Prepared in cooperation with the Kansas Department of Transportation and Federal Emergency Management Agency

Methods for Estimating Annual Exceedance-Probability Streamflows for Streams in Kansas Based on Data Through Water Year 2015

Scientific Investigations Report 2017–5063

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

Cover. Flooding of the Verdigris River near Independence, Kansas, on July 7, 2007. Photograph by Chris Moehring.

Methods for Estimating Annual Exceedance-Probability Streamflows for Streams in Kansas Based on Data Through Water Year 2015

By Colin C. Painter, David C. Heimann, and Jennifer L. Lanning-Rush

Prepared in cooperation with the Kansas Department of Transportation and
Federal Emergency Management Agency

Scientific Investigations Report 2017–5063

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

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

U.S. Geological Survey
William H. Werkheiser, Acting Director

U.S. Geological Survey, Reston, Virginia: 2017

Suggested citation:
Painter, C.C., Heimann, D.C., and Lanning-Rush, J.L., 2017, Methods for estimating annual exceedance-probability streamflows for streams in Kansas based on data through water year 2015: U.S. Geological Survey Scientific Investigations Report 2017–5063, 20 p., https://doi.org/10.3133/sir20175063.

ISSN 2328-0328 (online)

Contents

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

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

Supplemental Data

A water year is the 1-year period that begins October 1 and ends September 30 and is designated by the calendar year in which the period ends.

Abbreviations

AEP	annual exceedance probability
A_g	drainage area for a streamgage
A_u	drainage area for an ungaged site
CONTDA	contributing drainage area
CSG	crest-stage gage
CSL1085LFP	slope based on change in elevation divided by length between points 10 and 85 percent of distance along the longest flow path to the basin divide
DAR	drainage area ratio
DEM	digital elevation model
DRNAREA	total drainage area
ELEV	mean basin elevation
ELEVMAX	maximum basin elevation
EMA	expected moments algorithm annual exceedance-probability analysis
GIS	geographic information system
GLS	generalized least-squares regression
LFPLENGTH	length of longest flow path
LPIII	log-Pearson Type III
MGB	multiple Grubbs-Beck
MINBELEV	minimum basin elevation
MSE	mean-square error
NHD	National Hydrography Dataset
NWIS	National Water Information System
OLS	ordinary least squares
PILF	potentially influential low flow
PRECIPRIS90	mean annual 1961–90 precipitation
PRECIPRIS00	mean annual 1971–2000 precipitation
PRECIPRIS10	mean annual 1981–2010 precipitation
PRISM	Parameter Elevation Regressions on Independent Slopes Model
Pseudo-R^2	pseudo coefficient of determination
$Q_{50\%}$	annual exceedance-probability streamflow of 50 percent (2-year recurrence-interval flood streamflow)
$Q_{20\%}$	annual exceedance-probability streamflow of 20 percent (5-year recurrence-interval flood streamflow)
$Q_{10\%}$	annual exceedance-probability streamflow of 10 percent (10-year recurrence-interval flood streamflow)

$Q_{4\%}$	annual exceedance-probability streamflow of 4 percent (25-year recurrence-interval flood streamflow)
$Q_{2\%}$	annual exceedance-probability streamflow of 2 percent (50-year recurrence-interval flood streamflow)
$Q_{1\%}$	annual exceedance-probability streamflow of 1 percent (100-year recurrence-interval flood streamflow)
$Q_{0.5\%}$	annual exceedance-probability streamflow of 0.5 percent (200-year recurrence-interval flood streamflow)
$Q_{0.2\%}$	annual exceedance-probability streamflow of 0.2 percent (500-year recurrence-interval flood streamflow)
R^2	coefficient of determination
RELIEF	basin relief
RMSE	root mean square error
RRE	rural regression equation
SD	standardized distance
S_p	standard error of prediction
SME	standard model error
SOILPERM	mean soil permeability
SSURGO	Soil Survey Geographic database
SSURGOCLAY	mean clay content
SSURGOKSAT	mean saturated hydraulic conductivity
SSURGSAND	mean sand content
SSURGSILT	mean silt content
USGS	U.S. Geological Survey
WREG	weighted-multiple-linear regression

Methods for Estimating Annual Exceedance-Probability Streamflows for Streams in Kansas Based on Data Through Water Year 2015

By Colin C. Painter, David C. Heimann, and Jennifer L. Lanning-Rush

Abstract

A study was done by the U.S. Geological Survey in cooperation with the Kansas Department of Transportation and the Federal Emergency Management Agency to develop regression models to estimate peak streamflows of annual exceedance probabilities of 50, 20, 10, 4, 2, 1, 0.5, and 0.2 percent at ungaged locations in Kansas. Peak streamflow frequency statistics from selected streamgages were related to contributing drainage area and average precipitation using generalized least-squares regression analysis. The peak streamflow statistics were derived from 151 streamgages with at least 25 years of streamflow data through 2015. The developed equations can be used to predict peak streamflow magnitude and frequency within two hydrologic regions that were defined based on the effects of irrigation. The equations developed in this report are applicable to streams in Kansas that are not substantially affected by regulation, surface-water diversions, or urbanization. The equations are intended for use for streams with contributing drainage areas ranging from 0.17 to 14,901 square miles in the nonirrigation effects region and, 1.02 to 3,555 square miles in the irrigation-affected region, corresponding to the range of drainage areas of the streamgages used in the development of the regional equations.

Introduction

Peak streamflow magnitude and frequency probability estimates are important for the engineering design of hydraulic structures, geomorphological analyses of streams, and in ecological applications (Junk and others, 1989). Determination of long-term peak streamflow information is conducted and maintained by the U.S. Geological Survey (USGS) for current (2017) and discontinued streamgage locations throughout Kansas. These streamgage locations, however, represent a small part of the total stream reaches throughout the State. In order to estimate peak streamflow characteristics at locations without an existing streamgage, the available long-term streamflow record can be used to develop regression models to predict the magnitude and frequency of peak flows. A study was conducted by the USGS in cooperation with the Kansas Department of Transportation and Federal Emergency Management Agency using drainage basin characteristics to develop regression models to estimate peak flows of various annual exceedance probabilities (AEPs; 50, 20, 10, 4, 2, 1, 0.5, and 0.2 percent) of occurrence in any year. Estimated streamflow statistics from these equations may be used in the assessment of existing transportation structures and in the proper design of new or replacement structures.

Purpose and Scope

The purpose of this report is to present updated regression model results (Rasmussen and Perry, 2000) for the prediction of AEP streamflows at locations without streamgages on streams in Kansas. Regression models were generated using annual peak streamflows from gaged rivers in Kansas and surrounding States that were unregulated and without urbanization. The resulting models may be used to compute peak streamflow magnitude for eight selected AEPs of 50, 20, 10, 4, 2, 1, 0.5, and 0.2 percent ($Q_{50\%}$, $Q_{20\%}$, $Q_{10\%}$, $Q_{4\%}$, $Q_{2\%}$, $Q_{1\%}$, $Q_{0.5\%}$, and $Q_{0.2\%}$), which are equivalent to recurrence intervals of 2, 5, 10, 25, 50, 100, 200, and 500 years, respectively. Sites with at least 25 years of annual peak streamflow record were selected for the development of peak flow statistics. Peak streamflow record lengths ranged from water year 1885 to 2015 (a water year begins October 1 and ends September 30 and is designated by the calendar year in which the period ends). The contributing drainage area of selected streamgages ranged from 0.17 to 14,910 square miles (mi^2). Peak streamflow frequency statistics from selected streamgages were related to basin characteristics using generalized least squares (GLS) regression analysis. The developed equations can be used to predict peak streamflow magnitude and frequency at ungaged locations. The regression equations presented in this report can be included in the Kansas StreamStats tool, a Web-based map tool that provides a graphical means of interactively selecting a location and automatically calculating the associated peak streamflow statistics among other physical and hydrological basin characteristics (https://streamstats.usgs.gov/; Ries, 2007; Ries and others, 2004, 2008).

Description of Study Area

The study area includes the State of Kansas, an area of about 82,000 mi^2 and a 50-mile buffer extending into Nebraska, Missouri, Oklahoma, and Colorado (fig. 1). The physical geographic characteristics and climatic characteristics vary greatly across the study area resulting in a varied hydrologic response. Rivers and streams in Kansas generally flow from west to east in following with topography. Land use is predominantly agricultural, with pasture plus grassland (48.9 percent) and cropland (43.2 percent) accounting for most of the land use in the State (Fry and others, 2011). Forest land cover accounts for 3.8 percent and urban land use accounts for about 1.5 percent of the State (Fry and others, 2011) and, therefore, streams in Kansas generally are unaffected by urbanization. The major physiographic divisions in Kansas, roughly dividing the State in half, include the Great Plains physiographic province in the west and Central Lowlands in the east (Fenneman, 1946). Terrain in these provinces is diverse and includes flat plains, rolling hills, sandhills, and steep slopes (Moody and others, 1986).

The large west to east precipitation gradient in Kansas (fig. 2) is typical of the transitional Great Plains region of the central United States (Goodin and others, 2004). Mean annual precipitation amounts along the western border are only one-third of those in the southeast corner of the State. The Rocky Mountains cause a rain shadow effect, which produces semi-arid and arid conditions in the western half of the state. Humid air from the Gulf of Mexico, particularly in the spring and summer months, results in greater precipitation in the eastern part of the State (Goodin and others, 2004). Irrigation and associated groundwater withdrawals are common in the area west of the 25-inch precipitation contour line, which nearly divides the State in half (Goodin and others, 2004). These

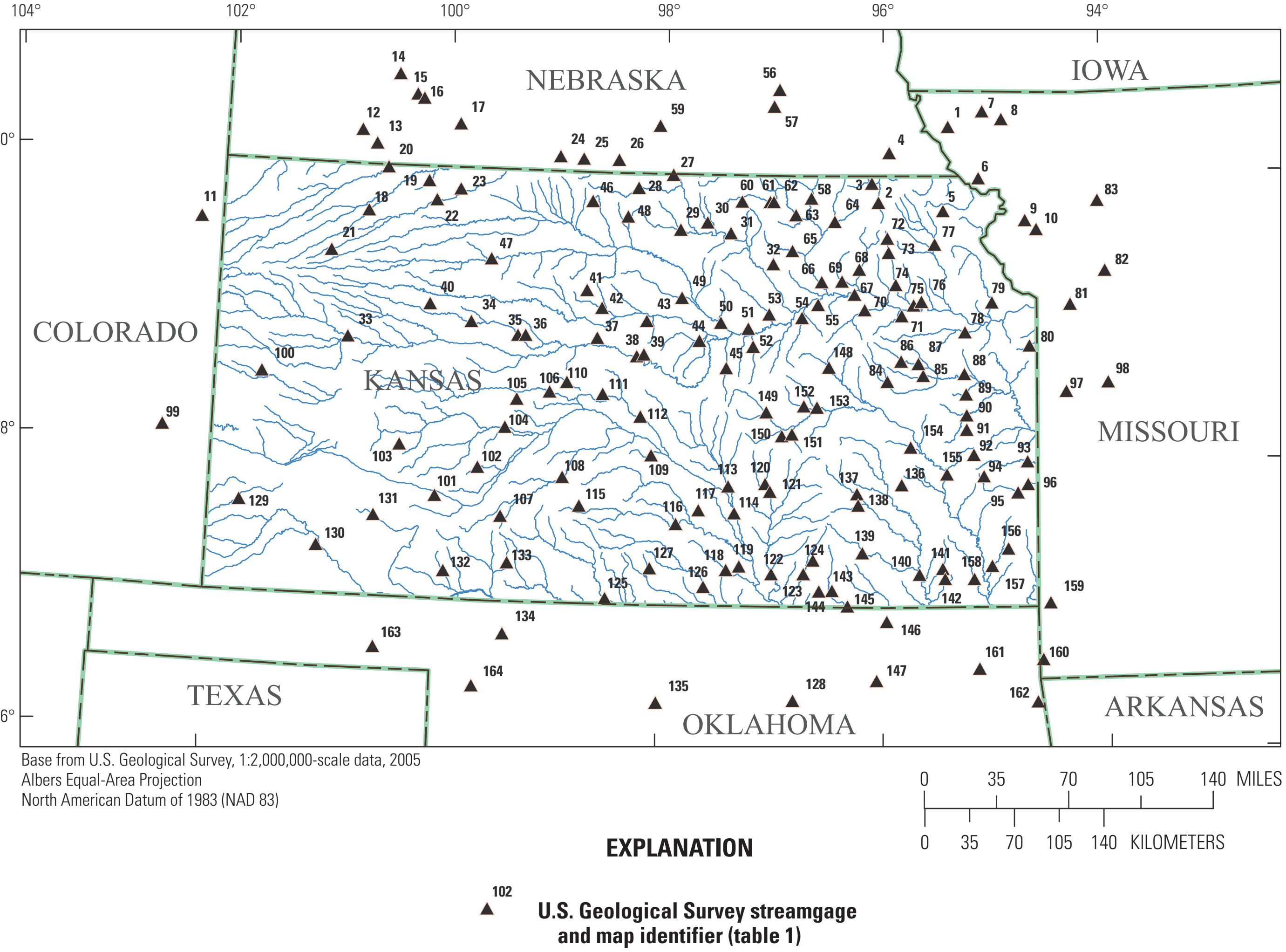
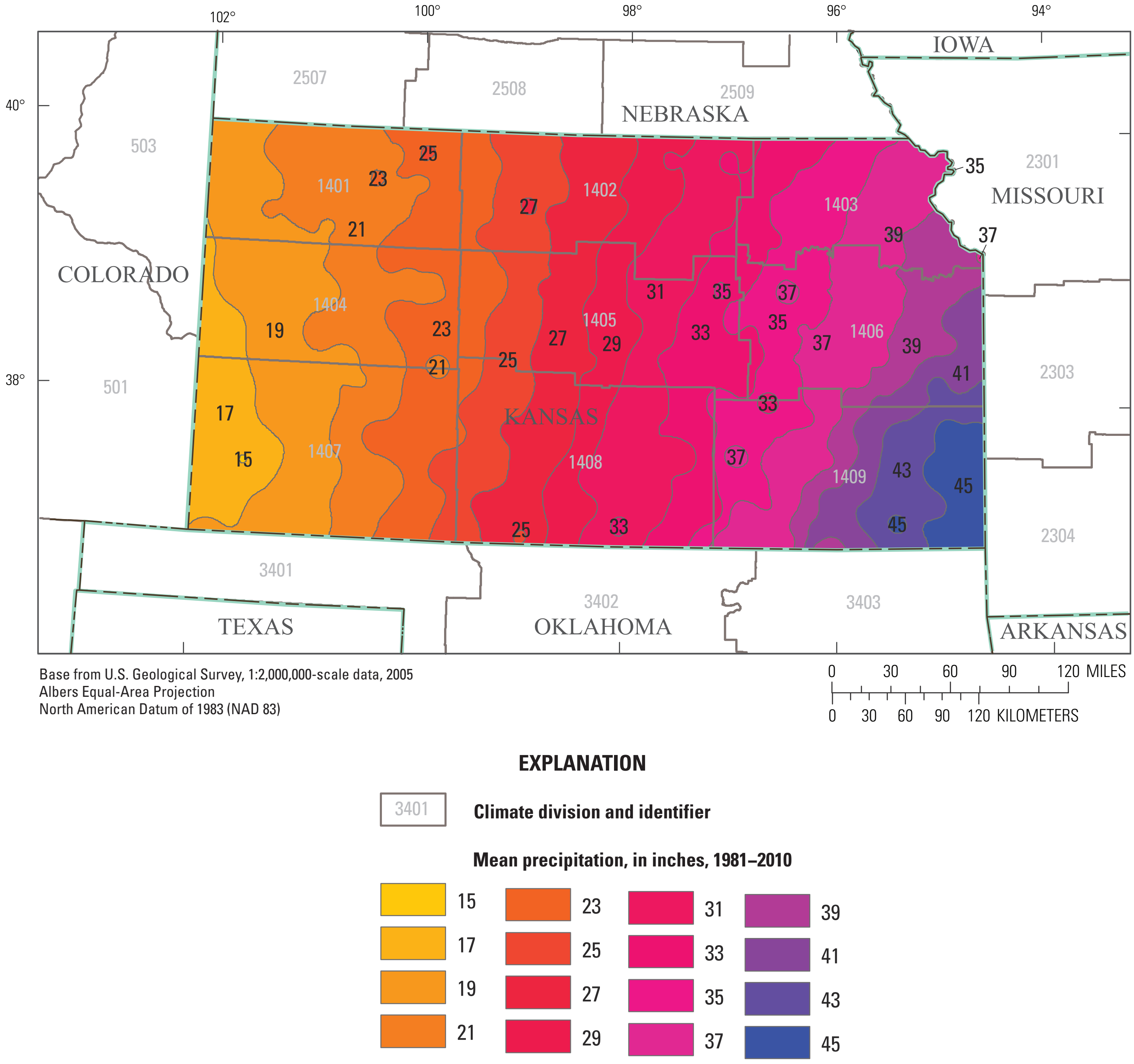


Figure 1. Selected streamgages in Kansas and surrounding States used in peak streamflow frequency statistics analyses.



Base from U.S. Geological Survey, 1:2,000,000-scale data, 2005
Albers Equal-Area Projection
North American Datum of 1983 (NAD 83)

Figure 2. Long-term mean (1981–2010) precipitation distribution in Kansas.

withdrawals are substantial and have the potential to affect surface waters in this part of the State including the magnitude and frequency of peak flows (Sophocleous and Wilson, 2000; Rasmussen and Perry, 2001).

The west to east gradient in physiographic and precipitation characteristics across Kansas also results in a west to east increase in mean annual surface-water runoff across the State. Mean annual runoff ranges from less than 0.2 inches near the Kansas-Colorado border to greater than 8 inches near the Kansas-Missouri border (Gebert and others, 1985). Streamflows at streamgages along the gradient have a consistent temporal hydrologic distribution with highest mean-monthly flows in the spring and early summer (Moody and others, 1986).

Previous Studies

Several previous studies have determined the magnitude and frequency of peak flows in Kansas. Ellis and Edelen (1960) used the index-flood method to estimate flood-frequency recurrence statistics for Kansas streams based on available peak flow data from 138 stations through 1956. Drainage areas of stations used in the analyses ranged from 111 to 45,000 mi^2 and developed relations were suitable for drainage areas greater than 150 mi^2. The product included estimates of peak flows corresponding to AEPs of 100 to 2 percent (recurrence intervals of 1 to 50 years).

Irza (1966) estimated the magnitude of floods with AEPs of 83, 43, 20, and 10 percent (recurrence intervals of 1.2, 2.33, 5, and 10 years). Peak streamflow data from 95 stations in Kansas and corresponding basin characteristics were used to develop multiple linear regression models to predict peak flows at ungaged sites. Basin factors included in the regression models were drainage area, channel slope, and the average number of wet days per year (days with greater than one inch of precipitation). The generated flood-frequency statistics filled in the data prediction gap for small drainage area sites documented in Ellis and Edelen (1960) because the input dataset included 8 years of record collected at 75 stations whose contributing-drainage areas ranged from 0.41 to 72.0 mi^2.

Patterson (1964) and Matthai (1968) used the index-flood method to estimate flood magnitudes in regional studies of the lower Mississippi River Basin and the Missouri River Basin downstream from Sioux City, Iowa, including Kansas. Hedman and others (1974) investigated the relation between active channel geometry of Kansas streams and mean flow and flood magnitudes. Equations were developed to predict flood-frequency characteristics of 50 to 1 percent AEP (2- to 100-year recurrence interval) flows using the active channel width, precipitation, and drainage area.

Jordan and Irza (1975) developed statewide regression equations for Kansas to determine flood magnitudes and frequencies using available data through 1972. The log-linear equations used contributing-drainage area and 2-year, 24-hour rainfall to estimate floods in unregulated drainages from 0.4 to 10,000 mi^2 with AEPs of 50, 20, 10, 2, and 1 percent.

Clement (1987) developed weighted least squares regression models using contributing drainage area, soil permeability, main-channel slope, and basin shape as the independent variables for determining the 50-, 20-, 10-, 4-, 2-, and 1-percent AEPs (2-, 5-, 10-, 25-, 50-, and 100-year recurrence interval) flows on unregulated streams. Data from 245 streamgages with at least 10 years of record through 1983 and drainage areas of 0.17 to 10,000 mi^2 were used to generate the models.

Rasmussen and Perry (2000) developed GLS regression models for estimating peak flow AEPs of 50, 20, 10, 4, 2, 1, and 0.5 for unregulated rural streams in Kansas using contributing drainage area, mean-annual precipitation, soil permeability, and slope of the main channel. Data from 253 streamgages with a minimum of 10 years of record and drainage areas ranging from 0.17 to 9,100 mi^2 were used in the development of the models and streamflow record extending through the 1997 water year. Perry and others (2004) used the equations determined from Rasmussen and Perry (2000) and interpolation from computed values at gage locations to determine peak streamflows for AEPs of 50, 20, 10, 4, 2, and 1 percent at 4,771 stream locations in Kansas.

Streamgage Selection and Data Analyses

The determination of peak streamflows for this study began with the selection of long-term streamgages in Kansas and surrounding States. The list of streamgages was filtered to eliminate redundant and nested streamgages and, therefore, to ensure that the stations used in the analyses were independent datasets. Peak streamflow frequency analyses were conducted to determine station skews and to test for significant temporal trends in peak streamflows. Stations with a long-term record were used in the determination of a generalized skew for two identified hydrologic regions. The generalized skew was weighted with the station skew to obtain a more accurate determination of peak streamflows for each region. Basin characteristics also were compiled for the streamgages to be used in regression analyses for the prediction of streamflow frequency statistics at ungaged locations.

Site Selection

Peak streamflow data used in this report were collected for 270 active and discontinued continuous-record streamgages located in Kansas and within 50 miles of the Kansas border in the surrounding States of Nebraska, Missouri, Oklahoma, and Colorado with a cumulative record extending from 1885 to 2015 (U.S. Geological Survey, 2015). Streamgages were selected with at least 25 years of annual peak flows that were unaffected by regulation, surface-water diversions, or urbanization. Sites then were screened for redundancy and nested basins before further analyses.

Redundant Sites

All selected streamgages were screened for redundancy to ensure that all sites used in flood-frequency analyses and regression analyses represented independent data points. To determine if peak streamflow data from streamgages were redundant (not independent), two characteristics were tested (1) the standardized distance (SD) of the basin centroids to determine if the basins were nested, and (2) the ratio of the basin drainage areas. The SD is defined as (Veilleux and Stedinger, 2013)

$$SD_{ij} = D_{ij}/(\sqrt{(0.5\ (DA_i + DA_j)})) \tag{1}$$

where

SD_{ij} is the standardized distance between centroids of basin i and basin j,

D_{ij} is the distance between centroids of basin i and basin j, and

DA_i and DA_j are the drainage areas at sites i and j.

The drainage area ratio (DAR) was used to determine if the basins were similar in size and was defined as (Veilleux, 2009)

$$DAR = \mathrm{MAX}\ [(DA_i/DA_j), (DA_j/DA_i)] \tag{2}$$

where

MAX is the maximum of values in brackets [ ], and

DA_i and DA_j are the drainage areas at sites i and j.

Site pairs that had an SD of less than or equal to 0.5 and a DAR of less than or equal to 5 were considered to be redundant. If the DAR value of site pairs was greater than 5, even if the SD was less than 0.5, the sites still were considered independent for the purposes of flood-frequency analyses. For site pairs with data that were considered redundant, one site from the pair was removed from the regional skew analyses. Sites from redundant data pairs were compared and those kept in analyses were selected to favor longer periods of record and smaller drainage areas. The minimum ratio for a period of record difference or drainage area difference was selected to be 1.333:1 and 0.667:1, respectively. If the periods of record or drainage areas were similar, or if one site fit the criteria of having a longer period of record but in turn had a larger drainage area, then the site with the most recent data was selected for analyses. Of the initial 270 streamgages, 106 subsequently were removed after redundancy screening, resulting in a total of 164 streamgages available for use in determination of the generalized skews for Kansas and in the flood-frequency analyses (fig. 1).

Annual Peak Streamflow Data

Annual peak streamflow data are available from USGS continuous streamgages (with or without crest-stage gages [CSG]) or from crest-stage only gages. A CSG provides information on peak stages that occurred between site visits or logged observations and is particularly useful in capturing peak stages of small (less than 100 mi^2) basins. The peak stages collected at the streamgages are converted to flows based on a stage-streamflow rating (Sauer and Turnipseed, 2010) and stored in the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2015) along with comments and codes documenting additional details of interest, if any, with each peak flow.

Peak Streamflow Frequency Analysis

The current (2017) standard methodology for the determination of flood-frequency statistics is Bulletin 17B of the U.S. Interagency Advisory Committee on Water Data (U.S. Interagency Advisory Committee on Water Data, 1982). The Bulletin 17B method fits a log-Pearson Type III (LPIII) distribution curve to the logarithms of annual peak streamflows at a given station using the method-of-moments to compute a mean, standard deviation, and station skew of the log-transformed peak streamflow data (U.S. Interagency Advisory Committee on Water Data, 1982). The user has the option to weigh the individual station skew estimate with a generalized/regional skew estimate, which typically improves the accuracy because skews tend to follow regional trends. At the time of this report, modest changes to Bulletin 17B, recommended by the Advisory Committee on Water Information (http://acwi.gov), are being drafted into Bulletin 17C. Modifications include the adoption of a generalized method-of-moments estimator, known as the expected moments algorithm (EMA) procedure (Cohn and others, 1997), and a generalized version of the Grubbs-Beck test for low outliers—the multiple Grubbs-Beck test (MGB; Cohn and others, 2013). The EMA is an updated method for fitting the LPIII frequency distribution that has been shown to be a more effective means of incorporating historical peak streamflow information into a flood-frequency analysis. The EMA can accommodate interval data, which simplifies analysis of datasets containing historic data, potentially influential low flows (PILFs), and uncertain data points while also providing enhanced confidence intervals on the estimated peak streamflows.

The USGS computer program PeakFQ (Flynn and others, 2006; Veilleux and others, 2014) version 7.1 (U.S. Geological Survey, 2014) was used to compute the flood-frequency estimates for the 164 streamgages used in the development of the generalized skew for the State of Kansas. The program automates many of the flood-frequency analyses procedures, including identifying and adjusting for high and low outliers and historical periods, and fitting the LPIII distribution to the streamflow data. The program includes the EMA procedure for flood-frequency analysis and MGB outlier screening, and the previous Bulletin 17B analysis also is still supported in the software. The Bulletin 17B analysis method utilizes systematic peaks (observed or estimated annual peaks during the systematic streamgaging program at a station) and historic peaks (peaks observed outside the range of the systematic streamgaging program; U.S. Interagency Advisory Committee on Water Data, 1982), whereas the EMA analysis method uses a more general description of the historical period, which includes both systematic and historic peaks. Flow intervals are used to describe the knowledge of the peak flow in each year and perception thresholds are used to describe the range of measurable potential streamflows in each year. Historic peaks were used to define the upper threshold of peak streamflow during periods of missing data between historic and systematic record. Flow intervals and perception thresholds are defined for every year of the historical period. Analysis results for EMA and Bulletin 17B should be the same for sites with no historical or censored data. The program also computes and reports the Kendall's tau (Kendall, 1938; 1975) parameters for the determination of monotonic trends in the systematic record. The generalized steps used in the determination of flood-frequency statistics for use in the determination of regression equations for the state of Kansas were:

1. Retrieve the annual peak streamflow data for selected streamgages from NWIS (U.S. Geological Survey, 2015);
2. Plot the annual-time series to find unusual observations that require further investigation;
3. Set lower and upper flow intervals for data gaps and estimated peaks and perception thresholds;
4. Run EMA/MGB analyses in the PeakFQ software using the station skew option to obtain initial at-site flood-frequency estimates and station skews for the streamgage;
5. Review the flood-frequency curve to determine if it adequately fits the annual peak data and evaluate the PILFs when identified by using the MGB test;
6. Determine if there are statistical trends in peak flows and precipitation;
7. Plot and assess station skew values and statistical trends for spatial or regional distribution patterns;
8. Calculate generalized skew and standard error of estimate for defined hydrologic regions; and
9. Run EMA/MGB analyses in the PeakFQ software using the weighted skew option, specify the generalized skew and standard error, and obtain the final at-site flood-frequency estimates to use along with basin characteristics in the development of regression models.

Trend Analysis

A general assumption of the peak flow data used in the development of flood-frequency analyses is that the peak streamflow series are stationary (that is, the statistical characteristics such as mean, variance, and skewness do not change with time). Rasmussen and Perry (2000) determined that the use of peak streamflow data with significant temporal trends caused a substantial (-20 to 70 percent) underestimation or overestimation of the magnitude of flood-frequency estimates of Kansas streams. The peak streamflow records of stations used in the flood-frequency analyses for this study were analyzed for statistically significant temporal trends using the Kendall's test (Kendall, 1938, 1975) and a significance level of 0.05. Temporal trends in peak flows can be the result of several factors including trends in precipitation, urbanization and other changes in land use, constructed impoundments and diversions, and groundwater withdrawals. An analysis of 80 streamgages in Kansas for significant trends in peak streamflows conducted by Rasmussen and Perry (2001) determined that 10 of 13 stations (located primarily in western Kansas) with significant trends were decreasing and likely the result of groundwater withdrawals with possible contributing effects of constructed impoundments.

Precipitation Trend Analysis

Precipitation is a major factor affecting the magnitude of peak flows, and trends in annual peak streamflows could be caused by trends in annual precipitation. Annual precipitation data from the 9 climatic divisions in Kansas (National Oceanic and Atmospheric Administration, 2016) and 11 divisions in adjacent States were analyzed for statistical trends during 3 selected periods using ordinary least squares (OLS) regression analyses (Helsel and Hirsch, 2002). Analysis periods included the approximate cumulative length of available streamflow record (1895 through 2015), the most common period of streamflow record availability for Kansas streamgages (1957 through 2015), and an approximate post-irrigation period (1978–2015) for the High Plains Aquifer (Smith and others, 2015). Three significant increasing trends in precipitation were determined and all were during the 1895–2015 period, and included two climate divisions in southeast Kansas (1408, 1409) and one in northeast Oklahoma (3403; fig. 2, table 1). No significant trends in annual precipitation were observed over the most common record period (1957–2015) or the post-irrigation period of record (1978–2015) (table 1); therefore, any trends observed in streamflow during these periods are not likely to have been related to trends in annual precipitation.

Annual Peak Streamflow Trend Analysis

Results of the Kendall's test for trends in streamflow indicated that 37 of the 164 streamgages used in the study had significant trends (*p*-value less than 0.05) during the

Table 1. Summary of ordinary least squares regression analyses to determine significant trends in precipitation for climatic divisions in Kansas and selected divisions in adjacent States for 1895–2015, 1957–2015, and 1978–2015.

[*p*-value, statistical probability level; *p*-values shaded in gray are below a statistical significance level of 0.05]

	Kansas climate division																	
	1401		**1402**		**1403**		**1404**		**1405**		**1406**		**1407**		**1408**		**1409**	
Analysis period	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**	**Slope**	***p*-value**
1895–2015	0.004	0.729	0.022	0.140	0.019	0.285	0.008	0.478	0.024	0.111	0.027	0.160	0.011	0.284	0.032	0.026	0.041	0.033
1957–2015	-0.010	0.732	-0.015	0.724	-0.025	0.651	-0.009	0.735	-0.012	0.779	-0.007	0.902	-0.008	0.764	0.024	0.551	0.053	0.357
1978–2015	-0.029	0.596	-0.027	0.739	-0.053	0.595	-0.013	0.789	0.028	0.729	0.001	0.995	-0.023	0.643	0.055	0.445	0.086	0.434
	Colorado climate division																	
	0501		**0502**		**0503**		**0504**		**0505**		**0506**		**0507**		**0508**		**0509**	
1895–2015	-0.003	0.656	--	--	-0.002	0.818	--	--	--	--	--	--	--	--	--	--	--	--
1957–2015	0.007	0.717	--	--	0.008	0.719	--	--	--	--	--	--	--	--	--	--	--	--
1978–2015	-0.037	0.321	--	--	-0.005	0.901	--	--	--	--	--	--	--	--	--	--	--	--
	Missouri climate division																	
	2301		**2302**		**2303**		**2304**		**2305**		**2306**		**2307**		**2308**		**2309**	
1895–2015	0.022	0.217	--	--	0.025	0.191	0.020	0.296	--	--	--	--	--	--	--	--	--	--
1957–2015	0.035	0.529	--	--	0.071	0.223	0.098	0.096	--	--	--	--	--	--	--	--	--	--
1978–2015	0.041	0.706	--	--	0.033	0.775	0.088	0.460	--	--	--	--	--	--	--	--	--	--
	Nebraska climate division																	
	2501		**2502**		**2503**		**2504**		**2505**		**2506**		**2507**		**2508**		**2509**	
1895–2015	--	--	--	--	--	--	--	--	--	--	--	--	0.005	0.671	0.014	0.300	0.015	0.332
1957–2015	--	--	--	--	--	--	--	--	--	--	--	--	-0.020	0.656	0.005	0.891	-0.020	0.656
1978–2015	--	--	--	--	--	--	--	--	--	--	--	--	0.025	0.658	-0.003	0.963	-0.041	0.636
	Oklahoma climate division																	
	3401		**3402**		**3403**		**3404**		**3405**		**3406**		**3407**		**3408**		**3409**	
1895–2015	0.005	0.658	0.030	0.054	0.041	0.039	--	--	--	--	--	--	--	--	--	--	--	--
1957–2015	-0.011	0.686	0.018	0.702	0.054	0.361	--	--	--	--	--	--	--	--	--	--	--	--
1978–2015	-0.048	0.358	0.002	0.977	0.056	0.599	--	--	--	--	--	--	--	--	--	--	--	--

available peak streamflow record period. Of the 37 stations, 27 had decreasing trends and 10 had increasing trends. Of the 27 stations with decreasing trends in peak streamflows, 20 were located in western Kansas, western Nebraska, or western Oklahoma (fig. 3). The lack of significant decreasing trends in precipitation in any of the study climate divisions and in any of the three tested periods along with the coincident distribution of highly irrigated land use, supports previous findings that the declines in peak streamflows likely are associated with documented changes in groundwater withdrawals for irrigation use (Rasmussen and Perry, 2001; Young and others, 2005; Whittemore and others, 2015; Juracek, 2015; Smith and others, 2015). Based on these results, the further analyses of the peak streamflow records and development of regression equations were regionalized by irrigation effects. An irrigation-affected region boundary was developed using the 25-inch mean precipitation contour from the 1981 to 2010 mean precipitation contours (Natural Resources Conservation Service, 2015) as a surrogate, along with a spatial data layer depicting the irrigated land cover (Jude Kastens, Kansas Applied Remote Sensing Program, written commun., 2016), the High Plains Aquifer boundary (Fischer and McGuire, 1999), and the distribution of streamgages with significant decreasing trends in peak streamflows (fig. 3). The resulting irrigation-affected region included 43 streamgages. The peak streamflow records at streamgages within the irrigation-affected area were divided into pre-irrigation and post-irrigation affected periods. The pre-irrigation (the “natural” record period of peak streamflows) period of record was used in the development of a generalized statewide skew and regression equations for region 1 (fig. 4), whereas the post-irrigation record was used in separate generalized skew and regression equation development for the defined irrigation-affected region (region 2, fig. 4). Based on Smith and others (2015), 1978 was selected as the first year of the post-irrigation effects period. Trend analyses again were conducted on the pre-irrigation or “natural” record (pre-1978) and post-irrigation (post-1978) period of record for the streamgages within the irrigated region (table 2).

Table 2. Summary of trend analyses of peak streamflows at streamgages used in the study.

[Table available for download at https://doi.org/10.3133/sir20175063]

The Kendall’s test is sensitive to multi-year sequences of high or low peak streamflow values near the beginning or end of the analyzed record (Wahl, 1998). These multi-year sequences of unusual conditions may result in a significant trend without a systematic change being present. Using the method described in Eash and others (2013), a few annual peaks were removed from the beginning or end of the record period of those streamgages with significant trends and the datasets were reanalyzed. The number of annual peaks removed from the beginning or end of record corresponded to a maximum of 6 percent of the total record in following with Eash and others (2013). The removal of as much as 6 percent of the record period at the beginning or end of peak streamflow record resulted in the removal of a significant trend in eight streamgages (table 2). These stations with adjusted record were included in the generalized skew (if remaining record length was at least 25 years) and regression analyses and removed only if they were a high leverage or high influence point in the regression model.

Regional Skew Analysis

Separate regional or generalized skews were generated for Kansas for two hydrologic regions determined based on irrigation effects (fig. 4). Development of a regional skew for region 1 included a peak streamflow record from those streamgages with at least 25 years of pre-1978 record (pre-irrigation) for the defined irrigation region and complete record of at least 25 years for stations east of the irrigation region. Record for 120 streamgages in Kansas and surrounding States were used in the determination of the region 1 generalized skew. The generalized skew for the irrigation-affected region, region 2, was generated using stations with at least 25 years of post-1978 (post-irrigation) record and included 14 stations (1 Colorado station and 13 Kansas stations) west of the defined irrigation boundary.

The station skews generated in the EMA analyses from long-term streamgages were used to calculate a more accurate generalized skew coefficient for the State of Kansas. Three accepted methods for the development of a generalized skew are described in Bulletin 17B: (1) plot station skew coefficients on a map and construct skew isolines, (2) use regression techniques to develop a skew-prediction equation relating station skew coefficients to selected basin characteristics, or (3) use the arithmetic mean of station skew coefficients from long-term streamgages (U.S. Interagency Advisory Committee on Water Data, 1982). The station skew values were mapped using geographic information system (GIS) software in following with method 1 and there was no geographic correlation of skew throughout the study region based on a visual assessment. Regression techniques also were explored (method 2) using OLS regression techniques and combinations of drainage area, elevation, mean clay content, mean annual precipitation, latitude, and longitude, but none of these basin factors indicated a significant correlation with station skew. Generalized skew coefficients using method 3 resulted in a region 1 generalized skew coefficient of -0.125, a standard error of 0.502, and a root mean square error (RMSE) of 0.252. Similarly, the generalized skew for the irrigation-affected region (region 2) was -0.478, the standard error was 0.459, and the RMSE was 0.210. The generalized skews were weighted with the station skews to produce final peak streamflow frequency statistics (tables 3 and 4) used in the development of regression equations that can be used for predicting peak streamflow frequency statistics at ungaged locations in Kansas (tables 3 and 4).

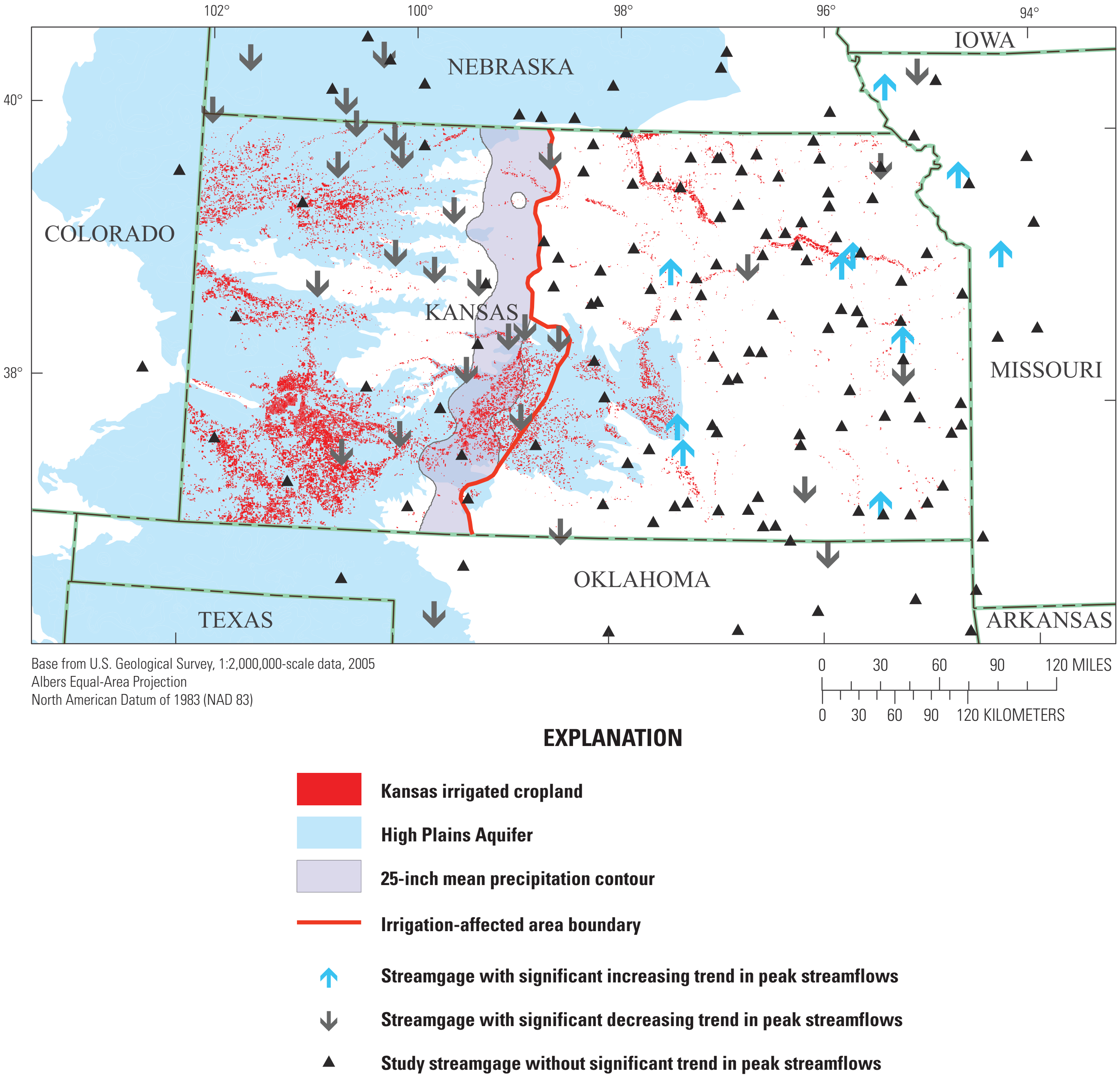


Figure 3. Distribution of streamgages with significant increasing or decreasing trends in peak streamflows.

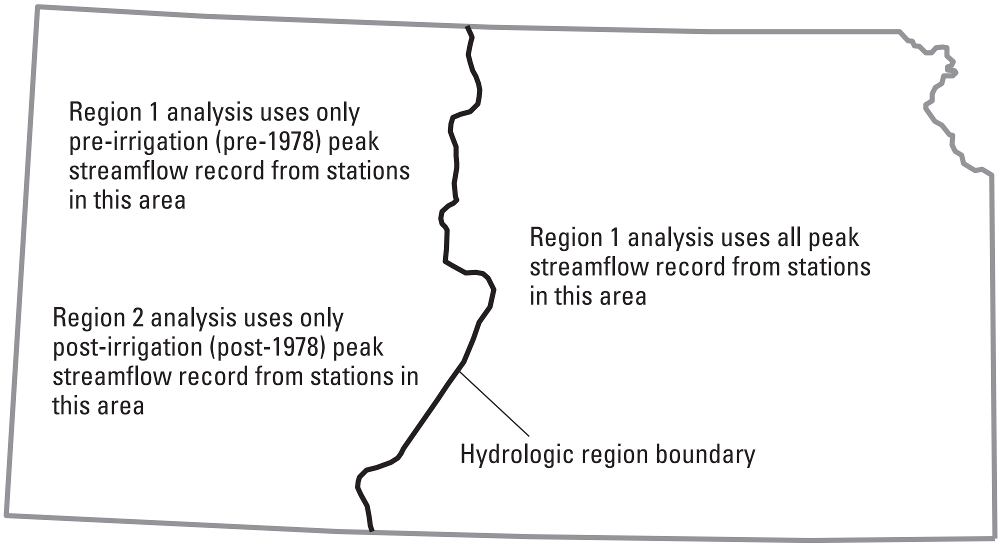


Figure 4. Hydrologic regions used in peak streamflow frequency analyses.

Table 3. Peak streamflow frequency statistics for streamgages used in the development of regression equations for hydrologic region 1 in Kansas.

[Table available for download at https://doi.org/10.3133/sir20175063]

Table 4. Peak streamflow frequency statistics for streamgages used in the development of regression equations for hydrologic region 2 in Kansas.

[Table available for download at https://doi.org/10.3133/sir20175063]

Basin Characteristics

Peak streamflows and resulting frequency statistics are affected by a number of physical (morphometric, soils) and climatic factors that define basin characteristics. The relation between basin characteristics and peak flows varies from one stream to another and from one region to another. In previous studies of Kansas peak flows (Irza, 1966; Jordan and Irza, 1975; Clement, 1987; Rasmussen and Perry, 2000), the basin characteristics determined to be significant factors in explaining the variability in peak flows were contributing drainage area; 2-year, 24-hour rainfall; mean-annual precipitation; average number of days with greater than 1 inch of precipitation; soil permeability; main-channel slope; and basin shape. Many of these factors commonly are used in the development of similar regression equations throughout the United States (Jennings and others, 1994).

For each selected streamgage used in the development of regression models, as many as 21 selected basin characteristics were computed using GIS software or obtained from established sources to be used as independent variables in peak streamflow frequency regression equations for Kansas (tables 5, 6). The terminology used in defining the basin characteristics was consistent with that used in the USGS StreamStats application (U.S. Geological Survey, 2016). Morphometric characteristics, including selected elevation (ELEV, MINBELEV, MAXELEV, RELIEF), slope (CSL1085LFP), and length (LFPLENGTH) characteristics, were derived from a USGS digital elevation model (DEM; Horizon Systems Corporation, 2010) with a 30-meter resolution. The total drainage basin area upstream from the streamgage (DRNAREA) and the contributing drainage area (CONTDA) were obtained from the USGS NWIS (U.S. Geological Survey, 2015). Soil characteristics including mean soil permeability (SOILPERM), mean saturated hydraulic conductivity (SSURGOKSAT), mean sand content (SSURGSAND), mean silt content (SSURGSILT), and mean clay content (SSURGOCLAY) were computed from the Natural Resource Conservation Service Soil Survey Geographic (SSURGO) database (Natural Resources Conservation Service, 2015). Mean annual precipitation for the 1961–90 (PRECPRIS90), 1971–2000 (PRECPRIS00), and 1981–2010 (PRECPRIS10) were obtained from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) Climate Group (Parameter-Elevation Regressions on Independent Slopes Model Climate Group, 2015).

Table 5. Basin characteristics of streamgages selected for use in the development of generalized least squares regression analyses of peak streamflow frequency statistics for Kansas.

[Table available for download at https://doi.org/10.3133/sir20175063]

Table 6. Description of terms and methods used in the development of basin characteristics for selected streamgages used in the study.

[SSURGO, soil survey geographic database; PRISM, parameter elevation regressions on independent slopes model]

Short_id	Long_id	Definition	Units	Data source
STAID	Station_ID	Station identification number	Dimensionless	U.S. Geological Survey (2015)
STANAME	Station_Name	Station name	Dimensionless	U.S. Geological Survey (2015)
LNG_GAGE	Longitude	Longitude	Decimal degrees	U.S. Geological Survey (2015)
LAT_GAGE	Latitude	Latitude	Decimal degrees	U.S. Geological Survey (2015)
LAT_CENT	Latitude_of_Basin_Centroid	Latitude of basin centroid	Decimal degrees	Horizon Systems Corporation (2010)
LONG_CENT	Longitude_of_Basin_Centroid	Longitude basin centroid	Decimal degrees	Horizon Systems Corporation (2010)
DATUM	Datum_of_Latitude_Longitude	Datum used to determine the site coordinates	Dimensionless	Horizon Systems Corporation (2010)
DRNAREA	Drainage_Area	Area that drains to a point on a stream	Square miles	Horizon Systems Corporation (2010)
CONTDA	Contributing _Drainage_Area	Contributing area that drains to a point on a stream	Square miles	U.S. Geological Survey (2015)
PRECPRIS90	Mean_Annual_Precip_PRISM_1961_1990	Basin average mean annual precipitation for 1961 to 1990 from PRISM	Inches	Parameter-Elevation Regressions on Independent Slopes Model Climate Group (2015)
PRECPRIS00	Mean_Annual_Precip_PRISM_1971_2000	Basin average mean annual precipitation for 1971 to 2000 from PRISM	Inches	Parameter-Elevation Regressions on Independent Slopes Model Climate Group (2015)
PRECPRIS10	Mean_Annual_Precip_PRISM_1981_2010	Basin average mean annual precipitation for 1981 to 2010 from PRISM	Inches	Parameter-Elevation Regressions on Independent Slopes Model Climate Group (2015)
SOILPERM	Average_Soil_Permeability	Mean soil permeability	Inches per hour	Aggregation of soils characteristics from SSURGO database (Natural Resources Conservation Service, 2015)
SSURGOKSAT	SSURGO_Saturated_Hydraulic_Conductivity	Mean saturated hydraulic conductivity	Micrometers per second	Aggregation of soils characteristics from SSURGO database (Natural Resources Conservation Service, 2015)
SSURGOCLAY	Percent_Clay_from_SSURGO	Percentage of clay in soils, from SSURGO variables of area- and depth-weighted averages of select soil attributes	Percent	Aggregation of soils characteristics from SSURGO database (Natural Resources Conservation Service, 2015)
SSRGSILT	Percent_Silt_from_SSURGO	Percentage of silt in soils, from SSURGO variables of area- and depth-weighted averages of select soil attributes	Percent	Aggregation of soils characteristics from SSURGO database (Natural Resources Conservation Service, 2015)
SSURGSAND	Percent_Sand_from_SSURGO	Percentage of sand in soils, from SSURGO variables of area- and depth-weighted averages of select soil attributes	Percent	Aggregation of soils characteristics from SSURGO database (Natural Resources Conservation Service, 2015)
I24H100Y	24_Hour_100_Year_Precipitation	Maximum 24-hour precipitation that occurs on average once in 100 years	Inches	National Oceanic and Atmospheric Administration (2015)
MINBELEV	Minimum_Basin_Elevation	Minimum basin elevation	Feet	Horizon Systems Corporation (2010)
ELEVMAX	Maximum_Basin_Elevation	Maximum basin elevation	Feet	Horizon Systems Corporation (2010)
ELEV	Mean_Basin_Elevation	Mean basin elevation	Feet	Horizon Systems Corporation (2010)
LFPLENGTH	LFP_length	Length of longest flow path	Miles	Horizon Systems Corporation (2010)
ELEV85FT	Elev_85pct_LFP_from_DEM	Elevation at 85 percent from outlet along longest flow path slope using digital elevation model	Feet	Horizon Systems Corporation (2010)
ELEV10FT	Elev_10pct_LFP_from_DEM	Elevation at 10 percent from outlet along longest flow path slope using digital elevation model	Feet	Horizon Systems Corporation (2010)
CSL1085LFP	Stream_Slope_10_and_85_Longest_Flow_Path	Slope based on change in elevation divided by length between points 10 and 85 percent of distance along the longest flow path to the basin divide	Feet per mile	Horizon Systems Corporation (2010)

Regression Models to Predict the Magnitude and Frequency of Peak Flows at Ungaged Sites

The regression model in this study, as in the previous determination of peak flow statistics for ungaged streams in Kansas by Rasmussen and Perry (2000), used GLS models of the form:

$$\mathrm{Log}_{10}Y = b_0 + b_1\log_{10}X_1 + b_2\log_{10}X_2 + \ldots\ldots + b_n\log_{10}X_n, \quad (3)$$

equivalent to:

$$Y = 10^{b0}\,(X_1^{\,b1})\,(X_2^{\,b2})\ldots\ldots(X_n^{\,bn}) \quad (4)$$

where

Y is the dependent variable (peak streamflow for selected AEP),

b_0 to b_n are the regression model coefficients, and

X_1 to X_n are independent variables (basin characteristics).

All variables were transformed to base 10 logarithms for use in the analyses except those variables that represent a percentage (SSURGOCLAY, SSURGSAND, SSURGSILT) because these variables are constrained to a range of 0 to 100.

Regression Analyses

Regression models for predicting the magnitudes of various peak streamflow AEPs were developed using the USGS computer program Weighted-Multiple-Linear Regression (WREG; Eng and others, 2009). The software was used to transform dependent (peak streamflow) and independent (basin characteristics) variables, estimate regression model coefficients, graph performance metrics, and provide quantitative model performance statistics.

A correlation matrix was developed between the $Q_{1\%}$ and selected basin characteristics to determine potential significant independent variables and multicollinearity amongst the independent variables. Independent variables with a strong correlation with the $Q_{1\%}$ and limited multicollinearity were then tested for significance using OLS regression analysis methods to eliminate statistically insignificant (p-value greater than 0.05) independent variables. Initial independent variables selected based on maximum correlation with the $Q_{1\%}$ and minimal multicollinearity with other selected variables included CONTDA, PRECIPRIS10, MINBELEV, and SSURGOCLAY. These variables were verified as significant model variables using OLS in WREG. The final GLS model for region 1 (statewide including pre-1978 irrigation region streamgages) included CONTDA and PRECIPRIS10 and the final GLS model for region 2 (post-1978 irrigation-affected region) included CONTDA. Model variables were included based on low multicollinearity, statistical significance, maximized pseudo coefficient of determination (pseudo-R^2) of the selected model and minimized model errors, and the number of high leverage and high influence points.

The GLS regression analysis method was used to develop the final regression equations (tables 7, 8) using peak streamflow data from 151 streamgages (table 2). Of the 270 streamgages originally selected for the study, 106 were removed during redundancy and nested basin screening and 12 were removed as a result of significant temporal trends in peak

Table 7. Regression equations and performance metrics for estimating annual exceedance-probability streamflows for unregulated streams in hydrologic region 1 in Kansas.

(S_P, average standard error of prediction; pseudo-R^2, psuedo coefficient of determination; SME, standard model error; AVP, average variance of prediction; ft^3/s, cubic feet per second; mi^2, square miles; $Q_{x\%}$, annual excedance probability streamflow of x percent; CONTDA, contributing drainage area; PRECIPRIS10, average 1981–2010 precipitation]

Annual exceedance-probability equation	S_P (percent)	Psuedo-R^2 (percent)	SME (percent)	AVP (log ft^3/s)^2
Data from 143 streamgages with contributing drainage areas ranging from 0.17 to 14,901 mi^2 used to develop equations				
$Q_{50\%} = 0.0019\ (\mathrm{CONTDA})^{0.510}(\mathrm{PRECIPRIS10})^{3.36}$	50.6	89.6	49.3	0.043
$Q_{20\%} = 0.0275\ (\mathrm{CONTDA})^{0.497}(\mathrm{PRECIPRIS10})^{2.83}$	44.0	90.9	42.8	0.033
$Q_{10\%} = 0.100\ (\mathrm{CONTDA})^{0.493}(\mathrm{PRECIPRIS10})^{2.575}$	44.4	90.1	43.0	0.034
$Q_{4\%} = 0.3388\ (\mathrm{CONTDA})^{0.490}(\mathrm{PRECIPRIS10})^{2.336}$	46.9	88.6	45.4	0.036
$Q_{2\%} = 0.7244\ (\mathrm{CONTDA})^{0.489}(\mathrm{PRECIPRIS10})^{2.194}$	48.9	87.7	47.5	0.040
$Q_{1\%} = 1.41\ (\mathrm{CONTDA})^{0.486}(\mathrm{PRECIPRIS10})^{2.06}$	53.2	85.4	51.4	0.045
$Q_{0.5\%} = 2.57\ (\mathrm{CONTDA})^{0.486}(\mathrm{PRECIPRIS10})^{1.95}$	55.5	84.5	53.8	0.051
$Q_{0.2\%} = 5.13\ (\mathrm{CONTDA})^{0.485}(\mathrm{PRECIPRIS10})^{1.82}$	60.5	82.1	58.6	0.059

Table 8. Regression equations and performance metrics for estimating annual exceedance-probability streamflows for unregulated streams in hydrologic region 2 in Kansas.

(S_P, average standard error of prediction; pseudo-R^2, psuedo coefficient of determination; SME, standard model error; AVP, average variance of prediction; ft^3/s, cubic feet per second; mi^2, square miles; $Q_{x\%}$, annual excedance probability streamflow of x percent; CONTDA, contributing drainage area]

Annual exceedance-probability equation	S_P (percent)	Psuedo-R^2 (percent)	SME (percent)	AVP (log ft^3/s)^2
Data from 24 streamgages with contributing drainage areas ranging from 1.02 to 3,555 mi^2 used to develop equations				
$Q_{50\%} = 57.5\ (\text{CONTDA})^{0.313}$	97.5	47.2	90.9	0.126
$Q_{20\%} = 155\ (\text{CONTDA})^{0.344}$	65.5	67.6	60.7	0.068
$Q_{10\%} = 257\ (\text{CONTDA})^{0.352}$	62.0	71.0	57.0	0.062
$Q_{4\%} = 447\ (\text{CONTDA})^{0.354}$	65.0	70.2	59.4	0.067
$Q_{2\%} = 631\ (\text{CONTDA})^{0.353}$	69.3	68.2	63.2	0.075
$Q_{1\%} = 832\ (\text{CONTDA})^{0.351}$	74.6	65.6	67.8	0.084
$Q_{0.5\%} = 1{,}071\ (\text{CONTDA})^{0.350}$	79.4	63.3	72.1	0.093
$Q_{0.2\%} = 1{,}445\ (\text{CONTDA})^{0.347}$	86.3	60.1	78.1	0.106

streamflows. A total of 143 streamgages were used in the GLS regression analyses for region 1, 24 streamgages were used in the GLS regression analyses for region 2, and 16 streamgages were used in the GSL regression analyses for both region 1 and region 2. The flood flows used in developing the regression models have differing variances for different streamgages, depending on the streamgage record length. Sample estimates based on longer records are more reliable and will have lower variance than that of streamgages with less data. The GLS regression analysis method is used to assign weights to the peak flow data for each station to adjust not only for differences in record lengths but also for cross-correlation of the annual time series on which the peak streamflow frequency statistics are based (Stedinger and Tasker, 1985; Tasker and Stedinger, 1989).

The differentiated hydrologic regions in Kansas (fig. 4) were developed for this study to address the effects of irrigation on peak flows and the widespread decreasing trends in peak streamflows in western Kansas. The differences in the predicted values from the regression equations for region 1 (using a constant PRECIPRIS10 value of 20 inches) and region 2 capture the spatiotemporal effects of irrigation and groundwater declines on peak flows (fig. 5).

Several graphs of performance metrics are generated within the WREG computer program to identify datasets with possible errors (fig. 6). The graph of GLS model residuals allows for the examination of the distribution of variance of predicted flow values with change in the independent variables and provides verification of the assumption of homoscedasticity (equal distribution of variance over the range of independent variables). The leverage metric is used to measure how distant the values of independent variables at one streamgage are from the centroid of values of the same variables at remaining streamgages. The influence metric indicates whether data from a streamgage had a large effect on the estimated regression model coefficients (Eng and others, 2009). Data points identified as having high influence or high leverage were checked for possible errors. During the development of GLS regression equations for region 1, data from three streamgages (06825500, 06917100, 06917400; table 2) resulted in these streamgages being high leverage and high outlier points and, therefore, they were eliminated from the analyses. Data from an additional short record period (12 years) streamgage that was a high outlier (07234100) also were eliminated from the analyses. The development of the GLS regression equations for region 2 resulted in the elimination of two streamgages. One streamgage (06846500) was eliminated as it was an adjusted trend station (table 2) and a high influence point, whereas another streamgage (06863400) was a high residual and high influence point and a high outlier.

Accuracy and Limitations of Regression Equations

Regression equations that produce estimates of streamflow statistics are statistical models that minimize differences between regression-estimated and computed streamflow statistics (residual errors). The use of GLS regression analysis methods allows separation of the model error variance from the sampling error variance, and model accuracy depends on the combined error from these sources.

Model error is the variability in the dependent variable (peak streamflows for various probabilities) that is unexplained by the selected independent variables (basin characteristics). Sampling error measures the ability of a finite number of streamgages with a finite number of recorded annual peak flows to describe the true peak streamflow characteristics for

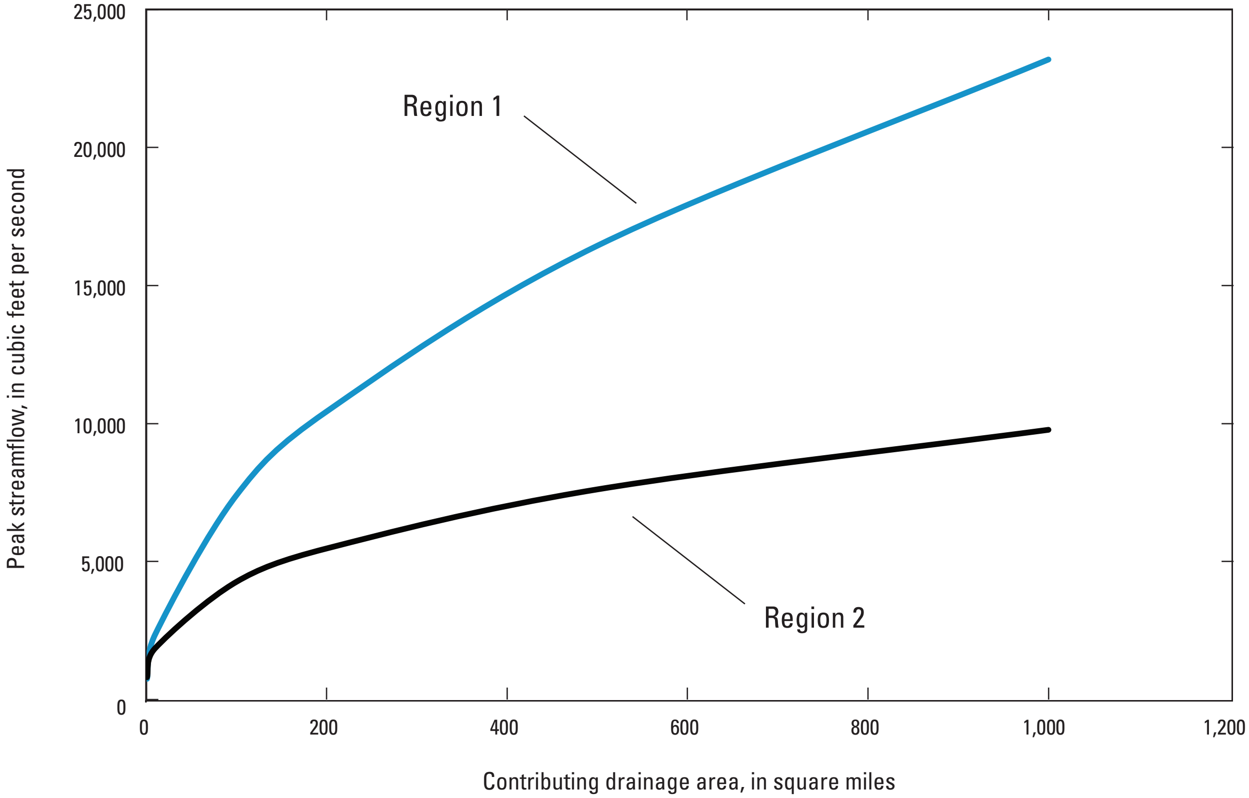


Figure 5. Comparison of predicted 1-percent annual exceedance-probability streamflows for hydrologic regions 1 and 2 used in the study.

a streamgage. The sampling error depends on the number of streamgages and record length of streamgages used in the analysis and decreases as either the number of streamgages or length of record increases (Helsel and Hirsch, 2002). The model error depends on the number and predictive power of the explanatory variables in a regression equation.

The WREG computer program provides three performance metrics in the model output that can be used to assess the accuracy of regression-estimated peak streamflow frequency statistics: the average standard error of prediction (S_p), the pseudo coefficient of determination (pseudo-R^2), and the standard model error (SME).

The average standard error of prediction measures the average accuracy of the regression equations when predicting values for ungaged sites. The standard error of prediction for region 1 equations ranged from 44.0 to 60.5 percent (table 7) and those for region 2 ranged from 62.0 to 97.5 percent (table 8). About two-thirds of the regression estimates for ungaged sites will have errors less than the given average standard errors of prediction, and about one-third of estimates will have errors larger than the given standard errors of prediction (Ries and others, 2008). The average standard errors of prediction generally range from 30 to 60 percent for most of the developed flood-peak equations for the United States (Jennings and others, 1994), although a lower range from near 15 percent or an upper range greater than 100 percent have been determined for some regions. The smallest flood-peak standard errors generally are for equations developed for the eastern United States, whereas the largest standard errors generally are for flood-peak equations developed for the western United States. The larger potential errors in western portions of the United States are attributed to greater at-site variability of the flood records, a more sparse data network, and shorter periods of station record (Ries and others, 2007). All three of these factors are present to some degree in the irrigation-affected region of western Kansas and account for the extended upper range in the standard error of prediction.

The coefficient of determination (R^2) is the proportion of the variability in the dependent variable in a regression model explained by the independent variable. Larger R^2 values indicate that a greater portion of the variability in the dependent variable is explained by the independent variable. The pseudo-R^2 (Griffis and Stedinger, 2007), however, is a more appropriate performance metric for GLS regressions and is a measure of the variability in the dependent variable explained by the regression after removing the effect of the time-sampling error (Eng and others, 2009). The pseudo-R^2 values for region 1 in Kansas ranged from 82.1 to

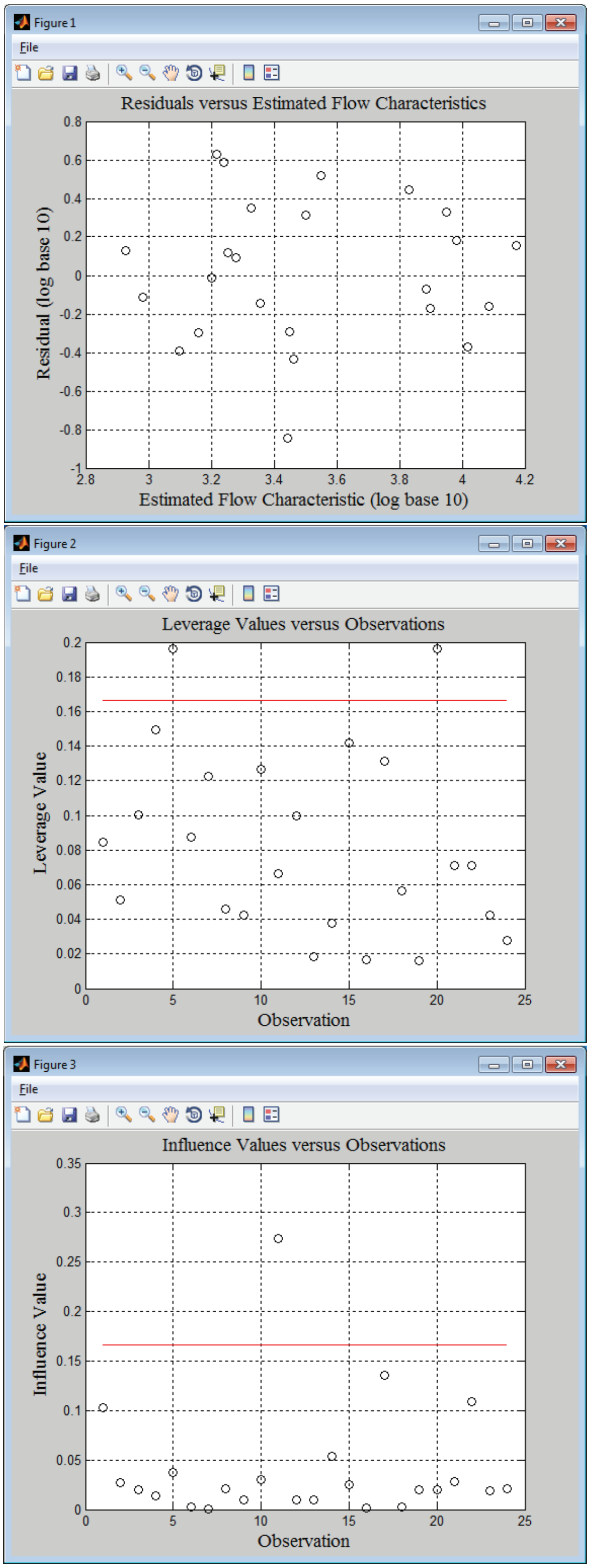


Figure 6. Performance metrics for the 1-percent annual exceedance-probability peak streamflows for region 2 generated from generalized-least squares regression analyses developed in the U.S. Geological Survey weighted-multiple-linear regression (WREG) software.

90.9 (table 7), whereas the values for region 2 ranged from 47.2 to 71.0 (table 8) and are comparable to pseudo-R^2 values determined for the panhandle of western Oklahoma (Smith and others, 2015) and the plains hydrologic region of eastern Colorado (Capesius and Stephens, 2009).

The SME is a measure of the ability of the independent model variables to estimate peak streamflow frequency statistics from the station records that were used to develop the equations. The SME is smaller than the standard error of prediction (Jennings and others, 1994) for the same peak streamflow frequency regression equation. The SMEs for the region 1 equations ranged from 42.8 to 58.6 percent (table 7), whereas the SMEs for the irrigation-affected region (region 2) ranged from 57.0 to 90.9 percent (table 8). The SMEs for equations developed for other regions of the High Plains Aquifer using similar techniques ranged from 76 to 174 percent (Capesius and Stephens, 2009; Smith and others, 2015).

Application of Regression Equations

Three methods are presented below to estimate AEP streamflows depending on whether the location is a long-term (10 or more years of record) streamgage station used in this study or an ungaged site. The methods include (1) weighting the EMA-derived AEP streamflow with the AEP streamflow derived from regression equations at gaged locations, (2) use of the drainage-area weighted method for ungaged sites located on a stream with a long-term streamgage with computed AEP streamflows, and (3) use of the regression equations at ungaged locations on streams without a streamgage used in this study.

Streamgage Locations

The EMA estimate (at-site estimate) of AEP streamflow can be improved by weighting the EMA estimate with the rural regression equation (RRE) estimate (Rasmussen and Perry, 2000). The average variance of prediction (AVP) is a measure of the AEP streamflow uncertainty provided by the WREG computer program (tables 7 and 8) and is used as the weighting factor. The EMA and RRE estimates are assumed to be independent and the variance of the weighted estimate will be less than the variance of either of the independent estimates. Once the variances have been obtained from the PeakFQ and WREG analyses output, the two independent AEP estimates can be weighted using the following equation (Cohn and others, 2012; Southard and Veilleux, 2014):

$$logQ_{P(g)w} = \frac{VP_{P(g)r}\,logQ_{P(g)s} + VP_{P(g)s}\,logQ_{P(g)r}}{VP_{P(g)s} + VP_{P(g)r}} \tag{5}$$

where

$Q_{P(g)w}$ is the weighted independent estimate of annual peak flow for the selected P-percent AEP for a streamgage, g, in cubic feet per second;

$VP_{P(g)r}$ is the variance of prediction at the streamgage derived from the applicable regional-regression equations for the selected P-percent AEP (from tables 7 and 8), in log units;

$Q_{P(g)s}$ is the at-site estimate from the EMA for the selected P-percent AEP (from tables 3 and 4) for a streamgage, g, in cubic feet per second;

$VP_{P(g)s}$ is the variance of prediction at the streamgage from the EMA for the selected P-percent AEP (from PeakFQ output), in log units; and

$Q_{P(g)r}$ is the peak streamflow estimate for the selected P-percent AEP at the streamgage derived from the applicable regional-regression equations (from tables 3 and 4), in cubic feet per second.

The weighted AEP streamflow estimates that were computed from equation 5 are listed in tables 3 and 4. The average variance of prediction values for the streamgages included in the WREG analyses are listed in tables 7 and 8.

Drainage-Area Ratio Method

For sites on streams that have long-term streamgages upstream or downstream from the site of interest, the drainage-area ratio method is a more accurate method for estimating peak streamflow frequency statistics than the regression equations (Guimaraes and Bohman, 1992; Stamey and Hess, 1993). This method is applicable to ungaged locations within a ratio of 0.5 to 1.5 times the drainage area of the corresponding streamgage. The weighted streamgage estimate ($Q_{T(G)w}$) (weighted estimate value in tables 3 and 4) is used to obtain an estimate for the ungaged site that is based on the flow per unit area at the streamgage ($Q_{T(U)g}$) by use of the following equation:

$$Q_{T(U)g} = \left[\frac{A_u}{A_g}\right]^b Q_{T(G)w}\,, \tag{6}$$

where

A_u is the drainage area for the ungaged site,

A_g is the drainage area for the upstream or downstream streamgage, and

b is the exponent of drainage area (CONTDA) from the appropriate P-percent AEP regional equation (tables 7 and 8).

Within the Streamstats application (Ries and others, 2004, 2008), the weighting equation gives full weight to the regression estimates when the drainage area for the streamgage is less than 0.5 or greater than 1.5 times the drainage area for the ungaged site and increases weight to the streamgage-based estimates as the drainage area ratio approaches 1. The weighting procedure is not to be applied when the drainage area is less than 0.5 or greater than 1.5.

Regional Regression Equations

The regression equations developed in this report (tables 7 and 8) are applicable to streams in Kansas that are not substantially affected by regulation, impoundments, or urbanization and that do not meet the criteria for the drainage-area ratio method. The region 1 regression equations are intended for use for stream sites with contributing drainage areas ranging from 0.17 to 14,901 mi^2, and the region 2 equations are intended for locations within the irrigation-affected area of western Kansas with contributing drainage areas of 1.02 to 3,555 mi^2, corresponding to the drainage areas of the streamgages used in the development of the equations.

Summary

A study was conducted by the U.S. Geological Survey, in cooperation with the Kansas Department of Transportation and the Federal Emergency Management Agency, to develop regression models to estimate peak flows of various annual exceedance probabilities (50, 20, 10, 4, 2, 1, 0.5, and 0.2 percent) of occurrence in any year. Regression models were generated using record from gaged rivers in Kansas and surrounding States that were unregulated and without urbanization. Streamgages with a minimum of 25 years of record were selected for the development of peak streamflows at current (through 2015) and discontinued streamgages with peak streamflow record ranging from water year (October through September) 1885 to 2015. Contributing drainage areas of selected streamgages ranged from 0.17 to 14,190 square miles (mi^2). Peak streamflow frequency statistics from selected streamgages were related to basin characteristics using generalized least squares regression analysis. The developed equations can be used to predict peak streamflow magnitude and frequency at ungaged locations.

Peak streamflow data used in this report were collected for 270 active and discontinued continuous-record streamgages located in Kansas and within 50 miles of the Kansas border in the surrounding States of Nebraska, Missouri, Oklahoma, and Colorado with cumulative record extending from 1885 to 2015. Record at selected streamgages was unaffected by regulation, surface-water diversions, or urbanization. Following redundancy screening and temporal trend analyses, peak streamflow frequency data from 151 streamgages were available for use in the regression analyses.

Peak streamflows and resulting frequency statistics are affected by a number of physical and meteorological factors that define basin characteristics. For each selected streamgage, as many as 21 selected basin characteristics were computed using geographic information system (GIS) software or obtained from established sources to be used as independent variables in peak streamflow frequency regression equations for Kansas.

Separate generalized skews were developed for two hydrologic regions in Kansas determined based on irrigation effects. Region 1 included those streamgages with at least 25 years of pre-1978 record (pre-irrigation) for a defined irrigation region in western Kansas and a complete record of at least 25 years for streamgages east of the irrigation region. Records for 120 streamgages in Kansas and surrounding States were used in the determination of the region 1 generalized skew. The generalized skew for the irrigation-affected region, region 2, was generated using streamgages with at least 25 years of post-1978 (post-irrigation) record and included 1 Colorado station and 13 Kansas streamgages west of the defined irrigation boundary. Generalized skew coefficients, generated using the arithmetic mean of the station skews, resulted in a region 1 generalized skew coefficient of -0.125, a standard error of 0.502, and a root mean square error of 0.252. Similarly, the generalized skew for the irrigation-affected region (region 2) was -0.478, the standard error was 0.459, and the root mean square error was 0.210. The generalized skews were weighted with the station skews to produce final flood-frequency statistics used in the development of regression equations for predicting flood-frequency statistics at ungaged locations in Kansas.

The generalized least squares regression analysis method was used to formulate the final regression equations using peak streamflow data from 151 streamgages. The final regression model for region 1 (statewide including pre-1978 irrigation region streamgages) included contributing drainage area and the 1981–2010 mean precipitation as independent variables and the final model for region 2 (post-1978 irrigation-affected region) included contributing drainage area. Three performance metrics in the model output were used to assess the accuracy of regression-estimated peak streamflow frequency statistics including the average standard error of prediction, the pseudo coefficient of determination (pseudo-R^2), and the standard model error. The standard error of prediction for region 1 regression equations ranged from 44.0 to 60.5 percent and those for region 2 ranged from 62.0 to 97.5 percent. The pseudo-R^2 values for region 1 in Kansas ranged from 82.1 to 90.9, whereas the values for region 2 ranged from 47.2 to 71.0 percent. The standard model errors for the region 1 equations ranged from about 42.8 to 58.6 percent, whereas the standard model errors for the irrigation-affected region ranged from 57.0 to 90.9 percent.

Three methods are presented for computing estimates of annual exceedance probabilities at a site. If the site is at a streamgage with 10 or more years of record, improved estimates for the site can be obtained by weighting the annual exceedance probability log-Pearson Type III estimate with the regression-equation estimate by weighting the variance of prediction of each estimate. For sites on streams that have streamgages with 10 or more years of record upstream or downstream from the site of interest, the drainage-area ratio method is a more accurate method for estimating peak streamflow frequency statistics than the regression equations. If a site does not meet these two conditions, the regression equations presented in this report may be used to estimate the annual exceedance probability streamflows. The regression equations developed in this report are applicable to streams in Kansas that are not substantially affected by regulation, surface-water diversions, or urbanization. The region 1 regression equations are intended for use for stream sites with contributing drainage areas ranging from 0.17 to 14,901 mi^2, and the region 2 equations are intended for stream sites within the irrigation-affected area of western Kansas with contributing drainage areas of 1.02 to 3,555 mi^2.

References Cited

Capesius, J.P., and Stephens, V.C., 2009, Regional regression equations for estimation of natural streamflow statistics in Colorado: U.S. Geological Survey Scientific Investigations Report 2009–5136, 46 p.

Clement, R.W., 1987, Floods in Kansas and techniques for estimating their magnitude and frequency on unregulated streams: U.S. Geological Survey Water-Resources Investigations Report 87–4008, 50 p.

Cohn, T.A., Berenbrock, Charles, Kiang, J.E., and Mason, R.R., Jr., 2012, Calculating weighted estimates of peak streamflow statistics: U.S. Geological Survey Fact Sheet 2012–2038, 4 p., accessed December 12, 2016, at http://pubs.usgs.gov/ fs/2012/3038/.

Cohn, T.A., England, J.F., Berenbrock, C.E., Mason, R.R., Stedinger, J.R., and Lamontagne, J.R., 2013, A generalized Grubbs-Beck test statistic for detecting multiple potentially influential low outliers in flood series: Water Resources Research, v. 49, no. 8, p. 5047–5058.

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

Eash, D.A., Barnes, K.K., and Veilleux, A.G., 2013, Methods for estimating annual exceedance-probability discharges for streams in Iowa, based on data through water year 2010: U.S. Geological Survey Scientific Investigations Report 2013–5086, 63 p. with appendix.

Ellis, D.W., and Edelen, G.W., Jr., 1960, Kansas streamflow characteristics, part 3, Flood frequency: Kansas Water Resources Board Technical Report No. 3, 221 p.

Eng, Ken, Chen, Yin-Yu, and Kiang, J.E., 2009, User's guide to the weighted-multiple-linear-regression program (WREG version 1.0): U.S. Geological Survey Techniques and Methods, book 4, chap. A8, 21 p.

Fenneman, N.M., 1946, Physical divisions of the United States: Washington, D. C., U. S. Geological Survey special map.

Fischer, B.C., and McGuire, V.L., 1999, Digital data set of water-level changes in the High Plains aquifer in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming, 1980–1996: U.S. Geological Survey Open-File Report 99–198, accessed December 10, 2016, at https://water.usgs.gov/lookup/getspatial?ofr99-198_wlc80_96.

Flynn, K.M., Kirby, W.H., and Hummel, P.R., 2006, User's manual for program PeakFQ, Annual flood frequency analysis using Bulletin 17B guidelines: U.S. Geological Survey Techniques and Methods, book 4, chap. B4, 42 p.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N., and Wickham, J., 2011, Completion of the 2006 National Land Cover Database for the Conterminous United States: Photogrammetric Engineering and Remote Sensing, v. 77, no. 9, p. 858–864.

Gebert, W.A., Graczyk, D.J., and Krug, W.R., 1985, Average annual runoff in the United States, 1951–80: U.S. Geological Survey Open-File Report 85–627, scale 1:2,000,000.

Goodin, D.G., Mitchell, J.E., Knapp, M.C., and Bivens, R.E., 2004, Climate and weather atlas of Kansas: Lawrence, Kans., Kansas Geological Survey Educational Series 12, accessed December 17, 2016, at http://www.kgs.ku.edu/Publications/Bulletins/ED12/KGS_ED12.pdf.

Griffis, V.W., and Stedinger, J.R., 2007, The use of GLS regression in regional hydrologic analyses: Journal of Hydrology, v. 344, p. 82–95.

Guimaraes, W.B., and Bohman, L.R., 1992, Techniques for estimating magnitude and frequency of floods in South Carolina: U.S. Geological Survey Water-Resources Investigations Report 92–4040, 174 p.

Hedman, E.R., Kastner, W.M., and Hejl, H.R., 1974, Kansas streamflow characteristics—part 10, Selected streamflow characteristics as related to active-channel geometry of streams in Kansas: Kansas Water Resources Board Technical Report No. 10, 21 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, 522 p.

Horizon Systems Corporation, 2010, National Hydrography Dataset Plus (NHDPlus): Horizon Systems Corporation, accessed November 1, 2010, at http://www.horizon-systems.com/nhdplus/.

Irza, T.J., 1966, Preliminary flood-frequency relations for small streams in Kansas: U.S. Geological Survey Open-File Report 66–67, 19 p.

Jennings, M.E., Thomas, W.O., Jr., and Riggs, H.C., 1994, Nationwide summary of U.S. Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993: U.S. Geological Survey Water-Resources Investigations Report 94–4002, 196 p., floppy disk.

Jordan, P.R., and Irza, T.J., 1975, Kansas streamflow characteristics—Magnitude and frequency of floods in Kansas, unregulated streams: Kansas Water Resources Board Technical Report No. 11, 34 p.

Junk, W.J., Bayley, P.B., and Sparks, R.E., 1989, The flood pulse concept in river-floodplain systems, *in* Dodge, D.P., ed., Proceedings of the International Large River Symposium: Canadian Journal of Fisheries and Aquatic Sciences, v. 106, p. 110–127.

Juracek, K.E., 2015, Streamflow characteristics and trends at selected streamgages in southwest and south-central Kansas: U.S. Geological Survey Scientific Investigations Report 2015–5167, 20 p., accessed January 5, 2017, at http://dx.doi.org/10.3133/sir20155167.

Kendall, M.G., 1938, A new measure of rank correlation: Biometrika, v. 30, p. 81–93.

Kendall, M.G., 1975, Rank correlation methods (4th ed.): London, U.K., Charles Griffin, 272 p.

Matthai, H.F., 1968, Magnitude and frequency of floods in the United States—Part 6B, Missouri River Basin below Sioux City, Iowa: U.S. Geological Survey Water-Supply Paper 1681, 636 p.

Moody, D.W., Chase, E.B., and Aronson, D.A., comps., 1986, National water summary 1985—Hydrologic events and surface-water resources: U.S. Geological Survey Water-Supply Paper 2300, 506 p.

National Oceanic and Atmospheric Administration, 2015, Hydrometeorological Design Studies Center, Precipitation Frequency Data Server, accessed February 12, 2015, at http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_gis.html.

National Oceanic and Atmospheric Administration, 2016, U.S. Climate Divisions: accessed December 27, 2016, at http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-divisions.php.

Natural Resources Conservation Service, 2015, Geospatial data gateway: U.S. Department of Agriculture, accessed March 30, 2015, at http://datagateway.nrcs.usda.gov/.

Parameter-Elevation Regressions on Independent Slopes Model Climate Group (PRISM), 2015, Normal annual precipitation grid for the conterminous United States: accessed March 15, 2015, at http://www.prism.oregonstate.edu/explorer/.

Patterson, J.L., 1964, Magnitude and frequency of floods in the United States—Part 7, Lower Mississippi River Basin: U.S. Geological Survey Water-Supply Paper 1691, 636 p.

Perry, C.A., Wolock, D.M., and Artman, J.A., 2004, Estimates of flow duration, mean flow, and peak-discharge frequency values for Kansas stream locations: U.S. Geological Survey Scientific Investigations Report 2004–5033, 219 p.

Rasmussen, P.P., and Perry, C.A., 2000, Estimation of peak streamflows for unregulated rural streams in Kansas: U.S. Geological Survey Water-Resources Investigations Report 2000–4079, 33 p.

Rasmussen, T.J., and Perry, C.A., 2001, Trends in peak flows of selected streams in Kansas: U.S. Geological Survey Water-Resources Investigations Report 01–4203, accessed October 3, 2016, at https://ks.water.usgs.gov/pubs/reports/wrir.01-4203.html.

Ries, K.G., III, 2007, The national streamflow statistics program—A computer program for estimating streamflow statistics for ungaged sites: U.S. Geological Survey Techniques and Methods, book 4, chap. A6, 37 p.

Ries, K.G., III, Guthrie, J.G., Rea, A.H., Steeves, P.A., and Stewart, D.W., 2008, StreamStats—A water resources web application: U.S. Geological Survey Fact Sheet 2008–3067, 6 p. [Also available at http://pubs.usgs.gov/fs/2008/3067/.]

Ries, K.G., III, Steeves, P.A., Coles, J.D., Rea, A.H., and Stewart, D.W., 2004, StreamStats—A U.S. Geological Survey web application for stream information: U.S. Geological Survey Fact Sheet 2004–3115, 4 p.

Sauer, V.B., and Turnipseed, D.P., 2010, Stage measurement at gaging stations: U.S. Geological Survey Techniques and Methods, book 3, chap. A7, 28 p.

Smith, S.J., Lewis, J.M., and Graves, G.M., 2015, Methods for estimating the magnitude and frequency of peak streamflows at ungaged sites in and near the Oklahoma Panhandle: U.S. Geological Survey Scientific Investigations Report 2015–5134, 35 p., accessed August 23, 2016, at http://dx.doi.org/10.3133/sir20155134.

Sophocleous, M.A., and Wilson, B.B., 2000, Surface water in Kansas and its interaction with groundwater: accessed December 12, 2016, at http://www.kgs.ku.edu/HighPlains/atlas/atswqn.htm.

Southard, R.E., and Veilleux, A.G., 2014, Methods for estimating annual exceedance-probability discharges and largest recorded floods for unregulated streams in rural Missouri: U.S. Geological Survey Scientific Investigations Report 2014–5165, 39 p., accessed June 27, 2016, at http://dx.doi.org/10.3133/sir20145165.

Stamey, T.C., and Hess, G.W., 1993, Techniques for estimating magnitude and frequency of floods in rural basins of Georgia: U.S. Geological Survey Water-Resources Investigations Report 93–4002, 75 p.

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

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

U.S. Geological Survey, 2014, PeakFQ: Computer program, ver. 7.1., accessed June 8, 2016, at http://water.usgs.gov/software/PeakFQ/.

U.S. Geological Survey, 2015, National Water Information System—Web interface: accessed January 12, 2015, at http://dx.doi.org/10.5066/F7P55KJN.

U.S. Geological Survey, 2016, Welcome to StreamStats: accessed November 12, 2016, at https://streamstatsags.cr.usgs.gov/ss_defs/basin_char_defs.aspx.

U.S. Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flow frequency: Reston, Va., U.S. Geological Survey, Office of Water Data Coordination, Hydrology Subcommittee Bulletin 17B [variously paged].

Veilleux, A.G., 2009, Bayesian GLS regression for regionalization of hydrologic statistics, floods and Bulletin 17 skew: Cornell, Cornell University, M.S. Thesis, 155 p.

Veilleux, A.G., and Stedinger, J.R., 2013, Regional Skewness Regression, *in* Eash, D.A., Barnes, K.K, and Veilleux, A.G., Methods for estimating annual exceedance-probability discharges for streams in Iowa, based on data through water year 2010: U.S. Geological Survey Scientific Investigations Report 2013–5086, 63 p. with appendix.

Veilleux, A.G., Cohn, T.A., Flynn, K.M., Mason, R.R., Jr., and Hummel, P.R., 2014, Estimating magnitude and frequency of floods using the PeakFQ 7.0 program: U.S. Geological Survey Fact Sheet 2013–3108, 2 p., accessed July 17, 2015, at http://pubs.usgs.gov/fs/2013/3108/

Wahl, K.L., 1998, Sensitivity of non-parametric trend analyses to multi-year extremes, *in* Proceedings of the Western Snow Conference, April 20–23, 1998: Snowbird, Utah, Western Snow Conference, p. 157–160.

Whittemore, D.O., Butler, J.J., Jr., and Wilson, B.B., 2015, Assessing the major drivers of water-level declines—New insights into the future of heavily stressed aquifers: Hydrological Sciences Journal, 12 p., accessed October 2015 at http://dx.doi.org/10.1080/02626667.2014.959958.

Young, D.P., Macfarlane, P.A., Whittemore, D.O., and Wilson, B.B., 2005, Hydrogeologic characteristics and hydrologic changes in the Cimarron River Basin, southwestern Kansas: Kansas Geological Survey Open-File Report 2005–26, 41 p.

Publishing support provided by:
Rolla and Pembroke Publishing Service Centers

For additional information concerning this publication, contact:
Director, USGS Kansas Water Science Center
4821 Quail Crest Place
Lawrence, KS 66049
(785) 842-9909

Or visit the Kansas Water Science Center website at:
https://ks.water.usgs.gov

Painter and others—**Methods for Estimating Annual Exceedance-Probability Streamflows for Streams in Kansas, Water Year 2015**—Scientific Investigations Report 2017–5063

ISSN 2328-0328 (online)
https://doi.org/10.3133/sir20175063