USGS
science for a changing world

Estimating Basin Lagtime and Hydrograph-Timing Indexes Used to Characterize Stormflows for Runoff-Quality Analysis

Scientific Investigations Report 2012–5110

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

U.S. Department of Transportation
Federal Highway Administration

This page has been left blank intentionally.

Estimating Basin Lagtime and Hydrograph-Timing Indexes Used to Characterize Stormflows for Runoff-Quality Analysis

By Gregory E. Granato

Scientific Investigations Report 2012–5110

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

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2012

Contents

Figures

Tables

Conversion Factors and Abbreviations

Multiply	By	To obtain
Length		
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
acre	4,047	square meter (m^2)
acre	0.4047	hectare (ha)
square foot (ft^2)	0.09290	square meter (m^2)
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
Slope		
foot per mile (ft/mi)	0.18939	meter per kilometer (m/km)
Volume		
cubic foot (ft^3)	28.32	liter (L)
cubic foot (ft^3)	0.02832	cubic meter (m^3)
acre-foot (acre-ft)	0.001233	cubic hectometer (hm^3)
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]
Mass		
pound, avoirdupois (lb)	0.4536	kilogram (kg)

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

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

Acronyms and Abbreviations

ASEE	average standard error of the estimate
ASEP	average standard error of the prediction
BDF	basin development factor (USGS Streamstats variable name)
BLF	basin lag factor (variable name) which is the basin length divided by the square root of the slope
BMP	best management practice
CN	curve number
DCIA	directly connected impervious area
DQOs	data-quality objectives
EIA	effective impervious area
FHWA	Federal Highway Administration
GAGES	geospatial attributes of gages for evaluating streamflow
GIS	geographic information system
HIS	historical instantaneous streamflow
IDA	instantaneous data archive
IMPERV	total imperviousness, the USGS Streamstats variable name for TIA
LID	low impact development
MAD	median absolute deviation
MPV	most probable value
MS4	municipal separate storm sewer system
NHD	National Hydrography Dataset
NLCD	National Land Cover Dataset
NPDES	National Pollutant Discharge Elimination System
NWISWeb	National Water Information System Web
OLS	ordinary least squares
PDA	percent developed area
PRESS	prediction error sum of squares
RDB	relational database
RMSE	root mean square error
SCS	Soil Conservation Service
SELDM	Stochastic Empirical Loading and Dilution Model
TIA	total impervious area, the USGS Streamstats variable name for TIA is *IMPERV*
USGS	U.S. Geological Survey

Estimating Basin Lagtime and Hydrograph-Timing Indexes Used to Characterize Stormflows for Runoff-Quality Analysis

By Gregory E. Granato

Abstract

A nationwide study to better define triangular-hydrograph statistics for use with runoff-quality and flood-flow studies was done by the U.S. Geological Survey (USGS) in cooperation with the Federal Highway Administration. Although the triangular hydrograph is a simple linear approximation, the cumulative distribution of stormflow with a triangular hydrograph is a curvilinear S-curve that closely approximates the cumulative distribution of stormflows from measured data. The temporal distribution of flow within a runoff event can be estimated using the basin lagtime, (which is the time from the centroid of rainfall excess to the centroid of the corresponding runoff hydrograph) and the hydrograph recession ratio (which is the ratio of the duration of the falling limb to the rising limb of the hydrograph). This report documents results of the study, methods used to estimate the variables, and electronic files that facilitate calculation of variables.

Ten viable multiple-linear regression equations were developed to estimate basin lagtimes from readily determined drainage basin properties using data published in 37 stormflow studies. Regression equations using the basin lag factor (*BLF*, which is a variable calculated as the main-channel length, in miles, divided by the square root of the main-channel slope in feet per mile) and two variables describing development in the drainage basin were selected as the best candidates, because each equation explains about 70 percent of the variability in the data. The variables describing development are the USGS basin development factor (*BDF*, which is a function of the amount of channel modifications, storm sewers, and curb-and-gutter streets in a basin) and the total impervious area variable *(IMPERV)* in the basin. Two datasets were used to develop regression equations. The primary dataset included data from 493 sites that have values for the *BLF*, *BDF*, and *IMPERV* variables. This dataset was used to develop the best-fit regression equation using the *BLF* and *BDF* variables. The secondary dataset included data from 896 sites that have values for the *BLF* and *IMPERV* variables. This dataset was used to develop the best-fit regression equation using the *BLF* and *IMPERV* variables.

Analysis of hydrograph recession ratios and basin characteristics for 41 sites indicated that recession ratios are random variables. Thus, recession ratios cannot be estimated quantitatively using multiple linear regression equations developed using the data available for these sites. The minimums of recession ratios for different streamgages are well characterized by a value of one. The most probable values and maximum values of recession ratios for different streamgages are, however, more variable than the minimums. The most probable values of recession ratios for the 41 streamgages analyzed ranged from 1.0 to 3.52 and had a median of 1.85. The maximum values ranged from 2.66 to 11.3 and had a median of 4.36.

Introduction

For runoff-quality modeling, information about the timing of runoff from a site of interest and from the upstream basin of the receiving stream at the location of the stormflow outfall is necessary to estimate the quantity of the upstream flow that occurs concurrently with runoff from the site of interest. The focus of planning-level analyses of runoff quality has traditionally been on event-mean concentrations and total storm loads for the entire runoff event rather than on processes that occur during events. Differences in the locations, sizes, and drainage characteristics of the site of interest and the upstream basin, however, may cause differences in the timings and durations of runoff from each area. If the drainage area of the site of interest is small and the runoff drains directly to the stream, the duration of appreciable runoff from the site of interest may be approximated by the duration of the precipitation event. If the drainage area of the upstream basin is relatively large and more pervious than the drainage area of the site of interest, the duration of appreciable runoff from the basin may continue for hours or days longer than runoff from the site of interest. In this case, only a small proportion of the upstream runoff may be available to dilute runoff constituents from the site of interest in the receiving waters. However, if a

structural best management practice (BMP) is used at the site of interest to attenuate and extend the runoff hydrograph, then much more of the upstream runoff may be available to dilute runoff constituents in the receiving waters. This concept is demonstrated schematically in figure 1. In this hypothetical example, the triangular runoff hydrograph for the upstream basin is superimposed on a rectangular representation of the prestorm base flow (fig. 1A). The durations of runoff hydrographs from a highway or small urban area with and without BMP modification are labeled "Duration 1" and "Duration 2," respectively. As indicated in the figure, a small increase in the duration of runoff from the site of interest 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. 1B).

Granato (2010) demonstrated that triangular runoff hydrographs commonly are used to model intra-event stormflows in hydraulic and water-quality models and are adequate for producing planning-level estimates for dilution analyses. The triangular hydrograph commonly is used in hydrology. For example, the Soil Conservation Service (SCS) triangular hydrograph commonly is selected because it 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). The triangular distribution is commonly used as a synthetic unit hydrograph to estimate runoff flows from within-storm precipitation-excess increments (Granato, 2010). For planning-level analyses, however, the entire precipitation event may be characterized by a single increment. The triangular hydrograph has been shown to provide a good representation of the cumulative amount of stormflow that occurs during a storm. 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. Similarly, Guo and Adams (1998) 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 hydrograph can be fully parameterized with the total runoff volume, the start of runoff (T_0), the end of runoff (T_e), and the time to peak (T_p) (fig. 2). With this information, the cumulative volume of runoff within a given time interval is simple to compute with a triangular hydrograph. Although the triangular hydrograph is a simple linear approximation, the cumulative distribution of stormflow with a triangular hydrograph is a curvilinear S-curve that will closely approximate the cumulative distribution of stormflows
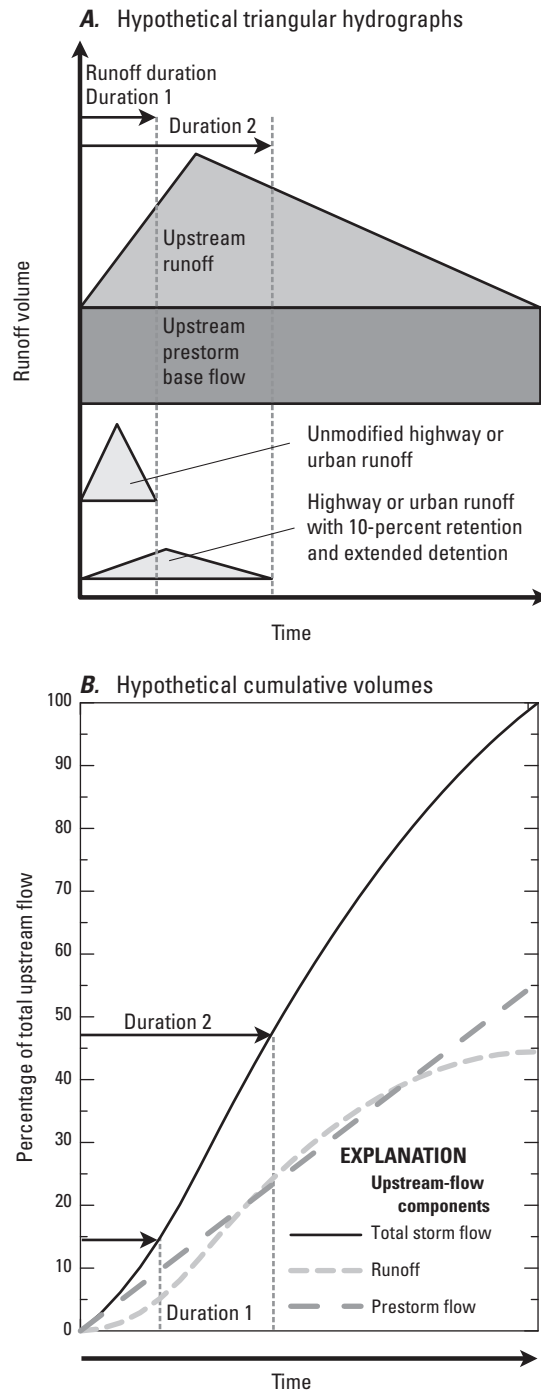
Figure 1.   Simplified schematic diagrams showing *A*, hypothetical triangular hydrographs and *B*, the hypothetical cumulative upstream storm volume that would occur concurrently with the duration of unmodified runoff from a highway or urban site of interest and with the duration of runoff from an extended detention structure. This diagram shows the hypothetical runoff event with two upstream flow components (runoff and prestorm base flow), an unmodified runoff hydrograph, and a runoff hydrograph from a site of interest with retention and detention.

from measured data. The proportion of total runoff at time T_i from the beginning of the storm for a triangular hydrograph is expressed as

$$R_c = \frac{(T_i - T_o)^2}{(T_e - T_o) \times (T_p - T_o)} \text{ if } T_o <= T_i <= T_p \text{ and} \qquad (1)$$

$$R_c = 1 - \frac{(T_e - T_i)^2}{(T_e - T_o) \times (T_e - T_p)} \text{ if } T_p <= T_i <= T_e \qquad (2)$$

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_0 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 (T_e) is equal to the duration of the runoff hydrograph (T_b) (fig. 2).

The time to peak is commonly calculated as one-half the precipitation duration ($D/2$) plus a basin lagtime (*LAGTIME*, in hours) that depends on basin characteristics. Although there are many definitions of the basin lagtime in the literature (Rao and Delleur, 1974; Linsley and others, 1975; Chow and

others, 1988; Fang and others, 2005), the Stochastic Empirical Loading and Dilution Model (SELDM) uses the basin lagtime defined as the time from the center of mass (centroid) of rainfall excess to the centroid of the corresponding runoff hydrograph (Granato, 2010). This definition was selected because U.S. Geological Survey (USGS) runoff studies commonly use this definition. The relation between the time to peak T_p (in hours) and the time to the centroid T_c (in hours) of the runoff hydrograph is a function of the unitless ratio of the duration of the falling to the rising limb of the hydrograph (R_f). For a triangular hydrograph this may be calculated as:

$$T_p = 3 \times \left(\frac{\dfrac{D}{2} + LAGTIME}{R_f + 2} \right) \qquad (3)$$

Thus, *LAGTIME* is an important factor for quantifying the time response of runoff in a given basin. This factor is used in hydraulic design and analysis and in many precipitation runoff models (Eagleson, 1962; Leopold, 1968; Rao and Delleur, 1974; Linsley and others, 1975; Laenen, 1980; Chow and others, 1988; Sutherland, 1988; Fang and others, 2005; Ries, 2007; Simas and Hawkins, 2011). A number of formulas have been developed for calculating the basin lagtime from basin characteristics (Carter, 1961; Chow, 1964; Kent, 1973; Laenen, 1980; Sauer and others, 1983; Franklin, 1984; Chow

D Duration of rainfall-excess increment

BL Basin *lagtime*

T_p Time to peak (rising-limb time)

T_f Recession (falling-limb time)

T_b Hydrograph base time

T_0 Begin time of the runoff hydrograph

T_e End time of the runoff hydrograph
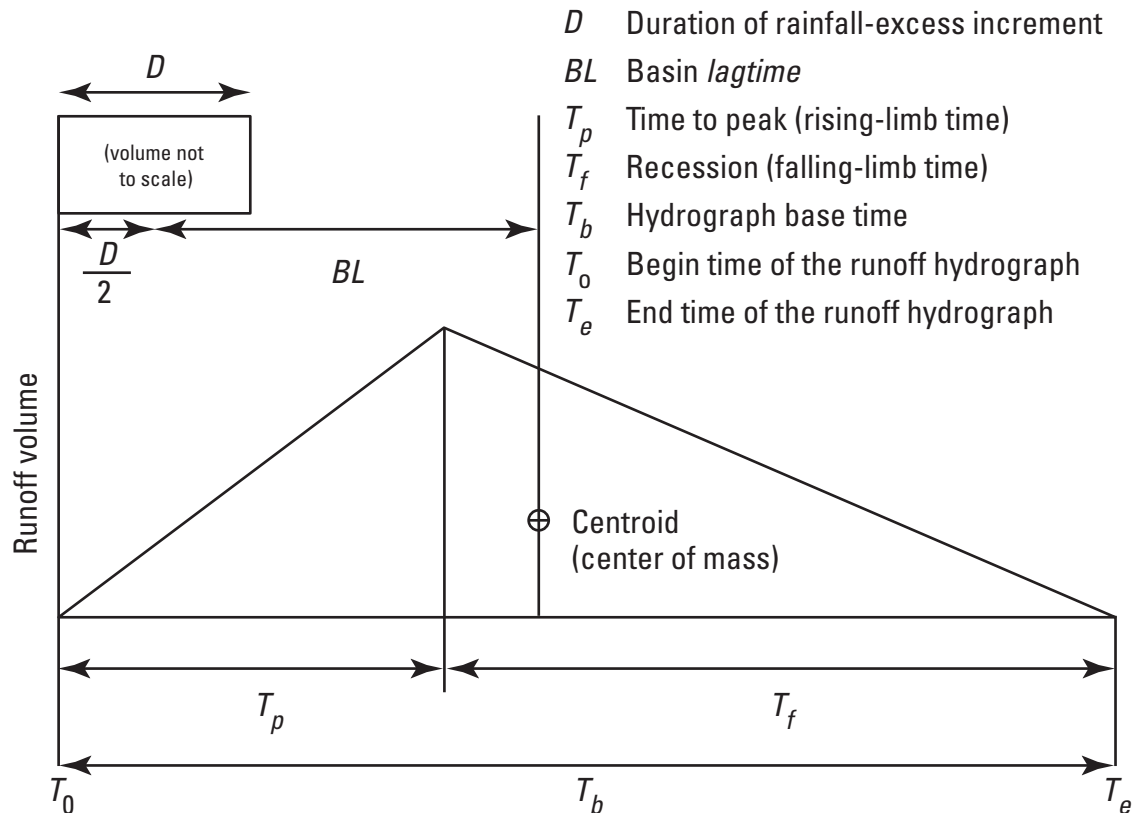
Figure 2.   Schematic diagram showing time factors for a triangular storm-event hydrograph (modified from Kent, 1973).

and others, 1988; Sutherland, 1988; Wanielista, 1990; Muzik, 1992; Pilgrim and Cordery, 1993; Wanielista and Yousef, 1993; Ries, 2007). 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, SCS curve number (CN), or a channel roughness factor. Some equations account for storm characteristics (usually rainfall intensity), but basin lagtime 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 with limited ranges in basin characteristics. For example, the Kirpich equation (1940) is based on data from only seven rural basins (Pilgrim and Cordery, 1993; Chow and others, 1988). Carter (1961) used 24 suburban basins with total impervious area (TIA) values ranging from 0 to 12 percent. Chow (1962) used data from 20 basins in Illinois, but did not use land-use factors. Schulz and Lopez (1974) used data from 9 urban basins. McCuen and others (1984) used data from 48 urban basins with drainage areas less than 6.25 square miles (mi^2). Watt and Chow (1985) used data from 44 basins, but did not include any land-use information. Sauer and others (1983) had the most comprehensive study to date. They used data from 170 basins throughout the United States with drainage areas ranging from 0.2 to 100 mi^2, basin lengths from 0.47 to 88.1 miles (mi), main channel slope from 3 to 500 feet per mile (ft/mi), total impervious area (*IMPERV*, USGS Streamstats variable for TIA) values from 3 to 50 percent of the basin area, and basin development factors from 0 to 12 to develop basin lagtime and flood-flow equations.

Although the basin lagtime equations provided by Sauer and others (1983) were recognized as being the most comprehensive available, there were several concerns about the application and use of these equations. The first concern was about extending use of the USGS equations beyond the limits of the input dataset (Andrew H. McDaniel, North Carolina Department of Transportation, written commun., 2011). Specifically, extending its use to very small sites (less than 0.2 mi^2 or about 128 acres) and to highly impervious sites greater than 50 percent *IMPERV*. The second concern was about properly specifying a basin development factor for use in the equation. The third was about the feasibility (or lack thereof) for developing an automated method for calculating a basin development factor (Peter A. Steeves, written commun. U.S. Geological Survey, October, 2010).

The basin lagtime in equation 3 commonly is defined as a characteristic of the basin rather than a characteristic of individual storms, so this variable is fixed in the SELDM water-quality analyses. However, SELDM uses Monte Carlo methods to generate a random population of each of the other hydrograph properties in equation 3, so T_p is calculated as a random variable, which is consistent with observed hydrographs. Precipitation event durations (*D*) are generated

using synoptic precipitation statistics (selected or input on the synoptic storm event precipitation statistics form) with a two-parameter exponential distribution, and the upstream hydrograph recession values (R_f) are generated using user-defined values (Granato, 2010). Thus, estimates of R_f are needed to calculate T_p (equation 3), which is used with R_f to calculate T_e as:

$$T_e = T_p + T_p \times R_f \qquad (4)$$

If T_o is set to the beginning of the runoff event, then the proportion of total upstream stormflow that occurs concurrent to discharge from the site of interest, or concurrent to discharge from a BMP treating runoff from the site of interest, is calculated using either equation 1 or equation 2. The total volume for dilution is the sum of stormflow and prestorm base flow that occurs during the period of concurrent flow (Granato, 2010).

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 square miles) raised to the power 0.2, commonly is 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 rational method is based on the assumption that the runoff hydrograph is an isosceles triangle with equal rising- and falling-limb durations (Linsley and others, 1975; Pilgrim and Cordery, 1993). Thus, the rational method hydrograph-recession ratio is equal to one. 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) provides hydrograph-recession ratios ranging from 1.25 to 12 and qualitatively links his recession ratio values to basin slope and land use. However, the underlying interpretation, data, and basin characteristics used for derivation of these ratios are not published (Wanielista, 1990; Wanielista and Yousef, 1993). Other studies provide the information and data that are needed to estimate recession ratios but do not provide actual recession-ratio values (Craig and Rankl, 1978; Stricker and Sauer, 1982; Franklin, 1984; Inmann, 1986; Gamble, 1989; Neely, 1989; Becker, 1986, 1990; Bohman, 1990; Bohman, 1992; Holnbeck and Parrett, 1996; Mason and Bales, 1996; Liscum and others, 1997; Liscum, 2001; McCuen and others, 2002; Weaver, 2003; Shamir and others, 2005; Shuster and others, 2008).

Purpose and Scope

This report documents data and methods that are used to estimate basin lagtimes and hydrograph-recession ratios. 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. This study was done by the USGS in cooperation with the Federal Highway Administration (FHWA) to better define triangular-hydrograph statistics for use with runoff-quality and flood-flow studies. The statistics developed in this analysis are intended, primarily, for use with SELDM. However, regression estimates of the basin lagtime may be used to develop planning-level estimates of the effect of development on stormflows from a basin and for evaluating the potential effectiveness of low-impact-development (LID) practices to mitigate such stormflows. These methods meet data-quality objectives (DQOs) for developing planning-level water-quality estimates at unmonitored sites. The methods and statistics that are described in this report should be useful for other stormwater analyses. Detailed information about the sites (including location, storm events, and sources of data), geographic information system (GIS) files, computer programs, and regression results are documented in the digital media accompanying this report.

Estimating Basin Lagtimes

The regression equations developed in this report can be used to estimate the basin lagtime, which is defined as the time from the center-of-mass of rainfall excess to the center-of-mass of the corresponding runoff. There are many definitions of the basin lagtime in the literature (Rao and Delleur, 1974; Linsley and others, 1975; Chow and others, 1988; Fang and others, 2005). For this study, the center-of-mass definition was selected because the data used in the regression analysis were primarily derived from USGS runoff studies with basin lagtimes that were calculated using the center-of-mass definition. Information and data from 37 studies including 21 USGS studies (which accounted for about 95 percent of the basin lag values) were used to formulate the regression equations (table 1). The data in the report by Sauer and others (1983) includes data for 57 metropolitan areas and was compiled from 36 earlier USGS studies

To be included in the dataset for regression analysis, sites were selected from reports with documented values of the basin lagtimes, drainage areas, basin lengths and channel slopes (or the associated basin lag factor). Each selected report also had documented values of the *IMPERV* or the basin development factor *(BDF)*. Some data were obtained from ancillary sources such as related reports, written communication with the author, or written communication with the surface-water specialist in the local USGS Water

Science Center. Once the data were assembled from the published documents and ancillary sources, two datasets were selected for analysis. The primary dataset includes all sites with a basin lagtime, drainage area, basin length, channel slope, and both *IMPERV* and *BDF* values; 22 of the 37 studies (table 1) had one or more sites that were used in the primary dataset. This primary dataset included 495 sites from different areas of the United States, but two sites (database site-identification numbers 741 and 742) were identified as high-leverage far outliers and were not used for regression analysis (fig. 3). The secondary dataset, assembled using data from all 37 studies (table 1), includes 896 sites (fig. 3) with a basin lagtime, drainage area, basin length, channel slope, and *IMPERV*. The basin characteristics, basin lagtime values, and source citations for all the sites in the primary and secondary datasets are recorded in the database "Compilation.mdb" on the digital media accompanying this report.

Definition of Selected Basin Characteristics for Estimating Basin Lagtimes

Physiographic, land-cover, and land-use factors are considered as the primary basin variables controlling factors that define the characteristics of storm hydrographs for a given basin (Eagleson, 1962; Linsley and others, 1975; Spencer and Alexander, 1978; Laenen, 1980; Subramanya, 1984; Chow and others, 1988; Wanielista, 1990; Wanielista and Yousef, 1993). Commonly used physiographic variables include drainage area, basin shape, basin length, basin slope (commonly, the main channel slope), drainage density, channel length, and channel slope. Commonly used land-cover variables include soil type, vegetative cover, and percent of surface storage (lakes, ponds, and wetlands). Commonly used land-use variables include the percent developed area, *IMPERV*, directly connected (or effective) impervious area, and percent sewered area (percentage of an area drained by storm sewers). It is also recognized that climatic factors such as temperature, potential evapotranspiration, rainfall intensity, rainfall duration, and (for large basins) primary direction of storm movement can affect average characteristics of storm hydrographs for a given basin, as well as storm-to-storm characteristics. However, most studies focus only on basin properties for estimating average-storm hydrograph characteristics because of the complexities involved in quantifying characteristic climatic factors.

Seven basin characteristics were selected as possible explanatory variables in the regression analysis for estimating the basin lagtime based on the availability of these variables in the existing hydrologic studies used to compile the basin lagtime dataset. These basin characteristics were considered sufficient for this study because they have consistently been shown to be important hydrologic variables for describing the basin lagtime (Benson, 1962; Linsley and others, 1975; Harley, 1978; Sauer and others, 1983; Masch, 1984; Chow and others, 1988; Sutherland, 1988; Wanielista, 1990; Wanielista

Table 1. Studies used to compile data for developing basin lagtime equations.

[Report citations in **bold** were used in the primary dataset. mi², square miles; USGS, U.S. Geological Survey; *IMPERV*, total impervious area (percent); *BDF*, basin development factor (unitless); *STORAGE*, storage (percent)]

Study number	Report citation	Drainage area (mi²)		Number of sites with the variable			USGS streamgages	USGS study
		Minimum	Maximum	*IMPERV*	*BDF*	*STORAGE*		
1	**Anderson, 1970**	0.00034	570	71	5	22	57	Y
2	Bailey and others, 1989	1.06	5.98	13	0	5	13	Y
3	**Becker, 1990**	0.14	38.9	61	61	25	61	Y
4	Beighley and others, 2009	0.0232	3.14	9	0	0	0	N
5	**Bhat and others, 2007**	0.293	9.77	5	5	0	0	N
6	**Bohman, 1992**	0.18	9.05	30	30	0	30	Y
7	Burns and others, 2005	0.174	0.216	2	0	0	2	N
8	**Cristina and Sansalone, 2003**	0.000116	0.000116	1	1	1	0	N
9	**Espey and others, 1969**	0.00033	72.8	16	6	0	0	N
10	**Franklin and Losey, 1984**	0.21	15.9	15	15	15	15	Y
11	Gambel, 1989	0.043	19.4	32	0	0	32	Y
12	Holnbeck and Parrett, 1996	58.7	1477	12	12	3	12	Y
13	Hood and others, 2007	0.00656	0.0212	3	0	0	0	N
14	Hubbard, 1992	0.03	21.4	7	0	7	7	Y
15	Inman, 2000	0.04	19.1	69	0	0	69	Y
16	**Laenen, 1983**	0.27	12.6	18	18	18	18	Y
17	**Leopold, 1991**	0.25	81.7	28	11	7	10	Y
18	**Liscum, 2001**	0.5	95.1	26	26	0	26	Y
19	Martens, 1968	6.98	41	6	0	0	6	Y
20	**Martin and others, 1997**	1.66	64	13	13	13	13	Y
21	**Mason and Bales, 1996**	0.04	14.8	50	29	2	50	Y
22	**McEnroe and Zhao, 1999**	0.83	10.3	19	1	17	19	N
23	McEnroe and Zhao, 2001	0.28	28.3	14	0	0	4	N
24	**Meierdiercks and others, 2010**	0.494	0.741	3	3	0	0	N
25	**Neely, 1984**	0.05	19.4	27	1	0	27	Y
26	**Neely, 1989**	14.2	535	2	2	0	2	Y
27	**Rao and Delleur, 1974**	0.0455	19.3	13	4	5	7	N
28	**Robbins, 1984**	0.21	24.3	22	22	0	22	Y
29	**Robbins and Pope, 1996**	0.06	2.85	17	17	0	17	Y
30	**Sauer and others, 1983**	0.12	630	208	208	208	208	Y
31	**Sherwood, 1994**	0.026	4.09	1	30	0	30	Y
32	Straub and others, 2000	0.06	37	9	0	4	9	Y
33	Veenhuis and Gannett,1986	2.31	51.3	14	0	0	14	Y
34	**Viessman, 1968**	0.000617	0.00149	4	4	4	0	N
35	Weaver, 2003	0.123	92.4	24	0	1	24	Y
36	Wibben, 1976	0.17	64	29	0	29	29	Y
37	Williams, 1980	1.62	28.2	3	0	1	3	N

Base from U.S. Geological Survey National Map
Ecoregions from U.S. Environmental Protection Agency
Geographic projection
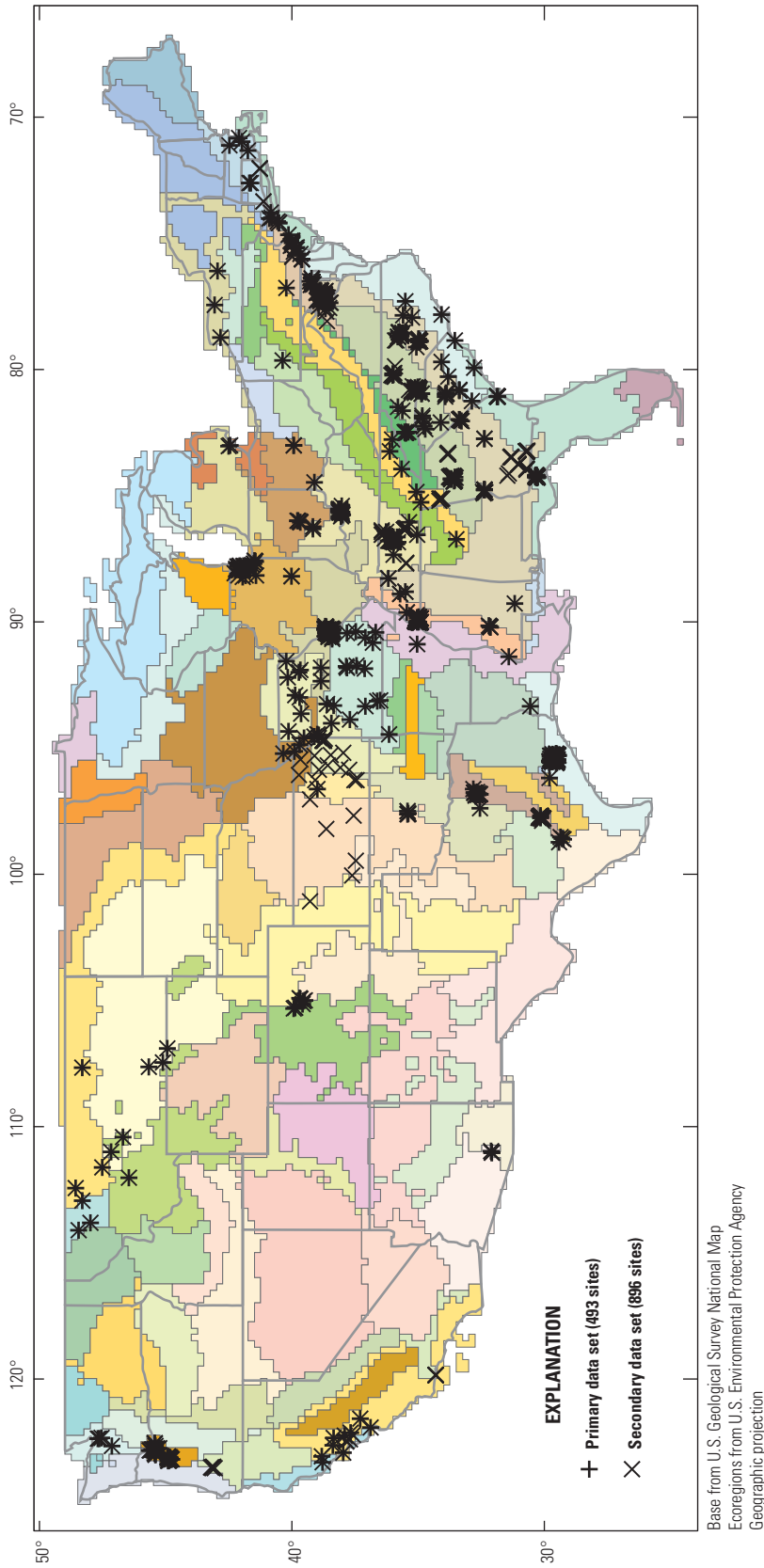
Figure 3. Map showing spatial distribution of 493 monitoring sites in the primary basin lagtime dataset and 896 sites in the secondary basin lagtime dataset with respect to U.S. Environmental Protection Agency (2003) Level III ecoregions (colored polygons) that have been discretized to a 15-minute grid in the conterminous United States. Many stations are clustered within local study areas. Ecoregions are identified on the plate "useco.pdf" in the digital media accompanying this report.

and Yousef, 1993; McCuen and others, 2002; Fang and others, 2005). Four primary physiographic variables, drainage area, basin length, main channel slope, and percent storage were compiled from available basin lagtime studies for use in this analysis. One secondary physiographic variable, the basin lag factor, was calculated from the compiled basin-length and channel-slope data and used in this analysis. Two anthropogenic basin properties, the *IMPERV* and the *BDF* also were compiled and used to develop regression equations.

Physiographic Basin Characteristics

Basin Drainage Area (*DRNAREA*): The basin drainage area is defined as the area of a river basin, measured in a horizontal plane that is enclosed by a topographic divide, such that direct surface runoff from precipitation normally would drain by gravity into the river basin (U.S. Geological Survey, 1977). To determine the *DRNAREA* manually, mark the point of interest, mark the high points on the topographic map around the stream and its tributaries, and draw lines that follow the ridges to connect the high points all around the basin. The curvilinear segments of the resulting basin-delineation polygon between the high points should cross contour lines perpendicularly at the point where the bulge follows the ridge toward the next point. Runoff flows perpendicular to the contours so the curvilinear segments between high points should split bulges in the contours that point downhill. Any arrow drawn perpendicular to the basin-delineation segments into the resulting polygon should represent a flow path toward the stream or its tributaries. Amman and Stone (1991) provide a step-by-step example with illustrations. Once the watershed divide has been delineated, the area of the polygon can be determined using a grid, a planimeter, a digitizer, or GIS software. The contributing drainage area is calculated by subtracting areas within the watershed that do not drain to the main stem or a tributary. Noncontributing areas may include topographic depressions and areas where runoff is diverted across a topographic divide. The USGS StreamStats Web application can be used for the States for which it is available to quickly and easily delineate basins and determine the drainage area at any point on a stream defined in the National Hydrography dataset (Ries, 2007; Ries and others, 2008; U.S. Geological Survey, 2011).

Drainage areas in urban areas and in very small, highly impervious sites can be difficult to determine because of substantial effects of microtopography and drainage diversions on the amount of runoff that may flow to one outlet or another (Strecker and others, 2001; Church and others, 2003; Lee and Heaney, 2003; Richards and Brenner, 2004; Smith and Granato, 2010; Liu and others, 2011). Strecker and others (2001) noted that it is difficult to accurately delineate a small low-slope catchment, because small surface features have an inordinate effect on drainage patterns in these catchments. In such catchments the drainage area may change from storm to storm. For example, Smith and Granato (2010) noted that periodic bypass flows from neighboring drainage areas

along ruts in the roadway and along the road edge around neighboring catch basins occurred during some storm events with high-intensity rainfall at one site in Massachusetts. Lee and Heaney (2003) noted that topographic contours of about 2 feet (ft) and manual inspection of drainage basins were necessary to delineate basins on the scale of a city block (about 0.02 mi², which is about 14 acres) in a suburban area with storm sewers. Differences between natural topographic divides and anthropogenic divides can be substantial even in relatively large urban basins. For example, Richards and Brenner (2004) noted that sewer-drainage delineation for Mallets Creek in Michigan more than doubled the drainage area from about 3.9 to 8 mi². Liu and others (2011) used digital datasets with GIS technology to delineate small urban catchments and discovered substantial differences in drainage areas (greater than 20 percent) in 17 of 18 basins based on substantial differences between expected and measured median runoff coefficients. In this study, several order-of-magnitude differences were confirmed by obtaining detailed drainage plans and by doing site inspections

Basin Length *(LENGTH)*: The basin length, also known as the main-channel length (fig. 4), is the total distance in miles from the point of interest to the highest point on the basin boundary following the main-channel route (Benson, 1962, U.S. Geological Survey, 1977, 1980; Sauer and others, 1983; Federal Emergency Management Agency, 2001a). The main-channel length may be much longer than the straight-line length between the point of interest and the selected point on the basin divide if the channel is sinuous. The primary method for identifying the main channel at each bifurcation is by following the fork that has the largest drainage area (Benson, 1962, U.S. Geological Survey, 1977). A secondary, but acceptable method is to follow the forks that have the longest watercourses (Langbein, 1947; U.S. Geological Survey, 1977). The upstream end of the system is determined by extending the main channel from the end of the mapped representation of the stream (blue line on a USGS 1:24,000 map) to the basin divide. To do this, the channel is extended so that it crosses crenulations in topographic contours that point uphill perpendicularly. If the length is measured manually then a minimum chord length of 0.1 mi is recommended for maps with scales greater than or equal to 1:24,000, and a chord length of 0.25 mi is recommended for maps with scales less than or equal to 1:250,000 (U.S. Geological Survey, 1977). The USGS StreamStats Web application has several tools that can be used to determine the natural basin length and will output a table of distance and elevations along the path of interest (Ries and others, 2008; U.S. Geological Survey, 2011). In sewered areas, the designated main channel may follow the storm-drain system from the outfall to the basin divide if the upstream drainage area of the storm-drain system is larger than the drainage area obtained by following the natural channel upstream.

Main Channel Slope (*CSL10_85*): The main channel slope (also known as the basin slope) is the average slope of the main channel of the stream upstream from the point of

Basin length *(LENGTH)* in miles on a topographic map from the outlet along the main channel to the drainage divide.
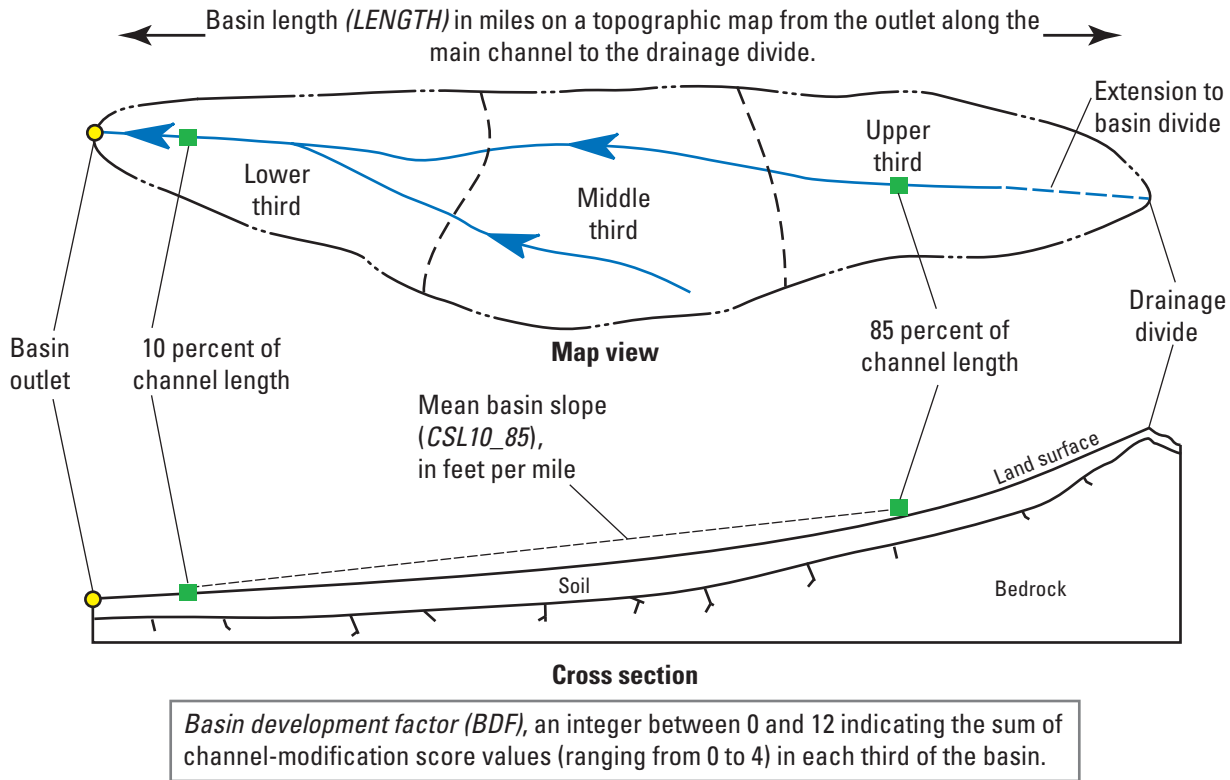
Figure 4. Simplified schematic diagram showing the physiographic basin characteristics used by Sauer and others (1983) to estimate basin lagtime.

interest (Benson, 1962; U.S. Geological Survey, 1977, 1980; Sauer and others, 1983; Federal Emergency Management Agency, 2001a). There are a number of different measures of basin slope and channel slope (U.S. Geological Survey, 1977), and it is not always clear which method has been used in a given report because the terms are commonly used interchangeably. The main channel slope selected for this analysis is known as the 10-85 slope because it is calculated by determining the locations and elevations of points at 10- and 85-percent along the main channel from the point of interest to the basin divide and then by dividing the difference in elevations by the distance in miles between these points (fig. 4). Historically, this was done by delineating the main channel as described above and selecting the topographic contours that were closest to the selected points or by visual interpolation between contours in areas of very low slope. The main-channel slope commonly is reported as feet per mile (ft/mi). Benson (1962) selected the 10-85 slope for several reasons. He postulated that the steep headwaters near and above the uppermost part of the stream probably affect the slope out of proportion to the volume of water furnished by the headwater area. He also postulated that the low slopes commonly found at the downstream end of monitored drainage basins may not represent flows from steeper slopes in the majority of the monitored basin. Benson (1962) did exhaustive tests to determine the best measure of slope using

data from 164 streamgages with drainage areas ranging from 1.64 to 9,661 mi². He determined that the 10-85 slope yielded the minimum standard error and maximum accuracy for regression models for different flood sizes among different measures of slope. USGS flood studies commonly use the 10-85 slope based on Benson's findings. As such, all the slopes from the source documents that were reported in feet per mile were assumed to be the 10-85 main channel slope. The USGS StreamStats Web application will calculate the 10-85 slope of the natural channel if this slope was used to develop regional regression models; if this variable was not used this slope can be derived using the output of the stream-network- and basin-profiling tools (U.S. Geological Survey, 2011).

Percent Storage *(STORAGE)***:** The *STORAGE* is defined as the percent of contributing area occupied by the surface of lakes, ponds, swamps, and wetlands (Langbein, 1947; Benson, 1962, U.S. Geological Survey, 1977; Sauer and others, 1983; Federal Emergency Management Agency, 2001b). *STORAGE* is measured using the same basic methods that are used for other areal measurements. Areas of interest are identified and delineated on a map, areas are measured using a grid, a planimeter, a digitizer, or GIS software and all such areas are summed and divided by the total basin area (Benson, 1962; U.S. Geological Survey, 1977; Federal Emergency Management Agency, 2001b).

Basin Lag Factor *(BLF)***:** The *BLF* is used to calculate a number of hydrograph-timing variables including the basin lagtime, the time to peak, and the time of concentration in many hydrologic studies (Benson, 1962; Linsley and others, 1975; Sauer and others, 1983; Masch, 1984; Chow and others, 1988; McCuen and others, 2002; Fang and others, 2005). The *BLF* is a secondary physiographic variable because it is calculated from primary variables as the main channel length (in mi) divided by the square root of the channel slope (in ft/mi). The *BLF* was available in most of the datasets compiled for this study or could be calculated from information documented in the reports used to compile the data.

Anthropogenic Basin Characteristics

Total Impervious Area *(IMPERV)***:** The total impervious area, commonly abbreviated as "TIA" is defined as the percentage of contributing area covered by anthropogenic impervious surfaces, which primarily consist of paved surfaces and roofs. *IMPERV* can be expressed as the fraction (0–1) or the percentage (0–100) of the total drainage area. *IMPERV* is an important explanatory variable for characterizing rainfall-runoff transformations because impervious areas can rapidly convey runoff towards the channel network. Impervious areas that drain directly to streams or storm drains with sewer outfall are known as effective impervious area (EIA) or directly connected impervious area (DCIA). Some parts of the *IMPERV*, however, may convey runoff to adjacent pervious (or semipervious) areas that may retain or retard this part of runoff. Granato (2010, appendix 6) provided an overview of methods used to estimate *IMPERV* and DCIA in a drainage basin. Granato (2010) tabulated *IMPERV* values by land-cover category in 30 studies from the literature and compiled 11 equations from the literature to estimate *IMPERV* using various land-cover characteristics. Granato (2010) developed a regression equation to estimate *IMPERV* from the percent developed area (*PDA*); commonly defined as area being covered by at least 30 percent of constructed materials, which includes pavement, rooftops, and other structures) within a basin using data from 262 stream basins in 10 metropolitan areas of the conterminous United States with drainage areas ranging from 0.35 to 216 mi^2 and *PDA* values ranging from 0.16 to 99.06 percent. Granato (2010) developed a multisegment regression equation for estimating *IMPERV* from population density using data from 6,255 stream basins in the United States with drainage areas ranging from 0.62 to 19,229 mi^2. Granato (2010) also developed two multisegment regression equations to correct bias in *IMPERV* estimates derived using the 2001 National Land Cover Dataset (NLCD). If *IMPERV* is to be determined manually from maps or aerial photographs (for example, using Google Earth) then calculating *IMPERV* from the percent developed area is most feasible. Digital datasets of land cover or land use data are more commonly analyzed using GIS software to estimate TIA. The USGS Streamstats application can be used to calculate the *PDA* or the *IMPERV* of a delineated basin in states

where these values are used in regional regression equations. The USGS StreamStats application also can produce a GIS coverage of the delineated basin that may be used with other GIS coverages to estimate *IMPERV* in the basin of interest.

Basin Development Factor *(BDF)***:** The *BDF* is defined as an index of urbanization and the prevalence of engineered drainage features (Sauer and others, 1983; Masch, 1984; Federal Emergency Management Agency, 2001c; McCuen and others, 2002). The *BDF* is an important explanatory variable for characterizing rainfall-runoff transformations because it is an index that quantifies directly connected impervious areas and the characteristics of the drainage-conveyance network. 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. 4) and applying four binary criteria to each third of the basin. Each third of the basin 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. Precise definition of the basin thirds is not considered necessary, because it will not have much effect on the final value of *BDF*. Therefore, the boundaries between basin thirds can be drawn by eye without precise measurements. Once the basin is divided, the analyst must assign a score of 1 or 0 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 defined by Sauer and others (1983) are:

1. *Channel improvements*—Channel improvements are defined as straightening, enlarging, deepening, and clearing the channel to reduce flow resistance. If at least 50 percent of the main drainage channels and principal tributaries (those that drain directly into the main channel) have been straightened, enlarged, deepened, or cleared, then a code of 1 is assigned. If such channel improvements have not been made, then a code of 0 is assigned. The 50-percent criterion should be applied using the total length of the main channel and principal tributaries in each third of the basin.

2. *Channel Linings*—Channel linings are defined as smooth impervious surfaces such as concrete. If more than 50 percent of the length of the main channels and principal tributaries has been lined with an impervious surface, then a code of 1 is assigned for this criterion; otherwise, a code of 0 is assigned. This criterion also is applied on the basis of the total length of the main channel and principal tributaries in each third of the basin. If a section of the basin can be classified as meeting the channel-lining criterion then, by default, this third of the basin must meet the channel-improvement criterion. If the main stem or principle tributaries are routed through pipe or box culverts, then these segments should also be considered as lined channels.

3. *Storm Drains or Storm Sewers—*Storm drains and storm sewers are defined as those enclosed drainage structures (usually pipes) that convey runoff directly from streets, parking lots, or roofs to the main channels or tributaries. This criterion is applied to the length of secondary tributaries in each third of the basin. If more than 50 percent of the secondary tributaries are storm sewers, then a code of 1 is assigned for this criterion; otherwise, a code of 0 is assigned.

4. *Curb-and Gutter-Streets—*Curb-and-gutter streets are defined as paved areas that are constructed to collect sheet flow and route runoff along their edges to a drainage area. Drainage from curb-and-gutter streets commonly empties into storm drains, but this criterion also may apply to areas that route runoff directly to streams or to disconnected pervious areas. If more than 50 percent of the subarea (third) is urbanized (covered with residential, commercial, and (or) industrial development), and if more than 50 percent of the streets and highways in the subarea are constructed with curbs and gutters, then a code of 1 is assigned to this aspect; otherwise, a code of 0 is assigned.

Under these definitions, the *BDF* will be calculated as an integer in the range from 0 to 12. The *BDF* is considered to be a fairly easy index to estimate manually for a given basin, because each criterion is assigned a 1 or a 0 (Sauer and others, 1983; Masch, 1984; Federal Emergency Management Agency, 2001c; McCuen and others, 2002; Granato, 2010). Sauer and others (1983) indicate that this binary four-category ranking system seems to produce consistent scores among similar basins by different analysts. It is clear that a basin with a *BDF* of 12 is highly developed. A *BDF* score of 0, however, does not necessarily indicate that a basin is undeveloped. A basin with a *BDF* of 0 may have a substantial amount of development and a relatively high value of *IMPERV*.

The *BDF*, however, is not currently (2012) in common use. Use of the *BDF* has declined with increases in the use of GIS for stream-basin analysis, because the *BDF* ranking system has been more difficult to automate than methods to estimate *IMPERV*. Use of the *BDF* is not readily characterized using a single GIS coverage. 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. First, the stream reaches would have to be coded to identify areas with channel improvements, channel linings, and storm sewers. This may be possible with continuing improvements to the National Hydrography dataset

(NHD; http://nhd.usgs.gov). In some areas, such data may be available because National Pollutant Discharge Elimination System (NPDES) Phase II rules require sewer-system mapping for many municipal separate storm sewer-systems (MS4s), which includes the collection of information about the location of intakes, major pipes, and outfalls to waters of the United States (U.S. Environmental Protection Agency, 2000). Second, algorithms would need to be developed to delineate a basin into three equal sections that may cut across subwatershed boundaries.

Values of the Basin Lagtime and Explanatory Basin Characteristics

Four hydrologic variables developed from the basin characteristics were used to develop regression equations to estimate the basin lagtime. The explanatory variables used were the drainage area, the basin lag factor, an adjusted imperviousness factor, and the basin development factor.

Basin Lagtime *(LAGTIME)***:** Both the primary and secondary datasets cover a wide range of *LAGTIME* values (table 2). In both datasets the minimum *LAGTIME* is 0.0235 hours (about 1.41 min) and the maximum *LAGTIME* is 105 hours. The average, geometric mean, and median values for the primary dataset (4.6, 2.1, and 2.0 hours, respectively) are larger than for the secondary dataset (3.5, 1.7, and 1.6 hours, respectively). Although there are fewer values in the primary dataset, the standard deviations of the arithmetic and logarithmic values are higher than for the secondary dataset (table 2).

Basin Drainage Area *(DRNAREA)***:** Both the primary and secondary datasets cover a wide range of *DRNAREA*s (table 2, fig. 5A). The minimum *DRNAREA* is 0.000116 mi^2 (about 0.074 acre) and the maximum *DRNAREA* is 1,477 mi^2 (945,280 acres) in both datasets. The average, geometric mean, and median values for the primary dataset are larger than for the secondary dataset, but the geometric mean and median values are about the same (table 2). There are 11 sites in the primary dataset and 22 sites in the secondary dataset that have *DRNAREA*s that are less than about 0.016 mi^2 (10 acres). There are 110 sites in the primary dataset and 245 sites in the secondary dataset that have *DRNAREA*s that are less than 1 mi^2 (640 acres). The probability plots of these datasets (fig. 5A) show that differences primarily occur among values toward the center of the distribution.

BasinLag Factor *(BLF)***:** As with drainage areas, the *BLF*s for the primary and secondary datasets share the same range with a small positive bias in the primary dataset in comparison to the secondary dataset (table 2, fig. 5B). The *BLF* is the basin length in miles divided by the square root of the channel slope in feet per mile. *BLF*s in both datasets range from 0.0012 to 85.57. The basin length ranged from 0.0124 mi (about 65.5 ft) to 114 mi in both datasets (fig. 5C). The main channel slope ranged from 0.9 to 1,220 ft/mi in both datasets (fig. 5D).

Table 2.   Ranges of values, population statistics, standardized beta coefficients, and standard deviations of the logarithm of the basin lagtime (*LAGTIME)* and each explanatory variable in the regression models for predicting the *LAGTIME* from basin properties.

[*BDF* basin development factor, unitless *(13−BDF* is used in the regression models); *BLF*, basin lag factor, the main channel length *(LENGTH)*, in miles divided by the square root of the main channel slope *(CSL10_85)*, in feet per mile; *BPE*, basin perviousness estimator *(100−0.99×IMPERV)*, where *IMPERV* is the total impervious area, in percent of the basin area; *DRNAREA*, basin drainage area, in square miles; *LAGTIME*, basin lagtime, the time between the centroid of the excess rainfall and the centroid of the runoff hydrograph, in hours; --, no data]

Variable	Minimum	Maximum	Mean		Median	Standard deviation		Standardized beta coefficient
			Arithmetic	Geometric		Arithmetic	Logarithm	
Primary dataset (Imperviousness and *BDF*) N = 493 sites								
LAGTIME	0.0235	105	4.6	2.1	2.0	7.93	0.5502	1.00
DRNAREA	0.000116	1,477	30.4	3.8	4.1	135	0.9255	1.68
BLF	0.0012	85.57	2.0	0.6	0.6	5.85	0.6684	1.21
IMPERV	0	100	21.3	--	20.0	18.4	--	--
BPE (100−0.99×IMPERV)	1	100	78.9	71.7	80.2	18.2	0.2986	0.54
BDF	0	12	5.2	--	5.0	3.63	--	--
13−BDF	1	13	7.8	6.7	8.0	3.63	0.2787	0.51
Secondary dataset (No *BDF*) N = 896 sites								
LAGTIME	0.0235	105	3.5	1.7	1.6	6.29	0.5377	1.00
DRNAREA	0.000116	1,477	21.0	2.8	3.8	103	0.9451	1.76
BLF	0.0012	85.57	1.5	0.5	0.5	4.50	0.6565	1.22
IMPERV	0	100	21.6	--	20.0	18.2	--	--
BPE (100−0.99×IMPERV)	1	100	78.7	73.5	80.2	18.0	0.2374	0.44

Basin Perviousness Estimator *(BPE)***:** The *IMPERV* values in the dataset were used to calculate the *BPE*. The *BPE* is 100 minus 0.99 times the *IMPERV* (in percent). The *BPE* was used for three reasons. Relations between basin properties and basin lagtimes were best characterized by regression equations developed using the logarithms of data; sites with *IMPERV* values of zero would have to be excluded without an adjustment factor. In many cases a small positive arbitrary number is added to every value to include zero values in such regression equations, but the objective in developing the *BPE* was to maintain values between 0 and 100. *IMPERV* was converted to an estimate of perviousness so that the factor would increase with decreasing development and, presumably, with increasing lagtimes; this is consistent with the approach taken by Sauer and others (1983) for using the *BDF* in lagtime regression equations. In both datasets *IMPERV* ranges from 0 to 100 percent and the *BPE* ranges from 0 to 100 percent (table 2). Both datasets also have similar mean (about 21 percent) and median (about 20 percent) *IMPERV* values. The probability plots of the datasets are very similar, but the secondary dataset has many more tied values in the lower *IMPERV* range (fig. 5E). This may be an artifact of the methods used to estimate *IMPERV* in some studies or may reflect the extra effort in basin characterization needed to develop the *BDF* values in the primary dataset.

Basin Development Factor *(BDF)***:** By definition, the *BDF* values are integers in the range 0 to 12, and the term used in the regression equations, *13−BDF*, ranges from 1 to 13. Sauer and others (1983) used the *BDF* on a reverse scale, *13−BDF*, so that the factor would increase with decreasing development and, presumably, with increasing lagtimes. The average and median *BDF* values in the primary dataset were 5.2 and 5.0 respectively (table 2). About 19 percent of the sites have a *BDF* value of zero and about 5 percent have a *BDF* of 12 (fig. 5F). In comparison, about 19 percent of sites have *IMPERV* values less than 5 percent. The *BDF* and *IMPERV* are correlated, but because the *BDF* is strongly influenced by engineered modifications to receiving-water channels, the *IMPERV* does not determine the *BDF*. The wide variation in *IMPERV* values for a given *BDF* is apparent in figure 6.

Percent Storage *(STORAGE)***:** *STORAGE* was not used for estimating basin lagtime and hydrograph-timing indexes in this study because (1) the number of sites with storage values were limited; (2) the range of values were skewed toward zero; (3) the location of lakes, ponds, and wetlands in the basin could determine the effect of storage; and (4) concerns about the consistency of map delineation between different studies. Only 61 percent of the 493 sites in the primary dataset have an associated storage value, and only 43 percent of the 896 sites in the secondary dataset have an associated storage

Figure 5. Scatterplots showing probability distribution of basin properties for sites in the primary (493 sites) and secondary (896 sites) dataset including *A,* the drainage area; *B,* the basin lag factor, which is the basin length divided by the square root of the channel slope; *C,* the basin length in miles; *D,* the main channel slope in feet per mile; *E,* the total impervious area, in percent of the basin area; and *F,* the basin development factor, unitless.

value. About 38 percent of storage values in both datasets are equal to zero, which limits the utility of percent storage as a predictive variable. The positions of storage areas is a concern because the positions of natural storage areas in the watershed are almost as important as the total area of storage. The positions determine the percentage of stormwater that would be routed through a storage area and the travel time between the outlet of a storage area and the site of interest (Langbien, 1947). Concerns about the consistency of map delineation of percent storage among different studies are primarily focused on wetland delineation. Langbien (1947) noted that the areal extent of swamps as reported are affected by the hydrologic conditions under which the topographic surveys were made. Swamp areas would be greater if surveys were made in wet years or wet seasons rather than if the surveys were made in dry years or dry seasons. Benson (1962) noted

that areas of swamps as delineated on different topographic maps of the same area varied considerably (by as much as 200–300 percent), so he used a storage parameter that did not include wetlands. National standards for wetland mapping have only been published recently, and these standards still allow for a substantial amount of deviation from the approved methods in the standards (Wetlands Subcommittee of the Federal Geographic Data Committee, 2009).

Furthermore, *STORAGE* does not include engineered or accidental in-channel and detention storage in the reports used to compile the database. Sauer and others (1983) defined this storage as temporary water storage in planned detention areas (for example, behind such structures as detention dams) or unplanned detention areas (for example, behind highway or railroad embankments with under-sized culverts). They indicated that unplanned detention was difficult to quantify

Figure 5. Scatterplots showing probability distribution of basin properties for sites in the primary (493 sites) and secondary (896 sites) dataset including *A*, the drainage area; *B*, the basin lag factor, which is the basin length divided by the square root of the channel slope; *C*, the basin length in miles; *D*, the main channel slope in feet per mile; *E*, the total impervious area, in percent of the basin area; and *F*, the basin development factor, unitless.—Continued

and required subjective determination based on examination of available high-water profile data and maps, bridge and highway plans, surveys, and on field inspections. Unplanned detention also would vary considerably based on storm size and duration. Sauer and others (1983) found that this variable had a significant effect on flood peaks but did not examine the effects on lagtime. Meierdiercks and others (2010) used monitoring data and watershed modeling to show that the effects of engineered detention on peak flows and basin response times were substantial for small watersheds. Sauer and others (1983) indicated the presence of engineered storage in basins by using a "Y" for yes, indicated the absence of engineered storage with a "N" for no, and used a "U" for unknown. This classification was not sufficient to develop regression equations. None of the other 36 studies (with the exception of Meierdiercks and others, 2010) provided detention storage data.

Analytical Procedures for Regression Analysis

Ordinary least squares (OLS) regression techniques were used to relate the selected basin characteristics to the basin lagtime. Multiple linear regression analysis were used to develop regression models that incorporated one or more physiographic variables to represent the effects of basin properties and one or more anthropogenic properties to represent the effects of development on basin lagtimes. Multiple linear regression analysis using OLS techniques provides a systematic method for estimating the coefficients of a mathematical equation of the relation between a response variable (basin lagtime) and two or more explanatory variables (basin characteristics). In this regression analysis, basin lagtime (the dependent variable) for a group of data-collection stations is statistically related to one or more physical characteristics of the drainage areas for the stations (the
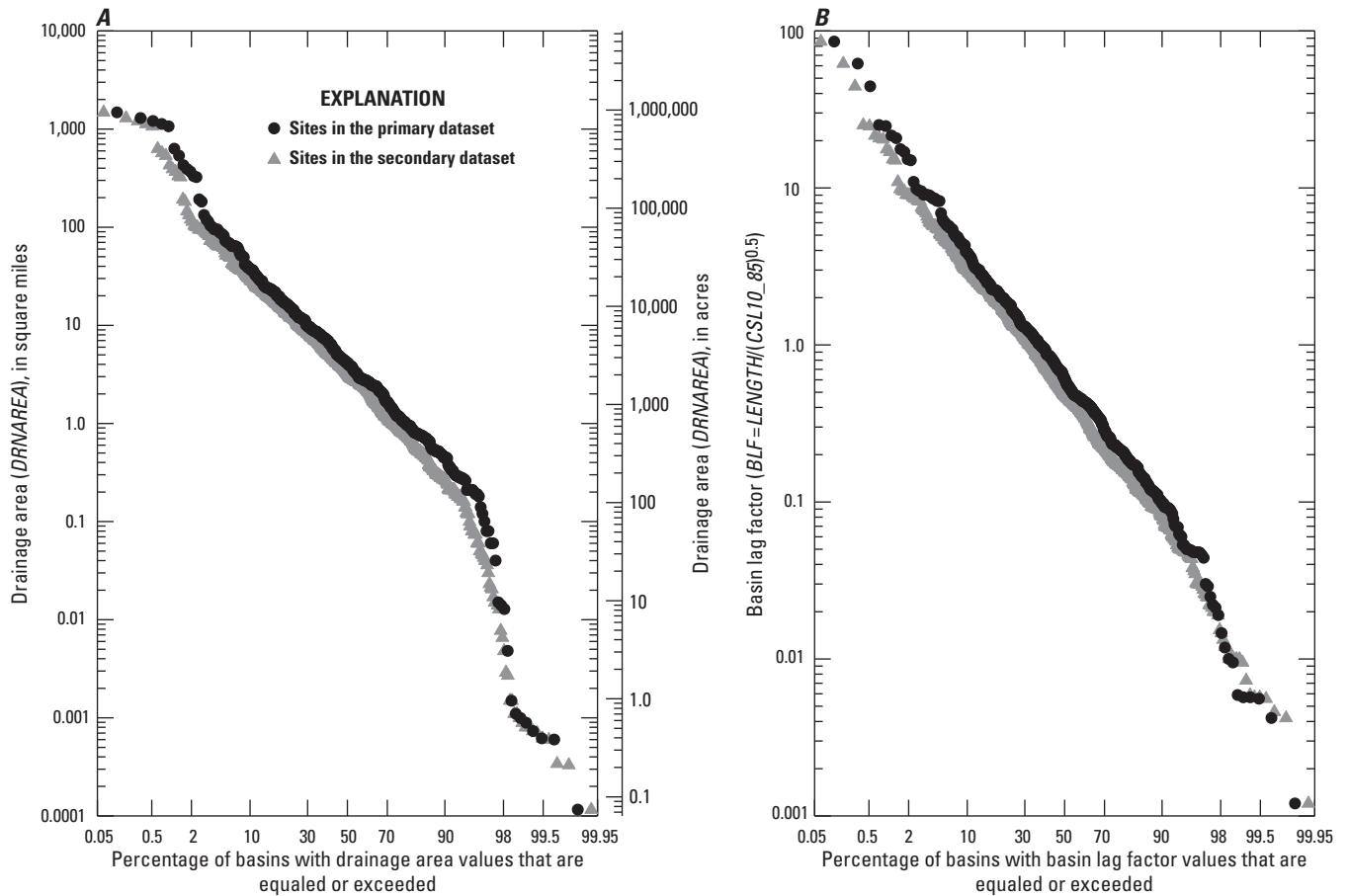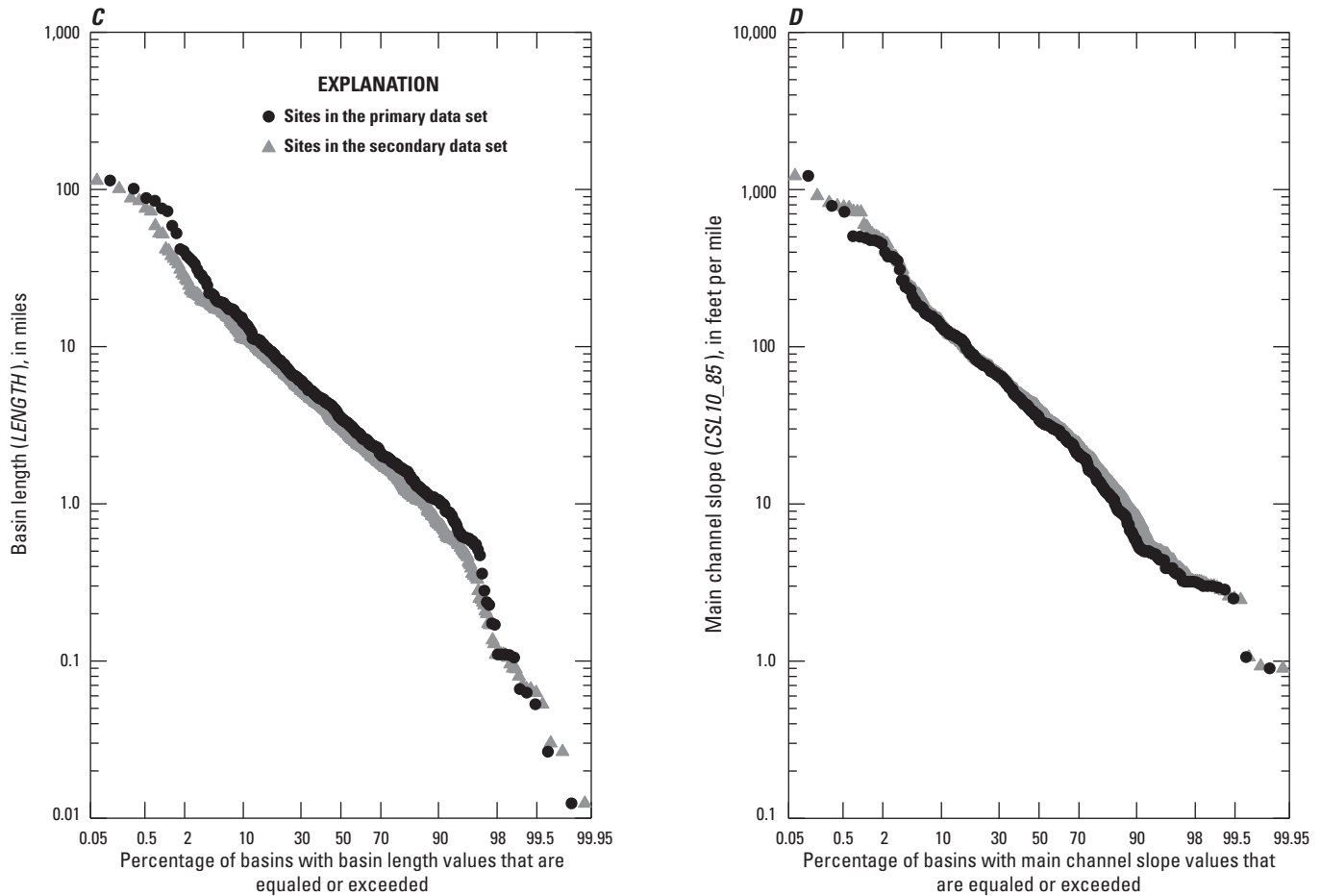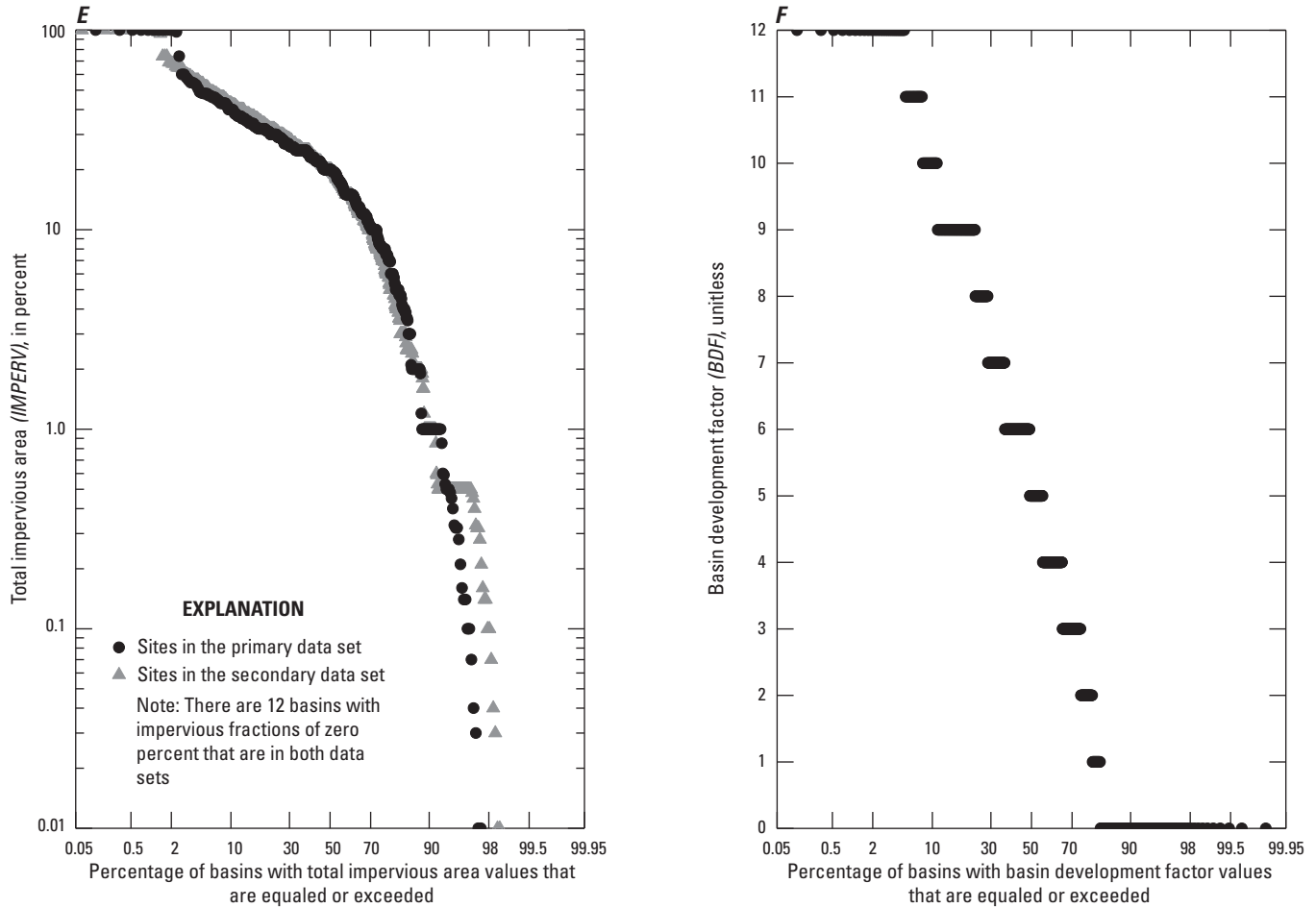
Figure 5. Scatterplots showing probability distribution of basin properties for sites in the primary (493 sites) and secondary (896 sites) dataset including *A*, the drainage area; *B,* the basin lag factor, which is the basin length divided by the square root of the channel slope; *C*, the basin length in miles; *D,* the main channel slope in feet per mile; *E,* the total impervious area, in percent of the basin area; and *F,* the basin development factor, unitless.—Continued

independent variables). This results in an equation that can be used to estimate the basin lagtime for sites where concurrent precipitation and streamflow data are not available. Linear regression analysis was used to develop regression models that incorporated either the drainage area or the basin lag factor. Multiple regression techniques were used to add the *BPE* and *13−BDF* variables to the equations.

The basin lagtime and all the potential explanatory variables were transformed to their common (base 10) logarithms to develop the regression models. The regression equations were developed with the logarithms of data because basin lagtimes, drainage areas, basin lengths, channel slope, and imperviousness all have a lower bound of zero and vary over one or more orders of magnitude. Transforming the variables to logarithms increased the linearity of relations and reduced heteroscedasticity (nonconstant variance of residuals). Logarithmic multiple linear regression equations commonly are reported in the following form:

$$log\left(Y_p\right) = \beta_0 + \beta_1 log\left(X_1\right) + \beta_2 log\left(X_2\right), \quad (5)$$
$$+...+ \beta_n log\left(X_n\right)+\varepsilon$$

where

Y_p	is the predicted value, in this case the basin lagtime;
β_0	is the estimated value of the linear intercept;
n	is the number of explanatory variables used in the equation;
X_1 to X_n	are the values of each of the n explanatory variables;
β_1 to β_n	are the slopes associated with each of the n explanatory variables; and
ε	is the random error component (scatter of measured points around the regression estimates).
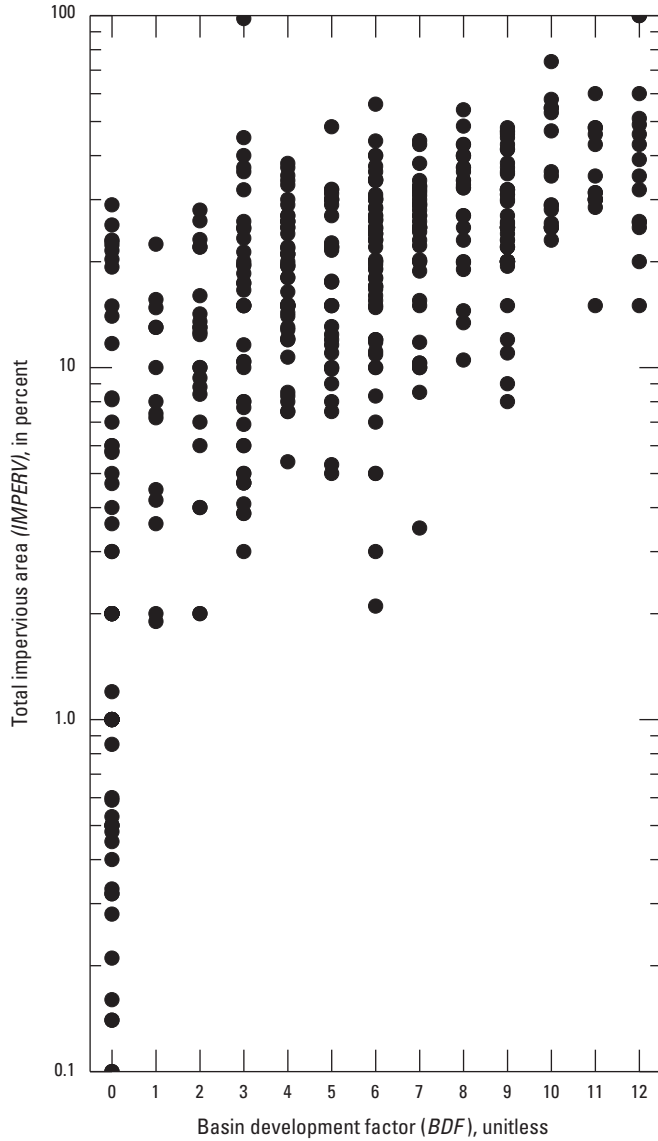
Figure 6. Scatterplot showing the relation between the basin development factor *(BDF)* and the total impervious area (*IMPERV*, in percent of basin area) for the 493 sites in the primary dataset.

Once the coefficients are calculated, equation 5 is transformed to a power equation in arithmetic space because the basin lagtime is of interest (rather than the logarithm of the basin lagtime) and the arithmetic values of the explanatory variables are more intuitive to use. Addition in log space is equivalent to multiplication in arithmetic space, and multiplication in log space is equivalent to exponentiation in arithmetic space. If retransformed from the common (base 10) logarithms, equation 5 becomes:

$$Y_p = 10^{\beta_0} \times X_1^{\beta_1} \times X_2^{\beta_2} \times ... \times X_n^{\beta_n} \times 10^{\varepsilon} , \qquad (6)$$

where the terms of the equations have the same meaning as in equation 5. In regression analysis, the mean of the errors (e) should be zero in log space; when transformed this term would equal one. However, the mean of the errors in logarithmic space is the geometric mean of errors in arithmetic space. As a consequence, the best-fit line in logarithmic space tends to provide a biased estimate of the dependent variable in arithmetic space, and a bias correction factor is needed to compensate. The nonparametric smearing estimator (the average of the retransformed log-regression residuals) proposed by Duan (1983) was selected as a bias correction factor for the regression equations developed in this study because it performs reasonably well and is not sensitive to statistical assumptions regarding residual-population characteristics (Gilroy and others, 1990; Crawford, 1991; Hirsch and others, 1993; Helsel and Hirsch, 2002). The Duan (1983) *BCF* is called the smearing estimator because the method applies or "smears" the average retransformed error over all measurements. A generalized expression of the smearing estimator, applicable for any log-based transformation, is:

$$BCF = \frac{\sum_{i=1}^{n} G(e_i)}{N_{XY}} , \qquad (7)$$

where

BCF	is the bias correction factor;
G	is the retransformation function;
e_i	is the residual error, which is calculated by subtracting the regression estimate from the data value for each data point (Y_i); and
N_{XY}	is the number of *XY* data pairs.

For the common logarithms, *G* is:

$$G(e_i) = 10^{e_i} , \qquad (8)$$

Assuming that the average error of the logarithms is zero it is the magnitude of the error, which determines the quality of the fit of the regression equation. There are a number of statistics that commonly are used to assess the strength of a regression equation (Haan, 1977; Gilroy and Tasker, 1989; Helsel and Hirsch, 2002). The coefficient of determination, commonly known as the *R*-squared (R^2) indicates the fraction of the variance explained by explanatory variables in the model; the adjusted R^2 also indicates the fraction of the variance explained by the explanatory variables in the model; however, in contrast to R^2, the adjusted R^2 is adjusted for the number of variables in the model and the size of the sample. In general, larger R^2 values indicate a better fit to the available data; a perfect linear fit would have an R^2 value of one. The Prediction Error Sum of Squares *(PRESS)* statistic is the sum of the squared errors calculated for each value (*i*)

using a version of the regression equation developed with all the other $N_{xy}-1$ values. In theory, the equation with the minimum *PRESS* value will produce the least error when making predictions with new data. The average standard error of the estimate *(ASEE)* is also known as the mean-square error, and the variance of the residuals (σ^2); it is the sum of squared residual errors divided by $N_{xy}-2$. In this case, it is a measure of the variation between the basin lagtime estimates and the observed basin lagtimes from the data-collection stations used in deriving the models. The square root of the *ASEE* is known as the root mean-square error *(RMSE)* or the standard deviation of residuals (σ). In general, regression models with smaller *ASEE* values will produce the least error among the available data. *ASEE* values are calculated within log space and are commonly converted to percent errors to facilitate model evaluation and comparison. The *ASEE* statistic is converted from a logarithm to a percent error using the equation:

$$ASEE_{Percent} = 100 \times \left(e^{(\sigma^2 \times (\ln(10))^2)} - 1 \right)^{0.5} , \qquad (9)$$

Theoretically, about two thirds of the estimates obtained from the equations for the sites used in the regression analysis will have errors less than the calculated standard errors of estimate. The average standard error of the prediction (*ASEP*) is similar to the *ASEE* but it is a measure of the predictive ability of the regression model. The *ASEP* can be estimated using the *PRESS* statistic (Gilroy and Tasker, 1989; Koltun and Whitehead, 2002) using the equation:

$$ASEP_{est} = \sqrt{\frac{PRESS}{N_{XY}}} , \qquad (10)$$

The *ASEP* can be converted from logarithms to a percent error using the method shown in equation 9. In theory, regression models with smaller *ASEP* values will produce the least error for predictions using new data. Theoretically about two-thirds of the estimates obtained from the equations for unmonitored sites will have errors less than the calculated standard errors of prediction. Another measure of accuracy of the regression models is the median absolute deviation (*MAD*) of residuals. The *MAD* of residuals is analogous to the *ASEE*, but it is a robust measure that is not affected by a few outliers. In theory the value of the *MAD* should be about two-thirds of the *ASEE* if the residuals are normally distributed without large outliers (Helsel and Hirsch, 2002). About half of the estimates obtained from the equations for the sites used in the regression analysis will have absolute errors less than the calculated *MAD* value. In general, regression models with smaller *MAD* values will produce the least error among the available data.

Development of Basin Lagtime Regression Equations

The objective of this analysis was to provide a number of regression equations so that a scientist or engineer could produce a basin lagtime estimate with available data describing the basin characteristics of a site of interest. Regression statistics from models with different numbers of explanatory variables were calculated to help assess potential benefits of more thorough watershed characterization. For example, if a regression equation based on drainage area provides an acceptable amount of accuracy and precision for a given application, then the more complex basin characteristics do not need to be calculated for that application. A multiple stepwise regression was done in which independent variables were added one at a time to obtain the final equation. In stepwise regression the automated selection program finds the explanatory variable with the highest correlation (R^2) with the response variable (basin lagtime); it then tries each of the remaining variables until it finds the two explanatory variables with the highest R^2; then it tries all of them again until it finds the three variables with the highest R^2, and so on. The goal of the stepwise procedure is to ensure that only statistically significant independent variables are added to an equation and to exclude the variables that account for a very small part of the variance in the dependent variable.

Two correlation matrixes, one with the Pearson product-moment correlation coefficient (Pearson's r) the other with Spearman's rank correlation coefficient (rho) were used to asses relations between the logarithms of explanatory variables and the logarithms of basin lagtime and relations among the explanatory variables (table 3). This was done to guide variable selection and minimize use of highly correlated variables in the same regression model. The two physiographic variables, drainage area (*DRNAREA*) and basin lag factor *(BLF)*, are strongly correlated (correlation coefficients exceed 0.7) with basin lagtime (*LAGTIME)*, and the correlations between each of these variables and *LAGTIME* are equivalent. The sign of the correlations are logical, indicating that *LAGTIME* increases as the *DRNAREA* and *BLF* increase. The *DRNAREA* and *BLF* are highly correlated (correlation coefficients exceed 0.9) with each other in the primary and secondary datasets, so these variables were not used in the same regression model. Because of this, separate regression models were developed using each of these physiographic variables. Linear correlations between the physiographic variables (*DRNAREA* and *BLF)* and the anthropogenic variables *(13−BDF* and *BPE)* are moderate to low. Correlations between the *LAGTIME* and the anthropogenic variables *13−BDF* and *BPE* are moderate and are similar indicating that these variables explain about half of the variation in *LAGTIME*. The sign of the correlations is logical because *LAGTIME* is expected to increase as *13−BDF*

Table 3. Correlation matrixes showing the parametric (Pearson's r) and nonparametric (Spearman's rho) correlation coefficients among the logarithms of variables used to formulate the regression equations.

[Correlations are among the common logarithms (base 10) of all the variables, which were used to develop the regression equations. *BDF*, basin development factor, unitless (*13−BDF* is used in the regression models); *BLF*, basin lag factor, the basin length *(LENGTH)*, in miles divided by the square root of the main channel slope *(CSL10_85)*, in feet per mile; *BPE*, basin perviousness estimator *(100−0.99×IMPERV)*, in percent where *IMPERV* is the total impervious area, in percent of the basin area; *DRNAREA*, basin drainage area, in square miles; *LAGTIME*, basin lagtime, the time between the centroid of the excess rainfall and the centroid of the runoff hydrograph, in hours; --, no data. All p-values were less than 0.005, except for the rank correlation between *13−BDF* and *BLF* (p = 0.067)]

| | Pearson's r | | | | | Spearman's rho | | | | |
	DRNAREA	*BLF*	*13−BDF*	*BPE*	*LAGTIME*	*DRNAREA*	*BLF*	*13−BDF*	*BPE*	*LAGTIME*
	Primary dataset (Imperviousness and *BDF*) N = 493 sites									
DRNAREA	1	0.924	0.355	0.608	0.786	1	0.927	0.141	0.311	0.734
BLF	0.924	1	0.251	0.444	0.781	0.927	1	0.082	0.266	0.741
13−BDF	0.355	0.251	1	0.565	0.519	0.141	0.082	1	0.73	0.415
BPE	0.608	0.444	0.565	1	0.518	0.311	0.266	0.73	1	0.518
LAGTIME	0.786	0.781	0.519	0.518	1	0.734	0.741	0.415	0.518	1
	Secondary dataset (No *BDF*) N = 896 sites									
DRNAREA	1	0.921	--	0.519	0.809	1	0.934	--	0.35	0.785
BLF	0.921	1	--	0.385	0.809	0.934	1	--	0.291	0.791
13−BDF	--	--	--	--	--	--	--	--	--	--
BPE	0.519	0.385	--	1	0.478	0.35	0.291	--	1	0.525
LAGTIME	0.809	0.809	--	0.478	1	0.785	0.791	--	0.525	1

and *BPE* increase (as the *BDF* and *IMPERV* decrease, respectively). Correlations between the two anthropogenic variables, *BPE* and *13−BDF*, are moderate to strong (with an r value of 0.565 and a rho value of 0.73). The positive, but nonlinear, correlation between these variables is apparent in figure 6.

Coefficients of the regression equations and regression statistics for predicting the basin lagtime from basin characteristics are shown in table 4. The coefficients are for the retransformed version of the regression equation (equation 6). The regression equations and regression statistics were calculated using the Minitab™ software (version 15). Plots of the equation values and residuals were made and examined to ensure that predicted values were reasonable and that residuals were unbiased and had constant variance. Two sites in the primary dataset (database site identification number 741 and 742) were identified as high-leverage far outliers and were not used for regression analysis. Analyses of regression residuals for both the primary and secondary datasets indicate that there are no geographic biases in the residuals from the regression equations. Comparison between regression equations developed with the primary and secondary datasets for the *DRNAREA* (equations RE01 and RE10), the *BLF* (equations RE05 and RE12), the *DRNAREA* and *BPE* (equations RE02 and RE11), the *BLF* and *BPE* (equations RE06 and RE13) all show similar regression coefficients that are statistically equivalent within a 95-percent confidence limit (table 4).

The USGS basin lag equation RE09 by Sauer and others (1983) also is included in table 4 for comparison to results of the current study. Equations RE07 and RE09 are based on the same explanatory variables and equation RE09 was developed with a subset of the sites used to develop RE07. The R^2, *ASEE*, and *MAD* statistics were calculated for both equations using all the data in the primary dataset. Equation RE09 had comparable regression fit statistics to equation RE07, but as expected, equation RE07 has a higher R^2 and lower error statistics than RE09 because it was developed using all the data in the primary dataset. Equation RE07 has a *BLF* slope that is similar to the *BLF* slope in equation RE09, but the slope of the *13−BDF* term is substantially higher than in equation RE09. The larger slope in equation RE07 indicates a greater difference in lagtimes with changes in *BDF* than in the USGS basin lag equation (RE09). The slope of both the *BLF* and *13−BDF* terms for these equations also are significantly different at a 95-percent confidence limit.

Multicollinearity is the condition where one or more predictor variables are highly correlated. If highly correlated variables are used, then the multiple linear regression analysis may produce faulty regression coefficients and may result in non-optimal predictions for some input values (Helsel and Hirsch, 2002). The effects of multicollinearity among predictor variables were demonstrated using the equations that use both the *IMPERV* and *BDF* (equations RE04 and RE08; three-variable models). The correlation between the associated regression variables *13−BDF* and *BPE* are fairly strong with

Table 4. Coefficients of the regression equations and regression statistics for predicting basin lagtime (*LAGTIME*), in hours from basin properties.

[*ASEE*, average standard error of the estimate; *ASEP*, average standard error of the prediction; *BDF*, basin development factor, unitless $(13-BDF$ is used in the regression models); *BCF*, bias correction factor; *BLF*, basin lag factor the main channel slope $(LENGTH)$, in miles divided by the square root of the main channel slope $(CSL10\_85)$, in feet per mile; *BPE*, basin perviousness estimator $(100-0.99 \times MPERV)$, in percent where *IMPERV* is the total impervious area, in percent of the basin area; *DRNAREA*, basin drainage area, in square miles; *LAGTIME*, basin lagtime, the time between the centroid of the excess rainfall and the centroid of the runoff hydrograph, in hours; *MAD*, median absolute deviation of residuals; *PRESS*, prediction error sum of squares; --, no data. **Bold** indicates the best equation for each dataset]

Regression equation	Intercept	DRNAREA drainage area (mi)	BLF flow factor (mi/(ft/mi)^0.5)	BPE perviousness (percent)	BDF (13–BDF) (unitless)	Adjusted R-squared (unitless)	PRESS (log(hours)^2)	ASEE (percent)	ASEP (percent)	MAD (percent)	BCF (hr)
				Primary dataset (Imperviousness and *BDF*) N = 493 sites							
RE01	1.120	0.467	--	--	--	61.8	57.26	92.0	92.3	57.9	1.397
RE02	0.698	0.445	--	0.117	--	61.9	57.12	91.6	92.1	58.9	1.390
RE03	0.431	0.409	--	--	0.542	68.3	47.59	81.3	81.8	49.7	1.317
RE04	0.851	0.441	--	-0.209	0.631	68.9	46.74	80.3	80.8	48.1	1.307
RE05	2.880	--	0.643	--	--	60.9	58.60	93.4	93.7	57.9	1.410
RE06	0.514	--	0.565	0.394	--	64.5	53.27	87.5	87.9	57.2	1.360
RE07	**0.760**	--	**0.571**	--	**0.681**	**72.0**	**41.96**	**75.2**	**75.5**	**45.8**	**1.272**
RE08	0.647	--	0.565	0.0482	0.656	72.0	42.07	75.1	75.6	46.0	1.270
RE09[a]	0.850	--	0.62	--	0.47	64.8	--	87.3	--	45.76	1.769
				Secondary dataset (No *BDF*) N = 896 sites							
RE10	1.039	0.46	--	--	--	65.3	89.72	83.6	83.7	49.6	1.340
RE11	0.493	0.437	--	0.179	--	65.9	88.89	82.9	83.2	50.5	1.334
RE12	2.679	--	0.663	--	--	65.3	89.87	83.6	83.8	46.5	1.346
RE13	**0.382**	--	**0.601**	**0.443**	--	**68.6**	**81.81**	**78.5**	**78.9**	**44.0**	**1.306**

[a] USGS basin lagtime equation (Sauer and others, 1983).

Pearson R values of 0.565 and Spearman's rho values of 0.73 (table 3). For example, three-variable regression equation RE04 (table 4), which uses *DRNAREA*, *IMPERV*, and *BDF* to estimate basin lagtimes, has substantially better *PRESS*, *ASEE*, *ASEP*, and *MAD* values than two-variable regression equation RE02, which uses only *DRNAREA* and *IMPERV*. It has slightly better statistics than the equation using only *DRNAREA* and *BDF* (RE03). However, the coefficient of the *BPE* variable in RE04 is negative, which would indicate decreasing lagtimes with increasing perviousness (or increasing lagtimes with increasing *IMPERV*). Similarly, three-variable regression equation RE08 has much better regression statistics than equation RE06 and similar statistics as equation RE07, but the coefficient of the *BPE* term is very low (table 4). Thus, equations RE04 and RE08 should not be used to estimate *LAGTIME* values.

Application of Basin Lagtime Regression Equations

Regression lines calculated using equations RE07, RE06, and RE13 are shown with data points from the primary and secondary datasets in figures 7A, B, and C, respectively. All three graphs indicate a good fit to both datasets and substantial decreases in basin lagtimes with increasing development. The graphs, which are plotted on log scale, all show that the effect of development becomes more pronounced as the *BDF* exceeds 6 and the *IMPERV* exceeds about 50 percent. Similarly, Granato (2010) calculated a break in slope in the relations between *IMPERV* and average runoff coefficients at an *IMPERV* value of 55 percent. Comparison among these graphs indicates that an *IMPERV* value of 100 in equations RE06 and RE13 will produce a lower value of basin lagtime than a *BDF* of 12 in equation RE07. Conversely, an *IMPERV* value of 0 in equations RE06 and RE13 will produce substantially lower value of basin lagtime than a *BDF* of 0 in equation RE07. Visually, equation RE13 seems to produce a better estimate of lagtimes for small, highly impervious basins, and equation RE07 seems to produce a better estimate of lagtimes for larger less developed basins (fig. 7). It should be noted, however, that differences among regression estimates for small, highly developed basins are on the order of a few minutes, whereas the differences between lagtimes for large undeveloped basins are on the order of a few hours to tens of hours (fig. 7).

Each regression equation produces a unique value of the basin lagtime for a unique combination of the values of the explanatory variables, but there is uncertainty associated with each estimate (Driver and Tasker, 1990). Prediction-interval estimates commonly are used to define the degree of uncertainty for estimating the basin lagtime for any given site. Because the equations were developed with the logarithms of data, the prediction intervals are multiplicative. The prediction interval can be calculated by

$$\frac{1}{T}\frac{BLag_i}{BCF} < BLag_i < \frac{BLag_i}{BCF}T \; , \tag{11}$$

for an unmonitored site i where the value of T is the value of the $100(1-\alpha)$ prediction interval, $BLag_i$ is the basin lagtime estimate for the site, and BCF is the bias correction factor from table 4. Because the regression equations were developed using the common (base 10) logarithms of the input variables, the value of T is calculated as:

$$T = 10^{\left[t_{(\alpha/2,n-p)} \times \left(V_{pi}\right)^{0.5}\right]} \; , \tag{12}$$

where $t_{(\alpha/2,n-p)}$ is the critical value of the Student's t distribution for the selected prediction interval for $n-p$ degrees of freedom with n data points in the regression analysis dataset and p parameters in the regression equation. V_{pi} is the variance of prediction for site i and is calculated as:

$$V_{pi} = \sigma^2 \left(1 + x_i U x_i' \right), \tag{13}$$

where σ is the standard error of the estimate for the selected regression equation, U is the variance-covariance matrix for the regression equation, and x_i and x_i' are the row and column vector of the basin characteristics of site i, respectively. For example, the row vector (x_i) for equation RE07 in table 4 would be:

$$x_i = \left[1, \log(BLF_i), \log(13 - BDF_i)\right] \; , \tag{14}$$

which includes the constant 1, and the logarithms of the *BLF* and $13-BDF$ terms. The column vector is made up of the same three values transposed. The variable U in equation 13 is the variance-covariance matrix for the regression equation. For example, the variance-covariance matrix for the regression equation RE07 in table 4 would be:

$$U = \begin{pmatrix} 0.02234 & 0.00347 & -0.02369 \\ 0.00347 & 0.00485 & -0.00292 \\ -0.02369 & -0.00292 & 0.02792 \end{pmatrix} , \tag{15}$$

The critical Student's t value for the 90-percent prediction interval, the model error variance, and the variance covariance matrix for selected regression equations are shown in table 5.

The procedure to calculate the basin lagtime is explained by an example computation following the method of Driver and Tasker (1990). The basin properties of USGS streamgage 008730025 Big Branch Tributary at Wingate Drive, Raleigh, North Carolina, were selected for this example because the contributing area of this streamgage is a very small, highly developed basin. This streamgage has a drainage area of 0.08 mi^2, *BLF* of about 0.05 mi/(ft/mi)$^{0.5}$, an *IMPERV* value of 41.7 percent, and a *BDF* of 9 (Robbins and Pope, 1996). Using equation RE07, the basin lagtime (*BLag*) is:

Figure 7. Scatterplots showing basin lagtime data and regression equations developed using the basin lag factor *(BLF)* and *A*, the basin development factor *(BDF)* with the primary dataset (equation RE07); *B*, the total impervious area *(IMPERV)* with the primary dataset (equation RE06); *C, IMPERV* with the secondary dataset (equation RE13).
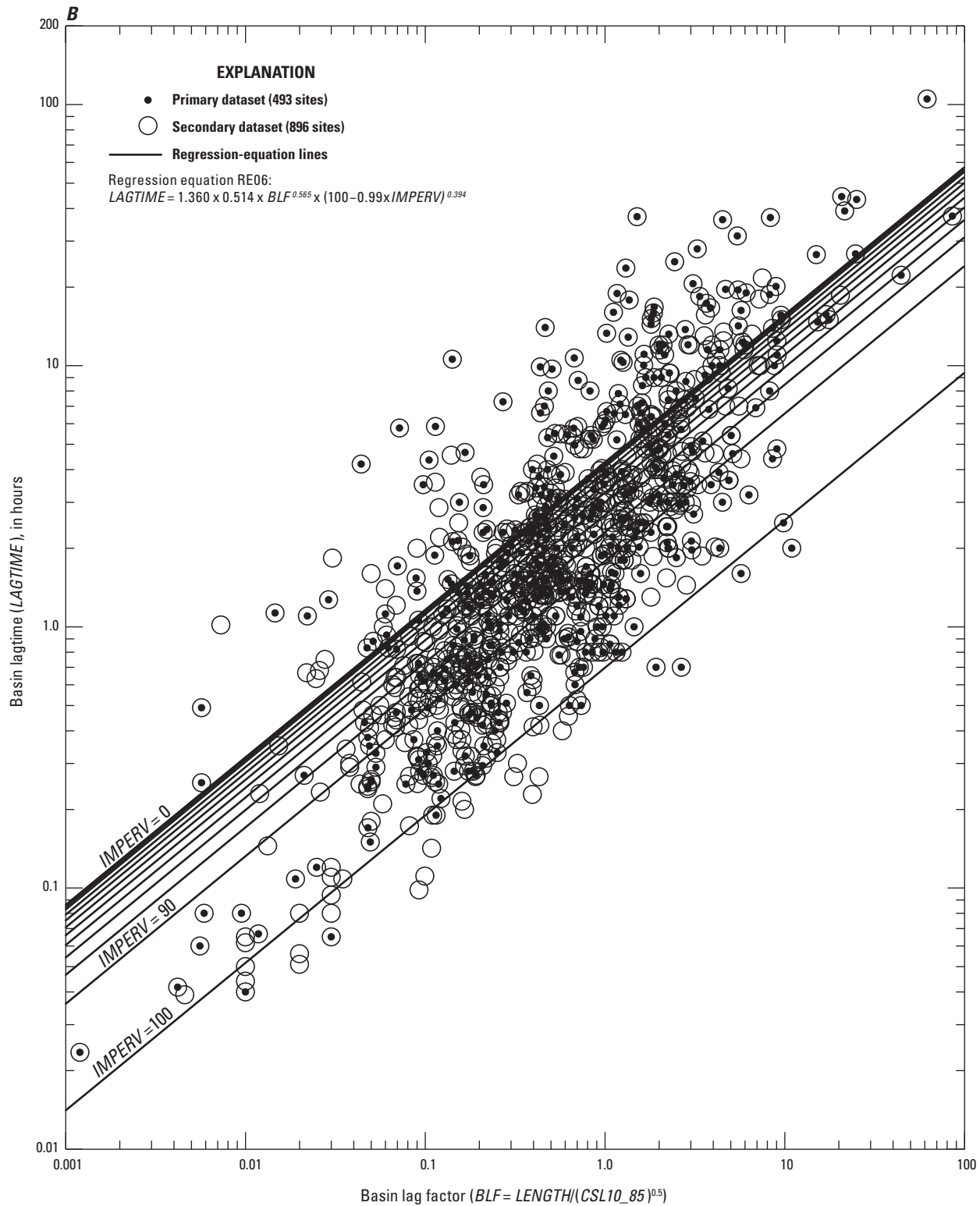
Figure 7. Scatterplots showing basin lagtime data and regression equations developed using the basin lag factor *(BLF)* and *A*, the basin development factor *(BDF)* with the primary dataset (equation RE07); *B*, the total impervious area *(IMPERV)* with the primary dataset (equation RE06); *C*, *IMPERV* with the secondary dataset (equation RE13).—Continued
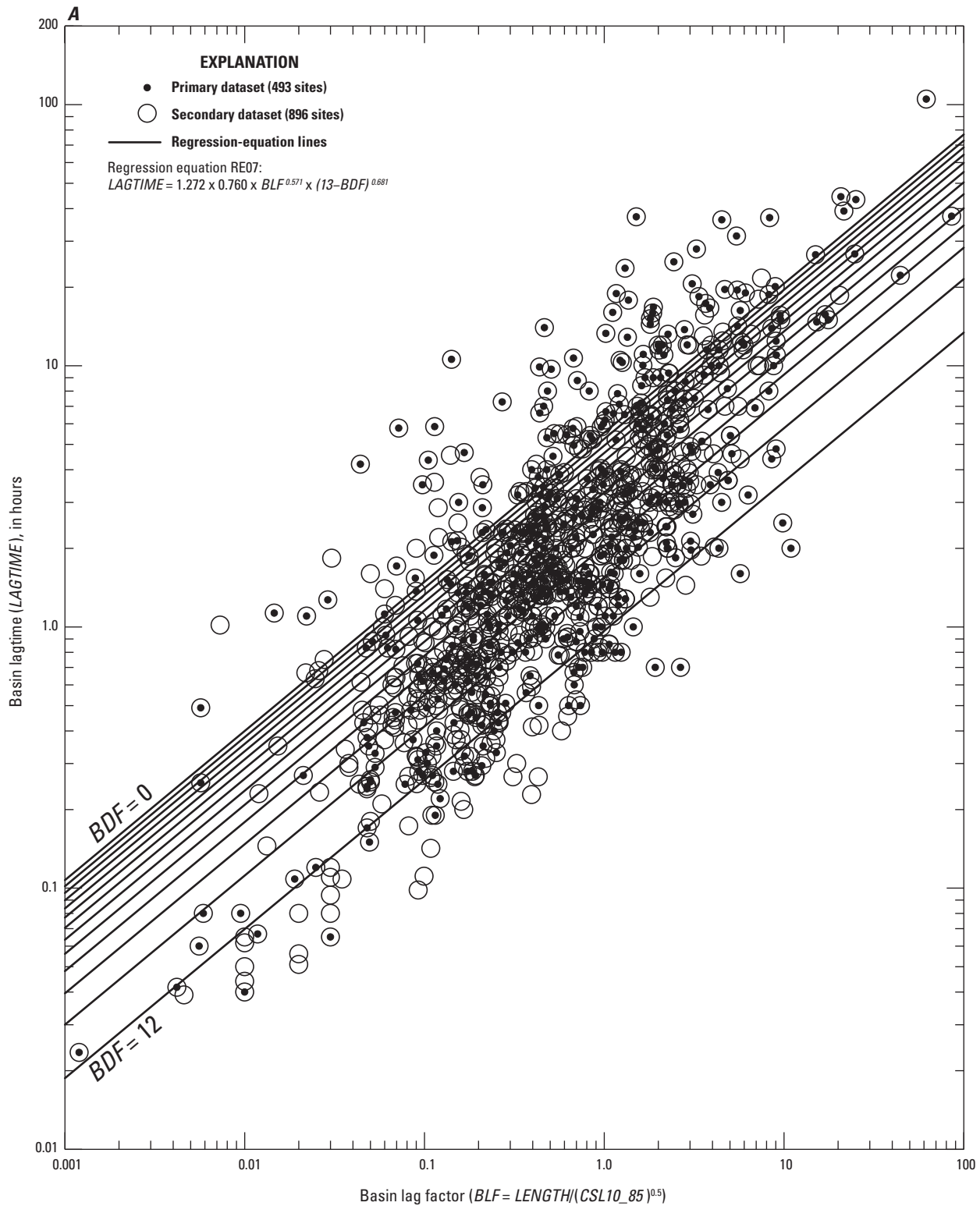
Figure 7.   Scatterplots showing basin lagtime data and regression equations developed using the basin lag factor *(BLF)* and *A*, the basin development factor *(BDF)* with the primary dataset (equation RE07); *B*, the total impervious area *(IMPERV)* with the primary dataset (equation RE06); *C*, *IMPERV* with the secondary dataset (equation RE13).—Continued
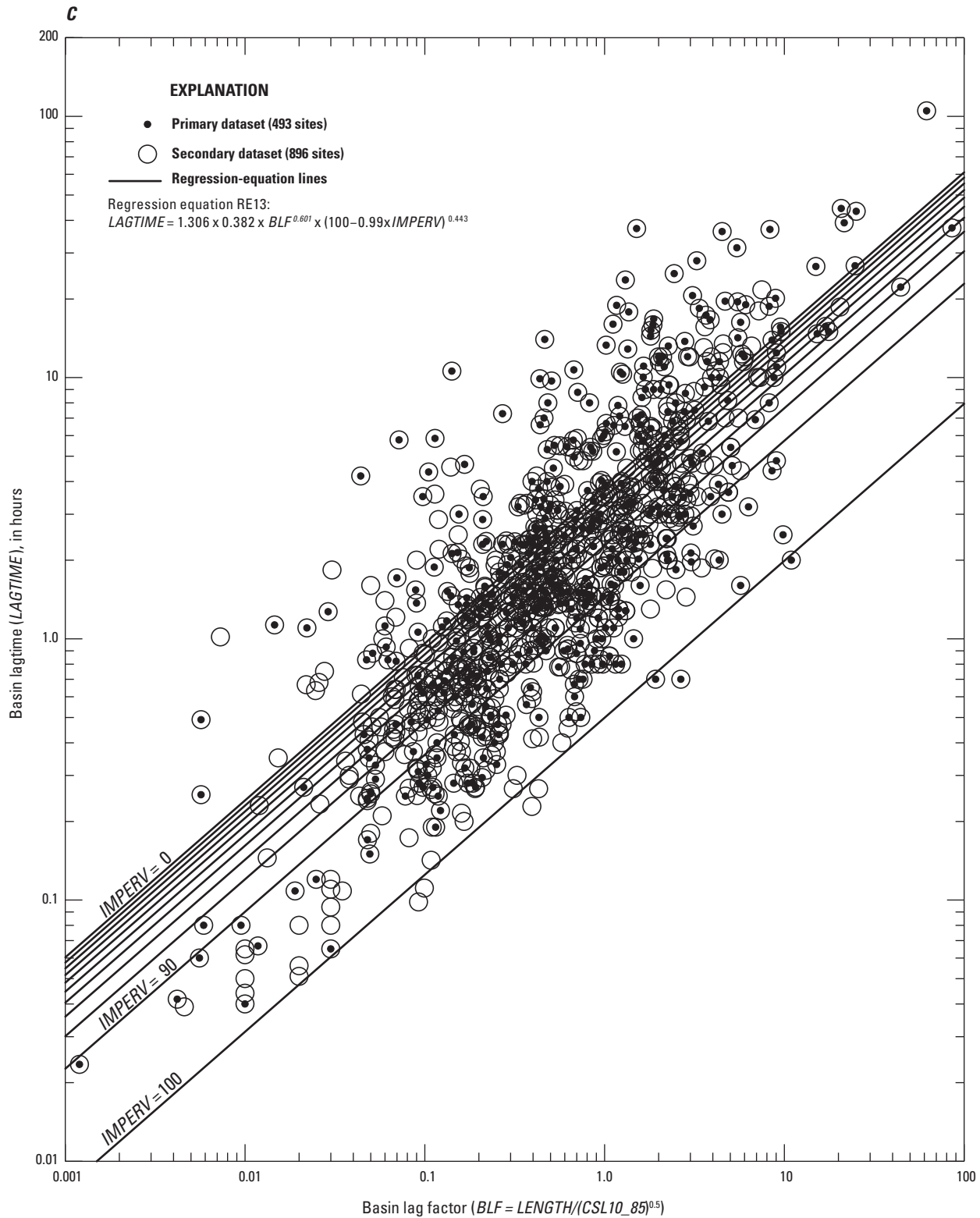
Table 5.   Values needed to determine 90-percent prediction intervals for selected basin lagtime regression equations in table 4.

[Regression model statistics developed using the common logarithms (base 10) of data values. *BDF*, basin development factor, unitless (*13−BDF* is used in the regression models); *BLF*, basin lag lactor, the basin length *(LENGTH)*, in miles divided by the square root of the main channel slope *(CSL10_85)*, in feet per mile; *BPE*, basin perviousness estimator *(100−0.99×IMPERV)*, in percent where *IMPERV* is the total impervious area, in percent of the basin area; CI, confidence interval; *DRNAREA*, basin drainage area, in square miles]

Regression equation	Number of sites	Critical student's *t* for 90-percent CI	Model error variance	Variance-covariance matrix			
Primary dataset (Imperviousness and BDF) N = 493 sites							
RE01	493	1.648	0.1158		Constant	*DRNAREA*	
				Constant	0.00282367	-0.0013737	
				DRNAREA	-0.0013737	0.00237293	
RE02	493	1.648	0.1150		Constant	*DRNAREA*	*BPE*
				Constant	0.11256004	0.01098566	-0.0629923
				DRNAREA	0.01098566	0.00376495	-0.0070947
				BPE	-0.0629923	-0.0070947	0.03615962
RE03	493	1.648	0.0958		Constant	*DRNAREA*	*13−BDF*
				Constant	0.02028623	0.00107217	-0.0228655
				DRNAREA	0.00107217	0.00271552	-0.0032027
				13−BDF	-0.0228655	-0.0032027	0.02994001
RE05	493	1.648	0.1184		Constant	*BLF*	
				Constant	0.00224496	0.00099258	
				BLF	0.00099258	0.00454919	
RE06	493	1.648	0.1072		Constant	*BLF*	*BPE*
				Constant	0.1045443	0.01166781	-0.0538735
				BLF	0.01166781	0.00566319	-0.0056219
				BPE	-0.0538735	-0.0056219	0.02837115
RE07	493	1.648	0.0845		Constant	*BDA*	*13−BDF*
				Constant	0.02234	0.00347	-0.02369
				DRNAREA	0.00347	0.00485	-0.00292
				13−BDF	-0.02369	-0.00292	0.02792
Secondary dataset (No BDF) N = 896 sites							
RE10	896	1.647	0.0999		Constant	*DRNAREA*	
				Constant	0.00135869	-0.0005509	
				DRNAREA	-0.0005509	0.00125087	
RE11	896	1.647	0.0986		Constant	*DRNAREA*	*BPE*
				Constant	0.09013766	0.00584647	-0.0490835
				DRNAREA	0.00584647	0.00171186	-0.0035369
				BPE	-0.0490835	-0.0035369	0.02713693
RE12	896	1.647	0.1000		Constant	*BLF*	
				Constant	0.00137348	0.00081689	
				BLF	0.00081689	0.00259239	
RE13	896	1.647	0.0906		Constant	*BLF*	*BPE*
				Constant	0.08626083	0.00699653	-0.0444448
				BLF	0.00699653	0.00304226	-0.0032355
				BPE	-0.0444448	-0.0032355	0.02327019

$$BLag = 1.272 \times 0.76 \times (0.05)^{0.571} \times (13-9)^{0.681} = 0.45 , \quad (16)$$

Given:

$n–p$	is 493–3 = 490 (degrees of freedom);
α	is 0.1 (90-percent prediction interval);
$t_{(\alpha/2,n-p)}$	is 1.648 (from table 5, or a statistical text);
x_i	is [1, log(0.05), log(13−9)], vector of basin characteristics (equation 10);
U	is equivalent to equation 11 (provided for the other regression equations in table 5);
σ^2	is 0.0845, the variance of the estimate (from table 5); and
BCF	is 1.272 (from table 4).

Calculate:

V_{pi}	0.0845 $(1+x_i U x_i')$ = 0.0845×1.0385 = 0.08775;
T	$10^{[1.648 (0.08775)^{\wedge}0.5]}$ = 3.077; and

$$\frac{1}{3.077} \times \frac{0.45}{1.272} < 0.45 < 3.077 \times \frac{0.45}{1.272} .$$

Thus, the estimated basin lagtime for this basin is 0.45 hours with a 90-percent prediction interval from 0.11 to 1.09 hours using equation RE07 (table 4). In comparison, Robbins and Pope (1996) estimated the basin lagtime to be 0.26 hours. The spreadsheet "ExamplePredictionIntervals.xls" on the digital media accompanying this report is designed to allow the user to calculate basin lagtimes and associated prediction intervals for any site within the limits used to develop the regression equations. The spreadsheet can be used to calculate basin lagtimes and prediction intervals using regression equations RE01, RE02, RE03, RE05, RE06, RE07, RE10, RE11, RE12 and RE13.

Limitations of the Analysis

The physiographic and anthropogenic explanatory variables in the regression models must be computed or estimated from maps, observations, and other data, which are subject to errors in measurement and interpretation (Driver and Tasker, 1990). The amount of variance in residuals of these regression models and the large prediction intervals may be the result of at least four factors. These factors include random variation in individual errors caused by physical basin properties not characterized by the regression equations, regional differences in precipitation characteristics that affect the basin lagtime, random errors in the measurement of hydrologic variables and basin characteristics within each of the 37 studies used to compile the database, and systematic error introduced by variations in the methods and source maps used to determine the basin characteristics in the different studies. These limitations are inherent in the effort to develop national regression equations without a comprehensive and systematic national initiative to characterize basin lagtimes in different regions of the country.

The potential effect of random variation in individual errors caused by physical basin properties not characterized by the regression equations can be inferred by comparing the *PRESS* statistics, standard error statistics and *MAD* statistics between regression models based on one physiographic variable (regression models RE01 and RE05 in table 4) and the regression models with one or more anthropogenic variables. In each case, substantial reductions in error occur by including the anthropogenic variables. There are many other potential variables that affect basin lag; for example, Hood and others (2007) and Meierdiercks and others (2010) both documented the effects of engineered detention storage on basin lagtimes of small drainage areas. In the current study, however, selection of explanatory variables was limited to those variables commonly documented within the source documents.

Regional differences in precipitation characteristics also may affect the amount of variation in residuals from the basin lagtime regression equations. The geographic distributions of sites in these datasets are comprehensive (fig. 3). For example, the regional average within-storm precipitation distribution probably has a systematic effect on basin lagtime, but these datasets include sites in all four SCS design storm rainfall distribution regions (McCuen and others, 2002). However, a number of the flood hydrograph studies in these datasets develop regression equations and unit hydrographs for multiple regions within the same State (or within a cluster of adjacent States) to better account for systematic variation in storm event characteristics (for example, Gamble, 1989; Bohman, 1990, 1992; Mason and Bales, 1996; Inman, 2000). Use of these regions to improve regression relations may account for spatial variations in precipitation caused by elevation and location. Thus, use of a regional variable or the selection of sites by region may reduce the variance of errors in regression equations developed for local or regional use.

Random errors occur in all measurement processes and may affect relations between variables. Although standard methods for basin characteristics set acceptable error limits at 5 to 10 percent of measured values (U.S. Geological Survey, 1977), actual errors can be substantially higher. For example, in a review of hydrologic uncertainties Granato (2010, appendix 1) found that about 15 percent of all Virginia basins and 23 percent of all Rhode Island basins meet the 5-percent accuracy criterion, but with the exception of a few outliers, the smaller basins have the larger errors. About 49 percent of the smallest basins (less than 1 mi^2) in Virginia and about 30 percent of the smallest basins in Rhode Island exceeded the 5-percent accuracy criterion. Measurement errors that were as high as 200 percent of actual drainage area were noted among the small basins. Random misspecification of *IMPERV* and *BDF* may occur if changes in land use and drainage features occur in some basins during the period used to develop the average basin lagtime.

Systematic errors in the specification of basin characteristics and basin lagtimes may have been introduced by compiling a large number of different studies that

were completed during a 41 year period (1969–2010). The maps and methods used to compile the data can be sources of systematic errors. For example, Granato (2010, appendix 1) compiled results from several studies that examined automatic and manual drainage-area assessments and found that manual methods (probably used in most early studies) were more accurate than GIS delineations. Several studies also looked at the effect of GIS dataset resolution on accuracy of drainage-area delineation and found that finer resolution datasets produced considerably better estimates. Conceivably, the accuracy of drainage-area delineations and related physiographic variables may have been high when manual methods and field surveys were used, declined with early adoption of GIS methods, and then improved as more powerful computers and lower storage costs facilitated development and use of finer resolution GIS datasets. There also may be considerable differences among different methods used to determine *IMPERV*. Granato (2010, appendix 1) indicates that differences in *IMPERV* estimates made using different methods and different datasets commonly are on the order of 25 to 40 percent. However, it would be suspected that systematic differences in specified *BDF*s would be less than for other explanatory variables, because everyone uses the same criteria. Specifying the *BDF* is done using 12 binary choices based on gross measures of development, and the method is resistant to automation and detailed knowledge of the study basin is required to do the analysis. Another source of systematic error may be differences in the levels of skill and temperament among the hydrologists compiling the data for each study.

Potential effects of misspecification of explanatory variables were assessed using standardized beta coefficients (Driver and Tasker, 1989). Sensitivity tests indicate the effects of measurement and interpretive errors on estimation of the response variables in regression models. Standardized beta coefficients for all the variables in the regression models are listed in table 2 to facilitate comparisons between regression models and sensitivity testing. The standardized beta coefficient is the standard deviation of the explanatory variable divided by the standard deviation of the response variable. This coefficient (table 2) reflects the change in the mean response per unit change in the explanatory variable (in units of standard deviations of the log of both variables) when all other explanatory variables are held constant. Therefore, misspecification of a variable with a high standardized beta coefficient is likely to introduce more error than an equivalent error in a variable with a smaller beta coefficient. Some explanatory variables have more natural variance than other explanatory variables. For example, the drainage areas in the dataset vary by more than seven orders of magnitude. In comparison, the basin lag factor, which is a ratio, varies by about five orders of magnitude in the same dataset (table 2). Spacing of the observations on the explanatory variables, which may be somewhat arbitrary, also affects the standardized beta coefficients (Driver and Tasker, 1989).

Potential effects of misspecification of explanatory variables also can be assessed visually using the plots of equations in figure 7. Percent errors caused by misspecification of the *BDF* by 1 are less than 10 percent if the *BDF* is less than 5, but it can be as high as 20 to 50 percent if the *BDF* is greater than 9. Similarly, the percent errors caused by misspecification of *IMPERV* increase with increasing *IMPERV*, but the absolute values caused by minor misspecification errors are small.

Estimating Hydrograph-Timing Indexes Using Recession-Ratio Statistics

Analysis of instantaneous streamflow data is necessary to estimate triangular-hydrograph recession ratios, which are used with the precipitation duration and the basin lagtime to estimate various hydrograph timing indexes for each storm (equations 3 and 4). Many hydrologic studies have focused on the basin lagtime for analysis of flood peaks, and the regression equations in this report are sufficient to develop planning-level estimates for that variable, but recession-ratio values are not commonly available and correlations to potential explanatory variables are weak. Some literature values are available and some values can be interpreted from results of existing studies. 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 steeply 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. Furthermore, the underlying interpretation, data, and basin characteristics used for derivation of these ratios are not published (Wanielista, 1990; Wanielista and Yousef, 1993). 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. These hydrograph-recession equations indicate that, in the Houston area, the storm falling-limb time is about 3.6 times the basin lagtime for an undeveloped basin, about 5 times the basin lagtime for a developed basin, and about 13 times the basin lagtime for a fully developed basin. These ratios increase with increasing development because the reduction in the basin lagtime is much greater than the reduction in duration of the falling-limb time. 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). Shuster and others (2008) analyzed streamflow data from eight predominantly agricultural basins in southwestern Ohio with small drainage areas (ranging from 6 to 23 mi^2); their data indicate that basin-average hydrograph-recession ratios were between 1.8 and 5 (with a median of about 3.55).

Methods and data are needed to estimate triangular-hydrograph recession ratios that may be used for local hydrograph analyses. SELDM calculates the recession ratios and therefore models hydrograph durations as random variables using a triangular distribution. Thus estimates of the minimum value, most probable value, and maximum value of recession ratios are needed to represent the timing of stormflows from the upstream basin. Recession ratios may be estimated using published hydrographs or data from USGS water-data Web servers (U.S. Geological Survey, 2012a,b).

Estimating Values of the Triangular-Hydrograph Recession Ratio from Published Curvilinear Hydrographs

Triangular-hydrograph recession ratios may be estimated from published values of curvilinear flood hydrographs developed from instantaneous stormflow data from multiple storms and multiple streamgages. Curvilinear stormflow hydrographs for different regions of the country available from 11 USGS flood studies (table 6) were used to estimate triangular-hydrograph recession ratios. The flow values from the SCS curvilinear hydrograph documented in the FHWA handbook of hydrology (McCuen and others, 2002) also were used to estimate a triangular-hydrograph recession ratio. The Microsoft Excel® solver tool available in the analysis tool pack was used to find the optimal fit of the cumulative distribution of flow of a triangular hydrograph to the cumulative distribution of stormflow of these published curvilinear hydrographs. The optimal fit was calculated by minimizing the least-squares difference between the cumulative distributions of unit flows values. Least-squares optimization was used to fit the triangular hydrograph, because it has been shown to be effective for fitting data to the triangular distribution (Johnson, 1997; Back and others, 2000). These spreadsheets (USGS01.xls, USGS02.xls, and USGS03.xls) are available in the "RecessionRatio" directory on the digital media accompanying this report. First, the curvilinear hydrographs were normalized so that the first stormflow value at the beginning of the hydrograph was equal to zero and the last stormflow value was equal to 1 percent of the last tabularized flow value. The area under the hydrograph was normalized to one and the triangular hydrograph was fit to the entire

curvilinear hydrograph. The hydrograph-recession ratio for a triangular hydrograph must be adjusted to preserve the total unit flow volume and to approximate the cumulative-mass curve of a curvilinear hydrograph with a straight-line recession segment (fig. 8). Thus, the triangular hydrograph must truncate the tail of the curvilinear hydrograph. The duration of the tail of the curvilinear hydrograph may be a substantial part of the total duration of stormflow, but the mass of flow in the tail of the curvilinear runoff hydrograph commonly is a small percentage of the total stormflow. Flow values in the tails of 60 percent of these curvilinear runoff hydrographs had to be truncated so that the peaks of the curvilinear and triangular runoff hydrographs would coincide. In these cases, the tails were truncated, the total stormflow was unitized, and a new triangular hydrograph was fit using the least-squares method. This process was done iteratively until the peaks of the runoff hydrographs matched. It was not necessary to trim the national average stormflow hydrograph developed by Stricker and Sauer (1982) because the triangular hydrograph provides an excellent fit to this curvilinear hydrograph. Although the peaks are aligned, the triangular hydrograph produces flow values that slightly lag (under predicts) flows in the rising limb of the curvilinear hydrograph and slightly lead (over predicts) flows in the falling limb of the curvilinear hydrograph (fig. 8). In this case, the recession ratio is 1.75, and the duration of the triangular stormflow hydrograph is about 96 percent of the duration of the curvilinear hydrograph (table 6, fig. 8).

Analysis of available flood-flow hydrographs resulted in a population of triangular-hydrograph recession ratios that may be used as planning-level estimates for use with SELDM in many areas of the country. The recession ratios of the trimmed hydrographs ranged from 1.3 to 2.35 with a median of 1.67 and an average of 1.68 (table 6). The time-base ratios, which are the ratios of the total durations of the triangular hydrographs to the total durations of the trimmed curvilinear hydrographs, ranged from 0.56 to 0.96 with a median of 0.9 and an average of 0.86. The recession ratios of the full (untrimmed) hydrographs ranged from 1.3 to 8.75 with a median of 1.75 and an average of 2.49. The time-base ratios of the untrimmed triangular and curvilinear hydrographs ranged from 0.43 to 0.96 with a median of 0.84 and an average of 0.9. The values of the trimmed hydrographs may be best suited for estimating the most probable value of the recession ratios used in SELDM. The triangular hydrographs that are fit to the trimmed hydrographs better represent the distribution of flows in the early part of a storm when a highway site is discharging to the stream than the triangular hydrographs that are fit to the untrimmed hydrographs. This is because the straight-line triangular approximation to the curvilinear tail of the untrimmed hydrograph results in an overestimate of the very small proportion of total stormflow that occurs in the long tail of a runoff hydrograph.

Table 6.   Best-fit triangular-hydrograph recession ratios estimated from published curvilinear unit hydrographs.

[Trimmed, the published hydrograph was trimmed to align curvilinear and triangular-hydrograph peaks; Full, the published curvilinear hydrograph with asymptotic tail. The recession ratio is the recession time divided by the time to peak of the triangular hydrograph. The time base ratio is the end-time of the triangular hydrograph divided by the end time of the curvilinear hydrograph]

Citation	Trimmed recession ratio	Time-base ratio	Full recession ratio	Time-base ratio	Dimensionless hydrograph developed for:
Becker, 1990	1.67	0.83	1.67	0.83	Missouri
Bohman, 1990	1.99	0.92	3.35	0.79	Blue ridge region of South Carolina (rural-basin study)
Bohman, 1990	1.30	0.87	1.30	0.87	Piedmont region of South Carolina (rural-basin study)
Bohman, 1990	1.47	0.94	1.47	0.94	Coastal plain region of South Carolina (rural-basin study)
Bohman, 1992	1.71	0.85	2.26	0.85	Piedmont and upper coastal plain regions of South Carolina (urban-basin study)
Bohman, 1992	2.01	0.91	2.80	0.84	Lower coastal plain region of South Carolina
Craig and Rankl, 1978	2.35	0.75	3.02	0.61	Pritchard Draw near Lance Creek, Wyoming
Gamble, 1989	1.43	0.89	1.43	0.89	West Tennessee
Holnbeck and Parrett, 1996	1.66	0.90	8.75	0.43	Montana
Inman,1986	1.43	0.90	1.73	0.84	Georgia
Mason and Bales, 1996	1.41	0.90	1.62	0.87	North Carolina
McCuen and others, 2002	1.98	0.56	1.98	0.56	Soil Conservation Service curvilinear hydrograph
Neely, 1989	1.36	0.90	1.73	0.84	Arkansas
Stricker and Sauer, 1982	1.75	0.96	1.75	0.96	Nationwide urban sites
Weaver, 2003	1.73	0.86	2.51	0.75	Mecklenburg County, North Carolina
Minimum	1.30	0.56	1.30	0.43	
Median	1.67	0.90	1.75	0.84	
Average	1.68	0.86	2.49	0.79	
Maximum	2.35	0.96	8.75	0.96	

Estimating Values of the Triangular-Hydrograph Recession Ratio from Instantaneous Streamflow Data

Limitations in the availability of published recession ratios and stormflow hydrographs necessitated development of methods for facilitating recession ratios from available instantaneous unit streamflow data and for generating a dataset of recession ratios that would be sufficient for identifying potential explanatory variables for estimating recession ratios using commonly used basin characteristics. Nationally, data are available to estimate recession ratios at a site of interest from nearby hydrologically similar basins. Currently (2012), the USGS National Water Information System Web (NWISWeb) has daily data from about 26,000 streamgages (U.S. Geological Survey, 2012a), the USGS instantaneous data archive (IDA) has unit streamflow data for about 11,000 streamgages (U.S. Geological Survey, 2012b), and the USGS GAGES database (Falcone and others, 2010) has basin properties data for about 8,500 streamgages. There

are about 6,100 streamgages that are in NWISWeb, the IDA database, and the GAGES database. Methods were developed to use available data to estimate recession ratios for a given streamgage, and a dataset of basin properties and recession ratios was developed to assess the potential for developing predictive recession-ratio regression equations for use at ungaged sites.

Methods

A multistep process (fig. 9) was developed for estimating the recession ratios for USGS streamgages with data that are available from USGS water-data Web servers (U.S. Geological Survey, 2012a,b). Historical instantaneous streamflow (HIS) data, which are the flow values estimated from instantaneous measurements of stream stage recorded every few minutes, are used to analyze the hydrographs because the commonly reported daily mean discharge record is too coarse for hydrograph analysis for small basins. HIS data may be recorded every minute for small or highly impervious basins, but the most common recording rate is once every
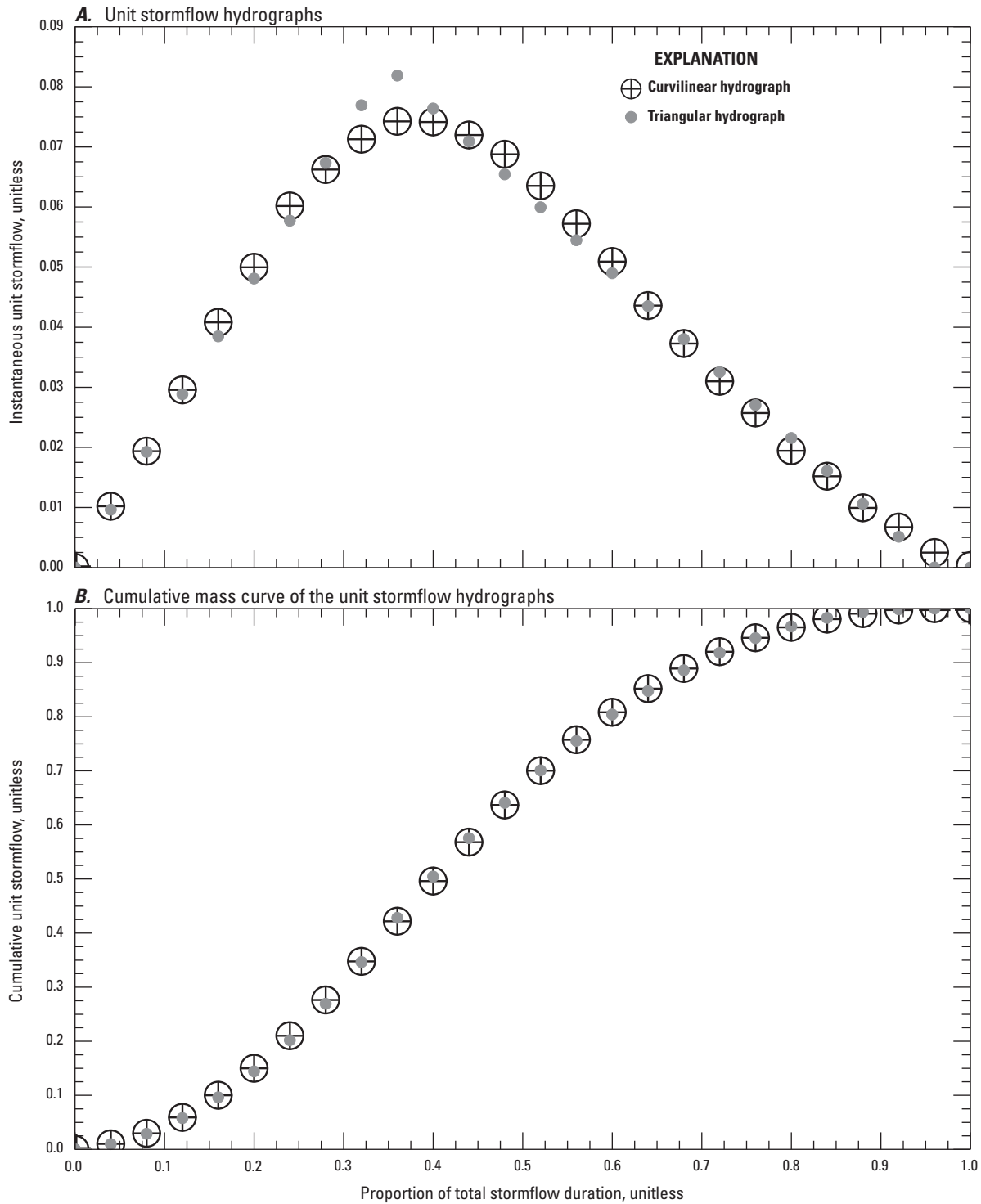
Figure 8. Graphs showing *A*, the unit stormflow hydrographs and *B*, the cumulative-mass curve of the curvilinear hydrograph published by Stricker and Sauer (1982), which is the average of unitless hydrograph values from 62 streamgages across the conterminous United States, and the best-fit triangular hydrograph.
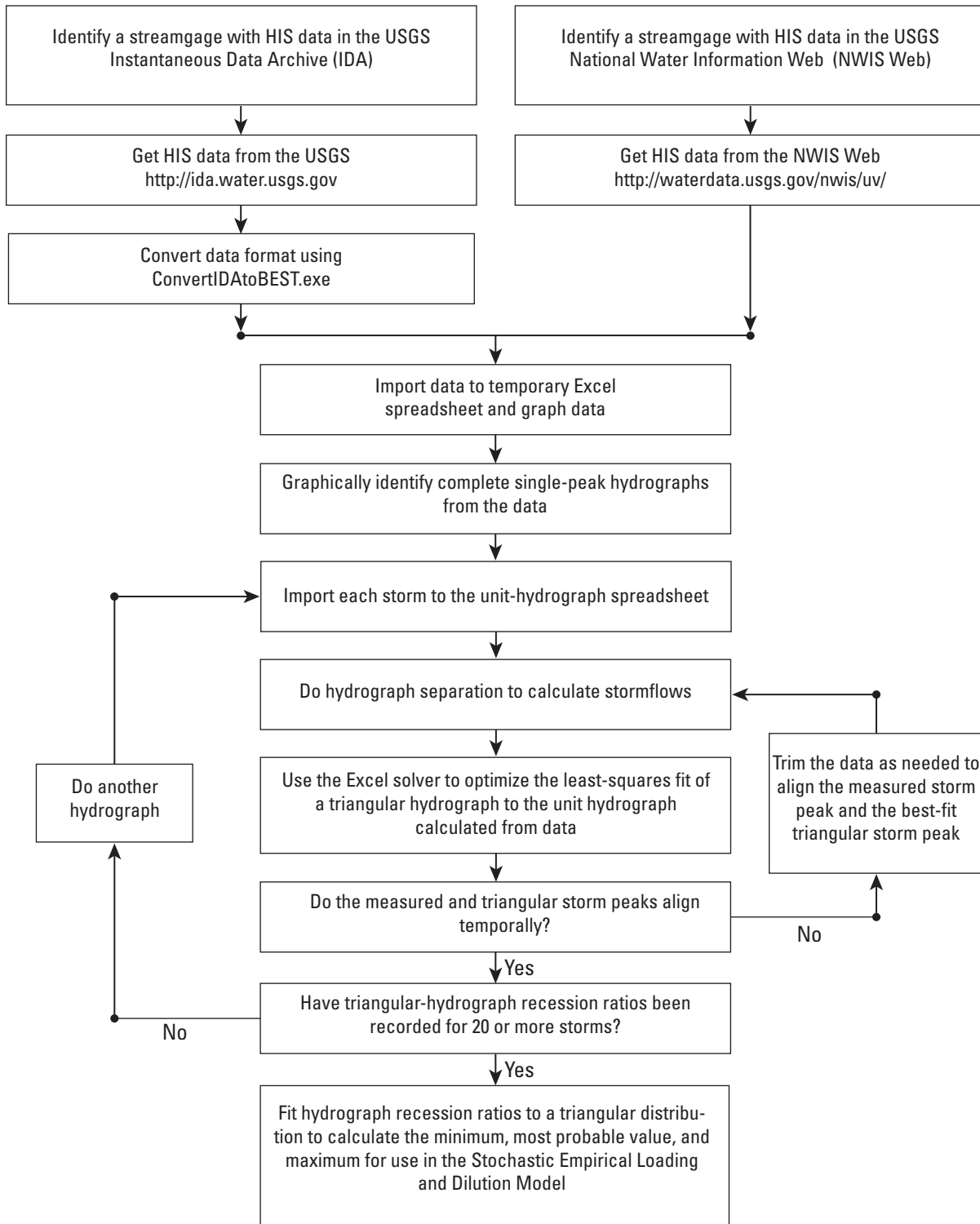
Figure 9.   Process-flow diagram for estimating the hydrograph recession-ratio statistics from historical instantaneous streamflow (HIS) data.

15 minutes. The process for estimating recession ratios is designed to identify data, process the data to identify suitable storm hydrographs, develop triangular storm hydrographs, and estimate the minimum, most probable value (MPV), and maximum recession ratios for use with SELDM. A computer program and several spreadsheets have been developed to facilitate the process; these files are available on the digital media accompanying this report. At least 20 storms are necessary to estimate robust recession-ratio statistics. The process for identifying and analyzing data from 20 to 21 storms for each streamgage may take from 2 to 4 hours.

The process for obtaining data for the hydrograph analysis depends on the source of HIS data. Before March 2012, all data were stored in the Instantaneous Data Archive (IDA). In March 2012, HIS data from October 2007 forward was transferred to the USGS National Water Information System Web (NWIS Web), and the older HIS data will be moved to NWIS Web in 2014 (Joseph Nielsen, USGS, written commun., March 2012). Thus, IDA data were used to develop the recession-ratio statistics in this report.

The first step (fig. 9) is to identify a gage that is hydrologically similar to the site of interest with HIS data in the IDA or in NWIS Web. In both cases, the NWIS Web site-inventory search page (http://waterdata.usgs.gov/nwis/inventory/) is used to identify all streamgages within a user-defined range of latitude and longitude. To select sites, choose the "Lat-Long box" and "Site type" options and click the "Submit" button. On the site selection page, select the "Site type" options, enter coordinates for a "Lat-Long box" around the ungaged site of interest, and select the "Detailed descriptions with links" option under "Retrieve Site inventory data for Selected Sites" and click the "Submit" button. The resulting page will list all the nearby sites. If data are available in the IDA or in the NWIS Web HIS database, a link to the dataset will appear with the site-inventory information. The geographic boundaries may be increased or decreased to help locate a site of interest.

The second step (fig. 9) is to get the HIS data for the streamgage of interest from the IDA or NWIS Web. The fundamental process is the same but the IDA and NWIS Web interfaces have different interface designs. In IDA enter the date range in the "Retrieve data from:" and "to:" input boxes, select the "Save to file" option in the "Tab Delimited data" combobox, click the "Retrieve Data" button and select the file name. In NWIS Web select the "00060 Discharge" option from the available parameters, select the "Tab-separated" option from the "Output format" menu, enter the "Begin date" and "End date" values, and click the "Go" button. When the file loads use the browser's File Save As menu to save the results as a text file.

The HIS data files can be very large because they contain instantaneous streamflow data that have been collected at 1 to 15 minute intervals. If there are no missing data, a 1-minute-interval file will include 525,600 data values, a 5-minute-interval file will include 105,120 data values, and a 15-minute-interval file will include 35,040 data values. The relational

database (RDB) files with the instantaneous values include a substantial amount of metadata. Microsoft Excel® 2003 has a limit of about 64,000 rows and the graphing interface has a limit of 32,000 values. Thus, it is best to retrieve and process data for 1 year at a time.

The date format in the IDA RDB files are complex and must be converted to a simpler tab-delimited text file with a date format that is recognized by Excel®. A small file-format translation program (ConvertIDAtoBEST.exe) was written to do this conversion (fig. 10). The program files, an installation package, example files, and the executable program are available in the "ConvertIDAtoBEST" directory on the digital media accompanying this report. The program has four buttons, "1. Specify Input File," "2. Specify Output File," "3. Convert File" (only visible once a file has been selected and processed), and "Exit" (fig. 10). To use the program, first click the "1. Specify Input File" button and select the RDB file using the standard Windows® common dialog box that appears. Then specify the output-file name in the input box that appears when the "2. Specify Output File," button is clicked. The default option is the input file name with ".Out.txt" appended to the end. The input and output filenames appear in the textbox under each button after the names are selected. Clicking the "3. Convert File" button using the default "Export date File" option will process the file and produce the desired output. This operation may take a few seconds to a few minutes, depending on the computer and the amount of data to be processed; an hourglass cursor will appear while the file is being processed. The user can repeat the process by clicking buttons one through three again and can exit at any time the program is not processing an input file. Once a file is processed, selected file statistics will appear in the text boxes at the bottom of the form. Once this process is complete, the data can be imported into a Microsoft Excel® spreadsheet or another program to graph the data and identify suitable stormflow hydrographs.

Once the data are in a suitable format, the next step (fig. 9) is to copy the reformatted data into a spreadsheet and graph it. It may be advantageous to convert the format of the date-value column to numbers to more easily identify the beginning and end of the storm. If the data are measured on a 15-minute interval, two decimal places will suffice; if the data are measured on a 5-minute interval, three decimal places are needed. If flow varies widely, a logarithmic scale for the flow axis can be useful for examining the hydrographs. Excel® has difficulty handling zero values with logarithmic scales; adding a very small positive number to all the values is a simple way to address this issue. The USGS commonly uses a value of 0.01 ft^3/sec as the minimum nonzero flow that can be reported. Therefore, adding a value of 0.0001 ft^3/sec to all numbers and setting the minimum value of the flow axis to 0.001 ft^3/sec will solve the problem and identify zero-value flows. It is the timing of stormflows rather than the magnitude of stormflows that are of interest for calculating recession ratios; adding the flow offset will not affect the analysis. Furthermore, base-flow separation is done in later steps and base flows are subtracted

A.   Initial view of the program interface

B.   View of the interface once a file has been selected and processed

Figure 10.   Screen captures of the interface of the "ConvertIDAtoBEST" program developed for converting U.S. Geological Survey Instantaneous Data Archive relational database files to a simple tab-delimited format. The screen captures show *A*, the initial view of the program interface and *B*, the view once a file has been selected and processed.

before normalizing the total stormflow so any offset will be removed in this step. The maximum number of points that can be plotted in Microsoft Excel® 2003 and Excel® 2007 is 32,000; the maximum number of points that can be plotted in later versions of Excel® is only limited by the capabilities of the analyst's computer. In any case, however, it may be prudent to set up a template spreadsheet with two graph pages if the analyst is doing a multistation analysis. Otherwise, the analyst can graph one range of data and then graph the rest of the data.

The next step (fig. 9) is to graphically identify complete, independent, single-peak hydrographs from the graph in the spreadsheet. This step was done manually because "good" hydrographs can be readily identified visually but are difficult

to unambiguously define using a simple computer algorithm. (Developing an automated program for identifying and processing storm events is beyond the scope of the current study.) An example of this process is shown in figure 11 using data measured during the spring of 2000 at USGS streamgage 01096000 Squannacook River near West Groton, Mass., which has a drainage area of 65.9 mi^2, a contributing drainage area of 63.7 mi^2, and about 2 percent *IMPERV*. This period of the record contains three "good" runoff events and two "bad" events for the purposes of recession-ratio analysis. These hydrographs are easy to identify because flow varies by a factor of 1.5 to 3, and random noise caused by the measurement process is small in comparison to the measured values. It can be more difficult to identify the beginning and
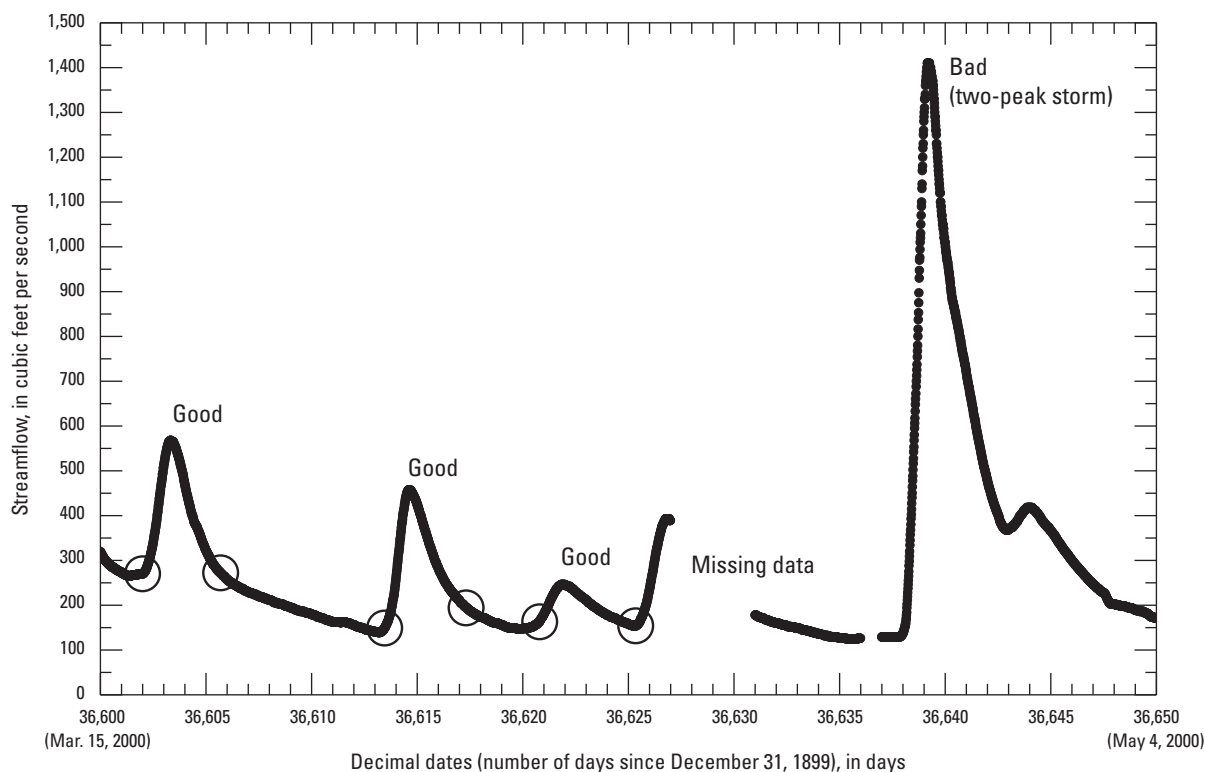
Figure 11.   Graph showing example data from U.S. Geological Survey streamgage 01096000 Squannacook River near West Groton, Massachusetts, showing the hydrographs for three "good" runoff events, defined herein as a complete, independent, single-peak event, and two "bad" events for the purposes of recession-ratio analysis.

end of smaller runoff events and runoff events from small basins in which the flows are small enough to be affected by measurement variations.

The next step (fig. 9) is to import the data for each selected runoff event into the unit-hydrograph spreadsheet and to do hydrograph separation. Hydrograph separation is necessary for normalizing the total amount of runoff that occurs during the storm. The hydrograph separation for each runoff event was done visually using principles described by Chow and others (1988). The end of the storm was loosely defined at the point where the falling limb of the hydrograph approached an asymptotic value. The spreadsheet calculates the beginning base-flow value as equal to the first line in the streamflow sample, and the ending base-flow value as defined in the heading rows of the spreadsheet (fig. 12). The runoff value at each time step is equal to the total streamflow minus the base flow. Two Microsoft Excel® spreadsheet templates—"Hydrograph01.xls," which can be used to process storms with as many as 480 data points (5 days of 15 minute data) and "Hydrograph02.xls," which can be used to process storms with as many as 880 data points (9 days of 15 minute data)—are available in the "RecessionRatio" directory on the digital media accompanying this report. These spreadsheets were sufficient to process runoff events for basins as large as 258 mi², which should be well beyond the range of basins that would be substantially affected by a stormwater outfall from a

single small site. However, if it is necessary to analyze longer storms, these spreadsheets can be modified by copying the equations downward to include more rows.

The next step (fig. 9) is to use the Microsoft Excel® solver tool to optimize the least-squares fit of a triangular hydrograph to the unit hydrograph calculated from data. The Microsoft Excel® solver tool should be installed with Excel®, but this tool must be activated using the Microsoft Excel® "Add-Ins" menu. The process for fitting a triangular hydrograph to the measured runoff data is the same as that used to fit the published curvilinear hydrograph (fig. 8). To calculate a representative recession ratio for a given storm the peak of the measured hydrograph and the best-fit hydrograph must be (approximately) aligned. It may be necessary to trim data from the initial estimated runoff period so that the peaks are aligned.

If the peaks are not aligned it is necessary to trim the data as needed to align the measured and triangular storm peaks and to use the Microsoft Excel® solver tool to generate a new best-fit hydrograph (fig. 9). If the triangular peak lagged the curvilinear peak then one or more flow values were trimmed from the beginning of the runoff period and the hydrograph separation, normalization, and fitting process was repeated. If the triangular peak preceded the curvilinear peak then one or more flow values were trimmed from the end of the runoff period and the hydrograph separation, normalization, and fitting process was repeated (fig. 9). The total volume
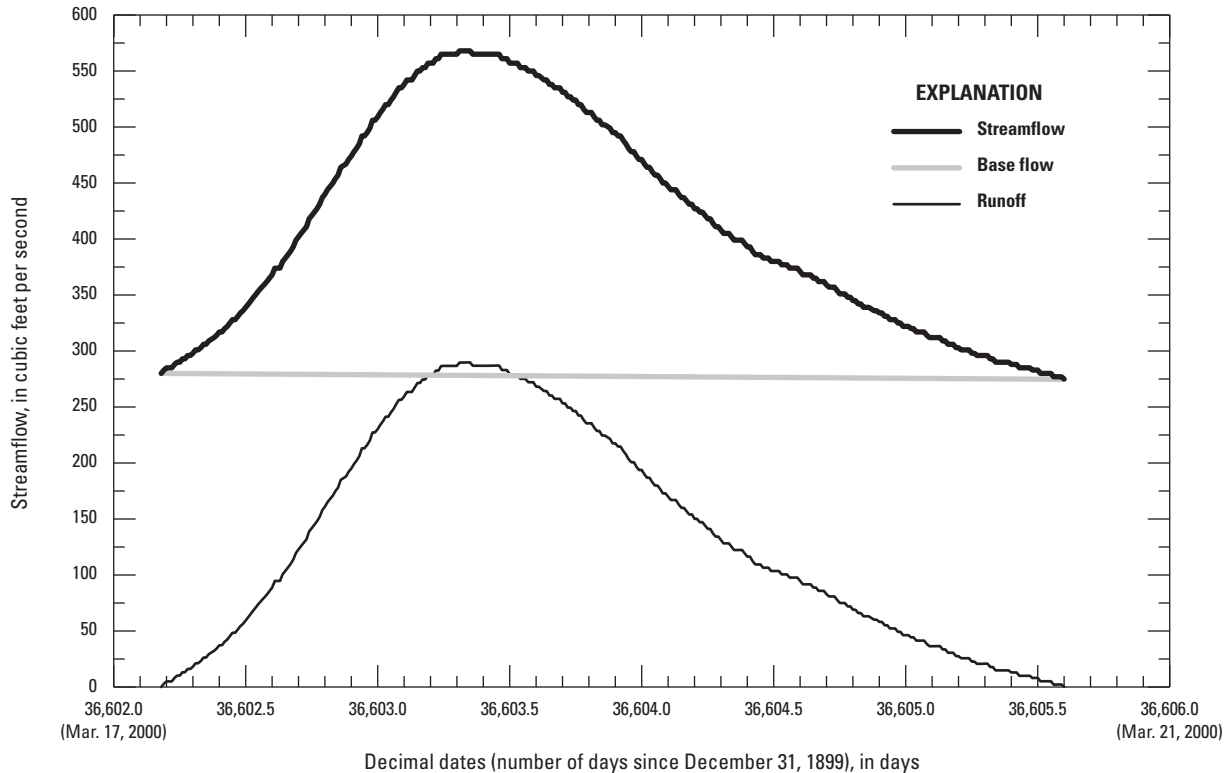
Figure 12.   Graph showing example data from USGS streamgage 01096000 Squannacook River near West Groton, Massachusetts, showing the streamflow, base-flow, and runoff values to demonstrate hydrograph separation for a runoff event that occurred during March 2000.

and duration of the trimmed period commonly was a small proportion of the initial runoff period estimate. If this analysis process resulted in a recession ratio of less than one, indicating that the falling limb of the hydrograph was shorter than the rising limb the storm was rejected or reevaluated because the theoretical lower limit of the recession ratio is one. The hydrograph separation process and the process for fitting the triangular hydrograph commonly was iterative. If data points were trimmed from the end of the hydrograph then the final value was used to redo base-flow separation and recomputed the total and unit storm volumes (fig. 9).

The next step is to repeat the separation and fitting process for at least 20 storms to provide values that could be used to determine the statistical characteristics of recession ratios for a given streamgage (fig. 9). A minimum of 20 storms was selected to produce a robust dataset for estimating recession-ratio statistics with the storms available during a period that would be representative of potential explanatory basin characteristics. Confidence intervals for parametric and nonparametric statistics commonly are expressed as a function of one divided by the square root of the number of values in a sample (Haan, 1977; McGill and others, 1978). The value of this sample-number multiplier is about 0.45, 0.32, 0.26, 0.22, 0.20, 0.18, and 0.16 for 5, 10, 15, 20, 25, 30, and 40 storms, respectively. Thus, doing the analysis for each additional

storm beyond 20 storms does not necessarily contribute a proportional amount of information to the statistical analysis. Furthermore, adding additional "good" storms required selection of storms from water years that were further and further from the year in which the potential explanatory land-cover variables were measured (in this study, that year was 2001).

The final step for each selected streamgage is to fit the calculated recession ratios from the storms to a triangular distribution (fig. 9). This process is done to estimate values for the minimum, most probable value, and maximum recession ratios used by SELDM to generate a stochastic population of runoff-event durations. To do this analysis, the recession-ratio results on the "Falling/Rising" column on the "Summary" tab of Hydrograph01.xls or Hydrograph02.xls file are copied to "FitTriangular.xls" (which is available in the "FitTriangular" directory on the digital media accompanying this report). The Microsoft Excel® solver tool also is used to fit the sample of recession ratios to the distribution. The solver-tool variables are set so that the minimum value must be greater than one. Recession ratios must be entered (or copied) into the worksheet labeled "Input-Output" in ascending order. The Microsoft Excel® solver tool is used on the worksheet labeled "Calculations" to find the best-fit minimum, MPV, and maximum values for a triangular distribution. These values

are shown on the worksheet labeled "Input-Output" with summary statistics and percentiles for the input data and the best-fit triangular distribution. The spreadsheet developed for this purpose also has a graph labeled "Chart1" to show the relation between the input values and the theoretical triangular distribution defined by the calculated statistics.

Values of the Triangular-Hydrograph Recession Ratios

The minimum, MPV, and maximum recession ratios were calculated for 41 USGS streamgages to provide a population of triangular-hydrograph recession ratio statistics that can be used as planning-level estimates for use with the SELDM (table 7). In this study, basin properties from the 2001 NLCD were used to examine potential relations between recession ratios and basin properties, so data from water year 2001 and surrounding years were used in the analysis to ensure that the land-cover variables would be consistent with the properties of the basin generating runoff. A total of 32 streamgages for basins draining parts of Massachusetts (the Massachusetts dataset) were selected for analysis. The selected gages had data in the USGS GAGES database and had data in the IDA database for several years before and after water year 2001. Streamgages for basins draining parts of Massachusetts were selected to do an initial analysis while keeping climatic variables fairly constant. Nine streamgages in other areas of the United States (the non-Massachusetts dataset) were selected from the basin lag dataset to form an initial equation-verification dataset to assess the potential for using the recession-ratio regression equations in other areas of the country.

The minimum recession ratios are well approximated by a value of one (table 7, fig. 13). The minimum recession ratios for the Massachusetts dataset ranged from 1 to 1.77 with a median of 1.05 and an average of 1.18. The minimum recession ratios for the non-Massachusetts dataset ranged from 1 to 1.27 with a median of 1.0 and an average of 1.05. The minimum recession ratios for the combined dataset of 41 gages ranged from 1 to 1.77 with a median of 1.02 and an average of 1.15. A rank sum test (Helsel and Hirsch, 2002) done using the minimum recession-ratio data indicates that the medians of the Massachusetts and non-Massachusetts datasets are not significantly different with a 95th percentile confidence limit.

The MPV of recession ratios varied substantially, and a rank sum test (Helsel and Hirsch, 2002) indicates that the medians of the Massachusetts and non-Massachusetts datasets are significantly different with a 95th percentile confidence limit (table 7, fig. 13). The MPV of recession ratios for the Massachusetts dataset ranged from 1 to 2.87 with a median of 2.01 and an average of 1.94. The MPV of recession ratios for the non-Massachusetts dataset ranged from 1 to 3.52 with a median of 1.16 and an average of 1.56. The MPV of recession ratios for the combined dataset of 41 gages ranged

from 1 to 3.52 with a median of 1.85 and an average of 1.85. These estimates of the MPV of recession ratios compare well with the values derived using the trimmed average stormflow hydrographs from the literature, which had a range of 1.3 to 2.35 a median of 1.67 and an average of 1.68 (table 6).

The maximum recession ratios varied the most among these three input statistics (table 7, fig. 13). The maximum recession ratios for the streamgages for the Massachusetts dataset ranged from 2.66 to 9.67 with a median of 4.51 and an average of 4.89. The maximum recession ratios for the non-Massachusetts dataset ranged from 2.97 to 11.31 with a median of 4.05 and an average of 5.03. The maximum recession ratios for the combined dataset of 41 streamgages ranged from 2.66 to 11.31 with a median of 4.36 and an average of 4.92. A rank sum test (Helsel and Hirsch, 2002) done using the maximum recession ratio data indicates that the medians of the Massachusetts and non-Massachusetts datasets are not significantly different with a 95th percentile confidence limit.

Correlations to Potential Explanatory Basin Characteristics

Correlations between the triangular-hydrograph recession-ratio statistics for the streamgages and selected basin characteristics were calculated to explore the potential for developing predictive multiple linear regression equations that could be used to estimate these statistics from basin properties. Standard methods for doing correlation and regression analyses were used (Haan, 1977; Helsel and Hirsch, 2002). These efforts were unsuccessful for developing meaningful predictive equations with commonly used basin characteristics, but the results do provide information for developing planning-level estimates of ratio statistics for use with SELDM. Correlations do not necessarily imply causation, but the signs (positive or negative) of the statistically significant correlations with potential explanatory basin characteristics are logically consistent with potential effects of those variables on the timing of runoff flows. The rank correlation coefficient (Spearman's rho) was selected because this nonparametric method indicates the correlation between variables regardless of the linearity of the relation, the units used for each variable, or data transformations needed to linearize the relation between variables (Helsel and Hirsch, 2002). Correlations were calculated using data from the entire 41 streamgage multistate dataset.

The values of basin properties in the 41-streamgage multistate triangular-hydrograph dataset compare well with values of basin properties in the primary, 493-streamgage basin lagtime dataset. The drainage areas of streamgages in the triangular-hydrograph dataset ranged from 0.6 to 258 mi^2 with a median of 27.8 mi^2 (table 7). This range includes the drainage areas for about 84 percent of drainage basins in the primary basin lagtime dataset (fig. 5). The *BLFs* of streamgages in the triangular-hydrograph dataset ranged from

Table 7. Best-fit triangular-hydrograph recession ratios estimated from 20 or more storm-event hydrographs at each listed streangage.

[*DRNAREA*, drainage area in square miles; *BLF*, basin lag factor, which is the basin length *(LENGTH)* in miles divided by the square root of the channel slope *(CSL10_85)* in feet per mile; *LENGTH*, main channel length in miles; *CSL10_85*, main channel slope in feet per mile; Max, maximum; Min, minimum; MPV, most probable value; *IMPERV* total impervious area in percent. MA, Massachusetts; RI, Rhode Island; CT, Connecticut; PA, Pennsylvania; MD, Maryland; NC, North Carolina; KY, Kentucky; CA, California]

Streamgage number	Name	Hydrograph-recession ratios			Basin properties (and basin-lag equation variables)				
		Min	MPV	Max	DRNAREA	LENGTH	CSL10_85	BLF	IMPERV
	Streamgages for basins draining areas in Massachusetts								
01094400	North Nashua River at Fitchburg, MA	1.00	2.23	5.00	63.4	17.8	40.7	2.78	5.98
01094500	North Nashua River near Leominster, MA	1.00	2.83	4.27	110	25.6	32.6	4.49	10.6
01095220	Stillwater River near Sterling, MA	1.66	2.01	4.05	30.4	11.5	39.3	1.83	1.52
01096000	Squannacook River near West Groton, MA	1.03	1.95	2.66	64.4	18.3	41.7	2.83	2.13
01097000	Assabet River at Maynard, MA	1.00	1.83	5.04	116	28.1	4.69	13.0	11.1
01097300	Nashoba Brook near Acton, MA	1.16	1.16	3.18	12.9	5.83	8.62	1.99	7.85
01100600	Shawsheen River near Wilmington, MA	1.08	1.34	3.08	36.5	16.2	8.61	5.52	25.3
01102500	Aberjona River at Winchester, MA	1.00	1.00	5.11	24.1	10.3	9.64	3.32	40.6
01103280	Charles River at Medway, MA	1.16	2.34	9.67	65.7	21.4	7.83	7.65	14.2
01105500	East Branch Neponset River at Canton, MA	1.25	2.20	6.23	27.2	8.32	23.4	1.72	20.0
01105600	Old Swamp River near South Weymouth, MA	1.00	1.00	3.39	4.47	4.76	10.3	1.49	25.5
01105730	Indian Head River at Hanover, MA	1.77	1.85	4.62	30.2	13.3	9.92	4.24	14.8
01108000	Taunton River near Bridgewater, MA	1.00	1.58	5.44	258	33.5	3.63	17.6	9.71
01109000	Wading River near Norton, MA	1.02	2.22	3.99	43.3	19.6	7.55	7.12	9.22
01109060	Threemile River at North Dighton, MA	1.00	1.22	5.56	84.3	32.5	5.91	13.3	10.8
01109070	Segreganset River near Dighton, MA	1.41	1.46	4.75	10.6	7.36	8.67	2.50	3.94
01111200	West River Below West Hill Dam, nr Uxbridge, MA	1.00	2.68	4.47	27.8	13.3	13.7	3.61	2.55
01111300	Nipmuc River near Harrisville, RI	1.00	2.53	5.73	16.0	7.79	30.4	1.41	1.08
01162500	Priest Brook near Winchendon, MA	1.20	2.51	4.56	19.2	15.2	19.0	3.49	0.52
01163200	Otter River at Otter River, MA	1.48	2.25	3.48	34.1	12.3	16.5	3.02	9.14
01169000	North River at Shattuckville, MA	1.18	2.87	3.24	89.9	22.6	49.0	3.23	0.56
01169900	South River near Conway, MA	1.00	1.38	3.81	24.1	14.6	58.1	1.91	0.89
01170100	Green River near Colrain, MA	1.00	2.07	4.09	41.3	19.1	59.4	2.48	0.24
01171500	Mill River at Northampton, MA	1.37	1.50	3.90	54.0	18.0	76.1	2.06	1.94
01173500	Ware River at Gibbs Crossing, MA	1.00	1.00	4.42	197	43.7	25.1	8.72	1.22

Table 7. Best-fit triangular-hydrograph recession ratios estimated from 20 or more storm-event hydrographs at each listed streamgage.—Continued

[DRNAREA, drainage area in square miles; BLF, basin lag factor, which is the basin length (LENGTH) in miles divided by the square root of the channel slope (CSL10_85) in feet per mile; LENGTH, main channel length in miles; CSL10_85, main channel slope in feet per mile; MPV, most probable value; Min, minimum; Max, maximum; IMPERV total impervious area in percent. MA, Massachusetts; RI, Rhode Island; CT, Connecticut; PA, Pennsylvania; MD, Maryland; NC, North Carolina; KY, Kentucky; CA, California]

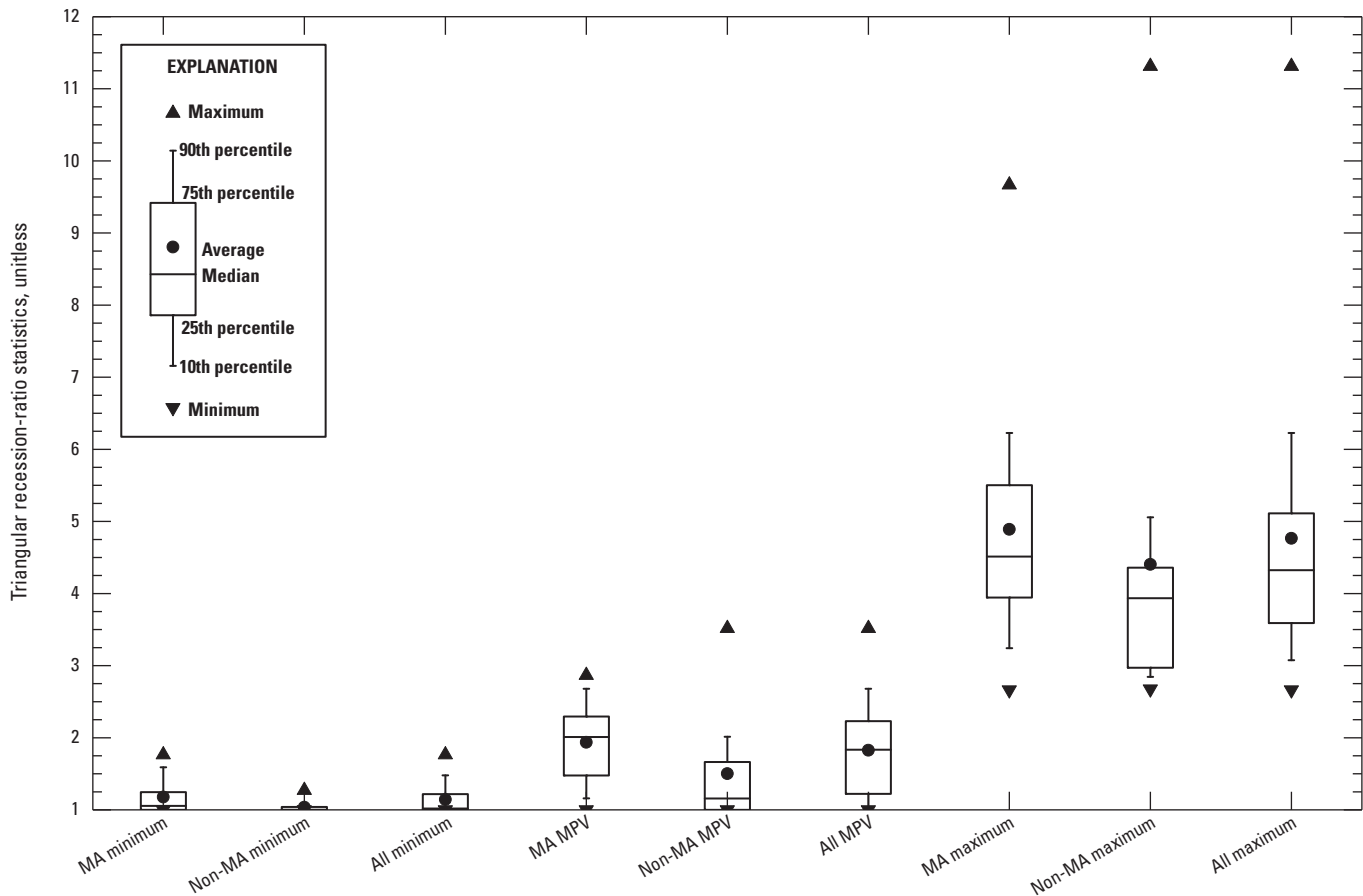| Streamgage number	Name	Hydrograph-recession ratios			Basin properties (and basin-lag equation variables)				
		Min	MPV	Max	DRNAREA	LENGTH	CSL10_85	BLF	IMPERV
01174565	West Branch Swift River near Shutesbury, MA	1.00	2.64	6.22	12.6	7.83	61.0	1.00	0.24
01174600	Cadwell Creek near Pelham, MA	1.21	2.02	4.02	0.6	1.91	129	0.17	0.43
01174900	Cadwell Creek near Belchertown, MA	1.00	2.77	4.56	2.89	3.95	135	0.34	0.17
01175670	Sevenmile River near Spencer, MA	1.59	1.67	8.60	8.69	7.95	39.4	1.27	0.71
01181000	West Branch Westfield River at Huntington, MA	1.22	2.13	6.11	93.7	23.3	44.2	3.50	0.43
01187300	Hubbard River near West Hartland, CT	1.67	1.67	9.13	19.9	10.4	67.5	1.26	0.18
01331500	Hoosic River at Adams, MA	1.24	2.10	4.12	46.7	14.4	10.3	4.49	1.52
	Minimum	1.00	1.00	2.66	0.60	1.91	3.63	0.17	0.17
	Median	1.05	2.01	4.51	32.2	14.5	24.2	2.93	2.34
	Average	1.18	1.94	4.89	52.2	16.0	34.3	4.17	7.35
	Maximum	1.77	2.87	9.67	258	43.7	135	17.6	40.6
Streamgages for basins draining different areas within the United States									
01568000	Sherman Creek at Shermans Dale, PA	1.04	1.63	3.59	207	45.3	8.23	15.8	0.64
01589100	East Branch Herbert Run at Arbutus, MD	1.00	1.66	3.94	2.47	3.30	92.4	0.34	15.0
02146409	Little Sugar Creek at Medical Center Dr at Charlotte, NC	1.00	1.00	4.05	11.8	7.26	16.9	1.77	45.7
02146700	Mcmullen Creek at Sharon View Rd near Charlotte, NC	1.00	1.00	7.43	6.95	5.47	24.2	1.11	31.6
03292500	South Fork Beargrass Creek at Louisville, KY	1.07	1.16	4.36	17.2	6.73	19.4	1.53	32.6
03293000	Main Fork Beargrass Creek at Old Cannons Lane at Louisville, KY	1.00	1.00	2.97	18.4	7.34	20.0	1.64	28.8
03302000	Pond Creek near Louisville, KY	1.27	2.01	3.33	64	12.7	11.7	3.72	35.1
11181008	Castro Valley Creek at Hayward, CA	1.00	3.52	11.31	5.51	5.10	136	0.44	40.0
11182500	San Ramon Creek at San Ramon, CA	1.02	1.02	4.32	5.89	5.72	102	0.57	1.20
	Minimum	1.00	1.00	2.97	2.47	3.30	8	0.34	0.64
	Median	1.00	1.16	4.05	11.8	6.73	20.0	1.53	31.6
	Average	1.05	1.56	5.03	37.7	11.0	47.9	2.99	25.6
	Maximum	1.27	3.52	11.31	207	45.3	136	15.8	45.7
Statistics for the combined data set									
	Minimum	1.00	1.00	2.66	0.60	1.91	3.63	0.17	0.17
	Median	1.02	1.85	4.36	27.8	12.7	23.4	2.50	5.98
	Average	1.15	1.85	4.92	49.0	14.9	37.3	3.91	11.4
	Maximum	1.77	3.52	11.31	258	45.3	136	17.6	45.7

Figure 13. Boxplot showing the distribution of the minimum, most probable value (MPV), and maximum of the best-fit triangular-hydrograph recession-ratio statistics estimated from 20 or more storm-event hydrographs from 32 streamgages for basins that drain areas in Massachusetts (MA), from 9 non-Massachusetts (non-MA) streamgages, and from the combined dataset of 41 streamgages.

0.17 to 17.6 with a median of 2.5 (table 7). This range includes the *BLF*s for about 81 percent of drainage basins in the primary basin lagtime dataset (fig. 5). The *IMPERV* values of streamgages in the triangular-hydrograph dataset ranged from 0.17 to 45.7 percent with a median of 5.98 percent (table 7). This range includes the *IMPERV* values for about 89 percent of drainage basins in the primary basin lagtime dataset (fig. 5). The triangular-hydrograph dataset may not fully characterize climatic conditions in many areas of the United States, but it includes representative ranges of basin properties as compared to the extensive (493 streamgage) nationwide basin lagtime dataset.

The recession ratio analysis indicates that correlations between basin characteristics and triangular-hydrograph recession-ratio statistics are weak or nonexistent (fig. 14, table 8). The selected basin properties, which were defined and delineated by Falcone and others (2010), are briefly described in table 8. Rank correlation coefficients between these explanatory variables and the associated 95-percent confidence limits are shown for the Massachusetts dataset and the entire multistate dataset in figure 14. If the correlation coefficients

are statistically significant at the 95-percent confidence limit, then the lines will not touch the zero line. The graph indicates that correlations are weak and most are not significantly different from zero. With one exception (percent water and wetlands), the correlation coefficients for the Massachusetts dataset and the entire multistate dataset each are within the confidence limit of the correlation coefficients for the other dataset. This indicates that the differences in coefficients between datasets are not statistically significant at this confidence limit. The streamgages in the Massachusetts dataset represent about 78 percent of the streamgages in the entire multistate dataset. The fact that many correlations switch sign with the addition of other streamgages reinforces the evidence for the lack of correlation. None of the basin characteristics have statistically significant correlations with the minimum recession ratios. This is because most of the minimum recession ratios are equal to 1 (fig. 13).

Seven variables have statistically significant correlations with the MPV of recession ratios (fig. 14). However, two variables (the percent impervious area and the subsurface flow contact time index) are significant for the multistate
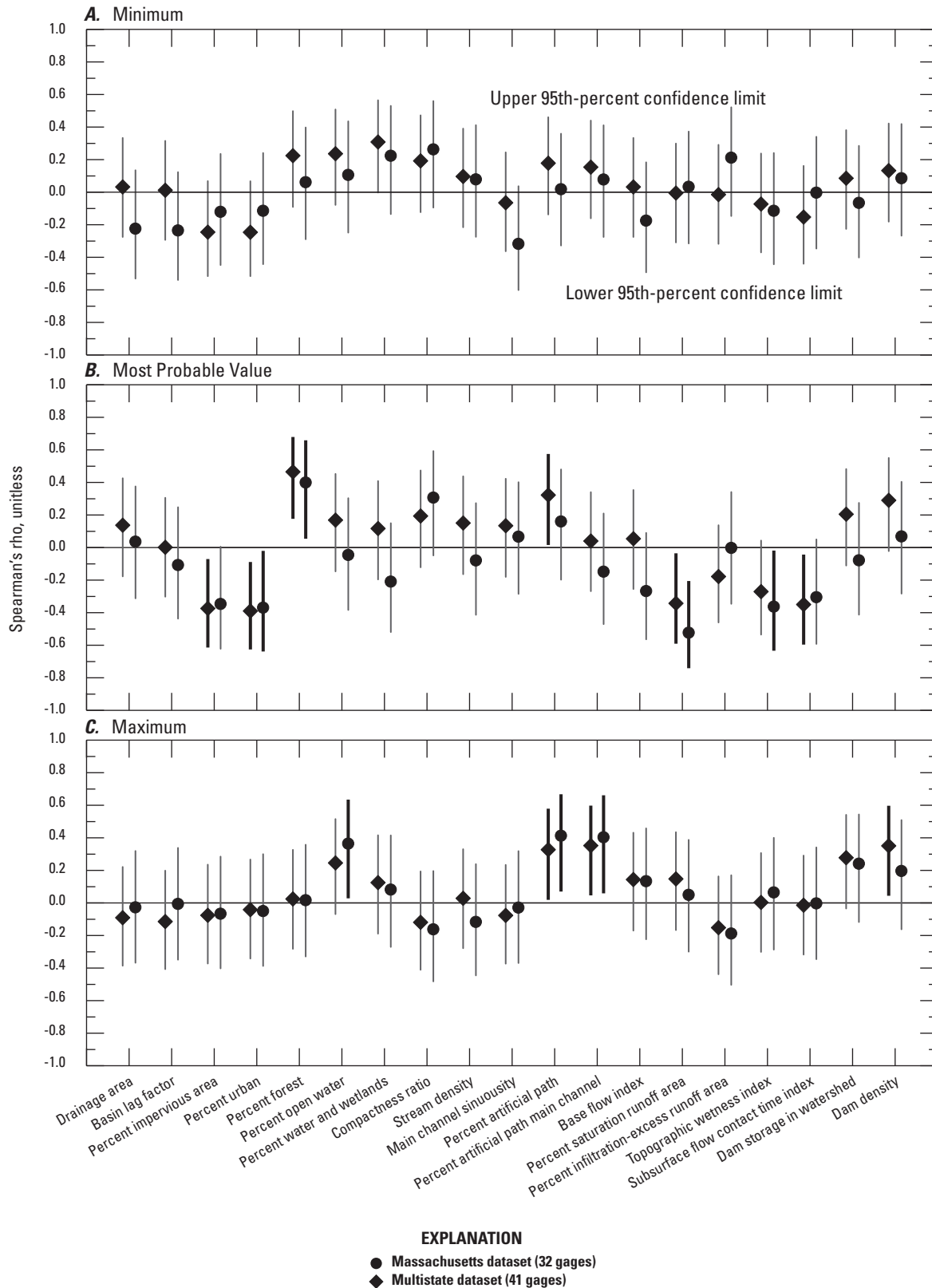
Figure 14.   Graphs showing rank correlation coefficients calculated for the *A*, minimum, *B*, most probable value (MPV), and *C,* maximum ratio of the duration of the falling to the rising limb of the hydrograph using 19 selected basin characteristics commonly used to model streamflow. Rank correlation coefficients are calculated for a dataset that includes 32 streamgages for basins draining parts of Massachusetts and for a multistate dataset that includes these streamgages and an additional 9 streamgages from different areas of the United States.

Table 8.   Rank correlation coefficients (Spearman's rho) between commonly selected basin properties compiled by Falcone and others (2010) and the maximum, minimum, and most probable value (MPV) of the triangular-hydrograph recession ratios for each streamgage in the multistate runoff-hydrograph dataset.

[Max, the maximum hydrograph-recession ratio fit to the values for each gage; Min, the minimum hydrograph-recession ratios fit to the values for each gage; MPV, the most probable value of the hydrograph-recession ratios fit to the values for each gage; NLCD 2001, National Land Cover Dataset compiled from sattlite data collected in 2001; NHD, National Hydrologic Dataset; NID, National Inventory of Dams; USACE, U.S. Army Corps of Engineers. The rank correlation coefficient is Spearman's rho. **Bold** values are (statistically) significantly different from zero at the 95-percent confidence limit]

Potential explanatory variable	Rank correlation coefficients			Variable definition
	Min	**MPV**	**Max**	
Drainage area (*DRNAREA*)	0.033	0.137	-0.090	Basin drainage area within the topographic divide that drains to the stream gage.
Basin lag factor *(BLF)*	0.013	0.002	-0.114	Main channel length *(LENGTH)* in miles divided by the square-root of the channel slope (*CSL10_85*) in feet per mile.
Percent impervious area *(IMPERV)*	-0.245	**-0.374**	-0.075	Percentage of total basin impervious area *(IMPERV)* derived from the NLCD 2001.
Percent urban	-0.246	**-0.390**	-0.041	Sum of the percentage of developed land covers derived from the NLCD 2001.
Percent forest	0.225	**0.465**	0.025	Sum of the percentage of different forest types derived from the NLCD 2001.
Percent open water	0.237	0.168	0.246	Percentage of open water derived from the NLCD 2001.
Percent water and wetlands	0.309	0.117	0.126	Sum of the percentage of open water and wetlands derived from the NLCD 2001.
Compactness ratio	0.193	0.194	-0.119	Basin drainage area divided by the squared length of the basin perimeter.
Stream density	0.097	0.151	0.029	Total length of streams divided by the basin drainage area derived from the NHD.
Main channel sinuousity	-0.064	0.134	-0.076	Curvilinear length of the main channel divided by the straight-line distance between the end points of the channel derived from the NHD.
Percent artifical path	0.179	**0.322**	**0.328**	Percent of stream lengths that are represented by polygons (impounded areas).
Main channel artifical path	0.155	0.040	**0.353**	Percent of the main channel length that is represented by polygons (impounded areas).
Base flow index	0.033	0.054	0.145	Estimated percentage of total streamflow that is base flow.
Percent saturation runoff area	-0.005	**-0.343**	0.148	Percentage of the basin contributing saturation overland flow (also know as Dunne overland flow).
Infiltration-excess runoff area	-0.014	-0.178	-0.151	Percentage of the basin contributing infiltration-excess overland flow (also know as Hortonian overland flow).
Topographic wetness index	-0.072	-0.270	0.004	Natural log of the upslope area per unit contour length divided by the slope at that point.
Subsurface flow contact time index	-0.152	**-0.350**	-0.014	Estimated average age of groundwater discharging into the stream.
Dam storage in watershed	0.086	0.204	0.279	Estimated volume of water stored behind dams per unit watershed area from the USACE 2006 NID.
Dam density	0.134	0.290	0.351	Number of dams per 100 square kilometers from the from the USACE 2006 NID.

dataset but not for the Massachusetts dataset. One variable (the topographic wetness index) is significant for the Massachusetts dataset but not the multistate dataset. The percent urban land covers are negatively correlated with the MPV of recession ratios, and the percent forest land covers are positively correlated with the MPV of recession ratios in both datasets. The signs of these correlation coefficients are logical if development decreases the duration of runoff. The correlations between the MPV and the land cover variables *IMPERV,* percent urban, and percent forest (−0.374, −0.390, and 0.465, respectively) probably are weak because increases in development tend to decrease the time to peak as well as the total duration of the hydrograph. The disproportionate decrease in the basin lagtime in comparison to the runoff duration was noted by Liscum (2001). The percent of stream length classified as artificial path in the NHD also has a weak positive correlation (0.322; table 8). A positive correlation between the artificial path and the MPV of recession ratios also is logical because natural release of temporary runoff storage from impounded areas would tend to increase the total duration of runoff. The percent saturation runoff area has a weak positive correlation with the MPV of recession ratios (−0.343; table 8). This variable represents the percentage of basin area in which the groundwater table rises to the surface during storm events to discharge groundwater and route within-event rainfall to the stream. The potential effects of this variable on recession ratios are two-fold. First, this type of runoff commonly occurs from pervious areas along the stream network. Second, discharge of lateral flow and bank storage as quickflow from such areas would tend to extend the duration of runoff and blur the boundary between runoff and post-storm base-flow recession. The subsurface flow contact time index also has a weak negative correlation coefficient with the MPV of recession ratios (−0.350; table 8). In this multistate dataset the subsurface flow contact time index may serve as an indicator for the extent of sand and gravel valley aquifers. Such areas have more permeable soils, which may extend the time to peak, and may have a greater percentage of saturation overland flow, which may extend the duration of flow.

Four variables have statistically significant correlations with the maximum of recession ratios (fig. 14). However, one variable (the percent open water) is significant for the Massachusetts dataset but not for the multistate dataset, and one variable (dam density) is significant for the multistate dataset but not for the Massachusetts dataset. The two variables, percent of stream length and percent of main channel length classified as artificial path in the NHD have weak positive correlation (0.328 and 0.353, respectively; table 8). As with the MPV of recession ratios, a positive correlation with the maximum of recession ratios is logical because the natural release of temporary runoff storage from impounded areas would tend to increase the total duration of runoff.

The weak results of these correlation analyses are consistent with the limited number of studies that quantify the timing of storm-runoff falling-limb properties. 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) also used nonparametric rank correlation coefficients (Spearman's rho) and their results indicated weak positive associations between the recession ratios and drainage areas (0.44), the percentage of forested area (0.42), and channel slope (0.28). In their study, the percentage of urban area had a moderately strong negative correlation (−0.67), and the percentage of agricultural area had a weak negative correlation (−0.26) with recession ratios.

Limitations of the Analysis

The analysis of the triangular-hydrograph recession ratios in this report and potential application of calculated ratios to other gaged and ungaged sites have several limitations. For gaged sites, the process for selecting "good" hydrographs, doing hydrograph separation, and fitting curvilinear hydrographs to a triangular hydrograph is highly interpretive. Thus, different analysts may obtain different results, and the same analysts may obtain different results if different water years are selected for analysis. Many USGS streamgages have instantaneous flow data, but few have concurrent precipitation data. Precipitation data were not included in the IDA database used in this study. Although, "good" hydrographs from single-peak hydrographs (fig. 11) were selected for analysis, recession ratios may be affected by the temporal and spatial distribution of precipitation within storms. For example, the hydrograph from a localized thunderstorm in the headwaters of a large basin may be smeared out as the runoff travels from the headwater to the streamgage. Also, a long, low-intensity storm may produce a runoff hydrograph with a different runoff ratio than a short, high-intensity storm. Furthermore, differences in seasonal conditions and antecedent precipitation may affect each basin's temporal response to precipitation. For example, the interpretive process may be affected if some storms occur during a period of increasing base flow from storm to storm and other storms occur during a period of decreasing base flow from storm to storm. It is important to represent a variety of hydrologic conditions and the selected events. Although recession ratios from 20 to 21 storms should characterize results of a random process, this sample size may not be sufficient for characterizing systematic effects of these types of hydrologic variation from storm to storm.

The size of the sample and the range of explanatory variables may not be sufficient to establish relations between the potential explanatory variables and the recession-ratio statistics. The scope of the current study limited the sample size to 41 streamgages because it takes two to four hours to complete the data processing and analysis effort for each streamgage. A larger more diverse dataset may improve correlations.

The recession-ratio statistics may largely be random variables, which may preclude development of predictive equations that can be used to estimate recession-ratio statistics at ungaged sites using measured basin characteristics. The analysis by Liscum (2001) indicates that increasing development may decrease the basin lagtime (and therefore the time to peak) at a greater rate than it decreases the falling-limb duration. Thus, these competing effects may serve to randomize recession ratios with respect to development and other related variables that affect the timing of runoff.

Summary

This nationwide study to better define triangular-hydrograph statistics for use with runoff-quality and flood-flow studies was done by the U.S. Geological Survey (USGS) in cooperation with the Federal Highway Administration. For runoff-quality modeling, information about the timing of runoff from a site of interest and from the upstream basin of the receiving stream at the location of the storm-flow outfall is necessary to estimate the quantity of the upstream flow that occurs concurrently with runoff from the site of interest. Triangular runoff hydrographs commonly are used to model stormflows and are adequate for producing planning-level estimates for discharge dilution analyses. Although the triangular hydrograph is a simple linear approximation, the cumulative distribution of stormflow with a triangular hydrograph is a curvilinear S-curve that will closely approximate the cumulative distribution of stormflows from measured data. The triangular hydrograph can be fully parameterized with the total runoff volume, the time to peak flow, and the duration of runoff. The time to peak is commonly calculated as one-half the precipitation duration plus a basin lagtime that depends on basin characteristics. In this report and in the Stochastic Empirical Loading and Dilution Model (SELDM), the basin lagtime is defined as the time from the center-of-mass (centroid) of rainfall excess to the centroid of the corresponding runoff hydrograph, which is consistent with most USGS runoff studies. The total runoff duration can be estimated from the basin lagtime using a triangular-hydrograph recession ratio, which is the ratio of the duration of the falling limb to rising limb of the hydrograph. Thus, the basin lagtime and recession ratio statistics can be used to model the distribution of runoff flows during a runoff event.

Ten viable multiple linear regression equations were developed to estimate basin lagtimes from readily determined drainage basin properties using data published in 37 stormflow studies. These equations were developed to update the equations developed by Sauer and others (1983) to address concerns about the range of drainage areas and impervious fractions in the original dataset and the need for equations that do not use the basin development factor *(BDF)*, which is simple to estimate manually, but difficult to calculate using automated methods. The database used to develop basin lag equations in this report includes sites with drainage areas ranging from 0.000116 to 1,477 square miles and total impervious areas *(IMPERV)* ranging from 0 to 100 percent. Methods for determining the explanatory variables, using the equations, and estimating prediction limits for the basin lagtimes are documented. Several equations were developed to estimate basin lagtime, but two primary equations provided best-fit estimates. One equation uses the basin lag factor (*BLF*, which is the main channel length *(LENGTH)* divided by the square root of the main-channel slope (*CSL_10_85*)) and the USGS *BDF*, which is a function of the amount of channel modifications, storm sewering, and curb-and-gutter streets; this equation explains about 72 percent of the variability in the basin lagtime. A primary dataset, which included data from 493 sites that have values for *BLF, BDF,* and *IMPERV,* was used to develop the best-fit regression equation using the *BLF* and *BDF*. If automated methods are developed for estimating the *BDF* using GIS methods this equation may be readily implemented using automated methods such as the USGS Streamstats application. Currently (2012), however, such methods have not been developed. A secondary dataset, which included data from 896 sites that have values for *BLF* and *IMPERV,* was used to develop the best-fit regression equation using these two variables. This equation explains about 68 percent of the variability in the lagtime. Presumably, the *BDF* equation produces more accurate estimates of basin lagtime than equations based on *IMPERV*, but *IMPERV* can be estimated using readily available geographic-information files and existing automated data-processing techniques. Thus the equation developed using *IMPERV* may be better suited for implementation as an automated method. An example spreadsheet (ExamplePredictionIntervals.xls) is provided on the digital media accompanying this report to facilitate calculation of regression estimates and associated prediction limits using the all the equations derived in this study. More accurate and precise equations may be developed if lagtimes and basin characteristics are systematically determined as part of a systematic national initiative to characterize basin lagtimes in different regions of the country. The basin characteristics, basin lagtime values, and source citations for all the sites in the primary and secondary datasets are recorded in the database "Compilation.mdb" on the digital media accompanying this report.

Triangular-hydrograph recession ratio statistics were developed using instantaneous streamflow data from 32 USGS streamgages for basins draining parts of Massachusetts (the Massachusetts dataset) and 9 USGS streamgages in other areas of the United States (the non-Massachusetts dataset). Correlations between recession ratios and basin characteristics were weak, which precluded development of meaningful predictive equations. The minimum recession ratios are well characterized using a value of 1. The median of calculated values of the most probable value (MPV) of recession ratios is about 1.85, which compares well with the median ratio of 1.67 developed using average curvilinear flood hydrographs from USGS studies in different areas of the United States.

The median of calculated values of the maximum recession ratios is about 4.36. The weak correlations do not support quantitative predictions, but comparison of the selected basin characteristics with streamgages in the recession ratio dataset may inform the choice for selecting recession-ratio statistics that are greater than or less than these median values. The methods for fitting triangular hydrographs and for calculating recession-ratio statistics are described in detail and example spreadsheets are provided on the digital media accompanying this report so that the reader may calculate recession-ratio statistics using data from nearby, hydrologically similar streamgages.

Acknowledgments

References Cited

Amman, Alan, and Stone, A.L., 1991, How to read a topographic map and delineate a watershed—Appendix E *in* The method for the comparative evaluation of nontidal wetlands in New Hampshire: Natural Resources Conservation Service, accessed April 28, 2012, at http://www.nh.nrcs.usds.gov.

Anderson, D.G., 1970, Effects of urban development on floods in Northern Virginia: U.S. Geological Survey Water-Supply Paper 2001–C, 22 p.

Back, W.E., Boles, W.W., and Fry, G.T., 2000, Defining triangular probability distributions from historical cost data: Journal of Construction Engineering, v. 126, no. 1, p. 29–37.

Bailey, J.F., Thomas, W.O., Jr., Wetzel, K.L., and Ross, T.J., 1989, Estimation of flood-frequency characteristics and the effects of urbanization for streams in the Philadelphia, Pennsylvania Area: U.S. Geological Survey Water-Resources Investigations Report 87–4194, 71 p.

Becker, L.D., 1986, Techniques for estimating flood-peak discharges from urban basins in Missouri: U.S. Geological Survey Water-Resources Investigations Report 86–4322, 38 p.

Becker, L.D., 1990, Simulation of flood hydrographs for small basins in Missouri: U.S. Geological Survey Water-Resources Investigations Report 90–4045, 40 p.

Beighley, R.E., Kargar, Maryam, and He, Yiping, 2009, Effects of impervious area estimation methods on simulated peak discharges: Journal of Hydrologic Engineering, v. 14, no. 4, p. 388–398.

Benson, M.A., 1962, Factors influencing the occurrence of floods in a humid region of diverse terrain: U.S. Geological Survey Water-Supply Paper 1580–B, 62 p.

Bhat, Shirish, Hatfield, Kirk, Jacobs, J.M., and Graham, W.D., 2007, Relationships between military land use and storm-based indices of hydrologic variability: Ecological Indicators, v. 7, p. 553–564.

Bohman, L.R., 1990, Determination of flood hydrographs for streams in South Carolina—Volume 1, Simulation of flood hydrographs for rural watersheds in South Carolina: U.S. Geological Survey Water-Resources Investigations Report 89–4087, 53 p.

Bohman, L.R., 1992, Determination of flood hydrographs for streams in South Carolina—Volume 2, Estimation of peak-discharge frequency, runoff volumes, and flood hydrographs for urban watersheds: U.S. Geological Survey Water-Resources Investigations Report 92–4040, 79 p.

Burns, Douglas, Vitvar, Tomas, McDonnell, Jeffrey, Hassett, James, Duncan, Jonathan, and Kendall, Carol, 2005, Effects of suburban development on runoff generation in the Croton River basin, New York, USA: Journal of Hydrology, v. 311, p. 266–281.

Carter, R.W., 1961, Magnitude and frequency of floods in suburban areas, in short papers in the geologic and hydrologic sciences: U.S. Geological Survey Professional Paper 424–B, p. B9–B11.

Chow, V.T., 1962, Hydrologic determination of waterway areas for the design of drainage structures in small drainage basins: University of Illinois Engineering Experiment Station Bulletin No. 462, Urbana, Ill., 104 p.

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. 14.1–14.54.

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. 47–79.

Craig, G.S., Jr., and Rankl, J.G., 1978, Analysis of runoff from small drainage basins in Wyoming: U.S. Geological Survey Water-Supply Paper 2056, 70 p.

Crawford, C.G., 1991, Estimation of suspended-sediment rating curves and mean suspended-sediment loads: Journal of Hydrology, v. 129, p. 331–348.

Cristina, C.M., and Sansalone, J.J., 2003, Kinematic wave model of urban pavement rainfall-runoff subject to traffic loadings: Journal of Environmental Engineering, v. 129, no. 7, p. 629–636.

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.

Duan, Naihua, 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610.

Eagleson, P.E., 1962, Unit hydrograph characteristics for sewered areas: American Society of Civil Engineers, Proceedings, v. 88, no. HY 2, p. 1–25.

Espey, W.H., Jr., Morgan, C.W., and Masch, F.D., 1969, A study of some effects of urbanization on storm runoff from a small watershed: Texas Water Development Board Report 23, 110 p.

Falcone, J.A., Carlisle, D.M., Wolock, D.M., and Meador, M.R., 2010, GAGES—A stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecology, v. 91, p. 621, Ecological Archives E091-045-D1 accessed April 28, 2012, at http://esapubs.org/archive/ecol/E091/045/default.htm.

Fang, Xing, Cleveland, Theodore, Garcia, C.A., Thompson, David, and Malla, Ranjit, 2005, Literature review on timing parameters for hydrographs: Texas Department of Transportation Research Technical Report 0–4696–1, 83 p.

Federal Emergency Management Agency, 2001a, 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, accessed April 28, 2012, at http://www. floodmaps.fema.gov/listserv/nf_sep01.shtml.

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 2: National Flood Frequency Bulletin, Issue No. 6, accessed April 28, 2012, at http://www. floodmaps.fema.gov/listserv/nf_oct01.shtml.

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 3: National Flood Frequency Bulletin, Issue No. 7, accessed April 28, 2012, at http://www. floodmaps.fema.gov/listserv/nf_nov01.shtml.

Franklin, M.A., 1984, Magnitude and frequency of flood volumes for urban watersheds in Leon County, Florida: U.S. Geological Survey Water-Resources Investigations Report 84–4233, 20 p.

Franklin, M.A., and Losey, G.T., 1984, Magnitude and frequency of flood volumes for urban streams in Leon County, Florida: U.S. Geological Survey Water-Resources Investigations Report 84–4004, 37 p.

Gamble, C.R., 1989, Techniques for simulating flood hydrographs and estimating flood volumes for ungaged basins in east and west Tennessee: U.S. Geological Survey Water-Resources Investigations Report 89–4076, 40 p.

Gilroy, E.J., Hirsch, R.M., and Cohn, T.A., 1990, Mean-square error of regression-based constituent transport estimates: Water Resources Research, v. 26, no. 9, p. 2069–2077.

Gilroy, E.J., and Tasker, G.D., 1989, Multicollinearity and influential observations in hydrologic model selection, *in* Berk, Kenneth, and Malone, Linda, eds., Computing Science and Statistics—Proceedings of the 21st Symposium on the Interface: Alexandria, Va., American Statistical Association, p. 350–354.

Granato, G.E., 2010, Methods for development of planning-level estimates of stormflow at unmonitored sites in the conterminous United States: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, FHWA–HEP–09–005, 90 p. with CD-ROM.

Guo, Yiping, and Adams, B.J., 1998, 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.

Harley, B.M., 1978, Research on the effects of urbanization on small streamflow quantity: Federal Highway Administration, FHWA–RD–78–88, 56 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., Helsel, D.R., Gilroy, E.J., and Cohn, T.A., 1993, Chapter 17—Statistical analysis of hydrologic data, in Maidment, D.R., ed., Handbook of hydrology: New York, McGraw-Hill Book Company, p. 17.1–17.55.

Holnbeck, S.R., and Parrett, Charles, 1996, Procedures for estimating unit hydrographs for large floods at ungaged sites in Montana: U.S. Geological Survey Water-Supply Paper 2420, 60 p.

Hood, M.J., Clausen, J.C., and Warner, G.S., 2007, Comparison of stormwater lagtimes for low impact and traditional residential development: Journal of the American Water Resources Association, v. 43, no. 4, p. 1036–1046.

Hubbard, L.E., 1992, Urban storm runoff in the Roseburg area, Oregon, as related to urban flood characteristics of the Willamette Valley: U.S. Geological Survey Water-Resources Investigations Report 91–4063, 28 p.

Inman, E.J., 1986, Simulation of flood hydrographs for Georgia streams: U.S. Geological Survey Water-Resources Investigations Report 86–4004, 41 p.

Inman, E.J., 2000, Lagtime relations for urban streams in Georgia: U.S. Geological Survey Water-Resources Investigations Report 00–4049, 12 p.

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, Inc., 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.

Kirpich, Z.P., 1940, Time of concentration of small agricultural watersheds: Civil Engineering, v. 10, no. 6, p. 362.

Koltun, G.F., and Whitehead, M.T., 2002, Techniques for estimating selected streamflow characteristics of rural, unregulated streams in Ohio: U.S. Geological Survey Water-Resources Investigations Report 02–4068, 50 p.

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

Laenen, Antonius, 1980, Storm runoff as related to urbanization in the Portland, Oregon-Vancouver, Washington area: U.S. Geological Survey Open-File Report 80–689, 71 p.

Laenen, Antonius, 1983, Storm runoff as related to urbanization based on data collected in Salem and Portland, and generalized for the Willamette Valley, Oregon: U.S. Geological Survey Water-Resources Investigations Report 83–4143, 88 p.

Langbein, W.B., and others, 1947, Topographic characteristics of drainage basins: U.S. Geological Survey Water-Supply Paper 968–C, p. 125–157.

Lee, J.G., and Heaney J.P., 2003, Estimation of urban imperviousness and its impacts on storm water systems: Journal of Water Resources Planning and Management, v. 129, no. 5, p. 419–426.

Leopold, L.B., 1968, Hydrology for urban land planning—A guidebook on the hydrologic effects of urban land use: U.S. Geological Survey Circular 554, 18 p.

Leopold, L.B., 1991, Lagtimes for small drainage basins: Catena, v. 18, p. 157–171.

Linsley, R.K., Jr., Kohler, M.A., and Paulhus, J.L.H., 1975, Hydrology for Engineers, 2 ed.: New York, McGraw-Hill, Inc., 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 96–250, 51 p.

Liu, Yang, Soonthornnonda, Puripus, Li, Jin, and Christensen, E.R., 2011, Stormwater runoff characterized by GIS determined source areas and runoff volumes: Environmental Management, v. 47, p. 201–217.

Martens, L.A., 1968, Hydrologic effects of urban growth—Flood inundation and effects of urbanization in metropolitan Charlotte, North Carolina: U.S. Geological Survey Water-Supply Paper 1591–C, 44 p.

Martin, G.R., Ruhl, K.J., Moore, B.L., and Rose, M.F., 1997, Estimation of peak-discharge frequency of urban streams in Jefferson County, Kentucky: U.S. Geological Survey Water-Resources Investigations Report 97–4219, 40 p.

Masch, F.D., 1984, Hydrology: Federal Highway Administration Report FHWA-IP-84-015, 342 p.

Mason, R.R., and Bales, J.D., 1996, Estimating flood hydrographs for urban basins in North Carolina: U.S. Geological Survey Water-Resources Investigations Report 96–4085, 19 p.

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

McCuen, R.H., Wong, S.L., and Rawls, W.J., 1984, Estimating urban time of concentration: Journal of Hydraulic Engineering, v. 110, no. 7, p. 887–904.

McEnroe, B.M., and Zhao, H., 1999, Lagtimes and peak coefficients for small rural watersheds in Kansas: Kansas Department of Transportation Report No. K-TRAN: KU–98-1, 37 p.

McEnroe, B.M., and Zhao, H., 2001, Lagtimes of urban and developing watersheds in Johnson County, Kansas: Kansas Department of Transportation Report No. K-TRAN: KU–99–5, 30 p.

McGill, Robert, Tukey, J.W., Larsen, W.A., 1978, Variations of box plots: The American statistician, v. 32, no. 1, p. 12–16.

Meierdiercks, K.L., Smith, J.A., Baeck, M.L., and Miller, A.J., 2010, Analyses of urban drainage network structure and its impact on hydrologic response: Journal of the American Water Resources Association, v. 46, no. 5, p. 932–943.

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.

Neely, B.L., Jr., 1984, Flood frequency and storm runoff of urban areas of Memphis and Shelby County, Tennessee: U.S. Geological Survey Water-Resources Investigations Report 84–4110, 51 p.

Neely, B.L., Jr., 1989, Estimating flood hydrographs for Arkansas streams: U.S. Geological Survey Water-Resources Investigations Report 89–4109, 19 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, Inc., p. 21.1–21.97.

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.

Rao, R.A., and Delleur, J.W., 1974, Instantaneous unit hydrographs, peak discharges, and time lags in urban basins: Hydrological Sciences Journal, v. 19, no. 2, p. 185–198.

Richards, P.L., and Brenner, A.J., 2004, Delineating source areas for runoff in depressional landscapes: Implications for hydrologic modeling: Journal of Great Lakes Research, v. 30, no. 1, p. 9–21.

Ries III, K.G., 2007, The national streamflow statistics program—A computer program for estimating streamflow statistics for ungaged sites: U.S. Geological Survey Techniques and Methods 4–A6, 37 p., at http://pubs.usgs. gov/tm/2006/tm4a6/.

Ries, K.G., III, Guthrie, J.G., Rea, A.H., Steeves, P.A., and Stewart, D.W., 2008, StreamStats—A water resources Web application: U.S. Geological Survey Fact Sheet 2008–3067, 6 p., at http://pubs.usgs.gov/fs/2008/3067/.

Robbins, C.H., 1984, Synthesized flood frequency for small urban streams in Tennessee: U.S. Geological Survey Water-Resources Investigations Report 84–4182, 24 p.

Robbins, J.C., and Pope, B.F., 1996, Estimation of flood-frequency characteristics of small urban streams in North Carolina: U.S. Geological Survey Water-Resources Investigations Report 96–4084, 21 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.

Schulz, E.F., and Lopez, O.G., 1974, Determination of urban watershed response time: Fort Collins, Colo., Colorado State University, Hydrology Paper Number 71, 80 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.

Sherwood, J.M., 1994, Estimation of peak-frequency relations, flood hydrographs, and volume-duration-frequency relations of ungaged small urban streams in Ohio: U.S. Geological Survey Water-Supply Paper 2432, 52 p.

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.

Simas, M.J., and Hawkins, R.H., 2011, Lagtime characteristics for small watersheds in the U.S.: United States Department of Agriculture, Natural Resources Conservation Service, West National Technology Support Center, 7 p., accessed April 28, 2012, at http://www.nrcs.usda.gov/Internet/FSE_ DOCUMENTS/stelprdb1044199.pdf.

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.

Smith, K.P., and Granato, G.E., 2010, Quality of stormwater runoff discharged from Massachusetts highways, 2005–2007: U.S. Geological Survey Scientific Investigations Report 2009–5269, 198 p. with CD–ROM, at http://pubs.usgs.gov/sir/2009/5269/.

Spencer, D.W., and Alexander, T.W., 1978, Technique for estimating the magnitude and frequency of floods in St. Louis County, Missouri: U.S. Geological Survey Water-Resources Investigations Report 78–139, 23 p.

Straub, T.D., Melching, C.S., and Kocher, K.E., 2000, Equations for estimating Clark unit-hydrograph parameters for small rural watersheds in Illinois: U.S. Geological Survey Water-Resources Investigations Report 00–4184, 30 p.

Strecker, Eric, Mayo, Lynn, Quigley, Marcus, and Howell, James, 2001, Guidance manual for monitoring highway runoff water quality: Washington, D.C., U.S. Department of Transportation, Federal Highway Administration, FHWA-EP-01-022, 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.

Subramanya, K., 1984, Engineering hydrology, 2 ed.: New Delhi, Tata McGraw-Hill, 391 p.

Sutherland, R.C., 1988, Methodology for estimating lagtime of natural, partially urbanized and urban watershed based on published USGS data for watersheds throughout the metropolitan areas of Portland and Salem, Oregon: Oregon City Drainage Master Plan, appendix C, 10 p.

U.S. Environmental Protection Agency, 2000, Storm water phase II compliance assistance guide: U.S. Environmental Protection Agency EPA 833–R–00–002, 96 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., accessed April 28, 2012, at http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm.

U.S. Geological Survey, 1977, Physical basin characteristics for hydrologic analyses—Chapter 7 *in* National handbook for recommended methods for water-data acquisition: Reston, Va., U.S. Geological Survey, Office of Water Data Coordination, 38 p.

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, 2011, StreamStats user instructions: U.S. Geological survey Web site, accessed April 28, 2012, at http://water.usgs.gov/osw/streamstats/.

U.S. Geological Survey, 2012a, Instantaneous data archive: U.S. Geological survey Web site, accessed January 1, 2012, at http://ida.water.usgs.gov/ida/.

U.S. Geological Survey, 2012b, USGS water data for the Nation: U.S. Geological survey Web site, accessed January 1, 2012, at http://waterdata.usgs.gov/nwis/.

Veenhuis, J.E., and Gannett, D.G., 1986, The effects of urbanization on floods in the Austin metropolitan area, Texas: U.S. Geological Survey Water-Resources Investigations Report 86–4069, 66 p.

Viessman, Warren, Jr., 1968, Runoff estimation for very small drainage areas: Water Resources Research, v. 4, no. 1, p. 87–93.

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.

Watt, W.E., and Chow, K.C.A., 1985, A general expression for basin lagtime: Canadian Journal of Civil Engineering, v. 12, p. 291–300.

Weaver, J.C., 2003, Methods for estimating peak discharges and unit hydrographs for streams in the city of Charlotte and Mecklenburg County, North Carolina: U.S. Geological Survey Water-Resources Investigations Report 03–4108, 50 p.

Wetlands Subcommittee of the Federal Geographic Data Committee, 2009, Wetlands Mapping Standard: Federal Geographic Data Committee Document Number FGDC-STD-015-2009, 35 p.

Wibben, H.C., 1976, Effects of urbanization on flood characteristics in Nashville-Davidson County, Tennessee: U.S. Geological Survey Water-Resources Investigations Report 76–121, 33 p.

Williams, K.K., 1980, Calibration of federal agency storm water runoff quantity models for Oklahoma City: Proceedings of the Oklahoma Academy of Science, v. 60, p. 75–81, accessed April 28, 2012, at http://digital.library.okstate.edu/OAS/oas_pdf/v60/p75_81.pdf.

This page has been left blank intentionally.