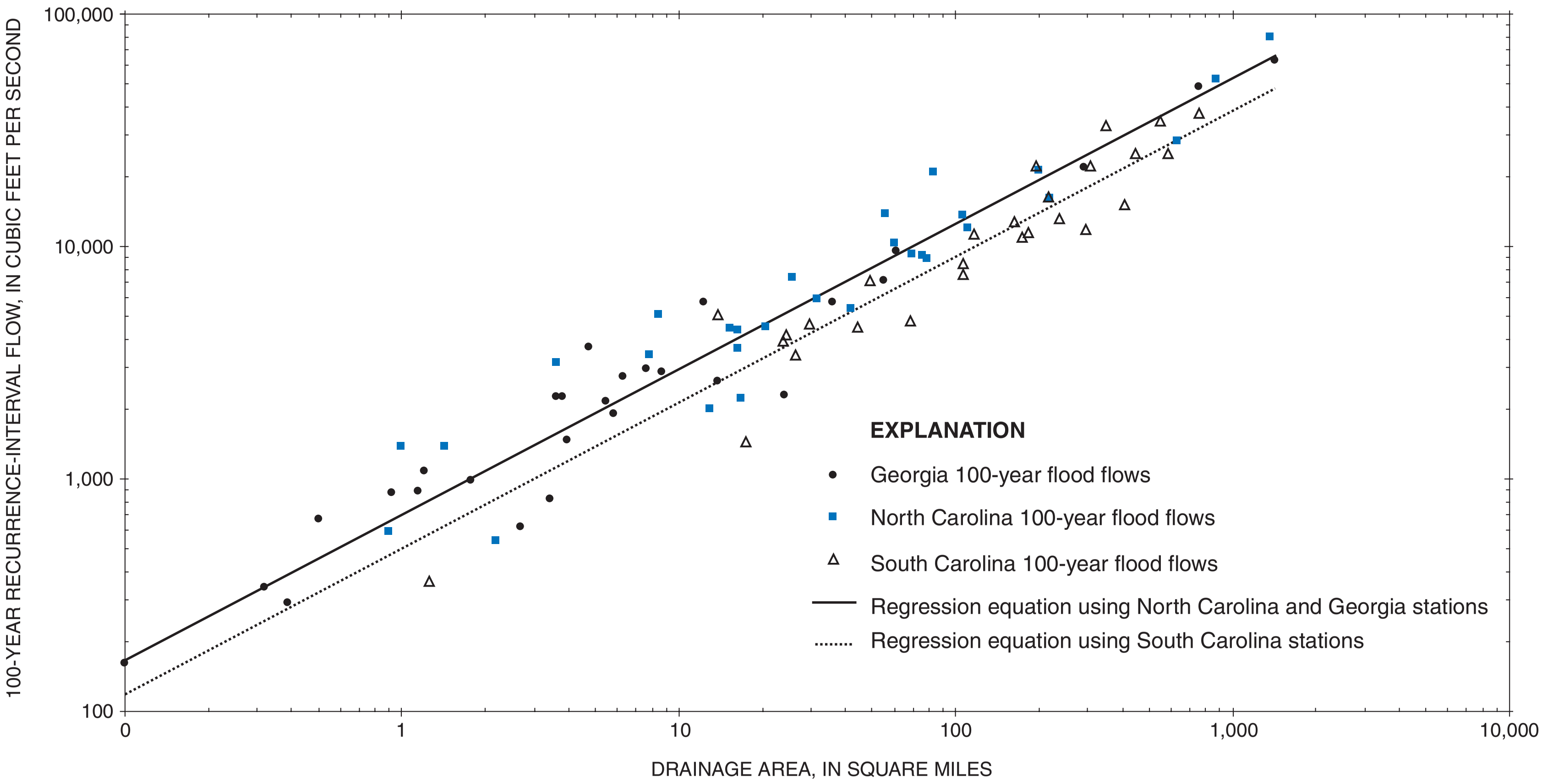


Figure 4. Relation between drainage area and 100-year recurrence-interval flow for stations on streams draining the Piedmont physiographic province in South Carolina, North Carolina, and Georgia.

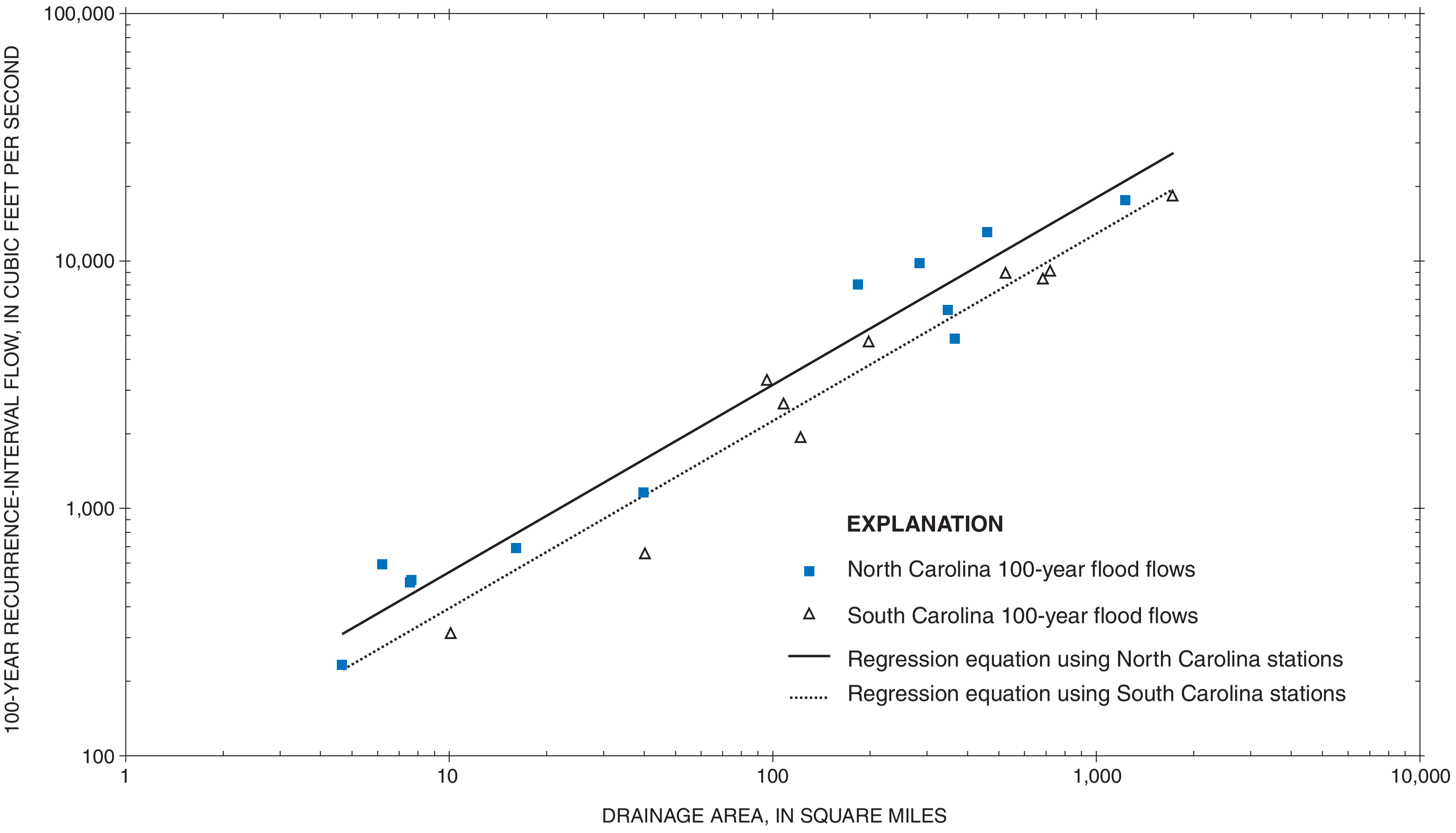


Figure 5. Relation between drainage area and 100-year recurrence-interval flow for stations on streams draining the upper Coastal Plain physiographic province in South Carolina and North Carolina.

better estimate of the accuracy of the equations than OLS regression equations when streamflow records at gaging stations are of different and widely varying lengths and when concurrent flows at different stations are correlated. Generalized least squares regression techniques give less weight to streamflow-gaging stations that have shorter periods of record than other stations. Less weight also is given to those stations where concurrent peak flows are correlated with other stations (Hodgkins, 1999). Table 8 shows the peak-flow regression equations for recurrence intervals of 2, 5, 10, 25, 50, 100, 200, and 500 years that resulted from the GLS regression analysis for South Carolina.

Accuracy of the Method

As shown in table 8, the regional regression analyses for South Carolina resulted in the development of a set of equations that allow the user to estimate the 2-, 5-, 10-, 25-, 50-, 100-, 200-, or 500-year recurrence-interval flood at an ungaged, unregulated stream in a rural drainage basin. When applying these equations, users should not interpret these empirical results as exact. These regression equations are statistical models that should be interpreted and applied within the limits of the data and with the understanding that the results are best-fit estimates with an associated scatter or variance.

One measure of how well the regression equations estimate the peak flows at an ungaged site is the standard error of prediction (S_p). The S_p is the square root of the mean square error of prediction, MSEp. The MSEp is the sum of two components--the mean square error resulting from the model, γ^2, and the sampling mean square error, $MSE_{s,i,}$ which results from estimating the model parameters from samples of the population. The mean square model error, γ^2, is a characteristic of the model, is constant for all sites, and cannot be reduced by additional data collection. The mean square sample error, $MSE_{s,i,}$ for a given site, however, depends on the values of the explanatory variables, in this case drainage area (DA), used to develop the flow estimate at that site. Consequently, the sampling error can be reduced by additional data collection at existing stations, or by installing new stations in the same physiographic province, or some combination of both. The standard error of prediction for a site, *i*, is computed as:

$$S_{p,\,i} = (\gamma^2 + MSE_{s,\,i})^{\frac{1}{2}} \,, \tag{7}$$

(variables previously defined) and varies from site to site. Assuming the explanatory variables for the gaged sites in the regression are a representative sample of all sites in the region, the average accuracy of prediction for the regression model can be determined by computing the average standard error of prediction:

Table 8. South Carolina rural flood-frequency equations

[DA, drainage area, in square miles. Result will be in cubic feet per second]

Rural flood recurrence interval (years)	Physiographic province			
	Blue Ridge	Piedmont	Upper Coastal Plain	Lower Coastal Plain
2	116 (DA) $^{0.757}$	111 (DA) $^{0.678}$	21.2 (DA) $^{0.754}$	51.7 (DA) $^{0.663}$
5	204 (DA) $^{0.734}$	197 (DA) $^{0.656}$	32.4 (DA) $^{0.752}$	96.6 (DA) $^{0.653}$
10	281 (DA) $^{0.720}$	262 (DA) $^{0.647}$	40.6 (DA) $^{0.753}$	136 (DA) $^{0.646}$
25	395 (DA) $^{0.705}$	353 (DA) $^{0.637}$	51.5 (DA) $^{0.754}$	197 (DA) $^{0.639}$
50	491 (DA) $^{0.695}$	426 (DA) $^{0.631}$	60.1 (DA) $^{0.755}$	251 (DA) $^{0.634}$
100	595 (DA) $^{0.688}$	503 (DA) $^{0.626}$	69.1 (DA) $^{0.757}$	312 (DA) $^{0.631}$
200	707 (DA) $^{0.681}$	584 (DA) $^{0.621}$	78.7 (DA) $^{0.759}$	381 (DA) $^{0.628}$
500	868 (DA) $^{0.675}$	698 (DA) $^{0.615}$	92.1 (DA) $^{0.761}$	486 (DA) $^{0.624}$

$$S_p = \left\{ \gamma^2 + \frac{1}{n} \sum_{i=1}^{n} MSE_{s,\,i} \right\}^{\frac{1}{2}} , \qquad (8)$$

where n is the number of observations and all other variables are previously defined.

The standard error of the model ($SE_{(model)}$) can be converted from log (base 10) units to percentage error by using the transformation formulas,

$$+\,PercentSE_{(model)} = 100[10^{\gamma} - 1] \qquad (9)$$

and

$$-\,PercentSE_{(model)} = 100[10^{-\gamma} - 1] \ , \qquad (10)$$

where the variables were previously defined.

Similarly, the average standard error of prediction (S_p) can be transformed to positive or negative percentage error by substituting $S_p{}^2$ for γ^2 in equations 6 and 7, respectively. Computation of $S_{p,i}$ for a given ungaged site, i, involves fairly complex matrix algebra; therefore, a computer program that computes the standard error of prediction for any study site has been developed. Appendix C provides details of the GLS regression method and calculation of prediction errors and intervals.

Another overall measure of how well regression equations can be used to estimate flood peaks when applied to ungaged basins is the PRESS (“PRediction Error Sum of Squares”) (Helsel and Hirsch, 1992) statistic. The PRESS statistic is a validation-type statistic. To compute the PRESS statistic, one gaging station is removed from the stations used to develop the regression equation, then the value of the one removed is predicted. The difference between the predicted value from the regression equation and the observed peak flow at that station is computed. The gaging station that is removed is then changed and the above process repeated until every station has been left out once. The prediction errors are then squared and summed (Helsel and Hirsch, 1992). PRESS/n is analogous to the average variance of prediction, and the square root of PRESS/n is analogous to the average standard error of prediction. Values of the square root of PRESS/n close to the values of the average standard error of prediction provide some measure of validation of the regression equations (Hodgkins, 1999).

A third measure of the overall accuracy of the regression equations is the average equivalent years of record. This measure represents the average number of years of peak-flow data needed to provide an estimate by using log-Pearson Type III techniques that would be equal in accuracy to an estimate made by using regional methods (table 9). The average equivalent years of record is a function of the accuracy of the regression equations, the recurrence interval, and the average variance and skew of the water-year peak flows at gaging stations (Hardison, 1971).

Region-of-Influence Method

A new technique for estimating flood frequency at ungaged sites is the *region-of-influence* (ROI) method (Tasker and Slade, 1994; Tasker and others, 1996; Asquith and Slade, 1999; Pope and others, 2001). In this method a regression equation is estimated for each ungaged site and for each recurrence-interval peak flow. The regression equation for a site is computed using data from a unique region called the *region of influence* by Burn (1990a, 1990b) and suggested by Acreman and Wiltshire (1987). The subset of gaging stations that make up the *region of influence* for each ungaged site is made up of the *N* nearest neighbors. In this method, the nearness of two neighbors is measured, not by the physical distance between the sites, but by a distance defined in terms of the watershed characteristics. The *distance* between any two sites, indexed by *i* and *j*, is determined by the Euclidean distance metric:

$$d_{ij} = \left(\sum_{k=1}^{p} \left(\frac{x_{ik} - x_{jk}}{sd(X_k)} \right)^2 \right)^{1/2} , \qquad (11)$$

where

d_{ij}	is the distance between sites *i* and *j* in the Cartesian product space of the watershed characteristics;
p	is the number of watershed characteristics needed to calculate d_{ij} ;
X_k	represents the *k*th watershed characteristic;
$sd(X_k)$	is the sample standard deviation for X_k ;
x_{ik}	is the value of X_k at the *i*th streamflow-gaging station; and
x_{jk}	is the value of X_k at the ungaged site, *j*.

For this method to work, the value of N, the number of streamflow-gaging stations, should be large enough to have enough degrees of freedom in the regression to estimate two or three parameters.

To adapt the ROI method to South Carolina sites, the parameters p and N must be determined along with the identity of the basin characteristics that are used to compute d_{ij}. Selection of the number of gaged sites, N, and the number and identity of the basin characteristics that will define the region of influence for South Carolina was done by systematically evaluating the PRESS (Helsel and Hirsch, 1992) statistic for N and p and for different basin characteristics. As previously mentioned, the PRESS statistic is a validation-type estimator of error that uses all observations except one to develop the equation, then estimates the value of the one left out. It then selects another observation to be left out, and repeats the process for each observation. The difference between the predicted and observed logarithms of flood peaks are squared and summed to compute PRESS. After testing several combinations of N and p with different combinations of basin characteristics, it was concluded that the best combination for South Carolina sites is, N=25, p=2, with basin characteristics log(elevation) and log(slope) used to define the distance metric based on minimizing the PRESS. Further, it was found that if the prediction site was in the Blue Ridge or Piedmont Province, the search for similar sites was limited to the Blue Ridge and Piedmont sites; If the prediction site was in one of the Coastal Plain Provinces, the search was limited to Coastal Plain Provinces.

To estimate recurrence-interval flows at an ungaged rural site, the ROI method performs GLS regression using estimation data from the 25 most similar sites in terms of elevation and slope. The ROI-GLS regression for each site where an estimate is needed uses log-transformed T-year flood peaks as response variables and log(AREA) and two qualitative variables as explanatory variables. The qualitative variables (STATE and PIED) are zero-one variables defined as follows:

STATE = 1 if site is in South Carolina and STATE = 0 if not in South Carolina, and

PIED = 1 if site is in the Piedmont Province and PIED = 0 if not in Piedmont.

Table 9. Accuracy statistics for the regional regression equations, presented by physiographic province and recurrence interval

[PRESS, PRediction Error Sum of Squares; n, number of sites]

Recurrence interval (years)	Mean standard error of prediction (percent)	$(PRESS/n)^{1/2}$ (percent)	Average equivalent years of record
Blue Ridge			
2	-27 to 37	-28 to 39	3.1
5	-24 to 32	-27 to 36	5.6
10	-23 to 31	-27 to 36	8.6
25	-23 to 30	-27 to 38	12.8
50	-23 to 30	-28 to 40	15.6
100	-24 to 31	-30 to 42	17.8
200	-25 to 33	-31 to 45	19.2
500	-26 to 35	-33 to 49	20.1
Piedmont			
2	-30 to 43	-31 to 46	2.6
5	-27 to 38	-29 to 41	4.1
10	-27 to 36	-29 to 40	5.8
25	-26 to 35	-29 to 40	8.3
50	-26 to 35	-29 to 41	10.1
100	-26 to 36	-30 to 43	11.7
200	-27 to 36	-31 to 44	13.0
500	-27 to 38	-32 to 46	14.3
Upper Coastal Plain			
2	-16 to 19	-18 to 22	7.8
5	-17 to 21	-19 to 24	10.3
10	-19 to 23	-21 to 27	11.7
25	-21 to 27	-24 to 32	12.6
50	-23 to 31	-26 to 36	12.8
100	-25 to 34	-28 to 40	12.8
200	-27 to 38	-31 to 44	12.6
500	-30 to 43	-33 to 50	12.3
Lower Coastal Plain			
2	-22 to 27	-25 to 34	7.8
5	-20 to 25	-24 to 32	13.3
10	-21 to 27	-26 to 35	15.8
25	-24 to 32	-29 to 41	16.9
50	-27 to 36	-32 to 47	17.0
100	-29 to 41	-35 to 53	16.7
200	-31 to 46	-37 to 59	16.4
500	-34 to 52	-40 to 68	15.8

Comparison of Methods

The predictive abilities of the Regional Regression Equations (RRE) and ROI methods were compared using the ratio of the PRESS computed from the ROI method to the PRESS computed from the RRE. A ratio of less than one indicated that the ROI method is a better predictor at an ungaged site than the RRE method. Table 10 shows that the ROI method performed systematically better only in the Blue Ridge, and therefore, should be used only for that Province.

Limitations

The following limitations should be recognized when using either the regional regression equations or the region-of-influence method.

1. The methods should be used only for ungaged sites where the basin drainage area is between 0.6 and 945 mi^2 for the Blue Ridge, 0.1 to 1,430 mi^2 for the Piedmont, 4.7 to 1,720 mi^2 for the upper Coastal Plain, and 1.1 to 1,250 mi^2 for the lower Coastal Plain.
2. The methods should not be used for sites where the watershed is substantially affected by regulation from impoundments, channelization, levees, or other man-made structures.
3. The methods should not be used for sites on streams in urban areas unless the effects of urbanization are insignificant.
4. The methods do not apply where flooding is influenced by extreme tidal events.
5. The methods should be used with caution in areas where the streamflow characteristics have not been sufficiently defined by high-flow measurements (fig. 1). In York and Chester Counties, the regional equations tend to produce flow results that may be significantly lower than those obtained using flow records (Guimaraes and Bohman, 1991). C.L. Sanders, Jr. (U.S. Geological Survey, written commun., Nov. 1993) observed that this area can be defined by the extent of the Iredell-Mecklenburg and Wilkes-Winnsboro-Mecklenburg soil types shown on the "General Soil Map, South Carolina" (Smith and Hollbick, 1979). Sanders showed that the regressions equations developed by Gunter and others (1987) tend to produce more reasonable results in this area than do the equations developed by Guimaraes and Bohman (1991). Until the streamflow characteristics in this area are sufficiently defined by higher flow measurements, the North Carolina rural flood-frequency equations by Pope and others (2001) and the North Carolina urban flood-frequency equations by Robbins and Pope (1996) can be used to estimate the magnitude and frequency of floods in the York and Chester County area.
6. The methods should be used with caution in the Upper Three Runs Basin in Aiken and Barnwell Counties in the western part of the upper Coastal Plain physiographic province. The regional equations tend to produce flows that may be substantially higher than those obtained using observed flow records for the Upper Three Runs Basin. To obtain flows for the Upper Three Runs Basin, either the station recurrence-interval flow should be used (appendix A) or the flow at an ungaged site should be adjusted by drainage area, using an appropriate Upper Three Runs gaging station as an index station.

Table 10. Ratio of PRESS (ROI) statistic to PRESS (RRE) statistic for each physiographic province

[PRESS, PRediction Error Sum of Squares]

Recurrence interval (years)	Physiographic province			
	Blue Ridge	Piedmont	Upper Coastal Plain	Lower Coastal Plain
2	0.87	1.17	1.53	0.91
5	0.79	1.13	1.17	1.08
10	0.75	1.09	1.40	1.12
25	0.73	1.04	1.36	1.14
50	0.73	1.02	1.32	1.10
100	0.74	1.00	1.30	1.08
200	0.75	0.98	1.28	1.07
500	0.77	0.97	1.25	1.05

Computer Software

A computer software program was developed that estimates flood frequency at rural ungaged sites in South Carolina. The computer application includes an executable program file and three supporting data files. All of these files must be located in a common directory for the computer program to work properly. Flood-frequency estimates provided by the computer program can be used alone or weighted with data from a nearby gaging station, as described in the Application of Methods section of this report.

Each time the computer program is executed, it produces flood-frequency estimates using the RRE method and, if the site is in the Blue Ridge Province, the ROI method. The computer program produces on-screen summary results and generates one or two output files containing the results of flood-frequency estimates at rural ungaged sites in South Carolina. The first output file contains flood-magnitude predictions, standard error of the predictions, and 90 percent prediction intervals for each recurrence-interval flow. The second output file is produced if the ROI method is used and contains detailed diagnostic information for the ROI method, including a listing of the ungaged site’s "region of influence," as well as for each recurrence-interval flow, the unique regression equations, and overall prediction error statistics for the method. Appendix C provides details of the GLS regression method and calculation of prediction errors and intervals.

The computer program and necessary data files can be downloaded to a personal computer from the USGS website. To download the computer files using an internet browser, enter the following address: http://sc.water.usgs.gov/SCFFREQ/. From there, download the compressed (or zipped) file named scff.zip to a designated directory on your personal computer. Extract the following six files from scff.zip: (1) scff.for is the program source code; (2) scff.exe is the executable file; (3) sc.cmn is a common block file used with the source code; (4) sc115.txt, (5) scroi.cr, and (6) scroi.rl are data files for the region-of-influence method. Once the files have been extracted, open a DOS window and type scff to run the program.

APPLICATION OF METHODS

The methods presented in this report can be used to estimate the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence-interval flows at ungaged sites and to improve estimates at gaged sites for unregulated, rural streams in South Carolina. To improve the estimated flow for a selected recurrence-interval flood at a gaged site, the flow at the gaged site should be weighted with the regional flow estimate using the number of years of station record and the accuracy of the regional flood-frequency relation, expressed in equivalent years of record (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982). The accuracy of the regional flood-frequency relations is the same as would be determined by gaging a stream for the number of equivalent years of record listed in table 9. Accuracy for the weighted estimate is the sum of the accuracy of each estimate in equivalent years of record, assuming the estimates are independent.

For ungaged sites, the regional peak-flow estimate for a selected recurrence interval can be improved if the site is located near a gaged site on the same stream that has at least 10 years of peak-flow record. Flows for ungaged sites are weighted according to the relative proximity, with respect to drainage area, of the ungaged site to the gaged site. Gaged sites having less than 10 years of peak-flow record should be treated as ungaged sites.

Flood Frequency at Gaged Sites for Streams Draining One Physiographic Province

Flood-frequency estimates at gaged sites can be improved by combining the estimate determined by regional methods with the estimate determined by fitting the log-Pearson Type III distribution to the peak-flow record at the gaged site. The procedure for estimating a selected recurrence-interval flow at gaged sites can be computed using the following equation (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982):

$$LogQ_{T(w)} = \frac{N(\log Q_{T(s)}) + EY(\log Q_{T(r)})}{N + EY}, \qquad (12)$$

where

$Q_{T(w)}$	is the weighted flow for the selected T-year recurrence interval, in cubic feet per second;
N	is the number of years of record used to compute $Q_{T(s)}$;
$Q_{T(s)}$	is the estimated flow at the station from the log-Pearson Type III analysis for the selected T-year recurrence interval, in cubic feet per second;
EY	is the equivalent years of record for $Q_{T(r)}$ from table 9; and
$Q_{T(r)}$	is the regional flow at the station for the selected T-year recurrence interval computed using the applicable regional relation from table 8, in cubic feet per second.

The log-Pearson Type III flood-frequency data and weighted flows for the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence interval floods for the South Carolina streamflow sites used in this report are presented in appendix B. The log-Pearson Type III flood-frequency data also are presented in appendix A.

The flood-frequency data were computed using available records through the 1999 water year and supersede the values presented by Guimaraes and Bohman (1991).

Flood Frequency Near Gaged Sites on the Same Stream Draining One Physiographic Province

Peak flows for a selected recurrence interval at an ungaged site on a stream can be estimated by determining in which physiographic province (fig. 1) the basin is located, computing the drainage area for the basin, and computing the flow using the appropriate regional flood-frequency relation from table 8. The regional estimate can be improved if the site is located on the same stream as a gaged site having at least 10 years of record and the drainage area of the ungaged site is within 50 percent of the drainage area of the gaged site. The weighted flow, $Q_{T(w)}$, at the gaged site can be transferred to the ungaged site using the following equation:

$$Q_u = \left(\frac{A_u}{A_g}\right)^b Q_{T(w)} \qquad (13)$$

and then a weighted flow value can be computed by the equation

$$Q_{u(w)} = \left(\frac{2|\Delta A|}{A_g}\right) Q_r + \left(1 - \frac{2|\Delta A|}{A_g}\right) Q_u \qquad (14)$$

where

Q_u is the flow at the ungaged site transferred from the gaged site in cubic feet per second (ft^3/s);

A_u is the drainage area of the ungaged site, in square miles (mi^2);

A_g is the drainage area of the gaged site, in square miles;

b is the exponent of the drainage area term of the regional flood-frequency relation for the applicable physiographic province and recurrence interval, from table 8;

$Q_{T(w)}$ the weighted flow at the gaged site for the selected T-year recurrence interval, in cubic feet per second;

$Q_{u(w)}$ is the weighted flow at the ungaged site for the selected T-year recurrence interval, in cubic feet per second;

$|\Delta A|$ is the absolute value of difference in drainage areas of the gaging station and the ungaged site, in square miles; and

Q_r is the regional flow for the selected recurrence interval at the ungaged site computed, using the applicable regional relation from table 8, in cubic feet per second.

This procedure should only be used to adjust the RRE estimate of flow if the drainage area of the ungaged site on the same stream as a nearby gaged site is within 50 percent of the drainage area of the gaged site. If it is not, use of the regional regression equations as described previously will provide the best estimate of flow.

Flood Frequency at Ungaged Sites on Streams Draining More Than One Physiographic Province

For an ungaged site that drains more than one physiographic province, the selected recurrence-interval flow may be computed by solving the appropriate equations for each physiographic province as though the drainage area were located entirely in each province, and then weight the flow as described below. For example, if the drainage area is 80 mi^2 with 60 percent of the basin located in the Piedmont and 40 percent located in the Blue Ridge, computation of the 100-year flow is as follows:

Piedmont: $Q_{100} = 503\ (80)^{0.626} = 7{,}810\ ft^3/s$,

Weighted Piedmont $Q_{100} = 7{,}810 \times 0.60 = 4{,}690\ ft^3/s$,

Blue Ridge: $Q_{100} = 595\ (80)^{0.688} = 12{,}100\ ft^3/s$,

Weighted Blue Ridge $Q_{100} = 12{,}100 \times 0.40 = 4{,}840\ ft^3/s$, and

Final weighted Q_{100} at site $= 4{,}690 + 4{,}840 = 9{,}530\ ft^3/s$.

Flooding along the Waccamaw River due to Hurricane Floyd at railway station on S.C. Highway 905. Photo by Andy Caldwell on September 29, 1999.

FLOOD FREQUENCY AT GAGED SITES ON STREAMS DRAINING MORE THAN ONE PHYSIOGRAPHIC PROVINCE

Flood frequencies for gaged sites along the mainstem of the Broad, Little Pee Dee, Lynches, Edisto, and Salkehatchie Rivers (fig. 6) were not determined using the regional flood-frequency analysis because these sites are located on streams draining more than one physiographic province. The flood-frequency relations at such sites are not representative of the other sites used to develop the regional relations. Some of the streams listed above are subject to minor regulation at low to medium flows that do not significantly affect water-year peak flows. The logarithm of water-year peak flows for these stations was fitted to a log-Pearson Type III distribution to determine flows for selected recurrence intervals and are described in unnumbered text tables that follow.

Broad River

The Broad River originates in the Blue Ridge Province of North Carolina and flows eastward and southward through the Piedmont to Columbia, S.C., where it merges with the Saluda River to form the Congaree River (fig. 6). Floods on the Broad River consist primarily of unregulated runoff, although there is some regulation at low to medium flows (Guimaraes and Bohman, 1991).

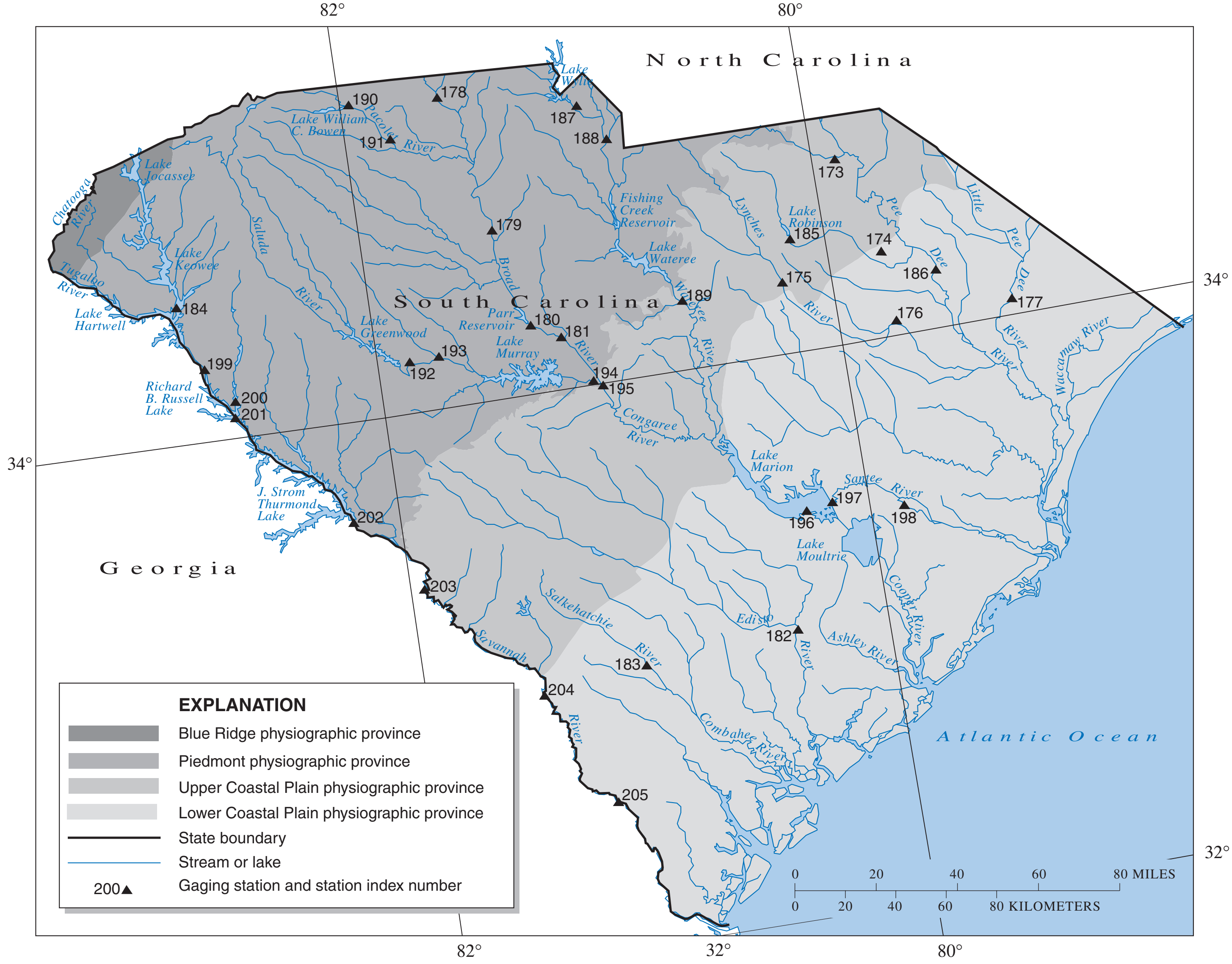


Figure 6. Locations of streamflow gaging stations on regulated streams and gaging stations on streams draining more than one physiographic province in South Carolina.

Flows were computed for Station 02153500, Broad River near Gaffney, S.C., for the period 1939-90 (map index number 178); Station 02156500, Broad River near Carlisle, S.C., for the period 1926-99 (map index number 179); and Station 02161500, Broad River near Richtex, S.C., for the period 1897-1907 and 1926-99 (map index number 181) (appendix C). The drainage area at Station 02161500 is 4,850 mi^2 and the period of record is 1926-83. Stage and flow data also were collected from 1897-1907 and 1981-99 at Station 02161000, Broad River at Alston, S.C. (map index number 180), which has a drainage area of 4,790 mi^2. Because the difference in the size of these two stations is small (approximately 1 percent), the data were combined. The combined period of record (1897-1907 and 1926-99) was used to compute the flood frequencies at Station 02161500.

Significant floods were recorded at Station 02161500 during water years 1928, 1930, and 1936. To incorporate these historical floods into the frequency analysis at Station 02156500, which has an actual period of record for water years 1939-99, the water-year peak flows at Station 02156500 were regressed against the water-year peak flows at Station 02161500 for the concurrent peaks. The R^2 for this regression was 0.85 and the standard error of estimate was 15 percent. Therefore, the 1928, 1930, and 1936 peak floods were estimated at Station 02156500 from this regression. The peaks were then included in the Station 02156500 data as historical peaks and a historical period of 74 years (1926-99) was used in the log-Pearson analysis.

A similar regression analysis was made for Station 02153500 with Station 02161500. The R^2 from that regression was 0.56 and the standard error of estimate was 28 percent. Based on the R^2, the standard error of estimate, and the difference in the drainage areas between the two stations (1,490 mi^2 and 4,850 mi^2, respectively), the log Pearson analysis at Station 02153500 was computed using only the systematic record (1939-90).

An additional analysis was made using the period from 1939-83 for which data were simultaneously collected at Stations 02153500, 02156500, and 02161500. The relation between the drainage area and the 100-year flow at the three stations is shown in figure 7. For the analysis using the complete period of record, the graph indicates a relatively linear relation with a slight increase in the slope between Stations 02156500 and 02161500. The plot of the simultaneous period from 1939-83 shows a similiar pattern with a slope between Stations 02153500 and 02156500 that is flatter but with a similiar increase in slope between Stations 02156500 and 02161500. Figure 7 also shows that the flood frequencies are sensitive to the larger flows (1928, 1930, and 1936). Given the sensitivity to the larger flows and the similar change in slopes between Stations 02156500 and 02161500, the data suggest that using the estimated floods at Station 02156500 results in a more accurate estimate of the flood frequency.

The flows for selected recurrence intervals are tabulated below for Station 02153500, Broad River near Gaffney, S.C., Station 02156500, Broad River near Carlisle, S.C., and Station 02161500, Broad River at Richtex, S.C.

Map index No.	Station name and number	Drainage area (square miles)	Flow, in cubic feet per second, for indicated recurrence interval, in years							
			2	5	10	25	50	100	200	500
178	Broad River near Gaffney, S.C. (02153500)	1,490	31,500	47,900	59,000	73,300	84,100	94,800	106,000	120,000
179	Broad River near Carlisle, S.C. (02156500)	2,790	42,000	63,400	79,300	101,000	119,000	137,000	157,000	186,000
181	Broad River at Richtex, S.C. (02161500)	4,850	58,800	88,700	112,000	147,000	177,000	210,000	247,000	304,000

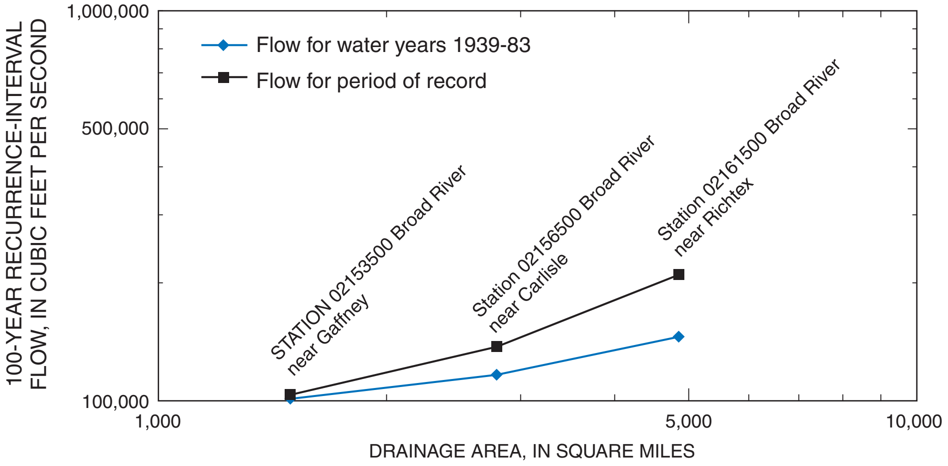


Figure 7. Relation between drainage area and 100-year recurrence-interval flows for period of record and water years 1939-83 for Station 02153500, Broad River near Gaffney, S.C., Station 02156500, Broad River near Carlisle, S.C., and Station 02161500, Broad River at Richtex, S.C.

Little Pee Dee River

The Little Pee Dee River and the Lumber River are unregulated streams that originate in the Sandhill region of North Carolina and flow southward and eastward to the lower Coastal Plain of South Carolina (fig. 6). Soon after entering South Carolina, the Lumber River merges with the Little Pee Dee River and continues to be known as the Little Pee Dee River below the confluence. Station 02132500, Little Pee Dee River near Dillon, S.C., was included in the regional analysis because the drainage area is predominately located in one physiographic province. Flows for selected recurrence intervals are tabulated below for Station 02135000, Little Pee Dee River at Galivants Ferry, S.C. (map index number 177, period of record 1942-99).

Map index No.	Station name and number	Drainage area (square miles)	Flow, in cubic feet per second, for indicated recurrence interval, in years							
			2	5	10	25	50	100	200	500
177	Little Pee Dee River at Galivants Ferry, S.C. (02135000)	2,790	12,400	18,000	21,700	26,300	29,600	33,000	36,300	40,600

Lynches River

The Lynches River originates in the Piedmont of North Carolina, and flows southeastward through the upper Coastal Plain to the lower Coastal Plain of South Carolina (fig. 6). Flows for selected recurrence intervals are tabulated below for Station 02131500, Lynches River near Bishopville, S.C. (map index number 175, period of record 1943-99), and Station 02132000, Lynches River at Effingham, S.C. (map index number 176, period of record 1928-99). Bulletin 17B, appendix 7 (Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982), outlines a procedure for adjusting the logarithmic mean and standard deviation of a short-record site on the basis of a regression analysis with a nearby long-record site. This procedure was used to adjust the logarithmic mean and standard deviation at Station 02131500. The flows for the selected recurrence intervals listed in the table below are smaller for the downstream station, Lynches River near Effingham, S.C., because as the Lynches River flows from the upper to lower Coastal Plain, the flood peaks are attenuated due to increased storage in the floodplain.

Map index No.	Station name and number	Drainage area (square miles)	Flow, in cubic feet per second, for indicated recurrence interval, in years							
			2	5	10	25	50	100	200	500
175	Lynches River near Bishopville, S.C. (02131500)	675	7,710	11,900	14,900	18,700	21,600	24,500	27,500	31,500
176	Lynches River near Effingham, S.C. (02132000)	1,030	5,640	9,170	11,800	15,600	18,700	22,000	25,500	30,600

Edisto River

The North Fork Edisto River and South Fork Edisto River originate in the upper Coastal Plain of South Carolina and flow southeastward to the lower Coastal Plain. The South and North Forks of the Edisto River merge in the lower Coastal Plain to form the Edisto River (fig. 6). Stations 02172500, South Fork Edisto near Montmorence, S.C., 02173000, South Fork Edisto near Denmark, S.C., and 02173500, North Fork Edisto at Orangeburg, S.C. were included in the regional analysis because their respective drainage areas are located in one physiographic province. In addition, Station 02174000, Edisto River near Branchville, S.C. was included in the regional analysis because the drainage area is predominately located in one physiographic province. Flows for selected recurrence intervals for Station 02175000, Edisto River near Givhans, S.C. are tabulated below (map index number 182, period of record 1939-99).

Map index No.	Station name and number	Drainage area (square miles)	Flow, in cubic feet per second, for indicated recurrence interval, in years							
			2	5	10	25	50	100	200	500
182	Edisto River near Givhans, S.C. (02175000)	2,730	9,770	15,000	18,400	22,800	26,000	29,200	32,400	36,600

Salkehatchie River

The Salkehatchie River originates in the upper Coastal Plain of South Carolina and flows southeastward to the lower Coastal Plain (fig. 6). Flows for selected recurrence intervals for Station 02175500, Salkehatchie River near Miley, S.C. are tabulated below (map index number 183, period of record 1951-99).

Map index No.	Station name and number	Drainage area (square miles)	Flow, in cubic feet per second, for indicated recurrence interval, in years							
			2	5	10	25	50	100	200	500
183	Salkehatchie River near Miley, S.C. (02175500)	341	1,510	2,230	2,760	3,480	4,060	4,670	5,320	6,240

FLOOD FREQUENCY AT GAGED SITES ON REGULATED STREAMS

Many of South Carolina's streams, especially the larger ones, are regulated by reservoirs. Flows from reservoirs are regulated to satisfy requirements for in-stream water-use downstream from the reservoirs, power generation, maintenance of lake levels for recreation, and flood control. Regulation procedures may change as flow requirements change. During extremely large floods, the relative effects of storage are diminished and operations are directed more toward preventing dam failure than toward flood control and protection of downstream property. Consequently, flood frequencies for regulated streams are dependent on many more factors than unregulated streams and are, therefore, quite complex and beyond the scope of this report.

For informational purposes, water-year peak-flow data for stations on regulated streams in South Carolina are provided in appendix C of this report. The locations of these gaging stations are shown on figure 5. Other useful information pertaining to regulated streams for which data are available is provided as published in the previous report by Guimaraes and Bohman (1991).

Pee Dee River

The Pee Dee River originates in the Blue Ridge Province of North Carolina and flows through the Piedmont, upper Coastal Plain, and lower Coastal Plain physiographic provinces (fig. 6). Through most of North Carolina, the Pee Dee River is known as the Yadkin River.

Three reservoirs on the Yadkin River and two reservoirs on the Pee Dee River are used for hydroelectric power generation. A sixth reservoir (W. Kerr Scott Reservoir), located on the Yadkin River, is used for flood control and water supply. These reservoirs are located in North Carolina and are not shown on figure 6. Selected data on these reservoirs are listed below (Ruddy and Hitt, 1990).

Name of reservoir	Date of completion	Name of stream	Drainage area (square miles)	Flood-storage capacity (acre-feet)
W. Kerr Scott	1963	Yadkin	350	112,000
High Rock Lake	1927	Yadkin	4,000	64,400
Tuckertown	1962	Yadkin	4,120	0
Badin Lake	1917	Yadkin	4,180	75,800
Lake Tillery	1928	Pee Dee	4,600	29,500
Blewett Falls Lake	1912	Pee Dee	6,830	7,000

Water-year peak-flow data are available for one gaging station on the Pee Dee River downstream from the lakes: Station 02131000, Pee Dee River at Peedee, S.C. (map index number 186, period of record 1939-99) (appendix C).

Catawba River

The Catawba River originates in the Blue Ridge Province of North Carolina, flows through the Piedmont province of South Carolina, and becomes the Wateree River below Lake Wateree (fig. 6). Ten reservoirs, six in North Carolina (not shown on fig. 6) and four in South Carolina (fig. 6), are located on the Catawba River and are storage facilities for hydroelectric power generation. Selected data on nine of these reservoirs are presented below (Ruddy and Hitt, 1990).

Name of reservoir	Date of completion	Drainage area (square miles)	Flood-storage capacity (acre-feet)
Lake James	1919	380	67,900
Rhodhiss Lake	1925	1,090	50,200
Lake Hickory	1928	1,310	61,140
Lookout Shoals Lake	1915	1,450	6,240
Lake Norman	1963	1,790	182,000
Mountain Island	1923	1,860	23,300
Lake Wylie	1925	3,020	33,000
Fishing Creek	1916	3,810	0
Lake Wateree	1919	4,750	45,100

Water-year peak-flow data are available for Station 02146000, Catawba River near Rock Hill (map index number 187, period of record 1896-1903, 1942-99) and Station 02147000, Catawba River near Catawba (map index number 188, period of record 1968-91) (appendix C).

Wateree River

The Wateree River originates at the outflow of Lake Wateree (fig. 6). The Wateree River flows southeastward through the upper Coastal Plain where it merges with the Congaree River to form the Santee River (fig. 6).

Water-year peak-flow data are available for Station 02148000, Wateree River near Camden (map index number 189, period of record 1930-99) (appendix C).

Saluda River

The Saluda River originates in the Blue Ridge Province and flows southeastward through the Piedmont province of South Carolina where it merges with the Broad River near Columbia, S.C., forming the Congaree River (fig. 6). Two reservoirs, Lake Greenwood and Lake Murray, are located on the Saluda River. Selected data for these reservoirs are presented below (Ruddy and Hitt, 1990).

Name of reservoir	Date of completion	Drainage area (square miles)	Flood-storage capacity (acre-feet)
Lake Greenwood	1940	1,150	147,000
Lake Murray	1930	2,420	125,000

The Saluda River above Lake Greenwood is not significantly affected by regulation. Flows for sites along the Saluda River above Lake Greenwood can be determined, by using methods described in previous sections of this report, by estimating flow at an ungaged site, or by estimating flow at or near a gaged site. The three gaging stations on the Saluda River upstream from Lake Greenwood (fig. 1) are Station 02162500, Saluda River near Greenville, S.C. (map index number 25, period of record 1942-99), Station 02163000, Saluda River near Pelzer, S.C. (map index number 26, period of record 1930-93), and Station 02163500, Saluda River near Ware Shoals, S.C. (map index number 27, period of record 1939-99). Records also are available for three stations downstream from Lake Greenwood as shown on figure 6. These are Station 02167000, Saluda River at Chappells, S.C. (map index number 192, period of record 1927-99), Station 02167500, Saluda River near Silverstreet, S.C. (map index number 193, period of record 1928-65), and Station 02169000, Saluda River near Columbia, S.C. (map index number 194, period of record 1926-99) (appendix C).

Congaree River

The Congaree River is formed at the confluence of the Broad River and Saluda River at Columbia, S.C. The Congaree River flows southeastward and joins the Wateree River to form the Santee River (fig. 6). The Broad River Basin makes up about two-thirds of the drainage area of the Congaree River. Flow of the Congaree River is highly regulated by Lake Murray, which is located on the Saluda River. Water-year peak-flow data are available for the Station 02169500, Congaree River at Columbia, S.C. (map index number 195, period of record 1939-99) (appendix C).

Santee River

Formed at the confluence of the Congaree and Wateree Rivers, the Santee River flows directly into Lake Marion - the largest reservoir by surface area in South Carolina (Ruddy and Hitt, 1990).

Name of reservoir	Date of completion	Drainage area (square miles)	Flood-storage capacity (acre-feet)
Lake Marion	1941	14,680	255,000

From 1941 to 1986, most of the flow from Lake Marion was diverted to the Cooper River through a diversion canal to Lake Molutrie (fig. 6). A rediversion canal, completed in 1986, restored approximately 80 percent of the previously diverted flow back to the Santee River. Due to the extensive hydrologic modification, which has taken place over the last 50 to 60 years, frequency computations were not made for stations along the Santee River. However, water-year peak flows for the three stations with 10 or more years of record are listed in appendix C; Station 02170000, Santee River at Ferguson, S.C. (map index number 196, period of record 1908-41), Station 02171500, Santee River near Pineville, S.C. (map index number, 197, period of record 1943-88), Station 02171650, Santee River below St. Stephens, S.C. (map index number 198, period of record 1971-81).

Savannah River

The Tugaloo River and Seneca River, which was inundated by Lake Keowee and Lake Jocasse, originate in the Blue Ridge Province and converge to form the Savannah River (fig. 6). The Savannah River is the State boundary between Georgia and South Carolina and is regulated by three reservoirs along its main stem. The reservoirs are operated by the U.S. Army Corps of Engineers for flood control, power generation, and navigation. Data pertaining to the reservoirs are listed below (Sanders and others, 1990).

Name of reservoir	Date of completion	Drainage area (square miles)	Flood-storage capacity (acre-feet)
Lake Hartwell	1960	2,090	293,000
Richard B. Russell Lake	1984	2,900	140,000
J. Strom Thurmond Lake	1953	6,150	390,000

Sanders and others (1990) showed that the period after 1951 for Station 02197000, Savannah River at Augusta, Ga., was significantly free of extremely large floods, based on an unusually long period of flood data (1796-1985). A flood-frequency relation was established for the site using peak flows computed by the routing of synthesized inflow hydrographs through the reservoirs to the site based on operating conditions at the time of the study. The frequency data from that study for Station 02197000 are tabulated below.

Map index No.	Station name and number	Drainage area (square miles)	Flow, in cubic feet per second, for indicated recurrence interval, in years					
			2	5	10	25	50	100
203	Savannah River at Augusta, Ga. (02197000)	7,510	34,500	51,500	69,000	105,000	140,000	180,000

Water-year peak-flow data for five stations with 10 or more years of record on the Savannah River are listed in appendix C: Station 02187500, Savannah River near Iva, S.C. (map index number 199, period of record 1950-81), Station 02189000, Savannah River near Calhoun Falls, S.C. (map index number 201, period of record 1897, 1900-80), Station 02195000, Savannah River near Clarks Hill, S.C. (map index number 202, period of record 1940-54), Station 02197500, Savannah River at Burtons Ferry Bridge near Millhaven, Ga. (map index number 204, period of record 1940-99), and Station 02198500, Savannah River near Clyo, Ga. (map index number 205, period of record 1925-99).

Pacolet River

The Pacolet River originates in the Blue Ridge Province and flows southeastward to the Piedmont Province where it merges with the Broad River (fig. 6). The South Pacolet River is regulated by Lake William C. Bowen, and the North Pacolet River is unregulated. Selected data for the reservoir are presented below (Ruddy and Hitt, 1990).

Name of reservoir	Date of completion	Drainage area (square miles)	Flood-storage capacity (acre-feet)
William C. Bowen	1956	79.4	9,600

Water-year peak-flow data are available for Station 02155500, Pacolet River near Fingerville, S.C. (map index number 190, period of record 1931-99) and Station 02156000, Pacolet River near Clifton, S.C. (map index number 191, period of record 1940-78) (appendix C).

SUMMARY

Flood magnitude and frequency estimates are important factors in the design of bridges, highway embankments, culverts, levees, and other structures near streams. This report describes methods for determining flood magnitude and frequency at rural streamflow-gaging stations and rural, ungaged sites in South Carolina. For this study, 167 streamflow-gaging stations in or near South Carolina were used in the regional regression analysis (54 in South Carolina, 65 in North Carolina, and 48 in Georgia). Stations used for this study have 10 years or more of water-year peak-flow data within a rural basin that is not significantly affected by regulation. The database has drainage areas that range from 0.1 to 1,720 square miles.

Peak flow data were analyzed for the mean, standard deviation, and skew using a log-Pearson Type III distribution. Using the station skew at sites with 25 years or more of data, generalized skew coefficients were developed for South Carolina. The skews were computed using a weighted arithmetic mean with the weight being based on record length. No statistically significant difference was found among the station skews in the Blue Ridge, upper Coastal Plain, and lower Coastal Plain. Station skews for the Piedmont, however, showed a statistically significant difference. Consequently, two generalized skew coefficients were computed for South Carolina and used in the final log-Pearson Type III analysis. The regional skew was weighted with the station skew and the weighted skew was used within the log-Pearson Type III analysis to determine the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence interval flows at each streamflow-gaging station.

Regional regression analysis, using generalized least squares regression, was used to develop a set of predictive equations that can be used to estimate the 2-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence-interval flows for rural ungaged basins in the Blue Ridge, Piedmont, upper Coastal Plain, and lower Coastal Plain physiographic provinces of South Carolina. The predictive equations are all functions of drainage area. Average errors of prediction for these regression equations ranged from -16 to 19 percent for the 2-year recurrence-interval flow in the upper Coastal Plain to -34 to 52 percent for the 500-year recurrence-interval flow in the lower Coastal Plain.

A region-of-influence method also was developed that interactively estimates the recurrence-interval flows for rural ungaged basins in South Carolina. The predictive errors from the regional regression method were compared with those from the region-of-influence method. From the comparison, it was concluded that the region-of-influence method only produced lower predictive errors in the Blue Ridge physiographic province; therefore, the region-of-influence method should be used only in the Blue Ridge. A computer program was developed that computes the selected recurrence-interval flows listed in the previous paragraph using the appropriate method as chosen by the user.

The methods described in this report are suitable only for use on unaltered, rural streams. These methods should not be used where dams, flood-detention structures, or other anthropogenic factors may significantly affect the peak-flow data. Furthermore, the models should be used within the range of the drainage areas used during model development for each physiographic province.

Appendix A provides the data for the 54 South Carolina gaging stations used in the frequency analyses. The appendix includes peak stages and flows, frequency and statistical data from the log-Pearson Type III analysis, and a station description. Appendix D provides peak stages and flows and a station description for regulated stations having 10 or more years of record. In addition, appendix D includes peak stages and flows, frequency and statistical data from the log-Pearson Type III analysis, and a station description for stations that drain more than one physiographic province but that are not significantly affected by regulation and have 10 or more years of record.

SELECTED REFERENCES

Acreman, M.C., and Wiltshire, S.E., 1987, Identification of regions for regional flood frequency analysis [abs.], EOS, v. 68, no. 44.

Asquith, W.H., and Slade, Jr., R.M., 1999, Site-specific estimation of peak-streamflow frequency using generalized least squares regression for natural basins in Texas: U.S. Geological Survey Water-Resources Investigations Report 99-4172, 19 p.

Bloxham, W.M., 1976, Low-flow characteristics of streams in the inner Coastal Plain of South Carolina: South Carolina Water Resources Commission Report No. 5, 31 p.

-----1979, Low-flow frequency and flow duration of South Carolina streams: South Carolina Water Resources Commission Report No. 11, 48 p.

-----1981, Low-flow characteristics of ungaged streams in the Piedmont and lower Coastal Plain of South Carolina: South Carolina Water Resources Commission Report No. 14, 48 p.

Bodhaine, G.L., 1968, Measurement of peak discharge at culverts by indirect methods: U.S. Geological Survey Techniques of Water-Resources Investigations Book 3, chap. A3, 60 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 Techniques of Water-Resources Investigations Report 92-4040, 79 p.

Burn, D.H., 1990a, An appraisal of the “region of influence” approach to flood frequency analysis: Hydrogeological Sciences Journal, v. 35, no. 24, p. 149-165.

----- 1990b, Evaluation of regional flood frequency analysis with a region of influence approach: Water-Resources Research, v. 26, no. 10, p. 2257-2265.

Cooke, C.W., 1936, Geology of the Coastal Plain of South Carolina: U.S. Geological Survey Bulletin 867, 196 p.

Cooney, T.W., Drewes, P.A., Ellisor, S.W., and Melendez, Frank. 1999, Water resources data, South Carolina water year 1999, U.S. Geological Survey Water-Data Report SC-99-1, 632 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.

Ensminger, P.A., Floods in Louisiana, magnitude and frequency, 1998 (5th ed.): U.S. Geological Survey Water-Resources Technical Report No. 60, 333 p.

Guimaraes, W.B., and Bohman, L.R., 1991, Techniques for estimating magnitude and frequency of floods in South Carolina, 1988: U.S. Geological Survey Water-Resources Investigations Report 91-4157, 174 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 Investigations Report 87-4096, 52 p.

Hardison, C.H., 1971, Prediction error of regression estimates of streamflow characteristics at ungaged sites: U.S. Geological Survey Professional Paper 750-C, p. 228-236.

Helsel, R.M., Hirsch, D.R., 1992, Studies in environmental science 49--Statistical methods in water resources, Amsterdam, Elsevier Science, 529 p.

Hodge, S.A., and Tasker, G.D., 1995, Magnitude and frequency of floods in Arkansas: U.S. Geological Survey Water-Resources Investigations Report 95-4224, 52 p.

Hodgkins, G.A., 1999, Estimating the magnitude of peak flows for streams in Maine for selected recurrence intervals: U.S. Geological Survey Water-Resources Investigations Report 99-4008, 45 p.

Hydrology Subcommittee of the Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood frequency: U.S. Geological Survey Bulletin 17B, Office of Water Data Collection, Reston, Va., 183 p.

Lanier, T.H., 1996, Determination of the 100-year flood plain on Upper Three Runs and selected tributaries, and the Savannah River at the Savannah River Site, South Carolina, 1995: U.S. Geological Survey Water-Resources Investigations Report 96-4014, 65 p.

Lichty, R.W., and Liscum, F., 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.

Pope, B.F., and Tasker, G.D., 1999, Estimating the magnitude and frequency of floods in rural basins of North Carolina, U.S. Geological Survey Water-Resources Investigations Report 99-4114, 44 p.

Pope, B.F., Tasker, G.D., and Robbins, J.C., 2001, Estimating magnitude and frequency of floods in rural basins of North Carolina (revised): U.S. Geological Survey Water-Resources Investigations Report 01-4207, 44 p.

Rantz, S.E., and others, 1982, Measurement and computation of streamflow: Volume 2. Computation of discharge: U.S. Geological Survey Water-Supply Paper 2175, 347 p.

Riggs, H.C., 1973, Regional analysis of streamflow characteristics: Techniques of Water- Resources Investigations of the U.S. Geological Survey, Book 4, chap. B3, 15 p.

Robbins, J.C., and Pope, B.F., 1996, Estimation of flood-frequency characteristics of small urban streams in North Carolina: U.S. Geological Survey Water-Resources Investigations Report, 96-4084, 12 p.

Ruddy, B.C., and Hitt, K.J., 1990, Summary of selected characteristics of large reservoirs in the United States and Puerto Rico, 1988: U.S. Geological Survey Open-File Report 90-163, 295 p.

Sanders, Jr., C.L., Kubik, H.E., Hoke, J.T., Jr., Kirby, W.H., 1990, Flood frequency of the Savannah River at Augusta, Georgia: U.S. Geological Survey Water-Resources Investigations Report, 90-4024, 87 p.

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

Seaber, P.R., Kapinos, F.P., and Knapp, G.L., 1987, Hydrologic unit maps: U.S. Geological Survey Water-Supply Paper 2294, 63 p.

Smith, B.R., and Hollbick, D.C., 1979, General soil map, South Carolina: South Carolina Agricultural Experiment Station, Clemson University, Soil Map 48.

South Carolina Water Resources Commission, 1983, South Carolina State water assessment: South Carolina Water Resources Commission Report No. 140, p. 367.

Speer, P.R., and Gamble, C.R., 1964, Magnitude and frequency of floods in the United States, Part 2-A, South Atlantic Slope basins, James River to Savannah River: U.S. Geological Survey Water-Supply Paper 1673, p. 329.

Stamey, T.C., and Hess, G.W., 1993, Techniques for estimating magnitude and frequency of floods in rural basins of Georgia: U.S. Geological Survey Water-Resources Investigations Report 93-4016, 75 p.

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

----- 1986a, Correction to regional hydrologic analysis 1. Ordinary, weighted, and generalized least square compared: Water-Resources Research, v. 22, no. 5, p. 844.

----- 1986b, Correlation to regional hydrologic analysis, 1. Ordinary, weighted and generalized least squares compared: Water Resources Research, v. 22, no. 5, 844 p.

Tasker, G.D., 1978, Flood frequency analysis with a generalized skew coefficient: Water Resources Research, v. 14, no. 2, p. 373-376.

Tasker, G.D., Hodge, S.A., and Barks, C.S., 1996, Region of influence regression for estimating the 50-year flood at ungaged sites: Water Resources Bulletin, v. 32, no. 1, p. 163-170.

Tasker, G.D., and Slade, R.M., 1994, An interactive regional regression approach to estimating flood quantiles, in Fontane, D.G., and Tuvel, H.N., eds., Proceedings of the Twenty-First Annual Conference, Water policy and management, solving the problems, May 23-26, 1994: Denver, Colo., Denver Society of American Society of Civil Engineers, p. 782-785.

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, 1976, Guidelines for determining flood flow frequency: U.S. Water Resources Council Bulletin 17, 26 p.

Whetstone, B.H., 1982, Techniques for estimating magnitude and frequency of floods in South Carolina: U.S. Geological Survey Water-Resources Investigations Report 82-1, p. 78.

Printed on recycled paper

Techniques for Estimating the Magnitude and Frequency of Floods in Rural Basins of South Carolina, 1999

FEASTER

USGS WRIR 02-4140