

Prepared in cooperation with the Arkansas Department of Environmental Quality, Southwestern Energy, the Arkansas Natural Resources Commission, and the Arkansas Game and Fish Commission

Dry Season Mean Monthly Flow and Harmonic Mean Flow Regression Equations for Selected Ungaged Basins in Arkansas

Scientific Investigations Report 2015–5031 Version 1.1, July 2015

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

By Brian K. Breaker

Prepared in cooperation with the Arkansas Department of Environmental Quality, Southwestern Energy, the Arkansas Natural Resources Commission, and the Arkansas Game and Fish Commission

Scientific Investigations Report 2015–5031 Version 1.1, July 2015

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: 2015 First release: 2015 Revised: July 2015 (ver 1.1)

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>/.

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:

Breaker, B.K., 2015, Dry season mean monthly flow and harmonic mean flow regression equations for selected ungaged basins in Arkansas (ver. 1.1, July 2015): U.S. Geological Survey Scientific Investigations Report 2015–5031, 25 p., http://dx.doi.org/10.3133/sir20155031. ISSN 2328-0328 (online)

Contents

Figures

Tables

Conversion Factors

Inch/Pound to International System of Units

Datum

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

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

Abbreviations

By Brian K. Breaker

Abstract

The U.S. Geological Survey, in cooperation with the Arkansas Department of Environmental Quality, Southwestern Energy, the Arkansas Natural Resources Commission, and the Arkansas Game and Fish Commission, developed regression equations for estimation of dry season mean monthly flows and harmonic mean flows that are representative of natural streamflow conditions at selected ungaged basins in Arkansas. Observed values of dry season mean monthly flow and harmonic mean flow computed from daily-mean flow data were used with basin characteristics to identify significant explanatory variables for multiple-linear-regression equations to estimate predicted values of dry season mean monthly flow and harmonic mean flow. Five dry season mean monthly flow regression equations and two harmonic mean flow regression equations were developed using dry season mean monthly flows and harmonic mean flows established for 91 and 93 U.S. Geological Survey continuous-record streamflowgaging stations, respectively. The dry season in Arkansas is defined as the months of July through November for this study. For harmonic mean flow calculations and regression equations, the study area is composed of the Springfield-Salem Plateaus (Arkansas and Missouri), Boston Mountains, Arkansas Valley, Ouachita Mountains (Arkansas and Oklahoma), and West Gulf Coastal Plain (Arkansas) physiographic sections. All continuous-record streamflow-gaging stations used to compute dry season mean monthly flows were located within Arkansas.

Equations for two regions were found to be statistically significant for developing regression equations for estimating harmonic mean flows at ungaged basins; thus, equations are applicable only to streams in those respective regions in Arkansas. Regression equations for dry season mean monthly flows are applicable only to streams located throughout Arkansas. All regression equations are applicable only to unaltered streams where flows were not significantly affected by regulation, diversion, or urbanization. The median number of years used for dry season mean monthly flow calculation was 43, and the median number of years used for harmonic mean flow calculations was 34 for region 1 and 43 for region 2.

Introduction

Water use in the State of Arkansas was estimated to be about 11,500 million gallons per day in 2010 (A.L. Pugh and Terrance W. Holland, U.S. Geological Survey, written commun., 2014). Groundwater and surface-water sources comprised 69 percent and 31 percent, respectively, of total water use. Total water use increased in Arkansas by 435 percent between 1965 and 2010 (A.L. Pugh and Terrance W. Holland, U.S. Geological Survey, written commun., 2014). As population and agriculture in Arkansas continue to increase, more stress is placed on streams in the State. The Arkansas Department of Environmental Quality (ADEQ) protects and regulates water resources of Arkansas through various programs (*<http://www.adeq.state.ar.us/Default.htm>*) involving permitting. Through issuance of National Pollutant Discharge Elimination System (NPDES) permits, the ADEQ is responsible for ensuring waters of the State of Arkansas are suitable for sustaining diverse biological communities and do not simultaneously pose threats to human health. The NPDES permits are required for industrial, municipal, or other facilities that discharge treated wastewater directly to surface waters. The U.S. Environmental Protection Agency (EPA) recommends that the long-term harmonic mean flow be used for assessing potential human health effects because it provides a more conservative estimate than the arithmetic mean flow. The harmonic mean flow is determined by taking the reciprocal of the mean value of the reciprocal of individual values.

The U.S. Geological Survey (USGS), in cooperation with the ADEQ, Southwestern Energy, the Arkansas Natural Resources Commission, and the Arkansas Game and Fish Commission, developed regional regression equations for estimation of dry season mean monthly flows and harmonic mean flows that are representative of natural streamflow conditions, defined as streamflows that are not affected by regulation, diversion, or urbanization (referred to hereinafter as unaltered) at ungaged basins in Arkansas. A continuousrecord streamflow-gaging station (referred to hereinafter as a streamflow gage) is a location on a stream where gage height is recorded continuously and for which daily-mean streamflow is computed (Funkhouser and others, 2008).

Dry season mean monthly flows and harmonic mean flows are routinely needed for ungaged streams for water-quality regulation, stream-related structural design, wastewater management, and stream-hazard identification. These two streamflow statistics, in particular, are useful for setting criteria for wastewater-treatment plant effluent and allowable pollutant loads to meet water-quality standards for human health criteria, irrigation, recreation, and wildlife conservation. For the estimation of these two statistics at ungaged streams, regional regression equations were developed using statistical relations that exist between streamflow data collected at streamflow gages and geologic, climatic, physical, and statistical variables for watersheds that contribute flow to a streamflow gage (Eash and Barnes, 2012).

Purpose and Scope

The purpose of this report is to present regression equations for estimation of dry season mean monthly flows and harmonic mean flows (QAH) at ungaged basins in Arkansas. Dry season mean monthly flows and QAH computed from USGS streamflow gages are also presented. Equations developed during this study also are intended for delivery in the USGS StreamStats program (U.S. Geological Survey, 2012a,b). The StreamStats program allows users to select a point on a stream in an interactive map and then automatically delineates the watershed for that point and computes selected statistics for the watershed (Eash and Barnes, 2012). The StreamStats program will provide users the ability to estimate dry season mean monthly flow, QAH, and associated 90 percent prediction intervals for ungaged streams in Arkansas.

Previous Studies

Six studies have been conducted for estimation of lowflow characteristics in Arkansas beginning with Hines (1965). The most recent study by Funkhouser and others (2008) used data collected through the 2005 water year, defined as the period from October 1 of a given year to September 30 of the following year designated by the calendar year in which it ends. Previous reports for low-flow statistics in Arkansas

have focused primarily on development of regional regression equations for low-flow frequencies, such as the 7-day, 2-year low flow $(Q_{7,2})$ and the 7-day,10-year low flow $(Q_{7,10})$. No previous studies in Arkansas have developed regional regression equations to estimate dry season mean monthly flow or QAH.

Methods of Analysis for Data from U.S. Geological Survey Continuous-Record Streamflow Gages

Data used for this report are from 113 streamflow gages in Arkansas, Oklahoma, and Missouri (fig. 1); however, streamflow gages from neighboring States were used only for QAH because explanatory variables used in regression equations for dry season mean monthly flow were spatially limited to within Arkansas. Streamflow gages from neighboring States were used to improve the representativeness of QAH and basin characteristics found in the Arkansas border areas and to provide better estimates of the error of the regression equations for ungaged sites near the Arkansas border. Streamflow gages located on unaltered streams (fig. 1; app. 1) with a minimum of 15 water years of daily-mean flows were initially selected for evaluation in the study. However, some gages with 12 water years of daily-mean flow data were added to enhance the spatial distribution of streamflow gages used to develop regression equations. Daily-mean flow data collected through the 2013 water year were retrieved for the 113 streamflow gages from the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2001).

Basin Characteristics

For this study, 47 basin characteristics (app. 2) were evaluated as potential explanatory variables for use in regression equations for estimating dry season mean monthly flow and QAH. The geographic information system (GIS) derived characteristics were selected to represent the varying geologic, climatic, physical, and statistical properties of the watersheds for the 113 streamflow gages.

Figure 1. U.S. Geological Survey streamflow gages used for dry season mean monthly flow and harmonic mean flow regressions.

Regression Techniques

Ordinary-least-squares (OLS) regressions were used to develop initial multiple-linear-regression (MLR) equations that were used for the initial analysis of streamflow data. Final equations for regional regressions were developed using weighted-least-squares (WLS) regressions. Logarithmic transformations (base 10) were used on all response variables (mean monthly flow and QAH) and tested for select explanatory variables (table 1) for the final regression equations. Logarithmic transformations were applied based on graphical comparisons of response variables and explanatory variables. Only drainage area was log transformed for the final regression equations. Other transformations were used for select explanatory variables (percent Ordovician and Mississippian) as necessary to increase linearity between the response variable and explanatory variables. The response variable was assumed to be a linear function of the explanatory variables.

MLR equations followed the general form

where

$$
Y_i
$$
 is the response variable for site *i*,
\n X_1 to X_n are the n explanatory variables for site *i*,
\n b_0 to b_n are the n + 1 regression model coefficients,
\nand
\ne is the residual error for site *i*

 $Y_i = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n + e_i$ (1)

 e_i is the residual error

A base-10 logarithmic transformation of the MLR has the form of

$$
log Y_i = b_0 + b_1 log X_1 + b_2 log X_2 + \dots + b_n log X_n + e_i \tag{2}
$$

where

log is the base-10 logarithmic transformation of the variable.

The algebraic equivalence of equation 2 after transformation back to its original units is

$$
Y_i = 10^{b0} X_1^{b1} X_2^{b2} \dots X_n^{bn} 10^{ei}
$$
 (3)

Tasker (1980) reported that OLS regression assumes that the time-sampling variance in the response-variable estimates is the same for each streamflow gage used in the analysis. As a result, this would imply that all observations of the response variable are equally reliable. This assumption is not always satisfied in hydrologic regressions because of the reliability of the record available for computation of the response variable based on the length of the observed streamflow record (Tasker, 1980; Eash and Barnes, 2012). The WLS regression adjusts for the variation in the reliability of the response-variable estimates by using a weight for each streamflow gage to account for differences in the period of record available for computation of the response variable (Eng and others, 2009; Eash and Barnes, 2012).

Table 1. Range of basin characteristic values used to develop dry season mean monthly flow and harmonic mean flow regression equations for unregulated streams in Arkansas.

[Slope1085, slope of channel in feet per mile between points at 10 and 85 percent of the longest flow path from the outlet]

Candidate regressions were selected for WLS by testing all possible combinations of the 42 explanatory variables (app. 2) using the *allReg()* function contained within the "USGSwsStats" package (Lorenz, 2013b) for R, an integrated suite of software facilities for data manipulation, calculation, and graphical display (R Core Team, 2014). The best candidate regression equations were selected based on the lowest values of Mallow's C_p and analysis of multicollinearity using the variance inflation factor (Helsel and Hirsch, 2002). The WLS regressions were repeatedly performed in R (R Core Team, 2014) to reduce the number of explanatory variables to those significant at the 95-percent confidence interval. Regression diagnostics, including residual standard error, coefficient of determination (R^2) , multicollinearity, Cook's D, leverage, and graphical relations (Helsel and Hirsch, 2002), were used to evaluate the performance of WLS regressions.

Dry Season Mean Monthly Flow Data

Daily-mean streamflow data for 91 USGS streamflow gages were used to compute monthly-mean flows. A mean monthly flow is computed as the arithmetic mean of all monthly-mean flows for a given month of the year for a selected period of record for a streamflow gage. A monthlymean flow is computed as the arithmetic mean of the dailymean flows for a given month of the year. Regression analysis was performed with ensembles for five critical, dry season months, July through November.

Five regression equations were developed for the calculation of dry season mean monthly flows during the months of July through November. All five equations contain at least one significant explanatory variable (table 1) derived from a previous runoff study in Arkansas (Pugh and Westerman, 2014). The median number of years of record used for calculation of dry season mean monthly flow at the 91 streamflow gages was 43 (app. 1). Equations for estimation of dry season mean monthly flow are applicable only at ungaged, unaltered stream locations in Arkansas.

Harmonic Mean Flow Data

Equations for two regions were found to be statistically significant for developing regression equations for estimating harmonic mean flow at selected ungaged basins in Arkansas. Of the 93 streamflow gages used for QAH (fig. 1; app. 1), 33 and 60 streamflow gages were used for the first and second regions, respectively. The median number of years of data used for QAH calculations in the WLS regression analysis was 34 for region 1 and 43 for region 2 (app. 1). The WLS regression equations used to estimate QAH are applicable only to ungaged, unaltered streams in region 1 and region 2.

For QAH calculations (app. 3), the study area is composed of the Springfield-Salem Plateaus (Arkansas and Missouri), Boston Mountains, Arkansas Valley, Ouachita Mountains (Arkansas and Oklahoma), and West Gulf Coastal Plain (Arkansas) physiographic sections (fig. 2). A QAH value was calculated for all streamflow gages used in the regression equations using the USGS "DVstats" package (Lorenz, 2013a) in R software (R Core Team, 2014), which is based on the EPA's computer program for estimating design flows for use in water-quality studies (DFLOW) (Rossman, 1990b). The QAH statistic generally is smaller than the arithmetic mean statistic, gives more weight to lower flows, and is corrected for daily flow values of zero (Rossman, 1990a). The QAH is calculated as

$$
QAH = \left(\frac{N_{nz}}{N_t}\right)\left(\frac{N_{nz}}{\sum_{i=1}^{N_{nz}}\left(\frac{1}{Q_i}\right)}\right)
$$
(4)

where

Qi is the daily mean streamflow, in cubic feet per second,

 N_{nz} is the number of nonzero days, and N_{nz} is the total number of *Q* values. is the total number of Q_i values.

If N_{n_z} equals N_{ρ} QAH is equal to the reciprocal of the mean of the reciprocals of all Q_i .

Regionalization of Harmonic Mean Flow Data

Regionalization is a statistical framework used to estimate statistics at ungaged stream locations from statistics calculated at USGS streamflow gages (Ries, 2007). Regionalization techniques are used because streamflow statistics can vary substantially between regions because of differences in geology, climate, and physical characteristics. Regionalization of low-flow statistics in Arkansas was first attempted by Hines (1965). Subsequently, Ludwig and Tasker (1993) divided the western two-thirds of Arkansas into three low-flow regions of well sustained and poorly sustained low flow. Minor changes were made to the existing low-flow statistical regions by Funkhouser and others (2008) attributable to enhanced GIS capabilities for spatial data processing. No regional methods for mean flow estimation have been developed for Arkansas prior to this study.

Streamflow gages used for QAH were located on streams in Arkansas, Oklahoma, or Missouri. An initial OLS regression was performed on all gages selected for inclusion in a statewide regression analysis. Residuals from the statewide OLS regression were spatially analyzed and indicated the need for equations in two regions.

EXPLANATION

Figure 2. Physiographic sections of Arkansas.

Region 1 includes a small part of Arkansas that is located exclusively in the northern part of the State (fig. 1), mostly in the Springfield-Salem Plateaus (figs. 1 and 2). The region is underlain by a series of limestone and dolomite units and therefore exhibits numerous karst features that affect regional hydrology and groundwater/surface-water interaction (Ludwig and Tasker, 1993). Streams in region 1 often are sustained by numerous springs. The well-sustained flows from numerous springs are indicative of a regional source of water that is supplemented during extended periods of precipitation by a local component of groundwater recharge (Ludwig and Tasker, 1993).

Region 2 includes the rest of the State excluding the Mississippi Alluvial Plain. The northern part of region 2 (fig. 1) includes the Boston Mountains, Arkansas Valley, Ouachita Mountains, and parts of the Springfield Plateau physiographic sections (fig. 2; Fenneman, 1946). Region 2 is underlain or mantled by consolidated rocks consisting primarily of sandstones and shales. The primary porosity and permeability of these sandstones and shales have been greatly reduced by compaction and deep burial; therefore, only limited amounts of groundwater are available from secondary openings including joints and fractures. The fractures, however, do not supply the base flows of streams to the extent that numerous springs do in region 1 (Ludwig and Tasker, 1993).

The southern part of region 2 is located in the West Gulf Coastal Plain physiographic section, which is underlain by unconsolidated deposits composed of sand, silt, and clay. The streams in the southern part of region 2 generally do not have sustained base flow because (1) the stream channels are not incised deeply enough to intercept the water table, and (2) the surficial deposits typically have low permeability and porosity (Ludwig and Tasker, 1993).

Dry Season Mean Monthly Flow and Harmonic Mean Flow Regression Equations

Observed values of dry season mean monthly flows and QAH (app. 3) computed from daily-mean flow data were used with basin characteristics (app. 2) to identify significant explanatory variables (table 1) for multiple-linear-regression equations (tables 2 and 3) to estimate predicted values of dry season mean monthly flows and QAH. Predicted values were compared to observed values to evaluate the performance of the WLS regression equations. The widely used residual standard error (RSE) and adjusted correlation coefficient (Helsel and Hirsch, 2002) diagnostics of the regression analysis are provided for their respective regression equations (tables 2 and 3). Normalized root mean square error (NRMSE) and percent bias (PBIAS) were additional metrics computed for all of the regression equations. Values for normalized root mean square error ranged from 33.6 to 38.4 percent for dry season mean monthly flow (table 2) and from 29.2 to 45.8 percent for QAH (table 3). The NRMSE is computed as

$$
NRMSE = 100 \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (predicted_i - observed_i)^2}}{sd\left(observed_i\right)}
$$
 (5)

where

sd is standard deviation,

N is the total number of comparisons, and

 i is the individual value.

and PBIAS as

$$
PBIAS = 100 \frac{\sum_{i=1}^{N} (predicted_i - observed_i)}{\sum_{i=1}^{N} observed_i}
$$
 (6)

Of the 47 basin characteristics (app. 2), 5 were found to be significant explanatory variables for dry season mean monthly flow and QAH regressions for regions 1 and 2: (1) drainage area; (2) base-flow index; (3) slope 1085; the stream slope computed as the change in elevation in feet between points at 10 and 85 percent of the length in miles of the longest flow path from the outlet, divided by length between the points; (4) mean dry season total runoff; and (5) percent Mississippian and Ordovician surficial geology (app. 3). An example regression equation for QAH for region 2 is:

$$
logQAH = A + blogDA + cOM + dBFI
$$
 (7)

where

Or, in the algebraic equivalence of equation 7, the regression equation can be transformed back to the original units

$$
QAH = 10^{4}(DA)^{b}10^{c(OM)} 10^{d(BFI)}
$$
 (8)

Table 2. Regression equations and ancillary regression diagnostics for dry season mean monthly flow estimation in Arkansas.

[DRNAREA, drainage area; DryTotRun, mean dry season total runoff; RSE, residual standard error; R², correlation coefficient; NRMSE, normalized root mean square error; PBIAS, percent bias]

Table 3. Regression equations and ancillary regression diagnostics for harmonic mean flow estimation in two regions identified in Arkansas.

[QAH, harmonic mean flow; DRNAREA, drainage area; BFI, base flow index; Slope1085, slope of channel in feet per mile between points at 10 and 85 percent of the longest flow path from the outlet; ORDOMISS, percent surficial geology Ordovician and Mississippian rocks divided by 100 plus 1; RSE, residual standard error; R², correlation coefficient; NRMSE, normalized root mean square error; PBIAS, percent bias]

Limitations

The final regression equations are applicable only to streams in Arkansas that are unaltered. The range of values for explanatory variables used to develop the final regression equations for this report is listed in table 1. A measure of the uncertainty associated with the regression of dry season mean monthly flows and QAH is the prediction interval. A prediction interval is the probability that the actual value of the estimated statistic will be within a specific margin of error (Helsel and Hirsch, 2002). For a 90-percent prediction interval, the true streamflow statistic has a 90-percent probability of being within the margin of error. The following equation described by Eash and Barnes (2012), which is modified from Tasker and Driver (1988), can be used to compute the 90-percent prediction interval for the true value of a streamflow statistic for an ungaged site:

 $\frac{Q}{T}$ < Q < QT (9)

where

Q is the dry season mean monthly flow or QAH predicted for the ungaged site from the regression equation, and *T* is computed as

is the critical value from the student's

t-distribution at alpha level α (α =0.10 for

$$
T = 10^{\left[t_{(a/2,n-p)}S_i\right]}
$$
 (10)

where

 $t_{(a/2, n-t)}$

the 90-percent prediction intervals; critical values may be obtained in many statistics textbooks (Iman and Conover, 1983),

- $n-p$ is the degree of freedom with *n* streamflow gages included in the regression analysis and *p* parameters in the equation (the number of explanatory variables plus one), and
	- *Si* is the standard error of prediction for site *i* and is computed as

$$
S_i = [\text{MEV} + X_i U X_i']^{0.5} \tag{11}
$$

where

- *MEV* is the mean squared error from WLS regression equations developed for this study using a user-defined weighting matrix;
	- X_i is the row vector for the streamflow gage *i*, starting with the number 1, followed by the logarithmic values of the basin characteristics used in the regression;
	- *U* is the covariance matrix for the annual or seasonal regression coefficients; and
	- *X'i* is the matrix algebra transpose of the *Xi* (Ludwig and Tasker, 1993; Ries and Friesz, 2000).

The $X_i U X'_i$ in equation 11 also is referred to as the sampling error variance. The values of $t_{(a/2,n-p)}$ and *U* needed to determine prediction intervals for estimates obtained by the regression equations in tables 2 and 3 are listed in table 4.

Table 4. Values needed to determine the 90-percent prediction intervals for estimates obtained from regression equations using covariance matrices in A rkansas.

 $[t_{(\alpha_2,n-p)}$, the critical value from Students t-distribution for the 90-percent probability used in equation 10; MEV, regression model error variance used in equation 11; U, covariance matrix as used in equation 10; logDA, log 10 transformed drainage area; DryTotRun, mean dry season total runoff; BFI, base-flow index; Slope1085, slope of channel in feet per mile between points at 10 and 85 percent of the longest flow path from the outlet; OM, percent surficial geology Ordovician and Mississippian divided by 100 plus 1]

Summary

The U.S. Geological Survey (USGS), in cooperation with the Arkansas Department of Environmental Quality (ADEQ), Southwestern Energy, the Arkansas Natural Resources Commission, and the Arkansas Game and Fish Commission, developed regression equations for estimation of dry season mean monthly flows and harmonic mean flows (QAH) that are representative of natural streamflow conditions at selected ungaged basins in Arkansas. Observed values of dry season mean monthly flow and QAH computed from daily-mean flow data were used with basin characteristics to identify significant explanatory variables for multiple-linear-regression equations to estimate predicted values of dry season mean monthly flow and QAH. These two streamflow statistics are routinely needed for ungaged streams for water-quality regulation, stream-related structural design, wastewater management, and stream-hazard identification. Dry season mean monthly flow and QAH are useful for setting criteria for wastewater-treatment plant effluent and allowable pollutant loads to meet water-quality standards for human health criteria, irrigation, recreation, and wildlife conservation.

Five dry season mean monthly flow regression equations and two QAH regression equations were developed using dry season mean monthly flow and QAH established for 91 and 93 USGS continuous-record streamflow gaging stations, respectively. The dry season in Arkansas is defined as the months of July through November. For QAH calculations and regression equations, the study area is composed of the Springfield-Salem Plateaus (Arkansas and Missouri), Boston Mountains, Arkansas Valley, Ouachita Mountains (Arkansas and Oklahoma), and West Gulf Coastal Plain (Arkansas) physiographic sections. All streamflow-gaging stations used to compute dry season mean monthly flows were located within Arkansas. Continuousrecord streamflow-gaging stations used for this study had a minimum of 12 water years of daily-mean flow data. Basin characteristics that included geologic, climatic, physical, and statistical variables were computed for each basin.

The median number of years used for dry season mean monthly flow calculation was 43. Regression equations for mean monthly flow were applicable only to stream sites located throughout Arkansas. The regression equations for dry season mean monthly flow and QAH are applicable only to unaltered streams where flows were not significantly affected by regulation, diversion, or urbanization.

Equations for two regions were found to be statistically significant for developing regression equations for estimating QAH at ungaged basins in Arkansas. Of the 93 continuousrecord streamflow-gaging stations used for QAH, 33 and 60 streamflow-gaging stations were used for the first and second regions, respectively. The median number of years of data used for QAH calculations in the weighted-least-squares regression analysis was 34 for region 1 and 43 for region 2. The weightedleast-squares regression equations used to estimate QAH were applicable only to streams in their respective regions, region 1 and region 2.

Residual standard error, adjusted correlation coefficient, normalized root mean square error, and percent bias were used to evaluate the performance of the regression equations developed for this study. Values for normalized root mean square error ranged from 33.6 to 38.4 percent for dry season mean monthly flow and from 29.2 to 45.8 percent for QAH. Equations developed during this study also are intended for delivery in the USGS StreamStats program. StreamStats will provide users the ability to estimate dry season mean monthly flow, QAH, and 90 percent prediction intervals for ungaged streams in Arkansas.

Selected References

- Acreman, W.C., and Wiltshire, S.E., 1987, Identification of regions for regional flood frequency analysis [abs.]: EOS, v. 68, no. 44, p. 1–262.
- Eash, D.A., and Barnes, K.K., 2012, Methods for estimating selected low-flow frequency statistics and harmonic mean flows for streams in Iowa: U.S. Geological Survey Scientific Investigations Report 2012–5171, 99 p. (Also available at http://pubs.usgs.gov/sir/2012/5171/.)
- Eng, Ken, Chen, Yin-Yu, and Kiang, J.E., 2009, User's guide to the weighted-multiple-linear-regression program (WREG version 1.0): U.S. Geological Survey Techniques and Methods, book 4, chap. A8, 21 p. (Also available at [http://pubs.](http://pubs.usgs.gov/tm/tm4a8/) [usgs.gov/tm/tm4a8/](http://pubs.usgs.gov/tm/tm4a8/)*.*)
- Falcone, J.A., Carlisle, D.M., Wolock, D.M., and Meador, M.R., 2010, GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecology, v. 91, no. 2, p. 621, a data paper in Ecological Archives E091–045–D1.
- Fenneman, N.M., 1946, Physical divisions of the United States: U.S. Geological Survey map, scale 1:7,000,000, 1 sheet.
- Funkhouser, J.E., Eng, Ken, and Moix, M.W., 2008, Low-flow characteristics and regionalization of low-flow characteristics for selected streams in Arkansas: U.S. Geological Survey Scientific Investigations Report 2008–5065, 161 p. (Also available at<http://pubs.usgs.gov/sir/2008/5065/>*.*)
- Haley, B.R., Glick, E.E., Bush, W.V., Clardy, B.F., Stone, C.G., Woodward, M.B., and Zachry, D.L., 1993 Geologic map of Arkansas: U.S. Geological Survey, 1 sheet, scale 1:500,000.
- 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/](http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html) [pdf_new.html](http://pubs.usgs.gov/twri/twri4a3/html/pdf_new.html)*.*)
- Hines, M.S., 1965, Water-supply characteristics of selected Arkansas streams: Arkansas Geological Commission Water Resources Circular 9, 43 p.

Iman, R.L., and Conover, W.J., 1983, A modern approach to statistics: New York, John Wiley and Sons, Inc., 497 p.

Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., and Xian, G. 2013, [A comprehensive change detection method for](http://www.mrlc.gov/downloadfile2.php?file=Preferred_NLCD11_citation.pdf) [updating the National Land Cover Database to circa 2011](http://www.mrlc.gov/downloadfile2.php?file=Preferred_NLCD11_citation.pdf): Remote Sensing of Environment, v. 132, p. 159–175. (Also available at http://www.mrlc.gov/nlcd2011.php.)

Lorenz, D.L., 2013a, DVstats—An R package for managing daily-values data, version 0.1: U.S. Geological Survey, accessed March 1, 2014, at [https://github.com/USGS-R/](https://github.com/USGS-R/DVstats) [DVstats](https://github.com/USGS-R/DVstats)*.*

Lorenz, D.L., 2013b, USGSwsStats—An R package for the analysis of hydrologic data, version 0.6: U.S. Geological Survey, accessed March 1, 2014, at https://github.com/ USGS-R/USGSwsStats.

Ludwig, A.H., and Tasker, G.D., 1993, Regionalization of low-flow characteristics of Arkansas streams: U.S. Geological Survey Water-Resources Investigations Report 93–4013, 26 p.

PRISM Climate Group, 2012, PRISM climate data: Oregon State University, accessed on May 30, 2014, at http://www. prism.oregonstate.edu/.

Pugh, A.L., and Westerman, D.A., 2014, Mean annual, seasonal, and monthly precipitation and runoff in Arkansas, 1951–2011: U.S. Geological Survey Scientific Investigations Report 2014–5006, 40 p. (Also available at http:// dx.doi.org/10.3133/sir20145006.)

R Core Team, 2014, The R project for statistical computing: Vienna, Austria, The R Foundation for Statistical Computing, accessed March 1, 2014, at <http://www.R-project.org/>.

Reed, J.C., and Bush, C.A., 2005, Generalized geologic map of the United States, Puerto Rico, and the U.S. Virgin Islands: U.S. Geological Survey National Atlas of the United States. (Also available at http://pubs.usgs.gov/atlas/ geologic/.)

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. (Also available at <http://pubs.usgs.gov/tm/2006/tm4a6/pdf/tm4a6.pdf>*.*)

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 2000–4135, 81 p. (Also available at http://pubs.usgs.gov/ wri/wri004135/*.*)

Rossman, L.A., 1990a, Design stream flows based on harmonic means: Journal of Hydraulic Engineering, v. 116, no. 7, p. 946–950

Rossman, L.A., 1990b, DFLOW user's manual: Cincinnati, Ohio, U.S. Environmental Protection Agency, Risk Reduction Engineering Laboratory, 26 p., accessed May 21, 2014, at [http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [TXT?ZyActionD=ZyDocument&Client=EPA&Index=1](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [986+Thru+1990&Docs=&Query=600890051%20or%20](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [dflow%20or%20user's%20or%20manual%20or%20](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [rossman&Time=&EndTime=&SearchMethod=1&TocRestr](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [ict=n&Toc=&TocEntry=&QField=pubnumber%5E%22600](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [890051%22&QFieldYear=&QFieldMonth=&QFieldDay=&](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [UseQField=pubnumber&IntQFieldOp=1&ExtQFieldOp=1](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [%5C86thru90%5CTxt%5C00000005%5C30001JEH.txt&U](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [ser=ANONYMOUS&Password=anonymous&SortMethod=](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [h%7C&MaximumDocuments=10&FuzzyDegree=0&Image](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [Quality=r75g8/r75g8/x150y150g16/i425&Display=p%7Cf](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActi](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [onS&BackDesc=Results%20page&MaximumPages=1&Zy](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user) [Entry=1&SeekPage=x&ZyPURL](http://nepis.epa.gov/Exe/ZyNET.exe/30001JEH.TXT?ZyActionD=ZyDocument&Client=EPA&Index=1986+Thru+1990&Docs=&Query=600890051%20or%20dflow%20or%20user)*.*

Schwarz, G.E., and Alexander, R.B., 1995, STATe Soil GeOgraphic (STATSGO) database for the conterminous United States: U.S. Geological Survey Open-File Report 95–449, 95 p.

Tasker, G.D., 1980, Hydrologic regression with weighted least squares: Water Resources Research, v. 16, no. 6, p. 1107– 1113.

Tasker, G.D., and Driver, N.E., 1988, Nationwide regression models for predicting urban runoff water quality at unmonitored sites: Water Resources Bulletin, v. 24, no. 5, p. 1091– 1101, accessed December 12, 2014, at http://onlinelibrary. wiley.com/doi/10.1111/j.1752-1688.1988.tb03026.x/pdf.

U.S. Department of Agriculture, 2001, STATe Soil GeOgraphic (STATSGO) database—Data use information (revised July 1994): Natural Resources Conservation Service, National Soil Survey Center, Miscellaneous Publication No. 1492, 110 p.

U.S. Geological Survey, 2001,USGS Water data for the Nation: National Water Information System: Web Interface, accessed March 21, 2014, at http://waterdata.usgs.gov/ nwis/.

U.S. Geological Survey, 2011, National elevation dataset: U.S. Geological Survey, accessed March 21, 2014, at [http://ned.](http://ned.usgs.gov/) [usgs.gov/](http://ned.usgs.gov/).

U.S. Geological Survey, 2012a, Welcome to StreamStats—The StreamStats Program: U.S. Geological Survey, accessed May 30, 2013, at [http://water.usgs.gov/osw/streamstats/](http://water.usgs.gov/osw/streamstats/index.html) [index.html](http://water.usgs.gov/osw/streamstats/index.html)*.*

U.S. Geological Survey, 2012b, Welcome to StreamStats— Basin characteristic definitions: U.S. Geological Survey, accessed May 30, 2013, at http://water.usgs.gov/osw/ streamstats/bcdefinitions1.html.

Appendixes 1–3

Appendix 1

Table 1–1. U.S. Geological Survey streamflow gages used for regression analysis for dry season mean monthly flow and harmonic mean flow.

[ft3/s, cubic feet per second; USGS, U.S. Geological Survey; Ark., Arkansas; Mo., Missouri; Okla., Oklahoma; --, gage was not used for the regression in the column heading for that station]

ភ

Table 1–1. U.S. Geological Survey streamflow gages used for regression analysis for dry season mean monthly flow and harmonic mean flow.—Continued

[ft³/s, cubic feet per second; USGS, U.S. Geological Survey; Ark., Arkansas; Mo., Missouri; Okla., Oklahoma; --, gage was not used for the regression in the column heading for that station]

 $\overline{\mathbf{u}}$

Appendix 2. Definitions of Basin Characteristics Evaluated as Response Variables for Inclusion in the Regression Analysis

Annual precipitation, in inches, is the average annual precipitation for the drainage basin as computed from PRISM (Prism Climate Group, 2012) data for 1951–2011.

August average precipitation, in inches, is the basin average for the month of August averaged from PRISM (Prism Climate Group, 2012) data for 1981–2010.

Average basin elevation, in feet, is the average elevation of the basin as determined from the National Elevation Dataset (U.S. Geological Survey, 2011) 10 meter grid.

Base flow Index (BFI), dimensionless, is the mean ratio of base flow to annual streamflow from the USGS kriged BFI grid (Falcone and others, 2010) as an averaged value for the basin.

Basin perimeter distance, in miles, is the distance around the boundary of the basin.

Basin shape factor, dimensionless, is the ratio of the total drainage area to the basin length.

Drainage area, in square miles, is the area measured in a horizontal plane that is enclosed by a drainage divide.

Forest, in percent, calculated from the National Land Cover Database as the percentage of the basin that is mixed forested (Jin and others, 2013).

July average precipitation, in inches, as a basin average for the month of July averaged from PRISM (Prism Climate Group, 2012) data for 1981–2010.

Longest flow path length, in miles, is the maximum flow distance within a basin from the start of overland flow to the outlet.

Maximum basin elevation, in feet, is the maximum elevation of the basin computed from the National Elevation Dataset (U.S. Geological Survey, 2011) 10 meter grid.

Mean annual groundwater runoff, in inches, is the portion of the total runoff at the outlet from seepage of water from the ground into a stream channel as computed from geographic information system (GIS) grid from Pugh and Westerman (2014).

Mean annual precipitation 1951–2011, in inches, at the basin outlet is the average annual precipitation as determined from PRISM (Prism Climate Group, 2012) data for 1951–2011.

Mean annual precipitation 1971–2000, in inches, at the basin centroid and averaged for the basin is the average annual precipitation as determined from PRISM (Prism Climate Group, 2012) data for 1971–2000.

Mean annual surface runoff, in inches, is the portion of the total runoff at the outlet that travels over the land surface into the stream channel as computed from GIS grid data from Pugh and Westerman (2014).

Mean annual total runoff, in inches, is the total runoff at the outlet that travels over the land surface and from seepage of water from the ground into the stream channel as computed from GIS grid data from Pugh and Westerman (2014).

Mean dry season groundwater runoff, in inches, is the portion of the total runoff at the outlet from seepage of water from the ground into a stream channel for the months of June through November as computed from GIS grid data from Pugh and Westerman (2014).

Mean dry season precipitation, in inches, is the average precipitation for the months of July through November averaged over the drainage basin as computed from PRISM (Prism Climate Group, 2012) data for 1951–2011.

Mean dry season surface runoff, in inches, is the portion of the total runoff at the outlet that travels over the land surface into the stream channel for the months of June through November as computed from GIS grid data from Pugh and Westerman (2014).

Mean dry season total runoff, in inches, is the total runoff at the outlet that travels over the land surface and from seepage of water from the ground into the stream channel for the months of June through November as computed from GIS grid data from Pugh and Westerman (2014).

Mean wet season groundwater runoff, in inches, is the portion of the total runoff at the outlet from seepage of water from the ground into a stream channel for the months of December through May as computed from GIS grid data from Pugh and Westerman (2014).

Mean wet season precipitation, in inches, is the average precipitation for the months of December through June averaged over the drainage basin as computed from PRISM (Prism Climate Group, 2012) data 1951–2011.

Mean wet season surface runoff, in inches, is the portion of the total runoff at the outlet that travels over the land surface into the stream channel for the months of December through May as computed from GIS grid data from Pugh and Westerman (2014).

Mean wet season total runoff, in inches, is the total runoff at the outlet that travels over the land surface and from seepage of water from the ground into the stream channel for the months of December through May as computed from GIS grid data from Pugh and Westerman (2014).

Minimum basin elevation, in feet, is the maximum elevation of the basin computed from the National Elevation Dataset (U.S. Geological Survey, 2011) 10 meter grid.

November average precipitation, in inches, as a basin average for the month of November averaged from PRISM (Prism Climate Group, 2012) data for 1981–2010.

October average precipitation, in inches, as a basin average for the month of October averaged from PRISM (Prism Climate Group, 2012) data for 1981–2010.

Outlet elevation, in feet, is the elevation at the gage location computed from the National Elevation Dataset (U.S. Geological Survey, 2011) 10 meter grid.

Percent Cretaceous (K), is the percentage of the basin in which Cretaceous sedimentary rocks dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Percent Paleozoic (lPz), is the percentage of the basin in which lower Paleozoic sedimentary rocks dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Percent Mississippian, is the percentage of the basin in which Mississippian-age rocks dominate the surface of the basin as computed from GIS grid data from Haley and others (1993).

Percent Middle Paleozoic (mPz), is the percentage of the basin in which Middle Paleozoic sedimentary rocks dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Percent Ordovician, is the percentage of the basin in which Ordovician-age rocks dominate the surface of the basin as computed from GIS grid data from Haley and others (1993).

Percent Ordovician and Mississippian, is the percentage of the basin in which Mississippian and Ordovician-age rocks dominate the surface of the basin as computed from GIS grid data from Haley and others (1993).

Percent Paleogene (pgT), is the percentage of the basin in which Paleogene sedimentary rocks dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Percent Quaternary (Q), is the percentage of the basin in which Quaternary deposits dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Percent Upper Paleozoic (uPz), is the percentage of the basin in which Upper Paleozoic sedimentary rocks dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Percent Middle Proterozoic (Yv), is the percentage of the basin in which Middle Proterozoic volcanic rocks dominate the surface of the basin as computed from GIS grid data created by Reed and Bush (2005).

Relief, in feet, is the maximum basin elevation minus the minimum basin elevation computed from the National Elevation Dataset (U.S. Geological Survey, 2011) 10 meter grids.

September average precipitation, in inches, as a basin average for the month of September averaged from PRISM (Prism Climate Group, 2012) data for 1981–2010.

Slope, in feet per mile, of the longest flow path through the basin.

Slope 1085, in feet per mile, is the stream slope computed as the change in elevation between points at 10 and 85 percent of length along the longest flow path from the outlet, determined by GIS, divided by length between the points.

Soil Permeability, in inches per hour, is the rate at which water flows through soil as computed from STATe Soil GeOgraphic (STATSGO) grid data (Schwarz and Alexander, 1995; U.S. Department of Agriculture, 2001).

Soil hydrologic group, dimensionless, percentage of drainage basin in hydrologic soil group as computed from STATe Soil GeOgraphic (STATSGO) grid data (Schwarz and Alexander, 1995; U.S. Department of Agriculture, 2001).

Streamflow-variability Index, dimensionless, is a measure of the steepness of the slope of a streamflow duration curve as computed at the outlet.

Sum of stream lengths for basin, in miles, is the total length of all streams in the basin combined.

Urban, percent, calculated from the National Land Cover Database as the percentage of the basin that is urban (Jin and others, 2013).

Appendix 3

Table 3–1. Dry season mean monthly flow, harmonic mean flow, and explanatory variable values for final regression equations.

Fable 3–1. Dry season mean monthly flow, harmonic mean flow, and explanatory variable values for final regression equations.
USGS, U.S. Geological Survey; ft³/s, cubic feet per second; Slope1085, slope of channel betwe [USGS, U.S. Geological Survey; ft³/s, cubic feet per second; Slope1085, slope of channel between points at 10 and 85 percent of the longest flow path from the outlet; mi², square mile; ft/mi, foot per mile; in., inches; --, gage was not used for the regression in the column heading for that station; Ark., Arkansas; Mo., Missouri; Okla., Oklahoma]

71

Table 3–1. Dry season mean monthly flow, harmonic mean flow, and explanatory variable values for final regression equations.—Continued

[USGS, U.S. Geological Survey; ft³/s, cubic feet per second; Slope1085, slope of channel between points at 10 and 85 percent of the longest flow path from the outlet; mi², square mile; ft/mi, foot per mile; in., inches; --, gage was not used for the regression in the column heading for that station; Ark., Arkansas; Mo., Missouri; Okla., Oklahoma]

23

Publishing support provided by Lafayette Publishing Service Center

ISSN 2328-0328 (online) <http://dx.doi.org/10.3133/sir20155031>