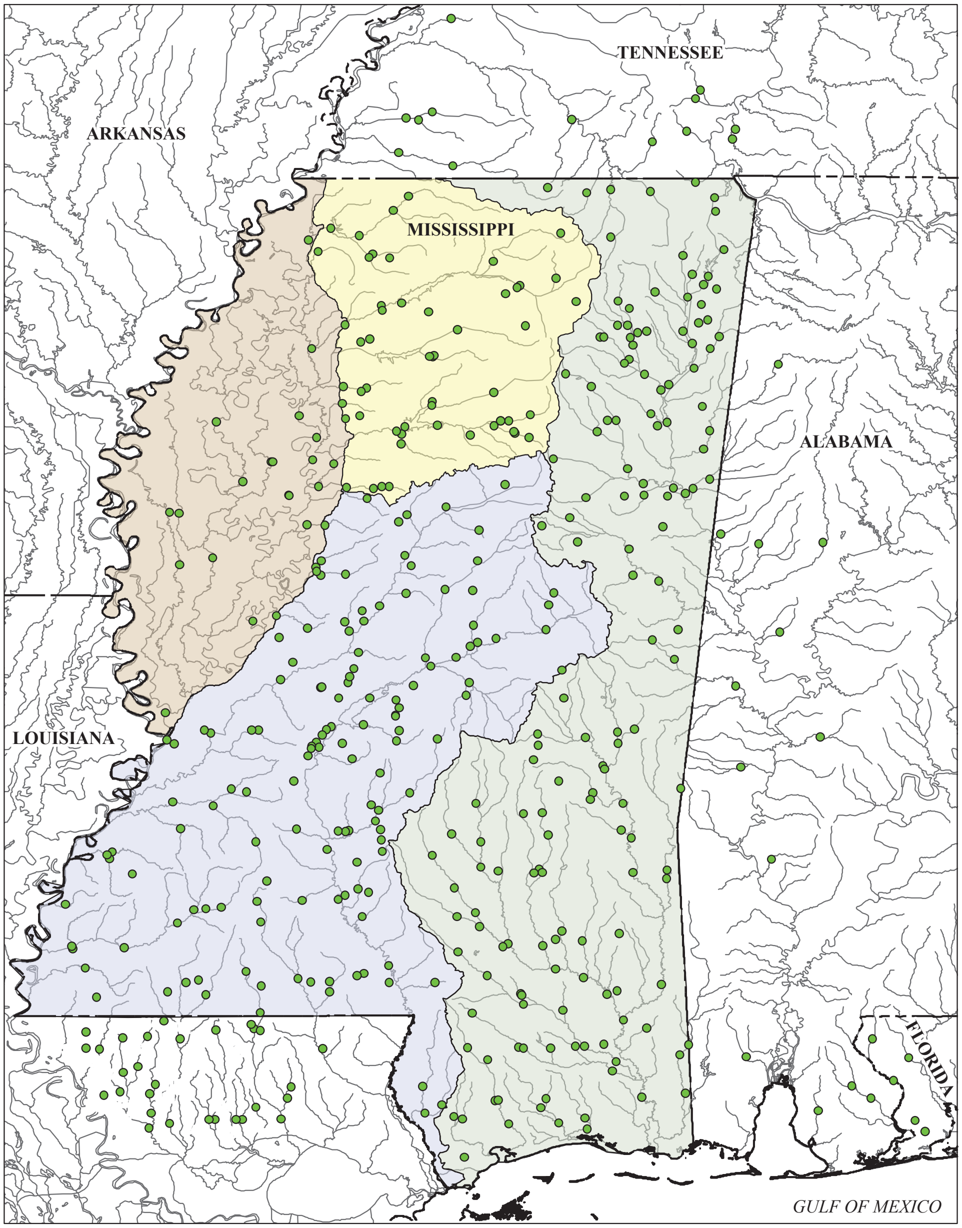
Prepared in cooperation with the Mississippi Department of Transportation

Flood Frequency of Rural Streams in Mississippi, 2013



Scientific Investigations Report 2018–5148

Supersedes USGS Water-Resources Investigations Report 91–4037

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

Flood Frequency of Rural Streams in Mississippi, 2013

By Brandon T. Anderson

Prepared in cooperation with the Mississippi Department of Transportation

Scientific Investigations Report 2018–5148
Supersedes USGS Water-Resources Investigations Report 91–4037

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

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
James F. Reilly II, Director

U.S. Geological Survey, Reston, Virginia: 2018
Supersedes USGS Water-Resources Investigations Report 91–4037

Suggested citation:
Anderson, B.T., 2018, Flood frequency of rural streams in Mississippi, 2013: U.S. Geological Survey Scientific Investigations Report 2018–5148, 12 p., https://doi.org/10.3133/sir20185148. [Supersedes USGS Water-Resources Investigations Report 91–4037.]

ISSN 2328-031X (print)
ISSN 2328-0328 (online)

ISBN 978-1-4113-4264-4

Acknowledgments

The following U.S. Geological Survey personnel are acknowledged for their assistance in developing the computations for this project and helping to complete the report. T. Scott Hedgecock, Hydrologist in the U.S. Geological Survey Lower Mississippi-Gulf Water Science Center (USGS LMGWSC) Montgomery, Alabama, office, provided mentorship during the project. Andrea Veilleux, Hydrologist in the USGS Integrated Modeling and Prediction Division, Analysis and Prediction Branch, Reston, Virginia, assisted in the development of the study-specific skew and provided project support. Brian Breaker, Hydrologist in the USGS LMGWSC, Little Rock, Arkansas, office helped with statistical software. Stephanie Kula, Biologist in the USGS Ohio-Kentucky-Indiana Water Science Center, provided geographic information system (GIS) data for the development of StreamStats and provided GIS-derived basin characteristics.

Contents

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
	Hydraulic gradient	
foot per mile (ft/mi)	0.1894	meter per kilometer (m/km)

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C = (°F - 32) / 1.8.$$

Datum

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).

Vertical coordinate information is referenced to the North American Vertical Datum of 1988 (NAVD 88) or the National Geodetic Vertical Datum of 1929 (NGVD 29).

Elevation, as used in this report, refers to distance above the vertical datum.

Abbreviations

AEP	Annual exceedance probability
AVP	Average variance of prediction
EMA	Expected moments algorithm
GIS	Geographic information system
GLS	Generalized least squares
HUC	Hydrologic unit code
LP3	Log-Pearson type III
MAP	Mississippi Alluvial Plain
MGB	Multiple Grubbs-Beck test
MSE	Mean square error
MSEp	Mean square error of prediction
NWIS	National Water Information System
OLS	Ordinary least squares
PILFs	Potentially influential low floods
RRE	Regional regression equation
SEM	Standard error of the model
SEP	Standard error of prediction
USGS	U.S. Geological Survey
WREG	Weighted-multiple-linear regression program

Flood Frequency of Rural Streams in Mississippi, 2013

By Brandon T. Anderson

Abstract

To improve flood-frequency estimates at rural streams in Mississippi, annual exceedance probability flows at gaged streams and regional regression equations used to estimate annual exceedance probability flows for ungaged streams were developed by using current geospatial data, new analytical methods, and annual peak-flow data through the 2013 water year. The regional regression equations were derived from statistical analyses of peak-flow data and basin characteristics for 281 streamgages and incorporated a newly developed study-specific skew coefficient at streamgages located in five subregional watersheds (Middle Tennessee-Elk, Mobile-Tombigbee, Lower Mississippi-Big Black, Pearl, and Pascagoula) in Mississippi. Three flood regions—A, B, and C—were identified based on residuals from the regional regression analyses and contain sites with similar basin characteristics. Analysis was not conducted for the fourth flood region, the Mississippi Alluvial Plain, because of insufficient long-term streamflow data and poorly defined basin characteristics.

Introduction

Improved flood-frequency information is important for the effective management of flood plains, including the safe and economic design of bridges, culverts, dams, levees, and other structures near streams. The last flood-frequency study for Mississippi was published more than 25 years ago (Landers and Wilson, 1991). Since that time, improvements in estimation techniques have increased the accuracy of flood-frequency estimates. One such technique is the expected moments algorithm (EMA), which allows historical data to be represented as intervals (Cohn and others, 1997). Another is the multiple Grubbs-Beck (MGB) test, which increases the accuracy of peak-flow statistics by objectively and systematically detecting and removing low, highly variable peak flows, and can be used with the EMA (Cohn and others, 2013). In 1982, the Interagency Advisory Committee on Water Data presented a generalized regional skew coefficient map for the United States that was developed using peak-flow data collected through 1976 (Interagency Advisory Committee on Water Data, 1982). Since then, an additional 37 years of peak-flow data have been collected and can be used to develop updated skew coefficients.

In 2000, the U.S. Geological Survey (USGS), in cooperation with the Mississippi Department of Transportation, began a study to update the regional regression equations (RRE) and annual exceedance probability (AEP) flows for rural streams in Mississippi using recent geospatial data, new analytical methods, and additional annual peak-flow data through the 2013 water year.[1] Results of the study, including the updated AEP flows and RRE information, would be published and also incorporated into the USGS StreamStats application, which is an online tool that provides flood plain planners and water managers in Mississippi with basin characteristics and estimates of flow statistics at gaged and ungaged stream sites (http://water.usgs.gov/osw/streamstats/).

Purpose and Scope

The purposes of this report are to (1) describe the development of a study-specific skew coefficient that was applied to sites located in five subregional watersheds in Mississippi; (2) update flood regions for all areas in Mississippi except the Mississippi Alluvial Plain (MAP); (3) update the annual exceedance probability flows at gaged locations; and (4) update regional flood-frequency equations (regional regression equations) for use at ungaged locations in Mississippi. The MAP was not included in this analysis because of insufficient long-term streamflow data and poorly defined basin characteristics. All data used in support of the analysis presented in this report, including geographic information system (GIS) data for the four flood regions in Mississippi, are available from Anderson (2018). This report supersedes the flood-frequency analysis published by Landers and Wilson (1991).

Description of Study Area

The study area (fig. 1) includes the State of Mississippi and selected locations in Louisiana, Alabama, Tennessee, and Florida. Rainfall in Mississippi generally is associated

[1]The water year is the annual period from October 1 through September 30 and is designated by the year in which the period ends. For example, the 2013 water year is from October 1, 2012, through September 30, 2013.

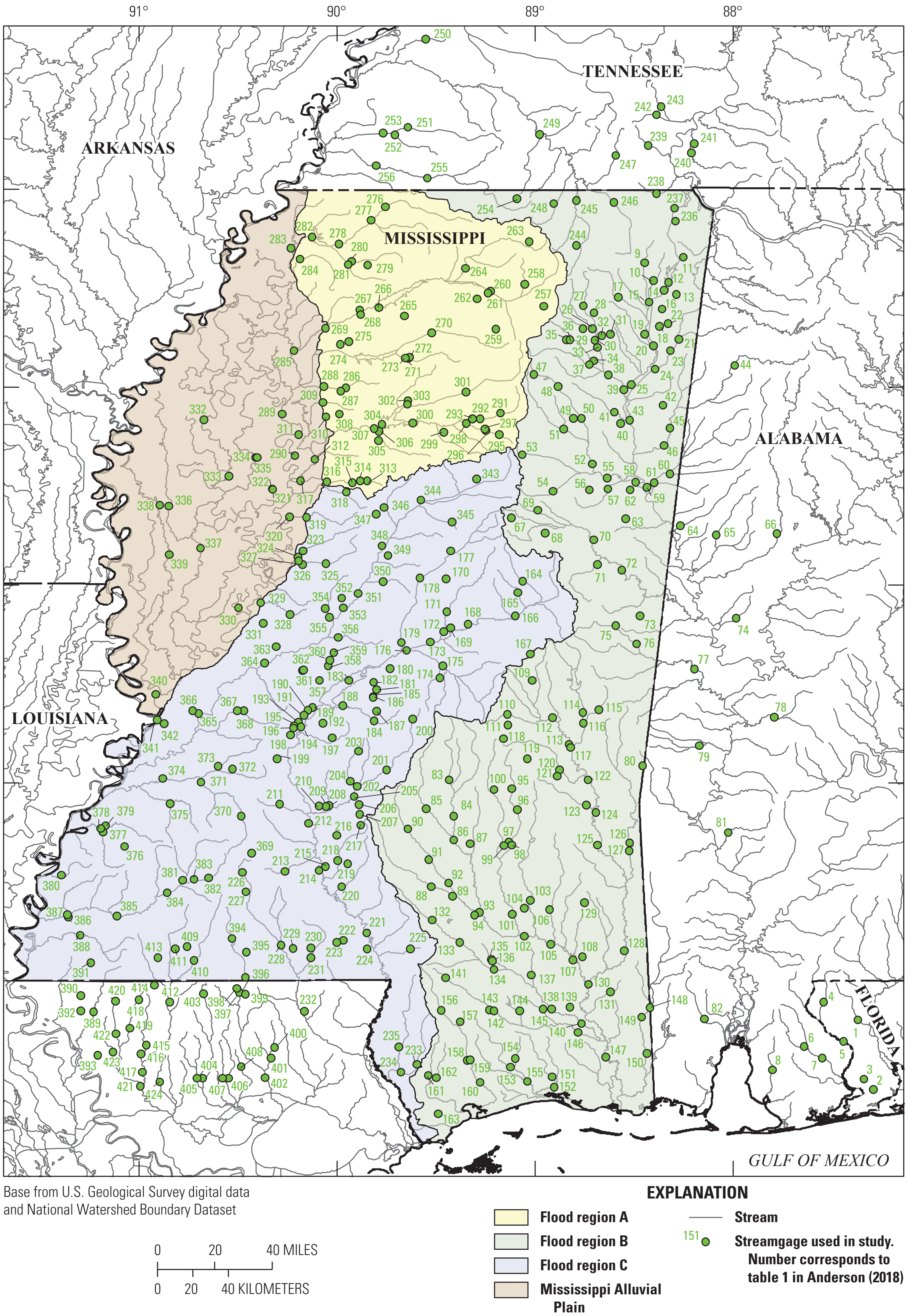


Figure 1. Locations of streamgages included in this study and the four newly developed flood regions in Mississippi, 2013.

with the movement of warm and cold fronts across the State from November through April and isolated thunderstorms from May through October. From June to September, tropical storms or hurricanes occasionally enter the State along the gulf coast and produce unusually heavy amounts of rainfall. The average annual precipitation for Mississippi is 54.16 inches (U.S. Climate Data, 2017). The average annual high and low temperatures are 75.5 and 53.6 degrees Fahrenheit, respectively (U.S. Climate Data, 2017).

Previous Investigations

Wilson and Trotter (1961) developed techniques for estimating the magnitude and frequency of floods at streams in Mississippi. Colson and Hudson (1976) used a multiple-linear regression model to update those techniques. Landers (1985) used linear regressions to develop updated techniques for estimating the magnitude and frequency of floods at streams located in the alluvial plain of the lower Mississippi River. Landers and Wilson (1991) used a generalized least-squares regression to develop updated techniques for estimating the magnitude and frequency of flows in streams located in Mississippi. From the residuals in the 1991 analysis, Landers and Wilson defined three flood regions in Mississippi (East, West, and the MAP).

Data Compilation

USGS streamgages in Mississippi, Louisiana, Alabama, Florida, and Tennessee that have 10 or more years of annual peak-flow data were used in this analysis. Streamgages with annual peak-flow data are either continuous-record gages or crest-stage gages. Continuous-record streamgages are equipped with instrumentation to record the water-surface elevation, or stage, of the water body at fixed-time intervals. The stage data are transmitted by satellite to the local USGS offices and are applied to a stage-discharge rating to determine flow for the given stage value. Crest-stage gages record only the peak stage of a flood; the peak stage is then applied to a stage-discharge rating to determine the associated flow. Hereafter, these two types of gages are referred to as streamgages.

Peak-Flow Data and Basin Characteristics

This analysis was conducted by using data from 424 streamgages (341 in Mississippi and 83 from adjacent States) that had 10 or more years of annual peak-flow data (systematic and historic peaks) collected through the end of the 2013 water year (fig. 1). Peak-flow data for the streamgages were downloaded from the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2017a). The drainage areas for the 424 streamgages range from 0.04 to 1,144,500 square miles (mi^2). The drainage areas of the 341 streamgages in Mississippi (excluding the Mississippi River at Vicksburg, MS, streamgage) range from 0.05 to 13,400 mi^2. Annual peak-flow data were analyzed for trends using the Mann-Kendall test. Trends in annual peak-flow data could bias the AEP flow analyses because an assumption of probability analyses is that annual peak flows are independent and stationary through time. However, annual peak-flow data can exhibit some serial correlation that can cause the Mann-Kendall trend test to indicate a significant trend when there is none, especially at streamgages with less than 30 years of peak-flow data (Hodgkins and Martin, 2003). For this reason, the Mann-Kendall trend test was not considered for streamgages with less than 30 years of peak-flow data. A statistically significant trend ($p \leq 0.05$) in peak flow was shown at 27 streamgages, representing approximately 6 percent of the total sites. A positive trend was indicated for 11 and a negative trend indicated for 16 streamgage annual peak flows. These sites represent a minority of the total sites, are distributed almost evenly between positive and negative trends, and are believed to be chance occurrences rather than true trends; therefore, no sites were removed from the analyses.

Twenty-one basin characteristics for each site included in this study were obtained by using the USGS map-based web application StreamStats (U.S. Geological Survey, 2017b). Because soil types and land use can influence runoff, the GIS-derived basin characteristics from StreamStats were compiled for each basin by using the 2011 National Land Cover Database (Homer and others, 2015) (table 1).

Analysis of Flow at Gaged Locations

Flood-frequency estimates for gaged locations were computed by fitting a Log-Pearson Type III (LP3) mathematical probability distribution to the series of annual peak flows as described in Bulletin 17B (Interagency Advisory Committee on Water Data, 1982). The LP3 distribution is a three-parameter distribution that requires estimates of the mean, standard deviation, and skew coefficient of the population of logarithms of annual peak flow at each gaged site (Parrett and others, 2011). The EMA method produces estimates of the three LP3 statistics that are identical to those produced by the standard LP3 method described in Bulletin 17B. The EMA method improves upon the standard LP3 method by allowing for the analysis of historical peak datasets containing censored observations, historical data, low outliers, and uncertain data points and accommodates the interval data by using perception thresholds and flow intervals (Cohn and others, 1997).

Table 1. Basin characteristics used in the regional regression analysis for determining flood frequency of rural streams in Mississippi, 2013.

[NLCD, National Land Cover Database from Homer and others (2015); A21, area of developed open land from NLCD class 21; A22, area of developed land, low intensity, from NLCD class 22; A23, area of developed land, medium intensity, NLCD 2011 class 23; A24, area of developed land, high intensity, NLCD class 24]

Basin characteristic definition	Units	Abbreviation	Source(s)
Area that drains to a point on a stream	Square miles	DRNAREA	StreamStats, version 3.0
Change in elevation between points 10 and 85 percent of length along main channel to basin divide divided by length between points ft per mi	Feet per mile	CSL10_85fm	StreamStats, version 3.0
Latitude of Basin Centroid	Decimal degrees	LAT_CENT	StreamStats, version 3.0
Longitude of Basin Centroid	Decimal degrees	LONG_CENT	StreamStats, version 3.0
Maximum basin elevation	Feet	ELEVMAX	StreamStats, version 3.0
Minimum basin elevation	Feet	MINBELEV	StreamStats, version 3.0
Length of longest flow path	Mile	LFPLENGTH	StreamStats, version 3.0
Area of developed land-use from NLCD 2011 classes 21-24	Square miles	LC11ADEV	StreamStats, version 3.0; NLCD
Percentage of developed (urban) land from NLCD 2011 classes 21-24	Percent	LC11DEV	StreamStats, version 3.0; NLCD
Area of developed open land from NLCD 2011 class 21	Square miles	LC11ADVOPN	StreamStats, version 3.0; NLCD
Area of developed land, low intensity, from NLCD 2011 class 22	Square miles	LC11ADEVLO	StreamStats, version 3.0; NLCD
Area of developed land, medium intensity, NLCD 2011 class 23	Square miles	LC11ADEVMD	StreamStats, version 3.0; NLCD
Area of developed land, high intensity, NLCD 2011 class 24	Square miles	LC11ADEVHI	StreamStats, version 3.0; NLCD
Percentage drainage area that is in low to high developed land-use classes 22-24 from NLCD 2011	Percent	LC11DEVLMH	StreamStats, version 3.0; NLCD
Impervious percentage computed as ((.10*A21+.25*A22+.65*A23+.90*A24)/DA)*100 from NLCD 2011	Percent	LC11DINT	StreamStats, version 3.0; NLCD
Percentage of forest from NLCD 2011 classes 41-43	Percent	LC11FOREST	StreamStats, version 3.0; NLCD
Area of water from NLCD 2011 class 11	Square miles	LC11AWATER	StreamStats, version 3.0; NLCD
Percent of open water, class 11, from NLCD 2011	Percent	LC11WATER	StreamStats, version 3.0; NLCD
Area of wetlands from NLCD 2011 classes 90 and 95	Square miles	LC11AWETL	StreamStats, version 3.0; NLCD
Percentage of wetlands, classes 90 and 95, from NLCD 2011	Percent	LC11WETLND	StreamStats, version 3.0; NLCD
Percentage of water bodies and wetlands determined from the NLCD 2011	Percent	LC11STOR	StreamStats, version 3.0; NLCD

The basic equation for fitting the LP3 distribution to a measured series of annual peak flows is

$$\log Q_p = \bar{X} + K_p S, \tag{1}$$

where

Q_p is the P-percent AEP flow in cubic feet per second;

$\bar{X}$ is the mean of the logarithms of the annual peak flows;

K_p is a factor based on the skew coefficient and the given percentage of annual exceedance probability, which can be obtained from appendix 3 of Bulletin 17B (Interagency Advisory Committee on Water Data, 1982); and

S is the standard deviation of the logarithms of the annual peak flows.

In previous USGS reports about floods in Mississippi, the term "recurrence interval, in years" has been used to characterize flood frequency (50-year flood, and so forth). The USGS and other Federal agencies now refer to the P-percent chance of occurrence as an annual exceedance probability (AEP). For example, the 0.02 percent ($Q_{2\%}$) AEP flood has a 2-percent chance of occurring in any given year and corresponds to a recurrence interval of 50 years (reciprocal of the AEP; table 2) (Griffis and Stedinger, 2007). An increase in the number of years of peak-flow record at a streamgage increases the level of confidence in computed AEP flows. For example, a stream with 30 years of historical peak-flow data will have a lower variance than a stream with 10 years of record, thus increasing the confidence of the estimated AEP flow.

Table 2. T-year recurrence intervals with corresponding annual exceedance probability and P-percent chance exceedance for flood-frequency flow estimates (Feaster and others, 2009).

T-year recurrence interval	Annual exceedance probability	P-percent annual exceedance probability
2	0.5	50
5	0.2	20
10	0.1	10
25	0.04	4
50	0.02	2
100	0.01	1
200	0.005	0.5
500	0.002	0.2

The MGB test, a generalization of the Grubbs-Beck method, provides a standard procedure for identifying low-flow outliers and multiple potentially influential low floods (PILFs) (Cohn and others, 2013). PILFs are annual peaks that meet three criteria: (1) their magnitude is much smaller than the flood quantile of interest; (2) they occur below a statistically significant break in the flood-frequency plot; and (3) they have excessive influence on the estimated frequency of large floods. The USGS PeakFQ software version 7.1, available for download from https://water.usgs.gov/software/PeakFQ/ (U.S. Geological Survey, 2014), was used to conduct the EMA/MGB test.

Development of a Study-Specific Skew Coefficient

The skew coefficient measures the asymmetry of the probability distribution of a set of annual peak flows, which is affected by the presence of high or low outliers (Wagner and others, 2016). Peak-flow data available through water-year 2013 were used to develop a skew coefficient, hereafter referred to as the study-specific skew coefficient, for five subregional watersheds in Mississippi: Middle Tennessee-Elk, Mobile-Tombigbee, Lower Mississippi-Big Black, Pearl, and Pascagoula (fig. 2).

Two criteria were required to ensure consistency with the skew analysis outlined in Bulletin 17B (Interagency Advisory Committee on Water Data, 1982) and compatibility with previous flood-frequency studies across the region. Streamgage sites selected for the study-specific skew coefficient development could not have substantial urbanization, regulation, diversion, or tidal influence. Station skew coefficients at altered sites could be location specific based on the type and degree of alteration and have the potential to affect the skew analysis. The other criterion was that a streamgage station needed 25 years or more of systematic peak-flow record. Based on these criteria, 93 rural streamgage stations were selected for use in computing the study-specific skew coefficient and mean square error (MSE). Station skew coefficient statistics were compiled for each of the 93 candidate stations (table 3) using the software PeakFQ, version 7.1 (https://water.usgs.gov/software/PeakFQ/; U.S. Geological Survey, 2014). The arithmetic mean of the 93 individual station skew coefficients yielded a study-specific skew coefficient of −0.08, and the MSE was computed to be 0.22. This study-specific skew coefficient replaced the generalized skew coefficient from Bulletin 17B (Interagency Advisory Committee on Water Data, 1982) in the five subregional watersheds that contained all 93 streamgages and was weighted with the station skew coefficient at each site to calculate the AEP flows. In all other areas of Mississippi and areas of adjacent States (fig. 2), the Bulletin 17B generalized skew coefficient was used for station weighting. It is important to note that the study-specific skew coefficient cannot be applied in other subregional watersheds in Mississippi.

Redundant sites were not considered for the study-specific skew coefficient analysis because of potential bias. Redundancy occurs when drainage basins of two gaged sites are nested, meaning that one is contained inside the other, and the sizes of the two basins are similar (Parrett and others, 2011). Thus, instead of providing two independent spatial observations that relate drainage basin characteristics to skew coefficient (or flood quantiles), these two basins likely have the same hydrologic response to a given storm and represent only one spatial observation (Parrett and others, 2011). When sites are considered redundant, statistical analyses using both gaged sites incorrectly represent the information in the regional dataset (Gruber and Stedinger, 2008). Sites were considered redundant if gages on the same stream possessed drainage area ratios within 50 percent of each other. When evaluating redundancy between two sites in a reach, the site having the longer period of record was chosen to represent that reach or sub-reach.

Regional Regression Analysis

Of the 424 streamgages considered for inclusion in this study, 281 were used for the regional regression analysis (fig. 1). Streamgage information and additional data related to the regional regression analysis are provided in the data release associated with this study (Anderson, 2018). Streamgages were evaluated for backwater, regulation, diversion, channelization, and urbanization by compiling historical data and inspecting NWIS peak-flow data qualification codes; nested sites and sites having any of these attributes were removed. Basins of sites having more than 20 percent development also were removed. Redundant sites were not considered for the RRE analysis because of potential bias. A statewide ordinary least squares (OLS) regression equation was developed for the State, excluding the MAP, by using drainage area as the only explanatory variable for the 1 percent ($Q_{1\%}$) AEP. The residuals for each streamgage were evaluated to detect any geographic biases or clusters. Residuals for the streamgages represent the difference between observed values and the predicted values of streamflow using the RREs.

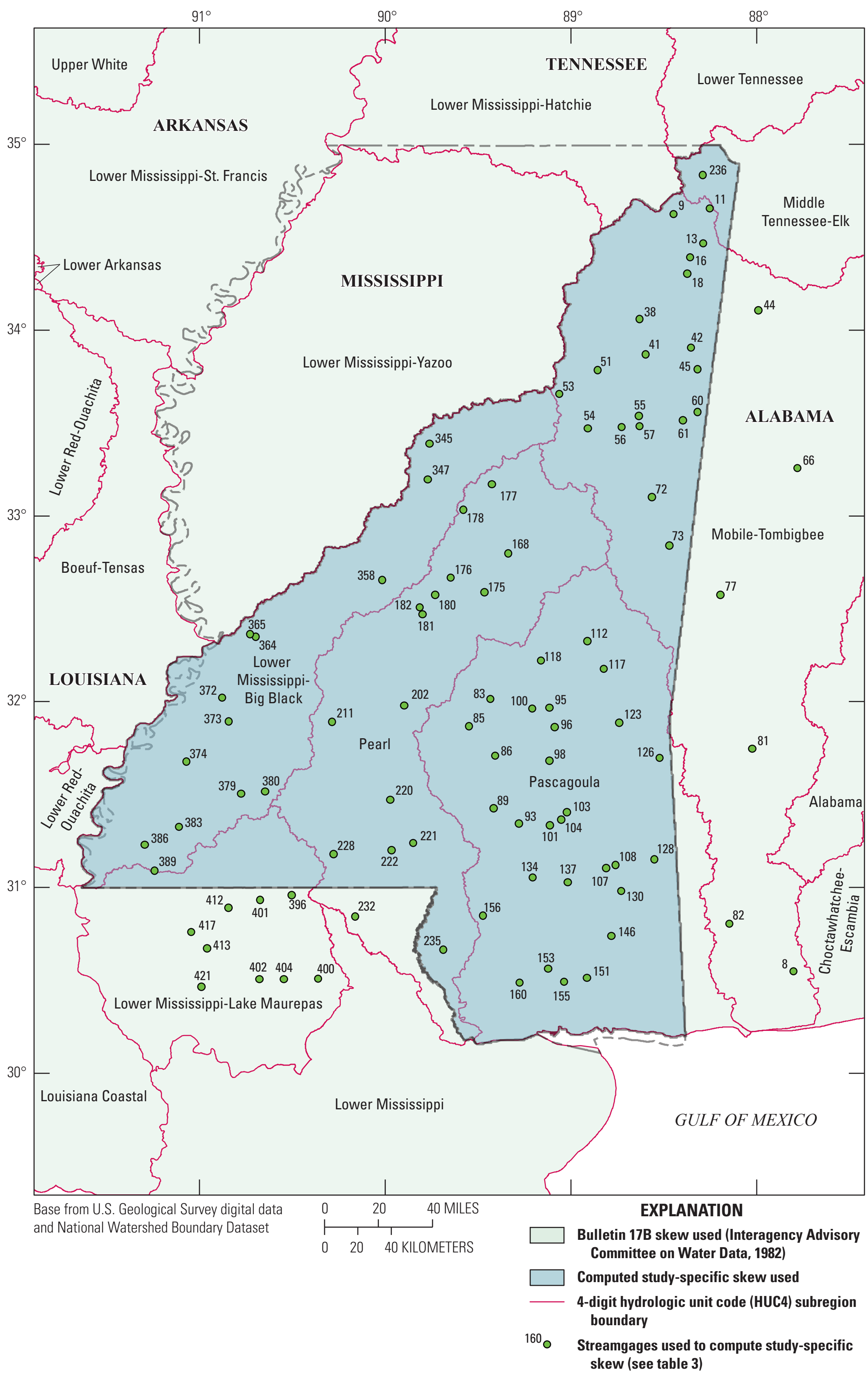


Figure 2. Locations of streamgages used to compute the study-specific skew coefficient in select subregional watersheds in Mississippi.

Table 3. Streamgages used to compute the study-specific skew in select subregional watersheds in Mississippi—the Middle Tennessee-Elk, Mobile-Tombigbee, Lower Mississippi-Big Black, Pearl, and Pascagoula.

[MSE, mean square error in $\log_{10}$ cubic feet per second]

Site number (fig. 2)	Station number	Station skew	MSE	Site number (fig. 2)	Station number	Station skew	MSE	Site number (fig. 2)	Station number	Station skew	MSE
8	02378500	0.063	0.019	96	02473480	0.07	0.021	202	02487500	0.073	0.022
9	02429900	–0.632	0.310	98	02473500	–0.023	0.003	211	02487900	0.219	0.087
11	02429980	0.978	1.110	100	02473850	–0.388	0.098	220	02488700	0.076	0.023
13	02430085	0.121	0.039	101	02474500	0.316	0.153	221	02489000	0.457	0.283
16	02430615	–0.395	0.102	103	02474600	–0.099	0.001	222	02489030	0.229	0.093
18	02430880	0.425	0.250	104	02474650	0.001	0.006	228	02490500	–0.771	0.484
38	02436500	0.847	0.851	107	02475000	0.178	0.064	232	02491500	–0.598	0.239
41	02437300	–0.742	0.444	108	02475050	–0.282	0.043	235	02492360	0.229	0.114
42	02437550	0.098	0.030	112	02475500	0.096	0.029	236	03592718	–0.873	0.636
44	02438000	–0.35	0.075	117	02477000	–0.005	0.005	345	07289265	–1.41	1.781
45	02439400	–0.559	0.234	118	02477050	–0.794	0.516	347	07289350	–0.123	0.002
51	02440400	0.284	0.129	123	02477330	0.041	0.014	358	07289600	–0.287	0.045
53	02440600	0.182	0.066	126	02477990	0.2	0.076	364	07290000	0.074	0.022
54	02440800	–0.611	0.287	128	02478500	0.573	0.420	365	07290005	–0.134	0.003
55	02441000	–0.729	0.427	130	02479000	0.119	0.038	372	07290650	0.051	0.016
56	02441220	–0.253	0.032	134	02479130	0.719	0.631	373	07290690	–0.116	0.002
57	02441300	0.188	0.069	137	02479155	0.233	0.095	374	07290830	0.146	0.049
60	02443000	–0.774	0.488	146	02479300	–0.138	0.004	379	07291000	–0.655	0.336
61	02443500	–0.538	0.214	151	02480500	0.237	0.098	380	07291250	–0.245	0.029
66	02446500	–0.434	0.129	153	02481000	–0.356	0.079	383	07292500	–0.553	0.228
72	02448000	0.237	0.098	155	02481130	–0.154	0.006	386	07295000	–0.61	0.286
73	02448620	0.807	0.779	156	02481400	0.35	0.181	389	07373550	–0.107	0.001
77	02467500	0.441	0.267	160	02481510	0.108	0.034	396	07375307	–0.538	0.214
81	02469800	0.693	0.590	168	02482000	–0.111	0.001	400	07375500	–0.174	0.010
82	02471001	0.113	0.035	175	02483000	–0.073	0.000	401	07375800	–0.177	0.010
83	02471100	–1.001	0.857	176	02483500	0.952	1.055	402	07376000	–0.132	0.003
85	02471500	–0.649	0.329	177	02483890	0.121	0.039	404	07376500	–0.264	0.036
86	02472000	0.175	0.063	178	02484000	0.156	0.054	412	07377000	–0.454	0.143
89	02472500	0.561	0.405	180	02484600	0.552	0.394	413	07377210	–0.735	0.435
93	02473000	0.387	0.214	181	02484750	0.008	0.007	417	07377500	–0.937	0.742
95	02473460	0.001	0.000	182	02484760	0.464	0.215	421	07378500	–0.672	0.356

The flood magnitudes obtained from station LP3 frequency distributions, which were obtained through analysis of streamflow records at gaged sites, were related to basin characteristics (table 1) by using generalized least squares (GLS) multiple-linear regression analysis. Ordinary least squares was used in the initial regression analysis to evaluate the statistical significance of each basin characteristic (Wagner and others, 2016). The USGS weighted-multiple-linear regression program (WREG) version 1.05 (https://water.usgs.gov/software/WREG/) was used to complete the GLS regression analysis (Eng and others, 2009; U.S. Geological Survey, 2013b), a method that weights streamgages in the regression according to differences in streamflow record length, the variance of streamflow measurements in the record, and spatial cross correlations of concurrent flows among streamgages. The estimates of streamflow at gaged sites were weighted with the regional estimates of streamflow to compute a final set of AEP flows.

Other regression diagnostics were reviewed to identify streamgages that have high leverage and (or) high influence metrics. The leverage metric is used to compare the values of independent variables at one streamgage to the values of the same variables at all other streamgages. The influence metric indicates if data from a streamgage have a high influence on the estimated regression metric values (Eng and others, 2009). A streamgage may have a high leverage metric, indicating that its independent variables are substantially different from those at all other streamgages, but the same streamgage may not have a high influence on the regression metrics. Conversely, a streamgage with a high influence may not have a high leverage metric. Measurement or transposing errors in reported values of some independent variables sometimes produce high leverage or influence metrics. Data from streamgages with high influence or leverage were given additional review to determine if such errors had been made or if the data should be excluded for other reasons.

Three flood regions for Mississippi (fig. 1) were delineated based upon review of residual plots, previous reports, eight-digit hydrologic unit code maps, geologic maps, and physiographic maps. A GLS analysis was then performed to determine the importance of the explanatory variables in each region. For flood regions A, B, and C, drainage area was the significant variable in the RREs. Region B is similar to the East region from the 1991 study but differs due to the fact that the Pearl River Basin was excluded. Regions A and C combine to cover the area represented by the West region in the 1991 study but also include the Pearl River Basin.

Application of Methods

When applying regression equations, users are advised not to interpret the empirical results as exact. Regression equations are statistical models that must be interpreted and applied within the limits of the data and with the understanding that the results are best-fit estimates with an associated variance. Flood-frequency estimation methods for stream AEP flows in Mississippi, excluding sites located in the MAP, differed among gaged sites, ungaged locations on gaged streams, and ungaged streams.

Flood-Frequency Estimates at Gaged Sites in Mississippi

The estimates of the AEP flows at gaged sites were computed by using EMA and MGB to test for PILFs. Annual exceedance probability flows at gaging stations should be determined by weighting the station skew coefficient with either the generalized skew coefficient from Bulletin 17B (Interagency Advisory Committee on Water Data, 1982) or the study-specific skew coefficient in the five previously identified subregional watersheds. The accuracy of AEP flows at gaged sites can be further improved by weighting those flows with RRE-predicted AEP flows. The variance of prediction is a function of the regression equations and the independent variables used to develop the flow estimate. If the estimated AEP flows at gaged sites and RRE-predicted AEP flows are assumed to be independent and are weighted in inverse proportion to the associated variances, the variance of the weighted estimate will be less than the variance of either of the independent estimates. Once the variances have been computed, the two independent flow estimates can be weighted by using the following equation:

$$log_{10}Q_{P(g)w} = \frac{V_{p,P(g)r}*log_{10}Q_{P(g)s} + V_{p,P(g)s}*log_{10}Q_{P(g)r}}{V_{p,P(g)s} + V_{p,P(g)r}}, \quad (2)$$

where

$Q_{P(g)w}$ is the weighted-flow estimate for the selected recurrence interval, in cubic feet per second;

$V_{p,P(g)r}$ is the variance of prediction at the gaged station derived from the applicable RRE AEP, in log units;

$Q_{P(g)s}$ is the estimate of peak flow at the gaged station from the EMA analysis for the selected AEP, in cubic feet per second;

$V_{p,P(g)s}$ is the variance of prediction at the gaged station from the EMA analysis for the selected AEP, in log units; and

$Q_{P(g)r}$ is the estimate of peak flow at the gaged station from the RRE for the selected AEP, in cubic feet per second.

The weighted-flow estimates for the gaged stations in Mississippi are provided in the data release associated with this report (Anderson, 2018).

Flood-Frequency Estimates at Ungaged Locations on Gaged Streams

The AEP flows for a gaged site on a stream can be transferred to an ungaged location by using the area-weighting method; however, equation 3 does not weight the ungaged AEP flows with the RRE AEP flows for the ungaged site. This procedure is suggested if the drainage area at an ungaged site is within 50 percent of the drainage at the gaged site (drainage area ratio is more than 0.5 or less than 1.5) (Ries and Dillow, 2006). The AEP flows at a gaged site can be transferred to an ungaged site by using the following equation:

$$Q_{P(u)} = \left(\frac{A_{(u)}}{A_{(g)}} \right)^{b} Q_{P(g)w} \ , \qquad (3)$$

where

- $Q_{P(u)}$ is the estimate of flood flow for the selected P-percent AEP for the ungaged site, u, in cubic feet per second;
- $A_{(u)}$ is the drainage area of the ungaged site, in square miles;
- $A_{(g)}$ is the drainage area of the upstream or downstream streamgage, in square miles;
- $Q_{P(g)w}$ is the weighted estimate of flood flow for the selected P-percent AEP for the upstream or downstream streamgage, in cubic feet per second; and
- b is the exponent of drainage area from the appropriate RRE.

The AEP flows at the gaged site can be transferred and weighted with RRE AEP flows for the ungaged site by using the following equation:

$$Q_{P(u)w} = \left(\frac{2|\Delta A|}{A_g} \right) Q_{P(u)r} + \left(1 - \frac{2|\Delta A|}{A_g} \right) Q_{P(u)} \ , \qquad (4)$$

where

- $Q_{P(u)w}$ is the weighted-flow estimate at the ungaged site after transferring the weighted peak flow from the gaged site, in cubic feet per second;
- $Q_{P(u)r}$ is the RRE flood estimate at the ungaged site for the selected recurrence interval, in cubic feet per second;
- A_g is the drainage area of the gaged site, in square miles;
- $|\Delta A|$ is the absolute difference in drainage areas between the ungaged site and the gaged site, in square miles; and
- $Q_{P(u)}$ is the estimate of flood flow for the selected P-percent AEP for the ungaged site, u, in cubic feet per second.

If the drainage area at an ungaged site differs by more than 50 percent from that of the gaged site, the RRE estimates should be used. If an ungaged site is between two gaged sites on the same stream, the suggested limitations should be applied to determine which gaged site should be used to estimate weighted AEP flows at the ungaged site. In general, the site with the smallest drainage area ratio and longest period of record should be used (Sauer, 1974).

Flood-Frequency Estimates at Locations on Ungaged Streams

For locations on ungaged streams, the flood region should be determined by using figure 1 or StreamStats (http://water.usgs.gov/osw/streamstats/). The RREs for the flood region should then be used to compute flows for the desired range of AEPs (table 4). The standard error of prediction (SEP) is a measure of how well the regression relation estimates flood magnitudes when applied to ungaged basins. The given SEPs for the range of AEPs are listed in table 4.

Accuracy and Limitations of Regional Regression Equations

The RREs only apply to rural streams that can be described using the basin characteristics listed in table 1 and with drainage areas within the ranges of drainage areas listed in table 5 that were used to develop the equations. These methods should not be used for sites in a watershed that are affected substantially by regulation from impoundments, channelization, levees, or other man-made structures. The methods also should not be used for sites on streams in urban areas (impervious area greater than 10 percent) unless the effects of urbanization are insignificant. The methods do not apply where flooding is influenced by extreme ocean storm surge or tidal events. Reliability of the regression relations for values outside of the flood region limits are unknown. Because the RREs were computed by using basin characteristics that were derived from GIS layers in StreamStats (http://water.usgs.gov/osw/streamstats/), RREs should only be computed by using basin characteristics from StreamStats.

The accuracy of a flood-frequency relation traditionally has been expressed in two ways—as the mean standard error of the model (SEM) or as the mean SEP. The SEM is a measure of how well the regression equation fits the data used to derive the relation and represents the standard deviation of the differences between station data and the corresponding values computed from the regression equation. The SEP is a measure of how well the regression relation estimates flood magnitudes when applied to ungaged basins. The SEP is the square root of the mean square error of prediction (MSEp). The MSEp is the sum of two components—the MSE resulting from the model and the sampling MSE, which results from estimating the model parameters from samples of the population.

Table 4. Final regional regression equations for estimating annual exceedance probability flows and generalized least squares model diagnostics for unregulated streams in Mississippi.

[MSE, mean square error in $\log_{10}$ cubic feet per second; AVP, Average Variance of Prediction in $\log_{10}$ cubic feet per second; SEP, standard error of prediction; Pseudo-R^2, Pseudo coefficient of determination; Q, flood quantiles, in cubic feet per second; A, contributing drainage area in square miles]

Regional regression equation	MSE (log ft³/s)	AVP (log ft³/s)²	SEP (percent)	Pseudo-R^2 (percent)
Flood region A (27 streamgages)				
$Q_{50\%}=661(A)^{0.552}$	0.018	0.016	30	91
$Q_{20\%}=955(A)^{0.564}$	0.015	0.015	28	92
$Q_{10\%}=1,148(A)^{0.571}$	0.016	0.015	28	92
$Q_{4\%}=1,380(A)^{0.580}$	0.017	0.016	29	92
$Q_{2\%}=1,514(A)^{0.587}$	0.018	0.016	30	92
$Q_{1\%}=1,660(A)^{0.593}$	0.020	0.017	31	92
$Q_{0.5\%}=1,820(A)^{0.599}$	0.022	0.018	32	92
$Q_{0.2\%}=1,995(A)^{0.606}$	0.025	0.019	33	91
Flood region B (134 streamgages)				
$Q_{50\%}=269(A)^{0.601}$	0.030	0.027	39	94
$Q_{20\%}=455(A)^{0.608}$	0.023	0.019	33	96
$Q_{10\%}=593(A)^{0.612}$	0.022	0.017	31	97
$Q_{4\%}=776(A)^{0.614}$	0.024	0.018	31	97
$Q_{2\%}=912(A)^{0.617}$	0.027	0.021	34	96
$Q_{1\%}=1,052(A)^{0.619}$	0.032	0.022	35	96
$Q_{0.5\%}=1,197(A)^{0.621}$	0.037	0.026	38	95
$Q_{0.2\%}=1,390(A)^{0.624}$	0.046	0.032	43	94
Flood region C (120 streamgages)				
$Q_{50\%}=337(A)^{0.607}$	0.036	0.032	43	94
$Q_{20\%}=512(A)^{0.628}$	0.027	0.023	36	96
$Q_{10\%}=632(A)^{0.638}$	0.027	0.022	35	96
$Q_{4\%}=787(A)^{0.648}$	0.030	0.023	36	96
$Q_{2\%}=908(A)^{0.654}$	0.034	0.027	39	96
$Q_{1\%}=1,026(A)^{0.660}$	0.039	0.029	41	95
$Q_{0.5\%}=1,148(A)^{0.664}$	0.045	0.033	43	95
$Q_{0.2\%}=1,309(A)^{0.670}$	0.053	0.037	46	94

Table 5. Number of streamgages used in regional regression analysis and ranges of drainage area in each flood region of Mississippi.

[mi^2, square miles]

Region (see fig. 1)	Number of streamgages	Range of drainage area
A	27	1.41 to 612 mi^2
B	134	0.15 to 1,750 mi^2
C	120	0.05 to 1,010 mi^2

Summary and Conclusions

Flood-frequency estimates for rural Mississippi streams had not been updated in over 25 years (Landers and Wilson, 1991), and since that time, improvements in estimation techniques and additional streamflow data were expected to improve the accuracy these estimates. Thus, the U.S. Geological Survey performed analyses to estimate annual exceedance probability (AEP) flows at gaged streams and regional regression equations (RREs) to estimate AEPs at ungaged locations. Twenty-one basin characteristics were analyzed, and those that were significant were used in the RREs. In addition, the RREs incorporated a newly developed study-specific skew coefficient (−0.08) at 93 streamgages located in five subregional watersheds of Mississippi: the Middle Tennessee-Elk, Mobile-Tombigbee, Lower Mississippi-Big Black, Pearl, and Pascagoula. RREs established for the other subregional watersheds in the State utilized the generalized skew coefficient published in Bulletin 17B (Interagency Advisory Committee on Water Data, 1982).

Using the regression residuals and basin characteristics derived from geographic information system data, this study defined four flood regions (A, B, C, and the Mississippi Alluvial Plain) for the State of Mississippi, which differed from that of Landers and Wilson (1991) who identified three flood regions (East, West, and the Mississippi Alluvial Plain) in the previous flood-frequency study. The West region from the 1991 study was separated into two regions in the current study—flood region A and flood region C that includes the Pearl River Basin.

Extreme flow events have the potential for devastating impacts to the economy, infrastructure, and the landscape. Keeping flood-frequency analysis current and updated will offer water-resource managers the information needed to make educated decisions regarding flood-response planning. By broadening the regional approach for the development of flood-frequency RREs across State lines, flood-frequency estimates have the potential to be more accurate and applicable to a larger study area. Finally, providing the analysis results on a publicly accessible web interface, StreamStats, allows a user to select a site of interest and obtain the AEP flows for that site at any time.

References Cited

Anderson, B.T., 2018, Regions and Tables for Mississippi Flood Frequency, Data through 2013: U.S. Geological Survey data release, https://doi.org/10.5066/F7ZP45B8.

Cohn, T.A., England, J.F., Berenbrock, C.E., Mason, R.R., Stedinger, J.R., and Lamontagne, J.R., 2013, A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series: Water Resources Research, v. 49, p. 5047–5058, accessed October 13, 2014, at http://dx.doi.org/10.1002/wrcr.20392.

Cohn, T.A., Lane, W.L., and Baier, W.G., 1997, An algorithm for computing moments-based flood quantile estimates when historical flood information is available: Water Resources Research, v. 33, no. 9, p. 2089–2096, https://doi.org/10.1029/97WR01640.

Colson, B.E., and Hudson, J.W., 1976, Flood frequency of Mississippi streams: Mississippi State Highway Department, RO-76-014-PR, 34 p.

Eng, Ken, Chen, Yin-Yu, and Kiang, J.E., 2009, User’s guide to the weighted-multiple-linear regression program (WREG version 1.0): U.S. Geological Survey Techniques and Methods, book 4, chap. A8, 21 p. [Also available at http://pubs.usgs.gov/tm/tm4a8/.]

Feaster, T.D., Gotvald, A.J., and Weaver, J.C., 2009, Magnitude and frequency of rural floods in the Southeastern United States, 2006—Volume 3, South Carolina: U.S. Geological Survey Scientific Investigations Report 2009–5156, 226 p.

Griffis, V.W., and Stedinger, J.R., 2007, Log-Pearson type 3 distribution and its application in flood frequency analysis. II. Parameter estimation methods: Journal of Hydrologic Engineering, v. 12, no. 5, p. 492–500, https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(492).

Gruber, A.M., and Stedinger, J.R., 2008, Models of LP3 regional skew, data selection, and Bayesian GLS regression, Paper 596, *in* Babcock, R., and Walton, R., eds., World Environmental and Water Resources Congress, Honolulu, Hawaii, May 12–16, 2008: American Society of Civil Engineers.

Hodgkins, G.A., Martin, G.R., 2003, Estimating the magnitude of peak flows for streams in Kentucky for selected recurrence intervals: U.S. Geological Survey Water-Resources Investigations Report 03–4180, 68 p.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., 2015, Completion of the 2011 National Land Cover Database for the conterminous United States—Representing a decade of land cover change information: Photogrammetric Engineering and Remote Sensing, v. 81, no. 5, p. 345–354.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood-flow frequency: Bulletin 17B, 183 p.

Landers, M.N., 1985, Floodflow frequency of streams in the alluvial plain of the lower Mississippi River in Mississippi, Arkansas, and Louisiana: U.S. Geological Survey Water-Resources Investigations Report 85–4150, 21 p.

Landers, M.N., and Wilson, K.V., Jr., 1991, Flood characteristics of Mississippi streams: U.S. Geological Survey Water-Resources Investigations Report 91–4037, 82 p.

Parrett, C., Veilleux, A., Stedinger, J.R., Barth, N.A., Knifong, D.L., and Ferris, J.C., 2011, Regional skew for California, and flood frequency for selected sites in the Sacramento–San Joaquin River Basin, based on data through water year 2006: U.S. Geological Survey Scientific Investigations Report 2010–5260, 94 p.

Ries, K.G., III, and Dillow, J.J.A., 2006, Magnitude and frequency of floods on nontidal streams in Delaware: U.S. Geological Survey Scientific Investigations Report 2006–5146, 59 p.

Sauer, V.B., 1974, Flood characteristics of Oklahoma streams techniques for calculating magnitude and frequency of floods in Oklahoma, with compilations of flood data through 1971: U.S. Geological Survey Water-Resources Investigations Report 73–52, 307 p.

U.S. Climate Data, 2017, Climate data for Jackson, Mississippi, accessed February 7, 2017, at https://www.usclimatedata.com/climate/mississippi/united-states/3194.

U.S. Geological Survey, 2007, Facing tomorrow's challenges—U.S. Geological Survey science in the decade 2007–2017: U.S. Geological Survey Circular 1309, 70 p.

U.S. Geological Survey, 2013a, Cooperative Water Program—Priority Activities for FY 14, accessed June 1, 2014, at http://water.usgs.gov/coop/about/CWP.science.priorities.pdf.

U.S. Geological Survey, 2013b, WREG, weighted-multiple-linear regression program, accessed June 1, 2014, at https://water.usgs.gov/software/WREG/.

U.S. Geological Survey, 2014, PeakFQ, accessed July 3, 2013, at https://water.usgs.gov/software/PeakFQ/.

U.S. Geological Survey, 2017a, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed July 3, 2013, at http://dx.doi.org/10.5066/F7P55KJN. [Peak-flow data directly accessible at https://nwis.waterdata.usgs.gov/usa/nwis/peak.]

U.S. Geological Survey, 2017b, Welcome to StreamStats, accessed July 3, 2013, at https://water.usgs.gov/osw/streamstats/.

Wagner, D.M., Krieger, J.D., and Veilleux, A.G., 2016, Methods for estimating annual exceedance probability discharges for streams in Arkansas, based on data through water year 2013: U.S. Geological Survey Scientific Investigations Report 2016–5081, 136 p., accessed July 30, 2014, at https://doi.org/10.3133/sir20165081.

Wilson, K.V., and Trotter, I.L., Jr., 1961, Floods in Mississippi, magnitude and frequency: U.S. Geological Survey Open-File Report, 326 p.

For more information about this publication, contact
Director, Lower Mississippi-Gulf Water Science Center
U.S. Geological Survey
640 Grassmere Park, Suite 100
Nashville, Tennessee 37211

For additional information, visit
https://www.usgs.gov/centers/lmg-water

Publishing support provided by
Lafayette Publishing Service Center

USGS

Anderson—**Flood Frequency of Rural Streams in Mississippi, 2013**—SIR 2018–5148

ISBN 978-1-4113-4264-4

9 781411 342644

ISSN 2328-031X (print)
ISSN 2328-0328 (online)
https://doi.org/10.3133/sir20185148
Supersedes USGS WRIR 91–4037

Printed on recycled paper