

Prepared in cooperation with the Indiana Department of Environmental Management

Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows for Ungaged, Unregulated Streams in Indiana



Scientific Investigations Report 2016–5102

Version 1.1, October 2016

Cover. Bridge over Little Otter Creek in Jefferson County, Indiana. The photograph is courtesy of the U.S. Fish and Wildlife Service, Big Oaks National Wildlife Refuge.

Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows for Ungaged, Unregulated Streams in Indiana

By Gary R. Martin, Kathleen K. Fowler, and Leslie D. Arihood

Prepared in cooperation with the Indiana Department of Environmental Management

Scientific Investigations Report 2016–5102
Version 1.1, October 2016

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, Director

U.S. Geological Survey, Reston, Virginia: 2016

First release: 2016

Revised: October 2016 (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://store.usgs.gov/>.

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

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

Suggested citation:

Martin, G.R., Fowler, K.K., and Arihood, L.D., 2016, Estimating selected low-flow frequency statistics and harmonic-mean flows for ungaged, unregulated streams in Indiana (ver 1.1, October 2016): U.S. Geological Survey Scientific Investigations Report 2016–5102, 45 p., <http://dx.doi.org/10.3133/sir20165102>.

ISSN 2329-132X (online)

Contents

Acknowledgments	vii
Abstract	1
Introduction	2
Purpose and Scope	2
Previous Studies	2
Description of Study Area	2
Physiography and Geology	4
Climate	5
Development of Datasets for Streamgages	5
Source of Low-Flow Frequency and Harmonic-Mean Flow Statistics	5
Sources of Basin Characteristics	12
Development of Equations for Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows at Ungaged, Unregulated Stream Sites	17
Logistic-Regression Analysis.....	17
Multiple-Linear-Regression Analyses	18
Limitations and Accuracy of the Estimating Equations	21
Prediction Intervals	27
Example Applications of the Estimating Equations	28
Example 1	28
Example 2	28
Example 3	29
Results and Discussion of Regression Analyses	30
Summary.....	30
References Cited.....	31
Glossary.....	34
Appendix 1. Classification Tables for Logistic-Regression Equations for Estimating the Probability of Zero Flow for Selected Low-Flow Frequencies in Indiana	35
Appendix 2. Values Needed to Determine 90-Percent Prediction Intervals for Multiple-Linear-Regression Estimates of Low-Flow Statistics in Indiana	38
Appendix 3. Values of Basin Characteristics That Were Significant Explanatory Variables in the Regression Equations	40

Figures

1. Map showing major rivers and river basins in Indiana.....	3
2. Map showing physiographic regions of Indiana and streamgages used in regression analyses.....	4
3. Map showing grid of texture-based average horizontal hydraulic conductivity of the first 70 feet below land surface of unconsolidated deposits in Indiana	13
4. Map showing residuals of a 7-day 10-year recurrence interval streamflow regression equation for the entire study area.....	20
5. Screenshot of the weighted-multiple-linear regression program smoothing function for generalized-least squares correlation of the time series of annual minimum 7-day mean flows as a function of distance between 78 streamgages with 30 years of concurrent flow in Indiana.....	21
6. Graphs showing comparison of selected low-flow frequencies estimated from measured streamflow and from regression equations for unregulated streams in Indiana, observed and predicted values for annual minimum 1-day mean, 10-year low flows; annual minimum 7-day mean, 10-year low flows; and annual minimum 30-day mean, 10-year low flows	23
7. Graphs showing comparison of harmonic-mean flows estimated from measured streamflow and from regression equations for unregulated streams in three low-flow regions of Indiana	25

Tables

1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana	6
2. Description of basin characteristics evaluated as potential explanatory variables in regression equations for estimating selected low-flow statistics in Indiana	14
3. Logistic-regression equations for estimating the annual probability of zero flow for selected low-flow frequencies for ungaged, unregulated streams statewide in Indiana.....	18
4. Generalized-least-squares regression equations for estimating selected nonzero low-flow frequencies for ungaged, unregulated streams statewide in Indiana.....	20
5. Weighted-least-squares regression equations for estimating harmonic-mean flow for ungaged, unregulated streams in each low-flow region in Indiana	22
6. Basin-characteristic ranges for the logistic-regression equations for estimating the annual probability of zero flow for selected low-flow frequencies and generalized-least-squares regression equations for estimating selected nonzero low-flow frequencies for ungaged, unregulated streams statewide in Indiana.....	22
7. Basin-characteristic ranges for the weighted-least-squares regression equations for estimating harmonic-mean flows for ungaged, unregulated streams in each low-flow region in Indiana	22

Conversion Factors

Inch-Pound to International System of Units (SI)

Multiply	By	To obtain
Length		
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)
Hydraulic conductivity		
foot per day (ft/d)	0.3048	meter per day (m/d)
Transmissivity		
foot squared per day (ft ² /d)	0.092903	meter squared per day (m ² /d)

Datums

Vertical coordinate information is referenced to either (1) stage, the height above an arbitrary datum established at a streamgage, or (2) elevation, the height above North American Vertical Datum of 1988 (NAVD 88).

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

Abbreviations

BSLDEM10M	average basin slope in percent
DFLOW	computer program developed by the U.S. Environmental Protection Agency (Rossman, 1990b), which was used to compute harmonic-mean flow.
DRNAREA	total basin drainage area in square miles
FOREST	percent forest cover
GIS	geographic information system
GLS	generalized least squares
K1	average hydraulic conductivity (feet/per day) for the top 70 ft of unconsolidated deposits
K2	average hydraulic conductivity (feet/per day) of the full thickness of unconsolidated deposits
LAT_OUT	latitude of the basin outlet in decimal degrees
M1D10Y	1-day mean 10-year low flow in cubic feet per second
M7D10Y	7-day mean 10-year low flow in cubic feet per second
M30D10Y	30-day mean 10-year low flow in cubic feet per second
OLS	ordinary least squares
QSSPERMTHK	index of permeability and thickness of Quaternary surficial sediments in feet
QAH	harmonic-mean flow in cubic feet per second
R ²	coefficient of determination
ST2	average transmissivity (square feet per day) of the full thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel
T2	average transmissivity (square feet per day) of the full thickness of unconsolidated deposits
USGS	United States Geologic Survey
WLS	weighted least squares
WREG	USGS Weighted-multiple-linear REGression Program

Acknowledgments

The authors wish to thank the many local, State, and Federal agencies that have cooperated to fund the operation and maintenance of the gages used for this study, especially the Indiana Department of Natural Resources, Division of Water, the Indiana Department of Environmental Management, and the U.S. Army Corps of Engineers—Louisville and Chicago Districts, the Indiana Department of Transportation, and the City of Indianapolis, Department of Public Works; their long-term support made this data analysis possible. Special thanks are given to Indiana Department of Environmental Management for their support of this study.

Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows for Ungaged, Unregulated Streams in Indiana

By Gary R. Martin, Kathleen K. Fowler, and Leslie D. Arihood¹

Abstract

Information on low-flow characteristics of streams is essential for the management of water resources. This report provides equations for estimating the 1-, 7-, and 30-day mean low flows for a recurrence interval of 10 years and the harmonic-mean flow at ungaged, unregulated stream sites in Indiana. These equations were developed using the low-flow statistics and basin characteristics for 108 continuous-record streamgages in Indiana with at least 10 years of daily mean streamflow data through the 2011 climate year (April 1 through March 31). The equations were developed in cooperation with the Indiana Department of Environmental Management.

Regression techniques were used to develop the equations for estimating low-flow frequency statistics and the harmonic-mean flows on the basis of drainage-basin characteristics. A geographic information system was used to measure basin characteristics for selected streamgages. A final set of 25 basin characteristics measured at all the streamgages were evaluated to choose the best predictors of the low-flow statistics.

Logistic-regression equations applicable statewide are presented for estimating the probability that selected low-flow frequency statistics equal zero. These equations use the explanatory variables total drainage area, average transmissivity of the full thickness of the unconsolidated deposits within 1,000 feet of the stream network, and latitude of the basin outlet. The percentage of the streamgage low-flow statistics correctly classified as zero or nonzero using the logistic-regression equations ranged from 86.1 to 88.9 percent.

Generalized-least-squares regression equations applicable statewide for estimating nonzero low-flow frequency statistics use total drainage area, the average hydraulic conductivity of the top 70 feet of unconsolidated deposits, the slope of

the basin, and the index of permeability and thickness of the Quaternary surficial sediments as explanatory variables. The average standard error of prediction of these regression equations ranges from 55.7 to 61.5 percent.

Regional weighted-least-squares regression equations were developed for estimating the harmonic-mean flows by dividing the State into three low-flow regions. The Northern region uses total drainage area and the average transmissivity of the entire thickness of unconsolidated deposits as explanatory variables. The Central region uses total drainage area, the average hydraulic conductivity of the entire thickness of unconsolidated deposits, and the index of permeability and thickness of the Quaternary surficial sediments. The Southern region uses total drainage area and the percent of the basin covered by forest. The average standard error of prediction for these equations ranges from 39.3 to 66.7 percent.

The regional regression equations are applicable only to stream sites with low flows unaffected by regulation and to stream sites with drainage basin characteristic values within specified limits. Caution is advised when applying the equations for basins with characteristics near the applicable limits and for basins with karst drainage features and for urbanized basins. Extrapolations near and beyond the applicable basin characteristic limits will have unknown errors that may be large. Equations are presented for use in estimating the 90-percent prediction interval of the low-flow statistics estimated by use of the regression equations at a given stream site.

The regression equations are to be incorporated into the U.S. Geological Survey StreamStats Web-based application for Indiana. StreamStats allows users to select a stream site on a map and automatically measure the needed basin characteristics and compute the estimated low-flow statistics and associated prediction intervals.

¹Scientist Emeritus, U.S. Geological Survey.

Introduction

The need for developing methods to understand, manage, preserve and protect the many water assets of Indiana has been identified by policy research from legislative and business representatives in Indiana (Indiana Legislative Services Agency, 2013; Wittman, 2014). Knowledge of the magnitude and frequency of low streamflow events (low flows) is essential for water-supply planning and design, waste-load allocation, reservoir design, maintenance of aquatic life, and understanding the quantity and quality of water that is used for irrigation and recreation. Industrial, municipal, and other facilities must obtain National Pollutant Discharge Elimination System (NPDES) permits if their discharges go directly to surface waters (U.S. Environmental Protection Agency, 2014). As authorized by the Clean Water Act, the NPDES permit program limits water pollution by regulating point sources that discharge pollutants into waters of the United States. In Indiana, the NPDES permit program is administered by the Indiana Department of Environmental Management based on low-flow statistics (Indiana Department of Environmental Management, 2015).

The U.S. Geological Survey (USGS), in cooperation with State, local, and Federal agencies, has collected daily streamflow data in Indiana since 1928. This continually growing dataset is important in many aspects of water-resources investigations in Indiana.

Low-flow characteristics have been determined by the USGS at 272 streamgages in Indiana (Fowler and Wilson, 2015). Low-flow statistics investigated in that report include the 1-, 7-, and 30-day mean low flows for a recurrence interval of 10 years (variable names M1D10Y, M7D10Y, and M30D10Y, respectively) and the harmonic-mean flow (QAH). There are many stream sites in Indiana, however, that do not have continuous records of streamflow on which to base estimates of these low-flow statistics. This study began in 2014 in cooperation with the Indiana Department of Environmental Management, to develop methods for estimating low-flow statistics at ungaged stream sites. These methods effectively transfer information from the streamgage network to ungaged stream sites with similar hydrology in Indiana.

Purpose and Scope

This report presents methods for estimating low-flow statistics for ungaged sites on unregulated stream reaches in Indiana. (Examples of regulated reaches are those with flow controlled by reservoirs and those that have other major modifications to the flow regime.) Multiple regression techniques were used to develop equations to estimate the M1D10Y, M7D10Y, and M30D10Y low flows, the probability that these statistics may be zero, and the QAH statewide. Regression analysis was used to identify basin characteristics that are significant predictors of low-flow statistics for gaged sites. Geographic information system datasets of basin

characteristics for Indiana were used for this regression analysis. Data from 108 continuous-record streamgages were used to develop the estimating equations. The streamgages were selected based on the following criteria: (1) a drainage area of less than 1,000 square miles (mi²), (2) a period of record of at least 10 years, and (3) little or no regulation of streamflow in the drainage area. The low-flow frequencies and harmonic-mean flows of the 108 streamgages were based on streamflow data ending on or before the 2011 climate year.

Previous Studies

Arihood and Glatfelter (1986) presented equations for estimating certain streamflow statistics for ungaged sites in the central and northern areas of Indiana based on flow characteristics published in Stewart (1983). Vj g'Arihood and Glatfelter equations y g'g'wugf "vq estimatg'vj g low-flow characteristics M7D2Y (seven-day mean, two-year low flow) and M7D10Y0 Equations for estimating QAH, M1D10Y, and M30D10Y have not been published previously for Indiana. Arvin (1989) presented statistical summaries for streamflow data in Indiana that included flow-duration tables and annual low flows for selected consecutive days.

Description of Study Area

The State of Indiana has an area of 36,418 mi² and is in the east-central United States. The study area (fig. 1) includes the entire State of Indiana. The major drainage basins in Indiana (fig. 1) are the Great Lakes Basin, which includes the St. Joseph Basin, the Lake Michigan Basin, and the Maumee River Basin; the Upper Mississippi River Basin, which includes the Kankakee River Basin; the Wabash River Basin, which includes the upper, middle, and lower Wabash River Basins, the West Fork White River Basin, the East Fork White River Basin, and the Patoka River Basin; and the Ohio River tributary drainage basin, which includes the Whitewater River Basin, and minor tributaries to the Ohio River (Ohio River Basin).

Nearly 80 percent of Indiana is drained by streams that discharge into the Ohio River, which is the southern border of Indiana (fig. 1). The largest river in Indiana, the Wabash River, drains 32,910 mi². The White River, a tributary to the Wabash River, has two subbasins of nearly equal size, the main stem (West Fork White River) and the East Fork White River, with a total drainage area of 11,349 mi². The Whitewater River, which drains 1,369 mi², eventually discharges into the Ohio River. The Kankakee and Iroquois Rivers are part of the Illinois River drainage; they drain about 7 percent of Indiana (2,581 mi²) and flow westward into Illinois. Approximately 10 percent of Indiana is drained by three rivers that are part of the Great Lakes Basin. The Little Calumet and St. Joseph Rivers drain into Lake Michigan, and the Maumee River drains into Lake Erie (not shown). Streams in the extreme south and southeast area of the State drain directly into the Ohio River (Hoggatt, 1975).

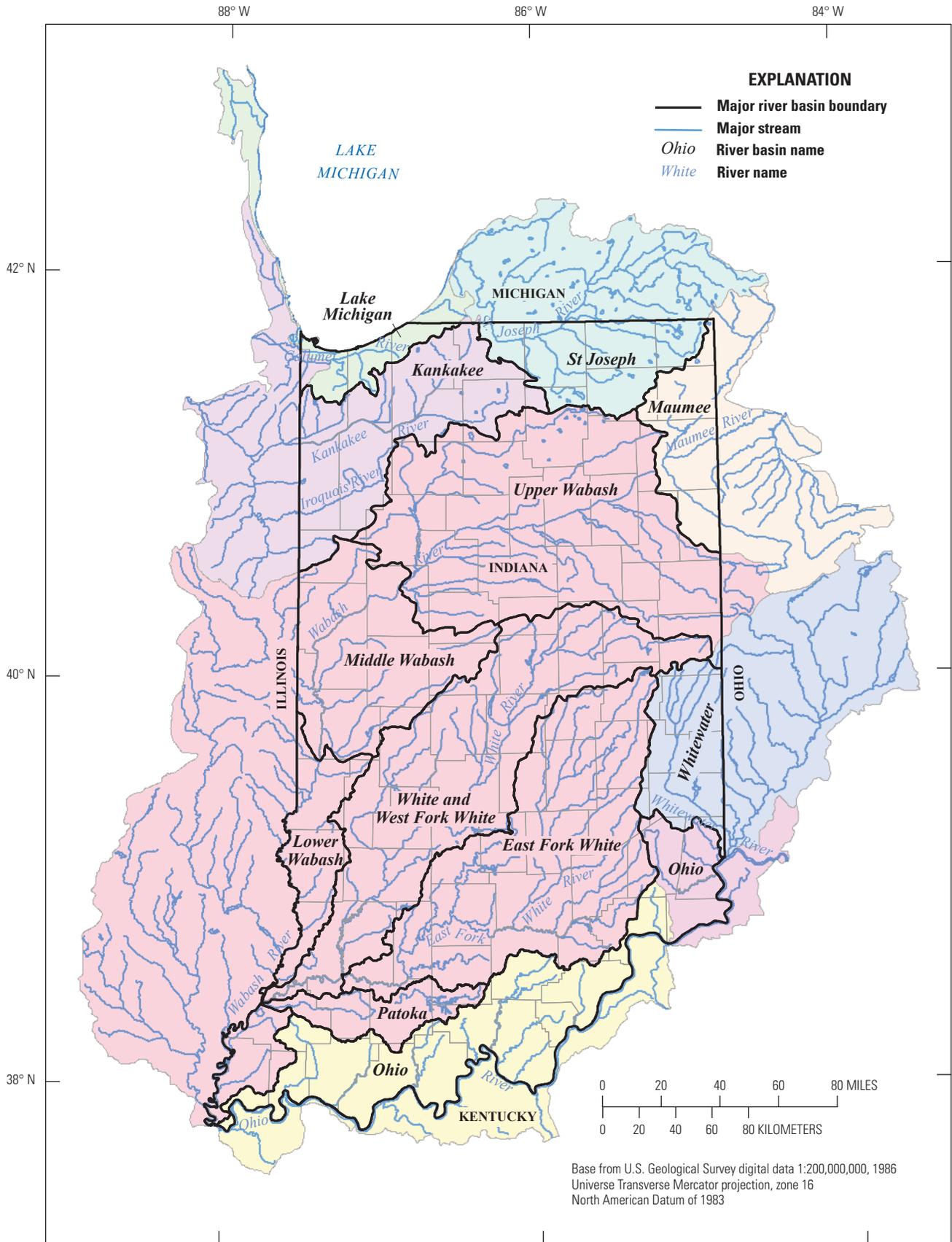


Figure 1. Major rivers and river basins in Indiana.

Physiography and Geology

Indiana can be divided into four broad physiographic regions that are based on similarities in topography and geology (Gray, 2000) (fig. 2). The Northern Moraine and Lake Region (hereinafter Northern region) is mostly glacial in origin and generally has more relief than the central zone. The Central Till Plain Region (hereinafter Central region) is a depositional plain of low relief underlain by thick glacial till that has been modified by post-glacial stream erosion. The small eastern zone, the Maumee Lake Plain Region, is part of a larger region

extending into Ohio that is fairly flat with a few low ridges of silt and sand. Landforms in the southern zone, the Southern Hills and Lowlands Region (hereinafter Southern region), are formed from degradational processes, such as weathering and stream erosion. A detailed description of the subdivisions within each region can be found in Indiana Geological Survey Special Report 61, Physiographic divisions of Indiana, (Gray, 2000).

Groundwater storage is generally abundant in the north and central areas of Indiana where there are glacial deposits. Underlying bedrock with shallow soils limits groundwater storage in much of south central Indiana (Scheeringa, 2002).

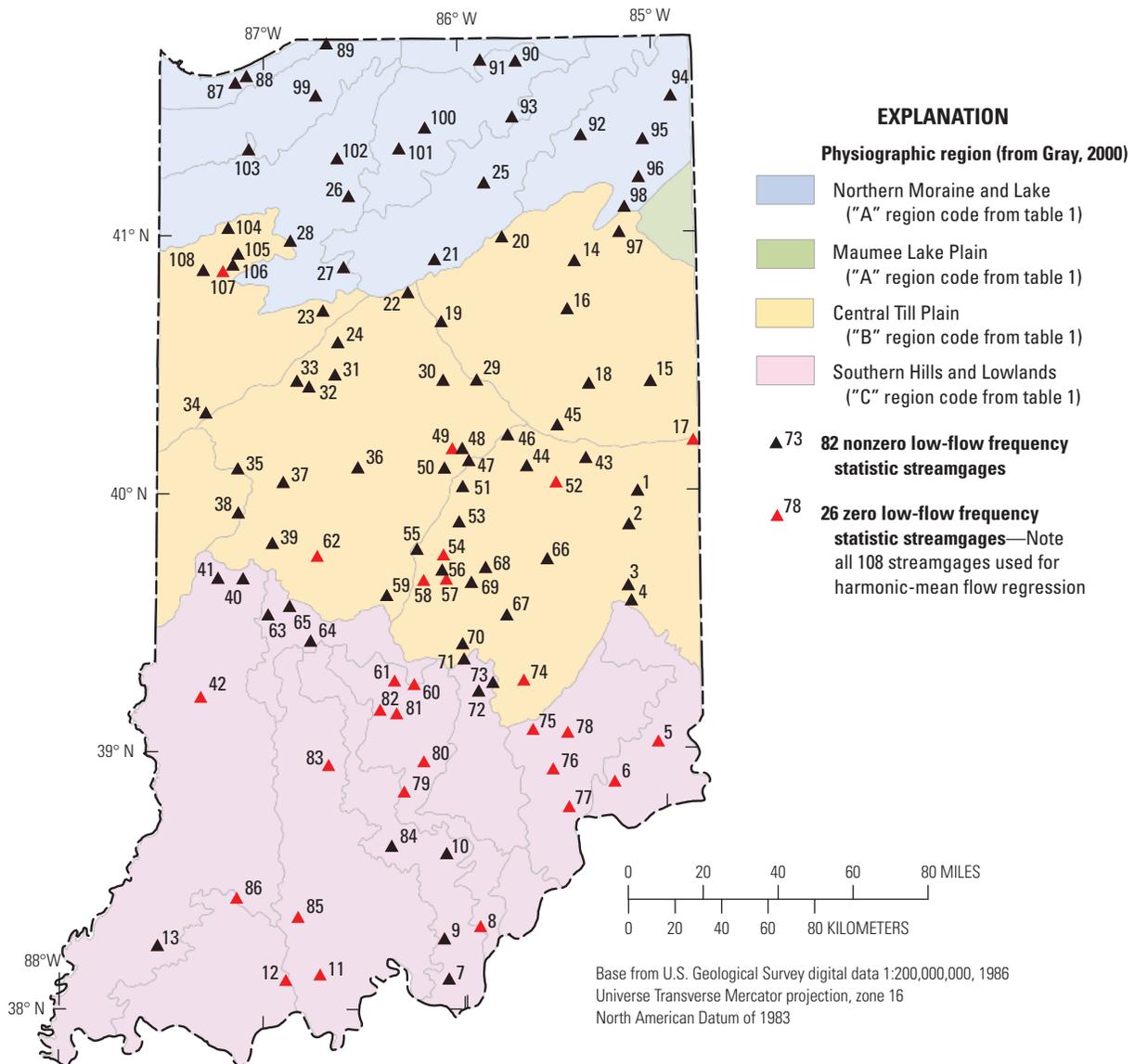


Figure 2. Physiographic regions of Indiana and streamgages used in regression analyses.

Climate

Indiana climate is classified as continental, warm, and humid; the summers tend to be hot and humid, whereas winters tend to be cold and damp. The transitional seasons of spring and fall have frequent changes in weather. The growing season extends approximately from April through October. A well-defined, north-south climatic gradient across the State provides a cool, temperate climate in the north and a warm, temperate climate in the south. High humidity and frequent variations in temperature are characteristic (Scheeringa, 2002).

Average annual temperatures from 1981 to 2010 ranged from 49.9 degrees Fahrenheit (°F) in northeast Indiana to 54.9 °F in southwest Indiana (National Oceanic and Atmospheric Administration, 2014). January is typically the coldest month of the year with normal daily maximum temperatures ranging from 31 to 38 °F north to south across Indiana. Normal January minimums range between 15 and 21 °F north to south. July is the warmest month with daily maximums averaging 80 to 83 °F and minimums 63 to 65 °F north to south (Scheeringa, 2002). The average annual temperatures have slightly increased from the period 1961–90 to the period 1981–2010. Temperature has increased in the northern divisions by about 0.6 °F, in the central divisions by about 0.8 °F, and in the southern divisions by about 0.1 °F (Fowler and Wilson, 2015).

Average annual precipitation from 1981 to 2010 ranged from 38.1 inches in northeast Indiana to 46.6 inches in the south central area (National Oceanic and Atmospheric Administration, 2014). May is typically the wettest month of the year with average rainfall between 4 and 5 inches across the State. Monthly average rainfall decreases slightly as summer progresses. Autumn months are drier with 3 inches of rainfall typical in each month. Indiana winters are the driest time of year with less than 3 inches of precipitation commonly received each month. February is the driest month of the year statewide, then precipitation increases in March and April as the spring soil moisture recharge season begins.

Annual precipitation is usually adequate for water needs, but an uneven distribution in the summer can limit crops. Mild to severe droughts occasionally occur in the summer when evapotranspiration is highest and dependence on rainfall is greatest for crops. The average annual precipitation has increased from the period 1961 to 1990 (Fowler and Wilson, 1996), to the period 1981 to 2010. On average, annual precipitation has increased about 1.7 inches in the northern divisions, 2.2 inches in the central divisions, and 1.3 inches in the southern divisions (Fowler and Wilson, 2015).

Development of Datasets for Streamgages

Regression analyses were used to relate low-flow statistics estimated from streamgage data to drainage-basin physical and climatic characteristics. The low-flow statistics at the streamgages were used as the dependent variables in the regressions, and the streamgage drainage-basin characteristics were used as independent, or explanatory, variables. The regression analyses were iterative processes used to identify optimal combinations of explanatory variables considering predictive power of the variables and other criteria as described further in this report.

Source of Low-Flow Frequency and Harmonic-Mean Flow Statistics

The low-flow statistics M1D10Y, M7D10Y, M30D10Y, and QAH were calculated by Fowler and Wilson (2015) for 272 continuous-record streamgages in Indiana. The low-flow frequency statistics were determined by frequency analysis. In low-flow investigations, frequency curves relate the minimum average discharge for a given number of consecutive days (*D*-day) to the recurrence interval in years (*T*-year). To compute return period low-flow values, such as M7D10Y, an annual time series of average consecutive 7-day minimum discharges was calculated for all continuous-record streamgages. For example, the 7-day, 10-year low flow (M7D10Y) is the minimum average discharge for 7 consecutive days, which has a 0.1 probability of not being exceeded in a given year. The recurrence interval is the reciprocal of the probability of recurrence. The recurrence interval is an estimated period, averaged over a very long period of time, and it is not a prediction of when a particular flow will happen. For example, the M7D10Y low could happen 3 years in a row, and it may not recur during any given 20-year period. Values of these low-flow statistics at streamgages (table 1) can be accessed in the Indiana StreamStats Web-based application (U.S. Geological Survey, 2015).

The harmonic-mean flow statistic, the QAH, can serve as a design flow for human health criteria that are based on lifetime exposures because the QAH can be used to calculate the average exposure concentration of a contaminant for an average contaminant loading rate (Rossman, 1990a). The exposure concentration will be greater and more harmful on days with low flow than on days with high flow. The QAH statistic computed from a streamflow record generally is

6 Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows for Ungaged, Unregulated Streams in Indiana

Table 1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana.

[mi², square mile; physiographic region code: A, northern region; B, central region; C, southern region; ft³/s, cubic feet per second; M1D10Y, 1-day, 10-year low flow; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow]

Station number	Station name	Map number (fig. 2)	Latitude	Longitude	Drainage area (mi ²) from StreamStats ¹	Physio-graphic region code
			(decimal degrees)			
03274650	Whitewater River near Economy, Indiana	1	40.00422	-85.11562	10.36	B
03274750	Whitewater River near Hagerstown, Indiana	2	39.87365	-85.16317	59.49	B
03274950	Little Williams Creek at Connersville, Indiana	3	39.63780	-85.17199	9.30	B
03275000	Whitewater River near Alpine, Indiana	4	39.57946	-85.15734	521.85	B
03276700	South Hogan Creek near Dillsboro, Indiana	5	39.02953	-85.03817	38.14	C
03291780	Indian-Kentuck Creek near Canaan, Indiana	6	38.87794	-85.25744	27.47	C
03302220	Buck Creek near new Middletown, Indiana	7	38.12021	-86.08818	65.35	C
03302300	Little Indian Creek near Galena, Indiana	8	38.32211	-85.93134	17.09	C
03302500	Indian Creek near Corydon, Indiana	9	38.27611	-86.11002	128.93	C
03302680	West Fork Blue River at Salem, Indiana	10	38.60563	-86.09471	19.13	C
03303300	Middle Fork Anderson River at Bristow, Indiana	11	38.13883	-86.72104	39.66	C
03303400	Crooked Creek near Santa Claus, Indiana	12	38.11810	-86.88997	7.98	C
03322011	Pigeon Creek near Fort Branch	13	38.25227	-87.51962	32.91	C
03324000	Little River near Huntington, Indiana	14	40.89862	-85.41325	263.39	B
03324200	Salamonie River at Portland, Indiana	15	40.42769	-85.03902	85.60	B
03324300	Salamonie River near Warren, Indiana	16	40.71244	-85.45402	425.54	B
03325311	Little Mississinewa River at Union City, Indiana	17	40.19639	-84.82944	9.76	B
03326070	Big Lick Creek near Hartford City, Indiana	18	40.42202	-85.35107	28.96	B
03327520	Pipe Creek near Bunker Hill, Indiana	19	40.66835	-86.09554	158.42	B
03328000	Eel River at North Manchester, Indiana	20	40.99456	-85.78258	419.36	B
03328430	Weesau Creek near Deedsville, Indiana	21	40.90949	-86.12660	9.26	A
03328500	Eel River near Logansport, Indiana	22	40.78264	-86.26444	788.89	B
03329400	Rattlesnake Creek near Patton, Indiana	23	40.71293	-86.69694	7.44	B
03329700	Deer Creek near Delphi, Indiana	24	40.59032	-86.62178	275.54	B
03331110	Walnut Creek near Warsaw, Indiana	25	41.20485	-85.86995	19.73	A
03331500	Tippecanoe River near Ora, Indiana	26	41.15680	-86.56383	856.37	A
03332300	Little Indian Creek near Royal Center, Indiana	27	40.88143	-86.59056	35.21	A
03332400	Big Monon Creek near Francesville, Indiana	28	40.98431	-86.86181	153.37	A
03333450	Wildcat Creek near Jerome, Indiana	29	40.44090	-85.91870	149.07	B
03333600	Kokomo Creek near Kokomo, Indiana	30	40.44119	-86.08902	25.26	B
03334000	Wildcat Creek at Owasco, Indiana	31	40.46475	-86.63651	395.56	B
03334500	South Fork Wildcat Creek near Lafayette, Indiana	32	40.41815	-86.76832	242.93	B
03335000	Wildcat Creek near Lafayette, Indiana	33	40.44096	-86.82967	794.28	B
03335700	Big Pine Creek near Willamsport, Indiana	34	40.31756	-87.29029	323.34	B
03339108	East Fork Coal Creek near Hillsboro, Indiana	35	40.10172	-87.12877	32.52	B
03339280	Prairie Creek near Lebanon, Indiana	36	40.10468	-86.52276	33.28	B
03339500	Sugar Creek at Crawfordsville, Indiana	37	40.04892	-86.89994	510.16	B
03340000	Sugar Creek near Byron, Indiana	38	39.93054	-87.12597	669.76	B

Table 1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana.—Continued

[mi², square mile; physiographic region code: A, northern region; B, central region; C, southern region; ft³/s, cubic feet per second; M1D10Y, 1-day, 10-year low flow; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow]

M1D10Y ² (ft ³ /s)	M7D10Y ² (ft ³ /s)	M30D10Y ² (ft ³ /s)	QAH— Harmonic- mean flow ² (ft ³ /s)	Probability of zero flow for M1D10Y flow	Probability of zero flow for M7D10Y flow	Probability of zero flow for M30D10Y flow	Period of record	Number of climatic years used to compute statistic
0.2	0.3	0.3	1.60	0.03	0.03	0.00	1971–2011	40
6.2	6.8	8.4	28.00	0.00	0.00	0.00	1971–2003	32
0.2	0.3	0.3	2.00	0.05	0.05	0.00	1969–1991	22
48	54	62	203.00	0.00	0.00	0.00	1929–2011	82
0	0	0	0.60	0.55	0.32	0.19	1962–1993	31
0	0	0	0.50	0.51	0.49	0.27	1970–2011	41
0.7	0.8	1.3	8.60	0.00	0.00	0.00	1970–2011	41
0	0	0	0.60	0.35	0.32	0.09	1969–2003	34
0	0	0.2	3.40	0.24	0.16	0.00	1944–1993	49
0	0	0.1	0.80	0.13	0.13	0.03	1971–2011	40
0	0	0	0.90	0.55	0.45	0.14	1962–2011	49
0	0	0	0.20	0.91	0.79	0.27	1970–2003	33
0.4	0.5	0.7	3.50	0.00	0.00	0.00	1987–2001	14
4.4	5.4	7.5	37.00	0.00	0.00	0.00	1945–2011	67
0.6	1	1.4	6.10	0.00	0.00	0.00	1960–1993	33
6.8	8.1	10	50.00	0.00	0.00	0.00	1957–2011	54
0	0	0	0.30	0.21	0.21	0.14	1983–1997	14
0.3	0.5	0.7	3.00	0.00	0.00	0.00	1972–2004	31
4.1	4.7	5.6	28.00	0.00	0.00	0.00	1969–2003	34
32	39	47	142.00	0.00	0.00	0.00	1930–2011	81
0.3	0.4	0.5	2.50	0.00	0.00	0.00	1971–2001	30
98	105	117	330.00	0.00	0.00	0.00	1944–2011	67
0.1	0.1	0.1	1.20	0.00	0.00	0.00	1969–1993	24
11	12	15	66.00	0.00	0.00	0.00	1944–2011	67
0.5	0.6	0.9	4.80	0.00	0.00	0.00	1970–2003	33
133	138	153	471.00	0.00	0.00	0.00	1944–2011	67
0.6	0.7	1.2	7.10	0.00	0.00	0.00	1960–1973	13
9.1	9.9	13	55.00	0.00	0.00	0.00	1960–1973	13
1.2	1.4	1.8	13.00	0.00	0.00	0.00	1962–2011	49
0.1	0.2	0.3	2.00	0.00	0.00	0.00	1960–2011	51
18	21	25	104.00	0.00	0.00	0.00	1944–1972, 1989–2011	51
19	20	23	79.00	0.00	0.00	0.00	1944–2011	67
59	63	74	260.00	0.00	0.00	0.00	1955–2011	56
7.6	8.1	10.2	52.00	0.00	0.00	0.00	1956–1987	31
3.1	3.4	4.3	14.00	0.00	0.00	0.00	1969–1991	23
0.3	0.5	1.2	6.70	0.00	0.00	0.00	1988–2011	23
8	8.9	12	76.00	0.00	0.00	0.00	1939–2011	72
21	22	28	130.00	0.00	0.00	0.00	1941–1971	30

8 Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows for Ungaged, Unregulated Streams in Indiana

Table 1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana.—Continued

[mi², square mile; physiographic region code: A, northern region; B, central region; C, southern region; ft³/s, cubic feet per second; M1D10Y, 1-day, 10-year low flow; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow]

Station number	Station name	Map number (fig. 2)	Latitude	Longitude	Drainage area (mi ²) from StreamStats ¹	Physio-graphic region code
			(decimal degrees)			
03340800	Big Raccoon Creek near Fincastle, Indiana	39	39.81256	-86.95403	138.74	B
03341000	Big Raccoon Creek at Mansfield, Indiana	40	39.67549	-87.10162	247.58	C
03341200	Little Raccoon Creek near Catlin, Indiana	41	39.67720	-87.22724	133.36	C
03342100	Busseron Creek near Hymera, Indiana	42	39.21510	-87.31154	16.87	C
03347500	Buck Creek near Muncie, Indiana	43	40.13478	-85.37329	35.12	B
03348000	White River at Anderson, Indiana	44	40.10625	-85.67210	406.19	B
03348020	Killbuck Creek near Gaston, Indiana	45	40.26262	-85.51489	24.34	B
03348350	Pipe Creek at Frankton, Indiana	46	40.22739	-85.76635	113.02	B
03348500	White River near Noblesville, Indiana	47	40.12948	-85.96284	827.53	B
03349500	Cicero Creek near Arcadia, Indiana	48	40.17622	-85.99584	130.81	B
03349700	Little Cicero Creek near Arcadia, Indiana	49	40.17562	-86.04623	40.32	B
03350100	Hinkle Creek near Cicero, Indiana	50	40.10143	-86.08634	18.29	B
03350700	Stony Creek near Noblesville, Indiana	51	40.02887	-85.99543	51.05	B
03351400	Sugar Creek near Middletown, Indiana	52	40.04092	-85.52509	5.70	B
03352200	Mud Creek at Indianapolis, Indiana	53	39.89192	-86.01614	42.10	B
03353160	Pleasant Run at Brookville Road at Indianapolis, Indiana	54	39.76449	-86.09536	11.23	B
03353600	Little Eagle Creek at Speedway, Indiana	55	39.78755	-86.22796	18.81	B
03353620	Lick Creek at Indianapolis, Indiana	56	39.70577	-86.10391	15.02	B
03353630	Little Buck Creek near Southport, Indiana	57	39.66974	-86.08263	5.95	B
03353637	Little Buck Creek near Indianapolis, Indiana	58	39.66653	-86.19638	17.07	B
03353800	White Lick Creek at Mooresville, Indiana	59	39.60829	-86.38209	210.98	B
03354500	Beanblossom Creek at Beanblossom, Indiana	60	39.26258	-86.24801	14.62	C
03355000	Bear Creek near Trevlac, Indiana	61	39.27750	-86.34594	6.95	C
03357350	Plum Creek near Bainbridge, Indiana	62	39.76161	-86.72921	2.99	B
03357500	Big Walnut Creek near Reelsville, Indiana	63	39.53568	-86.97644	326.42	C
03358000	Mill Creek near Cataract, Indiana	64	39.43353	-86.76338	244.43	C
03359500	Deer Creek near Putnamville, Indiana	65	39.56804	-86.86791	59.04	C
03361000	Big Blue River at Carthage, Indiana	66	39.74399	-85.57533	183.97	B
03361500	Big Blue River at Shelbyville, Indiana	67	39.52904	-85.78218	420.44	B
03361650	Sugar Creek at New Palestine, Indiana	68	39.71434	-85.88567	93.97	B
03361850	Buck Creek at Acton, Indiana	69	39.65702	-85.95770	78.86	B
03362000	Youngs Creek near Edinburgh, Indiana	70	39.41884	-86.00470	107.62	B
03362500	Sugar Creek near Edinburgh, Indiana	71	39.36092	-85.99821	474.15	B
03363900	Flatrock River at Columbus, Indiana	72	39.23521	-85.92712	532.71	C
03364200	Haw Creek near Clifford, Indiana	73	39.26786	-85.85650	47.39	C
03364500	Clifty Creek at Hartsville, Indiana	74	39.27489	-85.70185	91.32	B

Table 1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana.—Continued

[mi², square mile; physiographic region code: A, northern region; B, central region; C, southern region; ft³/s, cubic feet per second; M1D10Y, 1-day, 10-year low flow; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow]

M1D10Y ² (ft ³ /s)	M7D10Y ² (ft ³ /s)	M30D10Y ² (ft ³ /s)	QAH— Harmonic- mean flow ² (ft ³ /s)	Probability of zero flow for M1D10Y flow	Probability of zero flow for M7D10Y flow	Probability of zero flow for M30D10Y flow	Period of record	Number of climatic years used to compute statistic
1.6	1.9	2.5	17.00	0.00	0.00	0.00	1958–2011	53
2.4	3.1	3.6	28.00	0.00	0.00	0.00	1940–1958	18
4.2	4.5	5.3	20.00	0.00	0.00	0.00	1958–1971	13
0	0	0	0.30	0.39	0.25	0.06	1967–2003	36
6.7	7.5	8.5	22.00	0.00	0.00	0.00	1955–2003	48
25	29	36	135.00	0.00	0.00	0.00	1927–2011	65
1	1.1	1.3	6.90	0.00	0.00	0.00	1969–1991	22
3.6	4	4.9	23.00	0.00	0.00	0.00	1969–2003	34
60	66	76	252.00	0.00	0.00	0.00	1917–1974	55
0.8	1.1	1.5	10.00	0.00	0.00	0.00	1955–1976	21
0	0	0	2.00	0.25	0.20	0.15	1956–1976	20
0.1	0.2	0.3	2.20	0.00	0.00	0.00	1956–1976	20
2.4	2.6	3.2	15.00	0.00	0.00	0.00	1968–2011	43
0	0	0.1	0.40	0.00	0.00	0.00	1969–1989	20
0.4	0.5	0.6	5.50	0.00	0.00	0.00	1959–1976	17
0	0	0.1	1.10	0.38	0.33	0.05	1960–1981	21
0.1	0.2	0.6	2.90	0.10	0.04	0.00	1960–2011	51
0.1	0.2	0.5	2.60	0.00	0.00	0.00	1971–2011	40
0	0	0	0.20	0.70	0.40	0.00	1990–2000	10
0	0	0	1.50	0.81	0.62	0.14	1991–2011	21
3	3.9	5.6	36.00	0.00	0.00	0.00	1958–2011	53
0	0	0	0.50	0.68	0.56	0.29	1952–1993	41
0	0	0	0.30	0.95	0.75	0.30	1953–1973	20
0	0	0	0.10	0.51	0.49	0.24	1970–2011	41
3.8	4.6	5.6	54.00	0.00	0.00	0.00	1950–2002	52
1.4	1.8	2.9	21.00	0.00	0.00	0.00	1950–2011	61
0.12	0.12	0.21	2.50	0.00	0.00	0.00	1955–1965, 1968–1972	14
23	25	33	98.00	0.00	0.00	0.00	1952–2006	53
40	43	48	174.00	0.00	0.00	0.00	1944–2011	67
1.7	2.3	3.1	20.00	0.00	0.00	0.00	1968–2011	43
0.8	1.4	2.4	14.00	0.00	0.00	0.00	1968–2011	43
1.3	1.7	2.3	14.00	0.00	0.00	0.00	1943–2011	68
18	20	24	121.00	0.00	0.00	0.00	1943–2011	68
28	29	32	163.00	0.00	0.00	0.00	1968–2011	43
0.5	0.6	0.8	5.80	0.00	0.00	0.00	1968–1991, 2010–2011	23
0	0	0	1.60	0.43	0.33	0.16	1948–2011	63

10 Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows for Ungaged, Unregulated Streams in Indiana

Table 1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana.—Continued

[mi², square mile; physiographic region code: A, northern region; B, central region; C, southern region; ft³/s, cubic feet per second; M1D10Y, 1-day, 10-year low flow; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow]

Station number	Station name	Map number (fig. 2)	Latitude	Longitude	Drainage area (mi ²) from StreamStats ¹	Physio-graphic region code
			(decimal degrees)			
03365000	Sand Creek near Brewersville, Indiana	75	39.08392	-85.65922	154.48	C
03366000	Graham Creek near Vernon, Indiana	76	38.92978	-85.56247	77.20	C
03366200	Harberts Creek near Madison, Indiana	77	38.78200	-85.48552	9.25	C
03368000	Brush Creek near Nebraska, Indiana	78	39.07019	-85.48632	11.32	C
03371520	Back Creek at Leesville, Indiana	79	38.84669	-86.30161	24.11	C
03371600	South Fork Salt Creek at Kurtz, Indiana	80	38.96330	-86.20365	38.13	C
03372000	North Fork Salt Creek near Belmont, Indiana	81	39.14975	-86.33683	119.81	C
03372300	Stephens Creek near Bloomington, Indiana	82	39.16371	-86.41830	10.83	C
03373200	Indian Creek near Springville, Indiana	83	38.95061	-86.67554	60.71	C
03373530	Lost River near Leipsic, Indiana	84	38.63617	-86.36559	35.07	C
03375800	Hall Creek near St. Anthony, Indiana	85	38.36264	-86.82870	21.75	C
03376260	Flat Creek near Otwell, Indiana	86	38.43672	-87.13112	21.35	C
04094000	Little Calumet River at Porter, Indiana	87	41.62193	-87.08708	65.97	A
04094500	Salt Creek near McCool, Indiana	88	41.59684	-87.14453	75.19	A
04096100	Galena River near Laporte, Indiana	89	41.74761	-86.67508	17.87	A
04099808	Little Elkhart River at Middlebury, Indiana	90	41.67532	-85.70016	97.56	A
04099850	Pine Creek near Elkhart, Indiana	91	41.68105	-85.88304	30.23	A
04100295	Rimmell Branch near Albion, Indiana	92	41.38515	-85.37054	10.96	A
04100377	Solomon Creek near Syracuse, Indiana	93	41.45828	-85.72044	36.22	A
04177720	Fish Creek at Hamilton, Indiana	94	41.53206	-84.90378	37.42	A
04179500	Cedar Creek at Auburn, Indiana	95	41.36570	-85.05200	87.32	A
04180000	Cedar Creek near Cedarville, Indiana	96	41.21867	-85.07678	269.50	A
04182590	Harber Ditch at Fort Wayne, Indiana	97	41.00763	-85.18256	21.89	B
04182810	Spy Run Creek at Fort Wayne, Indiana	98	41.10499	-85.15345	13.94	A
05515400	Kingsbury Creek near Laporte, Indiana	99	41.54679	-86.73051	6.33	A
05516000	Yellow River near Bremen, Indiana	100	41.41976	-86.17101	134.66	A
05516500	Yellow River at Plymouth, Indiana	101	41.34031	-86.30433	293.85	A
05517000	Yellow River at Knox, Indiana	102	41.30282	-86.62057	435.07	A
05517890	Cobb Ditch near Kouts, Indiana	103	41.33865	-87.07503	30.62	A
05521000	Iroquois River at Rosebud, Indiana	104	41.03317	-87.18028	38.14	B
05522500	Iroquois River at Rensselaer, Indiana	105	40.93352	-87.12870	204.66	B
05523500	Slough Creek near Collegeville, Indiana	106	40.89151	-87.15483	83.51	B
05524000	Carpenter Creek at Egypt, Indiana	107	40.86620	-87.20547	44.84	B
05524500	Iroquois River near Foresman, Indiana	108	40.87022	-87.30644	448.74	B

¹USGS StreamStats is a Web based application that provides streamflow statistics for streams in Indiana (<http://water.usgs.gov/osw/streamstats/indiana.html>).

²Data from Fowler and Wilson, 2015.

Table 1. Description of U.S. Geological Survey streamgages and low-flow frequency statistics, harmonic-mean flows, and probability of zero flows used in the development of regional regression equations for Indiana.—Continued

[mi², square mile; physiographic region code: A, northern region; B, central region; C, southern region; ft³/s, cubic feet per second; M1D10Y, 1-day, 10-year low flow; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow]

M1D10Y ² (ft ³ /s)	M7D10Y ² (ft ³ /s)	M30D10Y ² (ft ³ /s)	QAH— Harmonic- mean flow ² (ft ³ /s)	Probability of zero flow for M1D10Y flow	Probability of zero flow for M7D10Y flow	Probability of zero flow for M30D10Y flow	Period of record	Number of climatic years used to compute statistic
0	0	0	6.30	0.16	0.13	0.11	1948–1986	38
0	0	0	1.00	0.59	0.41	0.18	1956–1973	17
0	0	0	0.30	0.68	0.44	0.12	1969–2003	34
0	0	0	0.30	0.82	0.69	0.33	1956–2011	55
0	0	0	0.60	0.47	0.25	0.06	1971–2003	32
0	0	0	0.40	1.00	0.90	0.50	1961–1971	10
0	0	0	1.80	0.33	0.21	0.13	1947–1971	24
0	0	0	0.40	0.30	0.25	0.05	1971–1991	20
0	0	0	2.20	0.18	0.18	0.18	1962–1973	11
0.5	0.6	0.7	6.10	0.00	0.00	0.00	1993–2001, 2010–2011	10
0	0	0	0.50	0.40	0.27	0.13	1971–2001	30
0	0	0	0.80	0.53	0.47	0.06	1965–1982	17
20	21	23	48.00	0.00	0.00	0.00	1946–2011	65
18	19	22	46.00	0.00	0.00	0.00	1946–1991	45
7.5	8.1	9.4	20.00	0.00	0.00	0.00	1971–2004	33
27	29	32	69.00	0.00	0.00	0.00	1980–2003	23
2.9	3.4	4.6	12.00	0.00	0.00	0.00	1980–2003	23
0.1	0.1	0.2	1.10	0.00	0.00	0.00	1980–2001	20
7.3	8.3	10	25.00	0.00	0.00	0.00	1988–2003	15
0.9	1.1	1.6	7.90	0.00	0.00	0.00	1970–2011	41
1.3	1.8	2.3	12.00	0.00	0.00	0.00	1944–1973	29
19	21	24	79.00	0.00	0.00	0.00	1947–2011	64
0.1	0.1	0.2	1.10	0.00	0.00	0.00	1965–1991	26
1.1	1.4	2	5.00	0.00	0.00	0.00	1984–2001	17
1.1	1.2	1.3	3.20	0.00	0.00	0.00	1971–1986	15
6	6.3	6.9	25.00	0.00	0.00	0.00	1956–1973	17
21	23	27	94.00	0.00	0.00	0.00	1949–2011	62
64	71	83	226.00	0.00	0.00	0.00	1944–2011	67
8.6	9.4	10	21.00	0.00	0.00	0.00	1969–2003	34
2.2	2.4	3	10.00	0.00	0.00	0.00	1949–2003	54
5.8	6.7	8.6	45.00	0.00	0.00	0.00	1949–2011	62
1.3	1.4	1.8	11.00	0.00	0.00	0.00	1949–1950, 1953–1982	31
0	0	0	1.60	0.26	0.16	0.06	1949–1951, 1953–1981	31
11	12	15	86.00	0.00	0.00	0.00	1949–2011	62

smaller than the corresponding arithmetic mean discharge, is adjusted for the days with zero flow, and gives greater weight to low daily mean discharges than high daily mean discharges. The QAH was determined from the daily mean discharge record by using a computer program based on DFLOW, a computer program developed by the U.S. Environmental Protection Agency (Rossman, 1990b). The QAH streamflow statistic is calculated as:

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

where

- Q_i is the mean streamflow for a given day,
- N_{nz} is the number of non-zero daily mean streamflows (Q_i) and
- N_t is the total number of daily mean streamflows (Q_i).

If no zero-flow days are in the record, the harmonic-mean flow is equal to the reciprocal of the mean of the reciprocals of the daily mean streamflow data.

The original set of 272 streamgages in Indiana (Fowler and Wilson, 2015) were screened for applicability in estimating the flow statistics used in this study so as to include only streamgages with less than 1,000 square miles drainage area, a period of record of at least 10 years, and unregulated, unmodified flows not affected by anthropogenic activities. This initial screening resulted in a beginning regression dataset of 153 streamgages. Subsequent review of the streamgage locations indicated 13 streamgages were redundant because these gages were located on the same stream reach or were located close together and nested in the same drainage basin. These redundant gages were removed from the regression dataset. During the regression analyses, residuals plots and influence and leverage values revealed 32 streamgages that were considered outliers, which on closer examination of drainage basin conditions were deemed unsuitable for use in the regression analysis. The modified flows were attributed to (1) surface-water or groundwater withdrawals for water supplies or irrigation, (2) wastewater discharges, (3) upstream impoundments, and (4) channelization. After screening, the remaining 108 long-term continuous-record streamgages used in the regression had low flows that were considered to be minimally controlled by anthropogenic effects.

Sources of Basin Characteristics

The utility of basin characteristics that describe physical, hydrologic, geologic, soils, land cover, climatic, and location features of the streamgage basins in estimating low-flow characteristics was investigated in the regression analyses. The basin characteristics were measured and calculated by means of a geographic information system (GIS) in a fashion

consistent with automated methods used to measure basin characteristics in the Indiana StreamStats application (U.S. Geological Survey, 2015). A comprehensive set of basin characteristics measured recently for use in a study to regionalize flow-duration statistics in Illinois and Indiana (Over and others, 2014) was available for many of the selected streamgages. In addition, several basin characteristics estimated from digital well-log records to describe aquifer properties for use in groundwater modeling (Arihood, 2009), were used in the regression analyses. Applicability of the basin characteristics to estimate low-flow statistics in Indiana was assessed using correlations and exploratory regressions. A subset of the basin characteristics (table 2) was determined to have the greatest potential as explanatory variables in the regression analyses.

Ultimately, nine basin characteristics were used in equations for estimating low-flow statistics in Indiana. Those characteristics included total drainage area of each basin (DRNAREA), the average hydraulic conductivity of the top 70 feet (ft) of unconsolidated deposits within each basin (K1), the average hydraulic conductivity of the entire thickness of the unconsolidated deposits within each basin (K2), the average transmissivity of the entire thickness of the unconsolidated deposits within each basin (T2), the average transmissivity of the entire thickness of the unconsolidated deposits within 1,000 ft from the streams (ST2), the index of permeability and thickness of the Quaternary surficial sediments (QSSPERMTHK), the slope of the basin (BSLDEM10M), the percent of forested land (FOREST), and the latitude of the basin outlet (LAT_OUT). The hydraulic conductivity and transmissivity of the unconsolidated deposits were evaluated in relation to these low-flow statistics because groundwater discharge makes up most of the low flows. The basin characteristics K1, K2, T2, and ST2 and QSSPERMTHK are described in further detail below.

Several basin characteristics derived from characteristics in well logs from the Indiana Water-Well Record (Indiana Department of Natural Resources, 2002) were tested to evaluate their potential use to explain the variability of low-flow statistics in Indiana streams. One of the basin characteristics, called K1, estimates the relative hydraulic conductivity of the first 70 ft of unconsolidated deposits within the basin based on the coarse or fine grained texture of these deposits as described in water well logs. K1 was included as a characteristic to describe the potential for precipitation to infiltrate into the unconsolidated deposits and flow through those deposits and discharge into streams. The K1 characteristic has a potential advantage over characteristics derived from soil or surface geologic characteristics because the latter measures assume that the hydraulic conductivity of unconsolidated deposits does not vary with depth and is identical to the permeability of soil and surface geologic characteristics.

The K1 characteristic was derived from a grid of texture-based estimates of average horizontal hydraulic conductivity for the unconsolidated deposits in each basin and can range from 1 to 100 feet per day (ft/d) (fig. 3). The K1 characteristic

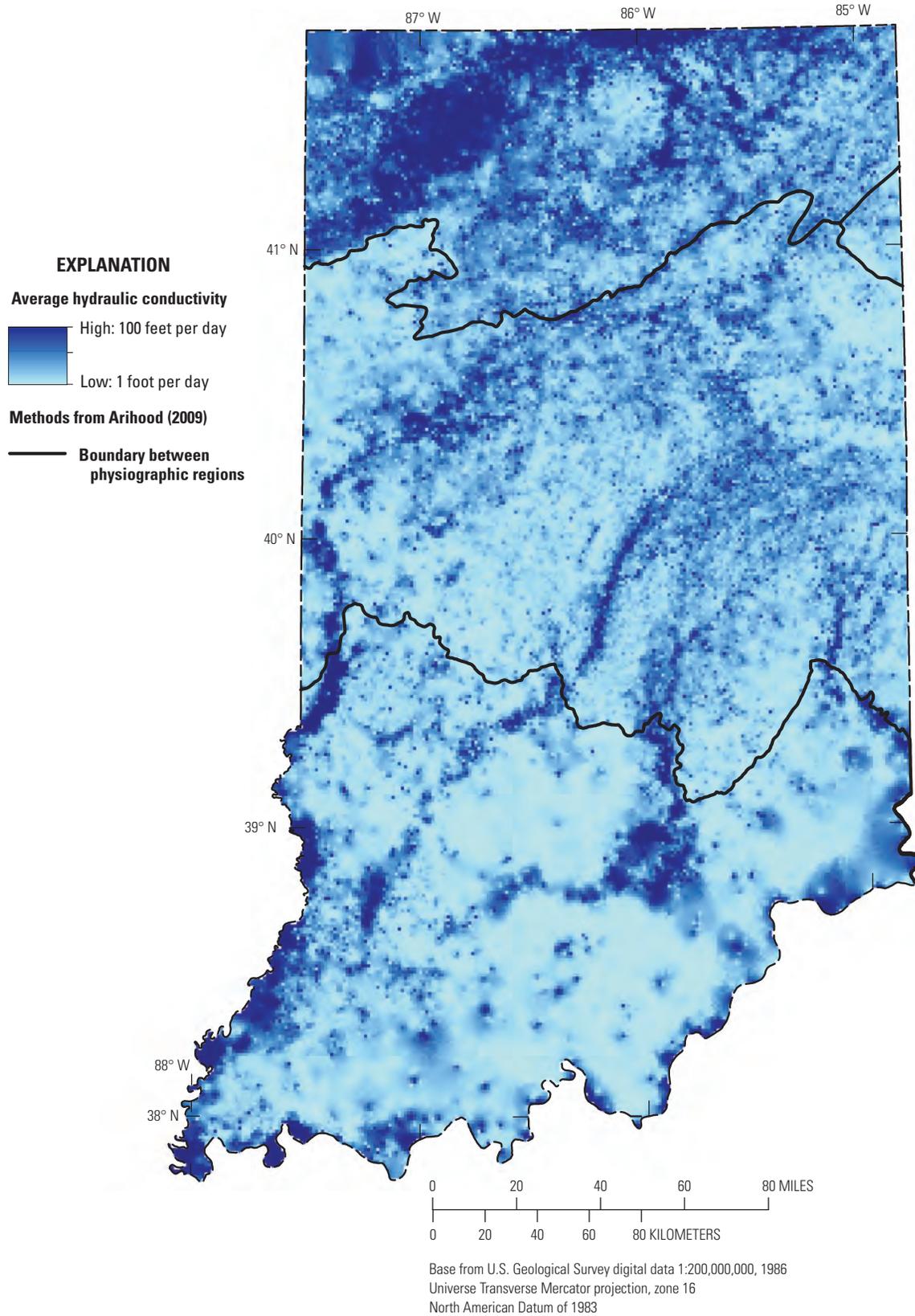


Figure 3. Grid of texture-based average horizontal hydraulic conductivity of the first 70 feet below land surface of unconsolidated deposits (K1) in Indiana.

Table 2. Description of basin characteristics evaluated as potential explanatory variables in regression equations for estimating selected low-flow statistics in Indiana.[mi², square miles; ft, feet; ft², feet squared; ft/d, feet per day; ft²/day, feet squared per day; %, percent; in., inches]

Basin characteristic	Description	Basin characteristic included in final regression equations
Total drainage area (DRNAREA) ¹	The total drainage area (mi ²) for a stream or river basin is measured in a horizontal plane enclosed by the drainage divide outlining the basin. Drainage areas for this report were measured in USGS StreamStats (http://water.usgs.gov/osw/streamstats/indiana.html).	Yes
Average hydraulic conductivity (K1)	Average hydraulic conductivity (ft/d) for the top 70 ft of unconsolidated deposits below land surface. It is a characteristic based on textural descriptions from water well driller logs obtained from state well log databases. Indiana Department of Natural Resources, Division of Water digital well database (http://www.in.gov/dnr/water/3595.htm). The values were converted into a statewide grid. Characteristic was calculated using method described in this report, as modified from Arihood (2009).	Yes
Average hydraulic conductivity (K2)	Average hydraulic conductivity (ft/d) of the full depth of unconsolidated deposits. It is a characteristic based on textural descriptions from water well driller logs obtained from state well log databases. Indiana Department of Natural Resources, Division of Water digital well database (http://www.in.gov/dnr/water/3595.htm). The values were converted into a statewide grid. Characteristic was calculated using method described in this report, as modified from Arihood (2009).	Yes
Average transmissivity near streams (ST2)	Average transmissivity (ft ² /d) of the full depth of unconsolidated deposits within 1,000 ft of the basin's stream channel. It is a characteristic based on textural descriptions from water well driller logs obtained from state well log databases. Indiana Department of Natural Resources, Division of Water digital well database (http://www.in.gov/dnr/water/3595.htm). The values were converted into a statewide grid. Characteristic was calculated using method described in this report, as modified from Arihood (2009).	Yes
Average transmissivity (T2)	Average transmissivity (ft ² /d) of the full depth of unconsolidated deposits. It is a characteristic based on textural descriptions from water well driller logs obtained from state well log databases. Indiana Department of Natural Resources, Division of Water digital well database (http://www.in.gov/dnr/water/3595.htm). The values were converted into a statewide grid. Characteristic was calculated using method described in this report, as modified from Arihood (2009).	Yes
Standard deviation of transmissivity	Standard deviation for the transmissivity of the entire thickness of unconsolidated deposits. It is an aquifer permeability characteristic based on textural descriptions from water well driller logs obtained from state well log databases. Indiana Department of Natural Resources, Division of Water digital well database (http://www.in.gov/dnr/water/3595/htm).	No
Index of permeability and thickness (QSSPERMTHK)	Index of permeability and thickness of Quaternary unconsolidated sediments in feet. Source data from Soller and Packard (1998), DDS-38, Digital Representation of a Map Showing the Thickness and Character of Quaternary Sediments in the Glaciated United States East of the Rocky Mountains (http://pubs.usgs.gov/dds/dds38/).	Yes
Percentage of coarse-grained stratified sediment	Coarse-grained stratified sediment (%), Multi-State: Surficial Quaternary sediments; Soller and Packard, (1998), DDS-38, Digital Representation of a Map Showing the Thickness and Character of Quaternary Sediments in the Glaciated United States East of the Rocky Mountains (http://pubs.usgs.gov/dds/dds38/). Source: IndianaMap, http://maps.indiana.edu/LayerGallery.html .	No
Index of Quaternary sediment permeability	$QSS_PermB = 100 * \text{fraction coarse-grained stratified sediment} + \text{fraction fine-grained stratified sediment} + \text{fraction glacial till} + 0.1 * \text{fraction exposed bedrock or sediment not of glacial origin}$ (Soller, D.R., and Berg R.C., 1992, Using regional geologic information to assess relative aquifer contamination potential—An example from the central United States: U.S. Geological Survey Open-File Report 92-694, map, scale 1:1,000,000).	No

Table 2. Description of basin characteristics evaluated as potential explanatory variables in regression equations for estimating selected low-flow statistics in Indiana.—Continued[mi², square miles; ft, feet; ft², feet squared; ft/d, feet per day; ft²/day, feet squared per day; %, percent; in., inches]

Basin characteristic	Description	Basin characteristic included in final regression equations
Drainage quality, wettest condition	Drainage quality (wettest condition), Source: CONUS-SOIL (Miller and White, 1998; Miller, D.A. and White, R.A., 1998; A conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling: Earth Interactions, v. 2, no. 2, p. 1–26).	No
Index of Drainage quality, wettest condition	Index = fraction “very poorly drained” + 2*fraction “poorly drained” + 4*fraction “somewhat poorly drained” + 8*fraction “moderately well-drained” + 16*fraction “well-drained” + 32*fraction “excessively drained”. Source: CONUS-SOIL (Miller, D.A. and White, R.A., 1998; A conterminous United States multi-layer soil characteristics data set for regional climate and hydrology modeling: Earth Interactions, v. 2, no. 2, p. 1–26).	No
Palustrine Emergent Wetlands	Percent palustrine emergent (PEM) wetlands, Source: (NWI, Ducks Unlimited, National Wetlands Inventory) Cowardin, L.M., Carter, V., Golet, F.C., and LaRoe, E.T., 1979, Classification of wetlands and deepwater habitats of the United States: Washington, D.C., U.S. Department of the Interior, Fish and Wildlife Service, Washington, DC., FWS/OBS–79/31.	No
Base flow index	Base-flow index grid for the conterminous United States, David M. Wolock, U.S. Geological Survey Open-File Report 03–263.	No
Slope (BSLDDEM10M)	Average basin slope (in percent) computed from U.S. Geological Survey 10-meter Digital Elevation Model (https://www.sciencebase.gov/catalog/item/4f70aa9fe4b058caae3f8de5).	Yes
Crop land	Percent crop land-National Land-Cover database, U.S. Geological Survey, 2007—National Land-Cover Database, 2001 (http://www.mrlc.gov).	No
Pasture	Percent pasture-National Land-Cover database, U.S. Geological Survey, 2007—National Land-Cover Database, 2001 (http://www.mrlc.gov).	No
Vegetated area	Percent vegetated NLCD 2001, classified as deciduous, evergreen, and mixed forest, developed open space (lawn grasses), shrub/scrub, and grassland/herbaceous (codes 41, 42, 43, 21, 52, and 71, respectively), U.S. Geological Survey, 2007—National Land-Cover Database, 2001 (http://www.mrlc.gov).	No
Forest (FOREST)	Percent forested NLCD 2001, classified as deciduous, evergreen, and mixed forest (codes 41, 42, and 43, respectively), U.S. Geological Survey, 2007—National Land-Cover Database, 2001 (http://www.mrlc.gov).	Yes
Sinkholes (INSINKHOLE)	Percent Sinkhole drainage area per basin, KARST_MM65_IN: Sinkhole Areas and Sinking-Stream Basins in Southern Indiana (Indiana Geological Survey, 1:126,720, Polygon Shapefile).	No
Sinking streams (INSINKING)	Percent Sinking stream drainage area, KARST_MM65_IN: Sinkhole Areas and Sinking-Stream Basins in Southern Indiana (Indiana Geological Survey, 1:126,720, Polygon Shapefile).	No
Precipitation	Mean annual precipitation (in.) 1981–2010, PRISM Climate Group at Oregon State University, United States Average Annual Precipitation, 1981–2010 (http://www.prism.oregonstate.edu/normals/).	No
Physiographic region (LOWREG)	Physiographic region categories: North Moraine (Northern region), Central Till Plain and the Maumee Lake Plain (Central region), Southern Hills Lowlands (Southern region), Gray, H.H., 2000, Physiographic Divisions of Indiana, Special Report 61, Indiana University, Indiana Geological Survey.	No
Urban land cover (URBAN)	Percent Urban land cover. Measured in USGS StreamStats, http://water.usgs.gov/osw/streamstats/indiana.html .	No
Impervious land cover (LC11IMP)	Percent Impervious land cover, from USGS National Land Cover Database (NLCD), 2011 (http://www.mrlc.gov).	No
Latitude (LAT_OUT)	Latitude of the basin outlet in decimal degrees. Measured in USGS StreamStats, http://water.usgs.gov/osw/streamstats/indiana.html .	Yes

¹Names in parentheses denote the corresponding name used in USGS StreamStats.

was computed by averaging the K1 grid values contained within each basin boundary. The statewide grid of 3,855 feet by 3,855 feet square was created by means of an inverse-distance-weighting interpolation of point values of texture-based average horizontal hydraulic conductivity. The point values are located at water wells and are based on the lithologic description recorded by the well driller at the time of well installation. Water-well records used to compute the gridded data are from the published, on-line Indiana Water-Well Record Database (Indiana Department of Natural Resources, 2002). The hydraulic-conductivity grids were created using 82,588 wells in Indiana (an average density of 2.7 wells per mi²) that penetrate at least 50 percent of the unconsolidated deposits.

Well records used for this study had been processed to classify texture-based lithology according to procedures described in Arihood (2009). The descriptions of lithology for each well record consisted of phrases that describe the depth range and type of unconsolidated deposits recorded during drilling, such as “sand from 10 to 20 ft “clay from 20 to 30 ft,” and “sand and gravel from 20 to 23 ft.” Depth ranges were converted into thicknesses of unconsolidated deposits. The lithologic descriptions were converted into a texture-based value of horizontal hydraulic conductivity at each well by the following procedure. All coarse-grained deposits (for example, sand, sand and gravel, gravel) were assumed to have a horizontal hydraulic conductivity of 100 ft/d, and all fine-grained deposits (for example, clay, silt, silty sand) were assumed to have a horizontal hydraulic conductivity of 1 ft/d. The two horizontal hydraulic conductivity values used to represent coarse-grained deposits and fine-grained deposits were considered sufficiently similar to real-world values and different from each other to represent a contrast in the aquifer and nonaquifer material. Next, using all unconsolidated lithologies for the first 70 ft below land surface at the well site, a thickness-weighted average of the two hydraulic conductivities was calculated by multiplying 100 ft/d times the thickness of the coarse-grained deposits, adding 1 ft/d multiplied by the thickness for the fine-grained deposits and dividing the total by the 70 ft thickness of the defined “top layer;” this depth was modified from the 100 ft depth used by Arihood (2009) to better approximate the upper part of coarse-grained unconsolidated deposits that were most likely to contribute flow to and discharge into streams. If less than 70 ft of unconsolidated deposits were present, then K1 was based on the thickness of deposits that were present. A more complete description of the procedure is provided in Arihood (2009, p. 5–10).

The K2 characteristic computed for each basin is the average thickness-weighted horizontal hydraulic conductivity of the entire sequence of unconsolidated deposits. K2 was calculated for each grid cell the same way that K1 is derived, except that the entire thickness of unconsolidated deposits (not just the top 70 ft) described in water well logs between

land surface and bedrock was used in the computations. The K2 characteristic was also computed by averaging the K2 grid values contained within each basin boundary. Values of K2 used to represent each basin in the regression equations are listed in table 3–1.

The T2 characteristic computed for each basin represents the average texture-based transmissivity of the entire thickness of unconsolidated deposits within a basin. T2 was calculated for each grid cell by multiplying the average texture-based horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits (K2) by the total thickness of unconsolidated deposits. The T2 characteristic was also computed by averaging the T2 grid values contained within each basin boundary. Values of T2 used to represent each basin in the regression equations are listed in table 3–1.

The ST2 characteristic computed for each basin represents the average texture-based transmissivity of the entire thickness of unconsolidated deposits within a 1,000-ft buffer zone around the stream network. The ST2 was calculated for each grid cell by multiplying the average texture-based horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits within the buffer zone around the stream network by the total thickness of unconsolidated deposits within the buffer zone. Values of ST2 used to represent each basin in the regression equations are listed in table 3–1.

A QSSPERMTHK characteristic that describes permeability and thickness of Quaternary surface sediments reported by Over and others (2014; referred to as PermBXThick in that report) was considered as an explanatory variable to estimate low-flow statistics in the regression analyses. The QSSPERMTHK characteristic was derived by combining qualitative descriptions of a property (“permeability”) into a continuous numerical index (Over and others, 2014). The QSSPERMTHK characteristic was computed as

$$\text{QSSPERMTHK} = \text{QSS_PermB} * \text{QSS_Thick} \quad (2)$$

where

The QSS_PermB characteristic is an area- and permeability-weighted index of the permeability of Quaternary surficial sediments computed using the areal fractions of the drainage basin in the following aquifer classes as:

$$\begin{aligned} \text{QSS_PermB} = & 100 * (\text{fraction coarse-grained stratified} \\ & \text{sediment}) + 1 * (\text{fraction fine-grained stratified sediment}) \\ & + 1 * (\text{fraction glacial till}) + 0.1 * (\text{fraction exposed bedrock} \\ & \text{or sediment not of glacial origin}) \end{aligned} \quad (3)$$

where

the weights (100, 1, and 0.1) are estimated relative permeability values (Soller and Berg, 1992).

The QSS_Thick characteristic is an area-weighted average of the thickness of the Quaternary surficial sediments, computed as follows:

$$\begin{aligned} \text{QSS_Thick} = & 25 * (\text{fraction } 0\text{--}50 \text{ feet thick}) + 75 * \\ & (\text{fraction } 50\text{--}100 \text{ feet thick}) + 150 * (\text{fraction } 100\text{--}200 \text{ feet} \\ & \text{thick}) + 300 * (\text{fraction } 200\text{--}400 \text{ feet thick}) + 500 * \\ & (\text{fraction } 400\text{--}600 \text{ feet thick}) \end{aligned} \quad (4)$$

where

the drainage-basin areal fractions of Quaternary surficial sediment types and thicknesses were computed in ArcMap from USGS Digital Data Series DDS 38 (Soller and Packard, 1998).

The definition of QSSPERMTHK is similar to transmissivity in that the index was computed as a permeability multiplied by a thickness of surficial sediments. As such, the index may seem to be correlated to and a duplication of the basin characteristic T2. The two characteristics are calculated differently, however, and the data for the calculation were obtained from different sources of information. QSSPERMTHK included the permeability of exposed bedrock or sediment not of glacial origin, whereas T2 does not. Also, the percent of fine- and coarse-grained deposits used to calculate QSSPERMTHK was based mostly on the texture of surficial deposits (Soller, 1992), whereas T2 is based on a dense distribution of well logs that are used to calculate the coarse- and fine-grain percentages from land surface to the bedrock surface.

Development of Equations for Estimating Selected Low-Flow Frequency Statistics and Harmonic-Mean Flows at Ungaged, Unregulated Stream Sites

Regression analyses were completed to develop equations to estimate selected low-flow statistics in Indiana. The regression equations relate the flow statistic or the probability of zero flow to selected basin characteristics. Multiple-linear least-squares regression methods including ordinary-least-squares (OLS), weighted-least-squares (WLS), and generalized-least-squares (GLS) methods were applied to develop equations to estimate the nonzero low-flow statistics. A logistic regression analysis was done to estimate the probability of selected low-flow frequencies being zero.

Logistic-Regression Analysis

Logistic-regression equations were developed to estimate the probability of the M1D10Y, M7D10Y, and M30D10Y being zero. The proportion of streamgages that have zero

low-flow statistics for M1D10Y, M7D10Y, and M30D10Y varies from 24 to 28 percent, which is within the appropriate range for use of logistic regressions—20 to 50 percent (Helsel and Hirsch, 2002). None of the QAH values were zero, so logistic-regression equations were not developed for QAH. Procedures and assumptions of logistic regression are described in detail elsewhere (Tasker, 1989; Ludwig and Tasker, 1993; SAS Institute, Inc., 1995; Allison, 1999; Hosmer and Lemeshow, 2000; Helsel and Hirsch, 2002; Hortness, 2006; Bent and Steeves, 2006; Eash and Barnes, 2012). Explanatory variables tested in the logistic regressions included numerous physical and climatic characteristics of each gaged basin (table 2). The dependent variables were the empirical probabilities, or observed frequencies, of zero values for the 1-, 7-, and 30-day annual low flows at each of the long-term streamgages, expressed as a decimal value ranging from zero to one. The observed frequencies of zero flows at a given streamgage were computed as the number of climate years in which the 1-, 7-, and 30-day low flows were zero, divided by the total number of climate years of record at the streamgage. The final logistic-regression model is expressed in the following form:

$$\text{PROBZ}_D = \frac{1}{1 + e^{(-b_0 - b_1(x_1) - b_2(x_2) \dots - b_n(x_n))}} \quad (5)$$

where

- PROBZ_D is the probability of the annual low-flow statistic being zero;
- D is the notation for the 1-, 7-, or 30-day, 10-year low flow: 1Q10, 7Q10, or 30Q10;
- e is the exponential constant, approximately equal to 2.71828;
- n is the number of basin-characteristic explanatory variables in the regression;
- x_i is the value of the i^{th} basin characteristic; and
- $b_0, b_1, \text{ and } b_n$ are the coefficients determined by maximizing the log-likelihood function.

The logistic-regression equations, which apply statewide, were developed using SAS statistical software (SAS Institute, Inc., 1995 and 2004). Weights were applied to each of the 108 streamgages included in the regressions. The period of record at each streamgage (table 1), which is the total number of climatic years, was used to compute the weight because the record length is a measure of the reliability of the estimated low-flow characteristic, in this case, the annual probability of zero flow. The weights were centered, or standardized, by dividing each streamgage period of record by the average period of record (38 years) for the 108 streamgages included in the logistic regression. The mean value of these standardized weights was, therefore, one (Stuckey, 2006).

Alternative model selections were done by iterative step-forward, backward-elimination, stepwise, and score procedures to evaluate numerous alternative combinations of the explanatory variables, both log-10 transformed and untransformed. The final logistic-regression equations (table 3) included DRNAREA, total drainage area; ST2, the average transmissivity, in ft²/d of the entire sequence of unconsolidated deposits within 1,000 ft of the basin drainage channel; and LAT_OUT, basin outlet latitude in degrees, minus 36, as the explanatory variables (table 2). All the regression coefficients are statistically different from zero at a level of significance of 0.05.

Analysis of the estimates provided by the logistic-regression estimating equations indicated performance in terms of the percentage of streamgages correctly classified as having zero or nonzero flow for each low-flow frequency statistic (M1D10Y, M7D10Y, and M30D10Y) varied depending on the particular ‘cutpoint’ probability used. Classification tables (SAS Institute, Inc., 1995; Hosmer and Lemeshow, 2000) (appendix 1, tables 1–1, 1–2 and 1–3) were developed by determining the numbers of correct and incorrect classifications of events (zero flow) and nonevents (nonzero flow) by comparison of the estimated probabilities to the zero and nonzero status of each low-flow statistic at the various probability cutpoints. Optimal cutpoints were selected as the cutpoint value that maximized the percentage correctly classified. The optimal cutpoints for all three low-flow frequencies were at or near 0.5 and are shown in bold numbers in the classification tables. A cutpoint of 0.5 was adopted for all the low-flow frequencies in this study.

The logistic-regression equations are applied to determine if the statistic of interest is either zero or nonzero. The estimated probabilities of zero flows are compared to the selected probability cutpoint of 0.5. If the estimated probability of zero flow

is greater than 0.5, then the low-flow statistic is estimated to be zero. If the probability of zero flow is less than or equal to 0.5, then the low-flow frequency should be estimated by using the multiple-linear-regression equations for estimating nonzero low-flow statistics described in the “Multiple-Linear-Regression Analyses” section. Zero values of the low-flow frequency statistics were not observed for the streamgages in Indiana used in this study (table 1) with more than 200 mi² total drainage area.

Multiple-Linear-Regression Analyses

Multiple-linear-regression analyses included exploratory use of OLS and WLS regressions to identify appropriate explanatory variables to include in the regression models. WLS regression was used to compensate for differences in the length of record (time-sampling error) of low-flow statistics. Regression equations for estimating QAH were finalized using WLS regression. Regression equations for the low-flow frequency statistics (M1D10Y, M7D10Y, and M30D10Y) were finalized using GLS regression, which compensates for differences in both the variability and reliability of, and correlation among, the low-flow frequency statistics at the streamgages included in the analysis.

The initial exploratory regressions incorporated the new estimated properties of the unconsolidated deposits described above in combination with other basin characteristics measured in a previous study (Over and others, 2014) that were all available for a subset of 54 of the streamgages distributed statewide. Iterative sequential exploratory regressions were done using subsets of the prospective explanatory variables that were all available at this subset of 54 streamgages. A final subset of 25 of the basin characteristics (table 2) that explained the most variance were measured for all of the 108 streamgages,

Table 3. Logistic-regression equations for estimating the annual probability of zero flow for selected low-flow frequencies for ungaged, unregulated streams statewide in Indiana.

[Low-flow statistic: M1D10Y, 1-day, 10-year low flow in cubic feet per second; M7D10Y, 1-day, 10-year low flow in cubic feet per second; M30D10Y, 30-day, 10-year low flow in cubic feet per second. Equation: PROBZ1Q10, probability the 1-day, 10-year low flow equals zero; *e*, exponential constant, approximately equal to 2.71828; DRNAREA, total drainage area in square miles; ST2, average transmissivity of full thickness of unconsolidated deposits within 1,000 feet of basin stream channel in square feet per day; LAT_OUT, latitude of the basin outlet in decimal degrees; PROBZ7Q10, probability the 7-day, 10-year low flow equals zero; PROBZ30Q10, probability the 30-day, 10-year low flow equals zero]

Low-flow statistic	Zero-flow probability equation	Number of gages used	Percent correctly classified
M1D10Y	$PROBZ1Q10 = \frac{1}{1 + e^{(3.57 \log_{10} DRNAREA + 5.25 \log_{10} ST2 + 2.38(LAT_OUT - 36) - 30.7)}}$	108	88.9
M7D10Y	$PROBZ7Q10 = \frac{1}{1 + e^{(3.62 \log_{10} DRNAREA + 4.64 \log_{10} ST2 + 2.33(LAT_OUT - 36) - 28.6)}}$	108	88.9
M30D10Y	$PROBZ30Q10 = \frac{1}{1 + e^{(2.76 \log_{10} DRNAREA + 5.42 \log_{10} ST2 + 1.69(LAT_OUT - 36) - 26.8)}}$	108	86.1

and these 25 basin characteristics were compared further in subsequent exploratory regressions.

Inspection of scatterplots depicting relations among dependent and explanatory variables and plots of residuals from initial linear regressions indicated that logarithmic (base 10) transformation of the dependent and most of the explanatory variables would be appropriate. This transformation generally helped improve the relations and increase the uniformity of variance of residuals about the regression line than before transformation. The relations between dependent and explanatory variables after transformation were consistent with the assumed linear form of the model. Only data for streamgages with nonzero low-flow statistics were included in the multiple linear regressions, because the logarithmic transformation is not possible for streamgages with low-flow statistics equaling zero.

The general form of the multiple linear regression models developed in this study follows:

$$\text{Log}(DQ_T) = b_0 + b_1 \log X_1 + b_2 \log X_2 + \dots + b_n \log X_n + \varepsilon, \quad (6)$$

where

- DQ_T is a low-flow statistic of D-days annual duration and T-years average return interval,
- b_0 is a constant (y-intercept),
- b_i (for $i=1$ to n) is the regression coefficient for the i^{th} explanatory variable,
- X_i (for $i=1$ to n) is the i^{th} explanatory variable,
- ε is a random error component equal to model plus time-sampling error, and
- n is the total number of explanatory variables.

The algebraically equivalent form when the log (base-10) transformation is used and when the equation is retransformed to the original units is:

$$DQ_T = 10^{b_0} X_1^{b_1} X_2^{b_2} \dots X_n^{b_n} \quad (7)$$

The alternative regression models were generated using several model-selection methods including all-possible-regression, forward-selection, backward-elimination, and stepwise-regression procedures (SAS Institute, Inc., 2004) by using the prospective explanatory variables (table 2). Various factors were considered in evaluating alternative regression models including (1) the coefficient of determination, the proportion of the variation in the response variable explained by the regression equation; (2) the standard error of the estimate, a measure of model-fitting error; (3) the prediction sum of squares statistic, commonly known as PRESS, a measure of model-prediction error; (4) the statistical significance of each alternative explanatory variable; (5) potential multicollinearity as indicated by the correlation of explanatory variables and the value of the variance-inflation factor (Montgomery and Peck, 1982) and condition index (Belsley and others, 1980); (6) the

effort and modeling benefit of determining the values of each additional explanatory variable; and (7) the hydrologic validity of the signs and magnitudes of the regression exponents.

The distributions of residuals from statewide drainage-area-only regressions indicated that defining low-flow regions of the State may improve regression-model performance. The physiographic regions of Indiana described by Gray (2000) (fig. 2) were tested and subsequently adopted as low-flow regions in this study according to the following classification based on the streamgage location.

- Northern region includes the Northern Moraine and Lake physiographic region.
- Central region includes the Central Till Plain and the Maumee Lake Plain physiographic regions.
- Southern region includes the Southern Hills and Lowlands physiographic region.

The Northern, Central, and Southern low-flow regions had statistically significant differences in the median values of key basin characteristics such as K1, QSSPERMTHK, BSLDEM10M and mean annual precipitation and in QAH yields. Regional differences were examined using region-indicator variables added to the statewide regressions that included terms to identify significantly different regression intercepts as well as interaction terms that identify significantly different regression slopes (coefficients) between regions. Generally when three or more explanatory variables were included in these trial regressions, the region-indicator variables were no longer statistically significant, as regional differences were explained by the information in the additional explanatory variables. No geographic biases were apparent for the residuals from the finalized regression models, as depicted in residuals (logarithm base-10 cubic foot per second [ft^3/s]) from the M7D10Y regression equation (fig. 4).

The regression models for estimating low-flow-frequency statistics (table 4) were finalized using GLS regression techniques (Stedinger and Tasker, 1985; and Tasker and Stedinger, 1989), which were implemented in the computer program WREG (Eng and others, 2009). Two of the major assumptions of OLS regression commonly are violated in regression of streamflow statistics because (1) the precision of the estimates of the statistics varies between streamgages (in part because of differences in record length), and (2) the statistics for a given streamgage frequently are correlated with those from other streamgages, because the same climatic conditions and weather events generally simultaneously affect most of the streams within a hydrologic region. Stedinger and Tasker (1985, 1986) have shown that where streamflow record lengths vary widely and flows (and therefore the flow statistics) at different streamgages are highly correlated, then GLS regression provides more accurate estimates of the regression coefficients, better estimates of the accuracy of the regression coefficients, and almost unbiased estimates of the model error when compared to OLS regression. GLS regression gives more weight to long-term streamgages than short-term

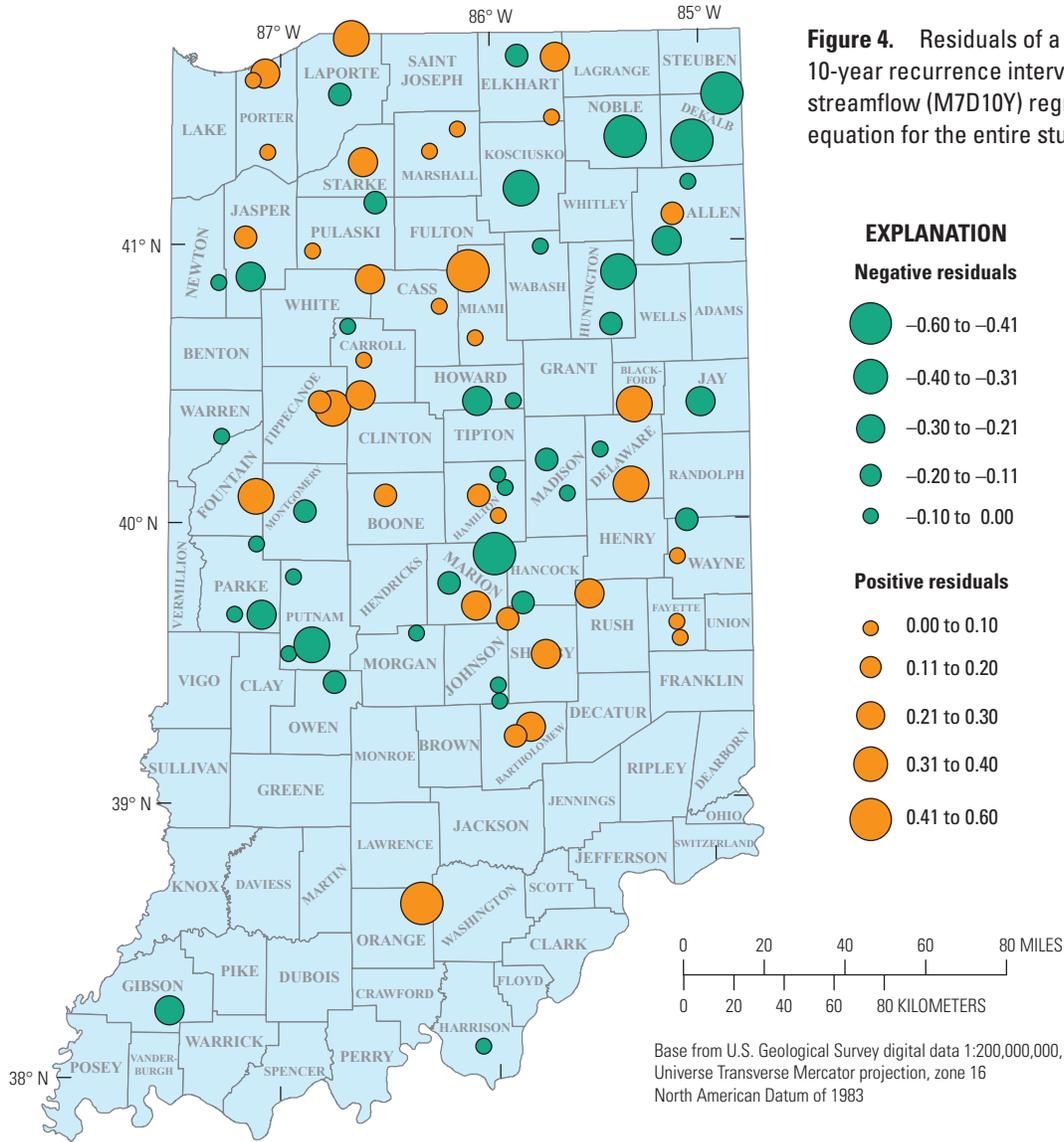


Figure 4. Residuals of a 7-day 10-year recurrence interval streamflow (M7D10Y) regression equation for the entire study area.

Table 4. Generalized-least-squares regression equations for estimating selected nonzero low-flow frequencies for ungaged, unregulated streams statewide in Indiana.

[R² pseudo, coefficient of determination after removing effects of time-sampling error; Sp, average standard error of prediction in log₁₀ cubic feet per second; log₁₀ ft³/s, log base-10 cubic feet per second; M1D10Y, 1-day, 10-year low flow in cubic feet per second; M7D10Y, 7-day, 10-year low flow in cubic feet per second; M30D10Y, 30-day, 10-year low flow in cubic feet per second; DRNAREA, total drainage area in square miles; K1, average horizontal hydraulic conductivity of first 70 feet of unconsolidated deposits in feet per day; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet; BSLDEM10M, average basin slope computed from the 10-meter digital-elevation model in percent]

Equation	Number of gages used	R ² pseudo (percent)	Sp (average standard error of prediction)		Standard model error (percent)
			(log ₁₀ ft ³ /s)	(percent)	
M1D10Y = 0.0000118 DRNAREA ^{1.22} K1 ^{1.76} BSLDEM10M ^{0.679} (QSSPERMTHK+1) ^{0.140}	78	92.7	0.2365	58.8	55.7
M7D10Y = 0.0000292 DRNAREA ^{1.16} K1 ^{1.61} BSLDEM10M ^{0.624} (QSSPERMTHK+1) ^{0.147}	78	92.7	0.2256	55.7	52.7
M30D10Y = 0.000166 DRNAREA ^{1.09} K1 ^{1.18} BSLDEM10M ^{0.541} (QSSPERMTHK+1) ^{0.173}	82	91.3	0.2457	61.5	58.6

streamgages and more weight to the streamgages where flows are least correlated to flows at other streamgages. The cross-correlations were estimated by means of a smoothing function relating correlation between streamflow time series to distance between streamgages having at least 30 years of concurrent record. The smoothing function was defined through an iterative graphical fit in the WREG application (fig. 5).

About 90 percent of the total-error variance in the regional low-flow frequency equations developed in this study was model-error variance. GLS regression in WREG enables partitioning of total regression error into model error and sampling error, which consists of both time- and space-sampling error. Model error arises from limitations of the model formulation and includes only the variance caused by the imperfect model, and it cannot be reduced by additional data collection. Time- and space-sampling error, however, are reduced through additional data collection by extending the period of data collection and by expanding the variety of basin characteristics of the sites where data are collected, respectively.

The final regression equations for estimating the low-flow statistics M1D10Y, M7D10Y, and M30D10Y that apply statewide for estimating nonzero flows include the variables DRNAREA, K1, BSLDEM10M, and QSSPERMTHK (table 4). A significance level of 0.05 for the F statistic of explanatory variables was used as the criteria for an explanatory variable to enter or stay in the model. These equations are used when the logistic equations indicate that nonzero flow is likely. Procedures for use of these equations are described in “Example Applications of the Estimating Equations.”

The QAH estimating equations were finalized using WLS regression with the weights set proportional to the length of streamgage record, standardized by dividing the length of record by the average length of record for all 108 gages (weighting was done as in the logistic-regression equations). The lowest QAH error was attained by defining separate equations for each low-flow region (table 5). Note the region in which the stream site is located determines which regional estimating equation to use—upstream parts of the basin may drain from a different low-flow region. The signs of regression coefficients were consistent with literature values.

Limitations and Accuracy of the Estimating Equations

The regional regression equations apply to unregulated streams in Indiana. The regression equations were developed using basin characteristics in a certain range of values (tables 6 and 7). Application of these equations in basins with characteristics beyond these ranges would be an extrapolation of the equations, and the magnitude of errors is unknown outside these ranges and may be large. Also, measurement of the basin characteristics should be done using the same GIS methods and layers as applied in this study. The appropriate GIS methods and layers are used in the Indiana StreamStats Web application (U.S. Geological Survey, 2015).

Caution is warranted in applying the equations for basins with characteristics approaching the maximum and minimum

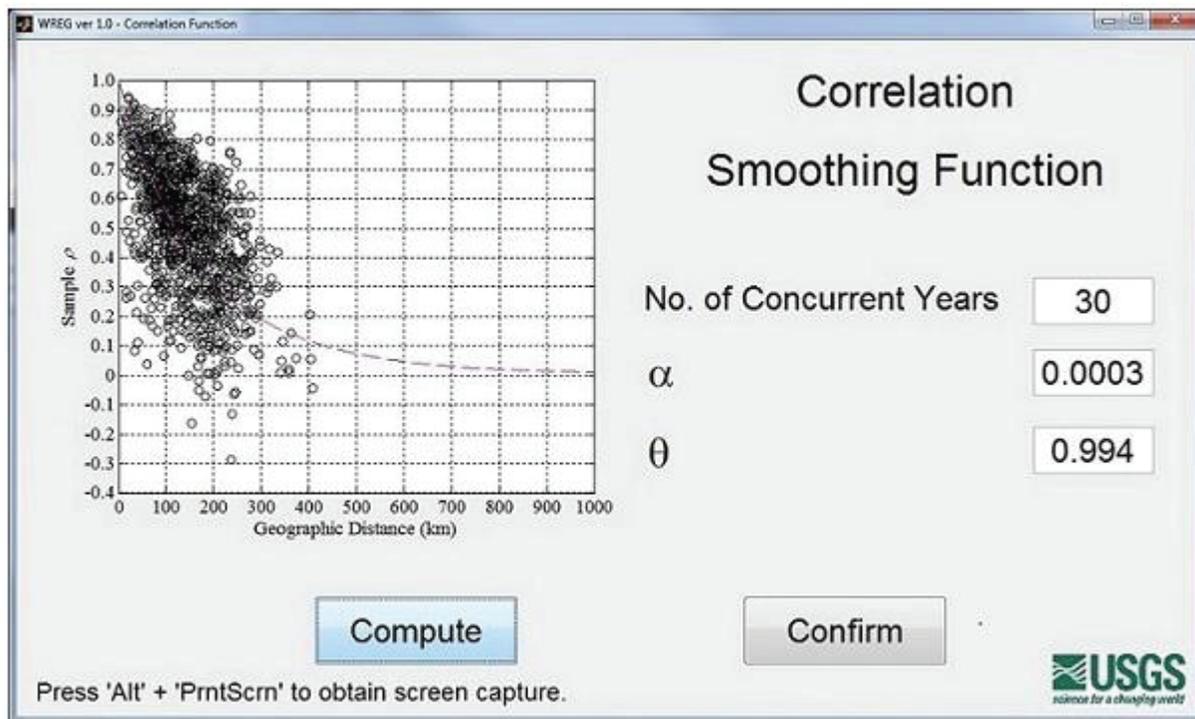


Figure 5. Screenshot of the weighted-multiple-linear regression program (WREG) smoothing function for generalized-least squares (GLS) correlation of the time series of annual minimum 7-day mean flows as a function of distance between 78 streamgages with 30 years of concurrent flow in Indiana.

Table 5. Weighted-least-squares regression equations for estimating harmonic-mean flow for ungaged, unregulated streams in each low-flow region in Indiana.

[R², coefficient of determination adjusted for error degrees of freedom in percent; log₁₀ ft³/s, log base-10 cubic feet per second; QAH, harmonic-mean flow in cubic feet per second; DRNAREA, total drainage area in square miles; T2, transmissivity of the entire thickness of unconsolidated deposits in square feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits feet per day; QSSPERMTHK, index of permeability and thickness of Quaternary surface sediment in feet; FOREST; portion of the drainage basin covered in forest in percent.]

Equation	Number of gages used	Adjusted R ² (percent)	Average standard error of estimate		Average standard error of prediction	
			(log ₁₀ ft ³ /s)	(percent)	(log ₁₀ ft ³ /s)	(percent)
Northern region						
QAH=0.00000332 DRNAREA ^{0.956} T2 ^{1.47}	21	92.1	0.1941	47.0	0.2243	55.3
Central region						
QAH=0.000736 DRNAREA ^{1.02} K2 ^{1.17} (QSSPERMTHK+1) ^{0.274}	55	96.5	0.1540	36.6	0.1645	39.3
Southern region						
QAH=0.233 DRNAREA ^{1.22} (FOREST+1) ^{-0.718}	32	93.1	0.2408	60.0	0.2636	66.7

Table 6. Basin-characteristic ranges for the logistic-regression equations for estimating the annual probability of zero flow for selected low-flow frequencies and generalized-least-squares regression equations for estimating selected nonzero low-flow frequencies for ungaged, unregulated streams statewide in Indiana.

[Statistic: PROBZ1Q10, probability the 1-day, 10-year low flow equals zero; PROBZ7Q10, probability the 7-day, 10-year low flow equals zero; PROBZ30Q10, probability the 30-day, 10-year low flow equals zero; M1D10Y, 1-day, 10-year low flow in cubic feet per second; M7D10Y, 7-day, 10-year low flow in cubic feet per second; M30D10Y, 30-day, 10-year low flow in cubic feet per second. DRNAREA, total drainage area in square miles; ST2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin’s stream channel in square feet per day; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; BSLDEM10M, average basin slope computed from the 10-meter digital-elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediment in feet; LAT_OUT, latitude of basin outlet in decimal degrees; Min, minimum; Max, maximum; —, not applicable]

Statistic	Number of gages	DRNAREA		ST2		K1		BSLDEM10M		QSSPERMTHK		LAT_OUT	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
PROBZ1Q10 PROBZ7Q10 PROBZ30Q10	108	2.99	856	409	7,650	—	—	—	—	—	—	38.1181	41.7476
M1D10Y	78	6.33	856	—	—	5.78	76.9	0.916	7.80	0	30,000	—	—
M7D10Y	78	6.33	856	—	—	5.78	76.9	0.916	7.80	0	30,000	—	—
M30D10Y	82	5.70	856	—	—	1.00	76.9	0.916	9.13	0	30,000	—	—

Table 7. Basin-characteristic ranges for the weighted-least-squares regression equations for estimating harmonic-mean flows for ungaged, unregulated streams in each low-flow region in Indiana.

[DRNAREA, total drainage area in square miles; T2, average transmissivity of the entire thickness of unconsolidated deposits in square feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet; FOREST, proportion of drainage basin covered in forest in percent; Min, minimum; Max, maximum; —, not applicable]

Region	Number of gages	DRNAREA		T2		K2		QSSPERMTHK		FOREST	
		Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Northern	21	6.33	856	1,700	7,590	—	—	—	—	—	—
Central	55	2.99	828	—	—	6.36	45.9	43.8	5,400	—	—
Southern	32	6.95	533	—	—	—	—	—	—	7.3	91.3

limits (tables 6 and 7), which are near the periphery of the multidimensional explanatory-variable sample spaces and may be extrapolations. Inconsistent results such as the M1D10Y regression estimate exceeding the less extreme M7D10Y estimate may be obtained. The same model form with four variables for all three low-flow frequencies was used to help minimize possible inconsistencies. Only the streamgages with nonzero low-flow statistics can be included in these log-transformed regressions. Consequently, the applicable sample spaces differ somewhat between the low-flow frequencies. The more extreme low-flow events have progressively fewer streamgages included in these regressions than the less extreme low-flow frequencies—the M1D10Y and M7D10Y regressions included 78 streamgages, whereas the M30D10Y regression used 82 streamgages (fig. 2).

The M1D10Y regression estimate exceeded the M7D10Y regression estimate at two streamgages with an average difference of 1.6 percent of the M7D10Y estimate, and the M7D10Y regression estimate exceeded the M30D10Y regression estimate at 12 streamgages with an average difference of 8.0 percent of the M30D10Y estimate, which is about 15 percent of the standard error of prediction of these equations (table 4). In cases where inconsistent results are obtained, the user should adopt the smaller value for both statistics given the differences are likely to be substantially less than the standard error of prediction.

Misclassification error for the logistic-regression equations (table 3) for the low-flow frequencies analyzed ranged from 11 to 14 percent (appendix 1 tables 1–1 to 1–3). The average prediction errors for the GLS-regression equations for estimating M1D10Y, M7D10Y, and M30D10Y were 58.8, 55.7, and 61.5 percent, respectively (table 4). The average standard error of prediction of the WLS-regression equations for estimating QAH for the Northern, Central, and Southern regions were 55.3, 39.3, and 66.7 percent, respectively (table 5).

Scatterplots of the M1D10Y, M7D10Y, and M30D10Y computed from flows measured at the streamgage and the values estimated from the regression equations are shown in fig. 6, and scatterplots of the QAH computed from measured flow and the values estimated from the WLS regression equations are shown in fig. 7. The upper range of values in these scatterplots shows some reduced residuals and a slight tendency to underestimate low flows, greater than approximately 50 ft³/s for M7D10Y, for example. The larger basins tend to have lower variability in streamflow than smaller basins.

The equations should be applied with caution in areas where streamflows are affected by hydrologic discontinuities such as large springs and sinks common to karst terrain in areas underlain by limestone in parts of southern Indiana. Streamflows in these areas may vary unpredictably in karst drainageways, and karst drainage features commonly are

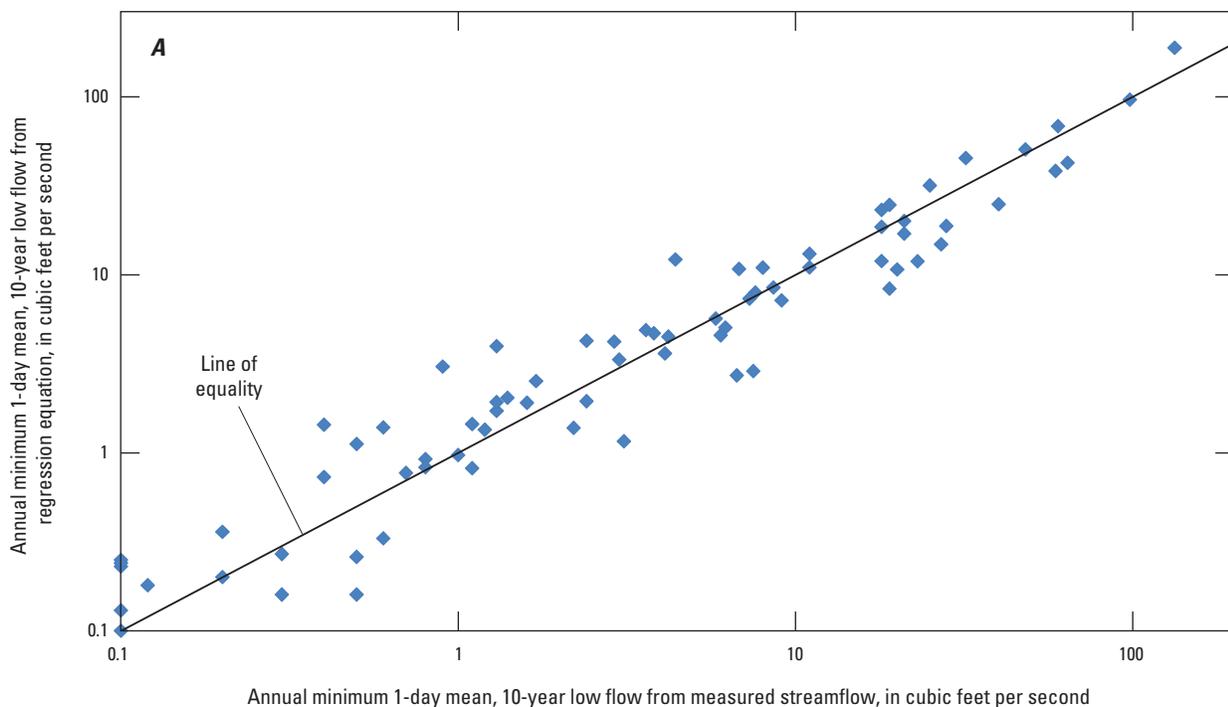


Figure 6. Comparison of selected low-flow frequencies estimated from measured streamflow and from regression equations for unregulated streams in Indiana, observed and predicted values for the following: *A*, annual minimum 1-day mean, 10-year low flows; *B*, annual minimum 7-day mean, 10-year low flows; and *C*, annual minimum 30-day mean, 10-year low flows.

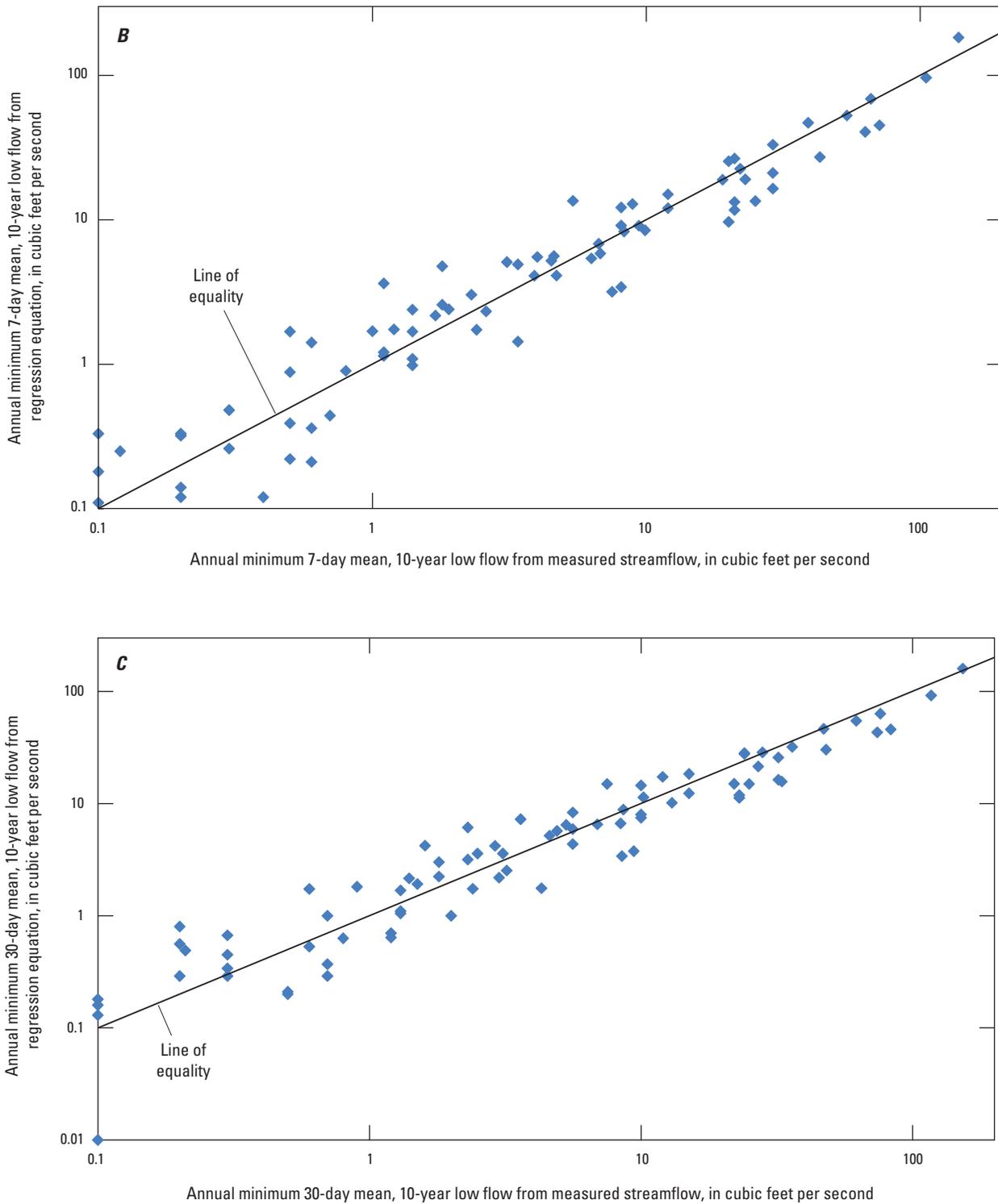


Figure 6. Comparison of selected low-flow frequencies estimated from measured streamflow and from regression equations for unregulated streams in Indiana, observed and predicted values for the following: *A*, annual minimum 1-day mean, 10-year low flows; *B*, annual minimum 7-day mean, 10-year low flows; and *C*, annual minimum 30-day mean, 10-year low flows.—Continued

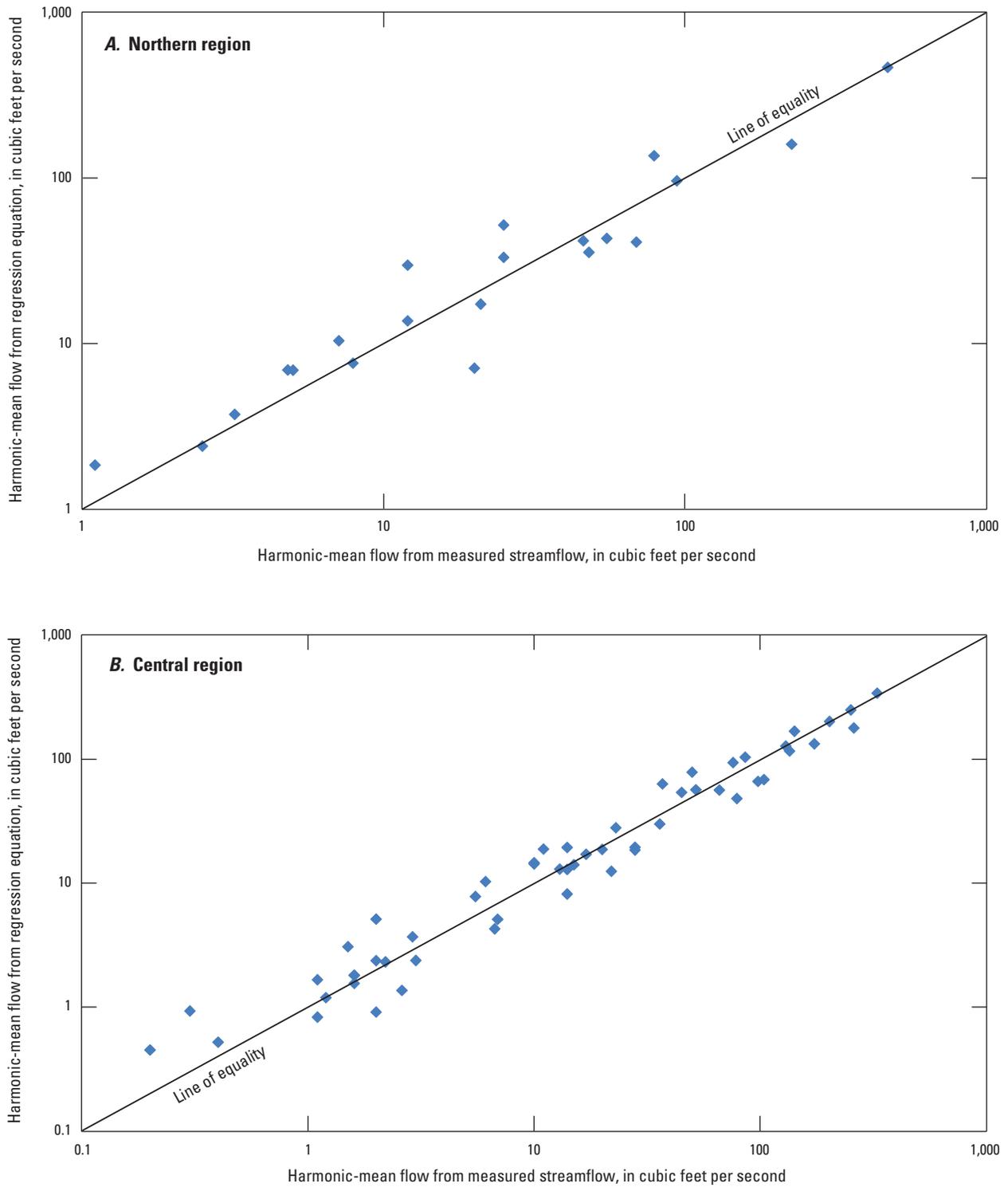


Figure 7. Comparison of harmonic-mean flows estimated from measured streamflow and from regression equations for unregulated streams in three low-flow regions of Indiana. *A*, Northern low-flow region; *B*, Central low-flow region; and *C*, Southern low-flow region.

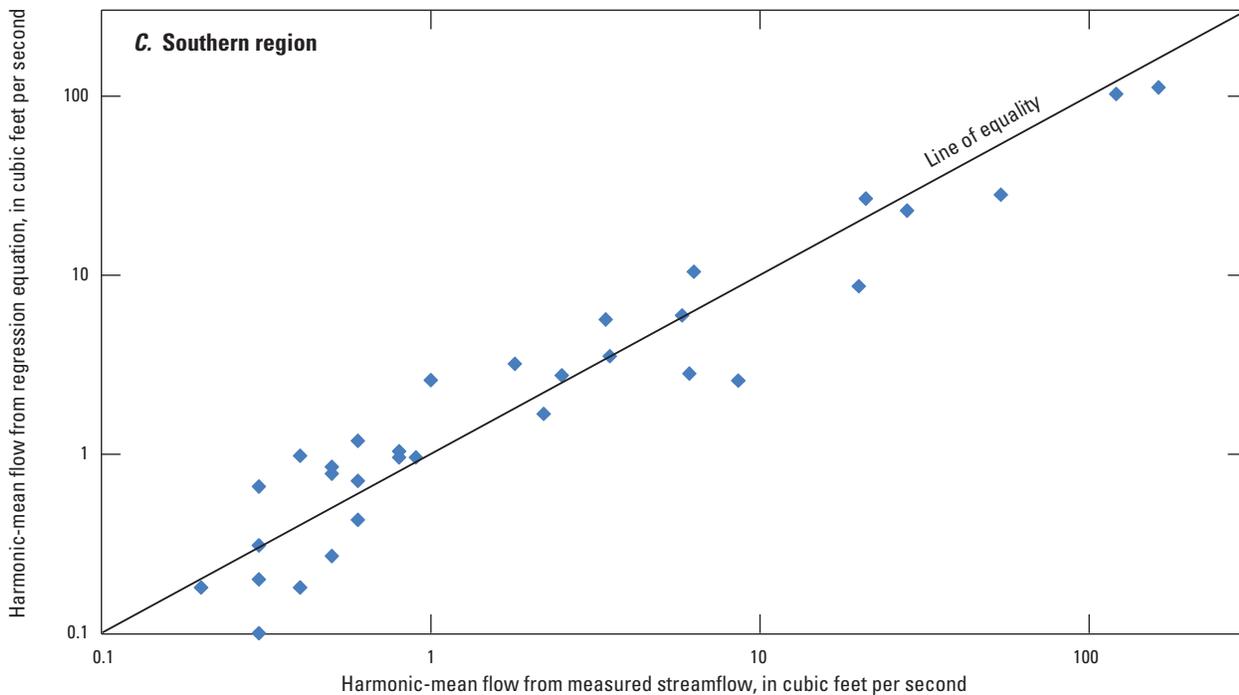


Figure 7. Comparison of harmonic-mean flows estimated from measured streamflow and from regression equations for unregulated streams in three low-flow regions of Indiana. *A*, Northern low-flow region; *B*, Central low-flow region; and *C*, Southern low-flow region.—Continued

unique to a particular stream site. An accurate basin drainage area in karst terrain may be difficult to determine solely on the basis of topographic divides (Bayless and others, 2014, 1994). Karst drainage features were present at 15 of 108 basins used for the QAH and probability of zero low-flow frequencies equations and 5 of 78 basins used for the M7D10Y equations. As much as 22 percent of basin area is sinkhole drainage area (variable *INSINKHOLE* in table 2) and as much as 98 percent of basin area is sinking stream drainage area (variable *INSINKING* in table 2) for all the low-flow statistics in the Southern region. Most of the basins had less than 10 percent of basin area with sinkholes and sinking streams, whereas three basins ranged from 22 to 98 percent of basin with these karst features. The increased QAH equation standard error in the Southern region is likely associated with the presence of karst drainage there. Values of the basin characteristics *INSINKHOLE* and *INSINKING* at the stream sites of interest in the Southern region should be considered, along with assessment of existing karst flow patterns (by field reconnaissance or literature review), in determining if the estimating equations are representative of basin conditions and appropriate to apply.

Urbanized basins were among the set of 108 streamgages included in the final regressions, located primarily in central Indiana. Urban development, which may include water pipelines, stormwater sewers, sanitary sewers, stormwater detention facilities along with impervious surfaces that can be distributed in widely varied patterns across a basin, can each potentially affect low flows differently—either decreasing or

increasing flows. Impervious surfaces tend to increase surface runoff and impede groundwater recharge, whereas impoundments may increase recharge, unless lined with an impermeable barrier.

There were 8 urbanized basins with the percent of basin area in impervious cover (*LC11IMP* in table 2) exceeding 10 percent, which ranged from 11 to 43 percent. Low flows at these streamgages appeared minimally affected by the urbanization based on (1) examination of basin conditions, checking for obvious alterations of flows as described above, (2) the residuals in the exploratory regressions and (3) comparisons of the QAH yields among basins. Some of the initial exploratory statewide regressions for QAH did include percent urban land cover (*URBAN* in table 2) as an explanatory variable; however it entered the equation as the fourth or fifth variable, and it did not substantially improve model performance. A one-variable regression equation for QAH yield as a function of percent impervious had a positive slope that was statistically significant (p -value 0.0108), but little of the variance was explained (0.05 adjusted R^2). In this increasing linear relation, with intercept of -1.05 log base-10 $\text{ft}^3/\text{s}/\text{mi}^2$ and slope of 0.339, the QAH yield for a basin with 40 percent impervious exceeds the QAH yield of a basin with 5 percent impervious by approximately 100 percent. The central tendency observed in Indiana is increased QAH with increased urbanization, though there is wide variability in the data as indicated in the little variance explained. This positive correlation is consistent with the relation of urban

development to low flows observed in Pennsylvania (Stuckey, 2006). In the final low-flow estimating equations, however, the trial explanatory variables LC11IMP and URBAN coefficients were not statistically different from zero at a level of significance of 0.05. Therefore, LC11IMP and URBAN were not used in the final equations.

The basin urbanization introduces some potential additional uncertainty in application of the regression equations. Increased urban development may increase the likelihood that facilities that appreciably modify low flows (wastewater discharges, withdrawals for water supply, and leaking water lines, for example) will be present. Three urban basins were removed from this dataset because of apparent alterations of low flows by withdrawals for irrigation (golf courses) and water supplies and channelization of streams. Careful investigation of basin conditions using maps of water-related infrastructure, field reconnaissance, and water-use data is required to identify facilities that may affect low flows. The uneven distribution of the urban basins across the State (most are in central Indiana) adds further uncertainty as to applicability of the equations in urban basins in other parts of the State.

The maximum percentage of basin area in developed land cover (URBAN in table 2) was 85 percent for the statewide logistic and M30D10Y GLS regressions and the QAH WLS regression in the Central region. Maximum URBAN values for basins used in the regressions were 49 percent for the MID10Y and M7D10Y statewide regressions, 35 percent for the QAH regression in the Northern region, and 8 percent for the QAH regression in the Southern region.

The maximum percentage of basin area in impervious cover (LC11IMP in table 2) was 43 percent for the statewide logistic and M30D10Y GLS regressions and the QAH WLS regression in the Central region. Maximum LC11IMP values for basins used in the regressions were 38 percent for the MID10Y and M7D10Y statewide regressions, 30 percent for the QAH regression in the Northern region, and 4 percent for the QAH regression in the Southern region.

The percent of basin area in impervious cover at the stream site of interest should be considered in determining if the estimating equations are representative of basin conditions and appropriate to apply. Given the uncertainties associated anthropogenic effects on low flows that may be encountered in urban basins, a limit of approximately 25 percent impervious area is suggested when applying the estimating equations.

Prediction Intervals

The prediction error is typically quantified as a probability that the true but unknown value of the flow statistic falls within specific bounds. For example, the 90-percent prediction interval defines the range of flow-statistic values within which there is a 90-percent probability the true value is located. The 90-percent prediction intervals are provided in the Indiana StreamStats Web application (U.S. Geological Survey, 2015).

Methods to estimate regression-equation prediction errors are outlined in the following equations and tables (Tasker

and Driver, 1988; Ludwig and Tasker, 1993; Ries and Dillow, 2006; Stuckey, 2006; Gotvald and others, 2009; Eash and Barnes, 2012). The 90-percent prediction interval for the true but unknown value of a flow statistic for an ungaged site is

$$\frac{Q}{T} < Q < QT \tag{8}$$

where

Q is the flow estimated for the ungaged site from the regression equation, and T is computed as follows:

$$T = 10^{\left[t_{(\alpha/2, n-p)} S_i \right]} \tag{9}$$

where

$t_{(\alpha/2, n-p)}$ is the critical value from the Student's t -distribution at α alpha level ($\alpha=0.10$ for 90-percent prediction intervals; critical t values may be obtained in many statistics references),

$n-p$ is the degrees of freedom with n streamgages included in the regression analysis; and

p is the number of explanatory variables in the equation plus 1 and,

S_i is the standard error of prediction for site i , S_i is computed as follows:

$$S_i = \left[MEV + X_i U X_i^T \right]^{0.5} \tag{10}$$

where

MEV is the model-error variance from GLS regression or the mean-square error from WLS regression equations developed in this study with user-defined weights;

X_i is the row vector for the streamgage 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 regression coefficients; and

X_i^T is the matrix algebra transpose of X_i (Ludwig and Tasker, 1993; Ries and Friesz, 2000; Gotvald and others, 2009; Eash and Barnes, 2012).

Similar to the average standard error of prediction (SEP), S_i is the square root of the sum of the model error variance and the sampling error variance for a single site i .

The $X_i U X_i^T$ term in equation 10 is referred to as the sampling error variance. The values of $t_{(\alpha/2, n-p)}$, MEV , and U required to determine prediction intervals for estimates obtained by the regression equations are given in appendix 2.

Example Applications of the Estimating Equations

Procedures for using the estimating equations are described in the following examples:

Example 1

An estimate of the M7D10Y low flow is needed for an ungaged stream site. The site has the following drainage-basin characteristics: Total drainage area (DRNAREA) is 40 mi²; and the average transmissivity in ft²/d of the full thickness of unconsolidated deposits within 1,000 ft of the basin's stream channel (ST2) is 1,300 ft²/d, and LAT_OUT is 39.000 degrees. First, confirm that the basin characteristics are in the acceptable ranges defined for this statistic (table 6), which is correct.

Next, compute the probability of the annual low flow (M7Q10) being zero using the logistic-regression equation (table 3):

$$PROBZ7Q10 = \frac{1}{1 + e^{(3.62 \log_{10} DRNAREA + 4.64 \log_{10} ST2 + 2.33(LAT_OUT - 36) - 28.6)}}$$

$$PROBZ7Q10 = \frac{1}{1 + e^{(3.62 \log_{10} (40) + 4.64 \log_{10} (1,300) + 2.33(39.000 - 36) - 28.6)}}$$

$$PROBZ7Q10 = \frac{1}{1 + e^{(-1.36)}}$$

$$PROBZ7Q10 = 0.796 > 0.5 \text{ cutpoint used in this study.}$$

Therefore, M7D10Y is estimated as zero.

Example 2

An estimate of the M7D10Y low flow is needed for an ungaged stream site. The site has the following drainage-basin characteristics: DRNAREA is 200 mi²; and ST2 is 1,300 ft²/d; LAT_OUT is 39.500 degrees; K1 is 23 ft/d; BSLDEM10M is 2.15 percent; and QSSPERMTHK is 3,200 ft. First, confirm that the basin characteristics are in the acceptable ranges defined for this statistic (table 6), which is correct.

Next, compute the probability of the M7D10Y low flow being zero using the logistic-regression equation (table 3):

$$PROBZ7Q10 = \frac{1}{1 + e^{(3.62 \log_{10} (200) + 4.64 \log_{10} (1,300) + 2.33(39.500 - 36) - 28.6)}}$$

$$PROBZ7Q10 = \frac{1}{1 + e^{(2.33)}}$$

$$PROBZ7Q10 = 0.088 < 0.5$$

Therefore, M7D10Y is nonzero and should be estimated using the GLS-regression equations (table 4).

$$M7D10Y = 0.0000292 DRNAREA^{1.16} K1^{1.61} BSLDEM10M^{0.624} (QSSPERMTHK + 1)^{0.147}$$

$$M7D10Y = 0.0000292 (200)^{1.16} (23)^{1.61} (2.15)^{0.624} (3200 + 1)^{0.147}$$

$$M7D10Y = 11.2 \text{ ft}^3/\text{s}$$

Example 3

Estimate the QAH and 90-percent prediction interval for a stream site in the Northern region with a drainage area of 17.9 mi² and a transmissivity T2 of 3,100 ft²/d. First, confirm that the basin characteristics are in the acceptable ranges defined for this statistic (table 7), which is correct.

Then apply the equation for the Northern region (table 5).

$$\text{QAH} = 0.00000332 \text{ DRNAREA}^{0.956} \text{T2}^{1.47}$$

$$\text{QAH} = 0.00000332 (17.9)^{0.956} (3,100)^{1.47}$$

$$\text{QAH} = 7.10 \text{ ft}^3/\text{s}.$$

To estimate the prediction interval, beginning with equation 10, the X_i vector is

$$X_i = \{1, \log_{10}(17.9), \log_{10}(3,100)\} = \{1, 1.252853, 3.491362\}$$

The model-error variance (MEV) for the Northern region from table 2–1 (appendix 2) is 0.0377, and the covariance matrix (U) from table 2–1 is

	Intercept	DRNAREA	T2
Intercept	1.751295583	0.02367123	-0.510231162
DRNAREA	0.023671229	0.00565783	-0.009847560
T2	-0.510231162	-0.00984756	0.150429478

Using matrix algebra, the sampling-error variance (the product $X_i U X_i^T$) is determined by multiplying the covariance matrix, U by X_i^T (transpose of X_i) to obtain $U X_i^T$, and then multiplying X_i by $U X_i^T$. In this example, $X_i U X_i^T$ is equal to 0.0042.

The standard error of prediction for this site as computed from equation 10 is

$$S_i = [0.0377 + 0.0042]^{0.5} = 0.2047,$$

and T from equation 9 is

$$T = 10^{(1.7341)(0.2047)} = 2.264,$$

where

the critical value ($t_{(\alpha/2, n-p)}$) from the Student's t -distribution for the 90-percent prediction interval is 1.7341 (table 2–1, appendix 2) for the regression equation in the Northern region. The 90-percent prediction interval is estimated from equation 8 as

$$(7.1)/(2.264) < \text{QAH} = 7.10 < (7.1)(2.264), \text{ or,}$$

the QAH equals 7.10 ft³/s, with a 90-percent prediction interval of 3.14 to 16.1 ft³/s.

An approximate estimate of the 90-percent prediction interval can be made based on the average standard error of prediction for the regression equation given in table 5 of 0.2243 log₁₀ units for the Northern region. Substituting into equation 9 gives

$$T = 10^{(1.7341)(0.2243)} = 2.449, \text{ and substituting this into equation 8 gives}$$

$$(7.10)/(2.449) < \text{QAH} = 7.10 < (7.10)(2.449), \text{ or,}$$

the 90-percent prediction interval is approximately 2.90 to 17.4 ft³/s.

Results and Discussion of Regression Analyses

Nine basin characteristics including DRNAREA, K1, K2, T2, ST2, QSSPERMTHK, BSLDEM10M, FOREST, and LAT_OUT were used in the final regression equations for estimating low-flow statistics in Indiana. The new texture-based estimates of aquifer properties K1, K2, T2 and ST2, along with QSSPERMTHK, worked well in combination with the other commonly used explanatory variables.

The K1 characteristic explained the most variance in M1D10Y, M7D10Y, and M30D10Y after DRNAREA, whereas T2 explained the most variance in QAH in the Northern region after DRNAREA. A possible reason for the predictive success of K1 and T2 may be related to the physical features represented by the two characteristics. K1 represents the overall hydraulic conductivity of the first 70 ft of unconsolidated deposits within the basin through which groundwater infiltrates and flows. In areas where the total unconsolidated thickness is 70 ft or less, K1 represents the hydraulic conductivity of the entire thickness of unconsolidated deposits. K1 measures the hydraulic conductivity of the shallow groundwater flow paths within the basin, but it does not consider the thickness of the flow paths. T1 was expected to be a superior predictor of low flows, because T1 factors in not only the K1 variable, but also the thickness of deposits. Thickness is relevant because theoretically, the thicker the deposits, the greater the capacity to store and transmit groundwater to streams. In practice, however, T1 is essentially the same number as K1, because most values of T1 in the State were derived by multiplying K1 by the same thickness, which is 70 ft. Therefore, no additional information is provided by the transmissivity characteristic T1, and it was not used for the final regressions.

The ST2 characteristic is most successful at estimating the probability of zero flows for M1D10Y, M7D10Y, and M30D10Y, possibly because ST2 includes information on hydraulic conductivity in the deeper parts of the unconsolidated deposits. When streamflow declines to zero flow, discharge of groundwater from all aquifers to surface water have been exhausted. The deep aquifers (commonly deeper than 70 ft in Indiana) typically are the last of the aquifers to stop discharging to streams, and as such, play an important role in determining whether or not streamflow is sustained.

The statewide low-flow frequency M1D10Y, M7D10Y, and M30D10Y estimating equations and the QAH estimating equations for the Central region include the new texture-based estimates of hydraulic conductivity of the unconsolidated deposits derived from the digital well-log records (K1 or K2) combined with the estimate of sediment permeability based on surficial deposits (QSSPERMTHK) defined by Soller and

Packard (1998). Although these estimated hydraulic conductivities have some similarities in the calculation methods, they were estimated independently from different source data—K1 and K2 from observed characteristics of unconsolidated deposits at depth, whereas QSSPERMTHK is mostly based on characteristics of the surficial deposits. These separate estimates of hydraulic conductivities bring complementary information into the regression models. In the final models, correlation of these variables was moderate (Pearson coefficient of 0.19 to 0.58), and harmful collinearity of these terms was not indicated.

Other new basin characteristics estimated from the well log records included total sand and gravel thickness and the vertical hydraulic conductivity. These characteristics were not as useful in explaining low flows, probably because these do not effectively represent or measure the ability of the entire flow system to transmit groundwater discharge to surface water as well as K1, K2, T2, and ST2 do.

FOREST was inversely related to QAH in the Southern region. Decreased low flows with increased forest cover has previously been reported (Price, 2011), likely because of the association of increased evapotranspiration with increased forest cover.

Other basin characteristics INSINKHOLE, INSINKING, URBAN, and LC11IMP were not included in the final estimating equations, because little or no information was added for the sampled basins. Elevated values for these ancillary variables outside the range sampled, however, may potentially affect low flows, and values of these basin characteristics should be considered in determining if the equations are representative of basin conditions and appropriate to use.

Summary

Knowledge of low-flow characteristics is needed for effective management of water resources in Indiana. Decisions related to waste-load allocations, discharge and withdrawal permits, water-supply planning, and in-stream flow requirements depend on estimates of low flows, and methods for estimating low-flow frequencies at ungaged stream sites are part of this need. The U.S. Geological Survey, in cooperation with the Indiana Department of Environmental Management, has developed regional regression equations for estimating selected low-flow frequencies and harmonic-mean flow using data at 108 streamgages in Indiana with data through the 2011 climate year.

Statewide logistic-regression equations are presented for estimating the annual probability of the selected low-flow frequency statistics being equal to zero. These equations use the explanatory variables total drainage area, average transmissivity of the full thickness of the unconsolidated deposits

within 1,000 feet of the basin's stream channel, and latitude of the basin outlet. The percentage of the streamgage low-flow statistics correctly classified as zero or nonzero using the logistic-regression equations ranged from 86.1 to 88.9 percent.

The statewide generalized-least-squares regression equations for estimating the nonzero low-flow frequency statistics used total drainage area, the average hydraulic conductivity of the top 70 feet of unconsolidated deposits, the index of permeability and thickness of the Quaternary surface sediments, and the slope of the basin as explanatory variables. The average standard errors of prediction of these regression equations ranged from 55.7 to 61.5 percent.

Weighted-least-squares regression equations were developed for estimating the harmonic-mean flows in three low-flow regions. The Northern region used total drainage area and the average transmissivity of the entire thickness of unconsolidated deposits as explanatory variables. The Central region used total drainage area, the average hydraulic conductivity of the entire thickness of unconsolidated deposits, and the index of permeability and thickness of the Quaternary surface sediments. The Southern region used total drainage area and the percent of the basin covered by forest. The average standard error of prediction for these equations ranged from 39.3 to 66.7 percent.

Advances in techniques for this report include use of the new texture-based aquifer properties and use of the logistic regression to estimate the probability of zero low-flow frequency statistics. With the exception of drainage area, the texture-based hydraulic conductivity characteristics explained more variance in the low-flow statistics than most of the other basin characteristics. The reason why may be that texture-based hydraulic conductivity characteristics are measures of fundamental properties controlling low flow that determine the ease with which groundwater can flow into and through the stream basin.

The regression equations apply only to stream sites with low flows unaffected by regulation and to drainage basins with values of basin characteristics within specified ranges. Caution is advised when applying the equations for basins with characteristics near the applicable limits, for basins with karst drainage features, and for urbanized basins. Extrapolations near and beyond the applicable basin characteristic limits will have unknown errors that may be large. The regression equations will be implemented in the U.S. Geological Survey StreamStats Web-based application for Indiana. StreamStats allows users to select a stream site on a map and automatically measure the needed basin characteristics and compute the estimated low-flow statistics and associated prediction intervals.

References Cited

- Allen, D.M., 1971, Mean square error of prediction as a criterion for selecting prediction variables: *Technometrics*, v. 13, p. 469–475.
- Allen, D.M., 1974, The relationship between variable selection and data augmentation and a method for prediction: *Technometrics*, v. 16, no. 1, p. 125–127.
- Allison, P.D., 1999, *Logistic regression using SAS—Theory and application*: Cary, N.C., SAS Institute, Inc., 288 p.
- Arihood, L.D., 2009, Processing, analysis, and general evaluation of well-driller logs for estimating hydrogeologic parameters of the glacial sediments in a ground-water flow model of the Lake Michigan Basin: U.S. Geological Survey Scientific Investigations Report 2008–5184, 26 p. [Also available at <http://pubs.usgs.gov/sir/2008/5184/>.]
- Arihood, L.D., and Glatfelter, D.R., 1986, Method for estimating low-flow characteristics of ungaged streams in Indiana: U.S. Geological Survey Open-File Report 86–323, 32 p.
- Arvin, D.V., 1989, Statistical summary of streamflow data for Indiana: U.S. Geological Survey Open-File Report 89–62, 964 p.
- Bayless, E.R., Cinotto, P.J., Ulery, R.L., Taylor, C.J., McCombs, G.K., Kim, M.H., and Nelson, H.L., Jr., 2014, Surface-water and karst groundwater interactions and streamflow-response simulations of the karst-influenced upper Lost River watershed, Orange County, Indiana: U.S. Geological Survey Scientific Investigations Report 2014–5028, 39 p. [Also available at <http://dx.doi.org/10.3133/sir20145028>.]
- Bayless, E.R., Taylor, C.J. and Hopkins, M.S., 1994, Directions of ground-water flow and locations of ground-water divides in the Lost River watershed near Orleans, Indiana: U.S. Geological Survey Water-Resources Investigations Report 94–4195, 25 p.
- Belsley, D.A., Kuh, E., and Welsch, R.E., 1980, *Regression diagnostics*: New York, N.Y., John Wiley and Sons, Inc., 292 p.
- Bent, G.C., and Steeves, P.A., 2006, A revised logistic regression equation and an automated procedure for mapping the probability of a stream flowing perennially in Massachusetts: U.S. Geological Survey Scientific Investigations Report 2006–5031, 107 p.

- Choquette, A.F., 1988, Regionalization of peak discharges for streams in Kentucky: U.S. Geological Survey Water-Resources Investigations Report 87–4209, 105 p.
- 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.
- 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., accessed September 1, 2015, at <http://pubs.usgs.gov/tm/tm4a8/>.
- Fowler, K.K., and Wilson, J.T., 1996, Low-flow characteristics of Indiana streams: U.S. Geological Survey Water-Resources Investigations Report 96–4128, 316 p.
- Fowler, K.K., and Wilson, J.T., 2015, Low-flow characteristics for selected streams in Indiana: U.S. Geological Survey Scientific Investigations Report 2014–5242, 353 p. [Also available at <http://dx.doi.org/10.3133/sir20145242>.]
- Gray, H.H., 2000, Physiographic divisions of Indiana: Indiana University, Indiana Geological Survey, Special Report 61, 14 p.
- Gotvald, A.J., Feaster, T.D., and Weaver, J.C., 2009, Magnitude and frequency of rural floods in the southeastern United States, 2006—Volume 1, Georgia: U.S. Geological Survey Scientific Investigations Report 2009–5043, 120 p.
- 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.
- Hardison, C.H., 1971, Prediction error of regression estimates of streamflow characteristics at ungaged sites: U.S. Geological Survey Professional Paper 750–C, 9 p.
- Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 510 p.
- Hoggatt, R.E., 1975, Drainage areas of Indiana streams: U.S. Geological Survey, 231 p.
- Hortness, J.E., 2006, Estimating low-flow frequency statistics for unregulated streams in Idaho: U.S. Geological Survey Scientific Investigations Report 2006–5035, 31 p.
- Indiana Department of Environmental Management, 2015, Water Quality in Indiana, Municipal Permits Overview: Indianapolis, Indiana, accessed December 13, 2015, at <http://in.gov/idem/cleanwater/2433.htm>.
- Indiana Department of Natural Resources, Division of Water, Water Rights and Use Section, Water Well Record Database, 2002: Website accessed June 17, 2016, at <http://www.in.gov/dnr/water/3595.htm>.
- Indiana Legislative Services Agency, 2013, Final report of the Water Resources Study Committee, November 2013: Indianapolis, Indiana, Indiana Legislative Services Agency, 9 p., accessed December 13, 2015, at <http://www.in.gov/legislative/interim/committee/reports/WRSCGB1.pdf>.
- Hosmer, D.W., and Lemeshow, S., 2000, Applied logistic regression: New York, Wiley, 373 p.
- 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, 19 p., accessed September 1, 2015, at http://ar.water.usgs.gov/LOCAL_REPORTS/WRI93-4013.pdf.
- Marquardt, D.W., 1970, Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation: *Technometrics*, v. 12, p. 591–612.
- Montgomery, D.C., and Peck, E.A., 1982, Introduction to linear regression analysis: New York, Wiley, 504 p.
- National Oceanic and Atmospheric Administration, 2014, Climatological data, annual summary, Indiana, 2010: Asheville, N.C., National Climatic Data Center.
- Over, T.M., Riley, J.D., Sharpe, J.B., and Arvin, Donald, 2014, Estimation of regional flow-duration curves for Indiana and Illinois: U.S. Geological Survey Scientific Investigations Report 2014–5177, 24 p. and additional downloads, tables 2–5, 8–13, and 18, accessed September 1, 2015, at <http://dx.doi.org/10.3133/sir20145177>.
- Price, Katie, 2011, Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions—A review—*Progress in Physical Geography*, v. 35, no. 4, p. 465–492.
- Ries, K.G., III, and Dillow, J.A., 2006, Magnitude and frequency of floods on nontidal streams in Delaware: U.S. Geological Survey Scientific Investigations Report 2006–5146, 57 p.
- Ries, K.G., III, and Friesz, P.J., 2000, Methods for estimating low-flow statistics for Massachusetts streams: U.S. Geological Survey Water-Resources Investigations Report 00–4135, 81 p.
- 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.
- SAS Institute, Inc., 1985, SAS User's Guide—Statistics: Cary, N.C., SAS Institute Inc., 956 p.

- SAS Institute, Inc., 1995, Logistic regression examples using the SAS system, (1st ed.), version 6: Cary, N.C., SAS Institute, Inc., 163 p.
- SAS Institute, Inc., 2004, SAS OnlineDoc® 9.1.2: Cary, N.C., SAS Institute Inc. [Available at https://support.sas.com/documentation/onlinedoc/91pdf/sasdoc_912/access_mysql_8139.pdf]
- Scheeringa, Kenneth, 2002, About Indiana Climate: Indiana State Climate Office at Purdue University, accessed May 5, 2016, at <http://iclimate.org/narrative.asp>.
- Soller, D.R., 1992, Text and references to accompany “Map showing thickness and character of Quaternary sediments in the glaciated United States east of the Rocky Mountains”: U.S. Geological Survey Bulletin 1921, 55 p. [Also available at <http://pubs.usgs.gov/bul/1921/report.pdf>.]
- Soller, D.R., and Berg R.C., 1992, Using regional geologic information to assess relative aquifer contamination potential—an example from the central United States: U.S. Geological Survey Open-File Report 92–694, map, scale 1:1,000,000.
- Soller, D.R., and Packard, P.H., 1998, Digital representation of a map showing thickness and character of Quaternary sediments in the glaciated United States east of the Rocky Mountains: U.S. Geological Survey Digital Data Series DDS–38. [Also available at <http://pubs.usgs.gov/dds/dds38/>.]
- Stedinger, J.R., and Tasker, G.D., 1985, Regional hydrologic regression, 1. Ordinary, weighted and generalized least squares compared: Water Resources Research, v. 21, no. 9, p. 1421–1432.
- Stedinger, J.R., and Tasker, G.D., 1986, Correlation to regional hydrologic analysis—Ordinary, weighted and generalized least squares compared: Water Resources Research, v. 22, no. 5, 844 p.
- Stewart, J.A., 1983, Low-flow characteristics of Indiana streams: U.S. Geological Survey Open-File Report 82–1007, 277 p.
- Stuckey, M.H., 2006, Low-flow, base-flow, and mean-flow regression equations for Pennsylvania streams: U.S. Geological Survey Scientific Investigations Report 2006–5130, 64 p.
- Tasker, G.D., 1989, Regionalization of low flow characteristics using logistic and GLS regression, in Kavvas, M.L., ed., New directions for surface water modeling: International Association of Hydrological Sciences Publication No. 181, p. 323–331.
- 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 September 1, 2015, at <http://onlinelibrary.wiley.com/doi/10.1111/j.1752-1688.1988.tb03026.x/pdf>.
- 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. Environmental Protection Agency, 2014, Clean Water Act National Pollutant Discharge Elimination System compliance monitoring: accessed September 2015 at <https://www.epa.gov/compliance/clean-water-act-national-pollutant-discharge-elimination-system-compliance-monitoring>.
- U.S. Geological Survey, 2015, The StreamStats program—Indiana, version 3: accessed December 13, 2015, at <http://water.usgs.gov/osw/streamstats/indiana.html>.
- Wittman, Jack, 2014, Water and economic development in Indiana—Modernizing the State’s approach to a critical resource, August 2014: prepared by Intera, Inc., Austin, Texas for the Indiana Chamber of Commerce, Indianapolis, Indiana, 92 p., accessed December 13, 2015, at <http://share.indianachamber.com/media/WaterStudyReport2014LoRes.pdf>.

Glossary

coefficient of multiple determination (R^2) The proportion of the variation in the dependent variable explained by the variables in a fitted regression model. Reported values are adjusted for error degrees of freedom.

climate year The 12-month period from April 1 through March 31. The climate year is designated by the calendar year in which the period ends.

cutpoint A cutpoint, also known as a cut-off point, is a probability selected to provide the largest number of correct classifications from a logistic regression. In this report, a cutpoint probability of 0.5 was selected to minimize the number of false negative classifications of non-events (non-zero flow) when actual flow was zero and secondly to minimize the number of false positive classifications of events (zero flow) when actual flow was non-zero.

level of significance The selected maximum probability of making a Type I error, or rejecting a true null hypothesis. Hypothesis tests were used to determine if statistically significant relations existed between dependent and explanatory variables of regression models.

local diversion A localized transfer of water, such as a water-supply withdrawal or wastewater releases, that artificially increase or decrease streamflow in a reach.

multicollinearity A high correlation (near linear dependencies) between two or more explanatory variables of a regression. Multicollinearity causes instability in the estimates of the least-squares regression coefficients.

multiple linear-regression A method used to model the linear relation between a dependent variable and one or more independent variables.

ordinary-least-squares regression A method of fitting a regression model in which the sum of squared residuals (see residual) is minimized.

prediction sum of squares (PRESS) statistic A measure of model-prediction error useful in regression-model selection. PRESS is computed by summing the square of the prediction residuals resulting from the series of predictions of each observation by regressions defined by using all other observations. Thus, each observation is left out of the regression dataset and is not used in prediction of that observation. This process simulates prediction with new data and is a form of data splitting useful for model validation (Allen, 1971 and 1974; and Montgomery and Peck, 1982).

regulated streamflow Streamflow controlled by upstream hydraulic structures such as dams, water-supply withdrawals, and wastewater discharges.

residual Difference between the observed value of the dependent variable and the value estimated using the regression. In this report, it is the difference between the value of the streamflow statistic computed using streamgauge data and value estimated using a regression model.

R^2 pseudo R^2 pseudo is a measure of variability explained by the regression, after removing the effects of time-sampling error (Griffis and Stedinger, 2007; Eng and others, 2009)

standard error of estimate A measure of model-fitting error; it is the standard deviation of the residuals of a regression adjusted for error degrees of freedom. Percentage values in this report were estimated using model root-mean-square error, or the square root of the sum of the squares of the residuals divided by the error degrees of freedom— $n-p$, where n is the number of observations and p is the number of explanatory variables in the regression plus 1 (Statistical Analysis System Institute, Inc., 1985) and information from Hardison (1971).

standard error of prediction A measure of model-prediction error; it was estimated as the square root of the PRESS divided by the degrees of freedom for error (Statistical Analysis System Institute, Inc., 1985; Montgomery and Peck, 1982; and Choquette, 1988). (See Prediction Sum of Squares (PRESS) Statistic.)

streamflow Discharge, measured as the volume of water that passes a given point in a given period of time (ft^3/s), that flows in a channel whether or not it is affected by diversion or regulation.

streamgauge An installation that provides systematic observations of streamflow, or river stage or index velocity, from which streamflow is computed. Continuous-record streamgages have equipment to make automated measurements generally at 15-minute to 1-hour intervals.

variance inflation factor (VIF) An indicator of multicollinearity; it is a measure of the combined effect of the dependencies among explanatory variables on the variance of each term in a regression model (Marquardt, 1970; and Montgomery and Peck, 1982).

water year The 12-month period from October 1 through September 30. The water year is designated by the calendar year in which the period ends.

Appendix 1. Classification Tables for Logistic-Regression Equations for Estimating the Probability of Zero Flow for Selected Low-Flow Frequencies in Indiana

Table 1–1. Classification table for the logistic-regression equation for estimating zero-flow probability for the 1-day, 10-year low flow in Indiana.

[Correct, number of observations correctly classified; incorrect, number of observations incorrectly classified; event, zero flow; nonevent, nonzero flow; percent correct, the frequency with which the equation correctly classifies the low-flow statistic for each probability cutpoint; percent sensitivity, the ratio of correctly classified events to the total number of events; percent specificity, the ratio of correctly classified nonevents to the total number of nonevents; percent false positive, the ratio of the number of nonevents incorrectly classified as events to the sum of all observations classified as events; percent false negative, the ratio of the number of events incorrectly classified as nonevents to the sum of all observations classified as nonevents; source; Statistical Analysis System Institute, Inc. 1995, p. 45–50; **bold** cutpoint is optimal probability level (0.5) adopted for use in this study]

Probability level (cutpoint)	Correct		Incorrect		Percent				
	Event	Nonevent	Event	Nonevent	Correct	Sensitivity	Specificity	False positive	False negative
0.05	30	50	28	0	74.1	100.0	64.1	48.3	0.0
0.1	29	57	21	1	79.6	96.7	73.1	42.0	1.7
0.2	28	63	15	2	84.3	93.3	80.8	34.9	3.1
0.3	28	67	11	2	88.0	93.3	85.9	28.2	2.9
0.4	26	68	10	4	87.0	86.7	87.2	27.8	5.6
0.5	25	71	7	5	88.9	83.3	91.0	21.9	6.6
0.6	24	72	6	6	88.9	80.0	92.3	20.0	7.7
0.7	21	72	6	9	86.1	70.0	92.3	22.2	11.1
0.8	18	75	3	12	86.1	60.0	96.2	14.3	13.8
0.9	14	76	2	16	83.3	46.7	97.4	12.5	17.4
0.99	2	78	0	28	74.1	6.7	100.0	0.0	26.4

Table 1–2. Classification table for the logistic-regression equation for estimating zero-flow probability for the 7-day, 10-year low flow in Indiana.

[Correct, number of observations correctly classified; incorrect, number of observations incorrectly classified; event, zero flow; nonevent, nonzero flow; percent correct, the frequency with which the equation correctly classifies the low-flow statistic for each probability cutpoint; percent sensitivity, the ratio of correctly classified events to the total number of events; percent specificity, the ratio of correctly classified nonevents to the total number of nonevents; percent false positive, the ratio of the number of nonevents incorrectly classified as events to the sum of all observations classified as events; percent false negative, the ratio of the number of events incorrectly classified as nonevents to the sum of all observations classified as nonevents; source; Statistical Analysis System Institute, Inc. 1995, p. 45–50; **bold** cutpoint is optimal probability level (0.5) adopted for use in this study]

Probability level (cutpoint)	Correct		Incorrect		Percent				
	Event	Nonevent	Event	Nonevent	Correct	Sensitivity	Specificity	False positive	False negative
0.05	30	51	27	0	75.0	100.0	65.4	47.4	0.0
0.1	28	57	21	2	78.7	93.3	73.1	42.9	3.4
0.2	28	63	15	2	84.3	93.3	80.8	34.9	3.1
0.3	27	67	11	3	87.0	90.0	85.9	28.9	4.3
0.4	26	69	9	4	88.0	86.7	88.5	25.7	5.5
0.5	25	71	7	5	88.9	83.3	91.0	21.9	6.6
0.6	22	72	6	8	87.0	73.3	92.3	21.4	10.0
0.7	18	74	4	12	85.2	60.0	94.9	18.2	14.0
0.8	18	75	3	12	86.1	60.0	96.2	14.3	13.8
0.9	13	77	1	17	83.3	43.3	98.7	7.1	18.1
0.99	1	78	0	29	73.1	3.3	100.0	0.0	27.1

Table 1–3. Classification table for the logistic-regression equation for estimating zero-flow probability for the 30-day, 10-year low flow in Indiana.

[Correct, number of observations correctly classified; incorrect, number of observations incorrectly classified; event, zero flow; nonevent, nonzero flow; percent correct, the frequency with which the equation correctly classifies the low-flow statistic for each probability cutpoint; percent sensitivity, the ratio of correctly classified events to the total number of events; percent specificity, the ratio of correctly classified nonevents to the total number of nonevents; percent false positive, the ratio of the number of nonevents incorrectly classified as events to the sum of all observations classified as events; percent false negative, the ratio of the number of events incorrectly classified as nonevents to the sum of all observations classified as nonevents; source; Statistical Analysis System Institute, Inc. 1995, p. 45–50; **bold** cutpoint is optimal probability level (0.5) adopted for use in this study]

Probability level (cutpoint)	Correct		Incorrect		Percent				
	Event	Nonevent	Event	Nonevent	Correct	Sensitivity	Specificity	False positive	False negative
0.05	25	50	32	1	69.4	96.2	61.0	56.1	2.0
0.1	25	61	21	1	79.6	96.2	74.4	45.7	1.6
0.2	23	67	15	3	83.3	88.5	81.7	39.5	4.3
0.3	21	69	13	5	83.3	80.8	84.1	38.2	6.8
0.4	18	73	9	8	84.3	69.2	89.0	33.3	9.9
0.5	17	76	6	9	86.1	65.4	92.7	26.1	10.6
0.6	14	77	5	12	84.3	53.8	93.9	26.3	13.5
0.7	11	80	2	15	84.3	42.3	97.6	15.4	15.8
0.8	10	80	2	16	83.3	38.5	97.6	16.7	16.7
0.9	4	82	6	16	79.6	20.0	93.2	60.0	16.3
0.99	0	82	0	26	75.9	0.0	100.0	0.0	24.1

Appendix 2. Values Needed to Determine 90-Percent Prediction Intervals for Multiple-Linear-Regression Estimates of Low-Flow Statistics in Indiana

Table 2-1. Values needed to determine the 90-percent prediction intervals using covariance matrices for estimates obtained from regional regression equations for ungaged, unregulated streams in Indiana.

[*n*, number of streamgages used in regression; *p*, number of parameters in regression, which equals the number of variables plus 1; *DF*, error degrees of freedom; *t*, the critical value from the Student's *t*-distribution for the 90-percent probability used in equation 9; *MEV*, regression model-error variance used in equation 10; *U*, covariance matrix as used in equation 10; MID10Y, 1-day, 10-year low flow, in cubic feet per second; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow; Intercept, y-axis intercept of the regression equation; DRNAREA, total drainage area; K1, average horizontal hydraulic conductivity of first 70 feet of unconsolidated deposits; BSLDEM10M, average basin slope computed from the 10-meter digital-elevation model; QSSPERMTHK, index of permeability and thickness of Quaternary surface sediments; T2, transmissivity of the full thickness of unconsolidated deposits; K2, average horizontal hydraulic conductivity of full thickness of unconsolidated deposits; FOREST, portion of the drainage basin covered in forest; Intercept, DRNAREA, K1, BSLDEM10M, QSSPERMTHK, T2, K2, and FOREST correspond to the parameters in the regression equations for the indicated flow statistics]

Flow statistic	<i>n</i>	<i>p</i>	<i>DF</i>	<i>t</i>	<i>MEV</i>	<i>U</i>					
MID10Y	78	5	73	1.6660	0.0509	Intercept	DRNAREA	K1	BSLDEM10M	QSSPER-MTHK+1	
						Intercept	0.067003107	-0.007410885	-0.031446359	-0.015454791	0.000194034
						DRNAREA	-0.007410885	0.002654278	0.001632632	0.001632180	-0.000353081
						K1	-0.031446359	0.001632632	0.024024953	0.003764889	-0.002390015
						BSLDEM10M	-0.015454791	0.001632180	0.003764889	0.021240063	-0.000385696
						QSSPER-MTHK+1	0.000194034	-0.000353081	-0.002390015	-0.000385696	0.001298481
M7D10Y	78	5	73	1.6660	0.0463	Intercept	DRNAREA	K1	BSLDEM10M	QSSPER-MTHK+1	
						Intercept	0.062062529	-0.006787552	-0.028842160	-0.014534454	0.000048980
						DRNAREA	-0.006787552	0.002401577	0.001483403	0.001551703	-0.000303679
						K1	-0.028842160	0.001483403	0.021899377	0.003555768	-0.002144432
						BSLDEM10M	-0.014534454	0.001551703	0.003555768	0.019626656	-0.000353075
						QSSPER-MTHK+1	0.000048980	-0.000303679	-0.002144432	-0.000353075	0.001190526
M30D10Y	82	5	77	1.6649	0.0558	Intercept	DRNAREA	K1	BSLDEM10M	QSSPER-MTHK+1	
						Intercept	0.050009472	-0.006420908	-0.021881637	-0.014770691	0.000025225
						DRNAREA	-0.006420908	0.002581803	0.001231275	0.001119220	-0.000327818
						K1	-0.021881637	0.001231275	0.018830663	0.004651446	-0.002707958
						BSLDEM10M	-0.014770691	0.001119220	0.004651446	0.020151268	-0.000306366
						QSSPER-MTHK+1	0.000025225	-0.000327818	-0.002707958	-0.000306366	0.001448410

Table 2-1. Values needed to determine the 90-percent prediction intervals using covariance matrices for estimates obtained from regional regression equations for ungaged, unregulated streams in Indiana.—Continued

[*n*, number of streamgages used in regression; *p*, number of parameters in regression, which equals the number of variables plus 1; *DF*, error degrees of freedom; *t*, the critical value from the Student's *t*-distribution for the 90-percent probability used in equation 9; *MEV*, regression model-error variance used in equation 10; *U*, covariance matrix as used in equation 10; M1D10Y, 1-day, 10-year low flow, in cubic feet per second; M7D10Y, 7-day, 10-year low flow; M30D10Y, 30-day, 10-year low flow; Intercept, y-axis intercept of the regression equation; DRNAREA, total drainage area; K1, average horizontal hydraulic conductivity of first 70 feet of unconsolidated deposits; BSLDEM10M, average basin slope computed from the 10-meter digital-elevation model; QSSPERMTHK, index of permeability and thickness of Quaternary surface sediments; T2, transmissivity of the full thickness of unconsolidated deposits; K2, average horizontal hydraulic conductivity of full thickness of unconsolidated deposits; FOREST, portion of the drainage basin covered in forest; Intercept, DRNAREA, K1, BSLDEM10M, QSSPERMTHK, T2, K2, and FOREST correspond to the parameters in the regression equations for the indicated flow statistics]

Flow statistic	<i>n</i>	<i>p</i>	<i>DF</i>	<i>t</i>	<i>MEV</i>	<i>U</i>				
Northern region										
QAH	21	3	18	1.7341	0.0377	Intercept	DRNAREA	T2		
						Intercept	1.751295583	0.023671230	-0.510231162	
						DRNAREA	0.023671230	0.005657832	-0.009847560	
						T2	-0.510231162	-0.009847560	0.150429478	
Central region										
QAH	55	4	51	1.6753	0.0237	Intercept	DRNAREA	K2	QSSPER-MTHK+1	
						Intercept	0.031998980	0.000280751	-0.020473881	-0.001710286
						DRNAREA	0.000280751	0.001535885	-0.001423533	-0.000551321
						K2	-0.020473881	-0.001423533	0.019993001	-0.001111597
						QSSPER-MTHK+1	-0.001710286	-0.000551321	-0.001111597	0.001514689
Southern region										
QAH	32	3	29	1.6991	0.0580	Intercept	DRNAREA	FOREST+1		
						Intercept	0.149095715	-0.031711111	-0.061882227	
						DRNAREA	-0.031711111	0.009230730	0.010543220	
						FOREST+1	-0.061882227	0.010543220	0.029217493	

Appendix 3. Values of Basin Characteristics That Were Significant Explanatory Variables in the Regression Equations

Table 3–1. Values of basin characteristic that were significant explanatory variables in the regression equations.

[Locations of streamgages are shown in figure 2. DRNAREA, drainage area in square miles; LAT_OUT, latitude of the basin outlet in decimal degrees; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; T2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel in square feet per day; BSLDEM10M, average basin slope computed from 10-meter digital elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet]

Station name	Map number (fig. 2)	DRNAREA from USGS Stream- Stats ¹	LAT_OUT
Whitewater River near Economy, Indiana	1	10.36	40.004
Whitewater River near Hagerstown, Indiana	2	59.49	39.874
Little Williams Creek at Connersville, Indiana	3	9.30	39.638
Whitewater River near Alpine, Indiana	4	521.85	39.579
South Hogan Creek near Dillsboro, Indiana	5	38.14	39.030
Indian-Kentuck Creek near Canaan, Indiana	6	27.47	38.878
Buck Creek near new Middletown, Indiana	7	65.35	38.120
Little Indian Creek near Galena, Indiana	8	17.09	38.322
Indian Creek near Corydon, Indiana	9	128.93	38.276
West Fork Blue River at Salem, Indiana	10	19.13	38.606
Middle Fork Anderson River at Bristow, Indiana	11	39.66	38.139
Crooked Creek near Santa Claus, Indiana	12	7.98	38.118
Pigeon Creek near Fort Branch	13	32.91	38.252
Little River near Huntington, Indiana	14	263.39	40.899
Salamonie River at Portland, Indiana	15	85.60	40.428
Salamonie River near Warren, Indiana	16	425.54	40.712
Little Mississinewa River at Union City, Indiana	17	9.76	40.196
Big Lick Creek near Hartford City, Indiana	18	28.96	40.422
Pipe Creek near Bunker Hill, Indiana	19	158.42	40.668
Eel River at North Manchester, Indiana	20	419.36	40.995
Weesau Creek near Deedsville, Indiana	21	9.26	40.909
Eel River near Logansport, Indiana	22	788.89	40.783
Rattlesnake Creek near Patton, Indiana	23	7.44	40.713
Deer Creek near Delphi, Indiana	24	275.54	40.590
Walnut Creek near Warsaw, Indiana	25	19.73	41.205
Tippecanoe River near Ora, Indiana	26	856.37	41.157
Little Indian Creek near Royal Center, Indiana	27	35.21	40.881
Big Monon Creek near Francesville, Indiana	28	153.37	40.984
Wildcat Creek near Jerome, Indiana	29	149.07	40.441
Kokomo Creek near Kokomo, Indiana	30	25.26	40.441
Wildcat Creek at Owasco, Indiana	31	395.56	40.465
South Fork Wildcat Creek near Lafayette, Indiana	32	242.93	40.418
Wildcat Creek near Lafayette, Indiana	33	794.28	40.441
Big Pine Creek near Willamsport, Indiana	34	323.34	40.318
East Fork Coal Creek near Hillsboro, Indiana	35	32.52	40.102

Table 3-1. Values of basin characteristic that were significant explanatory variables in the regression equations.

[Locations of streamgages are shown in figure 2. DRNAREA, drainage area in square miles; LAT_OUT, latitude of the basin outlet in decimal degrees; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; T2, Average transmissivity of the entire thickness of unconsolidated deposits in square feet per day; ST2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel in square feet per day; BSLDEM10M, average basin slope computed from 10-meter digital elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet]

K1	K2	T2	ST2	FOREST	BSLDEM10M	QSSPERMTHK
24	15	2,055	2,316	8.8	2.88	3,916.15
28	24	1,721	1,757	14.0	4.10	4,779.60
21	21	1,262	1,677	19.8	4.43	99.98
25	27	1,585	1,724	12.0	3.24	4,314.09
10	8	433	409	49.0	8.40	18.24
2	7	1,846	1,770	50.8	9.46	16.78
14	14	3,015	2,838	41.7	7.80	0.00
2	2	1,684	1,658	52.8	10.20	213.15
4	5	1,710	1,682	44.4	9.13	123.58
1	1	1,356	1,356	17.8	5.55	0.00
3	4	1,567	1,422	71.0	15.09	175.35
20	20	2,660	2,556	47.9	9.60	3.68
24	23	1,507	1,494	7.6	2.74	32.91
23	26	1,721	1,721	8.4	2.25	906.16
19	23	1,324	1,247	9.3	1.77	127.60
19	24	1,836	1,818	8.0	1.67	447.92
17	19	2,284	1,861	2.9	0.76	150.00
12	17	1,488	1,756	7.6	1.63	116.63
26	28	1,464	1,498	2.5	1.23	48.89
28	31	3,692	3,779	9.9	3.20	2,881.51
13	11	2,269	2,071	6.6	3.33	345.50
28	31	3,356	3,432	10.0	3.10	3,326.34
24	30	2,350	2,121	1.6	1.16	150.00
28	30	1,806	1,769	3.7	1.38	276.63
26	24	2,854	2,997	12.7	3.44	6,474.33
38	37	4,317	4,326	10.4	2.58	11,468.38
17	23	2,585	2,542	2.5	1.33	144.81
28	29	2,616	2,582	11.5	1.08	8,323.34
17	19	1,366	1,436	1.6	0.92	79.83
22	21	1,762	1,795	2.9	1.00	77.35
21	24	1,703	1,786	4.7	1.55	387.81
20	19	2,393	2,429	6.1	1.87	1,910.19
22	23	2,077	2,213	6.0	1.80	1,032.31
17	22	2,188	2,233	6.5	2.13	568.97
25	28	1,254	1,064	6.4	2.56	854.47

Table 3-1. Values of basin characteristic that were significant explanatory variables in the regression equations.—Continued

[Locations of streamgages are shown in figure 2. DRNAREA, drainage area in square miles; LAT_OUT, latitude of the basin outlet in decimal degrees; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; T2, Average transmissivity of the entire thickness of unconsolidated deposits in square feet per day; ST2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel in square feet per day; BSLDEM10M, average basin slope computed from 10-meter digital elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet]

Station name	Map number (fig. 2)	DRNAREA from USGS Stream- Stats ¹	LAT_OUT
Prairie Creek near Lebanon, Indiana	36	33.28	40.105
Sugar Creek at Crawfordsville, Indiana	37	510.16	40.049
Sugar Creek near Byron, Indiana	38	669.76	39.931
Big Raccoon Creek near Fincastle, Indiana	39	138.74	39.813
Big Raccoon Creek at Mansfield, Indiana	40	247.58	39.675
Little Raccoon Creek near Catlin, Indiana	41	133.36	39.677
Busseron Creek near Hymera, Indiana	42	16.87	39.215
Buck Creek near Muncie, Indiana	43	35.12	40.135
White River at Anderson, Indiana	44	406.19	40.106
Killbuck Creek near Gaston, Indiana	45	24.34	40.263
Pipe Creek at Frankton, Indiana	46	113.02	40.227
White River near Noblesville, Indiana	47	827.53	40.129
Cicero Creek near Arcadia, Indiana	48	130.81	40.176
Little Cicero Creek near Arcadia, Indiana	49	40.32	40.176
Hinkle Creek near Cicero, Indiana	50	18.29	40.101
Stony Creek near Noblesville, Indiana	51	51.05	40.029
Sugar Creek near Middletown, Indiana	52	5.70	40.041
Mud Creek at Indianapolis, Indiana	53	42.10	39.892
Pleasant Run at Brookville Road at Indianapolis, Indiana	54	11.23	39.764
Little Eagle Creek at Speedway, Indiana	55	18.81	39.788
Lick Creek at Indianapolis, Indiana	56	15.02	39.706
Little Buck Creek near Southport, Indiana	57	5.95	39.670
Little Buck Creek near Indianapolis, Indiana	58	17.07	39.667
White Lick Creek at Mooresville, Indiana	59	210.98	39.608
Beanblossom Creek at Beanblossom, Indiana	60	14.62	39.263
Bear Creek near Trevlac, Indiana	61	6.95	39.277
Plum Creek near Bainbridge, Indiana	62	2.99	39.762
Big Walnut Creek near Reelsville, Indiana	63	326.42	39.536
Mill Creek near Cataract, Indiana	64	244.43	39.434
Deer Creek near Putnamville, Indiana	65	59.04	39.568
Big Blue River at Carthage, Indiana	66	183.97	39.744
Big Blue River at Shelbyville, Indiana	67	420.44	39.529
Sugar Creek at New Palestine, Indiana	68	93.97	39.714
Buck Creek at Acton, Indiana	69	78.86	39.657
Youngs Creek near Edinburgh, Indiana	70	107.62	39.419
Sugar Creek near Edinburgh, Indiana	71	474.15	39.361
Flatrock River at Columbus, Indiana	72	532.71	39.235
Haw Creek near Clifford, Indiana	73	47.39	39.268

Table 3-1. Values of basin characteristic that were significant explanatory variables in the regression equations.

[Locations of streamgages are shown in figure 2. DRNAREA, drainage area in square miles; LAT_OUT, latitude of the basin outlet in decimal degrees; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; T2, Average transmissivity of the entire thickness of unconsolidated deposits in square feet per day; ST2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel in square feet per day; BSLDEM10M, average basin slope computed from 10-meter digital elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet]

K1	K2	T2	ST2	FOREST	BSLDEM10M	QSSPERMTHK
12	12	1,466	1,531	1.7	1.38	3,157.93
15	17	1,575	1,519	3.9	1.54	2,132.14
15	17	1,478	1,447	7.5	2.17	2,015.10
13	16	1,698	1,984	6.4	2.36	731.95
11	14	1,539	1,724	18.5	4.19	433.95
16	19	1,846	1,933	25.4	5.17	448.42
6	6	817	800	27.6	3.53	39.84
34	33	2,700	2,557	10.6	3.23	1,540.36
31	29	2,506	2,482	7.6	1.99	1,058.64
45	46	1,619	1,580	4.4	1.27	57.00
32	35	2,582	2,425	4.3	1.54	295.69
31	32	2,471	2,461	6.0	1.74	835.56
12	17	2,085	2,077	1.8	1.06	396.22
13	21	2,496	2,388	2.1	1.34	248.55
9	15	1,614	1,628	6.9	2.37	979.35
34	35	1,739	1,925	3.7	1.23	474.10
27	16	2,266	2,321	4.1	1.52	298.95
34	39	2,047	2,288	5.2	1.64	70.63
10	15	1,371	1,337	0.5	2.03	156.40
16	20	1,615	1,685	4.1	1.77	1,724.40
11	17	1,468	1,644	4.9	2.34	194.96
9	13	1,599	1,778	5.5	2.54	300.00
17	20	1,725	1,860	3.7	2.35	1,194.00
13	17	1,404	1,512	8.6	2.40	876.25
8	9	1,189	1,262	75.6	11.75	11.24
16	18	892	498	90.3	18.06	0.00
4	7	1,432	1,838	8.1	3.10	69.16
10	12	1,471	1,448	22.5	4.34	420.97
9	12	1,511	1,524	14.4	2.73	473.74
6	6	1,165	1,187	31.7	4.98	32.38
24	24	1,994	2,048	10.5	2.79	5,395.32
23	24	1,832	1,833	8.1	2.15	3,242.19
26	26	2,601	2,621	5.3	1.26	503.40
13	18	2,216	2,211	4.7	1.52	1,075.50
15	19	1,424	1,403	5.3	1.86	1,203.95
24	27	2,203	2,237	6.3	1.51	2,286.84
18	19	1,258	1,254	6.9	1.91	2,327.03
10	13	1,097	1,140	6.7	1.88	217.70

Table 3-1. Values of basin characteristic that were significant explanatory variables in the regression equations.—Continued

[Locations of streamgages are shown in figure 2. DRNAREA, drainage area in square miles; LAT_OUT, latitude of the basin outlet in decimal degrees; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; T2, Average transmissivity of the entire thickness of unconsolidated deposits in square feet per day; ST2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel in square feet per day; BSLDEM10M, average basin slope computed from 10-meter digital elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet]

Station name	Map number (fig. 2)	DRNAREA from USGS Stream- Stats ¹	LAT_OUT
Clifty Creek at Hartsville, Indiana	74	91.32	39.275
Sand Creek near Brewersville, Indiana	75	154.48	39.084
Graham Creek near Vernon, Indiana	76	77.20	38.930
Harberts Creek near Madison, Indiana	77	9.25	38.782
Brush Creek near Nebraska, Indiana	78	11.32	39.070
Back Creek at Leesville, Indiana	79	24.11	38.847
South Fork Salt Creek at Kurtz, Indiana	80	38.13	38.963
North Fork Salt Creek near Belmont, Indiana	81	119.81	39.150
Stephens Creek near Bloomington, Indiana	82	10.83	39.164
Indian Creek near Springville, Indiana	83	60.71	38.951
Lost River near Leipsic, Indiana	84	35.07	38.636
Hall Creek near St. Anthony, Indiana	85	21.75	38.363
Flat Creek near Otwell, Indiana	86	21.35	38.437
Little Calumet River at Porter, Indiana	87	65.97	41.622
Salt Creek near McCool, Indiana	88	75.19	41.597
Galena River near Laporte, Indiana	89	17.87	41.748
Little Elkhart River at Middlebury, Indiana	90	97.56	41.675
Pine Creek near Elkhart, Indiana	91	30.23	41.681
Rimmell Branch near Albion, Indiana	92	10.96	41.385
Solomon Creek near Syracuse, Indiana	93	36.22	41.458
Fish Creek at Hamilton, Indiana	94	37.42	41.532
Cedar Creek at Auburn, Indiana	95	87.32	41.366
Cedar Creek near Cedarville, Indiana	96	269.50	41.219
Harber Ditch at Fort Wayne, Indiana	97	21.89	41.008
Spy Run Creek at Fort Wayne, Indiana	98	13.94	41.105
Kingsbury Creek near Laporte, Indiana	99	6.33	41.547
Yellow River near Bremen, Indiana	100	134.66	41.420
Yellow River at Plymouth, Indiana	101	293.85	41.340
Yellow River at Knox, Indiana	102	435.07	41.303
Cobb Ditch near Kouts, Indiana	103	30.62	41.339
Iroquois River at Rosebud, Indiana	104	38.14	41.033
Iroquois River at Rensselaer, Indiana	105	204.66	40.934
Slough Creek near Collegeville, Indiana	106	83.51	40.892
Carpenter Creek at Egypt, Indiana	107	44.84	40.866
Iroquois River near Foresman, Indiana	108	448.74	40.870

¹USGS StreamStats is a web based application that provides streamflow statistics for streams in Indiana (<http://water.usgs.gov/osw/streamstats/indiana.html>).

Table 3-1. Values of basin characteristic that were significant explanatory variables in the regression equations.

[Locations of streamgages are shown in figure 2. DRNAREA, drainage area in square miles; LAT_OUT, latitude of the basin outlet in decimal degrees; K1, average texture-based horizontal hydraulic conductivity for the top 70 feet of unconsolidated deposits below land surface in feet per day; K2, average horizontal hydraulic conductivity of the entire thickness of unconsolidated deposits in feet per day; T2, Average transmissivity of the entire thickness of unconsolidated deposits in square feet per day; ST2, Average transmissivity of the entire thickness of unconsolidated deposits within 1,000 feet of the basin's stream channel in square feet per day; BSLDEM10M, average basin slope computed from 10-meter digital elevation model in percent; QSSPERMTHK, index of permeability and thickness of Quaternary surficial sediments in feet]

K1	K2	T2	ST2	FOREST	BSLDEM10M	QSSPERMTHK
6	6	790	800	8.3	2.42	42.80
8	10	1,214	1,246	25.2	4.12	91.70
10	18	1,936	2,215	55.4	3.93	20.89
3	13	2,640	2,811	53.3	2.39	25.00
8	14	1,175	1,238	39.8	5.41	24.83
4	5	2,943	3,102	46.1	9.63	2.87
16	16	1,376	1,367	64.9	13.51	162.40
7	7	770	754	87.6	18.30	90.28
5	5	1,548	1,249	81.7	21.03	0.00
3	3	2,100	2,211	67.4	10.65	106.33
12	13	1,612	1,526	12.1	3.16	0.00
21	21	1,607	1,650	29.7	11.85	208.59
15	16	3,121	3,633	24.4	4.08	25.87
41	47	3,966	3,745	22.6	3.79	4,733.41
58	60	4,063	3,961	16.0	3.50	1,223.62
40	53	3,100	3,030	29.9	4.28	25,100.39
45	38	3,393	3,532	3.7	1.71	21,061.27
42	40	3,444	3,573	5.9	2.87	14,230.44
16	13	1,704	1,436	13.7	3.51	9,738.16
54	38	7,588	7,652	6.1	2.20	22,697.96
29	30	2,008	2,507	6.9	3.89	5,591.18
21	24	2,927	2,984	11.7	2.98	4,951.33
27	31	3,961	4,015	13.7	3.06	3,599.53
15	18	1,208	1,171	3.8	1.42	75.00
39	41	3,572	3,780	10.2	3.20	169.39
77	73	3,925	3,570	7.8	1.75	30,000.00
23	31	2,378	2,327	8.3	1.62	1,489.37
26	30	2,955	2,945	7.4	1.53	5,296.52
31	33	3,239	3,279	11.6	1.81	7,202.86
72	78	4,005	3,993	12.3	2.48	4,323.54
29	28	2,007	2,049	9.2	1.24	3,833.57
20	21	1,481	1,460	10.7	1.33	3,225.27
24	24	857	880	8.8	1.05	1,049.13
10	10	830	856	2.4	1.10	49.23
20	21	1,349	1,318	8.4	1.24	1,971.03

Manuscript approved July 18, 2016

Prepared by the USGS Science Publishing Network

Editing by Teri Losano
Rolla Publishing Service Center

Figures and layout by Caryl J. Wipperfurth
Reston Publishing Service Center

For more information concerning this report,
please contact:

Director, Indiana-Kentucky Water Science Center
U.S. Geological Survey
5957 Lakeside Blvd
Indianapolis, IN 46278
<http://in.water.usgs.gov>

