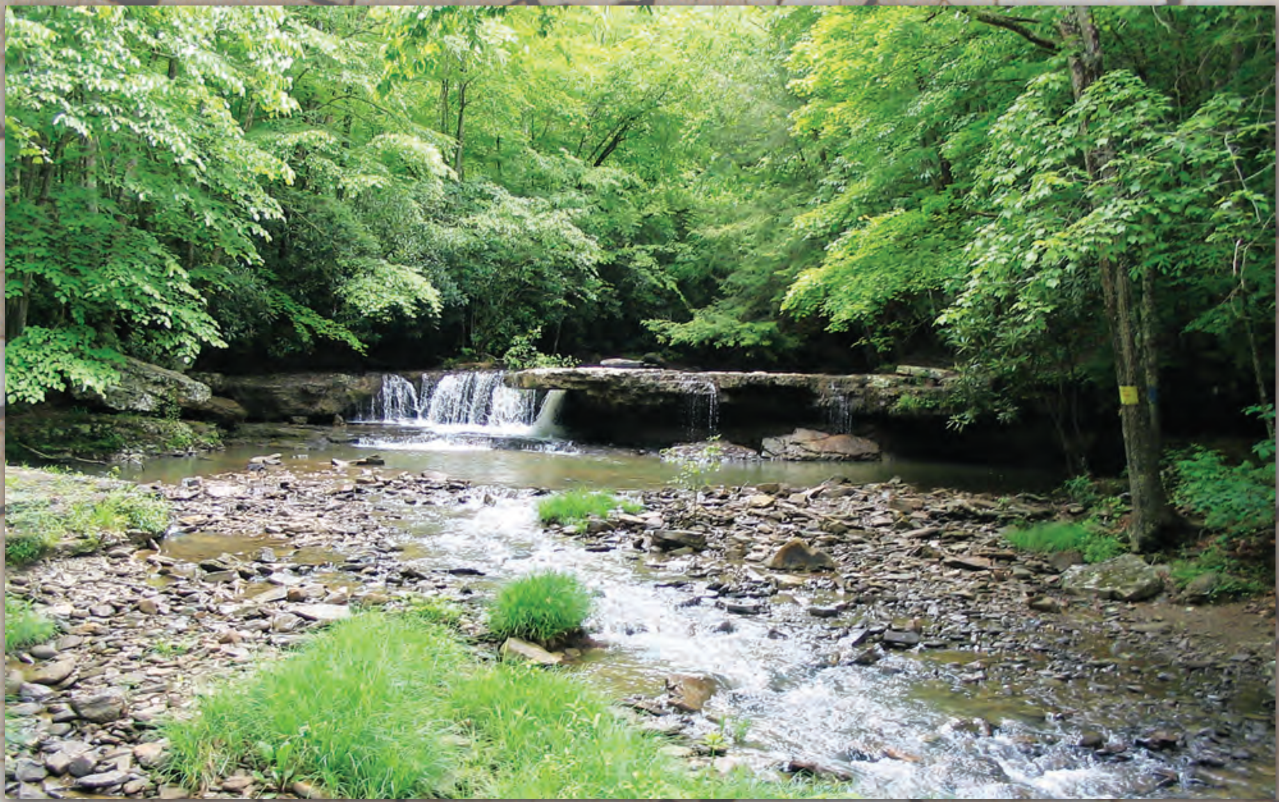
USGS
science for a changing world

In cooperation with the West Virginia Department of Environmental Protection, Division of Water and Waste Management

Estimating Selected Streamflow Statistics Representative of 1930–2002 in West Virginia



Scientific Investigations Report 2008–5105

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

Cover. Mash Fork at Camp Creek State Forest, Mercer County, West Virginia. Photo by Terence Messinger, U.S. Geological Survey.

Estimating Selected Streamflow Statistics Representative of 1930–2002 in West Virginia

By Jeffrey B. Wiley

In cooperation with the West Virginia Department of Environmental Protection,
Division of Water and Waste Management

Scientific Investigations Report 2008–5105

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

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2008

For product and ordering information:
World Wide Web: http://www.usgs.gov/pubprod
Telephone: 1-888-ASK-USGS

For more information on the USGS--the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:
World Wide Web: http://www.usgs.gov
Telephone: 1-888-ASK-USGS

Suggested citation:
Wiley, J.B.., 2008, Estimating selected streamflow statistics representative of 1930–2002 in West Virginia: U.S. Geological Survey Scientific Investigations Report 2008–5105, 24 p.

ISBN 978-1-4113-2260-8

Contents

Figures

Tables

Conversion Factors, Acronyms, and Abbreviations

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
acre	4,047	square meter (m^2)
square foot (ft^2)	0.09290	square meter (m^2)
square mile (mi^2)	2.590	square kilometer (km^2)
	Volume	
cubic foot (ft^3)	0.02832	cubic meter (m^3)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

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

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

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

Acronyms

LOESS	Locally weighted regression
MOVE.1	Maintenance of variance extension, type 1
NRAC	National Resource Analysis Center
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey
WCMS	Watershed Characterization Management System
WVDEP	West Virginia Department of Environmental Protection
WVDEP, DMR	West Virginia Department of Environmental Protection, Division of Mining and Reclamation
WVDEP, DWWM	West Virginia Department of Environmental Protection, Division of Water and Waste Management

Abbreviations

1B3	1-day, 3-year biologically based flow
1Q2	1-day, 2-year hydrologically based flow
1Q5	1-day, 5-year hydrologically based flow
1Q10	1-day, 10-year hydrologically based flow

3Q2	3-day, 2-year hydrologically based flow
3Q5	3-day, 5-year hydrologically based flow
3Q10	3-day, 10-year hydrologically based flow
4B3	4-day, 3-year biologically based flow
7Q2	7-day, 2-year hydrologically based flow
7Q5	7-day, 5-year hydrologically based flow
7Q10	7-day, 10-year hydrologically based flow
14Q2	14-day, 2-year hydrologically based flow
14Q5	14-day, 5-year hydrologically based flow
14Q10	14-day, 10-year hydrologically based flow
30Q2	30-day, 2-year hydrologically based flow
30Q5	30-day, 5-year hydrologically based flow
30Q10	30-day, 10-year hydrologically based flow
A	Agriculture cover
A_U	Drainage area at the location of the unknown streamflow
A_K	Drainage area at the location of the known streamflow
A_{US}	Drainage area at the upstream location
A_{DS}	Drainage area at the downstream location
B	Barren land cover
BOr	Basin orientation
BP	Basin perimeter
BR	Basin relief
BS	Basin slope
BW	Basin width
CR	Compactness ratio
CL	Channel length
CM	Channel maintenance
CS	Channel slope
D10	10-percent-duration flow
D25	25-percent-duration flow
D50	50-percent-duration flow
D75	75-percent-duration flow
D90	90-percent-duration flow
DA	Drainage area
E	Mean basin elevation

ER	Elongation ratio
EX	Exponent for drainage-area ratios
F	Forest cover
G	Grassland cover
GLS	Generalized least squares
HM	Harmonic-mean flow
I	Impervious cover
I24-2	24-hour, 2-year rainfall
JANMIN	January minimum temperature
LAT_C	Latitude of the basin centroid
LAT_O	Latitude of the basin outlet
$LONG_C$	Longitude of the basin centroid
$LONG_O$	Longitude of the basin outlet
M_P	Mean of the log_{10}-transformed streamflows at the partial-record station
M_G	Mean of the log_{10}-transformed streamflows at the gaging station
OLS	Ordinary least squares
P	Annual precipitation
Q_{DS}	Value of the streamflow statistic at the downstream location
Q^{G}	Log_{10}-transformed value of the streamflow at the gaging station
Q_K	Known value of the streamflow statistic
Q_{KE}	Regional equation evaluated at the location of the known value of the streamflow statistic
Q_P	Value of the streamflow statistic at the partial-record station
Q_U	Unknown value of the streamflow statistic
Q_{UE}	Regional equation evaluated at the location of the unknown value of the streamflow statistic
Q_{US}	Value of the streamflow statistic at the upstream location
RB	Rotundity of basin
R^{DS}	Downstream limit of the ratio of drainage areas
RN	Ruggedness number
RR	Relative relief
$R_{U/K}$	Ratio of the drainage area at the location of the unknown streamflow to the drainage area at the location of the known streamflow
R_{US}	Upstream limit of the ratio of drainage areas
S	Annual snow depth
SD	Stream density

SF	Shape factor
S_G	Standard deviation of the concurrent $\log_{10}$-transformed streamflows at the gaging station
SIR	Sinuosity ratio
SL	Stream length
SLR	Slope ratio
SP	Slope proportion
S_P	Standard deviation of the concurrent $\log_{10}$-transformed streamflows at the partial-record station
U	Urban land cover
VL	Valley length
W	Wetland cover
Wa	Open-water cover
WLS	Weighted least squares

Estimating Selected Streamflow Statistics Representative of 1930–2002 in West Virginia

By Jeffrey B. Wiley

Abstract

Regional equations and procedures were developed for estimating 1-, 3-, 7-, 14-, and 30-day 2-year; 1-, 3-, 7-, 14-, and 30-day 5-year; and 1-, 3-, 7-, 14-, and 30-day 10-year hydrologically based low-flow frequency values for unregulated streams in West Virginia. Regional equations and procedures also were developed for estimating the 1-day, 3-year and 4-day, 3-year biologically based low-flow frequency values; the U.S. Environmental Protection Agency harmonic-mean flows; and the 10-, 25-, 50-, 75-, and 90-percent flow-duration values.

Regional equations were developed using ordinary least-squares regression using statistics from 117 U.S. Geological Survey continuous streamflow-gaging stations as dependent variables and basin characteristics as independent variables. Equations for three regions in West Virginia—North, South-Central, and Eastern Panhandle—were determined. Drainage area, precipitation, and longitude of the basin centroid are significant independent variables in one or more of the equations.

Estimating procedures are presented for determining statistics at a gaging station, a partial-record station, and an ungaged location. Examples of some estimating procedures are presented.

Introduction

Streamflow statistics are used in the development and management of surface- and ground-water resources in West Virginia, including assessing the availability of water for municipal, industrial, and irrigation supplies; recreation; aquatic-life, and wildlife conservation; and disposal of liquid wastes. Flow statistics also are useful for forecasting low streamflows, as indicators of the amount of ground-water inflow to streams, and as legal indexes for maintaining water-quality standards.

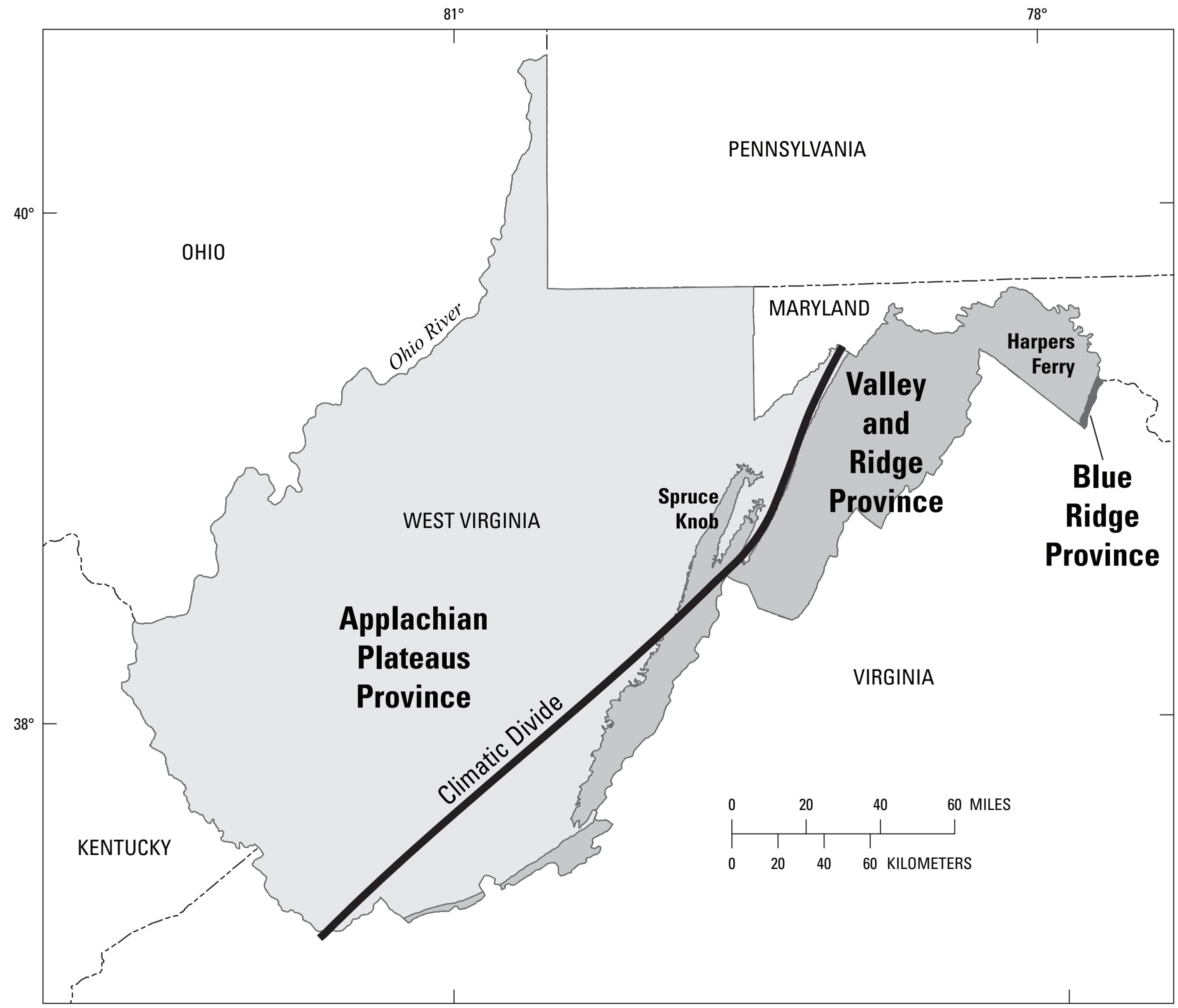
This report documents the development of regional equations and estimating procedures for determining the 1-, 3-, 7-, 14-, and 30-day 2-year (1Q2, 3Q2, 7Q2, 14Q2, and 30Q2, respectively); 1-, 3-, 7-, 14-, and 30-day 5-year (1Q5, 3Q5, 7Q5, 14Q5, and 30Q5, respectively); and, 1-, 3-, 7-, 14-, and 30-day 10-year (1Q10, 3Q10, 7Q10, 14Q10, and 30Q10, respectively) hydrologically based low-flow frequency values for unregulated streams in West Virginia. Equations and procedures are also developed for estimating the 1-day, 3-year (1B3) and 4-day, 3-year (4B3) biologically based low-flow frequency values; the U.S. Environmental Protection Agency (USEPA) harmonic-mean flows (HM); and the 10-, 25-, 50-, 75-, and 90-percent flow-duration values (D10, D25, D50, D75, and D90, respectively).

Description of Study Area

West Virginia contains parts of three physiographic provinces—the Appalachian Plateaus, Valley and Ridge, and Blue Ridge (Fenneman, 1938) (fig. 1). Two climatic regions, defined by the movement of air masses across the State, are separated by a line known as the Climatic Divide (Wiley and others, 2000) (fig. 1).

Generally, the part of the State west of the Climatic Divide is in the Appalachian Plateaus Province, where altitudes decrease northwestward from about 2,500 to 4,861 ft (Spruce Knob) along the Climatic Divide to about 550 to 650 ft along the Ohio River. The part of West Virginia east of the Climatic Divide is in the Valley and Ridge Province, except for the extreme eastern tip of the State, which is in the Blue Ridge Province. Altitudes decrease eastward from the Climatic Divide to 274 ft at Harpers Ferry in the Eastern Panhandle (U.S. Geological Survey, 1990, 2006; National Oceanic and Atmospheric Administration, 2006a).

The Appalachian Plateaus Province consists of consolidated, mostly noncarbonate sedimentary rocks that slope gently from southeast to northwest near the Climatic Divide and are nearly flat-lying along the Ohio River. One exception is in the northeastern part of the province (west of the Climatic Divide), where the rocks are gently folded and some carbonate rock crops out (Fenneman, 1938). The rocks in the Appala-



Base from U.S. Geological Survey 1:100,00 Digital Line Graphics for state boundaries; physiographic provinces from Fenneman, 1938. All data are in the Universal Tranverse Mercator projection, zone 17, NAD 83 projection.

Figure 1. Physiographic provinces and Climatic Divide in West Virginia.

chian Plateaus Province have been eroded to form steep hills and deeply incised valleys; drainage patterns are dendritic.

The Valley and Ridge Province in West Virginia consists of consolidated carbonate and noncarbonate sedimentary rocks that are sharply folded and extensively faulted (Fenneman, 1938). Northeast-trending valleys and ridges parallel the Climatic Divide; drainage patterns are trellis.

The Blue Ridge Province in West Virginia consists predominantly of metamorphosed sandstone and shale (Fenneman, 1938). The province is characterized by high relief between mountains and wide valleys that parallel the Climatic Divide.

The climate of West Virginia is primarily continental, with mild summers and cold winters. Major weather systems generally approach from the west and southwest, although polar continental air masses of cold, dry air that approach from the north and northwest are not unusual. Air masses from the Atlantic Ocean sometimes affect the area east of the Climatic Divide and less frequently affect the area west of the Climatic Divide. Generally, hot, dry tropical continental air masses from the southwest affect the climate west of the Climatic Divide. Warm, moist tropical maritime air masses from the Gulf of Mexico affect the climate east of the Climatic Divide. Evaporation from local and upwind land surfaces, lakes, and reservoirs is an additional source of moisture that affects the State's climate (U.S. Geological Survey, 1991; National Oceanic and Atmospheric Administration, 2006a).

Annual precipitation averages about 42 to 45 in. statewide; about 60 percent is received from March through August. July is the wettest month, and September through November are the driest. Annual average precipitation in the State generally decreases northwestward from about 50 to 60 in. along the Climatic Divide to about 40 in. along the Ohio River, and increases from about 30 to 35 in. east of the Climatic Divide to about 40 in. at the extreme eastern tip of the State. Precipitation is greater along and west of the Climatic Divide than east of it as a consequence of the higher elevations along the Divide and the general movement of weather systems approaching from the west and southwest. Annual average snowfall follows the general pattern of annual average precipitation, decreasing northwestward from about 36 to 100 in. along the Climatic Divide to about 20 to 30 in. along the Ohio River. Annual average snowfall ranges from 24 to 36 in. east of the Climatic Divide (U.S. Geological Survey, 1991; Natural Resources Conservation Service, 2006; National Oceanic and Atmospheric Administration, 2006a, 2006b).

Previous Studies

Selected statistics for U.S. Geological Survey (USGS) streamflow-gaging stations representative of conditions during 1930–2002 were determined by Wiley (2006). In that study, a criterion-based sample of the record period was used to determine statistics representative of the period, rather than the entire record period and (or) record-extension techniques. The selected statistics included hydrologically and biologically based low-flow frequency values, harmonic means, and flow-duration values (including variability index). Statistics published by Wiley (2006) were used in the current study to develop procedures for estimating statistics at ungaged locations in West Virginia.

The values for 7Q2 and 7Q10 at 100 streamflow-gaging stations and 296 partial-record sites were determined by Friel and others (1989), who revised values determined by Frye and Runner (1970). Friel and others (1989) developed equations, with drainage area and variability index as independent variables, and procedures for estimating the values for 7Q2 and 7Q10 at ungaged locations for two regions of the State. The results of the current study supersede these earlier low-flow frequency values, variability indexes, and estimating procedures primarily because of inclusion of more than 20 years of additional data.

Flood-frequency values at unregulated USGS streamflow-gaging stations and estimating equations were determined by Wiley and others (2000, 2002); these reports revised results of previous studies by Frye and Runner (1969, 1970, and 1971) and Runner (1980). Wiley and others (2000, 2002) developed procedures for estimating peak streamflows at the 1.1-, 1.2-, 1.3-, 1.4-, 1.5-, 1.6-, 1.7-, 1.8-, 1.9-, 2-, 2.5-, 3-, 5-, 10-, 25-, 50-, 100-, 200-, and 500-year recurrence intervals at unregulated, ungaged locations. Estimating equations were determined for three regions of the State, with drainage area as the independent variable. The regional equations developed by Wiley and others (2000, 2002) are not superseded by the current study; however, the estimating procedures presented in this report can be substituted for those in Wiley and others (2000, 2002) by using the information provided in appendix 1.

Some of the statistics and estimating equations developed by Wetzel and Bettandorff (1986), applicable to much of West Virginia, have been superseded by the flood frequencies determined by Wiley and others (2000) and the results of this study.

Selected Streamflow Statistics

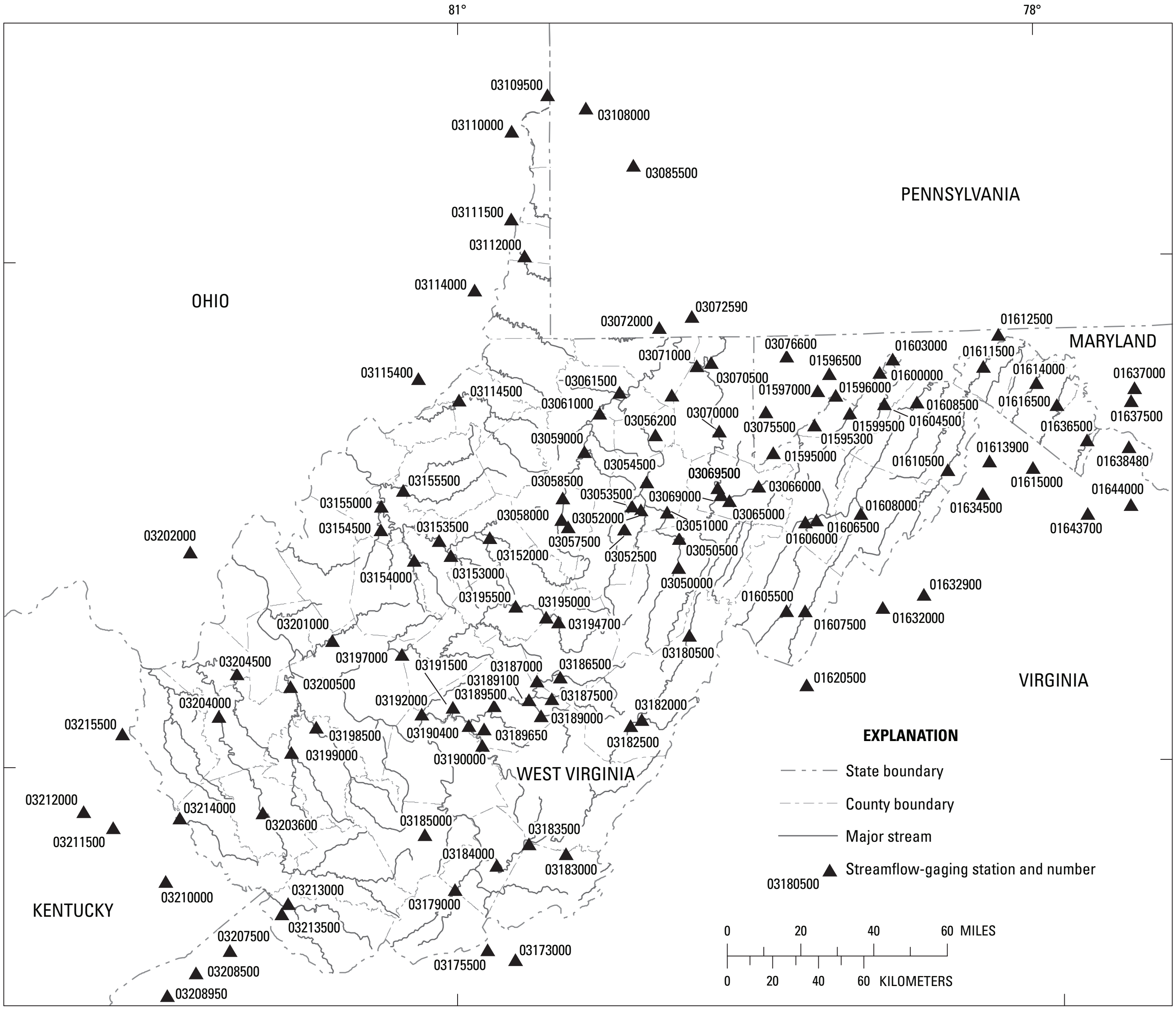
In this study, estimating procedures were developed for the following statistics: 1Q2, 1Q5, 1Q10, 3Q2, 3Q5, 3Q10, 7Q2, 7Q5, 7Q10, 14Q2, 14Q5, 14Q10, 30Q2, 30Q5, 30Q10, 1B3, 4B3, HM, D10, D25, D50, D75, and D90. These statistics were available in Wiley (2006) for 77 streamflow-gaging stations in West Virginia and 40 stations in adjacent states (fig. 2).

Hydrologically based low-flow frequency values were determined by Wiley (2006) using methods described by Riggs (1972). A series of the annual minimum n-day (number of consecutive days) daily mean low streamflows are fitted to a log-Pearson Type III probability curve. A plot of the probability curve and data is reviewed for fitness. Other probability distributions are considered or a smooth curve is constructed through the data if the data do not fit the log-Pearson Type III probability curve. The frequency of the n-day streamflow is computed from the fitted probability curve or read from the smooth curve constructed through the data. In Wiley (2006), all data fit the log-Pearson Type III probability curve. For example, the value of “7Q10” would be the minimum average streamflow for 7 consecutive days expected on the average of once every 10 years.

Biologically based low-flow frequency values were determined by Wiley (2006) using methods described by the U.S. Environmental Protection Agency (1986). A daily series of average n-day daily mean streamflows are computed for a station record. The n-day series is evaluated for the desired frequency of occurrence of excursions based on the number of years of record. An excursion is a low-flow period that is determined to be hydrologically separate from other low-flow periods, typically separated by a minimum of 120 days. For example, the value for “4B3” would be the minimum average streamflow for 4 consecutive days expected on the average of once every 3 years.

The USEPA harmonic-mean flows (HM) were determined by Wiley (2006) using methods described by Rossman (1990). The average of the reciprocals of the daily mean flows is computed for a station record. The harmonic-mean flow is the reciprocal of that average. The USEPA harmonic-mean flow is the weighted average of the harmonic mean of the non-zero flows and the arithmetic mean of the zero flows (i.e., zero). The harmonic mean of the non-zero flows is weighted by the number of non-zero values and the arithmetic mean of the zero flows is weighted by the number of zero values.

Flow-duration values were determined by Wiley (2006) using methods described by Searcy (1959). A yearly record of daily mean flows is divided into 20 to 30 classes of average streamflows. Every complete year of record at a station is divided into the same classes. The number of days in each



Base from U.S. Geological Survey 1:100,00 Digital Line Graphics for state boundaries; 1:100,000 digital data for streams; and West Virginia Department of Enviromental Protection 1:24,000 digital data for county boundaries.
All data are in the Universal Tranverse Mercator projection, zone 17, NAD 83 projection.

Figure 2. U.S. Geological Survey streamflow-gaging stations used in this study that are in West Virginia and adjacent states.

class is computed for the entire record period and the percentage of the time a streamflow is in each class is determined. A particular flow-duration value is extrapolated from the class percentiles, and a log-probability plot of the class-percentile flow values is a flow-duration curve. For example, the flow for "D50" is equaled or exceeded 50 percent of the time.

Development of Equations for Estimating Selected Streamflow Statistics

Ordinary least squares (OLS) regressions of the selected gaging-station statistics determined by Wiley (2006) with basin characteristics as independent variables were used in the current study to develop regional equations for estimating statistics at ungaged locations. Regression procedures were performed using the computer program S-PLUS 7.0 (Insightful Corporation, 2005), a commercially available statistical computing package. Dependent and independent variables were $\log_{10}$-transformed, and both transformed and untransformed independent variables were used in the regression procedures. A correlation matrix of independent variables was assessed to eliminate highly correlated independent variables from the equations, and data plots were assessed to ensure linearity between dependent and independent variables.

Generalized least squares (GLS) regression (Stedinger and Tasker, 1985; Tasker and Stedinger, 1989) was not used for this study. GLS regression requires an annual time series of data, and is more accurate than OLS regression for hydrologic purposes, primarily when record lengths vary between stations (Tasker and Stedinger, 1989, p. 363). Record lengths for the individual stations are all representative of 1930-2002 (73 years) regardless of the actual individual record lengths or actual years of record used to determine the statistics for this study (Wiley, 2006). GLS regression was not used because biological frequency values, USEPA harmonic-mean flows, and flow-duration values do not have annual time series, and because the representative record periods for all stations are identical.

Estimating procedures were not developed using a weighted least squares (WLS) regression because the results would likely be less accurate than those derived from the procedures developed using the criterion-based sample of data and OLS regression for the present study. The weights for each gaging station could be based on a comparison between the average annual minimum flows for the period of record at the station and the average annual minimum flows for the region where the station is located, where the averages are normalized by either drainage area or standard deviations. The entire record period at a station could be used, rather than the criterion-based sample of the record period. The time-sampling error would not relate to the length of the record, but to the departure of low flows at a station from the 1930–2002 low flows in a region. The WLS regression could account for part of the time-sampling error, but the positive bias in flow data resulting from the operation of more stations during a wetter-than-average time period (Wiley, 2006, p. 22) would remain. WLS regression weights based simply on the record length are not appropriate because longer records do not necessarily provide more accurate low-flow estimates representative of 1930–2002; a 15-year record could include a wetter-than-average period and should be given less weight than a 10-year record that does not include a wetter-than-average period.

Basin Characteristics

Basin characteristics for 117 streamflow-gaging stations in West Virginia and adjacent states (Paybins, 2008) were available for use as independent variables for regression. The basin characteristics included the following: drainage area (DA), in square miles; latitude of basin outlet (LAT_O), in decimal degrees; longitude of basin outlet ($LONG_O$), in decimal degrees; latitude of the basin centroid (LAT_C), in decimal degrees, longitude of the basin centroid ($LONG_C$), in decimal degrees; basin perimeter (BP), in miles; basin slope (BS), in feet per mile; basin relief (BR), in feet; basin orientation (BOr), in degrees; channel length (CL), in miles; valley length (VL), in miles; channel slope (CS), in feet per mile; stream length (SL), in miles; mean basin elevation (E), in feet; 24-hour 2-year rainfall (I24-2), in inches; annual precipitation (P), in inches; January minimum temperature (JANMIN), in degrees Fahrenheit; annual snow depth (S), in inches; forest cover (F), in percent; grassland cover (G), in percent; barren land cover (B), in percent; urban land cover (U), in percent; wetland cover (W), in percent; open-water cover (Wa), in percent; agriculture cover (A), in percent; impervious cover (I), in percent; basin width (BW), in miles; shape factor (SF), dimensionless; elongation ratio (ER), dimensionless; rotundity of basin (RB), dimensionless; compactness ratio (CR), dimensionless; relative relief (RR), in feet per mile; sinuosity ratio (SIR), dimensionless; stream density (SD), in miles per square mile; channel maintenance (CM), in square miles per mile; slope proportion (SP), dimensionless; ruggedness number (RN), in feet per mile; and slope ratio (SLR), dimensionless.

Variability index was considered as a basin characteristic for use as an independent variable for regression. The variability index was determined using methods described by Lane and Lei (1950). The variability index is a measure of the slope of the flow-duration curve. The variability index is the standard deviation of the logarithms of flow durations at 10-percent intervals from 5 to 95 percent. A variation of the method used to compute the variability index was necessary when the 95-percent flow duration was 0.00 ft^3/s; the 95- and 5-percent flow durations were excluded from the computation and the standard deviation was computed for the logarithms of flow durations at 10-percent intervals from 15 to 85 percent. Variability indexes are available in Wiley (2006).

The base-flow recession time constant was not considered for use as an independent variable for regression. The base-flow recession time constant is the characteristic time constant of exponential decay of streamflow long after a storm, and can be a significant dependent variable for regression of low-flow characteristics (Eng and Milly, 2007). The constant, unlike the variability index, does not require the mapping of the variable for use at ungaged locations, but it does require at least one pair of base-flow measurements for a single recession to estimate the value. The constant was not computed because the USGS West Virginia Water Science Center does not operate a network of partial-record sites where paired base-flow measurements could be made, and available data from networks of partial-record sites measured by state agencies and private consultants do not include the necessary pair of base-flow measurements.

Data Correlation

The basin characteristics were $\log_{10}$ transformed, and both transformed and untransformed variables were evaluated for linear correlation among themselves by using a Pearson coefficient, or Pearson's r (Helsel and Hirsch, 2002). The integer 1 was added to values of G, B, U, W, WA, A, and I to ensure that values were greater than zero for $\log_{10}$ transformation. To decrease values for regression analysis, 77 was subtracted from $LONG_C$ and 37 was subtracted from LAT_C.

Variables were considered highly correlated where the absolute value of the Pearson coefficient was greater than or equal to 0.80. Untransformed and transformed variables except DA, CS, CR, SP, and SLR were highly correlated. Untransformed DA, CS, CR, SP, and SLR were not highly correlated with any transformed variables. Transformed DA, CL, SL, CR, SP, and variability index were highly correlated; transformed VL and DA were highly correlated with this group of variables but were not highly correlated themselves. Transformed CS, SP, and SLR; BR and RN; CS and RR; F and I; U and I; and BW and SIR were highly correlated.

In addition to being highly correlated, the absolute value of the Pearson coefficient was equal to 1 (singularity) for $\log_{10}$-transformed SF, ER, and RB; for transformed SD and CM; for transformed DA, BW, SF, and RB; and for transformed SL, CR, RR, SD, and RN. The variables ER, RB, CM, and BW were removed from consideration for regression because of singularity, and additional variables having singularity were removed if they became significant in the regression analysis.

There was a very high correlation (0.963) between $\log_{10}$-transformed DA and variability index. DA and variability index were the only two significant independent variables in the previous study by Friel and others (1989). The correlation for $\log_{10}$-transformed DA and variability index in the previous study was 0.260 (calculated from values presented in Friel and others, 1989, table 1). DA and variability index probably were not highly correlated in the previous study because calculations of variability index probably had time-sampling errors caused by effects associated with climate variability, similar to the difference of 1Q10, 7Q10, 30Q5, 1B3, 4B3, and USEPA harmonic-mean flows calculated for various time periods by Wiley (2006, table 2, p. 15). The criterion-based sample of data used to calculate variability index probably accounted for most of the time-sampling error, resulting in the high correlation with DA in the current study. Variability index was removed from consideration for regression in this study because of its high correlation with DA.

Based on the correlation analysis, all $\log_{10}$-transformed variables except ER, RB, CM, BW, and variability index, and only untransformed DA, CS, CR, SP, and SLR, were considered for regression. Some $\log_{10}$-transformed variables that were not highly correlated on a statewide basis were highly correlated on a regional basis. S and P, and S and $LONG_C$ were highly correlated in the South-Central Region. S and $LONG_C$, and U and I were highly correlated in the North Region. All but one highly correlated value was eliminated from consideration when more than one highly correlated value became a significant independent variable in a regional regression equation.

Regional Regression Analysis

Multiple and simple regression techniques were used to determine regional boundaries. The selected statistics for 117 streamflow-gaging stations in West Virginia and adjacent states were $\log_{10}$-transformed and regressed with the basin characteristics as independent variables. Log_{10}-transformed DA was a significant independent variable for all regressions. Plots of residuals by latitude and longitude of basin centers from the simple regressions of selected statistics and $\log_{10}$-transformed DA indicated a regional boundary between the western part of the state and the Eastern Panhandle. A plot of residuals from regression of statistics and DA for the western part of the state indicated a regional boundary similar to that determined in the study by Friel and others (1989). This boundary in the western part of the state was described previously as approximately the outcrop of the base of the Upper Pennsylvanian (Conemaugh Group) rocks (Friel and others, 1989, p. 11); in the current study, it was delineated following topographic features (basin boundaries). Plots of residuals by latitude and longitude of basin centers from the simple regressions of selected statistics and DA did not indicate any additional regions. The following 14 stations in adjacent states were removed from the analysis because the absolute values of their residuals were greater than those of the residuals for stations within the state, indicating these stations were not representative of hydrologic conditions in West Virginia: 01612500, 01632000, 01643700, 01644000, 03085500, 03108000, 03109500, 03110000, 03111500, 03202000, 03210000, 03211500, 03212000, and 03215500.

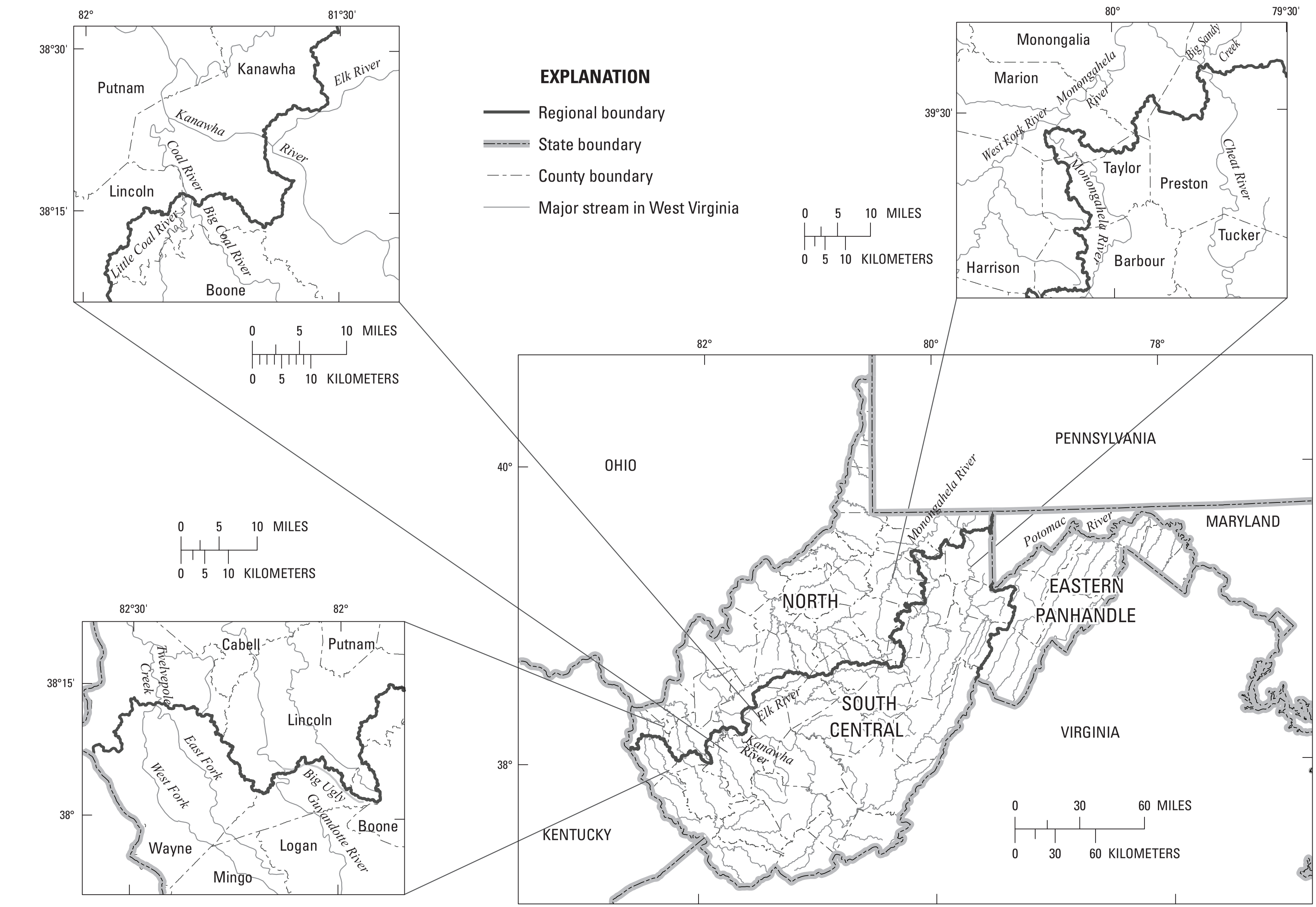
The three regions—North, South-Central, and Eastern Panhandle (fig. 3)—are separated by topographic features. The boundary between the Eastern Panhandle and South-Central Regions follows the Potomac River Basin boundary. The South-Central Region (from northeast to southwest) is the area upstream from the confluence of Big Sandy Creek (excluding Big Sandy Creek) on the Cheat River, upstream from the confluence of West Fork River (excluding West Fork River) on the Monongahela River, upstream from the confluence of the Elk and Kanawha Rivers, upstream from the confluence of the Big and Little Coal Rivers, upstream from the confluence of Big Ugly Creek and Guyandotte River, upstream from the confluence of East and West Forks of Twelvepole Creek, and upstream from the confluence of Tug Fork on the Big Sandy River; the main stems of the Cheat, Monongahela, Kanawha, Coal, Guyandotte, and Big Sandy Rivers and Twelvepole Creek downstream from the regional boundary are included in the South-Central Region. The North Region consists of the remainder of the State.

The regional equations were evaluated subjectively. Equations for hydrologically based flows with recurrence intervals of 2 years, flow-duration values, and HM were developed by the minimum number of independent variables determined when additional independent variables did not increase the coefficient of determination (r^2) by at least 0.05 or decrease the standard error of the estimate by at least 5 percent. Equations for statistics other than hydrologically based flows with recurrence intervals of 2 years, flow-duration values, and HM were evaluated with consideration of additional independent variables with the added requirement that inclusion did not result in unreasonable solutions (higher recurrence interval flows greater than lower recurrence interval flows). Unreasonable solutions were found, particularly in the North Region, where the number of observations was less than in the other regions and F, U, and I were significant in some equations (U and I are highly correlated in this region); therefore, F, U, and I were not included in the final equations. F was significant in some equations in the South-Central Region but was not included in the final equations because of unreasonable solutions.

Three independent variables, DA, $LONG_C$, and P, were significant in equations for determining the selected statistics for the three regions in West Virginia (table 1). DA was significant and ranged from 16.3 to 1, 516 mi^2 in the North Region, 2.78 to 1,619 mi^2 in the South-Central Region, and 8.83 to 3,041 mi^2 in the Eastern Panhandle Region. $LONG_C$ was significant and ranged from 79.618 to 82.023 decimal degrees in the North Region. P was significant and ranged from 42.3 to 61.4 in. in the South-Central Region.

The standard errors of the estimate for equations developed in this study were compared to the errors determined in the previous study by Friel and others (1989). In the previous study, the standard error for 7Q2 and 7Q10 ranged from 43 to 57 percent and from 82 to 83 percent, respectively; in comparison, in the current study, the standard error for 7Q2 and 7Q10 ranged from 45.1 to 53.5 percent and from 82.9 to 151 percent, respectively. The results of the current study are considered more precise than those of the previous study even though the standard errors for 7Q2 and 7Q10 are similar to or greater than those of the previous study because (1) more than 20 years of additional data were used in the current study, (2) an additional unmeasured error in the previous study is associated with the determination of variability index by producing a map of values for ungaged locations, (3) the variability index was not used as an independent variable in the current study because of the high correlation with drainage area resulting from the criterion-based sample of available data that removed the time-sampling errors associated with climate variability, (4) the estimates developed in this study are representative of the period from 1930 to 2002 but no period is specified in the previous study, and (5) 18 streamflow-gaging stations (03051500, 03059500, 03062400, 03062500, 03151400, 03151500, 03152500, 03178500, 03179500, 03182700, 03187300, 03201410, 03202400, 03202750, 03203000, 03206600, 03207000, and 03207020) in West Virginia that were used in the regression in the previous study were excluded from the current study because their records were considered to have a positive bias during the period from 1930 to 2002 (Wiley, 2006, p. 22).

Investigators comparing the results of this study to those of other studies should consider the removal of variables from the final equations because of unreasonable solutions (discussed above). The final equation for 7Q10 in the North Region had an error of 151 percent, but equations that did not include independent variables that resulted in unreasonable solutions were a six-variable equation with a 79.2-percent error, a five-variable equation with a 90.9-percent error, a four-variable equation with a 106-percent error, and a three-variable equation with a 102-percent error.



Base from U.S. Geological Survey 1:100,00 Digital Line Graphics for state boundaries; 1:100,000 digital data for streams;
and West Virginia Department of Enviromental Protection 1:24,000 digital data for county boundaries.
All data are in the Universal Tranverse Mercator projection, zone 17, NAD 83 projection.

Figure 3. The North, South-Central, and Eastern Panhandle Regions of West Virginia for which equations for estimating selected streamflow statistics were developed. (Insets provided to clarify location of the regional boundary.)

Table 1. Estimating equations for selected streamflow statistics for the North, South-Central, and Eastern Panhandle Regions of West Virginia.

[xQy, x-day y-year hydrologically based flow, in cubic feet per second; xBy, x-day y-year biologically based flow, in cubic feet per second; HM, U.S. Environmental Protection Agency harmonic-mean flow, in cubic feet per second; Dn, n-percent-duration flow, in cubic feet per second; DA, drainage area, in square miles; $LONG_C$, longitude of basin centroid, in decimal degrees from North American Datum of 1983; P, average annual precipitation, in inches]

Equation	Regression coefficient (r^2), unitless	Average standard error of estimate, in percent	Number of observations used to develop equation
North Region (range in DA from 16.3 to 1,516; range in $LONG_C$ from 79.618 to 82.023)			
$1Q2 = 7.33 \times 10^{-1} DA^{1.35} / (LONG_C - 77)^{4.79}$	0.90	48.9	23
$1Q5 = DA^{1.50} / (LONG_C - 77)^{6.67}$	.83	82.2	18
$1Q10 = 7.39\ DA^{1.59} / (LONG_C - 77)^{9.30}$	.63	233	16
$3Q2 = 7.63 \times 10^{-1} DA^{1.33} / (LONG_C - 77)^{4.64}$	.91	45.4	23
$3Q5 = 1.57\ DA^{1.41} / (LONG_C - 77)^{6.47}$	.86	69.9	19
$3Q10 = 7.68\ DA^{1.55} / (LONG_C - 77)^{9.06}$	.63	217	17
$7Q2 = 1.06\ DA^{1.31} / (LONG_C - 77)^{4.66}$	.91	45.1	23
$7Q5 = 2.31\ DA^{1.45} / (LONG_C - 77)^{6.88}$	.78	101	21
$7Q10 = 4.84\ DA^{1.41} / (LONG_C - 77)^{7.84}$	.66	151	18
$14Q2 = 1.01\ DA^{1.37} / (LONG_C - 77)^{4.65}$	.92	42.8	23
$14Q5 = 2.32\ DA^{1.32} / (LONG_C - 77)^{6.14}$	.75	95.3	22
$14Q10 = 7.58\ DA^{1.55} / (LONG_C - 77)^{8.66}$	.66	179	20
$30Q2 = 1.07\ DA^{1.31} / (LONG_C - 77)^{3.99}$	.94	35.8	23
$30Q5 = 5.93 \times 10^{-1} DA^{1.61} / (LONG_C - 77)^{5.90}$	.80	104	23
$30Q10 = 4.82\ DA^{1.40} / (LONG_C - 77)^{7.08}$	.80	91.2	21
$1B3 = 6.33\ DA^{1.13} / (LONG_C - 77)^{7.07}$	.73	88.4	15
$4B3 = 3.69\ DA^{1.19} / (LONG_C - 77)^{6.69}$	.72	92.7	15
$HM = 8.86\ DA^{1.18} / (LONG_C - 77)^{4.83}$	.83	63.0	23
$D10 = 2.89\ DA^{1.03}$	.98	16.9	23
$D25 = 4.19\ DA^{1.08} / (LONG_C - 77)^{1.15}$	.99	13.4	23
$D50 = 2.84\ DA^{1.11} / (LONG_C - 77)^{1.86}$	.98	17.9	23
$D75 = 1.32\ DA^{1.21} / (LONG_C - 77)^{2.72}$	.96	25.5	23
$D90 = 4.14 \times 10^{-1} DA^{1.38} / (LONG_C - 77)^{3.64}$	.95	32.8	23

Table 1. Estimating equations for selected streamflow statistics for the North, South-Central, and Eastern Panhandle Regions of West Virginia.—Continued

[xQy, x-day y-year hydrologically based flow, in cubic feet per second; xBy, x-day y-year biologically based flow, in cubic feet per second; HM, U.S. Environmental Protection Agency harmonic-mean flow, in cubic feet per second; Dn, n-percent-duration flow, in cubic feet per second; DA, drainage area, in square miles; $LONG_C$, longitude of basin centroid, in decimal degrees from North American Datum of 1983; P, average annual precipitation, in inches]

Equation	Regression coefficient (r^2), unitless	Average standard error of estimate, in percent	Number of observations used to develop equation
South-Central Region (range in DA from 2.78 to 1,619; range in P from 42.3 to 61.4)			
$1Q2 = 1.86 \times 10^{-2} DA^{1.19}$	0.89	54.9	51
$1Q5 = 6.06 \times 10^{-3} DA^{1.25}$	.79	90.5	51
$1Q10 = 3.25 \times 10^{-3} DA^{1.27}$	.63	135	50
$3Q2 = 2.18 \times 10^{-2} DA^{1.18}$	.89	53.3	51
$3Q5 = 7.59 \times 10^{-3} DA^{1.22}$	.80	85.9	51
$3Q10 = 4.48 \times 10^{-3} DA^{1.23}$	.63	125	50
$7Q2 = 2.71 \times 10^{-2} DA^{1.16}$	.90	51.0	51
$7Q5 = 1.00 \times 10^{-2} DA^{1.20}$	.81	81.2	51
$7Q10 = 6.88 \times 10^{-3} DA^{1.18}$	.65	113	50
$14Q2 = 4.10 \times 10^{-2} DA^{1.12}$	.90	48.6	51
$14Q5 = 1.26 \times 10^{-2} DA^{1.19}$	.82	75.3	51
$14Q10 = 5.58 \times 10^{-3} DA^{1.25}$	.76	103	51
$30Q2 = 6.77 \times 10^{-2} DA^{1.10}$	.90	47.5	51
$30Q5 = 3.16 \times 10^{-2} DA^{1.09}$	.83	65.7	51
$30Q10 = 2.04 \times 10^{-2} DA^{1.09}$	.77	84.2	51
$1B3 = 9.78 \times 10^{-3} DA^{1.08}$	.66	113	49
$4B3 = 1.18 \times 10^{-2} DA^{1.09}$	.68	106	49
$HM = 1.10 \times 10^{-1} DA^{1.14}$	.86	61.5	51
$D10 = 7.16 \times 10^{-5} DA^{0.987} P^{2.79}$	.99	12.7	51
$D25 = 1.26 \times 10^{-5} DA^{1.00} P^{3.03}$	.99	13.9	51
$D50 = 9.89 \times 10^{-7} DA^{1.01} P^{3.43}$	.98	18.8	51
$D75 = 1.06 \times 10^{-7} DA^{1.08} P^{3.64}$	.96	30.9	51
$D90 = 1.75 \times 10^{-5} DA^{1.13} P^{2.04}$	.92	45.4	51

Table 1. Estimating equations for selected streamflow statistics for the North, South-Central, and Eastern Panhandle Regions of West Virginia.—Continued

[xQy, x-day y-year hydrologically based flow, in cubic feet per second; xBy, x-day y-year biologically based flow, in cubic feet per second; HM, U.S. Environmental Protection Agency harmonic-mean flow, in cubic feet per second; Dn, n-percent-duration flow, in cubic feet per second; DA, drainage area, in square miles; $LONG_C$, longitude of basin centroid, in decimal degrees from North American Datum of 1983; P, average annual precipitation, in inches]

Equation	Regression coefficient (r^2), unitless	Average standard error of estimate, in percent	Number of observations used to develop equation
Eastern Panhandle Region (range in DA from 8.83 to 3,041)			
$1Q2 = 3.85 \times 10^{-2} DA^{1.13}$	0.91	53.3	29
$1Q5 = 1.53 \times 10^{-2} DA^{1.22}$	.88	71.8	29
$1Q10 = 8.78 \times 10^{-3} DA^{1.27}$	.85	84.4	29
$3Q2 = 4.06 \times 10^{-2} DA^{1.13}$	.91	53.5	29
$3Q5 = 1.69 \times 10^{-2} DA^{1.21}$	.88	70.9	29
$3Q10 = 1.00 \times 10^{-2} DA^{1.26}$	.85	86.3	29
$7Q2 = 4.57 \times 10^{-2} DA^{1.12}$	.91	53.5	29
$7Q5 = 1.98 \times 10^{-2} DA^{1.19}$	.88	69.0	29
$7Q10 = 1.22 \times 10^{-2} DA^{1.24}$	.86	82.9	29
$14Q2 = 5.40 \times 10^{-2} DA^{1.10}$	.91	51.6	29
$14Q5 = 2.38 \times 10^{-2} DA^{1.17}$	.89	66.0	29
$14Q10 = 1.49 \times 10^{-2} DA^{1.22}$	.86	78.5	29
$30Q2 = 7.08 \times 10^{-2} DA^{1.08}$	.92	47.5	29
$30Q5 = 3.39 \times 10^{-2} DA^{1.14}$	.90	58.6	29
$30Q10 = 2.26 \times 10^{-2} DA^{1.17}$	.88	68.4	29
$1B3 = 5.43 \times 10^{-3} DA^{1.34}$	.81	115	29
$4B3 = 9.17 \times 10^{-3} DA^{1.27}$	.83	101	29
$HM = 1.72 \times 10^{-1} DA^{1.08}$	.93	43.0	29
$D10 = 3.54 DA^{0.931}$	.95	33.1	29
$D25 = 1.70 DA^{0.937}$	.94	36.6	29
$D50 = 6.21 \times 10^{-1} DA^{0.969}$	.93	39.1	29
$D75 = 1.74 \times 10^{-1} DA^{1.04}$	.94	40.4	29
$D90 = 6.94 \times 10^{-2} DA^{1.09}$	.93	45.9	29

Procedures for Estimating Selected Streamflow Statistics

Estimating procedures were developed for streamflow statistics at a gaging station, at a partial-record station, and at an ungaged location. For gaging stations having records of less than 10 years or not representative of the period 1930–2002, a partial record can be developed from base-flow measurements made at the stations (or published daily mean streamflows during base-flow conditions as surrogates for measurements).

At a Gaging Station

Streamflow statistics for a gaging station were published in Wiley (2006, table 11). Not all of those statistics were selected for developing estimating procedures in the current study. A minimum of 10 years of record representative of the period 1930–2002 was required for a station to be included in the current study.

At a Partial-Record Station

A minimum of eight base-flow measurements at a partial-record station made across a wide range of base flows in more than 1 year are compared to concurrent streamflows at a nearby gaging station to develop a relation to transfer the selected flow statistics from the gaging station to the partial-record station (Riggs, 1972). The gaging station used for comparison should be within the same basin and have similar geology to meet the assumptions that the base flows are sufficiently correlated and the relation is linear. The mean daily streamflow can be used as the concurrent streamflow at the gaging station under base-flow conditions because the change in streamflow over the day is insignificant. A log-log plot of flow data should be viewed to ensure the relation is linear. The "maintenance of variance extension, type 1" (MOVE.1), also referred to as "line of organic correlation," is developed between the measurements at the partial-record station and the concurrent streamflows at the gaging station using methods described by Hirsch (1982), Hirsch and Gilroy (1984), and Helsel and Hirsch (2002). MOVE.1 was developed for extending streamflow records, but is used in the current study for transferring flow statistics according to the procedures described by Riggs (1972). The means (M) and standard deviations (S) of the concurrent $\log_{10}$–transformed streamflows at the partial-record station and the gaging station are determined. The value of the streamflow statistic at the partial-record station is computed by evaluating MOVE.1 at the flow value for the statistic at the gaging station, by using the following equation:

$$Q_P = 10^{[M_P + (S_P / S_G)(Q_G - M_G)]} , \qquad (1)$$

where

- Q_P is the value of the streamflow statistic at the partial-record station, in ft^3/s;
- M_P is the mean of the concurrent $\log_{10}$–transformed streamflows at the partial-record station, in ft^3/s;
- S_P is the standard deviation of the concurrent $\log_{10}$–transformed streamflows at the partial-record station, in ft^3/s;
- S_G is the standard deviation of the concurrent $\log_{10}$–transformed streamflows at the gaging station, in ft^3/s;
- Q_G is the $\log_{10}$–transformed streamflow of the statistic at the gaging station, in ft^3/s; and
- M_G is the mean of the concurrent $\log_{10}$–transformed streamflows at the gaging station, in ft^3/s.

A graphical procedure is used to estimate statistics if the relation between base-flow measurements at a partial-record station and concurrent streamflows at a nearby gaging station is not linear (Riggs, 1972). The untransformed streamflows are plotted on log-log graph paper with streamflows for the gaging station plotted on the x-axis, streamflows for the partial-record station plotted on the y-axis, and a smooth line constructed through the streamflow points. The streamflow for the statistic of interest for the gaging station is projected parallel to the y-axis from the value on the x-axis to the smooth line, and then projected parallel to the x-axis from the smooth line to the corresponding partial-record streamflow on the y-axis. The extrapolated streamflow on the y-axis is the estimated value for the statistic of interest at the partial-record station.

Statistics at partial-record stations are limited to estimates at and below the streamflow of 50-percent duration (median) because concurrent streamflows for the gaging station and partial-record station are generally at the same base-flow condition (same flow duration). Concurrent streamflows for the gaging station and partial-record station above the streamflow of 50-percent duration typically change rapidly and are not under base-flow conditions at one or both locations. The D10 and D25 at a partial-record station can be estimated by applying the regional equations.

The USGS West Virginia Water Science Center does not operate a network of partial-record stations. However, measurements at partial-record networks have been made by State agencies and private consultants in West Virginia in the recent past. A private consultant made monthly streamflow measurements for the West Virginia Department of Environmental Protection (WVDEP), Division of Mining and Reclamation (DMR), at a network of about 240 locations in the coal-mining region of the State. WVDEP, Division of Water and Waste Management (DWWM), measures streamflow at a network of partial-record stations as part of a 5-year cycle of hydrologic assessment of basins in the State. Streamflow statistics for partial-record stations in these two networks are estimated by WVDEP, DWWM, using base-flow measurements and extrapolating statistics from nearby gaging stations.

At an Ungaged Location

Four different procedures are used to determine streamflow statistics at an ungaged location: (1) when the ungaged location is upstream from a gaging station or partial-record station; (2) when the ungaged location is downstream from a gaging station or partial-record station; (3) when the ungaged location is between two gaging stations and (or) partial-record stations on the same stream; and (4) when the ungaged location is not on the same stream as a gaging station or partial-record station. Two locations were considered to be on the same stream when the stream path from the downstream location to the basin divide followed the stream segment with the largest drainage area at each stream confluence and passed through the upstream location.

It is necessary to determine if the ungaged location is near a gaging station or partial-record station, and arithmetic methods are used to quantify the definition of “near.” A drainage-area-ratio method for estimating statistics at ungaged locations has been used by several researchers, including Hayes (1991), Ries and Friesz (2000), and Flynn (2003). Ratio of the drainage areas ($R_{U/K}$) is defined as the ratio of the drainage area where the value of the statistic is unknown (A_U) to the drainage area where the value of the statistic is known (A_K). These three researchers use arithmetic methods for determining the upstream and downstream limits of the range of drainage-area ratios over which streamflow statistics can be accurately estimated from those at a gaging station using an equation similar to the following:

$$Q_U = Q_K \, (R_{U/K})^{EX} \, , \qquad \textbf{(2)}$$

where

- Q_U is the value of the unknown streamflow statistic, in ft^3/s;
- Q_K is the value of the known streamflow statistic, in ft^3/s;
- $R_{U/K}$ is the ratio of the drainage area at the location of the unknown streamflow (A_U) to the drainage area at the location of the known streamflow (A_K), unitless; and
- EX is the exponent for the particular statistic, unitless.

The 1, 3, 7, 14, and 30Q2; 1, 3, 7, 14, and 30Q5; 1, 3, 7, 14, and 30Q10; 1B3; 4B3; HM; and D10, 25, 50, 75, and 90 statistics at 26 pairs of gaging stations located on the same stream in West Virginia and adjacent states (table 2) were evaluated to quantify the definition of “near” for application of the drainage-area-ratio method in this study. Two computations, one upstream and one downstream, were made for each pair of stations. Ratios of the drainage areas for the 52 computations ranged from 0.21 to 4.76 (table 3). Equation 2 was solved for the exponent EX as the dependent variable:

$$\log_{10}(Q_U / Q_K) = EX\,(\log_{10}(R_{U/K})) \, . \qquad \textbf{(3)}$$

The exponent was evaluated for the 26 pairs of gaging stations for each statistic using simple linear regression with no intercept (regression line goes through the graph origin). The values of the exponent (EX) for the flow durations were about 1.0 and the remaining values ranged from 1.1 to 1.6, and averaged about 1.3 (table 3).

The upstream and downstream limits for application of drainage-area ratios used to quantify the definition of “near” were determined by plotting the drainage-area ratio against the absolute percent difference between the value of the flow statistic at the gaging station determined from the station record and the value of the statistic estimated by applying (1) the regional equations, and (2) the drainage-area-ratio method. S-PLUS 7.0 (Insightful Corporation, 2005), a commercially available statistical computing package, was used to construct locally weighted regression (LOESS) curves (a data-smoothing technique) through differences between flow statistics computed from gaging-station records and the estimated values (selected LOESS parameters were “span = 0.4–to–0.5; degree = one, locally-linear fitted; family = Symmetric, no feature for handling outlier distortions, strictly applying locally-linear fitting”). For example, the absolute percent differences for estimates of D25 made by applying the drainage-area-ratio method are lower than the estimates made by applying the regional equation at drainage-area ratios greater than about 2.5 and less than about 0.4 (fig. 4). Between ratios of 0.4 and 2.5, there is no significant difference between estimates determined from regional equations and estimates determined from drainage-area ratios. The upstream and downstream limits used to quantify the definition of “near” for D25 are set to the minimum and maximum ratios studied—0.21 and 4.76, respectively (table 3)—because drainage-area ratios provide estimates that are equal to or better than those provided by the regional equations over the entire range of ratios investigated, and no information is available outside these ratios with which to make an assessment.

The LOESS curves for 3Q5 (fig. 5) indicate the drainage-area ratios provide a better estimate than the equations when the ratios are between the upstream limit of about 0.5 and the downstream limit of about 2.0 (table 3). LOESS curves for most of the statistics are similar to those for 3Q5, with “near” quantified as drainage-area ratios with upstream and downstream limits equal to those of 3Q5 (table 3). The LOESS curve for 1Q2 indicated upstream and downstream limits equal to 0.5 and 2.0, respectively. The LOESS curve for D10 indicated the same upstream and downstream limits as for D25 (0.21 and 4.76, respectively). LOESS curves for D50 and D75 indicated upstream and downstream limits equal to 0.21 and 2.0, respectively.

Table 2. Pairs of U.S. Geological Survey streamflow-gaging stations in West Virginia or adjacent states that were evaluated to quantify the definition of “near” for application of the drainage-area-ratio method in this study.

[MD, Maryland; WV, West Virginia; VA, Virginia; PA, Pennsylvania; OH, Ohio; KY, Kentucky]

Pair number	Upstream station			Downstream station		
	Station number	Station name	State	Station number	Station name	State
1	01600000	NORTH BRANCH POTOMAC RIVER AT PINTO	MD	01603000	NORTH BRANCH POTOMAC RIVER NEAR CUMBERLAND	MD
2	01607500	SOUTH FORK SOUTH BRANCH POTOMAC RIVER AT BRANDYWINE	WV	01608000	SOUTH FORK SOUTH BRANCH POTOMAC RIVER NEAR MOOREFIELD	WV
3	01606500	SOUTH BRANCH POTOMAC RIVER NEAR PETERSBURG	WV	01608500	SOUTH BRANCH POTOMAC RIVER NEAR SPRINGFIELD	WV
4	01615000	OPEQUON CREEK NEAR BERRYVILLE	VA	01616500	OPEQUON CREEK NEAR MARTINSBURG	WV
5	01637500	CATOCTIN CREEK NEAR MIDDLETOWN	MD	01638480	CATOCTIN CREEK AT TAYLORSTOWN	VA
6	01643700	GOOSE CREEK NEAR MIDDLEBURG	VA	01644000	GOOSE CREEK NEAR LEESBURG	VA
7	03050000	TYGART VALLEY RIVER NEAR DAILEY	WV	03050500	TYGART VALLEY RIVER NEAR ELKINS	WV
8	03050500	TYGART VALLEY RIVER NEAR ELKINS	WV	03051000	TYGART VALLEY RIVER AT BELINGTON	WV
9	03058500	WEST FORK RIVER AT BUTCHERVILLE	WV	03059000	WEST FORK RIVER AT CLARKSBURG	WV
10	03059000	WEST FORK RIVER AT CLARKSBURG	WV	03061000	WEST FORK RIVER AT ENTERPRISE	WV
11	03069500	CHEAT RIVER NEAR PARSONS	WV	03070000	CHEAT RIVER AT ROWLESBURG	WV
12	03070000	CHEAT RIVER AT ROWLESBURG	WV	03071000	CHEAT RIVER NEAR PISGAH	WV
13	03152000	LITTLE KANAWHA RIVER AT GLENVILLE	WV	03153500	LITTLE KANAWHA RIVER AT GRANTSVILLE	WV
14	03153500	LITTLE KANAWHA RIVER AT GRANTSVILLE	WV	03155000	LITTLE KANAWHA RIVER AT PALESTINE	WV
15	03180500	GREENBRIER RIVER AT DURBIN	WV	03182500	GREENBRIER RIVER AT BUCKEYE	WV
16	03182500	GREENBRIER RIVER AT BUCKEYE	WV	03183500	GREENBRIER RIVER AT ALDERSON	WV
17	03183500	GREENBRIER RIVER AT ALDERSON	WV	03184000	GREENBRIER RIVER AT HILLDALE	WV
18	03187000	GAULEY RIVER AT CAMDEN ON GAULEY	WV	03189100	GAULEY RIVER NEAR CRAIGSVILLE	WV
19	03189100	GAULEY RIVER NEAR CRAIGSVILLE	WV	03189500	GAULEY RIVER NEAR SUMMERSVILLE	WV
20	03190000	MEADOW RIVER AT NALLEN	WV	03190400	MEADOW RIVER NEAR MOUNT LOOKOUT	WV
21	03189500	GAULEY RIVER NEAR SUMMERSVILLE	WV	03192000	GAULEY RIVER ABOVE BELVA	WV
22	03194700	ELK RIVER BELOW WEBSTER SPRINGS	WV	03195000	ELK RIVER AT CENTRALIA	WV
23	03195500	ELK RIVER AT SUTTON	WV	03197000	ELK RIVER AT QUEEN SHOALS	WV
24	03203600	GUYANDOTTE RIVER AT LOGAN	WV	03204000	GUYANDOTTE RIVER AT BRANCHLAND	WV
25	03210000	JOHNS CREEK NEAR META	KY	03211500	JOHNS CREEK NEAR VAN LEAR	KY
26	03213000	TUG FORK AT LITWAR	WV	03214000	TUG FORK NEAR KERMIT	WV

Upstream From a Gaging Station or Partial-Record Station

This procedure is used when there is a gaging station or partial-record station downstream from the ungaged location but none upstream on the same stream. The hydrologic assumption for this circumstance is that the conditions affecting streamflow, such as lithology, structure of rock formations, and evapotranspiration, are unchanged upstream from those at the gaging station or partial-record station. Mathematically, the value of a statistic is proportioned by drainage area. It is suggested to establish a partial-record station when $R_{U/K}$ is less than or equal to the upstream limit of the ratio of drainage areas (R_{US}) (table 3). The following equation is used to estimate the value of a statistic:

$$Q_U = Q_K \left(R_{U/K}\right)^{EX}, \qquad (4)$$

where

Q_U is the value of the unknown streamflow statistic, in ft^3/s;

Q_K is the value of the known streamflow statistic, in ft^3/s;

$R_{U/K}$ is the ratio of the drainage area at the location of the unknown streamflow (A_U) to the drainage area at the location of the known streamflow (A_K), unitless, and

EX is the exponent for the particular statistic (table 3), unitless.

In this method, it is not assumed that the conditions affecting streamflow change toward the regional tendency in the upstream direction. The critical situation for this assumption is where the unknown location approaches the headwaters of a stream. This assumption might be acceptable if the streamflow at the known location is greater than that estimated by applying the regional equation at the known location because the upstream estimate of streamflow would be lower, and therefore more conservative from an availability or regulatory perspective, than that determined using the method presented. However, this assumption would be unacceptable if the streamflow at the known location is less than that estimated by applying the regional equation because the upstream estimate of streamflow would be greater than that determined using the method presented, thus requiring an assumption of additional unit inflow. The method presented requires establishing a partial-record station in order to increase unit inflow when the known streamflow is less than that estimated using the regional equation.

Downstream From a Gaging Station or Partial-Record Station

This procedure is used when there is a gaging station or partial-record station upstream from the ungaged location but none downstream on the same stream. The hydrologic assumption for this circumstance is that the conditions affecting streamflow at the gaging station or partial-record station change, in the downstream direction, toward those of the regional tendency. The conditions affecting streamflow in the vicinity of the gaging station or partial-record station could be an aquifer, land use, or diversion that would likely not significantly affect streamflow if the drainage area were larger and, therefore, conditions were more similar to the regional tendency. Mathematically, the value of a statistic is changed to that estimated by applying the regional equation as $R_{U/K}$ approaches the downstream limit of the ratio of drainage areas (R_{DS}) (table 3). It is suggested to establish a partial-record station when $R_{U/K}$ is greater than or equal to R_{DS}. The value of the

Table 3. Values of the exponent (EX), and upstream (R_{US}) and downstream (R_{DS}) limits of the drainage-area ratios used to quantify the definition of “near” for estimating selected statistics at ungaged locations in West Virginia and adjacent states.

[xQy, x-day y-year hydrologically based flow, in cubic feet per second; xBy, x-day y-year biologically based flow, in cubic feet per second; HM, U.S. Environmental Protection Agency harmonic-mean flow, in cubic feet per second; Dn, n-percent-duration flow, in cubic feet per second]

Statistic	Exponent (EX), unitless	Upstream limit (R_{US}), unitless	Downstream limit (R_{DS}), unitless
1Q2	1.30	0.5	2
1Q5	1.34	.5	2
1Q10	1.37	.5	2
3Q2	1.28	.5	2
3Q5	1.30	.5	2
3Q10	1.33	.5	2
7Q2	1.26	.5	2
7Q5	1.34	.5	2
7Q10	1.54	.5	2
14Q2	1.19	.5	2
14Q5	1.49	.5	2
14Q10	1.41	.5	2
30Q2	1.10	.5	2
30Q5	1.34	.5	2
30Q10	1.60	.5	2
1B3	1.29	.5	2
4B3	1.29	.5	2
HM	1.25	.5	2
D10	.95	.21	4.76
D25	.95	.21	4.76
D50	.96	.21	2
D75	1.01	.21	2
D90	1.15	.5	2

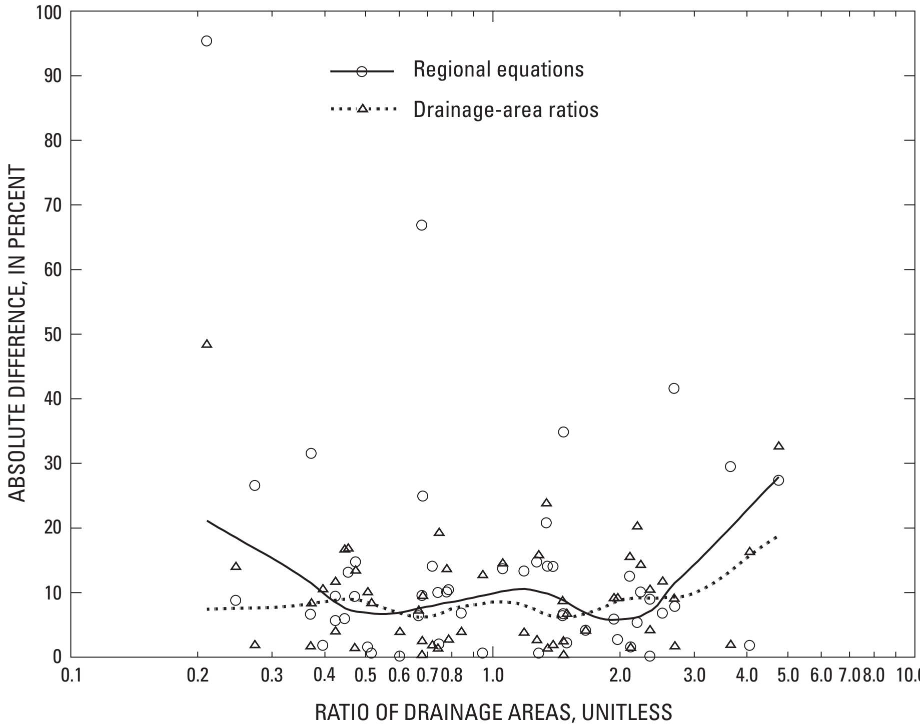


Figure 4. LOESS curves of absolute differences between 25-percent-duration flow (D25) determined at streamflow-gaging stations and values estimated from (1) regional equations, and (2) drainage-area ratios.

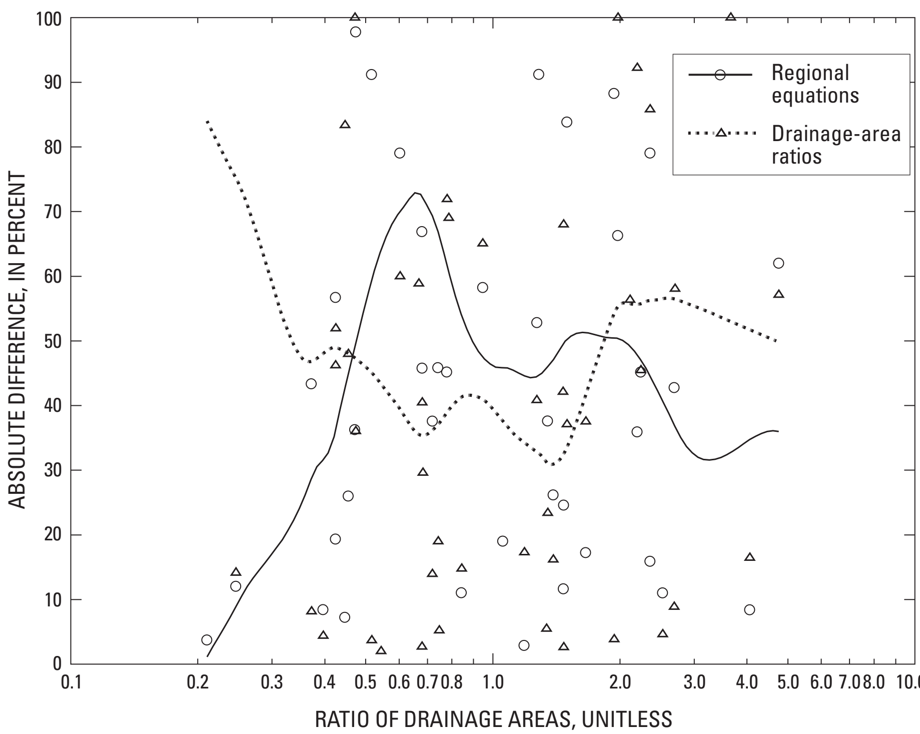


Figure 5. LOESS curves of absolute differences between 3-day, 5-year hydrologically based flow (3Q5) determined at streamflow-gaging stations and values estimated from (1) regional equations, and (2) drainage-area ratios. (Some values greater than 100 percent are not shown.)

statistic is estimated by applying the regional equation when $R_{U/K}$ is greater than or equal to R_{DS}, and the following equation is used to estimate the value of a statistic when $R_{U/K}$ is less than R_{DS}:

$$Qu = Q_{UE} + (Q_K - Q_{KE})(R_{DS} - R_{U/K}) / (R_{DS} - 1), \quad (5)$$

when

$R_{U/K} < R_{DS}$; and

where

- Q_U is the value of the unknown streamflow statistic, in ft^3/s;
- Q_{UE} is the regional equation evaluated at the location of the unknown value of the streamflow statistic, in ft^3/s;
- Q_K is the value of the known streamflow statistic, in ft^3/s;
- Q_{KE} is the regional equation evaluated at the location of the known value of the streamflow statistic, in ft^3/s;
- R_{DS} is the downstream limit of the ratio of drainage areas (table 3), unitless; and
- $R_{U/K}$ is the ratio of the drainage area at the location of the unknown streamflow (A_U) to the drainage area at the location of the known streamflow (A_K), unitless.

Between Gaging Stations and (or) Partial-Record Stations

This procedure is used when there are gaging stations or partial-record stations both upstream and downstream from the ungaged location on the same stream. The hydrologic assumption for this circumstance is that the conditions affecting streamflow are changing on the basis of the relation between the streamflows at the gaging stations or partial-record stations, the ratios of the drainage areas, and differences between regional hydrologic conditions and those that affect streamflows at the stations. It is suggested to establish a partial-record station when $R_{U/K}$ is greater than R_{DS} and less than R_{US}, and when one of the values of the streamflow statistic at the upstream and downstream locations is greater than and one of the values is less than that estimated by applying the regional equation. Two alternative hydrologic assumptions are described in detail below.

Hydrologic Conditions Change Linearly between Stations

This hydrologic assumption is that the conditions affecting streamflow at the upstream gaging station or partial-record station change linearly with drainage area to those at the downstream location when (1) both gaging stations or partial-record stations are near the ungaged location, or (2) both gaging stations or partial-record stations are not near the ungaged location, but the hydrologic conditions affecting streamflow at the stations and ungaged location are consistent. The conditions affecting streamflow between the upstream and downstream locations are well defined by streamgages or partial-record stations when both are near the ungaged location; the conditions affecting streamflow are consistent from the upstream to the downstream location when neither is near the ungaged location and the conditions at both stations are more similar to each other than to the regional hydrologic conditions. Mathematically, the value of a statistic changes linearly with respect to drainage area from the upstream to the downstream value when (1) $R_{U/K}$ is less than R_{DS} at the upstream location and $R_{U/K}$ is greater than R_{US} at the downstream location; or (2) $R_{U/K}$ is greater than or equal to R_{DS} at the upstream location, $R_{U/K}$ is less than or equal to R_{US} at the downstream location, and the values of the statistic at the upstream and downstream locations are both greater than or both less than those estimated by applying the regional equation. The following equation is used to estimate the value of a statistic under the limitations described above:

$$Q_U = [Q_{US}(A_{DS} - A_U) + Q_{DS}(A_U - A_{US})] / (A_{DS} - A_{US}), \quad (6)$$

when

$R_{U/K} < R_{DS}$ at upstream location and $R_{U/K} > R_{US}$ at downstream location,

or when

$R_{U/K} \geq R_D$ at upstream location and $R_{U/K} \leq R_{US}$ at downstream location; and
Q_{KE} at the upstream location $> Q_{US}$ and Q_{KE} at the downstream location $> Q_{DS}$ or
Q_{KE} at the upstream location $< Q_{US}$ and Q_{KE} at the downstream location $< Q_{DS}$; and

where

- Q_U is the value of the unknown streamflow statistic, in ft^3/s;
- Q_{US} is the value of the streamflow statistic at the upstream location, in ft^3/s;
- Q_{DS} is the value of the streamflow statistic at the downstream location, in ft^3/s;
- Q_{KE} is the regional equation evaluated at the location of the known value of the streamflow statistic, in ft^3/s;
- A_U is the drainage area at the location of the unknown value of the streamflow statistic, in mi^2;
- A_{US} is the drainage area at the upstream location, in mi^2;
- A_{DS} is the drainage area at the downstream location, in mi^2;
- R_{US} is the upstream limit of the ratio of drainage areas (table 3), unitless; and
- R_{DS} is the downstream limit of the ratio of drainage areas (table 3), unitless.

Hydrologic Conditions Change Linearly to the Regional Hydrologic Conditions between Stations

This hydrologic assumption is that the conditions affecting streamflow at the gaging station or partial-record station change linearly with drainage area to those represented by the regional equation when neither the gaging station nor the partial-record station is near the ungaged location and the hydrologic conditions affecting streamflow at the two stations are inconsistent. The conditions affecting streamflow can be significantly different as a result of factors such as input from a productive aquifer, input from a tributary stream that is hydrologically different from the stream on which the station is located or from regional hydrologic conditions, water withdrawal for domestic or industrial use, an outcrop of impervious rock strata, transfer of water to deeper rock strata, or transfer of water to or from an underground mine (possibly into or out of the basin). Mathematically, the hydrologic conditions are inconsistent if one of the values of the statistic at the upstream and downstream locations is greater than and one of the values is less than those estimated by applying the regional equation—that is, the streamflow changes from a value greater than the value estimated by applying the regional equation at the upstream location to a value equal to the value estimated from the regional equation, and then to a value less than the value estimated from the regional equation at the downstream location, or the reverse. The regional equation is applied to estimate the value of a statistic if $R_{U/K}$ is greater than or equal to R_{DS} at the upstream location and $R_{U/K}$ is less than or equal to R_{US} at the downstream location. The equation used in the procedures described in the section "Downstream from a Gaging Station or Partial-Record Station" is used to estimate the value of a statistic if $R_{U/K}$ is less than R_{DS} at the upstream location and $R_{U/K}$ is less than or equal to R_{US} at the downstream location. The following equation is used to estimate the value of a statistic if $R_{U/K}$ is greater than or equal to R_{DS} at the upstream location and $R_{U/K}$ is greater than R_{US} at the downstream location:

$$Q_U = Q_{UE} + (Q_K - Q_{KE})\,(R_{U/K} - R_{US})\,/\,(1 - R_{US})\ , \qquad \textbf{(7)}$$

when

$R_{U/K} \geq R_{DS}$ at upstream location and
$R_{U/K} > R_{US}$ at downstream location; and

where

Q_U is the value of the unknown streamflow statistic, in ft^3/s;

Q_{UE} is the regional equation evaluated at the location of the unknown value of the streamflow statistic, in ft^3/s;

Q_K is the value of the known streamflow statistic, in ft^3/s;

Q_{KE} is the regional equation evaluated at the location of the known value of the streamflow statistic, in ft^3/s;

R_{DS} is the downstream limit of the ratio of drainage areas (table 3), unitless;

R_{US} is the upstream limit of the ratio of drainage areas (table 3), unitless; and

$R_{U/K}$ is the ratio of the drainage area at the location of the unknown streamflow (A_U) to the drainage area at the location of the known streamflow (A_K), unitless.

Not on the Same Stream as a Gaging Station or Partial-Record Station

This procedure is used when there is no gaging station or partial-record station on the same stream as the ungaged location. The hydrologic assumption for this circumstance is that the conditions affecting streamflow are those represented by the regional equation. A partial-record station could be established when there is no gaging station or partial-record station on the same stream as the ungaged location. Mathematically, the value of a statistic is estimated by applying the regional equation.

Example Applications of Estimating Procedures

The example applications of the estimating procedures are presented for manual computations of statistics at theoretical locations considering only the statistics from the gaging stations. The USGS West Virginia Water Science Center, National Resource Analysis Center (NRAC), and West Virginia Department of Environmental Protection (WVDEP) are incorporating the estimating procedures into the Watershed Characterization Management System (WCMS) for electronic computation of statistics for all streams in West Virginia. This system will incorporate the statistics from gaging stations and partial-record stations, and will encompass statistics available from various sources for regulated streams. WCMS is a map-based Web applications system developed by the NRAC (associated with West Virginia University) for the WVDEP. WCMS will allow WVDEP to add and revise partial-record and regulated locations and statistics. WCMS is similar to the USGS "StreamStats" program (Ries and others, 2004), and is used by government agencies in managing the natural resources of West Virginia.

- **Example 1:** The annual 30Q5 at an ungaged location with a drainage area of 84.2 mi^2 (A_U) upstream from the gaging station 01607500, South Fork South Branch Potomac River at Brandywine, can be estimated using **Equation 4**. This equation can be used because there are no additional gaging stations upstream on the same stream (fig. 2) and the size of the drainage area is within the limits for which the estimate can be made using drainage-area ratios. The annual 30Q5 at the gaging station is 4.53 ft^3/s (Q_K) (Wiley, 2006, p. 73), and the drainage area at the gaging stations is 103 mi^2 (A_K) (Paybins, 2008). The ratio of drainage areas ($R_{U/K}$) is A_U divided by A_K, or 84.2 mi^2 divided by 103 mi^2, which is 0.817. The exponent (E) for 30Q5 is 1.34 (table 3). Substituting into **Equation 4**, the 30Q5 at the ungaged location (Q_U) is 3.46 ft^3/s. The value of $R_{U/K}$ of 0.817 is greater than the upstream limit (R_{US}) of 0.21 (table 3), indicating the establishment of a partial-record station is not necessary.

- **Example 2:** The annual D75 at an ungaged location with a drainage area of 1,450 mi^2 (A_U) downstream from the gaging station 03183500, Greenbrier River at Alderson, with a drainage area of 1,364 mi^2 (Paybins, 2008) (A_{US}) and upstream from the gaging station 03184000, Greenbrier River at Hilldale, with a drainage area of 1,619 mi^2 (Paybins, 2008) (A_{DS}) can be estimated using **Equation 6**. This equation can be used because the ratio of drainage areas ($R_{U/K}$ or A_U divided by A_K) is equal to 1,450 mi^2 divided by 1,364 mi^2, or 1.06, which is less than the R_{DS} of 4.76 (table 3) at the upstream location, and the $R_{U/K}$ (A_U divided by A_K, or 1,450 mi^2 divided by 1,619 mi^2) of 0.896 is greater than the R_{US} of 0.21 (table 3) at the downstream location (fig. 2). The annual D75 at Alderson (Q_{US}) is 301 ft^3/s (Wiley, 2006, p. 116) and at Hilldale (Q_{DS}) is 342 ft^3/s (Wiley, 2006, p. 117). By substituting these values into **Equation 6**, the D75 at the ungaged location (Q_U) is 315 ft^3/s.

Accuracy and Limitations of Estimating Procedures

The estimating procedures presented in this report are applicable only to unregulated streams in West Virginia, and estimates are representative of the period 1930–2002. The procedures are not applicable to streams regulated by large lakes, ponds, and navigation dams. Equations are applicable only within the specified limits of dependent variables. The statistics for streamflow-gaging stations from surrounding states used in this study do not supersede values determined for the particular state.

Caution should be used when applying the estimating procedures in areas of underground mining and karst terrain, where water can be transferred between basins and streams can lose water. A partial-record station can be established where there is some streamflow for estimating statistics, but estimating procedures for ungaged locations should not be applied without first determining that the streams involved are not losing or gaining water to or from underground mines or karst geology.

Estimating procedures for ungaged locations are applicable only to perennial streams. The median drainage area upstream from the location where an intermittent stream becomes perennial was determined to be 40.8 acres (0.064 mi^2). This value ranged from 10.2 to 150.1 acres (0.016 to 0.235 mi^2) in a limited study of 36 sites conducted in the southern coal fields of West Virginia (Paybins, 2003), and differed by region, with a median of 66.1 acres (0.103 mi^2) in the northeastern part of the southern coal fields and 34.8 acres (0.054 mi^2) in the southwestern part. Estimating procedures should not be applied to drainage areas less than 0.05 mi^2 because the streams are likely not perennial; procedures should be applied to drainage areas less than 0.25 mi^2 only when there is some determination (such as a field observation at low streamflow) that the stream is perennial.

The estimating procedures presented in this report, unlike the methods developed by Hayes (1991), are not conservative at the confluence of streams. The value of the statistic estimated downstream from the confluence of two streams will not equal the summation of the values of the statistic estimated upstream from the confluence. Low streamflows can be affected by mining (Hobba, 1981; Puente and Atkins, 1989; Borchers and others, 1991; Wiley and others, 2001), which can result in differences in timing and magnitude of base-flow conditions between nearby locations. Streamflows may be reduced in streams that are “dewatered” by underlying underground mines, or may be increased in streams that are downdip from flooded underground mines. Water also can be transferred between basins by drainage through coal mines, and low streamflows can be increased by drainage from valley-fill deposits. Streamflow at outflow points of large basins that are stratigraphically below mined coal beds likely would be increased from the pre-mining condition, except where large interbasin transfer of water occurs. The variability of the effects caused by mining and other conditions is accounted for within the accuracy of the non-conservative estimating procedures developed in this study.

The estimating procedures could be applied to other streamflow statistics as well as those mentioned in this report. For example, flood statistics could easily be transferred using the methods of this study rather than those in Wiley and others (2000, 2002). Appendix A contains the drainage-area-ratio exponents and upstream and downstream limits for transferring published flood statistics. Seasonal streamflow statistics published by Wiley (2006) also could be transferred if regional equations were developed.

Summary

The U.S. Geological Survey, in cooperation with the West Virginia Department of Environmental Protection, Division of Water and Waste Management, developed procedures for estimating annual streamflow statistics on unregulated streams in West Virginia.

Regional equations were developed for estimating the 1-, 3-, 7-, 14-, and 30-day 2-year; 1-, 3-, 7-, 14-, and 30-day 5-year; and 1-, 3-, 7-, 14-, and 30-day 10-year hydrologically

based low-flow frequency values for unregulated streams in West Virginia. Equations and procedures for the 1-day, 3-year and 4-day, 3-year biologically based low-flow frequency values; the U.S. Environmental Protection Agency harmonic-mean flows; and the 10-, 25-, 50-, 75-, and 90-percent flow-duration values also were developed. Regional equations were developed using ordinary least squares regression using flow statistics from 117 streamflow-gaging stations as dependent variables with basin characteristics for these gaging stations as independent variables. Generalized least squares regression was not used because biological frequencies, USEPA harmonic-mean flows, and flow durations do not have annual time series, and because the record periods for all stations represent equal periods of 73 years (1930–2002).

Equations were developed for three hydrologic regions—North, South-Central, and Eastern Panhandle. Drainage area, precipitation, and longitude of the basin centroid were significant independent variables in one or more of the regional regression equations.

Estimating procedures are presented for determining statistics at gaging stations, partial-record stations, and ungaged locations, including (1) an ungaged location upstream from a gaging station or partial-record station, (2) an ungaged location downstream from a gaging station or partial-record station, (3) an ungaged location on a stream other than the one on which the gaging station or partial-record station is located, and (4) an ungaged location between two gaging stations or partial-record stations. The procedures are based on a comparison of estimates made at 26 pairs of gaging stations by using drainage-area ratios and estimates made using the regional regression equations. Example applications of estimating procedures are presented.

Acknowledgments

The author thanks USGS colleagues Rodney E. Southard and Donald C. Hayes for their technical reviews of this report. The author is also grateful to William P. Bartlett, USGS Surface-Water Specialist for the Northeast and Midwest Areas, for suggesting areas for additional discussion. The technical assistance of USGS West Virginia Water Science Center colleagues Mark D. Kozar, John T. Atkins, Jr., Katherine S. Paybins, Ronald D. Evaldi, and Hugh E. Bevans also is appreciated.

References Cited

Borchers, J.W., Ehlke, T.A., Mathes, M.V., and Downes, S.C., 1991, The effects of coal mining on the hydrologic environment of selected stream basins in southern West Virginia: U.S. Geological Survey Water-Resources Investigations Report 84-4300, 119 p.

Eng, K., and Milly, P.C.D., 2007, Relating low-flow characteristics to the base flow recession time constant at partial record stream gages: Water Resources Research, v. 43, W01201, doi:10.1029/2006WR005293, 8 p.

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

Flynn, R.H., 2003, Development of regression equations to estimate flow durations and low-flow-frequency statistics in New Hampshire streams: U.S. Geological Survey Water-Investigations Report 02-4298, 66 p.

Friel, E.A., Embree, W.N., Jack, A.R., and Atkins, J.T., Jr., 1989, Low-flow characteristics of streams in West Virginia: U.S. Geological Survey Water-Resources Investigations Report 88-4072, 34 p.

Frye, P.M., and Runner, G.S., 1969, Procedure for estimating magnitude and frequency of floods in West Virginia: Charleston, West Virginia, West Virginia State Road Commission design directive, 10 p.

Frye, P.M., and Runner, G.S., 1970, A proposed streamflow data program for West Virginia: U.S. Geological Survey Open-File Report, 38 p.

Frye, P.M., and Runner, G.S., 1971, A preliminary report on small streams flood frequency in West Virginia: Charleston, West Virginia, West Virginia Department of Highways design directive, 9 p.

Hayes, D.C., 1991, Low-flow characteristics of streams in Virginia: U.S. Geological Survey Water-Supply Paper 2374, 69 p.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations Report, book 4, chap. A3, 510 p.

Hirsch, R.M., 1982, A comparison of four streamflow record extension techniques: Water Resources Research, v. 18, no. 4, p. 1081-1088.

Hirsch, R.M., and Gilroy, E.J., 1984, Methods of fitting a straight line to data: Examples in water resources: Water Resources Bulletin, v. 20, no. 5, p. 706–711.

Hobba, W.A., Jr., 1981, Effects of underground mining and mine collapse on the hydrology of selected basins in West Virginia: West Virginia Geologic and Economic Survey Report of Investigations RI-33, 77 p.

Insightful Corporation, 2005, S-PLUS 7.0 Professional edition software for Windows: Insightful Corporation, Seattle, Washington.

Lane, E.W., and Lei, Kai, 1950, Stream flow variability: American Society of Civil Engineers Transactions, v. 115, p. 1084–1134.

National Oceanic and Atmospheric Administration, 2006a, Climate of West Virginia, accessed September 30, 2006, at URL *http://cdo.ncdc.noaa.gov/climatenormals/clim60/states/Clim_WV_01.pdf*

National Oceanic and Atmospheric Administration, 2006b, Total precipitation in inches by month for climate divisions, accessed September 30, 2006, at URL *http://www.cdc.noaa.gov/USclimate/pcp.state.19712000.climo.html*

Natural Resources Conservation Service, 2006, West Virginia Precipitation data/maps, accessed September 30, 2006, at URL *http://www.ncgc.nrcs.usda.gov/products/datasets/climate/data/precipitation-state/wv.html*

Paybins, K.S., 2003, Flow origin, drainage area, and hydrologic characteristics for headwater streams in the mountaintop coal-mining region of southern West Virginia, 2000–01: U.S. Geological Survey Water-Resources Investigations Report 02-4300, 20 p.

Paybins, K.S., 2008, Basin characteristics for selected gaging stations in and near West Virginia: U.S. Geological Survey Scientific Investigations Report 2008-1087, 9 p.

Puente, C., and Atkins, J.T., 1989, Simulation of rainfall-runoff response in mined and unmined watersheds in coal areas of West Virginia: U.S. Geological Survey Water-Supply Paper 2298, 48 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.

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

Riggs, H.C., 1972, Low-flow investigations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. B1, 18 p.

Rossman, L.A., 1990, DFLOW user's manual: U.S. Environmental Protection Agency, 26 p.

Runner, G.S.,1980, Runoff studies on small drainage areas (Technique for estimating magnitude and frequency of floods in West Virginia), West Virginia Department of Highways Research Project 16: U.S. Geological Survey Open-File Report 80-1218, 44 p.

Searcy, J.K., 1959, Flow-duration curves—Manual of hydrology—part 2, Low-flow techniques: U.S. Geological Survey Water-Supply Paper 1542-A, 33 p.

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

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

U.S. Environmental Protection Agency, 1986, Technical guidance manual for performing wasteload allocations—Book VI, Design conditions, chapter 1, Stream design flow for steady-state modeling (EPA 440/4-86-014): U.S. Environmental Protection Agency, Office of Water, Washington, D.C., variously paged.

U.S. Geological Survey, 1990, National water summary 1987—Hydrologic events and water supply and use: U.S. Geological Survey Water Supply Paper 2350, 553 p.

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

U.S. Geological Survey, 2006, The National Map, accessed September 30, 2006, at URL *http://nationalmap.gov*

Wetzel, K.L., and Bettandorff, J.M., 1986, Techniques for estimating streamflow characteristics in the Eastern and Interior Coal Provinces of the United States: U.S. Geological Survey Water-Supply Paper 2276, 80 p.

Wiley, J.B., 2006, Low flow analysis and selected flow statistics representative of the period 1930–2002 for streamflow-gaging stations in or near West Virginia: U.S. Geological Survey Scientific Investigations Report 2006-5002, 190 p.

Wiley, J.B., Atkins, J.T., Jr., and Newell, D.A., 2002, Estimating the magnitude of annual peak discharges with recurrence intervals between 1.1 and 3.0 years for rural, unregulated streams in West Virginia: U.S. Geological Survey Water-Resources Investigations Report 02-4164, 73 p.

Wiley, J.B., Atkins, J.T., Jr., and Tasker, Gary D., 2000, Estimating magnitude and frequency of peak discharges for rural, unregulated streams in West Virginia: U.S. Geological Survey Water-Resources Investigations Report 00-4080, 93 p.

Wiley, J.B., Evaldi, R.D., Eychaner, J.H., and Chambers, D.B., 2001, Reconnaissance of stream geomorphology, low streamflow, and stream temperature in the mountaintop coal-mining region, southern West Virginia, 1999-2000: U.S. Geological Survey Water-Resources Investigations Report 01-4092, 34 p.

Appendix 1. Transferring Flood Statistics to Ungaged Locations

The estimating procedures presented by Wiley and others (2000, 2002) can be replaced by the estimating procedures in the present study.

The exponents and the upstream and downstream limits for application of the drainage-area ratios to quantify the definition of "near" (table A1) were determined from the drainage-area ratios at the same 26 pairs of stations used in the present study, and previously developed flood statistics and regional equations (Wiley and others, 2000, 2002). Only 21 pairs of stations were available for computing ratios for the (1.1–1.9)-, 2.5-, and 3-year frequency values because these statistics were not published for stations outside West Virginia. The exponents for the (1.1–1.9)-, 2.5-, and 3-year frequency values were estimated based on the relation between the exponents determined from the 26 pairs of stations and the exponents determined from the 21 pairs of stations between the 1.9- and 5-year frequencies; 0.1 was subtracted from the values of the exponents determined from the 21 pairs of stations to estimate values expected from the 26 pairs of stations.

The upstream and downstream limits of the equations were determined by plotting the drainage-area ratio against the absolute percent difference between the frequency values determined at the gaging stations and the frequency values (1) estimated by applying the regional equations, and (2) estimated by applying the drainage-area-ratio method. S-PLUS 7.0 (Insightful Corporation, 2005), a commercially available statistical computing package, was used to construct locally weighted regression (LOESS) curves (a data-smoothing technique) through the differences (selected LOESS parameters were "span = 0.4–to–0.5; degree = one, locally-linear fitted; family = Symmetric, no feature for handling outlier distortions, strictly applying locally-linear fitting"). Estimates of the 2-year recurrence-interval value made by applying the drainage-area-ratio method were associated with lower absolute percent differences than estimates made by applying the regional equation at drainage-area ratios greater than about 0.4 (fig. A1), and were associated with greater absolute percent differences at drainage-area ratios less than about 0.4. There was no significant difference between the estimates of the 50-year recurrence-interval value made by applying the drainage-area-ratio method and applying the regional equation (fig. A2). There was no significant difference between the estimates of the 100-year recurrence-interval value from applying the drainage-area ratio method and applying the regional equation at drainage-area ratios greater than about 0.4 (fig. A3), and estimates from applying the drainage-area-ratio method were associated with a greater absolute percent difference at drainage-area ratios less than about 0.4. Relations were similar to those of the 2-year recurrence-interval value for intervals less than 25 years, similar to those of the 50-year recurrence-interval value for the 25-year interval, and similar to those of the 100-year recurrence-interval value for intervals greater than 50 years. The upstream limit for applying the drainage-area-ratio method was determined to be 0.40 by plotting the estimates for all recurrence intervals less than 25 years (fig. A4). The upstream limit for applying the drainage-area-ratio method with intervals from 25 to 500 years was set at 0.21, which was the minimum drainage-area ratio considered in the analysis. Applying the drainage-area-ratio method resulted in less or little absolute difference in the estimates than applying the regional equations, and no information below a ratio of 0.21 was available to make an assessment. The downstream limit was set at 4.76, the maximum drainage-area ratio considered in the analysis, because no information above that ratio was available to make an assessment.

Table A1. Values of the exponent (EX), and upstream (R_{US}) and downstream (R_{DS}) limits of the drainage-area ratios used to quantify the definition of "near" for estimating selected flood statistics at ungaged locations in West Virginia and adjacent states.

[Q(n), discharge in cubic feet per second for the (n)-year recurrence interval; all values are unitless; estimates for some exponents were made from the calculated value indicated in parentheses]

Flood statistic	EX	R_{US}	R_{DS}	Flood statistic	EX	R_{US}	R_{DS}
Q(1.1)	0.68 (0.78)	0.40	4.76	Q(2.5)	0.66 (0.76)	0.40	4.76
Q(1.2)	.68 (.78)	.40	4.76	Q(3)	.66 (.76)	.40	4.76
Q(1.3)	.68 (.78)	.40	4.76	Q(5)	.65	.40	4.76
Q(1.4)	.68 (.78)	.40	4.76	Q(10)	.64	.40	4.76
Q(1.5)	.67 (.77)	.40	4.76	Q(25)	.63	.21	4.76
Q(1.6)	.67 (.77)	.40	4.76	Q(50)	.63	.21	4.76
Q(1.7)	.67 (.77)	.40	4.76	Q(100)	.63	.21	4.76
Q(1.8)	.67 (.77)	.40	4.76	Q(200)	.62	.21	4.76
Q(1.9)	.67 (.77)	.40	4.76	Q(500)	.61	.21	4.76
Q(2)	.66	.40	4.76				

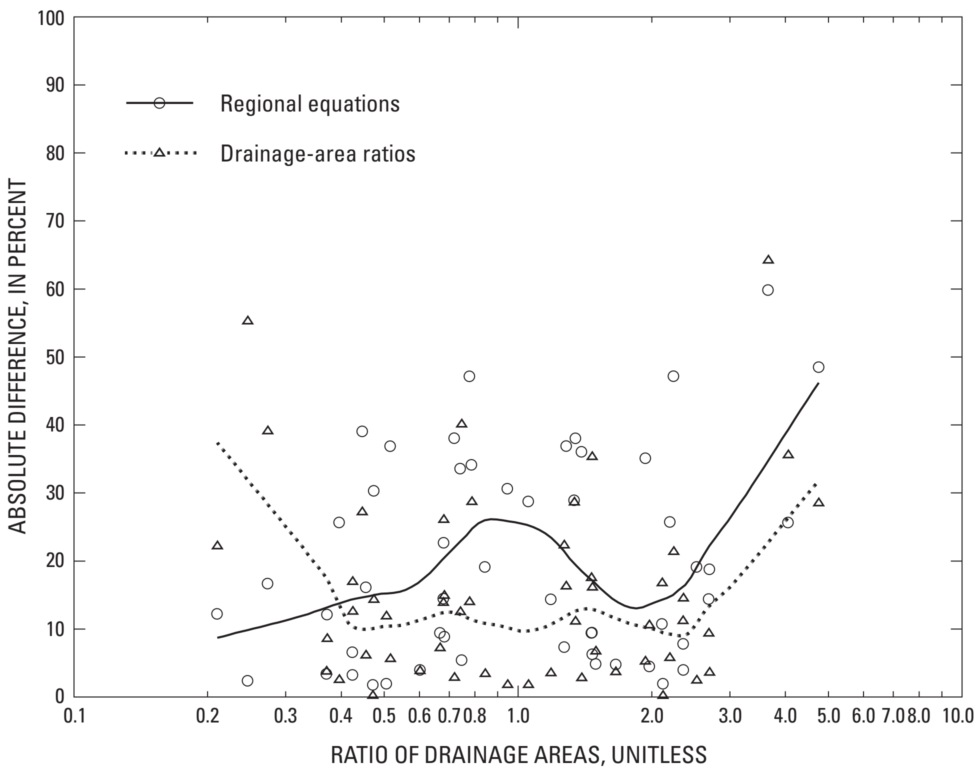


Figure A1. LOESS curves of absolute differences between the 2-year recurrence-interval values (Q(2)) determined at streamflow-gaging stations and values estimated from (1) regional equations, and (2) drainage-area ratios.

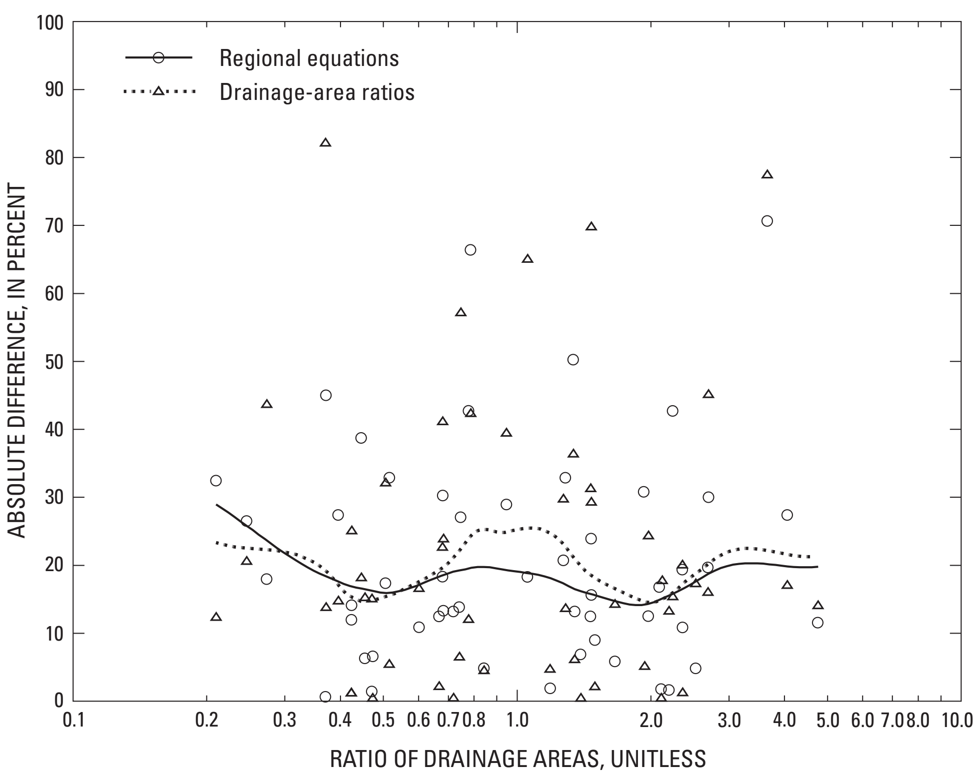


Figure A2. LOESS curves of absolute differences between the 50-year recurrence-interval values (Q(50)) determined at streamflow-gaging stations and values estimated from (1) regional equations, and (2) drainage-area ratios.

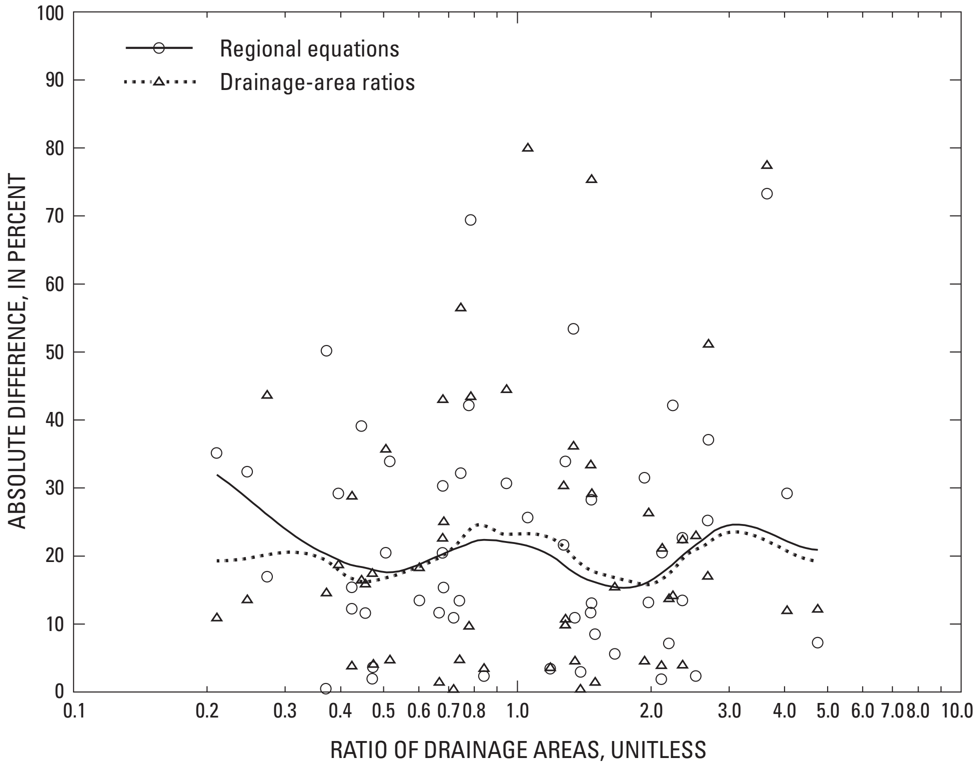


Figure A3. LOESS curves of absolute differences between the 100-year recurrence-interval values (Q(100)) determined at streamflow-gaging stations and values estimated from (1) regional equations, and (2) drainage-area ratios.

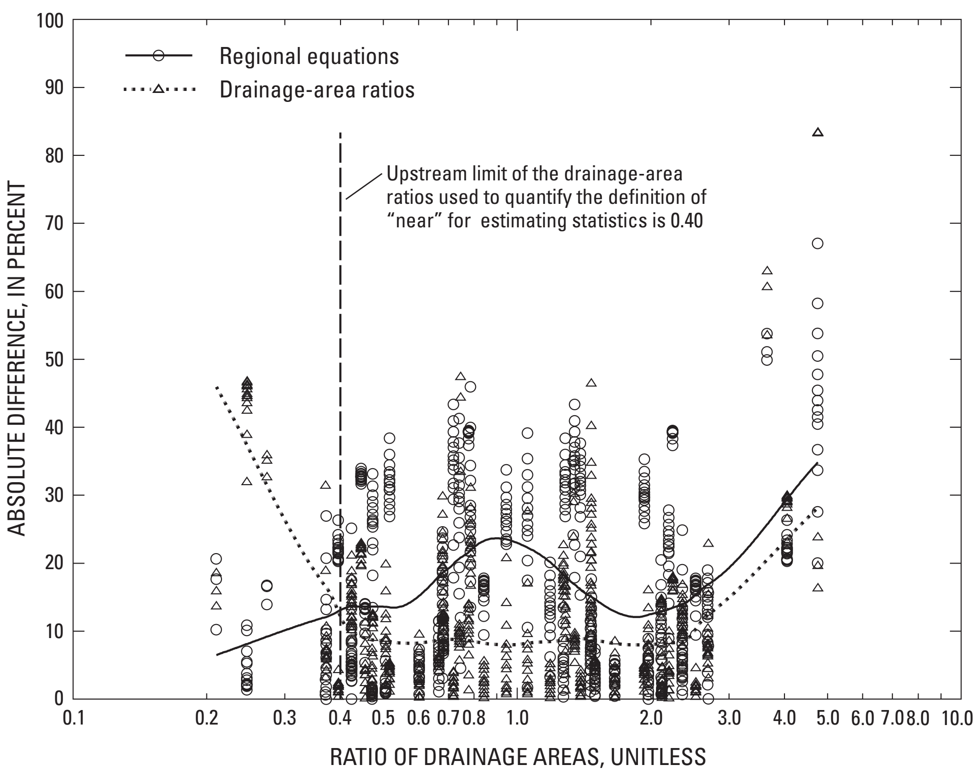


Figure A4. LOESS curves of absolute differences between the 1.1- and 10-year recurrence-interval values (Q(1.1)–Q(10)) determined at streamflow-gaging stations and values estimated from (1) regional equations, and (2) drainage-area ratios.

For additional information, write to:
Director
U.S. Geological Survey
West Virginia Water Science Center
11 Dunbar St
Charleston, WV 25301

or visit our Web site at:
http://wv.usgs.gov/

Document prepared by the West Trenton Publishing Service Center

USGS

Wiley—**Estimating Selected Streamflow Statistics Representative of 1930–2002 in West Virginia**—Scientific Investigations Report 2008–5105

ISBN 978-1-4113-2260-8

9 781411 322608

Printed on recycled paper