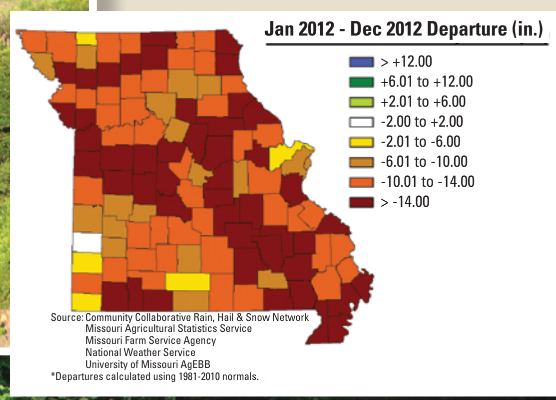
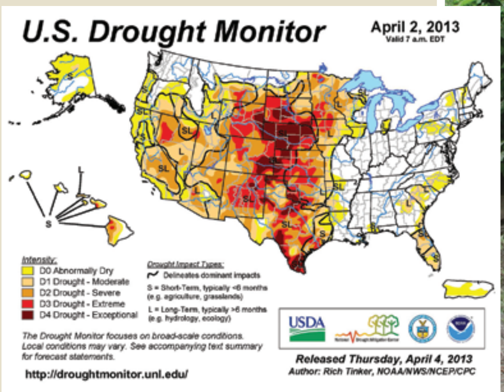
Prepared in cooperation with the Missouri Department of Natural Resources

Computed Statistics at Streamgages, and Methods for Estimating Low-Flow Frequency Statistics and Development of Regional Regression Equations for Estimating Low-Flow Frequency Statistics at Ungaged Locations in Missouri





Scientific Investigations Report 2013–5090

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

Cover. **Upper left photograph:** Big River near Desloge, Missouri, June 28, 2001.
Lower right photograph: Sebago Creek near Rock Hill, Missouri, May 19, 2003.
Upper right graphic: Rainfall departure from normal (1971–2000) in 2012 calendar year furnished by University of Missouri Agricultural Electronic Bulletin Board.
Lower left graphic: U.S. Drought Monitor for April 2, 2013, furnished by the National Drought Mitigation Center at the University of Nebraska-Lincoln, the U.S. Department of Agriculture, and the National Oceanic and Atmospheric Administration.

Computed Statistics at Streamgages, and Methods for Estimating Low-Flow Frequency Statistics and Development of Regional Regression Equations for Estimating Low-Flow Frequency Statistics at Ungaged Locations in Missouri

By Rodney E. Southard

Prepared in cooperation with the Missouri Department of Natural Resources

Scientific Investigations Report 2013–5090

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

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2013

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit http://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit http://www.usgs.gov/pubprod

To order this and other USGS information products, visit http://store.usgs.gov

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:
Southard, R.E., 2013, Computed statistics at streamgages, and methods for estimating low-flow frequency statistics and development of regional regression equations for estimating low-flow frequency statistics at ungaged locations in Missouri: U.S. Geological Survey Scientific Investigations Report 2013–5090, 28 p., http://pubs.usgs.gov/sir/2013/5090/.

ISBN 978 1 4113 3594 3

Contents

Figures

Tables

Note: The following tables are available in Excel format on CD at the back of the report and at *http://pubs.usgs.gov/sir/2013/5090/downloads/.*

Tables

Note: The following tables are available in Excel format on CD at the back of the report and at *http://pubs.usgs.gov/sir/2013/5090/downloads/.* Page numbers in this section refer to first mention in text.

Conversion Factors

Inch/Pound to SI

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
mile, nautical (nmi)	1.852	kilometer (km)
	Area	
square mile (mi^2)	259.0	hectare (ha)
mile per square mile (mi/mi^2)	0.621	kilometer per square kilometer (km/km^2)
square mile per mile (mi^2/mi)	1.609	square kilometer per kilometer (km^2/km)
	Flow rate	
foot per second (ft/s)	0.3048	meter per second (m/s)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Vertical coordinate information is referenced to North American Vertical Datum of 1988 (NAVD 88).

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

Water year is defined as the 12-month period from October 1 through September 30 of the following year.

Abbreviations

Adj-R^2	Adjusted coefficient of determination
AMLE	Adjusted maximum-likelihood estimation
BASINS	Better Assessment Science Integrating point & Non-point Sources
CCC	Criterion continuous concentration
CMC	Criterion maximum concentration
CROPSNLCD01	Cultivated crops from National Land Cover Dataset 2001
DAR	Drainage-area ratio
DEM	Digital elevation model
DRNAREA	GIS-determined drainage area
e	Base of the natural logarithm, approximately equal to 2.7183
EPA	U.S. Environmental Protection Agency
GIS	Geographic Information System
HUC	Hydrologic unit code
KSATSSURG	Average soil permeability or saturated hydraulic conductivity of soil
LFPLENGTH	Longest flow length
LOWESS	Locally-Weighted Scatterplot Smoothing
M1D10Y	Annual 1-day mean low flow for a recurrence interval of 10 years
M2D10Y	Annual 2-day mean low flow for a recurrence interval of 10 years
M3D10Y	Annual 3-day mean low flow for a recurrence interval of 10 years
M7D10Y	Annual 7-day mean low flow for a recurrence interval of 10 years
M10D10Y	Annual 10-day mean low flow for a recurrence interval of 10 years
M30D10Y	Annual 30-day mean low flow for a recurrence interval of 10 years
M60D10Y	Annual 60-day mean low flow for a recurrence interval of 10 years
Mallow's *Cp*	Measure of the total squared error for a subset model containing n independent variables
MLE	Maximum-likelihood estimation
NAWQA	National Water-Quality Assessment
NHD	National hydrography dataset
NRCS	Natural Resources Conservation Service
NWIS	National Water Inventory System
OLS	Ordinary-least-squares regression
POR	Period of record
PRESS	Predicted REsidual Sum of Squares
R^2	Coefficient of determination

RMSE	Root mean square error, also referred to as SEE
SEE	Average standard error of estimate, also referred to as RMSE
SHAPE	Basin shape
SOILASSURGO	Hydrologic soil type A
SOILBSSURGO	Hydrologic soil type B
SOILCSSURGO	Hydrologic soil type C
SSURGO	NRCS Soil Survey Geographic database
StreamStats	USGS Web-based GIS tool (*http://water.usgs.gov/osw/streamstats/index.html*)
STREAM_VAR	Streamflow-variability index
SWSTAT	USGS Surface-Water Statistics computer program
USDA	U.S. Department of Agriculture
USGS	U.S. Geological Survey
1B3	Biological annual 1-day 3-year criterion continuous concentration
4B3	Biological annual 4-day 3-year criterion maximum concentration

Computed Statistics at Streamgages, and Methods for Estimating Low-Flow Frequency Statistics and Development of Regional Regression Equations for Estimating Low-Flow Frequency Statistics at Ungaged Locations in Missouri

By Rodney E. Southard

Abstract

The weather and precipitation patterns in Missouri vary considerably from year to year. In 2008, the statewide average rainfall was 57.34 inches and in 2012, the statewide average rainfall was 30.64 inches. This variability in precipitation and resulting streamflow in Missouri underlies the necessity for water managers and users to have reliable streamflow statistics and a means to compute select statistics at ungaged locations for a better understanding of water availability. Knowledge of surface-water availability is dependent on the streamflow data that have been collected and analyzed by the U.S. Geological Survey for more than 100 years at approximately 350 streamgages throughout Missouri. The U.S. Geological Survey, in cooperation with the Missouri Department of Natural Resources, computed streamflow statistics at streamgages through the 2010 water year, defined periods of drought and defined methods to estimate streamflow statistics at ungaged locations, and developed regional regression equations to compute selected streamflow statistics at ungaged locations.

Streamflow statistics and flow durations were computed for 532 streamgages in Missouri and in neighboring States of Missouri. For streamgages with more than 10 years of record, Kendall's *tau* was computed to evaluate for trends in streamflow data. If trends were detected, the variable length method was used to define the period of no trend. Water years were removed from the dataset from the beginning of the record for a streamgage until no trend was detected. Low-flow frequency statistics were then computed for the entire period of record and for the period of no trend if 10 or more years of record were available for each analysis.

Three methods are presented for computing selected streamflow statistics at ungaged locations. The first method uses power curve equations developed for 28 selected streams in Missouri and neighboring States that have multiple streamgages on the same streams. Statistical estimates on one of these streams can be calculated at an ungaged location that has a drainage area that is between 40 percent of the drainage area of the farthest upstream streamgage and within 150 percent of the drainage area of the farthest downstream streamgage along the stream of interest. The second method may be used on any stream with a streamgage that has operated for 10 years or longer and for which anthropogenic effects have not changed the low-flow characteristics at the ungaged location since collection of the streamflow data. A ratio of drainage area of the stream at the ungaged location to the drainage area of the stream at the streamgage was computed to estimate the statistic at the ungaged location. The range of applicability is between 40- and 150-percent of the drainage area of the streamgage, and the ungaged location must be located on the same stream as the streamgage. The third method uses regional regression equations to estimate selected low-flow frequency statistics for unregulated streams in Missouri. This report presents regression equations to estimate frequency statistics for the 10-year recurrence interval and for the N-day durations of 1, 2, 3, 7, 10, 30, and 60 days.

Basin and climatic characteristics were computed using geographic information system software and digital geospatial data. A total of 35 characteristics were computed for use in preliminary statewide and regional regression analyses based on existing digital geospatial data and previous studies. Spatial analyses for geographical bias in the predictive accuracy of the regional regression equations defined three low-flow regions with the State representing the three major physiographic provinces in Missouri. Region 1 includes the Central Lowlands, Region 2 includes the Ozark Plateaus, and Region 3 includes the Mississippi Alluvial Plain. A total of 207 streamgages were used in the regression analyses for the regional equations. Of the 207 U.S. Geological Survey streamgages, 77 were located in Region 1, 120 were located in Region 2, and 10 were located in Region 3. Streamgages located outside of Missouri were selected to extend the range of data used for the independent variables in the regression

analyses. Streamgages included in the regression analyses had 10 or more years of record and were considered to be affected minimally by anthropogenic activities or trends. Regional regression analyses identified three characteristics as statistically significant for the development of regional equations. For Region 1, drainage area, longest flow path, and streamflow-variability index were statistically significant. The range in the standard error of estimate for Region 1 is 79.6 to 94.2 percent. For Region 2, drainage area and streamflow variability index were statistically significant, and the range in the standard error of estimate is 48.2 to 72.1 percent. For Region 3, drainage area and streamflow-variability index also were statistically significant with a range in the standard error of estimate of 48.1 to 96.2 percent.

Limitations on the use of estimating low-flow frequency statistics at ungaged locations are dependent on the method used. The first method outlined for use in Missouri, power curve equations, were developed to estimate the selected statistics for ungaged locations on 28 selected streams with multiple streamgages located on the same stream. A second method uses a drainage-area ratio to compute statistics at an ungaged location using data from a single streamgage on the same stream with 10 or more years of record. Ungaged locations on these streams may use the ratio of the drainage area at an ungaged location to the drainage area at a streamgage location to scale the selected statistic value from the streamgage location to the ungaged location. This method can be used if the drainage area of the ungaged location is within 40 to 150 percent of the streamgage drainage area. The third method is the use of the regional regression equations. The limits for the use of these equations are based on the ranges of the characteristics used as independent variables and that streams must be affected minimally by anthropogenic activities.

Introduction

Missouri is located in the Midwest region of the conterminous United States. In 2008, the State of Missouri had the largest above normal departure in precipitation for any state in the United States, averaging 16.58 inches above normal with an average rainfall of 57.34 inches (National Climatic Data Center, 2013). The rainfall in 2008 was the wettest calendar year on record since 1895 for Missouri while temperature ranked the 19th coldest. Flooding was widespread in the Midwest with 147 peak-of-record streamflows recorded at U.S. Geological Survey (USGS) streamgages (Holmes and others, 2010). Conversely, in 2012, the Midwest experienced one of the driest and hottest years on record. On July 23, 2012, Missouri Governor Jay Nixon signed Executive Order 12-07 declaring a state of emergency because of prolonged heat and drought (Office of Missouri Governor, 2012). Rainfall deficits continued throughout the remainder of 2012 with the National Climatic Data Center (2013) reporting a precipitation deficit of 10.12 inches for Missouri from January to December 2012 with an average rainfall of 30.64 inches. Thus, 2012 ranks as the 7th driest 12-month period since 1895 and the warmest 12-month period since 1895 (National Climatic Data Center, 2013).

These extremes in precipitation underlie the necessity for Federal, State, local, and private entities to have a good understanding of water availability by having access to reliable streamflow statistics and a means to compute selected streamflow statistics at ungaged locations. The USGS, in cooperation with the Missouri Department of Natural Resources, defined periods of historically persistent drought, computed streamflow statistics at streamgages, and developed relations to estimate selected streamflow statistics for the State of Missouri. To place the 2012 drought in Missouri in historical context, the annual mean streamflows for 2012 for selected streamgages were compared to historical drought periods. This information and methodology to estimate selected streamflow statistics will provide a consistent means to help manage the State's water resources by water users for water supply and water quality.

The USGS has collected daily streamflow data at approximately 350 streamgages throughout Missouri. Analyses of the streamgage data provide users with reliable duration and frequency estimates of streamflow at that location; however, it is not possible to operate a streamgage at every location where streamflow statistics are needed, so a method is needed to estimate selected streamflow statistics at ungaged locations. Equations were developed by the USGS using statistically significant basin characteristics to estimate the streamflow characteristics at ungaged stream locations. The basin characteristics are used in the equations as independent variables and were derived from digital geospatial datasets using geographic information system (GIS) software to measure the value of the variables. This methodology allows for the automated computation of the characteristics to estimate the selected statistic at an ungaged location. The data and techniques implemented in this study were developed for future use in a web-based GIS tool named Streamstats to display streamgage data and automate the computation of selected streamflow statistics.

Purpose and Scope

The purpose of the report is to present streamflow statistics computed at USGS streamgages in and adjacent to Missouri, and to describe the development and application of regression equations for estimating the magnitude and frequency of the annual N-day, 10-year low-flow statistic. The N-day durations selected for the regression equations are 1-, 2-, 7-, 10-, 30-, and 60-day periods.

The report also describes the site selection, statistical analyses of streamgage data, the development of two methods to estimate streamflow statistics at ungaged locations on streams where streamgages exist, and the development of regional regression equations for selected N-day durations for gaged or ungaged streams. The report presents the limitations of the methods that can be used to estimate low-flow frequency statistics at ungaged locations.

Previous Studies

This is the fifth in a series of reports that describe low-flow characteristics for Missouri streams. The first report (Skelton, 1966) contained information on low-flow frequency estimates for 262 continuous-record and partial-record streamgages, and flow duration for 63 selected continuous-record streamgages. Streamflow data were analyzed through the 1964 water year. A water year is defined as the 12-month period from October 1 through September 30 of the following year. Thus, the 1964 water year ends on September 30, 1964. Skelton (1966) only presented low-flow information for streamgages and no attempt was made to develop a methodology to estimate low-flow statistics at ungaged locations. The second report (Skelton, 1970) contained information on the results of hydrograph analyses of base-flow recessions and a statistical study of seasonal low-flow information for Missouri streams. Base-flow recession characteristics for 116 streamgages and seasonal low-flow frequency analyses (May through October) for 215 streamgages are contained in the report. Skelton (1970) presented a graphical methodology for estimating low-flow characteristics at an ungaged site. Streamflow measurements at the ungaged location are collected, preferably during several years, and plotted on graph paper with concurrent flows at a nearby streamgage. The relation defined by the points plotted can be used to estimate the statistic for the ungaged location using the computed statistic at the streamgage. The third report (Skelton, 1974) contained regression equations for estimating selected statistics for the Ozarks Plateaus physiographic province (Fenneman, 1938). Regression equations were developed for the annual 7-day mean low-flow for recurrence intervals of 2 years (M7D2Y), 10 years (M7D10Y), and 20 years (M7D20Y). The equations used the cross-sectional area of a stream at an ungaged location as the independent variable. The cross-section survey is to be completed during the time period of late August through November. The standard error for estimating the M7D10Y statistic was 60 percent for the regression equation. The fourth report (Skelton, 1976) contained low-flow frequency data for 460 streamgages and flow-duration data for 84 streamgages. The report contains a tabulation of low-flow frequency characteristics at approximately 260 additional streamgages not included in Skelton (1974) and frequency analyses of seepage-run data compiled through the 1974 water year.

Site Selection

Data used in the analyses for this report were collected for 532 active and inactive continuous-record streamgages located in Missouri and in the neighboring States of Iowa, Nebraska, Kansas, Oklahoma, and Arkansas (table 1 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_1.xlsx*; fig. 1). Location of these streamgages are shown in figure 1; 343 are in Missouri, 38 are in Iowa, 7 are in Nebraska, 70 are in Kansas, 18 are in Oklahoma, and 56 are in Arkansas. The list of 532 streamgages included streamgages located on unregulated and regulated streams and on streamflow originating from springs. Data from streamgages on regulated streams below reservoirs were included in the statistical analyses for completeness. Likewise, data from streamgages on spring streamflow were included in the statistical analyses. Streamflow data affected by reservoirs or springs were removed for use in regression analyses to prevent biasing the computations of the selected low-flow frequency statistics. For Missouri, all streamgages with more than 1 year of continuous record were included in the statistical data summaries. Streamgages were required to have at least 10 years of continuous record for frequency computations. Selection of streamgages in neighboring States was based on basins having similar basin and flow characteristics to basins in Missouri. Basic statistics such as minimum, maximum, and mean streamflows, monthly and annual flow durations, and trend analyses were computed for all streamgages, if there was sufficient record length. Low-flow frequency analyses were implemented on streamgages with at least 10 complete years of daily mean discharges. Streamgages, with streamflow trends detected, had statistics computed for the entire period of record and for the period of no trend. Streamgages from neighboring states were included to improve the representativeness of selected low-flow frequency statistics and basin characteristics indicated in Missouri border areas, and to provide better estimates of the error of the regression equations for ungaged sites near the State border. Daily mean discharge data collected through the 2010 water year (through September 30, 2010) were retrieved for the 532 streamgages from the USGS National Water Inventory System (NWIS; USGS, 2011) database for use in computing selected low-flow frequency statistics.

Streamflow data at a streamgage were considered for use in the development of the regression equations if data had a minimum of 10 years of record and if the data represented natural flow conditions or were affected minimally by anthropogenic activities. Anthropogenic activities that may affect low-flow frequency statistics include, but are not limited to, regulation, irrigation, diversions, storage, and urbanization. In general, all streamgages with data affected by upstream regulations or diversions during typical low-flow periods were deleted from the regression analyses data set. Streamgages were removed from consideration if they were affected by structures such as dams or diversions for irrigation. Streamgages also were removed if considerable storage from small impoundments and water-supply lakes were located in the basin upstream from the streamgage. Streamgages in urban areas generally were excluded from the analyses because of channel improvements, impervious area, and basin development. Streamgages were removed if backwater conditions existed at the site and low flow computations were considered

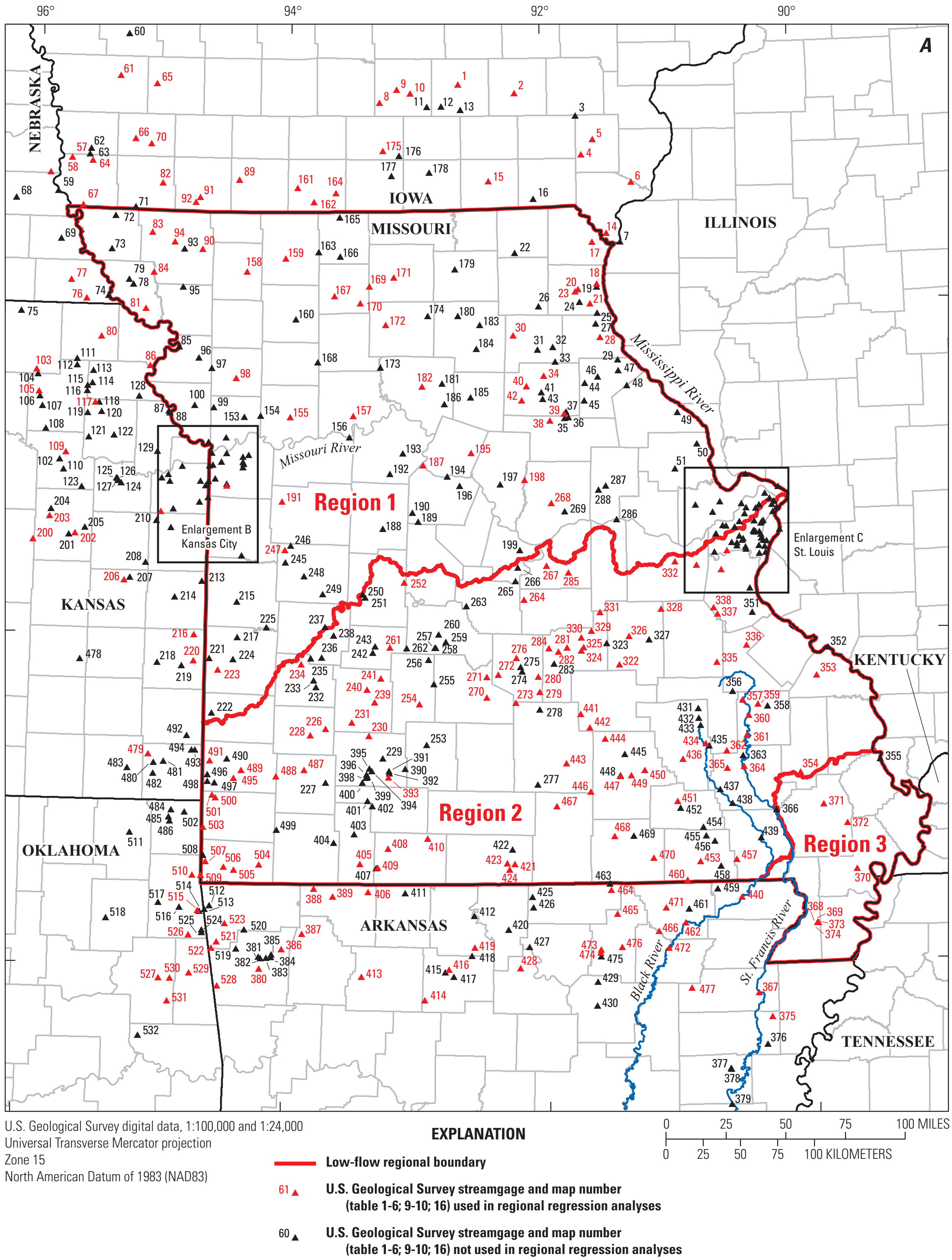
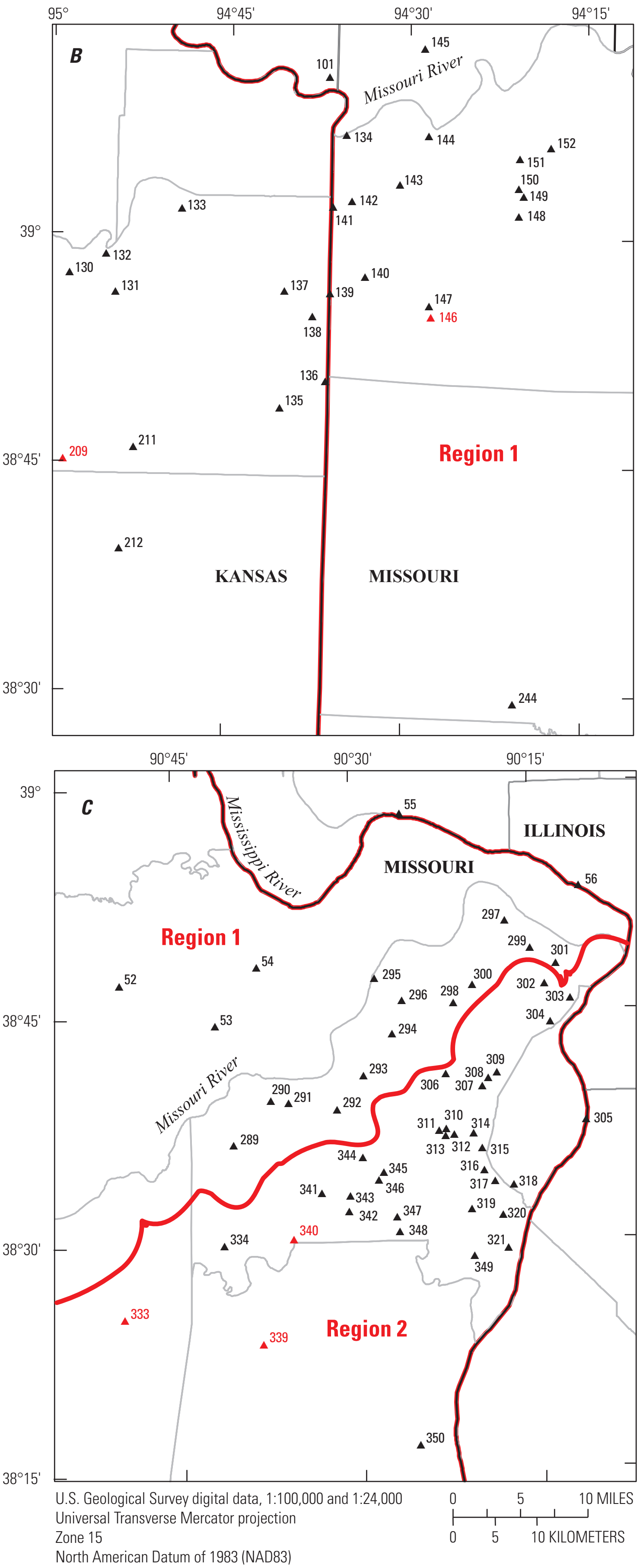


Figure 1. Location of streamgages in *A*, Missouri and neighboring States of Missouri; *B*, an enlargement of Kansas City, Missouri, area; *C*, an enlargement of St. Louis, Missouri, area.



EXPLANATION

146 ▲ U.S. Geological Survey streamgage and map number (table 1–11; 13; 16) used in regional regression analyses

101 ▲ U.S. Geological Survey streamgage and map number (table 1–11; 13; 16) not used in regional regression analyses

EXPLANATION

Low-flow regional boundary

333 ▲ U.S. Geological Survey streamgage and map number (table 1–11; 13; 16) used in regional regression analyses

52 ▲ U.S. Geological Survey streamgage and map number (table 1–11; 13; 16) not used in regional regression analyses

Figure 1. Location of streamgages in *A*, Missouri and neighboring States of Missouri; *B*, an enlargement of Kansas City, Missouri, area; *C,* an enlargement of St. Louis, Missouri, area.—Continued

poor in quality. Decisions on inclusion or exclusion of streamgage data were made using hydrologic judgment and the above criteria. Streamgages known to be affected minimally by regulation or diversion were left in the regression, evaluated for bias during the regression analyses, and removed if a bias was detected. The screening process resulted in 207 streamgages being used in the statewide regression analyses.

Historical Drought Periods

Missouri has a history of extreme floods and extended droughts. The USGS (Waite, 1991) documented historical floods and droughts for Missouri through water year 1988. Waite (1991) defined four drought periods in Missouri after evaluating data from 1930 to 1988: 1930–41, 1952–57, 1962–69, and 1975–82. An evaluation of annual streamflow data was done for this study with respect to droughts through water year 2010 using flow-duration statistics from the 207 streamgages (table 2 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_2.xlsx*). From table 2, 24 streamgages, with more than 40 years of continuous record each, were selected for evaluation of drought periods. The selected streamgages were distributed spatially throughout Missouri and neighboring States (table 3 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_3.xlsx*; fig. 2).

At each streamgage, annual mean streamflows were compared to the mean for the entire period of record (POR) retrieved from the NWIS database through the 2010 water year. Annual means above the POR mean were considered a wet year and annual means below the POR annual mean were considered a dry year. Continuous periods of record of 3 or more years with a mean streamflow that appeared to be substantially lower than the mean streamflow for the POR were designated qualitatively as periods of persistent drought in table 3. An example graph showing the annual mean streamflows by water year compared to the POR mean streamflow is shown in figure 3 for Fox River at Wayland (streamgage 05495000). Several years were observed where drought may have existed in 1- or 2-year time spans throughout the period of record for a streamgage such as water years 1934, 1988, and 2000 at streamgage 05495000; however, near or above average runoff existed before or immediately after the dry year. The definition for a drought period of 3 or more years reduced the number of droughts and provided a means to identify areas of the State that experienced persistent drought conditions (table 3).

By assessing the location of the streamgages spatially and by common periods of substantially below POR mean streamflow, the State of Missouri was broken into four areas (East Central, North, Southeast, and West Central) with similar drought periods (fig. 2). The East Central area appears to have the most drought periods with 3 years or more of consecutive below normal runoff: 1931–34, 1953–56, 1962–66, 1976–78, and 1999–2001. The North area had the fewest drought periods: 1922–25, 1936–41, and 1953–57. The mean streamflows of the drought periods were compared to the POR mean streamflow using the Wilcoxon Rank Sum test to test if the annual mean flow values for the drought periods were different statistically from the annual mean streamflows for the POR (table 3). The *p*-values from the Wilcoxon Rank Sum test are listed in table 3 for the drought periods for the 24 streamgages. *P*-values less than 0.05 indicate the drought period mean streamflow is different statistically than the POR mean streamflow. Mean streamflow for a drought period may be substantially below the period of record mean streamflow, but may not be statistically different in magnitude. However, adverse effects on wildlife and fish habitat, water supply, and agriculture can occur during these drought periods. The 1950's drought was the only drought period where all 24 streamgages had statistically different annual mean streamflows for the drought period compared to the POR annual mean streamflows. The 1930's drought had the next highest number of streamgages with statistically significant flows for 14 out of 19 streamgages, excluding the southeast area where no drought period was defined for the 1930's. The drought of the 1950's affected the entire State and the last drought to occur in east central Missouri was from 1999 to 2001.

To keep the 2012 streamflows in perspective, the annual mean for water year 2012 also is included in table 3 to compare 2012 annual mean streamflow values to historical drought mean streamflows and the POR mean streamflows at the 24 streamgages evaluated for drought periods. The 2012 water year was a dry year with below normal runoff for 21 of 22 streamgages that were in operation in water year 2012, with the exception of streamgage 06899500 (map number 167). As an example, streamgage 05514500 (map number 51) had an annual mean streamflow of 226 cubic feet per second (ft^3/s) in water year 2012 and the lowest computed drought period mean streamflow was 266 ft^3/s from 1953 to 1957 (table 3).

Computed Statistics at Streamgages

The availability of a stream to supply water needed for human use commonly is evaluated on the basis of statistical analyses of historical data from a streamgage or network of streamgages. Further analyses of the statistical data are often used to predict the probability of occurrence of low streamflow for an annual period. For this study, historical streamflow data through water year 2010 were included in the statistical analyses. For the basic statistics and the flow-duration analyses of historical streamflow, the water year period from October 1 through September 30 of the following year was used, and for N-day low-flow frequency analyses, the climatic year that begins on April 1 and ends on March 31 of the following year was used.

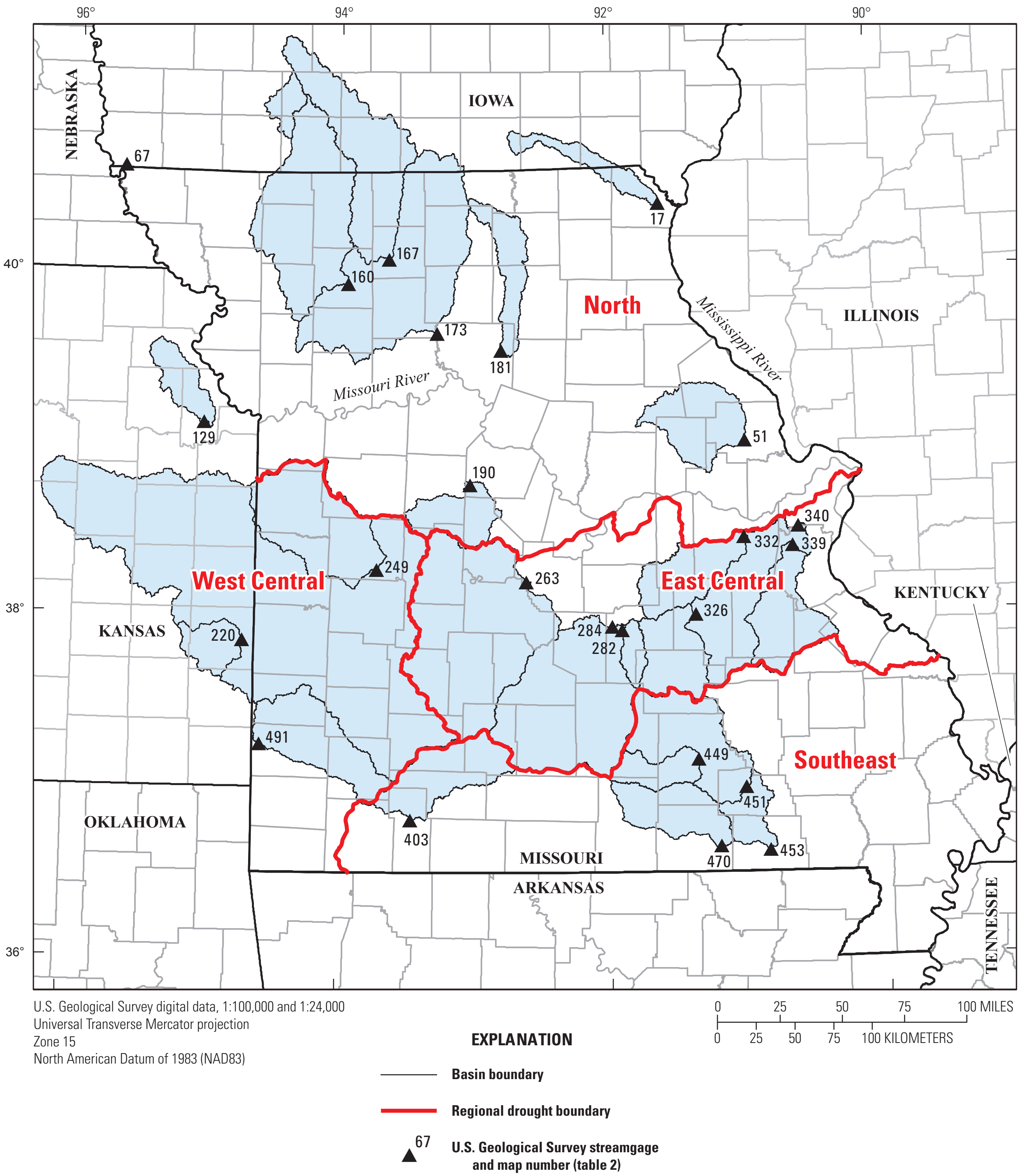


Figure 2. Drought areas and location of streamgages used to define drought periods.

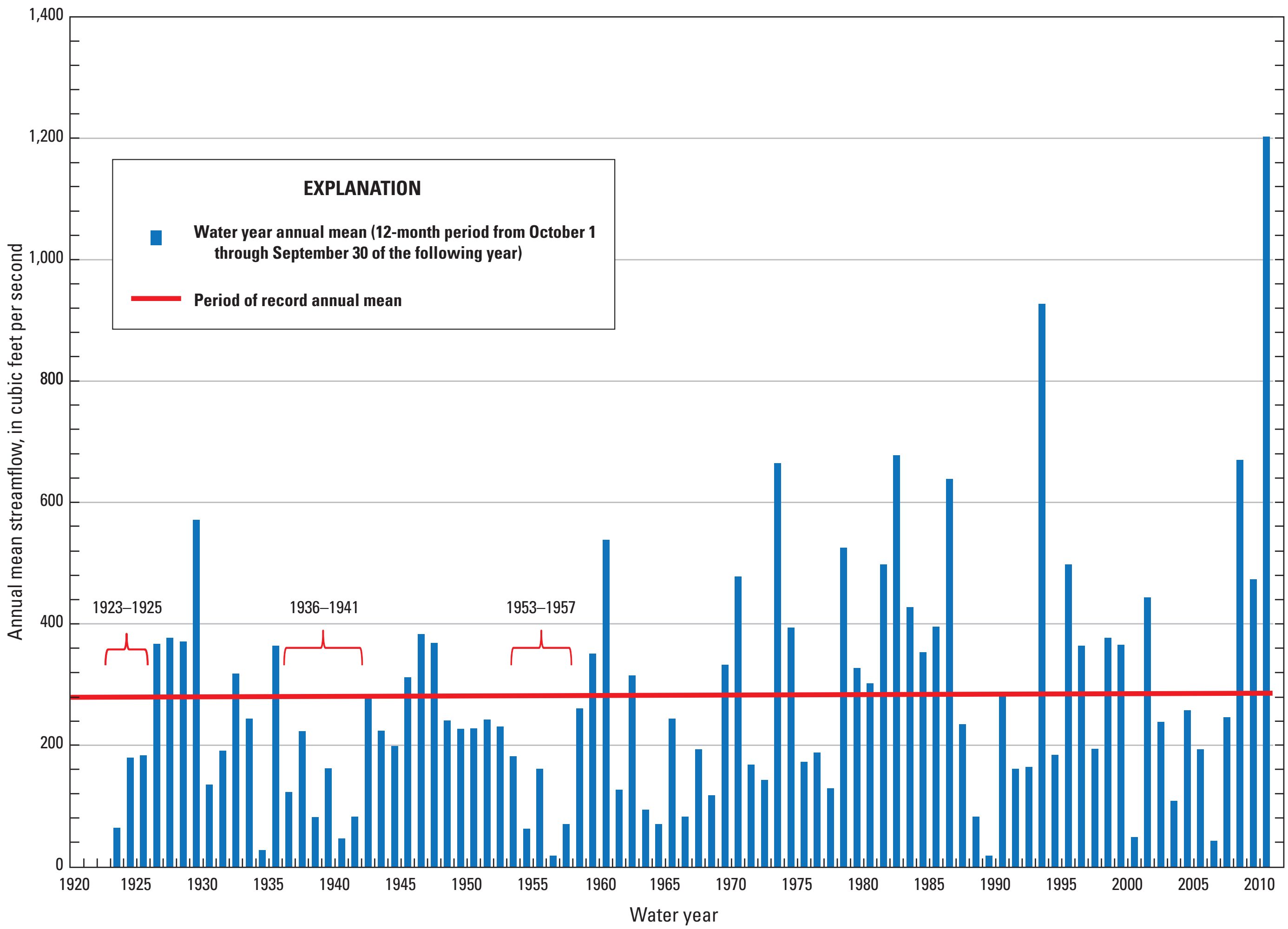


Figure 3. Annual mean and period of record annual mean streamflows for Fox River at Wayland, Missouri (streamgage 05495000, map number 17).

Basic Statistics

Daily-mean streamflow data were downloaded from the USGS NWIS database using U.S. Environmental Protection Agency (EPA) Better Assessment Science Integrating point and Non-point Sources (BASINS) version 4.0 software (2009). BASINS is a multipurpose environmental analysis system that integrates a GIS capability that allows a user to retrieve and identify USGS streamgages in a user defined area and retrieve selected streamflow data from the USGS NWIS database (U.S. Environmental Protection Agency, 2009). Streamgage selection was based on three criteria. First, all streamgages within Missouri were selected. Second, streamgages within basins that drain into Missouri were included. And third, streamgages in neighboring states with basin characteristics similar to basins in Missouri were added to improve the range in basin characteristics and provide additional coverage on streams draining out of Missouri.

The basic statistics for each of the 532 streamgages were computed using a utility within BASINS. The minimum, maximum, mean, and standard deviation of all daily mean streamflows for the period of record approved by the USGS within NWIS are presented in table 4 (available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_4.xlsx*).

BASINS also has the capability to determine low-flow statistics that can be related to the stress that in-stream aquatic organisms may experience. The biologically based design flows of annual 1-day, 3-year (1B3) for the criterion continuous concentration (CCC) and of annual 4-day, 3-year (4B3) for criterion maximum concentration (CMC) also were computed within BASINS [EPA, 2009; tables 5 and 6 (available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_5.xlsx; http://pubs.usgs.gov/sir/2013/5090/downloads/table_6.xlsx*)]. A description of the differences of hydrological-based design flows and biologically based design flows can be found at (U.S. Environmental Protection Agency, 2012). The definition of the 1B3 and 4B3 biologically based design flows from the EPA website is given below.

- The biological method examines all low flow events within a period of record, even if several occur in one year. The biologically based design flow is intended to examine the actual frequency of biological exposure.

The method directly uses site-specific durations (i.e., averaging periods) and frequencies specified in the aquatic life criteria (e.g., 1 day and 3 years for CMC and 4 days and 3 years for CCC).

- Since biologically based design flows are based on durations and frequencies specified in water quality criteria for individual pollutants and whole effluents, they can be based on the available biological, ecological, and toxicological information concerning the stresses that aquatic organisms, ecosystems, and their uses can tolerate. The biologically based calculation method is flexible enough to make full use of special averaging periods and frequencies that might be selected for specific pollutants (e.g., ammonia) or site-specific criteria. This method is empirical, not statistical, because it deals with the actual flow record itself, not with a statistical distribution that is intended to describe the flow record.

To compute the flow duration, trends, and N-day low-flow frequency statistics, the computer program Surface Water Statistics (SWSTAT; Lumb and others, 1990) module by the USGS within BASINS was implemented. The USGS has established standard methods for estimating low-flow frequency statistics for streamgages (Riggs, 1972).

Flow-Duration Statistics

Daily mean streamflows for all complete water years of record for a streamgage were used to compute flow-duration statistics and flow-duration curves. Flow-duration curves (fig. 4) are graphical representations of the percentage of time that streamflow for a given time step is equaled or exceeded during a specified period. The time step used in this study was a daily time step. Flow-duration statistics are points along a flow-duration curve that represent the discharge that is equaled and exceeded for a given percentage of time (table 7 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_7.xlsx*). Flow-duration statistics commonly are denoted in the form of (discharge) D_{xx} where the streamflow "D" is subscripted by the exceedance probability the streamflow is equaled or exceeded. For example, the flow-duration statistic D_{90} is a streamflow that is equaled or exceeded 90 percent of the time. As an example on figure 4 D_{90} would be about 1.5 ft^3/s for streamgage 05506800 and about 21 ft^3/s for streamgage 06918440.

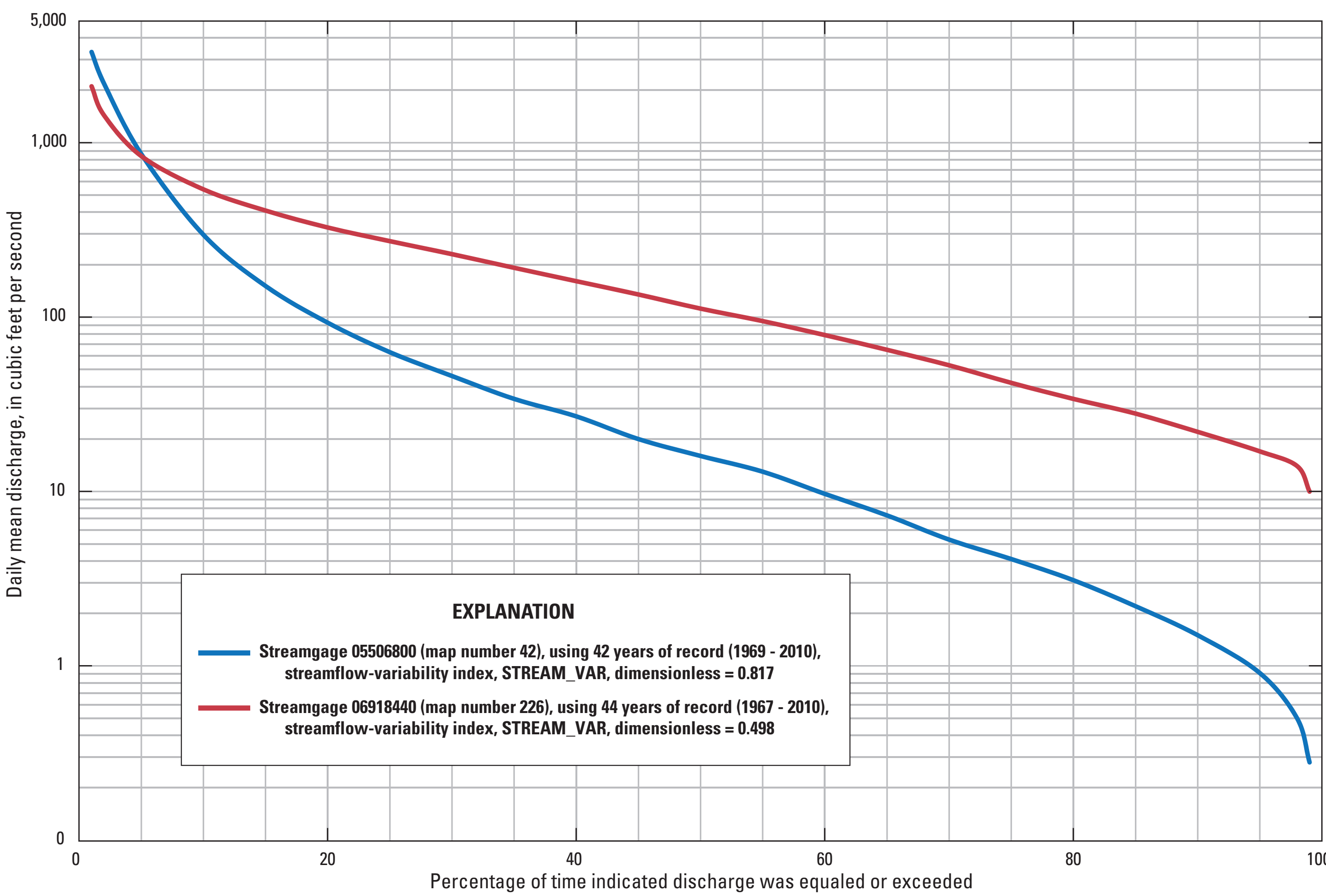


Figure 4. Flow-duration curves for Elk Fork Salt River near Madison, Missouri, (streamgage 05506800, map number 42) and Sac River near Dadeville, Missouri (streamgage 06918440, map number 226).

Flow-duration statistics can be computed for any period of time. Flow-duration statistics were computed using SWSTAT (Lumb and others, 1990) for daily mean streamflows for complete water years (October 1 to September 30) for the period of record for 532 streamgages (table 2). Flow duration statistics also were computed by month, and these results are presented in table 8 (available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_8.xlsx*). The procedure in SWSTAT used to compute flow-duration statistics in this study is similar to the procedures used in Lewis and Esralew (2009).

In figure 4, two flow-duration curves are shown representing streamgage 05506800 (map number 42), which is located in north Missouri (Region 1), and streamgage 06918440 (map number 226), which is located in west-central Missouri (Region 2). The general slope of the flow-duration curve for streamgage 06918440 is flatter than the flow-duration curve for streamgage 05506800. Also, about 95 percent of the time streamgage 06918440 has a greater flow than streamgage 05506800 for a given exceedance probability. As the indicated exceedance probability increases (moving from left to right on the graph in figure 4), the difference in flows also increases. Streamgage 06918440 is located in the Ozark Plateaus physiographic province of the State where extensive areas of karst exist. This area is known to contain one of the largest concentrations of springs in the United States (Imes and Emmett, 1994). The basin upstream from streamgage 06918440 has 116 springs identified compared to zero springs identified for streamgage 05506800 (table 9 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_9.xlsx*). Springs are less susceptible than streams to dry years because of a groundwater component supporting the spring flows. In some basins, springs contribute substantially to base flow during low flow periods. Thus, streams with substantial base flow from spring inflow tend to have flatter duration curves and greater sustained flows.

Trend Statistics

An assumption of frequency analyses is that annual low flows are independent and stationary with time. Trends in data could introduce a bias into the low-flow frequency analyses from climatic or anthropogenic effects. N-day data calculated for annual climatic years were analyzed for the entire period of record for trends using the Kendall's *tau* hypothesis test in the SWSTAT program (Lumb and others, 1990). The Kendall's *tau* test was used to compute the monotonic relation between N-day values (discharge) and time in climatic years (Helsel and Hirsch, 2002). A *p*-value threshold of 5 percent (α=0.05) was used in this study for the Kendall's *tau* test, and *p*-values less than or equal to 5 percent were flagged as having statistically significant trends (positive or negative).

The Kendall's *tau* test was implemented on all stations with 10 or more years of record. If trends were detected, the longest and most recent period of record without a significant trend was used. This variable-length of record method allows for longer records to be used for frequency analyses for many streamgages compared to using a common period of record for all streamgages. Trend analyses were then computed by decreasing the length of record by 1-year increments sequentially from the beginning of the record until a significant trend was not detected. This procedure was used for each streamgage to determine the beginning year of the longest period of recent record without a significant trend for any of the N-day durations tested. The 1-, 2-, 3-, 7-, 10-, 30-, 60-, 90-, 183-, and 365-day durations were tested for trends. Results of the trend analyses indicated a geographic bias in which more streamgages in Region 1 have a trend than streamgages in Regions 2 and 3. The results of the trend analyses determined that shorter periods of record were available for use in low-flow frequency analyses in the northern part of the State compared to the southern part of the State. The effects of springs sustaining base flow and contributing substantially to base flow in the southern part of the State may mask the climatic and anthropogenic effects on low flow. Results of the trend analyses are shown in table 10 (available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_10.xlsx*).

The trend analysis was computed for all 532 streamgages; however, no attempt was made to determine a period of no trend for stations with less than 10 years of record. Streamgages with period of records lengths less than 10 years were analyzed and included in the table for completeness, but the trend analyses may be biased by extreme dry or wet years. Review of streamgage record lengths resulted in 357 streamgages with 10 or more years of record. From the subset of 357 streamgages, trends existed at 134 streamgages. Results of the trend analyses do not affect the computation of annual duration statistics (table 7); however, they do affect the computation of monthly duration statistics for months that were included in years outside of the no trend period (table 8). Monthly duration data for only the period of no trend were recomputed and included in table 8. In general, flow values for the period of no trend were greater than the flow values for the entire period of record. A review of the monthly data for streamgage 05471500 shows about a 21-percent increase in flow for the POR median streamflow (D_{50}) for the month of October from 197 ft^3/s (1946 to 2010 water years) to 238 ft^3/s (1961 to 2010 water years) for the no trend period (table 8).

Low-Flow Frequency Statistics

The low-flow frequency statistics are based on the N-day, Y-year frequency statistic of daily mean streamflow for the climatic year annual period. Low-flow frequency statistics only were computed for the 357 streamgages with 10 or more years of record. These statistics are the minimum consecutive N-day mean streamflow expected to occur once in any Y-years. Y-years is the recurrence interval or return period signifying the average number of years between nonexceedances of a selected low-flow magnitude. Probability is the reciprocal of the annual nonexceedance recurrence interval expressed as 1/Y. For example, the M7D10Y low-flow statistic is the annual minimum mean streamflow for 7 consecutive days that is expected to not be exceeded once on the average during any 10-year period, or in probability terms, has a 0.10 probability of not being exceeded in a given year. For this report, low-flow frequencies were estimated for annual N-day durations of 1-, 2-, 3-, 7-, 10-, 30-, 60-, 90-, 183-, and 365-days. Low-flow recurrence interval (return period) ranged from 1.05 to 500 years and corresponding probabilities ranged from 0.95 to 0.002 (table 11 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_11.xlsx*). Statistics are provided in table 11 for the period of record and for the period of no trend, if applicable.

Each N-day, Y-year frequency statistic was based on an annual series of the lowest mean discharge for a pre-defined consecutive number of days. The annual series for the determination of frequency statistics uses the climatic year of April 1 to March 31 instead of the water year (October 1 to September 30). In Missouri, the months of August, September, and October typically are the months of lowest flows in streams (table 8). Using the climatic year time period prevents the analysis from possibly including the same low-flow event in the fall for 2 consecutive years.

A frequency analysis was implemented on each N-day annual series using a log-Pearson Type III distribution (Riggs, 1972). For this study, SWSTAT was used to determine the annual series of minimum mean low flows to fit them to a log-Pearson Type III distribution, and to plot the computed low-flow frequency curve through the annual values. More specific information about the log-Pearson Type III distribution can be found in Interagency Advisory Committee on Water Data (1982). A visual inspection of the fit of the low-flow frequency curve to the data, and the computed low-flow frequency values to be used, were reviewed for quality assurance.

Methods for Estimating Low-Flow Frequency Statistics at Ungaged Locations in Missouri

Of the 532 streamgages included in this study, the USGS has collected daily mean streamflow data at approximately 350 streamgages during the last 100 years in Missouri. Despite the network of streamgages, State and local agencies have a need to determine select statistics at ungaged stream sites. Three methods that can be used to estimate selected statistics at ungaged sites in Missouri include (1) power curve relations for selected streams with multiple streamgages on the same stream, (2) a drainage-area ratio calculation on streams with a streamgage, and (3) the regional regression equations relating streamflow statistics to drainage-basin characteristics independent of the existence of a streamgage on the stream.

Power Curve Equations for Selected Streams with Multiple Streamgages Method

Within the study area, 28 basins have 2 or more streamgages located on the same stream with a M7D10Y statistic value greater than zero. Most of these streams only had two streamgages located on the same stream, but the St. Francis and Black Rivers each have six streamgages located on the same stream. To improve the estimation of N-day duration frequency estimates at ungaged locations along these streams, a power curve was fit to a log-log plot of drainage area and a selected low-flow frequency statistic. The power curve equations are listed in table 12 (available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_12.xlsx*) by stream name and selected low-flow frequency statistic. The objective of developing the power curve relations (equations) between drainage area and a selected low-flow statistic is to improve the estimate of the statistic compared to other low-flow estimation methods for ungaged locations on selected streams with multiple streamgages. This method does not apply to tributary streams and can only be applied to streams where the selected statistic is nonzero. The applicable extent of the power curve equations upstream from the most upstream streamgage and downstream from the most downstream streamgage is subjective and was based on the findings of the drainage-area ratio method discussed in the regional regression-equation method section later in the report.

The 82 streamgages used in this analysis ranged in drainage area from 5.63 to 14,628 square miles (mi^2). Drainage-area ratios only were computed for streamgages on the same stream. The range of drainage-area ratios was 0.053 to 19.03 for 96 pairs in the 28 selected streams with multiple streamgages (table 13 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_13.xlsx*). The power curve equation developed for each selected stream was used to compute the estimated M7D10Y statistic for each streamgage on the main stem. Absolute differences, in percent, between the observed streamflow (table 14) and the estimates computed from the power curve equations, which were determined for the M7D10Y statistic and the median and standard deviation of the 96 pairs of estimates, are listed in table 14. A smoothed curve calculated from a LOWESS (Locally Weighted Scatterplot Smoothing) algorithm computed using Spotfire S+ statistical software (TIBCO Software Inc. 2008) for the range of drainage-area ratios from 0.05 to 20 is shown in figure 5. The absolute percent differences curve for the LOWESS curve for the power curve equations slopes slightly downward from a drainage-area ratio of 0.06 to 0.5 and 3 to 20, and is relatively flat from 0.4 to 3 (fig. 5).

Drainage-Area Ratio Method

The drainage-area ratio (DAR) method also can be used to estimate selected low-flow statistics for an ungaged location on a gaged stream. The DAR method is based on the

Table 14. Medians and standard deviations of absolute differences between annual mean 7-day low flow for a recurrence interval of 10 years (M7D10Y) statistic using observed streamflow and mutliple-streamgage relations, drainage-area ratio method, and regional regression equations for selected streamgages in Missouri and in neighboring States of Missouri.

[<, actual value is less than value shown; >, actual value is greater than value shown]

Group	Drainage-area ratio range	Number in group	Median absolute difference (percent)	Standard deviation
Power curve	ALL	192	9.9	47.1
	< 0.4 and > 1.5	106	12.3	60.7
	0.4 to 1.5	86	8.8	18.5
Drainage-area ratio method	ALL	192	37.6	301.0
	< 0.4 and > 1.5	106	52.1	228.2
	0.4 to 1.5	86	26.6	373.0
Regional regression equations	ALL	192	35.2	152.4
	< 0.4 and > 1.5	106	34.5	150.5
	0.4 to 1.5	86	44.3	155.7

assumption that streamflow at an ungaged location is the same per unit area as that for a streamgage located upstream or downstream from the ungaged location. This method is applicable to any stream with a streamgage and is not dependent on multiple streamgages being located on the same stream. The accuracy of the drainage-area ratio method is dependent on how close the streamgage and ungaged location are to each other, similarities in drainage area, and other physical and climatic characteristics of the drainage basins (Ries and Friesz, 2000).

The following equation is the DAR method calculation:

$$Q_{DARu=}\left[\frac{DA_u}{DA_g}\right]Q_{og} \tag{1}$$

where

Q_{DARu} is the DAR low-flow frequency estimate of the ungaged site,

DA_u is the drainage area of the ungaged site,

DA_g is the drainage area of the streamgage, and

Q_{og} is the low-flow frequency estimate computed from the observed streamgage record.

The DAR method has been used in previous studies. In Idaho, Hortness (2006) found that the drainage-area ratio limits of paired streamgages were 0.5 to 1.5. In Ohio, Koltun and Schwartz (1986) determined the DAR range was applicable from 0.85 to 1.15, and in Iowa, Eash and Barnes (2012) found the DAR range to be from 0.5 to 1.4. In each study, regression equations were recommended outside of the published DAR ranges.

Similar to the power curve equations method, the absolute differences between DAR and observed estimates of M7D10Y, in percent, were computed for the DAR method. The LOWESS plot of the DAR method also is shown on figure 5. The absolute percent differences for the DAR method is minimized at approximately 0.85 drainage-area ratio and the absolute percent differences increase for drainage-area ratios less than 0.8. For drainage-area ratios greater than 0.9, the absolute percent differences increase gradually to a ratio of 10 and remains relatively flat for ratios of 10 to 20. Absolute differences were greater for the DAR method compared to the power curve equation method for the entire range of drainage-area ratios. The minimal difference between plots is about 15 percent for drainage-area ratios in the 0.8 to 0.9 range. Absolute differences, in percent, between the observed streamflow and the estimates computed from the DAR method were determined for the M7D10Y statistic and the median and standard deviation of the 96 pairs are shown in table 14.

Regional Regression-Equation Method

A common methodology to estimate selected statistics at ungaged locations on streams with or without streamgage data is to relate a streamflow statistic at a streamgage to basin and climatic characteristics. The dependent variable (flow statistic) commonly is log-transformed along with the independent variables (basin and climatic characteristics) to improve linearity. When computing low-flow statistics, the computed or observed flow at a streamgage may be zero for a selected statistic. Numerous streamgages in Missouri have computed low-flow statistics of zero flow. For the N-day durations of 1 to 60 days, the percentage of zero flow values ranged from 28 to 13 percent from an initial data set of 258 streamgages being evaluated for use in the regression analyses. Fifty-one streamgages had low-flow frequency statistics that resulted in residuals that were considered to be outliers in the initial regression data set. Further evaluation of the streamgage data and the basins upstream was implemented to evaluate the applicability of

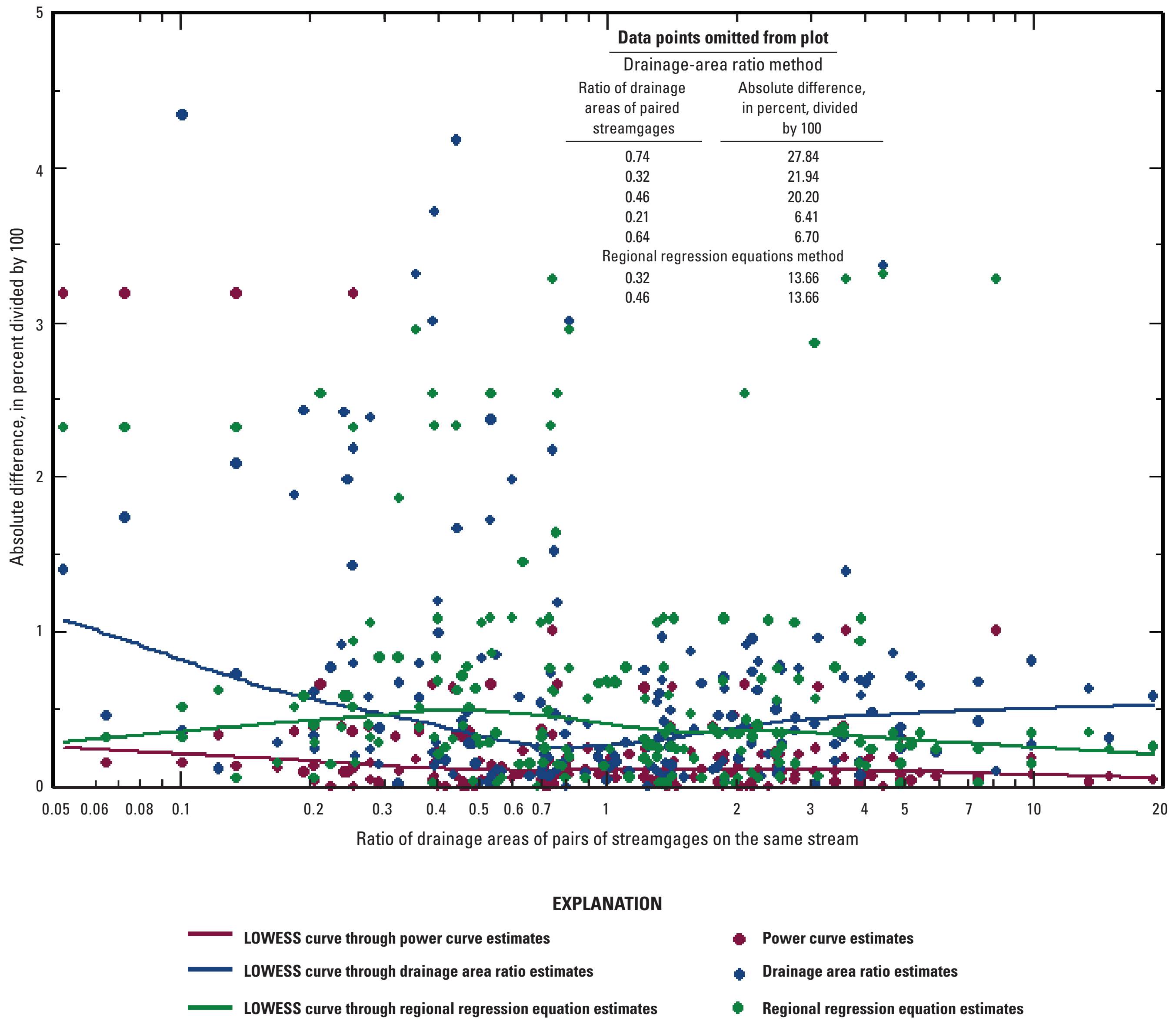


Figure 5. Relation of drainage-area ratio to absolute percent difference in annual 7-day mean low-flow for a recurrence interval of 10 years (M7D10Y) statistic between estimates computed from observed streamflow and estimates derived from the power curve equations method, drainage-area ratio method, and regional regression equations method.

the streamgage data for use in the regression analyses. These 51 streamgages were removed from use in the regression analyses for 1 or more of the following reasons: (1) backwater from downstream confluences, (2) substantial urbanization in the basin, (3) diversion for public water supply, (4) diversion for irrigation, (5) point source inflow from wastewater treatment plants, (6) water withdrawals for industrial use, and (7) storage of small impoundments and water-supply lakes in the basin. The remaining 207 streamgages were considered to be affected minimally by the above anthropogenic effects and were used to develop the regression equations.

Estimates of zero flow computed from observed streamflow often are considered to be censored data (Kroll and Stedinger, 1996; Kroll and Vogel, 2002), and the use of multiple-linear regression is not recommended for censored data (Helsel and Hirsch, 2002). For data sets with zero flow percentages of this magnitude, left-censored regression analyses are appropriate to use (J.E. Kiang, U.S. Geological Survey, written commun., 2012) for regional investigations. Left-censored regression analyses were implemented to allow for the use of a censoring threshold in the development of equations to estimate seven low-flow frequency statistics (M1D10Y, M2D10Y, M3D10Y, M7D10Y, M10D10Y, M30D10Y, and M60D10Y). A weighted left-censored regression technique using the inverse of the variance of the streamgage N-day frequency analyses was selected to weight the streamgage data.

Censored and uncensored streamflow values may be used in a left-censored regression. For this study, zero flows and observed flows less than 0.1 ft^3/s were replaced with a censored value of 0.1 ft^3/s in the regression analyses. Because a censored value is used for zero flows and small magnitude flows only, this method is referred to as left-censored regression. Censored regression is similar to multiple-linear regression, except the regression coefficients are fit by maximum-likelihood estimation (MLE) (Helsel and Hirsch, 2002). MLE uses a probability distribution to match the observed data, assuming the residuals are distributed normally around the regression line for the estimation of the slope and intercept. The variance of the range of predicted values is assumed to be constant. An adjusted maximum-likelihood estimation (AMLE) procedure, implemented in the USGS computer-program library version 4.0 (Lorenz and others, 2011) for Spotfire S+ statistical software (TIBCO Software Inc., 2008), was used to develop the left-censored regression equations in this study. The AMLE computation is a first-order bias adjustment that removes the bias in censored regression estimates in this type of regression (Cohn, 1988).

The 0.1 ft^3/s censored value was used for this study to be consistent with a low-flow study by Eash and Barnes (2012) for streams that flow from Iowa into Missouri. Anthropogenic effects on streamflow are uncertain at some streamgage locations downstream from cities and communities. The base flow may be artificially supported by point and nonpoint sources. Also, streamflow may be affected by temperatures at or below freezing at a streamgage in winter months. Considerable difficulty exists when estimating flows on days when freezing and thawing is occurring. These issues support the use of a censoring threshold in the magnitude of 0.1 ft^3/s.

The final regression equations presented in the report were selected primarily on the basis of minimizing the standard error of estimate. The improvement in the standard error of estimate with the addition of the next statistically significant basin characteristic also was evaluated. If an independent variable did not substantially improve the standard error of estimate, the variable was not added to the equation. Absolute differences, in percent, between the observed streamflow and the estimates computed from the regional regression equations were determined for the M7D10Y statistic, and the median and standard deviation of the 96 pairs are shown in table 14. The LOWESS plot of the regression equation method is shown on figure 5.

The absolute differences for the power curve equations for selected streams with multiple streamgages method is less than the absolute differences for the regression equation method throughout the range of drainage-area ratios shown in figure 5 (0.05 to 20). The absolute differences for the DAR method is less than the regression equation method for drainage-area ratios from 0.4 to 1.5. These drainage-area ratio limits are similar to the drainage-area ratio limits of 0.5 to 1.4 for streams in Iowa (Eash and Barnes, 2012).

Median absolute differences, in percent, and the standard deviation for the power curve equations for selected streams with multiple streamgages and DAR methods for the 96 pairs of streamgages are shown in table 14, for drainage-area ratios outside of the range of 0.4 to 1.5, and for drainage-area ratios within the range of 0.4 to 1.5. The power curve equations for selected streams with multiple streamgage method for drainage-area ratios with the range of 0.4 to 1.5 has a median absolute difference of 8.8 percent and the DAR method has a median absolute difference of 26.6 percent. The standard deviation for the power curve method is 18.5, whereas the DAR method is 373.0 for this range. The statistics indicate the power curve method may provide more accurate estimates of streamflow compared to the DAR method for selected statistics on the streams with multiple streamgages. The power curve equations for selected streams with multiple streamgages method and the DAR method had smaller median absolute differences than the regional regression equation of 44.3 percent. The standard deviation for the power curve equations for selected streams with multiple streamgages method substantially was lower than the regional regression equation standard deviation of 155.7. The DAR method standard deviation was higher at 373.0 than the regional regression equation.

The Wilcoxon signed-rank test was used to determine the statistical difference between the medians of the power curve and DAR methods compared to the regional regression equation method. When the DAR is between 0.4 and 1.5, this test indicates the power curve median of 8.8 (table 14) is statistically less ($p=0$) than the regional regression median of 44.3. The DAR median of 26.6 also is statistically less ($p= 0.048$) than the regional regression median of 44.3. The test also was implemented on the median differences for DAR less than 0.4 or greater than 1.5. The power curve median of 12.3 is statistically less ($p=0$) than the regional regression median of 34.5, and the DAR median of 52.1 is statistically greater ($p=0.026$) than the regional regression median of 34.5.

On the basis of the Wilcoxon signed-rank test, the power curve and DAR methods generally would provide estimates of M7D10Y that are better than estimates obtained using the regional regression equation, when the DAR is between about 0.4 and 1.5. These methods also may provide better estimates for the six other selected low-flow frequency statistics presented in this report, when the DAR is between 0.4 and 1.5 based on the results of the M7D10Y test. The power curve method also may provide better estimates than the regional regression method for DAR less than 0.4 or greater than 1.5; however, power curve estimates for the selected low-flow statistics for ungaged locations with DAR less than 0.4 may have increased uncertainty as the estimate approaches zero flow. Also, for DAR greater than 1.5, limited data exist in Missouri to define the upper limit of the DAR because the majority of selected streams had drainage areas less than the 1.5 times the drainage area of the farthest downstream streamgage. With the uncertainty in computing zero flow for selected low-flow statistics and the limited data sets for DAR greater than 1.5, the power curve method should be limited to the DAR range of 0.4 to 1.5.

Development of Regional Regression Equations for the Estimation of Low-Flow-Frequency Statistics

The most commonly used method to estimate selected low-flow statistics at an ungaged location is the use of regression equations. With the advancement of software and availability of digital geospatial data, basin characteristics can be measured from digital geospatial data quickly and efficiently for use in the development of regression equations. Regression equations were developed using frequency data from 207 streamgages and the results from a statewide analysis were evaluated for regional bias. Spatial analysis of the statewide regression residuals indicated that the State could be divided into three low-flow regions similar to the flood regions defined by Alexander and Wilson (1995).

Basin Characteristics

Computation of basin characteristics using GIS software and an increasing number of digital geospatial data have substantially added to the number of possible independent variables for use in regression analyses. All basin and climatic characteristics evaluated for use in this study were measured from digital geospatial data to allow for the automated computation of the characteristic. Review of previous low-flow studies in Missouri and in other States was completed to denote which characteristics were likely to be statistically significant in regression equations. A list of characteristics was compiled and additional characteristics that could be compiled from the same data source were included in the list. The completed list included 35 basin and climatic characteristics (table 15 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_15.xlsx*). Digital geospatial data were assembled to cover most of the 532 streamgages listed in table 1. All 35 characteristics were computed for each streamgage except for springs and the Missouri and Mississippi River streamgages. For some digital geospatial data types, the data were not available in States bordering Missouri. Every effort was made to make each digital geospatial data type as complete as possible by compiling additional non-digital data where available and appended to the existing digital geospatial data. The basin and climatic characteristics can be categorized into four categories: morphometric (physical or shape) characteristics, hydrologic characteristics, pedologic (soils)/geologic/land-use characteristics, and climatic characteristics (table 15).

Morphometric characteristics were derived from a USGS digital elevation model (DEM) with a 10-meter resolution (1/3 arc-second National Elevation Dataset) available in 2011. The DEM data are updated on a regular basis as more recent and accurate elevation data become available (L.A. Phillips, U.S. Geological Survey, oral commun., 2011). The latest DEM data set may be retrieved from the USGS National Map Viewer (U.S. Geological Survey, 2012). Definition of the characteristics and the source(s) used to compute the characteristic are listed in table 15. With higher resolution and more accurate DEM data sets, there are slight differences in the computation of drainage area at some streamgages compared to previously published data. For consistency purposes, the GIS-derived drainage area values were used in the regression analyses.

The Mississippi Alluvial Plain in southeastern Missouri is a relatively flat area that is drained by a series of man-made drainage ditches. Existing (2011) DEM data are not accurate enough to automatically define surface and channel features from the data. Thus, the basin boundaries may be less accurately located in that area. To improve the interpretation of the basin boundaries, the 1:24,000-scale USDA/NRCS Watershed Boundary Dataset (USGS and USDA/NRCS, 2009) using 12-digit hydrologic unit codes (HUCs) and the 1:24,000-scale USGS National Hydrography Dataset (Simley and Carswell, 2009) were implemented to define the basin boundaries for select streamgages in the Mississippi River Alluvial Plain.

The hydrologic characteristic streamflow-variability index (STREAM_VAR) is a measure of the steepness of the slope of the duration curve and is dimensionless (Koltun and Whitehead, 2002). The STREAM_VAR characteristic is computed by (1) computing a flow-duration curve using daily mean discharge data to obtain discharge values at 5-percent exceedance intervals from 5 to 95 percent, and (2) calculating the standard deviation of the logarithms of the 19 discharge values corresponding to the 5-percent exceedance intervals from 5 to 95 percent (Searcy, 1959). The flow-duration curve is a cumulative frequency curve that shows the percentage of time that a specific discharge is equaled or exceeded (fig. 4).

The STREAM_VAR statistic is calculated as:

$$STREAM\_VAR = \sqrt{\frac{\sum_{i=5,5}^{95} (log(Q_{ci}) - \overline{log(Q_c)})^2}{18}} \quad (2)$$

where

$log(Q_{ci})$ is the base 10 logarithm of the *i*-percent duration streamflow (*i*=5, 10, 15, 20… 95), and

$\overline{log(Q_c)}$ is the mean of the base 10 logarithms of the 19 streamflow values at 5-percent intervals from 5 to 95 percent on the flow-duration curve of daily mean discharges.

If an *i*-percent duration streamflow value is zero (which cannot be log-transformed), the log (Q_{ci}) value was set to zero in equation 2 to allow all nineteen 5-percent intervals to be included in the calculation of STREAM_VAR.

The magnitude of the STREAM_VAR value is related inversely to the capacity of a basin to sustain base flow in a stream. The smaller the STREAM_VAR value, the flatter the slope of the flow-duration curve, which is indicative of higher sustained flows. Conversely, the larger the STREAM_VAR value, the steeper the slope of the flow-duration curve, which is indicative of less capacity of the basin to sustain base flows (fig. 4). The nearest neighbor interpolation technique in ArcGIS was used to create a STREAM_VAR grid for the State of Missouri. The grid was based on the observed STREAM_VAR point data from the 207 streamgages that were included in the regional regression analyses (fig. 6). The STREAM_VAR grid was then used to interpolate a mean STREAM_VAR basin characteristic value for a streamgage.

Pedologic (soils)/geologic/land-use characteristics were computed from the NRCS Soil Survey Geographic (SSURGO) Database (NRCS, 2012; Multi-Resolution Land Characteristics Consortium, National Land Cover Dataset (2012), and geospatial data obtained from the Missouri Environmental Geology Atlas (Missouri DNR, 2007). Additional information about springs from an Arkansas Geological Survey report on springs (Branner, 1937) was added to the Missouri springs digital geospatial data to provide more complete information on spring locations in the study area. The basin characteristics from these sources were processed using the National Water-Quality Assessment (NAWQA) Area-Characterization Toolbox (NACT.tbx) developed by Price and others (2010).

Mean annual precipitation was obtained from the PRISM Climate Group (PRISM Climate Group, 2008). Climatic characteristics were digitized and rectified from Technical Paper No. 40 by Hershfield (1961). The digitized contours were converted to a raster surface for processing by the NACT.tbx in ArcGIS version 9.3 (Environmental Systems Research Institute, 2001, 2009).

Regression Analyses

Missouri has three major physiographic provinces: Central Lowlands, Ozark Plateaus, and the Mississippi Alluvial Plain (fig. 7; Fenneman, 1938). The effect of these physiographic provinces on the streamflow characteristics has been discussed and documented by Skelton (1970, 1976) and Alexander and Wilson (1995). Skelton and Homyk (1970) provided a summary of the low-flow characteristics of streams in each physiographic province:

> Low-flow potential of most streams in the Plains (Central Lowlands) is poor because of the low hydraulic conductivity of the clays and shales of the area… Storage Reservoirs are required for effective utilization of surface-water supplies in this region.
>
> The streams of the Ozarks (Ozark Plateaus) generally have the best sustained low flows in the state because of inflow from extensive natural underground reservoirs in the soluble carbonate rocks. However, the low-flows of some streams in this region are affected by the underground solution cavities, resulting in water losses and non-conformance to areal patterns.
>
> Low flows in the Lowlands (Mississippi Alluvial Plain) are second in magnitude to those of the Plateaus and are sustained by groundwater inflow from the extensive alluvial aquifer. This is a relatively flat region where major manmade channels have been constructed for drainage of the excellent farmland. Since construction of the ditches, groundwater releases from the alluvium have been generally larger, and this accounts, at least in part, for the well-sustained low flows of the region.
>
> Minimum streamflow in Missouri usually occurs in fall or late summer. More minimum flows have occurred at long-time gaging stations (streamgages) in August, September and October than in any other period.

The differences in flow-duration curves computed for a streamgage in the Central Lowlands (05506800, map number 42) and a streamgage in the Ozark Plateaus (06918440, map number 226) are illustrated in figure 4. In the Ozark Plateaus province, springs can provide a major part of the base flow at a streamgage, which results in greater base flows when compared to base flows that are observed at streamgages in the Central Lowlands and Mississippi Alluvial Plain of comparable sized basins.

A preliminary statewide regression analysis was implemented using ordinary-least-squares (OLS) regression and selected streamgages with 10 or more years of record. The low-flow frequency statistic M7D10Y was chosen for all regression analyses because it is a commonly used statistic for surface-water and water-quality regulation in Missouri. Residual values (differences between low-flow frequency statistics computed from observed streamflow and those predicted from the regression equations) from the preliminary statewide regression analyses were mapped at streamgage locations to identify spatial trends in the predictive accuracy of the regression equation. Residuals from the statewide analyses confirmed that the physiographic provinces affected the predictive accuracy of the preliminary statewide M7D10Y regression equation. OLS regression analyses implemented using subsets of the statewide data set were computed separately for each major physiographic province to compare regional and statewide predictive accuracies. An improvement in accuracy was indicated using a subset of streamgage data for the Central Lowlands and Ozark Plateaus regions. On the basis of previous studies findings and the results of the OLS analyses, the three major physiographic provinces were then evaluated for regional regression analyses.

Regional Regression-Equation Development

Regression equations were developed for each of the physiographic provinces using the same methodology for each region. In cooperation with Missouri Department of Natural Resources, the statistics selected for analyses were the 10-year frequency with N-day durations of 1, 2, 3, 7, 10, 30, and 60 days. Similar to the statewide regression analyses, the selection of basin and climatic characteristics and the evaluation

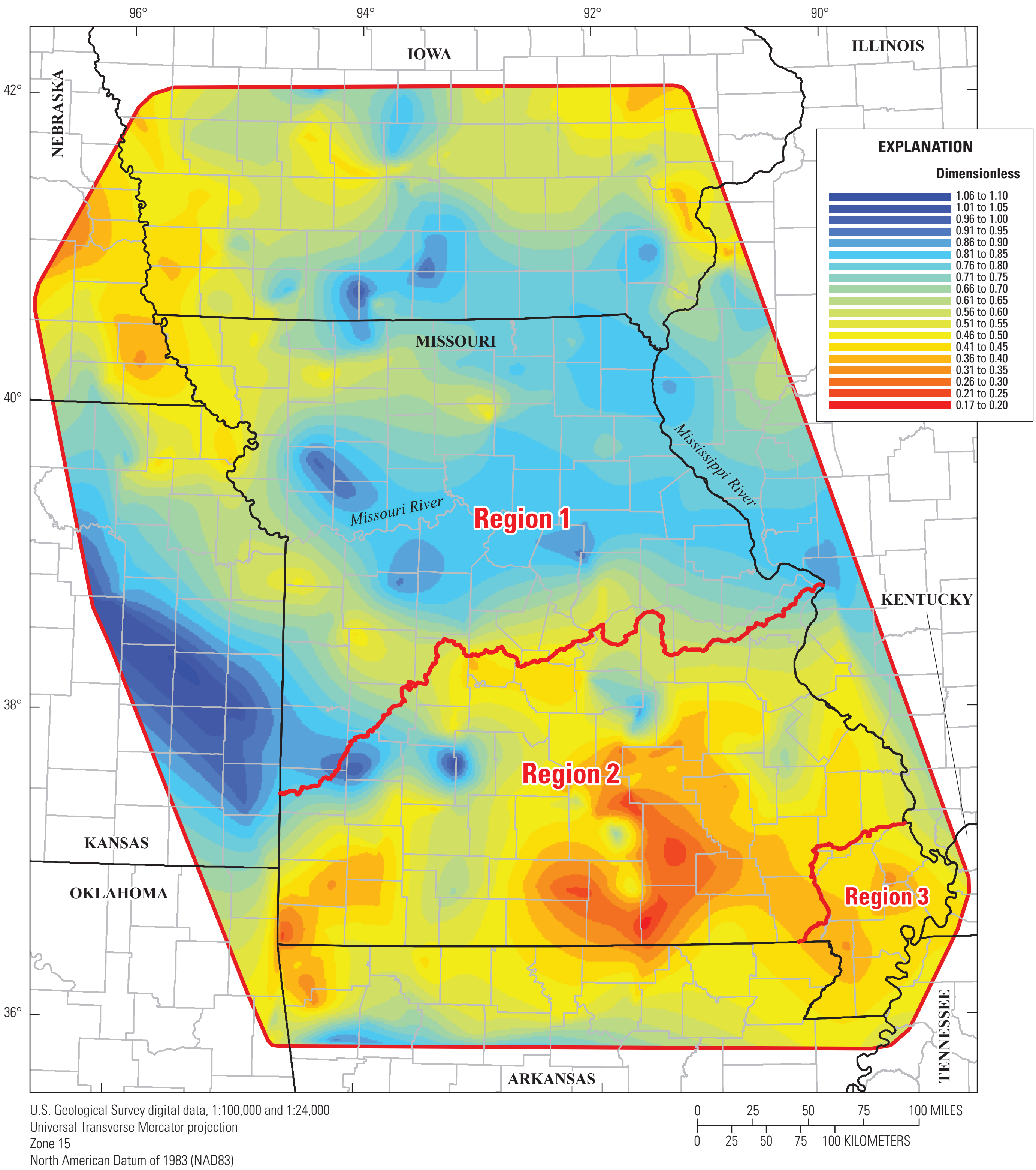


Figure 6. Streamflow-variability index (STREAM_VAR) grid and low-flow regions for Missouri.

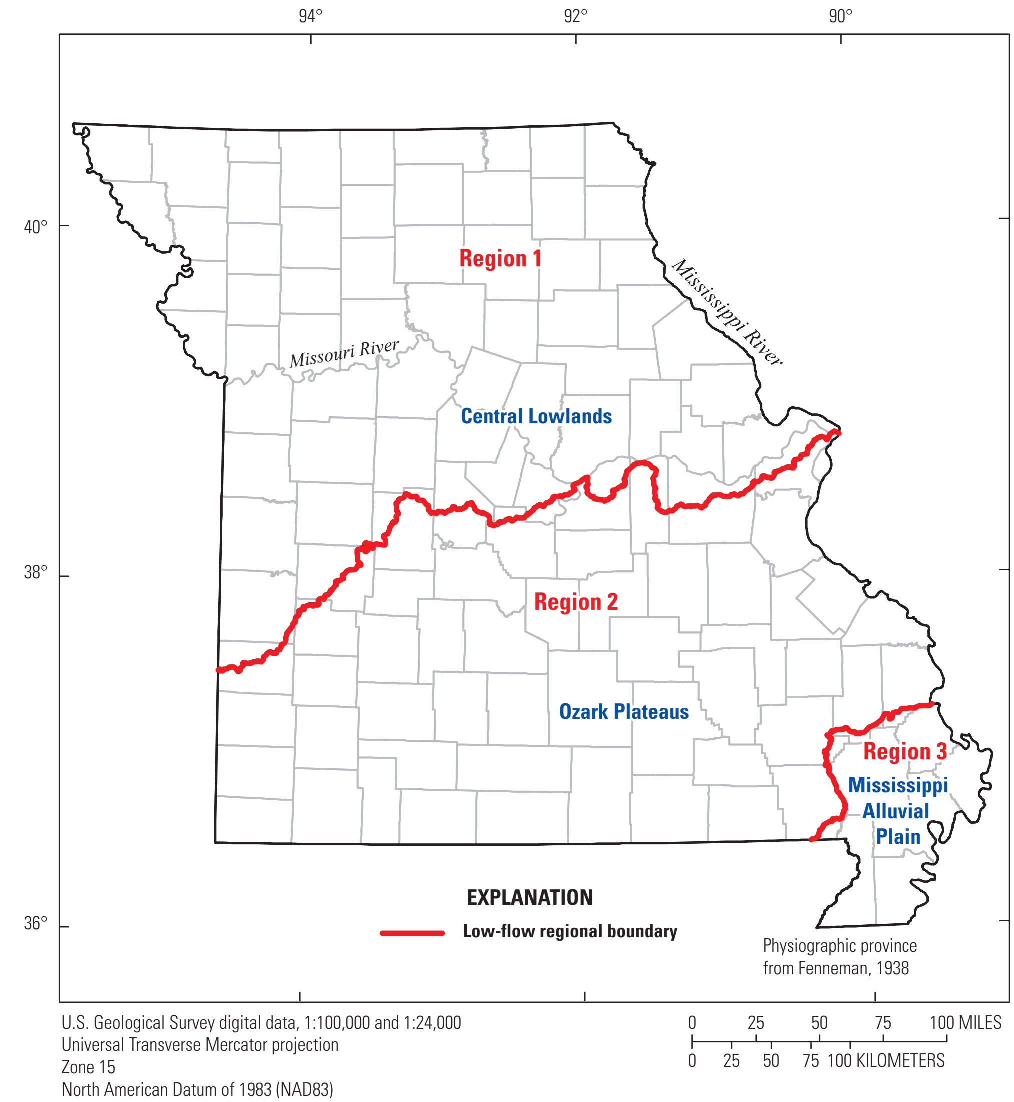


Figure 7. Regions for low-flow frequency equations and physiographic provinces in Missouri.

of the accuracy of the regional equations was based on the M7D10Y statistic. The streamgages were subdivided into separate data sets by physiographic province with consideration given to major drainage basin boundaries at the 4-digit hydrologic unit code level. The physiographic provinces were referenced using the terminology of Alexander and Wilson (1995) with the Central Lowlands as Region 1, Ozark Plateaus as Region 2, and Mississippi Alluvial Plain as Region 3 for the development of the regression equations (fig. 7). Boundaries of Region 1, 2, and 3 approximate the location of the physiographic province boundaries but the two are not coincident. Streamgages were identified by region. Streamgages with minimal anthropogenic effects, sufficient record length, and appreciably unaffected by backwater conditions were selected for regional analyses. The number of streamgages selected for Region 1, 2, and 3 were 77, 120, and 10, respectively.

To identify which basin characteristics are statistically significant for inclusion in the regression analyses, the Efroymson stepwise-selection method (Efroymson, 1960) was used to define potential explanatory variables from the list of 35 characteristics. The procedure is similar to forward selection, which tests basin characteristics one by one and identifies those that are statistically significant; however, as each new basin characteristic is identified as being significant, partial correlations are checked to see if any previously identified variables can be deleted (Ahearn, 2010). Highly correlated characteristics were included in the Efroymson selection method one at a time to avoid problems with multicolliearity. Significant characteristics were defined for each region in Missouri. The statistical analyses were implemented using Spotfire S+ statistical software (TIBCO Software Inc., 2008).

For Region 1, the characteristics of drainage area (DRNAREA), longest flow length (LFPLENGTH), basin shape (SHAPE), soil type A (SOILASSURGO), cultivated crops (CROPSNLCD01), and streamflow-variability index (STREAM_VAR) were found to be statistically significant (table 15). To evaluate which combination of characteristics to use for the left-censored regression, a linear model subset selection was used to identify the “best” three linear regression model combinations for each of the one-variable to six-variable regression equations. The variables used in the analyses were transformed using a natural log-transformation except for SOILASSURGO and CROPSNLCD01. These two variables were not log-transformed because they represent a percentage with numerical limits of 0 to 100. STREAM_VAR also was not log-transformed because it is based on flow durations in percent. The final regression model was based on the following performance metrics:

- Adjusted *R* Squared (Adj-R^2) is an alternative to *R*-Squared (R^2) in which the percentage of variation in the dependent variable (M7D10Y) can be explained by the variation of the independent variables in the model. In contrast to R^2, Adj-R^2 is adjusted for the number of parameters in the model [number of streamgages and number of independent variables (basin characteristics)] (Freund and Littell, 2000);

- Mallow's *Cp* statistic is a measure of the total squared error for a subset model containing *n* independent variables (Freund and Littell, 2000). Mallow's *Cp* is an indicator of model bias (Cavalieri and others, 2000). Models with a large *Cp* are biased because they contain independent variables that are not important in the population;

- Predicted REsidual Sum of Squares (PRESS) statistic is the sum of squares of residuals using models obtained by estimating the equation with all observations except for the i^{th} observation (Freund and Littell, 2000) and is an estimate of the prediction error sum of squares. The PRESS statistic measures how well the regression model predicts the i^{th} observation as though it were a new observation (Cavalieri and others, 2000).

The Adj-R^2 statistic is maximized and the Mallow's *Cp* and PRESS statistics are minimized with better combinations of independent variables in a regression model that explain more of the variance in the dependent variable. Incremental improvements in the performance metrics also were evaluated with the addition of another independent variable to the model. The linear model subset results indicated that three independent variables provided the best model to use in the left-censored regression analyses based on the above criteria; however, the best three models were close in statistical performance, so the three sets of models were each evaluated in the left-censored regression analyses.

The combination of DRNAREA, LFPLENGTH, and STREAM_VAR resulted in the lowest standard error of estimate of 80.8 percent for the M7D10Y statistic for Region 1 (table 16). Equations for Region 1 for the 1-, 2-, 3-, 7-, 10-, 30-, and 60-day 10-year frequency statistics are presented in table 16. The standard error of estimate ranged from 79.6 percent for the M30D10Y statistic to 94.2 percent for the M1D10Y statistic for Region 1.

For Region 2,. the most statistically significant independent variables from the Efroymson selection method were DRNAREA, LFPLENGTH, main channel slope measured between the 10- and 85-percent points along the longest flow path (CSL1085LFP), and STREAM_VAR. The linear model subset results indicated that a two-variable equation was the most efficient model to use in the left-censored regression analyses. The two-variable equation with the lowest standard error of estimate included the independent variables of DRNAREA and STREAM_VAR. These two variables resulted in a standard error of estimate of 51.0 percent for the M7D10Y and a range from 48.2 percent for the M30D10Y statistic to 72.1 percent for the M1D10Y statistic (table 16).

For Region 3, a limited number of streamgages (10) with sufficient record length were available for regression analyses. Linear model subset results were unable to define statistically significant variables. The variables found to be statistically significant in the Region 1 and 2 analyses were evaluated in the left-censored regression for Region 3. Only DRNAREA and STREAM_VAR were found to be statistically significant. The standard error of estimate was more than 100 percent for one-variable models. When the variables DRNAREA and STREAM_VAR were included in a two-variable model, the standard error of estimate was substantially lowered to 74.2 percent for the M7D10Y statistic. Because these two variables are included in the final left-censored equations for the two other regions of the State, and substantial improvement in the predictive accuracy of the equations was obtained with a two-variable equation compared to either one-variable equation, final regression equations for Region 3 were developed using two variables with only 10 streamgages in the data set. For region 3, the standard error of estimate ranges from 48.1 percent for the M60D10Y statistic to 96.2 percent for the M2D10Y statistic (table 16).

The observed values were plotted against the estimated values from the left-censored equations presented in table 16 for the M7D10Y statistic in figure 8. Region 1 shows a fairly uniform distribution around the line of equality. M7D10Y flow values of zero were not plotted on the log-log plots. Region 2 shows a fairly uniform distribution around the line of equality except for streamgages 07070500 and 07061900 (map numbers 468 and 436), which appear to be overestimated by the regression equations. Streamgage 07070500 is in the Eleven Point River Basin, which includes a major tributary (Hurricane Creek) that is a losing stream (Kleeschulte and others, 2008) to Big Spring in the Current River Basin. Similarly, Logan Creek (streamgage 07061900) is a losing stream to Blue Spring in the Current River Basin (Kleeschulte and others, 2008).

It is difficult to quantify the magnitude of streamflow lost in losing streams or the magnitude of streamflow gained in gaining streams on a regional basis. Interestingly, the standard error of estimate is the lowest for Region 2 compared to the rest of the State. This may be due, in part, to the stability and sustained base flows to area streams within Region 2. For Regions 1 and 3, the scatter around the line of equality is fairly uniform. Using the regional regression equations, predicted low-flow frequency statistics were computed for

Table 16. Regional-regression equations for the 1-, 2-, 3-, 7, 10-, 30-, and 60-day durations with a recurrence interval of 10 years on unregulated streams in Missouri.

[MLE, Maximum likelihood estimation; SEE, Standard error of estimate; e, Exponential function; M1D10Y, annual 1-day mean low flow with a recurrence interval of 10 years; M2D10Y, annual 2-day mean low flow with a recurrence interval of 10 years; M3D10Y, annual 3-day mean low flow with a recurrence interval of 10 years; M7D10Y, annual 7-day mean low flow with a recurrence interval of 10 years; M10D10Y, annual 10-day mean low flow with a recurrence interval of 10 years; M30D10Y, annual 30-day mean low flow with a recurrence interval of 10 years; M60D10Y, annual 60-day mean low flow with a recurrence interval of 10 years; DRNAREA, GIS drainage area, square miles; LFPLENGTH, Length of longest flow path, miles; STREAM_VAR, streamflow-variability index,dimensionless]

Statistic	Number of streamgages used to develop equation	Statistic equation	MLE SEE (percent)	MLE SEE (unbiased percent)	SEE (percent)
Region 1					
M1D10Y	77	$M1D10Y=0.029*(DRNAREA)^{2.596}*(LFPLENGTH)^{-1.903}*e^{STREAM\_VAR*-5.909}$	0.776	0.797	94.2
M2D10Y	77	$M2D10Y=0.032*(DRNAREA)^{2.645}*(LFPLENGTH)^{-1.982}*e^{STREAM\_VAR*-5.904}$	0.773	0.794	93.7
M3D10Y	77	$M3D10Y=0.036*(DRNAREA)^{2.593}*(LFPLENGTH)^{-1.912}*e^{STREAM\_VAR*-5.913}$	0.754	0.774	90.6
M7D10Y	77	$M7D10Y=0.057*(DRNAREA)^{2.379}*(LFPLENGTH)^{-1.554}*e^{STREAM\_VAR*-6.650}$	0.690	0.709	80.8
M10D10Y	77	$M10D10Y=0.047*(DRNAREA)^{2.580}*(LFPLENGTH)^{-1.911}*e^{STREAM\_VAR*-5.758}$	0.698	0.717	82.0
M30D10Y	77	$M30D10Y=0.266*(DRNAREA)^{2.174}*(LFPLENGTH)^{-1.557}*e^{STREAM\_VAR*-6.192}$	0.682	0.700	79.6
M60D10Y	77	$M60D10Y=0.389*(DRNAREA)^{2.458}*(LFPLENGTH)^{-2.072}*e^{STREAM\_VAR*-5.487}$	0.684	0.702	79.9
Region 2					
M1D10Y	120	$M1D10Y=1.605*(DRNAREA)^{1.285}*e^{STREAM\_VAR*-10.972}$	0.639	0.647	72.1
M2D10Y	120	$M2D10Y=1.624*(DRNAREA)^{1.276}*e^{STREAM\_VAR*-10.743}$	0.490	0.496	52.8
M3D10Y	120	$M3D10Y=1.751*(DRNAREA)^{1.266}*e^{STREAM\_VAR*-10.740}$	0.481	0.487	51.8
M7D10Y	120	$M7D10Y=2.197*(DRNAREA)^{1.244}*e^{STREAM\_VAR*-10.807}$	0.475	0.481	51.0
M10D10Y	120	$M10D10Y=2.314*(DRNAREA)^{1.236}*e^{STREAM\_VAR*-10.730}$	0.467	0.473	50.1
M30D10Y	120	$M30D10Y=2.392*(DRNAREA)^{1.215}*e^{STREAM\_VAR*-10.240}$	0.451	0.457	48.2
M60D10Y	120	$M60D10Y=2.232*(DRNAREA)^{1.200}*e^{STREAM\_VAR*-9.515}$	0.465	0.471	49.8
Region 3					
M1D10Y	10	$M1D10Y=11.462*(DRNAREA)^{1.041}*e^{STREAM\_VAR*-13.217}$	0.665	0.795	93.9
M2D10Y	10	$M2D10Y=13.343*(DRNAREA)^{1.029}*e^{STREAM\_VAR*-13.351}$	0.677	0.809	96.2
M3D10Y	10	$M3D10Y=8.224*(DRNAREA)^{1.004}*e^{STREAM\_VAR*-11.640}$	0.572	0.684	77.2
M7D10Y	10	$M7D10Y=8.182*(DRNAREA)^{0.989}*e^{STREAM\_VAR*-11.228}$	0.554	0.662	74.2
M10D10Y	10	$M10D10Y=8.109*(DRNAREA)^{0.981}*e^{STREAM\_VAR*-11.007}$	0.559	0.668	75.0
M30D10Y	10	$M30D10Y=4.792*(DRNAREA)^{0.965}*e^{STREAM\_VAR*-8.990}$	0.396	0.473	50.1
M60D10Y	10	$M60D10Y=2.149*(DRNAREA)^{0.948}*e^{STREAM\_VAR*-6.366}$	0.382	0.457	48.1

the streamgages used in the regression analyses (table 17 available on CD and at *http://pubs.usgs.gov/sir/2013/5090/downloads/table_17.xlsx*).

A comparison of the three regions for the M7D10Y frequency statistic is shown graphically in figure 9. The curves in figure 9 are from the final regional regression equations using representative basin characteristics such as a range of drainage areas from 120 to 2,200 mi^2, a constant streamflow variability index of 0.5, and a range of stream lengths from 36 to 140 miles dependent on drainage area size. The highest M7D10Y estimates are shown for Region 2 for a given size of drainage area. Region 3 is slightly higher than Region 1 for drainage areas of less than about 1,000 mi^2 and is lower than Region 1 for basins greater than about 1,000 mi^2. Use of different combinations of characteristics may produce slightly different results. Region 2 has greater low flows statistics because of the inflow from springs in the karst areas of the Ozark Plateaus and there appears to be a cumulative effect of increasing base flows with increasing drainage area size compared to base flows from the other two regions.

A plot of the basin characteristics and residuals of the M7D10Y frequency statistic for each region is shown in figure 10. The magnitude and numerical sign of the residuals were checked for possible regional biases and none were

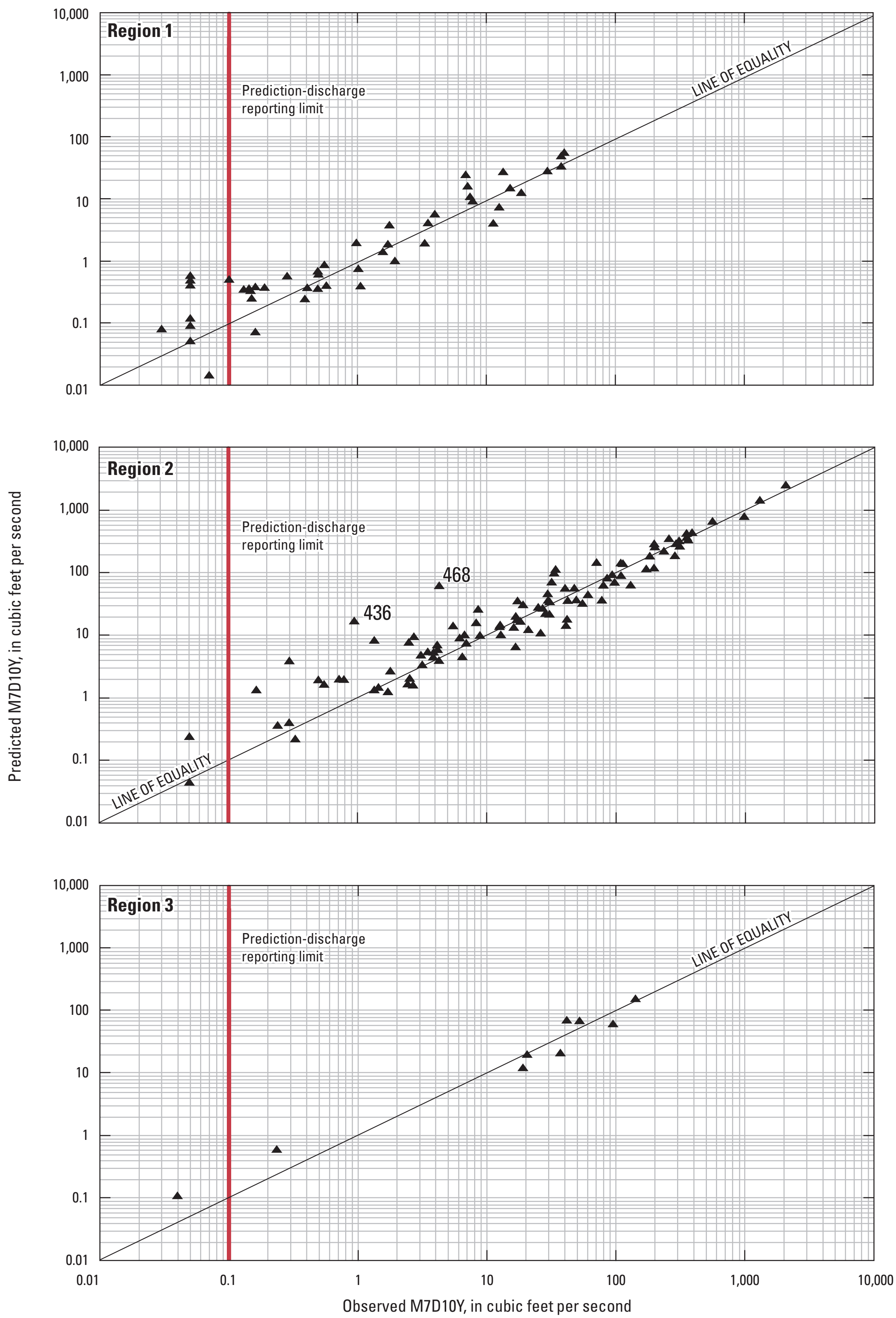


Figure 8. Relation between the annual 7-day mean low-flow discharges for a recurrence interval of 10 years (M7D10Y) computed from observed streamflow and those predicted from regression equations for low-flow regions in Missouri.

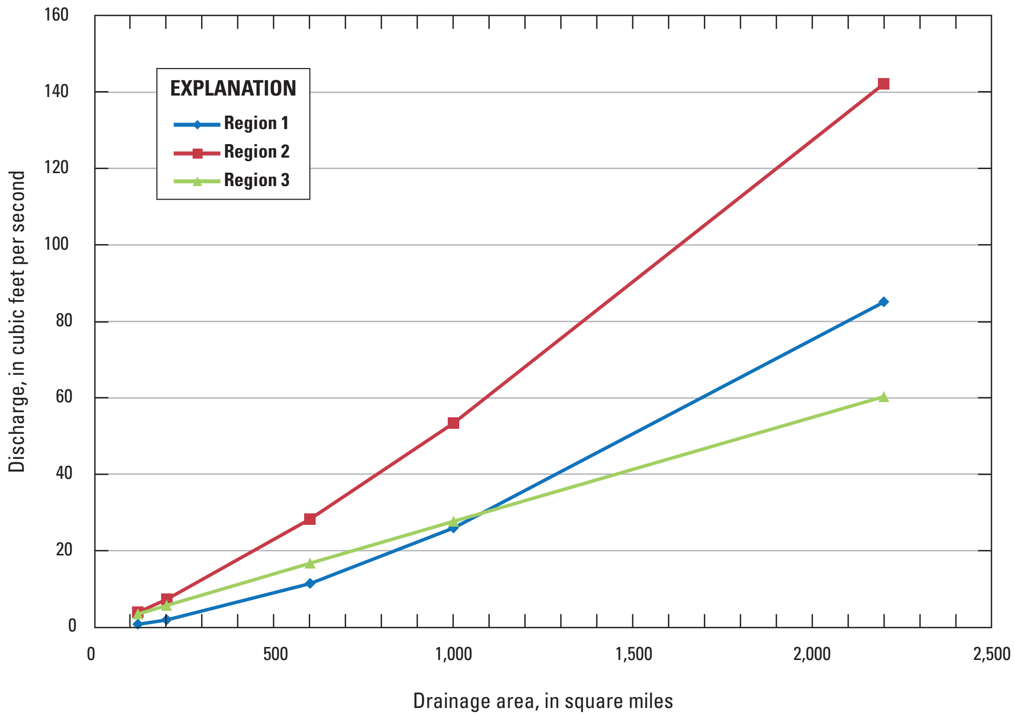


Figure 9. Curves from regression equations for the 7-day mean low flow for a recurrence interval of 10 years (M7D10Y) statistic for each low-flow region for a given set of basin characteristics.

found. The random scatter of the points above and below the zero reference line indicates that the models were satisfactorily meeting the assumption of multiple regression techniques. The residuals for Region 2 indicate streamgage 07072500 (map number 472) is over predicted substantially. This streamgage is on the Black River at Black Rock, Arkansas and the observed M7D10Y value is 2,056 ft^3/s (table 11) and the predicted value is 2,552 ft^3/s using the regional regression equation for the M7D10Y statistic in Region 2 in table 16 and basin characteristics from table 4; however, the difference between the observed and predicted values is less than 25 percent.

Application and Limitations on the Estimation of Low-Flow Frequency Statistics

The report presents three methods for computing selected low-flow frequency statistics at ungaged stream sites in Missouri. Because the accuracy degrades in the predicted low-flow frequency statistics in the three methods presented in this report, the user may apply the predictive methods in the following order: (1) the power curve equations for selected streams with multiple streamgage method should be employed if the ungaged location meets the criteria for use of this method, (2) the DAR method should be used if the ungaged location meets the criteria for use of this method and if the power curve method cannot be used, and (3) the regional regression equations should be used if the ungaged location meets the criteria for use of this method, and if neither the power curve equations for selected streams with multiple streamgage method nor the DAR method can be used.

The first method is based on the 28 streams investigated in this study with multiple streamgages; the power curve method appears to the most reliable method to estimate low-flow frequency estimates because it has the lowest absolute percent differences of the methods evaluated. Equations for selected frequency statistics are shown in table 12 along with the range of drainage areas for each stream for which the equations are applicable. The range of drainage areas were computed by multiplying the drainage area of the streamgage with the smallest drainage area by 0.4 and multiplying the drainage area of the streamgage with the largest drainage area by 1.5. Where the 1.5 multiplier is larger than the total drainage area of the stream at its mouth, the method is applicable only to the mouth of the stream. If an ungaged location is selected on one of the streams listed in table 12, the user may use the power curve equation to compute the desired statistic if the computed drainage area is within the drainage area range listed in table 12. The 0.4 to 1.5 drainage-area ratio limitation is based on the results of the drainage area method and Eash and Barnes (2012). An example of a power curve equation and its range of applicability are shown in figure 11 for the St Francis River; however, this restriction may be conservative because the power curve has lower absolute percent differences than the regression equation method throughout the drainage area range of 0.1 to 10 (fig. 5).

The second method is based on drainage-area ratios for ungaged locations on streams with a streamgage that has low-flow frequency statistics computed. This method is applicable to any stream where more than 10 years of streamgage data are available and the flow data represents existing basin conditions. If the ungaged location has a drainage-area ratio between 40 and 150 percent of the drainage area of the

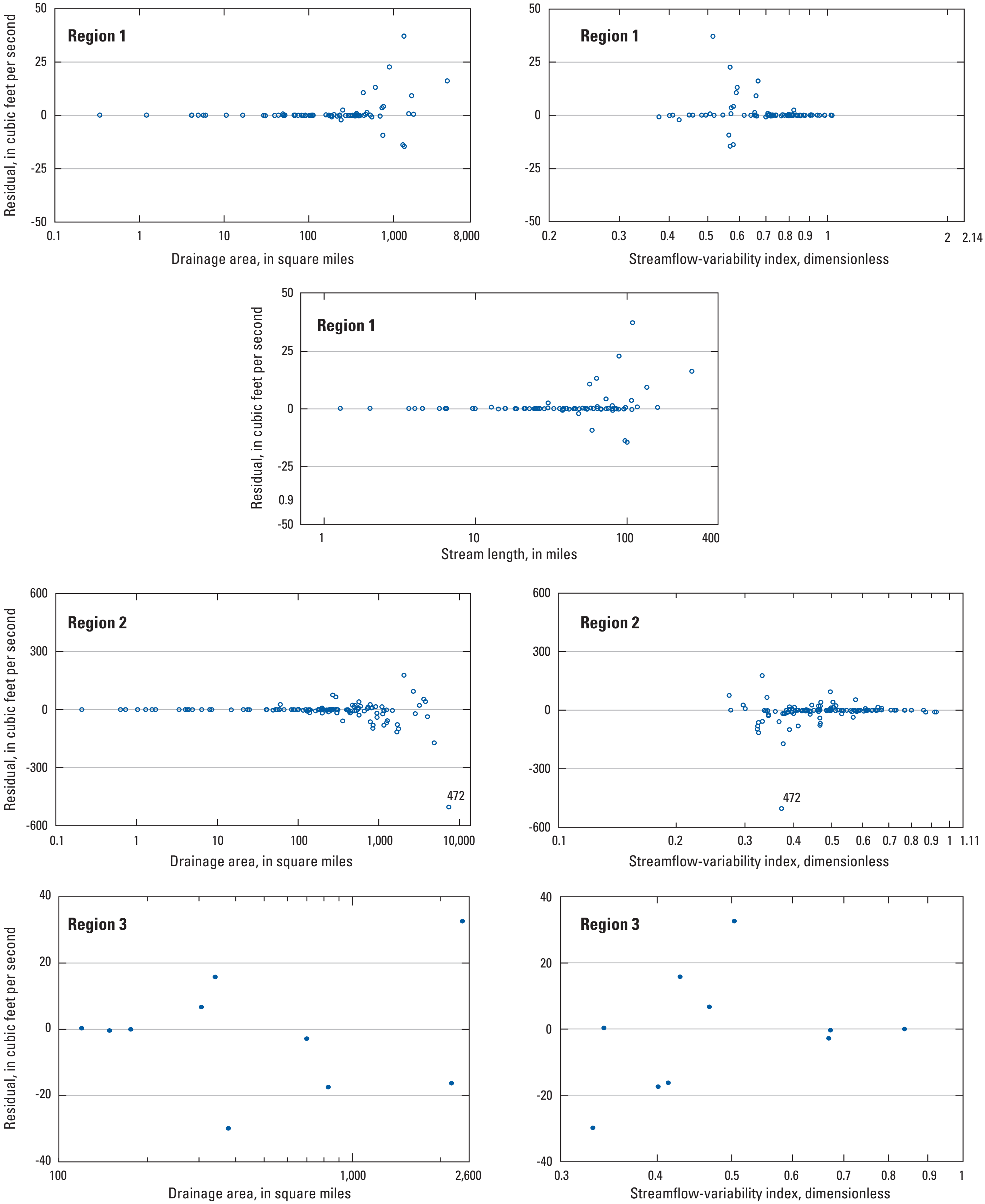


Figure 10. Relation of basin characteristics and residual from regression analyses for each region for the 7-day mean low flow for a recurrence interval of 10 years (M7D10Y) statistic.

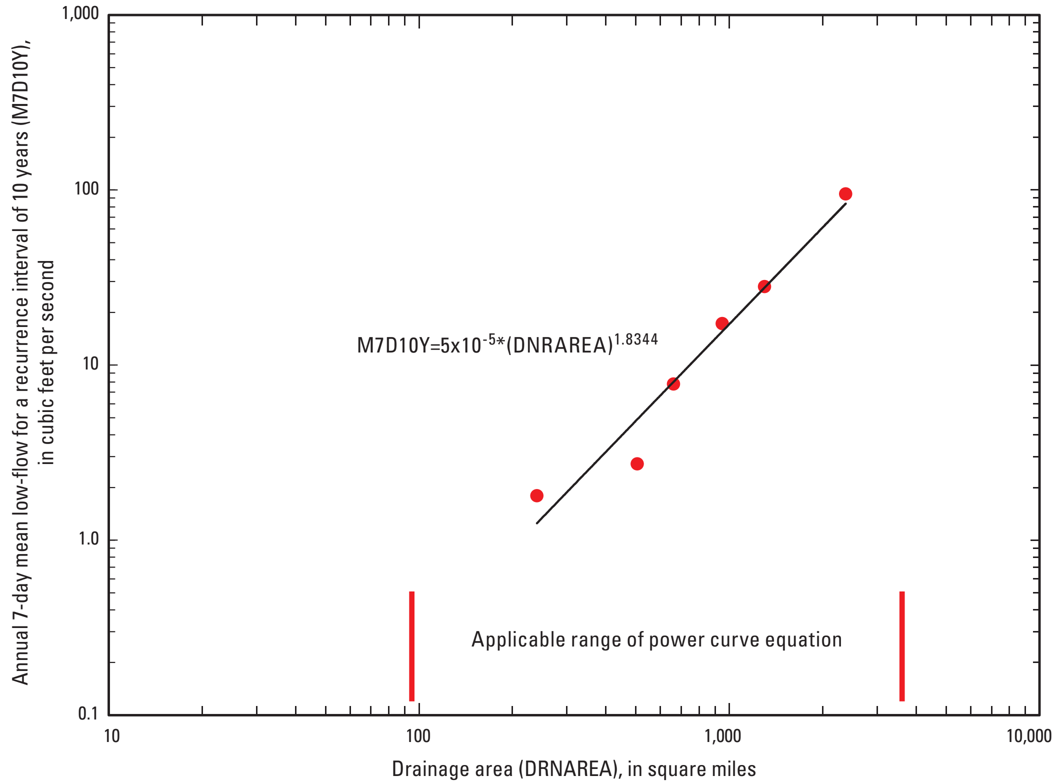


Figure 11. Application of the power curve equation method for the St. Francis River.

streamgage, equation 1 of the report may be used to compute a selected low-flow frequency statistic. The user may need to evaluate the credibility of the computed statistic if the streamgage has a minimal period of record of 10 years or if the streamgage record is affected unduly by extreme periods of drought or excessive rainfall.

The third method is the use of the regional regression equations. These equations are applicable to streams minimally affected by anthropogenic activities. The applicable range of basin characteristics for the equations for each region is listed in table 18. These equations should be used with caution for the determination of statistics at ungaged locations for which the basin characteristics are outside the range of those used to develop the regression equations. Region 1 has three basin-characteristic ranges for applying the regional equations. For Region 1, the applicable range for drainage area is from 0.34 to 4,316.33 mi^2, for longest flow path is from 1.28 to 267.8 miles, and for streamflow-variability index is from 0.377 to 1.026. Regions 2 and 3 each have two basin characteristics ranges for applying the regional equations. For Region 2, the applicable range for drainage area is from 0.21 to 7,372.43 mi^2 and for streamflow-variability index is from 0.273 to 0.925. For Region 3, the applicable range for drainage area is from 119.56 to 2,372.42 mi^2 and for streamflow-variability index is from 0.33 to 0.84.

Summary

The U.S. Geological Survey, in cooperation with the Missouri Department of Natural Resources, computed low-flow statistical data sets at selected streamgages, presented three different methods to estimate selected low-flow statistics for ungaged locations, and developed regional regression equations for use throughout Missouri. Missouri has experienced large departures from normal in precipitation and temperature in 2008 and 2012. The rainfall for 2008 was 16.58 inches above normal rainfall and in 2012 the rainfall was 10.12 inches below normal rainfall. The State of Missouri needs reliable low-flow statistics and a consistent methodology to estimate selected low-flow frequency statistics at ungaged locations to support the proper management of water resources in a sustainable manner for the benefit of water users and the environment.

As part of the low-flow study, an evaluation of streamflow during potential drought periods for 24 streamgages was done to identify historical drought periods. Annual mean streamflows during potential drought periods were compared to the mean streamflow for the period of record. Any continuous period of 3 or more years of substantially lower mean streamflows than the period of record mean streamflow was designated as a drought period. Comparing the location of the streamgages and the drought periods, the State of Missouri was divided into four similar drought areas with similar drought periods (East Central, North, Southeast, and West Central). The drought in the 1950's affected the entire State.

Table 18. Range of basin-characteristic values used to develop selected low-flow frequency regression equations for unregulated streams in Missouri.

[DRNAREA, GIS drainage area, in square miles; LFPLENGTH, length of longest flow path, in miles; STREAM_VAR, streamflow-variability index, dimensionless; NA, not applicable]

Statistic equation	DRNAREA				LFPLENGTH				STREAM_VAR			
	Minimum	Mean	Median	Maximum	Minimum	Mean	Median	Maximum	Minimum	Mean	Median	Maximum
Region 1	0.34	422.42	231.64	4316.33	1.28	54.20	45.79	267.8	0.377	0.716	0.730	1.026
Region 2	0.21	645.19	257.74	7372.43	NA	NA	NA	NA	0.273	0.496	0.467	0.925
Region 3	119.56	754.56	359.39	2372.42	NA	NA	NA	NA	0.330	0.507	0.448	0.840

The latest persistent drought to affect the State existed in East Central Missouri from 1999 to 2001. Annual mean streamflow for 2012 also was presented to compare recent streamflow conditions to historical drought periods and the period of record mean streamflow.

Statistical analyses through the 2010 water year were computed on flow data collected at 532 streamgages located in Missouri and in the neighboring States of Iowa, Nebraska, Kansas, Oklahoma, and Arkansas. Statistical analyses for streamgages included basic statistics such as minimum, maximum, and mean; monthly and annual flow durations; and trend analyses. Low-flow frequency statistics were computed for streamgages with 10 or more years of record. Basic statistics and flow-duration analyses were computed for streamgages on an annual water year period of October 1 to September 30. N-day low-flow frequency analyses were based on the climatic year that begins on April 1 and ends on March 31 of the following year. Low-flow frequencies were computed for the N-day durations of 1, 2, 3, 7, 10, 30, 60, 90, 183, and 365 days. Low-flow frequency probabilities ranged from 0.95 to 0.002.

The Kendall's *tau* test was implemented on all streamgages to determine if a trend existed in the streamgage data. For streamgages with a statistically significant trend, a variable-length record method was used to define the longest no trend period by decreasing the period of record from the beginning of the record. Trends results also were presented for the period of no trend if the streamgage had 10 or more years of record. For streamgages with trends identified, basic statistics, flow durations, and frequency values were computed for the no-trend period of time. The no-trend period as well as the period of record had to be 10 years or greater in length for frequency computations. More streamgages in Region 1 had trends than streamgages in Region 2 or 3.

Three methods are presented to estimate selected low-flow frequency statistics at ungaged locations. First, a power curve method was evaluated for streams with more than one streamgage on the same stream. In the study area, 28 streams were identified that have multiple streamgages on the same stream where a power curve equation could be developed to estimate flows on the stream at an ungaged location. For this method to be used, the drainage area of the ungaged site needs to be within 40 percent of the drainage area of the most upstream streamgage and 150 percent of the drainage area of the most downstream streamgage. Second, a drainage-area ratio method is presented and compared to the regional regression equations. The absolute percentage differences for the drainage-area ratio method were smaller than the absolute percentage differences using regional regression equations for the range of drainage-area ratios between 0.4 to 1.5. If the computed 1.5 drainage-area ratio exceeds the drainage area of the stream, the method is only applicable to the mouth of the stream. And third, a regional-regression-equation method is presented to estimate selected low-flow frequency statistics for ungaged locations on any stream with minimal anthropogenic effects. Analyses of the streamgage frequency statistics determined that a left-censored regression technique was appropriate for use in all regions of the State. Seven low-flow frequency statistics were used in the regression analyses (M1D10Y, M2D10Y, M3D10Y, M7D10Y, M10D10Y, M30D10Y, and M60D10Y). A streamflow value of 0.1 ft^3/s was used as the censoring threshold in the regression analyses for sites with flow less than 0.1 ft^3/s. The final regression equations were selected primarily on the basis of minimizing the standard error of estimate and the improvement in the standard error of estimate when considering an additional independent variable in the equation.

Basin and climatic characteristics investigated in the regression analyses were computed using geographic information system software and digital geospatial data. A total of 35 characteristics were evaluated for use in Missouri and were classified into four classes: morphometric, hydrologic, pedologic/geologic/land-use, and climatic. The Efroymson stepwise-selection method was used to define potential explanatory variables from the list of 35 characteristics. A preliminary statewide ordinary-least squares regression was implemented using the potential explanatory variables from the Efroymson analysis.

The preliminary regression analyses indicated that the physiographic provinces of Central Lowlands (Region 1), Ozark Plateaus (Region 2), and the Mississippi Alluvial Plain (Region 3) affect the low-flow characteristics of local streams. Flow-duration curves for streams in Region 2 are, in general, flatter than streams in Region 1 or 3 because of the probable base-flow contribution from springs. Regression analyses for Region 1 indicated that characteristics drainage area (DRNAREA), longest flow length (LFPLENGTH), and streamflow-variability index (STREAM_VAR) were the most statistically significant variables for regression analyses. For Region 2, regression analyses indicated that DRNAREA and STREAM_VAR were the most significant independent variables. And for Region 3, DRNAREA and STREAM_VAR were the most significant independent variables.

Low-flow frequency equations were determined for the 10-year frequency and N-day durations of 1, 2, 3, 7, 10, 30, and 60 days. The standard error of estimate was the lowest for Region 2 and ranged from 48.2 to 72.1 percent. The range in standard error of estimate was 79.6 to 94.2 percent for Region 1, and 48.1 to 96.2 percent for Region 3. Of the 207 USGS streamgages used in the regression analyses, 77 were used in Region 1, 120 were used in Region 2, and 10 were used in Region 3. Streamgages outside of Missouri were used to extend the range of data used for the independent variables. The limits for the use of these equations are based on the ranges of the characteristics used as independent variables and that streams must be affected minimally by anthropogenic activities.

References Cited

Ahearn, E.A., 2010, Regional regression equations to estimate flow-duration statistics at ungaged stream sites in Connecticut: U.S. Geological Survey Scientific Investigations Report 2010–5052, 45 p. (Also available at *http://pubs.usgs.gov/sir/2010/5052/*.)

Alexander, T.W., and Wilson, G.L., 1995, Techniques for estimating the 2- to 500-year flood discharges on unregulated streams in rural Missouri: U.S. Geological Survey Water-Resources Investigations Report 95–4231, 33 p.

Branner, G.C., 1937, Data on springs in Arkansas: Little Rock, Arkansas, Arkansas Geological Survey, 163 p.

Cavalieri, P., Jayawickrama, J., Luca, R., Patetta, M., Scott, K., and Walsh, S., 2000, Statistics I—Introduction to ANOVA, Regression, and Logistic Regression: Cary, North Carolina, SAS Institute, Inc., 504 p.

Cohn, T.A., 1988, Adjusted maximum likelihood estimation of the moments of lognormal populations from type I censored samples: U.S. Geological Survey Open-File Report 88–350, 34 p.

Eash, D.A., and Barnes, K.K., 2012, Methods for estimating selected low-flow frequency statistics and harmonic mean streamflows for streams in Iowa: U.S. Geological Survey Scientific Investigations Report 2012–5171, 94 p.

Efroymson, M.A., 1960, Multiple regression analysis, *in* Ralston, A., and Wilf, H.S., eds., Mathematical methods for digital computers: New York, John Wiley and Sons, Inc., p. 191–203.

Environmental Systems Research Institute, Inc., 2001, ArcGIS geostatistical analyst, Statistical tools for data exploration, modeling, and advanced surface generation: Redlands, California, Environmental Systems Research Institute, Inc., 23 p., accessed April 7, 2011, at *http://www.esri.com/library/whitepapers/pdfs/geostat.pdf*.

Environmental Systems Research Institute, Inc., 2009, ArcGIS desktop help: accessed April 7, 2011, at *http://webhelp.esri.com/arcgisdesktop/9.3*.

Fenneman, N.M., 1938, Physiography of eastern United States: New York, McGraw-Hill, 714 p.

Freund, R.J., and Littell, R.C., 2000, SAS system for regression (3rd ed.): Cary, North Carolina, 235 p.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 510 p. (Also available at *http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html*.)

Hershfield, D.M., 1961, Rainfall frequencies atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years: Washington, D.C., U.S. Weather Bureau,Technical Paper No. 40, 115 p.

Holmes, R.R., Jr., Koenig, T.A., and Karstensen, K.A., 2010, Flooding in the United States Midwest, 2008: U.S. Geological Survey Professional Paper 1775, 64 p.

Hortness, J.E., 2006, Estimating low-flow frequency statistics for unregulated streams in Idaho: U.S. Geological Survey Scientific Investigations Report 2006–5035, 31 p. (Also available at *http://pubs.usgs.gov/sir/2006/5035/*.)

Imes, J.L., and Emmet, L.F., 1994, Geohydrology of the Ozark Plateaus aquifer system in parts of Missouri, Arkansas, Oklahoma, and Kansas: U.S. Geological Survey Professional Paper 1414-D, 127 p.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency: Reston, Virginia, Hydrology Subcommittee Bulletin 17B, 28 p. and appendixes, accessed April 7, 2011, at *http://water.usgs.gov/osw/bulletin17b/dl_flow.pdf*.

Kleeschulte, M.J., ed., 2008, Hydrologic investigations concerning lead mining issues in southeastern Missouri: U.S. Geological Survey Scientific Investigations Report 2008–5140, 238 p.

Koltun, G.F., and Schwartz, R.R., 1986, Multiple-regression equations for estimating low flows at ungaged stream sites in Ohio: U.S. Geological Survey Water-Resources Investigations Report 86–4354, 39 p., 6 pls. (Also available at *http://pubs.er.usgs.gov/publication/wri864354*.)

Koltun, G.F., and Whitehead, M.T., 2002, Techniques for estimating selected streamflow characteristics of rural, unregulated streams in Ohio: U.S. Geological Survey Water-Resources Investigations Report 02–4068, 50 p. (Also available at *http://oh.water.usgs.gov/reports/wrir/wrir02-4068.pdf*.)

Kroll, C.N., and Stedinger, J.R., 1996, Estimation of moments and quantiles using censored data: Water Resources Research, v. 32, no. 4, p. 1005–1012, accessed April 8, 2011, at *http://www.esf.edu/ere/kroll/estimation_of_moments_cencosred_data.pdf*.

Kroll, C.N., and Vogel, R.M., 2002, Probability distribution of low streamflow series in the United States: Journal of Hydrologic Engineering, v. 7, no. 2, p. 137–146, accessed April 8, 2011, at *http://engineering.tufts.edu/cee/people/vogel/publications/probability-low.pdf*.

Lewis, J.M., and Esralew, R.A., 2009, Statistical summaries of streamflow in and near Oklahoma through 2007: U.S. Geological Survey Scientific Investigations Report 2009–5032, 36 p.

Lorenz, D.L., and others, 2011, USGS library for S–PLUS for windows—Release 4.0: U.S. Geological Survey Open-File Report 2011–1130, accessed February 7, 2012, at *http://water.usgs.gov/software/S-PLUS/*.

Lumb, A.M., Kittle, J.L., Jr., and Flynn, K.M., 1990, User's manual for ANNIE, a computer program for interactive hydrologic analyses and data management: U.S. Geological Survey Water-Resources Investigations Report 89–4080, 236 p. (Also available at *http://water.usgs.gov/software/ANNIE/*.)

Missouri Department of Natural Resources, 2007, Missouri environmental geology atlas: Rolla, Missouri, Division of Geology and Land Survey, CD-ROM.

Multi-Resolution Land Characteristics Consortium (MRLC), 2012, National Land Cover Database (NLCD): accessed May 1, 2012, at *http://www.mrlc.gov/index.php*.

National Climatic Data Center, 2012, Climate of Iowa, accessed April 30, 2012, at *http://www.crh.noaa.gov/images/dvn/downloads/Clim_IA_01.pdf*.

National Climatic Data Center, 2013, Climatological Rankings, accessed January 20, 2013, at *http://www.ncdc.noaa.gov/temp-and precip/ranks.php?periods%5B%5D=12&parameter=pcp&year=2008&month=12&state=23&div=0*.

National Climatic Data Center, 2013, Climatological Rankings, accessed January 20, 2013, at *http://www.ncdc.noaa.gov/temp-and-precip/ranks.php?periods%5B%5D=12&parameter=pcp&year=2012&month=12&state=23&div=0*.

Natural Resources Conservation Service, 2012, Geospatial data gateway: U.S. Department of Agriculture, accessed March 30, 2012, at *http://datagateway.nrcs.usda.gov/*.

Office of Missouri Governor, 2012, Executive Order 12-07, accessed November 20, 2012, at *http://governor.mo.gov/newsroom/2012/Gov_Nixon_declares_state_of_emergency_because_of_concerns_from_prolonged_severe_heat_and_drought*.

Price, C.V., Nakagaki, Naomi, and Hitt, K.J., 2010, National Water-Quality Assessment (NAWQA) area-characterization toolbox, release 1.0: U.S. Geological Survey Open-File Report 2010–1268, accessed November 20, 2012 at *http://pubs.usgs.gov/of/2010/1268*.

PRISM Climate Group, 2008, Normal annual precipitation grid for the conterminous United States, accessed March 20, 2012, at *http://www.prism.oregonstate.edu/state_products/maps.phtml?id=US*, also see *http://www.prism.oregonstate.edu/pub/prism/docs/prisguid.pdf*.

Riggs, H.C., 1972, Low-flow investigations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. B1, 18 p., accessed April 7, 2011, at *http://pubs.usgs.gov/twri/twri4b1/index.html*.

Ries, K.G., and Friesz, P.J., 2000, Methods for estimating low-flow statistics for Massachusetts streams: U.S. Geological Survey Water-Resources Investigations Report 00–4135, 81 p. (Also available at *http://pubs.usgs.gov/wri/wri004135/*.)

Searcy, J.M., 1959, Flow-duration curves, manual of hydrology—Part 2, low-flow techniques: U.S. Geological Survey Water-Supply Paper 1542–A, 33 p. (Also available at *http://pubs.er.usgs.gov/publication/wsp1542A*.)

Simley, J.D., and Carswell, W.J., Jr., 2009, *The National Map*—Hydrography: U.S. Geological Survey Fact Sheet 2009–3054, 4 p. (Also available at *http://pubs.usgs.gov/fs/2009/3054/*, also see *http://nhd.usgs.gov/*.)

Skelton, J., 1966, Low-flow characteristics of Missouri streams: Missouri Geological Survey and Water Resources, Water Resources Report 20, 95 p.

Skelton, J., 1970, Base flow recession characteristics and seasonal low-flow frequency characteristics for Missouri streams: Missouri Geological Survey and Water Resources, Water Resources Report 25, 43 p.

Skelton, J., 1974, Estimating low-flow frequency for perennial Missouri Ozarks streams: U.S. Geological Survey Water-Resources Investigations Report 73–59, 19 p.

Skelton, J., 1976, Missouri stream and springflow characteristics: low-flow frequency and flow duration: Missouri Department of Natural Resources, Geological Survey, Water Resources 32, 76 p.

Skelton, J., and Homyk, A., 1970, A proposed streamflow data program for Missouri: U.S. Geological Survey Open-File Report, 77 p.

TIBCO Software Inc., 2008, TIBCO Spotfire S+ 8.1 for Windows® User's Guide: Palo Alto, California, 582 p., accessed May 1, 2012, at *http://www.msi.co.jp/splus/support/download/V81/getstart.pdf*.

U.S. Environmental Protection Agency, 2009, BASINS 4.0 Climate Assessment Tool (CAT)—Supporting documentation and user's manual (final report): Washington, D.C., U.S. Environmental Protection Agency, EPA/600/R-08/088F.

U.S. Environmental Protection Agency, 2012, Water: DFLOW, accessed November 20, 2012, at *http://water.epa.gov/scitech/datait/models/dflow/flow101.cfm*.

U.S. Geological Survey and U.S. Department of Agriculture, Natural Resources Conservation Service, 2009, Federal guidelines, requirements, and procedures for the National Watershed Boundary Dataset: U.S. Geological Survey Techniques and Methods 11–A3, 55 p. (Also available at *http://pubs.usgs.gov/tm/tm11a3/*, also see *http://datagateway.nrcs.usda.gov/*).

U.S. Geological Survey, 2012, National hydrography dataset: U.S. Geological Survey, accessed April 30, 2012, at *http://nhd.usgs.gov/*.

U.S. Geological Survey, 2011, National elevation dataset: U.S. Geological Survey, accessed April 30, 2012, at *http://ned.usgs.gov/*.

U.S. Geological Survey, 2011, National Water Information System: U.S. Geological Survey, accessed April 1, 2011, at *http://waterdata.usgs.gov/nwis/sw*.

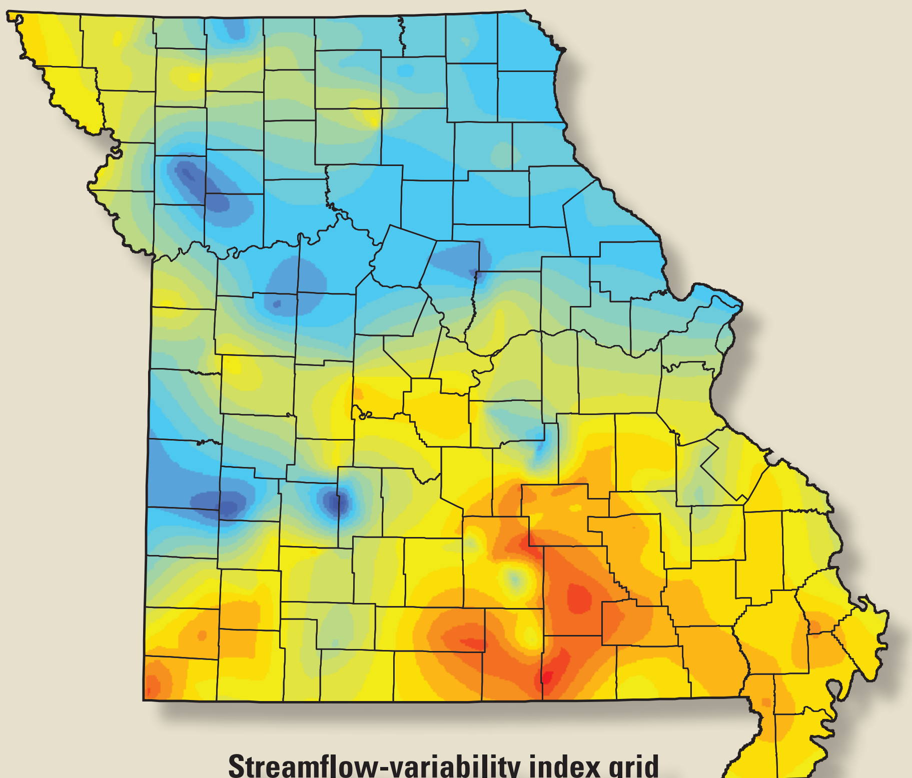
U.S. Geological Survey, 2012, The National Map Viewer, accessed November 20, 2012, at *http://viewer.nationalmap.gov/viewer/*.

Waite, L.A., 1991, National water summary, 1988–89—Hydrologic events and floods and droughts: U.S. Geological Survey Water-Supply Paper 2375, 591 p.

Publishing support provided by:
Rolla Publishing Service Center

For more information concerning this publication, contact:
Director, USGS Missouri Water Science Center
1400 Independence Road
Rolla, MO 65401
(573) 308–3667

Or visit the Missouri Water Science Center Web site at:
http://mo.water.usgs.gov/



Streamflow-variability index grid

Southard—**Low-Flow Frequency Statistics at Streamgages and Methods for Estimating Statistics at Ungaged Locations in Missouri**—SIR 2013–5090

ISBN 978-1-4113-3594-3