

Prepared in cooperation with the Minnesota Pollution Control Agency

Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams in Minnesota



Scientific Investigations Report 2015–5170

Cover. Redwood River near Redwood Falls, Minnesota. Photograph by the U.S. Geological Survey.

Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams in Minnesota

By Jeffrey R. Ziegeweid, David L. Lorenz, Chris A. Sanocki, and Christiana R. Czuba

Prepared in cooperation with the Minnesota Pollution Control Agency

Scientific Investigations Report 2015–5170

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

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

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

U.S. Geological Survey, Reston, Virginia: 2015

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

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

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

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

Suggested citation:

Ziegeweid, J.R., Lorenz, D.L., Sanocki, C.A., and Czuba, C.R., 2015, Methods for estimating flow-duration curve and low-flow frequency statistics for ungaged locations on small streams in Minnesota: U.S. Geological Survey Scientific Investigations Report 2015–5170, 23 p., <http://dx.doi.org/10.3133/sir20155170>.

ISSN 2328-0328 (online)

Contents

Abstract.....	1
Introduction.....	1
Purpose and Scope	2
Description of Study Area	2
Methods for Data Development	3
Trend Analyses.....	3
Flow-Duration Curves.....	5
Low-Flow Frequency	6
Basin Characteristics.....	6
Regression Analysis.....	8
Regional Regression Analysis	9
Flow-Duration Curve Regression Equations	9
Low-Flow Frequency Regression Equations	9
Drainage-Area Ratio Calculations	13
Limitations of Regression Equations	16
Constructing Consistent Flow-Duration Curves	16
StreamStats	18
Summary.....	19
Acknowledgments	20
References Cited.....	20
Appendix 1. Detailed Streamgauge Information Used to Conduct Analyses.....	23

Figures

1. Map showing hydrologic regions in Minnesota and streamgages in study area	4
2. Graph showing flow-duration curve developed using 13 exceedance-probability quantiles for the South Fork Crow River near Mayer, Minnesota, for water years 1941–79	5
3. Graphs showing relation between observed 7-day annual low-flow values and values estimated using log-Pearson type III, three-parameter lognormal, and the log-generalized extreme value distributions for computing a 7-day annual low-flow frequency graph for the South Fork Crow River near Mayer, Minnesota	7
4. Graph showing relation between drainage-area ratio and the observed 0.5 flow-duration curve exceedance probabilities for the 25 pairs of streamgages listed in table 4	15
5. Graph showing example of how to construct a consistent flow-duration curve from regional regression estimates.....	18

Tables

1. Basin characteristics tested as explanatory variables in the development of low-flow regression equations.....	8
2. Regression equations for estimating selected flow-duration curve exceedance probabilities for unregulated streams in Minnesota	10
3. Regression equations for estimating selected low-flow frequency statistics for unregulated streams in Minnesota.....	12
4. Paired streamgages selected to develop within-basin equations that use drainage-area ratios to estimate streamflow statistics at ungaged locations on streams with a streamgage in another location.....	14
5. Standard errors of selected flow-duration curve and low-flow frequency statistics based on statewide drainage area-ratio equations developed in this study	15
6. Maximum, minimum, and mean values of the basin characteristics used in the regional regression analysis of each region in this study.....	17

Conversion Factors

Inch/Pound to International System of Units

Multiply	By	To obtain
Length		
inch (in.)	2.54	centimeter (cm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
square mile (mi ²)	259.0	hectare (ha)
square mile (mi ²)	2.590	square kilometer (km ²)
Flow rate		
foot per second (ft/s)	0.3048	meter per second (m/s)
cubic foot per second (ft ³ /s)	0.02832	cubic meter per second (m ³ /s)
Hydraulic conductivity		
foot per day (ft/d)	0.3048	meter per day (m/d)
Hydraulic gradient		
foot per mile (ft/mi)	0.1894	meter per kilometer (m/km)

Water year is the 12-month period of October 1 through September 30 and is designated by the calendar year in which it ends.

Climate year is the 12-month period of April 1 through March 31 and is designated by the calendar year in which it ends.

Datums

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.

Abbreviations

BSLDEM10M	mean basin slope, in feet per foot
CROPSNLCD01	percent area of cultivated land
DAR	drainage-area ratio
DRNAREA	drainage area, in square miles
FDC	flow-duration curve
FOREST	percent area of forested land
GIS	geographic information system
LAKEAREA	percent area of lakes
LFF	low-flow frequency
NRCS	Natural Resources Conservation Service
PFLATLOW	percent mean basin low-lying flatland
PFLATUP	percent mean basin upland flatland
PMPE	mean basin precipitation minus potential evapotranspiration, in millimeters
PPCC	probability plot correlation coefficient
SEE	standard error of the estimate
SOILA	percent streamgage drainage basin area consisting of hydrologic soil group A
SOILB	percent streamgage drainage basin area consisting of hydrologic soil group B
SOILC	percent streamgage drainage basin area consisting of hydrologic soil group C
SOILD	percent streamgage drainage basin area consisting of hydrologic soil group D
SSURGCLAY	percent mean basin clay
SSURGO	NRCS Soil Survey Geographic database
SSURGOKSAT	mean basin hydraulic conductivity of saturated soils, in micrometers per second

SSURGOM	percent mean basin organic matter content
SSURGSAND	percent mean basin sand
SSURGSILT	percent mean basin silt
STORAGE	percent streamgage drainage basin area consisting of lakes and wetlands
USGS	U.S. Geological Survey

Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams in Minnesota

By Jeffrey R. Ziegeweid, David L. Lorenz, Chris A. Sanocki, and Christiana R. Czuba

Abstract

Knowledge of the magnitude and frequency of low flows in streams, which are flows in a stream during prolonged dry weather, is fundamental for water-supply planning and design; waste-load allocation; reservoir storage design; and maintenance of water quality and quantity for irrigation, recreation, and wildlife conservation. This report presents the results of a statewide study for which regional regression equations were developed for estimating 13 flow-duration curve statistics and 10 low-flow frequency statistics at ungaged stream locations in Minnesota. The 13 flow-duration curve statistics estimated by regression equations include the 0.0001, 0.001, 0.02, 0.05, 0.1, 0.25, 0.50, 0.75, 0.9, 0.95, 0.99, 0.999, and 0.9999 exceedance-probability quantiles. The low-flow frequency statistics include annual and seasonal (spring, summer, fall, winter) 7-day mean low flows, seasonal 30-day mean low flows, and summer 122-day mean low flows for a recurrence interval of 10 years. Estimates of the 13 flow-duration curve statistics and the 10 low-flow frequency statistics are provided for 196 U.S. Geological Survey continuous-record streamgages using streamflow data collected through September 30, 2012.

The study area includes 196 streamgages located within Minnesota and 50 miles beyond the State's borders in North Dakota, South Dakota, Iowa, and Wisconsin. The study area was divided into five regions that were developed in a previous study using the concept of hydrologic landscape units. Geographic information system software was used to calculate 18 characteristics investigated as potential explanatory variables in regression analyses for each streamgage drainage basin. Trend analyses indicated statistically significant trends in summer 7-day low flows that were not related to precipitation patterns for 19 streamgages. For 16 of these streamgages, the streamflow record was subset using structural change analysis to identify the most recent period of record without a significant trend. The three remaining streamgages with significant trends were excluded from the final analysis because the effective period of record without a significant trend was less than 10 years.

Because several streams in this study have zero flow as their minimum reported flow, weighted left-censored regression was used to analyze the flow data in an unbiased manner, with weights based on the number of years of record. A total of 115 regression equations were developed in this study to calculate flow-duration curve and low-flow frequency statistics for ungaged locations in the study area. In addition, data from 25 pairs of streamgages were used to develop drainage-area ratio equations that can be used to estimate streamflow statistics at ungaged locations on streams that have a streamgage in another location. Streamflow statistics estimated using regional regression and drainage-area ratio equations were compared among regions. For regions A, D, and E, drainage-area ratio equations were more accurate than regional regression equations for flows, but regional regression equations were more accurate for high flows. For region F, regional regression equations were consistently more accurate than drainage-area ratio equations. For region BC, the pattern in accuracies of regional regression and drainage-area ratio equations between low flows and high flows was not consistent.

Equations developed in this study apply only to stream locations where flows are not substantially affected by regulation, diversion, or urbanization. All equations presented in this study will be incorporated into StreamStats, a web-based geographic information system tool developed by the U.S. Geological Survey. StreamStats allows users to obtain streamflow statistics, basin characteristics, and other information for user-selected locations on streams through an interactive map.

Introduction

Low flow can be defined as “the flow of water in a stream during prolonged dry weather” (U.S. Environmental Protection Agency, 2013). Knowledge of the magnitude and frequency of low flows for streams is fundamental for water-supply planning and design; waste-load allocation; reservoir storage design; and maintenance of water quality and quantity for irrigation, recreation, and wildlife conservation (Eash and Barnes, 2012). Because low-flow frequency (LFF) statistics

2 Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams

indicate the probable availability of water in streams during times when conflicts between water supply and demand are most prevalent, LFF statistics can be used by Federal, State, tribal, and local agencies for water-quality regulatory activities, water-supply planning, and water management. For example, LFF statistics can be used as thresholds when setting wastewater-treatment plant effluent limits and allowable pollutant loads to meet water-quality regulations. In addition, LFF statistics can be used by commercial, industrial, and hydroelectric facilities to determine availability of water for water supply, waste discharge, and power generation. Finally, LFF statistics can be used in ecological research because low-flow conditions can disturb ecosystems and create biological responses that include changes in habitat, reduced populations of aquatic species, and shifts in the relative distribution of species (Miller and Golladay, 1996).

In 2014, 1,649 stream reaches in Minnesota were designated as impaired (Minnesota Pollution Control Agency, 2014) and are scheduled to have pollutant loads analyzed and maximum loading rates established by total maximum daily load (TMDL) assessments (U.S. Environmental Protection Agency, 2012). Reliable estimates of expected streamflow are needed for specific periods of the year when determining the maximum allowable load of a pollutant in a stream. Estimates of expected streamflow are especially important for low-flow periods when agencies need to determine waste-load allocations (WLAs) for National Pollution Discharge Elimination System (NPDES; U.S. Environmental Protection Agency, 2012) discharge permits for municipalities, industries, and other entities with facilities that release treated wastewater into a stream. A WLA is the loading capacity or maximum quantity of a pollutant each point-source discharger is allowed to release into a particular stream (U.S. Environmental Protection Agency, 2012), and WLAs are used to establish water-quality-based limits for point-source discharges.

The U.S. Geological Survey (USGS) operates a network of streamgages in Minnesota and throughout the Nation that are used to compute streamflow data for a variety of purposes. Flow-duration curve (FDC) and LFF statistics are calculated using streamflow data collected at these locations; however, streamgages are not available for every stream, and therefore, methods are needed for estimating FDC and LFF statistics at ungaged stream locations. The USGS initiated a statewide study in 2012 in cooperation with the Minnesota Pollution Control Agency (MPCA) to estimate FDC and LFF statistics for ungaged stream locations in Minnesota using streamflow data collected through September 30, 2012. Major components of the study included (1) computing 13 selected FDC statistics and 10 selected LFF statistics at 196 continuous-record streamgages within Minnesota and adjacent States, (2) measuring 18 basin characteristics for each streamgage, (3) developing 115 regional regression equations to estimate 23 selected statistics for ungaged stream locations based on basin characteristics, and (4) developing drainage-area ratio equations to provide estimates of low flow at ungaged locations on gaged streams in Minnesota. The regression equations

developed for this study will be incorporated into StreamStats, a web-based geographic information system (GIS) tool developed by the USGS (Ries and others, 2008).

This study builds upon the work presented in several previous USGS reports (Warne, 1978; Arntson and Lorenz, 1987; Winterstein and others, 2007; Lorenz and others, 2010; Eash and Barnes, 2012). Warne (1978) presented a map in atlas form to demonstrate variability in LFF data for streamgages in Minnesota. Arntson and Lorenz (1987) and Winterstein and others (2007) compiled LFF characteristics for continuous-record streamgages in Minnesota. Lorenz and others (2010) used data from streamgages in Minnesota and adjacent States to develop techniques for estimating the magnitude and frequency of peak flows on small, ungaged streams in Minnesota, and the results of their study were incorporated into StreamStats. Finally, Eash and Barnes (2012) used data from continuous-record streamgages in Iowa and adjacent States to estimate selected LFF statistics and harmonic mean flows for ungaged stream locations in Iowa.

Purpose and Scope

This report presents methods used to develop regional regression equations for use in estimating FDC and LFF statistics at ungaged locations on unregulated streams in Minnesota. The regression equations relate selected FDC and LFF statistics to physical and hydrologic soil characteristics of gaged drainage basins for five defined hydrologic regions. The regression equations can be used to estimate 13 selected FDC statistics and 10 selected LFF statistics for ungaged locations on unregulated streams in Minnesota. The 13 FDC statistics estimated by regression equations include the 0.0001, 0.001, 0.02, 0.05, 0.1, 0.25, 0.50, 0.75, 0.9, 0.95, 0.99, 0.999, and 0.9999 exceedance-probability quantiles. The LFF statistics include annual and seasonal (spring, summer, fall, winter) 7-day mean low flows, seasonal 30-day mean low flows, and summer 122-day mean low flows for a recurrence interval of 10 years.

The FDC and LFF statistics were computed using streamflow data collected through September 30, 2012. Names of LFF statistics used in this report were selected to maintain consistency with names used within StreamStats (<http://water.usgs.gov/osw/streamstats/StatisticsDefinitions.html>). All statistics reported in this study were computed for each streamgage using the longest, most recent period of record without a significant trend in low flow related to factors other than precipitation patterns. The accuracy and limitations of the regression equations and the methodology used to develop the equations are described in the report.

Description of Study Area

The study area (fig. 1) includes the State of Minnesota and a 50-mile (mi) buffer around Minnesota in the neighboring States of Iowa, North Dakota, South Dakota, and Wisconsin.

Canadian portions of the basins were not included because many of the characteristics were not available in the same format as those for U.S. portions of the basins. The study area is divided into five hydrologic regions (fig. 1) developed using a combination of statistical analyses from previous studies (Jacques and Lorenz, 1987; Lorenz and others, 1997; Lorenz and others, 2010) and the concept of hydrologic landscape units (Winter, 2001; Wolock and others, 2004). Region BC represents the combined regions B and C from Lorenz and others (2010) because not enough streamgages with low-flow data were available in region C to develop regional regression equations. Analyses of residuals from regressions confirmed that regions B and C could be combined without affecting the ability of regional regression equations to accurately estimate FDC and LFF statistics. Streamgages included in this study were assigned a map number, and information associated with each map number (streamgage) can be found in table 1–1 of the appendix.

Differences among the five hydrologic regions are addressed using the hydrologic landscape unit information presented in Lorenz and others (2010). Region A has the most heterogeneous landscapes, with generally low slopes that become more moderately sloped near the boundary. Most of region BC is dominated by sandy soils and low to moderate slopes, but the northeastern part of region BC along the north shore of Lake Superior and the northeastern Canadian border generally is high in slope. Low slopes near the center of region D change to moderate slopes and less sandy soils around the regional boundaries. Region E is similar to region D, with moderate slopes and low sand content in the soils; however, there are distinct differences in the drainage patterns between regions D and E (Lorenz and others, 2010). Finally, high slope areas in region F change to moderately sloped areas along the western regional boundary.

Methods for Data Development

Data used in this report were collected for 196 active and inactive continuous-record streamgages located in Minnesota or within 50 miles of Minnesota in the neighboring States of Iowa, North Dakota, South Dakota, and Wisconsin (fig. 1; table 1–1). Selected streamgages met the following criteria: (1) at least 10 complete water years of daily mean streamflows (a water year is the 12-month period from October 1 through September 30 and is designated by the calendar year in which it ends); (2) streamflows unaffected by regulation, diversion, or urbanization; and (3) drainage areas less than 3,000 square miles. For 5 of the 196 streamgages, additional streamflow records available from the Minnesota Department of Natural Resources (MDNR; Minnesota Department of Natural Resources, 2013) were appended to the USGS streamflow data. Streamflow data from MDNR streamgages 01063003 (April 1, 2008, through September 30, 2012), 01092001 (September 1, 2008, through September 30, 2012), 70018001

(September 1, 2007, through September 30, 2012), 41043001 (July 1, 2009, through September 30, 2012), and 40016001 (April 1, 2008, through September 30, 2012) were appended to streamflow data from USGS streamgages 04012500, 04014500, 05095000, 05374900, and 05376800, respectively.

Streamflow data from USGS streamgages in neighboring States were used to improve the representativeness of selected FDC and LFF statistics and to provide better estimates of the error associated with the regression equations for ungaged locations near the Minnesota border. Nationwide, USGS streamflow data are collected using consistent methods and instrumentation specifications to ensure data are comparable (Rantz and others, 1982). Daily mean streamflow data collected through water year 2012 (September 30, 2012) were retrieved for the 196 streamgages from the USGS National Water Information System (U.S. Geological Survey, 2013) and used in computing selected FDC and LFF statistics. All statistical analyses presented in this report were performed using R version 3.1.3 (Venables and others, 2010).

Trend Analyses

Trends in data can introduce bias into the calculation of FDC and LFF statistics because the statistical analysis is based on the assumption that statistical parameters associated with the data do not change over time (National Institute of Standards and Technology, 2013). Such parameters include the percentiles of flow, which are used in FDCs for the distribution of low flows and for determining the 10 selected LFF statistics. For this study, trends were evaluated based on the summer (June through September) 7-day low flows. An initial assessment of trends using the Mann-Kendall trend test (Helsel and Hirsch, 2002) determined that 69 streamgages had significant trends at a significance level of 0.05 (p -values less than 0.05); however, precipitation patterns have a considerable influence on streamflow.

To further evaluate those trends in streamflow, a linear regression model was used to relate the summer 7-day low flow, minimum monthly summer precipitation (June through September); the total antecedent precipitation in the year (January through September); the total precipitation from the previous calendar year; and the year of the observation (to indicate the time trend). Mean monthly precipitation data were available by climate division for the period of January 1895 to December 2012 from the National Climatic Data Center (National Oceanic and Atmospheric Administration, 2013). After accounting for precipitation effects on streamflow, 50 streamgages had significant trends at a significance level of 0.05. To reduce the statistical chance of making a type I error (identifying a trend when it is not present in all of the trend analyses), p -values from the linear regression analyses were adjusted for a false discovery rate (Benjamini and Hochberg, 1995); 19 streamgages had significant trends at a false discovery level of 0.05.

4 Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams

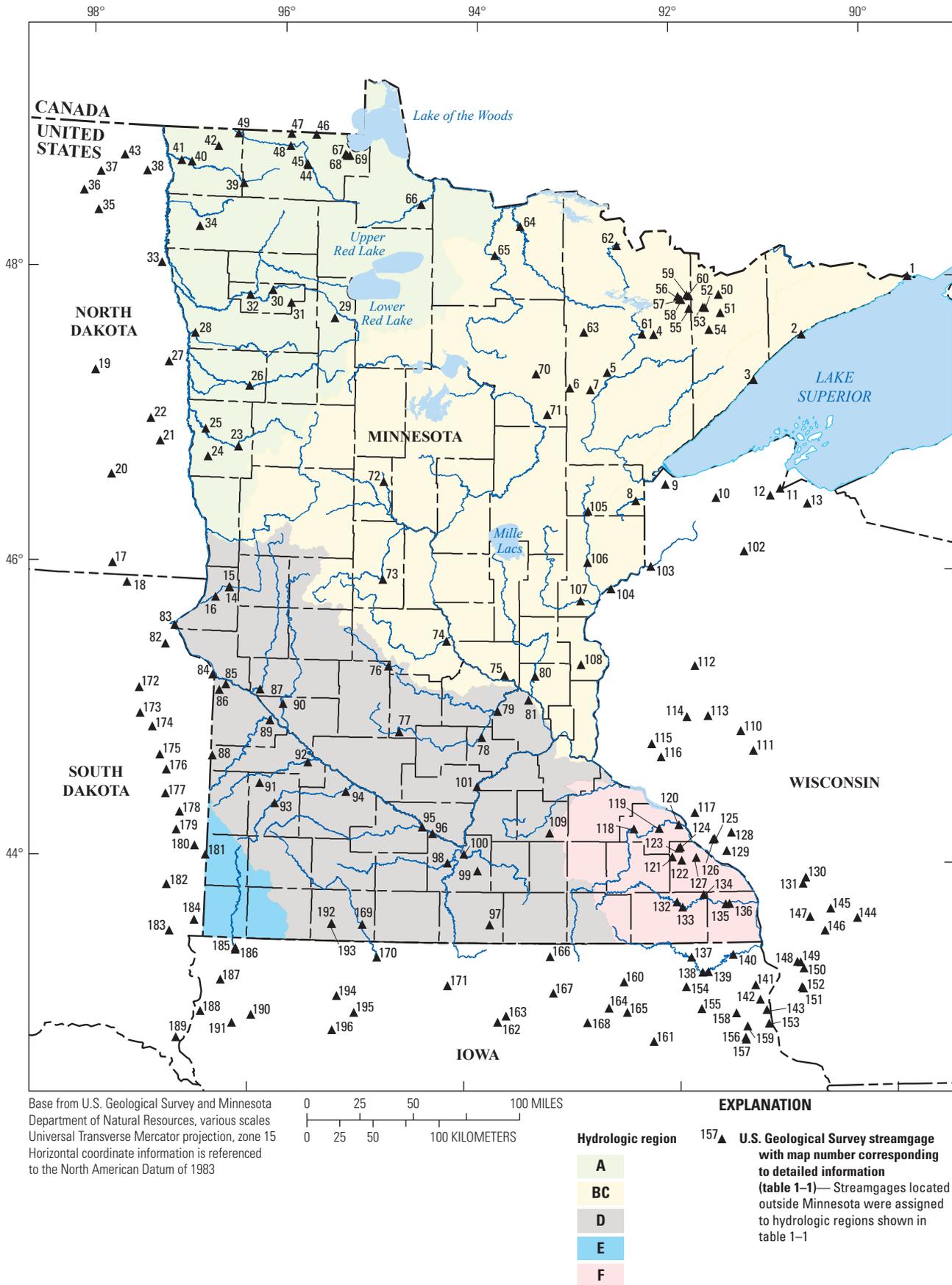


Figure 1. Hydrologic regions in Minnesota and streamgages in study area.

The 19 streamgages with significant trends not related to precipitation patterns could indicate changes within the basin due to human activities, such as changes in cropping patterns, drainage, wetland restoration, and other factors. Such temporal changes would affect the relation between the basin characteristic and the FDC and LFF statistics. For example, previous studies attributed increases in streamflow in many rivers to increasing precipitation beginning in the 1940s (Houghton and others, 1996; Griggs and Noguera, 2002); however, Zhang and Schilling (2006) demonstrated that increases to streamflow and base flow in the Mississippi River basin actually resulted from reduced evapotranspiration and increased groundwater recharge associated with conversion of land from perennial vegetation to seasonal row crops. For 16 of the 19 streamgages, the streamflow record was subset using structural change analysis (Zeileis and others, 2002), which identified changes in the relation between the summer low flows and precipitation patterns, to define the most recent period of record of at least 10 years without a significant trend at a significance level of 0.05 (table 1–1). Three streamgages were excluded from the final analysis because their effective period of record was less than 10 years; the three excluded streamgages were not included in the 196 streamgages listed in table 1–1.

Flow-Duration Curves

For a given streamgage, the FDC was computed from daily mean streamflows from complete water years for the

entire streamflow period of record (Searcy, 1959), except for the 16 streamgages with significant trends identified in the streamflow data not attributable to precipitation patterns and for which a subset streamflow record was used, as described previously in the “Trend Analyses” section. The FDCs were computed by sorting daily-value data and assigning exceedance probabilities to each value by means of the plotting position formula:

$$p_i = (i - a) / (n + 1 - 2a), \tag{1}$$

where

- p_i is the non-exceedance probability,
- i is the rank (1 to n , smallest to largest),
- n is the number of values, and
- a is a constant, taken here as 0.4 (Helsel and Hirsch, 2002).

The 0.0001, 0.001, 0.02, 0.05, 0.1, 0.25, 0.50, 0.75, 0.9, 0.95, 0.99, 0.999, and 0.9999 exceedance-probability quantiles were obtained from the daily streamflow data. The computed FDC quantiles are provided in table 1–1 of the appendix, along with information on the period of record used to compute the quantiles. The FDC for USGS streamgage 05279000 (map number 78; fig. 1), the South Fork Crow River near Mayer, Minnesota, is shown in figure 2 as an example. Figure 2 also illustrates that flow values of zero prevent the FDC from reaching 100 percent exceedance.

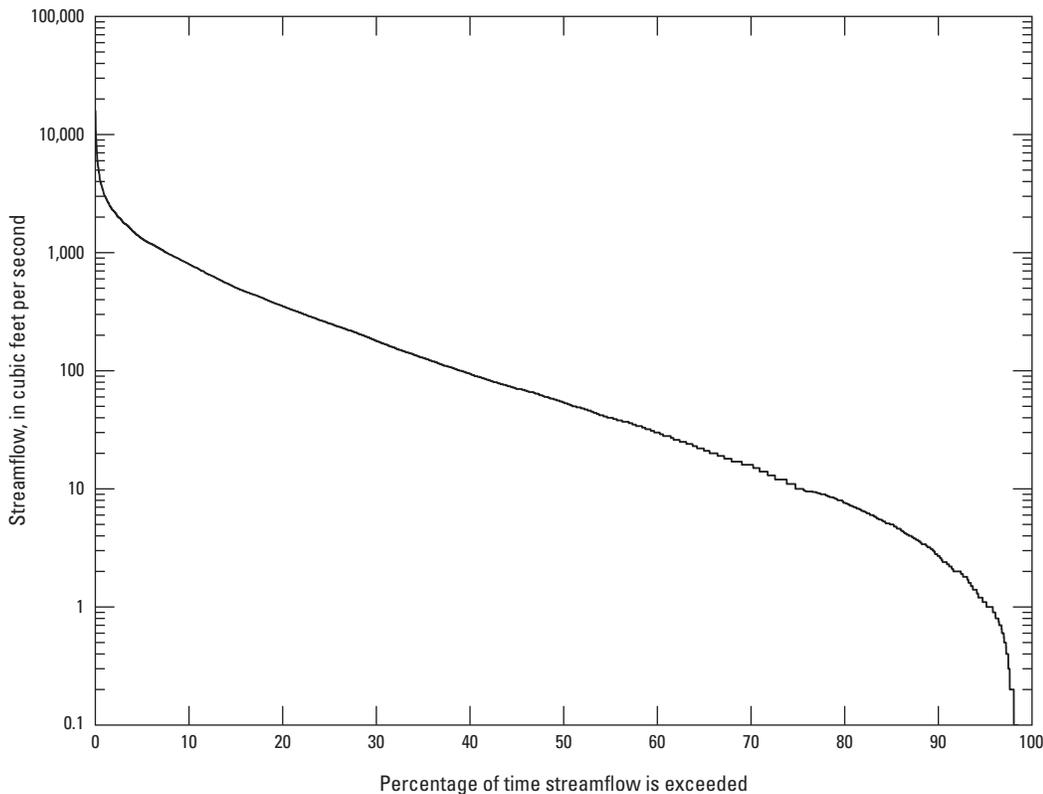


Figure 2. Flow-duration curve developed using 13 exceedance-probability quantiles for the South Fork Crow River near Mayer, Minnesota, (U.S. Geological Survey streamgage 05279000) for water years 1941–79.

Low-Flow Frequency

A frequency analysis was conducted to estimate 10 selected low-flow statistics at the 196 streamgages used in the study. The LFF statistics typically consist of mean low flows for a specified number of days within a season or climate year for a specified recurrence interval. A climate year is defined as the 12-month period from April 1 through March 31 and is designated by the calendar year in which the period ends. Recurrence interval, in years, is the reciprocal of the annual or seasonal non-exceedance probability. The MPCA uses a recurrence interval of 10 years for regulatory protocols (Minnesota Pollution Control Agency, 2008), which was used for all LFF analyses in this study.

Climate years were used to divide annual and seasonal low-flow data because the start of the climate year (April 1) generally corresponds with high flow in the spring, so low-flow conditions do not span across multiple years. Low-flow frequencies were estimated for the annual (climate year) 7-day mean low flows, the seasonal (June through September, October through November, December through March, and April through May) 7- and 30-day mean low flows, and the summer seasonal (June through September) 122-day mean low flows. More simply, 7-day mean low flow with a recurrence interval of 10 years is the lowest 7-day mean flow that occurs (on average) once every 10 years (U.S. Environmental Protection Agency, 2013).

For a given streamgage, the occurrence of each low-flow value was computed over the climate year or season by identifying consecutive days with the lowest 7-, 30-, or 122-day mean streamflows. If there were fewer than six nonzero values, then an empirical fit was used to fit the data. Otherwise, the magnitude and frequency for each of the LFF statistics were computed by fitting each nonzero value to one of three distributions: log-Pearson type III, three-parameter lognormal, or log-generalized extreme value (Vogel and Kroll, 1989; Helsel and Hirsch, 2002). The distribution used to determine LFF statistics at each streamgage is listed in table 1–1 of the appendix. The log-Pearson type III distribution was used unless the data fit the distribution poorly. Fit of the data to the log-Pearson type III distribution was assessed by using the probability plot correlation coefficient (PPCC) from the observed nonzero data and the predicted values (Helsel and Hirsch, 2002); the closer the linear correlation coefficient was to 1.0, the better the data fit the distribution.

In this study, the log-Pearson type III distribution automatically was used when computed linear correlation coefficients were greater than 0.985. Otherwise, PPCCs were computed for all three distributions (log-Pearson type III, three-parameter lognormal, log-generalized extreme value), and the distribution that resulted in the highest PPCC was selected. In addition, a conditional probability adjustment for zero-flow values (appendix 5 in U.S. Interagency Advisory Committee on Water Data, 1982) was used for streamgages with one or more annual 7-, 30-, or 122-day low-flow values

of zero. The 7-day annual LFF graph for USGS streamgage 05279000 (map number 78; fig. 1), South Fork Crow River near Mayer, Minn., is shown in figure 3 to provide a comparison of the three distributions used to compute LFF statistics in this study. The log-Pearson type III analysis was used for the 7-day annual LFF analysis of the 33 nonzero values for streamgage 05279000, and a conditional probability adjustment was included in the analysis for the 5 zero-flow years in the annual 7-day low-flow record.

Basin Characteristics

Low-flow characteristics of streams are related to the physical, geologic, and climatic properties of drainage basins (Smakhtin, 2001). In most basins, drainage area is a statistically significant variable in explaining low-flow variability (Funkhouser and others, 2008; Kroll and others, 2004). Basin characteristics investigated in this study as potential explanatory variables in regression analyses were selected on the basis of their theoretical relation to low flows, results of previous studies in similar hydrologic areas, and the ability to quantify the basin characteristics using GIS technology. The use of GIS enables the automation of the basin-characteristic calculations and solution of the regional regression equations using StreamStats (<http://water.usgs.gov/osw/streamstats/index.html>).

Using GIS technology, 18 basin characteristics (table 1) were calculated for each of the 196 streamgage basins in this study and investigated as potential explanatory variables in regression analyses with 23 selected low-flow statistics as response variables (table 1–1). Basin characteristics generally can be divided into three categories: (1) physical measurements, (2) land cover, and (3) soils. Physical measurement characteristics include drainage area, in square miles (DRNAREA); mean basin slope, in feet per foot (BSLDEM10M); percent streamgage drainage basin area consisting of lakes and wetlands (STORAGE); precipitation minus potential evapotranspiration, in millimeters (PMPE); and percent mean basin low-lying flatland (PFLATLOW) and upland flatland (PFLATUP). Land-cover characteristics include percent areas of lakes (LAKEAREA), forested land (FOREST), and cultivated land (CROPSNLCD01). Soil characteristics include percent mean basin organic matter content (SSURGOM), sand (SSURGSAND), silt (SSURGSILT), clay (SSURGCLAY); mean basin hydraulic conductivity of saturated soils, in micrometers per second (SSURGOKSAT); and percent streamgage drainage basin areas consisting of hydrologic soil groups A (SOILA), B (SOILB), C (SOILC), and D (SOILD). Sources of data were the National Elevation Dataset (<http://ned.usgs.gov/>; Gesch and others, 2009), National Wetlands Inventory (<http://www.fws.gov/wetlands/Data/Data-Download.html>); U.S. Fish and Wildlife Service, 2013), hydrologic landscape units (David M. Wolock, U.S. Geological Survey, unpub.data, 2003), National Land Cover Dataset (<http://www.mrlc.gov/nlcd2006.php>; Homer and others, 2012),

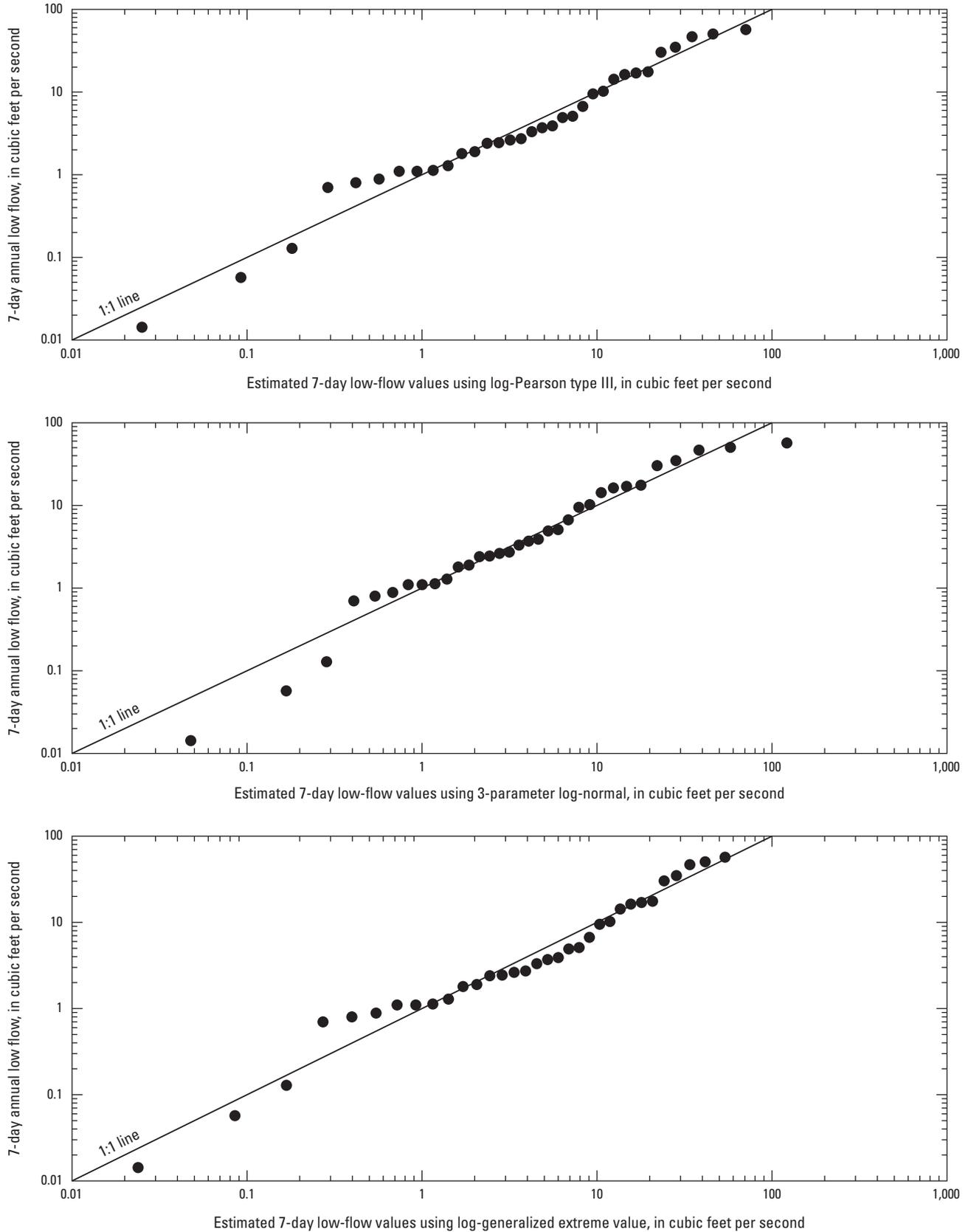


Figure 3. Relation between observed 7-day annual low-flow values and values estimated using log-Pearson type III, three-parameter lognormal, and the log-generalized extreme value distributions for computing a 7-day annual low-flow frequency graph for the South Fork Crow River near Mayer, Minnesota (U.S. Geological Survey streamgage 05279000).

8 Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams

Table 1. Basin characteristics tested as explanatory variables in the development of low-flow regression equations.

[StreamStats and NHDPlus are web-based tools used to measure areas of watersheds. Basin mean is the mean of all raster values in the basin area. Percent drainage area is the area of drainage basin covered by feature, in percent. DRNAREA, drainage area; BSLDEM10M, mean basin slope; STORAGE, area of lakes and wetlands; PMPE, precipitation minus potential evapotranspiration; PFLATLOW, low-lying flatland; PFLATUP, upland flatland; LAKEAREA, area of lakes; FOREST, area of forested land; CROPSNLCD01, area of cultivated land; SSURGOM, organic matter; SSURGSAND, sand; SSURGSILT, silt; SSURGCLAY, clay; SSURGOKSAT, hydraulic conductivity of saturated soils; SOILA, soil in hydrologic group A; SOILB, soil in hydrologic group B; SOILC, soil in hydrologic group C; SOILD, soil in hydrologic group D]

StreamStats name	Units	Method	Source data
Physical measurements			
DRNAREA	square miles	StreamStats NHDPlus	http://water.usgs.gov/osw/streamstats/index.html http://www.horizon-systems.com/NHDPlus/NHDPlusV1_data.php
BSLDEM10M	feet per foot	Basin mean	http://ned.usgs.gov/
STORAGE	percent	Percent drainage area	http://www.fws.gov/wetlands/Data/Data-Download.html
PMPE	millimeters	Basin mean	David M. Wolock, U.S. Geological Survey, unpub.data, 2003
PFLATLOW	percent	Basin mean	David M. Wolock, U.S. Geological Survey, unpub.data, 2003
PFLATUP	percent	Basin mean	David M. Wolock, U.S. Geological Survey, unpub.data, 2003
Land cover			
LAKEAREA	percent	Percent drainage area	http://www.fws.gov/wetlands/Data/Data-Download.html
FOREST	percent	Percent drainage area	http://www.mrlc.gov/nlcd2006.php
CROPSNLCD01	percent	Percent drainage area	http://www.mrlc.gov/nlcd2006.php
Soils			
SSURGOM	percent	Basin mean	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SSURGSAND	percent	Basin mean	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SSURGSILT	percent	Basin mean	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SSURGCLAY	percent	Basin mean	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SSURGOKSAT	micrometers per second	Basin mean	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SOILA	percent	Percent drainage area	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SOILB	percent	Percent drainage area	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SOILC	percent	Percent drainage area	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/
SOILD	percent	Percent drainage area	http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/

Soil Survey Geographic Database (<http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/>; Soil Survey Staff, 2012), StreamStats (<http://water.usgs.gov/osw/streamstats/index.html>; Ries and others, 2008) and the National Hydrography Dataset Plus (http://www.horizon-systems.com/NHDPlus/NHDPlusV1_data.php; Bondelid and others, 2006). Data for the 18 basin characteristics will be incorporated into StreamStats for use in calculating FDC and LFF statistics at ungaged locations on small streams in Minnesota.

Regression Analysis

Because some streams in this study have values of zero for selected FDC or LFF statistics, special regression techniques are required to analyze these statistics in an unbiased manner. Estimates of zero flow computed from observed streamflow commonly are treated as left-censored data (Kroll and Stedinger, 1996; Kroll and Vogel, 2002), and the use of

multiple-linear regression is not recommended for censored data (Helsel and Hirsch, 2002). For this study, weighted left-censored regression was used for all regression analyses, with weights based on the number of years of record (Lorenz, 2014). When no data are censored, weighted left-censored regression provides the same results as weighted least-squares regression (Helsel and Hirsch, 2002). Because of the uncertainty in measuring low flows and estimating FDC and LFF statistics less than 0.1 cubic foot per second (ft³/s), the censoring threshold used to develop the left-censored regression equations was set at 0.1 ft³/s.

Regional regression equations were developed for the five regions shown in figure 1, and residuals from regression analyses were analyzed to ensure that the regional divisions were supported by the hydrologic data collected at the streamgages within the study area. Residuals were required to be approximately normally distributed, homoscedastic (having constant variance throughout the range of fitted values), and

show no regional pattern. Region BC (fig. 1) represents the combined regions B and C from Lorenz and others (2010) because the number of streamgages in region C is too small to develop regional regression equations. A consistent set of explanatory variables was used in all of the regression equations for each region to make consistent estimations at the selected FDC exceedance probabilities or LFF statistics. A set of consistent explanatory variables that had logical coefficients and were statistically significant at the 0.05 level was selected for the regression equations for all FDC exceedance probabilities or LFF statistics within that region. This approach guarantees consistency for uncensored regressions but does not guarantee consistency when censored data are included. In this report, “inconsistency” refers to situations when an estimated low flow for a smaller number of consecutive days is greater than an estimated low flow for a larger number of consecutive days. For example, an estimated 7-day low flow may be larger than the estimated 30-day low flow.

Regional Regression Analysis

In a regional regression study, dividing a large study area into subregions that are relatively homogeneous in terms of hydrology typically helps to reduce error in the regression equations. Hydrologic regions were developed for peak flows in Minnesota by Lorenz and others (2010) using the concept of hydrologic landscape units (Winter, 2001; Wolock and others, 2004) and not peak-flow residuals. For this study, the regions of Lorenz and others (2010) were used as the initial hydrologic regions, but residual analyses based on the regressions were used to verify that the hydrologic regions form reasonably homogenous FDC and LFF regression equations. As previously mentioned, regions B and C were combined.

Flow-Duration Curve Regression Equations

The weighted left-censored regression method (Helsel and Hirsch, 2002) was used to develop regression equations for each exceedance probability for each region in figure 1. One streamgage (05412100, Roberts Creek above Saint Olaf, Iowa) was excluded from the analysis for development of regression equations in region F because it was an extreme outlier, but regression results for the gage are provided for completeness (table 1–1). The regression equations and corresponding standard errors of the estimate (SEEs) are listed in table 2. The number of explanatory variables was minimized in regression equations to maximize residual degrees of freedom and minimize multicollinearity, which is the condition where at least one explanatory variable is closely related to one or more other explanatory variables (Helsel and Hirsch, 2002).

Nine of the 18 examined basin characteristics were used as explanatory variables in the development of FDC regional regression equations (table 2). Explanatory variables used in

regression equations for region A were DRNAREA, SSUR-GOM, and BSLDEM10M. Explanatory variables used in regression equations for region BC were DRNAREA, STORAGE, and SOILC. Explanatory variables used in regression equations for region D were DRNAREA, PMPE, and PFLATLOW. Explanatory variables used in regression equations for region E were DRNAREA and CROPSNLCD01. Finally, explanatory variables used in regression equations for region F were DRNAREA, FOREST, and PFLATLOW. For each region, selected explanatory variables minimized SEEs and produced consistent FDCs. Generally, SEEs (expressed as percentage-based values) were largest for the largest exceedance probabilities and generally increased with increased censored data.

Low-Flow Frequency Regression Equations

The weighted left-censored regression method was used to develop regression equations for each LFF statistic for each region shown in figure 1. One streamgage (05412100, Roberts Creek above Saint Olaf, Iowa; map number 159; fig. 1) was excluded from the analysis for development of regression equations in region F because it was an extreme outlier, but regression results for the gage are provided for completeness (table 1–1). The regression equations and corresponding SEEs are listed in table 3. The number of explanatory variables were minimized in regression equations to maximize residual degrees of freedom and minimize multicollinearity (Helsel and Hirsch, 2002).

Eight of the 18 examined basin characteristics were used as explanatory variables in the development of LFF regional regression equations (table 3). Explanatory variables used in regression equations for region A were DRNAREA, SSUR-GOM, and BSLDEM10M. Explanatory variables used in regression equations for region BC were DRNAREA, STORAGE, and SOILC. Explanatory variables used in regression equations for region D were DRNAREA, PMPE, and PFLATLOW. Explanatory variables used in regression equations for region E were DRNAREA and CROPSNLCD01. Finally, explanatory variables used in regression equations for region F were DRNAREA, CROPSNLCD01, and BSLDEM10M. For each region, explanatory variables were selected to minimize SEEs, produce consistent LFF statistics, and minimize multicollinearity (Helsel and Hirsch, 2002). Ranges of SEEs varied by region, with regions BC and F having smaller SEEs compared to regions A, D, and E. Among all regions, SEEs ranged from 23 to 665 percent. Fewer censored values occurred in regions BC and F, and SEEs generally increased as the number of censored values increased among regions.

The number of censored values in each analysis varied by region. Annually, across regions A, BC, D, and E, most of the censored values occurred in the winter (December through March) and the summer (June through September) seasonal regression analyses. Region F had the fewest number of censored flow values; many of the streams in region F

10 Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams

Table 2. Regression equations for estimating selected flow-duration curve exceedance probabilities for unregulated streams in Minnesota.

[SEE, standard error of the estimate; number censored, number of streamflow values less than or equal to 0.1 cubic foot per second; D0_01, streamflow with an exceedance probability of 0.0001; DRNAREA, drainage area, in square miles; SSURGOM, percentage organic matter estimated using the Soil Survey Geographic Database; BSLDEM10M, mean basin slope computed from a 10-meter digital elevation model; D0_1, streamflow with an exceedance probability of 0.001; D2, streamflow with an exceedance probability of 0.02; D5, streamflow with an exceedance probability of 0.05; D10, streamflow with an exceedance probability of 0.10; D25, streamflow with an exceedance probability of 0.25; D50, streamflow with an exceedance probability of 0.50; D75, streamflow with an exceedance probability of 0.75; D90, streamflow with an exceedance probability of 0.90; D95, streamflow with an exceedance probability of 0.95; D99, streamflow with an exceedance probability of 0.99; D99_9, streamflow with an exceedance probability of 0.999; D99_99, streamflow with an exceedance probability of 0.9999; STORAGE, the percentage of area of water storage in lakes, ponds, reservoirs, and wetlands; SOILC, percentage of area of hydrologic soil type C; PMPE, precipitation minus potential evapotranspiration, in millimeters; PFLATLOW, percentage of low-lying flatland; CROPSNLCD01, percentage of area of cultivated land; FOREST, percentage of area covered by forest]

Regression equation	SEE (percent)	Number censored
Region A (35 streamgages)		
$D0_01 = 118.4 (DRNAREA)^{0.6475} (SSURGOM + 1)^{-0.08101} (BSLDEM10M)^{-0.2437}$	51	0
$D0_1 = 47.16 (DRNAREA)^{0.7086} (SSURGOM + 1)^{-0.08814} (BSLDEM10M)^{-0.1047}$	30	0
$D2 = 2.881 (DRNAREA)^{0.8819} (SSURGOM + 1)^{0.2393} (BSLDEM10M)^{-0.08488}$	32	0
$D5 = 0.5466 (DRNAREA)^{0.9751} (SSURGOM + 1)^{0.4639} (BSLDEM10M)^{0.04280}$	45	0
$D10 = 0.1105 (DRNAREA)^{1.047} (SSURGOM + 1)^{0.7049} (BSLDEM10M)^{0.2260}$	58	0
$D25 = 0.006626 (DRNAREA)^{1.180} (SSURGOM + 1)^{1.040} (BSLDEM10M)^{0.6323}$	71	2
$D50 = 0.0007381 (DRNAREA)^{1.203} (SSURGOM + 1)^{1.279} (BSLDEM10M)^{1.234}$	79	6
$D75 = (6.929 \times 10^{-5}) (DRNAREA)^{1.299} (SSURGOM + 1)^{1.553} (BSLDEM10M)^{1.716}$	95	9
$D90 = (2.892 \times 10^{-6}) (DRNAREA)^{1.436} (SSURGOM + 1)^{2.146} (BSLDEM10M)^{2.305}$	150	13
$D95 = (6.658 \times 10^{-7}) (DRNAREA)^{1.435} (SSURGOM + 1)^{2.470} (BSLDEM10M)^{2.765}$	238	17
$D99 = (1.256 \times 10^{-7}) (DRNAREA)^{1.363} (SSURGOM + 1)^{2.557} (BSLDEM10M)^{3.735}$	604	22
$D99_9 = (2.835 \times 10^{-8}) (DRNAREA)^{1.599} (SSURGOM + 1)^{0.9545} (BSLDEM10M)^{4.882}$	675	28
$D99_99 = (4.183 \times 10^{-7}) (DRNAREA)^{1.143} (SSURGOM + 1)^{1.187} (BSLDEM10M)^{4.296}$	420	28
Region BC (49 streamgages)		
$D0_01 = 37.88 (DRNAREA)^{0.7700} (STORAGE + 1)^{-0.1406} (SOILC + 1)^{0.2069}$	65	0
$D0_1 = 16.58 (DRNAREA)^{0.7986} (STORAGE + 1)^{-0.06526} (SOILC + 1)^{0.2098}$	50	0
$D2 = 4.068 (DRNAREA)^{0.8549} (STORAGE + 1)^{0.02530} (SOILC + 1)^{0.1813}$	39	0
$D5 = 2.458 (DRNAREA)^{0.8785} (STORAGE + 1)^{0.06252} (SOILC + 1)^{0.1305}$	35	0
$D10 = 1.227 (DRNAREA)^{0.9333} (STORAGE + 1)^{0.03388} (SOILC + 1)^{0.1166}$	27	0
$D25 = 1.673 (DRNAREA)^{0.9455} (STORAGE + 1)^{-0.1367} (SOILC + 1)^{-0.02305}$	30	0
$D50 = 2.436 (DRNAREA)^{1.001} (STORAGE + 1)^{-0.4477} (SOILC + 1)^{-0.1631}$	38	0
$D75 = 3.903 (DRNAREA)^{1.021} (STORAGE + 1)^{-0.6862} (SOILC + 1)^{-0.2812}$	54	0
$D90 = 4.759 (DRNAREA)^{1.061} (STORAGE + 1)^{-0.8323} (SOILC + 1)^{-0.4041}$	68	1
$D95 = 5.226 (DRNAREA)^{1.113} (STORAGE + 1)^{-0.9446} (SOILC + 1)^{-0.5080}$	80	3
$D99 = 7.486 (DRNAREA)^{1.147} (STORAGE + 1)^{-1.133} (SOILC + 1)^{-0.6666}$	112	7
$D99_9 = 8.96 (DRNAREA)^{1.231} (STORAGE + 1)^{-1.303} (SOILC + 1)^{-0.9244}$	217	8
$D99_99 = 9.861 (DRNAREA)^{1.218} (STORAGE + 1)^{-1.271} (SOILC + 1)^{-1.037}$	261	8
Region D (47 streamgages)		
$D0_01 = 513.1 (DRNAREA)^{0.7935} (1.003)^{PMPE} (PFLATLOW + 1)^{-0.7232}$	43	0
$D0_1 = 210.9 (DRNAREA)^{0.8329} (1.003)^{PMPE} (PFLATLOW + 1)^{-0.7058}$	38	0
$D2 = 18.45 (DRNAREA)^{0.9443} (1.005)^{PMPE} (PFLATLOW + 1)^{-0.5820}$	35	0
$D5 = 6.698 (DRNAREA)^{0.9972} (1.006)^{PMPE} (PFLATLOW + 1)^{-0.5603}$	41	0

Table 2. Regression equations for estimating selected flow-duration curve exceedance probabilities for unregulated streams in Minnesota.—Continued

Regression equation	SEE (percent)	Number censored
Region D (47 streamgages)—Continued		
D10 = 3.279 (DRNAREA) ^{1.042} (1.007) ^{PMPE} (PFLATLOW + 1) ^{-0.6111}	51	0
D25 = 2.171 (DRNAREA) ^{1.109} (1.009) ^{PMPE} (PFLATLOW + 1) ^{-0.9597}	91	1
D50 = 1.650 (DRNAREA) ^{1.215} (1.011) ^{PMPE} (PFLATLOW + 1) ^{-1.500}	108	3
D75 = 0.5481 (DRNAREA) ^{1.404} (1.013) ^{PMPE} (PFLATLOW + 1) ^{-1.955}	157	9
D90 = 0.4237 (DRNAREA) ^{1.387} (1.014) ^{PMPE} (PFLATLOW + 1) ^{-2.135}	182	13
D95 = 0.6143 (DRNAREA) ^{1.260} (1.016) ^{PMPE} (PFLATLOW + 1) ^{-2.216}	255	14
D99 = 0.09255 (DRNAREA) ^{1.215} (1.021) ^{PMPE} (PFLATLOW + 1) ^{-1.983}	497	19
D99_9 = 0.02581 (DRNAREA) ^{1.578} (1.023) ^{PMPE} (PFLATLOW + 1) ^{-2.750}	1,248	25
D99_99 = 0.04036 (DRNAREA) ^{1.512} (1.023) ^{PMPE} (PFLATLOW + 1) ^{-2.831}	1,301	26
Region E (20 streamgages)		
D0_01 = 2.034 (DRNAREA) ^{0.6332} (CROPSNLCD01 + 1) ^{0.9775}	56	0
D0_1 = 0.4470 (DRNAREA) ^{0.6875} (CROPSNLCD01 + 1) ^{1.113}	40	0
D2 = 0.03186 (DRNAREA) ^{0.9674} (CROPSNLCD01 + 1) ^{0.9967}	39	0
D5 = 0.01140 (DRNAREA) ^{1.069} (CROPSNLCD01 + 1) ^{0.9321}	53	0
D10 = 0.002035 (DRNAREA) ^{1.146} (CROPSNLCD01 + 1) ^{1.090}	65	0
D25 = (3.855 x 10 ⁻⁵) (DRNAREA) ^{1.226} (CROPSNLCD01 + 1) ^{1.659}	84	0
D50 = (1.941 x 10 ⁻⁷) (DRNAREA) ^{1.340} (CROPSNLCD01 + 1) ^{2.449}	107	0
D75 = (1.078 x 10 ⁻¹⁰) (DRNAREA) ^{1.557} (CROPSNLCD01 + 1) ^{3.578}	187	3
D90 = (1.707 x 10 ⁻¹⁹) (DRNAREA) ^{1.635} (CROPSNLCD01 + 1) ^{7.858}	104	6
D95 = (1.070 x 10 ⁻²²) (DRNAREA) ^{1.665} (CROPSNLCD01 + 1) ^{9.343}	102	8
D99 = (6.127 x 10 ⁻³⁰) (DRNAREA) ^{1.999} (CROPSNLCD01 + 1) ^{12.33}	128	11
D99_9 = (1.368 x 10 ⁻³³) (DRNAREA) ^{2.364} (CROPSNLCD01 + 1) ^{13.35}	366	14
D99_99 = (3.975 x 10 ⁻³⁸) (DRNAREA) ^{3.312} (CROPSNLCD01 + 1) ^{14.12}	747	15
Region F (44 of 45 streamgages)		
D0_01 = 255.5 (DRNAREA) ^{0.9331} (FOREST + 1) ^{-0.5002} (PFLATLOW + 1) ^{-0.3321}	31	0
D0_1 = 68.29 (DRNAREA) ^{0.9256} (FOREST + 1) ^{-0.4113} (PFLATLOW + 1) ^{-0.05605}	28	0
D2 = 4.081 (DRNAREA) ^{1.102} (FOREST + 1) ^{-0.2792} (PFLATLOW + 1) ^{-0.1492}	33	0
D5 = 1.313 (DRNAREA) ^{1.122} (FOREST + 1) ^{-0.1566} (PFLATLOW + 1) ^{-0.08057}	34	0
D10 = 0.7140 (DRNAREA) ^{1.106} (FOREST + 1) ^{-0.06831} (PFLATLOW + 1) ^{-0.0492}	33	0
D25 = 0.2575 (DRNAREA) ^{1.073} (FOREST + 1) ^{0.1421} (PFLATLOW + 1) ^{0.004680}	31	0
D50 = 0.09218 (DRNAREA) ^{1.036} (FOREST + 1) ^{0.3885} (PFLATLOW + 1) ^{0.06342}	30	0
D75 = 0.03823 (DRNAREA) ^{1.018} (FOREST + 1) ^{0.5786} (PFLATLOW + 1) ^{0.1189}	30	0
D90 = 0.01424 (DRNAREA) ^{1.018} (FOREST + 1) ^{0.7888} (PFLATLOW + 1) ^{0.1919}	27	0
D95 = 0.008009 (DRNAREA) ^{1.030} (FOREST + 1) ^{0.8961} (PFLATLOW + 1) ^{0.2253}	28	0
D99 = 0.004000 (DRNAREA) ^{1.040} (FOREST + 1) ^{1.022} (PFLATLOW + 1) ^{0.2531}	32	1
D99_9 = 0.002113 (DRNAREA) ^{1.028} (FOREST + 1) ^{1.151} (PFLATLOW + 1) ^{0.2955}	42	1
D99_99 = 0.002002 (DRNAREA) ^{0.9861} (FOREST + 1) ^{1.152} (PFLATLOW + 1) ^{0.3661}	55	1

12 Methods for Estimating Flow-Duration Curve and Low-Flow Frequency Statistics for Ungaged Locations on Small Streams

Table 3. Regression equations for estimating selected low-flow frequency statistics for unregulated streams in Minnesota.

[SEE, standard error of the estimate; number censored, number of streamflow values less than or equal to 0.1 cubic feet per second; M7D10Y, annual 7-day low flow for a recurrence interval of 10 years; DRNAREA, drainage area, in square miles; SSURGOM, percentage organic matter estimated using the Soil Survey Geographic Database; BSLDEM10M, mean basin slope computed from a 10-meter digital elevation model; M7D10Y_FAL, fall (October and November) 7-day low flow for a recurrence interval of 10 years; M30D10Y1011, fall 30-day low flow for a recurrence interval of 10 years; M7D10Y1203, winter (December through March) 7-day low flow for a recurrence interval of 10 years; M30D10Y123, winter 30-day low flow for a recurrence interval of 10 years; M7D10Y0405, spring (April and May) 7-day low flow for a recurrence interval of 10 years; M30D10Y45, spring 30-day low flow for a recurrence interval of 10 years; M7D10Y0609, summer (June through September) 7-day low flow for a recurrence interval of 10 years; M30D10Y69, summer 30-day low flow for a recurrence interval of 10 years; M122D10Y69, summer 122-day low flow for a recurrence interval of 10 years; STORAGE, percentage of area of water storage in lakes, ponds, reservoirs, and wetlands; SOILC, percentage of area of hydrologic soil type C; PMPE, precipitation minus potential evapotranspiration, in millimeters; PFLATLOW, percentage of low-lying flatland; CROPSNLCD01, percentage of area of cultivated land]

Regression equation	SEE (percent)	Number censored
Region A (35 streamgages)		
M7D10Y = (1.897 x 10 ⁻⁸) (DRNAREA) ^{1.684} (SSURGOM + 1) ^{2.546} (BSLDEM10M) ^{3.664}	665	22
M7D10Y_FAL = (8.463 x 10 ⁻⁶) (DRNAREA) ^{1.222} (SSURGOM + 1)2.065 (BSLDEM10M) ^{2.472}	147	15
M30D10Y1011 = (1.596 x 10 ⁻⁵) (DRNAREA) ^{1.275} (SSURGOM + 1) ^{1.93} (BSLDEM10M) ^{2.157}	109	13
M7D10Y1203 = (3.166 x 10 ⁻⁷) (DRNAREA) ^{1.449} (SSURGOM + 1) ^{2.536} (BSLDEM10M) ^{3.004}	245	19
M30D10Y123 = (3.968 x 10 ⁻⁷) (DRNAREA) ^{1.457} (SSURGOM + 1) ^{2.508} (BSLDEM10M) ^{2.917}	229	18
M7D10Y0405 = (5.596 x 10 ⁻⁵) (DRNAREA) ^{1.33} (SSURGOM + 1) ^{1.592} (BSLDEM10M) ^{1.776}	125	10
M30D10Y45 = (1.536 x 10 ⁻³) (DRNAREA) ^{1.162} (SSURGOM + 1) ^{1.259} (BSLDEM10M) ^{0.9878}	80	5
M7D10Y0609 = (2.390 x 10 ⁻⁷) (DRNAREA) ^{1.461} (SSURGOM + 1) ^{2.372} (BSLDEM10M) ^{3.188}	336	19
M30D10Y69 = (2.907 x 10 ⁻⁶) (DRNAREA) ^{1.399} (SSURGOM + 1) ^{1.828} (BSLDEM10M) ^{2.485}	171	14
M122D10Y69 = (3.847 x 10 ⁻⁴) (DRNAREA) ^{1.200} (SSURGOM + 1) ^{1.409} (BSLDEM10M) ^{1.187}	88	6
Region BC (49 streamgages)		
M7D10Y = (5.491) (DRNAREA) ^{1.152} (STORAGE + 1) ^{-1.106} (SOILC + 1) ^{-0.5761}	95	5
M7D10Y_FAL = (4.823) (DRNAREA) ^{1.066} (STORAGE + 1) ^{-0.7602} (SOILC + 1) ^{-0.5446}	80	2
M30D10Y1011 = (4.979) (DRNAREA) ^{1.027} (STORAGE + 1) ^{-0.6926} (SOILC + 1) ^{-0.4697}	71	1
M7D10Y1203 = (5.295) (DRNAREA) ^{1.13} (STORAGE + 1) ^{-1.018} (SOILC + 1) ^{-0.5363}	93	5
M30D10Y123 = (6.096) (DRNAREA) ^{1.101} (STORAGE + 1) ^{-1.001} (SOILC + 1) ^{-0.508}	88	4
M7D10Y0405 = (3.741) (DRNAREA) ^{1.031} (STORAGE + 1) ^{-0.6304} (SOILC + 1) ^{-0.3486}	57	0
M30D10Y45 = (1.763) (DRNAREA) ^{0.9214} (STORAGE + 1) ^{-0.1277} (SOILC + 1) ^{-0.1699}	63	0
M7D10Y0609 = (4.391) (DRNAREA) ^{1.118} (STORAGE + 1) ^{-0.9835} (SOILC + 1) ^{-0.5133}	86	5
M30D10Y69 = (4.493) (DRNAREA) ^{1.093} (STORAGE + 1) ^{-0.9214} (SOILC + 1) ^{-0.4692}	80	4
M122D10Y69 = (2.648) (DRNAREA) ^{0.8943} (STORAGE + 1) ^{-0.4049} (SOILC + 1) ^{-0.1264}	62	0
Region D (47 streamgages)		
M7D10Y = (2.161 x 10 ⁻³) (DRNAREA) ^{1.212} (1.022) ^{PMPE} (0.901315) ^{PFLATLOW}	498	19
M7D10Y_FAL = (1.263 x 10 ⁻²) (DRNAREA) ^{1.155} (1.016) ^{PMPE} (0.9029) ^{PFLATLOW}	242	13
M30D10Y1011 = (1.401 x 10 ⁻²) (DRNAREA) ^{1.193} (1.016) ^{PMPE} (0.9026) ^{PFLATLOW}	267	13
M7D10Y1203 = (7.681 x 10 ⁻³) (DRNAREA) ^{1.142} (1.017) ^{PMPE} (0.8970) ^{PFLATLOW}	329	17
M30D10Y123 = (1.077 x 10 ⁻²) (DRNAREA) ^{1.14} (1.017) ^{PMPE} (0.8953) ^{PFLATLOW}	289	15
M7D10Y0405 = (6.765 x 10 ⁻²) (DRNAREA) ^{1.177} (1.011) ^{PMPE} (0.9107) ^{PFLATLOW}	101	5
M30D10Y45 = (2.117 x 10 ⁻¹) (DRNAREA) ^{1.087} (1.011) ^{PMPE} (0.9125) ^{PFLATLOW}	98	4
M7D10Y0609 = (7.028 x 10 ⁻³) (DRNAREA) ^{1.162} (1.020) ^{PMPE} (0.8948) ^{PFLATLOW}	402	16
M30D10Y69 = (7.587 x 10 ⁻³) (DRNAREA) ^{1.237} (1.018) ^{PMPE} (0.8989) ^{PFLATLOW}	246	14
M122D10Y69 = (6.237 x 10 ⁻²) (DRNAREA) ^{1.132} (1.014) ^{PMPE} (0.9238) ^{PFLATLOW}	93	4

Table 3. Regression equations for estimating selected low-flow frequency statistics for unregulated streams in Minnesota.—
Continued

Regression equation	SEE (percent)	Number censored
Region E (20 streamgages)		
M7D10Y = (3.866 x 10 ⁻²⁵) (DRNAREA) ^{1.973} (CROPSNLCD01 + 1) ^{9.971}	115	11
M7D10Y_FAL= (2.036 x 10 ⁻²¹) (DRNAREA) ^{1.681} (CROPSNLCD01 + 1) ^{8.717}	110	7
M30D10Y1011 = (1.040 x 10 ⁻¹⁷) (DRNAREA) ^{1.631} (CROPSNLCD01 + 1) ^{6.940}	120	6
M7D10Y1203 = (3.350 x 10 ⁻²⁸) (DRNAREA) ^{1.978} (CROPSNLCD01 + 1) ^{11.54}	92	10
M30D10Y123 = (1.039 x 10 ⁻²⁷) (DRNAREA) ^{1.961} (CROPSNLCD01 + 1) ^{11.38}	89	10
M7D10Y0405 = (4.178 x 10 ⁻¹⁰) (DRNAREA) ^{1.616} (CROPSNLCD01 + 1) ^{3.149}	285	5
M30D10Y45 = (8.254 x 10 ⁻⁶) (DRNAREA) ^{1.403} (CROPSNLCD01 + 1) ^{1.346}	158	1
M7D10Y0609 = (4.648 x 10 ⁻²³) (DRNAREA) ^{1.612} (CROPSNLCD01 + 1) ^{9.548}	172	8
M30D10Y69 = (1.859 x 10 ⁻²²) (DRNAREA) ^{1.592} (CROPSNLCD01 + 1) ^{9.401}	122	7
M122D10Y69 = (1.844 x 10 ⁻¹⁰) (DRNAREA) ^{1.487} (CROPSNLCD01 + 1) ^{3.619}	182	2
Region F (44 of 45 streamgages)		
M7D10Y = (0.2560) (DRNAREA) ^{1.084} (0.9663) ^{CROPSNLCD01} (BSLDEM10M) ^{0.4765}	36	1
M7D10Y_FAL = (0.3163) (DRNAREA) ^{1.069} (0.9707) ^{CROPSNLCD01} (BSLDEM10M) ^{0.4283}	33	0
M30D10Y1011 = (0.3469) (DRNAREA) ^{1.068} (0.9712) ^{CROPSNLCD01} (BSLDEM10M) ^{0.4129}	30	0
M7D10Y1203 = (0.2940) (DRNAREA) ^{1.083} (0.9660) ^{CROPSNLCD01} (BSLDEM10M) ^{0.4536}	34	1
M30D10Y123 = (0.3243) (DRNAREA) ^{1.084} (0.9670) ^{CROPSNLCD01} (BSLDEM10M) ^{0.4410}	32	1
M7D10Y0405 = (0.4147) (DRNAREA) ^{1.079} (0.9751) ^{CROPSNLCD01} (BSLDEM10M) ^{0.2612}	30	0
M30D10Y45 = (0.6311) (DRNAREA) ^{1.082} (0.9750) ^{CROPSNLCD01} (BSLDEM10M) ^{0.1574}	30	0
M7D10Y0609 = (0.3130) (DRNAREA) ^{1.077} (.9710) ^{CROPSNLCD01} (BSLDEM10M) ^{0.3580}	37	0
M30D10Y69 = (0.3060) (DRNAREA) ^{1.061} (0.9742) ^{CROPSNLCD01} (BSLDEM10M) ^{0.3691}	31	0
M122D10Y69 = (0.4153) (DRNAREA) ^{1.065} (0.9798) ^{CROPSNLCD01} (BSLDEM10M) ^{0.2417}	23	0

are spring-fed as a result of the underlying karst topography (Ruhl, 1989), and the springs likely sustain streamflows during low-flow periods and reduce ice thicknesses in the winter time. Also, region F generally has higher-gradient streams than regions A, D, and E (Gesch and others, 2009), which corresponds to higher BSLDEM10M values (table 1–1). Region BC is characterized by high-gradient streams in the northeastern part of the region (Gesch and others, 2009), lake-fed streams in the central and northern parts of the region, and spring-fed streams in the northwestern part of the region. In contrast, regions A, D, and E generally consist of low-gradient streams (Gesch and others, 2009) that develop thick ice in the winter. Ice formation shrinks the effective depth of stream channels and reduces winter streamflows (Melcher and Walker, 1992). Summer flows in regions A, D, and E generally are low because potential evapotranspiration exceeds precipitation in most years, reducing recharge from groundwater (Delin and others, 2007; Lorenz and Delin, 2007). This hypothesis is further supported by small or negative PMPE values for regions A, D, and E (table 1–1).

Drainage-Area Ratio Calculations

For some ungaged locations on streams that have a streamgage at another location, a more accurate estimation of FDC or LFF statistics can be made based on ratios of the drainage areas between ungaged and gaged locations than from regional regression equations, provided landscapes did not change drastically within examined watersheds. For all FDC and LFF statistics computed in this study, the direct application of the drainage-area ratio (DAR) provides a good estimate of the value at the ungaged location for DARs ranging from 0.25 to 4. Of the 196 streamgages used in this statewide analysis, there were 25 pairs of streamgages from the same stream or tributary that were used to develop a DAR equation (table 4). The selected streamgages produced DARs greater than 0.25 and less than 1 (table 4).

For the 25 selected pairs of streamgages, the relation between DARs and ratios of corresponding values for FDC and LFF statistics was compared for selected FDC exceedance probabilities (0.02, 0.10, 0.50, 0.90, and 0.99) and LFF

Table 4. Paired streamgages selected to develop within-basin equations that use drainage-area ratios to estimate streamflow statistics at ungaged locations on streams with a streamgage in another location.

[Basin name, the river basin containing the paired streamgages; hydrologic region, the hydrologic region of the basin; upstream, the farthest upstream of the paired streamgages; downstream, the farthest downstream of the paired streamgages; DAR, drainage-area ratio calculated by dividing the drainage area at the upstream streamgage by the drainage area at the downstream streamgage]

Basin name	Hydrologic region (fig. 1)	Upstream	Downstream	DAR
Mustinka River	D	05047500	05049000	0.2536
Maple River	A	05059700	05060000	0.5737
Buffalo River	A	05061000	05062000	0.3403
Buffalo River	A	05061500	05062000	0.4673
Clearwater River	A	05078000	05078500	0.4078
Roseau River	A	05104500	05107500	0.4246
Roseau River	A	05107500	05112000	0.7140
Crow River	D	05279000	05280000	0.4363
Redwood River	D	05315000	05316500	0.4150
St. Croix River	BC	05333500	05336000	0.5219
Eau Galle River	F	05369945	05370500	0.5234
Zumbro River	F	05374000	05374500	0.8590
Zumbro River	F	05374500	05374900	0.9417
Whitewater River	F	05376000	05376800	0.3701
Whitewater River	F	05376800	05377500	0.9416
Root River	F	05384000	05385000	0.4927
Root River	F	05385000	05386000	0.8096
Upper Iowa River	F	05387500	05388250	0.6655
Kickapoo River	F	05407470	05408000	0.4407
Kickapoo River	F	05410000	05410490	0.8970
Turkey River	F	05411850	05412020	0.7079
Wapsipinicon River	D	05420560	05420680	0.2784
Des Moines River	D	05476000	05476500	0.8859
Skunk Creek	E	06481480	06481500	0.4226
Little Sioux River	D	06605600	06605850	0.8642

statistics (annual and seasonal 7-day mean low flows for a recurrence interval of 10 years; table 5). The relation was assessed visually and statistically using nonlinear least squares (Bates and Watts, 1988). None of the nonlinear least-squares results indicated that the coefficient for the DAR was significantly different from 1, resulting in the general DAR equation:

$$Stat_u = Stat_g * (DRNAREA_u/DRNAREA_g) \quad (2)$$

where

- $Stat_u$ is the selected FDC or LFF statistic for the ungaged stream location,
- $Stat_g$ is the selected FDC or LFF statistic for the gaged stream location,

- $DRNAREA_u$ is the drainage area of the ungaged stream location, and
- $DRNAREA_g$ is the drainage area of the gaged stream location.

Figure 4 shows the relation between the DARs and the observed 0.5 FDC exceedance-probability ratios for the 25 pairs of streamgages as an example, as well as a line that represents a 1:1 relation between DAR and 0.5 FDC exceedance-probability ratio. The SEEs for selected FDC and LFF statistics calculated using DAR equations for DARs ranging from 0.25 to 1 are reported in table 5. The relation between DARs and 0.5 exceedance-probability ratios for DARs ranging from 1 to 4 can be obtained by plotting the reciprocals of DARs and reciprocals of the associated 0.5 exceedance-probability ratios presented in figure 4.

Table 5. Standard errors of selected flow-duration curve and low-flow frequency statistics based on statewide drainage area-ratio equations developed in this study.

[SEE, standard error of the estimate; D2, streamflow with an exceedance probability of 0.02; D10, streamflow with an exceedance probability of 0.10; D50, streamflow with an exceedance probability of 0.50; D90, streamflow with an exceedance probability of 0.90; D99, streamflow with an exceedance probability of 0.99; M7D10Y, annual 7-day low flow for a recurrence interval of 10 years; M7D10Y_FAL, fall (October and November) 7-day low flow for a recurrence interval of 10 years; M7D10Y1203, winter (December through March) 7-day low flow for a recurrence interval of 10 years; M7D10Y0405, spring (April and May) 7-day low flow for a recurrence interval of 10 years; M7D10Y0609, summer (June through September) 7-day low flow for a recurrence interval of 10 years]

Statistic	SEE
D2	62
D10	63
D50	65
D90	69
D99	113
M7D10Y	79
M7D10Y_FAL	72
M7D10Y1203	73
M7D10Y0405	66
M7D10Y0609	79

The SEEs of the selected flow statistics calculated using DAR equations (table 5) were compared to SEEs of the same flow statistics calculated using the regional regression equations (tables 2 and 3). For regions A, D, and E, comparisons of the SEEs (tables 2 and 5) indicate that regional regression equations may be more accurate than DAR equations for higher flows with exceedance probabilities less than or equal to 0.10. However, DAR equations seem to be more accurate than regional regression equations for lower flows with exceedance probabilities of greater than or equal to 0.50 (tables 2 and 5). Furthermore, in regions A, D, and E, SEEs consistently were lower using DAR equations for annual and seasonal 7-day LFF statistics (tables 3 and 5).

Comparisons of SEEs for region BC were less consistent than comparisons for regions A, D, and E. For FDC statistics, SEEs were lower for regional regression equations than for DAR equations for exceedance probabilities of 0.02, 0.10, and 0.50. Standard errors for exceedance probabilities of 0.90 and 0.99 were within 1 percent but still slightly lower using regional regression equations (tables 2 and 5). In contrast, SEEs for annual and seasonal 7-day LFF statistics generally were lower using DAR equations than using regional regression equations (tables 3 and 5), with the exception of the spring (April and May) 7-day low flow. For the spring 7-day low flow, the SEE using the DAR equation was 66 percent (table 5), but the SEE using the regional regression equation was 57 percent (table 3).

For streamgages in region F, SEEs of FDC and LFF statistics indicate that regional regression equations consistently were more accurate than DAR equations. The observed differences between region F and other regions in the study may be related to the karst topography of region F (Ruhl, 1989). In karst topography, low flows could be affected substantially by gains from large springs and by losses from sinkholes (Eash and Barnes, 2012).

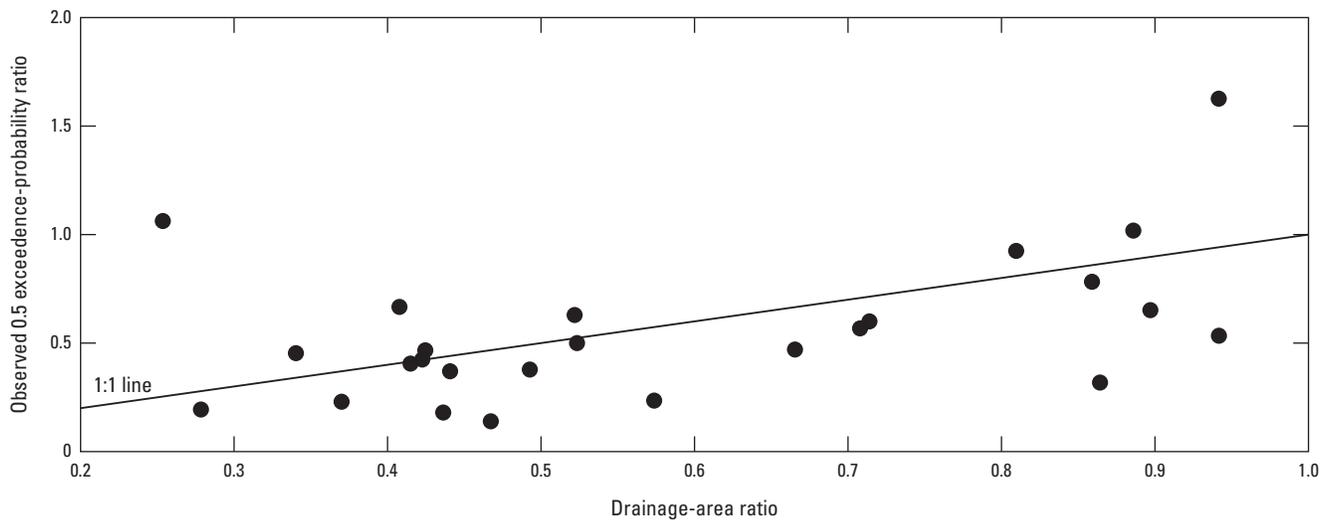


Figure 4. Relation between drainage-area ratio and the observed 0.5 flow-duration curve exceedance probabilities for the 25 pairs of streamgages listed in table 4.

Limitations of Regression Equations

Several limitations warrant consideration when using the regression equations presented in this report. The regional regression equations apply only to stream locations in Minnesota where flows are not substantially affected by regulation, diversion, or urbanization. The applicability and accuracy of the regional equations depend on whether the basin characteristics calculated for an ungaged stream location are within the range of the values for variables used to develop the regression equations. The acceptable range of basin characteristic values used to develop each regional regression equation are tabulated as maximum and minimum values in table 6. The mean is included to represent the most typical values for the range of the data; for PMPE, the mean is the arithmetic mean, and for all other basin characteristics, the mean represents geometric mean (table 6). Where the minimum value is 0, the geometric mean is adjusted by adding 1 to each value, computing the geometric mean of the adjusted values and subtracting 1 to produce the adjusted geometric mean. The applicability of the regional equations is unknown when any characteristic value calculated for an ungaged location is outside the acceptable range. In addition, basin-characteristic calculations at ungaged locations should be computed using the same GIS datasets and calculation methods used in this study. The USGS StreamStats Web-based GIS tool (<http://water.usgs.gov/osw/streamstats/index.html>; Ries and others, 2008) includes the same GIS data layers and calculation methods as used to develop the regression equations in this study.

The regression equations presented in this report should be used with caution for ungaged stream locations with basin-characteristic values approaching the minimum or maximum values (table 6) because of the potential for inconsistencies in the estimates. Although the occurrence of inconsistencies were minimized by using the same set of explanatory variables for each regional set of LFF equations, inconsistencies still may occur. Inconsistencies may occur because regional regression equations were developed separately and have variable estimation intervals depending on the size and variability of the datasets used to develop regression equations. If inconsistencies are obtained for an ungaged stream location, a comparison of all LFF estimates for the location and a check of streamgage data or other published data may help determine which LFF statistic is inconsistent (Eash and Barnes, 2012).

Several factors affect the estimation accuracy of the presented regression equations. Estimation accuracy depends on the sample size, the accuracy of each recorded streamflow, and how well the chosen distribution fits the actual distribution of the data (Lorenz and others, 2010). The accuracies of regression estimates are affected by errors in explanatory variables, and systematic errors in the computation of the response variable can bias estimates. Although streamgages used to develop the regression were weighted based on the number of years of record, the period of record also can bias estimates of response variables because of patterns in precipitation, flow conditions, other weather-related factors, and land use changes. Finally,

estimating flow statistics at a location immediately downstream from a lake or wetland also could affect the accuracy of estimates made using the presented regression equations.

The DAR equations to estimate streamflow statistics at ungaged locations on streams that have a streamgage were developed using paired streamgage data from all regions presented in this study (fig. 1; table 4). However, each region was not equally represented in the DAR equations. Regions BC and E each were represented only by one pair of streamgages. Regions A and D each were represented by 6 pairs of streamgages, and region F was represented by 11 pairs of streamgages. Region E is the smallest region in Minnesota, and region BC contains a small number of suitable streamgages because of a lack of streamgages along the high-gradient north shore of Lake Superior, extensive connected wetlands in the northern part of the region, and rivers affected by regulation or diversion in the southwestern part of the region.

The regression equations presented in this report also should be used with caution in areas where flows are affected by substantial gains as a result of large springs or substantial losses as a result of sinkholes common to karst topography. Region F, in southeastern Minnesota (fig. 1), contains karst areas where flows may vary considerably because of gaining or losing stream reaches (Eash and Barnes, 2012). User judgement may be required to decide if an ungaged location in a karst area may be affected by substantial gains or losses in flow, and regression estimates should be compared against streamgage data or other published data. If the equations are used at ungaged locations on regulated streams, or on streams affected by water-supply and agricultural withdrawals, then the estimates will need to be adjusted by the amount of regulation or withdrawal to estimate the actual flow conditions if that is of interest.

Special attention must be given to censored values and the number of significant figures used. Because of the uncertainty in measuring and estimating flows less than 0.1 ft³/s, the censoring threshold used to develop the left-censored regression equations was set at 0.1 ft³/s. Thus, any regression estimates that are 0.1 ft³/s or less should be reported as less than 0.1 ft³/s. Because the precision of response- and explanatory-variable data used to develop the equations commonly was limited to three significant figures, selected-statistic streamflows estimated from the regression equations also should be limited to three significant figures.

Constructing Consistent Flow-Duration Curves

Flow-duration curves are constructed from point estimates of streamflow for 13 exceedance probabilities using the regression equations developed for this report. The regression equations were developed to produce reasonably consistent FDCs, which means the magnitude of the predicted value

Table 6. Maximum, minimum, and mean values of the basin characteristics used in the regional regression analysis of each region in this study.

[DRNAREA, drainage area; BSLDEM10M, mean basin slope; STORAGE, area of lakes and wetlands; PMPE, precipitation minus potential evapotranspiration; PFLATLOW, low-lying flatland; PFLATUP, upland flatland; LAKEAREA, area of lakes; FOREST, area of forested land; CROPSNLCDD01, area of cultivated land; SSURGOM, organic matter; SSURGSAND, sand; SSURGSILT, silt; SSURGCLAY, clay; SSURGOKSAT, hydraulic conductivity of saturated soils; SOILA, soil in hydrologic group A; SOILB, soil in hydrologic group B; SOILC, soil in hydrologic group C; SOILD, soil in hydrologic group D]

Statistic	DRNAREA	BSLDEM10M	STORAGE	PMPE	PFLATLOW	PFLATUP	LAKEAREA	FOREST	CROPSNLCDD01	SSURGOM	SSURGSAND	SSURGSILT	SSURGCLAY	SSURGOKSAT	SOILA	SOILB	SOILC	SOILD
Region A																		
Maximum	1,440	5.99	95.3	38.7	89.0	62.7	6.02	44.0	92.4	52.2	67.4	77.2	31.2	66.7	98.9	68.9	66.9	7.95
Mean	221	1.55	12.4	-45.7	36.0	29.1	0.46	5.62	46.9	2.29	42.7	30.9	16.6	26.3	21.3	24.9	22.1	1.29
Minimum	11.2	0.46	0.52	-123	5.99	4.00	0	0.18	1.34	0.29	13.8	15.8	7.16	8.61	2.01	0.25	0.49	0
Region BC																		
Maximum	2,960	10.7	48.0	308	40.8	28.4	21.4	76.5	74.5	19.8	76.4	41.0	36.5	72.5	75.1	63.0	100	43.5
Mean	177	4.15	22.0	184	16.6	5.49	2.91	40.7	4.93	5.86	47.5	21.7	9.33	31.5	16.8	14.8	28.7	3.07
Minimum	2.98	1.61	2.44	43.7	1.44	0.49	0	8.41	0	0	15.6	5.30	3.65	8.34	0	0	2.61	0
Region D																		
Maximum	2,640	9.38	17.4	203	56.9	40.7	6.40	7.76	90.0	3.05	60.4	46.5	30.6	48.9	32.2	82.8	77.1	17.0
Mean	327	2.56	4.60	42.0	26.3	17.7	1.13	1.43	72.3	0.81	36.7	34.4	23.6	16.4	2.78	44.7	38.9	0.93
Minimum	7.65	1.18	0.09	-132	4.88	1.01	0	0.22	3.46	0.32	20.8	22.8	13.9	5.17	0.10	18.6	11.1	0
Region E																		
Maximum	2,400	4.76	10.5	34.4	24.3	23.8	3.81	1.27	89.5	1.33	47.1	64.3	28.9	114	2.79	97.0	99.0	10.8
Mean	237	2.96	1.62	-18.8	13.2	11.0	0.42	0.51	72.2	0.58	25.0	43.0	24.5	23.9	0	33.1	31.3	1.06
Minimum	49.0	2.08	0.02	-82.7	2.61	5.76	0	0.30	41.1	0.34	6.87	30.0	20.7	6.78	0	0.43	2.21	0
Region F																		
Maximum	1,540	21.1	5.80	222	15.2	12.9	0.24	48.9	89.0	1.43	74.0	63.0	28.5	60.6	42.3	99.1	78.5	28.8
Mean	159	9.78	1.05	188	2.46	2.48	0.03	18.2	60.4	0.52	28.6	38.2	18.7	19.2	2.08	54.3	18.1	1.72
Minimum	2.28	3.41	0.01	129	0	0	0	0.69	20.8	0.29	4.67	15.4	7.37	4.87	0	8.76	0.50	0

[†]For PMPE, the mean is the arithmetic mean, but for all other basin characteristics, the mean represents the geometric mean.

decreases with increasing exceedance probability. However, the regression equations for the largest exceedance probabilities are less stable than the equations for the smaller exceedance probabilities because regression equations for the largest FDC exceedance probabilities were developed using censored data. Decreased stability resulting from the use of censored data can result in inconsistent FDCs.

Inconsistencies can be removed from FDCs by forcing succeeding magnitudes of increasing exceedance probability to be less than or equal to the previous value. The process of removing inconsistencies in regression-equation estimates of FDC exceedance probabilities is illustrated in figure 5. In the example (fig. 5), the predicted flow for the 0.999 exceedance probability (0.143 ft³/s) is greater than the predicted value for the 0.99 exceedance probability (0.110 ft³/s), but the predicted flow for the 0.9999 exceedance probability (0.061 ft³/s) is less than the predicted flow for the 0.99 exceedance probability (0.110 ft³/s). Log-q-normal interpolation is used to compute a consistent flow value for the 0.999 exceedance probability. First, the logarithms of the flows for the 0.99 and 0.9999 exceedance probabilities are computed. Second, the value for the logarithm of the flow for the 0.999 exceedance probability is linearly interpolated along the normal quantiles for the 0.99, 0.999, and 0.9999 exceedance probabilities. The normal quantiles of the 0.99, 0.999, and 0.9999 exceedance probabilities in this example are 2.326, 3.090, and 3.719, respectively. Back-transformation of the interpolated logarithm results in an estimated flow of 0.079 ft³/s for the 0.999 exceedance probability. However, if the flow for the 0.9999

exceedance probability also had been larger than the flow for the 0.99 exceedance probability, then the flows for the 0.999 and 0.9999 exceedance probabilities would have been set to the flow for the 0.99 exceedance probability (0.110 ft³/s).

StreamStats

StreamStats is a USGS Web-based GIS tool (<http://water.usgs.gov/osw/streamstats/index.html>; Ries and others, 2008) that allows users to obtain streamflow statistics, basin characteristics, and other information for user-selected locations on streams. Users can select stream locations of interest from an interactive map and can obtain information for these locations. If a user selects the location of a USGS streamgage, the user will receive previously published information for the streamgage from a database. If a stream location is selected where no data are available (an ungaged location), a GIS program will estimate information for the location. The GIS program determines the boundary of the drainage basin upstream from the stream location, calculates the basin characteristics of the drainage basin, and solves the appropriate regression equations to estimate streamflow statistics for that location. The results are presented in a table and a map showing the basin-boundary outline. The estimates are applicable for stream locations not substantially affected by regulation, diversions, or urbanization. In the past, it could take an experienced person more than a day to estimate this information at an ungaged location. StreamStats reduces the effort to only a few minutes.

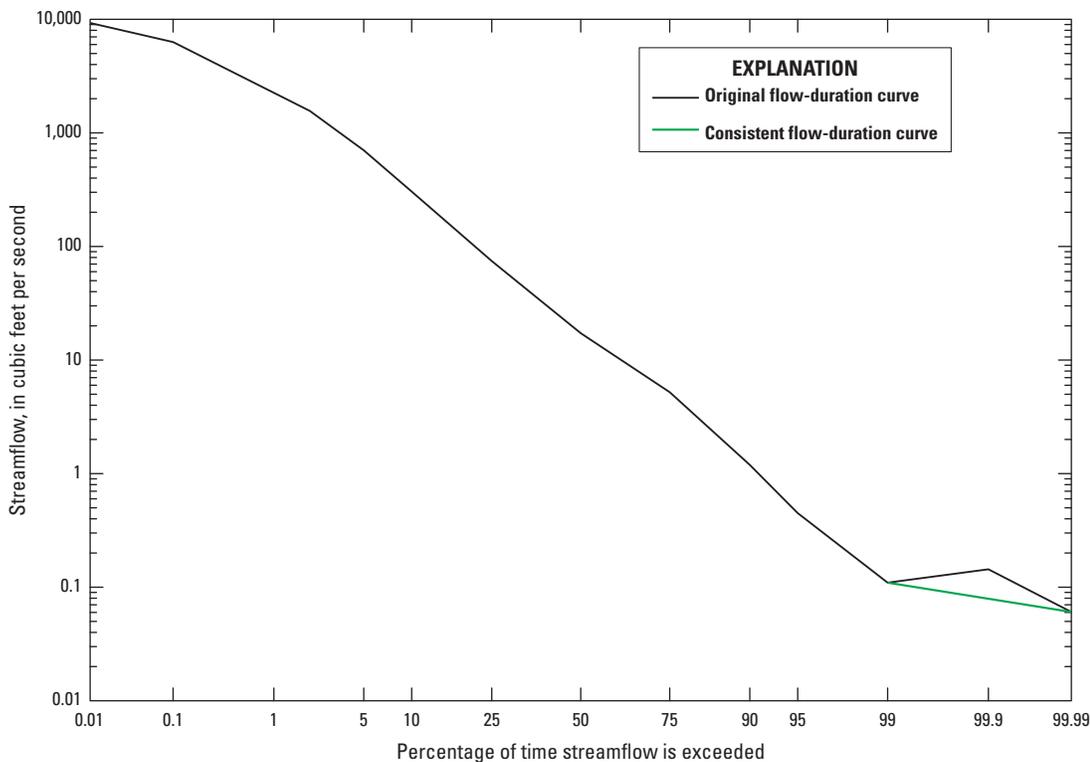


Figure 5. Example of how to construct a consistent flow-duration curve from regional regression estimates.

StreamStats makes the process of computing streamflow statistics for ungaged locations much faster, more accurate, and more consistent than previously used manual methods. Examples of streamflow statistics that can be provided by StreamStats include the 1-percent flood probability, the median annual flow, and the mean 7-day, 10-year low flow. Examples of basin characteristics include the drainage area, stream slope, mean annual precipitation, and percent of area underlain by hydrologic soil groups. Basin characteristics provided by StreamStats are the physical, geologic, and climatic properties that have been statistically related to the movement of water through a drainage basin to a stream location.

Use of the regression equations developed in this study to estimate FDC and LFF statistics will be facilitated through incorporation into the USGS StreamStats Web-based GIS tool (<http://water.usgs.gov/osw/streamstats/index.html>). Streamflow statistics can be estimated at any location along a stream and can assist with the following: (1) water-resources planning, management, and permitting; (2) design and permitting of facilities such as wastewater-treatment plants and water-supply reservoirs; and (3) design of structures such as roads, bridges, culverts, dams, locks, and levees. In addition, planners, regulators, and resource managers often need to know the physical, geologic, and climatic properties (basin characteristics) of the drainage basins upstream from locations of interest to help them understand the processes that control water availability and water quality at these locations.

Summary

Knowledge of the magnitude and frequency of low flows in streams, which are flows in a stream during prolonged dry weather, is fundamental for water-supply planning and design; waste-load allocation; reservoir storage design; and maintenance of water quality and quantity for irrigation, recreation, and wildlife conservation. The U.S. Geological Survey initiated a statewide study in 2012 in cooperation with the Minnesota Pollution Control Agency to estimate selected flow-duration curve and low-flow frequency statistics for stream locations in Minnesota.

Regional regression equations were developed to estimate 13 flow-duration curve statistics and 10 low-flow frequency statistics at ungaged stream locations in Minnesota. The 13 flow-duration curve statistics estimated by regression equations include the 0.0001, 0.001, 0.02, 0.05, 0.1, 0.25, 0.50, 0.75, 0.9, 0.95, 0.99, 0.999, and 0.9999 exceedance-probability quantiles. The low-flow frequency statistics include annual and seasonal (spring, summer, fall, winter) 7-day mean low flows, seasonal 30-day mean low flows, and summer 122-day mean low flows for a recurrence interval of 10 years. Estimates of the 13 flow-duration curve statistics and the 10 low-flow frequency statistics are provided for 196 U.S. Geological Survey continuous-record streamgages using streamflow data collected through September 30, 2012.

The study area includes 196 streamgages located within Minnesota and 50 miles beyond the Minnesota border in the states of North Dakota, South Dakota, Iowa, and Wisconsin. The study area was divided into five regions that were developed in a previous study using the concept of hydrologic landscape units. Geographic information system software was used to calculate 18 selected basin characteristics for each streamgage, which were investigated as potential explanatory variables in regression analyses. Trend analyses indicated statistically significant trends in summer 7-day low flows that were not related to precipitation patterns for 19 streamgages. For 16 of these streamgages, the streamflow record was subset using structural change analysis to identify the most recent period of record without a significant trend. The three remaining streamgages with significant trends were excluded from the final analysis because the effective period of record without a significant trend was less than 10 years.

Because several streamgages in this study have minimum reported flows of zero, weighted left-censored regression was used to analyze the flow data in an unbiased manner, with weights based on the number of years of record. Flow-duration curve and low-flow frequency estimates calculated from censored regression equations that are 0.1 cubic foot per second or less should be reported as less than 0.1 cubic foot per second to maintain a consistent prediction-streamflow-reporting limit for Minnesota. Geographic information system software is required to calculate the basin characteristics included as explanatory variables in the regression equations. A total of 115 regression equations were developed in this study to estimate 23 selected flow-duration curve and low-flow frequency statistics for ungaged locations in the study area.

In addition to regional regression equations, data from 25 pairs of streamgages were used to develop drainage-area ratio equations that can be used to estimate streamflow statistics at ungaged locations on streams or tributaries that have a streamgage in another location. Standard errors of selected streamflow statistics determined using drainage-area ratio and regional regression equations were compared among regions. For regions A, D, and E, drainage-area ratio equations were more accurate than regional regression equations for flows, but regional regression equations were more accurate for high flows. For region F, regional regression equations were consistently more accurate than drainage-area ratio equations. For region BC, there was not a consistent pattern in accuracies of regional regression and drainage-area ratio equations between low flows and high flows. All regional regression and drainage-area ratio equations presented in this study will be incorporated into StreamStats, a web-based geographic information systems tool developed by the U.S. Geological Survey. StreamStats allows users to obtain streamflow statistics, basin characteristics, and other information for user-selected locations on streams through an interactive map.

Several limitations warrant consideration when applying the results presented in this study. The regression equations developed in this study apply only to stream locations in Minnesota where flows are not substantially affected by

regulation, diversion, or urbanization. Furthermore, the regression equations developed in this study are not intended for use at ungaged stream locations in which the basin characteristics are outside the range of those used to develop the equations. Inconsistencies in estimates may result for flow-duration curve and low-flow frequency equation estimates if basin-characteristic values approach the minimum or maximum values of the range. The regression equations should be used with caution in areas where low flows are affected by substantial gains as a result of large springs or substantial losses as a result of sinkholes common to karst topography. If the equations are used at ungaged locations on regulated streams, or on streams affected by water-supply and agricultural withdrawals, then the estimates will need to be adjusted if actual flow conditions are of interest.

Acknowledgments

David Bender and Kristen O'Connor of the U.S. Geological Survey (USGS) South Dakota Water Science Center developed the initial regional regression equations. Erich Kessler and Sarah Elliott of the USGS Minnesota Water Science Center helped conduct statistical analyses. Scott Bennett, a contract employee at the USGS Minnesota Water Science Center, created the study area map and compiled the basin characteristics. David Eash of the USGS Iowa Water Science Center and James Fallon of the USGS Minnesota Water Science Center provided technical reviews.

References Cited

- Arntson, A.D., and Lorenz, D.L., 1987, Low-flow-frequency characteristics for continuous-record streamflow stations in Minnesota: U.S. Geological Survey Water-Resources Investigations Report 88-4353, 360 p. [Also available at <http://pubs.er.usgs.gov/publication/wri864353/>.]
- Bates, D.M., and Watts, D.G., 1998, Nonlinear regression analysis and its applications: New York, Wiley, 90 p.
- Benjamini, Yoav, and Hochberg, Yosef, 1995, Controlling the false discovery rate—A practical and powerful approach to multiple testing: *Journal of the Royal Statistical Society Series B* 57, p. 289–300. [Also available at <http://www.jstor.org/stable/2346101>.]
- Bondelid, T., Johnston, C., McKay, C., Moore, R., and Rea, A., 2006, NHDPlus User Guide: Horizons Systems Corporation, 115 p. [Also available at http://www.horizon-systems.com/NHDPlus/NHDPlusV1_documentation.php.]
- Delin, G.N., Healy, R.W., Lorenz, D.L., and Nimmo, J.R., 2007, Comparison of local- to regional-scale estimates of ground-water recharge in Minnesota, USA: *Journal of Hydrology*, v. 334, p. 231–249. [Also available at <http://dx.doi.org/10.1016/j.jhydrol.2006.10.010>.]
- Eash, D.A., and Barnes, K.K., 2012, Methods for estimating selected low-flow frequency statistics and harmonic mean flows for streams in Iowa: U.S. Geological Survey Scientific Investigations Report 2012-5171, 99 p. [Also available at <http://pubs.er.usgs.gov/publication/sir20125171/>.]
- Funkhouser, J.E., Eng, K., and Moix, M.W., 2008, Low-flow characteristics and regionalization of low-flow characteristics for selected streams in Arkansas: U.S. Geological Survey Scientific Investigations Report 2008-5065, 161 p. [Also available at <http://pubs.usgs.gov/sir/2008/5065/>.]
- Gesch, Dean, Evans, G., Mauck, J., Hutchinson, J., Carswell, W.J., Jr., 2009, The National Map—Elevation: U.S. Geological Survey Fact Sheet 2009-3053, 4 p. [Also available at <http://pubs.usgs.gov/fs/2009/3053/>.]
- Griggs, D.J., and Noguer, M., 2002, Climate change 2001—The scientific basis, Contribution of Working Group I to the third assessment report of the intergovernmental panel on climate change: *Weather*, v. 57, no. 8, p. 267–269. [Also available at <http://dx.doi.org/10.1256/004316502320517344>.]
- Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 524 p. [Also available at <http://pubs.er.usgs.gov/publication/twri04A3/>.]
- Homer, C.H., Fry, J.A., and Barnes, C.A., 2012, The National Land Cover Database: U.S. Geological Survey Fact Sheet 2012-3020, 4 p. [Also available at <http://pubs.usgs.gov/fs/2012/3020/>.]
- Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A., and Maskell, K., eds., 1996, Climate change 1995—The science of climate change: New York, Cambridge University Press, 572 p.
- Jacques, J.E., and Lorenz, D.L., 1987, Techniques for estimating the magnitude and frequency of floods in Minnesota: U.S. Geological Survey Water-Resources Investigations Report 87-4170, 48 p. [Also available at <http://pubs.er.usgs.gov/publication/wri874170/>.]
- Kroll, C.N., Luz, J., Allen, B., and Vogel, R.M., 2004, Developing a watershed characteristics database to improve low streamflow prediction: *Journal of Hydrologic Engineering*, v. 9, no. 2, p. 116–125. [Also available at [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2004\)9:2\(116\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2004)9:2(116)).]

- Kroll, C.N., and Stedinger, J.R., 1996, Estimation of moments and quantiles using censored data: *Water Resources Research*, v. 32, no. 4, p. 1005–1012. [Also available at <http://dx.doi.org/10.1029/95WR03294>.]
- Kroll, C.N., and Vogel, R.M., 2002, Probability distribution of low streamflow series in the United States: *Journal of Hydrologic Engineering*, v. 7, no. 2, p. 137–146. [Also available at [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2002\)7:2\(137\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2002)7:2(137)).]
- Lorenz, D.L., 2014, smwrQW—R functions to support water-quality data analysis for statistical methods in water resources: Geological Survey R Archive Network. [Also available at <https://github.com/USGS-R/smwrQW>.]
- Lorenz, D.L., Carlson, G.H., and Sanocki, C.A., 1997, Techniques for estimating peak flow on small streams in Minnesota: U.S. Geological Survey Water-Resources Investigations Report 97–4249, 44 p. [Also available at <http://pubs.er.usgs.gov/publication/wri974249/>.]
- Lorenz, D.L., and Delin, G.N., 2007, A regression model to estimate regional ground water recharge: *Ground Water*, v. 45, no. 2, p. 196–208. [Also available at <http://dx.doi.org/10.1111/j.1745-6584.2006.00273.x>.]
- Lorenz, D.L., Sanocki, C.A., and Kocian, M.J., 2010, Techniques for estimating the magnitude and frequency of peak flows on small streams in Minnesota based on data through water year 2005: U.S. Geological Survey Scientific Investigations Report 2009–5250, 54 p. [Also available at <http://pubs.er.usgs.gov/publication/sir20095250/>.]
- Melcher, N.B., and Walker, J.F., 1992, Evaluation of selected methods for determining streamflow during periods of ice effect: U.S. Geological Survey Water Supply Paper 2378, 47 p. [Also available at <http://pubs.er.usgs.gov/publication/wsp2378/>.]
- Miller, A.M., and Golladay, S.W., 1996, Effects of spates and drying on macroinvertebrate assemblages of an intermittent and a perennial stream: *Journal of North American Benthological Society*, v. 15, no. 4, p. 670–689, accessed June 7, 2011, at <http://dx.doi.org/10.2307/1467815>.
- Minnesota Department of Natural Resources, 2013, DNR/MPCA cooperative stream gaging: accessed April 3, 2013, at <http://www.dnr.state.mn.us/waters/csg/index.html>.
- Minnesota Pollution Control Agency, 2008, Dissolved oxygen TMDL protocols and submittal requirements: accessed July 9, 2015, at <http://www.pca.state.mn.us/index.php/view-document.html?gid=8529mn.us/index.php/view-document.html?gid=8529>.
- Minnesota Pollution Control Agency, 2014, Minnesota’s impaired waters list: accessed April 13, 2015, at <http://www.pca.state.mn.us/index.php/water/water-types-and-programs/minnesotas-impaired-waters-and-tmdls/impaired-waters-list.html>.
- National Institute of Standards and Technology, 2013, NIST/SEMATECH e-Handbook of Statistical Methods: accessed July 2015 at <http://www.itl.nist.gov/div898/handbook/pmc/section4/pmc442.htm>.
- National Oceanic and Atmospheric Administration, 2013, National Climatic Data Center division-averaged climate data time series: accessed February 5, 2013, at <http://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp>.
- Rantz, S.E., and others, 1982, Measurement and computation of streamflow—Volume 1, Measurement of stage and discharge, and Volume 2, Computation of discharge: U.S. Geological Survey Water-Supply Paper 2175, p. 1–631. [Also available at <http://pubs.usgs.gov/wsp/wsp2175/>.]
- Ries, K.G., III, Guthrie, J.D., Rhea, A.H., Steeves, P.A., and Stewart, D.W., 2008, StreamStats—A water resources web application: U.S. Geological Survey Fact Sheet 2008–3067, 6 p. [Also available at <http://pubs.usgs.gov/fs/2008/3067/>.]
- Ruhl, J.F., 1989, Flow of ground water through fractured carbonate rocks in the Prairie Du Chien-Jordan Aquifer, southeastern Minnesota: U.S. Geological Survey Open-File Report 89–253, 2 p. [Also available at <http://pubs.er.usgs.gov/publication/ofr89253/>.]
- Searcy, J.K., 1959, Flow-duration curves: U.S. Geological Survey Water-Supply Paper 1542–A, 33 p. [Also available at <http://pubs.er.usgs.gov/publication/wsp1542A/>.]
- Smakhtin, V.U., 2001, Low flow hydrology—A review: *Journal of Hydrology*, v. 240, no. 3–4, p. 147–186. [Also available at [http://dx.doi.org/10.1016/S0022-1694\(00\)00340-1](http://dx.doi.org/10.1016/S0022-1694(00)00340-1).]
- Soil Survey Staff, 2012, Natural Resources Conservation Service, U.S. Department of Agriculture, Soil Survey Geographic (SSURGO) Database: accessed April 3, 2013, at <http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/>.
- U.S. Environmental Protection Agency, 2012, Overview of impaired waters and total maximum daily loads program: accessed May 20, 2015, at <http://water.epa.gov/lawsregs/lawsguidance/cwa/tmdl/intro.cfm#section303>.
- U.S. Environmental Protection Agency, 2013, Flow 101: accessed July 7, 2015, at <http://water.epa.gov/scitech/datait/models/dflow/flow101.cfm>.

- U.S. Fish and Wildlife Service, 2013, National Wetlands Inventory Web site: Washington, D.C., U.S. Department of the Interior, Fish and Wildlife Service, accessed January 2013 at <http://www.fws.gov/wetlands/Data/Data-Download.html>.
- U.S. Geological Survey, 2013, USGS surface-water data for the Nation: U.S. Geological Survey National Water Information System, accessed April 3, 2013, at <http://waterdata.usgs.gov/nwis/sw>.
- U.S. Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood flow frequency: Reston, Virginia, U.S. Geological Survey, Office of Water Data Coordination, Hydrology Committee Bulletin 17B, 183 p.
- Venables, W.N., Smith, D.M., and the R Development Core Team, 2010, An introduction to R: Bristol, Network Theory 2008, 100 p.
- Vogel, R.M., and Kroll, C.N., 1989, Low-flow frequency analysis using probability-plot correlation coefficients: *Journal of Water Resources Planning and Management*, v. 115, p. 338–357. [Also available at [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1989\)115:3\(338\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1989)115:3(338)).]
- Warne, S.A., 1978, Map showing low-flow frequency of Minnesota streams: U.S. Geological Survey Water-Resources Investigations Report 78–132, 1 sheet. [Also available at <http://pubs.er.usgs.gov/publication/wri78132/>.]
- Winter, T.C., 2001, The concept of hydrologic landscapes: *Journal of the American Water Resources Association*, v. 37, p. 335–349. [Also available at <http://pubs.er.usgs.gov/publication/70023605/>.]
- Winterstein, T.A., Arntson, A.D., and Mitton, G.B., 2007, Methods used to compute low-flow frequency characteristics for continuous-record streamflow stations in Minnesota, 2006: U.S. Geological Survey Open-File Report 2007–1033, 22 p. [Also available at <http://pubs.er.usgs.gov/publication/ofr20071033/>.]
- Wolock, D.M., Winter, T.C., and McMahon, G., 2004, Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses: *Environmental Management*, v. 34, no. 1, p. S71–S88. [Also available at <http://dx.doi.org/10.1007/s00267-003-5077-9>.]
- Zeileis, Achim, Leisch, F., Hornik, K., and Kleiber, C., 2002, strucchange—An R package for testing for structural change in linear regression models: *Journal of Statistical Software*, v. 7, no. 2, p. 1–38. [Also available at <http://www.jstatsoft.org/v07/i02/>.]
- Zhang, Y.-K., and Schilling, K.E., 2006, Increasing streamflow and baseflow in Mississippi River since the 1940s—Effects of land use change: *Journal of Hydrology*, v. 324, p. 412–422. [Also available at <http://dx.doi.org/10.1016/j.jhydrol.2005.09.033>.]

Appendix 1. Detailed Streamgage Information Used to Conduct Analyses

Table 1–1 presents a summary of streamgages used in the analyses of this report. The map numbers presented in table 1–1 correspond to the map numbers listed for the streamgages presented in figure 1. Table 1–1 is presented as a Microsoft Excel® spreadsheet (<http://dx.doi.org/10.3133/sir20155170>).

Publishing support provided by:
Rolla Publishing Service Center

For more information concerning this publication, contact:
Director, USGS Minnesota Water Science Center
2280 Woodale Drive
Mounds View, Minnesota 55112
(763) 783-3100

Or visit the Minnesota Water Science Center Web site at:
<http://mn.water.usgs.gov/>

