U.S. Department
of Transportation

**Federal Highway
Administration**

science for a changing world

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

Publication No. FHWA-HEP-09-005
2010

Methods for Development of Planning-Level Estimates of Stormflow at Unmonitored Stream Sites in the Conterminous United States

Office of Project Development
and Environmental Review
1200 New Jersey Avenue, SE
Washington, DC 20590

Foreword

The mission of the Federal Highway Administration (FHWA) is to continually improve the quality of our Nation's highway system and intermodal connections in a manner that protects and enhances the natural environment and communities affected by transportation. In enacting the Intermodal Surface Transportation Efficiency Act of 1991 (ISTEA); the Transportation Equity Act for the 21st Century (TEA-21) in 1998; and the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) in 2005, the U.S. Congress has consistently emphasized the need for an integrated and multimodal transportation system that reflects environmental sensitivity and community values. Protecting and enhancing the environment and communities affected by transportation requires that principles of environmental stewardship be incorporated in all of the FHWA's policies, procedures, and decisions. This means that the FHWA responsibly considers and evaluates all aspects of the environment throughout the highway design, planning, and development process. Beyond its obligations embodied in environmental stewardship, the FHWA must demonstrate leadership on environmental matters in its collaboration with State and local agencies that implement transportation projects and programs throughout the country. The FHWA also has a responsibility to streamline the complex environmental stewardship process to ensure that highway projects are done in the most efficient and economical manner possible. To meet these goals, the FHWA must develop and disseminate research products that help FHWA and its partners implement surface transportation programs in a manner that protects and enhances the natural and human environment. More specifically, the Water and Ecosystems Team of the FHWA Office of Project Development and Environmental Review strives to develop and disseminate skills, tools, and information to redesign Federal environmental and transportation decisionmaking, and to ensure an integrated process at the Federal, State, tribal, and local levels. These tools, techniques and methods are designed to reduce direct and indirect adverse impacts of highways on water quality, habitat, and ecosystems to preserve and enhance human health, biological productivity, and ecological diversity.

This report, the associated computer applications, and data provide tools and techniques for developing planning-level estimates of prestorm streamflow, precipitation-event characteristics, and storm-event runoff at sites receiving highway runoff. This information is vital for assessing the potential for adverse effects of runoff on receiving waters throughout the Nation. Ready availability of methods, statistics, and computer applications for estimating upstream storm flows should provide transportation agencies with the tools and information necessary to improve project delivery without compromising environmental protection.

Patricia Cazenas
Highway Engineer
Office of Project Development and Environmental Review
Federal Highway Administration

Notice

Quality Assurance Statement

Technical Report Documentation Page

1. Report No.
FHWA-HEP-09-005

2. Government Accession No.

3. Recipients Catalog No.

4. Title and Subtitle
Methods for Development of Planning-Level Estimates of Stormflow at Unmonitored Stream Sites in the Conterminous United States

5. Report Date
December 2010

6. Performing Organization Code

7. Authors
Gregory E. Granato

8. Performing Organization Report No.

9. Performing Organization Name and Address
U.S. Geological Survey
12201 Sunrise Valley Drive
Reston, VA 20192

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DTFH61-02-Y-30079

12. Sponsoring Agency Name and Address
Federal Highway Administration
Office of Project Development and Environmental Review
1200 New Jersey Avenue, SE
Washington, DC 20590

13. Type of Report and Period Covered
Final Report
October 2003–September 2010

14. Sponsoring Agency Code

15. Supplementary Notes

Contracting Officer's Technical Representatives (COTR's): Patricia A. Cazenas

16. Abstract

This report documents methods for data compilation and analysis of statistics for stormflows that meet data-quality objectives for order-of-magnitude planning-level water-quality estimates at unmonitored sites in the conterminous United States. Statistics for prestorm streamflow, precipitation, and runoff coefficients are used to model stormflows for use with the Stochastic Empirical Loading and Dilution Model (SELDM), which is a highway-runoff model. SELDM is designed to better quantify the risk of exceeding water-quality criteria as precipitation, discharge, ambient water quality, and highway-runoff quality vary from storm to storm. Summary statistics also may be used to help estimate annual-average water-quality loads. Streamflow statistics are used to estimate prestorm flows. Streamflow statistics are estimated by analysis of data from 2,873 U.S. Geological Survey streamgages in the conterminous United States with drainage areas ranging from 10 to 500 square miles and at least 24 years of record during the period 1960–2004. Streamflow statistics are regionalized using U.S. Environmental Protection Agency Level III nutrient ecoregions. Storm-event precipitation statistics are estimated by analysis of data from 2,610 National Oceanic and Atmospheric Administration hourly-precipitation data stations in the conterminous United States with at least 25 years of data during the 1965–2006 period. Storm-event precipitation statistics are regionalized using U.S. Environmental Protection Agency rain zones. Statistics to characterize volumetric runoff coefficients are estimated using data from 6,142 storm events at 306 study sites. Runoff coefficient statistics are not regionalized, but are organized by total impervious area. All of the geographic information system files, computer programs, data files, and regression results developed for this study are included on the CD–ROM accompanying this report.

17. Key Words

Streamflow, storm event, prestorm flow, runoff, precipitation,
runoff coefficient, ecoregion, event mean concentration,
water quality,
Stochastic Empirical Loading and Dilution Model

18. Distribution Statement

No restrictions.

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
90

22. Price

Form DOT F 1700.7 (8-72) (computer generated)

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS					APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**	**Symbol**	**When You Know**	**Multiply By**	**To Find**	**Symbol**
LENGTH					**LENGTH**				
in	inches	25.4	millimeters	mm	mm	millimeters	0.039	inches	in
ft	feet	0.305	meters	m	m	meters	3.28	feet	ft
yd	yards	0.914	meters	m	m	meters	1.09	yards	yd
mi	miles	1.61	kilometers	km	km	kilometers	0.621	miles	mi
AREA					**AREA**				
in^2	square inches	645.2	square millimeters	mm^2	mm^2	square millimeters	0.0016	square inches	in^2
ft^2	square feet	0.093	square meters	m^2	m^2	square meters	10.764	square feet	ft^2
yd^2	square yard	0.836	square meters	m^2	m^2	square meters	1.195	square yards	yd^2
ac	acres	0.405	hectares	ha	ha	hectares	2.47	acres	ac
mi^2	square miles	2.59	square kilometers	km^2	km^2	square kilometers	0.386	square miles	mi^2
VOLUME					**VOLUME**				
fl oz	fluid ounces	29.57	milliliters	mL	mL	milliliters	0.034	fluid ounces	fl oz
gal	gallons	3.785	liters	L	L	liters	0.264	gallons	gal
ft^3	cubic feet	0.028	cubic meters	m^3	m^3	cubic meters	35.314	cubic feet	ft^3
yd^3	cubic yards	0.765	cubic meters	m^3	m^3	cubic meters	1.307	cubic yards	yd^3
	NOTE: volumes greater than 1000 L shall be shown in m^3								
MASS					**MASS**				
oz	ounces	28.35	grams	g	g	grams	0.035	ounces	oz
lb	pounds	0.454	kilograms	kg	kg	kilograms	2.202	pounds	lb
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")	Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
TEMPERATURE (exact degrees)					**TEMPERATURE (exact degrees)**				
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C	°C	Celsius	1.8C+32	Fahrenheit	°F
ILLUMINATION					**ILLUMINATION**				
fc	foot-candles	10.76	lux	lx	lx	lux	0.0929	foot-candles	fc
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2	cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS					**FORCE and PRESSURE or STRESS**				
lbf	poundforce	4.45	newtons	N	N	newtons	0.225	poundforce	lbf
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa	kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

(Revised March 2003)

Contents

Appendixes on CD–ROM:

Figures

26–27. Boxplots showing—

28–29. Probability plots showing—

33–34. Schematic diagrams showing—

Tables

Conversion Factors and Abbreviations

Multiply	By	To obtain
	Length	
inch (in.)	2.54	centimeter (cm)
foot (ft)	0.3048	meter (m)
	Area	
acre	0.4047	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
cubic foot per second per square mile [$(ft^3/s)/mi^2$]	0.01093	cubic meter per second per square kilometer [$(m^3/s)/km^2$]
	Basin slope	
foot per mile (ft/mi)	0.1894	meter per kilometer (m/km)

Abbreviations

1B3	1-day 3-year biological low flow
4B3	4-day 3-year biological low flow
7Q2	7-day 2-year streamflow
7Q10	7-day 10-year streamflow
AMC	antecedent moisture condition
ARC	antecedent runoff condition
ASCII	American Standard Code for Information Interchange
BASINS	Better Assessment Science Integrating point and Non-point Sources
BCF	bias correction factor
BDF	basin development factor
BMP(s)	best management practice(s)
CD–ROM	computer disk-read only memory
CN	Soil-Conservation Service Curve Number
COV	coefficient of variation
DCIA	directly connected impervious area
DEM	digital elevation model
DOS	Disk Operating System
DQOs	data quality objectives
EDNA	Elevation Derivatives for National Applications
EIA	effective impervious area
EMC	event mean concentration
EUSE	effects of urbanization on stream ecosystems
FHWA	Federal Highway Administration
GIS	geographic information system
GNWISQ	Get National Water Information System Streamflow (Q)

GUI	Graphical User Interface
HCDN	Hydro-Climatic Data Network
HUC	hydrologic unit code
HYSEP	computer program for streamflow hydrograph separation
IET	interevent time
KTRLine	Kendall-Theil robust line
Log	logarithm
Log10	common logarithm
MIA	mapped impervious area
MkDFlowF	Make U.S. Environmental Agency DFLOW3 batch input Files
MkPP	Make plotting position file
MOVE	maintenance of variance
MRLCC	Multi-Resolution Land Characteristics Consortium
NAWQA	National Water-Quality Assessment
NCDC	National Climatic Data Center
NOAA	National Oceanic and Atmospheric Administration
NLCD	National Land Cover Database
NRCS	Natural Resources Conservation Service
NURP	National Urban Runoff Program
NWIS	National Water Information System
NWISWeb	National Water Information System Web
PD	population density
PDA	percent developed area
PDF	portable document format
PRISM	precipitation-elevation regression on independent slopes model
Q	streamflow
QA	Quality assurance
QSTATS	Streamflow (Q) Statistics
Rv	runoff coefficient (volumetric)
SCS	Soil Conservation Service (now the NRCS)
SELDM	stochastic empirical loading and dilution model
SPAF	synoptic precipitation analysis facilitator
SREF	Streamflow Record Extension Facilitator
STATSGO	State Soil Geographic
SWQDM	Surface-Water Quality Data Miner Database
SYNOP	synoptic rainfall data analysis program
SYNPREP	synoptic rainfall data analysis preparation program
TIA	total impervious area
TMDL(s)	Total Maximum Daily Load(s)
URL	uniform resource locator (Internet or Web address)
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey
WDM	water data management
WDMUtil	water data management utilities program

Methods for Development of Planning-Level Estimates of Stormflow at Unmonitored Stream Sites in the Conterminous United States

By Gregory E. Granato

Abstract

This report documents methods that meet data-quality objectives for development of planning-level estimates of stormflow at unmonitored stream sites in the conterminous United States. Planning-level estimates are defined as the results of analyses that are recognized to include substantial uncertainties (commonly orders of magnitude). Planning-level estimates of stormflow for a site of interest can be made using statistics in the literature, regional statistics, statistics estimated using data collected at nearby hydrologically similar sites, or with statistics estimated using limited data collected at the site of interest. Estimates of total stormflow are derived from statistics for prestorm streamflow, precipitation, and runoff coefficients (calculated as the ratio of total runoff, in watershed inches, to rainfall, in inches). Streamflow statistics are used to estimate prestorm flows, precipitation statistics are used to estimate storm-event characteristics, and runoff coefficient statistics are used with precipitation statistics to estimate the volume of runoff from the highway and the upstream basin. The statistics developed in this analysis are intended for use with the Stochastic Empirical Loading and Dilution Model (SELDM).

The report documents selected methods for data compilation and analysis of statistics for prestorm streamflow, precipitation, and runoff. Each section of the report includes a description of data, methods, and software that can be used to estimate the necessary statistics. Appendixes to the report document reviews of previous investigations, give background information, and describe alternative methods for stormflow analysis. The geographic information system files, computer programs, data files, and regression results developed for this study are included on the CD–ROM accompanying this report.

The means, standard deviations, and skews of the logarithms of nonzero streamflows and the proportions of zero flows relative to all streamflow values are used in SELDM to generate a population of prestorm flows. These streamflow statistics were estimated by analysis of data from 2,873 U.S. Geological Survey (USGS) streamgages in the conterminous United States. Streamgages with drainage areas ranging from 10 to 500 square miles and at least 24 years of record during the period 1960–2003 were selected for analysis. In this study, streamflow statistics were regionalized according to U.S. Environmental Protection Agency Level III nutrient ecoregions. Initial estimates of prestorm flow statistics may be made using the drainage-area-ratio method with regional statistics. These estimates may be refined with statistics from nearby, hydrologically similar basins. This report was written to document methods for estimating statistics at ungaged sites, but site-specific statistics can be calculated if limited data are available by using software developed for use with SELDM. If a long-term record of daily mean flows is available, the statistics can be calculated with the existing record. If limited data are available (or are collected for analysis) from the site of interest, record extension or augmentation methods may be used to estimate the necessary statistics.

The lower bounds and the averages of the precipitation volume and duration and the time between storm midpoints are used in SELDM to generate a population of storm events. Storm-event precipitation statistics were estimated by analysis of data from 2,610 National Oceanic and Atmospheric Administration hourly-precipitation data stations in the conterminous United States. Precipitation-monitoring stations with at least 25 years of data during the 1965–2006 period were selected for analysis. Storm-event statistics were regionalized according to U.S. Environmental Protection Agency rain zones and Level III nutrient ecoregions. Initial estimates of storm-event statistics may be made using regional statistics. These estimates may be refined with statistics from nearby hourly-precipitation data stations.

The mean, standard deviation, and skew of runoff coefficients are used with storm-event statistics in SELDM to generate a population of runoff volumes. Runoff-coefficient statistics were estimated by analysis of data from 6,142 storm events at 306 study sites. Study sites included residential, commercial, industrial, institutional, agricultural, urban open space, and natural land uses in many areas within the conterminous United States. Runoff-coefficient statistics are not regionalized, but instead are analyzed using total impervious area. Regression equations were developed to

estimate the average, standard deviation, and skew of runoff coefficients on the basis of the estimated total impervious area.

Information about the duration of precipitation, runoff flows from the highway site, and runoff flows from the upstream basin also are needed to estimate the proportion of upstream stormflows that are available for diluting highway runoff. A triangular hyetograph is a better representation of the temporal distribution and peak intensity of precipitation during a storm event, but a rectangular distribution is sufficient for SELDM because it is a lumped-parameter model. A triangular approximation to the storm-event hydrograph, however, is necessary to calculate the proportion of upstream stormflows that are available for diluting highway runoff. The USGS basin-lag equation and hydrograph-recession equations from the literature are used in SELDM to represent the temporal distribution of upstream flows. Site-specific values for the basin-lag equation and the hydrograph-recession factor also may be entered in the SELDM model.

Introduction

A mass-balance approach (fig. 1) commonly is applied to estimate the concentrations and loads of water-quality constituents in receiving waters downstream of an urban or highway-runoff outfall during storm events (Driscoll and others, 1979; Warn and Brew, 1980; Di Toro, 1984; Driscoll, Shelley, and others 1989; Driscoll and others, 1990a,b). Storm events commonly are defined as independent statistical events characterized by a volume, intensity, duration, and time between storm midpoints for the purposes of planning, analysis, and sampling efforts (Driscoll and others, 1979; Athayde and others, 1983; Goforth and others, 1983; Driscoll, Palhegyi, and others 1989; Driscoll, Shelley, and others, 1989; U.S. Environmental Protection Agency, 1992; Wanielista and Yousef, 1993; Adams and Papa, 2000; Church and others, 2003). In a mass-balance model, the loads (the products of measured water discharges and concentrations) of the upstream stormflow and runoff components are added to calculate the discharges, concentrations, and loads in the fully mixed receiving water downstream of a discharge point. Statistics describing the frequency distributions of component discharges and concentrations are needed to estimate the statistics for downstream discharges, concentrations, and loads with a mass-balance model (Warn and Brew, 1980; DiToro, 1984). The resulting probability distribution of downstream event-mean concentrations (EMCs) indicates the potential for exceeding water-quality criteria and, therefore, the potential need for more information and data that may be used to identify suitable mitigation measures.

Estimates of stormflows are needed to use a mass-balance approach for predicting the discharges, concentrations, and loads of constituents of concern in runoff and receiving waters (Warn and Brew, 1980; Schwartz and Naiman, 1999). Upstream constituent concentrations may vary randomly or may be correlated with stormflows. Schwartz and Naiman (1999) demonstrate the importance of good stormflow estimates and the effect of correlation between concentrations and flow in receiving waters on the adequacy of planning-level estimates of water-quality constituent concentrations and loads from runoff. Water-quality transport curves, which are regression equations for estimating constituent concentrations from streamflow, can be used to define correlations between concentrations and flows. For example, Granato and others (2009) developed water-quality transport curves indicating that suspended sediment concentrations commonly increase with increasing streamflow, and dissolved constituents such as total hardness commonly decrease with increasing streamflow. In the Stochastic Empirical Loading and Dilution Model (SELDM), statistics describing the population of stormflows, concentrations of highway runoff (Granato and Cazenas, 2009), and the characteristics of the upstream basin are used to derive mass-balance estimates of the population of downstream stormflows and concentrations in a receiving water body.

Highway and urban runoff-quality assessments are based on storm-event analyses to characterize potential effects of stormwater discharges on receiving waters. The mass-balance approach (fig. 1) for storm-event analyses is based on estimates of upstream and highway-runoff discharges. The total upstream stormflow component comprises a prestorm streamflow and the upstream runoff. Runoff is a function of the storm-event characteristics and the rainfall-runoff transformation that occurs in the upstream basin (fig. 2). Similarly, the highway-runoff discharge is determined by storm-event characteristics and the rainfall-runoff transformation that occurs in the highway catchment. The relative importance of each component in determining downstream discharge, concentrations, and loads depends on upstream-basin characteristics, highway-catchment characteristics, and storm-event characteristics. At one extreme, runoff from a highway catchment may compose all of the downstream flow during a small storm during which rainfall is completely absorbed by soils in a pervious rural basin with an ephemeral stream. At another extreme, runoff from a highway catchment during a local thunderstorm in a large basin with a large perennial stream may cause undetectable changes in downstream discharge and water quality.

Stormflow estimates are needed to estimate potential effects of runoff in receiving waters. Specifically, estimates of stormflows that occur during the period of highway runoff to the stream are needed to calculate the potential dilution of highway runoff in the receiving stream. Information about stormflow-hydrograph characteristics is needed for estimating concurrent stormflow volumes from the highway and the upstream basin. The duration of highway runoff determines the proportion of the upstream stormflow hydrograph that contributes to the concurrent downstream flow. Therefore, best management practices (BMPs) that extend the duration of the highway-runoff hydrograph within or beyond the upstream storm-event hydrograph may increase the total concurrent flow and thus the dilution of runoff constituents in the receiving water.

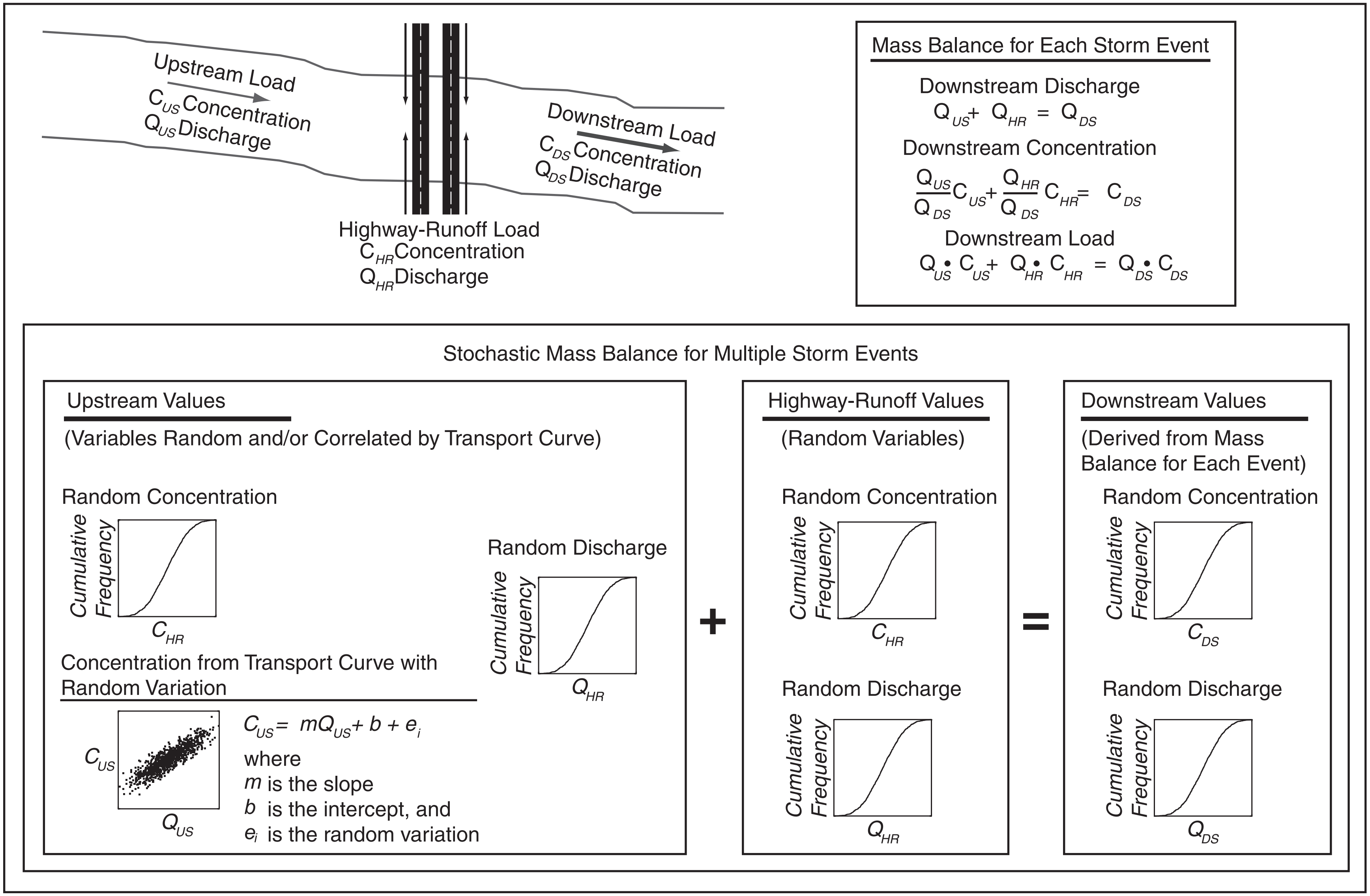


Figure 1. Schematic diagram showing the stochastic mass-balance approach for estimating discharge, concentration, and loads of water-quality constituents upstream of a highway-runoff outfall, from the highway, and downstream of the outfall.

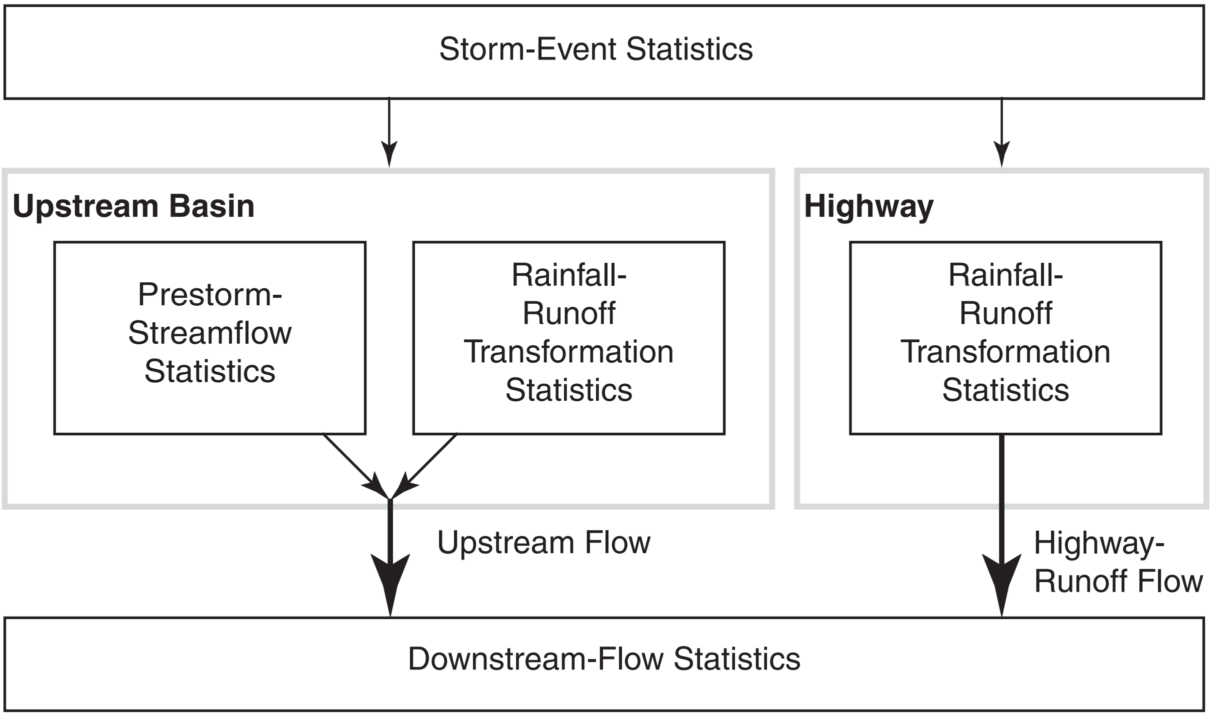


Figure 2. Schematic diagram showing the upstream-flow and highway-runoff components that must be estimated for mass-balance analysis of receiving-water quality.

The current study is designed to provide methods to derive planning-level estimates of upstream stormflows for unmonitored sites that may receive highway runoff in the conterminous United States. The stormflow estimates developed in this report are based on statistics for prestorm flows, storm-event characteristics, and runoff coefficients; each set of statistics is associated with a substantial amount of uncertainty. Planning-level estimates are commonly defined as the results of analyses used to evaluate broad policy measures. Planning-level estimates are recognized to include substantial uncertainties (commonly orders of magnitude) in all aspects of the decision process (Barnwell and Krenkel, 1982; Marsalek and Ng, 1989; Marsalek, 1991). Planning-level estimates may be based on statistics in the literature, regional statistics, statistics estimated using data collected at nearby hydrologicaly similar sites, or with statistics estimated using limited data collected at the site of interest. It may be expected that site-specific data would reduce uncertainties in planning-level estimates, but such data also may include many uncertainties (appendix 1).

Estimation of streamflow at unmonitored sites is considered to be one of the most difficult unsolved problems in hydrology (Sivapalan, 2003). Even at sites for which streamflow data are available, analysis and prediction of streamflow may be complicated by factors such as trends, step changes, seasonality, and serial correlation of flow values (Salas, 1993; Lettenmaier and Wood, 1993). The predictive abilities of conceptually based and statistical flow models for stormwater applications commonly are within an error band of one order of magnitude for any given storm (Lindner-Lunsford and Ellis, 1987; Driver and Tasker, 1990; Zarriello, 1998; Tasker and Granato, 2003). Stormflow consists of prestorm flow and runoff components of streamflow, which may include infiltration-excess overland flow, saturated overland flow, and rapid subsurface flow. Each component can vary in time and space within each storm. If the antecedent dry period is long (more than a few days, depending on the drainage area), prestorm flow may comprise only the groundwater discharge, commonly called base flow. If the antecedent dry period is short, however, prestorm flow may include flow from previous storms. Precipitation within a watershed also varies temporally and spatially, and this variation also may be scale dependent. Factors that affect the transformation of rainfall to runoff, recharge, or evapotranspiration also vary spatially and temporally.

The U.S. Geological Survey (USGS), in cooperation with the Federal Highway Administration (FHWA), began a study to develop SELDM in 2003. SELDM is a water-quality model that uses available data and stochastic Monte Carlo methods to generate planning-level estimates of EMCs, discharges, and loads from the highway and in the receiving waters upstream of the highway-runoff outfall. These values are then used to calculate the EMCs, discharges, and loads downstream of the highway-runoff outfall using mass balance methods. These estimates can be used to evaluate highway-runoff discharges as a potential source of water-quality constituents, the potential effects of runoff loads on receiving-water quality, and the potential effectiveness of BMPs for reducing the effects of highway runoff on receiving waters. Estimates of statistics for prestorm flows, precipitation, and runoff coefficients (calculated as the ratio of total runoff, in watershed inches, to rainfall, in inches) are needed for use with SELDM. Information about the duration of precipitation, runoff flows from the highway site, and flows from the upstream basin also are needed to estimate the proportion of upstream stormflows that are available for diluting highway runoff.

The effort, time, and expense needed to collect and analyze streamflow, precipitation, and runoff data limit the availability of such data for any given site. Therefore, methods to develop robust planning-level estimates of these data at unmonitored sites are needed. Initial estimates can be made on a regional basis with data that are available from a reliable source. Initial estimates of prestorm flows, storm-event characteristics, and runoff coefficients can be used for a screening-level analysis. A more detailed analysis may be warranted if initial screening indicates the potential for an unacceptable risk of adverse effects from highway runoff in receiving waters. The methods for collecting and processing data developed during this study are needed to refine regional statistics and improve estimates of local prestorm flow, storm-event characteristics, and runoff flows by using data from nearby hydrologically similar sites or from the site of interest.

Purpose and Scope

This report documents methods for compilation, analysis, and interpretation of statistics for three components of stormflow, including prestorm streamflow, precipitation, and runoff (estimated using runoff coefficients). The data, information, and statistics developed in this analysis are intended to facilitate stochastic planning-level analysis of the potential effects of highway runoff on receiving waters at unmonitored sites (or sites with limited monitoring data) in the conterminous United States. The statistics developed in this analysis are intended for use with the Stochastic Empirical Loading and Dilution Model (SELDM). Streamflow statistics are used to estimate prestorm flows, precipitation statistics are used to estimate storm-event characteristics, and runoff-coefficient statistics are used with precipitation statistics to estimate the volume of runoff from the highway and the upstream basin. These methods meet data-quality objectives (DQOs) for developing planning-level water-quality estimates at unmonitored sites in the conterminous United States.The methods and statistics that are described in this report should be useful for other stormwater analyses. For example, the “Simple Method” (Schueler, 1987; Chandler, 1994) commonly is used to develop estimates of long-term annual loads for initial screening-level runoff-quality analyses.

This report also describes methods that may be useful in obtaining and interpreting more quantitative site-specific data. If the regional estimates described in this report do not meet DQOs for a particular project, users may refine estimates by selecting and analyzing site-specific data. The report identifies potential sources of data for site-specific analyses, like the USGS National Water Information System (NWIS Web) (U.S. Geological Survey, 2009). If DQOs for a particular project require site-specific data, users may use the tools and methods described in this report to analyze their own data to refine statistical estimates for a given site. An extensive literature review is provided for each subject in this report to document source materials and to facilitate more detailed analyses.

Appendixes to the report document reviews of previous investigations, provide background information, and describe alternate methods for stormflow analysis. Detailed information about the sites (including location, storm events, and sources of data), geographic information system files, computer programs, and regression results are documented on the CD–ROM accompanying this report.

Data-Quality Objectives for Planning-Level Estimates

The FHWA has established a system of water-quality assessment and action plans that include different levels of interpretive analysis to determine potential environmental effects of highway runoff (Sevin, 1987; Cazenas and others, 1996; Federal Highway Administration, 1998). DQOs for these assessments depend on the level of interpretive analysis deemed necessary to evaluate conditions for a given site. The level of interpretive analysis may range from a completely qualitative initial assessment through an increasingly quantitative series of planning-level estimates. The compilation and interpretation of national prestorm-flow, precipitation, and runoff-coefficient statistics in the present study are designed to meet DQOs for the development of planning-level estimates of streamwater quality at unmonitored sites in the conterminous United States and to provide information and methods for refining such estimates.

The FHWA water-quality-assessment process is a step-by-step decision tree (Sevin, 1987; Cazenas and others, 1996; Federal Highway Administration, 1998). In the FHWA process, an initial assessment is completed to estimate the probability that the highway configuration being considered will produce unacceptable environmental effects. If the probable risk of an adverse effect is unacceptable to decisionmakers, the assessment is refined with more detailed data and analysis. The process is concluded if it can be demonstrated that there is a low probability that implementation of the highway design (including proposed best management practices) would produce unacceptable environmental effects. The decision rule for DQOs in this process is dependent on the sensitivity of the receiving waters, the presence of water supplies in the watershed, uncertainties in available data, and limitations of the analysis (Patricia Cazenas, U.S. Department of Transportation, Federal Highway Administration, oral commun., 2005).

The DQO process is designed to help evaluate the costs of data acquisition in relation to the consequences of a decision error caused by inadequate input data (U.S. Environmental Protection Agency, 1986, 1994, 1996; Granato and others, 2003). DQOs are meant to ensure that data and interpretations are useful for the intended purpose. In a review of hydrologic data collected by Federal, State, and local water-quality-monitoring entities, Hren and others (1987) defined five characteristics necessary to establish that data are useful. To be useful, data must be (1) representative of the system under study; (2) associated with sufficient quality assurance (QA) information to indicate the validity, reliability, and compatibility of data from different sources; (3) collected from a readily located sampling site (to assess data comparability and to interpret results of geographic/climatological variations); (4) available for public use as original data; and (5) available in useful computer files (to increase reliable compilation and manipulation of large volumes of data). The streamflow and precipitation data used in this study meet these criteria because they were collected using standard methods at well-defined monitoring sites and were drawn from extensive nationwide datasets. These data are available to the public on the Internet in standard electronic formats.

The data used to calculate runoff-coefficient statistics were compiled from different studies but meet DQOs for developing planning-level runoff-coefficient estimates. These data are representative because they were calculated with data from 6,142 storm events at 306 study sites in many areas within the conterminous United States. These sites also represent residential, commercial, industrial, institutional, agricultural, urban, open-space, and natural land uses. The study-site characteristics and the methods used for collecting and interpreting the site-specific rainfall and runoff data are documented in the original field studies.

This study also documents methods useful in obtaining and interpreting more quantitative site-specific estimates. If the regional stormflow estimates described in this report do not meet DQOs for a particular project, users may refine estimates by selecting and analyzing site-specific data. Regional estimates, however, may be more robust for predicting environmental variables at unmonitored sites than the results of a short-duration site-specific sampling program, unless this program characterizes the full range of discharge and is not affected by short-term natural or anthropogenic influences (Hughes and Larsen, 1988; Hosking and Wallis, 1997; Vogel and others, 1998; Robertson and others, 2001; Shirazi and others, 2001; Shirazi and others, 2003; Jenerette and others, 2002).

Prestorm Streamflow Statistics

SELDM uses streamflow statistics as the basis for the stochastic generation of random prestorm streamflows from the basin upstream of the highway-runoff outfall (fig. 2). Estimates of prestorm streamflow in receiving waters are important for assessing risks for adverse water-quality effects caused by runoff because prestorm flow can be a substantial proportion of total stormflow. Prestorm streamflow may include base flow (generally defined as groundwater discharge) and stormflow from a previous storm. Prestorm streamflow may be a substantial proportion of the total upstream stormflow during a storm event, especially in relatively undeveloped basins that are of greatest potential concern for maintaining aquatic and riparian ecosystems. For example, Winter and others (1998) examined daily streamflow data collected during the period 1961–1990 at 54 streamgages throughout the conterminous United States. Using hydrograph-separation techniques to estimate the groundwater contribution to streamflow, they found that the proportion of groundwater discharge in total annual streamflow may range from 14 percent (in a basin underlain by low-permeability silt and clay) to 90 percent (in a basin underlain by a highly permeable sand and gravel aquifer). The average groundwater contribution to streamflow at these sites was 52 percent, and the median was 55 percent (Winter and others, 1998). In some basins, however, streamflow data may indicate that the receiving stream is ephemeral or intermittent and thus may not have the dry-weather base flows necessary to maintain an aquatic ecosystem. In these basins, storm runoff may account for all the stormflow during many runoff events. In either case, estimates of prestorm flow (or the lack thereof) will indicate the potential dilution of upstream flows and consequently the risks for water-quality exceedences.

The population of prestorm flows is well represented by the complete population of daily mean streamflows. Prestorm flows commonly are associated with base flow because the occurrence of storm runoff defines the end of the base-flow recession period. Daily mean flow statistics, however, represent the full range of prestorm flows for many basins because of differences in the timing of discrete storm events and the storm-runoff hydrograph. For example, figure 3 indicates the wide range of prestorm flows (defined as the instantaneous flow at the beginning of a storm) if the definition of the minimum time between precipitation events is less than the stormflow-recession duration for a given basin. Approved streamflow data are reported as daily mean flows by the USGS (Mathey, 1998; U.S. Geological Survey, 2002, 2009). In comparison, independent storm events are commonly defined by using hourly data and by specifying an interevent time, which is the minimum number of dry hours between independent storm events (Driscoll and others, 1979; Athayde and others, 1983; Adams and others, 1986; Driscoll, Palhegyi, and others 1989; U.S. Environmental Protection Agency, 1992; Wanielista and Yousef, 1993; Guo and Adams, 1998a; Adams

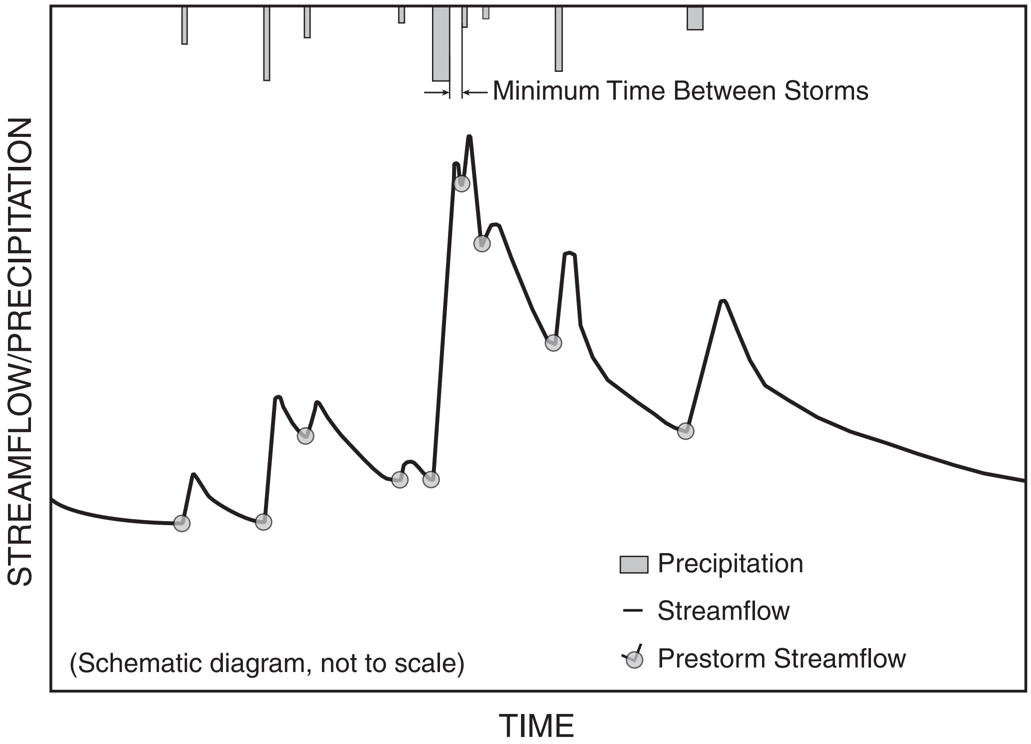


Figure 3. Schematic diagram showing the potential variability in prestorm flows that may occur if the definition of the minimum time between precipitation events is less than the duration of stormflow recession for a given basin. The minimum time between storms for highway and urban-runoff studies is 6 hours without measurable precipitation (Driscoll, Palhegyi, and others, 1989), whereas the stormflow recession for many basins may be greater than one or more days (Linsley and others, 1975).

and Papa, 2000). The minimum interevent time may differ considerably among regions but is generally approximated by an interval of about 6 hours (Driscoll, Palhegyi, and others 1989). Theoretically, there may be as many as four independent storm events with an event duration of one hour and a minimum interevent time of 6 hours in one 24-hour period used for reporting one daily mean streamflow value. Runoff events commonly are defined by the duration of the stormflow hydrograph (Linsley and others, 1975; Chow and others, 1988). Prestorm flows may include runoff from a previous storm because stormflow-recession durations for many basins commonly are longer than one or more days (Linsley and others, 1975; Sloto and Crouse, 1996). Despite the difference between the operational definitions of storm events and runoff events, daily mean streamflow statistics commonly are used as an approximation for receiving-water flow during storm events (Di Toro, 1984; Driscoll, Shelley, and others 1989; Driscoll and others, 1990a,b; Novotny, 2004).

Example data from different areas of the country were selected to indicate the potential suitability of daily mean streamflow statistics for estimating prestorm streamflow. Three streamgages that are associated with similar drainage areas and are within 20 mi of a long-term hourly precipitation monitoring station were chosen as examples to represent different climatic areas in the United States (table 1). The streamgages in Massachusetts and Washington State represent humid areas with different weather patterns. The streamgages in Massachusetts and Washington State are perennial streams. The streamgage in Arizona represents an arid area. This stream is an intermittent stream that is dry about 28 percent of the time. The medians of daily mean streamflow and associated median annual precipitation values indicate the differences in precipitation and runoff among these three areas (table 1).

Graphical examination of different streamflow values at these three streamgages (fig. 4) indicates the potential suitability of daily mean streamflow as for estimating prestorm streamflow. The boxplots in figure 4 indicate population statistics for the period 1972–1995 for all daily mean streamflows, daily mean streamflows on days before a day with measured precipitation, and daily mean streamflows on days with measured precipitation. These daily mean streamflows range from about 2 to about 800 ft^3/s at the streamgage in Massachusetts,

Table 1. Streamgages selected as examples for comparing streamflow in three areas of the country during the period from 1972 through 1995.

[USGS, U.S. Geological Survey; mi^2, square miles; ft^3/s/mi^2, cubic feet per second per square mile; median annual precipitation statistics from the National Climatic Data Center, 2002]

Station information								
USGS streamflow-gaging station					National Weather Service weather station			Ecoregion
Station name	Station number	Drainage area (mi^2)	Days with zero discharge (percent of total record)	Median streamflow (ft^3/s/mi^2)	Station number and location	Median annual precipitation 1971–2000, in inches	Distance to streamgage (miles)	Name and number
Neponset River at Norwood, MA	01105000	34.7	0	1.15	190770 Boston, MA	42.16	16.4	Northeastern Coastal Zone ecoregion 59
Sabino Creek near Tucson, AZ	09484000	35.5	27.5	0.04	028820 Tucson, AZ	11.83	15.0	Sonoran Basin and Range ecoregion 81
Skookumchuck River near Vail, WA	12025700	40.0	0	2.80	456114 Olympia, WA	50.22	19.9	Puget Lowland ecoregion 2

from zero to more than 3,000 ft^3/s at the station in Arizona, and from about 15 to more than 5,500 ft^3/s at the streamgage in Washington. At each site the population statistics for all daily mean streamflows are similar (the median of each is within the interquartile range of the others) to the statistics for daily mean streamflows on days before measured precipitation and daily mean streamflows on days with measured precipitation (fig. 4). In each case, nearly all statistics for daily mean streamflows on days before measured precipitation are higher than the equivalent statistics for all daily mean streamflows. These differences occur because the days before measured precipitation include many days with measured precipitation. This concept is illustrated hypothetically in figure 3, which shows examples of prestorm flows that are elevated by a previous storm. The boxplots in figure 4 also indicate that, in general, statistics for daily mean streamflows on days with measured precipitation are higher than equivalent statistics for all daily mean streamflows. This supports the assumption that statistics from all daily mean streamflows can be used to estimate prestorm flows, which are added to estimates of runoff statistics to estimate total stormflows. Furthermore, populations of all daily streamflows are readily obtained without concurrent precipitation measurements and thus are suitable for developing planning-level prestorm streamflow estimates at ungaged sites. These considerations explain the use of daily mean streamflow statistics by SELDM to stochastically generate a population of prestorm flows from basins upstream of a modeled highway-runoff outfall.

Methods of Estimating Streamflow at Ungaged Sites

Several methods, including streamflow maps, regression on basin characteristics, and drainage-area ratios are commonly used to estimate streamflows at sites without data. If limited data are available or can be collected at a site of interest, then record-extension methods can be used to estimate streamflow statistics at that site. The literature for each method is extensive; an overview of this literature is documented in Appendix 2 of this report.

The 1990 FHWA runoff-quality model (Driscoll and others, 1990a,b) calculated stormflows by using a national map of average annual streamflow values. This map shows 31 regions of average annual streamflow values that range from 0.05 to 5 ft^3/s/mi^2 (fig. 5). The 1990 FHWA runoff-quality model was a dilution model based on the assumption that upstream flows were lognormal and could be characterized by two parameters, the arithmetic average and coefficient of variation (COV). Driscoll and others (1990a,b) used streamflow data from 1,000 USGS streamgages to estimate a representative national COV value of 1.5. The 1990 FHWA runoff-quality model was based on methods developed by the National Urban Runoff Program (NURP), which also was based on the assumption of a lognormal distribution with an estimated COV of about 1.25 (Driscoll and others, 1979; Athayde and others, 1983). In each of these studies, the national map was provided for an initial

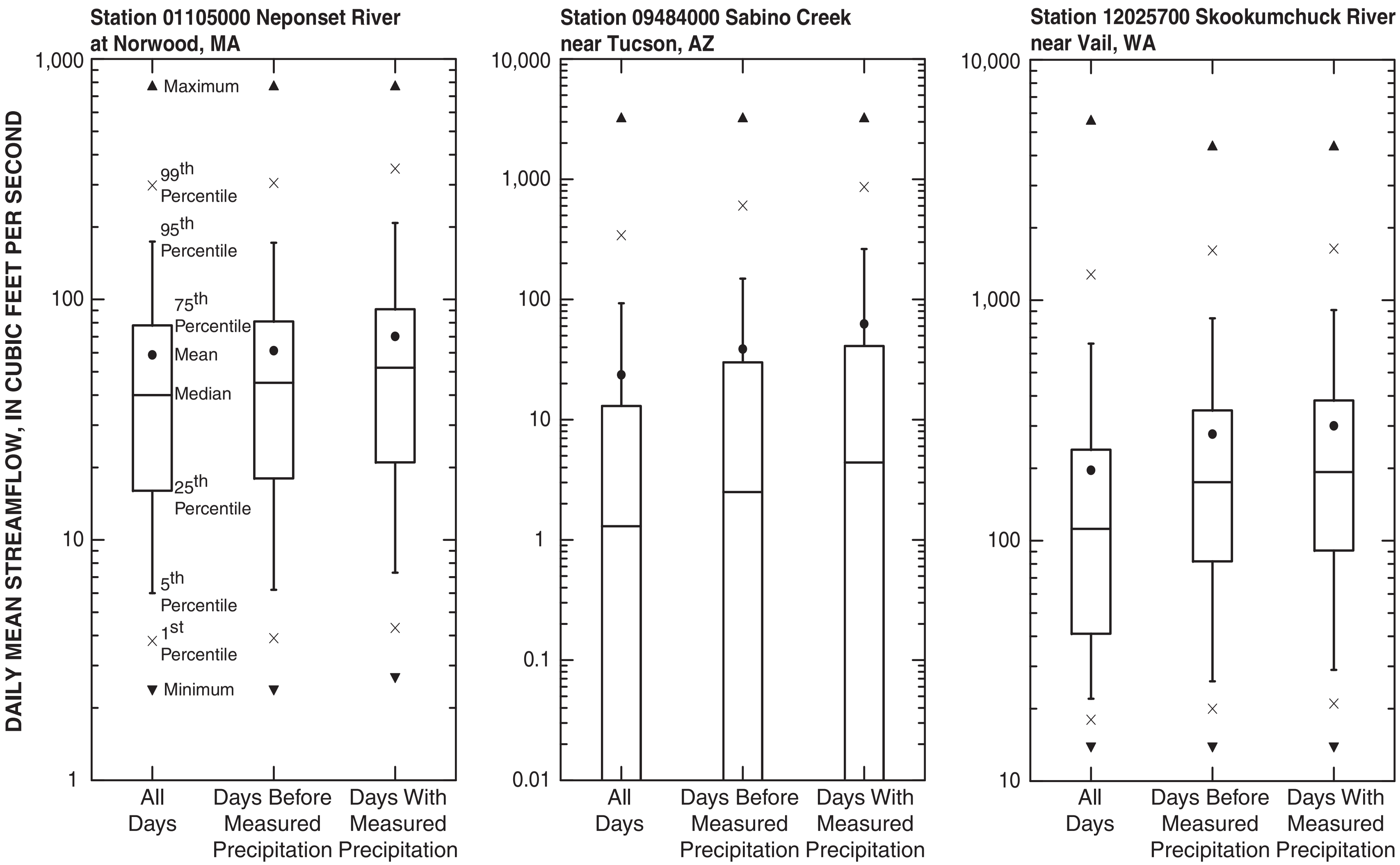
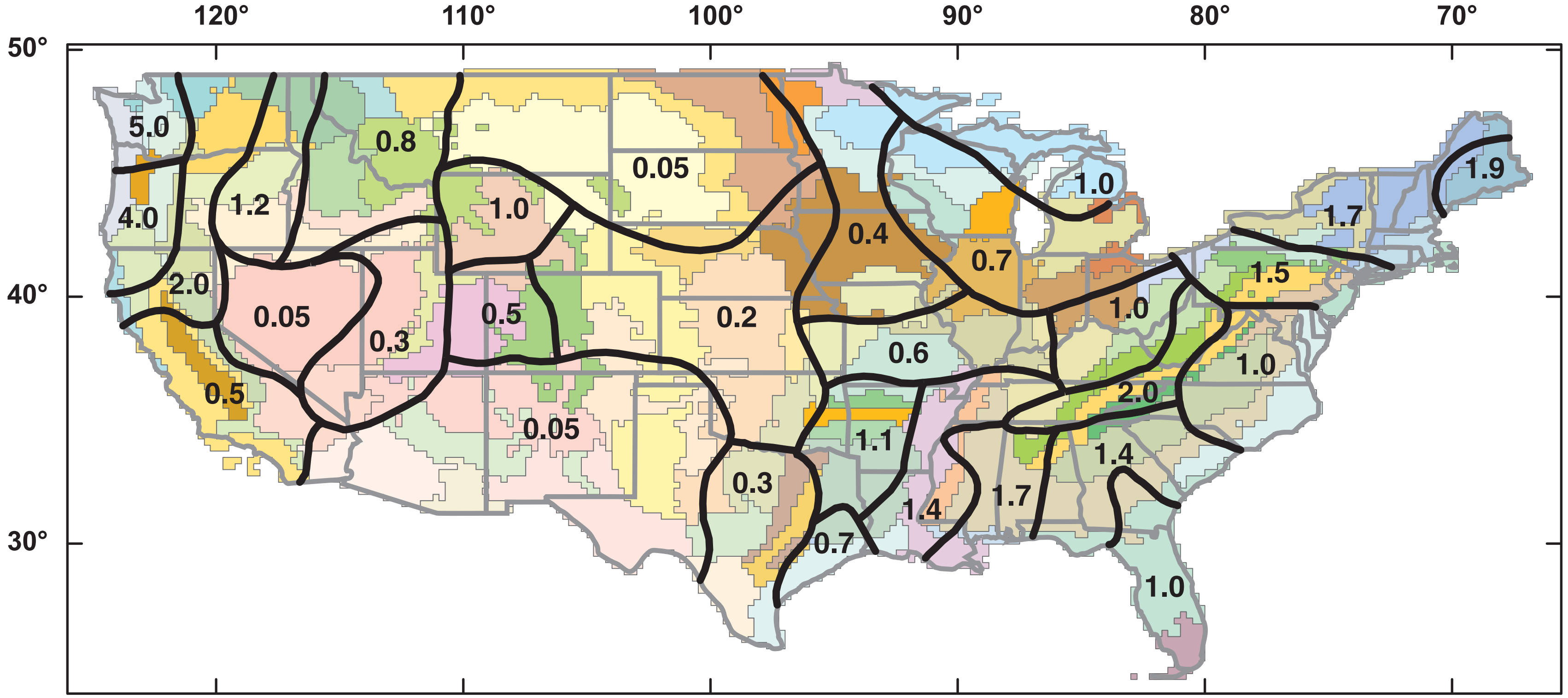


Figure 4. All mean-daily streamflows, daily mean streamflows on days before measured precipitation, and daily mean streamflows on days with measured precipitation at stations in three different climate regimes during the period from 1972 through 1995.



Base from U.S. Geological Survey National Map
Ecoregions from U.S. Environmental Protection Agency
Geographic projection

Figure 5. Regional estimates of average annual streamflow in cubic feet per second per square mile. (Modified from Driscoll and others, 1990a). These 1990 streamflow regions are shown superimposed on U.S. Environmental Protection Agency (2003) Level III ecoregions (colored polygons) that have been discretized to a 15-minute grid in the conterminous United States. Ecoregions are identified on the plate useco.pdf on the CD–ROM accompanying this report.

planning-level estimate, and was accompanied by the suggestion that local streamflow data are available from the USGS for refining these estimates.

The maps of runoff provided in the 1990 FHWA runoff-quality model (Driscoll and others, 1990a,b) and the NURP program (Driscoll and others, 1979; Athayde and others, 1983) were simplified from national maps of streamflow statistics provided in the National Atlas (Gerlach, 1970). Gerlach's contour map showed COVs that ranged from 0.20 in humid areas to 1.20 in arid areas of the southwest for annual average flows. These values were based on data collected during the period 1931–1960 and published in the USGS Hydrologic Atlas (Busby, 1966) for 8,400 streamgages with drainage areas less than 1,600 mi^2. Gebert and others (1987) updated this national streamflow map with data from the period 1951–1980.

The regression-on-basin-characteristics method, which can be used to refine map-based streamflow estimates, commonly is used to estimate streamflows at ungaged sites with data from multiple streamgages (Stedinger and others, 1993). It should be noted that implementation of the method requires a considerable effort to develop quantitative descriptions of basin characteristics and regression equations from streamflow data. Developing the required descriptions of basin characteristics for the site of interest also may be a considerable effort without knowledge of and access to the required geographic-information system (GIS) datasets. The USGS STREAMSTATS is an online application for estimating streamflow statistics by using regression-on-basin characteristics (Ries and others, 2004). Currently (2009), statewide STREAMSTATS applications have been implemented for some statistics and basin characteristics in 17 states, are being tested in 3 states, and are under development in 13 states. The existing applications currently do not have all of the basin characteristics and streamflow statistics necessary for use with SELDM, but this information can be developed.

The drainage-area-ratio method provides a simpler way to estimate daily mean streamflows and streamflow statistics at an ungaged location from a streamflow record collected from a hydrologically similar basin with a similar drainage area (Stedinger and others, 1993). The method is based on the assumption that streamflow statistics can be transferred to nearby hydrologically similar basins by adjusting the statistics to represent the differences in drainage areas. This simple method can provide estimates that are as good as estimates from more complex methods if the characteristics of the index site represent characteristics of the site of interest. Differences in basin characteristics, however, can substantially affect the representativeness of statistics from the index site. For example, Thomas (1966) developed flow-duration curves for 24 unregulated streams in a gently rolling glaciated terrain in Connecticut with 30 years of data (fig. 6). This graph shows substantial differences in streamflow statistics that may occur as a function of surficial geology.

If limited streamflow data are available for a site of interest or a nearby site, estimates of long-term streamflow data and streamflow statistics also may be generated by record-extension methods. Hirsch (1982) applied the technique called the "line of organic correlation" for streamflow-record extension with maintenance of variance (MOVE.1). Vogel and Stedinger (1985) adjusted and extended this technique (MOVE.3) for the possibility that the mean and variance of the common-record period used to develop the regression equation did not represent the full period to which the record would be extended. Both MOVE.1 and MOVE.3 are widely used for streamflow-record extension (Parrett and Cartier, 1990; Stedinger and others, 1993; Ries and Friesz, 2000).

Sources of Data for Estimating Prestorm Flows

The USGS streamflow data-collection program is designed to provide streamflow data at gaged sites and to provide information that can be used to estimate streamflows at almost any point along any stream in the United States (Benson and Carter, 1973; Wahl and others, 1995; National Research Council, 2004). The USGS National Water Information System (NWIS) internet application, NWIS Web, is a source of local, regional, or national streamflow data (Mathey, 1998; U.S. Geological Survey, 2002; 2009). The NWIS is a distributed network of computers and file servers used to store and retrieve hydrologic data. Continuous records of daily mean streamflow for periods of years to decades are available for more than 24,000 streamgages across the United States. The NWIS Web database can be searched for basic site characteristics, streamflow data, and streamflow statistics by latitude and longitude, by state, by hydrologic unit code (HUC), and by site number. Information and data from NWIS Web commonly are used to characterize streamflows at sites that have monitoring data and to predict streamflows at sites without monitoring data. These data are freely available to anyone with access to an Internet connection.

Many individual streamflow measurements that can be used for record extension (by MOVE.1 or MOVE.3) also are available from NWIS Web. The U.S. Geological Survey (2009) indicates that there are more than 46,000 sites with concurrent measurements of stage and flow. There also are data for more than 24,000 surface-water-quality monitoring stations in the conterminous United States with one or more paired measurements of stream discharge and concentration available in NWIS Web (Granato and others, 2009).

Software for Analyzing Streamflow Data

Reliable, efficient, and repeatable methods are needed to access and process streamflow information and data. Five computer programs and a database application were developed, utilized, and documented for obtaining and analyzing NWIS Web streamflow data to develop streamflow statistics for planning-level analyses to support SELDM (Granato, 2009). These computer programs include

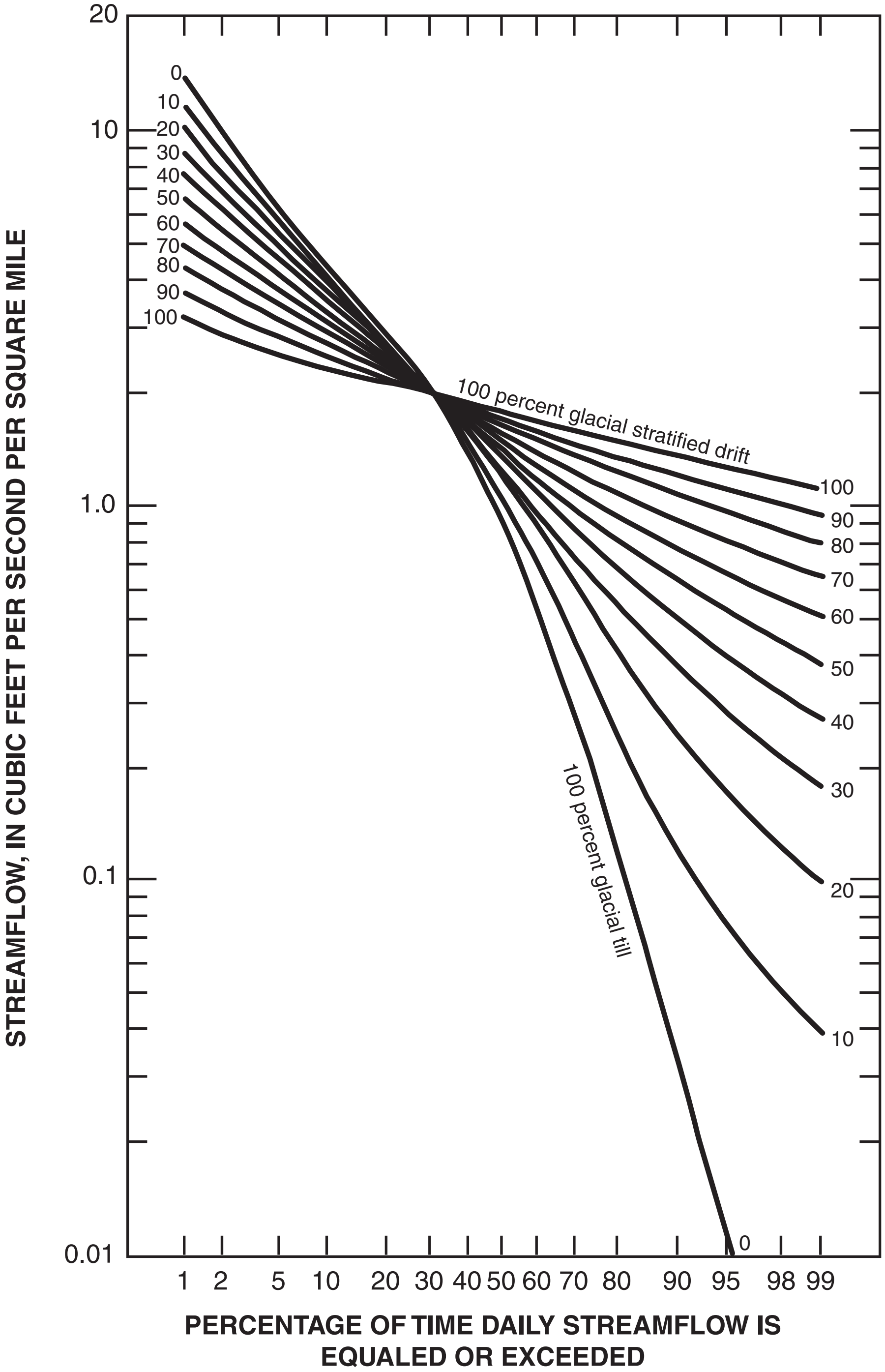


Figure 6. Flow-duration curves that indicate potential effects of the surficial geology of a drainage basin on variations in streamflow (from Thomas, 1966).

- Get National Water Information System Streamflow (Q) files (GNWISQ Version 1.0)—A program to facilitate the downloading of streamgage information and daily mean streamflow data files from the USGS NWIS Web site (Granato, 2009). The program also is designed to reformat the current (2009) NWIS Web text-file format for use with the other computer programs. The output from the GNWISQ program can be used to facilitate hydrologic-data analysis with multiple programs in individual or batch mode.

- Make plotting position file (MkPP Version 1.0)—A program for generating plotting positions and normal scores for daily mean streamflow files to facilitate visual analysis of daily mean streamflow data by plotting flow-duration curves (Granato, 2009). Visual analysis of streamflow data can be used to reveal similarities and differences in flow statistics at sites being considered as index sites. For example, the flow-duration curves in figure 6 indicate that surficial geology may be an important characteristic to consider for identifying hydrologic similarity among index sites in some areas of the country.

- Streamflow Record Extension Facilitator (SREF Version 1.0)—A program to provide an estimated long-term record of daily mean streamflows (record extension) or long-term estimates of streamflow statistics (record augmentation) at sites with limited data (Granato, 2009). SREF can be used to extend or augment limited data from a partial-record or short-term streamgage by using data from a representative long-term continuous-record streamgage. The user may output estimates of selected long-term statistics and an estimated record of daily mean streamflow data for the site of interest.

- Streamflow (Q) Statistics (QSTATS–Version 1.0)—A program to facilitate statistical analysis of daily mean streamflow data for one or multiple streamgages (Granato, 2009). The program can be used to calculate the average, standard deviation, skew, and median of daily mean streamflow values in arithmetic and common logarithmic (log10) space. The program provides the option to calculate probability-weighted moments and L-moments in arithmetic and log10 space. The program also calculates the total number of daily mean streamflow values, the number of gaps in the record (each of which may be one day to several decades long), and the fraction of zero flows recorded in the data file.

- Make U.S. Environmental Agency DFLOW3 batch input Files (MkDFlowF Version 1.0)—A program to facilitate the creation of batch input files for the U.S. Environmental Protection Agency (2004) DFLOW3 program (Granato, 2009). The input files are created automatically from a list of manually specified USGS streamgage numbers. The DFLOW3 program calculates low-flow statistics used for setting water-quality criteria and total maximum daily waste-load allocations.

The Surface Water-Quality Data Miner (SWQDM) database application was developed to facilitate the national synthesis of surface-water-quality data (Granato and others, 2009). The SWQDM database is a relational database that can be used to associate the streamflow statistics with an ecoregion or with any user-specified site location in the conterminous United States. Streamflow statistics are obtained by selecting the ecoregion of interest or by entering the latitude and longitude of the site of interest. Detailed instructions and database installation files for the SWQDM database are available on the CD–ROM accompanying the background water-quality report developed to support SELDM (Granato and others, 2009).

Regionalization of Streamflow Statistics

Streamflow statistics vary from site to site and are affected by a number of natural and anthropogenic factors (Langbein, 1949; Thomas and Benson, 1970; Lins, 1997; Poff and others, 2006). Therefore, a combination of sites or regionalization of data may be necessary to produce quantitative predictions of streamflow from available datasets for an unmonitored site. Regionalization is the process for reducing variability in a national dataset by identifying areas with common hydrological characteristics that may influence the data of interest. In this study, streamflow regionalization was used to identify hydrologically similar areas for estimation of prestorm streamflow at unmonitored sites. Methods for selecting regions included

- Use of geographical areas delineated by physical, political, or administrative boundaries, such as states or hydrologic units (U.S. Geological Survey, 1982; Seaber and others, 1987; Vogel and others, 1999; Wolock and others, 2004);

- Use of contour maps of the parameter(s) of interest (Langbein, 1949; Rainwater, 1962; Busby, 1966; Gerlach, 1970; Gerbert and others, 1987; Krug and others, 1989; Rochelle and others, 1989; Krug and others 1990; Driscoll and others 1990a; Bishop and Church, 1995; Lichty and Karlinger, 1995);

- Semiquantitative regionalization of explanatory variables for the parameter(s) of interest (Hughes and Larsen, 1988; Omernik and Bailey, 1997; U.S. Environmental Protection Agency, 2003; Wolock and others, 2004; Poff and others, 2006);

- Quantitative regionalization of explanatory variables for the parameter(s) of interest (Driver and Tasker, 1990; Poff, 1996; Robertson and others, 2001;

Griffith and others, 2002; Smith and others, 2003a; Wolock and others, 2004; Eng and others, 2005); and

- Detailed statistical analysis of the parameter(s) of interest (Langbein and Schumm, 1958; Hosking and Wallis, 1997; Lins, 1997; Robertson and others 2001; Jenerette and others, 2002).

Each method for regionalizing environmental data has benefits and limitations. For example, Lins (1997) concluded that the commonly used HUCs do not conform to any single pattern of streamflow variability identified by use of principal components analysis.

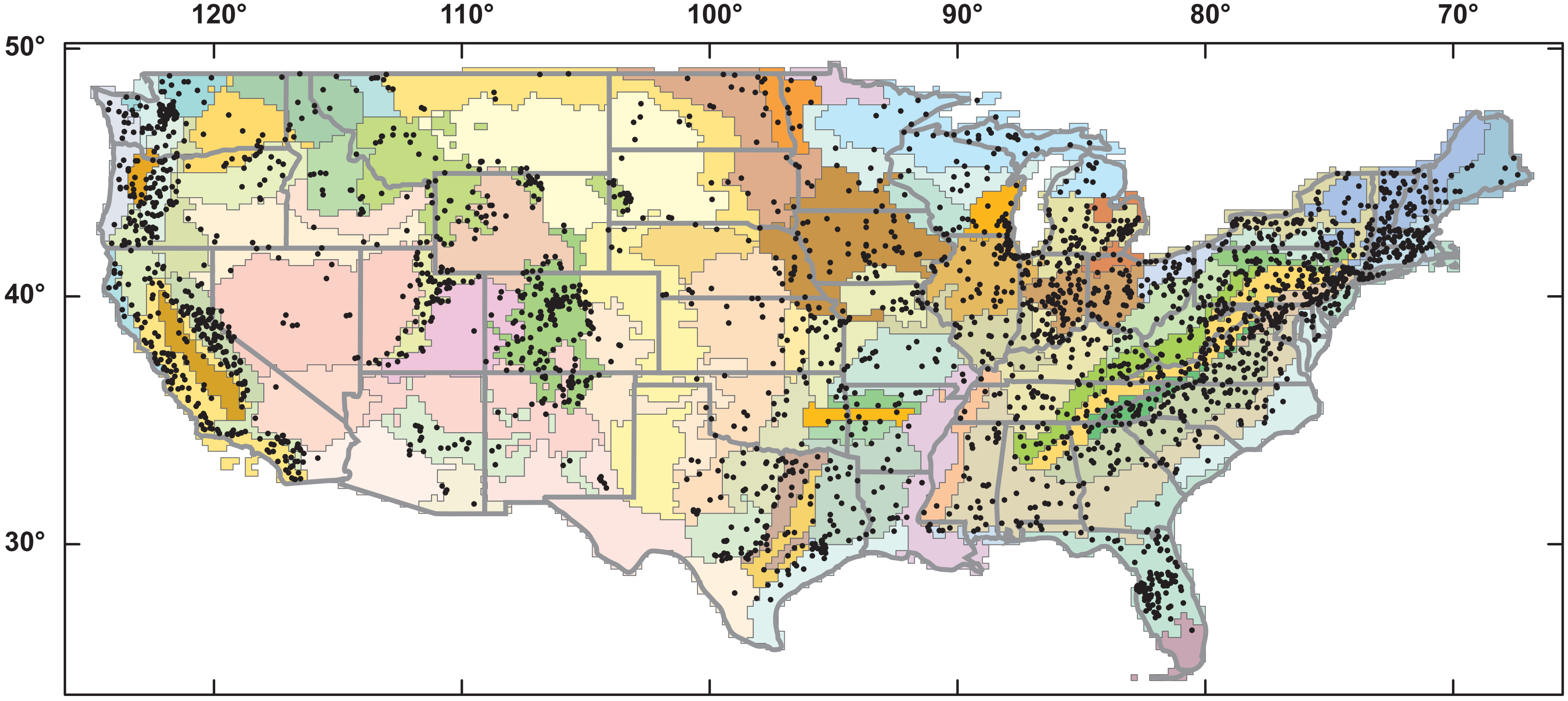
The U.S. Environmental Protection Agency (2003) Level III nutrient ecoregions were selected to regionalize streamflow and water-quality data (Granato and others, 2009) for the SELDM study. Ecoregions are defined as areas of relative homogeneity in ecological systems and their components (Omernik and others, 2000; Omernik, 2004). Level III nutrient ecoregions were chosed to provide a consistent national context for developing planning-level estimates of environmental conditions for runoff-quality analysis. Environmental-resource-management agencies in many states are increasingly using ecoregions to set water-quality criteria, develop biological criteria, and evaluate nonpoint-source management goals (Omernik and Bailey, 1997). Federal agencies that have missions for water-quality monitoring or management, like the USGS and the FHWA, also are using ecoregions as a spatial framework to organize and interpret environmental data (Intergovernmental Task Force on Monitoring Water Quality, 1995a,b; McMahon and others, 2001; Brown, 2006).

Delineation of ecoregions is a semiquantitative or qualitative process that uses information about geology, physiography, vegetation, climate, soils, land use, wildlife, and hydrology (Omernik, 1995; Omernik and others, 2000). Historically, these factors have been used as explanatory variables in models for predicting streamflow (Thomas and Benson 1970; Jennings and others, 1994; Vogel and others, 1998) and water quality (Langbein and Schumm, 1958; Syvitski and others, 2000; Shirazi and others, 2001; Griffith and others, 2002; Smith and others, 2003). Ecoregions are designed with different levels of detail; there are 15 Level I ecoregions, 52 Level II ecoregions, and 120 Level III ecoregions in North America (U.S. Environmental Protection Agency, 2003).

Ecoregions have been evaluated as a metric to estimate streamflow statistics. Bailey (1984) found substantial differences in monthly runoff at 87 percent of the USGS streamgages grouped by Level I ecoregions. He concluded that data from the remaining streamgages would be properly classified by use of finer scale regions. Carr and others (2000) found substantial differences in temperature, precipitation, growing degree days, and relations between annual average precipitation and temperature among Level II ecoregions. Wolock and others (2004) found that Level II USEPA ecoregions could explain about 79 percent of the variability in a statistic for precipitation minus potential evapotranspiration (a common estimate of streamflow) among streamgages in the conterminous United States and concluded that conceptually based hydrologic frameworks, such as ecoregions or hydrologic landscape regions, do provide advantages for interpreting data. Friesz and Ries (1998) indicated a substantial overlap in estimated mean annual streamflows among streamgages in different Level IV ecoregions in Massachusetts. Comparison of the mean annual streamflows estimated by Friesz and Ries (1998) by means of a rank-sum test (Helsel and Hirsch, 2002), however, indicates a statistically significant difference in the sample of mean annual streamflows for streamgages in non-adjacent subregions. The use of ecoregions for planning-level streamflow estimates also also has been demonstrated to be a useful indicator of regional water-quality characteristics (Heiskary and others, 1987; Hughes and Larsen, 1988; Robertson and others, 2001; Rohm and others, 2002; Simon and others, 2004).

Ecoregions may be a useful surrogate for identifying hydrologically similar areas used in developing regional-streamflow estimates at ungaged sites; however, use of site-specific explanatory variables, such as local land use or soil types commonly improves generalized predictions based on ecoregion-scale characteristics (Jenerette and others, 2002; Smith and others, 2003). For example, Poff and others (2006) found evidence for substantial anthropogenic alteration of hydrological characteristics of streamflows in different areas of the country. Detailed site-specific information, however, is not readily available in a uniform national dataset for each USGS streamgage for the period of record selected for this study (1960–2003). NWIS Web includes basic information about streamgages such as the location, altitude, drainage area, and topographic setting (U.S. Geological Survey, 2009), but not information about other explanatory variables that are used for ecoregion analysis such as local soils or land use in the basin of interest. The geographic coordinates of each streamgage, however, can be used to assign it to the appropriate ecoregion.

In the current investigation, the USEPA ecoregion coverage was discretized to the resolution of 15-minute (0.25 decimal-degree) latitude-longitude grid squares to facilitate analysis of ecoregion data with the SWQDM and SELDM database applications (fig. 7; table 2). The ecoregion with the highest proportion of area in each grid square was assigned as the ecoregion for that grid square. This discretization process did not violate the intent of the original delineation, because the 15-minute-grid square is relatively small in comparison to the scale of the Level III ecoregions. Although ecoregion boundaries are shown on maps as lines, transitions between regions actually occur over large amorphous zones along the edges of each region (J.M. Omernik, U.S. Environmental Protection Agency, written commun., 2004). Information about the USEPA (2003) Level III nutrient ecoregions, including a printable ecoregion map, a GIS coverage of the discretized version of these ecoregions, and a copy of the geographic-discretization grid, is provided on the CD–ROM accompanying this report.



Base from U.S. Geological Survey National Map
Ecoregions from U.S. Environmental Protection Agency
Geographic projection

Figure 7. Spatial distribution of 2,783 selected U.S. Geological Survey streamgages (black dots) with respect to U.S. Environmental Protection Agency (2003) Level III ecoregions (colored polygons) that have been discretized to a 15-minute grid in the conterminous United States. The selected U.S. Geological Survey streamgages have drainage areas between 10 and 500 square miles, with data for at least 24 years during the 1960–2003 period, and are not listed as being at, near, or immediately below a dam. Ecoregions are identified on the plate useco.pdf on the CD–ROM accompanying this report.

Table 2. Geographic Information System files used for regionalization and site selection.

[USEPA, U.S. Environmental Protection Agency; USGS, U.S. Geological Survey; NOAA, National Oceanic and Atmospheric Administration; --, not applicable]

Name of data layer	Scale	Feature type	Citation	Short description
gridtemplate	1:250,000	Polygon	--	Fishnet grid of 0.25-decimal-degree (15-minute) latitude-longitude grid cells
ecoregions	1:250,000	Polygon	USEPA, 2003	USEPA Level III nutrient ecoregions discretized to the 0.25-decimal-degree grid
rz15poly	1:250,000	Polygon	Smieszek and Granato, 2000	15 USEPA rain zones in the conterminous United States (filled polygons)
rz15line	1:250,000	Outline	Smieszek and Granato, 2000	15 USEPA rain zones in the conterminous United States (outlines)
states	1:250,000	Polygon	--	Conterminous United States
stormsites	1:250,000	Point	--	Rainfall-runoff data-collection sites
surfacewater	1:250,000	Point	--	Streamgages
synop2000	1:250,000	Point	--	NOAA hourly-precipitation monitoring stations with synoptic storm-event statistics for the 1965–2006 period

Selection of Streamgages

Streamgages were selected to provide the data and information necessary to estimate streamflow statistics at ungaged sites in each ecoregion. The site-selection process was constrained by several competing objectives, including ecoregion representation, drainage-area minimization, characterization of ecologically high-value perennial streams, use of a common hydrologic period, maximizing the length of record, and characterization of anthropogenic effects on streamflow. The spatial distribution of the 2,873 USGS streamgages that met the site-selection criteria is shown with respect to the 84 ecoregions of the conterminous United States in figure 7. A GIS coverage of these streamgages is provided in the GIS directory on the CD–ROM accompanying this report (table 2).

Minimizing the drainage area of selected streamgages is a consideration for streamflow and water quality. Smaller drainage areas are expected to be associated with fewer variations in the explanatory variables that affect streamflow and water quality. Streams from smaller drainage areas also are expected to be more sensitive to highway-runoff inputs. For example, based on a series of simplifying assumptions, Driscoll and others (1990b) estimated that streamflow from about 0.8 mi^2 of upstream drainage area per acre of pavement (or about 1.2 mi^2 per 12 ft-wide lane mile of pavement) might be sufficient to dilute runoff enough to minimize the potential for adverse effects of highway runoff. Larger basins also are more likely to include impoundments that alter the flow regime and affect water quality (Meade and Parker, 1985; Smith and others, 2003). Minimizing the drainage area for selected streamgages, however, competes with the objective to obtain multiple streamgages for estimating streamflow statistics that are characteristic of each ecoregion. Minimizing the drainage area also competes with the objective to characterize statistics for streamflow in perennial streams rather than intermittent or ephemeral streams. This is because larger drainage areas are necessary to support year-round flows in arid areas.

The focus of the streamflow analysis in this report is on perennial streams because they provide the best aquatic habitat. By definition, perennial streams will never go dry (Chow and others, 1988; Mosley and McKerchar, 1993). By practice and statute, however, perennial streams are defined as having an observable or measurable flow except during drought conditions (Bent and Archfield, 2002; Bent and Steeves, 2006). For example, Bent and Steeves (2006) point out that Idaho classifies stream reaches with a 7-day 2-year (7Q2) streamflow greater than 0.1 ft^3/s as perennial stream reaches. Intermittent streams commonly have a substantial proportion of nonzero prestorm flows. Ephemeral streams do not have prestorm base flows. Intermittent and ephemeral streams, however, may have nonzero prestorm flows between storms that occur in rapid succession. Intermittent and ephemeral streams have ecological value, but may not have a thriving aquatic ecosystem that would be affected by the quality of stormflow runoff. Intermittent and ephemeral streams were not exluded from this analysis beyond the application of a national minimum drainage-basin size threshold.

A minimum drainage area of 10 mi^2 was selected to provide a criterion for selecting perennial streams nationwide. This threshold was based on streamflows in Massachusetts, a relatively humid area of the country with relatively low hydrologic variability (Vogel and others, 1998), where streamflows at about 5 percent of sites with unaltered flow and drainage areas ranging from 2 to 11 mi^2 were intermittent (Bent and Archfield, 2002). This drainage-area threshold also was used to better represent general conditions in an ecoregion by including some spatial variation in basin characteristics.

In this report, a dry day is defined as a day with an daily mean streamflow that is less than detection limits—commonly about 0.01 ft^3/s (Rantz, 1982). About 30 percent of the 2,783 selected streamgages in the conterminous United States have one or more dry days during the period of record, and about 22 percent may be considered to be intermittent or ephemeral because they have an average of more than one dry day per year (fig. 8). The days with zero flow, however, are not evenly distributed from year to year, but are concentrated among drought years. For example, about 84 percent of the streamgages in the second to lowest category of dry days (0.27 to less than 1 percent of dry days) have a nonzero one-day

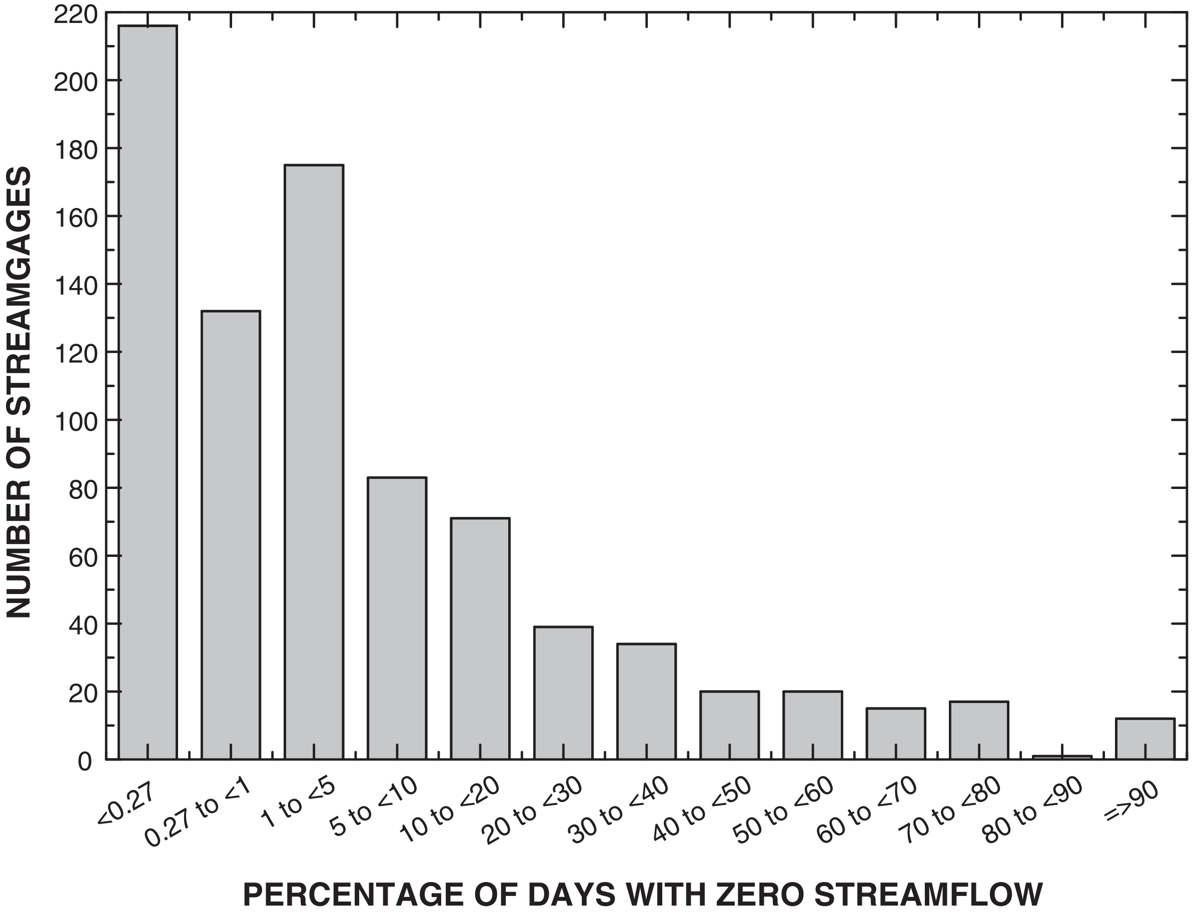


Figure 8. The percentage of dry days for the 835 of 2,783 selected U.S. Geological Survey streamgages in the conterminous United States at which one or more average daily flows were below detection limits during the 1960–2003 period (<, less than; =>, greater than or equal to).

three-year biological low flow (Rossman, 1990a,b), indicating that they are flowing streams under normal conditions. In the next category (1 to 5 percent of dry days), however, only about 30 percent of streamgages have a nonzero one-day three-year biological low flow.

Site selection by drainage area was an iterative process designed to represent streamflow statistics for perennial streams in each USEPA (2003) Level III ecoregion. The first objective was to select enough streamgages in each ecoregion to develop a dataset that is adequate for providing planning-level estimates of streamflow at ungaged sites. NWIS Web (U.S. Geological Survey, 2002; 2009) was queried to retrieve a list of all streamgages with drainage areas within the specified limits. After each NWIS Web query, a GIS query of the streamgage dataset was run to count the number of streamgages in each ecoregion. The specified maximum drainage-basin size was increased from an initial value of 50 to 500 mi^2 during this iterative process. The increase was necessary to identify one or more candidate streamgages in each Level III ecoregion. The distribution of drainage areas among the 2,387 streamgages with drainage areas ranging from 10 to 500 mi^2 is shown in figure 9.

The record length and hydrological period also were factors for site selection. An increase in record length increases both the accuracy and precision of estimates of hydrologic statistics and provides information to evaluate cycles or trends in hydrologic data (Haan, 1977; Stedinger and others, 1993; Helsel and Hirsch, 2002). Studies have shown that decades of rainfall and streamflow data are necessary to generate design-storm statistics for a catchment (Alley, 1977; Church and others, 2003). For example, in a national study of streamflow statistics for 176 streamgages in watersheds with minimal water withdrawals, Saunders and others (2004) found that about 20 years of record were needed to

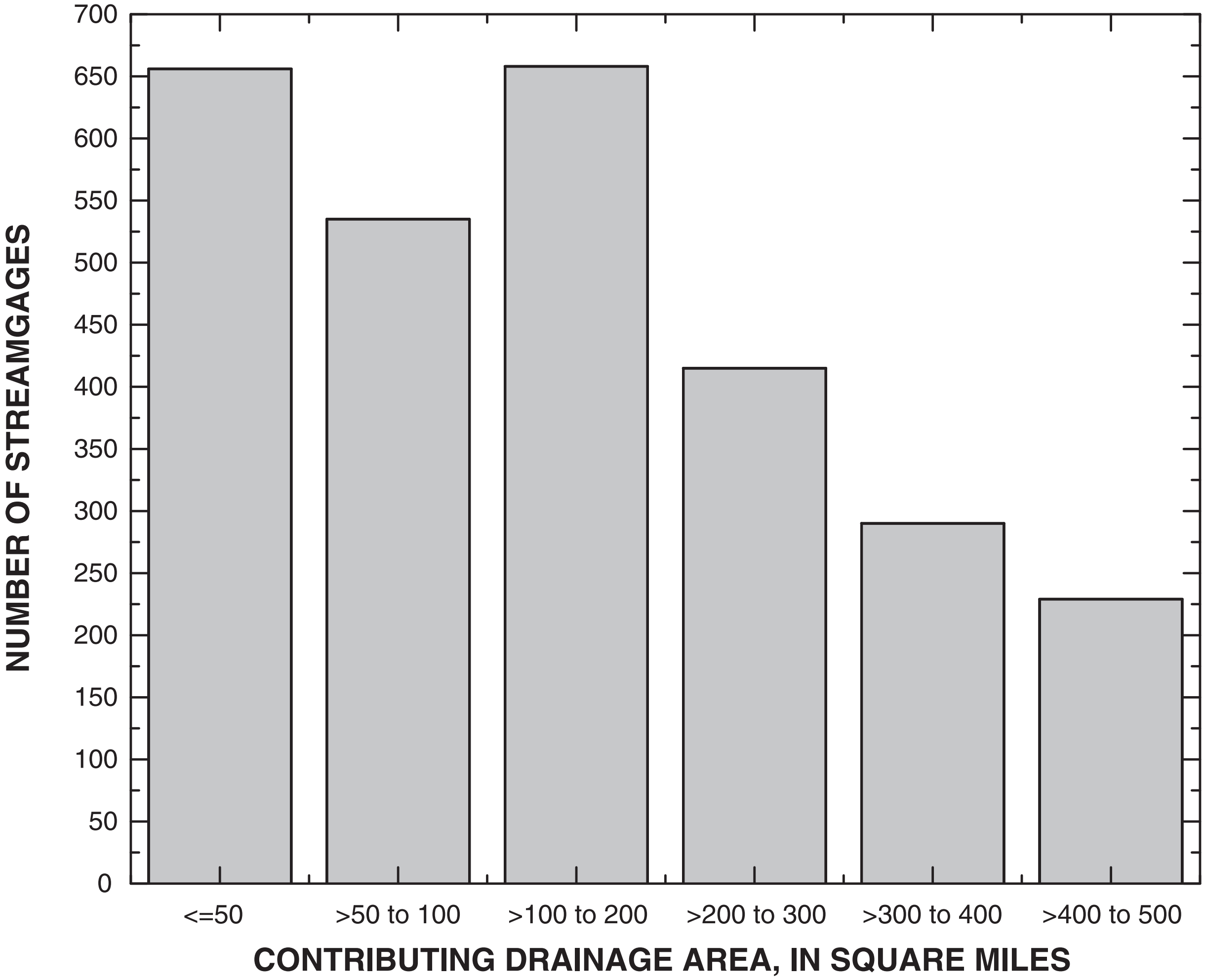


Figure 9. The distribution of drainage areas for 2,783 selected U.S. Geological Survey streamgages in the conterminous United States (<=, less than or equal to; >, greater than).

stabilize the variability and accuracy of estimates of low-flow statistics. Use of a common period of record maximizes the comparability of statistics among reference streamgages so that the effects of other explanatory variables such as basin characteristics may be more accurately quantified (Thomas and Benson, 1970; Haan, 1977; Stedinger and others, 1993). Record extension or augmentation (Hirsch, 1979, 1982; Vogel and Stedinger, 1985) is commonly used to increase the lengths of records for streamgages to a common period rather than eliminating streamgages or reducing record lengths for longer term streamgages to a common period. Because record extension is highly interpretive, requires great care in the selection of index sites, and requires a substantial analytical effort for each site (Hirsch, 1979, 1982; Helsel and Hirsch, 2002), these methods were not used to normalize data from all streamgages to a common period of record for this national evaluation of streamflow statistics by ecoregion.

The use of a common period of record and maximizing the length of record are competing objectives because the number of streamgages with common periods decrease as the period-length increases, and the periods during which data was collected for individual streamgages may be discontinuous during any given period. Streamgages with at least 24 years of record during the period 1960–2003 were selected to balance the needs for a long record, a common period of record, a suitable range of drainage areas, and the availability of data for each ecoregion. The distribution of the 2,387 streamgages by record length expressed in years (total days of record divided by 365.25 days) is shown in figure 10. The streamflow-analysis programs developed for the SELDM study, however, may be used with NWIS Web to collect and analyze streamflow data for different record lengths or hydroperiods if necessary for local analysis.

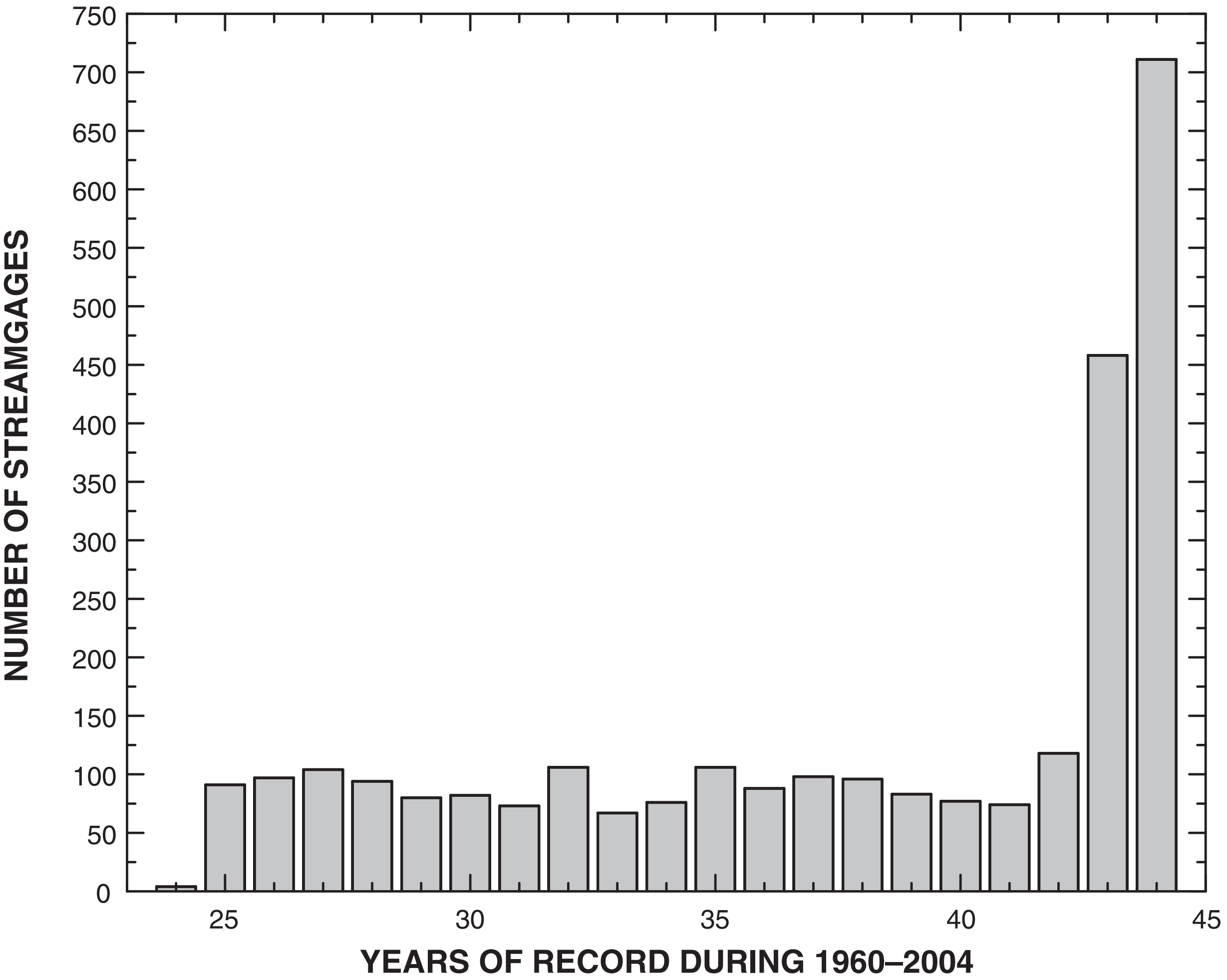


Figure 10. The distribution of record lengths for 2,783 selected U.S. Geological Survey streamgages in the conterminous United States.

The interpretation of streamflow statistics by ecoregion is limited without acknowledging the potential anthropogenic effects on streamflow. Some land-use characteristics are included in the delineation of USEPA Level III ecoregions (Omernik and others, 2000; U.S. Environmental Protection Agency, 2003; Omernik, 2004), but only in general terms. Increases in imperviousness can have a substantial effect on streamflow at the local scale (Jennings and Jarnagin, 2002). In general, increases in imperviousness have been related to increases in peak flows and reductions in low flows; these effects may increase the variability and skew of a long-term streamflow record. In large watersheds, increases in imperviousness commonly are gradual and continuous over long time periods. Water use also can have a substantial effect on streamflow at local and regional scales (Anning and Konieczki, 2005; Weiskel and others, 2007). The objective to characterize anthropogenic effects on streamflow would require multiple streamgages in ecoregions with hydrologically similar drainage basins but different water- and land-use characteristics. Alternatively, separating streamflow records into different time periods would be necessary for analyses of changes in land and water use through time. Attempting to achieve all of these objectives would split the population of streamgages in each ecoregion and reduce the record length for each streamgage. Although regional streamflow statistics calculated without regard to climatic and anthropological effects may be sufficient for initial screening-level analyses, estimates may be refined by selecting statistics for one or more nearby streamgages on the basis of drainage basins with similar hydrogeological characteristics, land use, and water use.

Basic site information from NWIS Web was useful for screening out some sites and interpreting data from other sites. NWIS Web includes basic information about streamgages, such as the station name, location, altitude, drainage area, and topographic setting (U.S. Geological Survey, 2009). All streamgages with names indicating that the streamgage is at, near, or immediately below a dam were excluded to reduce anthropogenic effects on regional streamflow statistics; dams tend to alter natural-flow statistics by reducing peaks, increasing minimum flows, and increasing the proportion of moderate flows in the streamflow record (Poff and others, 2006). Streamgage-information files from NWIS Web include the drainage area, which is defined by topographic surface-water divides, and the contributing drainage area, which may be delineated to reflect differing surface-water and groundwater contributing areas or water diversions. The contributing drainage areas are delineated by USGS data chiefs in each state on the basis of a knowledge of local conditions. Therefore, contributing drainage areas are used to interpret streamflow statistics, which are normalized by drainage area for interbasin comparison or extrapolation.

Results of the site-selection process are summarized in table 3. The median record lengths ranged from 28 years (in ecoregions 14, 26, and 79) to 44 years (in ecoregions 16, 28, 49, 52, 54, 66, and 82) (table 3). The spatial density of streamgages is higher in the more populous areas (fig. 7). There are relatively few streamgages in arid areas (for example, in ecoregions 12—the Snake River Basin/High Desert; 14—Southern Basin and Range; 24—Southern Deserts; and 79—Madrean Archipelago), rugged areas (for example, in ecoregions 38—Boston Mountains; and 41—Canadian Rockies), and areas that are predominantly wetlands (for example, in ecoregions: 49—Northern Minnesota Wetlands and 76—Southern Florida Coastal Plain) (table 3). The median drainage areas for streamgages in each ecoregion ranged from 35 mi^2 (in ecoregion 84) to 430 mi^2 (in ecoregion 49). The sum of monitored areas (gaged drainage areas) in each ecoregion ranged from 146 mi^2 (in ecoregion 76) to 19,604 mi^2 (in ecoregion 65). The sum of all 2,783 gaged drainage areas is about 461,000 mi^2, which is about 15 percent of the total land area of the conterminous United States. The actual percentage of total gaged drainage area, however, is probably less than 15 percent because one or more streamgages may be nested within the drainage area of another streamgage. In general, table 3 indicates that the selected streamgages may provide a robust representation of streamflow statistics in many areas of the conterminous United States.

Statistical Characterization of Streamflow

SELDM uses statistics that indicate the magnitude and variability of streamflow measurements at a site of interest. Statistics for the entire dataset and for nonzero streamflows were calculated for each streamgage by the QSTATS software (Granato, 2009). Statistics for the entire dataset include the proportion of zero flows, arithmetic average, standard deviation, skew, and median of all streamflows. Statistics for nonzero streamflows include the mean, standard deviation, skew, and median of the logarithms of the data. Additional statistics for specific values of interest in each dataset include the seven-day ten-year low flow (7Q10), the one-day three-year biological low flow (1B3), and the four-day three-year biological low flow (4B3). These low-flow statistics were calculated for each streamgage by the DFLOW3 software (Rossman, 1990a,b; U.S. Environmental Agency, 2004). Detailed information and statistics for each streamgage and each ecoregion are documented in the SWQDM database (Granato and others, 2009).

For an initial planning-level estimate, it is assumed that the median of the proportions of zero flows and the medians of the statistics for the logarithms of nonzero streamflows in each ecoregion can be used to model prestorm streamflows (table 4). The median of the proportions of zero flows and of the means, standard deviations, and skews of the logarithms of nonzero streamflow can be used for stochastic data generation of streamflows if conditional probability methods and a skew-adjusted frequency factor are used (Haan, 1977; Chow and others, 1988; Stedinger and others, 1993). Conditional probability methods may be used to adjust streamflow statistics to account for the proportion of zero streamflows in areas with a substantial number of zero-flow values. Streamflow statistics were calculated for each streamgage

Table 3. Summary information for selected U.S. Geological Survey streamgages within each of the 84 U.S. Environmental Protection Agency (2003) Level III nutrient ecoregions.

[The 2,783 gaging stations selected for this analysis have total drainage areas ranging from 10 to 500 square miles, contributing drainage areas ranging from 7.91 to 500 square miles, and at least 24 years of streamflow data collected during the period 1960–2003. No., number; mi^2, square miles. Monitored areas have not been adjusted for the potential nesting of areas. Record lengths are rounded to the nearest whole number of years. Ecoregions are identified on the plate *useco.pdf* on the CD–ROM accompanying this report]

Ecoregion		Number of stations	Median drainage area (mi^2)	Sum of monitored areas (mi^2)	Median record length, in years	Sum of record lengths, in years
No.	Name					
1	Coast Range	35	130	5,561	43	1,349
2	Puget Lowland	28	86	3,709	40	1,047
3	Willamette Valley	15	125	2,329	39	544
4	Cascades	72	108	10,577	42	2,732
5	Sierra Nevada	86	67	10,130	42	3,290
6	Southern and Central California Plains and Hills	113	92	15,099	37	4,051
7	Central California Valley	8	74	885	35	282
8	Southern California Mountains	17	36	1,444	43	666
9	Eastern Cascades Slopes and Foothills	11	79	1,815	32	380
10	Columbia Plateau	17	173	3,128	30	570
11	Blue Mountains	15	168	3,088	36	532
12	Snake River Basin/High Desert	3	335	988	40	120
13	Northern Basin and Range	30	72	4,121	36	1,074
14	Southern Basin and Range	6	63	465	28	192
15	Northern Rockies	11	170	1,916	38	417
16	Montana Valley and Foothill Prairies	14	178	2,921	44	555
17	Middle Rockies	48	122	8,196	39	1,785
18	Wyoming Basin	29	100	4,067	34	1,012
19	Wasatch and Uinta Mountains	40	61	3,731	39	1,492
20	Colorado Plateaus	25	105	3,969	37	905
21	Southern Rockies	114	90	14,427	39	4,242
22	Arizona/New Mexico Plateau	22	100	3,764	40	831
23	Arizona/New Mexico Mountains	19	120	3,477	38	724
24	Southern Deserts	5	285	1,564	39	191
25	Western High Plains	15	313	4,070	35	541
26	Southwestern Tablelands	14	273	3,934	28	451
27	Central Great Plains	35	307	10,036	38	1,283
28	Flint Hills	6	225	1,450	44	254
29	Central Oklahoma/Texas Plains	31	280	8,221	37	1,122
30	Edwards Plateau	13	116	1,986	38	493
31	Southern Texas Plains	7	241	1,941	43	285
32	Texas Blackland Prairies	28	132	4,491	41	1,076
33	East Central Texas Plains	11	239	2,751	36	399
34	Western Gulf Coastal Plains	29	88	3,595	33	987
35	South Central Plains	40	179	8,607	38	1,419
36	Ouachita Mountains	7	203	1,489	38	276
37	Arkansas Valley	4	205	983	38	149
38	Boston Mountains	6	288	1,558	36	214
39	Ozark Highlands	24	242	5,963	38	926
40	Central Irregular Plains	38	229	9,384	41	1,466
41	Canadian Rockies	3	65	371	42	114
42	Northwestern Glaciated Plains	12	235	2,561	30	380

Table 3. Summary information for selected U.S. Geological Survey streamgages within each of the 84 U.S. Environmental Protection Agency (2003) Level III nutrient ecoregions.—Continued

[The 2,783 gaging stations selected for this analysis have total drainage areas ranging from 10 to 500 square miles, contributing drainage areas ranging from 7.91 to 500 square miles, and at least 24 years of streamflow data collected during the period 1960–2003. No., number; mi², square miles. Monitored areas have not been adjusted for the potential nesting of areas. Record lengths are rounded to the nearest whole number of years. Ecoregions are identified on the plate *useco.pdf* on the CD–ROM accompanying this report]

Ecoregion		Number of stations	Median drainage area (mi^2)	Sum of monitored areas (mi^2)	Median record length, in years	Sum of record lengths, in years
No.	Name					
43	Northwestern Great Plains	27	148	4,865	38	995
44	Nebraska Sandhills	3	200	848	35	105
45	Piedmont	112	108	16,253	42	4,248
46	Northern Glaciated Plains	22	187	4,840	34	741
47	Western Corn Belt Plains	56	203	12,554	40	2,139
48	Lake Agassiz Plain	12	255	3,139	42	477
49	Northern Minnesota Wetlands	1	430	430	44	44
50	Northern Lakes and Forests	45	183	8,980	37	1,649
51	Northern Central Hardwood Forests	14	200	3,007	34	489
52	Driftless Area	13	142	1,985	44	501
53	Southeastern Wisconsin Till Plains	16	101	2,031	40	585
54	Central Corn Belt Plains	71	104	10,380	44	2,826
55	Eastern Corn Belt Plains	87	178	17,027	42	3,353
56	Southern Michigan/Northern Indiana Drift Plains	70	100	10,109	41	2,730
57	Huron/Erie Lake Plains	11	390	3,647	38	393
58	Northeastern Highlands	107	126	17,973	43	4,123
59	Northeastern Coastal Zone	79	64	8,328	43	3,082
60	Northern Appalachian Plateau and Uplands	31	108	5,040	43	1,198
61	Erie/Ontario Lake Hills and Plain	21	151	3,671	41	807
62	North Central Appalachians	35	136	5,884	43	1,407
63	Middle Atlantic Coastal Plain	23	75	2,757	42	875
64	Northern Piedmont	94	59	9,179	43	3,643
65	Southeastern Plains	100	175	19,604	36	3,570
66	Blue Ridge Mountains	35	104	4,805	44	1,414
67	Central Appalachian Ridges and Valleys	105	134	17,706	43	4,125
68	Southwestern Appalachians	15	199	3,531	34	493
69	Central Appalachians	54	159	10,354	42	2,074
70	Western Allegheny Plateau	54	177	11,546	43	2,168
71	Interior Plateau	59	170	10,671	35	2,126
72	Interior River Lowland	33	148	5,714	38	1,242
73	Mississippi Alluvial Plain	7	270	2,072	43	284
74	Mississippi Valley Loess Plains	7	180	1,288	37	263
75	Southern Coastal Plain	93	107	12,721	40	3,498
76	Southern Florida Coastal Plain	1	146	146	33	33
77	North Cascades	15	103	1,810	43	583
78	Klamath Mountains	34	146	6,364	35	1,198
79	Madrean Archipelago	3	79	581	28	85
80	Northern Basin and Range	13	180	2,548	34	471
81	Sonoran Basin and Range	11	36	783	41	426
82	Laurentian Plains and Hills	11	227	2,609	44	450
83	Eastern Great Lakes and Hudson Lowlands	43	113	6,501	38	1,609
84	Atlantic Coastal Pine Barrens	34	35	1,575	42	1,315

Table 4. Medians of selected streamflow statistics for the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions in the conterminous United States calculated by using daily mean streamflow data from 2,783 selected U.S. Geological Survey streamgages for the period 1960–2003. Statistics include the proportion of zero flows and the median, mean, standard deviation, and skew of the logarithms of nonzero mean daily streamflow measurements.

[No., number; $\mathrm{ft^3/s/mi^2}$, cubic foot per second per square mile; SD, standard deviation]

Ecoregion		Median of streamflow statistics for each ecoregion				
		Days with zero discharge (percent of total record)	Statistics for the common logarithms of nonzero discharges			
No.	Name		Median ($\mathrm{ft^3/s/mi^2}$)	Geometric mean ($\mathrm{ft^3/s/mi^2}$)	Geometric SD (dimensionless)	Coefficient of skew (dimensionless)
1	Coast Range	0.00	1.88	1.77	3.99	0.11
2	Puget Lowland	0.00	2.19	2.14	2.37	0.26
3	Willamette Valley	0.00	1.41	1.21	4.22	-0.03
4	Cascades	0.00	3.03	2.72	2.59	0.06
5	Sierra Nevada	0.00	0.42	0.50	3.92	0.41
6	Southern and Central California Plains and Hills	5.46	0.08	0.08	7.18	0.19
7	Central California Valley	40.59	0.09	0.08	7.46	-0.15
8	Southern California Mountains	15.32	0.04	0.04	7.20	0.02
9	Eastern Cascades Slopes and Foothills	0.00	1.60	1.67	2.25	0.43
10	Columbia Plateau	0.00	0.25	0.22	4.90	0.06
11	Blue Mountains	0.00	0.46	0.64	2.91	0.50
12	Snake River Basin/High Desert	45.56	0.06	0.06	5.00	-0.19
13	Northern Basin and Range	0.00	0.24	0.23	2.75	0.75
14	Southern Basin and Range	0.02	0.07	0.07	4.50	0.20
15	Northern Rockies	0.00	0.55	0.72	3.02	0.58
16	Montana Valley and Foothill Prairies	0.00	0.34	0.46	2.76	0.91
17	Middle Rockies	0.00	0.38	0.39	2.91	0.63
18	Wyoming Basin	0.00	0.23	0.28	3.61	0.70
19	Wasatch and Uinta Mountains	0.00	0.29	0.37	2.83	0.87
20	Colorado Plateaus	0.00	0.20	0.20	2.75	0.70
21	Southern Rockies	0.00	0.29	0.39	3.13	0.72
22	Arizona/New Mexico Plateau	0.00	0.11	0.13	2.71	0.64
23	Arizona/New Mexico Mountains	0.00	0.07	0.08	2.92	0.82
24	Southern Deserts	98.32	0.03	0.03	18.40	0.08
25	Western High Plains	0.00	0.15	0.10	2.49	0.37
26	Southwestern Tablelands	0.00	0.03	0.03	3.32	0.25
27	Central Great Plains	0.97	0.03	0.03	5.34	0.18
28	Flint Hills	0.40	0.12	0.09	5.60	0.07
29	Central Oklahoma/Texas Plains	10.20	0.03	0.04	8.06	0.31
30	Edwards Plateau	2.39	0.14	0.14	4.86	-0.31
31	Southern Texas Plains	58.27	0.01	0.01	8.29	0.48
32	Texas Blackland Prairies	8.77	0.06	0.06	9.07	0.10
33	East Central Texas Plains	5.90	0.05	0.05	8.26	0.00
34	Western Gulf Coastal Plains	0.00	0.15	0.20	5.08	0.56
35	South Central Plains	0.22	0.23	0.23	6.75	-0.01
36	Ouachita Mountains	2.67	0.34	0.23	9.16	-0.29
37	Arkansas Valley	2.82	0.28	0.20	10.17	-0.54
38	Boston Mountains	1.65	0.43	0.30	8.20	-0.65
39	Ozark Highlands	0.00	0.42	0.43	3.39	0.35
40	Central Irregular Plains	0.58	0.10	0.10	7.52	0.06
41	Canadian Rockies	0.02	2.07	1.30	3.57	0.07
42	Northwestern Glaciated Plains	0.35	0.18	0.21	3.38	0.16

Table 4. Medians of selected streamflow statistics for the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions in the conterminous United States calculated by using daily mean streamflow data from 2,783 selected U.S. Geological Survey streamgages for the period 1960–2003. Statistics include the proportion of zero flows and the median, mean, standard deviation, and skew of the logarithms of nonzero mean daily streamflow measurements.—Continued

[No., number; $ft^3/s/mi^2$, cubic foot per second per square mile; SD, standard deviation]

Ecoregion		Median of streamflow statistics for each ecoregion				
			Statistics for the common logarithms of nonzero discharges			
No.	Name	Days with zero discharge (percent of total record)	Median ($ft^3/s/mi^2$)	Geometric mean ($ft^3/s/mi^2$)	Geometric SD (dimensionless)	Coefficient of skew (dimensionless)
43	Northwestern Great Plains	0.02	0.11	0.12	3.05	0.68
44	Nebraska Sandhills	0.00	0.21	0.24	1.75	1.76
45	Piedmont	0.00	0.65	0.66	2.73	0.20
46	Northern Glaciated Plains	32.98	0.02	0.02	10.81	0.13
47	Western Corn Belt Plains	0.02	0.20	0.19	4.27	0.01
48	Lake Agassiz Plain	3.96	0.03	0.03	6.19	0.07
49	Northern Minnesota Wetlands	1.25	0.05	0.05	9.19	-0.38
50	Northern Lakes and Forests	0.00	0.65	0.69	2.13	0.74
51	Northern Central Hardwood Forests	0.00	0.31	0.34	2.65	0.60
52	Driftless Area	0.00	0.45	0.49	1.88	1.27
53	Southeastern Wisconsin Till Plains	0.00	0.37	0.39	2.84	0.36
54	Central Corn Belt Plains	0.00	0.40	0.39	3.65	0.04
55	Eastern Corn Belt Plains	0.00	0.35	0.34	4.25	0.24
56	Southern Michigan/Northern Indiana Drift Plains	0.00	0.56	0.53	2.41	0.13
57	Huron/Erie Lake Plains	0.00	0.21	0.25	4.13	0.36
58	Northeastern Highlands	0.00	1.06	1.09	2.90	0.09
59	Northeastern Coastal Zone	0.00	1.13	1.02	2.90	-0.16
60	Northern Appalachian Plateau and Uplands	0.00	0.69	0.68	3.20	0.11
61	Erie/Ontario Lake Hills and Plain	0.00	0.62	0.65	3.20	0.19
62	North Central Appalachians	0.00	1.07	1.03	2.98	-0.07
63	Middle Atlantic Coastal Plain	0.00	0.62	0.51	3.85	-0.00
64	Northern Piedmont	0.00	0.76	0.78	2.58	0.31
65	Southeastern Plains	0.00	0.65	0.64	3.04	0.28
66	Blue Ridge Mountains	0.00	1.86	1.84	1.99	0.36
67	Central Appalachian Ridges and Valleys	0.00	0.73	0.75	2.73	0.39
68	Southwestern Appalachians	0.00	0.80	0.67	4.87	0.02
69	Central Appalachians	0.00	0.89	0.78	3.63	-0.05
70	Western Allegheny Plateau	0.00	0.53	0.54	4.08	-0.18
71	Interior Plateau	0.00	0.48	0.48	4.70	0.02
72	Interior River Lowland	0.55	0.16	0.16	7.72	0.02
73	Mississippi Alluvial Plain	0.00	0.60	0.64	4.32	0.09
74	Mississippi Valley Loess Plains	0.00	0.48	0.64	2.73	1.52
75	Southern Coastal Plain	0.20	0.30	0.29	4.55	-0.20
76	Southern Florida Coastal Plain	58.48	3.38	2.40	4.43	-1.44
77	North Cascades	0.00	4.90	4.67	2.44	0.11
78	Klamath Mountains	0.00	0.72	0.80	3.68	0.22
79	Madrean Archipelago	3.57	0.004	0.004	3.62	0.32
80	Northern Basin and Range	0.00	0.06	0.10	3.37	0.40
81	Sonoran Basin and Range	27.52	0.10	0.08	8.10	-0.11
82	Laurentian Plains and Hills	0.00	1.12	1.07	2.86	0.16
83	Eastern Great Lakes and Hudson Lowlands	0.00	0.75	0.71	3.04	0.18
84	Atlantic Coastal Pine Barrens	0.00	1.01	1.04	1.96	0.11

and normalized by drainage area so that the values could be applied to basins of different size within each ecoregion. The median value of each statistic from all the streamgages in the region was selected to represent the ecoregion. These median ecoregion statistics do not necessarily characterize streamflow for any particular streamgage within each ecoregion, but the use of the median of each statistic for each ecoregion was expected to produce a population of streamflow values that would represent a typical basin within that ecoregion. The streamflow statistics in table 4 were based on the data available for the period 1960–2003 without adjusting for potential trends in the dataset. Trends may be considered for a more detailed analysis (appendix 2), but these period-of-record statistics were considered sufficient for planning-level estimates of prestorm streamflows for water-quality analysis.

The use of regional statistics to estimate prestorm streamflows at an ungaged site depends on the assumption that ecoregions provide an effective classification scheme for identifying hydrologically similar areas within different areas of the country. The nonparametric rank-sum test (Helsel and Hirsch, 2002) was used to examine the assumption that streamflow statistics are different in neighboring Level III ecoregions. The nonparametric rank-sum test was selected to compare statistics for all the selected gages within each ecoregion because this is a distribution-free test that can be used to detect differences between two groups even if there are different numbers of observations within each sample (Helsel and Hirsch, 2002). Streamflow statistics from adjacent ecoregions in three areas of the United States were used as examples. The Northeastern and the Northwestern United States were selected to represent humid areas with different weather patterns. The Southwestern United States was selected to represent conditions in an arid area. Differences in the geometric means, geometric standard deviations, and skews of the common logarithms of nonzero streamflow values were calculated for six Level III ecoregions within each of these three areas (table 5). These ecoregions are shown on the USEPA ecoregion map on the CD–ROM accompanying this report (U.S. Environmental Protection Agency, 2003).

If streamflow statistics among neighboring ecoregions are drawn from the same population, then the use of ecoregions may not provide a quantitative distinguishing factor that can be used to indicate areas of hydrologic similarity. The values in table 5 generated by the nonparametric rank-sum test represent the probabilities that the streamflow statistics for gages in each ecoregion are drawn from the same population. The grey-shaded values in the table indicate differences that are statistically significant in a two-sided test with a 95-percent confidence interval (a 5-percent *p*-value). Within each area, medians for each statistic for each of the six ecoregions were paired with the respective medians for each of the other five ecoregions. In total, 45 tests were done within each area of the nation. Overall, the results in table 5 indicate significant differences in statistics among nearby ecoregions for 42 percent of tests for the geometric means, 42 percent of tests for the geometric standard deviation, and 56 percent of tests for the geometric coefficient of skew. These differences in the statistics for streamgages in nearby ecoregions indicate that these ecoregions may be used to indicate hydrologic similarity within larger areas with similar climates.

Distinctions among neighboring regions may depend on methods used for stochastic data generation. If it is assumed that streamflows can be approximated by a two-parameter lognormal distribution, which is quantified by the geometric mean and standard deviation, then 67 percent of the combined tests indicate significant differences in one or the other of these statistics among nearby ecoregions. If it is assumed that streamflows can be approximated by a three parameter log-Pearson Type III distribution, which is quantified by the geometric mean, standard deviation, and skew, then 84 percent of the combined tests indicate significant differences in one of the three of these statistics among nearby ecoregions. For example, in the Northeastern area, the geometric means and standard deviations for flows at streamgages in ecoregion 82 are not significantly different from statistics in ecoregions 58 or 59 (table 5). However, the geometric means of 58 and ecoregion 59 are significantly different. Also, the geometric coefficients of skew differ significantly between ecoregion 58 and ecoregion 59 and between ecoregions 82 and ecoregion 59.

The results of these tests indicate that ecoregion delineation has some discerning power for identifying hydrologically similar sites within broad geographic regions of the United States that have somewhat similar climates. This is not to say, however, that hydrologically similar conditions do not exist for streamgages near the borders of neighboring ecoregions, or that streamflow statistics in neighboring regions are substantially different in magnitude. In fact, streamflow statistics in neighboring ecoregions are commonly of similar magnitude, as is indicated by the medians of streamflow statistics by ecoregion (table 4). Furthermore, statistics can vary considerably within an ecoregion. For example, figure 11 indicates variability in the normalized geometric means of nonzero streamflows both within and between ecoregions. The greatest within-ecoregion variability in the normalized geometric mean occurs in relatively arid ecoregions, which commonly have fewer streamgages.

SELDM uses the log-Pearson Type III distribution to stochastically generate a population of prestorm flows with statistics for the common logarithms of daily mean flow values. Several distributions commonly are used for modeling streamflow values in statistical surface-water-quality models (appendix 2). The lognormal distribution is most commonly used for highway- and urban-runoff studies. The log-Pearson Type III distribution is equivalent to the lognormal distribution if the logarithms of streamflow have zero skew, but is more flexible because it can represent nonzero skew in the logarithms of a streamflow dataset (Haan, 1997; Chow and others, 1988; Bobee and Ashkar, 1991).

The log-Pearson Type III distribution was selected for use with SELDM because data from many streamgages have nonzero coefficients of skew (table 4, fig. 12). These

Table 5. Results of the rank-sum test (Helsel and Hirsch, 2002) indicating the probability that streamflow statistics for different U.S. Environmental Protection Agency Level III nutrient ecoregions are drawn from the same population with examples from three areas of the conterminous United States.

[Streamflow statistics include measures of location (geometric mean of daily streamflows per unit area), spread (geometric standard deviation), and skewness (coefficient of skew of the log-transformed populations). No., number; <, less than; --, same ecoregion; gray boxes indicate values that are considered statistically significant with a probability of less than five percent (0.05) that the samples from different ecoregions are from the same population]

			Probability of obtaining the test statistic if the populations of geometric means of daily flows (per unit area) are the same for two ecoregions					Probability of obtaining the test statistic if the populations of geometric standard deviations of daily flows are the same for two ecoregions					Probability of obtaining the test statistic if the populations of geometric skew coefficients of daily flows are the same for two ecoregions				
								Northeastern United States									
Ecoregion		**Number of stations**	**Ecoregion number**					**Ecoregion number**					**Ecoregion number**				
No.	**Name**		**58**	**59**	**60**	**82**	**83**	**58**	**59**	**60**	**82**	**83**	**58**	**59**	**60**	**82**	**83**
58	Northeastern Highlands	107	--					--					--				
59	Northeastern Coastal Zone	79	0.035	--				0.856	--				<0.001	--			
60	Northern Appalachian Plateau and Uplands	31	<0.001	<0.001	--			0.011	0.008	--			0.713	<0.001	--		
82	Laurentian Plains and Hills	11	0.643	0.693	0.001	--		0.523	0.693	0.037	--		0.683	0.015	0.423	--	
83	Eastern Great Lakes and Hudson Lowlands	43	<0.001	<0.001	1.000	0.001	--	0.188	0.164	0.483	0.150	--	0.315	<0.001	0.859	0.764	--
84	Atlantic Coastal Pine Barrens	34	0.189	0.019	0.031	0.240	0.032	<0.001	<0.001	<0.001	0.006	<0.001	0.468	0.001	0.424	0.926	0.700
								Southwestern United States									
Ecoregion		**Number of stations**	**Ecoregion number**					**Ecoregion number**					**Ecoregion number**				
No.	**Name**		**6**	**8**	**20**	**22**	**23**	**6**	**8**	**20**	**22**	**23**	**6**	**8**	**20**	**22**	**23**
6	Southern and Central California Plains and Hills	113	--					--					--				
8	Southern California Mountains	17	0.105	--				0.761	--				0.649	--			
20	Colorado Plateaus	25	0.007	0.001	--			<0.001	<0.001	--			0.013	0.023	--		
22	Arizona/New Mexico Plateau	22	0.115	0.217	0.190	--		<0.001	<0.001	0.543	--		0.009	0.026	0.806	--	
23	Arizona/New Mexico Mountains	19	0.641	0.397	0.005	0.043	--	<0.001	0.002	0.192	0.441	--	0.005	0.019	0.394	0.327	--
81	Sonoran Basin and Range	11	0.937	0.190	0.149	0.390	0.796	0.902	0.925	<0.001	0.004	0.039	0.090	0.132	0.021	0.007	0.009
								Northwestern United States									
Ecoregion		**Number of stations**	**Ecoregion number**					**Ecoregion number**					**Ecoregion number**				
No.	**Name**		**1**	**2**	**4**	**10**	**15**	**1**	**2**	**4**	**10**	**15**	**1**	**2**	**4**	**10**	**15**
1	Coast Range	35	--					--					--				
2	Puget Lowland	28	0.403	--				<0.001	--				0.130	--			
4	Cascades	72	0.104	0.227	--			<0.001	0.207	--			0.126	0.004	--		
10	Columbia Plateau	17	<0.001	<0.001	<0.001	--		0.938	0.012	0.021	--		0.258	0.109	0.514	--	
15	Northern Rockies	11	<0.001	<0.001	<0.001	0.030	--	0.029	0.001	0.165	0.300	--	0.005	0.048	<0.001	0.010	--
77	North Cascades	15	0.021	0.018	0.007	<0.001	<0.001	<0.001	0.467	0.657	0.045	0.011	0.582	0.150	0.707	0.497	0.038

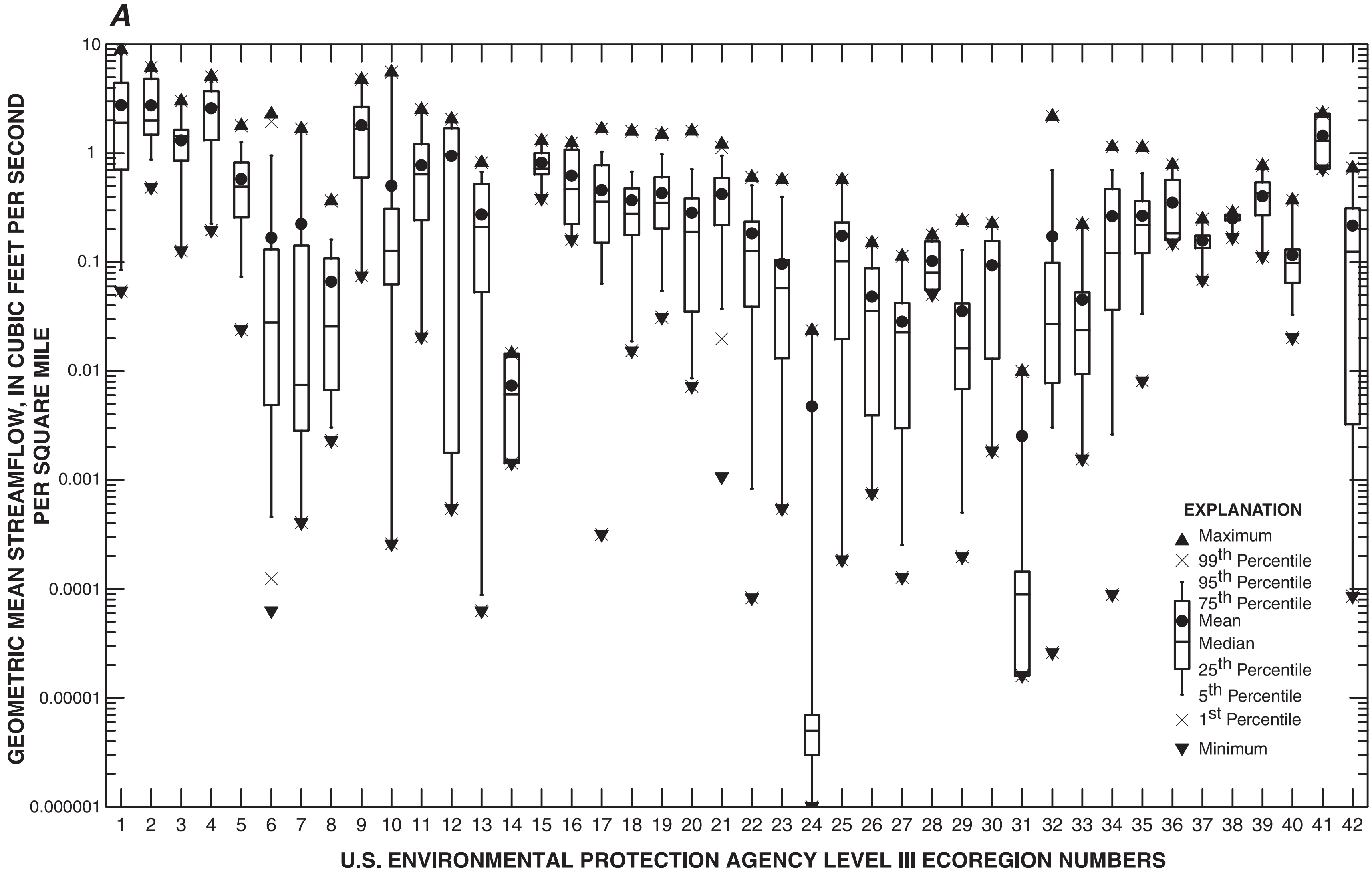


Figure 11. Populations of normalized geometric-mean streamflow values for different streamgages within each U.S. Environmental Protection Agency (2003) Level III ecoregion in the conterminous United States. See table 3 for a list of names of the numbered ecoregions. Ecoregions are shown on the map useco.pdf on the CD–ROM accompanying this report.

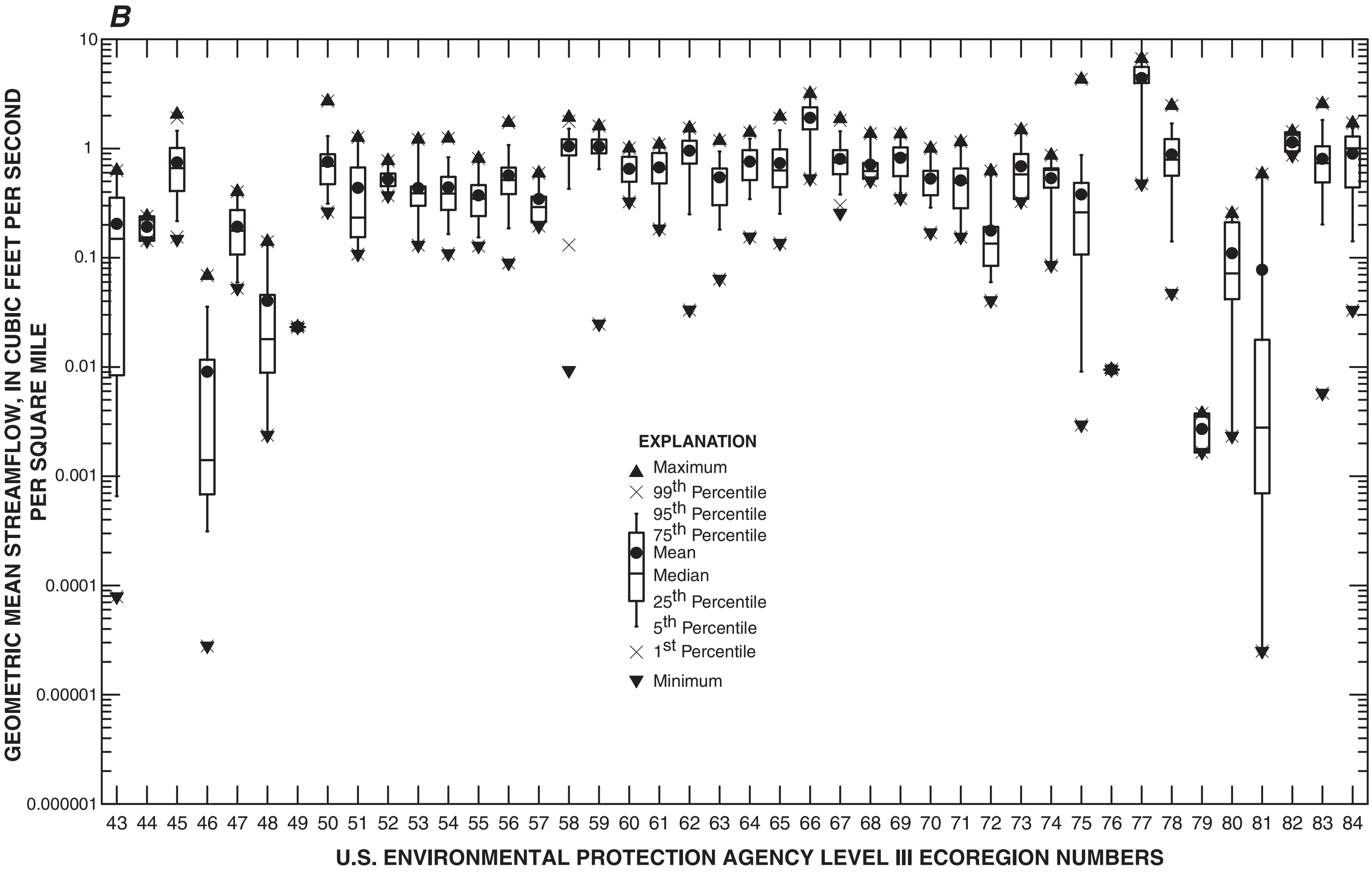


Figure 11. Populations of normalized geometric-mean streamflow values for different streamgages within each U.S. Environmental Protection Agency (2003) Level III ecoregion in the conterminous United States. See table 3 for a list of names of the numbered ecoregions. Ecoregions are shown on the map useco.pdf on the CD–ROM accompanying this report.—Continued

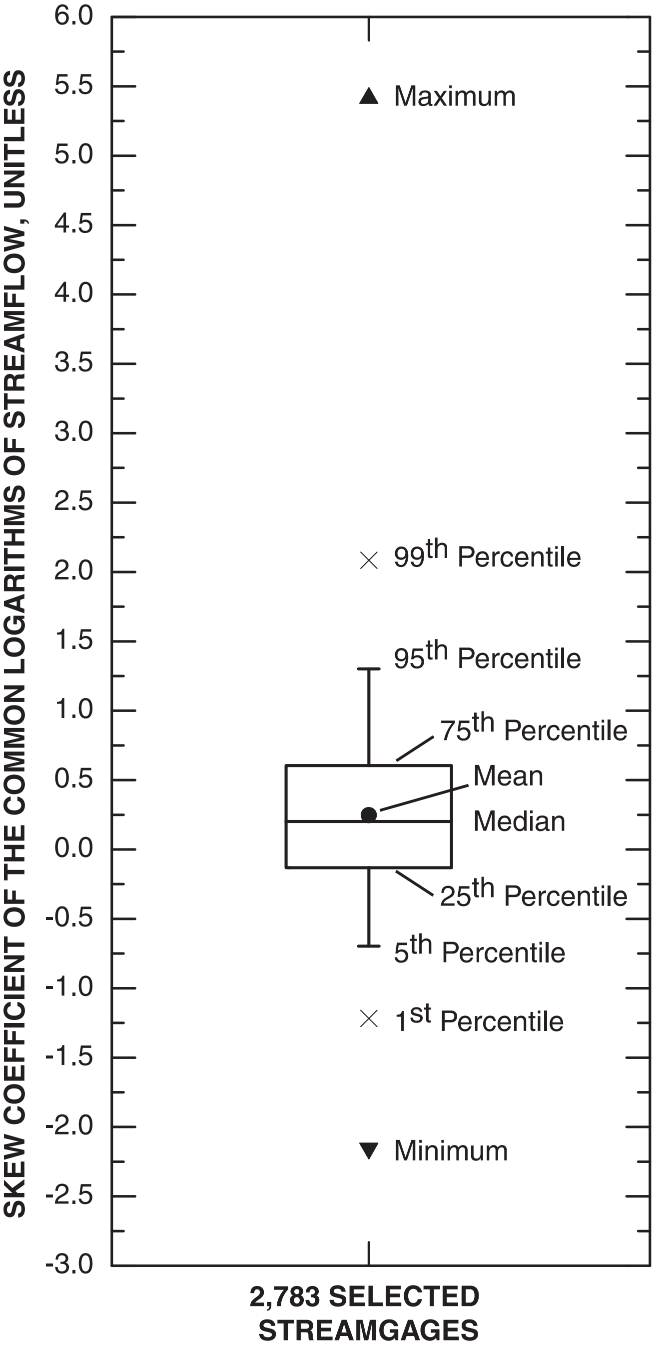


Figure 12. Skew coefficients of the common logarithms of average daily streamflow values for the 2,783 selected streamgages in the conterminous United States.

coefficients of skew range from -2.2 to 5.4 for the selected streamgages. The average coefficient of skew is about 0.25, and the median coefficient of skew is about 0.2 (fig. 12). Press and others (1992) indicate that the standard deviation of the coefficients of skew of subsamples from a normal distribution is calculated by dividing the number six by the sample size and taking the square root of the fraction. If the population of streamflows at each streamgage in the United States is lognormally distributed, then 95 percent of the calculated skew coefficients of the common logarithms of streamflow should fall within the interval of ±0.057 if the record lengths are greater than 20 years (7,305 daily mean values in this study). The data from only about 9 percent of the 2,783 streamgages in this study, however, met this criterion for lognormality; 90 percent of the calculated skew coefficients are within the interval -0.7 to 1.3, and 95 percent are within an interval of -0.88 to 1.59 (fig. 12). These logarithmic skew coefficients, however, are well within the range (-9 to 9) that can be approximated by log-Pearson Type III frequency factors (appendix 2).

The ecoregion coefficients of skew in table 4 are less variable than the coefficients of skew in figure 12 because the values in table 4 are the medians of all streamgages in each ecoregion. The median geometric skew coefficients values for streamgages by ecoregion range from -1.44 (for ecoregion 76) to 1.76 (for ecoregion 44) with a median of 0.17 among ecoregions (table 4). Median geometric skew values were equal to zero for 2 ecoregions and were within the plus-or-minus 0.057 range for another 9 ecoregions. Thus, daily mean streamflow statistics for 11 of the 84 ecoregions could be approximated by a two-parameter lognormal distribution. The log-Pearson Type III distribution, however, may be used for all ecoregions because it reverts to the lognormal distribution as the coefficient of skew approaches zero.

SELDM uses the frequency factor method (appendix 2) to generate the population of daily mean prestorm flow values. The frequency factor method uses the mean and the standard deviation from a data sample to predict values from the underlying population by use of a distribution-specific frequency factor. For a lognormal or log-Pearson Type III distribution, the equation for the frequency factor method is

$$X_i = X_m + S \times K_i \tag{1}$$

where

X_i is the value of the logarithm of the i^{th} streamflow value,
X_m is the geometric mean of the set of streamflow values (in logarithmic space),
S is the standard deviation of the logarithms of the streamflow values, and
K_i is the distribution-specific frequency factor.

The distribution-specific frequency factor relates the probability of occurrence of a value to a multiple of the standard deviation above or below the mean value. The frequency factor equals the standard normal variate if the skew value of the population is zero and an adjusted log-Pearson Type III variate if nonzero skews are modeled. Several methods can be used to generate log-Pearson Type III frequency factors, some of which are well suited for manual calculations.

SELDM uses the modified Wilson-Hilferty algorithm developed by Kirby (1972) to generate log-Pearson Type III frequency factors. This algorithm was selected because it was designed for numerical implementation and it provides acceptable estimates of log-Pearson Type III frequency factors for samples with coefficients of skews within the range of about ±9.

SELDM uses conditional probability methods to account for the occurrence of prestorm flows equal to zero, which cannot be modeled using the logarithms of streamflows (appendix 2). With conditional probability methods, the stochastic data-generation algorithm for prestorm flows must account for the probability of zero flows and the entire sample space for the logarithms of nonzero flows (William Kirby, U.S. Geological Survey Office of Surface Water, written commun., 2005). In SELDM, a uniform random number between zero and one is generated to represent the total probability (plotting position) of the zero and nonzero prestorm flow for each storm event. If this number is less than or equal to the proportion of zero flows, then a prestorm streamflow value of zero is assigned for that storm event. If this number is greater than the proportion of zero flows, then the uniform random number is rescaled to generate a frequency factor that represents the prestorm flow within the probability distribution of the nonzero streamflows. The result for all storm events will be a stochastic sample of prestorm streamflows that approximates the proportion of zero flows and the statistics of the logarithms of nonzero flows for the site (or ecoregion) of interest.

Estimating Streamflow Statistics for Ungaged Sites

Streamflow statistics are needed to model planning-level estimates of prestorm streamflows for water-quality analysis at sites without streamflow data. Several methods, including the use of runoff maps, regression-on-basin characteristics, and drainage-area ratios (appendix 2) were considered for estimating long-term streamflow statistics at sites without monitoring data. The drainage-area-ratio method was selected because this method can be used for generating planning-level estimates of prestorm streamflows on the basis of available statistics and the drainage-basin area for the site of interest. The default option in SELDM is to use the drainage-area-ratio method with the average or median of each statistic by ecoregion. The method may also be used with statistics from nearby hydrologically similar sites or user-defined statistics estimated on the basis of some other method.

The drainage-area-ratio method is used to calculate streamflow values at a site of interest from streamflows measured at one or more hydrologically similar index sites. The assumption of hydrologic similarity is implicit in the application of the drainage-area-ratio method because basin characteristics are not explicitly included in the predictive equation. Natural factors (such as orographic effects, variations in soils, and geology) and anthropogenic factors (such as imperviousness and water use) should be considered in assessments of hydrologic similarity. In this study, ecoregions are used as an initial approximation for hydrologic similarity. The general equation for the drainage-area-ratio method is

$$Q_y = (A_y/A_x)^Z \times Q_x \,, \tag{2}$$

where

Q_y is the estimated streamflow at the site of interest,
Q_x is the streamflow at the index site,
A_y is the drainage area for the site of interest,
A_x is the drainage area for the index site, and
Z is an exponent to adjust for systematic differences in the ratio of drainage-area to flow.

Despite some potential limitations (appendix 2), the drainage-area-ratio method will be used for stochastic analysis of potential effects of highway runoff on receiving waters at sites without available streamflow data. The default method for generating prestorm flows for the site of interest using the drainage-area-ratio method is based on the assumption that the exponent Z is one. If paired sites are used, the geometric mean streamflow at the index site (in ft^3/s) and the drainage areas of both sites are used in equation 2, and the result is used in equation 1 to generate a population of streamflows for the site of interest. If regional values are used, however, then equation 2 is rearranged as

$$Q_y = A_y \times (Q_x/A_x). \tag{3}$$

In this case, the drainage area of the site of interest is multiplied by the normalized geometric mean streamflow for the ecoregion (Q_x/A_x in $\mathrm{ft}^3/\mathrm{s}/\mathrm{mi}^2$) in equation 3, and the result is used in equation 1 with regional values of the geometric standard deviation and coefficient of skew to generate a population of streamflows for the site of interest. With these assumptions, the SELDM user needs only the drainage area and the location (latitude and longitude) of the basin of interest to estimate streamflow statistics at any site in the conterminous United States. In some cases, however, use of the drainage-area-ratio method may be affected by uncertainties in the drainage area, possibly as a result of differences between nominal surface-water and groundwater drainage divides (appendix 1).

Use of the drainage-area-ratio method with mean or median streamflow statistics may provide sufficient planning-level estimates in ecoregions with small variations in streamflow statistics, but such estimates may need further refinement in ecoregions with large variations in streamflow statistics (fig. 11). These ecoregions are characterized by arid climate, large topographic variation, or large area (table 3; fig. 11). These characteristics may reduce the hydrologic similarity among basins within these ecoregions. If variations in statistics for different basins within an ecoregion are small, these statistics may represent streamflow populations at an ungaged site. If variations are large, however, more refined estimates may be needed. For example, runoff maps (for example, Gebert and others, 1987) can be used to select a streamgage or streamgages that are close to the site of interest in a drainage basin with similar annual average runoff values. Theoretically, such drainage basins should be more hydrologically similar to the site of interest than the entire ecoregion and, for this reason, could be used with with the drainage-area-ratio method. However, patterns of runoff on these national-scale maps are similar to the ecoregion delineations in many areas of the country.

The potential for variations in flow statistics that are a function of drainage area also may account for some variations in streamflow statistics within each ecoregion. Regression of the logarithms of geometric mean streamflows in cubic feet per second as a function of basin drainage areas in square miles indicates potential variations in geometric mean streamflow as a function of the logarithm of basin drainage area by ecoregion. The slope of the regression analysis also indicates whether the exponent (Z) of the drainage-area ratio in equation 2 is substantially different from one, and if the difference is statistically significant. Table 6 shows nonparametric regression-equation statistics for the mean, standard deviation, and skew of the logarithms of nonzero streamflows for streamgages in each ecoregion. The regression equations in this table may be applied within the range of drainage areas that were used to develop the equations. The drainage-area-ratio exponents range from about -0.08 to about 4.3 for all ecoregions and have a median of about 0.99 and an interquartile range of about 0.24. About 50 percent of the estimated exponents are substantially different from one. Examination of regression statistics, however, indicates that an exponent of one is outside the 90-percent confidence interval for only five of the 82 ecoregions (table 6). Thus, the nonparametric regression analysis of the geometric means of streamflow data indicates that the unadjusted drainage-area-ratio method may be sufficient for planning-level estimates of streamflows in most ecoregions.

Nonparametric regression by drainage area also was done to examine the dependence of the geometric standard deviation (unitless) and skew coefficients (unitless) on drainage area. For these variables, a slope of zero would indicate that the values were independent of drainage area. Slopes of the regression equations for the geometric standard deviation for each ecoregion ranged from about -0.82 to about 0.57 with a median of -0.03 and an interquartile range of about 0.13 (table 6). Regression statistics for the geometric standard deviation indicate that a slope of zero is outside the 90-percent confidence interval for only nine ecoregions. Slopes of the regression equations for the geometric coefficient of skew for each ecoregion ranged from about -1.2 to about 2 with a median of about 0.04 and an interquartile range of about 0.46 (table 6). Regression statistics for the geometric coefficient of skew indicate that a slope of zero is outside the 90-percent confidence interval for only 10 ecoregions. Thus, the nonparametric regression analysis of the geometric standard deviations and skews of streamflow data also indicates that the unadjusted drainage-area-ratio method may be sufficient for planning-level estimates of streamflows in most ecoregions.

Parametric (Pearson’s r) and nonparametric (Spearman’s rho) correlation tests (Helsel and Hirsch, 2002) also indicate that correlations between drainage area and streamflow statistics are relatively weak (table 7). Relations between geometric mean streamflow, normalized to drainage area in cubic feet per second per square mile, and drainage area in square miles are less visible in ecoregions with weaker correlation coefficients (that is, positive or negative correlation coefficients that are close to zero). A weak correlation between the geometric mean streamflow and drainage area indicates that an ecoregion-average or median geometric mean value that is normalized to area may be used to estimate streamflows from different basins in the ecoregion and that the drainage-area-ratio exponent *Z* will approach one. This is because the remaining variations in normalized geometric-mean streamflow are caused by random variations (or other unexamined variables) rather than a higher (or lower) order relation with drainage area. The geometric standard deviation and skew coefficient of the logarithms of nonzero streamflows are unitless, so a weak correlation indicates that there is little if any systematic variation with drainage area. The scatterplots in figure 13 illustrate variations in geometric mean streamflows with drainage area for ecoregions with different correlation coefficients. Low correlations between the geometric standard deviation and coefficients of skew also indicate that representative ecoregion values may be used. It should be noted, however, that these nonparametric regression equations may be used to refine estimates even if the correlation is relatively weak. If correlations are weak, nonparametric regression results will approximate the median of values for streamflows in an ecoregion because nonparametric regression results are not heavily influenced by outliers in the data (Helsel and Hirsch, 2002; Granato, 2006).

The regression statistics in table 6 may be used to refine estimates of the population of streamflows by providing estimates of streamflow statistics for basins of different sizes. Figure 14 indicates how the regional regression equations may be used with frequency factor methods (eq. 1) to estimate a population of streamflow values for different basins within each ecoregion. Statistics from a very humid area (ecoregion 2, the Puget Lowland), a humid area (ecoregion 59, the Northeastern Coastal Zone), and an arid area (ecoregion 81, the Sonoran Basin and Range) were selected as examples. A random number generator was used to provide a population of standard normal variates, which were used with estimates of the geometric mean, standard deviation, and coefficient of skew for different basin sizes in each ecoregion to provide a sample of estimated daily mean streamflow values. The resultant populations of streamflow were normalized by drainage area and plotted against flow frequency to generate flow-duration curves (fig. 14). The median absolute deviations of values above and below the regression equations for each ecoregion in table 6 and the spread of the curves in figure 14 indicate that there may be substantial variability in the location, slope, and shape of the flow-duration curves for different drainage basins within each ecoregion.

For ecoregion 2, the Puget Lowland, the slope of the regression equation for the geometric mean streamflows is greater than 1 (table 6); as a result, the estimated streamflows

Table 6. Nonparametric regression-equation statistics for calculating the mean, standard deviation, and skew of the common logarithms of nonzero streamflows from the logarithms of drainage area for streamgages in each ecoregion. These regression-equation statistics may be used to refine drainage-area ratio estimates for sites within each ecoregion.

[ft^3/s, cubic foot per second; SD, standard deviation; Log$_{10}$, common logarithms; mi^2, square miles; MAD, median absolute deviation; NPCIS, ninety-percent confidence interval for the slope; N, the slope is not different from the specified value based on the 90-percent confidence interval; Y, the slope is different from the specified value based on the 90-percent confidence interval; N1, not enough streamgages to determine a meaningful confidence interval; --, not enough streamgages to calculate statistics; S., southern; N., northern. Nonparametric regression equations were developed by using the Kendall-Theil Robust Line method (Granato, 2006)]

Ecoregion				Statistics for the common logarithms of nonzero streamflows											
No.	Name	Number of streamgages	Minimum drainage area (mi^2)	Geometric mean (ft^3/s)			NPCIS <>1	Geometric SD (unitless)			NPCIS <>0	Coefficient of skew (unitless)			NPCIS <>0
				Intercept	Slope	MAD		Intercept	Slope	MAD		Intercept	Slope	MAD	
1	Coast Range	35	11.1	-0.60449	1.41124	0.41845	N	0.61399	-6.29E-03	0.16392	N	0.25466	-0.06668	0.12342	N
2	Puget Lowland	28	12	-0.04786	1.23512	0.17319	N	0.44146	-3.42E-02	5.81E-02	N	0.64816	-0.19859	0.15553	N
3	Willamette Valley	15	26.8	0.31973	0.92189	0.11621	N	0.61054	7.17E-03	6.27E-02	N	-0.63633	0.28811	0.14736	N
4	Cascades	72	12.6	0.36498	1.00111	0.21430	N	0.69239	-0.13713	9.20E-02	Y	-0.67504	0.36324	0.19301	Y
5	Sierra Nevada	86	10.5	-0.11719	0.85219	0.26664	N	0.64730	-2.94E-02	1.06E-01	N	0.18970	0.11878	0.37685	N
6	Southern and Central California Plains and Hills	113	10.8	-1.17014	0.98197	0.44569	N	0.84828	3.87E-03	1.38E-01	N	0.53308	-0.17315	0.36451	N
7	Central California Valley	8	46.4	-2.71106	1.91877	0.49805	N	0.64670	0.11938	5.74E-02	N	1.02866	-0.63256	0.10443	N
8	Southern California Mountains	17	16.9	-0.76179	0.68495	0.27783	N	0.71122	9.42E-02	0.14500	N	0.71774	-0.45135	0.26906	N
9	Eastern Cascades Slopes and Foothills	11	21	0.65153	0.67288	0.23213	N	-0.22854	0.30610	6.39E-02	Y	1.07587	-0.34130	0.39134	N
10	Columbia Plateau	17	12	-0.23895	0.76624	0.39134	N	0.92457	-0.10475	0.23030	N	-0.74046	0.35967	0.47726	N
11	Blue Mountains	15	50.9	1.97152	-0.01567	0.25522	N	0.14943	0.14103	0.14134	N	2.83538	-1.04758	0.39920	N
12	Snake River Basin/High Desert	3	253	-9.36371	4.26805	0.32870	N1	0.30572	0.15569	1.90E-02	N1	-1.50635	0.52076	4.01E-02	N1
13	Northern Basin and Range	30	11.1	-0.56324	0.92908	0.39743	N	0.20897	0.12408	9.48E-02	N	1.79172	-0.55972	0.23764	Y
14	Southern Basin and Range	6	15.1	-0.66669	0.57633	0.46308	N	0.99422	-0.19355	0.20895	N	-1.91386	1.17717	0.58738	N
15	Northern Rockies	11	11.1	0.13001	0.82107	0.13177	N	0.53993	-2.66E-02	4.19E-02	N	0.81998	-0.10593	0.16635	N
16	Montana Valley and Foothill Prairies	14	29.1	-0.32448	1.06206	0.26465	N	0.55499	-5.07E-02	3.28E-02	N	0.39739	0.22733	0.17308	N
17	Middle Rockies	48	10.6	-0.46379	0.98204	0.36762	N	0.79270	-0.15786	0.07681	Y	2.99E-02	0.28911	0.37774	N
18	Wyoming Basin	29	24	-0.67631	1.05950	0.19142	N	0.64267	-4.24E-02	8.27E-02	N	1.95318	-0.62674	0.51291	N
19	Wasatch and Uinta Mountains	40	11.1	-0.65514	1.08142	0.26277	N	0.54154	-5.03E-02	7.79E-02	N	0.57684	0.16196	0.27803	N
20	Colorado Plateaus	25	10.6	-2.55E-02	0.67120	0.31436	N	0.53741	-4.85E-02	7.08E-02	N	0.19138	0.25298	0.66753	N
21	Southern Rockies	114	10.5	-0.36605	0.90159	0.29332	N	0.65426	-8.12E-02	7.63E-02	Y	0.88251	-8.56E-02	0.26902	N
22	Arizona/New Mexico Plateau	22	13	0.40758	0.40039	0.16464	Y	0.31551	5.90E-02	4.58E-02	N	1.11129	-0.23639	0.38403	N
23	Arizona/New Mexico Mountains	19	36.4	-0.62245	0.82978	0.22480	N	0.47976	-6.57E-03	0.12633	N	0.77606	2.17E-02	0.81859	N
24	Southern Deserts	5	220	-3.61153	1.84153	0.14815	N	3.15168	-0.76859	8.73E-02	N	-0.75309	0.34014	0.13326	N
25	Western High Plains	15	8.6	0.82900	0.17454	0.43696	Y	0.40658	-4.46E-03	8.77E-02	N	-2.33018	1.08288	0.68352	N
26	Southwestern Tablelands	14	62.3	-3.22007	1.64352	0.40359	N	1.02329	-0.20617	7.55E-02	N	1.98938	-0.71516	0.20379	N
27	Central Great Plains	35	24.6	-1.01025	0.70950	0.36193	N	0.80700	-3.21E-02	0.14771	N	0.10212	2.64E-04	0.39214	N
28	Flint Hills	6	110	-1.18161	1.04728	0.12364	N	-0.31136	0.45047	7.63E-02	N	2.89983	-1.20649	0.44379	N
29	Central Oklahoma/Texas Plains	31	24.2	-1.28142	0.90907	0.28758	N	0.93370	-1.12E-02	0.10819	N	1.17120	-0.35163	0.38403	N

Table 6. Nonparametric regression-equation statistics for calculating the mean, standard deviation, and skew of the common logarithms of nonzero streamflows from the logarithms of drainage area for streamgages in each ecoregion. These regression-equation statistics may be used to refine drainage-area ratio estimates for sites within each ecoregion.—Continued

[ft³/s, cubic foot per second; SD, standard deviation; Log$_{10}$, common logarithms; mi², square miles; MAD, median absolute deviation; NPCIS, ninety-percent confidence interval for the slope; N, the slope is not different from the specified value based on the 90-percent confidence interval; Y, the slope is different from the specified value based on the 90-percent confidence interval; N1, not enough streamgages to determine a meaningful confidence interval; --, not enough streamgages to calculate statistics; S., southern; N., northern. Nonparametric regression equations were developed by using the Kendall-Theil Robust Line method (Granato, 2006)]

Ecoregion				Statistics for the common logarithms of nonzero streamflows											
No.	Name	Number of streamgages	Minimum drainage area (mi²)	Geometric mean (ft³/s)			NPCIS <>1	Geometric SD (unitless)			NPCIS <>0	Coefficient of skew (unitless)			NPCIS <>0
				Intercept	Slope	MAD		Intercept	Slope	MAD		Intercept	Slope	MAD	
30	Edwards Plateau	13	15	-1.02842	1.04354	0.16699	N	1.33685	-0.31475	0.13232	N	-1.34710	0.50338	0.34574	N
31	Southern Texas Plains	7	149	-1.83446	0.96242	0.63733	N	2.88454	-0.82537	0.38124	N	2.03651	-0.65214	0.50961	N
32	Texas Blackland Prairies	28	17.3	-1.58E-02	0.42433	0.32217	N	0.31504	0.30324	0.17240	N	0.59123	-0.23299	0.21766	N
33	East Central Texas Plains	11	142	-3.97326	2.10344	0.11180	N	2.27650	-0.57151	8.94E-02	N	-1.76690	0.74454	0.36519	N
34	Western Gulf Coastal Plains	29	16.1	-0.10516	0.60960	0.41662	N	0.46663	0.12291	0.15622	N	2.10300	-0.79449	0.40982	Y
35	South Central Plains	40	23.4	-0.52862	0.97379	0.18680	N	0.75344	3.38E-02	0.16219	N	0.13505	-6.39E-02	0.37256	N
36	Ouachita Mountains	7	40.1	-0.69526	1.02573	0.12558	N	0.91662	1.97E-02	1.21E-02	N	-0.53206	0.10360	0.34613	N
37	Arkansas Valley	4	147	-0.33805	0.83499	0.16515	N1	-0.31611	0.57379	5.70E-02	N1	1.63238	-0.94170	0.21788	N1
38	Boston Mountains	6	40.6	5.74E-02	0.76308	3.32E-02	N	0.27629	0.25896	6.24E-02	N	0.16728	-0.33346	0.06953	N
39	Ozark Highlands	24	14.2	-1.10428	1.26164	0.14848	N	0.69923	-7.12E-02	0.10646	N	-0.26933	0.25844	0.30811	N
40	Central Irregular Plains	38	26.6	-1.13224	1.07603	0.10566	N	0.98951	-4.79E-02	0.07730	N	-0.83477	0.38041	0.15741	N
41	Canadian Rockies	3	30.9	0.52249	0.73589	0.23568	N1	0.45740	5.28E-02	3.33E-02	N1	0.49308	-0.23587	0.14876	N1
42	Northwestern Glaciated Plains	12	13	0.74539	7.20E-02	0.64629	N	6.70E-02	0.19467	0.16216	N	1.54636	-0.58618	0.23609	N
43	Northwestern Great Plains	27	37.8	-0.57255	0.81606	0.59219	N	0.61103	-5.82E-02	0.15645	N	0.13641	0.25204	0.63771	N
44	Nebraska Sandhills	3	190	-1.59764	1.41487	3.15E-02	N1	1.05825	-0.35422	1.05E-03	N1	4.28579	-1.09968	0.39570	N1
45	Piedmont	112	14.8	-0.32805	1.06696	0.19273	N	0.52549	-4.41E-02	0.11033	N	0.39514	-9.64E-02	0.39843	N
46	Northern Glaciated Plains	22	17.4	-1.34661	0.74894	0.21887	N	1.26435	-0.10151	0.10219	N	-0.16862	0.13298	0.15447	N
47	Western Corn Belt Plains	56	16.9	-0.80349	0.99811	0.21378	N	0.77385	-6.20E-02	0.13572	N	-0.60034	0.26554	0.33827	N
48	Lake Agassiz Plain	12	16.9	-2.11302	1.25036	0.20409	N	1.15254	-0.15054	0.21523	N	-0.25085	0.13376	0.31107	N
49	Northern Minnesota Wetlands	1	430	--	--	--		--	--	--		--	--	--	
50	Northern Lakes and Forests	45	13.2	-0.34646	1.0717	0.13526	N	0.51255	-8.18E-02	0.13179	N	0.41613	0.14249	0.59855	N
51	Northern Central Hardwood Forests	14	7.9	-0.16777	0.81133	0.32364	N	0.14036	0.12121	0.21050	N	0.33788	0.11271	0.48743	N
52	Driftless Area	13	17.1	-0.88772	1.26514	7.51E-02	N	0.34446	-3.30E-02	3.86E-02	N	2.19728	-0.42981	0.27243	N
53	Southeastern Wisconsin Till Plains	16	18.2	-0.28781	0.97976	0.10133	N	0.62523	-8.63E-02	0.05011	N	0.14011	0.10967	0.24108	N
54	Central Corn Belt Plains	71	11.2	-0.25372	0.95466	0.16544	N	0.51004	2.58E-02	0.11284	N	0.40043	-0.17641	0.32572	Y
55	Eastern Corn Belt Plains	87	10.4	-0.67422	1.06349	0.13777	N	0.80578	-7.89E-02	0.10768	N	-0.27787	0.22923	0.24559	Y
56	S. Michigan/N. Indiana Drift Plains	70	13	-0.55743	1.14398	8.58E-02	Y	2.60933	-1.97E-03	0.50610	Y	-0.25335	0.19192	0.29055	N
57	Huron/Erie Lake Plains	11	149	-0.78743	1.04562	9.43E-02	N	1.44876	-0.32137	0.11257	N	-1.99473	0.91002	0.21626	Y
58	Northeastern Highlands	107	10.9	-0.11796	1.04397	8.38E-02	N	0.63691	-8.34E-02	5.72E-02	Y	-0.38359	0.22695	0.24887	Y
59	Northeastern Coastal Zone	79	10.6	-0.15873	1.08206	5.44E-02	Y	0.61697	-0.08563	5.86E-02	Y	-0.42204	0.14325	0.24716	N

Table 6. Nonparametric regression-equation statistics for calculating the mean, standard deviation, and skew of the common logarithms of nonzero streamflows from the logarithms of drainage area for streamgages in each ecoregion. These regression-equation statistics may be used to refine drainage-area ratio estimates for sites within each ecoregion.—Continued

[ft^3/s, cubic foot per second; SD, standard deviation; Log$_{10}$, common logarithms; mi^2, square miles; MAD, median absolute deviation; NPCIS, ninety-percent confidence interval for the slope; N, the slope is not different from the specified value based on the 90-percent confidence interval; Y, the slope is different from the specified value based on the 90-percent confidence interval; N1, not enough streamgages to determine a meaningful confidence interval; --, not enough streamgages to calculate statistics; S., southern; N., northern. Nonparametric regression equations were developed by using the Kendall-Theil Robust Line method (Granato, 2006)]

Ecoregion				Statistics for the common logarithms of nonzero streamflows											
No.	Name	Number of streamgages	Minimum drainage area (mi^2)	Geometric mean (ft^3/s)			NPCIS <>1	Geometric SD (unitless)			NPCIS <>0	Coefficient of skew (unitless)			NPCIS <>0
				Intercept	Slope	MAD		Intercept	Slope	MAD		Intercept	Slope	MAD	
60	Northern Appalachian Plateau and Uplands	31	12.2	-9.67E-02	0.99718	0.11038	N	0.50134	1.81E-03	3.59E-02	N	-0.15639	0.13145	0.15607	N
61	Erie/Ontario Lake Hills and Plain	21	21.8	-0.66740	1.22778	8.88E-02	N	0.89611	-0.17929	0.05533	Y	-7.12E-02	0.12093	0.16415	N
62	North Central Appalachians	35	20	0.10573	0.95978	0.10287	N	0.47782	-1.65E-03	5.04E-02	N	-0.20757	6.29E-02	0.12061	N
63	Middle Atlantic Coastal Plain	23	12.7	-0.15429	0.92608	0.12007	N	0.31267	0.14508	0.11088	N	1.14853	-0.61353	0.30029	Y
64	Northern Piedmont	94	11	-8.43E-02	0.97806	0.14287	N	0.42173	-5.47E-03	8.02E-02	N	0.49124	-0.10242	0.33480	N
65	Southeastern Plains	100	11.1	-9.62E-02	0.94283	0.16661	N	0.31827	0.07324	0.14011	N	0.29405	-6.87E-03	0.40096	N
66	Blue Ridge Mountains	35	26.8	0.30935	0.92403	0.15385	N	0.36270	-3.12E-02	2.56E-02	N	0.43915	-3.99E-02	0.12195	N
67	Central Appalachian Ridges and Valleys	105	10.7	-0.20983	1.02393	0.11433	N	0.22665	7.49E-02	0.23831	N	0.22665	7.49E-02	0.23831	N
68	Southwestern Appalachians	15	53.8	-0.32459	1.11929	0.15676	N	0.71182	-1.05E-02	4.38E-02	N	-0.36116	0.16454	0.23743	N
69	Central Appalachians	54	11	-0.22797	1.04464	0.11503	N	0.74190	-8.27E-02	6.88E-02	Y	-0.55446	0.22994	0.17811	Y
70	Western Allegheny Plateau	54	12.2	-0.35927	1.03897	0.14102	N	0.82814	-9.68E-02	0.11197	N	-0.60578	0.18758	0.24939	N
71	Interior Plateau	59	12.8	-0.44812	1.01631	0.16449	N	0.69284	-9.37E-03	0.13535	N	-1.83E-02	1.69E-02	0.31971	N
72	Interior River Lowland	33	14.6	-0.75199	0.99265	0.15125	N	0.87853	4.06E-03	0.11584	N	0.1714	-7.12E-02	0.21573	N
73	Mississippi Alluvial Plain	7	207	-1.49493	1.56947	0.21495	N	-0.01639	0.26791	0.19603	N	-4.80821	2.01377	0.25762	N
74	Mississippi Valley Loess Plains	7	68.2	-0.98227	1.34887	4.26E-02	Y	0.76525	-0.14574	0.02728	N	-0.63099	0.95434	0.42662	N
75	Southern Coastal Plain	93	12.4	-0.66349	0.98452	0.26436	N	0.74742	-4.41E-02	0.18477	N	-0.19788	-5.50E-04	0.36774	N
76	Southern Florida Coastal Plain	1	146	--	--	--		--	--	--		--	--	--	
77	North Cascades	15	19	0.46212	1.03913	0.19574	N	0.37988	3.73E-03	4.55E-02	N	0.44515	-0.16447	0.26861	N
78	Klamath Mountains	34	14.2	-0.21938	1.04369	0.20176	N	0.94901	-0.17728	0.13883	N	-0.28267	0.23265	0.22833	N
79	Madrean Archipelago	3	44.8	0.27391	-8.40E-02	6.40E-02	N1	1.44020	-0.46454	0.35386	N1	-3.03892	1.33934	0.45420	N1
80	Northern Basin and Range	13	20.5	-0.96604	0.89414	0.27951	N	0.62043	-4.09E-02	9.97E-02	N	0.68550	-0.12620	0.30958	N
81	Sonoran Basin and Range	11	10.9	0.48727	-8.10E-03	0.32466	N	0.17072	0.47584	0.12901	N	1.41957	-0.98377	0.32954	N
82	Laurentian Plains and Hills	11	92.9	-0.23558	1.13904	8.74E-02	N	0.61488	-6.70E-02	3.29E-02	N	0.25535	-4.24E-02	0.1499	N
83	Eastern Great Lakes and Hudson Lowlands	43	10.4	0.20799	0.84981	0.12624	N	0.27237	0.10237	7.36E-02	N	0.82613	-0.31340	0.23518	Y
84	Atlantic Coastal Pine Barrens	34	10	-0.41022	1.23019	0.13104	N	0.43358	-0.09172	6.13E-02	N	3.16E-02	4.93E-02	0.25423	N

per unit area increase with increasing drainage area (fig. 14A). The slope of the regression line for the geometric standard deviation is small, so the estimated flow-duration curves for different drainage areas are approximately parallel. The intercept of the regression equation for the coefficient of skew is positive, and the slope is negative (table 6). The result is a set of concave-upward flow-duration curves that flatten with increasing drainage area.

For ecoregion 59, the Northeastern Coastal Zone, the slope of the regression equation for the geometric mean streamflow is about one (table 6); as a result, the population of estimated streamflows per unit area is about equal for different drainage areas (fig. 14B). As for ecoregion 2, the slope of the regression line for the geometric standard deviation is small, so the estimated flow-duration curves for different drainage areas are approximately parallel. The intercept of the regression equation for the coefficient of skew is negative, and the slope is positive (table 6). The result is a set of concave-downward flow-duration curves curves that flatten with increasing drainage area.

For ecoregion 81, the Sonoran Basin and Range, the slope of the regression equation for the geometric mean streamflow is slightly negative (table 6); as a result, the population of estimated streamflows per unit area decreases with increasing drainage area (fig. 14C). The slope of the regression line for the geometric standard deviation is relatively high (about 0.5), so that basins with larger drainage areas exhibit more variability, and the estimated flow-duration curves for different basins diverge. The intercept of the regression equation for the coefficient of skew is positive, but the slope is negative (table 6). The result is a set of flow-duration curves that transition from concave upward to concave downward with increasing drainage area.

The trend in the estimated regional flow-duration curves as a function of basin size for ecoregion 81, the Sonoran Basin and Range, may appear to be counterintuitive (fig. 14C), but hydrologic studies in this area support these results. This basin-and-range region is characterized by mountainous headwater basins that drain to desert valleys. These flow-duration curves indicate that normalized streamflows are higher in the headwater basins than in the desert valleys downstream. Carr and others (2000) indicate that annual precipitation totals are higher, and average annual temperatures are lower, in the headwater basins than in the adjacent valleys. The desert valleys are characterized by high potential evapotranspiration, extensive basin-fill aquifer systems composed of transmissive sediments, and proportionally high consumptive water use (Anning and Konieczki, 2005). Research on the timing, duration, and infiltration of streamflow in ephemeral streams in southeastern Arizona indicates that much, if not all, of the intermittent stormflows in these channels may be lost to groundwater recharge through the streambed within the first 6 to 15 mi from the mountain fronts in this area (Gungle, 2006; Coes and Pool, 2005). Vogel and others (1999) indicate a relatively strong positive association between drainage area and mean annual streamflow in the lower Colorado Basin, which includes ecoregion 81. Their regression for mean annual flows, however, is predominantly based on streamflow data from large drainage areas; 53 percent of the streamgages used by Vogel and others (1999) to estimate streamflow statistics in this area have drainage areas greater than 1,000 mi^2. In comparison, ephemeral stream sites studied by Hejl (1980) in New Mexico and by Coes and Pool (2005) in Arizona have drainage areas of 168 and 211 mi^2, respectively. The results of these studies suggest that losses to infiltration, evapotranspiration, and water use may decrease streamflows from headwater areas to the valley floor in some arid ecoregions.

Similar patterns of decreasing streamflows per unit area also characterize ecoregion 79 (the Madrean Archipelago) and ecoregion 11 (the Blue Mountains) (table 7). Each of these ecoregions is an area with geographic and hydrologic conditions similar to those in the Sonoran Basin and Range ecoregion. Also, compared to the humid Northeast (fig. 7, table 7) these areas are represented by relatively few USGS streamgages that have drainage areas in the range of 10 to 500 mi^2 and long periods of record.

The regression-based statistics in table 6 can be used with SELDM to refine initial ecoregion-based values. If regression-based estimates are assumed to be more representative than ecoregion values, the calculated statistics can be entered into SELDM by selecting a user-defined statistics option. The effect of selecting statistics can be assessed as was done in the examples shown in figure 14. Potential effects on water-quality loads and concentrations can be assessed by doing a sensitivity analysis with SELDM.

The occurrence of zero flows may be a concern for streamflow statistics in some stream basins (appendix 2). In some cases, zero-flow measurements may be made if the volume of streamflow is below the minimum detection limit (commonly 0.01 ft^3/s) or if flows in a portion of the channel cross-section that is not measured by a streamgage are extremely low (appendix 1). In many cases, however, a zero-flow measurement does reflect a lack of flow and potentially a dry streambed condition. Statistics for streamgages in arid areas (documented in the SWQDM database; Granato and others, 2009), generally do not indicate a quantitative relation between the ratio of days with measured streamflow to the total number of days in the streamflow record and drainage area. In many ecoregions, streams in small headwater basins may be intermittent or ephemeral. In some arid ecoregions, streams in small headwater basins may be perennial and become intermittent or ephemeral as they lose water to evapotranspiration and the underlying aquifer during flow into arid valleys. Locally, there may be a relation between drainage area and flow status, but on a national scale, drainage-basin size is not a reliable indicator of the flow status of a given stream. In SELDM, the proportion of zero flows can be selected on the basis of the average or median of such values by ecoregion, by selecting values from nearby hydrologically similar sites, or by entering user-defined statistics.

Table 7. Correlation of selected statistics between drainage area and daily mean streamflow values measured during the period 1960–2004 at 2,783 U.S. Geological Survey streamgages in the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions.

[ft^3/s/mi^2, cubic foot per second per square mile; SD, standard deviation; No., number; R, Pearson's parametric correlation coefficient; Rho, Spearman's non-parametric correlation coefficient (correlation of ranks); --, statistics not calculated; S., southern; N., northern; gray boxes indicate correlations with absolute values equal to or greater than 0.5; **bold** text indicates correlations with absolute values equal to or greater than 0.75; the statistical significance of a correlation coefficient is a function of sample size (Caruso and Cliff, 1997)]

Ecoregion		Number of streamgages	Statistics for the common logarithms of nonzero streamflows							
			Median (ft^3/s/mi^2)		Geometric mean (ft^3/s/mi^2)		Geometric SD (unitless)		Coefficient of skew (unitless)	
No.	Name		R	Rho	R	Rho	R	Rho	R	Rho
1	Coast Range	35	0.12	0.24	0.09	0.21	0.01	-0.05	-0.11	-0.15
2	Puget Lowland	28	0.28	0.41	0.27	0.40	-0.20	-0.17	-0.27	-0.30
3	Willamette Valley	15	-0.33	-0.25	-0.22	-0.12	-0.16	-0.04	0.50	0.48
4	Cascades	72	-0.08	-0.12	-0.01	-0.03	-0.28	-0.37	0.16	0.42
5	Sierra Nevada	86	-0.13	-0.19	-0.14	-0.17	-0.13	-0.10	0.17	0.09
6	Southern and Central California Plains and Hills	113	0.05	0.01	0.08	-0.02	0.05	0.00	-0.01	-0.13
7	Central California Valley	8	0.15	0.69	0.12	0.57	0.33	0.21	-0.39	-0.50
8	Southern California Mountains	17	-0.31	-0.35	-0.29	-0.42	0.20	0.18	-0.42	-0.27
9	Eastern Cascades Slopes and Foothills	11	-0.16	-0.35	-0.34	-0.41	0.47	0.67	-0.43	-0.23
10	Columbia Plateau	17	-0.33	-0.15	-0.31	-0.04	-0.22	-0.08	0.22	0.12
11	Blue Mountains	15	-0.04	-0.46	-0.14	-0.56	0.50	0.06	-0.36	-0.34
12	Snake River Basin/High Desert	3	--	--	--	--	--	--	--	--
13	Northern Basin and Range	30	-0.03	-0.02	-0.05	-0.07	-0.00	0.28	-0.37	-0.50
14	Southern Basin and Range	6	-0.23	-0.09	-0.18	-0.09	-0.23	-0.09	0.60	0.49
15	Northern Rockies	11	-0.53	-0.49	-0.57	-0.50	0.07	-0.15	0.01	-0.15
16	Montana Valley and Foothill Prairies	14	-0.12	0.17	-0.08	0.08	-0.14	-0.29	0.12	0.08
17	Middle Rockies	48	0.06	-0.06	0.06	-0.02	-0.35	-0.41	-0.07	0.16
18	Wyoming Basin	29	0.03	0.13	0.06	0.08	-0.19	-0.10	-0.10	-0.31
19	Wasatch and Uinta Mountains	40	0.05	0.15	-0.02	0.08	-0.25	-0.17	0.01	0.14
20	Colorado Plateaus	25	-0.27	-0.44	-0.33	-0.37	-0.10	-0.19	-0.05	0.08
21	Southern Rockies	114	-0.04	-0.06	-0.13	-0.14	-0.12	-0.30	-0.12	-0.08
22	Arizona/New Mexico Plateau	22	-0.50	-0.64	-0.46	-0.57	0.20	0.28	-0.14	-0.07
23	Arizona/New Mexico Mountains	19	-0.33	-0.14	-0.34	-0.21	-0.08	-0.03	-0.06	0.06
24	Southern Deserts	5	**0.79**	0.30	**0.79**	0.30	-0.27	-0.60	-0.06	-0.10
25	Western High Plains	15	**-0.78**	-0.60	-0.74	-0.60	0.15	0.01	0.33	0.36
26	Southwestern Tablelands	14	0.52	0.49	0.57	0.52	-0.17	-0.32	-0.25	-0.34
27	Central Great Plains	35	-0.22	-0.11	-0.16	-0.10	-0.14	-0.04	0.06	0.06
28	Flint Hills	6	0.11	0.14	-0.01	0.14	0.52	0.49	-0.40	-0.37
29	Central Oklahoma/Texas Plains	31	-0.06	-0.06	-0.05	0.01	0.05	0.02	-0.19	-0.21
30	Edwards Plateau	13	-0.28	-0.13	-0.13	0.04	-0.29	-0.36	0.31	0.17
31	Southern Texas Plains	7	0.15	0.07	0.17	-0.14	-0.14	-0.25	0.08	0.00
32	Texas Blackland Prairies	28	-0.16	-0.41	-0.17	-0.38	0.31	0.31	0.00	-0.22
33	East Central Texas Plains	11	-0.05	0.25	0.02	0.29	-0.29	-0.36	0.29	0.32
34	Western Gulf Coastal Plains	29	-0.20	-0.34	-0.16	-0.28	0.04	0.23	-0.43	-0.47
35	South Central Plains	40	-0.17	-0.10	-0.12	-0.00	-0.01	0.04	-0.11	-0.02
36	Ouachita Mountains	7	-0.28	-0.29	-0.21	-0.07	-0.02	0.14	-0.04	0.14
37	Arkansas Valley	4	--	--	--	--	--	--	--	--
38	Boston Mountains	6	0.11	0.03	-0.33	-0.37	**0.86**	**0.83**	-0.72	**-0.77**
39	Ozark Highlands	24	0.31	0.31	0.31	0.35	-0.05	-0.11	0.17	0.13
40	Central Irregular Plains	38	-0.34	-0.11	-0.27	0.11	-0.05	-0.10	0.25	0.32
41	Canadian Rockies	3	--	--	--	--	--	--	--	--
42	Northwestern Glaciated Plains	12	-0.36	-0.43	-0.31	-0.43	0.31	0.37	-0.29	-0.24

Table 7. Correlation of selected statistics between drainage area and daily mean streamflow values measured during the period 1960–2004 at 2,783 U.S. Geological Survey streamgages in the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions.—Continued

[$ft^3/s/mi^2$, cubic foot per second per square mile; SD, standard deviation; No., number; R, Pearson's parametric correlation coefficient; Rho, Spearman's non-parametric correlation coefficient (correlation of ranks); --, statistics not calculated; S., southern; N., northern; gray boxes indicate correlations with absolute values equal to or greater than 0.5; **bold** text indicates correlations with absolute values equal to or greater than 0.75; the statistical significance of a correlation coefficient is a function of sample size (Caruso and Cliff, 1997)]

Ecoregion		Number of streamgages	Statistics for the common logarithms of nonzero streamflows							
			Median ($ft^3/s/mi^2$)		Geometric mean ($ft^3/s/mi^2$)		Geometric SD (unitless)		Coefficient of skew (unitless)	
No.	Name		R	Rho	R	Rho	R	Rho	R	Rho
43	Northwestern Great Plains	27	-0.15	-0.10	-0.11	-0.10	0.09	-0.08	0.14	0.13
44	Nebraska Sandhills	3	--	--	--	--	--	--	--	--
45	Piedmont	112	0.10	0.12	0.09	0.10	-0.20	-0.13	-0.07	-0.06
46	Northern Glaciated Plains	22	0.06	-0.14	0.13	-0.10	-0.16	-0.27	0.20	0.24
47	Western Corn Belt Plains	56	-0.15	-0.13	-0.03	-0.01	-0.14	-0.14	0.10	0.18
48	Lake Agassiz Plain	12	0.21	0.05	0.25	0.22	-0.20	-0.17	0.36	0.20
49	Northern Minnesota Wetlands	1	--	--	--	--	--	--	--	--
50	Northern Lakes and Forests	45	-0.01	0.07	0.02	0.10	-0.24	-0.18	0.00	0.06
51	Northern Central Hardwood Forests	14	-0.38	-0.33	-0.34	-0.20	-0.01	0.25	-0.04	0.05
52	Driftless Area	13	0.32	0.38	0.34	0.38	-0.07	-0.19	-0.28	-0.32
53	Southeastern Wisconsin Till Plains	16	-0.05	0.04	-0.10	-0.00	-0.20	-0.33	-0.18	0.12
54	Central Corn Belt Plains	71	-0.16	-0.04	-0.21	-0.10	0.13	0.07	-0.19	-0.14
55	Eastern Corn Belt Plains	87	0.08	0.06	0.15	0.16	-0.24	-0.19	0.22	0.24
56	S. Michigan/N. Indiana Drift Plains	70	0.10	0.22	0.14	0.29	-0.24	-0.26	0.21	0.20
57	Huron/Erie Lake Plains	11	0.23	-0.23	0.28	0.00	-0.45	-0.32	0.68	0.65
58	Northeastern Highlands	107	-0.10	-0.01	-0.03	0.15	-0.05	-0.31	0.11	0.22
59	Northeastern Coastal Zone	79	0.17	0.25	0.27	0.37	-0.30	-0.37	0.13	0.14
60	Northern Appalachian Plateau and Uplands	31	-0.10	-0.02	-0.07	0.15	-0.03	0.02	0.22	0.16
61	Erie/Ontario Lake Hills and Plain	21	0.27	0.40	0.39	0.41	-0.55	-0.56	0.19	0.20
62	North Central Appalachians	35	-0.09	-0.17	-0.06	-0.12	-0.05	-0.02	0.03	0.13
63	Middle Atlantic Coastal Plain	23	0.07	0.02	-0.18	-0.16	0.28	0.26	-0.57	-0.54
64	Northern Piedmont	94	-0.07	-0.03	-0.06	-0.05	-0.01	-0.01	-0.04	-0.07
65	Southeastern Plains	100	-0.15	-0.19	-0.11	-0.12	0.17	0.17	-0.02	-0.02
66	Blue Ridge Mountains	35	-0.11	-0.14	-0.12	-0.16	-0.24	-0.29	-0.15	-0.11
67	Central Appalachian Ridges and Valleys	105	-0.05	0.02	-0.03	0.06	-0.04	0.01	0.08	0.08
68	Southwestern Appalachians	15	0.11	0.24	0.14	0.23	-0.18	-0.05	0.16	0.09
69	Central Appalachians	54	-0.03	0.06	0.06	0.11	-0.33	-0.33	0.31	0.31
70	Western Allegheny Plateau	54	-0.10	-0.03	-0.02	0.09	-0.19	-0.17	0.17	0.21
71	Interior Plateau	59	0.06	0.01	0.09	0.02	0.01	-0.02	0.01	0.02
72	Interior River Lowland	33	0.09	-0.02	0.13	-0.01	-0.07	-0.01	-0.04	-0.07
73	Mississippi Alluvial Plain	7	0.03	0.07	0.11	0.07	-0.11	0.04	0.56	0.46
74	Mississippi Valley Loess Plains	7	0.75	0.64	0.72	**0.79**	-0.49	-0.71	0.21	0.18
75	Southern Coastal Plain	93	-0.03	-0.04	-0.00	-0.01	-0.09	-0.09	0.04	-0.01
76	Southern Florida Coastal Plain	1	--	--	--	--	--	--	--	--
77	North Cascades	15	0.17	0.14	0.08	0.19	-0.11	0.12	-0.40	-0.11
78	Klamath Mountains	34	0.19	-0.02	0.22	0.04	-0.26	-0.26	0.35	0.21
79	Madrean Archipelago	3	--	--	--	--	--	--	--	--
80	Northern Basin and Range	13	-0.03	-0.04	-0.05	-0.02	-0.12	-0.04	-0.16	-0.16
81	Sonoran Basin and Range	11	-0.31	-0.44	-0.40	-0.58	0.14	0.49	-0.42	-0.40
82	Laurentian Plains and Hills	11	0.26	0.32	0.23	0.22	-0.20	-0.20	-0.01	0.07
83	Eastern Great Lakes and Hudson Lowlands	43	-0.15	-0.24	-0.19	-0.29	0.28	0.31	-0.28	-0.32
84	Atlantic Coastal Pine Barrens	34	0.28	0.28	0.18	0.26	-0.20	-0.24	-0.04	0.03

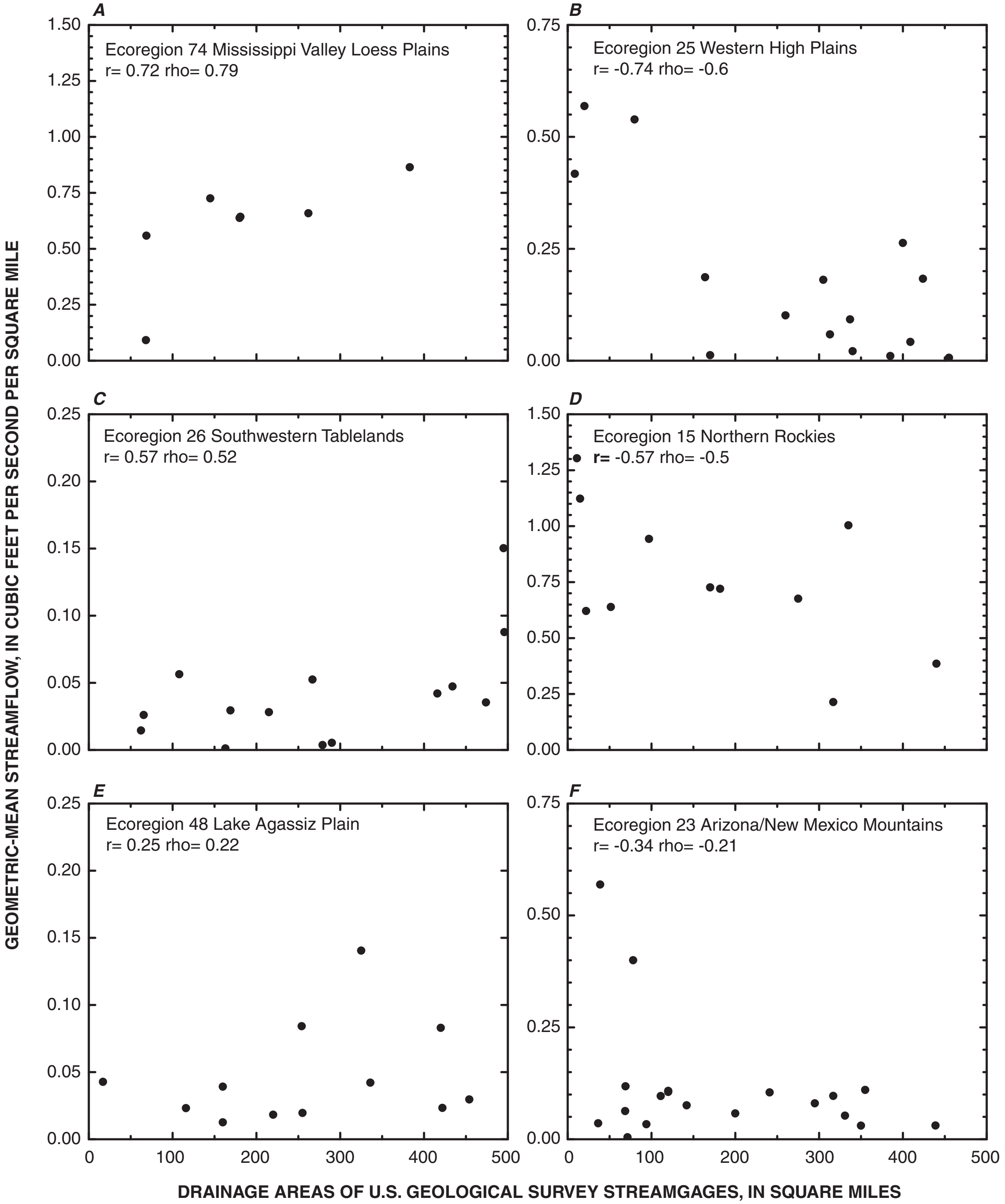


Figure 13. Examples of correlations between geometric mean streamflows and drainage areas for six ecoregions in the conterminous United States (including r, Pearson's parametric correlation coefficient and rho, Spearman's nonparametric correlation coefficient).

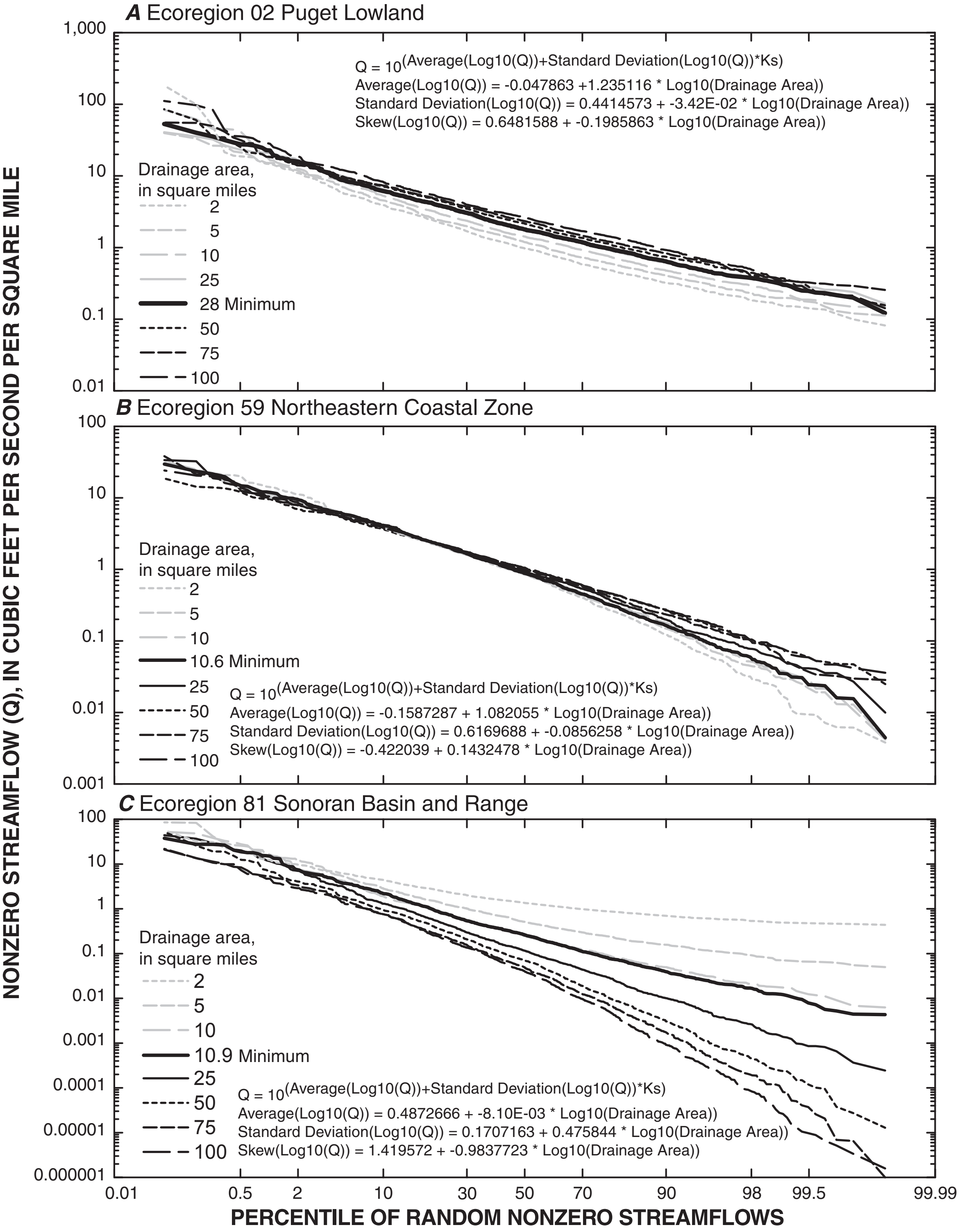


Figure 14. Flow-duration curves from streamflow data stochastically generated by using regional regression equations (from table 6) for the mean, standard deviation, and skew of the logarithms of nonzero streamflows in three selected ecoregions.

Estimating Streamflow Statistics at Sites with Limited Data

Streamflow correlation (appendix 2) was the method selected for estimating long-term streamflow statistics for sites with limited streamflow data. The SREF program (Granato, 2009) was developed as part of the SELDM study so that available site-specific data or data collected as part of a site-specific highway-runoff assessment could be used to estimate long-term streamflow statistics for use with with SELDM. For example, if streamflow data have been collected at a site of interest, regional planning-level prestorm-flow estimates can be refined by using this site-specific streamflow data.

Streamflow data from the continuous-record streamgages selected for this national synthesis may be used to refine estimates of streamflow statistics for many other sites. Streamflow estimates can be calculated by correlation methods using long-term continuous-record streamflow data and concurrent measurements available at or near the site of interest. Streamflow estimates for sites in the same basin as the site of interest could be used with the drainage-area-ratio method to refine estimates at the site of interest. Data from almost 25,000 continuous streamgages, from more than 46,000 other sites with concurrent measurements of stage and flow, and from more than 24,000 surface-water-quality monitoring stations with one or more measurements of stream discharge are available in the USGS NWIS Web (U.S. Geological Survey, 2009) database. Many of these sites may not meet the selection criteria for calculating long-term regional statistics, but they may be close to and hydrologically similar to a site of interest. These statistics can be input into SELDM by selecting the user-defined streamflow-statistics option.

Streamflow correlation may provide the best method for refining initial planning-level estimates necessary for stochastic data generation at a site of interest; however, this method can be very time and resource intensive if there is no preexisting data for the site of interest. At least one or two years of daily mean streamflow data are necessary to represent seasonality and flow variation. The cost of obtaining a sufficient number of instantaneous partial-record measurements to represent a wide range of streamflows may be about one to two orders of magnitude more than the cost of making a drainage-area-ratio estimate based on the statistics provided in this report. The cost of installing and running a short-term streamgage to improve the statistical estimates may be two to three orders of magnitude more than the cost of making a drainage-area-ratio estimate.

Storm-Event Precipitation Statistics

SELDM uses precipitation statistics as the basis for the stochastic generation of random storm events that produce runoff from the highway site and the associated upstream basin (fig. 2). To develop planning-level estimates of receiving-water flow and quality at a site of interest, it is necessary to characterize the precipitation statistics for storm events that generate runoff flows. Precipitation statistics can be used to estimate the number of storm events within a given interval, the time between successive storms, and the duration and total precipitation for each event. Precipitation statistics also may be used to differentiate between runoff-producing storm events and storm events that do not produce measurable runoff. Differences in the hydrologic characteristics of highway catchments and the associated upstream basins, however, result in differences in the occurrence, timing, and distribution of runoff flows from each of these areas for a given storm event.

Short-duration rainfall pulses in precipitation data are commonly aggregated to define discrete storm events. These storm events are characterized by a minimum interevent time (IET), the total event duration, the total event volume, and the average event intensity (fig. 15). This definition of a storm-event commonly is used for planning-level estimates of the quantity and quality of highway- and urban-runoff; the design and evaluation of runoff-quality BMPs; and the simulation of runoff flows (Driscoll and others, 1979; Goforth and others, 1983; Adams and others, 1986; Strecker and others, 2001; Driscoll and others, 1990a,b; U.S. Environmental Protection Agency, 1992; Adams and Papa, 2000; Asquith and others, 2006).

Methods for Analyzing Precipitation Statistics

Studies have been done to analyze precipitation statistics for stormwater sampling, defining storm-event characteristics, rainfall-runoff modeling, and BMP design and analysis purposes (appendix 3). The results of these studies have been used for the analysis of highway-and urban-runoff data. For example, the NURP used data from 40 sites to derive regional storm-event statistics (Athayde and others, 1983). Driscoll and others (1986) updated the original NURP statistics for planning-level runoff analysis and BMP evaluation by analyzing data from 62 sites to derive storm-event statistics for nine rain zones in the conterminous United States. Driscoll and others (1990a,b) used this national nine-zone map for characterizing storm-event statistics for use with the 1990 FHWA runoff-quality model throughout the conterminous United States. Driscoll, Palhegyi, and others (1989) updated the U.S. Environmental Protection Agency (USEPA) precipitation statistics with data collected at 136 sites during 1949–1987 to provide a map that is still in current use with 15 rain zones for characterizing storm events.

The report by Driscoll, Palhegyi, and others (1989) provides analysis of necessary statistics for runoff-generating events in 15 rain zones throughout the conterminous United States, and these rain-zone statistics are supported by the USEPA for planning and analysis purposes (table 8). This report was not published as a numbered USEPA report, but the results of this analysis have been published for use in urban-runoff-monitoring studies (U.S. Environmental Protection Agency, 1992) and for the planning and design of BMPs

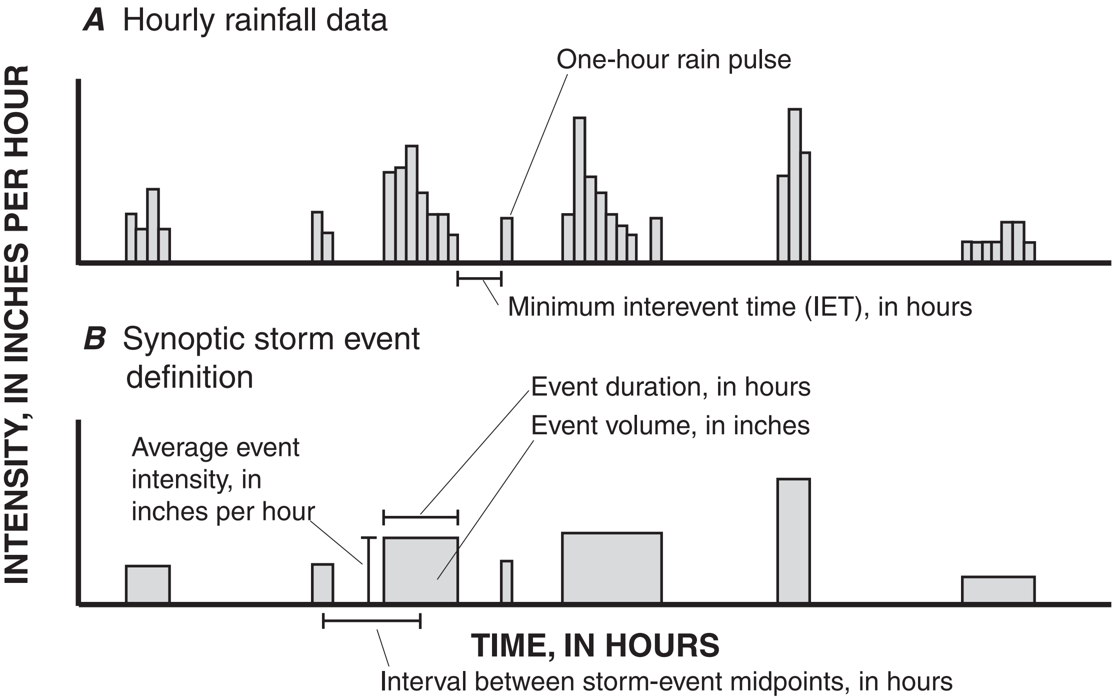


Figure 15. Schematic diagram showing the characterization of a synoptic storm event. (Modified from Driscoll, Palhegyi, and others, 1989)

for runoff control (U.S. Environmental Protection Agency, 2002; Clar and others, 2004). Although the report by Driscoll, Palhegyi, and others (1989) was not formally published by the USEPA, it is readily available through interlibrary loan. In addition, an Adobe PDF version of this document is provided in the SYNOPDOC subdirectory in the Precipitation directory on the CD–ROM accompanying this report; the methods developed by Driscoll, Palhegyi, and others (1989) are relevant because they are the same as the methods used in the SELDM development project.

Availability of Precipitation Data

A decade or more of hourly-precipitation data are necessary for developing representative storm-event statistics. The National Climatic Data Center (NCDC) is the primary source of most long-term hourly-precipitation data. Two NCDC datasets are of primary interest for use and interpretation of hourly precipitation. These datasets are the station-history dataset DSI–9767 (National Climatic Data Center, 2002; 2006; 2007) and the DSI–3240 hourly-precipitation dataset (National Climatic Data Center, 2003a; 2007). The NCDC sells hourly-precipitation data collected at more than 6,000 National Weather Service stations for the period 1948 to the present and 15-minute precipitation data collected at more than 3,400 National Weather Service stations for the period 1971 to the present (National Climatic Data Center, 2007). The NCDC provides precipitation-event statistics for hourly-precipitation data (National Climatic Data Center, 2005), but these statistics are based on a storm-event definition (National Climatic Data Center, 2003b) different from the one that is commonly used in runoff studies (Driscoll and others, 1979; Goforth and others, 1983; Adams and others, 1986; Strecker and others, 1989; Driscoll, Palhegyi, and others, 1989; Driscoll and others, 1990a,b).

A Microsoft Access database (SiteStormV01.mdb) was assembled for the SELDM-development project to facilitate the development of statistical storm-event estimates at ungaged sites. This database includes National Oceanic and Atmospheric Administration (NOAA) precipitation-monitoring-site characteristics and associated synoptic storm-event statistics. The regional statistics described in this report were derived from the individual station statistics stored in the SiteStormV01.mdb database available on the CD–ROM accompanying this report. The database includes statistics calculated in this study from data collected within the conterminous United States at 2,610 NOAA hourly-precipitation data stations for the 1965–2006 period. The database also includes statistics calculated by Driscoll, Palhegyi, and others (1989) from data collected at the 136 NOAA hourly-precipitation data stations for the 1949–1987 period.

Table 8. Summary synoptic precipitation-event statistics for the 15 rain zones within the conterminous United States defined by Driscoll, Palhegyi, and others (1989). The statistics are calculated from data collected at selected hourly-precipitation data stations in each rain zone during the 1949–1987 and 1965–2006 periods.

[COV, coefficient of variation defined as the standard deviation divided by the mean; storm events are defined as having a minimum volume of 0.1 inches of precipitation and a 6-hour minimum interevent time threshold]

Rain zone			Number of storm events (per year)		Annual precipitation volume (inches)		Storm-event volume (inches)		Storm-event duration (hours)		Time between storm-event midpoints (hours)	
Num-ber	Name	Number of stations	Mean	COV	Mean	COV	Mean	COV	Mean	COV	Mean	COV
Statistics for the 1949–1987 period (Driscoll, Palhegyi, and others, 1989)												
1	Northeast	11	70	0.13	34.61	0.18	0.50	0.95	11.2	0.81	126.3	0.94
2	Northeast Coastal	6	63	0.12	41.43	0.21	0.66	1.03	11.7	0.77	139.8	0.87
3	Mid-Atlantic	12	60	0.15	39.61	0.20	0.66	1.02	9.7	0.84	144.5	1.01
4	Central	7	68	0.14	41.87	0.19	0.62	1.00	9.2	0.85	132.7	0.99
5	North Central	19	55	0.16	29.81	0.22	0.55	1.01	9.5	0.83	166.7	1.17
6	Southeast	14	65	0.15	48.99	0.20	0.75	1.10	8.7	0.92	135.6	1.03
7	East Gulf	6	68	0.17	54.21	0.23	0.80	1.19	6.4	1.05	130.3	1.25
8	East Texas	9	41	0.22	31.22	0.29	0.76	1.18	8.0	0.97	212.7	1.28
9	West Texas	5	30	0.27	17.28	0.33	0.57	1.07	7.4	0.98	302.6	1.53
10	Southwest	10	20	0.30	7.40	0.37	0.37	0.88	7.8	0.88	473.1	1.46
11	West Inland	6	14	0.38	4.91	0.43	0.36	0.87	9.4	0.75	786.2	1.54
12	Pacific Southwest	3	19	0.36	10.23	0.42	0.54	0.98	11.6	0.78	476.0	2.09
13	Northwest Inland	19	31	0.23	11.50	0.29	0.37	0.93	10.4	0.82	303.8	1.43
14	Pacific Central	6	32	0.25	18.39	0.33	0.58	1.05	13.7	0.80	264.8	2.00
15	Pacific Northwest	3	71	0.15	35.73	0.19	0.50	1.09	15.9	0.80	123.0	1.50
Statistics for the 1965–2006 period (current study)												
1	Northeast	270	55	0.29	30.81	0.30	0.56	0.93	8.4	0.85	148.7	1.19
2	Northeast Coastal	61	53	0.26	37.59	0.30	0.71	1.00	9.3	0.81	153.9	1.04
3	Mid-Atlantic	122	49	0.29	32.67	0.31	0.67	0.99	7.8	0.88	166.2	1.26
4	Central	182	50	0.29	35.57	0.31	0.71	1.00	7.2	0.90	164.5	1.30
5	North Central	506	42	0.30	24.70	0.32	0.59	0.97	7.0	0.88	210.5	1.46
6	Southeast	227	53	0.25	43.19	0.28	0.82	1.04	6.8	0.93	158.8	1.25
7	East Gulf	84	57	0.23	47.56	0.27	0.84	1.11	5.7	1.00	150.1	1.36
8	East Texas	173	37	0.26	28.79	0.30	0.78	1.07	6.5	0.96	239.3	1.34
9	West Texas	63	24	0.30	14.95	0.35	0.62	1.00	5.6	1.00	375.4	1.72
10	Southwest	157	23	0.34	10.15	0.38	0.45	0.87	6.2	0.95	427.7	1.64
11	West Inland	72	17	0.43	11.48	0.51	0.60	0.99	8.6	0.91	675.7	1.70
12	Pacific Southwest	89	17	0.42	14.90	0.54	0.85	1.13	10.4	0.94	498.9	2.07
13	Northwest Inland	353	27	0.34	11.48	0.36	0.42	0.87	7.4	0.91	356.6	1.62
14	Pacific Central	134	35	0.34	29.97	0.41	0.84	1.10	12.0	0.92	255.1	2.05
15	Pacific Northwest	117	62	0.26	46.03	0.31	0.70	1.11	12.6	0.93	159.4	1.74

The USEPA Better Assessment Science Integrating Point and Non-Point Sources (BASINS) Program provides hourly-precipitation data in water-data-management (WDM) format (Lumb and others 1988; Flynn and others, 1995; Hummel and others, 2001). This hourly-precipitation dataset is available by state for the period 1970 through 1996 (U.S. Environmental Protection Agency, 2001a; 2007).

The USGS NWIS Web (available at http://waterdata.usgs.gov/nwis/) lists about 6,000 meteorological stations, many of which are described as rain gages. The NWIS Web interface allows the user to search for meteorological sites (and data) by state, by user-specified latitudes and longitudes, or by hydrologic region., Available data must be converted to hourly measurements in a standard format, however, to generate statistics that are consistent with historical synoptic storm-event precipitation statistics.

Software for Analyzing Precipitation Statistics

The Synoptic Rainfall Data-Analysis Program (SYNOP) program was developed by the USEPA to facilitate statistical analysis of precipitation-event statistics as part of area-wide-assessment procedures (U.S. Environmental Protection Agency, 1976). The SYNOP program reads hourly precipitation and aggregates this data into synoptic storm events (fig. 15) that meet user-defined criteria for the minimum storm-event size and the minimum IET. Driscoll and others (1986) updated the SYNOP program as part of the NURP program to iterate through hourly rainfall data to find an IET value that resulted in a COV of 1. This method was designed to make the distribution of IETs an exponential distribution so the storm events could be characterized as a Poisson process (Strecker and others, 1989; Driscoll and others, 1990a,b). The 1990 FHWA report series provided a version of the SYNOP model that was updated further and compiled for use on a personal computer (Strecker and others, 1989). This version enables the user to select calendar years or water years, do a seasonal analysis, exclude events that do not meet minimum precipitation criteria, and print output files. Published versions of the SYNOP program cannot process four-digit year-2000-compatible data formats, but the FHWA version of SYNOP was converted from a two-digit year-2000 format to a four-digit format (Eric Strecker, Geosyntec, written commun., 2003; Tarig Omer, Hydroqual, written commun., 2003). This version of SYNOP is designated as SYNOP2000 in this report.

The 1990 FHWA report series provided a SYNOP data preprocessor (SYNPREP) program for use with personal computers to process weather-data formats available at that time (Strecker and others, 1989; Driscoll and others, 1990a,b). SYNPREP provides the input-file format necessary for use with SYNOP. The SYNPREP program can be used to process the NOAA hourly-precipitation data in the NCDC variable-length (DSI-3240) text-file format (National Climatic Data Center, 2003a; 2007).

In the current SELDM-development project, a synoptic-precipitation-analysis facilitator (SPAF) was developed to facilitate batch processing of multistation hourly-precipitation datasets in the variable-length text-file format available from the NCDC (2003a; 2007). The SPAF program provides a Visual-Basic interface to run SYNPREP and SYNOP2000 in a command-prompt disc-operating system (DOS) shell by converting selections on the graphical-user interface (GUI) to input-file options and DOS-shell "run" commands. The user needs to obtain a NCDC station file, a list file, and a data file for each group of stations to be analyzed from the NCDC Website (http://cdo.ncdc.noaa.gov/). The NCDC station file is needed to identify station information. The list file is needed to identify the number of years of available record for each station. The variable-length data files are needed to provide precipitation data for each station. The SPAF program

- Copies the SYNPREP and SYNOP2000 computer programs to the data directory,
- Runs SYNPREP to reformat the data file,
- Creates SYNOP2000 input and control files for each station on the basis of user-selected options,
- Runs SYNOP2000 to calculate statistics for each selected station, and
- Writes the station information and summary statistics to a single output file.

Documentation, executable programs, and source code for SYNOP, SYNPREP and SPAF are in the directory named "Precipitation" on the CD–ROM accompanying this report. Scanned copies of available documentation in Adobe PDF format are included for the original SYNOP documentation in the USEPA Area Wide Waste Treatment Management And Planning Effort (United Environmental Protection Agency, 1976); the FHWA highway-runoff modeling project (Strecker and others, 1989; Driscoll and others, 1990a,b), and the 15-rain-zone-analysis effort completed as part of the USEPA NURP project (Driscoll, Palhegyi, and others, 1989). FORTRAN source-code files for SYNPREP, SYNOP, and SYNOP2000 are provided in American Standard Code for Information Interchange (ASCII) text-file format. The Visual Basic source code for SPAF is in Visual Basic project files in the SPAF subdirectory.

The Microsoft Access database (SiteStormV01.mdb) available on the CD–ROM accompanying this report includes seven tables (fig. 16) to store and facilitate use of the synoptic-precipitation statistics. In general, tables, fields, and example queries were named with whole words or acronyms to facilitate use of the database. The description property for each table, field, and query provides a short statement for defining each of these database objects. The table tblPrecipStationStatistics is the primary precipitation-statistics table in the database. This table comprises 19 data fields to identify sites and provide the statistics. The table tdxNOAAHPStations comprises 7 data fields to

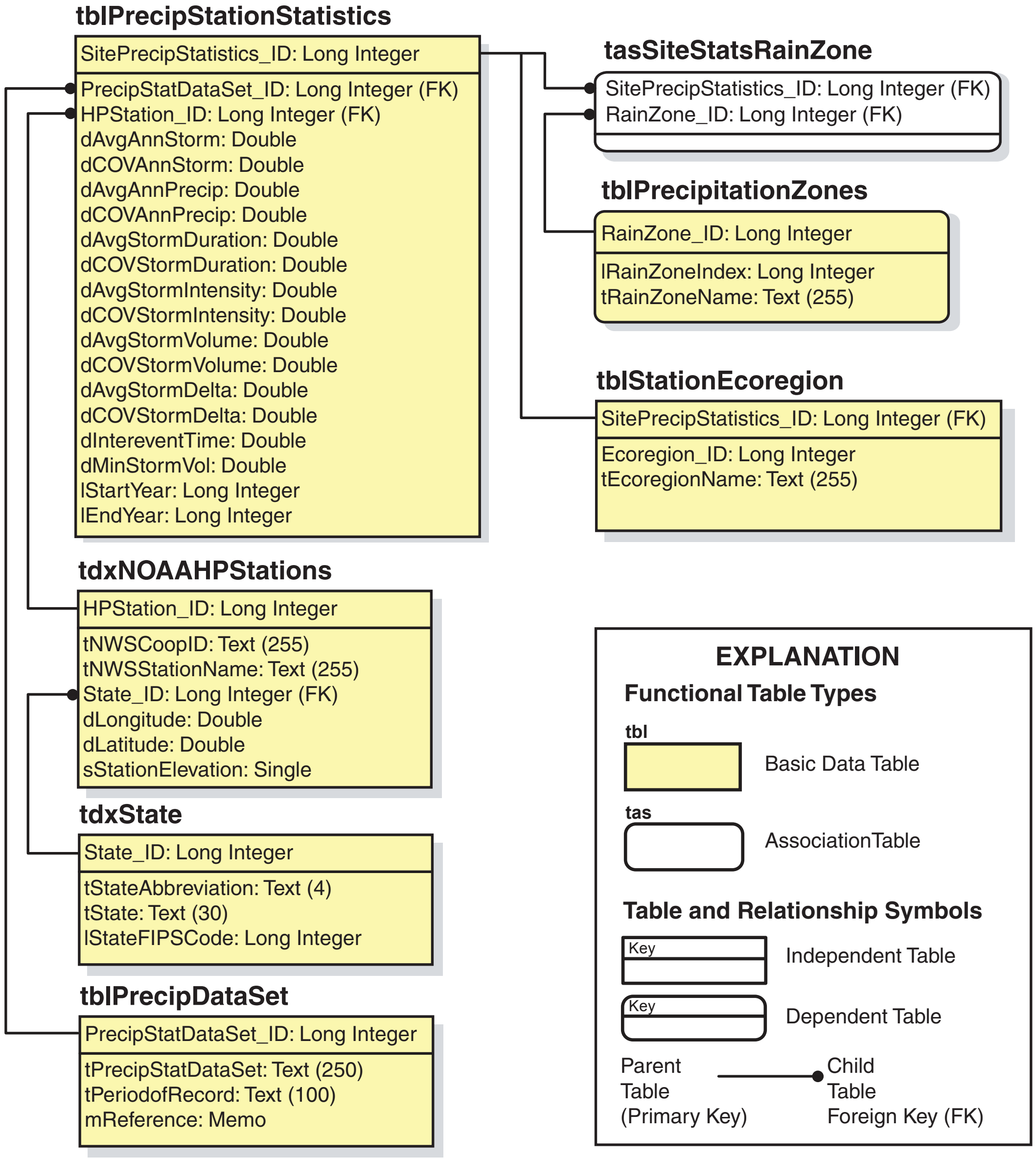


Figure 16. An entity-relationship diagram showing a graphical representation of tables, fields, and relationships of the data structure for precipitation data in the SiteStormV01.mdb database.

identify sites and provide the name, location, and elevation data that may be used to characterize each data-collection site. The table tblPrecipDataSet comprises 4 data fields to identify each synoptic precipitation-statistics dataset and the source documentation. The data table tasSiteStatsRainZone provides the link between precipitation statistics in table tblPrecipStationStatistics and the rain zones identified in table tblPrecipitationZones. These tables allow the user to link the site statistics to one or more rain zones. The data table tblStationEcoregion comprises 3 data fields and is used to identify the USEPA Level III nutrient ecoregion that includes each precipitation data-collection site.

The SiteStormV01.mdb database also includes example queries that were used to analyze the data. The query qryPrecipGetTable05 can be used to reproduce the precipitation statistics by USEPA rain zone in table 8. The query qryPrecipGetTableST-5 can be used to reproduce the precipitation statistics by USEPA Level III ecoregion in table 9. The query qryPrecipStatsbyLatLong prompts the user for the decimal latitude and longitude of a site of interest and provides list of stations with precipitation statistics that are ranked by proximity (measured by great-circle distance) to the entered coordinates.

The USEPA program WDMUtil (Hummel and others, 2001) may be used to reformat USEPA precipitation datasets compiled for the BASINS program (U.S. Environmental Protection Agency, 2001a, 2007). These precipitation datasets are in a binary WDM format (Lumb and others 1988; Flynn and others, 1995; Hummel and others, 2001). The WDMUtil program (Hummel and others, 2001) may be used to export precipitation data from WDM-file formats to text-file formats that are more suitable for use with an updated or redeveloped version of the SYNOP program.

Selection and Regionalization of Sites for Measuring Storm Precipitation

NOAA hourly-precipitation data stations were selected and regionalized to provide the data and information necessary for developing planning-level estimates of storm-event characteristics at ungaged sites. The primary criterion used for selecting NOAA hourly-precipitation data stations for synoptic storm-event analysis was the availability of a record of sufficient length within a common hydrological period. Maximizing the length of record and using a common hydrological period are competing objectives, however, because the number of potential sites with a common period of record decreases as the record length increases. The use of long record lengths, however, increases the accuracy and precision of estimates of hydrologic statistics and provides information to evaluate cycles or trends in hydrologic data (Haan, 1977; Stedinger and others, 1993; Helsel and Hirsch, 2002). Studies have shown that decades of precipitation data are necessary to generate representative design-storm statistics for a drainage basin (Alley, 1977; Church and others, 2003). The use of a common hydrological period maximizes the comparability of statistics among adjacent data-collection stations to limit the potential effects of climate variation or trends on the statistics of interest (Haan, 1977; Stedinger and others, 1993). In this study, 2,610 NOAA hourly-precipitation data stations with at least 25 years of record during the interval 1965–2006 were selected. About 12 percent of the stations have less than 30 years of record, 23 percent have 30 to 39 years of record, and about 65 percent have more than 39 years of record during the 42-year period. About 1.8 percent of the stations have records that end in 1989 or 1990. About 4.2, 4.3, and 6.0 percent of the stations have precipitation records that end in the intervals 1991–1995, 1996–2000, and 2001–2005, respectively. About 83.7 percent of the stations have precipitation records that include the year 2006.

The rain-zone regions developed by Driscoll, Palhegyi, and others (1989) and the USEPA Level III nutrient ecoregions developed by Omernik and others (2000) were selected to group stations within geographic regions to facilitate the selection of representative storm-event statistics for a site of interest. The rain-zone regions developed by Driscoll, Palhegyi, and others (1989) are commonly used to estimate storm-event statistics for runoff studies and BMP analysis. Use of the USEPA Level III nutrient ecoregions provides a consistent national context for the regionalization of environmental data because these ecoregions were defined on the basis of physiography, climate, hydrology, and other factors. The 15 rain-zone regions are shown with respect to the 84 USEPA Level III nutrient ecoregions and the state boundaries within the conterminous United States in figure 17. The number of ecoregions per rain zone averages about 11 with a range from 5 (for rain zone 12, the Pacific Southwest) to 21 (for rain zone 12, the Northwest Inland). Many ecoregions straddle multiple rain zones. The average areal density of the 2,610 NOAA hourly-precipitation data stations is about one station for every 280 mi^2. A GIS coverage of these NOAA hourly-precipitation data stations is provided in the GIS directory on the CD–ROM accompanying this report (table 2). Although this areal density results in an average radius of only about 9.4 mi per station, the stations are not evenly distributed throughout the Nation, but instead are clustered within populous areas (fig. 17).

Regional precipitation statistics may be sufficient for initial screening-level analyses, but the estimates may be refined by selecting statistics from one or more nearby precipitation-monitoring stations in areas with similar climatic characteristics; however, the closest monitoring stations (or station) may not be the best choices because they may not provide representative statistics. For example, Daly and others (1994) indicate that elevation and orientation to prevailing weather patterns can have a substantial effect on average annual precipitation values among nearby sites.

Table 9. Summary synoptic precipitation-event statistics for the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions in the conterminous United States. The statistics are calculated from data collected at selected hourly-precipitation data stations in each ecoregion during the period 1965–2006.

[COV, coefficient of variation defined as the standard deviation divided by the mean; S., southern; N., northern; storm events are defined as having a minimum volume of 0.1 inches of precipitation and a 6-hour minimum interevent time threshold]

Ecoregion			Number of storm events (per year)		Annual precipitation volume (inches)		Storm-event volume (inches)		Storm-event duration (hours)		Time between storm-event midpoints (hours)	
Number	Name	Number of stations	Mean	COV	Mean	COV	Mean	COV	Mean	COV	Mean	COV
1	Coast Range	42	62	0.27	57.35	0.34	0.94	1.19	14	0.97	144	1.96
2	Puget Lowland	14	63	0.25	37.93	0.29	0.59	1.06	12	0.90	136	1.58
3	Willamette Valley	14	64	0.20	36.23	0.27	0.57	1.06	12	0.92	136	1.78
4	Cascades	38	68	0.26	49.42	0.30	0.72	1.14	13	0.95	125	1.82
5	Sierra Nevada	37	30	0.37	30.73	0.45	0.98	1.19	12	0.97	308	1.90
6	Southern and Central California Plains and Hills	110	22	0.39	18.53	0.50	0.84	1.09	11	0.92	411	2.14
7	Central California Valley	13	23	0.34	11.75	0.41	0.49	0.91	10	0.83	421	2.10
8	Southern California Mountains	25	17	0.41	17.90	0.56	1.01	1.27	11	0.99	502	1.94
9	Eastern Cascades Slopes and Foothills	14	31	0.37	17.52	0.41	0.52	0.99	10	0.92	297	1.70
10	Columbia Plateau	26	29	0.31	11.76	0.34	0.36	0.78	8	0.78	351	1.43
11	Blue Mountains	15	30	0.36	12.34	0.38	0.38	0.83	8	0.84	308	1.51
12	Snake River Basin/High Desert	8	27	0.36	9.64	0.38	0.36	0.72	8	0.82	344	1.55
13	Northern Basin and Range	41	21	0.39	8.16	0.43	0.39	0.80	7	0.86	520	1.44
14	Southern Basin and Range	14	10	0.47	4.64	0.57	0.45	0.85	7	0.89	939	1.61
15	Northern Rockies	27	40	0.31	17.15	0.33	0.42	0.83	9	0.85	220	1.48
16	Montana Valley and Foothill Prairies	14	39	0.30	16.26	0.33	0.41	0.84	8	0.85	232	1.56
17	Middle Rockies	43	27	0.34	10.07	0.36	0.37	0.80	7	0.89	354	1.59
18	Wyoming Basin	24	17	0.40	6.19	0.40	0.37	0.81	7	0.92	559	1.57
19	Wasatch and Uinta Mountains	19	30	0.31	12.98	0.32	0.41	0.77	7	0.83	333	1.41
20	Colorado Plateaus	16	21	0.35	8.03	0.39	0.38	0.80	7	0.89	451	1.49
21	Southern Rockies	34	27	0.34	10.48	0.36	0.38	0.83	7	0.92	350	1.61
22	Arizona/New Mexico Plateau	25	20	0.35	8.02	0.37	0.39	0.80	6	0.93	459	1.62
23	Arizona/New Mexico Mountains	26	26	0.34	12.93	0.38	0.49	0.91	6	1.00	339	1.63
24	Southern Deserts	29	19	0.35	9.94	0.43	0.51	0.97	5	1.00	480	1.68
25	Western High Plains	69	24	0.32	12.70	0.36	0.54	0.98	6	0.99	388	1.78
26	Southwestern Tablelands	36	25	0.30	13.46	0.34	0.53	0.96	6	0.98	364	1.71
27	Central Great Plains	73	32	0.28	21.16	0.32	0.67	1.00	6	0.93	282	1.59
28	Flint Hills	9	38	0.26	27.47	0.30	0.72	1.01	6	0.87	228	1.48
29	Central Oklahoma/Texas Plains	47	37	0.24	29.05	0.29	0.78	1.03	6	0.93	233	1.30
30	Edwards Plateau	26	33	0.28	25.14	0.32	0.75	1.09	6	1.01	266	1.44

Table 9. Summary synoptic precipitation-event statistics for the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions in the conterminous United States. The statistics are calculated from data collected at selected hourly-precipitation data stations in each ecoregion during the period 1965–2006.—Continued

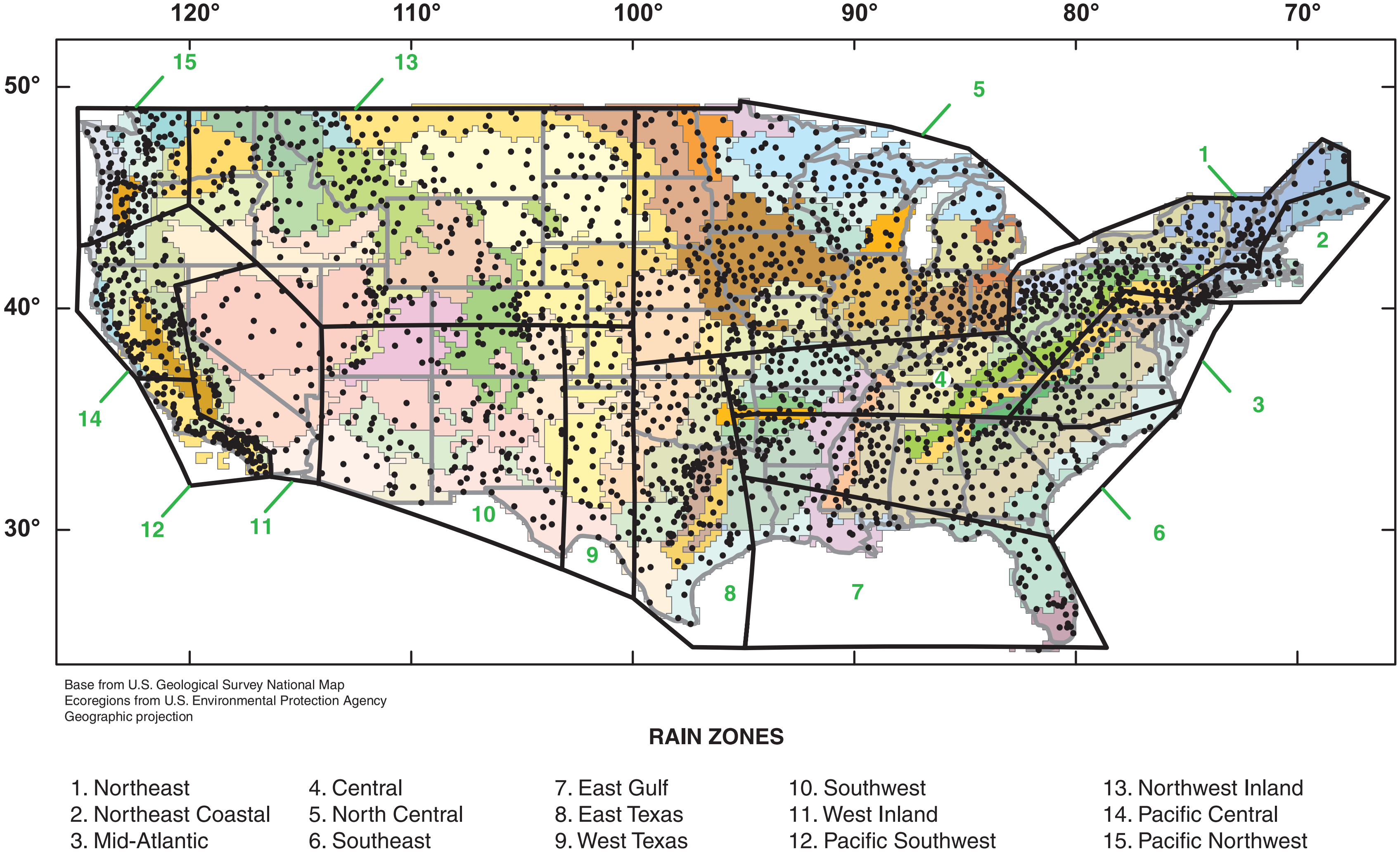
[COV, coefficient of variation defined as the standard deviation divided by the mean; S., southern; N., northern; storm events are defined as having a minimum volume of 0.1 inches of precipitation and a 6-hour minimum interevent time threshold]

Ecoregion			Number of storm events (per year)		Annual precipitation volume (inches)		Storm-event volume (inches)		Storm-event duration (hours)		Time between storm-event midpoints (hours)	
Number	Name	Number of stations	Mean	COV	Mean	COV	Mean	COV	Mean	COV	Mean	COV
31	Southern Texas Plains	7	22	0.37	16.70	0.42	0.75	1.11	6	1.06	393	1.58
32	Texas Blackland Prairies	27	39	0.25	32.02	0.30	0.81	1.06	7	0.96	217	1.26
33	East Central Texas Plains	8	38	0.27	30.23	0.32	0.80	1.06	6	0.96	226	1.28
34	Western Gulf Coastal Plains	20	47	0.23	39.53	0.28	0.84	1.22	7	1.02	195	1.23
35	South Central Plains	43	48	0.25	42.22	0.28	0.87	1.06	7	0.94	174	1.19
36	Ouachita Mountains	11	51	0.23	44.15	0.27	0.86	1.06	7	0.94	165	1.21
37	Arkansas Valley	11	49	0.23	39.93	0.29	0.81	1.04	7	0.93	174	1.22
38	Boston Mountains	9	48	0.24	38.40	0.29	0.80	1.03	7	0.92	176	1.20
39	Ozark Highlands	46	45	0.32	32.74	0.35	0.73	1.01	7	0.91	180	1.30
40	Central Irregular Plains	56	40	0.35	29.07	0.38	0.72	1.02	7	0.88	200	1.44
41	Canadian Rockies	8	41	0.29	18.49	0.32	0.45	0.98	9	0.92	214	1.36
42	Northwestern Glaciated Plains	33	22	0.35	10.29	0.40	0.46	0.96	7	0.96	431	1.86
43	Northwestern Great Plains	70	24	0.32	11.05	0.35	0.45	0.96	7	0.97	378	1.75
44	Nebraska Sandhills	10	29	0.31	15.54	0.34	0.54	0.95	7	0.97	315	1.79
45	Piedmont	65	52	0.26	38.83	0.28	0.75	1.01	7	0.94	161	1.24
46	Northern Glaciated Plains	29	27	0.30	14.86	0.35	0.54	0.95	6	0.93	331	1.84
47	Western Corn Belt Plains	97	39	0.29	24.40	0.33	0.63	0.99	7	0.88	223	1.56
48	Lake Agassiz Plain	11	27	0.34	14.68	0.37	0.54	0.99	6	0.90	324	1.96
49	Northern Minnesota Wetlands	5	32	0.39	16.02	0.40	0.50	1.06	8	0.87	262	1.82
50	Northern Lakes and Forests	54	43	0.32	21.45	0.33	0.51	0.93	8	0.86	195	1.40
51	Northern Central Hardwood Forests	26	39	0.33	21.35	0.34	0.55	0.96	7	0.87	219	1.56
52	Driftless Area	24	43	0.29	25.03	0.31	0.59	0.99	7	0.86	195	1.39
53	Southeastern Wisconsin Till Plains	11	48	0.24	25.99	0.26	0.54	0.97	8	0.84	182	1.35
54	Central Corn Belt Plains	41	48	0.25	29.36	0.28	0.61	0.96	7	0.87	176	1.27
55	Eastern Corn Belt Plains	58	52	0.25	30.63	0.27	0.59	0.91	7	0.86	159	1.18
56	S. Michigan/N. Indiana Drift Plains	33	50	0.26	26.89	0.28	0.54	0.95	8	0.86	166	1.27
57	Huron/Erie Lake Plains	9	48	0.31	24.18	0.30	0.51	0.93	8	0.85	173	1.20
58	Northeastern Highlands	61	56	0.27	33.93	0.29	0.61	0.93	9	0.82	148	1.09
59	Northeastern Coastal Zone	31	53	0.26	37.50	0.29	0.71	1.00	10	0.80	156	1.07
60	Northern Appalachian Plateau and Uplands	26	52	0.27	27.76	0.29	0.54	0.94	8	0.85	159	1.20

Table 9. Summary synoptic precipitation-event statistics for the 84 U.S. Environmental Protection Agency Level III nutrient ecoregions in the conterminous United States. The statistics are calculated from data collected at selected hourly-precipitation data stations in each ecoregion during the period 1965–2006.—Continued

[COV, coefficient of variation defined as the standard deviation divided by the mean; S., southern; N., northern; storm events are defined as having a minimum volume of 0.1 inches of precipitation and a 6-hour minimum interevent time threshold]

Ecoregion			Number of storm events (per year)		Annual precipitation volume (inches)		Storm-event volume (inches)		Storm-event duration (hours)		Time between storm-event midpoints (hours)	
Number	Name	Number of stations	Mean	COV	Mean	COV	Mean	COV	Mean	COV	Mean	COV
61	Erie/Ontario Lake Hills and Plain	29	58	0.24	30.73	0.26	0.53	0.92	8	0.85	144	1.20
62	North Central Appalachians	27	52	0.35	30.59	0.36	0.59	0.96	9	0.85	149	1.28
63	Middle Atlantic Coastal Plain	21	51	0.29	37.24	0.31	0.72	1.05	7	0.89	161	1.32
64	Northern Piedmont	29	50	0.31	34.63	0.33	0.70	1.03	9	0.85	161	1.12
65	Southeastern Plains	91	53	0.26	43.76	0.29	0.82	1.04	7	0.93	156	1.29
66	Blue Ridge Mountains	19	56	0.28	41.45	0.30	0.74	1.03	8	0.93	145	1.27
67	Central Appalachian Ridges and Valleys	62	52	0.28	33.56	0.30	0.64	0.97	8	0.87	159	1.24
68	Southwestern Appalachians	15	57	0.28	43.89	0.29	0.77	0.98	7	0.89	142	1.25
69	Central Appalachians	49	57	0.31	31.88	0.32	0.56	0.95	8	0.89	142	1.32
70	Western Allegheny Plateau	52	52	0.33	28.98	0.33	0.56	0.93	8	0.88	152	1.28
71	Interior Plateau	48	54	0.28	37.50	0.30	0.70	0.98	7	0.89	153	1.29
72	Interior River Lowland	48	47	0.29	31.45	0.31	0.67	0.97	7	0.88	175	1.31
73	Mississippi Alluvial Plain	27	51	0.26	43.52	0.30	0.86	1.06	6	0.95	165	1.31
74	Mississippi Valley Loess Plains	26	53	0.25	44.87	0.28	0.84	1.04	7	0.92	157	1.22
75	Southern Coastal Plain	44	57	0.22	45.52	0.25	0.81	1.10	6	1.01	152	1.35
76	Southern Florida Coastal Plain	14	60	0.25	43.22	0.28	0.73	1.17	5	1.03	146	1.63
77	North Cascades	15	59	0.32	47.99	0.37	0.77	1.18	12	0.91	150	1.78
78	Klamath Mountains	33	40	0.34	33.14	0.41	0.83	1.17	13	0.94	210	2.01
79	Madrean Archipelago	4	22	0.36	11.66	0.42	0.54	0.85	5	0.94	404	1.74
80	Northern Basin and Range	13	22	0.41	7.73	0.42	0.35	0.76	7	0.88	406	1.47
81	Sonoran Basin and Range	16	13	0.48	7.31	0.56	0.52	0.95	7	0.94	969	1.57
82	Laurentian Plains and Hills	12	54	0.27	33.33	0.31	0.62	0.94	9	0.82	154	1.08
83	Eastern Great Lakes and Hudson Lowlands	32	56	0.26	29.00	0.28	0.52	0.91	9	0.84	150	1.16
84	Atlantic Coastal Pine Barrens	16	52	0.25	36.03	0.29	0.70	0.98	9	0.80	158	1.02



Base from U.S. Geological Survey National Map
Ecoregions from U.S. Environmental Protection Agency
Geographic projection

RAIN ZONES

1. Northeast
2. Northeast Coastal
3. Mid-Atlantic
4. Central
5. North Central
6. Southeast
7. East Gulf
8. East Texas
9. West Texas
10. Southwest
11. West Inland
12. Pacific Southwest
13. Northwest Inland
14. Pacific Central
15. Pacific Northwest

Figure 17. Spatial distribution of 2,610 hourly-precipitation data stations (black dots) with respect to the U.S. Environmental Protection Agency rain zones (Driscoll, Palhegyi, and others, 1989) and the U.S. Environmental Protection Agency (2003) Level III ecoregions (colored polygons) that have been discretized to a 15-minute grid in the conterminous United States. Ecoregions are identified on the plate useco.pdf on the CD–ROM accompanying this report.

Statistical Characterization of Storms

SELDM uses the average, the standard deviation, and minimum-value threshold of storm-event parameters generated by SYNOP (fig. 14) to generate precipitation inputs. The implementation of the Poisson process to define storm-event precipitation statistics requires a consistent method to define the occurrence of independent events from available precipitation records (appendix 3). Storm events (fig. 14) are commonly defined in terms of a minimum IET that defines the wet and dry periods to aggregate hourly precipitation records into independent storm events (Driscoll and others, 1979; Goforth and others, 1983; Adams and others, 1986; Strecker and others, 1989; Adams and Papa, 2000; Asquith and others, 2006). Larger values of the minimum IET increase the statistical independence of subsequent events. Increasing the value of the minimum IET decreases the number of independent events in a precipitation record. The choice of IET also determines the best-fit probability distribution. The mean, standard deviation, and skew of the duration, volume, and average precipitation intensity of the storm events are affected by the choice of IET. Driscoll, Palhegyi, and others (1989) adopted an IET value of 6 hours in an analysis of 40 years of data from 136 sites as a consistent basis to define storm-event statistics in the 15 rain zones within the conterminous United States.

Storm events also are characterized by a minimum precipitation depth (Driscoll and others, 1979; Goforth and others, 1983; Adams and others, 1986; Schueler, 1987; Driscoll, Palhegyi, and others, 1989; Strecker and others, 1989; Driscoll and others, 1990a,b; Adams and Papa, 2000). A minimum precipitation threshold (commonly 0.01 in.) is needed to trigger a recorded measurement (Alley, 1977). In urban and highway-runoff studies, however, the emphasis is on runoff-generating events that affect the operation of BMPs and cause stormwater discharges into receiving waters. Thus, a minimum precipitation volume of 0.1 in. is commonly used to identify runoff-generating events from highly impervious highway and urban land. Selection of this non-zero minimum precipitation depth increases the event-mean volume, decreases the COV of event volumes, and decreases the number of events per year (Driscoll, Palhegyi, and others, 1989; Strecker and others, 1989). Use of the 0.1 in. threshold also reduces the mean annual precipitation volume because many small storms are not included in the annual total.

Results of the synoptic precipitation analysis for the 1965–2006 and 1949-1987 periods are included in table 8. The new precipitation statistics were developed by using the same analysis methods and the same storm-event definitions (a minimum storm volume of 0.1 in. and an IET of 6 hours) used by Driscoll, Palhegyi, and others (1989). However, the average rain-zone statistics for the datasets for the two periods differ substantially. For example, the medians of the percent differences for the number of storm events per year, the annual precipitation, storm-event volume, storm-event duration, and the time between storm-event midpoints for all stations (table 8) are about 17, 10, -11, 23, and -14 percent, respectively. If the same comparisons are made for these average rain-zone statistics on the basis of only the 129 stations that are common to both datasets, the medians of the percent differences for the same statistics are about -1.5, 0.24, -3.0, 8.4, and 1.4, respectively. This disparity indicates that the differences between the datasets in table 8 result primarily from the use of many more precipitation data stations in the current study.

The analysis of storm-event statistics for individual pairs of the 129 stations that are common to both datasets indicates that the differences between the datasets for the two time periods are statistically significant, but the magnitudes of the differences are not substantial. The results of the sign test (Helsel and Hirsch, 2002) for the paired-station data indicate that differences in the average annual precipitation, average storm-event intensity, volume, and the time between storm-event midpoints are statistically significant at the 95-percent confidence level, but the differences in the number of storms and the storm-event durations are not statistically significant at the 95-percent confidence level. The medians of the percent differences in individual paired-station statistics between the datasets are about 0.0, 0.67, -0.48, 1.4, and 0.64 percent for the number of storm events per year, the annual precipitation, storm-event volume, storm-event duration, and the time between storm-event midpoints, respectively. The scatterplot diagrams in figure 18 also indicate variations in statistics among these stations, but the two datasets have comparable values for most stations. Correlation coefficients between statistics for the 129 stations that are common to both datasets are about 0.99 for the number of storm events per year, the annual precipitation volume, and the storm-event precipitation volume; 0.98 for the time between storm-event midpoints; and about 0.94 for the storm-event durations. These results indicate that the large differences in rain-zone statistics between the 1949–1987 and the 1965–2006 periods in table 8 are primarily the result of the increased number of stations used to calculate these statistics for the 1965–2006 dataset. Rain-zone statistics from the 1965–2006 dataset developed in the current study may be more representative of conditions in each rain zone than statistics from the 1949–1987 dataset because the newer dataset has more than ten times the number of stations in each rain zone.

The statistics from the additional sites in the new dataset show that there may be large variations in storm-event statistics within each rain zone. Boxplots of the average number of storms per year, the average storm-event precipitation volume, and the average storm-event duration indicate that the 15-zone system is useful for categorizing some variability in storm-event statistics, but conditions at some sites in each rain zone may not be well characterized by the average values for that zone (fig. 19). For example, among different stations in rain-zone 1 (the Northeast), the average number of storm events per year ranges from 30 to 88 events per year, the average storm-event volumes range from 0.43 to 0.85 in., and the average storm-event durations range from about 5 to about 13 hours. Rain zones with the largest range in event statistics

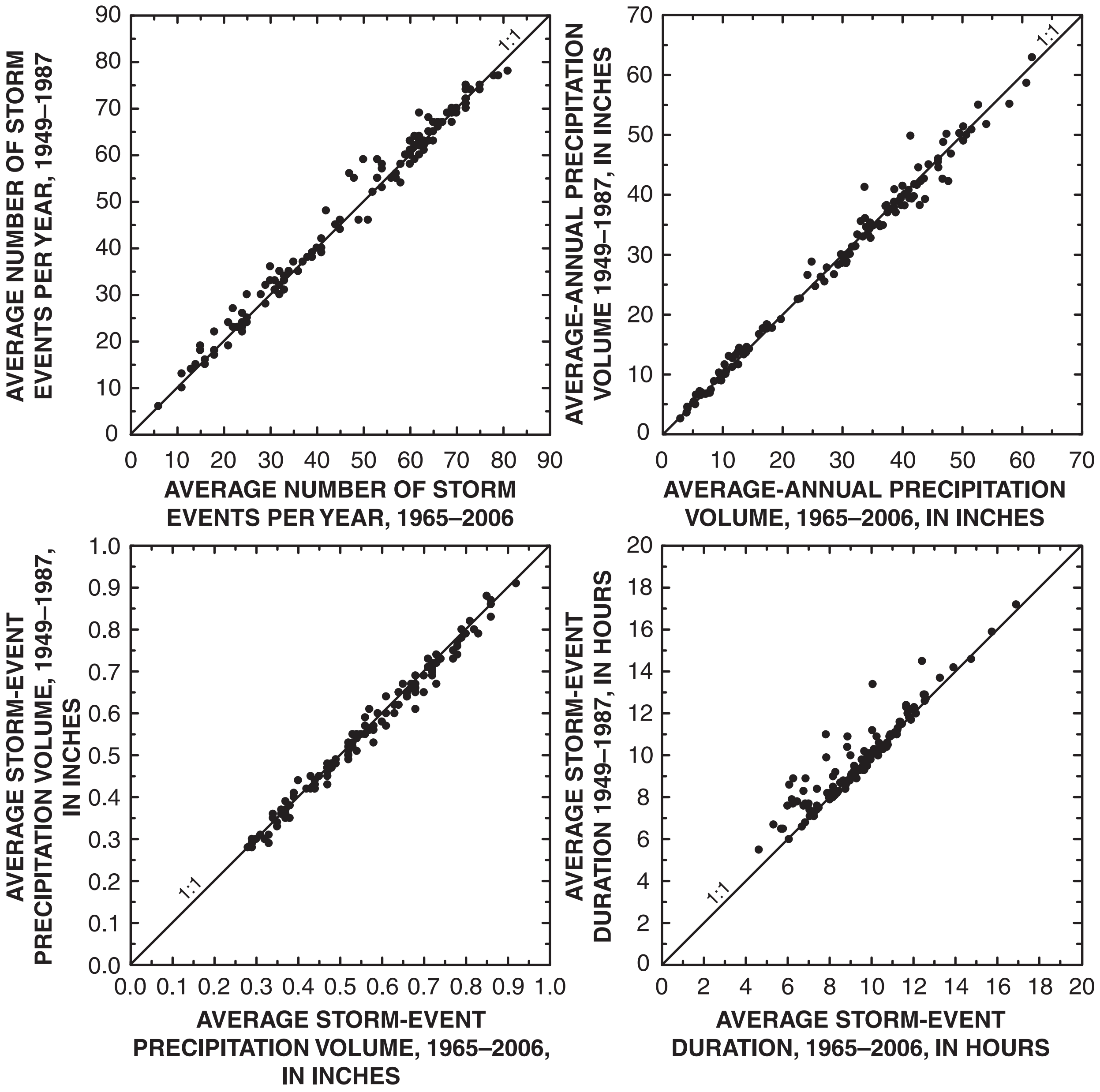


Figure 18. Scatterplot diagrams showing relations between storm-event statistics from 129 precipitation-gaging stations in the conterminous United States for the periods 1949–1987 and 1965–2006.

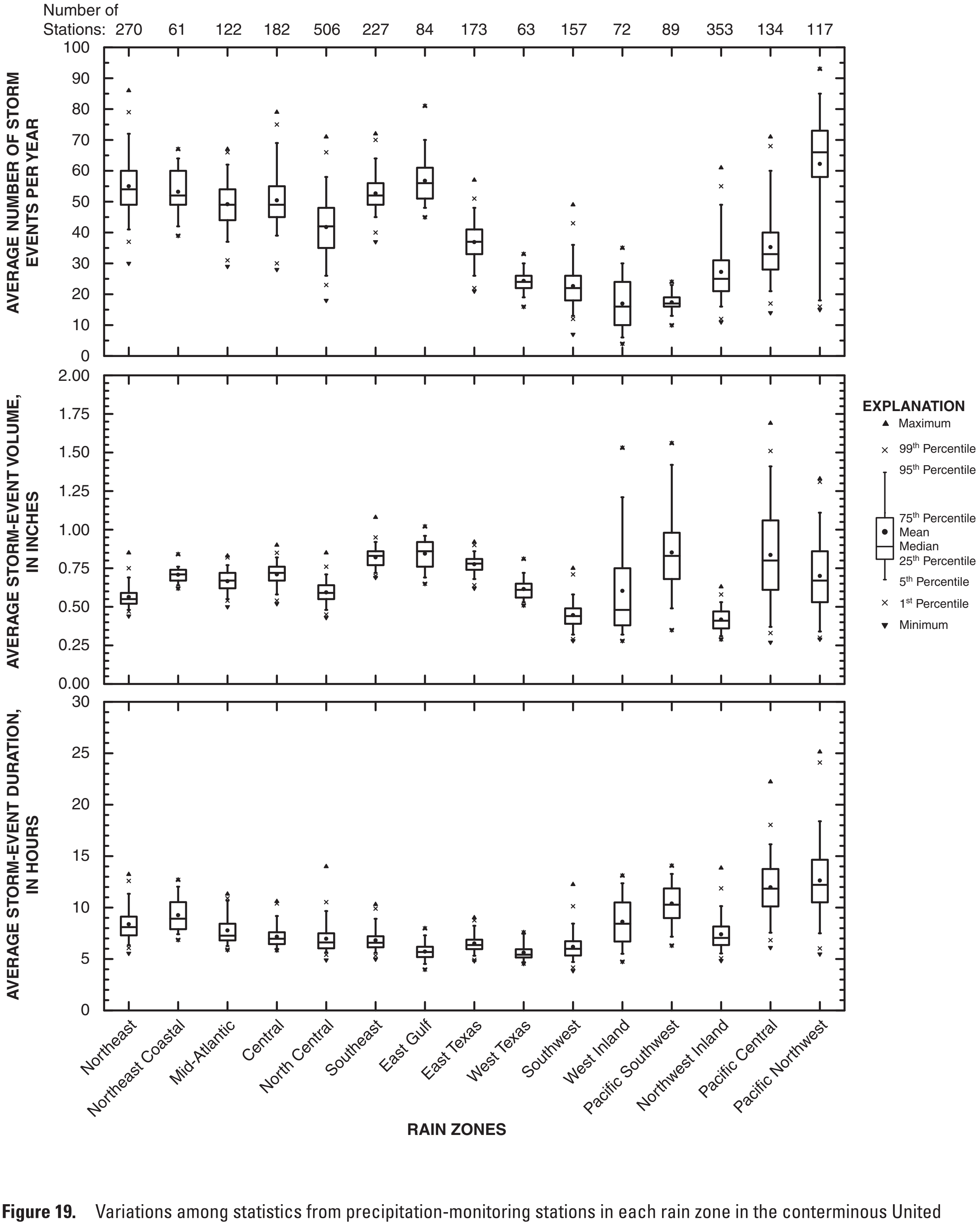


Figure 19. Variations among statistics from precipitation-monitoring stations in each rain zone in the conterminous United States for the average number of storms per year, the average storm-event volume, and the average storm-event duration during 1965–2006.

among stations are rain zone 15 (the Pacific Northwest) with averages of 15 to 93 events per year, rain zone 14 (the Pacific Central) with average storm-event volumes between 0.27 and 1.69 in., and rain zone 15 with average storm-event durations between about 5 and 25 hours.

Ecoregions also may be used to select stations that represent storm-event statistics at a site of interest (table 9). In theory, storm-event statistics from 84 ecoregions would be much more representative of local conditions near a site of interest than the 15 rain zones because ecoregions are smaller, and the delineation process for ecoregions includes several factors that may account for variations in precipitation at different sites within the 15 rain zones defined by Driscoll, Palhegyi, and others (1989). Statistical variations for different stations in each ecoregion, however, are not substantially less than statistical variations for different stations in each rain zone.

In SELDM, the user may select synoptic statistics by rain zone, by ecoregion, by selecting statistics from nearby monitoring stations, or by entering user-defined statistics. Regional estimates of storm-event statistics may be sufficient for a screening-level analysis, but a more focused approach may be necessary to refine these estimates for local conditions. Local variations in precipitation can introduce large uncertainties in the application of precipitation data (appendix 1). The use of a cluster analysis like the one done by Palecki and others (2005) or a regression analysis like the one done by Daly and others (1995) may improve local predictions. The results of the exploratory data analysis with the SYNOP statistics developed in this study, however, indicate that, for example, station elevation is not an effective predictor variable to improve estimates of storm-event statistics within a rain zone because station elevation is not highly correlated with any storm-event statistic within these large geographic areas. A detailed analysis to better characterize local variations in precipitation statistics is beyond the scope of the current study. If refined estimates of precipitation statistics are needed for a given site, other methods such as areal averaging, Thiessen polygons, contouring, or a reciprocal-distance-squared method may be used (Chow and others, 1988). However, Singh and Chowdhury (1986) indicate that the selection of a particular method for estimating precipitation statistics for a site within a network of nearby stations is not critical if the period of interest is a year or longer. Synoptic storm-event statistics for each station are available in the SiteStormV01.mdb database on the CD–ROM accompanying this report. The station name, latitude, longitude, and elevation are associated with the statistics so that local estimates can be based on statistics from nearby precipitation-monitoring stations.

Estimating Storm-Event Statistics at Ungaged Sites

Storm-event statistics indicating the number of storms per year, the volume of precipitation per storm, the duration of storm events, and the time between the midpoints of storm events are needed for stochastic analysis of potential effects of highway runoff on receiving waters. These storm-event statistics are used with runoff-coefficient statistics for the highway and the upstream watershed to estimate the proportion of stormflow that originates from each area (figs. 1, 2). Stochastic analysis of precipitation statistics and runoff-coefficient statistics provide a population of stormflow volumes.

Storm-event arrivals are commonly modeled as a Poisson process (appendix 3), which is a statistical method of modeling discrete events that occur on a continuous time scale (Haan, 1977). Examples of data generated stochastically with statistics provided by Driscoll, Palhegyi, and others (1989) for 1949–1987 for runoff-producing storm events were examined for two substantially different climatic regions—rain zone 1 (Northeast) in the humid Northeastern United States and rain zone 11 (West Inland) in the arid Southwestern United States. Figure 20 shows a stochastically generated population of 200 values that have a Poisson distribution for the number of events per year for each of the two rain zones. The histogram for rain zone 1 indicates that the number of runoff-producing events varies from about 45 to about 95 events per year with an average of 70 events per year. The histogram for rain zone 11 indicates that the number of runoff-producing events varies from about 5 to about 30 events per year with an average of 14 events per year. The frequency curve on figure 20 shows the theoretical Poisson distribution for each rain zone for the mean number of storms per year. The histograms are not a perfect representation of the theoretical frequency curve because the numerical data-generation process provides a stochastic sample of the theoretical distribution that would converge with the theoretical distribution only after thousands of random selections (Devroye, 1986; Saucier, 2000; Gentle, 2003).

A number of right-skewed probability distributions have been used to analyze and model storm-event statistics in the literature (appendix 3); each distribution may be considered an approximation for characterizing the properties of local storm-event statistics. Selection of an approximate probability distribution for stochastic data generation in planning-level runoff-quality simulations is a balance between theoretical rigor (for example, Asquith and others, 2006) and practical estimation techniques that will be acceptable to practitioners, planners, and decisionmakers (U.S. Environmental Protection Agency, 1997). Figure 21 shows boxplots of the storm-event volume, duration, and IET that were stochastically generated with five different right-skewed probability distributions for 1,500 storm events (fig. 21) and then compared on a theoretical and practical basis. Each of these stochastic samples was generated with the same average, COV, and minimum precipitation values (Driscoll, Palhegyi, and others, 1989). The one- and two-parameter exponential distributions, the two-parameter lognormal distribution, the two-parameter gamma distribution, and the Pearson Type III distribution were selected for stochastic simulation because they are commonly used to describe precipitation statistics (appendix 3) and because they can be modeled with the statistics—in particular, the mean and COV— provided by SYNOP (Driscoll, Palhegyi, and others,

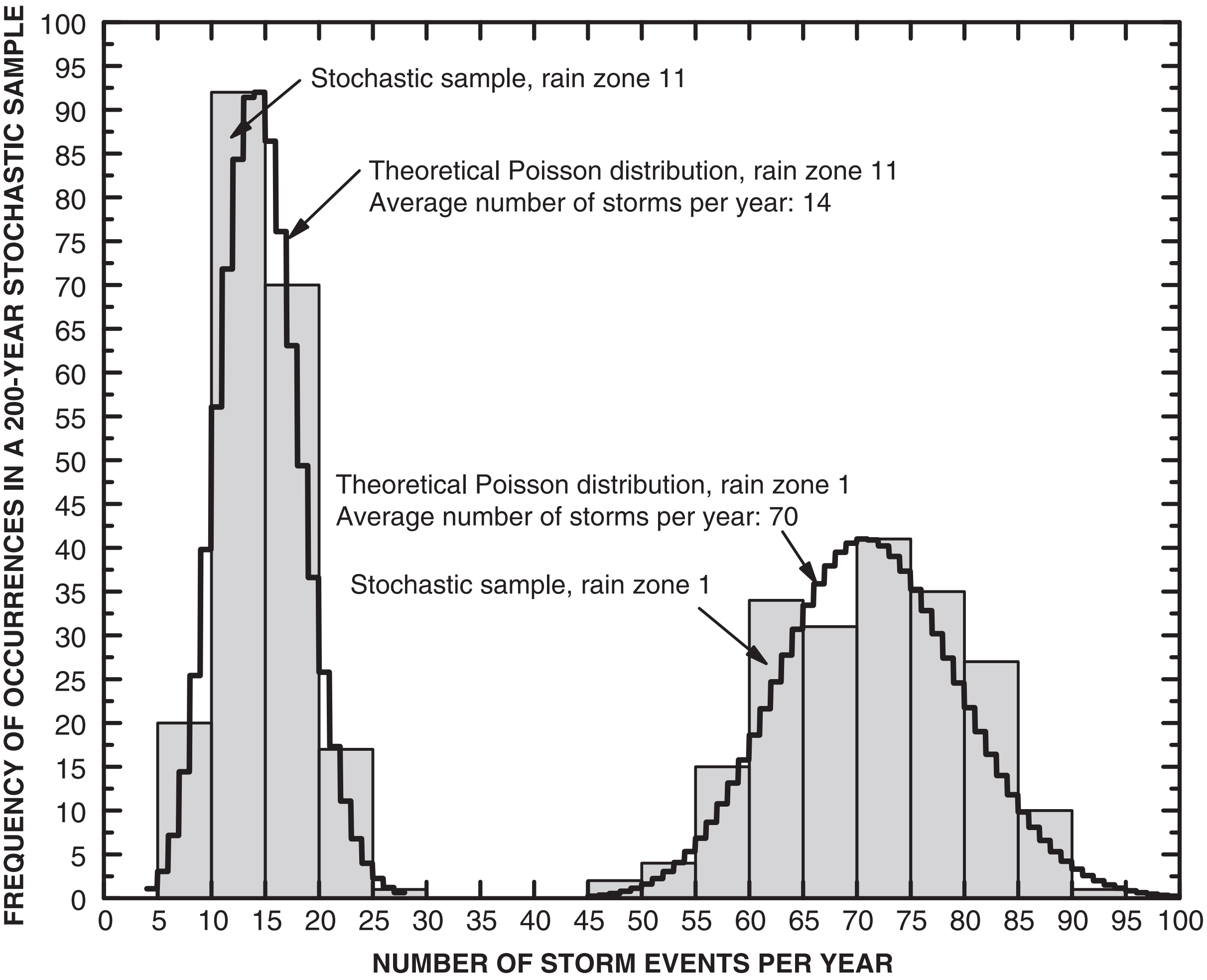


Figure 20. The theoretical Poisson distributions and stochastic samples of synoptic storm-event arrivals in the humid northeast (rain zone 1, Northeast) and in the arid southwest (rain zone 11, Western Inland) for storm events with a minimum duration of 1 hour, a minimum interevent time of 6 hours, and a minimum precipitation volume of 0.1 inch. (Statistics for the period 1949–87 from Driscoll, Palhegyi, and others, 1989)

1989). There are some differences among the interquartile ranges and symmetries of the distribution tails. Only the two-parameter exponential and Pearson Type III approximations preserve the lower limits of each storm statistic as defined by Driscoll and others (1989). The two-parameter exponential distribution was selected for use with SYNOP because it has the added benefit of being readily implemented in a stochastic data-generation algorithm (Devroye, 1986; Saucier, 2000; Gentle, 2003).

Storm-Event Hyetograph

A hyetograph is defined as the temporal distribution of rainfall within a storm event (Yen and Chow, 1980; 1983; Chow and others, 1988). Highway, urban, and BMP studies commonly use a rectangular hyetograph (appendix 3) to represent complete storm events rather than within-event processes for planning-level analysis. The use of a rectangular hyetograph provides an estimate of the average precipitation intensity during the storm as the quotient of total volume and total duration. In reality, however, precipitation rates are highly variable, and the average intensity is likely to substantially underrepresent the peak intensity. SELDM is a lumped-parameter, event-based model that does not evaluate within-storm precipitation characteristics. This approach is valid for developing planning-level runoff-quality estimates, but the lumped-parameter approach represented by the rectangular hyetograph should not be used for the hydraulic design of drainage structures. For example, the safety, hydraulic performance, and water-quality treatment provided by structural BMPs may be affected by variations in precipitation intensity and runoff flows during storms. Because the peak intensity shown on a triangular hyetograph that represents a given storm is twice that on the rectangular hyetograph for the same storm, the triangular hyetograph is a better approximation to use for hydraulic design (Yen and Chow, 1980; 1983).

Runoff-Coefficient and Stormflow-Hydrograph Statistics

SELDM uses runoff-coefficient statistics and basin characteristics as the basis for the stochastic generation of random runoff hydrographs from the highway site and the associated upstream basin (fig. 2). Runoff coefficients are used to estimate runoff flows from storm-event precipitation statistics. Basin characteristics are used with storm-event precipitation statistics to develop storm-event hydrographs. There are many methods, each with advantages and limitations, for estimating the variables that affect rainfall-runoff transformation processes and, therefore, storm-event flows. Many spatial and temporal complexities can influence net rainfall-runoff transformation processes; however, many of these hydrologic complexities cannot be quantified, even with complex models (Singh, 1977; Naef, 1981; Jakeman and Hornberger, 1993; Nix, 1994; Harremoës and Madsen, 1999; Shamir and others, 2005).

Runoff coefficients are calculated by dividing the total storm runoff (in watershed inches) by the basin-average precipitation (in inches) during a storm event. Runoff coefficients are needed to develop planning-level estimates of receiving water flow and quality at a site of interest. Runoff coefficients are a primary method for quantifying rainfall-runoff transformations for use in planning-level estimates of the effect of runoff on receiving waters (Athayde and others, 1983; Driscoll and others, 1986; Schueler, 1987; Driscoll and others, 1990b; Adams and Papa, 2000; U.S. Environmental Protection Agency, 2002). Runoff coefficients are used with SELDM because it is a lumped-parameter, event-based model that does not calculate the spatial distribution of precipitation and stormflow generation.

In this report, the term runoff will be used to describe storm-event flows (excluding prestorm base flow) regardless of the origin or mechanism of flow. Detailed characterization of different stormflow-generating mechanisms is not mathematically necessary for a statistical lumped-parameter model; however, a knowledge of different runoff mechanisms and the factors that affect runoff generation is important for selecting representative runoff-coefficient statistics. This knowledge also is important for selecting parameters that characterize the durations of the stormflow hydrographs from the highway site and the upstream basin. An overview of conceptual models for runoff-producing mechanisms is provided in Appendix 4 to help guide selection of rainfall-runoff parameters for use with SELDM.

SELDM conserves mass in the rainfall-runoff process for each individual storm event by limiting runoff-coefficient statistics to values between zero and one. The assumption is that runoff is defined as the total streamflow during a storm minus the prestorm flow. The prestorm flow will account for delayed discharge from previous events, and the calculated runoff coefficient will account for runoff from the current event. These assumptions should be sufficient for most conditions. Rain-on-snow events, however, may be an exception because the prestorm snowpack can contribute water that is not accounted for by prestorm streamflow and is in excess of input precipitation. A complex, spatially distributed energy and water-balance model for detailed characterization of stormflows and interevent flows would be needed to characterize such processes; however, substantial uncertainties in measuring and interpreting hydrologic processes (appendix 1) may overshadow the potential effects of relatively few rain-on-snow events. There also are substantial uncertainties in the application of streamflow (appendix 2) and precipitation (appendix 3) statistics to estimate stormflows at ungaged sites that overshadow potential effects of relatively few rain-on-snow events.

SELDM does not model within-storm processes, but information to develop storm-event hydrographs is necessary to estimate the proportion of upstream flows that occur concurrently with highway runoff. Efforts to develop planning estimates of the potential effects of highway runoff depend on the duration of runoff flows from the highway catchments rather than the duration of stormflows from the entire upstream basin because the dilution of runoff in the receiving waters is based on the amount of the concurrent flows that occur during the period of highway-runoff flow (or, if applicable, BMP discharge from the highway site). Rainfall-runoff statistics and knowledge of the hydraulic properties of the basin are needed to calculate the magnitude and timing of upstream runoff for each storm event.

Methods for Quantifying Rainfall-Runoff Processes

Many hydrologic studies have been designed to quantify rainfall-runoff transformation processes at gaged and ungaged sites. The complexities and variability in hydrologic processes that affect rainfall-runoff transformations limit the accuracy and precision of most estimation methods. Methods used to quantify rainfall-runoff processes include watershed-simulation models, curve number methods, statistical methods to estimate runoff coefficients, and regression-on-basin characteristics. SELDM is not a watershed simulation model; it is a lumped-parameter model. The Soil-Conservation Service Curve Number (CN) method commonly is used to design culverts, highway structures, and structural stormwater BMPs

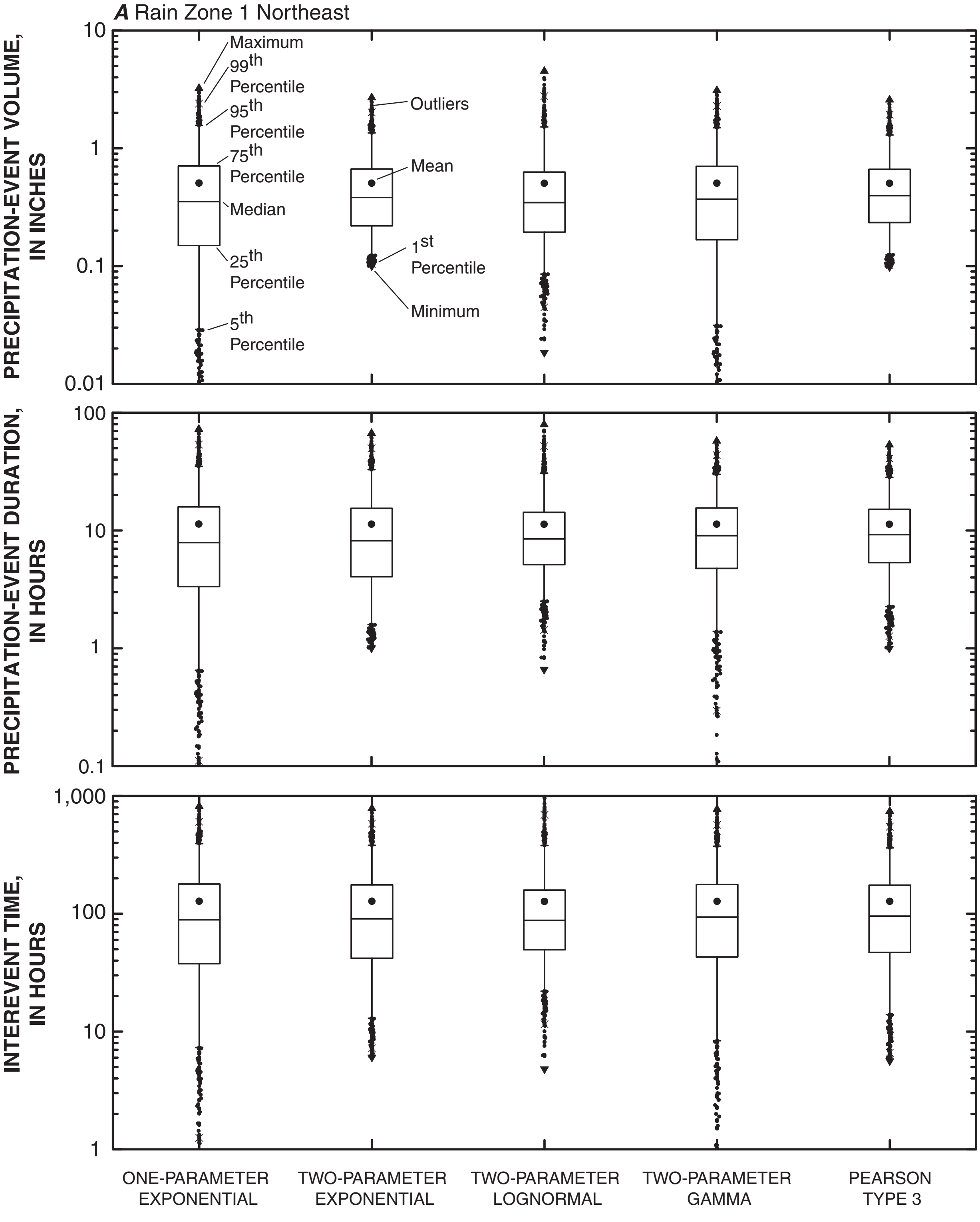


Figure 21. Examples of stochastic data generated using five right-skewed probability distributions for precipitation-event volume, duration, and the time between storms. Populations of 1,500 randomly generated data points are shown for *(A)* rain zone 1, Northeast, and *(B)* rain zone 11, West Inland. (Statistics from Driscoll, Palhegyi, and others, 1989, are documented in table 8.)

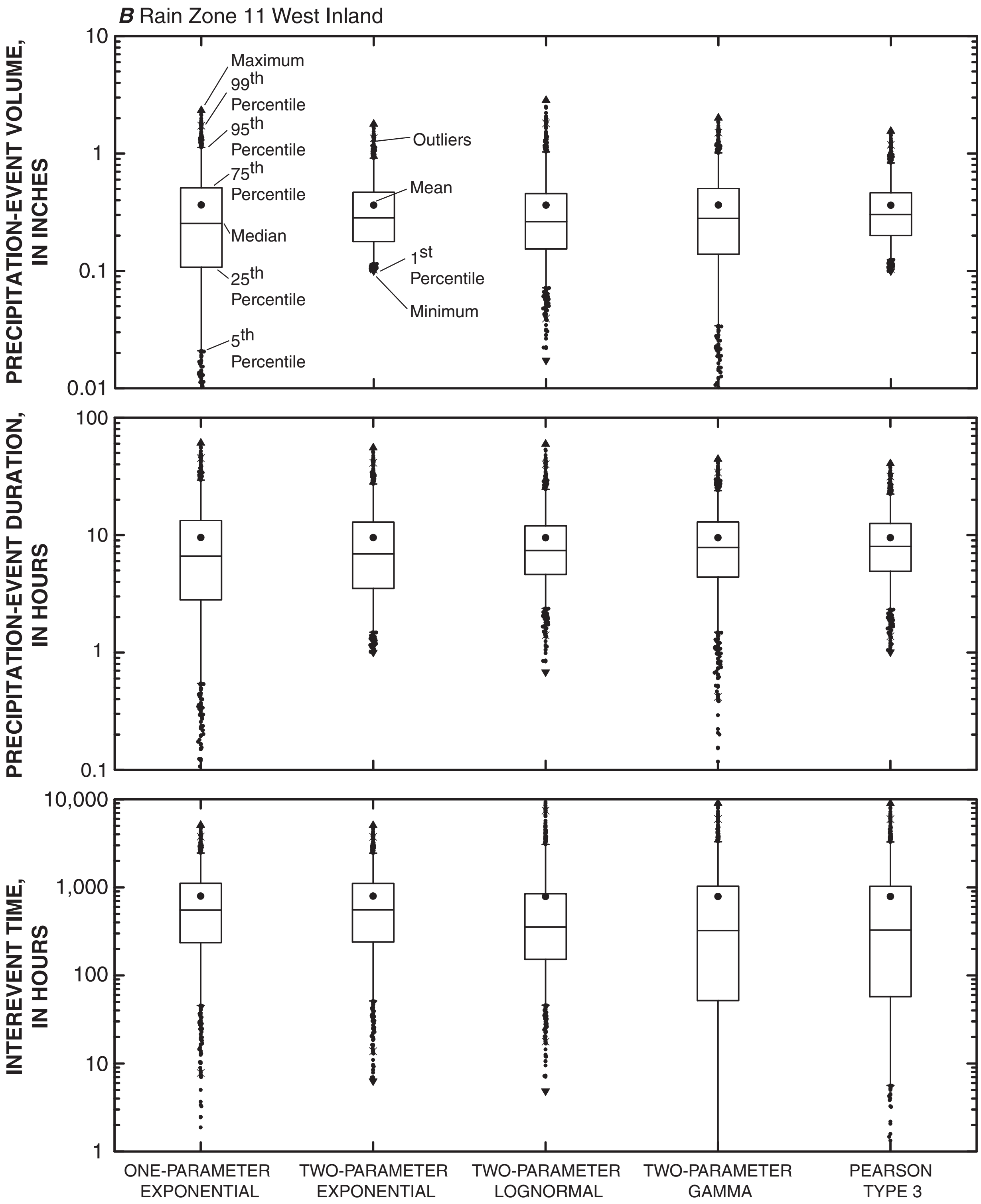


Figure 21. Examples of stochastic data generated using five right-skewed probability distributions for precipitation-event volume, duration, and the time between storms. Populations of 1,500 randomly generated data points are shown for *(A)* rain zone 1, Northeast, and *(B)* rain zone 11, West Inland. (Statistics from Driscoll, Palhegyi, and others, 1989, are documented in table 8.)—Continued

(McCuen and others, 2002). The CN method provides conservative runoff estimates for large storm events, but rainfall-runoff analyses in this study indicate that the CN method is not well suited for use with SELDM because runoff from frequent small storms (less than about 1 in.) is not well characterized by the CN method (appendix 5). Runoff generation in the SELDM model is based on user-defined statistics for generating a population of runoff coefficients. Finally, the use of statistics for all daily mean streamflows to model prestorm flows (shown schematically in fig. 3) can be used with selected runoff-coefficient statistics to represent specific hydrologic conditions.

Runoff coefficients are expected to vary randomly from storm to storm with variations in antecedent conditions and to vary from site to site as a function of impervious area (Duckstein and others, 1972; Driscoll and others, 1979; Athayde and others, 1983; Schueler, 1987; Driscoll and others, 1990b; Barrett and others, 1998; Becciu and Paoletti, 2000). It is widely recognized that runoff coefficients are random variables, but site-average values commonly are used for analysis and design purposes. If runoff-coefficient estimates are available for multiple sites, site characteristics such as land use or imperviousness may be used to transfer statistical runoff-coefficient estimates to similar ungaged sites.

Runoff coefficients have been approximated as random variables with positively skewed probability distributions with a lower bound of zero. For example, Athayde and others (1983) indicated that runoff coefficients at individual study sites in the NURP are well approximated by a lognormal distribution. Driscoll and others (1990b) also concluded that runoff coefficients from individual sites could be characterized as random lognormal variables. The lognormal distribution has a lower bound of zero but does not have a maximum limit. Duckstein and others (1972) used a gamma distribution to characterize the runoff coefficient as an independent random variable. Because the gamma distribution does not have an upper bound, Duckstein and others (1972) proposed the use of a truncated gamma distribution for stochastic rainfall-runoff modeling. Driscoll and others (1979) indicated that precipitation and runoff volume both follow a gamma distribution. As a result, the runoff coefficient may follow a beta distribution, which is the ratio of two variables that have gamma distributions. Several studies have used a beta distribution to model runoff coefficients because it can be defined with a lower bound of zero and an upper bound of one (Gottschalk and Weingartner, 1998; Franchini and others, 2005; Merz and others, 2006). The triangular distribution may be used because it is a simple and flexible alternative to the beta distribution (Johnson, 1997; Saucier, 2000). Becciu and Paoletti (2000) used a normal distribution truncated at zero and one to characterize the runoff coefficient as a random variable.

Identifying a single probability distribution to model runoff coefficients is complicated by several factors. The runoff coefficient for each storm is the quotient of basin-average rainfall and total runoff, but populations of each parameter may be characterized by different probability distributions. Identifying an appropriate probability distribution also may be complicated by random and systematic variations in the measurements of rainfall and runoff at different sites (appendix 1). Runoff coefficients commonly are derived as a component of urban-runoff-quality studies, but the cost of water-quality sampling and analyses commonly limits the number of storm events that can be monitored in many rainfall-runoff quality investigations. Limited sample size may confound efforts to identify a particular probability distribution with a high degree of confidence. At a given site, different storms may fit into different runoff-coefficient populations because different areas of the basin may be contributing runoff through different runoff-generating mechanisms (appendix 4). For example, Guo and Adams (1998a) and Chen and Adams (2005, 2007) used conditional probability methods with precipitation-volume statistics to estimate the proportion of precipitation events that would produce no runoff, only impervious-surface runoff, and both impervious- and pervious-surface runoff. Further complications may arise in the effort to identify a characteristic probability distribution of runoff coefficients because of the potential effects of large variations in climate and basin characteristics at different study sites.

Runoff-coefficient regression models are developed using site characteristics to estimate a site-specific runoff coefficient. These regression methods provide estimates of the average runoff coefficient for a given drainage basin on the basis of relatively few easily available site characteristics. Average runoff coefficients commonly are used to predict runoff volumes because of uncertainties in measurements made during individual storms (Strecker and others, 2001; Church and others, 2003).

The regression models that are used to estimate site-average runoff coefficients are commonly untransformed models that are based on the percentage or fraction of total impervious area (TIA) in the basin. Table 10 is a summary of runoff-coefficient regression equations from two highway-runoff studies (Driscoll and others, 1990b; Granato and Cazenas, 2009) and five urban-runoff studies (Goforth and others, 1983; Schueler, 1987; Urbonas and Guo, 1989; Barrett and others, 1998; Becciu and Paoletti, 2000). The drainage areas used to develop these equations ranged from 0.05 to 28,416 acres (0.00008 to 44.4 mi^2), and the TIA fraction varied from 0.01 to 1.

Comparison of runoff coefficients estimated with the different regression equations in table 10 indicates substantial variations over the full range of TIA fractions (fig. 22). For a TIA value of 0.0, the estimated average runoff coefficients range from 0.01 to 0.15 with a median of 0.08; for a TIA value of 0.5, the estimated average runoff coefficients range from 0.325 to 0.525 with a median of about 0.42; and for a TIA value of 1.0, the estimated average runoff coefficients range from 0.57 to 0.95 with a median of about 0.89. These differences may reflect real differences among datasets (for example, differences in the distributions of basin size, land-use characteristics, climate, and active runoff-generating mechanisms); the effects of different base-flow-separation techniques

Table 10. Regression equations developed to estimate runoff coefficient statistics from total impervious area.

[TIA, total impervious area fraction (0–1); Rv, runoff coefficient (volumetric); --, not reported; USEPA, United States Environmental Protection Agency; NURP, National Urban Runoff Program]

Reference	Number of sites	Drainage-area range (acres)	TIA range	Equation	Comments
Barrett and others, 1998	18	0.64 to 3,584	0.01 to 0.95	Average Rv = 0.0125 + 0.5677 TIA + 0.3428 TIA^{2}	Urban-runoff monitoring sites in Austin, TX
Becciu and Paoletti, 2000	21	4.67 to 626	0.06 to 0.97	Average Rv = 0.08 + 0.49 TIA Standard deviation Rv = 0.03 + 0.2 TIA	Experimental urban catchments in many locations worldwide
Driscoll and others, 1990b	15	0.28 to 106	0.27 to 1	Average Rv = 0.10 + 0.70 TIA	Highway-runoff-monitoring sites, nationwide
Goforth and others, 1983	--	--	--	Average Rv = 0.15 + 0.75 TIA	Storage, treatment, overflow, runoff model (STORM) default equation
Granato and Cazenas, 2009	83	0.05 to 106	0.27 to 1	Average Rv = 0.08 + 0.67 TIA	Highway-runoff-monitoring sites, nationwide
Schueler, 1987	76	1 to 28,416	0.01 to 1	Average Rv = 0.05 + 0.9 TIA	Urban-runoff data from the USEPA NURP study (Athayde, 1983)
Urbonas and Guo, 1989	76	1 to 28,416	0.01 to 1	Average Rv = 0.04 + 0.774 TIA – 0.780 TIA^{2} + 0.858 TIA^{3}	Urban-runoff data from the USEPA NURP study (Athayde, 1983)

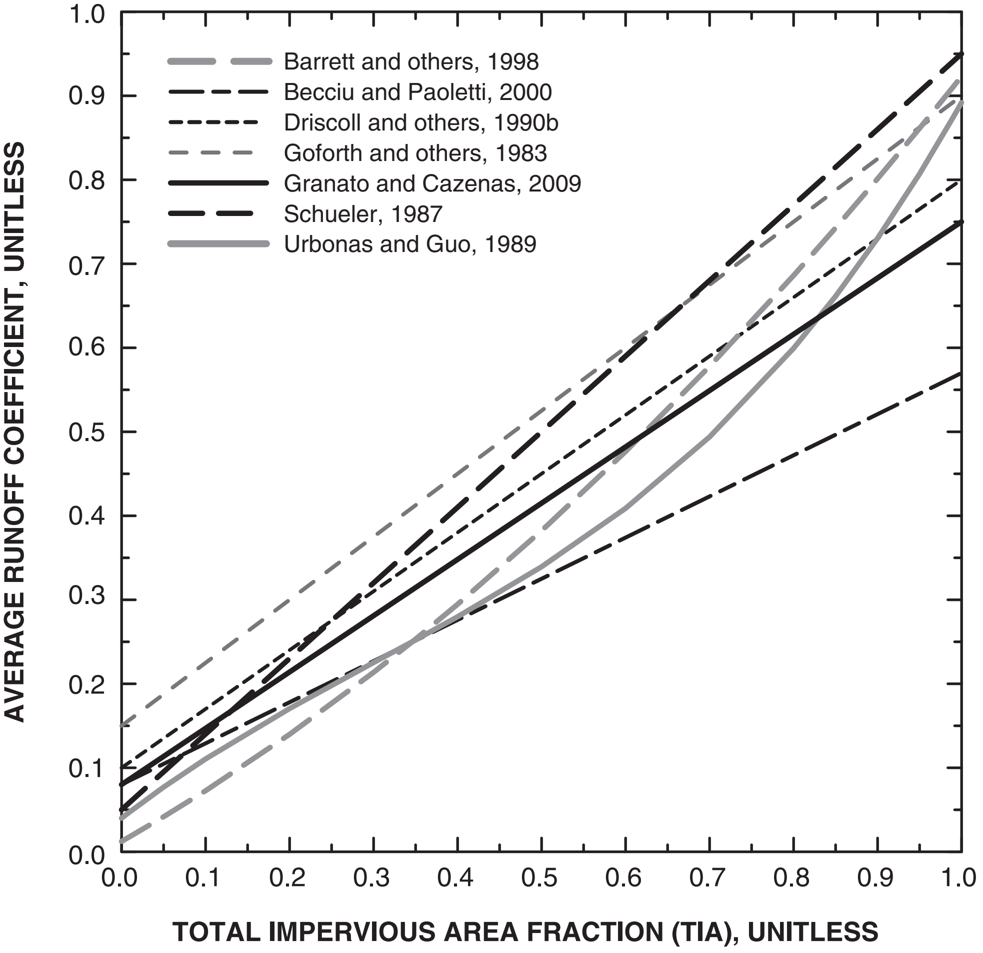


Figure 22. Regression equations for estimating average runoff coefficients from the total impervious area fraction. These regression equations are documented in table 10.

on calculated runoff amounts; differences in analytical methods that can introduce bias in some datasets (for example, calculating site-average runoff coefficients with or without storm-event runoff coefficients that are greater than one); and artifacts of the regression analysis (for example, the range or the distribution of TIA values within the different datasets). These differences and other potential sources of uncertainty may affect runoff-coefficient statistics and, therefore, these regression equations (appendix 1).

The standard deviation also is needed to characterize a population of runoff coefficients. Becciu and Paoletti (2000) published an equation for the standard deviation of runoff coefficients. The linear regression equation developed to predict the standard deviation of the runoff coefficients has a slope of 0.2 (times the TIA fraction) and an intercept of 0.03. This equation yields a standard deviation for the runoff coefficients of about 0.03 (COV 0.375) with a TIA fraction of 0.0, about 0.13 (COV 0.4) with a TIA fraction of 0.5, and about 0.23 (COV 0.4) with a TIA fraction of 1. Becciu and Paoletti (2000) was the only study found in the literature search to have published an equation for the standard deviation of runoff coefficients, but other studies in table 10 include information about the variability of runoff coefficients. Driscoll and others (1990b) published the medians and COVs of runoff coefficients for highway-runoff sites; the COVs ranged from 0.18 (for the site with a median runoff coefficient of 0.81) to 1.92 (for the site with a median runoff coefficient of 0.35). The median COV of the runoff coefficients in this study was 0.58. Athayde and others (1983) published the medians and COVs of runoff coefficients for the NURP sites (later used by Schueler, 1987, and Urbonas and Guo, 1989). The COVs of runoff coefficients for the NURP sites ranged from 0.19 (for the site with a median runoff coefficient of 0.82) to 6.64 (for the site with a median runoff coefficient of 0.17). Granato and Cazenas (2009) published the Highway Runoff Database, which includes queries to calculate the runoff-coefficient statistics. The COVs of runoff coefficients for these sites ranged from 0.007 (for the site with a median runoff coefficient of 0.7) to 1.97 (for the site with a median runoff coefficient of 0.56).

Sources of Data on Rainfall-Runoff Processes

Hydrologic-basin characteristics, information about land use and land cover, and rainfall-runoff data are necessary for characterizing rainfall-runoff transformation processes at gaged and ungaged sites. Hydrologic-basin characteristics such as drainage area, basin slope, and climate indicate the potential amount of runoff and the temporal distribution of stormflows from a given precipitation event. Information about land use and land cover indicates the proportion of rainfall that may generate storm flows from a given precipitation event and the temporal distribution of stormflows during a precipitation event. Rainfall-runoff data provide the information necessary to estimate runoff characteristics on the basis of these basin characteristics.

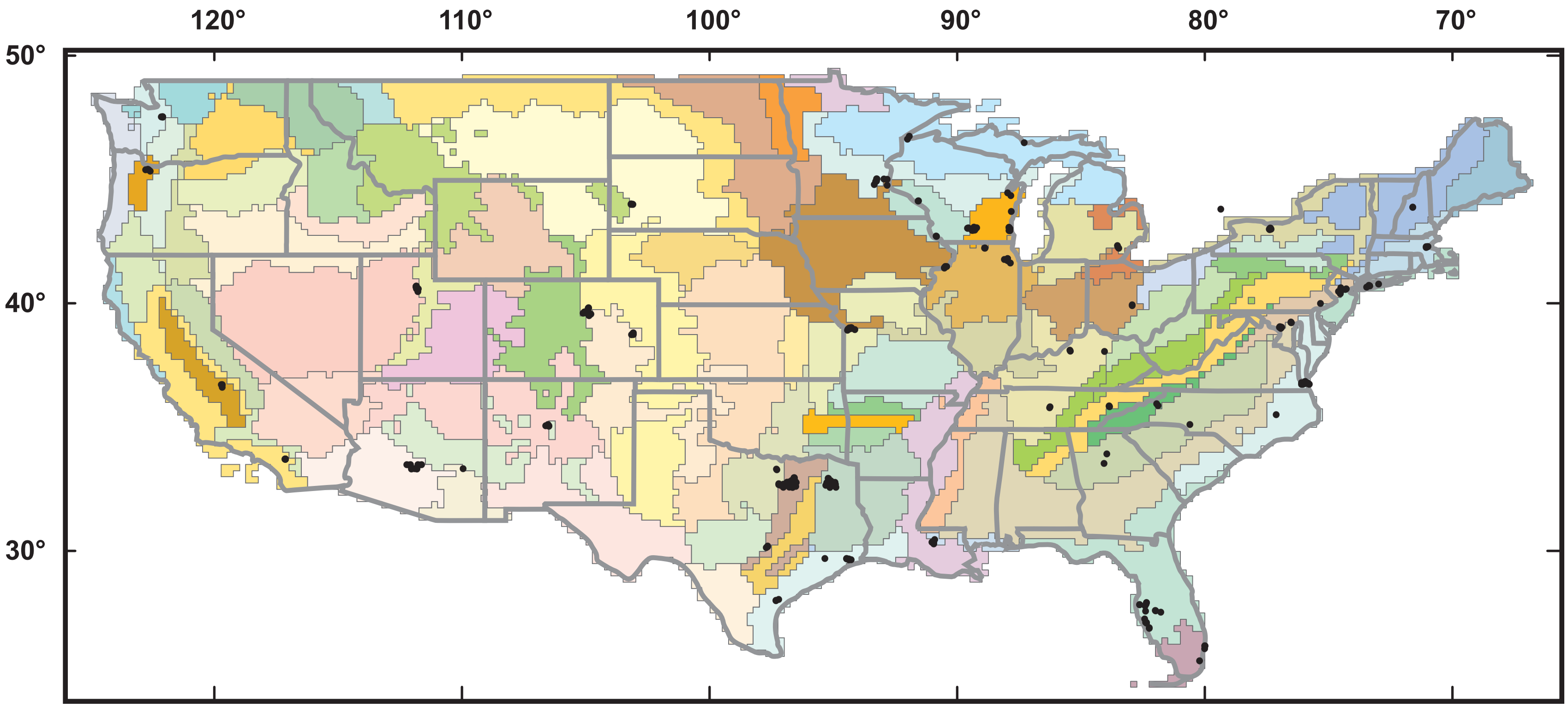
Hydrologic-basin characteristics and information about land use and land cover are geographic data that have commonly been derived on the basis of topographic maps and field surveys in the area of interest. Increasingly, the data necessary for basin characterization are being made available as online GIS datasets. For example, the USGS developed the Streamstats program (Ries and others, 2004; U.S. Geological Survey, 2007) as an Internet tool that allows the user to automatically delineate the contributing area for any point along a defined stream reach, query basin characteristics for the area, calculate streamflow statistics, and download GIS data for the selected basin. Many such datasets are being made available on the Internet for use with GIS software (appendix 6).

A database of site characteristics and accompanying rainfall-runoff data was assembled for this project to facilitate development of rainfall-runoff estimates at ungaged sites. The rainfall-runoff statistics described in this report were derived from the data in the Microsoft Access database (SiteStormV01.mdb) available on the CD–ROM accompanying this report. The data in the SiteStormV01.mdb database was compiled from 35 different hydrologic studies (Hornbeck, 1973; Yorke and Herb, 1978; Allen and Gray, 1984; Campbell, 1987; Mustard and others, 1987; Sloto, 1988; Metzker and others, 1993; Knutilla and Veenhuis, 1994; Schaap and Lucey, 1994; Schalk, 1994; Fossum and Davis, 1996; Guay, 1996; Outlaw, 1996; Trommer and others, 1996a,b; Waschbusch, 1996; Baldys and others, 1997; Steuer and others, 1997; Stumm and Ku, 1997; Demcheck and others, 1998; Duncker and Melching, 1998; Waschbusch, 1999; Waschbusch and others, 1999; Holmstrom and others, 2000; Martin and others, 2001; U.S. Environmental Protection Agency, 2001b; Walker and others, 2001; Breault and others, 2002; Ockerman, 2002; Zarriello and Barlow, 2002; Peters and others, 2003; Horwatich and others, 2004; Selbig and others, 2004; Maestre and Pitt, 2005; J.A. Horwatich, written commun., 2007; Selbig and Bannerman, 2007; 2008).

The SiteStormV01.mdb database includes data from 6,142 storm events monitored at 306 monitoring sites in the conterminous United States and southern Canada, 3 sites in Alaska, and 2 sites in Hawaii. Many of the sites are clustered within local or regional study areas (fig. 23). (A GIS coverage of these rainfall-runoff data collection sites is provided in the GIS directory on the CD–ROM accompanying this report (table 2)). The dataset includes a variety of climatic conditions among the data-site clusters and a variety of site characteristics within each cluster. The storms included in the database have precipitation volumes ranging from 0.01 to about 25 in. and runoff volumes ranging from 0 to about 21 in. The monitoring sites in the database are associated with drainage areas ranging from less than an acre to about 250 mi^2. The drainage basins to these monitoring sites have TIA fractions ranging from zero to one and represent a variety of land uses. Drainage areas among sites with low TIA fractions tend to be larger and more variable than drainage areas among sites with high TIA fractions (fig. 24). Many basins throughout the range of TIA fractions are relatively small (less than 1 mi^2), but relatively few basins with drainage areas larger than 1 mi^2 have TIA fractions greater than 40 percent. In general, this relation between TIA fraction and drainage area would be expected because smaller basins are more likely to have more homogenous land uses than larger basins. Information about site characteristics and runoff-coefficient summary statistics from the USEPA NURP study (Athayde and others, 1983) also are compiled in a table in the database to facilitate analysis of rainfall-runoff relations.

Software for Storage and Analysis of Rainfall-Runoff Data

The Microsoft Access database (SiteStormV01.mdb) available on the CD–ROM accompanying this report has five tables (fig. 25) to store rainfall-runoff datasets and facilitate use of this data and example queries that were used to analyze the data. In general, tables, fields, and queries were named with whole words or acronyms to facilitate use of the database. The description property for each table, field, and query provides a short statement for defining each of these database objects. The site table (tblSiteTable), the primary data table in the database, includes 2 key fields and 41 data fields. These data fields include the name, location, impervious fraction, land-use percentages, and hydrologic soil group for each site and associated drainage area. (These data were included in the database only if the data had been derived from source documents.) At a minimum, the site data must include a name and the impervious fraction, which is necessary for extending the data to predict runoff at ungaged sites. The storm-data table (tblStormData) includes 12 data fields that characterize each storm event. Each site may be associated with one or more storms. At a minimum, precipitation and runoff data must be associated with the storm and the site in the database. Additional storm-event characterization data are included in the database if the data are available in source documents. An association table (tasStudySource) is used to relate each site to studies and sources of data. This association table is used because each site may be associated with multiple source documents and, in theory, one source



Base from U.S. Geological Survey National Map
Ecoregions from U.S. Environmental Protection Agency
Geographic projection

Figure 23. Spatial distribution of 306 stormflow-monitoring stations (black dots) with respect to U.S. Environmental Protection Agency (2003) Level III ecoregions (colored polygons) that have been discretized to a 15-minute grid in the conterminous United States. Many stations are clustered within local study areas. Ecoregions are identified on the plate useco.pdf on the CD–ROM accompanying this report.

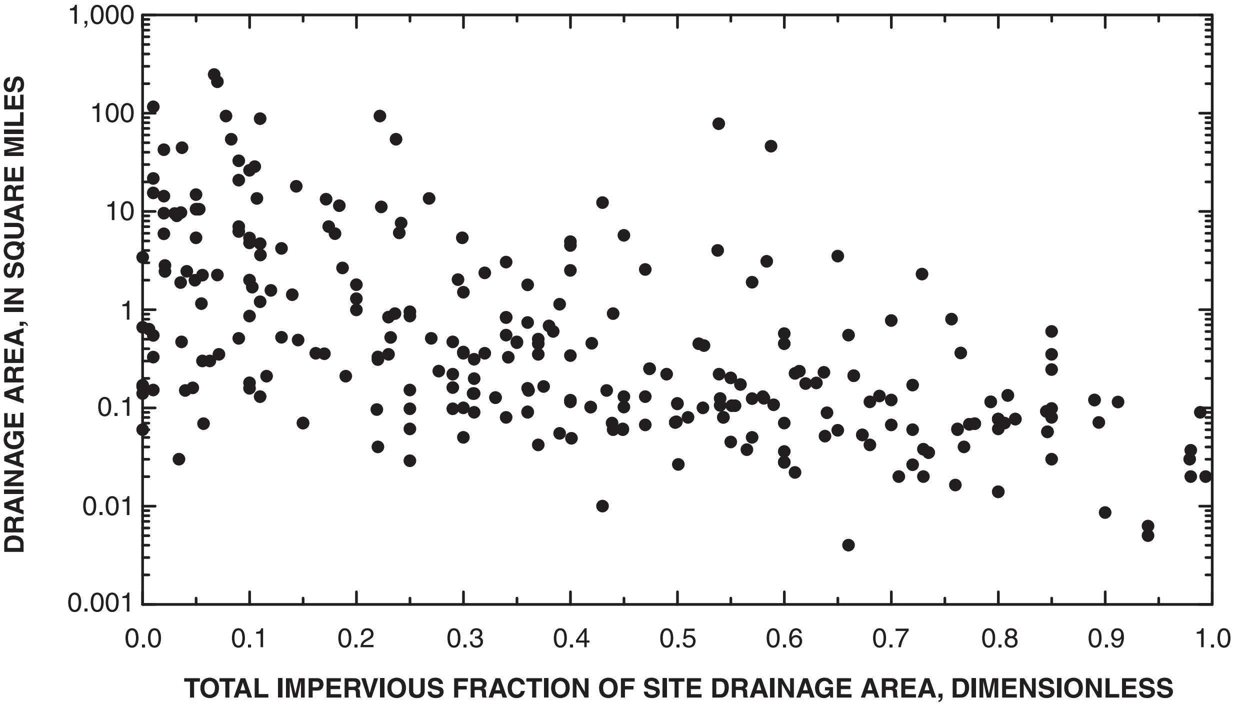


Figure 24. The drainage areas of the 306 runoff-monitoring sites in the SiteStormV01.mdb database with respect to the impervious fraction of the site drainage area.

document may summarize results from several studies. The data-source table (tblDataSource) includes four fields giving the bibliographic citation, an Internet reference to the source document, and a comment field. The USEPA NURP table (tblUSEPANURPSites) includes site information and runoff statistics documented by Athayde and others (1983).

The SiteStormV01.mdb database also includes example queries that were used to analyze the data. The query qryRvStorms can be used to calculate statistics for runoff coefficients that are less than or equal to one for all sites. The query qryRvImpervious can be used to calculate statistics for runoff coefficients that are less than or equal to one at sites at which data were collected for nine or more storms. The queries qryRvCountMoreThan6Baseflows and qryRvBaseflowCorrelation can be used to identify sites with more than six base-flow values and to provide these data for correlation analysis of runoff coefficients and base flow, respectively.

Statistical Characterization of Rainfall-Runoff Data

Rainfall-runoff data in the SiteStormV01.mdb are used to help characterize runoff-coefficient statistics. These statistics are needed to develop planning-level estimates of stormflows at ungaged sites on the basis of easily identifiable study-site characteristics. Statistics for estimated runoff coefficients also provide the information needed for stochastic runoff-modeling efforts.

Although there are 306 monitoring sites in the database, it may be difficult to select a single site from the database that is similar to an ungaged site of interest because many characteristics—including land use, imperviousness, drainage features, soils, depth to groundwater, and climate—may affect storm-event runoff production. Furthermore, the small sizes of datasets that were collected at some sites for only a few storm events may limit the transferability of site-specific

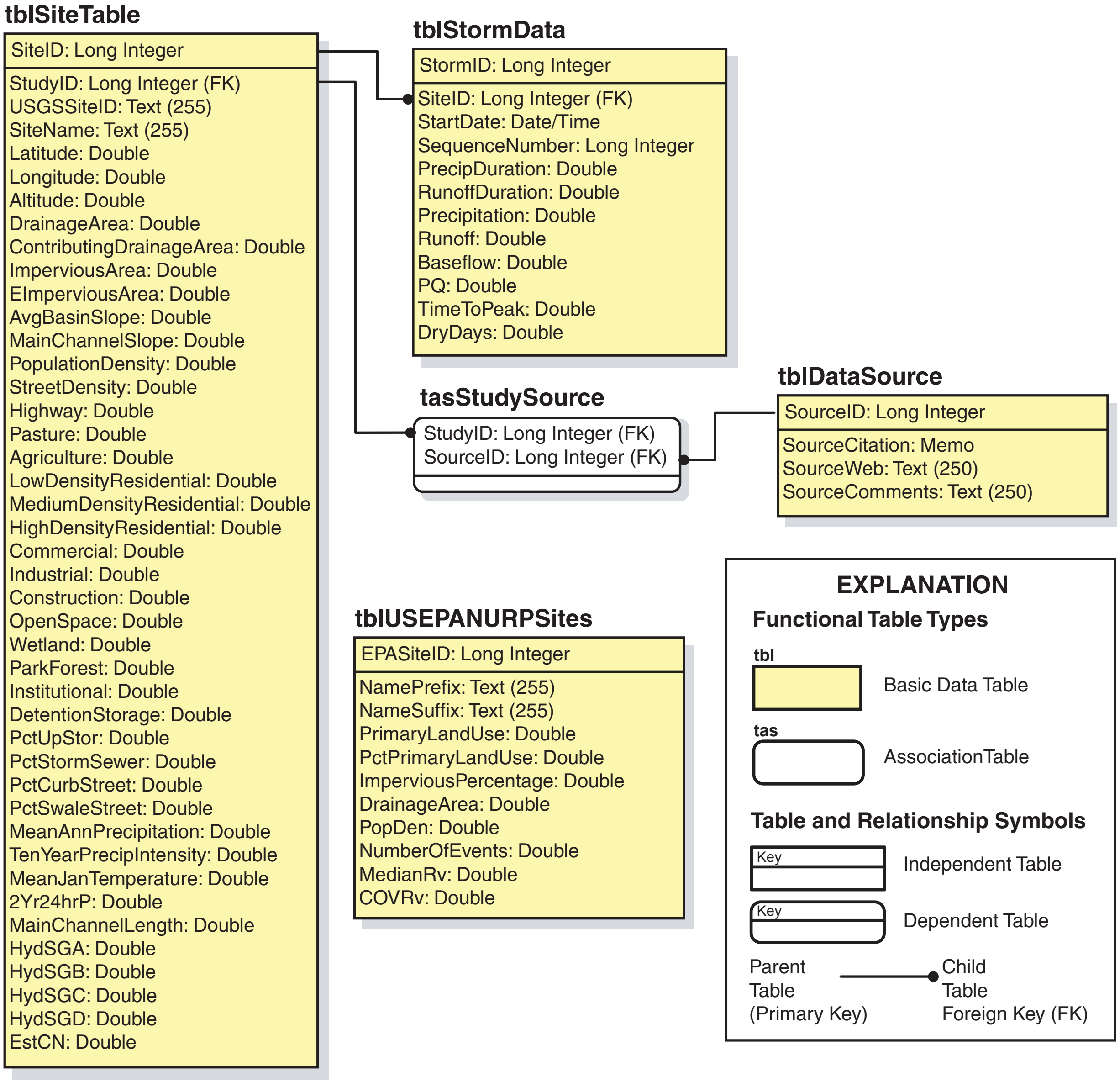


Figure 25. An entity-relationship diagram showing a graphical representation of tables, fields, and relationships of the data structure for the precipitation-runoff data in the SiteStormV01.mdb database.

runoff-coefficient statistics. In these cases, the use of data from multiple sites in the same impervious-fraction group may adequately represent the population of runoff coefficients at a site of interest. Runoff-coefficient values less than or equal to 1.0 for every storm in the SiteStormV01.mdb database are included in the boxplot in figure 26 to facilitate planning-level runoff-coefficient estimates for ungaged sites. The runoff-coefficient values in figure 26 were grouped among all sites within impervious-fraction increments of 0.1 (a range of 10 percent) to indicate the variability in runoff coefficients within each interval. The numbers of sites and storm events in each group are listed above each interval on the graph (fig. 26). About 52 percent of the sites (and 69 percent of the selected storms) are associated with TIA fractions less than or equal to 0.4; about 96 percent of the land area of the conterminous United States is characterized by this range of imperviousness (appendix 6).

Runoff-coefficient statistics were calculated for the 151 sites with data for nine or more storm events, and the results were grouped by TIA fraction (fig. 27). As on figure 26, the impervious-fraction groups are divided into intervals of about 0.1 (a range of 10 percent). Comparison with the number of sites and the number of storm events in figure 27 with that in figure 26 indicates that about half the sites in the database are associated with data from nine or more storm events. The boxplots of the average, COV, and coefficient of skew of the runoff coefficients for sites in each impervious-fraction group indicate substantial variation in these statistics even for sites with similar TIA fractions (fig. 27). For example, among sites with TIA fractions greater than 0.1 and less than or equal to 0.2, the site-average runoff coefficients range from about 0.06 to 0.64, the COVs range from 0.27 to 1.25, and the coefficients of skew of the runoff coefficients range from -0.06 to 2.5. It should be noted that the variability in individual runoff coefficients from all sites (fig. 26) is not substantially different than the variability among site-average runoff coefficients (fig. 27) within each impervious-fraction group. The boxplots indicate a substantial overlap in the average runoff coefficients among sites associated with TIA fractions greater than 0.2 and less than or equal to 0.7. Most sites with TIA fractions that

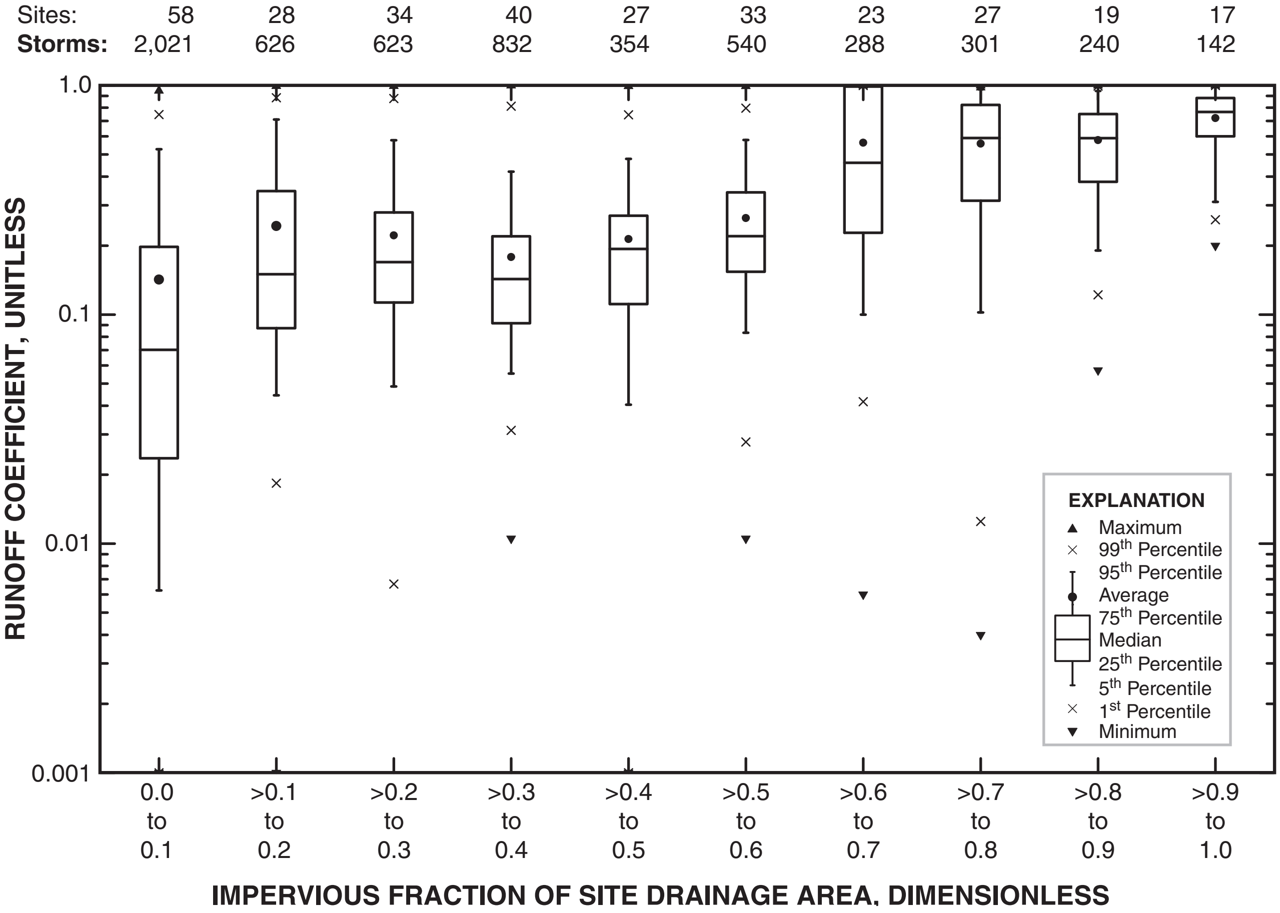


Figure 26. Runoff coefficients that are less than or equal to 1.0 for individual storm events from the 306 different sites that are grouped by impervious-fraction intervals (>, greater than).

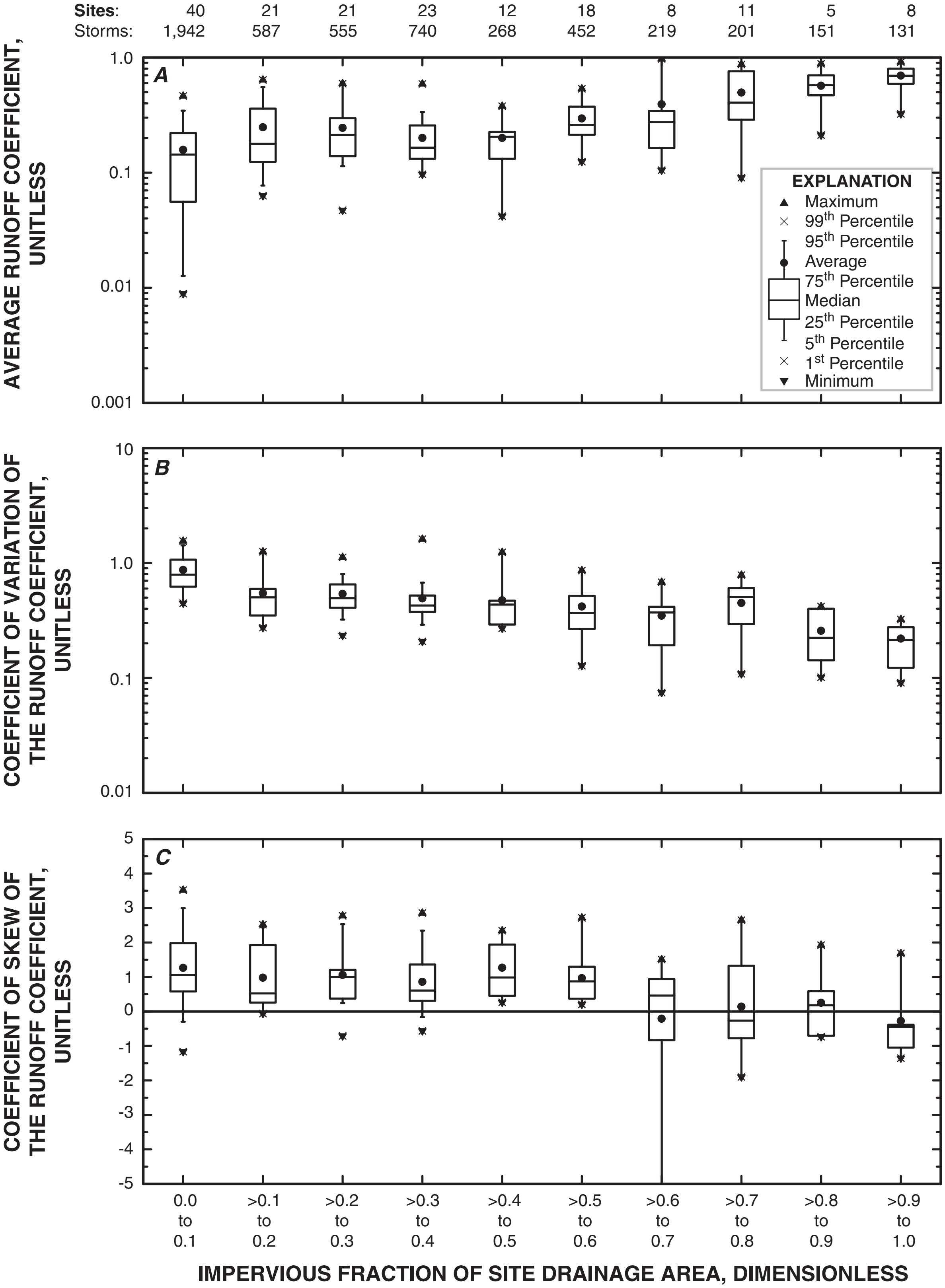


Figure 27. *(A)* The average, *(B)* coefficient of variation, and *(C)* coefficient of skew by site for runoff coefficients that are less than or equal to 1. The sites at which measurements were made for nine or more storm events were included and are grouped by impervious-fraction intervals (>, greater than).

are less than 0.7 have skew coefficients that are greater than 0.5; this indicates that relatively few storms have runoff coefficients that approach 1.0 for sites in this impervious-fraction group. Conversely, about 75 percent of sites with TIA fractions greater than 0.9 have negative skew coefficients (fig. 27); this indicates that many storms have runoff coefficients that are greater than or equal to 0.7, and few storms have low runoff coefficients for sites in this impervious-fraction group.

The nonparametric rank-sum test (Helsel and Hirsch, 2002) was applied to examine the population of site runoff-coefficient statistics to identify statistically significant differences in the averages and standard deviations among the ten impervious-fraction groups (table 11). The nonparametric rank-sum test was selected because it is a distribution-free test that can be used to detect differences between two groups even if the numbers of observations for the two groups are different (Helsel and Hirsch, 2002). The averages and standard deviations were calculated for each of the 151 sites with precipitation and stormflow data for nine or more storm events. The values in table 11 represent the probabilities that the runoff-coefficient statistics for the different impervious-fraction groups are drawn from the same population; the grey-shaded values indicate differences that are statistically significant in a two-sided test with a 95-percent confidence interval (a 5-percent *p*-value). In general, the rank-sum test statistics indicate no statistically significant differences between the site-average runoff coefficients between pairs of groups separated by one impervious-fraction interval but statistically significant differences between pairs of groups separated by more than one interval. The boxplots of site-average runoff coefficients in figure 27 support these results because the interquartile ranges of adjacent impervious-fraction groups overlap. The rank-sum test statistics for the standard deviations indicate that, for the most part, the

Table 11. Results of the rank-sum test (Helsel and Hirsch, 2002) indicating the probability that the runoff coefficient statistics for different impervious-fraction groups are drawn from the same population.

[IFR, impervious fraction range; >, greater than; <, less than; --, same impervious-fraction group; gray boxes indicate values that are considered statistically significant with a probability of less than 5 percent (0.05) that samples from different TIA groups are from the same population]

Probability of obtaining the test statistic when the populations of the averages of volumetric runoff-coefficients at sites in different impervious-fraction groups are the same

IFR	0.0 to 0.1	>0.1 to 0.2	>0.2 to 0.3	>0.3 to 0.4	>0.4 to 0.5	>0.5 to 0.6	>0.6 to 0.7	>0.7 to 0.8	>0.8 to 0.9
>0.1 to 0.2	0.055	--							
>0.2 to 0.3	0.018	0.546	--						
>0.3 to 0.4	0.099	0.963	0.347	--					
>0.4 to 0.5	0.155	0.985	0.694	0.543	--				
>0.5 to 0.6	<0.001	0.099	0.099	0.008	0.054	--			
>0.6 to 0.7	0.012	0.180	0.213	0.045	0.178	0.846	--		
>0.7 to 0.8	<0.001	0.010	0.009	0.001	0.002	0.041	0.265	--	
>0.8 to 0.9	0.001	0.011	0.019	0.005	0.010	0.040	0.272	0.571	--
>0.9	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.066	0.149	0.341

Probability of obtaining the test statistic when the populations of the standard deviations of volumetric runoff-coefficients at sites in different impervious-fraction groups are the same

IFR	0.0 to 0.1	>0.1 to 0.2	>0.2 to 0.3	>0.3 to 0.4	>0.4 to 0.5	>0.5 to 0.6	>0.6 to 0.7	>0.7 to 0.8	>0.8 to 0.9
>0.1 to 0.2	0.850	--							
>0.2 to 0.3	0.909	0.687	--						
>0.3 to 0.4	0.160	0.424	0.075	--					
>0.4 to 0.5	0.272	0.421	0.120	0.848	--				
>0.5 to 0.6	0.632	0.877	0.725	0.118	0.262	--			
>0.6 to 0.7	0.326	0.608	0.272	0.668	0.563	0.523	--		
>0.7 to 0.8	0.056	0.104	0.322	0.002	0.007	0.014	0.009	--	
>0.8 to 0.9	0.814	0.795	0.948	0.119	0.155	0.852	0.164	0.070	--
>0.9	0.516	0.393	0.196	0.014	0.041	0.192	0.018	0.201	0.421

standard deviations are not significantly different for most of the groups at a 95-percent confidence interval (table 11). This result indicates that the pattern of decreasing site-COV statistics with increasing imperviousness in figure 27 is the result of an increase in the values of the site-average runoff coefficients rather than a decrease in the variability of the runoff coefficients.

Variations in prestorm streamflow may be a factor contributing to the large variation in runoff coefficients (fig. 26) and runoff-coefficient statistics (fig. 27) at different sites. Rank correlation coefficients (Spearman's rho as described by Helsel and Hirsch, 2002) were calculated to evaluate potential relations between prestorm streamflow and runoff coefficients. Correlation does not necessarily imply causation; for example, antecedent precipitation may saturate soils and increase prestorm streamflow. In this case, higher prestorm streamflow may indicate wetter antecedent conditions, but not necessarily cause more runoff. In the storm-event database compiled for the SELDM study, 43 sites have at least 7 paired prestorm-streamflow and runoff-coefficient values. Figure 28 shows the Spearman's rho values for these datasets and the associated 95-percent confidence intervals, which are a function of sample size (Caruso and Cliff, 1997). Three sites had very weak negative correlations. Ten sites had positive rho values that were less than about 0.3, indicating that variations in prestorm streamflow may be associated with less than 30 percent of the variations in runoff coefficients from storm to storm at each of these sites. Seven sites have rho values between 0.3 and 0.5, indicating that prestorm streamflow may be associated with 30 to 50 percent of the variations in runoff coefficients at each of these sites. An additional 14 sites have rho values between 0.5 and 0.71, indicating that prestorm streamflow may be associated with 50 to 71 percent of variations in runoff coefficients at each of these sites. Nine sites have rho values that are greater than 0.71, indicating a moderate to strong correlation between prestorm streamflow and runoff coefficients at these sites. Differences in correlation coefficients among sites may reflect hydrologic-basin characteristics, artifacts in the assembled dataset, such as the use of different hydrograph-separation techniques in each study (appendix 1), or uncertainty in the samples. Seventeen sites had 95-percent confidence intervals that included zero; this result indicates that the true rho value may not be different from zero. Conversely, however, 35 of the sites have upper confidence limits that are greater than 0.5, which may indicate a substantial correlation between prestorm streamflow and the runoff coefficient.

Rank correlation coefficients also were calculated to evaluate potential relations between precipitation and runoff coefficients. In the storm-event database compiled for the SELDM study, 167 sites are listed with at least 9 paired precipitation and runoff-coefficient values. Figure 29 shows the Spearman's rho values for these datasets and the associated 95-percent confidence intervals, which are a function of sample size (Caruso and Cliff, 1997). Rho values range from -0.86 to 0.96 with about 20 percent of the sites having rank correlations below zero. Seventy-five sites (about 45 percent of the selected sites) have rho values between +0.3 and -0.3, indicating that variations in precipitation may be associated with less than 30 percent of variations in runoff coefficients from storm to storm at each of these sites. Six sites (about 4 percent of the selected sites) have rho values between -0.3 and -0.5, and 45 sites (about 27 percent of the selected sites) have rho values between 0.3 and 0.5, which indicate that precipitation may be associated with 30 to 50 percent of variations in runoff coefficients at each of these sites. Four sites (about 2 percent of the selected sites) have rho values between -0.5 and -0.71, and 21 sites (about 13 percent of the selected sites) have rho values between 0.5 and 0.71; these ranges indicate that precipitation may be associated with 50 to 71 percent of variations in runoff coefficients, respectively, at each of these sites. Only 14 sites (about 8 percent of the selected sites) have rho values greater than +0.71 or less than -0.71, indicating a moderate to strong correlation between precipitation and runoff coefficients for these sites.

Although the runoff coefficient is the ratio of runoff to rainfall, total storm-event rainfall may not be a good predictive variable for the runoff coefficient. For example, 105 sites (about 64 percent of the selected sites) have 95-percent confidence interval limits that included zero, which indicates that the true rho value may not be different from zero (fig. 29). This finding is consistent with the results of analysis of data from highway sites (Driscoll and others, 1990b) and urban-runoff-monitoring sites (Athayde and others, 1983). Mathematically, runoff coefficients would be expected to have a strong negative rank correlation with total storm-event rainfall because the runoff coefficient is calculated as the ratio of runoff to rainfall. Hydrologically, however, runoff coefficients would be expected to have a strong positive rank correlation with total storm-event rainfall because continuing increases in rainfall should eventually exceed initial abstractions and soil-infiltration rates, which would be expected to generate more runoff for each additional rainfall input. Differences in correlation coefficients between rainfall and runoff coefficients among sites may reflect hydrologic-basin characteristics, but the TIA fraction is not a substantial explanatory variable for the correlation coefficients. (In this study, the TIA fraction accounted for about 3 percent of the variation in rho values among the sites.)

Estimating Rainfall-Runoff Transformations at Ungaged Sites

Regression equations relating the average, standard deviation, and skew of runoff coefficients to the TIA fraction were developed to facilitate the selection of representative statistics for ungaged sites (fig. 30). The Kendall-Theil robust line, a nonparametric regression procedure, was used for developing these equations because this method is robust to the effects of outliers (Granato, 2006). The equations are based on data for drainage basins with TIA fractions ranging from

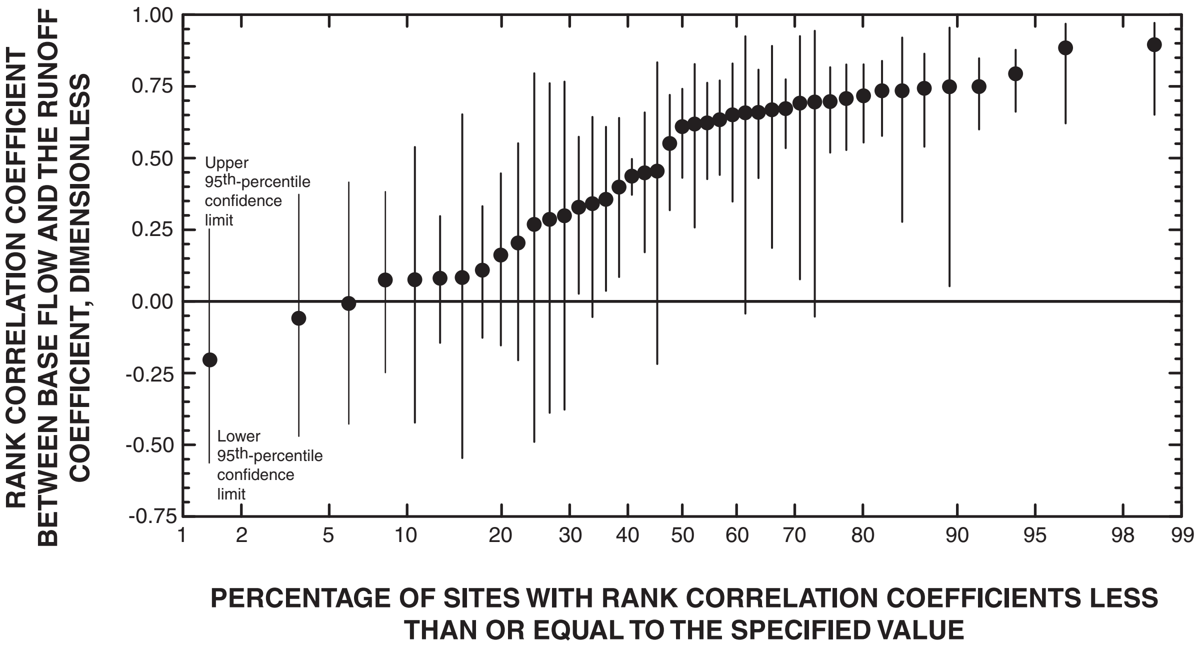


Figure 28. Probability plot showing the percentage of the 43 sites with 7 or more paired base-flow and runoff measurements that have nonparametric rank correlation coefficients (Spearman's rho) less than or equal to the specified value.

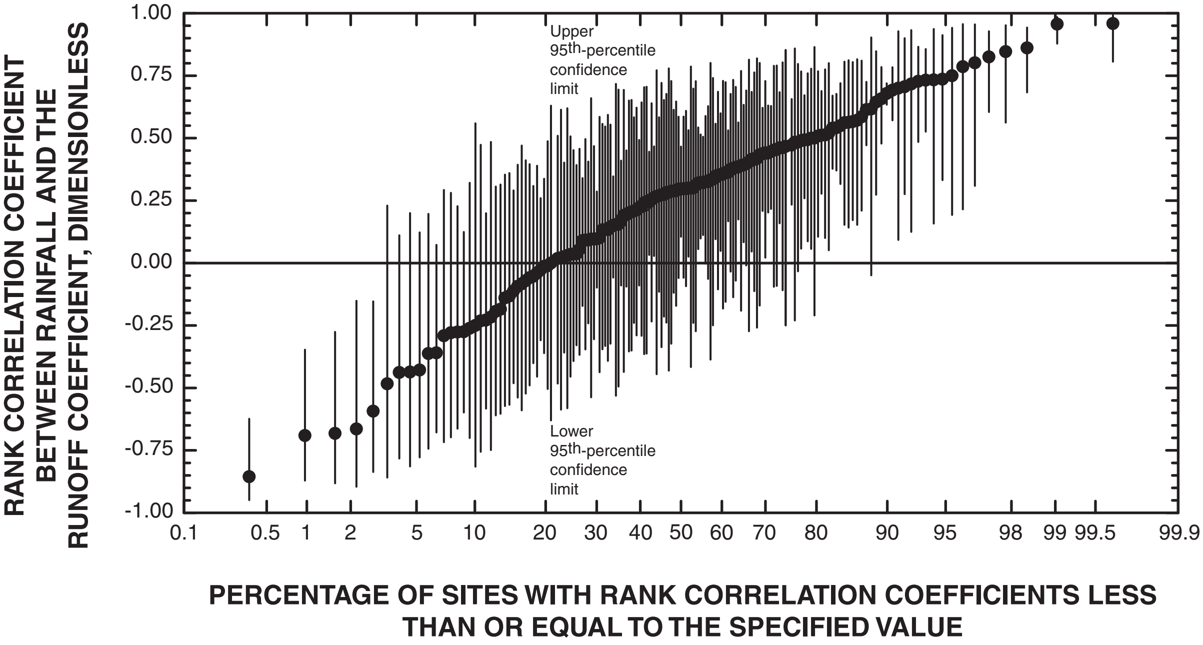


Figure 29. Probability plot showing the percentage of 165 sites with 7 or more paired storm-event rainfall and runoff measurements that have nonparametric rank correlation coefficients (Spearman's rho) less than or equal to the specified value.

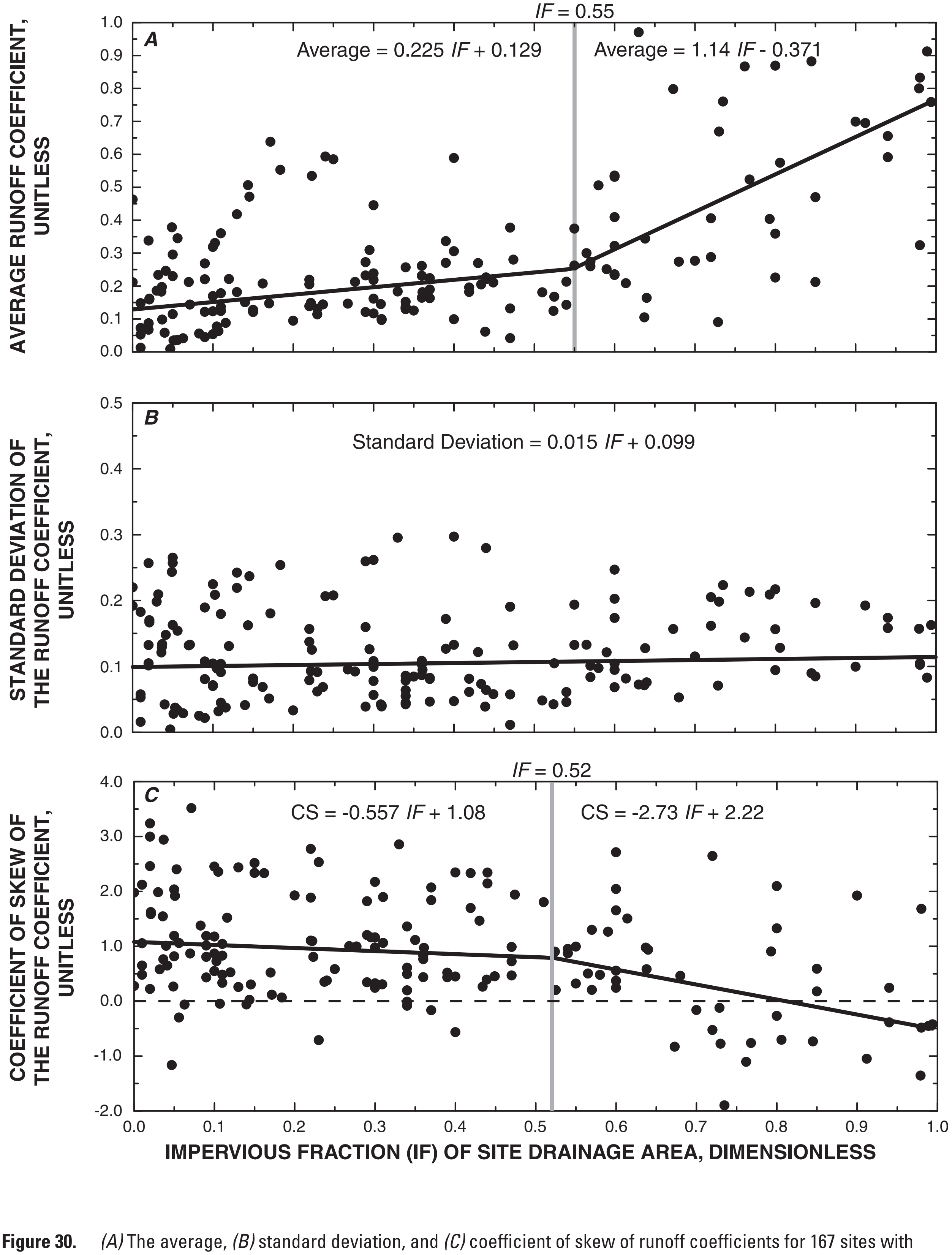


Figure 30. *(A)* The average, *(B)* standard deviation, and *(C)* coefficient of skew of runoff coefficients for 167 sites with 9 or more storm events. Nonparametric regression lines (Granato, 2006) indicate the relation between each statistic and the impervious fraction. Two-line regression models were developed for the average and coefficient of skew to better characterize the relation between these statistics and the impervious fraction.

0.01 to 99.4 percent of the drainage area. Appendix 6 provides an overview of methods and sources of data that can be used to estimate the TIA fractions needed for these regression equations. The regression equations should be viewed as approximate guidelines for picking representative values for a given TIA fraction at an ungaged site because the scatter in each statistic for different sites over the full range of TIA fractions is substantial. For sites with TIA fractions below 0.5, the scatter of the points and the low slopes of the regression lines for each statistic evident in figure 30 indicate that a precise TIA fraction estimate may not be critical for estimating runoff-coefficient statistics in this range.

A two-segment regression model was developed to estimate site-average runoff coefficients at ungaged sites from the estimated TIA fraction (fig. 30A). The two-segment regression model accounts for the steeper trend in site-average runoff coefficients above a TIA fraction of about 0.55. The slopes of both segments are significantly different from zero within a 95-percent confidence interval. This result indicates a relation between TIA fraction and site-average runoff coefficients; however, because of the scatter in values, the regression equation accounts for only about 23 percent of the variation in average runoff coefficients among sites. These equations yield a site-average runoff coefficient of about 0.13 with a TIA fraction of 0.0, about 0.24 with a TIA fraction of 0.5, and about 0.77 with a TIA fraction of 1.0. A site-average runoff coefficient of 0.77 may seem a bit low for a completely impervious area, but Wiles and Sharp (2008) indicate that 6–36 percent of storm-event precipitation may be lost to evaporation and infiltration through cracks and joints in paved surfaces.

A one-segment regression model was developed to estimate the standard deviation of runoff coefficients at ungaged sites from the estimated TIA fraction (fig. 30B). The slope of the standard deviation regression line is not significantly different from zero within a 95-percent confidence interval. For the standard deviation, use of the median value, which is about 0.1 regardless of the TIA fraction, has about the same predictive power as the associated regression equation. This nonparametric regression analysis confirms that the pattern of decreasing COV values with increasing average runoff coefficients (figs. 30A, B) is an artifact of the increased values of the average runoff coefficients rather than a decrease in the variability of runoff volumes with increasing impervious fractions.

A two-segment regression model was developed to estimate the skew of runoff coefficients at ungaged sites from the estimated TIA fraction (fig. 30C). The model accounts for the steeper trend in the coefficient of skew of runoff coefficients above a TIA fraction of about 0.52. The slope of the first segment is not significantly different from zero, but the slope of the second segment is significantly different from zero within a 95-percent confidence interval. This indicates that the skew coefficients vary randomly below a TIA fraction of 0.52 and generally decrease with increasing TIA fraction above this threshold. This two-segment regression model accounts for about 7 percent of the variation in the skew of runoff coefficients among all sites and about 11 percent of the variation among sites with TIA fractions greater than 0.52. These equations yield a coefficient of skew of runoff coefficients of about 1.0 with a TIA fraction of 0.0, about 0.8 with a TIA fraction of 0.5, and about -0.5 with a TIA fraction of 1.0.

A truncated Pearson Type III distribution was selected for generating stochastic planning-level estimates of runoff coefficients at ungaged sites because it is an extremely flexible distribution that can assume different shapes such as symmetrical, positively skewed, or negatively skewed (Haan, 1977; Chow and others, 1988; Bobee and Ashkar, 1991). Data from a Pearson Type III distribution also can be estimated with frequency factors (Haan, 1977; Interagency Advisory Committee on Water Data, 1982; Chow and others, 1988; Stedinger and others, 1993; Cheng and others, 2007). The mean and standard deviation of runoff-coefficient data are used to calculate the location and spread of the resultant runoff coefficients (eq. 1). The skew coefficient is used to adjust the standard normal variates to produce a representative sample of data. If the skew of a population equals zero, the frequency factor is the standard normal variate. As skew coefficients deviate from zero, the relation between exceedence probability and the associated frequency factor shifts to reflect the distribution of values above and below the median value. Almost 83 percent of the sites in figure 30 have skew coefficients within the range of ±2.0. Almost 99 percent of these sites have skew coefficients between -2.0 and 4.0. One outlier has an extreme skew coefficient equal to -5.15. The modified Wilson-Hilferty algorithm developed by Kirby (1972) provides acceptable estimates of Pearson Type III frequency factors for samples with coefficients of skew within the range of about ±9. Although the Pearson Type III distribution is not bounded by zero and one, standard acceptance–rejection methods for stochastic data generation can be used to limit results to values within this interval (Devroye, 1986; Saucier, 2000; Gentle, 2003).

Example datasets for impervious fractions of 0.0, 0.05, 0.20, 0.50, 0.80, 0.95, and 1.0 are presented in figure 31. These impervious fractions were selected to demonstrate the flexibility of the Pearson Type III algorithm for stochastic generation. The datasets were generated based on the averages, standard deviations, and coefficients of skew estimated using the regression equations in figure 30. The samples of generated values with IFs less than or equal to 0.50 are positively skewed, symmetrical with an impervious fraction of 0.80, and negatively skewed with impervious fractions of 0.95 and 1.0. Although runoff coefficient values are bounded by zero and one for individual precipitation events in SELDM (heavy black lines), the y-axis scale in figure 31 ranges from -0.2 to 1.2 to indicate the proportion of values that would be rejected from a stochastically generated dataset of 1,000 points. These datasets indicate the characteristics of runoff-coefficient populations that are based on the impervious fractions of upstream basins and may be used for rainfall-runoff analyses at ungaged sites.

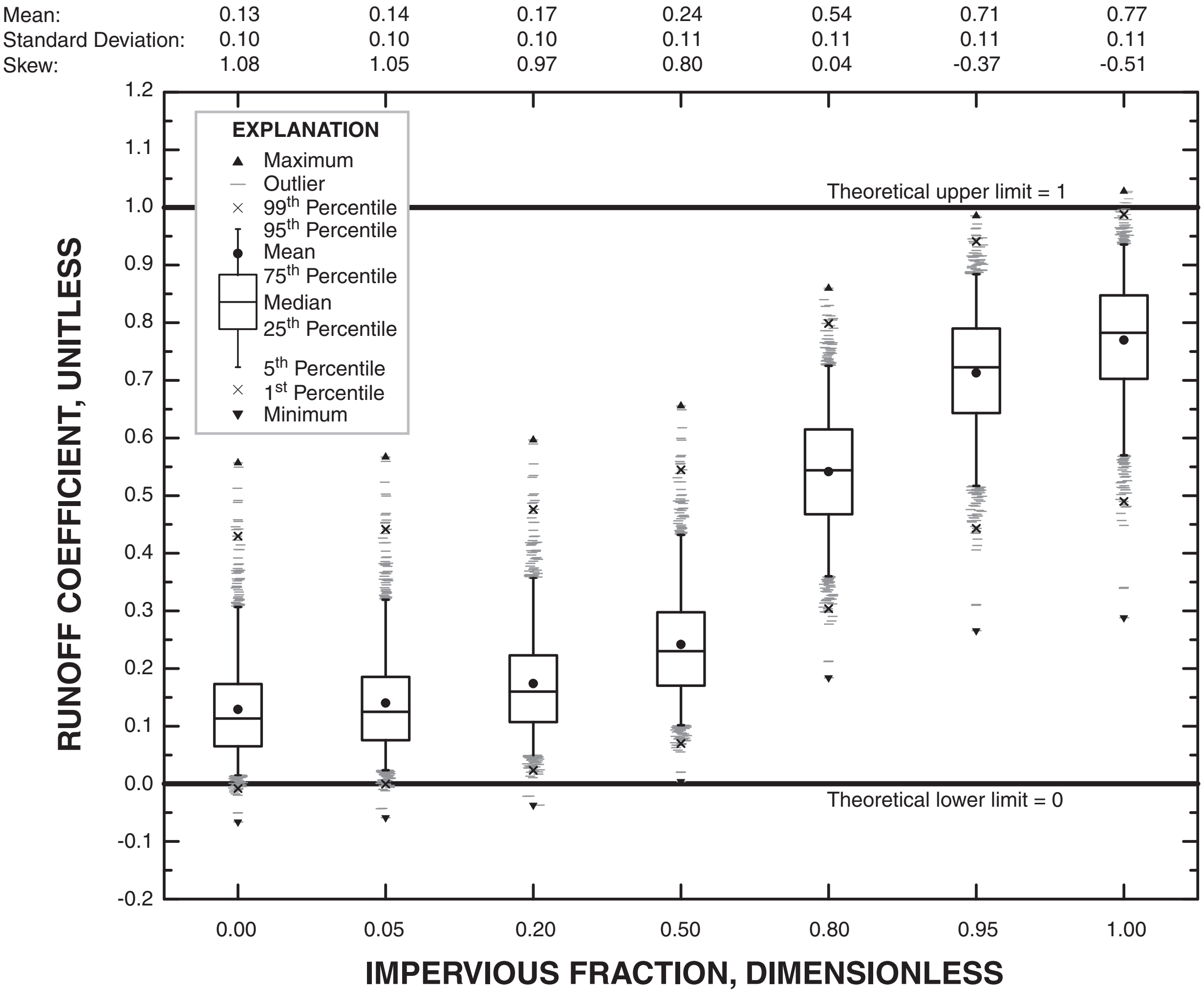
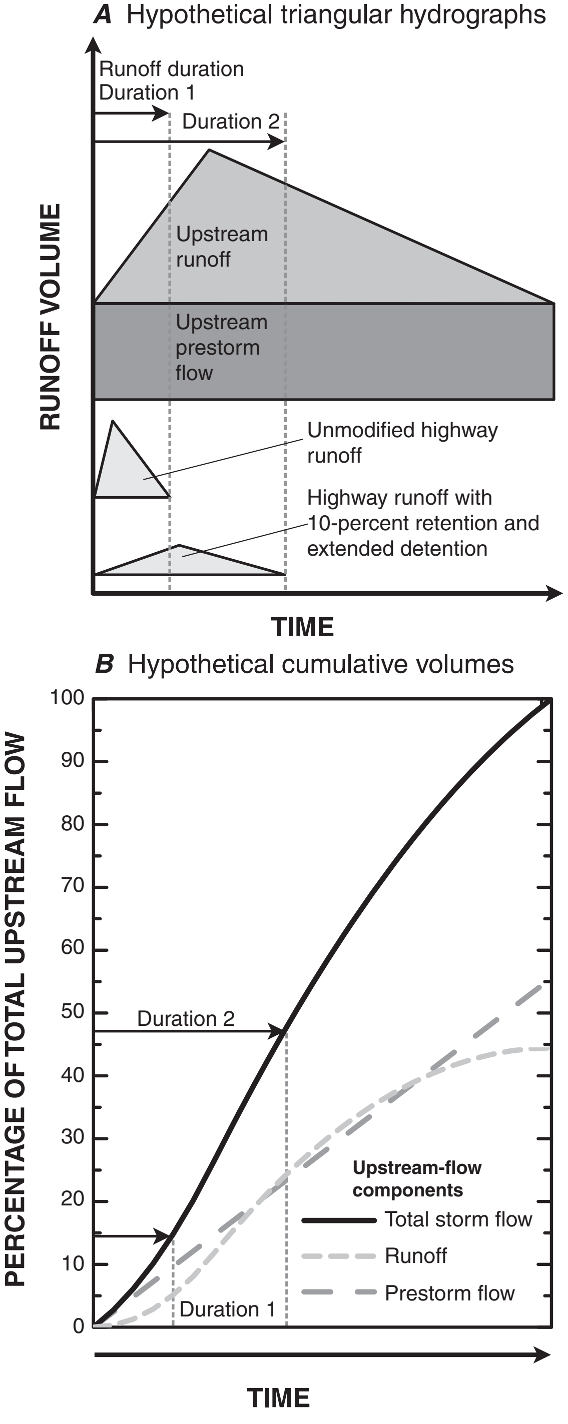


Figure 31. Runoff-coefficient samples that are stochastically generated with sample means, standard deviations, and coefficients of skew estimated using the impervious-fraction regression equations shown in figure 30.

Storm-Event Hydrograph

Information about the storm-event hydrograph for flow from the highway catchment and the upstream basin is necessary to estimate the quantity of the upstream flow that occurs concurrently with the highway runoff at the highway-runoff outfall (fig. 32). The focus of planning-level analyses of highway-runoff-quality has traditionally been on event-mean concentrations and total storm loads for the entire highway-runoff event rather than on processes during events. Differences in the locations, sizes, and drainage characteristics of the highway catchment and the upstream basin, however, may cause differences in the timings and durations of runoff from each area. If the highway catchment is small and the runoff drains directly to the stream, the duration of appreciable runoff from the highway catchment may be approximated by the duration of the precipitation event. If the upstream basin is relatively large and more pervious than the highway catchment, appreciable runoff from the basin may continue for hours or days longer than runoff from the highway catchment. In this case, only a small proportion of the upstream runoff may be available to dilute highway-runoff constituents in the receiving waters. If, however, a structural BMP is employed at the highway site to attenuate and extend the highway-runoff hydrograph, then much more of the upstream runoff may be available to dilute highway-runoff constituents in the receiving waters.

This concept is demonstrated schematically in figure 32. In this hypothetical example, the triangular runoff hydrograph for the upstream basin is superimposed on a rectangular representation of the prestorm base flow (fig. 32A). The durations of highway-runoff hydrographs with and without BMP modification are labeled “Duration 1” and “Duration 2,” respectively. As indicated in the figure, a small increase in the



A **Hypothetical triangular hydrographs**

B **Hypothetical cumulative volumes**

Figure 32. Simplified schematic diagram showing *(A)* hypothetical triangular hydrographs and *(B)* the hypothetical cumulative upstream storm volume that would occur concurrently with unmodified runoff from a highway and with runoff from an extended detention structure. This diagram shows the hypothetical runoff event with two upstream flow components (runoff and prestorm flow), an unmodified highway-runoff hydrograph, and a highway-runoff hydrograph with retention and detention.

duration of runoff from the highway may be accompanied by a large increase in the cumulative amount of concurrent runoff and base flow from the upstream basin, especially in the rising limb of the upstream storm-event hydrograph (fig. 32B).

Detailed characterization of within-storm processes are beyond the scope of a planning-level water-quality analysis, but a systematic method is necessary to estimate the duration of the highway-runoff hydrograph and the proportion of upstream flows that may be concurrent with a highway-runoff event. Unit-hydrograph methods are commonly used, but these methods can be computationally intensive and require detailed information about basin characteristics and the temporal distribution of precipitation within a storm (Linsley and others, 1975; Chow and others, 1988; Pilgrim and Cordery, 1993). Different right-skewed probability distributions such as the beta, Chi-square, gamma, lognormal, log-Pearson Type III, triangular, and Weibull have been used to represent runoff hydrographs (Koutsoyiannis and Xanthopoulos, 1989; Bhunya and others, 2007). Most of these distributions provide a good approximation for a continuous curvilinear runoff hydrograph, but do not have an upper bound that can be used to define the end of runoff (Bhunya and others, 2007). Many probability distributions do not have a simple analytical cumulative-distribution function to calculate the amount of runoff that occurs within a given time interval (Koutsoyiannis and Xanthopoulos, 1989). Parameterization of some probability distributions may need to be based on a substantial amount of data on within-storm precipitation and streamflow and characteristics from different basins so that these parameters can be used to estimate the runoff hydrograph at a site of interest. Uncertainties in input parameters such as areal rainfall estimates, flow measurements, and base-flow-separation estimates for different storms reduce the precision of parameterized runoff-hydrograph estimates (Koutsoyiannis and Xanthopoulos, 1989). Information from the disaggregation of within-storm precipitation, precipitation losses, and basin outflows at one outflow-measurement point can be ambiguous for heterogeneous river basins.

The triangular (or double-triangle) distribution was selected to develop planning-level estimates of cumulative runoff flows for sites in ungaged basins. The Soil Conservation Service (SCS) triangular hydrograph is easier to parameterize than other distributions, has an upper bound to define the end of runoff, and may provide results that are as accurate as a curvilinear hydrograph for ungaged basins (Jens and McPherson, 1964; Ogrosky and Mockus, 1964; Kent, 1973; Ward and others, 1981; Stricker and Sauer, 1982; Koutsoyiannis and Xanthopoulos, 1989; Wanielista, 1990; Wanielista and Yousef, 1993). For example, Guo and Adams (1998b) compared results calculated by a comprehensive watershed model and a simple stochastic model based on a triangular hydrograph for a 33-year period. They found that the simple triangular-hydrograph model provided runoff-population estimates that compared well with the watershed-modeling results. The triangular distribution can be fully parameterized with the area under the curve, a lower bound, an upper bound, and the location of the mode (Saucier, 2000). For a runoff hydrograph, these parameters are the total runoff volume, the start of runoff (T_0), the end of runoff (T_e), and the time to peak (T_p), respectively (fig. 33). With this information, the cumulative volume of runoff within a given time interval is simple to compute with a triangular hydrograph. The triangular distribution is commonly used as a synthetic unit hydrograph to estimate runoff flows from within-storm precipitation-excess increments. For planning-level analyses, however, the entire precipitation event may be characterized by a single increment. The triangular hyetograph may be used to estimate within-storm precipitation-excess increments, but the uncertainties in such an approach may not warrant the added complexity. For example, Naef (1981) indicated that many different unit-hydrograph shapes would produce similar levels of uncertainty, and that complex models may not provide substantial improvements for characterizing rainfall-runoff transformations.

Although the triangular distribution is a simple linear approximation to the hydrograph, the cumulative distribution of streamflow during a storm is a curvilinear S-curve. The proportion of total runoff at time T_i from the beginning of the storm for a triangular hydrograph is expressed as

$$R_c = (T_i - T_o)^2 / ((T_e - T_o) \times (T_p - T_o)) \text{ if } T_o \le T_i \le T_p \text{, and} \quad (4)$$

$$R_c = 1 - ((T_e - T_i)^2 / ((T_e - T_o) \times (T_e - T_p))) \text{ if } T_p \le T_i \le T_e \text{ ,} \quad (5)$$

where

R_c	is the cumulative proportion of the total runoff at time T_i,
T_e	is the end time of the runoff hydrograph,
T_i	is any selected time step within the runoff hydrograph,
T_o	is the begin time of the runoff hydrograph, and
T_p	is the peak time of the runoff hydrograph.

If the begin time is set to zero, the end time is equal to the duration of the runoff hydrograph T_b (fig. 33). The time to peak is commonly calculated as one-half the precipitation duration (D/2) plus a basin lag time (BL) that depends on basin characteristics.

Several formulas have been developed for calculating the basin lag time from basin characteristics (Chow, 1964; Kent, 1973; Sauer and others, 1983; Chow and others, 1988;

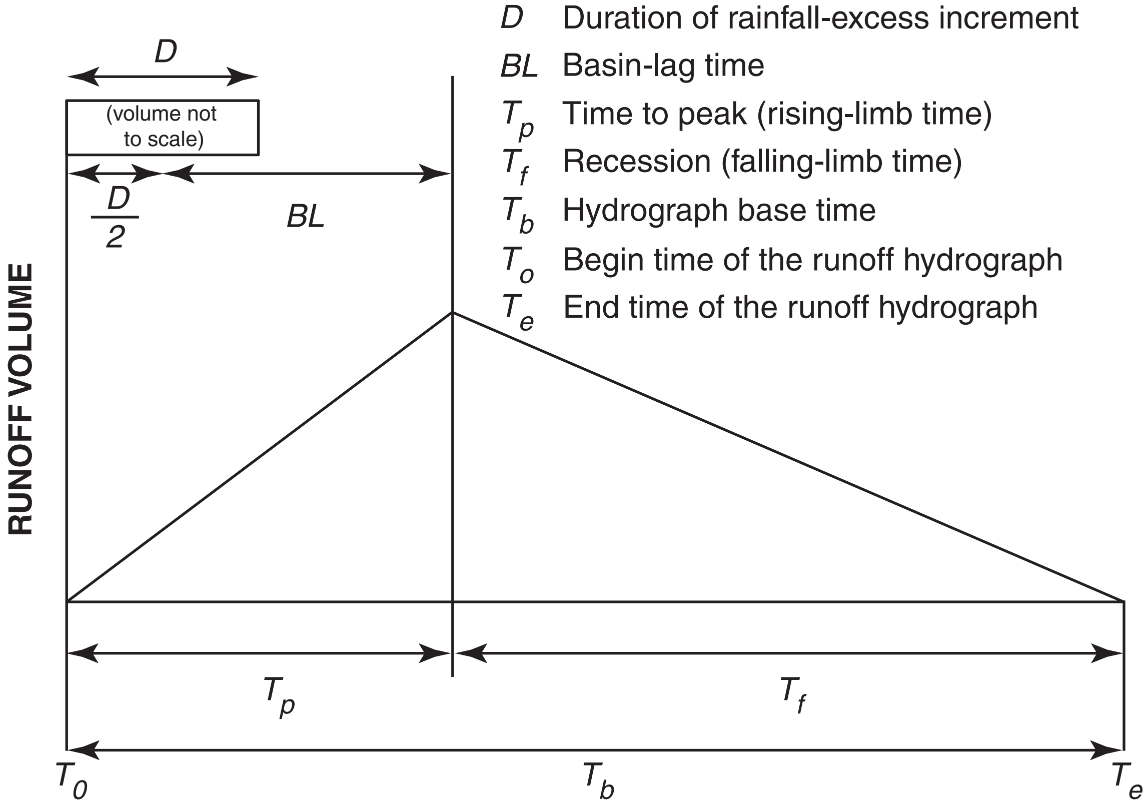


Figure 33. Schematic diagram showing time factors for a triangular storm-event hydrograph. (Modified from Kent, 1973)

Wanielista, 1990; Muzik, 1992; Pilgrim and Cordery, 1993; Wanielista and Yousef, 1993). Most basin lag equations include some measure of the basin slope and the length of flow along the main channel within the basin. Some equations also include factors that account for differences in overland or channel flow such as a runoff coefficient, CN, or a channel roughness factor. Some equations account for storm characteristics (usually rainfall intensity), but basin lag time is primarily associated with basin characteristics rather than storm characteristics (Sauer and others, 1983). Commonly used basin lag equations are based on data from a limited number of sites. For example, the Kirpich equation is based on data from only seven rural basins (Pilgrim and Cordery, 1993; Chow and others, 1988). It is recognized that the degree of uncertainty in the applicability of any basin lag formula for a given site is high.

The USGS basin lag equation developed by Sauer and others (1983) was selected for estimating a characteristic basin lag value from information available on topographic maps and aerial photographs (fig. 34). This equation is

$$BL = 0.85 \times (L / SL^{0.5})^{0.62} \times (13 - BDF)^{0.47}, \qquad (6)$$

where

BL is the basin lag time, in hours;
L is the basin length on a topographic map from the outlet to the drainage divide, in miles;
SL is the mean basin slope measured between points at 10 and 85 percent of the main channel length, in feet per mile (if SL is greater than 70 ft/mi, this value is used); and
BDF is the basin-development factor, an integer between 0 and 12 that describes the degree of drainage modification in the basin.

The basin lag equation was developed with data from 269 basins throughout the United States with drainage areas ranging from 0.2 to 100 mi^2, basin lengths from 0.47 to 88.1 mi, mean basin slopes from 3 to 500 ft/mi, TIA fractions from 3 to 50 percent of the basin area, and basin-development factors

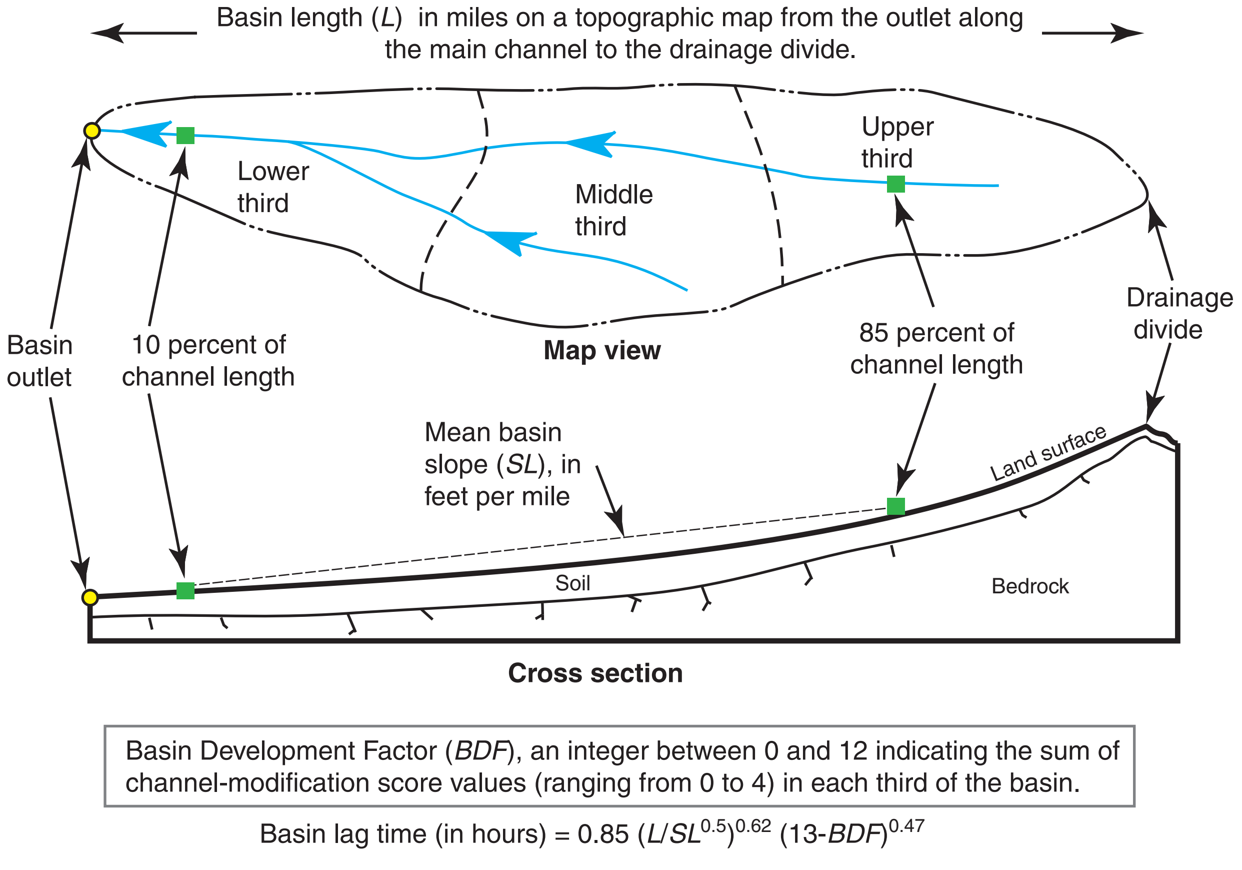


Figure 34. Schematic diagram showing the physical basin characteristics used by Sauer and others (1983) to estimate basin lag time. The equation was developed with data from 269 gaged basins with drainage areas ranging from 0.2 to 100 square miles, basin lengths from 0.47 to 88.1 miles, slopes from 3 to 70 feet per mile, basin development factors from 0 to 12, and impervious fractions from 3 to 50 percent.

from 0 to 12. Detailed descriptions of these basin characteristics are available in the National Handbook of recommended methods for water-data acquisition (U.S. Geological Survey, 1980) and other sources (Sauer and others, 1983; Federal Emergency Management Agency, 2001b; c; d; e; McCuen and others, 2002). A copy of this section of the National Handbook of recommended methods for water-data acquisition is provided on the CD–ROM accompanying this report. USGS basin lag equations are used by the FHWA (McCuen and others, 2002), the Federal Emergency Management Agency (2001a), and many state transportation agencies.

The BDF is an empirical factor characterizing urbanization and stream channelization that indicates the efficiency of the basin-drainage system (Sauer and others, 1983). The BDF is estimated by dividing the basin into equal-area thirds that drain the upper, middle, and lower parts of the drainage system (fig. 34). Each third may cut across one or more different tributary basins so that the travel distances among tributaries in each third of the basin are approximately equal. Once the basin is divided, the analyst must assign a score of one or zero to characterize each of four drainage-system components in each third of the basin. If more than 50 percent of the area in each third of the basin can be characterized as having one of the four drainage-system components, a score of 1 is given for that component in that third of the basin area.

The four drainage-system components are (1) channel improvements, (2) channel linings, (3) storm drains, and (4) curb-and-gutter streets. Channel improvements (defined in Sauer and others (1983) as improvements in flood-flow conveyance capacity rather than an improvement in ecological habitat) include straightening, enlarging, deepening, and clearing the main channel and principal tributaries. Channel linings include impervious, low-friction materials that replace natural streambed materials. Examples of channel linings include box or pipe culverts. If a stream reach has been lined, that reach also should be counted as in the channel-improvement score. Storm drains are defined as enclosed drainage structures that convey runoff from source areas to the main channel or principal tributaries. Curb-and-gutter streets are defined as roads or highways that collect and drain runoff using a conveyance system that drains to adjacent areas, storm drains, or tributary streams. To be assigned a score of one for the curb-and-gutter-street category, more than 50 percent of the area must be urbanized, and more than 50 percent of the streets in the area must be drained by curb and gutter.

Under these definitions, the BDF ranges from 0 (for a natural basin) to 12 (for an urbanized basin with storm drains and culverted streams). More detailed examples of the BDF scoring system are provided by Sauer and others (1983), the Federal Emergency Management Agency (2001b; c; d; e), and McCuen and others (2002). Sauer and others (1983) indicate that this binary four-category ranking system seems to produce consistent scores among similar basins by different analysts.

Use of the BDF requires some knowledge of the upstream basin. The BDF is the most interpretive term in equation 6, is not commonly reported in runoff studies, and is not readily characterized using a single GIS coverage. For example, the degree of imperviousness within a basin is readily derived from available GIS coverages and, in theory, should be a good predictor for the BDF. However, the rank correlation coefficient for the BDF and the impervious percentage based on basin-characteristics data from Sauer and others (1983) is 0.5. This rank correlation coefficient indicates a positive correlation between these variables with a substantial amount of scatter. The binary BDF classification system produces an integer scale for the BDF, and one particular score does not define a unique set of conditions for the basins. For example, a rural basin channelized for agricultural drainage may have a BDF of 3, which would exceed a BDF score of 2 for a basin with a lower third that is fully urbanized with curb-and-gutter streets and storm sewers that drain to a natural channel. The feasibility of automating the BDF scoring system is expected to increase as information about the degree of imperviousness (from land-use or land-cover data); flood-control features; and private, municipal, and transportation drainage systems becomes widely available in GIS formats.

The effects of errors in BDF specification on the calculated basin lag time depend on the BDF value itself. The basin lag time for an undeveloped basin (with a BDF of 0) is about 3.3 times the basin lag time for a completely developed basin (with a BDF of 12) with the same length and slope (eq. 6). Misspecification of the BDF score by 1 has a minor effect (of about 5 percent) on the basin lag time estimate below a BDF value of 6. As the BDF increases to 12, a specification error of 1 can change the basin lag time estimate by as much as 32 percent. Although the potential effects of BDF specification errors increase with increases in urbanization, this increase may be offset by the availability of more detailed drainage information for highly developed areas.

Once the basin lag time and, therefore, the time to peak is defined, the falling-limb time (T_f) must be estimated to determine the end time of the runoff hydrograph (T_e). Hydrograph recession-time studies are not common in the literature because most high-flow studies focus on the basin lag and magnitude of the peak flow to provide information for flood control. Several approximations are commonly used without supporting data. A rough recession-time approximation, in which the falling-limb duration (in days) is equal to the drainage area (in mi^2) raised to the power 0.2, is commonly used for base-flow separation (Linsley and others, 1975; Sloto and Crouse, 1996). This approximation, however, does not account for the basin slope or drainage features that affect the recession time. The falling-limb duration commonly is estimated using a hydrograph-recession ratio, defined as the ratio of the durations of the falling and rising limbs. The recession-time estimates used with the rational method are based on the assumption of an isosceles triangle with equal rising- and falling-limb durations (a hydrograph-recession ratio of 1). The falling-limb duration of the SCS triangular hydrograph has a standard hydrograph-recession ratio of 1.67 times the duration of the rising limb (Ogrosky and Mockus, 1964; Kent, 1973; Pilgrim and Cordery, 1993). Wanielista (1990) indicated that the hydrograph-recession ratio may be about 1.25 in steep-sloped urban drainage basins, 2.25 for mixed-use moderately sloped basins, 5.5 for rural basins with low slopes, and 12 for rural basins in flat areas. These ratios include the effects of slope and basin development but do not quantify the effect of each factor. For example, a flat basin with improved drainage may have a time to peak (T_p) equivalent to that of a higher slope basin with natural drainage. Although the time to peak may be similar for the two basins, the drainage structures in the more developed basin may attenuate runoff components such as throughflow that have longer response times than engineered drainage from impervious surfaces. Furthermore, the underlying interpretation, data, and basin characteristics used for derivation of these ratios are not published (Wanielista, 1990; Wanielista and Yousef, 1993).

Some studies do provide information that may be used to guide recession-time estimates. The effects of land use on recession times have been documented in several studies. For example, Shirmohammadi and others (1987) indicate that channelization decreases the storm-hydrograph base time. Stricker and Sauer (1982) developed a nomograph for estimating the hydrograph duration with flood data from 62 streamgages in different states. This method is based on the assumption that the dimensionless hydrograph is an isosceles triangle. This approach, however, requires the use of a rural-flood equation and an urban-adjustment regression equation for a given high-flow return period. Shamir and others (2005) examined data from 19 USGS streamgages in different areas of the country with drainage areas ranging from 86 to 1,850 mi^2 to determine rising and falling-limb densities for use in rainfall-runoff models. In this study, the basin-average rising-limb and falling-limb densities indicate that the hydrograph-recession ratio ranged from about 1.7 to about 3.5 (with a median of about 2.3). Shamir and others (2005) reported that average rising and falling-limb durations decreased with factors such as the ratio of flow length to basin area, the percentage of forest cover, daily mean precipitation, and minimum January temperature, but that these individual correlations were relatively weak. Shuster and others (2008) analyzed streamflow data from eight small drainage areas (ranging from 6 to 23 mi^2), predominantly agricultural basins in southwestern Ohio; their data indicate that basin-average hydrograph-recession ratios were between 1.8 and 5 (with a median of about 3.55). Nonparametric rank correlation coefficients (rho) indicate weak positive associations between the recession ratios and drainage areas (rho=0.44); the percentage of forested area, which ranged from 3 to 39.7 percent of total drainage-basin area (rho=0.42); and channel slope, which ranged from 0.4 to 2 percent (rho=0.28). The percentage of urban area, which ranged from 0.3 to 2.4, had a moderately strong negative correlation (rho=-0.67), and the percentage of agricultural area, which ranged from 42.5 to 93.6, had a weak negative correlation (rho=-0.26) with the recession ratio. Shuster and others (2008) used interbasin comparisons to indicate that hydrograph rise and fall rates are characteristics of the basin rather than characteristics of individual storm events.

Liscum (2001) developed regression equations to describe storm-discharge hydrographs with data collected at 42 sites from 1,089 storm events near Houston, Texas, during the period 1964 through 1989. The drainage area for these sites ranged from 0.13 to 182 mi^2, the BDFs ranged from 0 to 12, the percentage of developed area ranged from about 15 to 100 percent of the drainage-basin area, and mean basin slopes in this area ranged from about 2.5 to 8.8 ft/mi (Liscum, 1997). In comparison, the mean basin slopes reported by Sauer and others (1983) ranged from about 1 to 400 ft/mi. The mean basin slopes reported by Liscum and others (1997) are in the lowest 20th percentile of mean basin slopes reported by Sauer and others (1983). The equations developed by Liscum (2001) indicate that the basin lag time for a fully developed basin (BDF=12) would be about 11 percent of the basin lag time for the same basin if it were undeveloped. Similarly, the runoff duration in a fully developed basin is about 44 percent of the runoff duration in an undeveloped basin. The falling-limb time from the peak of the hydrograph for a fully developed basin is about 36 percent of the falling-limb time for an undeveloped basin. These hydrograph-recession equations indicate that, in the Houston area, the storm falling-limb time is about 3.6 times the basin lag time for an undeveloped basin (BDF=0), about 5 times for a developed basin (BDF=6), and about 13 times for a fully developed basin (BDF=12). These ratios increase with increasing BDF values because the reduction in the basin lag time is much greater than the increase in duration of the falling-limb time. These hydrograph-recession ratios were derived from regression equations developed with data from relatively flat basins and for curvilinear storm-event hydrographs.

The information provided by Liscum (2001) may be used to derive estimates of hydrograph-recession ratios for planning-level water-quality analyses in conjunction with estimated BDFs. The hydrograph-recession ratio for a triangular hydrograph must be adjusted to preserve the approximate recession-mass curve of a curvilinear hydrograph with a straight-line recession segment. The falling-limb time of the triangular hydrograph should be about 50 percent as long as the 99th percentile of the falling-limb time of the hydrograph recession ratio for a curvilinear hydrograph. Thus, halving the hydrograph-recession ratios that are derived from Liscum's (2001) regression equations may provide acceptable ratios for triangular runoff hydrographs. For example, one-half of the undeveloped (BDF=0) hydrograph-recession ratios derived from Liscum (2001) are about 1.8 times the basin lag time; this ratio approximates the SCS triangular hydrograph recession value of 1.67.

Summary

The U.S. Geological Survey (USGS), in cooperation with the Federal Highway Administration (FHWA), developed information and statistics to characterize prestorm flows, storm-event precipitation, and runoff coefficients for use with the stochastic empirical loading and dilution model (SELDM). SELDM is a water-quality model that uses a mass-balance approach with Monte Carlo methods to generate planning-level estimates of water-quality constituent concentrations, discharges, and loads in highway runoff and in the receiving stream upstream and downstream of the highway-runoff outfall at unmonitored sites in the conterminous United States. Planning-level estimates are defined as the results of analyses that are recognized to include substantial uncertainties (commonly orders of magnitude). Planning-level estimates of stormflow for a site of interest can be made using statistics in the literature, regional statistics, statistics estimated using data collected at nearby hydrologically similar sites, or with statistics estimated using limited data collected at the site of interest. Statistics describing the frequency distributions of component discharges and concentrations can be used to estimate the statistics for downstream discharges, concentrations, and loads with a mass-balance model. These statistics indicate the potential for exceeding water-quality criteria and, therefore, the potential need for more information and data that may be used to evaluate highway-runoff discharges as a potential source of water-quality constituents, the potential effects of runoff loads on receiving-water quality, and the potential effectiveness of best management-practice (BMP) structures at a site of interest. The information and data developed in this report also may be useful with other methods to estimate the effects of runoff on receiving-water quality. For example, the "Simple Method" (Schueler, 1987; Chandler, 1994) commonly is used to develop estimates of long-term annual loads for initial screening-level runoff-quality analyses.

The FHWA has established a system of water-quality-assessment and action plans that include different levels of interpretive analysis to determine the potential environmental effects of highway runoff. The data-quality objectives (DQOs) for these assessments depend on the level of interpretive analysis deemed necessary to evaluate conditions for a given site. This compilation and interpretation of national prestorm flow, precipitation statistics, and rainfall-runoff transformations by the USGS in cooperation with the FHWA are designed to meet DQOs for development and refinement of planning-level estimates of stream-water quality at unmonitored sites in the conterminous United States. The current study was designed to provide methods to derive planning-level estimates of storm-event upstream flows for unmonitored sites that may receive highway runoff. Such estimates are based on statistics for prestorm flows, storm-event characteristics, and rainfall-runoff transformation statistics, each with a substantial amount of uncertainty. This study also provides methods useful in obtaining and interpreting more precise site-specific estimates.

Prestorm streamflows, which are modeled using daily mean flow statistics, may be a large component of total storm-event flow in the receiving waters. Streamflow statistics were estimated by analysis of data from 2,873 USGS streamgages in the conterminous United States with drainage areas ranging from 10 to 500 mi^2 and at least 24 years of record during the period 1960–2003. Graphical and statistical examination of streamflow records indicates that long-term daily mean streamflow statistics can be used as planning-level estimates for prestorm streamflow. This is because, over a long period of time, storm events may occur after dry, wet, or normal antecedent flows. Five computer programs were developed for obtaining and analyzing this National Water Information System Web streamflow data. Streamflow statistics were regionalized according to U.S. Environmental Protection Agency (USEPA) Level III nutrient ecoregions (2003). Initial estimates of prestorm flow statistics were made by using the drainage-area-ratio method with regional statistics. Regression equations were developed to modify drainage-area-ratio estimates for regions with systematic changes in streamflow statistics with increasing drainage area. Initial estimates also may be refined with statistics from nearby hydrologically similar basins. If limited data are available from the site of interest, streamflow-correlation methods may be used to estimate site-specific statistics. If a long-term record on daily mean flows is available for a site of interest, then site-specific statistics can be calculated.

Many probability distributions have been used to characterize streamflow statistics. The lognormal distribution is used extensively in runoff-quality analysis; however, the skew coefficients of the logarithms of daily mean streamflow are substantially different from zero for many of the streamgages. The log-Pearson Type III distribution was selected to model prestorm streamflows because it is a very flexible distribution that can provide a good fit to many different types of hydrologic data even if the underlying population is not a pure log-Pearson Type III distribution. The mean, standard deviation, and skew of the logarithms of daily mean streamflow data can be used to stochastically generate a log-Pearson Type III distribution of values by means of standard frequency-factor methods and the modified Wilson-Hilferty approximation. If a stream is intermittent or ephemeral, standard conditional-probability methods may be used in the stochastic data-generation process to adjust prestorm streamflow statistics to account for the proportion of zero streamflows at a site of interest.

Storm-event precipitation statistics were estimated by analysis of data from 2,610 National Oceanic and Atmospheric Administration hourly-precipitation data stations in the conterminous United States with at least 25 years of data during the 1965–2006 period. The synoptic rainfall-data analysis program (SYNOP), the preprocessor SYNPREP, and the synoptic precipitation-analysis facilitator (SPAF) were compiled and used for this study. The results are recorded in the Microsoft Access database (SiteStormV01.mdb) on the CD–ROM accompanying this report. This database was designed to store the data

and facilitate analysis of site characteristics, precipitation, and runoff data. Storm-event statistics in the database were regionalized according to USEPA rain zones. Initial estimates based on regional statistics may be refined with statistics from nearby hourly-precipitation data stations.

Storm events that produce runoff are commonly defined by a minimum precipitation volume, a minimum interevent time, a total volume, a duration, and the average time between event midpoints. The minimum precipitation volume is used to determine if a storm will be included in the analysis of runoff-producing events. The minimum interevent time is used to define independent storm events. It is the minimum number of dry hours that must occur between precipitation measurements to define a new storm. The minimum precipitation volume was set at 0.1 in., and the minimum interevent time was set at 6 hours to be consistent with storm-event definitions currently used by the USEPA and the FHWA. The Poisson distribution was selected to model the number of discrete runoff-producing storm events per year because storm-event occurrence is commonly modeled as a Poisson process. The two-parameter exponential distribution was selected to model other storm-event characteristics because this distribution preserves the storm-event statistics themselves and the lower limit of each storm statistic, and the distribution is readily implemented in a stochastic data-generation algorithm.

Statistics to characterize runoff coefficients (calculated as the ratio of runoff, in watershed inches, to rainfall, in inches) were estimated using data from 6,142 storm events at 306 study sites. The results of this analysis are recorded in the Microsoft Access database (SiteStormV01.mdb) available on the CD–ROM accompanying this report. In this report, the term runoff is used to include all stormflow-generating mechanisms, including infiltration-excess overland flow, saturation overland flow, throughflow, near-stream discharge caused by groundwater ridging, and direct precipitation. Rainfall-runoff transformation statistics are not regionalized but are organized by total impervious area.

Many rainfall-runoff datasets include runoff-coefficient values that are greater than one. This is because there are many sources of systematic and random errors in precipitation and runoff measurements. Sources of systematic error may include drainage-area delineation errors, impervious-area characterization errors, bias in the representativeness of rain-gage monitoring location(s), bias in measurement errors, and disparate hydrograph-separation methods. Sources of random uncertainty may include variations in antecedent conditions, random variation in the representativeness of rain-gage monitoring location(s), variations in measured values, and variations in the accuracy of hydrograph-separation methods.

Regression equations were developed to estimate the average, standard deviation, and skew coefficient of runoff coefficients from the estimated total impervious area. The regression method was selected because it provides estimates that meet DQOs for planning-level runoff-quality analysis. The average, standard deviation, and skew coefficient of runoff coefficients can be used to stochastically generate a Pearson Type III distribution of runoff coefficients by using standard frequency-factor methods and the modified Wilson-Hilferty approximation. Standard acceptance-rejection methods can be used to discard runoff coefficients that are greater than one or less than zero.

Information about the storm-event hydrographs for runoff from the highway catchment and the upstream basin is necessary to estimate the quantity of upstream flow that occurs concurrently with the highway runoff. The focus of planning-level analyses of highway-runoff-quality analyses has traditionally been on event-mean concentrations and total storm loads for the entire event rather than on within-event processes. However, the differences in the locations, sizes, and drainage characteristics of the highway catchment and the upstream basin may cause differences in the timings and durations of runoff from each area. If the highway catchment is small and the runoff drains directly to the stream, the duration of appreciable runoff from the highway catchment may be approximated by the duration of the precipitation event. If the upstream basin is relatively large and more pervious than the highway catchment, appreciable runoff from the basin may continue for hours or days longer than runoff from the highway catchment. In this case, only a small proportion of the upstream runoff may be available to dilute highway-runoff constituents in the receiving waters. If, however, a structural BMP is employed at the highway site to attenuate and extend the highway-runoff hydrograph, then much more of the upstream runoff may be available to dilute highway-runoff constituents in the receiving waters. Detailed characterizations of within-storm processes are beyond the scope of a planning-level water-quality analysis, but a systematic method is necessary to estimate the duration of the highway-runoff hydrograph and the proportion of upstream flows that may occur during a highway-runoff event.

The triangular (or double-triangle) distribution was selected to develop planning-level estimates of cumulative runoff flows for sites in ungaged basins. The triangular hydrograph is easier to parameterize than other distributions, has an upper bound to define the end of runoff, and may provide results that are as accurate as a curvilinear hydrograph. The triangular distribution can be fully parameterized with the area under the curve, a lower bound, an upper bound, and the location of the mode. For a runoff hydrograph, these parameters are the total runoff volume, the start of runoff, the end of runoff, and the time to peak, respectively. The time to peak may be estimated by using the USGS basin lag equation, which is based on basin properties and a basin development factor. The runoff- hydrograph duration can be estimated on the basis of a user-selected hydrograph-recession ratio, defined as the ratio of the time from the peak to the end of runoff divided by the time to peak. Values of the hydrograph-recession ratio in the literature range from about 1 for steep highly developed basins to as high as 13 for low-gradient rural basins.

Acknowledgments

The author thanks William Kirby and Ken Eng of the U.S. Geological Survey for providing input and information on calculation and use of streamflow statistics. Eric Strecker and Tarig Omer, provided FORTRAN source code for an updated version of the FHWA PC-SYNOP program that was modified and used to calculate precipitation statistics. Patricia A. Cazenas and Susan C. Jones of the Federal Highway Administration and Leslie DeSimone, Tomas Over, Audrey L. Ishii, Elizabeth Murphy, David Soong, Mary Ashman, Kevin Breen, and Robert F. Breault, of the U.S. Geological Survey provided thoughtful and thorough technical and editorial reviews of this report, the software appendixes, and the associated CD–ROM.

References Cited

Adams, B.J., Fraser, H.G., Howard, C.D.D., and Hanafy, M.S., 1986, Meteorological data analysis for drainage system design, Journal of Environmental Engineering, v. 112, no. 5, p. 827–848.

Adams, B.J., and Papa, Fabian, 2000, Urban stormwater management planning with analytical probabilistic models: New York, John Wiley & Sons, Inc., 358 p.

Allen, H.E., Jr., and Gray, J.R., 1984, Runoff, sediment transport, and water quality in a northern Illinois agricultural watershed before urban development, 1979–81: U.S. Geological Survey Water-Resources Investigations Report 82–4073, 55 p.

Alley, W.M., 1977, Guide for collection, analysis, and use of urban stormwater data: A Conference Report, Easton, Md., November 28–December 3, 1976: American Society of Civil Engineers, 115 p.

Anning, D.W., and Konieczki, A.D., 2005, Classification of hydrogeologic areas and hydrogeologic flow systems in the Basin and Range Physiographic Province, southwestern United States: U.S. Geological Survey Professional Paper 1702, 37 p., 2 pls.

Asquith, W.H., Roussel, M.C., Cleveland, T.G., Fang, Xing, and Thompson, D.B., 2006, Statistical characteristics of storm interevent time, depth, and duration for eastern New Mexico, Oklahoma, and Texas: U.S. Geological Survey Professional Paper 1725, 299 p.

Athayde, D.N., Shelley, P.E., Driscoll, E.D., and Boyd, G., 1983, Results of the nationwide urban runoff program, volume I, final report: Washington, D.C., U.S. Environmental Protection Agency, WH-554, 186 p.

Bailey, R.G., 1984, Testing an ecosystem regionalization: Journal of Environmental Management, v. 19, p. 239–248.

Baldys, Stanley, III, Raines, T.H., Mansfield, B.L., and Sandlin, J.T., 1996, Hydrologic data for urban stormwater studies in the Dallas-Fort Worth area, Texas, 1992–94: U.S. Geological Survey Open-File Report 96–482, 256 p.

Barnwell, T.O., Jr., and Krenkel, P.A., 1982, Use of water quality models in management decision making: Water Science and Technology, v. 14, no. 9–11, p. 1095–1107.

Barrett, M.E., Quenzer, A.M., and Maidment, D.R., 1998, Water quality and quantity inputs for the urban creeks future needs assessment: University of Texas at Austin, Center for Research in Water Resources Bureau of Engineering Research Report 98–10, 74 p.

Becciu, Gianfranco, and Paoletti, Alessandro, 2000, Moments of runoff coefficient and peak discharge estimation in urban catchments: Journal of Hydraulic Engineering, v. 5, no. 2, p. 197–205.

Benson, M.A., and Carter, R.W., 1973, A national study of the streamflow data collection program: U.S. Geological Survey Water-Supply Paper 2028, 44 p.

Bent, G.C., and Archfield, S.A., 2002, A logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 2002–4043, 45 p.

Bent, G.C., and Steeves, P.A., 2006, A revised logistic regression equation and an automated procedure for mapping the probability of a stream flowing perennially in Massachusetts: U.S. Geological Survey Scientific Investigations Report 2006–5031, 107 p.

Bhunya, P.K., Berndtsson, R., Ojha, C.S.P, and Mishra, S.K., 2007, Suitability of gamma, Chi-square, Weibull, and Beta distributions as synthetic unit hydrographs: Journal of Hydrology, v. 334, p. 28–38.

Bishop, G.D., and Church, M.R., 1995, Mapping long-term regional runoff in the eastern United States using automated approaches: Journal of Hydrology, v. 169, p. 189–207.

Bobee, Bernard, and Ashkar, Fahim, 1991, The gamma family and derived distributions applied in hydrology: Littleton, Colo., Water Resources Publications, 203 p.

Breault, R.F., Sorenson, J.R., and Weiskel, P.K., 2002, Streamflow, water quality, and contaminant loads in the lower Charles River watershed, Massachusetts, 1999–2000: U.S. Geological Survey Water-Resources Investigations Report 02–4137, 139 p.

Brown, J.W. (ed.), 2006, Eco-logical—An ecosystem approach to developing infrastructure projects: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, FHWA-HEP-06-011, 96 p.

Busby, M.W., 1966, Annual runoff in the conterminous United States, 1951–80: U.S. Geological Survey Hydrologic Investigations Atlas HA–212, 1 pl.

Campbell, J.B., 1987, Rainfall-runoff data for Somerset County, New Jersey: U.S. Geological Survey Open-File Report 87–384, 161 p.

Carr, D.B., Olsen, A.R., Pierson, S.M., and Courbois, J.Y.P., 2000, Using linked micromap plots to characterize Omernik Ecoregions: Data Mining and Knowledge Discovery, v. 4, no. 1, p. 43–67.

Caruso, J.C., and Cliff, Norman, 1997, Empirical size, coverage, and power of confidence intervals for Spearman's Rho: Educational and Psychological Measurement, v. 57, no. 4, p. 637–654.

Cazenas, P.A., Bank, F.G., and Shrouds, J.M., 1996, Guidance on developing water-quality action plans: Washington, D.C., Department of Transportation, Federal Highway Administration, Office of Natural Environment, 9 p.

Chandler, R.D., 1994, Estimating annual urban nonpoint pollutant loads: Journal of Management in Engineering, v. 10, no. 6, p. 50–59.

Chen, Jieyun, and Adams, B.J., 2005, Analysis of storage facilities for urban stormwater quantity control: Advances in Water Resources, v. 28, p. 377–392.

Chen, Jieyun, and Adams, B.J., 2007, Development of analytical models for estimation of urban stormwater runoff: Journal of Hydrology, v. 336, p. 458–469.

Cheng, K.S., Ciang, J.L., and Hsu, C.W., 2007, Simulation of probability distributions commonly used in hydrological frequency analysis: Hydrological Processes, v. 51, p. 51–60.

Chow, V.T., 1964, Section 14–Runoff: *in* Chow, V.T. (ed.), Handbook of applied hydrology—A compendium of water-resources technology: New York, McGraw-Hill Book Co., p. 14.1–20.54.

Chow, V.T., Maidment, D.R., and Mays, L.W., 1988: Applied Hydrology, New York, McGraw-Hill, Inc., 572 p.

Church, P.E., Granato, G.E., and Owens, D.W., 2003, Basic requirements for collecting, documenting, and reporting precipitation and stormwater-flow measurements: *in* Granato, G.E., Zenone, C., and Cazenas, P.A. (eds.), National highway runoff water-quality data and methodology synthesis, v. I, Technical issues for monitoring highway runoff and urban stormwater: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, FHWA-EP-03-054, p. 113–163.

Clar, M.J., Barfield, B.J., and O'Connor, T.P., 2004, Stormwater best management practice design guide—Volume I–General considerations: U.S. Environmental Protection Agency Report EPA/600/R-04/121, 179 p.

Coes, A.L., and Pool, D.R., 2005, Ephemeral-stream channel and basin-floor infiltration and recharge in the Sierra Vista subwatershed of the Upper San Pedro basin, southeastern Arizona: U.S. Geological Survey Open-File Report 2005–1023, 84 p.

Daly, Christopher, Neilson, R.P., and Phillips, D.L., 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain: Journal of Applied Meteorology, v. 33, no. 2, p. 140–158.

Demcheck, D.K., Frederick, C.P., and Johnson, K.L., 1998, Water-quality characteristics of urban storm runoff at selected sites in East Baton Rouge Parish, Louisana, April 1993 through June 1995: U.S. Geological Survey Open-File Report 98–565, 59 p.

Devroye, Luc, 1986, Non-uniform random variate generation: New York, Springer-Verlag, 843 p.

DiToro, D.M., 1984, Probability model of stream quality due to runoff: Journal of Environmental Engineering, v. 110, no. 3, p. 607–628.

Driscoll, E.D., DiToro, D.M., Gaboury, D., and Shelley, P.E., 1986, Methodology for analysis of detention basins for control of urban runoff quality: U.S. Environmental Protection Agency Report EPA440/5-87-001, 74 p.

Driscoll, E.D., DiToro, D.M., and Thomann, R.V., 1979, A statistical method for assessment of urban runoff: Washington, D.C., U.S. Environmental Protection Agency, EPA 440/3-79-023, 200 p.

Driscoll, E.D., Palhegyi, G.E., Strecker, E.W., and Shelley, P.E., 1989, Analysis of storm event characteristics for selected rainfall gages throughout the United States: U.S. Environmental Protection Agency, OCLC 30534890, 43 p.

Driscoll, E.D., Shelley, P.E., Gaboury, D.R., and Salhotra, Atul, 1989, A probabilistic methodology for analyzing water quality effects of urban runoff on rivers and streams: U.S. Environmental Protection Agency, Office of Water, EPA 841-R89-101, 128 p.

Driscoll, E.D., Shelley, P.E., and Strecker, E.W., 1990a, Pollutant loadings and impacts from highway stormwater runoff, Volume I–Design procedure: U.S. Federal Highway Administration Final Report, FHWA-RD-88-006, 67 p.

Driscoll, E.D., Shelley, P.E., and Strecker, E.W., 1990b, Pollutant loadings and impacts from highway stormwater runoff, Volume III–Analytical investigation and research report: Federal Highway Administration Final Report FHWA-RD-88-008, 160 p.

Driver, N.E., and Tasker, G.D., 1990, Techniques for estimation of storm-runoff loads, volumes, and selected constituent concentrations in urban watersheds in the United States: U.S. Geological Survey Water-Supply Paper 2363, 44 p.

Duckstein, Lucien, Fogel, M.M., and Kisiel, C.C., 1972, A stochastic model of runoff-producing rainfall for summer type storms: Water Resources Research, v. 8, no. 2, p. 410–421.

Duncker, J.J., and Melching, C.S., 1998, Regional rainfall-runoff relations for simulation of streamflow for watersheds in Du Page County, Illinois: U.S. Geological Survey Water-Resources Investigations Report 98–4035, 80 p.

Eng, Ken, Tasker, G.D., and Milly, P.C.D., 2005, An analysis of region-of-influence methods: Journal of the American Water Resources Association, v. 41, no. 1, p. 135–143.

Federal Emergency Management Agency, 2001a, Use of USGS regression equations in the National Flood Insurance Program (NFIP): National Flood Frequency Bulletin, Issue No. 3, available at http://www.floodmaps.fema.gov/listserv/nf_jul01.htm.

Federal Emergency Management Agency, 2001b, Discussion on USGS regression equations and the NFF program—Unusual parameters of USGS regression equations and how to obtain them–Part 1: National Flood Frequency Bulletin Issue No. 5, available at http://www.floodmaps.fema.gov/listserv/nf_sep01.htm.

Federal Emergency Management Agency, 2001c, Discussion on USGS regression equations and the NFF program—Unusual parameters of USGS regression equations and how to obtain them–Part 2: National Flood Frequency Bulletin, Issue No. 6, available at http://www.floodmaps.fema.gov/listserv/nf_oct01.htm.

Federal Emergency Management Agency, 2001d, Discussion on USGS regression equations and the NFF Program—Unusual parameters of USGS regression equations and how to obtain them–Part 3: National Flood Frequency Bulletin, Issue No. 7, available at http://www.floodmaps.fema.gov/listserv/nf_nov01.htm.

Federal Emergency Management Agency, 2001e, Discussion on USGS Regression Equations and the NFF Program—Examples in which USGS regression equations are used for NFIP purposes: National Flood Frequency Bulletin, Issue No. 8, available at http://www.floodmaps.fema.gov/listserv/nf_dec01.htm.

Federal Highway Administration, 1998, Environmental flow charts: Washington, D.C., Department of Transportation, Federal Highway Administration, Office of Natural Environment, 45 p.

Flynn, K.M., Hummel, P.R., Lumb, A.M., and Kittle, J.L., Jr., 1995, User's manual for ANNIE, version 2, a computer program for interactive hydrologic data management: U.S. Geological Survey Water-Resources Investigations Report 95–4085, 211 p.

Fossum, K.D., and Davis, R.G., 1996, Physical, chemical, biological, and toxicity data from the study of urban stormwater and ephemeral streams, Maricopa County, Arizona, water years 1992–95: U.S. Geological Survey Open-File Report 96–394, 71 p.

Franchini, Marco, Galeati, Giorgio, and Lolli, Marco, 2005, Analytical derivation of the flood frequency curve through partial duration series analysis and a probabilistic representation of the runoff coefficient: Journal of Hydrology, v. 303, p. 1–15.

Friesz, P.J., and Ries, K.G., III, 1998, Mean annual streamflow at reference-stream sites in three ecological subregions of Massachusetts: U.S. Geological Survey Open-File Report 98–562, 13 p.

Gebert, W.A., Graczyk, D.J., and Krug, W.R., 1987, Average annual runoff map of the United States, 1951–80: U.S. Geological Survey Hydrologic Investigations Atlas HA–710, 1 pl.

Gentle, J.E., 2003, Random number generation and Monte Carlo Methods (2d ed.): New York, Springer Science+Business Media, Inc., 381 p.

Gerlach, A.C., (ed.) 1970, The National Atlas of the United States of America: Washington, D.C., United States Department of the Interior Geological Survey, 417 p.

Goforth, G.F., Heaney, J.P., and Huber, W.C., 1983, Comparison of basin modeling techniques: Journal of Environmental Engineering, v. 109, no. 5, p. 1082–1098.

Gottschalk, L., and Weingartner, R., 1998, Distribution of peak flow derived from a distribution of rainfall volume and runoff coefficient, and a unit hydrograph: Journal of Hydrology, v. 208, p. 148–162.

Granato, G.E., 2006, Kendall-Theil Robust Line (KTRLine–version 1.0)—A Visual Basic program for calculating and graphing robust nonparametric estimates of linear-regression coefficients between two continuous variables: Techniques and Methods of the U.S. Geological Survey, book 4, chap. A7, 31 p.

Granato, G.E., 2009, Computer programs for obtaining and analyzing daily mean streamflow data from the U.S. Geological Survey National Water Information System Web: U.S. Geological Survey Open-File Report 2008–1362, CD–ROM.

Granato, G.E., Bank, F.G., and Cazenas, P.A., 2003, Data quality objectives and criteria for basic information, acceptable uncertainty, and quality-assurance and quality-control documentation, *in* Granato, G.E., Zenone, C., and Cazenas, P.A. (eds.), National highway runoff water-quality data and methodology synthesis, Volume I–Technical issues for monitoring highway runoff and urban stormwater: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, FHWA-EP-03-054, p. 3–21.

Granato, G.E., Carlson, C.S., and Sniderman, B.S., 2009, Methods for development of planning-level estimates of water quality at unmonitored stream sites in the conterminous United States: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, Office of Natural and Human Environment, FHWA-HEP-09-003, 53 p.

Granato G.E., and Cazenas, P.A., 2009, Highway-Runoff Database (HRDB Version 1.0): A data warehouse and preprocessor for the stochastic empirical loading and dilution model: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, Office of Natural and Human Environment, FHWA-HEP-09-004, 57 p.

Griffith, J.A., Martinko, E.A., Whistler, J.L., and Price, K.P., 2002, Interrelationships among landscapes, NDVI, and stream water quality in the U.S. central plains: Ecological Applications, v. 12, no. 6, p. 1702–1718.

Guay, J.R., 1996, Effects of increased urbanization from 1970’s to 1990’s on storm-runoff characteristics in Perris Valley, California: U.S. Geological Survey Water-Resources Investigations Report 95–4273, 59 p.

Gungle, Bruce, 2006, Timing and duration of flow in ephemeral streams of the Sierra Vista Subwatershed of the Upper San Pedro Basin, Cochise County, Southeastern Arizona: U.S. Geological Survey Scientific Investigations Report 2005–5190, 47 p.

Guo, Yiping, and Adams, B.J., 1998a, Hydrologic analysis of urban catchments with event-based probabilistic models I–Runoff volume: Water Resources Research, v. 34, no. 12, p. 3421–3431.

Guo, Yiping, and Adams, B.J., 1998b, Hydrologic analysis of urban catchments with event-based probabilistic models II–Peak discharge rate: Water Resources Research, v. 34, no. 12, p. 3433–3443.

Haan, C.T., 1977, Statistical methods in hydrology: Ames, Iowa, Iowa State University Press, 378 p.

Harremoës, Poul, and Madsen, H., 1999, Fiction and reality in the modeling world—balance between simplicity and complexity, calibration and identifiability, verification and falsification: Water Science and Technology, v. 39, no. 9, p. 1–8.

Heiskary, S., Wilson, C.B., and Larsen, D.P., 1987, Analysis of regional patterns in lake water quality—Using ecoregions for lake management in Minnesota: Lake and Reservoir Management, v. 3, p. 337–344.

Hejl, H.R., Jr., 1980, Preliminary appraisal of ephemeral-streamflow characteristics as related to drainage area, active-channel width, and soils in northwestern New Mexico: U.S. Geological Survey Open-File Report 81–64, 19 p.

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

Hirsch, R.M., 1979, An evaluation of some record extension techniques: Water Resources Research, v. 15, no. 6, p. 1781–1790.

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

Holmstrom, B.K., Olson, D.L., and Ellefson, B.R., 2000, Water Resources Data, Wisconsin, Water Year 1999: U.S. Geological Survey Water Data Report WI-99-1, 608 p.

Hornbeck, 1973, Storm flow from hardwood-forested and cleared watersheds in New Hampshire: Water Resources Research, v. 9, no. 2, p. 346–354.

Horwatich, J.A., Corsi, S.R., and Bannerman, R.T., 2004, Effectiveness of a pressurized stormwater filtration system in Green Bay, Wisconsin: A study for the Environmental Technology Verification Program of the U.S. Environmental Protection Agency: U.S. Geological Survey Scientific Investigations Report 2004–5222, 19 p.

Hosking, J.R.M., and Wallis, J.R., 1997, Regional frequency analysis—An approach based on L-moments: New York, N.Y., Cambridge University Press, 224 p.

Hren, J., Chaney, T.H., Norris, J.M., and Childress, C.J.O., 1987, Water-quality data-collection activities in Colorado and Ohio Phase I—Inventory and evaluation of 1984 programs and costs: U.S. Geological Survey Water-Supply Paper 2295–A, 71 p.

Hughes, R.M., and Larsen, D.P., 1988, Ecoregions—An approach to surface water protection: Journal of the Water Pollution Control Federation, v. 60, no. 4, p. 486–493.

Hummel, P., Kittle, J., Jr., and Gray, M., 2001, WDMUtil Version 2.0—A Tool for Managing Watershed Modeling Time-Series Data User's Manual: U.S. Environmental Protection Agency Report, Contract No. 68-C-98-010, Work Assignment No. 2-05, 165 p.

Interagency Advisory Committee on Water Data, 1982, Guidelines for determining flood-flow frequency: Bulletin 17B of the Hydrology Subcommittee, Office of Water Data Coordination, U.S. Geological Survey, Reston, Va., 183 p., available at http://water.usgs.gov/osw/bulletin17b/bulletin_17B.html.

Intergovernmental Task Force on Monitoring Water Quality, 1995a, The strategy for improving water-quality monitoring in the United States: U.S. Geological Survey Final Report, 25 p.

Intergovernmental Task Force on Monitoring Water Quality, 1995b, The strategy for improving water-quality monitoring in the United States: U.S. Geological Survey Technical Appendixes, 117 p.

Jakeman, A.J., and Hornberger, G.M., 1993, How much complexity is warranted in a rainfall-runoff model: Water Resources Research, v. 29, no. 8, p. 2637–2649.

Jenerette, G.D., Lee, Jay, Waller, D.W., and Carlson, R.E., 2002, Multivariate analysis of the ecoregion delineation for aquatic systems: Environmental Management, v. 29, no. 1, p. 67–75.

Jennings, D., and Jarnagin, S.T., 2002, Changes in anthropogenic impervious surfaces, precipitation and daily streamflow discharge—A historical perspective in a mid-Atlantic sub-watershed: Landscape Ecology, v. 17, no. 5, p. 471–489.

Jennings, M.E., Thomas, W.O., Jr., and Riggs, H.C., 1994, Nationwide summary of U.S. Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993: U.S. Geological Survey Water-Resources Investigations Report 94–4002, 196 p. with 3 1/2 in. computer disk.

Jens, S.W., and McPherson, M.B., 1964, Section 20–Hydrology of urban areas: *in* Chow, V.T. (ed.), Handbook of applied hydrology—A compendium of water-resources technology, New York, McGraw-Hill Book Co., p. 20.1–20.45.

Johnson, David, 1997, The triangular distribution as a proxy for the Beta distribution in risk analysis: The Statistician, v. 46, no. 3, p. 387–398.

Kent, K.M., 1973, A method for estimating volume and rate of runoff in small watersheds: U.S. Department of Agriculture, Soil Conservation Service, SCS-TP-149, 64 p.

Kirby, W.H., 1972, Computer-oriented Wilson-Hilferty transformation that preserves the first three moments and the lower bound of the Pearson Type III distribution: Water Resources Research, v. 8, no. 5, p. 1251–1254.

Knutilla, R.L., and Veenhuis, J.E., 1994, Computer simulation of storm runoff for three watersheds in Albuquerque, New Mexico: U.S. Geological Survey Water-Resources Investigations Report 94–4143, 61 p.

Koutsoyiannis, Demetris, and Xanthopoulos, Themistocle, 1989, On the parametric approach to unit hydrograph identification: Water Resources Management, v. 3, p. 107–128.

Krug, W.R., Gebert, W.A., and Graczyk, D.J., 1989, Preparation of average annual runoff map of the United States, 1951–80: U.S. Geological Survey Open-File Report 87–535, 414 p.

Krug, W.R., Gebert, W.A., Graczyk, D.J., Stevens, D.L., Jr., Rochelle, B.P., and Church, M.R., 1990, Map of mean annual runoff for the Northeastern, Southeastern, and Mid-Atlantic United States, water years 1951–80: U.S. Geological Survey Water-Resources Investigations Report 88–4094, 11 p.

Langbein, W.B. (ed.), 1949, Annual runoff in the United States: U.S. Geological Survey Circular 52, 14 p., 1 pl.

Langbein, W.B., and Schumm, S.A., 1958, Yield of sediment in relation to mean annual precipitation: Transactions of the American Geophysical Union, v. 39, no. 6, p. 1076–1084.

Lettenmaier and Wood, 1993, Chapter 26—Hydrologic forecasting: *in* Maidment, D.R., ed., Handbook of Hydrology: New York, McGraw-Hill, Inc., p. 21.1–21.39.

Lichty, R.W., and Karlinger, M.R., 1995, Spatial trends in Pearson Type III statistical parameters: Journal of Hydraulic Engineering, v. 121, no. 9, p. 672–678.

Lindner-Lunsford, J.B., and Ellis, S.R., 1987, Comparison of conceptually based and regression rainfall-runoff models, Denver metropolitan area, Colorado, and potential applications in urban areas: U.S. Geological Survey Water-Resources Investigations Report 87–4104, 45 p.

Lins, H.F., 1997, Regional streamflow regimes and hydroclimatology of the United States: Water Resources Research, v. 33, no. 7, p. 1655–1667.

Linsley, R.K., Jr., Kohler, M.A., and Paulhus, J.L.H., 1975, Hydrology for Engineers (2d ed.): New York, McGraw-Hill Book Co., 482 p.

Liscum, Fred, 2001, Effects of urban development on stormwater runoff characteristics for the Houston, Texas, metropolitan area: U.S. Geological Survey Water-Resources Investigations Report 01–4071, 35 p.

Liscum, Fred, Brown, D.W., and Kasmarek, M.C., 1997, Summary of surface-water hydrologic data for the Houstin Metropolitan Area, Texas, water years 1964–89: U.S. Geological Survey Open-File Report, 51 p.

Lumb, A.M., Carsel, R.F., and Kittle, J.L., Jr., 1988, Data management for water-quality modeling development and use: Proceedings of the International Symposium on Water Quality Modeling of Agricultural Non-Point Sources, 14 p.

Maestre, Alex, and Pitt, Robert, 2005, The National Stormwater Quality Database, Version 1.1–A Compilation and Analysis of NPDES Stormwater Monitoring Information: Washington, D.C., U.S. Environmental Protection Agency, Office of Water, 447 p.

Marsalek, J.H., 1991, Pollutant loads in urban stormwater—Review of methods for planning-level estimates: Water Resources Bulletin, v. 27, no. 2, p. 283–291.

Marsalek, J.H., and Ng, Y.F., 1989, Evaluation of pollution loadings from urban nonpoint sources—Methodology and applications: Journal of Great Lakes Research, v. 15, no. 3, p. 444–451.

Martin, G.R., Zarriello, P.J., and Shipp, A.A., 2001, Hydrologic and water-quality characterization and modeling of the Chenoweth Run basin, Jefferson County, Kentucky: U.S. Geological Survey Water-Resources Investigations Report 2000–4239, 197 p.

Mathey, S.B., (ed.), 1998, National Water Information System: U.S. Geological Survey Fact Sheet FS–027–98, 2 p.

McCuen, R.H., Johnson, P.A., and Ragan, R.M., 2002, Highway hydrology—hydraulic design series number 2, Second Edition: Washington, D.C., Federal Highway Administration Report FHWA-NHI-02-001 HDS-2, 424 p.

McMahon, Gerard, Gregonis, S.M., Waltman, S.W., Omernik, J.M., Thorson, T.D., Freeouf, J.A., Rorick, A.H., and Keys, J.E., 2001, Developing a spatial framework of common ecological regions for the conterminous United States: Environmental Management, v. 28, no. 3, p. 293–316.

Meade, R.H., and Parker, R.S., 1985, Sediment in rivers of the United States, *in* U.S. Geological Survey National Water Summary 1984—Hydrologic events, selected water-quality trends, and ground-water resources: U.S. Geological Survey Water-Supply Paper 2275, p. 49–60.

Merz, R., Bloschl, G., and Parajka, J., 2006, Spatio-temporal variability of event runoff coefficients: Journal of Hydrology, v. 331, p. 591–604.

Metzker, K.D., Gold, R.L., and Thomas, R.P., 1993, Rainfall and runoff data for the Albuquerque, New Mexico, metropolitan area, 1984–1988: U.S. Geological Survey Open-File Report 92–653, 388 p.

Mosley, P.M., and McKerchar, A.I., 1993, Chapter 8–Streamflow: *in* Maidment, D.R., ed., Handbook of Hydrology: New York, McGraw-Hill, Inc., p. 8.1–8.39.

Mustard, M.H., Driver, N.E., Chyr, John, and Hansen, B.G., 1987, U.S. Geological Survey urban stormwater data base of constituent storm loads; characteristics of rainfall, runoff, and antecedent conditions; and basin characteristics: U.S. Geological Survey Water-Resources Investigations Report 87–4036, 328 p.

Muzik, Ivan, 1992, A physically based method for prediction of runoff frequencies: Environmental Monitoring and Assessment, v. 23, p. 129–135.

Naef, F., 1981, Can we model the rainfall-runoff process today? Hydrological Sciences Bulletin, v. 26, no. 3, p. 281–289.

National Climatic Data Center, 2002, Data documentation for dataset 9767A (DSI-9767A) Pre-1951 cooperative station histories at ftp://ftp.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9767a.pdf.

National Climatic Data Center, 2003a, Data documentation for dataset 3240 (DSI-3240) hourly precipitation data at ftp://ftp.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td3240.pdf.

National Climatic Data Center, 2003b, Data documentation for hourly precipitation data rainfall event statistics at http://www5.ncdc.noaa.gov/cdo/HPDStats_online_docs.pdf.

National Climatic Data Center, 2005, Hourly precipitation data rainfall event statistics at http://cdo.ncdc.noaa.gov/cgi-bin/HPD/HPDStats.pl.

National Climatic Data Center, 2006, Data documentation for dataset 9767B (DSI-9767B) master station history at ftp://ftp.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9767b.pdf.

National Climatic Data Center, 2007, Get/view online climate data at http://www.ncdc.noaa.gov/oa/climate/climatedata.html#hourly.

National Research Council, 2004, Assessing the National Streamflow Information Program: Washington, D.C., National Academies Press, 164 p.

Nix, S.J., 1994, Urban stormwater modeling and simulation: Ann Arbor, Mich., Lewis Publishers, 212 p.

Novotny, Vladimir, 2004, Simplified data based total maximum daily loads, or the world is log-normal: Journal of Environmental Engineering, v. 130, no. 6, p. 674–683.

Ockerman, D.J., 2002, Hydrologic conditions and quality of rainfall and storm runoff in agricultural and rangeland areas in San Patricio County, Texas, 2000–2001: U.S. Geological Survey Open-File Report 02–291, 20 p.

Ogrosky, H.O., and Mockus, Victor, 1964, Section 21–Hydrology of agricultural lands: *in* Chow, V.T. (ed.), Handbook of applied hydrology—A compendium of water-resources technology: New York, McGraw-Hill Book Co., p. 21.1–21.97.

Omernik, J.M., 1995, Ecoregions–A spatial framework for environmental management, chap. 5, *in* Davis, W.S., and Simon, T.P. (eds.), Biological assessment and criteria—Tools for water resource planning and decision making: Boca Raton, Fla., Lewis Publishers, p. 49–66.

Omernik, J.M., 2004, Perspectives on the nature and definition of ecological regions—Environmental Management, accessed on July 28, 2005, available at http://springerlink.metapress.com/openurl.asp?genre=article&id=doi:10.1007/s00267-003-5197-2.

Omernik, J.M., and Bailey, R.G., 1997, Distinguishing between watersheds and ecoregions: Journal of the American Water Resources Association, v. 33, no. 5, p. 935–949.

Omernik, J.M., Chapman, S.S., Lillie, R.A., and Dumke, R.T., 2000, Ecoregions of Wisconsin: Transactions of the Wisconsin Academy of Sciences, Arts, and Letters, v. 88, p. 77–103.

Outlaw, G.S., 1996, Flood-frequency and detention-storage characteristics of Bear Branch Watershed, Murfreesboro, Tennessee: U.S. Geological Survey Water-Resources Investigations Report 96–4005, 24 p.

Palecki, M.A., Angel, J.R., and Hollinger, S.E., 2005, Storm precipitation in the United States—Part I–Meteorological Characteristics: Journal of Applied Meteorology, v. 44, no. 6, p. 933–946.

Parrett, Charles, and Cartier, K.D., 1990, Methods for estimating monthly streamflow characteristics at ungaged sites in western Montana: U.S. Geological Survey Water-Supply Paper 2365, 30 p.

Peters, N.E., Freer, J., and Aulenbach, B.T., 2003, Hydrologic dynamics of the Panola Mountain Research Watershed, Georgia, U.S.A.: Ground Water, v. 41, no. 7, p. 973–988.

Pilgrim, D.H., and Cordery, Ian, 1993, Chapter 9–Flood Runoff: *in* Maidment, D.R., ed., Handbook of Hydrology: New York, McGraw-Hill, Inc., p. 9.1–9.42.

Poff, N.L., 1996, A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors: Freshwater Biology, v. 36, p. 71–91.

Poff, N.L., Bledsoe, B.P., and Cuhaciyan, C.O., 2006, Hydrologic variation with land use across the contiguous United States—Geomorphic and ecological consequences for stream ecosystems: Geomorphology, v. 39, p. 264–285.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., 1992, Numerical recipes in Fortran 77—The art of scientific computing (2d ed.): New York, Cambridge University Press, 992 p.

Rainwater, F.H., 1962, Stream composition of the conterminous United States: U.S. Geological Survey Hydrologic Investigations Atlas HA-61, 3 pls.

Rantz, S.E. (ed.), 1982, Measurement and computation of streamflow: Volume 1 and 2, Computation of Discharge: U.S. Geological Survey Water-Supply Paper 2175, 631 p.

Ries, K.G., III, and Friesz, P.J., 2000, Methods for estimating low-flow statistics for Massachusetts streams: U.S. Geological Survey Water-Resources Investigations Report 2000–4135, 81 p.

Ries, K.G., III, Steeves, P.A., Coles, J.D., Rea, A.H., and Stewart, D.W., 2004, StreamStats—A U.S. Geological Survey Web Application for Stream Information: U.S. Geological Survey Fact Sheet FS–2004–3115, 4 p., available at http://md.water.usgs.gov/publications/fs-2004-3115/html/index.html.

Robertson, D.M., Saad, D.A., and Wieben, A.M., 2001, An alternative regionalization scheme for defining nutrient criteria for rivers and streams: U.S. Geological Survey Water-Resources Investigations Report 2001–4073, 57 p.

Rochelle, B.P., Stevens, D.J., Jr., and Church, M.R., 1989, Uncertainty analysis of runoff estimates from a runoff contour map: Water Resources Bulletin, v. 25, no. 3, p. 491–498.

Rohm, C.M., Omernik, J.M., Woods, A.J., and Stoddard, J.L., 2002, Regional characteristics of nutrient concentrations in streams and their application to nutrient criteria development: Journal of the American Water Resources Association, v. 38, no. 1, p. 213–239.

Rossman, L.A., 1990a, Design stream flows based on harmonic means: Journal of Hydraulic Engineering, v. 116, p. 946–950.

Rossman, L.A., 1990b, DFLOW User's Manual. Risk Reduction Engineering Laboratory, Office of Research and Development: U.S. Environmental Protection Agency, Cincinnati, Ohio, 26 p.

Salas, J.D., 1993, Chapter 19–Analysis and modeling of hydrologic time series: *in* Maidment, D.R., ed., Handbook of Hydrology: New York, McGraw Hill, Inc., p. 19.1–19.72.

Saucier, Richard, 2000, Computer generation of statistical distributions: Army Research Laboratory ARL-TR-2168, 105 p.

Sauer, V.B., Thomas, W.O., Jr., Stricker, V.A., and Wilson, K.V., 1983, Flood characteristics of urban watersheds in the United States: U.S. Geological Survey Water-Supply Paper 2207, 63 p.

Saunders, J.F., III, Murphy, Marylee, Clark, Martyn, and Lewis, W.M., 2004, The influence of climate variation on the estimation of low flows used to protect water quality—A nationwide assessment: Journal of the American Water Resources Association, v. 40, no. 5, p. 1339–1349.

Schaap, B.D., and Lucey, K.J., 1994, Selected nutrients in stormwater runoff from Davenport, Iowa, 1992: U.S. Geological Survey Water-Resources Investigations Report 94–4130, 29 p.

Schalk, G.K., 1994, Quantity and quality of base flow and stormwater runoff in Independence, Missouri, October 1991 to February 1993: U.S. Geological Survey Open-File Report 93–495, 69 p.

Schueler, T.R., 1987, Controlling urban runoff—A practical manual for planning and designing urban BMP's: Washington, D.C., Metropolitan Washington Council of Governments, Department of Environmental Programs, 275 p.

Schwartz, S.S., and Naiman, D.Q., 1999, Bias and variance of planning-level estimates of pollutant loads: Water Resources Research, v. 35, no. 11, p. 3475–3487.

Seaber, P.R., Kapinos, F.P., and Knapp, G.L., 1987, Hydrologic unit maps: U.S. Geological Survey Water-Supply Paper 2294, 63 p.

Selbig, W.R., and Bannerman, R.T., 2007, Evaluation of street sweeping as a stormwater-quality-management tool in three residential basins in Madison, Wisconsin: U.S. Geological Survey Scientific Investigations Report 2007–5156, 103 p.

Selbig, W.R., and Bannerman, R.T., 2008, A comparison of runoff quantity and quality from two small basins undergoing implementation of conventional and low-impact-development (LID) strategies: Cross Plains, Wisconsin, water years 1999–2005: U.S. Geological Survey Scientific Investigations Report 2008–5008, 57 p.

Selbig, W.R., Jopke, P.L., Marshall, D.W., and Sorge, M.J., 2004, Hydrologic, ecologic, and geomorphic responses of Brewery Creek to construction of a residential subdivision, Dane County, Wisconsin, 1999–2002: U.S. Geological Survey Scientific Investigations Report 2004–5156, 33 p.

Sevin, A.F., 1987, Guidance for preparing and processing environmental and section 4(f) documents: Washington, D.C., Department of Transportation, Federal Highway Administration, Technical Advisory T6640.8A, 62 p.

Shamir, E., Imam, B., Morin, E., Gupta, H.V., and Sorooshian, S., 2005, The role of hydrograph indices in parameter estimation of rainfall-runoff models: Hydrological Processes, v. 19, p. 2187–2207.

Shirazi, M.A., Boersma, Larry, Johnson, C.B., and Haggerty, P.K., 2001, Predicting physical and chemical water properties from relationships with watershed soil characteristics: Journal of Environmental Quality, v. 30, p. 112–120.

Shirazi, M.A., Johnson, C.B., Omernik, J.M., White, D.A., Haggerty, P.K., and Griffith, G.E., 2003, Landscape and watershed processes—Quantitative soil descriptions for ecoregions of the United States: Journal of Environmental Quality, v. 32, no. 2, p. 550–561.

Shirmohammadi, Adel, Sheridan, J.M., and Knisel, W.G., 1987, Regional application of an approximate streamflow partitioning method: Water Resources Bulletin, v. 23, no. 1, p. 103–111.

Shuster, W.D., Zhang, Y., Roy, A.H., Daniel, F.B., and Troyer, M., 2008, Characterizing storm hydrograph rise and fall dynamics with stream stage data: Journal of the American Water Resources Association, v. 44, no. 6, p. 1431–1440.

Simon, Andrew, Dickerson Wendy, and Heins, Amanda, 2004, Suspended sediment transport rates at the 1.5-year recurrence interval for ecoregions of the United States—Transport conditions at the bankfull and effective discharge?: Geomorphology, v. 58, p. 243–262.

Singh, V.P., 1977, Sensitivity of some runoff models to errors in rainfall excess: Journal of Hydrology, v. 33, p. 301–318.

Singh, V.P., and Chowdhury, P.K., 1986, Comparing some methods of estimating mean areal rainfall: Water Resources Bulletin, v. 22, no. 2, p. 275–282.

Sivapalan, Murugesu, 2003, Prediction in ungaged basins—A grand challenge for theoretical hydrology: Hydrological Processes, v. 17, p. 3163–3170.

Sloto, R.A., 1988, Effects of urbanization on storm-runoff volume and peak discharge of Valley Creek, eastern Chester County, Pennsylvania: U.S. Geological Survey Water-Resources Investigations Report 87–4196, 32 p.

Sloto, R.A., and Crouse, M.Y., 1996, HYSEP: A computer program for streamflow hydrograph separation and analysis: U.S. Geological Survey Water-Resources Investigations Report 96–4040, 46 p.

Smieszek, T.W., and Granato, G.E., 2000, Geographic information for analysis of highway runoff-quality data on a national or regional scale in the conterminous United States: U.S. Geological Survey Open-File Report 00–432, 15 p.

Smith, R.A., Alexander, R.B., and Schwarz, G.E., 2003a, Natural background concentrations of nutrients in streams and rivers of the conterminous United States: Environmental Science and Technology, v. 37, no. 4, p. 3039–3047.

Stedinger, J.R., Vogel, R.M., and Foufouls-Georgiou, Efi, 1993, Chapter 18–Frequency analysis of extreme events; *in* Maidment, D.R., ed., Handbook of Hydrology: New York, McGraw-Hill, Inc., p. 18.1–18.66.

Steuer, Jeffrey, Selbig, William, Hornewer, N.J., and Prey, Jeffrey, 1997, Sources of contamination in an urban basin in Marquette, Michigan and an analysis of concentrations, loads, and data quality: U.S. Geological Survey Water-Resources Investigations Report 97–4242, 25 p.

Strecker, E.W., Driscoll, E.W., and Palhegyi, E.D., 1989, PC-SYNOP, a rainfall analysis tool in the proceedings of the stormwater and water-quality model user's group meeting, October 3–4, 1988, Denver, Colo., U.S. Environmental Protection Agency EPA/600/9-89/001, p. 161–172.

Strecker, Eric, Mayo, Lynn, Quigley, Marcus, and Howell, Jim, 2001, Guidance manual for monitoring highway runoff water quality, Federal Highway Administration, FHWA-EP-01-021, 206 p.

Stricker, V.A., and Sauer, V.B., 1982, Techniques for estimating flood hydrographs for ungaged urban watersheds: U.S. Geological Survey Open-File Report 82–365, 24 p.

Stumm, Frederick, and Ku, H.F., 1997, Urbanization and recharge in the vicinity of East Meadow Brook, Nassau County, New York; Part 2–Effect of urban runoff on the hydrology of the headwaters of East Meadow Brook, 1989–90: U.S. Geological Survey Water-Resources Investigations Report 97–4063, 24 p.

Syvitski, J.P., Morehead, M.D., Bahr, D.B., and Mulder, Thierry, 2000, Estimating fluvial sediment transport: Water Resources Research, v. 36, no. 9, p. 2747–2760.

Tasker, G.D., and Granato, G.E., 2003, Statistical approaches to interpretation of local, regional, and national highway-runoff and urban-stormwater data, *in* Granato, G.E., Zenone, C., and Cazenas, P.A. (eds.), National highway runoff water-quality data and methodology synthesis, Volume I–Technical issues for monitoring highway runoff and urban stormwater: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, FHWA-EP-03-054, p. 419–479.

Thomas, D.M., and Benson, M.A., 1970, Generalization of streamflow characteristics from drainage-basin characteristics: U.S. Geological Survey Water-Supply Paper 1975, 55 p.

Thomas, M.P., 1966, Effect of glacial geology upon the time distribution of stream flow in eastern and southern Connecticut: U.S. Geological Survey Professional Paper 550–B, p. 209–212.

Trommer, T.J., Loper, J.E., and Hammett, K.M., 1996a, Comparison of five estimating techniques for determining stormwater in west-central Florida: U.S. Geological Survey Open-File Report 96–129, 79 p.

Trommer, T.J., Loper, J.E., and Hammett, K.M., 1996b, Evaluation and modification of five techniques for estimating stormwater runoff for watersheds in west-central Florida: U.S. Geological Survey Water-Resources Investigations Report 96–4158, 37 p.

Urbonas, Ben, and Guo, J.C.Y., 1989, Sizing a capture volume for stormwater quality enhancement: Denver, Colo., Urban Drainage and Flood Control District Flood Hazard News, v. 19, no. 1, variously paged, available at http://www.udfcd.org/downloads/pdf/tech_papers/SIZING%20A%20CAPTURE%20VOLUME%20-1989%20FHN.pdf.

U.S. Environmental Protection Agency, 1976, Documentation for synoptic rainfall data analysis program-SYNOP Appendix E in Areawide assessment procedures manual: U.S. Environmental Protection Agency Report EPA-600/9-76-014, p. E.1–E.26.

U.S. Environmental Protection Agency, 1986, Development of data quality objectives: Washington, D.C., U.S. Environmental Protection Agency Quality Assurance Management Staff, 12 p.

U.S. Environmental Protection Agency, 1992, National pollutant discharge elimination system storm water sampling guidance document: Washington, D.C., U.S. Environmental Protection Report Agency, Office of Water, Report EPA 883-B-92-001, 177 p.

U.S. Environmental Protection Agency, 1994, Guidance for the data quality objectives process, EPA QA/G-4, Final Report: Washington, D.C., U.S. Environmental Protection Agency, EPA/600/R-96/055, 72 p.

U.S. Environmental Protection Agency, 1996, Guidance for data quality assessment, Practical methods for data analysis, EPA QA/G-9, Final Report: Washington, D.C., U.S. Environmental Protection Agency, EPA/600/R-96/084, 164 p.

U.S. Environmental Protection Agency, 1997, Guiding principles for Monte Carlo analysis: U.S. Environmental Protection Agency Report EPA/630/R-97/001, 39 p.

U.S. Environmental Protection Agency, 2001a, Better assessment science integrating point and nonpoint sources (BASINS) version 3.0–User's Manual: U.S. Environmental Protection Agency Report 823/B-01-001, 348 p.

U.S. Environmental Protection Agency, 2001b, International stormwater best management practices (BMP) database: U.S. Environmental Protection Agency, Office of Water, electronic product http://www.bmpdatabase.org.

U.S. Environmental Protection Agency, 2002, Considerations in the design of treatment best management practices (BMPs) to improve water quality: U.S. Environmental Protection Agency Report EPA/600/R-03/103, 185 p.

U.S. Environmental Protection Agency, 2003, Level III ecoregions of the continental United States: Washington, D.C., U.S. Environmental Protection Agency National Health and Environmental Effects Research Laboratory, 1 pl., available at http://www.epa.gov/wed/pages/ecoregions/level_iii.htm).

U.S. Environmental Protection Agency, 2004, Flow 101–Basic Information provides information on the development and uses of DFLOW and the differences between DFLOW 3.0 and earlier versions, accessed on November 24, 2004, at http://epa.gov/waterscience/models/dflow/flow101.htm.

U.S. Environmental Protection Agency, 2007, Better assessment science integrating point and nonpoint sources (BASINS) at http://www.epa.gov/waterscience/basins/.

U.S. Geological Survey, 1980, National handbook of recommended methods for water-data acquisition: Reston, Va., U.S. Geological Survey, variously paged.

U.S. Geological Survey, 1982, A U.S. Geological Survey data standard—Codes for the identification of hydrologic units in the United States and the Caribbean outlying areas: U.S. Geological Survey Circular 878–A, 115 p.

U.S. Geological Survey, 2002, NWISWeb: New site for the Nation's water data: U.S. Geological Survey Fact Sheet FS–128–02, 2 p.

U.S. Geological Survey, 2007, Welcome to StreamStats, accessed July 16, 2007, at http://water.usgs.gov/osw/streamstats/index.html.

U.S. Geological Survey, 2009, NWISWeb data for the Nation, accessed July 24, 2009, at http://waterdata.usgs.gov.nwis/.

Vogel, R.M., and Stedinger, J.R., 1985, Minimum variance streamflow record augmentation procedures: Water Resources Research, v. 21, no. 5, p. 715–723.

Vogel, R.M., Tsai, Y., and Limbrunner, J.F., 1998, The regional persistence and variability of annual streamflow in the United States: Water Resources Research, v. 34, no. 12, p. 3445–3459.

Vogel, R.M., Wilson, Ian, and Daly, Chris, 1999, Regional regression models of annual streamflow for the United States: Journal of Irrigation and Drainage Engineering, v. 125, no. 3, p. 148–157.

Wahl, K.L., Thomas, W.O., Jr., and Hirsch, R.M., 1995, Stream-gaging program of the U.S. Geological Survey: U.S. Geological Survey Circular 1123, 22 p., at http://pubs.er.usgs.gov/usgspubs/cir/cir1123/.

Walker, J.F., Graczyk, D.J., Corsi, S.R., Wierl, J.A., and Owens, D.W., 2001, Evaluation of nonpoint-source contamination, Wisconsin: Water year 1999: U.S. Geological Survey Open-File Report 2001–105, 37 p.

Wanielista, M.P., 1990, Hydrology and water-quality control: New York, John Wiley and Sons, 565 p.

Wanielista, M.P., and Yousef, Y.A., 1993, Stormwater management: New York, John Wiley and Sons, 579 p.

Ward, Andy, Bridges, Tom, and Wilson, Bruce, 1981, Simple procedure for developing a design-storm hydrograph: Water Resources Bulletin, v. 17, no. 2, p. 209–214.

Warn, A.E., and Brew, J.S., 1980, Mass balance: Water Research, v. 14, p. 1427–1434.

Waschbusch, R.J., 1996, Stormwater-runoff data, Madison, Wisconsin, 1993–94: U.S. Geological Survey Open-File Report 95–733, 33 p.

Waschbusch, R.J., 1999, Evaluation of the effectiveness of an urban stormwater treatment unit in Madison, Wisconsin, 1996–97: U.S. Geological Survey Water-Resources Investigations Report 99–4195, 49 p.

Waschbusch, R.J., Selbig, W.R., and Bannerman, R.T., 1999, Sources of phosphorus in stormwater and street dirt from two urban residential basins in Madison, Wisconsin, 1994–95: U.S. Geological Survey Water-Resources Investigations Report 99–4021, 47 p.

Weiskel, P.E., Vogel, R.M., Steeves, P.A., Zarriello, P.J., DeSimone, L.A., and Ries, K.G., 2007, Water use regimes—Characterizing direct human interaction with hydrologic systems: Water Resources Research, v. 43, p. W04402–W04413, doi:10.1029/2006WR005062.

Wiles, T.J., and Sharp, J.M., Jr., 2008, The secondary permeability of impervious cover: Environmental and Engineering Geoscience, v. 14, no. 4, p. 251–265.

Winter, T.C., Harvey, J.W., Franke, O.L., and Alley, W.M., 1998, Ground water and surface water—A single resource: U.S. Geological Survey Circular 1139, 79 p.

Wolock, D.M., Winter, T.C., and McMahon, Gerard, 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, p. s71–s88.

Yen, B.C., and Chow, V.T., 1980, Design hyetographs for small drainage structures: Journal of the Hydraulics Division, American Society of Civil Engineers, v. 106, no. HY6, p. 1055–1076.

Yen, B.C., and Chow, V.T., 1983, Local design rainfall, Volume II–Methodology and analysis: Federal Highway Administration Report FHWA/RD-82/064, 205 p.

Yorke, T.H., and Herb, W.J., 1978, Effects of urbanization on streamflow and sediment transport in the Rock Creek and Anacostia River Basins, Montgomery County, Maryland, 1962–74: U.S. Geological Survey Professional Paper 1003, 71 p.

Zarriello, P.J., 1998, Comparison of nine uncalibrated runoff models to observed flows in two small urban watersheds: *in* Proceedings of the first Federal Interagency Hydrologic Modeling Conference, April 19–23, 1998, Las Vegas, Nev., v. 2 of 2, p. 7.163–7.170.

Zarriello, P.J., and Barlow, L.K., 2002, Measured and simulated runoff to the lower Charles River, Massachusetts, October 1999–September 2000: U.S. Geological Survey Water-Resources Investigations Report 02–4129, 89 p.

Granato—**Methods for Development of Planning-Level Estimates of Stormflow at Unmonitored Stream Sites in the Conterminous United States**—FHWA-HEP-09-005