United States Department of the Interior

U.S. GEOLOGICAL SURVEY

IN REPLY REFER TO:

October 28, 2008

MEMORANDUM

To: Whom it may concern

From: U.S. Geological Survey, Massachusetts/Rhode Island Water Science Center

Date: October 28, 2008

Re: ERRATA in SIR 2006-5031, A Revised Logistic-Regression Equation and an Automated Procedure for Mapping the Probability of a Stream Flowing Perennially in Massachusetts, by Gardner C. Bent and Peter A. Steeves

---

After the publication of U.S. Geological Survey Scientific Investigations Report 2006-5031, “A Revised Logistic-Regression Equation and an Automated Procedure for Mapping the Probability of a Stream Flowing Perennially in Massachusetts,” by Gardner C. Bent and Peter A. Steeves, the following errors were found:

On page 23, equation #3 needs to be changed to:

$$
\begin{aligned}
\text{variance} &= 0.372729 + 0.034308\left(\ln\left(x_1\right)\right)^2 + 0.000032\left(x_2\right)^2 + 0.000057\left(x_3\right)^2 \\
&+0.102893\left(x_4\right)^2 + 2\left(\ln\left(x_1\right)\right)\left(0.048903\right) + 2\left(x_2\right)\left(-0.001272\right) + 2\left(x_3\right)\left(-0.004189\right) \\
&+2\left(x_4\right)\left(0.018564\right) + 2\left(\ln\left(x_1\right)\right)\left(x_2\right)\left(0.000131\right) + 2\left(\ln\left(x_1\right)\right)\left(x_3\right)\left(-0.000306\right) \\
&+2\left(\ln\left(x_1\right)\right)\left(x_4\right)\left(0.003397\right) + 2\left(x_2\right)\left(x_3\right)\left(0.000011\right) + 2\left(x_2\right)\left(x_4\right)\left(0.000163\right) \\
&+2\left(x_3\right)\left(x_4\right)\left(-0.000800\right)
\end{aligned}
$$

On page 26, the example calculation of the variance needs to be changed to:

$$
\begin{aligned}
\text{variance} &= 0.372729 + 0.034308\left(\ln\left(0.53\right)\right)^2 + 0.000032\left(0.00\right)^2 + 0.000057\left(78.76\right)^2 \\
&+0.102893\left(1\right)^2 + 2\left(\ln\left(0.53\right)\right)\left(0.048903\right) + 2\left(0.00\right)\left(-0.001272\right) + 2\left(78.76\right)\left(-0.004189\right) \\
&+2\left(1\right)\left(0.018564\right) + 2\left(\ln\left(0.53\right)\right)\left(0.00\right)\left(0.000131\right) + 2\left(\ln\left(0.53\right)\right)\left(78.76\right)\left(-0.000306\right) \\
&+2\left(\ln\left(0.53\right)\right)\left(1\right)\left(0.003397\right) + 2\left(0.00\right)\left(78.76\right)\left(0.000011\right) + 2\left(0.00\right)\left(1\right)\left(0.000163\right) \\
&+2\left(78.76\right)\left(1\right)\left(-0.000800\right) \\
&= 0.06
\end{aligned}
$$

A Revised Logistic Regression Equation and an Automated Procedure for Mapping the Probability of a Stream Flowing Perennially in Massachusetts

By Gardner C. Bent and Peter A. Steeves

Part 1.

A Revised Logistic Regression Equation for Estimating the Probability of a Stream Flowing Perennially in Massachusetts

By Gardner C. Bent

Part 2.

An Automated Procedure for Mapping Perennially Flowing Streams

By Peter A. Steeves, Gardner C. Bent, and Jennifer R. Hill (Horizon Systems Corporation)

In cooperation with the
Massachusetts Department of Environmental Protection
Bureau of Resource Protection
Wetlands and Waterways Program

Scientific Investigations Report 2006–5031

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

U.S. Department of the Interior
Gale A. Norton, Secretary

U.S. Geological Survey
P. Patrick Leahy, Acting Director

U.S. Geological Survey, Reston, Virginia: 2006

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

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

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

Contents

Figures

Tables

Conversion Factors, Datums, and Abbreviations

Multiply	By	To obtain
	Length	
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
acre	0.004047	square kilometer (km^2)
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C=(°F-32)/1.8$$

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

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

ABBREVIATIONS

CD-ROM	Compact Disk–Read Only Memory
CMR	Code of Massachusetts Regulations
DEMs	Digital Elevation Models
GIS	Geographic Information Systems
MDEP	Massachusetts Department of Environmental Protection
NHD	National Hydrography Dataset
Regulations	Commonwealth of Massachusetts Rivers Protection Act, Chapter 258 of the Acts of 1996
Revised Regulations	Commonwealth of Massachusetts Rivers Protection Act, Chapter 258 of the Acts of 1996 (revised December 20, 2002)
USGS	U.S. Geological Survey

A Revised Logistic Regression Equation and an Automated Procedure for Mapping the Probability of a Stream Flowing Perennially in Massachusetts

By Gardner C. Bent and Peter A. Steeves

Abstract

A revised logistic regression equation and an automated procedure were developed for mapping the probability of a stream flowing perennially in Massachusetts. The equation provides city and town conservation commissions and the Massachusetts Department of Environmental Protection a method for assessing whether streams are intermittent or perennial at a specific site in Massachusetts by estimating the probability of a stream flowing perennially at that site. This information could assist the environmental agencies who administer the Commonwealth of Massachusetts Rivers Protection Act of 1996, which establishes a 200-foot-wide protected riverfront area extending from the mean annual high-water line along each side of a perennial stream, with exceptions for some urban areas. The equation was developed by relating the observed intermittent or perennial status of a stream site to selected basin characteristics of naturally flowing streams (defined as having no regulation by dams, surface-water withdrawals, ground-water withdrawals, diversion, wastewater discharge, and so forth) in Massachusetts. This revised equation differs from the equation developed in a previous U.S. Geological Survey study in that it is solely based on visual observations of the intermittent or perennial status of stream sites across Massachusetts and on the evaluation of several additional basin and land-use characteristics as potential explanatory variables in the logistic regression analysis. The revised equation estimated more accurately the intermittent or perennial status of the observed stream sites than the equation from the previous study.

Stream sites used in the analysis were identified as intermittent or perennial based on visual observation during low-flow periods from late July through early September 2001. The database of intermittent and perennial streams included a total of 351 naturally flowing (no regulation) sites, of which 85 were observed to be intermittent and 266 perennial. Stream sites included in the database had drainage areas that ranged from 0.04 to 10.96 square miles. Of the 66 stream sites with drainage areas greater than 2.00 square miles, 2 sites were intermittent and 64 sites were perennial. Thus, stream sites with drainage areas greater than 2.00 square miles were assumed to flow perennially, and the database used to develop the logistic regression equation included only those stream sites with drainage areas less than 2.00 square miles. The database for the equation included 285 stream sites that had drainage areas less than 2.00 square miles, of which 83 sites were intermittent and 202 sites were perennial.

Results of the logistic regression analysis indicate that the probability of a stream flowing perennially at a specific site in Massachusetts can be estimated as a function of four explanatory variables: (1) drainage area (natural logarithm), (2) areal percentage of sand and gravel deposits, (3) areal percentage of forest land, and (4) region of the state (eastern region or western region). Although the equation provides an objective means of determining the probability of a stream flowing perennially at a specific site, the reliability of the equation is constrained by the data used in its development. The equation is not recommended for (1) losing stream reaches or (2) streams whose ground-water contributing areas do not coincide with their surface-water drainage areas, such as many streams draining the Southeast Coastal Region—the southern part of the South Coastal Basin, the eastern part of the Buzzards Bay Basin, and the entire area of the Cape Cod and the Islands Basins. If the equation were used on a regulated stream site, the estimated intermittent or perennial status would reflect the natural flow conditions for that site.

An automated mapping procedure was developed to determine the intermittent or perennial status of stream sites along reaches throughout a basin. The procedure delineates the drainage area boundaries, determines values for the four explanatory variables, and solves the equation for estimating the probability of a stream flowing perennially at two locations on a headwater (first-order) stream reach—one near its confluence or end point and one near its headwaters or start point. The automated procedure then determines the intermittent or perennial status of the reach on the basis of the calculated probability values and a probability cutpoint (a stream is considered to flow perennially at a cutpoint of 0.56 or greater for this study) for the two locations or continues to loop upstream or downstream between locations less than and greater than the cutpoint of 0.56 to determine the transition point from an

intermittent to a perennial stream. If the first-order stream reach is determined to be intermittent, the procedure moves to the next downstream reach and repeats the same process. The automated procedure then moves to the next first-order stream and repeats the process until the entire basin is mapped.

A map of the intermittent and perennial stream reaches in the Shawsheen River Basin is provided on a CD-ROM that accompanies this report. The CD-ROM also contains ArcReader 9.0, a freeware product, that allows a user to zoom in and out, set a scale, pan, turn on and off map layers (such as a USGS topographic map), and print a map of the stream site with a scale bar. Maps of the intermittent and perennial stream reaches in Massachusetts will provide city and town conservation commissions and the Massachusetts Department of Environmental Protection with an additional method for assessing the intermittent or perennial status of stream sites.

Introduction

The Commonwealth of Massachusetts Rivers Protection Act, Chapter 258 of the Acts of 1996 (The Commonwealth of Massachusetts, 1996), specifies that riverfront areas be protected on all rivers that flow perennially. The riverfront area is defined in 310 Code of Massachusetts Regulations (CMR) 10.58(2)(a) (hereafter referred to as the Regulations) (Massachusetts Department of Environmental Protection, 2002a, p. 393–402) as the 200-ft-wide area extending along the length of each side of perennial streams from the mean annual high-water line (determined from bankfull field indicators) on each side of perennial streams. Exceptions to the Regulations are provided for some urban areas. Streams that do not flow year round, intermittent streams, have no jurisdictional riverfront area along the stream. City or town conservation commissions are charged with administering the Regulations by determining the intermittent or perennial status of a stream site and by regulating work in the riverfront areas. The Massachusetts Department of Environmental Protection (MDEP) addresses appeals of decisions made by city or town conservation commissions concerning the intermittent or perennial status of stream sites. The logistic regression equation provides these agencies with an additional method for assessing the status of stream sites in Massachusetts.

The Regulations define a river as any natural flowing body of water that discharges into an ocean, lake, pond, or another river, and which flows throughout the year (Massachusetts Department of Environmental Protection, 2002a, p. 394). By this definition, perennial streams are rivers, but intermittent streams are not. When an intermittent stream is not flowing, surface water may be present in isolated pools or be absent. Rivers start at the point where an intermittent stream becomes perennial or at the point where a stream flows perennially from a spring, pond, or lake.

The revised Regulations of December 20, 2002 (Massachusetts Department of Environmental Protection, 2002b, p. 317–320; Massachusetts Department of Environmental Protection, 2002a, p. 394–395), specify that U.S. Geological Survey (USGS) topographic maps, or more recent maps provided by MDEP, will continue to be used for initial review of the intermittent or perennial status of a stream. Streams depicted as perennial on USGS topographic maps or more recent maps provided by MDEP will be classified as perennial. A stream site depicted as perennial, however, can be reclassified as intermittent with direct observations of no flow during any four days of any consecutive 12-month period. These observations cannot be made during a period of extended drought or on a stream measurably affected by withdrawals, impoundments, or other anthropogenic flow reductions or diversions. The definition of “extended drought” was amended to include the time periods during which the Massachusetts Drought Management Task Force declared an index level of “advisory, watch, warning, or emergency” (Massachusetts Executive Office of Environmental Affairs and Massachusetts Emergency Management Agency, 2001).

The revised Regulations state that streams depicted as intermittent, or those not shown as a stream on USGS topographic maps or more recent maps provided by MDEP, will be mainly classified on the basis of the drainage area size upstream from the stream site. If an intermittent stream site’s upstream drainage area is greater than or equal to 1.00 mi^2, the stream site will be classified as perennial. If an intermittent stream site’s upstream drainage area is less than 0.50 mi^2, the stream site will be classified as intermittent. If an intermittent stream site’s upstream drainage area is greater than or equal to 0.50 mi^2 and less than 1.00 mi^2, the stream site will be classified intermittent, with two exceptions. First, if the 99-percent flow duration estimated from low-flow statistics regression equations by the World Wide Web application STREAMSTATS[1] (Ries and others, 2000) at the stream site is greater than or equal to 0.01 ft^3/s, then the stream will be classified as perennial. Second, if the streamflow at the stream site cannot be estimated with STREAMSTATS (the stream is not shown on a USGS topographic map; or the stream is in the Buzzards Bay Basin, Cape Cod Basin, Islands Basin, North Coastal Basin, or Taunton River Basin) (fig. 1) and more than 75 percent of the drainage area comprises stratified deposits, then the stream will be classified as perennial.

In a previous study, Bent and Archfield (2002) developed a logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts. The equation was developed by relating the intermittent or perennial classification of a site on a naturally flowing stream (no regulations by dams, surface-water withdrawals, ground-water withdrawals, diversion, wastewater discharge, and so forth) to selected

[1] STREAMSTATS is a World Wide Web application that allows a user to estimate low-flow statistics for Massachusetts streams. The application allows a user to select a site on a stream. Then the application determines the drainage-basin boundary and basin characteristics (explanatory variables) for the selected site, and solves a number of regression equations for selected low-flow statistics. For more information on STREAMSTATS, please see Ries and others (2000).

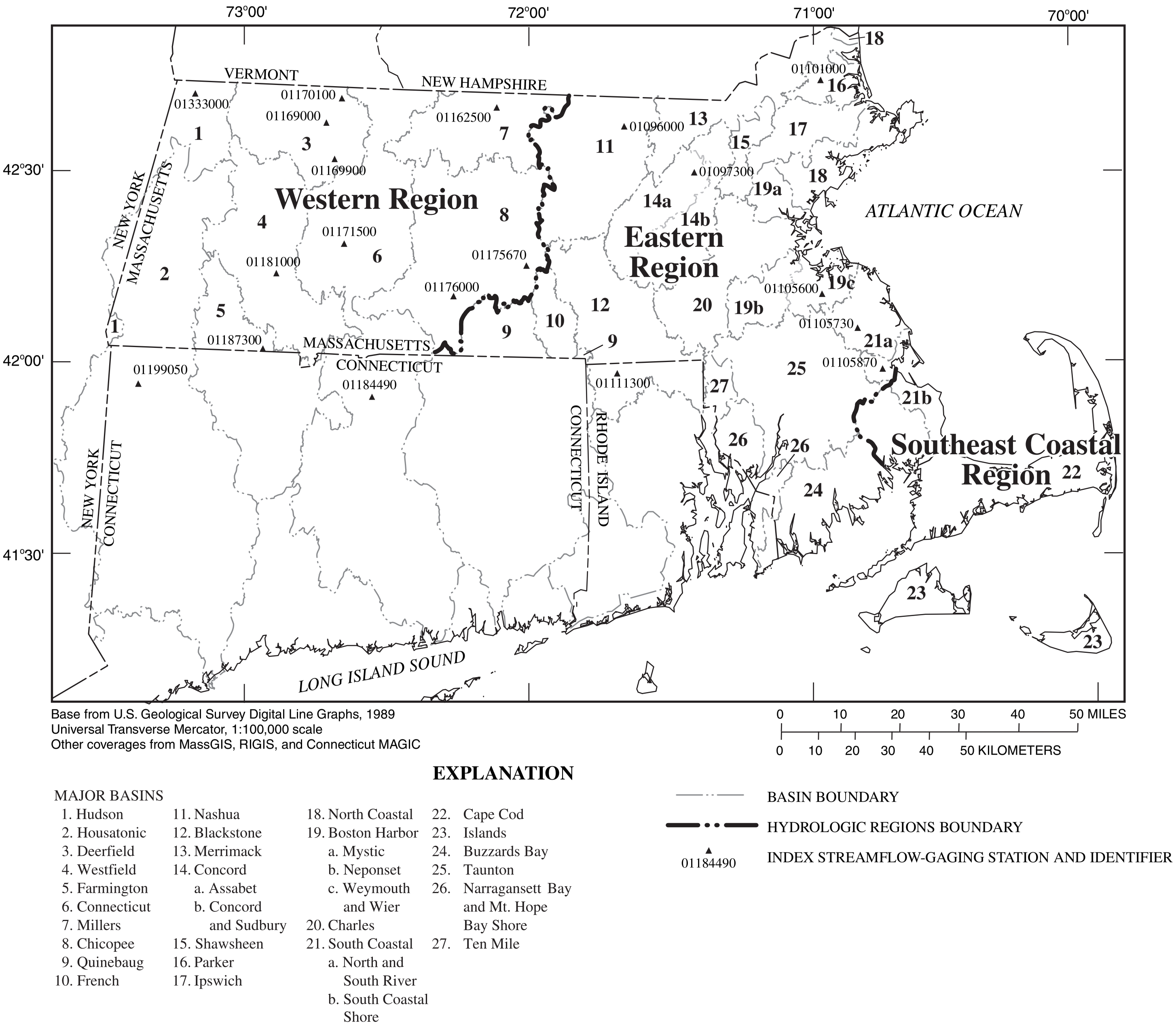


Figure 1. Locations of long-term continuous (index) streamflow-gaging stations used to estimate low-flow durations at nearby field-visited stream sites in the development and verification of the logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts, late July through early September 2001.

basin characteristics (drainage area, drainage density, areal percentage of stratified-drift deposits, and mean basin slope) and a location identifier (South Coastal Basin or the remainder of the state) (fig. 1). Stream sites used in the analysis data set were classified as intermittent or perennial on the basis of review of historical streamflow measurements at USGS data-collection sites throughout the state and on visual observation of sites in the South Coastal Basin, southeastern Massachusetts. Because of the limitations with this data set, Bent and Archfield (2002) suggested that the equation might be improved through visual observations at stream sites throughout the state. Additionally, they suggested that the equation might be improved by testing additional basin characteristics as potential explanatory variables in the analysis. Thus, the USGS, in cooperation with the MDEP, did a study in 2001–02 to develop a revised logistic regression equation to estimate the probability of a stream flowing perennially in Massachusetts more accurately and to test an automated procedure to map the intermittent and perennial stream reaches in the Shawsheen River Basin, northeastern Massachusetts (fig. 1).

Purpose and Scope

The development and application of a revised logistic regression equation to estimate the probability of a stream flowing perennially at a specific site with a drainage area less than or equal to 2.00 mi^2 in Massachusetts are described in the first part of this report. The equation is based on field observations throughout the state, except on Cape Cod and the Islands, during late July through early September 2001. Basin characteristics used in the analysis and digital data layers are described. Limitations of the logistic regression are discussed and areas for further study are presented. An automated procedure for mapping intermittent and perennial stream reaches is described in the second part of this report. The procedure works in conjunction with the logistic regression equation. An application of this procedure for the Shawsheen River Basin in northeastern Massachusetts is given on the CD-ROM.

Definitions of Intermittent and Perennial Streams

Langbein and Iseri (1960) defined intermittent and perennial streams as follows: "Intermittent or seasonal—one which flows only at certain times of the year when it receives water from springs or from surface sources such as melting snow in mountainous areas. Perennial—one which flows continuously." Meinzer's (1923) definitions of intermittent and perennial streams add more specific details to the definitions by Langbein and Iseri (1960).

> A spring-fed intermittent, or stretch of a stream, is one that flows only at certain times when it receives water from springs. The intermittent character of streams is generally due to fluctuations in the water table whereby the stream channel stands a part of time below and at part of the time above the water table. This is the ordinary type of intermittent stream. Perennial streams are generally fed in part by springs, and their upper surfaces generally stand lower than the water table in the localities through which they flow.

Several states, including Massachusetts, base regulations on a stream's ephemeral (a stream that flows in direct response to precipitation and whose channel is at all times above the water table), intermittent, or perennial status. For example, Paybins (2003) reports that, in West Virginia, valley-fill material from surface mining of coal generally can only be placed in ephemeral streams and not within 100 ft of intermittent and perennial streams, according to a U.S. District court's interpretation of Surface Mining Control and Reclamation Act and Clean Water Act Regulations. The state of Connecticut uses the USGS topographic maps' depiction of intermittent and perennial streams in decisions related to wellhead protection areas (C.R. Fitting, Connecticut Department of Environmental Protection, written commun., 2003). The North Carolina Water-Supply Protection Act instituted 30- and 100-ft buffers for low- and high-density development, respectively, along both sides of perennial streams in basins with water-supply withdrawals (North Carolina Department of Environment and Natural Resources, Division of Water Quality, Water Supply Protection Program, 2004). Fairfax County, Virginia is instituting 100-ft buffers along both sides of perennial streams as Resource Protection Areas required by the Chesapeake Bay Preservation Area Designation and Management Regulations (Virginia Department of Conservation and Recreation Chesapeake Bay Local Assistance, 2002). In some cases, the Regulations are based upon a stream's intermittent or perennial status as depicted on USGS topographic maps. Leopold (1994, p. 227–230) states that the depiction of intermittent and perennial streams on USGS topographic maps is not based solely on hydrologic criteria, but also specific topographic instructions to past USGS cartographers (U.S. Geological Survey, 1980). Thus, several states have started efforts to determine the intermittent or perennial status of stream reaches more accurately.

Related Intermittent and Perennial Stream Studies

Several studies have been done to define intermittent and perennial stream reaches as well as the transition points between intermittent and perennial stream reaches on the basis of basin characteristics, biological characteristics, climatic characteristics, physiographic provinces, and so forth. For example, Paybins (2003) found that the median drainage area upstream of the transition point from an intermittent to a perennial stream reach was about 0.06 mi^2 for 20 stream sites in southwestern West Virginia, where mean annual precipitation ranges from about 44 to 48 in. In the state of

Washington, Palmquist (2003) identified the transition points between intermittent and perennial stream reaches for three geographic regions (coastal, western, and eastern area of the state), and found that the median drainage area upstream of the transition points was about 0.09 mi^2 for 41 stream sites where the mean annual precipitation is less than 60 in. (Robert Palmquist, Northwest Indian Fisheries Commission, written commun., 2004). In the Robinson National Forest of eastern Kentucky, Fritz (2004) found that the drainage area upstream of the transition points ranged from 0.25 to 0.35 mi^2 for three stream sites where the mean annual precipitation is about 47 in. (K.M. Fritz, U.S. Environmental Protection Agency, written commun., 2004). Fritz (2005) also identified drainage area as a primary factor, and bankfull width, maximum pool depth, and channel entrenchment as secondary factors in classifying ephemeral, intermittent, and perennial streams. In northern Georgia, Rivenbark and Jackson (2004) found that drainage area upstream of the transition points ranged from about 0.02 to 0.05 mi^2, with an average of 0.03 mi^2, for 30 sites where mean annual precipitation ranges from about 60 to 80 in. (Spatial Climate Analysis Service—Oregon Climate Service, 2004).

In Fairfax County, Virginia, perennial streams are being mapped to address concerns that all perennial streams were not being protected under County Code (Fairfax County, Virginia, Public Works and Environmental Services, 2004). The Fairfax County intermittent/perennial stream field-identification protocol is based on a combination of hydrologic, physical, and biologic characteristics of the stream, and is similar to the field-identification protocol developed by the North Carolina Division of Water Quality (North Carolina Department of Environment and Natural Resources, Division of Water Quality, Wetlands/401, Water Quality Certification Unit, 2004). A study in North Carolina on the Upper Neuse River Basin using soils data from the Soil Survey Geographic (SSURGO) database (U.S. Department of Agriculture, Natural Resources Conservation Service, Soil Survey Division, 2004), land-use and land-cover data, elevation data from the National Elevation Dataset (NED) (U.S. Geological Survey, 2005a), LIght Detection And Ranging (LIDAR) data, hydrologic data from National Hydrography Dataset (NHD) (U.S. Geological Survey, 2005b), and Elevation Derivatives for National Application (EDNA) (U.S. Geological Survey, 2005c) to develop stream-classification strategies (Restrepo and Waisanen, 2004a and 2004b). Another study in North Carolina is identifying a predictive model for determining first-order streams using LIDAR data (Thomas Colson, North Carolina State University, written commun., 2004). A study in eastern Kentucky to define ephemeral, intermittent, and perennial stream reaches more accurately through measurements of ground-water levels in streambed piezometers, bankfull stream-geometry characteristics, and basin-characteristics data to develop a regression model to predict streamflow periodicity is described by Kolka and Stringer (2000). The study by Kolka and Stringer is being done as a result of requirements for streamside-management zones, which are a component of forestry best management practices. Idaho is developing maps of intermittent and perennial streams on the basis of an automated mapping procedure and a regression equation that estimates the low-flow statistic 7Q2 (7-day 2-year low flow). The automated mapping procedure solves the 7Q2 for headwater stream reaches and maps reaches with the 7Q2 less than 0.1 ft^3/s as intermittent and reaches with the 7Q2 equal to or greater than 0.1 ft^3/s as perennial. This work is being done by the USGS in cooperation with the Idaho Department of Environmental Quality (Jon Hortness, U.S. Geological Survey, written commun., 2005). Vermont is developing a predictive model that is similar to the model developed in this study and that uses observed intermittent and perennial stream status and basin characteristics, and is integrating the model with an automated procedure to map intermittent and perennial stream reaches. This work is being done by the USGS in cooperation with the Vermont Center for Geographic Information (Scott Olson, U.S. Geological Survey, written commun., 2005). Additional work in Massachusetts related to intermittent and perennial streams involves the evaluation of the hydrology, habitat characteristics, biological-community composition and structure of the tributaries of six perennial streams in central Massachusetts as the streams change from ephemeral to intermittent to perennial (Harvard Forest, Harvard University, 2004). Several other studies reported by the North American Benthological Society (2002, 2003, 2005a, and 2005b) are focused on the ecological functions of intermittent streams.

Part 1. A Revised Logistic Regression Equation for Estimating the Probability of a Stream Flowing Perennially in Massachusetts

By Gardner C. Bent

Description of Study Area

The geography, climate, and surficial geology of a basin upstream of a selected stream site can affect whether the stream at that site will be intermittent or perennial. In Massachusetts these factors, particularly the extent and type of surficial deposits, affect streamflow characteristics.

Massachusetts encompasses 8,093 mi^2 in the northeastern United States (fig. 1). Altitudes range from sea level in coastal areas to 3,491 ft above sea level in the northwest. Altitudes generally increase from eastern to western Massachusetts. The climate in Massachusetts is humid, with average annual precipitation ranging from about 40 to 45 in. in eastern Massachusetts to about 40 to 50 in. in western Massachusetts, where higher altitudes may cause orographic effects. Average annual temperature is about 50˚F in eastern Massachusetts and about 45˚F in western Massachusetts.

Surficial deposits that overlie bedrock in most of Massachusetts were deposited mainly during the last glacial period, but can include areas of recent flood-plain alluvium deposits. In this report, these surficial deposits are classified as either till (which includes till or bedrock, sandy till over sand, and end-moraine deposits) or stratified deposits (which includes sand and gravel, coarse sand, fine-grained sand, and flood-plain alluvium deposits). This classification of till and stratified deposits is consistent with characterizations of surficial deposits in several reports that discuss low-flow characteristics in Massachusetts (Ries, 1994a; 1994b; and 1997; Ries and Friesz, 2000). Till (also known as ground moraine) is an unsorted, unstratified mixture of clay, silt, sand, gravel, cobbles, and boulders deposited by glaciers commonly on top of bedrock throughout much of the state. Surficial till is primarily found in upland areas, but can also be found at depth in river valleys. Stratified deposits include sorted and layered glaciofluvial and glaciolacustrine deposits. Glaciofluvial deposits are material of all grain sizes (clay, silt, sand, gravel, and cobbles) deposited by glacial meltwater streams in outwash plains and river valleys. Glaciolacustrine deposits generally consist of clay, silt, and fine sand deposited in temporary lakes that formed after the retreat of the glacial ice sheet. Stratified deposits are more widespread in eastern Massachusetts than in western Massachusetts (Ries, 1994a, p. 6). In eastern Massachusetts, stratified deposits can be extensive outwash plains, particularly in the southeast. In other areas of the state, stratified deposits are more likely to be found in river valleys.

In southeastern Massachusetts (fig. 1), particularly in the Southeast Coastal Region—southern part of the South Coastal Basin, the eastern part of the Buzzards Bay Basin, and the Cape Cod and Islands Basin—the surficial geology is almost entirely stratified deposits (Simcox, 1992, p. 47, 51, and 52). In these areas the ground-water contributing area and the surface-water drainage area to a stream site can differ because ground water can flow from one surface-water basin into another. Thus, a logistic regression equation using surface-water drainage area as an explanatory variable may underestimate the probability of a stream flowing perennially for stream sites whose ground-water contributing areas are larger than their surface-water drainage areas. Conversely, for stream sites whose ground-water contributing areas are smaller than their surface-water drainage areas, the logistic regression equation may overestimate the probability of a stream flowing perennially. For these reasons, the Southeast Coastal Region (fig. 1) is not included in the study area.

Database Development

To develop an equation for estimating the probability of a Massachusetts stream flowing perennially at a specific site, a database of intermittent and perennial stream sites throughout the state was developed. Development of this database involved screening data to exclude stream sites affected by regulation, drought conditions, or other factors that may alter the intermittent or perennial status of streams. The sites were visited in late July through early September 2001 during low-flow conditions.

Factors that Affect the Intermittent or Perennial Status of Streams

For this study, a determination of the intermittent or perennial status of a stream site in Massachusetts was made by using data collected for naturally flowing streams (no regulation). Regulated streams are those affected by dams, surface-water withdrawals, ground-water withdrawals (pumping wells), diversions, wastewater discharges, and so forth. The intermittent or perennial status of a stream at a specific site cannot be evaluated accurately if the flows upstream or near the site are regulated, because the extent and type of regulation differ from site to site and the effects of the many regulations are not quantified or easily quantifiable.

Observations of the intermittent or perennial status of stream sites were made under selected hydrologic and climatic conditions. Abnormally dry and wet periods were avoided to make the most accurate determination of the intermittent or perennial status of a stream site. If observations were made during an abnormally dry period, a perennial stream might be observed to be intermittent, and during a wet period, an intermittent stream might be observed to be perennial. Field observations of the sites were made from late July through early September 2001 during low-flow conditions. Low-flow conditions for this study were defined as conditions during the months of July through September when streamflows were between about the 80- and 99-percent flow durations and when little to no precipitation had occurred for at least 3 to 5 days following a precipitation event totaling 0.10 in. or more. Flow durations for this study were determined by using the records of 19 long-term (greater than 10 years of record) continuous USGS streamflow-gaging stations minimally affected by regulation in and near Massachusetts (fig. 1 and table 1).

The revised Regulations of December 20, 2002 (Massachusetts Department of Environmental Protection, 2002a, p. 395) state that stream sites observed to be intermittent during a period of extended drought cannot be used for stream classification. The revised Regulations define a period of “extended drought” as a drought level of “Advisory” or more severe drought level (“Watch,” “Warning,” or “Emergency”) in the Massachusetts Drought Management Plan (Massachusetts Executive Office of Environmental Affairs and Massachusetts Emergency Management Agency, 2001, p. 14–18). The Massachusetts Drought Management Plan bases the drought level on seven indices: (1) Palmer Drought Index, (2) Crop Moisture Index, (3) fire-danger level, (4) precipitation, (5) ground-water levels, (6) streamflows, and (7) index reservoirs.

Although the revised Regulations definition of “extended drought” did not become effective until December 20, 2002 (Massachusetts Department of Environmental Protection, 2002b, p. 320), observations of the intermittent or perennial status of stream sites during late July through early September 2001 were done under normal conditions as defined by the original Regulations (Linda Marler, Massachusetts Department of Resource and Conservation, oral commun., 2003). The July through September 2001 period did not meet the original Regulations definition of “extended drought,” defined as a period of below-normal precipitation for that month and the three previous months, with at least three of the months having 75 percent or less of the normal precipitation and two months having 50 percent or less of normal precipitation. Additionally, ground-water levels and streamflows, which are indices (5) and (6) of the revised Regulations definition of “extended drought,” were reported to be in the normal range for most of the state during the months of July through September 2001 (Socolow and others, 2002). Because the revised and original Regulations have no operational definition of a wet month (above-normal precipitation conditions), stream sites were evaluated only during low-flow conditions, as previously defined for this study, to ensure that observations of stream sites were not made during wet periods (above-normal streamflow and precipitation conditions).

Site Selection

A total of 476 stream sites were selected for visual observation in the field from the most current USGS topographic maps available. About eight stream sites were generally picked from each metric-unit 7.5- by 15-minute USGS topographic map and about four stream sites from each English-unit 7.5- by 7.5-minute USGS topographic map. For USGS topographic maps of areas along the state borders or the Atlantic Ocean, the number of stream sites selected was weighted by the proportion of the map’s area within the state. No stream sites were selected for the Cape Cod and the Island Basins within the Southeast Coastal Region (fig. 1), as discussed previously. A few stream sites were selected in the southern part of the South Coastal Basin and eastern part of the Buzzards Bay Basin within the Southeast Coastal Region (fig. 1), because they were in areas where the ground-water contributing areas generally coincide with the surface-water drainage areas. Sites were also selected that were: (1) of varying drainage areas, basin elevation, basin slope, basin shape, wetland areas, water bodies, forest land, and urban land according to the USGS topographic maps; (2) representative of the number of intermittent and perennial stream sites shown on USGS topographic maps; and (3) accessible at road crossings (which permitted easier access and made it possible to avoid private property issues).

The visual inspection at each selected stream site included observations of whether the streambed was dry, had disconnected pools of water in the streambed, had water with no velocity (no flowing water), or had flowing water. Stream sites with a dry streambed or discontinuous pools of water were classified as intermittent. The stream sites were generally inspected along a reach at least two times the width of the bridge or culvert opening upstream of each road crossing to minimize any effect the bridge or culvert opening may have had on stream-channel characteristics, with the goal of observing a natural stream-channel section unaffected by human influences. Other observations recorded were the distance of the stream site upstream from the bridge; general land use in the area; any visible structures that may affect the stream status (such as public-water-supply wells, dams, cranberry bogs, and beaver dams); general cross-sectional information (width and depth) of the stream channel; general streambed material (clay, silt, sand, gravel, cobbles, boulders, or bedrock); and a crude estimate of the width, depth, and velocity of any flowing water in the stream channel. Additionally, the stream-channel section was documented with a photograph.

Table 1. Descriptions of long-term continuous index streamflow-gaging stations whose records were used to estimate low-flow durations at nearby field-visited stream sites in the development and verification of the logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts.

[**USGS station No.:** Streamflow-gaging stations shown on figure 1. **Latitude and longitude:** In degrees, minutes, and seconds. **Period of record:** p, present. No., number; mi², square mile; USGS, U.S. Geological Survey; --, none]

USGS station No.	Station name	Latitude ° ′ ″	Longitude ° ′ ″	Major river basin	Drainage area (mi²)	Period of record	Water years analyzed	Remarks
01096000	Squannacook River near West Groton, MA	42 38 03	71 39 30	Nashua	63.7	1949–p	1949–2001	Occasional regulation at low flow by mill upstream; regulation greater prior to 1961.
01097300	Nashoba Brook near Acton, MA	42 30 45	71 24 17	Concord	12.8	1963–p	1964–2001	Occasional regulation since 1967 by pond upstream.
01101000	Parker River at Byfield, MA	42 45 10	70 56 46	Parker	21.3	1945–p	1947–2001	Occasional regulation by mill and ponds upstream.
01105600	Old Swamp River near South Weymouth, MA	42 11 25	70 56 43	Boston Harbor	4.5	1966–p	1967–2001	--
01105730	Indian Head River at Hanover, MA	42 06 02	70 49 23	South Coastal	30.3	1966–p	1968–2001	Some regulation by mills and ponds upstream.
01105870	Jones River at Kingston, MA	41 59 27	70 44 03	South Coastal	15.7	1966–p	1967–2001	Flow regulated by pond upstream.
01111300	Nipmuc River near Harrisville, RI	41 58 52	71 41 11	Blackstone	16.0	1964–91, 1993–p	1965–91, 1994–2001	--
01162500	Priest Brook near Winchendon, MA	42 40 57	72 06 56	Millers	19.4	1916–p	1934, 1937–2001	Prior to 1962, occasional diurnal fluctuation at low flow by mill upstream.
01169000	North River at Shattuckville, MA	42 38 18	72 43 32	Deerfield	89.0	1939–p	1940–2001	Diurnal fluctuation at times by mill upstream.
01169900	South River near Conway, MA	42 32 31	72 41 39	Deerfield	24.1	1966–p	1967–2001	Diurnal fluctuation by small powerplant upstream since April 1982.
01170100	Green River near Colrain, MA	42 42 12	72 40 16	Deerfield	41.4	1967–p	1968–2001	--
01171500	Mill River at Northhampton, MA	42 19 05	72 39 21	Connecticut	54.0	1938–p	1940–2001	Flow regulated by mill upstream.
01175670	Sevenmile River near Spencer, MA	42 15 54	72 00 19	Chicopee	8.68	1960–p	1962–2001	Occasional regulation by ponds upstream since 1971.
01176000	Quaboag River at West Brimfield, MA	42 10 56	72 15 51	Chicopee	150	1912–p	1939–2001	Slight diurnal fluctuation at low flow by mill upstream prior to 1956; regulation much greater prior to 1938.
01181000	West Branch Westfield River at Huntington, MA	42 14 14	72 53 46	Westfield	94.0	1935–p	1936–2001	Prior to 1950, some diurnal fluctuation at low flow by mill upstream.
01184490	Broad Brook at Broad Brook, CT	41 54 50	72 33 00	Connecticut	15.5	1961–76, 1982–p	1967–76, 1983–2001	Flow regulated by reservoir and mill upstream.
01187300	Hubbard Brook near West Hartland, CT	42 02 14	72 56 22	Connecticut	19.9	1938–55, 56–p	1939–55, 1957–2001	--
01199050	Salmon Creek at Lime Rock, CT	41 56 32	73 23 29	Housatonic	29.4	1961–p	1962–2001	--
01333000	Green River at Williamstown, MA	42 42 32	73 11 50	Hudson	42.6	1949–p	1950–2001	Occasional slight diurnal fluctuation by mill upstream.

Of the 476 stream sites visited in the state, 125 were eliminated because of one or more of the following conditions: (1) observed or documented regulation of streamflows by dams, ground-water or surface-water withdrawals nearby, diversions, wastewater discharges, and so forth (evaluation of documented regulation at each stream site was done by using Geographic Information Systems (GIS) digital data layers available on the World Wide Web application STREAMSTATS (Ries and others, 2000)) or a Watershed Analyst Tool; (2) no visible stream channel observed in the field, although the site was shown as a stream on a USGS topographic map; (3) site not accessed because of private property, fences, or safety issues; (4) the drainage-basin boundary drawn by STREAMSTATS or a Watershed Analyst Tool did not match that determined from USGS topographic maps (most common for sites with small drainage basins in low-slope areas, generally with wetland areas); (5) STREAMSTATS or a Watershed Analyst Tool could not draw an accurate drainage-basin boundary because the stream-site location was too close to the end point of the centerline; (6) the centerline data in STREAMSTATS or the Watershed Analyst Tool started slightly downstream of the stream site; (7) it was determined that the ground-water contributing area and surface-water drainage area to the site did not coincide; and (8) observations were made when nearby streamflows were too high (flow durations were lower than 80 percent).

The application of these criteria resulted in the classification of 85 sites as intermittent and 266 sites perennial in the database (figs. 2 and 3). The stream sites observed to be intermittent had a minimum drainage area of 0.04 mi^2 (station PD-02), a mean drainage area of 0.61 mi^2, a median drainage area of 0.31 mi^2, and a maximum drainage area of 10.46 mi^2 (station AE-01) (table 8, at back of report). The sites observed to be perennial had a minimum drainage area of 0.06 mi^2 (station HF-05), a mean drainage area of 1.59 mi^2, a median drainage area of 0.92 mi^2, and a maximum drainage area of 10.96 mi^2 (station AN-01) (table 8, at back of report).

To confirm that field observations at the 85 intermittent- and 266 perennial-stream sites were made during low-flow conditions, the flow duration was estimated at each site on the day of the observation. The flow duration at each stream site was assumed to equal the calculated flow duration of the concurrent daily mean discharge on the day of the observation at the nearest long-term continuous USGS streamflow-gaging station (fig. 1 and table 1). All 351 stream sites were observed during flow durations between 80 and 99 percent, with 48 percent of the intermittent-site observations and 50 percent of the perennial-site observations between 85 and 95 percent, and 95 percent of all observations between 82 and 97 percent (fig. 4A and table 8, at back of report).

Comparison to Previous Stream-Status Observations

As a process of determining if the observed status of stream sites would be the same for different low-flow conditions and different field observers, 16 stream sites were visited during low-flow periods of two different years and by different personnel (table 2). In the Shawsheen River Basin, northeastern Massachusetts (figs. 2 and 3), two stream sites were observed on July 31, 2001 (table 2 and table 8, at back of report) and then again on July 31 or August 1, 2002. In the South Coastal Basin, southeastern Massachusetts (figs. 2 and 3), 14 stream sites that had been previously observed in mid-July through early September 1999 in the previous study by Bent and Archfield (2002, table 5) were observed in early August or early September 2001 (table 2 and table 8, at back of report). A different station-numbering system was used for stream sites in the South Coastal Basin during 1999 than during 2001 (table 2), but the stream sites are the same. In the Shawsheen River Basin, flow durations at the two stream sites were estimated to range between 85 and 92 percent on July 31 and August 1, 2002. In the South Coastal Basin, flow durations at the 14 stream sites were estimated to range between 73 and 99 percent (with 9 of the 14 sites between 89 and 95 percent) during mid-July through early September 1999.

All but 1 of the 16 stream sites in the Shawsheen River Basin and the South Coastal Basin were observed to have the same intermittent or perennial status during the two years. In the case of stream site number PE-03 in 2001 (table 2 and table 8, at back of report) and number PE-13 in 1999 (Bent and Archfield, 2002, table 5), the perennial observation in 1999 was made by looking upstream and downstream from the bridge. In 2001, water in the stream was observed at the upstream side of the bridge. About 125 ft upstream of the bridge (not visible from the bridge), however, the stream was found to consist of discontinuous puddles of water and was determined to be intermittent. Additionally, in 1999, the flow duration at the stream site was estimated to be at 73 percent on the day of the observation, while in 2001 the flow duration was estimated at 84 percent.

Overall consistency in the intermittent or perennial status of stream sites was also indicated for 22 of 31 sites (table 3) that were observed in 2001 and had been previously classified as intermittent or perennial by Bent and Archfield (2002, table 4) on the basis of the streamflow data in the USGS National Water Information System (NWIS). Differences in the intermittent or perennial stream status of the other nine stream sites may be a result of: (1) streamflow measurements or mean daily discharges that were recorded to be zero during a month that would meet the revised Regulations definition of “extended drought,” (2) measured or recorded streamflows less than 0.005 ft^3/s, which would be rounded down to 0.00 ft^3/s in the NWIS database, or (3) streamflow measurements recorded as “zero,” but noted as either “observed no flow” or “insufficient flow to measure”—implying that water was in the stream but that there may not have been sufficient depth or velocity to measure the streamflow. For these reasons, these nine stream sites may have been improperly designated as intermittent on the basis of data in the NWIS database.

Selection and Measurement of Basin Characteristics

Basin characteristics tested for use in the logistic regression analyses were selected on the basis of: (1) their theoretical relation to differences in the magnitude of low flows, (2) the results of previous studies involving estimation of the probability of a stream flowing perennially (Bent and Archfield, 2002) and selected low-flow statistics (Wandle and Randall, 1994; Ries, 1994a; 1994b; and 1997; Ries and Friesz, 2000; and Flynn, 2003), and (3) the ability to obtain consistent state-wide information on a characteristic, measure a characteristic, or both. Basin characteristics tested include drainage area; drainage density (the ratio of stream length to drainage area); mean basin slope and elevation; basin shape ratio (the ratio of length to width); and areal percentages of stratified deposits, sand and gravel deposits (excludes fine-grained deposits and flood-plain alluvium unless surrounded by sand and gravel deposits), water bodies, wetlands, urban land, and forest land.

Drainage-area boundaries were created and saved as shape files (.shp) by using a GIS-based Watershed Analyst Tool for all stream sites, unless the site was in the Buzzards Bay Basin, North Coastal Basin, or Taunton River Basin (figs. 2 and 3), where centerline data for the stream network are not available. Drainage areas in the Watershed Analyst Tool environment were determined from 1:25,000-scale digital-elevation models (DEMs) by the same procedure used by the program STREAMSTATS (Ries and others, 2000). Drainage-basin boundaries for stream sites in the Buzzards Bay Basin, North Coastal Basin, and Taunton River Basin were drawn in GIS with topographic maps as a backdrop to create shape files. The Watershed Analyst Tool also calculated the lengths of streams, areas of stratified deposits, and mean basin slopes by the same procedures used by the program STREAMSTATS for all stream sites outside of Buzzards Bay Basin, North Coastal Basin, and Taunton River Basin. For the stream sites in these three basins, lengths of streams, areas of stratified deposits, and mean basin slopes were calculated by using shape files, digital data layers, and the same procedures used by the Watershed Analyst Tool and STREAMSTATS. The lengths of streams were determined from centerline data for streams from a 1:25,000-scale hydrography digital data layer (MassGIS, 2004a). The areas of stratified deposits were determined from a 1:250,000-scale surficial-geology digital data layer (MassGIS, 2004b). Mean basin slopes were determined from 1:250,000-scale DEMs. The shape files created in the Watershed Analyst Tool environment and the procedures described previously for the Buzzards Bay Basin, North Coastal Basin, and Taunton River Basin were then used in conjunction with the areas of sand and gravel deposits, water bodies, wetlands, urban land, and forest land in digital data layers from MassGIS (2004a, 2004b, and 2004c). All the basin characteristics were calculated using scripts written in the Arc Macro Language (AML).

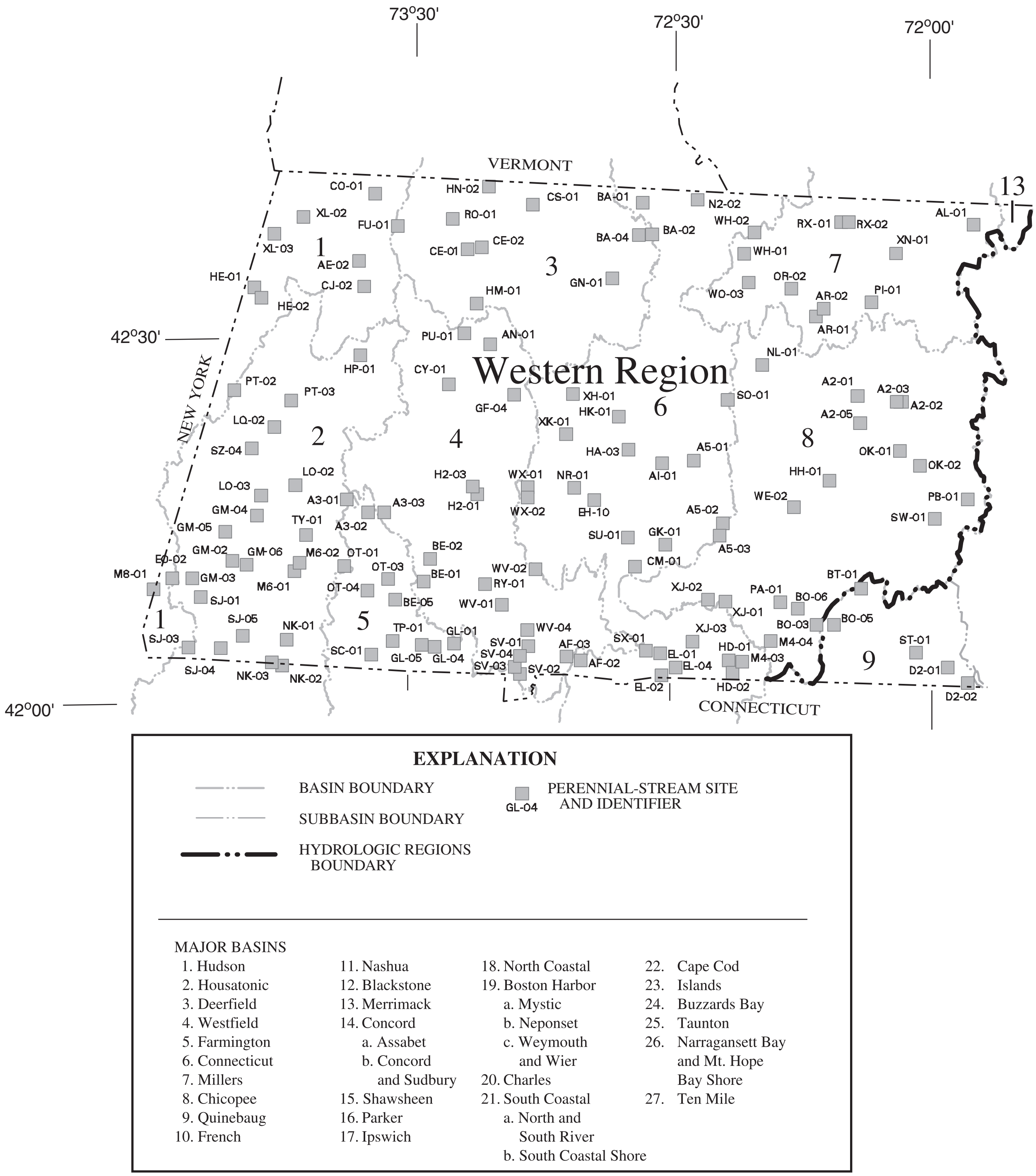
Drainage density and the percentages of the drainage area of each basin underlain by stratified deposits, sand and gravel deposits, water bodies, wetlands, urban land, and forest land were also determined for each stream site in developing a logistic regression equation. Drainage density represents the density of a network of streams available for drainage of runoff within a basin. Drainage density, mean basin slope, and basin shape likely would affect the efficiency with which water can be routed out of a basin. The areal percentage of stratified deposits, sand and gravel deposits, water bodies, wetlands, urban land, and forest land within a basin indicates which of these land covers and uses are dominant within a basin.

In the previous logistic regression analyses, Bent and Archfield (2002) tested the following potential explanatory variables: drainage area, drainage density, areal percentage of stratified deposits, mean basin slope, and a location variable to indicate whether a stream site was located in the South Coastal Basin or the remainder of the state. All five variables were found to be significant in the analyses. The cube root (1/3 power) of drainage area and the square root (1/2 power) of areal percentage of stratified deposits best represented these variables in that logistic regression equation. All of these variables, with the exception of the location variable, were also tested during this study as potential explanatory variables for the revised equation.

The extent and types of surficial deposits are important factors that explain flow characteristics of Massachusetts streams (Ries, 1994a; 1994b; and 1997; Ries and Friesz, 2000). During dry periods, the primary source of streamflow is ground-water discharge from the aquifer to the stream. Till and fine-grained stratified deposits generally have a lower infiltration capacity than medium- to coarse-grained stratified deposits. The lower infiltration capacity of these materials results in greater direct runoff of precipitation; therefore, less precipitation is available to infiltrate the soil and recharge the aquifer. Thus, basins underlain predominantly by till and fine-grained stratified deposits generally have a lower streamflow per unit area during dry periods than basins underlain predominantly by medium- to coarse-grained stratified deposits.

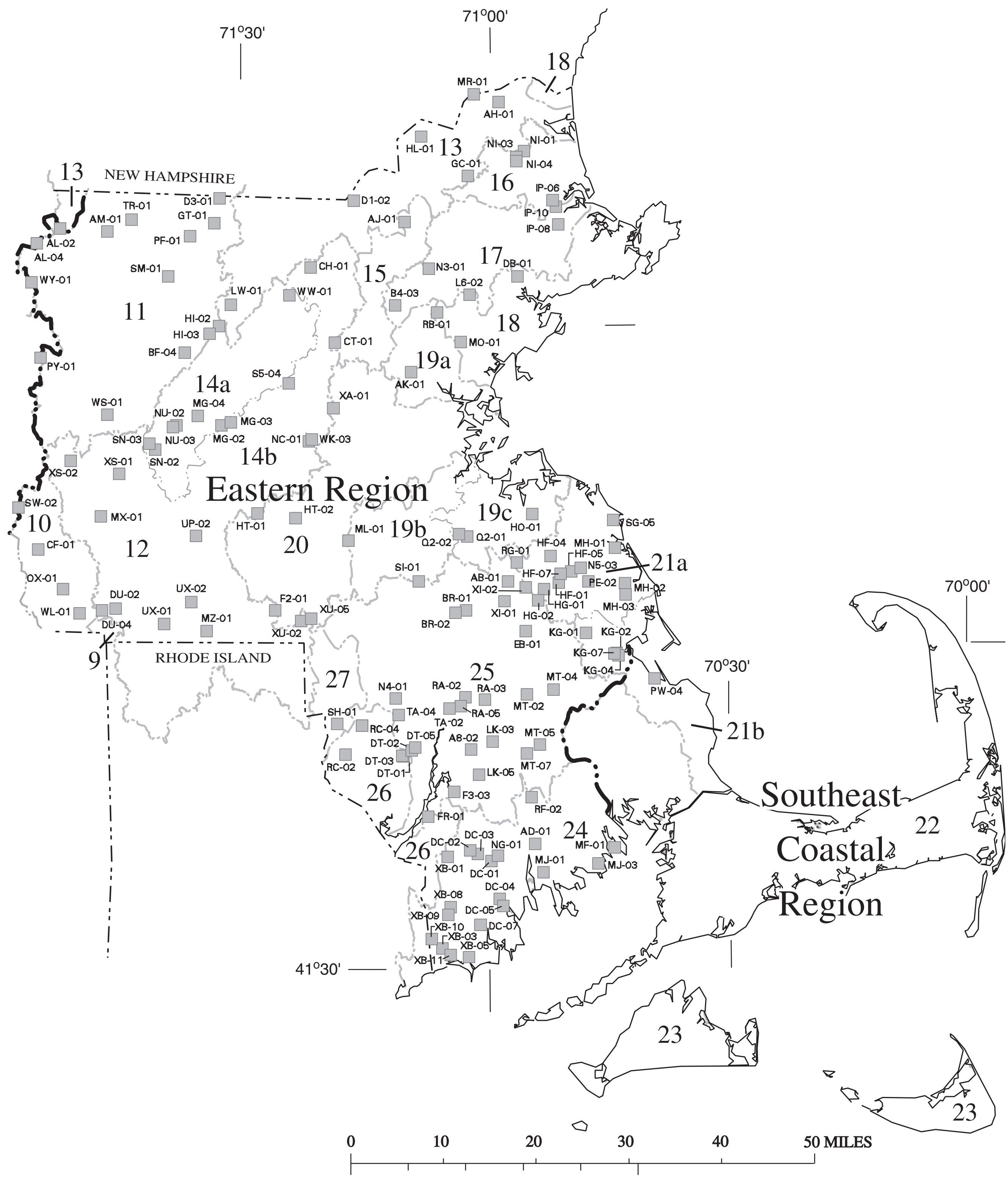
The MassGIS (2004b) 1:250,000-scale surficial-geology digital data layer consists of seven categories: (1) sand and gravel deposits, (2) till or bedrock outcrops, (3) sandy till over sand, (4) end moraines, (5) large sand deposits, distinguished from sand and gravel deposits, (6) fine-grained deposits, and (7) flood-plain alluvium. The areal extent of stratified deposits, tested in the logistic regression analyses, was calculated on the basis of categories 1, 5, 6, and 7. The areal extent of sand and gravel deposits, tested in the logistic regression analyses, was calculated on the basis of categories 1 and 5, and categories 6 and 7 if they were surrounded by categories 1, 5, or both.

The areal extent of water bodies, tested in the logistic regression analyses, was calculated on the basis of the water GIS digital data layer, which includes freshwater and coastal embayments (MassGIS, 2004c). The areal extent of wetlands, tested in the logistic regression analyses, was calculated on

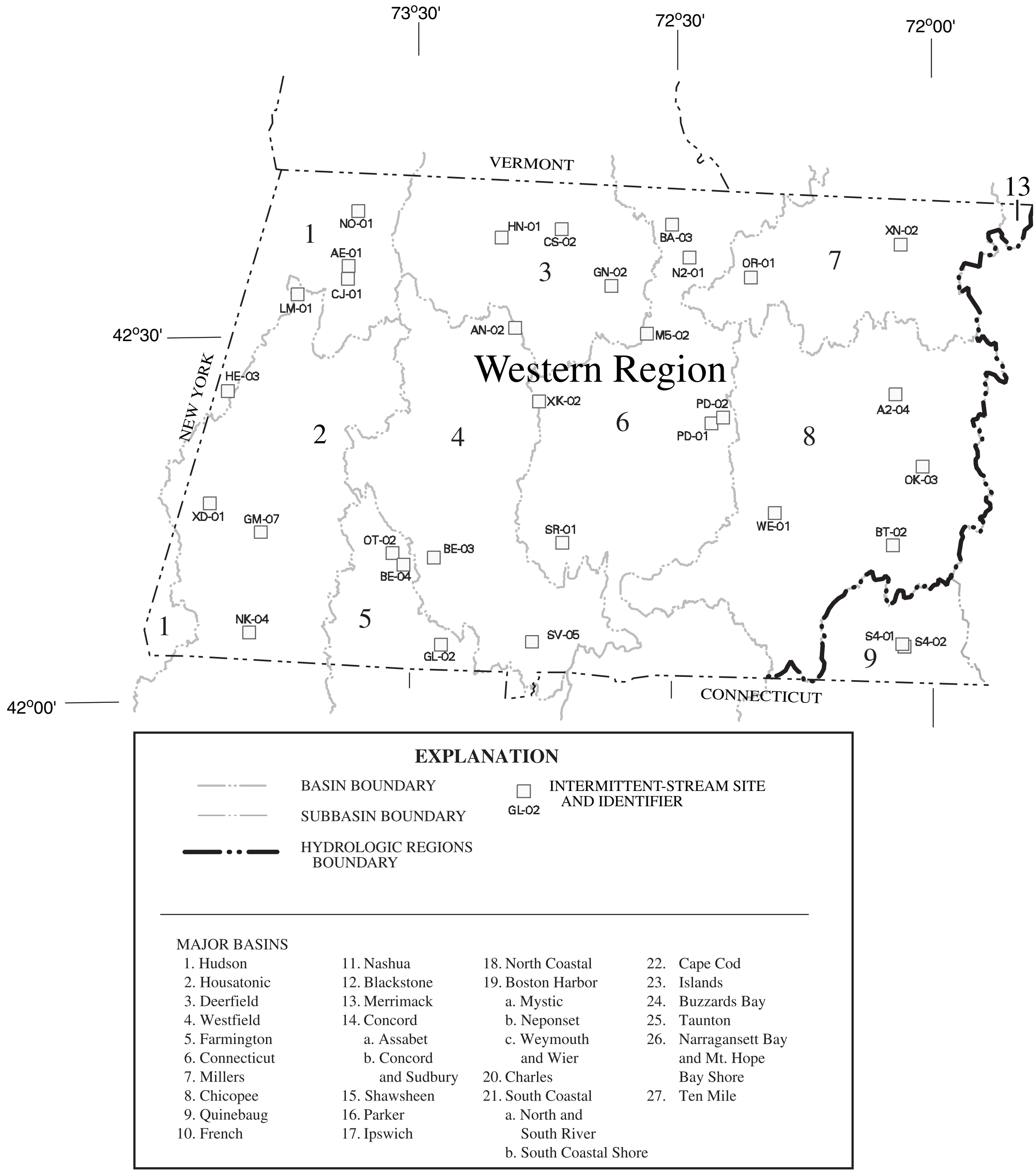


Map modified from Bent and Archfield, 2002
Additional coverages from University of Connecticut MAGIC and MassGIS databases
Base from U.S. Geological Survey digital data, 1:25,000, 1991
Lambert conformal conic projection. 1983 North American Datum
Massachusetts coordinate system mainland zone

Figure 2. Locations of field-visited stream sites designated as perennial streams used in development of the logistic regression

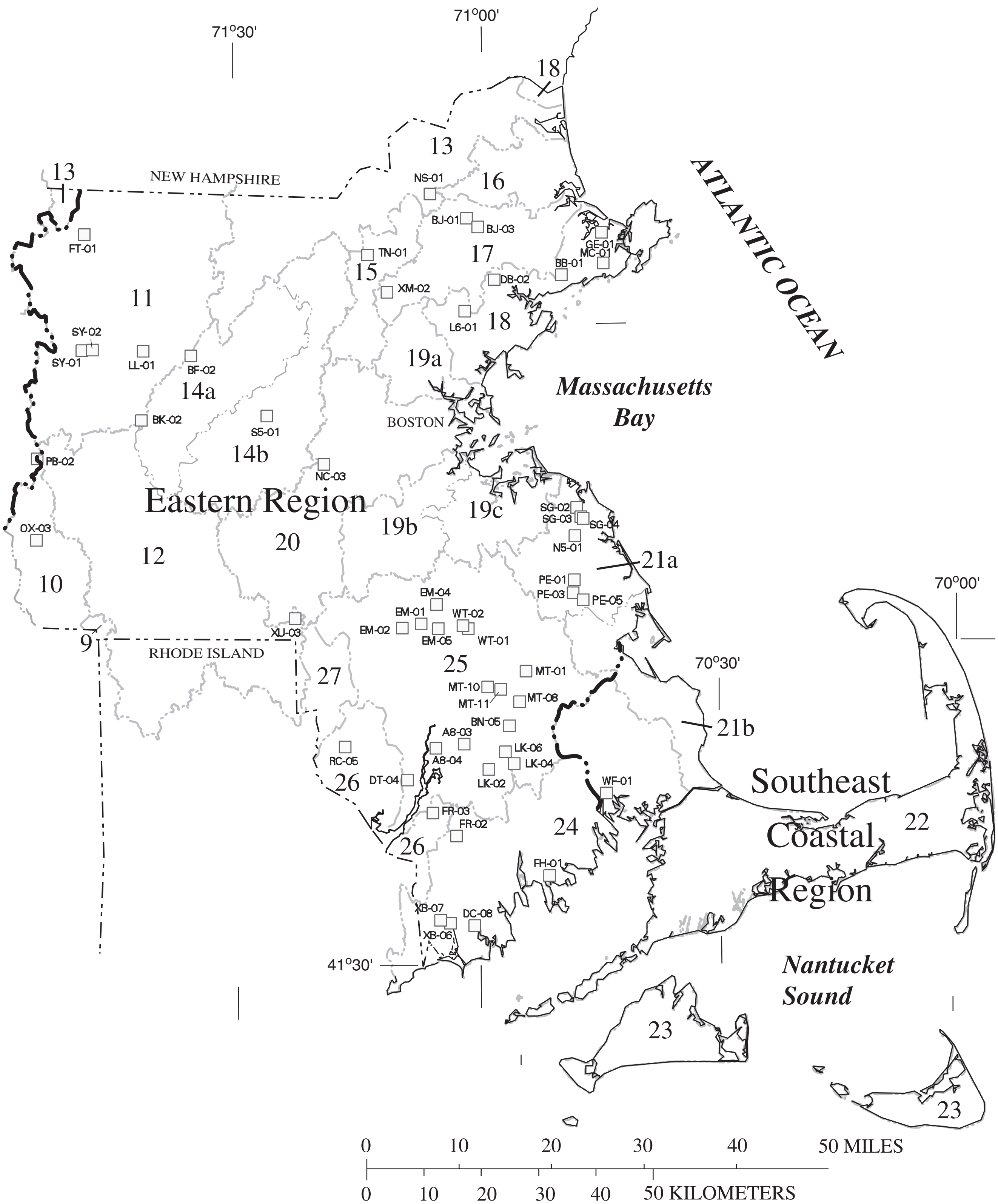


equation for estimating the probability of a stream flowing perennially in Massachusetts, late July through early September 2001.



Map modified from Bent and Archfield, 2002
Additional coverages from University of Connecticut MAGIC and MassGIS databases
Base from U.S. Geological Survey digital data, 1:25,000, 1991
Lambert conformal conic projection. 1983 North American Datum
Massachusetts coordinate system mainland zone

Figure 3. Locations of field-visited stream sites designated as intermittent streams used in development of the logistic regression



equation for estimating the probability of a stream flowing perennially in Massachusetts, late July through early September 2001.

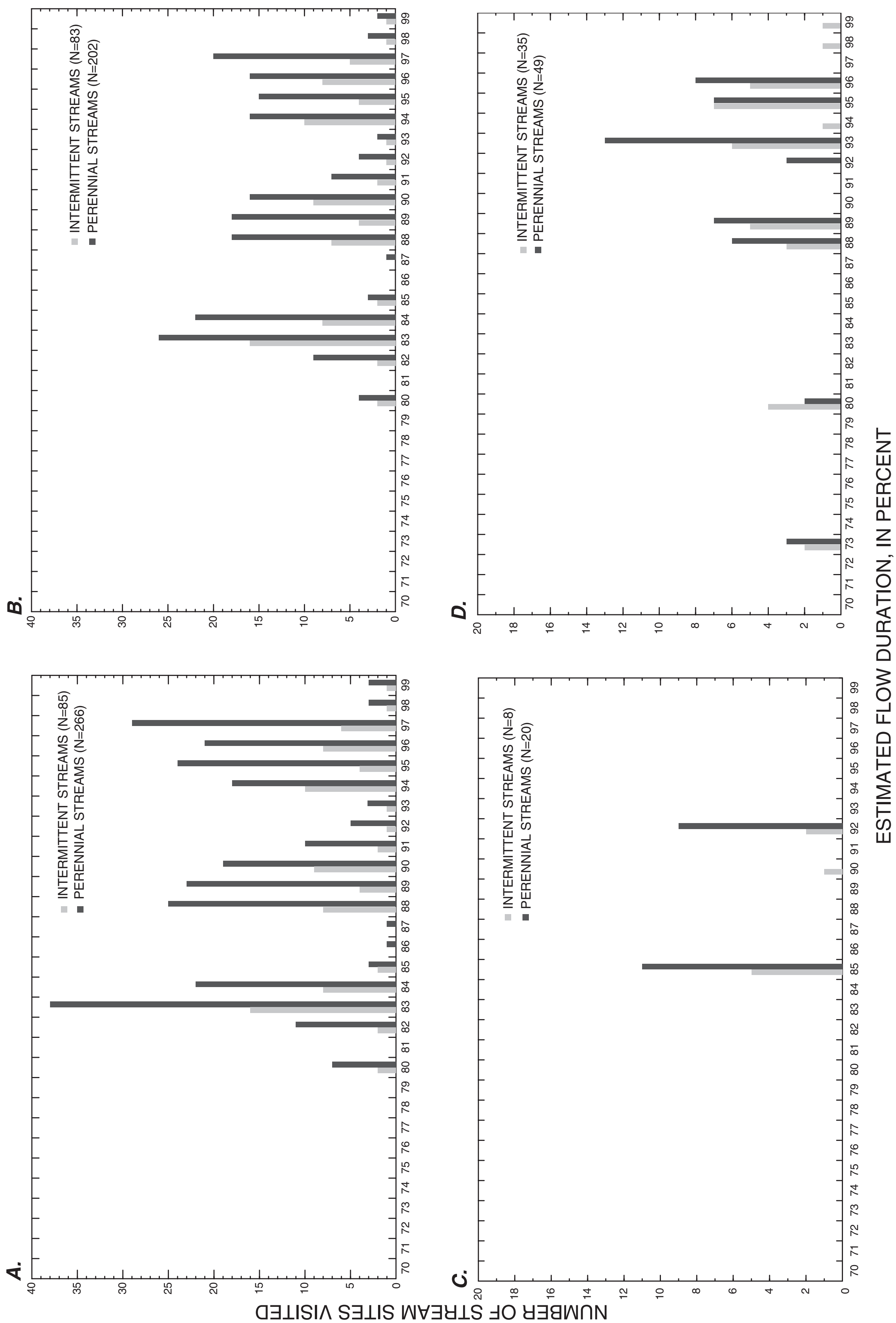


Figure 4. The distribution of the estimated flow durations on the dates of field visits at the *A*, 351 stream sites throughout Massachusetts; *B*, 285 stream sites with drainage areas less than 2.00 square miles throughout Massachusetts; *C*, 28 stream sites in the Shawsheen River Basin, northeastern Massachusetts; and *D*, 84 stream sites in the South Coastal Basin, southeastern Massachusetts.

Table 2. Comparison between intermittent and perennial field observations at stream sites visited in different years in the Shawsheen River Basin, northeastern Massachusetts, and in the South Coastal Basin, southeastern Massachusetts.

[**Observed status:** I, intermittent; P, perennial. No., number]

Station No.	Date observed	Observed status	Station No.	Date observed	Observed status
		Shawsheen River Basin			
	2002			2001[2]	
AJ-01	8-1-2002	P	AJ-01	7-31-2001	P
TN-01	7-31-2002	I	TN-01	7-31-2001	I
		South Coastal Basin			
	1999[1]			2001[2]	
HN-15	7-19-1999	P	HG-01	9-05-2001	P
HR-09	8-13-1999	P	HF-01	9-07-2001	P
HR-44	8-10-1999	P	HF-04	8-01-2001	P
HR-03	8-12-1999	P	HF-05	9-07-2001	P
HR-16	8-12-1999	P	HF-07	9-07-2001	P
KN-07	9-03-1999	P	KG-07	9-07-2001	P
NL-19	8-20-1999	I	N5-01	8-01-2001	I
PE-20	7-29-1999	P	PE-02	9-07-2001	P
PE-13	7-29-1999	P	PE-03	9-07-2001	I
PE-23	7-29-1999	I	PE-05	9-07-2001	I
RD-21	8-04-1999	P	RG-01	8-01-2001	P
SE-16	7-16-1999	I	SG-02	8-01-2001	I
SE-14	7-16-1999	I	SG-04	8-01-2001	I
SE-02	7-16-1999	P	SG-05	9-07-2001	P

[1] A different station numbering system was used for stream sites in the South Coastal Basin during 1999 than during 2001, but the stream sites are the same. 1999 data from Bent and Archfield, 2002, table 5.

[2] 2001 data in table 8. Stream sites observed as perennial in 2001–02 are shown on figure 2, and stream sites observed as intermittent in 2001–02 are shown on figure 3 under the station number for 2001.

the basis of the wetland (non-forested freshwater wetland) and cranberry bogs and salt wetland (salt marsh) GIS digital data layers (MassGIS, 2004c). Wandle and Randall (1994) found the areal extent of lakes (water bodies) and swamps (wetlands) to be inversely related to low flows in central New England, and suggested that this was the result of evaporation from lakes and evapotranspiration from swamps. Forty-seven of 71 studies involving the processes of wetlands in the hydrologic cycle found that wetlands reduce streamflow during dry periods (Bullock and Acreman, 2003). Bullock and Acreman (2003) reported that in 22 of 23 studies, this reduction was likely caused by greater evapotranspiration from wetland areas than from nonwetlands areas during dry periods. Thus, water bodies and wetlands may reduce streamflow during dry periods and may give a stream site a greater tendency to be intermittent during summer low-flow periods.

The areal extent of urban land, tested in the logistic regression analyses, was calculated on the basis of the residential (multifamily and smaller than 0.25 acre lots), commercial (general urban and shopping center), industrial (light and heavy industry), transportation (airports, docks, divided highway, freight, storage, railroads), and high-density residential housing (smaller than 0.25 acre lots) GIS digital data layers (MassGIS, 2004c). Urban areas have been found to reduce base flows (ground-water discharge to the stream) in streams (Simmons and Reynolds, 1982; Spinello and Simmons, 1992; Rose and Peters, 2001; and Calhoun and others, 2003). Less precipitation recharges the ground-water aquifer from impervious surface areas as runoff is routed to drainage structures, some of which may carry it outside the basin; less recharge of the aquifer also results from the use of septic tanks, because sanitary sewers route the wastewater to treatment facilities. The reduced base flow in urban areas may give a stream site a greater tendency to be intermittent during summer low-flow periods.

The areal extent of forest land, tested in the logistic regression analyses, was calculated on the basis of the forest GIS digital data layer (MassGIS, 2004c). Hornbeck and others (1993 and 1997) summarized the results of several studies on the effects of timber-management activities (cutting of trees) on water yield in the northeastern United States. They found that the cutting of trees increased streamflows during low-flow periods of the summer, because there is less evapotranspiration and less interception of precipitation by the tree canopy in the cut areas. Two studies in central Massachusetts determined that summer low flows, base flow, and ground-water recharge increased as a result of timber cutting (Mrazik and others, 1980, Bent, 1994; 2001; and Shanley and others, 1995). Conversely, more evapotranspiration and less ground-water recharge of the aquifer would be expected in forested areas. Thus, forested areas likely would produce reduced summer low flows and stream sites with greater tendencies to be intermittent.

Logistic Regression Equation

Logistic regression is a statistical technique in which the probability of a result being in one of two response groups (binary response) is modeled as a function of the magnitudes of one or more explanatory variables (Helsel and Hirsch, 1992, p. 393–402). For instance, the probability of whether a stream is intermittent or perennial at a specific site may be modeled as a function of the magnitudes of one or more basin characteristics. For this study, the response variable is 0 when the stream is intermittent and 1 when the stream is perennial.

Several other studies have used logistic regression to determine the intermittent or perennial status of streams or to investigate other water-resources issues. In a previous related study by Bent and Archfield (2002), logistic regression was used to estimate the probability of a stream flowing

Table 3. Comparison between intermittent and perennial stream status determined through field observations and estimated on the basis of measurements of daily mean discharge at stream sites by region and river basin in Massachusetts.

[**Station No.:** Stream sites observed as perennial shown in figure 2, and stream sites observed as intermittent shown on figure 3. **Date observed:** 2001 data in table 7; 2002 data in table 9. **Observed status:** I, intermittent; P, perennial. **Estimated status of stream site:** From Bent and Archfield, 2002, table 4. Based on streamflow measurements or daily mean discharge data. No., number; USGS, U.S. Geological Survey; ft^3/s, cubic feet per second.]

2001 or 2002			Various years				
Station No.	Date observed	Observed status	USGS station No.	Number of streamflow measurements or days of daily mean discharge data	Years of data	Estimated status of stream site	Remarks
Eastern Region							
Boston Harbor Basin							
AK-01	8-03-2001	P	01103015	27	1973–74, 1989–91, and 1999–2000	P	
Buzzards Bay Basin							
DC-07	9-05-2001	P	01105935	30	1972–74, 1991–93, and 2003	P	
DC-02	9-04-2001	P	01105937	34	1957, 1972–74, 1991–94, 1996, and 2003	P	
Merrimack River Basin							
D1-02	8-02-2001	P	01100050	2	1973–74	I	Zero flow 7-25-1974.
HL-01	8-02-2001	P	01100665	2	1973–74	I	Zero flow 7-23-1974.
MR-01	8-02-2001	P	01100800	2	1965 and 1974	I	Zero flow 8-20-1965 and 7-22-1974.
Narragansett Bay and Mt. Hope Bay Shore Basin							
RC-04	9-04-2001	P	[1]01109200	4,357	1962–74	I	Zero mean daily discharge recorded on one or more days in the months of 9-1964, 7-1965, 8-1965, 9-1965, 8-1966, and 8-1974.
Nashua River Basin							
D3-01	8-09-2001	P	01096505	23	1971–74, and 1991–93	P	
Shawsheen River Basin							
BC-04	7-31-2002	I	01100593	1	1974	I	Zero flow 7-31-1974.
TN-09	7-31-2002	P	01100608	15	1973-74, and 1994–96	P	
AJ-07	7-31-2002	P	01100618	1	1974	I	Zero flow 8-1-1974.
NS-03	8-01-2002	I	01100633	1	1974	I	Zero flow 8-1-1974.
Quinebaug River Basin							
BO-05	8-01-2001	P	01123161	10	1994–96	I	Observed no flow 8-11-1994 and 9-7-1994; insufficient flow to measure on 8-21-1995.
Western Region							
Chicopee River Basin							
A2-05	8-17-2001	P	[1]01173260	4,359	1962–74	I	Zero mean daily discharge recorded on one or more days in the months of 07-1963, 08-1963, 09-1963, 09-1964, 10-1964, 08-1965, 09-1965, 09-1968, and 08-1970.
NL-01	8-03-2001	P	[1]01174000	12,735	1947–82	P	
SW-01	8-17-2001	P	[1]01175670	15,644	1960–2003	P	Currently (2005) still operated as continuous streamflow-gaging station.

Table 3. Comparison between intermittent and perennial stream status determined through field observations and estimated on the basis of measurements of daily mean discharge at stream sites by region and river basin in Massachusetts.— Continued

[**Station No.:** Stream sites observed as perennial shown in figure 2, and stream sites observed as intermittent shown on figure 3. **Date observed:** 2001 data in table 7; 2002 data in table 9. **Observed status:** I, intermittent; P, perennial. **Estimated status of stream site:** From Bent and Archfield, 2002, table 4. Based on streamflow measurements or daily mean discharge data. No., number; USGS, U.S. Geological Survey; ft^3/s, cubic feet per second.]

2001 or 2002			Various years				
Station No.	Date observed	Observed status	USGS station No.	Number of streamflow measurements or days of daily mean discharge data	Years of data	Estimated status of stream site	Remarks
Western Region—Continued							
Connecticut River Basin							
NR-02	8-10-2001	P	[1]01171800	4,352	1962–74	P	
Deerfield River Basin							
BA-04	8-09-2001	P	01170240	14	1969 and 1994–96	P	
Housatonic River Basin							
HE-03	8-22-2001	I	01197080	4	1963–65	I	Zero flow on 8-05-1964 and 9-08-1965.
PT-03	8-22-2001	P	01197130	14	1963–65 and 1994–95	P	
LQ-02	8-22-2001	P	01197140	18	1963–65 and 1994–96	P	
LO-02	8-22-2001	P	01197180	16	1963–65, 1991–93, and 1995	P	
GM-04	8-23-2001	P	01197240	14	1964–65 and 1994–95	P	
SZ-04	8-23-2001	P	[1]01197300	4,322	1962–74	I	Zero mean daily discharge recorded on one or more days in the months of 9-1963, 8-1964, 9-1964, 8-1970, and 8-1972.
EO-02	8-23-2001	P	01198062	11	1994–95	P	Measurement on 8-23-1995 was 0.002 ft^3/s, but was published as 0.00 ft^3/s.
M6-02	8-23-2001	P	01198137	11	1994–95	I	Volumetric measurements on 8-22-1995 were 0.001 ft^3/s and on 9-5-1995 were 0.0004 ft^3/s, but were published as 0.00 ft^3/s.
NK-02	8-23-2001	P	01198260	11	1994–95	P	
Hudson River Basin							
XL-02	8-21-2001	P	01332900	15	1967–69 and 1994–96	P	
Millers River Basin							
WO-03	8-03-2001	P	[1]01166105	2,113	1985–91	P	
Westfield River Basin							
H2-01	8-10-2001	P	[1]01180000	10,621	1945–74	P	
A3-03	8-22-2001	P	[1]01180800	5,448	1962–77	P	

[1] Operated as a continuous streamflow-gaging station.

perennially in Massachusetts. In a study by Kliever (1996), logistic regression was used to determine the probability that streamflow would be zero for a particular low-flow statistic at a partial-record station when streamflow was at the same low-flow statistic at nearby index stations in northern Rhode Island. Tasker (1989) developed a logistic regression equation based on basin characteristics to estimate the probability of the annual N-day low flow at a network of low-flow stations in Florida being zero. Other applications of logistic regression analyses to water resources include those by Eckhardt and others (1989), Eckhardt and Stackelberg (1995), Mueller and others (1997), Nolan and others (1997), Tesoriero and Voss (1997), Koltun and Sherwood (1998), Rupert (1998), Squillace and others (1999), Nolan (2001), Nolan and others (2002), Battaglin and others (2003), Francy and others (2003), and Rupert (2003).

Development

Logistic regression analyses were used to estimate the probability of a stream flowing perennially by relating the observed intermittent or perennial status of a stream site to selected basin characteristics or regional characteristics of naturally flowing streams in Massachusetts. Of 351 stream sites visited, 64 of 66 (97 percent) with drainage areas greater than 2.00 and less than 11.00 mi^2 were observed to be perennial streams (fig. 5*A*). Thus, stream sites with drainage areas greater than 2.00 mi^2 were assumed to flow perennially, and the database used to develop the logistic regression equation included only those 285 stream sites (of which 83 were observed to be intermittent and 202 perennial) with drainage areas less than 2.00 mi^2.

The two intermittent stream sites (BA-03, Dry Brook in Bernardston and AE-01, Dry Brook in Adams) (fig. 3 and table 8) with drainage areas greater than 2.00 mi^2 (3.48 and 10.46 mi^2, respectively) are underlain by predominately till and bedrock. Short (length dimension) narrow areas of stratified deposits in river valleys composed only 9.82 and 3.37 percent (0.34 and 0.35 mi^2) of each drainage area, respectively. On the basis of this information as well as the fact that one stream site was completely dry and the other site had only a few discontinuous puddles of water (both sites exhibited no evidence of recent water flow in the stream channel), and that both are named “Dry Brook,” the sites were considered losing stream reaches. Losing streams are defined as streams or stream reaches of stream that lose water to the ground-water system (Winter and others, 1998, p. 9–10 and 16–17). Generally, stream reach is losing where the ground-water table does not intersect the streambed in the channel (water table is below the streambed) during low-flow periods. Losing stream reaches commonly begin at the juncture where the stream flows from an area of the basin underlain by till or bedrock onto an area underlain by stratified deposits (where hillsides meet river valleys in central and western Massachusetts). At this juncture, a stream can lose a substantial amount of water through its streambed.

Proportional measures of the basin characteristics drainage density and areal percentage of stratified-drift deposits, sand and gravel deposits, water bodies, wetlands, urban land, and forest land were tested in the logistic regression analyses over absolute measures of length of streams in miles and land-cover or land-use areas in square miles. Drainage density and areal percentages of land-cover and land-use characteristics provide a more equal comparison of basin characteristics for a wide range of drainage-area sizes, such as those in this study. Additionally, transformations (natural logarithm and the powers -2, -1, -0.5, 0.5, 2, and 3) of the 11 basin characteristics were tested as possible explanatory variables. Transforming data is a common procedure that makes the data more symmetric, linear, and constant in variance (homoscedasticity) (Helsel and Hirsch, 1992, p. 12–14).

Ries (1997) and Ries and Friesz (2000) found that low flows per unit area in the western region of Massachusetts are higher than in the eastern region of the state. The eastern region includes major river basins east of the Chicopee, Connecticut, and Millers River Basins; and the western region includes major river basins west of the Blackstone, French, Merrimack, Nashua, and Quinebaug River Basins (figs. 1–3). The percentage of perennial stream sites in each drainage-area range appears to differ between the eastern and western regions of the state (figs. 5*B* and *C*). In the western region, the percentages of perennial streams are lower than in the eastern region for drainage areas between 0.00 and 0.29 mi^2, but lower in the eastern region than in the western region for 7 of 10 drainage area ranges between 0.30 and 2.00 mi^2. Additionally, boxplots of the 11 basin characteristics grouped by region of the state and by stream status show differences in the 25th, 50th (median), and 75th percentiles for mean basin elevation and slope; basin shape ratio; and areal percentages of stratified deposits, sand and gravel deposits, wetlands, urban land, and forest land. No noticeable differences between the eastern and western regions were found for drainage area, drainage density, and the areal percentage of water bodies. Thus, an explanatory variable was included for testing in the logistic regression analyses to determine if any regional difference (between the eastern and western regions) was present between intermittent and perennial stream sites.

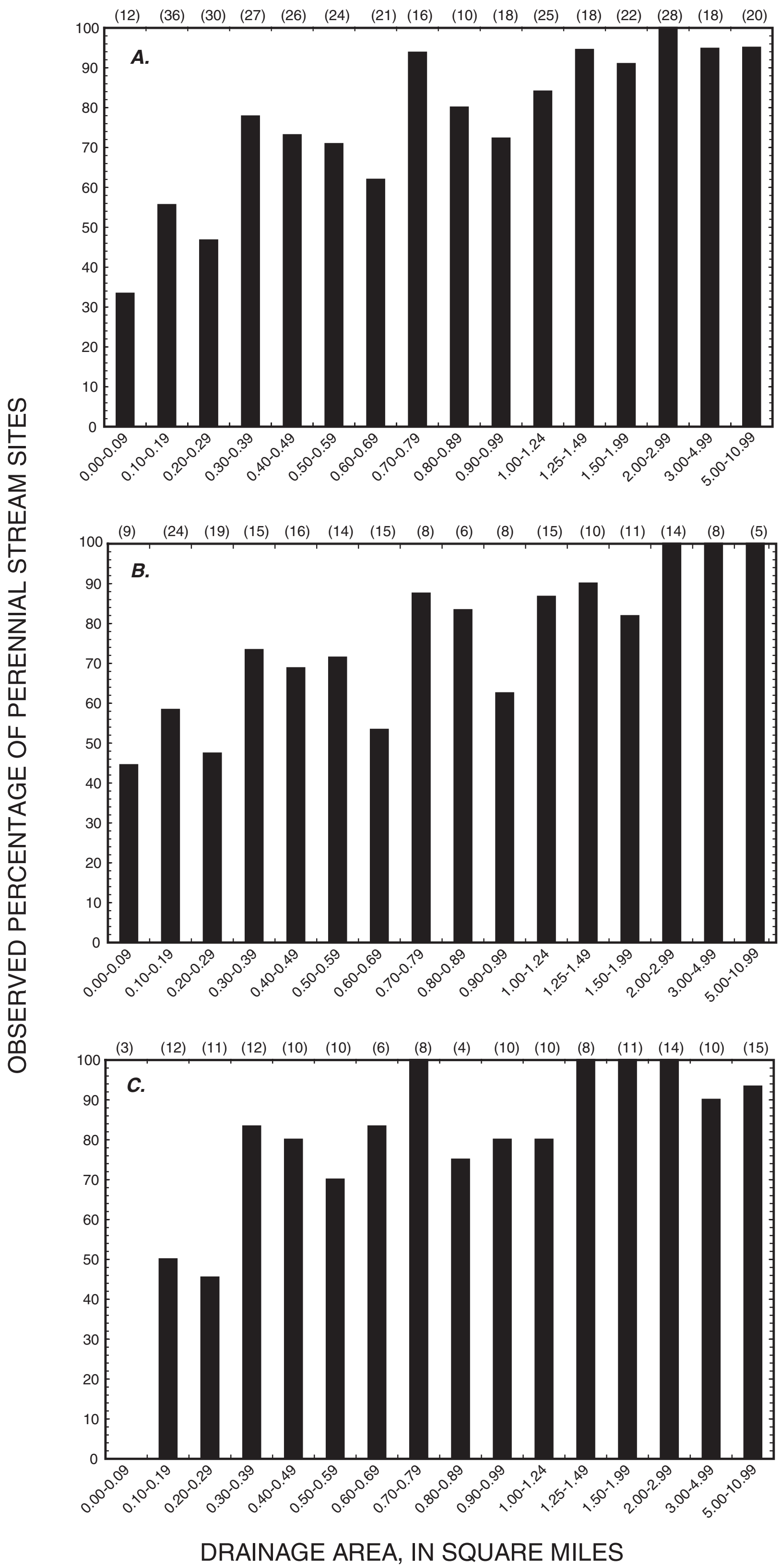


Figure 5. Observed percentage of perennial-stream sites by drainage-area range in Massachusetts *A*, statewide; *B*, eastern region; and *C*, western region. Number in parentheses at top of graph is the total number of sites in each drainage-area range. Regions are shown in figure 1.

The logistic regression analyses were computed with the Stata and SAS statistical software packages (Stata Corporation, 2003 and SAS Institute, Inc., 1989; 1995). The general form of a logistic regression equation is

$$P = \frac{exp\left(b_0 + b_1 x_1 + \ldots + b_i x_i\right)}{1 + exp\left(b_0 + b_1 x_1 + \ldots + b_i x_i\right)}, \quad (1)$$

where

- P is the probability of the condition being true;
- exp is the exponential function and is written as *exp(x)* or $e^{(x)}$ (where *e* is the base of the natural logarithm and is approximately equal to 2.7183);
- b_0 is the intercept;
- $b_{1 \ldots i}$ is the coefficient for explanatory variable *I*; and
- $x_{1 \ldots i}$ is the value of explanatory variable *i*.

More detailed information on logistic regression can be found in Collett (1991) and Hosmer and Lemeshow (2000).

All potential explanatory variables (basin characteristics; transformations of the basin characteristics; and location identifiers) were evaluated by using the procedures of forward selection, backward elimination, stepwise selection, and best subset selection to help determine the best possible logistic regression equations (SAS Institute, Inc., 1995, p. 51–65; Hosmer and Lemeshow, 2000, p. 91–142; and Cook, 2001, p. 81–144). A statistical significance level of 0.05 for *p*-values of explanatory variables was used for entry or retention in the equations. Each equation developed during the variable-evaluation process was used to predict the probability that its corresponding stream site was perennial.

The goodness-of-fit of each potential logistic regression equation was evaluated with the Hosmer and Lemeshow (2000) goodness-of-fit test that compares the observed to the predicted distribution of outcomes (SAS Institute, Inc., 1995, p. 67–72; Hosmer and Lemeshow, 2000, p. 147–156; and Cook, 2001, p. 158, 163, 165, and 172). The receiver-operating-characteristic (ROC) curves were also evaluated to assess the predictive accuracy of the logistic regression equation (SAS Institute, Inc., 1995, p. 87–92, Hosmer and Lemeshow, 2000, p. 160–164; and Cook, 2001, p. 159, 163, and 165). Finally, regression diagnostics of the equations were evaluated to determine how each observation affects the fit of the logistic regression equation (SAS Institute, Inc., 1995, p. 73–79, Hosmer and Lemeshow, 2000, p. 167–186; and Cook, 2001, p. 159–166).

The results given by the equations were summarized in classification tables (SAS Institute, Inc., 1995, p. 45–50; Hosmer and Lemeshow, 2000, p. 156–161). These tables provide information about the predictive accuracy of an equation by summarizing the frequency with which observations are correctly and incorrectly classified as events or nonevents for different probability cutpoints. Because the same data are used to develop the equation and to test its predictive accuracy, a method that approximates the unbiased jackknifing procedure was used to create the classification tables (SAS Institute, Inc., 1995, p. 45). Jackknifing minimizes bias caused when an independent set of observations is not available to test the predictive accuracy of the equation.

The best logistic regression equation determined from data in this study is

$$P = \frac{exp(2.8084 + 0.9884ln\left(x_1\right) + 0.0111x_2 - 0.0233x_3 + 0.7500x_4)}{1 + exp(2.8084 + 0.9884ln\left(x_1\right) + 0.0111x_2 - 0.0233x_3 + 0.7500x_4)}, \quad (2)$$

where

- P is the probability of a stream flowing perennially at a specific site;
- exp is approximately 2.7183;
- ln is the natural logarithm;
- x_1 is the drainage area of the basin (mi^2);
- x_2 is the areal percentage of sand and gravel deposits in the basin;
- x_3 is the areal percentage of forest land; and
- x_4 is an integer variable for the location of the stream site in the eastern region (0) or western region (1) (figs. 1–3).

Noticeable differences are present between intermittent and perennial stream sites throughout the state for the variables drainage area, areal percentage of sand and gravel deposits, and areal percentage of forest land (fig. 6); these differences support the identification of these three variables as significant explanatory variables. Additionally, the noticeable differences that exist between intermittent and perennial stream sites and between the two regions of the state for sand and gravel deposits and forest land support the identification of region as a significant explanatory variable. Results of the analysis of the maximum likelihood estimates of this equation are presented in table 4 (SAS Institute, Inc., 1995, p. 20–22). The *p*-value of each explanatory variable is less than 0.05, the value used as the statistical-significance level for entry or retention for the equation in this study. Summary statistics for the logistic regression analyses for this equation with four explanatory variables (parameters) and for other selected equations, which were determined to be the best of the equations tested with one to four variables, are presented in appendix 1. Data on the drainage area, areal percentage of sand and gravel deposits, and areal percentage of forest land for the 285 stream sites used in the equation development, as well as for the 66 stream sites with drainage areas greater than 2.00 mi^2, are provided in table 8 (back of report).

Confidence-Interval Estimation

An important adjunct to estimating the probability of a stream flowing perennially in Massachusetts by using a logistic regression equation is estimation of the confidence interval of each probability. The confidence intervals provide estimates of the upper and lower limits of the probability estimation of the equation. For example, the 95-percent confidence intervals are bounded by upper and lower limits between which the true estimate has a 95-percent chance of being found. The methods used to calculate the upper and lower limits of the 95-percent confidence intervals are outlined in Hosmer and Lemeshow (2000, p. 17–21, 40–42, and 85–88) and Cook (2001, p. 9–11, 29–30, 36–39, and 46–50).

The equations for variance and standard error are

$$\begin{aligned}\text{variance} = {} & 0.372729 + 0.034308\left(\ln\left(x_1\right)\right)^2 + 0.000032\left(x_2\right)^2 + 0.000057\left(x_3\right)^2 \\ & +0.102893\left(x_4\right)^2 + 2\left(\ln\left(x_1\right)\right)\left(0.048903\right) + 2\left(x_2\right)\left(-0.001272\right) + 2\left(x_3\right)\left(-0.004189\right) \\ & +2\left(x_4\right)\left(0.018564\right) + 2\left(\ln\left(x_1\right)\right)\left(x_2\right)\left(0.000131\right) + 2\left(\ln\left(x_1\right)\right)\left(x_3\right)\left(-0.000306\right) \\ & +2\left(\ln\left(x_1\right)\right)\left(x_4\right)\left(0.003397\right) + 2\left(x_2\right)\left(x_3\right)\left(0.000011\right) + 2\left(x_2\right)\left(x_4\right)\left(0.000163\right) \\ & +2\left(x_3\right)\left(x_4\right)\left(-0.000800\right),\end{aligned} \tag{3}$$

and

$$standard\ error = \left(variance\right)^{1/2}. \tag{4}$$

The equations for the lower and upper limits of the 95-percent confidence intervals are

$$lower\ limit = \frac{exp\left(2.8084 + 0.9884ln\left(x_1\right) + 0.0111x_2 - 0.0233x_3 + 0.7500x_4 - \left(1.96\right)\left(standard\ error\right)\right)}{1 + exp\left(2.8084 + 0.9884ln\left(x_1\right) + 0.0111x_2 - 0.0233x_3 + 0.7500x_4 - \left(1.96\right)\left(standard\ error\right)\right)}, \tag{5}$$

and

$$upper\ limit = \frac{exp\left(2.8084 + 0.9884ln\left(x_1\right) + 0.0111x_2 - 0.0233x_3 + 0.7500x_4 + \left(1.96\right)\left(standard\ error\right)\right)}{1 + exp\left(2.8084 + 0.9884ln\left(x_1\right) + 0.0111x_2 - 0.0233x_3 + 0.7500x_4 + \left(1.96\right)\left(standard\ error\right)\right)}. \tag{6}$$

In the lower and upper limit equations (5) and (6), the 1.96 value comes from the *Z*-distribution (standard normal distribution) table for $Z_{1-\alpha/2}$, where $\alpha = 0.05$ (1.0–0.95, where 0.95 is the confidence interval).

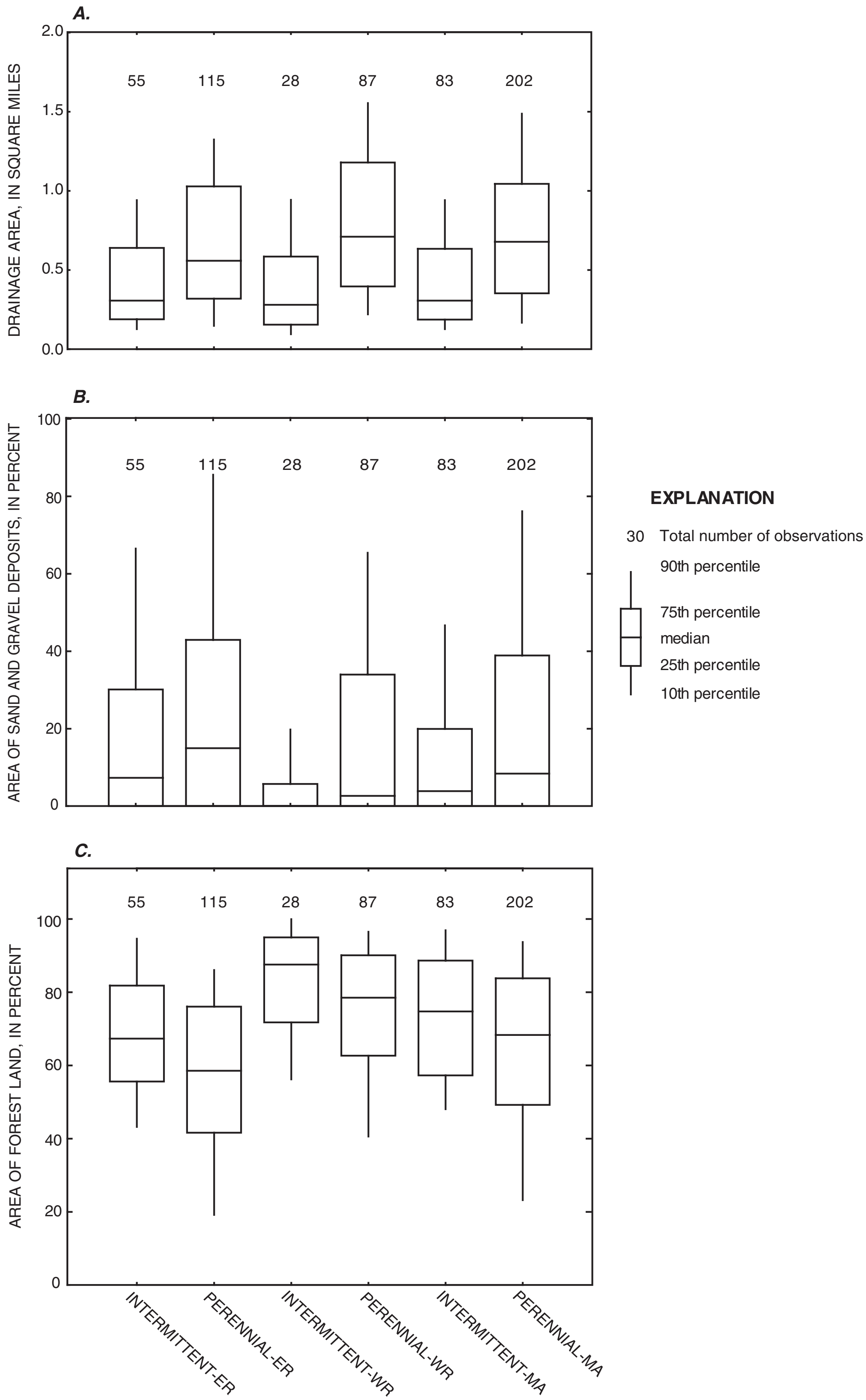


Figure 6. The distribution of *A*, drainage area; *B*, areal percentage of sand and gravel deposits; and *C*, areal percentage of forest land for intermittent- and perennial-stream sites in the eastern region of Massachusetts (ER), western region of Massachusetts (WR), and throughout Massachusetts (MA).

Table 4. Analysis of maximum likelihood estimates for the logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts.

[<, less than]

Explanatory variable	Degrees of freedom	Estimate	Standard error	Chi-square	*p*-value
Intercept	1	2.8084	0.6105	21.1608	<0.0001
Drainage area (natural logarithm – ln)	1	.9884	.1852	28.4758	<.0001
Areal percentage of sand and gravel deposits	1	.0111	.0056	3.8979	.0483
Areal percentage of forest land	1	-.0233	.0076	9.5072	.0020
Region[1]	1	.7500	.3208	5.4663	.0194

[1] Region is either the eastern or western region of Massachusetts. See figure 1 for region delineations.

Verification

In the Shawsheen River Basin to verify the logistic regression equation, 28 stream sites (not included in the development of the equation) with drainage areas less than 2.00 mi^2 were observed to determine their intermittent or perennial stream status on July 31, and August 1, 2002 (fig. 7 and table 9, at back of report). The 28 stream sites were observed during flow durations between 85 and 92 percent, which is within the range of flow durations for the 285 stream sites used to develop the equation (figs. 4*B* and 4*C*). Predictions of the intermittent and perennial status of the sites were compared to their observed stream status. The equation correctly estimated the intermittent or perennial status compared to the observed stream status by using a probability cutpoint of 0.56 for 20 of the 28 stream sites (71.4 percent). Of the 8 stream sites incorrectly classified, 5 sites observed to be intermittent were estimated to be perennial, and 3 sites observed to be perennial were estimated to be intermittent.

To verify the logistic regression equation further, predictions by the equation of the intermittent and perennial status of 84 stream sites with drainage areas less than 2.00 mi^2 in the South Coastal Basin were compared to their observed status (fig. 8 and table 10, at back of report). These sites were observed during mid-July through early September 1999 and were used in the study by Bent and Archfield (2002). The 84 stream sites were observed during flow durations between 73 and 99 percent, with 92 percent of the observations between 80 and 96 percent, which is within the range of flow durations for the 285 stream sites used to develop the equation (figs. 4*B* and 4*D*). The equation correctly estimated the intermittent or perennial status, compared to the observed stream status, by using a probability cutpoint of 0.56 for 53 of the 84 stream sites (63.1 percent). Of the 31 stream sites incorrectly classified, 23 sites observed to be intermittent were estimated to be perennial, and 8 sites observed to be perennial were estimated to be intermittent.

Application

An example application of the equation is provided for stream site GM-06 (fig. 2 and table 9, at back of report), Unnamed Tributary to Lake Buel at State Rt. 23 in Great Barrington, Massachusetts in the Housatonic River Basin. Flow was observed at this stream site on August 23, 2001, and the site is represented as intermittent on the USGS topographic map (Great Barrington, Mass.–N.Y.). The values of the explanatory variables used in the logistic regression equation were the natural logarithm of the drainage area of 0.53 mi^2 (x_1), the areal percentage of sand and gravel deposits of 0.00 percent (x_2), and the areal percentage of forest land of 78.76 percent (x_3). The integer location variable (x_4), was 1 because the stream site is located in the western region of the state.

The probability calculated by equation 2 that stream site GM-06 would be perennial is

$$P = \frac{exp\left(2.8084 + 0.9884\,ln(0.53) + 0.0111(0.00) - 0.0233(78.76) + 0.7500(1)\right)}{1 + exp\left(2.8084 + 0.9884\,ln(0.53) + 0.0111(0.00) - 0.0233(78.76) + 0.7500(1)\right)}$$
$$= 0.75.$$

The upper and lower limits of the 95-percent confidence interval calculated by equations 3–6 for the probability that stream site GM-06 would be perennial is

$$
\begin{aligned}
variance &= 0.0372729 + 0.034308(0.53)^2 + 0.000032(0.00)^2 + 0.000057(78.76)^2 \\
&+ 0.102893(1)^2 + 2(0.53)(0.048908) + 2(0.00)(-0.001272) + 2(78.76)(-0.004189) \\
&+ 2(1)(0.018564) + 2(0.53)(0.00)(0.000131) + 2(0.53)(78.76)(-0.000306) \\
&+ 2(0.53)(1)(0.003397) + 2(0.00)(78.76)(0.000011) + 2(0.00)(1)(0.000163) \\
&+ 2(78.76)(1)(-0.000800) \\
&= 0.06;
\end{aligned}
$$

$$standard\ error = (0.06)^{1/2} = 0.24;$$

$$lower\ limit = \frac{exp\left(2.8084 + 0.9884\,ln(0.53) + 0.0111(0.00) - 0.0233(78.76) + 0.7500(1) - (1.96)(0.24)\right)}{1 + exp\left(2.8084 + 0.9884\,ln(0.53) + 0.0111(0.00) - 0.0233(78.76) + 0.7500(1) - (1.96)(0.24)\right)}$$
$$= 0.65;$$

and

$$upper\ limit = \frac{exp\left(2.8084 + 0.9884\,ln(0.53) + 0.0111(0.00) - 0.0233(78.76) + 0.7500(1) + (1.96)(0.24)\right)}{1 + exp\left(2.8084 + 0.9884\,ln(0.53) + 0.0111(0.00) - 0.0233(78.76) + 0.7500(1) + (1.96)(0.24)\right)}$$
$$= 0.83.$$

Thus, for stream site GM-06, the probability of the stream flowing perennially is 0.75, and the lower and upper limits of the 95-percent confidence interval are 0.65 and 0.83.

Probabilities and the lower and upper limits for the 95-percent confidence intervals calculated by the logistic regression equation are given for the 351 stream sites visited throughout Massachusetts in 2001, the 28 sites visited in the Shawsheen River Basin in 2002, and 84 sites visited in the South Coastal Basin in 1999 in tables 8–10 (at back of report).

Probability Cutpoint

A probability cutpoint in logistic regression analyses is the probability level chosen as the boundary between the two response groups known as "event" or "nonevent." Computed probabilities equal to or greater than the cutpoint are classified as an "event" and those less than the cutpoint are classified as a "nonevent." In this study, an "event" is a classification of a stream site as perennial, and a "nonevent" is a classification of a site as intermittent. In general, the higher the probability cutpoint, the more likely a stream site would be classified as intermittent and the greater the likelihood that a perennial stream site could be incorrectly classified as intermittent. Conversely, the lower the cutpoint used, the more likely a stream site would be classified as perennial and the greater the likelihood that a intermittent stream site could be incorrectly classified as perennial. The determination of the probability cutpoint requires an evaluation of the accuracy of the equation in classifying events and nonevents, in balancing incorrect classification of events and nonevents, or both.

The evaluation of the predictive accuracy of the equation in this study was done with the classification table (table 5) produced during the logistic regression analyses. At a probability cutpoint of 0.56, the equation reaches its maximum level of correctness (75.1 percent), but if a stream site is incorrectly classified at this value, an intermittent site would be more likely to be incorrectly classified as perennial than a perennial site to be incorrectly classified as intermittent. Sensitivity is the ratio of correctly classified events to the total number of events. At a cutpoint of 0.56, the sensitivity is 87.6 percent,

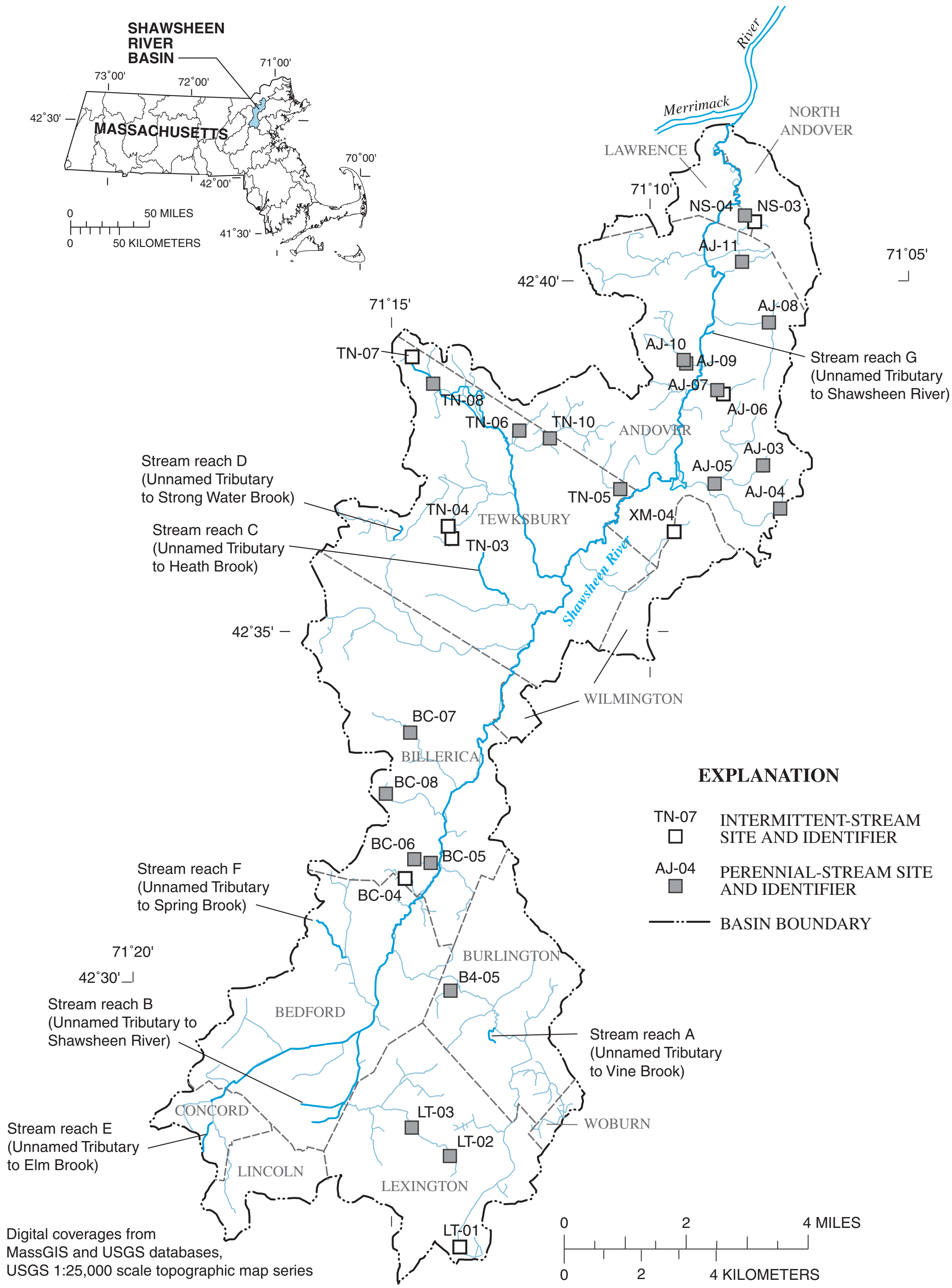


Figure 7. Locations of field-visited stream sites designated as intermittent or perennial used in verification of the logistic regression equation for estimating the probability of a stream flowing perennially in the Shawsheen River Basin, northeastern Massachusetts, late July and early August 2002. Seven selected headwater stream reaches are also shown.

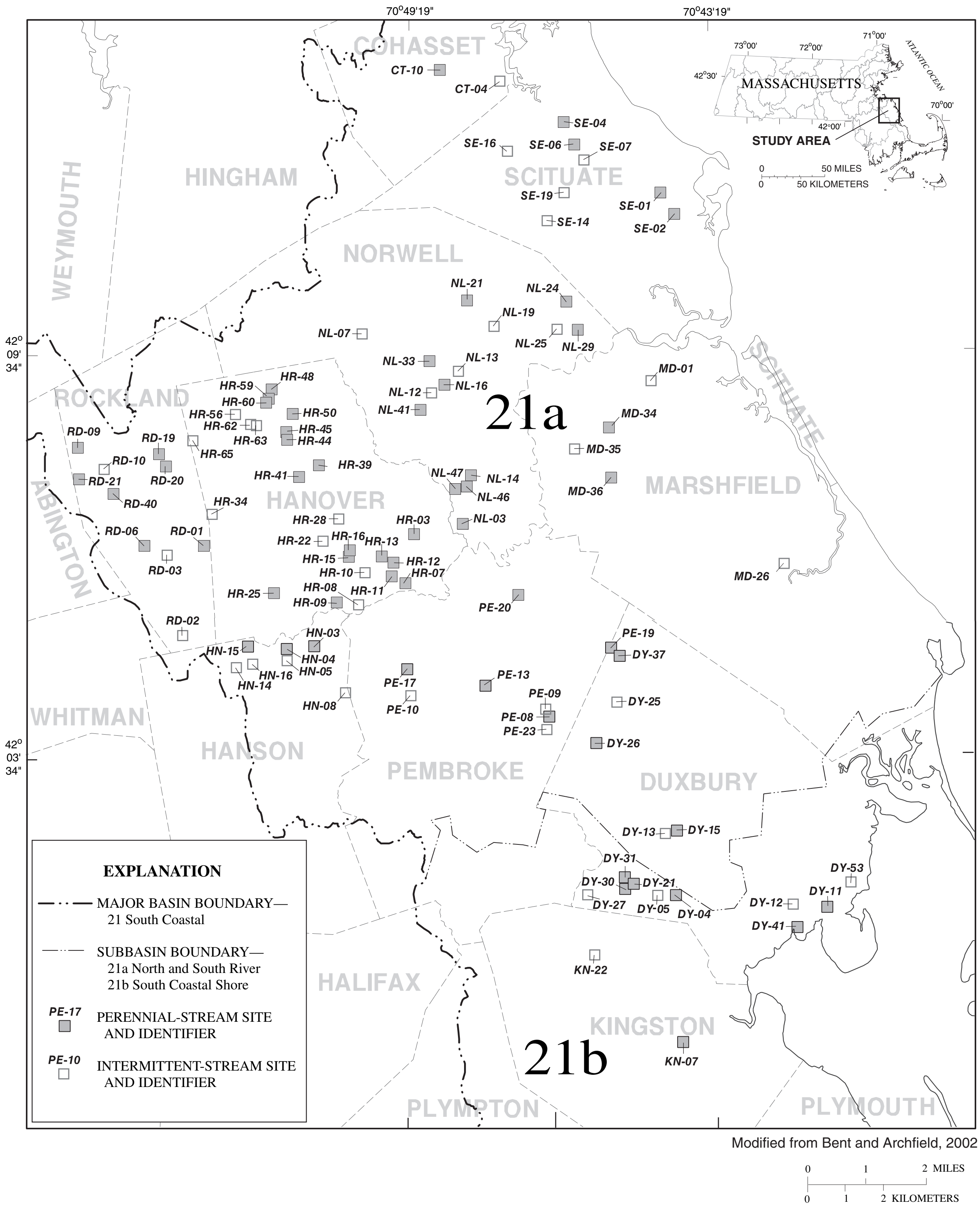


Figure 8. Locations of field-visited stream sites designated as intermittent or perennial used in verification of the logistic regression equation for estimating the probability of a stream flowing perennially in the South Coastal Basin, southeastern Massachusetts, mid-July through early September 1999.

Table 5. Classification table for probability levels for the logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts.

[**Correct:** The frequency with which observations are correctly classified. **Incorrect:** The frequency with which observations are incorrectly classified. **Event:** A perennial observation. **Nonevent:** An intermittent observation. **Percentages:** **Correct:** The probability that the equation correctly classifies the sample data for each probability cutpoint. **Sensitivity:** The ratio of correctly classified events to the total number of events. **Specificity:** The ratio of correctly classified nonevents to the total number of nonevents. **False positive:** The ratio of the number of nonevents incorrectly classified as events to the sum of all observations classified as events. **False negative:** The ratio of the number of events incorrectly classified as nonevents to the sum of all observations classified as nonevents. SAS Institute, Inc., 1995, p. 45–50]

Probability level (cutpoint)	Correct		Incorrect		Percentages				
	Event	Nonevent	Event	Nonevent	Correct	Sensitivity	Specificity	False positive	False negative
0.00	202	0	83	0	70.9	100.0	0.0	29.1	0.0
.02	202	0	83	0	70.9	100.0	.0	29.1	.0
.04	202	0	83	0	70.9	100.0	.0	29.1	.0
.06	202	0	83	0	70.9	100.0	.0	29.1	.0
.08	202	0	83	0	70.9	100.0	.0	29.1	.0
.10	202	0	83	0	70.9	100.0	.0	29.1	.0
.12	202	0	83	0	70.9	100.0	.0	29.1	.0
.14	202	1	82	0	71.2	100.0	1.2	28.9	.0
.16	202	2	81	0	71.6	100.0	2.4	28.6	.0
.18	201	2	81	1	71.2	99.5	2.4	28.7	33.3
.20	201	2	81	1	71.2	99.5	2.4	28.7	33.3
.22	201	2	81	1	71.2	99.5	2.4	28.7	33.3
.24	201	2	81	1	71.2	99.5	2.4	28.7	33.3
.26	199	5	78	3	71.6	98.5	6.0	28.2	37.5
.28	199	8	75	3	72.6	98.5	9.6	27.4	27.3
.30	199	9	74	3	73.0	98.5	10.8	27.1	25.0
.32	198	9	74	4	72.6	98.0	10.8	27.2	30.8
.34	196	10	73	6	72.3	97.0	12.0	27.1	37.5
.36	196	11	72	6	72.6	97.0	13.3	26.9	35.3
.38	196	11	72	6	72.6	97.0	13.3	26.9	35.3
.40	193	12	71	9	71.9	95.5	14.5	26.9	42.9
.42	192	13	70	10	71.9	95.0	15.7	26.7	43.5
.44	190	16	67	12	72.3	94.1	19.3	26.1	42.9
.46	190	20	63	12	73.7	94.1	24.1	24.9	37.5
.48	189	25	58	13	75.1	93.6	30.1	23.5	34.2
.50	186	25	58	16	74.0	92.1	30.1	23.8	39.0
.52	183	26	57	19	73.3	90.6	31.3	23.8	42.2
.54	180	34	49	22	75.1	89.1	41.0	21.4	39.3
.56	177	37	46	25	75.1	87.6	44.6	20.6	40.3
.58	173	39	44	29	74.4	85.6	47.0	20.3	42.6
.60	171	41	42	31	74.4	84.7	49.4	19.7	43.1
.62	164	43	40	38	72.6	81.2	51.8	19.6	46.9
.64	159	45	38	43	71.6	78.7	54.2	19.3	48.9
.66	155	49	34	47	71.6	76.7	59.0	18.0	49.0
.68	145	49	34	57	68.1	71.8	59.0	19.0	53.8
.70	135	52	31	67	65.6	66.8	62.7	18.7	56.3
.72	131	56	27	71	65.6	64.9	67.5	17.1	55.9
.74	124	58	25	78	63.9	61.4	69.9	16.8	57.4
.76	113	60	23	89	60.7	55.9	72.3	16.9	59.7
.78	102	63	20	100	57.9	50.5	75.9	16.4	61.3

Table 5. Classification table for probability levels for the logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts.—Continued

[**Correct:** The frequency with which observations are correctly classified. **Incorrect:** The frequency with which observations are incorrectly classified. **Event:** A perennial observation. **Nonevent:** An intermittent observation. **Percentages:** **Correct:** The probability that the equation correctly classifies the sample data for each probability cutpoint. **Sensitivity:** The ratio of correctly classified events to the total number of events. **Specificity:** The ratio of correctly classified nonevents to the total number of nonevents. **False positive:** The ratio of the number of nonevents incorrectly classified as events to the sum of all observations classified as events. **False negative:** The ratio of the number of events incorrectly classified as nonevents to the sum of all observations classified as nonevents. SAS Institute, Inc., 1995, p. 45–50]

Probability level (cutpoint)	Correct		Incorrect		Percentages				
	Event	Nonevent	Event	Nonevent	Correct	Sensitivity	Specificity	False positive	False negative
0.80	96	68	15	106	57.5	47.5	81.9	13.5	60.9
.82	86	73	10	116	55.8	42.6	88.0	10.4	61.4
.84	67	74	9	135	49.5	33.2	89.2	11.8	64.6
.86	59	75	8	143	47.0	29.2	90.4	11.9	65.6
.88	47	78	5	155	43.9	23.3	94.0	9.6	66.5
.90	38	80	3	164	41.4	18.8	96.4	7.3	67.2
.92	26	81	2	176	37.5	12.9	97.6	7.1	68.5
.94	17	83	0	185	35.1	8.4	100.0	.0	69.0
.96	6	83	0	196	31.2	3.0	100.0	.0	70.3
.98	2	83	0	200	29.8	1.0	100.0	.0	70.7
1.00	0	83	0	202	29.1	.0	100.0	.0	70.9

because 177 of the 202 events are correctly classified. Specificity is the ratio of correctly classified nonevents to the total number of nonevents. At a cutpoint of 0.56, the specificity is 44.6 percent, because 37 of 83 nonevents were correctly classified. Sensitivity differs from false positive, which is the ratio of the number of nonevents incorrectly classified as events to the sum of all observations classified as events. At a cutpoint of 0.56, the false positive is 20.6 percent, which is the ratio of 46 nonevents incorrectly classified as events to the sum of 223 events. Specificity differs from false negative, which is the ratio of the number of events incorrectly classified as nonevents to the sum of all observations classified as nonevents. At a cutpoint of 0.56, the false negative is 40.3 percent, which is the ratio of 25 events incorrectly classified as nonevents to the sum of 62 nonevents.

Other cutpoint probabilities could be used. For example, at a probability cutpoint of 0.54, the equation also had a level of correctness equal to 75.1 percent (table 5), but the difference between the sensitivity and specificity is 5.1 percent greater than at the probability cutpoint of 0.56. At a cutpoint of 0.72, the sensitivity (64.9 percent) and specificity (67.5 percent) are about equal, but the correctness of the equations drops to 65.6 percent.

If a probability cutpoint of 0.56 were used to classify stream sites in the eastern region of the state (region variable = 0), the maximum drainage area of an intermittent stream site would be about 0.79 mi^2 if the areal percentage of sand and gravel were 0 percent and the areal percentage of forest land were 100 percent. The minimum drainage area for a perennial stream site would be about 0.03 mi^2 if the areal percentage of sand and gravel were 100 percent and the areal percentage of forest land were 0 percent. Stream site DU-04 (Unnamed Tributary to Wallis Pond, State Route 16, Douglas) had the largest drainage area (0.56 mi^2) for sites that were classified intermittent (0.53 probability) in the eastern region of the state by using a probability cutpoint of 0.56 (table 8, at back of report). This stream site was observed to be perennial. Stream site ML-01 (Unnamed Tributary to Flynns Pond, High Street, Medfield) had the smallest drainage area (0.09 mi^2) for sites that were classified perennial (0.57 probability) in the eastern region. This stream site was observed to be perennial.

If a probability cutpoint of 0.56 were used to classify stream sites in the western region of the state (region variable = 1), the maximum drainage area of an intermittent stream site would be about 0.37 mi^2 if the areal percentage of sand and gravel were 0 percent and the areal percentage of forest land were 100 percent. The minimum drainage area for a perennial stream site would be about 0.02 mi^2 if the areal percentage of sand and gravel were 100 percent and the areal percentage of forest land were 0 percent. Stream site PD-01 (Unnamed Tributary to Amethyst Brook, North Valley Road, Pelham) had the largest drainage area (0.28 mi^2) for sites that were classified intermittent (0.52 probability) in the western region of the state by using a probability cutpoint of 0.56 (table 8, at back of report). This stream site was observed to be intermittent. Stream site SV-05 (Unnamed Tributary to Kellog Brook, State Route 10/State Route 202, Southwick) had the smallest drainage area (0.07 mi^2) for sites that were classified perennial (0.58 probability) in the western region. This stream site was observed to be intermittent.

Comparison of Equation to Other Methods of Determining the Intermittent or Perennial Status of Streams

To assess the accuracy of the logistic regression equation in estimating the intermittent or perennial status of stream sites, the observed stream statuses of the 351 sites were also compared to the status depicted on USGS topographic maps (original Regulations), the status predicted under the revised Regulations (December 20, 2002), the previously published logistic regression equation (Bent and Archfield, 2002), and selected low-flow statistics generated by STREAMSTATS (Ries and others, 2000, Ries and Friesz, 2000). A probability cutpoint of 0.56 was used for the equation determined for this study. For the previously published logistic regression equation, a probability cutpoint of 0.38 (representing the maximum level of correctness of about 72 percent for that equation) was used. The selected low-flow statistics evaluated by STREAMSTATS were the 99-, 98-, 95-, 90-, 85-, and 80-percent flow durations. If the streamflow was estimated to be 0.00 ft^3/s for any of the flow durations, then the stream site was classified as intermittent; if it was 0.01 ft^3/s or greater, then the stream site was classified as perennial. For stream sites in the Buzzards Bay Basin, North Coastal Basin, and Taunton River Basin (fig. 1) for which STREAMSTATS currently (2006) does not function, the explanatory variables (drainage area, area of stratified deposits, length of streams, mean basin slope, and region of the state) were determined by the same GIS methods used in STREAMSTATS.

The comparison among methods for correctly classifying the intermittent or perennial status of the 285 stream sites with drainage areas less than 2.00 mi^2 determined that the logistic regression equation used in this study was 76.5 percent accurate (table 6). The 76.5 percent is slightly greater than the 75.1 percent correct at the 0.56-probability level reported in table 5. This 1.4-percent difference was the result of differences in the calculation techniques. The correctness of the logistic regression equation in this study (table 6) was evaluated by determining the percentage of stream sites whose classification as intermittent or perennial with the 0.56-probability cutpoint matched the observed stream status. This was the only mechanism to evaluate the different methods for correctly classifying the intermittent or perennial status of stream sites within drainage-area ranges from 0.00 to 2.00 mi^2. Differences between the two calculation techniques are discussed in detail by SAS Institute, Inc. (1995, p. 35–36, 39–40, and 49–50). For the remaining methods, (1) USGS topographic maps (original Regulations) correctly depicted 62.8 percent of the 285 stream sites, (2) the revised Regulations (Rivers Protection Act, December 20, 2002) correctly classified 69.1 percent, (3) the previously published logistic regression equation (Bent and Archfield, 2002) correctly classified 37.5 percent, and (4) the selected low-flow statistics correctly classified 58.9 to 72.3 percent (table 6). Of the alternative methods, the 98- and 90-percent flow durations classified the most stream sites correctly (72.3 percent).

The logistic regression equation used in this study was as accurate or more accurate than the other methods in classifying the stream status in 7 of the 10 drainage-area ranges from 0.00 to 0.99 mi^2 (table 6). Only in the drainage-area ranges from 0.00 to 0.09, from 0.50 to 0.59, and from 0.80 to 0.89 mi^2 did the previously published equation by Bent and Archfield (2002) (by one stream site), the revised Regulations and the selected low-flow statistics (by two to three stream sites), and the revised Regulations (by one site), respectively, prove to be more accurate than the equation used here in matching the observed stream status. In the three drainage-area ranges from 1.00 to 1.99 mi^2, the logistic regression equation used in this study, the revised Regulations, and the selected low-flow statistics (except the 99-percent flow duration) matched the observed stream status from 84.0 to 94.4 percent of the time. In the three drainage-area ranges from 2.00 to 10.99 mi^2, all methods had the same level of accuracy, except the previously published logistic regression equation (Bent and Archfield, 2002) for two of the drainage-area ranges. None of the methods were 100-percent accurate in the 3.00–4.99 mi^2 and 5.00–10.99 mi^2 drainage-area ranges, because they could not correctly predict the observed intermittent stream status at sites BA-03, Dry Brook in Bernardston (drainage area of 3.48 mi^2) and AE-01, Dry Brook in Adams (drainage area of 10.46 mi^2), respectively.

Additional comparison of the logistic regression equation used in this study to other methods in classifying the stream status was done by using the verification sites in the Shawsheen River Basin and the South Coastal Basin (previous study by Bent and Archfield, 2000) (table 7). In the Shawsheen River Basin, the logistic regression equation used in this study correctly classified 71.4 percent of the 28 stream sites, and only the low-flow statistics for the 98- to 80-percent flow durations were as accurate or more accurate. The 98-percent flow-duration method estimated the statuses of most sites correctly (85.7 percent). In the South Coastal Basin, the logistic regression equation used in this study correctly classified 60.7 percent of the 84 stream sites; the previously published logistic regression equation (Bent and Archfield, 2002) and the 80- to 95-percent flow-duration methods gave more accurate estimates. The 95-percent flow-duration method estimated the statuses of the most stream sites correctly (67.9 percent).

Overall, the logistic regression equation used for this study was more accurate than the other methods considered in correctly classifying the intermittent or perennial status of a stream site for the range of drainage areas from 0.00 to 2.00 mi^2 over the entire state. The reason why the selected low-flow statistics between the 98- and 80-percent flow durations were as accurate or more accurate than the equation used in this study in some ranges of drainage areas and areas of the state (Shawsheen River Basin and South Coastal Basin) is likely that 95 percent of the observations of stream status were between the 97- and 82-percent flow durations. Additionally,

Table 6. Percentage of matches between several different classification methods and the field-observed intermittent or perennial status of stream sites in Massachusetts, late July through early September 2001.

[ft^3/s, cubic feet per second; mi^2, square miles; USGS, U.S. Geological Survey]

Drainage area (mi^2)	Number of stream sites	Percentage									
		Logistic regression equation used in this study	USGS topographic map	Revised Regulations	Bent and Archfield (2002) logistic regression equation	STREAMSTATS					
						99-percent flow duration	98-percent flow duration	95-percent flow duration	90-percent flow duration	85-percent flow duration	80-percent flow duration
0.00–0.09	12	66.7	58.3	58.3	75.0	66.7	66.7	58.3	66.7	50.0	33.3
0.10–0.19	36	72.2	55.6	55.6	47.2	41.7	61.1	52.8	55.6	55.6	55.6
0.20–0.29	30	63.3	53.3	53.3	56.7	50.0	46.7	46.7	46.7	46.7	46.7
0.30–0.39	27	81.5	74.1	74.1	40.7	37.0	74.1	74.1	77.8	77.8	77.8
0.40–0.49	26	80.8	50.0	50.0	26.9	38.5	69.2	73.1	73.1	73.1	73.1
0.50–0.59	24	62.5	54.2	70.8	41.7	75.0	70.8	70.8	70.8	70.8	70.8
0.60–0.69	21	61.9	47.6	52.4	47.6	47.6	61.9	61.9	61.9	61.9	61.9
0.70–0.79	16	93.8	56.3	87.5	25.0	56.3	93.8	93.8	93.8	93.8	93.8
0.80–0.89	10	80.0	80.0	90.0	10.0	70.0	80.0	80.0	80.0	80.0	80.0
0.90–0.99	18	72.2	55.6	66.7	33.3	66.7	72.2	72.2	72.2	72.2	72.2
1.00–1.24	25	84.0	68.0	84.0	20.0	76.0	84.0	84.0	84.0	84.0	84.0
1.25–1.49	18	94.4	88.9	94.4	22.2	83.3	94.4	94.4	94.4	94.4	94.4
1.50–1.99	22	90.9	90.9	90.9	27.3	90.9	90.9	90.9	90.9	90.9	90.9
2.00–2.99	28	100.0	100.0	100.0	32.1	100.0	100.0	100.0	100.0	100.0	100.0
3.00–4.99	18	94.4	94.4	94.4	77.8	94.4	94.4	94.4	94.4	94.4	94.4
5.00–10.99	20	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0
[1]0.00–1.99	285	[2]76.5	62.8	69.1	37.5	58.9	72.3	71.2	72.3	71.6	70.9
[3]0.00–10.99	351	80.3	69.2	74.4	42.5	66.1	76.9	76.1	76.9	76.4	75.8

[1] Combination of 0.00–1.99 mi^2 drainage areas.

[2] The 76.5 percent is slightly greater than the 75.1 percent "correct" at a 0.56 probability level reported in table 5. This 1.4-percent difference was the result of differences in the calculation techniques. Table 5 was created with the classification table function in the SAS Institute, Inc. (1995, p. 45–50), statistical software discussed in the "Logistic Regression Equation" section, "Development" subsection, of the report. In the classification table, the observed intermittent or perennial stream status was matched with the status estimated by using the probability of a stream flowing perennially and a 0.56-probability cutpoint. If the probability was less than 0.56, the stream site was classified as intermittent; and if the probability was equal to or greater than 0.56, the stream site was classified as perennial. Differences between the two calculation techniques are discussed in SAS Institute, Inc. (1995, p. 39–40 and 45–50).

[3] Combination of 0.00–10.99 mi^2 drainage areas.

Table 7. Percentage of matches between several different classification methods and the field-observed intermittent or perennial status of stream sites in the Shawsheen River Basin, northeastern Massachusetts, late July through early August 2002, and in the South Coastal Basin, southeastern Massachusetts, mid-July through early September 1999.

[mi^2, square mile; ft^3/s, cubic feet per second; USGS, U.S. Geological Survey]

Drainage area (mi^2)	Number of stream sites	Percentage									
		Logistic regression equation used in this study	USGS topographic map	Revised Regulations	Bent and Archfield logistic regression equation (2002)	STREAMSTATS					
						99-percent flow duration	98-percent flow duration	95-percent flow duration	90-percent flow duration	85-percent flow duration	80-percent flow duration
Shawsheen River Basin											
0.00–1.99	28	71.4	53.6	60.7	28.6	60.7	85.7	82.1	75.0	75.0	71.4
South Coastal Basin											
0.00–1.99	84	60.7	52.4	52.4	65.5	48.8	59.5	67.9	62.1	64.3	61.9

the previously published logistic regression equation developed by Bent and Archfield (2002) was slightly more accurate (by four stream sites or 4.8 percent) than the equation used in this study in correctly classifying the status of a stream in the South Coastal Basin. This greater accuracy of the previously published equation (Bent and Archfield, 2002) is likely the result of its specific location variable based on field observations at 132 stream sites in the South Coastal Basin, whereas the equation used in this study includes field observations at only 24 sites in the basin.

Limitations of the Logistic Regression Equation and Areas for Further Study

The logistic regression equation is applicable for stream sites with drainage areas between 0.04 and 2.00 mi^2 in Massachusetts, because this was the range of drainage areas used in equation development. The equation, which is based on data from naturally flowing streams, should be applicable to most stream sites because regulations, such as dams, surface-water withdrawals, ground-water withdrawals (pumping wells), diversions, wastewater discharges, and so forth, generally would not occur in basins smaller than 2.00 mi^2. If a stream site is regulated, the equation estimates the intermittent or perennial stream status as if the stream were naturally flowing. The equation is not applicable for losing stream reaches, because the equation would tend to overpredict the probability of a stream flowing perennially at a site. Examples of losing stream reaches are stream sites AE-01, Dry Brook in Adams, and BA-03, Dry Brook in Bernardston (table 8, at back of report).

The logistic regression equation is not applicable in areas of Massachusetts where ground-water contributing areas and surface-water drainage areas to stream sites differ appreciably. This condition may be present in southeastern Massachusetts, particularly for streams draining the Southeast Coastal Region—the southern part of the South Coastal Basin, the eastern part of the Buzzards Bay Basin, and the entire area of the Cape Cod and Islands Basins (fig. 1). In these areas, ground water can flow from one basin into another; therefore, in basins whose ground-water contributing areas are larger than their surface-water drainage areas, the equation would likely underpredict the probability that a stream is perennial. Conversely, in areas whose ground-water-contributing areas are smaller than their surface-water drainage areas, the equation would likely overpredict the probability that a stream is perennial.

The accuracy of the logistic regression equation is a function of the quality of the data used in the equation development. These data include the observed intermittent or perennial status of a stream site, the occurrence of unknown regulation above a site, and the measured basin characteristics. Basin characteristics of the stream sites used in the development of the regression logistic equation are limited by the accuracy of the digital data layers used. In the future, digital data layers (such as hydrography, wetlands, surficial geology, soils, DEMs, and land use) will likely become available at scales with better resolutions than are currently (2006) available. These digital data layers likely would improve the accuracy of the measured basin characteristics used as explanatory variables to predict the probability of a stream flowing perennially, but would require re-examination of the logistic regression equation. Digital data layers under development that could improve the equation include: (1) county-level soil-survey maps referred to as SSURGO database (U.S. Department of Agriculture, Natural Resources Conservation Service, Soil Survey Division, 2004); (2) 1970–2000 climate data available through the Parameter-elevation Regressions on

Independent Slopes Model (PRISM) climate-mapping system of Spatial Climate Analysis Service at Oregon State University (*http://www.ocs.oregonstate.edu/prism/*); and (3) statewide wetlands (1:12,000 scale) interpreted from stereo color-infrared photography (MassGIS, 2004d).

The logistic regression equation could be incorporated into a Web-based application of the USGS Office of Surface Water STREAMSTATS Program (*http://water.usgs.gov/osw/programs/streamstats.html*). A map-based interface is used in the Web-based application to allow a user to point and click on any stream site and have the application calculate selected streamflow statistics for an ungaged site or view available selected streamflow statistics for a gaged site. In a similar manner, a user could click on any stream site in Massachusetts and have the equation estimate the probability of a stream flowing perennially with the 95-percent confidence intervals.

Part 2. An Automated Procedure for Mapping Perennially Flowing Streams

By Peter A. Steeves [1], Gardner C. Bent [1], and Jennifer R. Hill [2]

Mapping the Intermittent and Perennial Reaches of Streams

The revised logistic regression equation was coupled with an automated procedure for mapping intermittent and perennial stream reaches in Massachusetts and tested for the Shawsheen River Basin in northeastern Massachusetts (fig. 7). The automated procedure utilizes ArcView GIS software including a toolkit developed for the National Hydrography Dataset (NHD). A map showing the transition points between intermittent and perennial stream reaches and the stream reaches in the Shawsheen River Basin is presented on a CD-ROM (appendix 2, in back pocket). The CD-ROM also contains ArcReader 9.0, a freeware product, that allows a user to zoom in and out, set a scale, pan, turn on and off map layers (such as a USGS topographic map), and print a map of the stream site with a scale bar. This CD-ROM provides an example of the map products that can be used by MDEP and city and town Conservation Commissions, as well as by others, for determining the intermittent or perennial status of a stream site in Massachusetts.

National Hydrography Dataset and Geographic Information System Tools

The automated procedure requires several preprocessing steps of the NHD (U.S. Geological Survey, 2005d). First, the NHD needs to be available at a 1:24,000 scale in a shapefile or coverage format. The NHD comprises surface-water features such as lakes, ponds, rivers, and streams; information about these features; and linkages between features (U.S. Geological Survey, 2005b). The linkage between features enables analysis and display of features in upstream or downstream order with use of GIS technology. These high-resolution stream-reach data are needed because of the spatial detail provided for headwater streams and for the associated small drainage area that generally defines the boundary between intermittent and perennial stream reaches. Second, elevation-derived data sets, like flow direction, flow accumulation, and catchment grids, need to be developed (U.S. Geological Survey, 2005d). Third, data layers and the values of the explanatory variables (basin characteristics) used in the logistic regression equation are required.

A set of GIS tools and applications has been specifically developed for use with NHD in ArcView. The tools are in the NHD ArcView Toolkit (U.S. Geological Survey, 2005e), and the applications include NHD Watershed and NHD Watershed Characteristics (U.S. Geological Survey, 2005f). The NHD ArcView Toolkit is a collection of ArcView extensions, written in Avenue programming language to assist in the understanding and use of NHD. One of the tools allows the user to navigate a hydrologic network (for example, to select reaches and other surface-water features that drain to a selected reach) (U.S. Geological Survey, 2005e). NHD Watershed allows users to delineate a drainage basin from any point on a stream (U.S. Geological Survey, 2005f). When a user selects a point, NHD Navigate (in the NHD ArcView Toolkit) first determines all reaches upstream of that point. NHD Watershed then selects the associated drainage basins of these reaches upstream of that point from a separate basin-boundary data layer. The reach from which the point was selected often requires further drainage-basin delineation from Digital Elevation Models (DEMs) data from that point to the next upstream reach. The drainage basin area of this upstream reach is then combined with the drainage basins of the reaches upstream of this reach to define the entire drainage basin to a selected point on a stream. NHD Watershed Characteristics summarizes watershed (basin) characteristics for the delineated drainage area (U.S. Geological Survey, 2005f). Stream reaches not shown on USGS topographic maps or upstream stream reaches which are not connected to a downstream reach are not part of NHD, and, thus, are not part of the automated mapping procedure.

Automated Mapping Procedure

The automated procedure steps through a selected basin by determining all starting (headwater) stream reaches for surface-water flow. The program then uses a search process to find the point along a stream reach where the flow status changes from intermittent to perennial. The following six steps are completed within the program (fig. 9):

[1]U.S. Geological Survey, MA–RI Water Science Center, Northborough, MA.

[2]Horizon Systems Corporation, Herndon, VA.

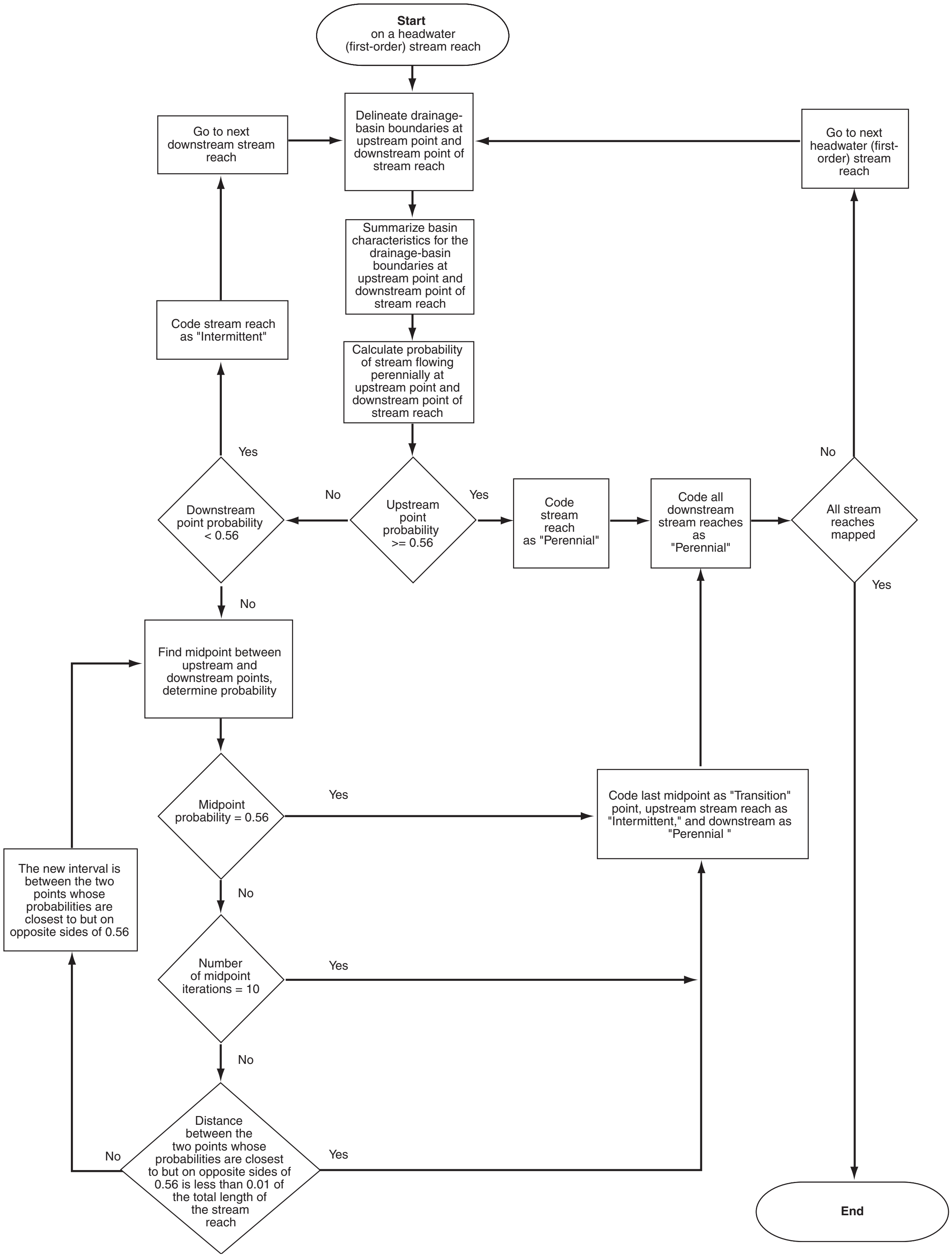


Figure 9. The steps of the automated mapping procedure for mapping intermittent and perennial stream reaches (I, intermittent; P, perennial; <, less than; >, greater than; =, equal to).

1. Delineate drainage-basin boundaries for an initial upstream point about 100 ft from the start of the stream reach and an initial downstream point about 100 ft from the end of the stream reach by using NHD Watershed. In the first iteration, the upstream point is near the headwaters of the reach, and the downstream point is near the mouth or the confluence with the next downstream reach.

2. Summarize basin characteristics (drainage areas, areal percentages of sand and gravel deposits, and areal percentages of forest land), as well as the region of the state (eastern or western) (figs. 1–3), for the upstream and downstream drainage basins by using NHD Watershed Characteristics.

3. Compute the probability of the stream flowing perennially at the upstream and downstream points of the stream reach by using the logistic regression equation.

4. Determine if the stream reach is intermittent, perennial, or changes from intermittent to perennial: (A) if the probability at the upstream point is greater than the probability cutpoint of 0.56, then the stream reach and all downstream stream reaches are coded "perennial" and the program goes to step 1 for the next first-order stream reach; (B) if the probability at the downstream point is less than 0.56, then the stream reach is coded "intermittent" and the program goes to step 1 for the next downstream stream reach; (C) if the probability at the downstream point is greater than 0.56 and the probability at the upstream point is less than 0.56, the transition point from intermittent to perennial is somewhere along the reach, and the program goes to step 5.

5. Find the midpoint between the upstream and downstream points of the stream reach and calculate the probability for that midpoint. Of the three points—upstream, midpoint, and downstream—choose the two whose probabilities are closest to but on opposite sides of 0.56. Then repeat steps 1, 2, and 3 for these two points, which define the new interval. The process iterates up and down as the two calculated probabilities approach 0.56 until: (A) the point whose probability is 0.56 is found; (B) 10 iterations are run; or (C) the distance between the two points is less than one-hundredth of the entire reach length. The last midpoint is coded as the "transition point from intermittent to perennial."

6. Go to step 1 and repeat steps until the entire basin is mapped.

Shawsheen River Basin Map Example

The automated mapping procedure was tested for the Shawsheen River Basin. The 78-mi^2 basin is in Essex and Middlesex Counties in northeastern Massachusetts (fig. 7), and includes parts of 11 towns and 1 city (Lawrence). The Shawsheen River flows northeastward for about 25 mi where it drains into the Merrimack River. Because the river loses only about 100 ft in altitude from its headwaters to its mouth (Gay and Delaney, 1980) and flows through wetlands for more than 50 percent of its length (Simcox, 1992) the river has a low gradient. Basin topography is characterized by low hills with altitudes from about 10 ft at its confluence with the Merrimack River to about 250 to 300 ft along its surface-water divide. Of its 18 lakes and ponds, 9 have areas greater than 10 acres (Simcox, 1992). The basin is about 61 percent underlain by stratified deposits, which are mainly in the lowland areas along the main stem of the Shawsheen River and its tributaries. The upland areas are primarily underlain by till and bedrock.

Results of applications of the NHD, NHD ArcView Toolkit, NHD Watershed, NHD Watershed Characteristics, and the logistic regression equation are provided for seven selected first-order stream reaches in the Shawsheen River Basin (fig. 7). The examples demonstrate how the probability of a stream flowing perennially, the basin characteristics used to solve the equation, and the 95-percent confidence intervals change along a reach (figs. 10 and 11). These examples also were done to test the automated mapping procedures developed for the study.

Seven different headwater (first-order) stream reaches (fig. 7) were selected for estimating the probability of a stream flowing perennially at incremental distances along the reach. Changes in the probability of a stream flowing perennially, the contributing drainage areas, the areal percentage of sand and gravel deposits, and the areal percentage of forest land are shown in relation to downstream distance from the headwaters for these stream reaches in figure 10. Of the seven first-order stream reaches, one reach is entirely intermittent (reach D—Unnamed Tributary to Strong Water Brook in Tewksbury), one reach is entirely perennial (reach E—Unnamed Tributary to Elm Brook in Concord), and five reaches change from intermittent to perennial on the basis of the probability cutpoint of 0.56. For first-order stream reaches A (Unnamed Tributary to Vine Brook in Burlington), B (Unnamed Tributary to Shawsheen River in Bedford), C (Unnamed Tributary to Heath Brook in Tewksbury), F (Unnamed Tributary to Spring Brook in Bedford), and G (Unnamed Tributary to Shawsheen River in Andover) there is a noticeable increase of at least 0.12 in the probability of the stream flowing perennially from less than 0.56 to greater than or equal to 0.56 over a small stream-reach distance from 0.01 to 0.03 mi. The noticeable increases in the probability for stream reaches A, B, and F result from increases in the contributing drainage areas of 0.17, 2.01, and 0.12 mi^2, respectively (fig. 10B). The large increase in drainage area for stream reach B is the result of most of the area of the runways and tarmac at Hanscom Air Force Base in Bedford, Massachusetts, becoming part of the contributing drainage area.

Typically, the transition from an intermittent to a perennial stream reach is caused by increases in the drainage area. Drainage area positively affects the probability of a stream flowing perennially; drainage area has the largest coefficient in

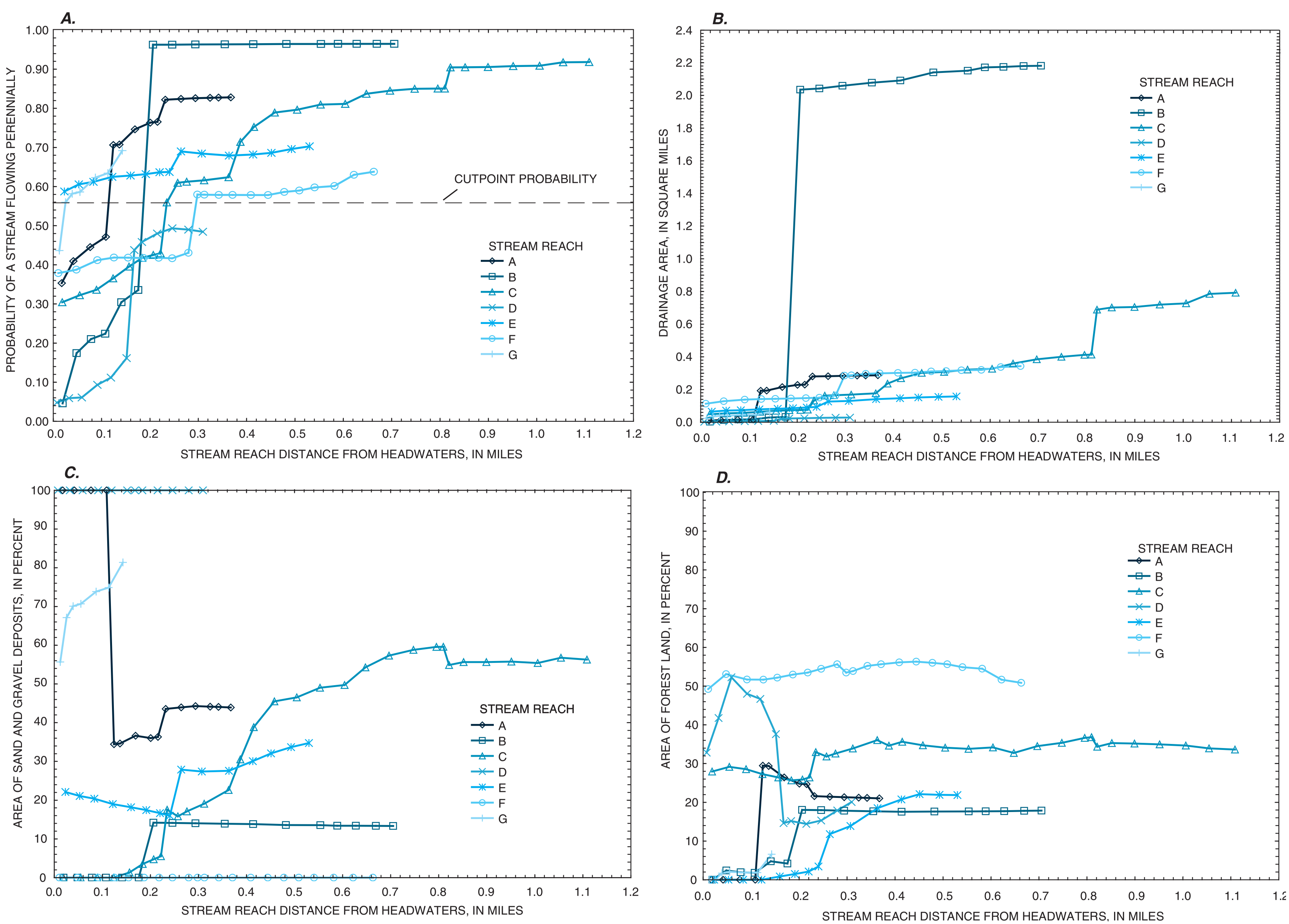


Figure 10. The calculated *A*, probability of a stream flowing perennially; *B*, drainage area; *C*, areal percentage of sand and gravel deposits; and *D*, areal percentage of forest land along seven selected headwater (first-order) stream reaches in the Shawsheen River Basin, northeastern Massachusetts. Stream reaches are shown in figure 7.

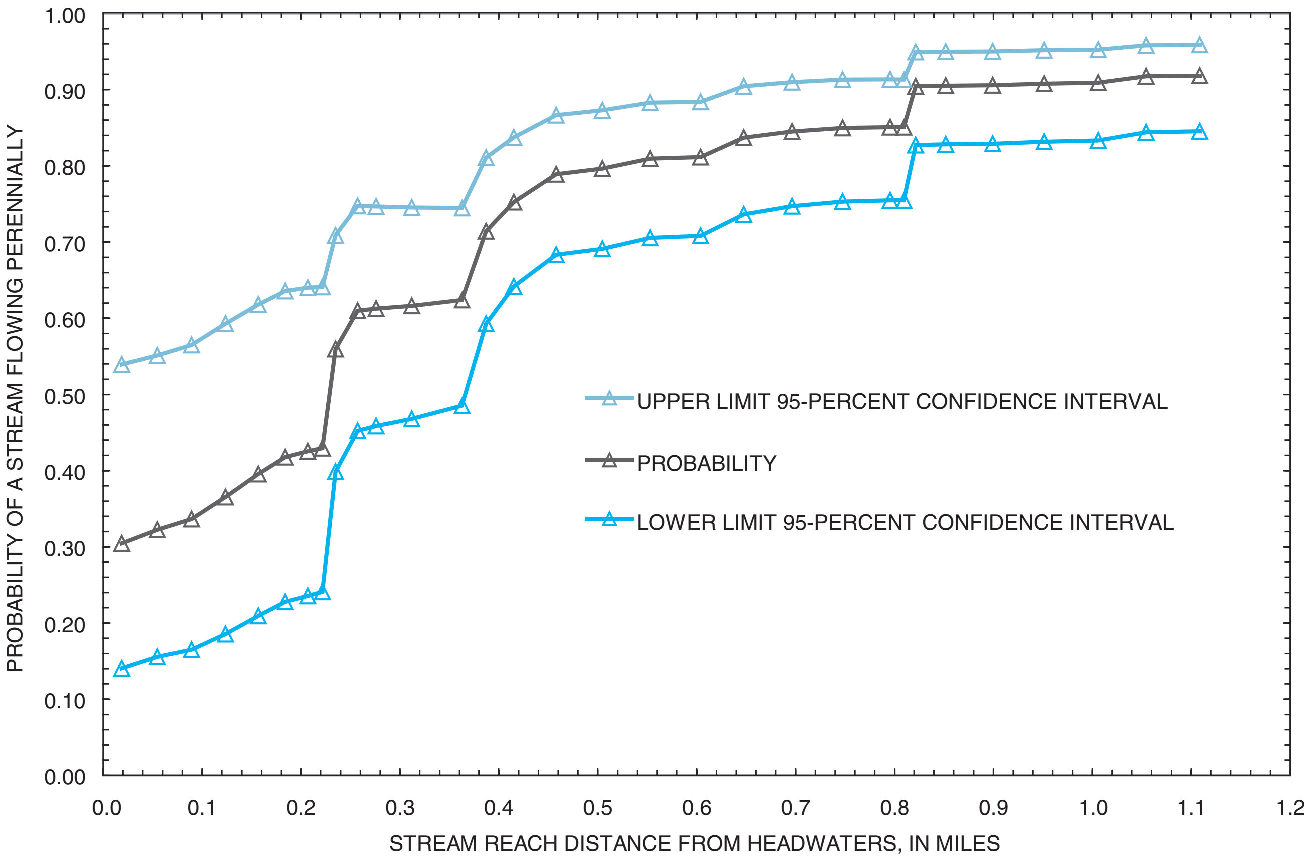


Figure 11. Probability of a stream flowing perennially and the upper and lower 95-percent confidence intervals along stream reach C in the Shawsheen River Basin, northeastern Massachusetts. Stream reach C shown in figure 7.

the logistic regression equation because it is the most significant explanatory variable in the equation, as shown for the one-, two-, three-, and four-variable equations in appendix 1. For stream reaches C and G, however, the drainage area only increases by 0.05 and 0.02 mi^2 (fig. 10B), respectively, as the probability increases from less than 0.56 to greater than or equal to 0.56. In the case of stream reaches C and G, the probability increases also appear to be related to increases of 12 and 11 percent, respectively, in the areal percentage of sand and gravel deposits (fig. 10C). The areal percentage of forest land (fig. 10D), which negatively affects the probability of a stream flowing perennially, appears to have a smaller effect on the probability than the areal percentage of sand and gravel deposits even though its coefficient in the logistic regression equation is larger than the coefficient of sand and gravel deposits. Generally, noticeable changes in the areal percentage of forest land are accompanied by drainage-area changes that outweigh those of the forest land. An exception to this general relation is illustrated by stream reach E from 0.26 to 0.36 mi, where the probability of the stream flowing perennially decreases slightly from about 0.69 to 0.68 mainly because of a 7-percent increase in the areal percentage of forest land.

The lower and upper 95-percent confidence intervals are also important in estimating the probability of a stream flowing perennially. The intervals can provide a range over which a stream reach is likely to change from intermittent to perennial for a given probability cutpoint. For example, on stream reach C in the Shawsheen River Basin (fig. 11), the upper limit of the 95-percent confidence interval for a probability cutpoint of 0.56 is calculated to be at about 0.08 mi from the headwaters, and the lower limit of the 95-percent confidence interval is calculated to be at about 0.38 mi from the headwaters. At this probability cutpoint, there is a 95-percent confidence that the transition point from an intermittent to a perennial stream reach is somewhere in this 0.30-mi segment of stream reach C.

In the Shawsheen River Basin, application of the automated mapping procedure indicated that 47 first-order stream reaches were entirely intermittent, 35 first-order stream reaches were entirely perennial, and 53 first-order stream reaches had a transition point from intermittent to perennial (fig. 12 and appendix 2). Of the 47 intermittent first-order stream reaches, 42 drain into a perennial stream reach and 5 have transition points from intermittent to perennial in downstream second- or third-order stream reaches.

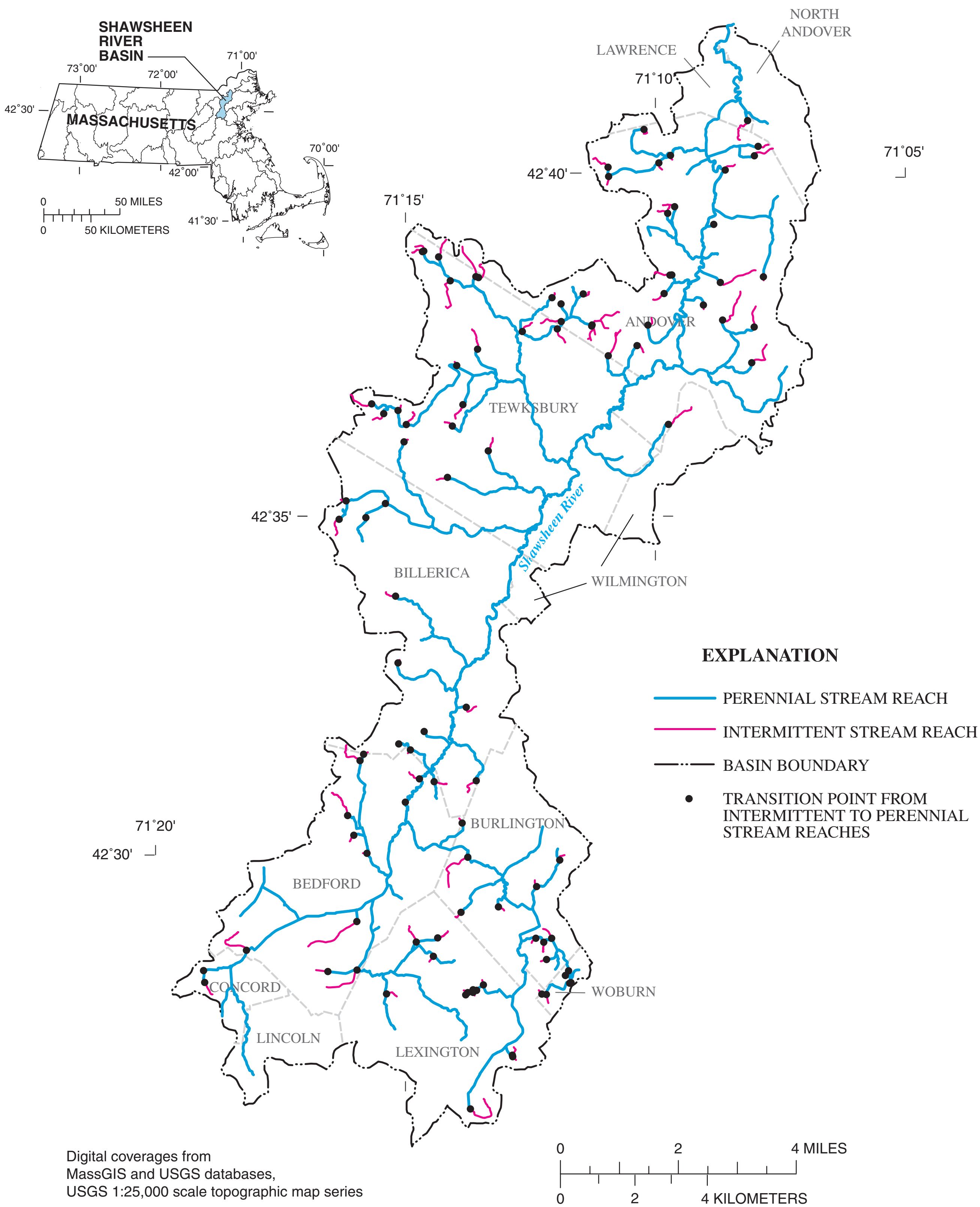


Figure 12. Intermittent and perennial stream reaches and transition points determined with the logistic regression equation and the automated mapping procedure, Shawsheen River Basin, northeastern Massachusetts.

Two types of errors were noted from visual quality assurance and quality control (QA/QC) of the map for the Shawsheen River Basin: (1) the drainage-basin boundary for an intermittent stream reach was delineated off (disconnected from) the centerline of the stream and (2) the drainage-basin boundary for an intermittent stream reach was delineated on the wrong stream. These errors, four of the first type and one of the second type, occurred only on five stream reaches. These errors were corrected manually independent of the automated mapping procedure.

Limitations

The maps produced with the automated mapping procedure of intermittent and perennial stream reaches and transition points between these reaches were determined by using the logistic regression equation and the probability cutpoint of 0.56. If this probability is calculated for a particular site, there is a 75-percent likelihood that the stream is correctly classified as perennial at that site. The lower and upper 95-percent confidence intervals can be used to indicate the most downstream and upstream points between which the transition point has a 95-percent confidence of being located. Although the maps provide fairly accurate depictions of intermittent and perennial stream reaches and transition points, the calculated statuses of all stream reaches should be checked against the results of field observations during summer low-flow periods to ensure the most accurate classification of the status of a stream reach. The limitations discussed previously in the "Limitations of the Logistic Regression Equation and Areas for Further Study" section (Part 1) should also be considered in using the maps. Stream reaches not depicted on USGS topographic maps are not a part of the NHD, and thus, their stream statuses and transitions from intermittent to perennial are not mapped with the automated mapping procedure or shown on the maps.

Summary and Conclusions

City and town conservation commissions and the Massachusetts Department of Environmental Protection (MDEP) are charged with protecting the riverfront areas of all rivers that flow perennially (year round) within Massachusetts, as specified in the Commonwealth of Massachusetts Rivers Protection Act of 1996. The 310 Code of Massachusetts Regulations (CMR) 10.58(2)(a) defines the riverfront area as the 200-ft-wide area extending along the length of each side of perennial streams from the mean annual high-water line (determined by bankfull field indicators), with exceptions for some urban areas.

The revised Regulations (December 20, 2002) specify that USGS topographic maps or more recent maps provided by MDEP will continue to be used for initial review of the intermittent or perennial status of a stream. Streams depicted as perennial on USGS topographic maps or on more recent maps provided by MDEP will be classified as perennial. A stream site depicted as perennial, however, can be reclassified as intermittent with direct observations of no flow during any 4 days in any consecutive 12-month period. The observations cannot be made during a period of "extended drought" as defined by the Massachusetts Drought Management Task Force, or if a stream is measurably affected by withdrawals, impoundments, or other anthropogenic flow reductions or diversions. Streams depicted as intermittent or not depicted on USGS topographic maps or more recent MDEP maps will be classified on the basis of their drainage-area sizes. An intermittent-stream site with a drainage area greater than or equal to 1.00 mi^2 will be classified as perennial. An intermittent-stream site with a drainage area greater than or equal to 0.50 mi^2 and less than 1.00 mi^2 will be classified as intermittent, with two exceptions. First, if the streamflow estimated with the World Wide Web application STREAMSTATS is equal to or greater than 0.01 ft^3/s at the 99-percent flow duration, then the stream site will be classified as perennial. Second, if STREAMSTATS cannot be used and more than 75 percent of the drainage area comprises stratified deposits, then the stream site will be classified as perennial. This second exception could occur if the stream is not depicted on USGS topographic maps or more recent maps provided by the MDEP; or if the stream is in the Buzzards Bay Basin, Cape Cod Basin, Islands Basin, North Coastal Basin, or Taunton River Basin, where the STREAMSTATS application does not function.

To assist city and town conservation commissions and the MDEP in determining whether a stream site is intermittent or perennial, a logistic regression equation was developed to estimate the probability of a stream flowing perennially at a specific site as a function of upstream basin characteristics. From late July through early September 2001, 476 stream sites throughout Massachusetts—except for the Cape Cod Basin, Islands Basin, southern part of the South Coastal Basin, and eastern part of the Buzzards Bay Basin—were observed during low-flow conditions to determine their intermittent or perennial status. Of the 476 stream sites visited in the state, 125 sites were omitted because of one or more of the following conditions: (1) regulation of streamflows; (2) no visible stream channel observed in the field; (3) site-access problems; (4) incorrect or no drainage-basin boundary drawn by STREAMSTATS or a Watershed Analyst Tool; (5) problems with STREAMSTATS or a Watershed Analyst Tool when the stream site was near the end point of the stream centerline; (6) the centerline data in STREAMSTATS or a Watershed Analyst Tool starting slightly downstream of the stream site visited; (7) the ground-water contributing area and surface-water drainage area not coinciding; and (8) observations made when nearby streamflows were too high (flow durations were lower than 80 percent).

The database included a total of 351 naturally flowing (no regulation) stream sites (85 intermittent and 266 perennial sites) with drainage areas that ranged from 0.04 to 10.96 mi^2. Of the 66 stream sites with drainage areas greater than

2.00 mi^2, 2 sites were intermittent and 64 sites were perennial. The 2 intermittent stream sites have drainage areas of 3.48 and 10.46 mi^2, and are underlain predominately by till and bedrock uplands with short narrow areas of stratified deposits in their river valleys. On the basis of this information, as well as the information from the field visit, the stream sites were classified as losing stream reaches. Thus, stream sites with drainage areas greater than 2.00 mi^2 were assumed to flow perennially, and the database used to develop the logistic regression equation included only the 285 stream sites—83 intermittent and 202 perennial sites—with drainage areas less than 2.00 mi^2.

Eleven different basin characteristics and a location variable, and transformations of these variables, were tested as potential explanatory variables for the logistic regression equation. The equation was determined to be a function of (1) drainage area (natural logarithm), (2) areal percentage of sand and gravel deposits, (3) areal percentage of forest land, and (4) region of the state (eastern region or western region). Equations were also formulated to calculate the upper and lower limits of the 95-percent confidence intervals for the estimated probability of a stream flowing perennially.

The logistic regression equation developed in this study estimated the intermittent or perennial stream status at the 285 observed stream sites throughout the state more accurately, in comparison to field observations of the stream status, than the USGS topographic maps (original Rivers Protection Act Regulations), revised Rivers Protection Act Regulations (December 20, 2002), the previously published logistic regression equation developed by the USGS in 2002, and selected low-flow statistics (99-, 98-, 95, 90, 85-, and 80-percent flow durations) estimated with STREAMSTATS. The equation used for this study was less accurate than the statistics for the 98- to 80-percent flow durations at estimating the intermittent or perennial stream status of the 28 verification sites in the Shawsheen River Basin. The equation used for this study was less accurate than the statistics for the 95- to 80-percent flow durations and the location-specific logistic regression equation for the South Coastal Basin in the previously published study at estimating the intermittent or perennial stream status of the 84 verification sites in the South Coastal Basin.

The logistic regression equation used for this study provides an objective means of determining the probability of a stream flowing perennially at a specific site; however, the reliability of the equation is affected by the data used to develop the equation. The equation is not recommended for (1) drainage areas less than 0.04 mi^2 in the state, (2) losing stream reaches, or (3) streams draining the Southeast Coastal Region—the southern part of the South Coastal Basin, the eastern part of the Buzzards Bay Basin, and the entire area of the Cape Cod and Islands Basins. If the equation were used on a regulated stream site, the estimated intermittent or perennial status would reflect the natural flow conditions for that site.

An automated mapping procedure was developed in ArcView for use with the National Hydrography Dataset (NHD). The NHD ArcView Toolkit and the NHD Watershed and NHD Watershed Characteristics applications were used in this automated procedure to determine the intermittent or perennial status of consecutive stream sites along the reaches in a given basin. The procedure starts at two locations on a headwater (first-order) stream reach, one near its most downstream point (about 100 ft upstream of its end point) and one near its most upstream point (about 100 ft downstream from its start point). The NHD Watershed application then delineates the drainage-area boundaries, the NHD Watershed Characteristics application determines values for the four explanatory variables of the logistic regression equation, and a project-specific script solves the equation for estimating the probability of a stream flowing perennially at the two locations. The automated procedure then determines the intermittent or perennial status of the reach on the basis of the calculated probability values for the two locations and the probability cutpoint (the set probability above which a stream is considered to flow perennially, 0.56 for this study), or continues to iterate upstream or downstream between locations to determine the location of the transition point between intermittent and perennial stream reaches. If the first-order stream reach was determined to be intermittent, the procedure moves to the next downstream reach, and repeats the process until it has covered the entire reach above a confluence. The automated procedure then moves to the next first-order stream, and repeats the process until the entire watershed is mapped.

The automated procedure was tested on the stream network in the Shawsheen River Basin in northeastern Massachusetts. The procedure classified 47 first-order stream reaches as entirely intermittent, 35 first-order stream reaches as entirely perennial, and 53 first-order stream reaches with a transition point from intermittent to perennial. Of the 47 intermittent first-order stream reaches, 42 were found to drain into a perennial-stream reach at their confluences and 5 to drain into an intermittent stream reach with a transition point in a downstream second- or third-order stream reach.

A map of the intermittent or perennial stream reaches in the Shawsheen River Basin is provided on a CD-ROM that accompanies this report. The CD-ROM also contains ArcReader 9.0, a freeware product that allows a user to zoom in and out, set a scale, pan, turn on and off map layers (such as a USGS topographic map), and print a map of the stream site with a scale bar. Maps of intermittent and perennial stream reaches in Massachusetts will provide city and town conservation commissions and the MDEP an additional method for assessing the intermittent and perennial status of stream sites.

Acknowledgments

The author thanks the MDEP's Perennial/Intermittent Advisory Committee; the MDEP, Bureau of Resource Protection, Wetlands and Waterways Program; and the Massachusetts Executive Office of Environmental Affairs, former Taunton River Basin Watershed Team for providing input on the State's perspective related to stream-classification issues. Thanks are due to USGS coworkers Stacey Archfield, Jaclyn McHugh, Kristen Pitocco, Christopher Ragnelli, and Timothy Thies for field assistance and data compilation; and to Phillip Zarriello for programs written to determine selected basin characteristics. Special thanks and appreciation are due to USGS volunteers Thomas Archfield, Edgar Bent, Kim Bent, Steven Mahood, and MDEP employees Heidi Davis and Thomas Maguire for field assistance.

References Cited

Battaglin, W.A., Ulery, R.L., Winterstein, Thomas, and Welborn, Toby, 2003, Estimating the susceptibility of surface water in Texas to nonpoint-source contamination by use of logistic regression modeling: U.S. Geological Survey Water-Resources Investigations Report 03–4205, 24 p.

Bent, G.C., 1994, Effects of timber cutting on runoff to Quabbin Reservoir, central Massachusetts, *in* Marston, R.A., and Hasfurther, V.R., eds., Proceedings, The effects of human-induced changes on hydrologic systems: Jackson Hole, WY, 1994 Annual Summer Symposium of the American Water Resources Association, June 26–29, 1994, p. 187–196.

Bent, G.C., 2001, Effects of forest-management activities on runoff components and ground-water recharge to Quabbin Reservoir, central Massachusetts: Forest Ecology and Management, v. 143, issue 1-3, p. 115–129.

Bent, G.C., and Archfield, S.A., 2002, A logistic regression equation for estimating the probability of a stream flowing perennially in Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 02–4043, 45 p.

Bullock, Andrew, and Acreman, Michael, 2003, The role of wetlands in the hydrologic cycle: Hydrology and Earth System Sciences, v. 7, no. 3, p. 358–389.

Calhoun, D.L., Frick, E.A., and Buell, G.R., 2003, Effects of urban development on nutrient loads and streamflow, upper Chattahoochee River Basin, Georgia, 1976–2001, *in* Hatcher, K.J., ed., Proceedings of the 2003 Georgia Water Resources Conference, Athens, GA, April 23–24, 2003, 5 p., accessed April 30, 2004, at *http://ga.water.usgs.gov/pubs/other/gwrc2003/pdf/Calhoun-GWRC2003.pdf*

Collett, D., 1991, Modelling binary data: New York, Chapman and Hall, 369 p.

Cook, E.D., 2001, Solutions manual to accompany *Applied Logistic Regression* (2d ed.), D.W. Hosmer and Stanley Lemeshow: New York, John Wiley and Sons, 270 p.

Eckhardt, D.V., Flipse, W.J., and Oaksford, E.T., 1989, Relation between land use and ground-water quality in the upper glacial aquifer in Nassau and Suffolk Counties, Long Island, New York: U.S. Geological Survey Water-Resources Investigations Report 86–4142, 35 p.

Eckhardt, D.V., and Stackelberg, P.E., 1995, Relation of ground-water quality to land use on Long Island, New York: Ground Water, v. 33, no. 6, p. 1019–1033.

Fairfax County, Virginia, Public Works and Environmental Services, 2004, Perennial stream mapping project, accessed January 12, 2005, at *http://www.co.fairfax.va.us/dpwes/watersheds/perennial.htm*

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

Francy, D.S., Gifford, A.M., and Darner, R.A., 2003, *Escherichia coli* at Ohio bathing beaches—distribution, sources, wastewater indicators, and predictive modeling: U.S. Geological Survey Water-Resources Investigations Report 02–4285, 117 p., accessed October 2, 2003, at *http://oh.water.usgs.gov/reports/Abstracts/wrir02-4285.html*

Fritz, K.M., 2004, Development of biological indicators, methods and assessment techniques for use in headwater intermittent streams, accessed March 23, 2004, at *http://www.epa.gov/region5/water/wqb/pdf/fritz_presentation.pdf*

Fritz, K.M., Johnson, B.R., and Walters, D.M., 2005, Factors controlling hydrologic permanence in headwater streams, *in* Abstracts, Annual North American Bethological Society Meeting, New Orleans, LA, 2005, accessed November 21, 2005, at *http://www.benthos.org/database/allnabstracts.cfm/db/NewOrleans2005abstracts/id/132*

Gay, F.B., and Delaney, D.F., 1980, Hydrology and water resources of the Shawsheen River Basin, Massachusetts: U.S. Geological Survey Hydrologic Investigations Atlas HA–614, 3 sheets, scale 1:48,000.

Harvard Forest, Harvard University, 2004, Headwater habitat streams in central Massachusetts, accessed February 24, 2004, at *http://harvardforest.fas.harvard.edu/data/p09/hf095/hf095.html*

Helsel, D.R., and Hirsch, R.M., 1992, Statistical methods in water resources: Studies in Environmental Science 49, Elsevier Science Publishing Company, Inc., New York, 522 p.

Hornbeck, J.W., Adams, M.B., Corbett, E.S., Verry, E.S., and Lynch, J.A., 1993, Long-term impacts of forest treatments on water yield—a summary for northeastern United States: Journal of Hydrology, v. 150, p. 323–344.

Hornbeck, J.W., Martin, C.W., and Eagar, C., 1997, Summary of water yield experiments at Hubbard Brook Experimental Forest, New Hampshire: Canadian Journal of Forest Research, v. 27, p. 2043–2052.

Hosmer, D.W., and Lemeshow, Stanley, 2000, *Applied Logistic Regression*, (2d ed.): New York, John Wiley and Sons, 375 p.

Kliever, J.D., 1996, Low-flow characteristics of selected streams in northern Rhode Island: U.S. Geological Survey Water-Resources Investigations Report 95–4299, 11 p.

Kolka, R.K., and Stringer, J.W., 2000, USGS Water Resources grant proposal—defining perennial, intermittent, and ephemeral channels in eastern Kentucky—application for forestry best management practices, accessed January 12, 2005, at *http://water.usgs.gov/wrri/00grants/KYchannels.pdf*

Koltun, G.F., and Sherwood, J.M., 1998, Factors related to the joint probability of flooding on paired streams: U.S. Geological Survey Water-Resources Investigations Report 98–4238, 32 p.

Langbein, W.B., and Iseri, K.T., 1960, General introduction and hydrologic definitions, Manual of Hydrology, Part 1, General Surface-Water Techniques: U.S. Geological Survey Water-Supply Paper 1541–A, p. 18.

Leopold, L.B., 1994, A view of the river: Cambridge, MA, Harvard University Press, 298 p.

Massachusetts Executive Office of Environmental Affairs and Massachusetts Emergency Management Agency, 2001, Working draft—Massachusetts drought management plan, accessed February 15, 2005, at *http://www.mass.gov/dcr/waterSupply/rainfall/droughtplan.doc*

Massachusetts Department of Environmental Protection, 2002a, Wetlands protection act regulations, 310 CMR 10.00, accessed January 10, 2005, at *http://www.state.ma.us/dep/brp/ww/files/310cmr10.pdf*

Massachusetts Department of Environmental Protection, 2002b, Preface to revisions to the Massachusetts wetlands regulation (310 CMR 10.00)—relating to the definition of “extended drought” and distinguishing “perennial rivers” from “intermittent streams,” accessed October 2, 2003, at *http://www.state.ma.us/dep/brp/ww/files/310cmr10.pdf*

MassGIS, 2004a, Datalayers/GIS Database—Hydrography (1:25,000)—February 2004, accessed May 3, 2004, at *http://www.state.ma.us/mgis/hd.htm*

MassGIS, 2004b, Datalayers/GIS Database—Surficial Geology—October 1999, scale 1:250,000, accessed May 3, 2004, at *http://www.state.ma.us/mgis/sg.htm*

MassGIS, 2004c, Datalayers/GIS Database—Land use—January 2002, scale 1:25,000, accessed April 30, 2004, at *http://www.state.ma.us/mgis/lus.htm*

MassGIS, 2004d, DEP wetlands (1:12,000)—October 27, 2004, accessed January 11, 2005, at *http://www.state.ma.us/mgis/wetdep.htm*

Meinzer, O.E., 1923, Outline of ground-water hydrology, with definitions: U.S. Geological Survey Water-Supply Paper 494, 71 p.

Mrazik, B.R., Mader, D.L., and MacConnell, W.P., 1980, Integrated watershed management—an alternative for the northeast: Amherst, MA, Massachusetts Agricultural Experiment Station, Research Bulletin 664, 50 p.

Mueller, D.K., Ruddy, B.C., and Battaglin, W.A., 1997, Logistic model of nitrate in streams of the upper-midwestern United States: Journal of Environmental Quality, v. 26, no. 5, p. 1223–1230.

Nolan, B.T., 2001, Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States: Ground Water, v. 39, no. 2, p. 290–299.

Nolan, B.T., Ruddy, B.C., Hitt, K.J., and Helsel, D.R., 1997, Risk of nitrate in ground waters of the United States—a national perspective: Environmental Science and Technology, v. 31, no. 8, p. 2229–2236.

Nolan, B.T., Hitt, K.J., and Ruddy, B.C., 2002, Probability of nitrate contamination of recently recharged ground waters in the conterminous United States: Environmental Science and Technology, v. 36, no. 10, p. 2138–2145.

North American Benthological Society, 2002, Ecological functions of intermittent streams, *in* Abstracts, 2002 Annual Meeting, Pittsburgh, PA, May 28–June 1, 2002, accessed February 24, 2004, at *http://www.benthos.org/meeting/nabs2002/displaysession.cfm/id/45*

North American Benthological Society, Abstracts, 2003, 2003 Annual Meeting, Athens, GA, May 27–May 31, 2003, accessed April 23, 2004, at *http://www.benthos.org/meeting/nabs2003/index.htm*

North American Benthological Society, 2005a, Temporary waters—ecological values, policy, and management I, *in* Abstracts, 2005 Annual Meeting, New Orleans, LA, May 22–May 27, 2005, accessed November 21, 2005, at *http://www.benthos.org/meeting/nabs2005/displaysession.cfm/id/16*

North American Benthological Society, 2005b, Temporary waters—ecological values, policy, and management II, *in* Abstracts, 2005 Annual Meeting, New Orleans, LA, May 22–May 27, 2005, accessed November 21, 2005, at *http://www.benthos.org/meeting/nabs2005/displaysession.cfm/id/24*

North Carolina Department of Environment and Natural Resources, Division of Water Quality, Water Supply Protection Program, 2004, NC DENR—Division water quality "redbook" surface waters and wetlands standards, accessed January 12, 2005, at *http://h2o.enr.state.nc.us/admin/rules/documents/rb080104.pdf*

North Carolina Department of Environment and Natural Resources, Division of Water Quality, Wetlands/401 Water Quality Certification Unit, 2004, Stream evaluation form, accessed January 12, 2005, at *http://h2o.enr.state.nc.us/ncwetlands/strmfrm.html*

Palmquist, Robert, 2003, Type N stream demarcation study, phase I—pilot results: final report, version 6.85, Department of Natural Resources, Washington, accessed March 23, 2004, at *http://www.dnr.wa.gov/forestpractices/adaptivemanagement/pipstudy/*

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

Restrepo, Miguel, and Waisanen, Pamela, 2004a, Strategies for stream classification using GIS, *in* Abstracts, 24th Annual ESRI International User Conference, San Diego, CA, August 9–13, 2004, accessed January 12, 2005, at *http://gis.esri.com/library/userconf/proc04/abstracts/a1869.html*

Restrepo, Miguel, and Waisanen, Pamela, 2004b, Strategies for stream classification using GIS, Earth Resources, Observations, and Science, 19 p., accessed January 12, 2005, at *http://edna.usgs.gov/Edna/pubs/Stream_Classification_Strategies_Paper.pdf*

Ries, K.G., 1994a, Estimation of low-flow duration discharges in Massachusetts: U.S. Geological Survey Water-Supply Paper 2418, 50 p.

Ries, K.G., 1994b, Development and application of generalized-least-squares regression models to estimate low-flow duration discharges in Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 94–4155, 33 p.

Ries, K.G., 1997, August median streamflows in Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 97–4190, 27 p.

Ries, K.G., and Friesz, P.J., 2000, Methods for estimating low-flow statistics for Massachusetts streams: U.S. Geological Survey Water-Resources Investigations Report 00–4135, 81 p.

Ries, K.G., Steeves, P.A., Freeman, Aleda, and Singh, Raj, 2000, Obtaining streamflow statistics for Massachusetts streams on the World Wide Web: U.S. Geological Survey Fact Sheet 104–00, 4 p.

Rivenbark, B.L., and Jackson, C.R., 2004, Average discharge, perennial flow initiation, and channel initiation—small southern Appalachian Basins: Journal of the American Water Resources Association, v. 40, no. 3, p. 639–646.

Rose, Seth, and Peters, N.E., 2001, Effects of urbanization on streamflow in the Atlanta area (Georgia, USA)—a comparative hydrologic approach: Hydrologic Processes, v. 15, p. 1441–1457.

Rupert, M.G., 1998, Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate ($NO_2 + NO_3 – N$) in ground water in the Idaho part of the upper Snake River Basin: U.S. Geological Survey Water-Resources Investigations Report 98–4203, 32 p., 1 pl., accessed October 2, 2003, at *http://idaho.usgs.gov/PDF/wri984203/index.html*

Rupert, M.G., 2003, Probability of detecting atrazine/desethyl-atrazine and elevated concentrations of nitrate in ground water in Colorado: U.S. Geological Survey Water-Resources Investigations Report 02–4269, 35 p., accessed October 2, 2003, at *http://water.usgs.gov/pubs/wri/wri02-4269/*

SAS Institute, Inc., 1989, SAS/STAT User's guide (4th ed.): Cary, N.C., version 6, v. 2, 846 p.

SAS Institute, Inc., 1995, Logistic regression examples using the SAS system (1st ed.): Cary, N.C., version 6, 163 p.

Shanley, J.B., Strause, J.L., and Risley, J.C., 1995, Effects of selective forest clearing, fertilization, and liming on the hydrology and water quality of a small tributary to the Quabbin Reservoir, central Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 95–4124, 57 p.

Simcox, A.C., 1992, Water resources of Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 90–4144, 94 p.

Simmons, D.L., and Reynolds, R.J., 1982, Effects of urbanization on base flow of selected south-shore streams, Long Island, New York: Water Resources Bulletin, v. 18, p. 797–805.

Socolow, R.S., Leighton, C.R., Whitley, J.F., and Ventetuolo, D.J., 2002, Water resources data—Massachusetts and Rhode Island, water year 2001: U.S. Geological Survey Water Data Report MA-RI-01-1, 307 p.

Spatial Climate Analysis Service—Oregon Climate Service, 2004, Precipitation—annual climatology (1971–2000): Oregon State University, accessed January 12, 2005, at *http://www.ocs.orst.edu/prism/products/viewer.phtml?file=/pub/prism/us/graphics/ppt/Normals/us_ppt_1971_2000.14.png&year=1971_2000&vartype=ppt&month=14&status=final*

Spinello, A.G., and Simmons, D.L., 1992, Base flow of 10 south-shore streams, Long Island, New York: U.S. Geological Survey Water-Resources Investigations Report 90–4205, 34 p.

Squillace, P.J., Moran, M.J., Lapham, W.W., Price, C.V., Clawges, R.M., and Zogorski, J.S., 1999, Volatile organic compounds in untreated ambient groundwater of the United States, 1985–1995: Environmental Science and Technology, v. 33, no. 23, p. 4176–4187.

Stata Corporation, 2003, Stata base reference manual: College Station, TX, release eight, reference volume 2 (G-M), 531 p.

Tasker, G.D., 1989, Regionalization of low flow characteristics using logistic and GLS regression, *in* Kavvas, M.L. ed., New directions for surface water modeling: IAHS Publication No. 181, p. 323–331.

Tesoriero, A.J., and Voss, F.D., 1997, Predicting the probability of elevated nitrate concentrations in the Puget Sound Basin—implications for aquifer susceptibility and vulnerability: Ground Water, v. 35, no. 6, p. 1029–1039.

The Commonwealth of Massachusetts, 1996, An act providing protection for the rivers of the Commonwealth, accessed January 10, 2005, at *http://www.state.ma.us/dep/brp/ww/files/riveract.htm*

U.S. Department of Agriculture, Natural Resources Conservation Service, Soil Survey Division, 2004, National SSURGO database—data, accessed April 29, 2004, at *http://www.ftw.nrcs.usda.gov/ssur_data.html*

U.S. Department of Commerce, National Oceanic and Atmospheric Administration, 2001, Climatological data: New England—August 2001, v. 118, no. 8, 56 p.

U.S. Geological Survey, 1980, Technical instructions of National Mapping Division, map editing, provisional chapter 5C1: U.S. Department of the Interior, U.S. Geological Survey, p. 16.

U.S. Geological Survey, 2005a, National Elevation Dataset, accessed January 12, 2005, at *http://ned.usgs.gov/*

U.S. Geological Survey, 2005b, National Hydrography Dataset, accessed January 12, 2005, at *http://nhd.usgs.gov/*

U.S. Geological Survey, 2005c, Elevation derivatives for national applications, accessed January 12, 2005, at *http://edna.usgs.gov/*

U.S. Geological Survey, 2005d, The National Hydrography Dataset, NHD Watershed Tool—Instructions for preprocessing supporting data layers, accessed April 26, 2005, at *http://nhd.usgs.gov/watershed/watershed_tool_inst_TOC.html*

U.S. Geological Survey, 2005e, National Hydrography Dataset Tools, accessed April 26, 2005, at *http://nhd.usgs.gov/tools.html*

U.S. Geological Survey, 2005f, National Hydrography Dataset Applications, accessed April 26, 2005, at *http://nhd.usgs.gov/applications.html*

Virginia Department of Conservation and Recreation, Chesapeake Bay Local Assistance, 2002, Chesapeake Bay preservation area of designation and management regulations, accessed January 12, 2005, at *http://www.cblad.virginia.gov/theregs.cfm*

Wandle, S.W., Jr., and Randall, A.D., 1994, Effects of surficial geology, lakes, and swamps, and annual water availability on low flows of streams in central New England and their use in low-flow estimations: U.S. Geological Survey Water-Resources Investigations Report 93–4092, 57 p.

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