

# DEVELOPMENT AND APPLICATION OF GENERALIZED-LEAST-SQUARES REGRESSION MODELS TO ESTIMATE LOW-FLOW DURATION DISCHARGES IN MASSACHUSETTS

***By Kernell G. Ries III***

---

U.S. GEOLOGICAL SURVEY  
Water-Resources Investigations Report 94-4155

Prepared in cooperation with the  
MASSACHUSETTS DEPARTMENT OF ENVIRONMENTAL MANAGEMENT,  
OFFICE OF WATER RESOURCES



Marlborough, Massachusetts  
1994

**U.S. DEPARTMENT OF THE INTERIOR**  
**BRUCE BABBITT, Secretary**

U.S. GEOLOGICAL SURVEY  
Gordon P. Eaton, Director

---

For additional information, write to:

District Chief,  
Massachusetts-Rhode Island District  
U.S. Geological Survey  
Water Resources Division  
28 Lord Road, Suite 280  
Marlborough, MA 01752

Copies of this report can be purchased from:

U.S. Geological Survey  
Earth Science Information Center  
Open-File Reports Section  
Box 25286, MS 517  
Federal Center  
Denver, CO 80225

# CONTENTS

Abstract .....	1
Introduction .....	1
Purpose and Scope .....	2
Previous Studies.....	2
Acknowledgments .....	4
Physical Setting.....	4
Characteristics of Massachusetts .....	4
Characteristics of the Study Basins .....	4
Data-Base Development.....	12
Computation of Selected Duration Discharges.....	12
Selection and Measurement of Basin Characteristics.....	16
Development of the Generalized-Least-Squares Regression Models .....	17
Generalized-Least-Squares Regression Analyses.....	17
Bias Correction .....	18
Final Equations for Predicting Long-Term Duration Discharges .....	18
Prediction Intervals .....	20
Application of the Generalized-Least-Squares Regression Models.....	21
Limitations of the Equations .....	24
Summary and Conclusions.....	24
References Cited .....	25

## FIGURES

1-6. Maps showing:	
1. Boundaries of the 27 major river basins in Massachusetts .....	3
2. Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the Concord River Basin .....	8
3. Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the North Coastal Basin .....	9
4. Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the South Coastal Basin .....	10
5. Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the Narragansett Bay Basins and Tenmile River .....	11
6. Type and locations of sites used to develop the regression models .....	13

## TABLES

1. Descriptions of selected sites for which low-flow estimates are provided.....	6
2. Description of streamflow-gaging stations and low-flow partial-record sites used to develop the regression models.....	14
3. Summary of generalized-least-squares regression equations used to estimate duration discharges for selected sites in Massachusetts, bias correction factors, and measures of model adequacy .....	19
4. Values needed to calculate 90-percent prediction intervals .....	21
5. Predicted 95-, 98-, and 99-percent duration discharges and prediction intervals for the selected sites in the study basins .....	22
6. Duration discharges and basin characteristics for sites used in the regression analyses.....	28
7. Basin characteristics for selected sites for which low-flow estimates are provided .....	32

## CONVERSION FACTORS, VERTICAL DATUM, AND ABBREVIATIONS AND DEFINITIONS

Multiply	By	To Obtain
cubic foot (ft <sup>3</sup> )	0.02832	cubic meter
cubic feet per second (ft <sup>3</sup> /s)	0.02832	cubic meter per second
foot (ft)	0.3048	meter
inch (in)	25.4	millimeter
mile (mi)	1.609	kilometer
square mile (mi <sup>2</sup> )	2.590	square kilometer

Temperature in degrees Fahrenheit (°F) can be converted to degrees Celsius (°C) as follows:  $^{\circ}\text{C} = 5/9 \times (^{\circ}\text{F} - 32)$ .

**Sea Level:** In this report “sea level” refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)--a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

### ABBREVIATIONS AND DEFINITIONS

#### Organizations

MOWR      Massachusetts Department of Environmental Management, Division of Resource Conservation, Office of Water Resources  
(formerly MDWR)

USGS      U.S. Geological Survey

#### Basin Characteristics

DAREA      Drainage area, in square miles

DENSITY      Drainage density -- Total length of streams in basin divided by drainage area

DRIFT      Area of stratified-drift deposits, in square miles

DRT/TST      Area of stratified drift per unit stream length

GWHEAD      Surrogate for head in the stratified-drift aquifer computed by subtracting the minimum basin altitude from the mean basin altitude

GWRELIEF      Surrogate for head in the stratified-drift aquifer computed by subtracting the minimum from the maximum altitude in the stratified-drift deposits

RELIEF      The maximum minus the minimum altitude in the basin.

#### Miscellaneous

GIS      Geographic information system computer software

GLS      Generalized-least-squares regression analysis

OLS      Ordinary-least-squares regression analysis

WLS      Weighted-least-squares regression analysis

# DEVELOPMENT AND APPLICATION OF GENERALIZED-LEAST-SQUARES REGRESSION MODELS TO ESTIMATE LOW-FLOW DURATION DISCHARGES IN MASSACHUSETTS

By Kernell G. Ries III

## Abstract

Physically based mathematical models were developed by use of generalized-least-squares regression analyses to estimate long-term 95-, 98-, and 99-percent duration discharges for ungaged streams in Massachusetts. Duration discharges for 61 sites were used in the regression analyses; 37 sites were streamflow-gaging stations and 24 sites were low-flow partial-record stations. The duration discharges were related to basin characteristics measured from digital data bases by use of geographic-information-systems computer software. Significant characteristics used in the models were drainage area, area underlain by stratified-drift deposits per unit of stream length in the basin, and a surrogate for the effective head on the aquifer in the stratified-drift deposits, computed by subtracting the minimum basin elevation from the mean basin elevation.

Standard errors of prediction were 57.5, 85.6, and 98.5 percent for models for the 95-, 98-, and 99-percent duration discharges, respectively. Model error variances were about 10 times the sampling error variances indicating that the precision of future models is likely to be improved more by obtaining new or improved measurements of basin characteristics or by adding data from new sites to the analyses, than by collecting more streamflow data at the sites presently used in the analyses.

The models were used to predict duration discharges for 35 selected sites in the Concord River, North Coastal, South Coastal, Narragansett, and Tenmile River Basins. Ninety-percent prediction intervals were computed for the estimates at each of

the sites, except at five sites where values of the independent variables were outside the ranges of those for the sites used in the regression analyses.

## INTRODUCTION

The Massachusetts Department of Environmental Management, Office of Water Resources (MOWR, formerly the Division of Water Resources), is required to develop water-management plans for each of the 27 major river basins in the State. The plans contain inventories and analyses of water availability and current demand, and alternatives for meeting projected future demand. The plans are used by communities, regional planners, and other State agencies to manage the water resources in the basins and to make decisions regarding permitting of new water withdrawals and interbasin transfers.

In 1988, the U.S. Geological Survey (USGS) began the first of a series of studies to estimate low-flow statistics for selected sites in Massachusetts. These studies, done in cooperation with the MOWR, are referred to as the Basin Yield studies. The low-flow statistics estimated by the Basin Yield studies are used by the MOWR to develop their basin-management plans. The estimates provide an indication of water availability during times when water-conservation practices are likely to be needed to protect instream flow uses.

This report describes the second Basin Yield study, in which models in the form of mathematical equations were developed for use in estimating low-flow statistics for streams in most of Massachusetts. Equations were developed to predict the discharges, in cubic feet per second, that are expected to be equalled or exceeded under essentially natural conditions 95, 98, and 99

percent of the time, assuming no climatic, physiographic or anthropogenic changes occur over the long term. These values are referred to as the long-term 95-, 98-, and 99-percent duration discharges.

Similar models were developed for the first Basin Yield study (Ries, 1993). The models for both studies were developed by use of regression analyses to determine linear relations between calculated low-flow statistics, for sites where streamflow data are available, and measured drainage-basin characteristics for the sites. Equations for the initial Basin Yield study (Ries, 1993) were developed by use of a weighted-least-squares (WLS) regression analysis procedure, which gave each site used in the analysis a different weight on the basis of differences in record length and differences in the variance of the regression residuals. The equations from the first Basin Yield study provided estimates of the 95-, 98-, and 99-percent duration discharges for a base period of 25 years (October 1, 1962 through September 30, 1987). Equations for the second study, on which this report is based, were developed with an expanded data base by use of a generalized-least-squares (GLS) regression analysis procedure. Weights for the sites used in the GLS analyses were adjusted for cross correlation between the concurrent streamflows of the sites and for differences in their record lengths.

The equations were used to estimate low-flow duration discharges for selected sites in five basins in eastern Massachusetts (fig. 1). The selected sites included any streamflow-gaging stations in the basins and other sites where water availability may be of interest to State or local agencies. A computerized geographic information system (GIS) was used to measure all of the basin characteristics. The basin characteristics were obtained from digitized map data available from various sources on a national scale, or developed on a statewide basis for this and other studies.

## Purpose and Scope

The purpose of this report is to provide (1) mathematical models for use in estimating the long-term 95-, 98-, and 99-percent duration discharges under natural flow conditions for streams in Massachusetts, and (2) estimates of the long-term duration discharges for selected sites in the Concord River, North Coastal, South Coastal, Narragansett Bay, and Tenmile River planning basins.

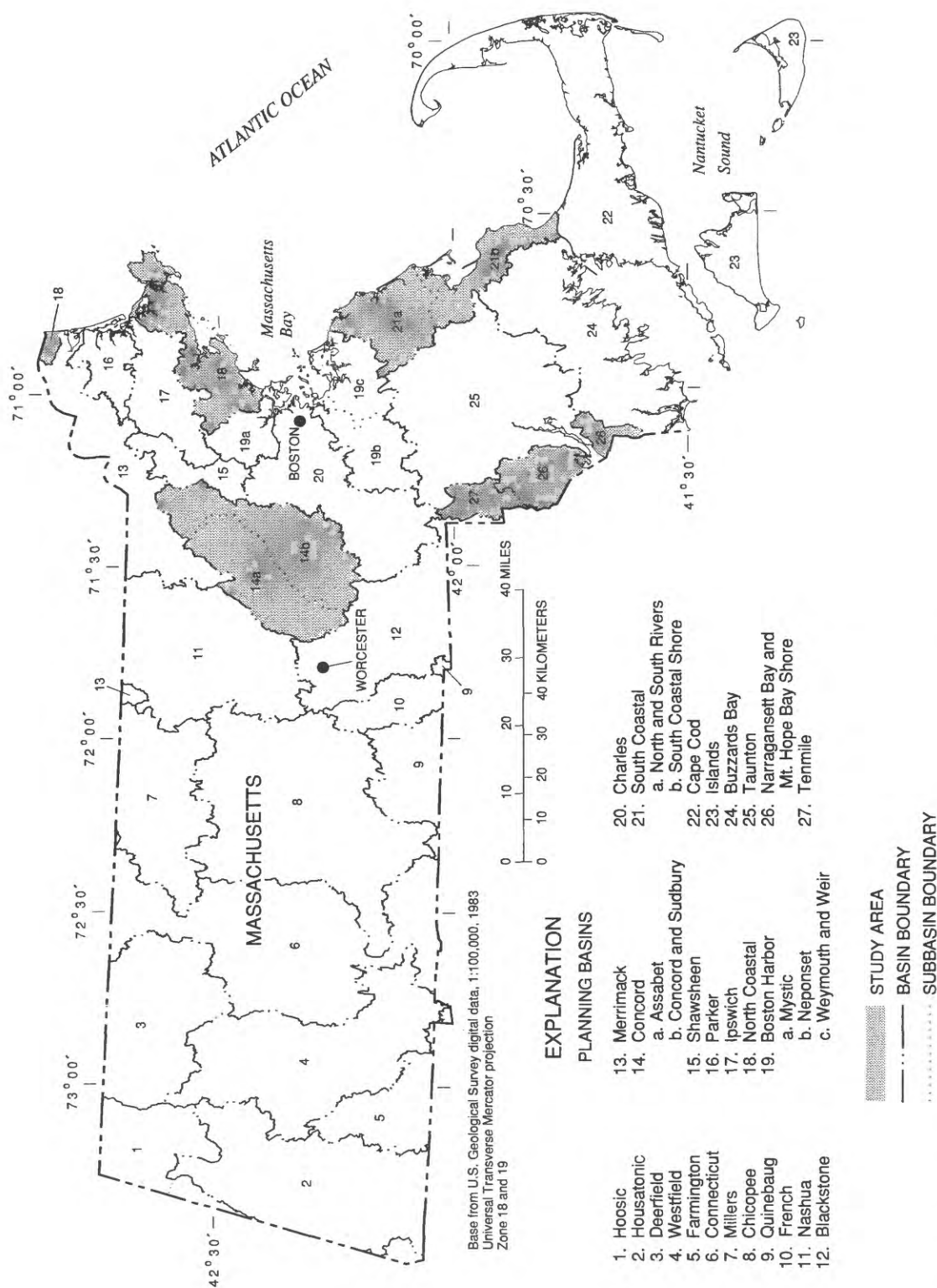
This report describes (1) the physical setting of Massachusetts and the study basins, (2) the development of the data base for the regression analyses, including methods used to compute streamflow statistics and basin characteristics, (3) the development and assessment of the GLS regression equations used to predict selected duration discharges, and (4) the application of regression equations to estimate long-term duration discharges for selected streams in the study basins.

## Previous Studies

Low-flow statistics for most streamflow-gaging stations and many low-flow partial-record stations in Massachusetts have been published by the USGS in a series of "Gazetteers of hydrologic characteristics of streams" published as Water-Resources Investigations reports, in Hydrologic Atlas reports, and in groundwater assessment reports published as Water-Resources Investigations reports (Olimpio and deLima, 1984; Lapham, 1988; Myette and Simcox, 1989; de Lima, 1991; Persky, 1993). References for the gazetteers and the Hydrologic Atlas reports were published in a compilation report by the U.S. Geological Survey (1987).

Several studies have attempted to regionalize low-flow statistics in the northeastern United States by use of regression analyses. The statistics selected for regionalization, methods of regression analysis, and precision of the resulting equations have varied considerably. Generally, equations developed by the more recent studies are more precise than those developed earlier. Improvements in the equations can be attributed to increased precision in the streamflow statistics due to longer record lengths, advances in regression analysis techniques, and greater access to and advances in the computers used to perform the analyses.

Most regionalization studies in the northeastern United States have developed equations to estimate low-flow frequency statistics, such as the 7-day, 10-year low flow—the annual 7-day mean low flow that occurs, on average, once in 10 years. Studies that regionalized low-flow frequency statistics were done for Connecticut (Cervione, 1982), Pennsylvania and New York (Ku and others, 1975), Maine (Parker, 1977), Massachusetts, New Hampshire, Rhode Island and Vermont (Johnson, 1970), southeastern Massachusetts (Tasker, 1972), and Massachusetts (Male and Ogawa, 1982; Vogel and Kroll, 1990).



**Figure 1.** Boundaries of the 27 major river basins in Massachusetts.

Studies that regionalized flow-duration statistics were done for Connecticut, New Hampshire, southeastern Massachusetts, and all of Massachusetts except for areas entirely underlain by stratified-drift deposits. Thomas (1966) used the area of coarse-grained stratified-drift deposits as the explanatory variable to estimate streamflow durations for Connecticut. Dingman (1978) used drainage areas and mean basin elevations to synthesize flow-duration curves for New Hampshire. Tasker (1972) related low streamflows in southeastern Massachusetts to drainage area and a "ground-water factor" computed by dividing the area of coarse-grained stratified drift by potential well yield determined from aquifer maps. A similar ground water factor was used by Male and Ogawa (1982), along with mean annual precipitation, drainage area and several other basin characteristics to estimate flow-duration statistics in Massachusetts. Fennessey and Vogel (1990) used drainage area and basin relief to synthesize the lower half of flow-duration curves in Massachusetts. Ries (1993) estimated the 95-, 98-, and 99-percent duration discharges for streams in Massachusetts for a 25-year base period of October 1, 1962 to September 30, 1987. Basin characteristics included in the equations were drainage area, area of stratified-drift deposits per unit length of streams, and a surrogate measure of effective aquifer head, computed by subtracting the basin-outlet elevation from the mean basin elevation.

## Acknowledgments

The author thanks Peter Phippen, of the MOWR, for his assistance in selecting sites for which low-flow duration discharges were estimated for this study. The work of everyone involved in digitizing, preparing, and retrieving data from the GIS data bases used in this study also is appreciated.

## PHYSICAL SETTING

Massachusetts has been subdivided by the MOWR into 27 separate basins for planning purposes (fig. 1). Planning basin boundaries generally follow major river basin boundaries. Some of the designated planning basins are actually parts of larger river basins with areas that extend into bordering states. For example, the Westfield, Deerfield, Millers, and Chicopee River Planning Basins are parts of the Connecticut River Basin. Several of the planning basins along the coastal plain of Massachusetts are not drained by a single major stream. These coastal planning basins generally were

comprised by grouping land areas that are drained by streams that discharge to the same receiving body of salt water, such as Boston Harbor and Buzzards Bay.

The following sections contain brief descriptions of physical characteristics affecting the hydrology of the State and the study basins.

## Characteristics of Massachusetts

Massachusetts encompasses an area of 8,093 mi<sup>2</sup>. Located in the northeastern United States, Massachusetts has a humid climate, with an average annual precipitation of about 45 in. and average annual temperatures ranging from 50°F in coastal areas to 45°F in the western mountains. Precipitation is fairly evenly distributed throughout the year. Average monthly temperatures in coastal areas range from about 30°F in February to about 71°F in July, with average monthly temperatures in the western parts of the State ranging from about 20°F in January to about 68°F in July (National Oceanic and Atmospheric Administration, 1989). Elevations range from sea level along the coast to almost 3,500 ft in the western mountains. Relief generally increases from east to west.

Except during and for a short time after storms, most flow in Massachusetts streams comes from ground water discharged by aquifers in unconsolidated deposits adjacent to the streams. The higher yielding aquifers usually are in coarse sand and gravel deposits along the valley floors of inland river basins and in coastal areas of southeastern Massachusetts. These materials, termed stratified drift, were sorted and laid down by meltwaters from retreating glaciers at the end of the last ice age. In addition to the high-yielding coarse sand and gravel deposits, the stratified drift commonly contains layers of fine sand and clay that yield little water to adjacent streams or to wells. The stratified-drift deposits are usually surrounded by upland areas underlain by till with exposed bedrock outcrops. Till is an unsorted glacial deposit that consists of material ranging in size from clay to large boulders. Till yields little water to adjacent streams or to wells in relation to yields from coarse-grained stratified drift.

## Characteristics of the Study Basins

The Concord River Basin, part of the Merrimack River Basin, drains 400 mi<sup>2</sup> of east-central Massachusetts. All or parts of 34 towns, including the largest in the State, Framingham, and the cities of Lowell and Marlborough are included in the basin. The population



of the basin is about 342,500 (U.S. Geological Survey, 1988). It is bounded on the west by the Nashua River Basin, on the north by the Merrimack River Basin, on the northeast by the Shawsheen River Basin, on the southeast by the Charles River Basin, and on the southwest by the Blackstone River Basin. Almost one-half of the basin is underlain by stratified-drift deposits.

The MOWR divides the basin into two subbasins, the Assabet and the Concord-Sudbury. The western half of the basin is drained by the Assabet River, a major tributary to the Concord River. The eastern half of the basin is drained in the north by the Concord River and in the south by its other major tributary, the Sudbury River. The Assabet River subbasin is characterized by low, rounded hills, with a maximum elevation of about 420 ft. Although the stream gradient in the entire basin is not large, the Concord-Sudbury River subbasin generally has a lower gradient than that of the Assabet River subbasin. The central part of the Concord-Sudbury River subbasin is nearly flat, with extensive wetlands adjacent to the river. Maximum elevation in the Concord-Sudbury River subbasin is 480 ft.

Streamflow from 93 mi<sup>2</sup> of the Concord-Sudbury River subbasin can be diverted from reservoirs for emergency use by the Greater Boston area. There are four streamflow-gaging stations in the basin. Names and descriptions for these stations, and for ten additional sites where estimated low-flow duration discharges are provided later in this report, are listed in table 1. Their locations are indicated in figure 2.

The North Coastal Basin has an area of 172 mi<sup>2</sup> that is drained by several small streams, the largest of which is the Saugus River, which discharge to Massachusetts Bay and the Atlantic Ocean. The basin contains all or part of 26 cities and towns, including the cities of Beverly, Lynn, Peabody, Salem, and Revere. The basin is bounded on the north and northwest by the Ipswich River Basin, and on the southwest by the Boston Harbor Basin. It has a population of about 428,500 (U.S. Geological Survey, 1988). Bedrock is at or near the surface in much of the basin, but there are several small stratified-drift aquifers in the southern half of the basin. Aquifer thicknesses generally are less than 50 ft. Stream gradients are low in the basin; where the topography consists of low hills with 350 ft maximum elevation. There are no streamflow-gaging stations in the basin. Estimated low-flow duration discharges are provided in this report for four sites in the basin whose locations are indicated in figure 3.

The South Coastal Basin has an area of about 240 mi<sup>2</sup> that is drained by several small streams, which discharge into Massachusetts and Cape Cod Bays. The basin is bounded to the north by the Boston Harbor Basin, to the west by the Taunton River and Buzzards Bay Basins, and to the south by Cape Cod. The basin is mostly flat and is characterized by low-gradient streams and many wetlands, lakes, ponds, and cranberry bogs. The largely residential area has a population of about 136,000. It has been subdivided into two parts, the North and South Rivers subbasin and the South Coastal Shore subbasin. The North and South Rivers subbasin has a drainage area of about 105 mi<sup>2</sup> and includes all or parts of 13 towns. About one-third of the subbasin is underlain by stratified-drift deposits. The North River is affected by tides for about 13 mi upstream from its mouth, and the South River is affected by tides for about 8 mi above its mouth. The South Coastal Shore subbasin has a drainage area of about 135 mi<sup>2</sup> and includes parts of seven towns. About 90 percent of the subbasin is underlain by stratified-drift deposits. The South Coastal Basin contains two streamflow-gaging stations—one in each of the subbasins. Both stations are affected by regulation by dams and cranberry bogs. Names and descriptions for these stations and for six sites where low-flow estimates are provided later in the report are listed in table 1. The sites are shown in figure 4.

The Narragansett Bay Basin has an area of 112 mi<sup>2</sup>. The land drains generally southward into Rhode Island, and eventually into Mount Hope Bay, a part of Narragansett Bay. Parts of seven towns and the city of Fall River are in the basin, which has a population of about 115,000 (U.S. Geological Survey, 1988). The basin is bounded to the northeast by the Taunton River Basin, which bisects the Narragansett Bay Basin near the mouth of the Taunton River. The Narragansett Bay Basin is bounded to the southeast by the Buzzards Bay Basin, to the northwest by the Tenmile River Basin, and to the southwest by the Rhode Island border. Topography is similar to that of the adjacent Tenmile River Basin. No streamflow-gaging stations are active in the basin, although there was a station on the West Branch Palmer River near Rehoboth from 1962 to 1974 (USGS station number 01109200, drainage area 4.35 mi<sup>2</sup>). Estimated low-flow duration discharges are provided in this report for four sites in the basin whose locations are indicated in figure 5. Names and descriptions for the sites are in table 1.

**Table I.** Descriptions of selected sites for which low-flow estimates are provided

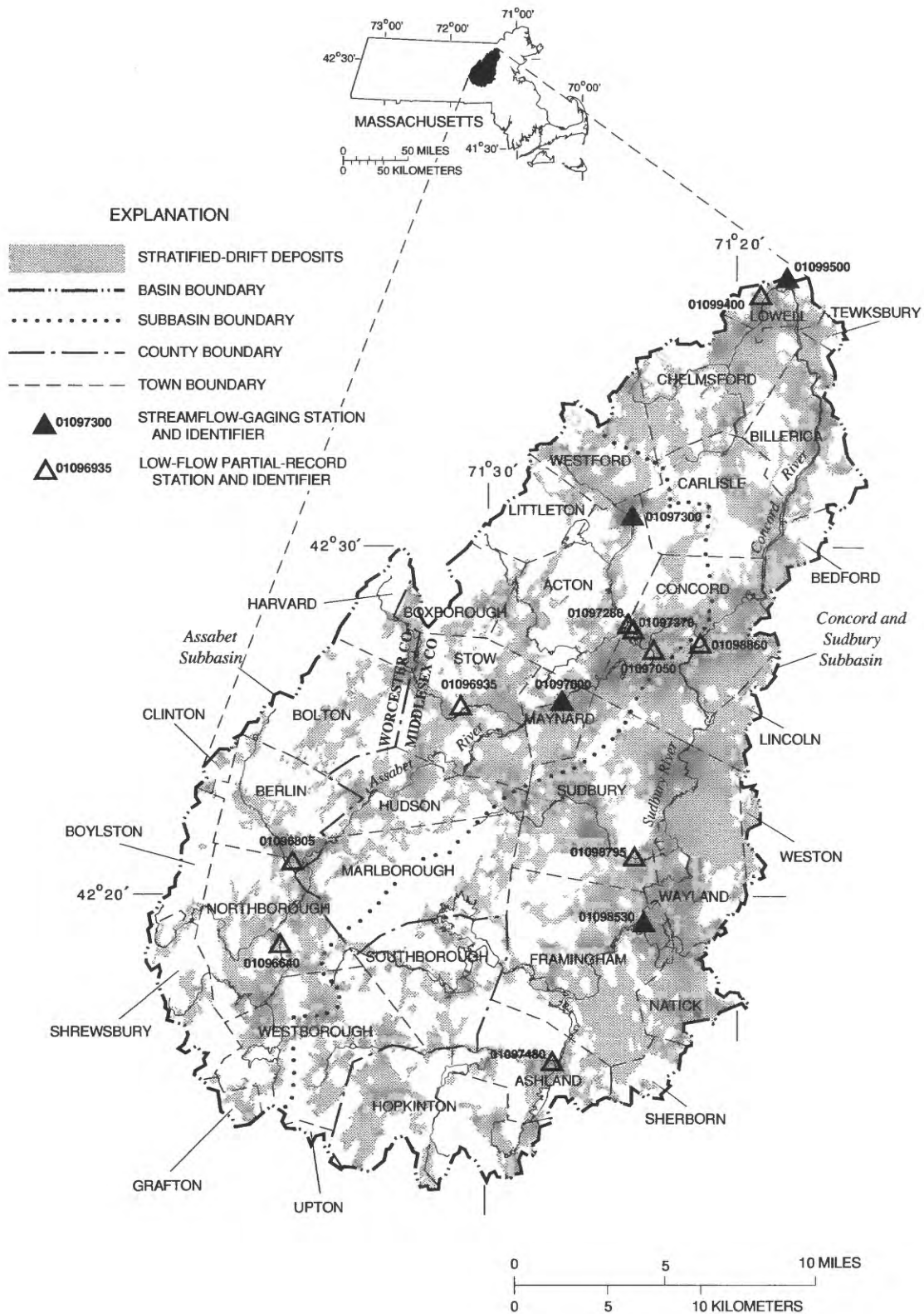
USGS station No.	Latitude	Longitude	Station name	Location
<b>CONCORD RIVER BASIN</b>				
01096640	42°18'55"	71°38'05"	Assabet River at Northborough	At culvert on railroad line east of School Street
01096805	42°21'15"	71°37'40"	North Brook at Berlin	At culvert on Whitney Street
01096935	42°25'47"	71°30'56"	Elizabeth Brook at Wheeler Street at Stow	At culvert on Wheeler Street
01097000	42°25'55"	71°27'01"	Assabet River at Maynard	150 feet upstream from bridge on State Route 27
01097050	42°27'23"	71°23'26"	Assabet River near Concord	At bridge in Main Street
01097280	42°28'07"	71°24'31"	Fort Pond at West Concord	400 feet upstream from confluence with Nashoba Brook
01097300	42°31'19"	71°24'15"	Nashoba Brook near Acton	1,500 feet downstream from dam at North Acton
01097370	42°28'06"	71°24'26"	Nashoba Brook at West Concord	200 feet downstream from State Highway 2
01097480	42°15'36"	71°27'25"	Sudbury River near Ashland	At culvert on Fountain Street
01098530	42°19'31"	71°23'53"	Sudbury River at Saxonville	At downstream side of Danforth Street Bridge
01098795	42°21'26"	71°24'13"	Wash Brook near Sudbury	At culvert on Landham Road
01098860	42°27'35"	71°21'35"	Sudbury River at Concord	At bridge on Nashawtuc Street
01099400	42°37'29"	71°19'11"	River Meadow Brook at Lowell	At bridge on Plain Street
01099500	42°38'12"	71°18'09"	Concord River below Meadow Brook at Lowell	300 feet downstream from Rogers Street bridge
<b>NORTH COASTAL BASIN</b>				
01102031	42°38'08"	70°36'38"	Sawmill Brook at Rockport	At bridge on State Route 127A
01102033	42°37'51"	70°39'14"	Alewife Brook at Gloucester	At outlet of Mill Pond
01102330	42°30'42"	71°02'11"	Saugus River near Lynnfield	At bridge on Salem Street
01102344	42°28'20"	71°00'26"	Saugus River at Saugus	At bridge on Elm Street
<b>SOUTH COASTAL BASIN</b>				
01105660	42°12'44"	70°49'18"	Bound Brook near Cohasset	At culvert on Doane Street
01105670	42°11'36"	70°43'37"	Satuit River at Scituate	200 ft behind senior center of Old Parish Road
01105730	42°06'02"	70°49'23"	Indian Head River at Hanover	At bridge on Elm Street
011058057	42°05'55"	70°48'37"	Confluence of Indian Head River and Herring Brook	About 0.7 mile upstream from bridge on State Routes 53 and 139
01105810	42°07'01"	70°48'35"	Third Herring Brook at Hanover	At culvert on River Street
01105820	42°09'36"	70°47'20"	Second Herring Brook at Norwell	At culvert on State Highway 123
01105845	42°05'32"	70°43'50"	South River at Marshfield	At culvert on Old Ocean Street
01105870	41°59'27"	70°44'03"	Jones River at Kingston	100 feet downstream from bridge on Elm Street

**Table 1.** Descriptions of selected sites for which low-flow estimates are provided--*Continued*

USGS station No.	Latitude	Longitude	Station name	Location
<b>NARRAGANSETT BAY BASIN</b>				
01109135	41°46'30"	71°11'57"	Cole River at Hortonville	At bridge on Hortonville Road
01109215	41°49'50"	71°16'55"	Palmer River near Rehoboth	At bridge on Summer Street
01109225	41°46'52"	71°15'03"	Rocky Run near Rehoboth	At culvert on Davis Street
01109280	41°48'59"	71°20'10"	Runnins River at Fall River Avenue at Seekonk	At bridge on Fall River Avenue
<b>TENMILE RIVER BASIN</b>				
01109375	41°56'59"	71°17'17"	Bungay River 0.5 mi upstream from mouth, near Attleboro	0.5 mile upstream from Mechanics Pond
01109391	41°53'41"	71°20'17"	Tenmile River at pond inlet at State line, near Attleboro	At inlet of pond 0.5 mile downstream from Pond Street
01109394	41°53'52"	71°20'25"	Sevenmile River at mouth, at Pawtucket, R.I.	At State line 0.35 mile downstream from County Road
01109397	41°52'20"	70°20'04"	Unnamed Brook at Newman Avenue, near Seekonk	At culvert on Newman Avenue
01109398	41°51'28"	71°19'48"	Coles Brook at Newman Avenue, near Seekonk	At culvert on Newman Avenue

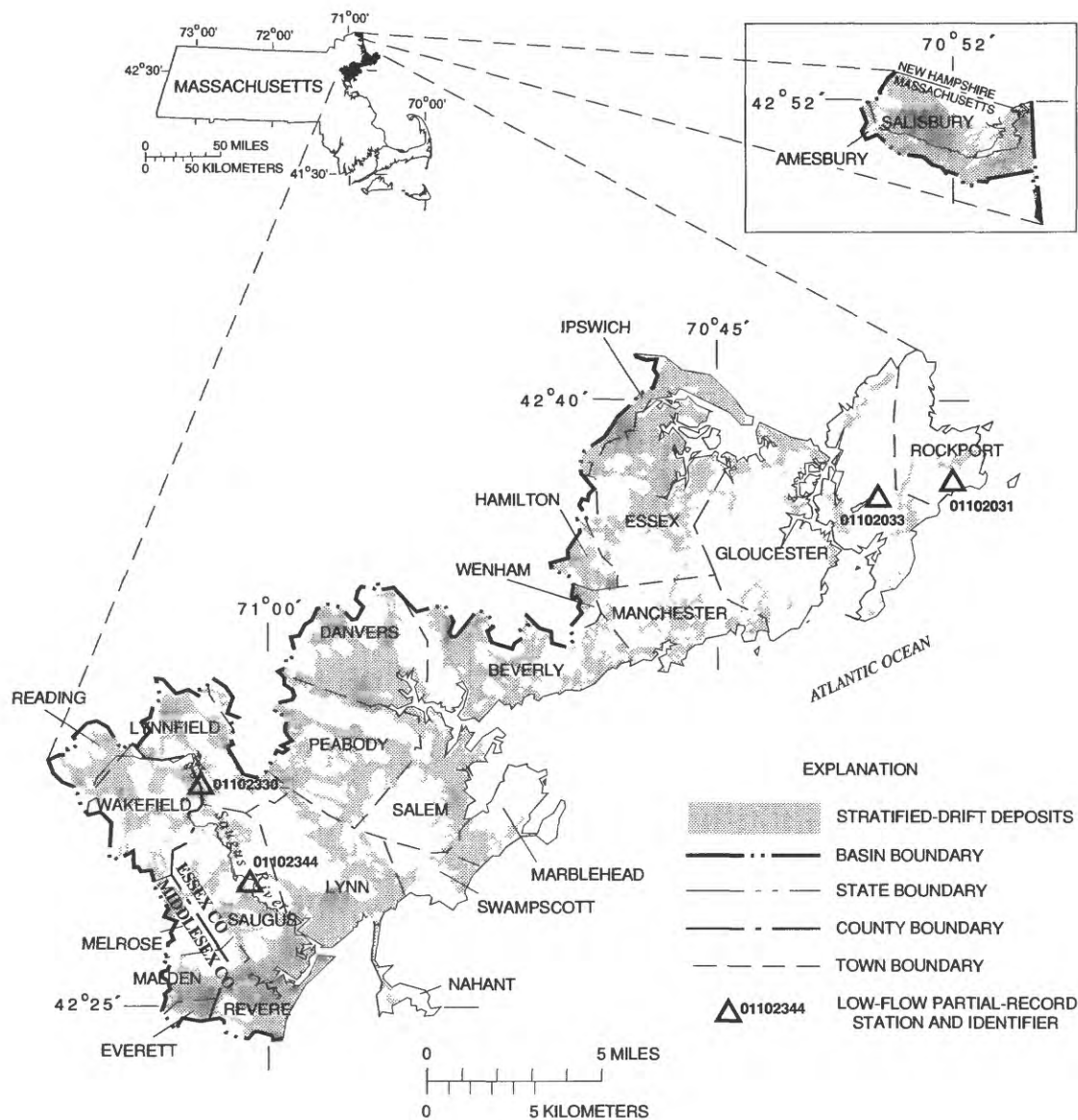
The Tenmile River Basin has an area in Massachusetts of 49 mi<sup>2</sup>, most of which is residential land. Parts of seven towns and the city of Attleboro are in the basin, which has a population of about 56,500 (U.S. Geological Survey, 1988). The basin is bounded on the west by the Blackstone River Basin, on the north by the Charles River Basin, on the east by the Taunton River Basin, and on the southeast by the Narragansett Bay Basin. The Massachusetts-Rhode Island border is on the southwest. The basin is mostly flat; some rounded hills in northern parts of the basin reach a maximum elevation of about 430 ft above sea level. Stratified-drift deposits underlie about one-half of the basin. Major tributaries to the Tenmile River are the Sevenmile and Bungay Rivers.

Streamflows in the tributaries and main stem of the Tenmile River are substantially affected by dam regulations and diversions for municipal water supplies. No streamflow-gaging stations are present in the Massachusetts part of the basin, although there is a streamflow-gaging station downstream from the State border at East Providence, Rhode Island (USGS station number 01109403, drainage area 53.1 mi<sup>2</sup>). The Tenmile River flows into Narragansett Bay about 0.5 mi downstream from this station. Estimated low-flow duration discharges are provided in this report for five sites in the basin whose locations are indicated in figure 5. Names and descriptions for the sites are in table 1.

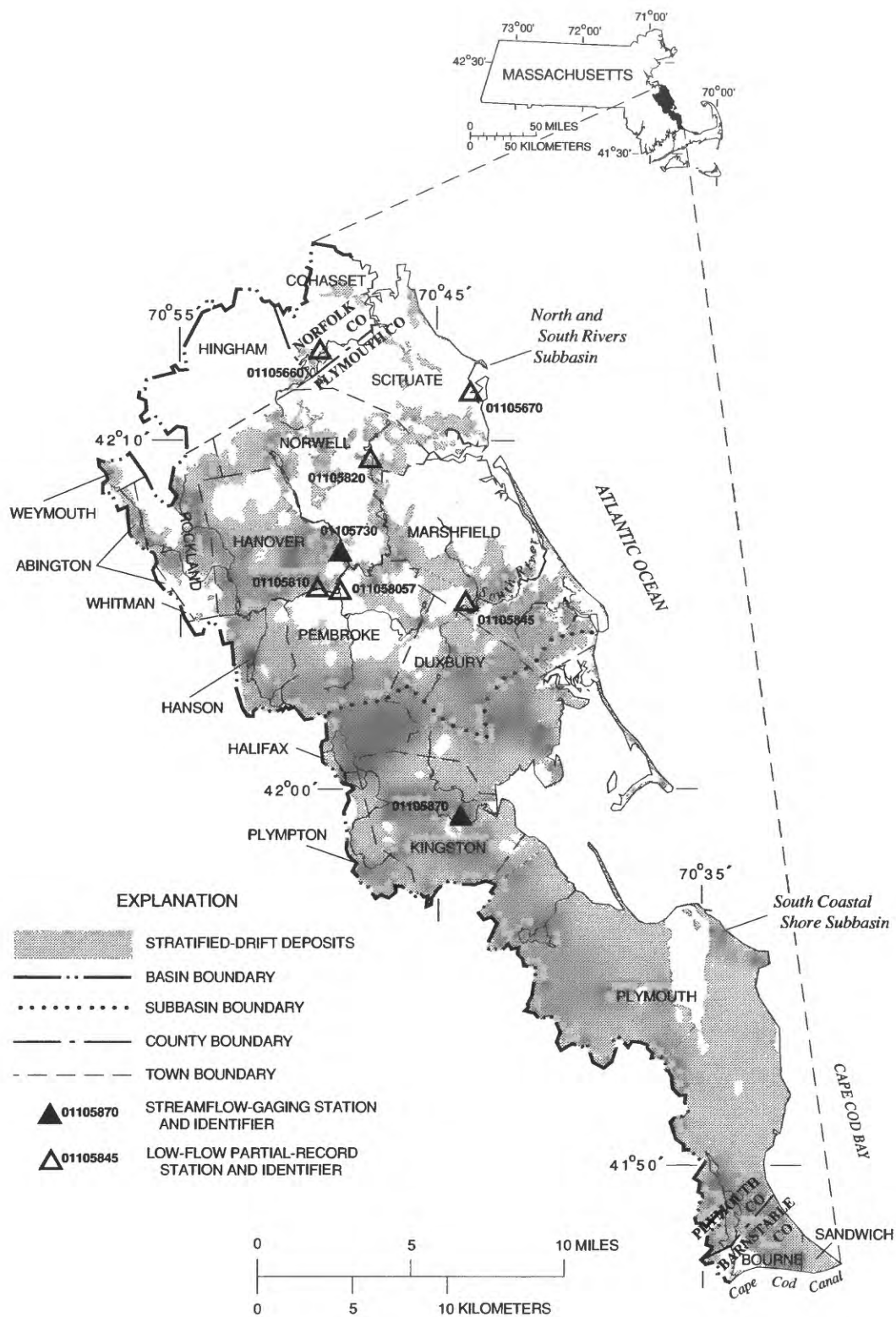


**Figure 2.** Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the Concord River Basin.

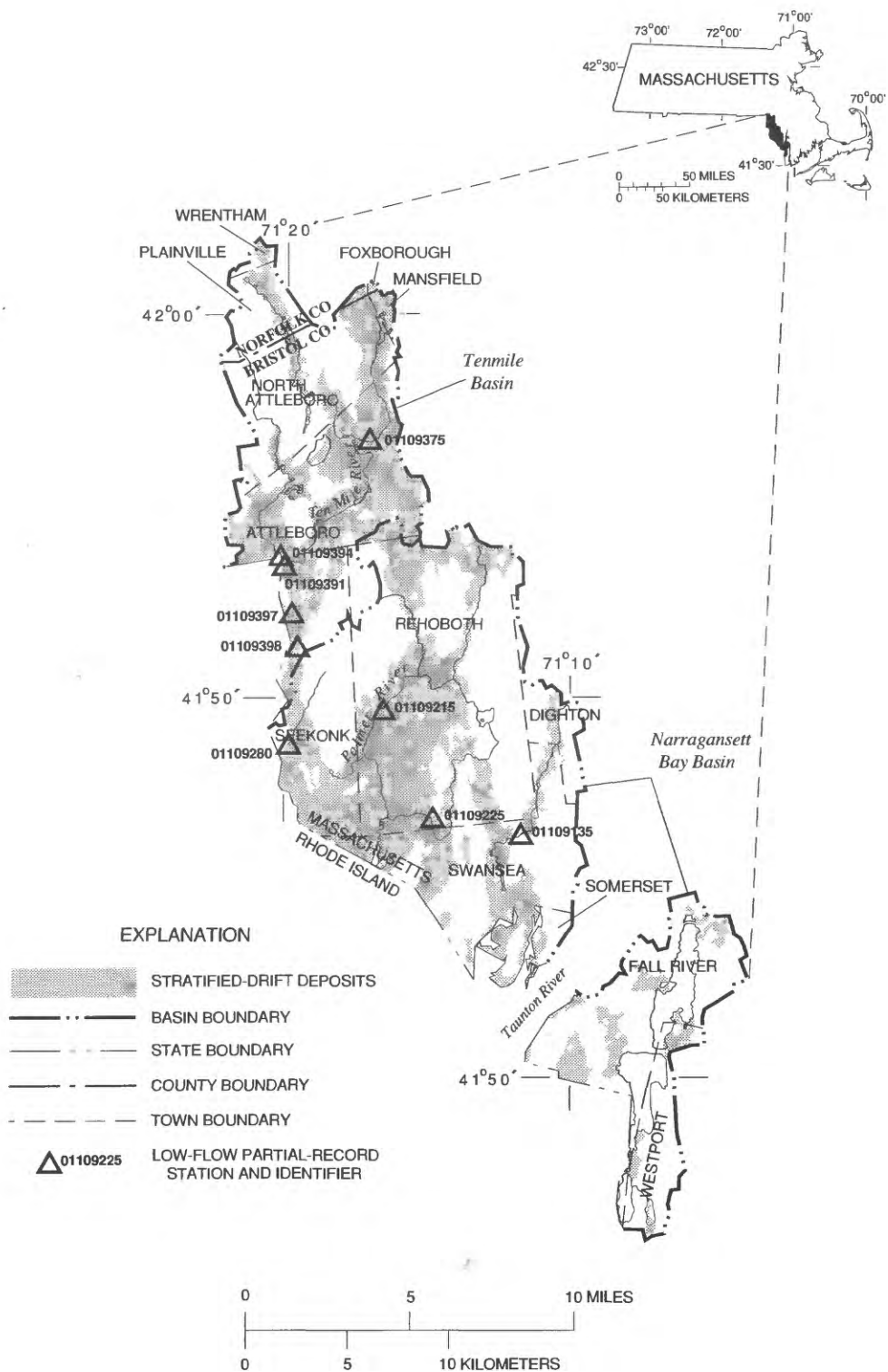




**Figure 3.** Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the North Coastal Basin.



**Figure 4.** Drainage boundaries, areas of stratified-drift deposits, and locations of selected sites in the South Coastal Basin.



**Figure 5.** Drainage boundaries, areas of stratified-drift deposits and locations of selected sites in the Narragansett Bay and Tenmile River Basins.

## DATA BASE DEVELOPMENT

The USGS operated 64 continuous streamflow-gaging stations in Massachusetts during water year 1991<sup>1</sup>, the latest water year for which data were available for this study. Periods of record for these stations ranged from 3 years (Quashnet River at Waquoit Village) to 88 years (Connecticut River at Montague City). In addition, at least one year of continuous data have been collected at 73 gaging stations that were not active during 1991. The USGS also has made at least one streamflow measurement at more than 1,000 miscellaneous-measurement and low-flow partial-record sites in the State. Data collected at the miscellaneous-measurement sites generally were used to provide information for specific studies, whereas data collected from networks of low-flow partial-record sites generally were used for low-flow investigations.

Discharge data for all streamflow-gaging stations and most miscellaneous-measurement and low-flow partial-record sites are stored in data bases at USGS offices in Albany, N.Y., and Reston, Va. Information regarding the availability of data and statistical analyses based on the data can be obtained from the U.S. Geological Survey, 28 Lord Road, Suite 280, Marlborough, MA 01752.

Most sites where streamflow data have been collected are not suitable for use in regional low-flow analyses because the low flows tend to be affected by dam regulations or diversions for municipal or industrial water supplies. Streamflows at the sites selected from the data base for use in this study were minimally affected by regulations and diversions. Streamflows at some of the selected sites were affected by regulations or diversions at higher flows, but the low flows occurred under essentially natural conditions. Sixty-one sites were selected from the available data for use in the analyses—37 streamflow-gaging stations and 24 low-flow partial-record sites. Of the 37 streamflow-gaging stations selected, 17 were operated during water year 1991, and 20 were discontinued. Locations and types of sites used in the analyses are in figure 6. Descriptions of the sites are in table 2.

<sup>1</sup>A water year is a period beginning October 1 of the previous calendar year, and ending September 30 of the year specified.

## Computation of Selected Duration Discharges

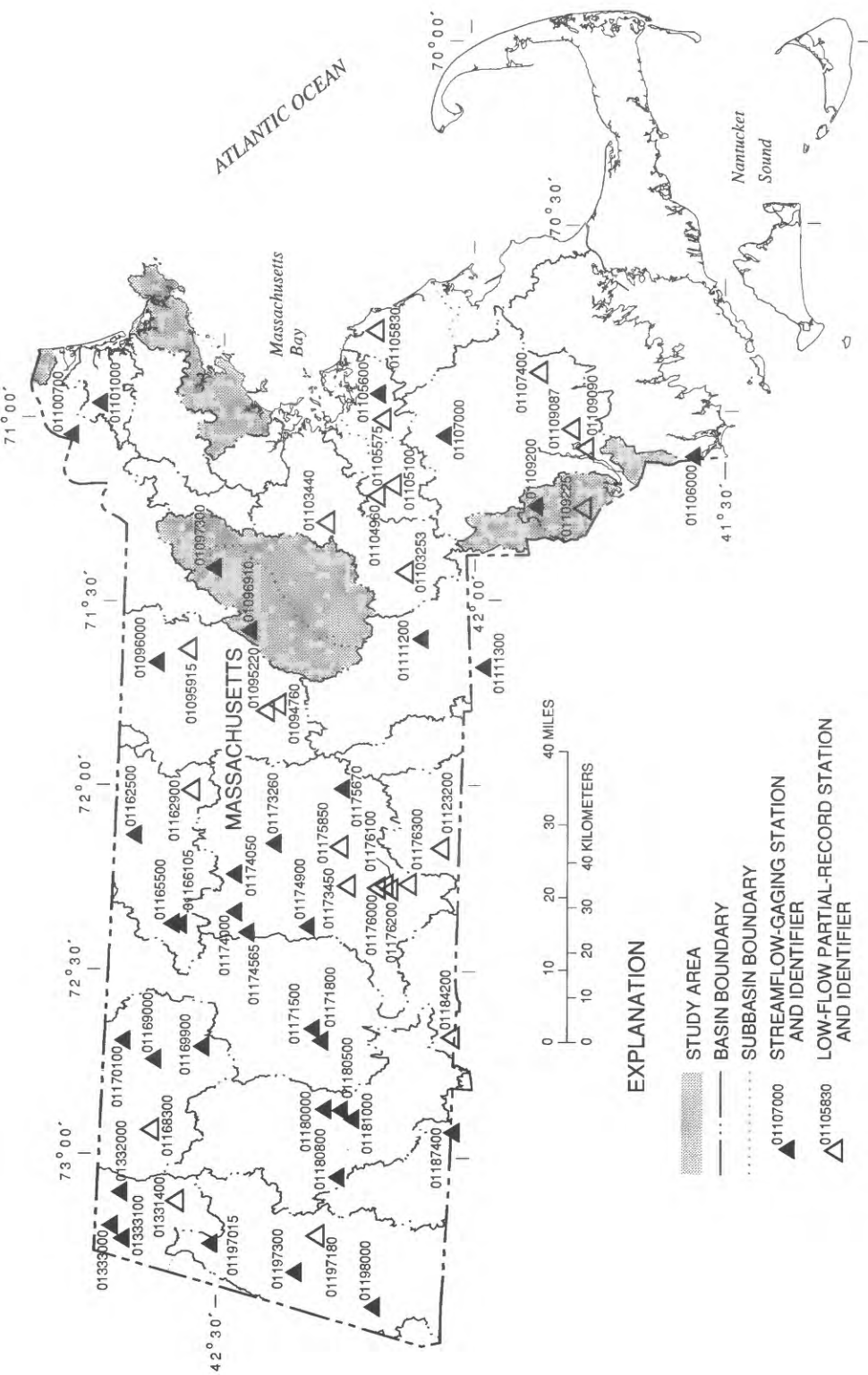
The 95-, 98-, and 99-percent duration discharges were used as the dependent variables in the regression analyses. These values were computed for each streamflow-gaging station used in the regression analyses from the available daily discharges for their periods of record. Only complete climatic years, which begin on April 1, were included in the computations. Periods of record for the streamflow-gaging stations ranged from 2 to 80 years, with a mean record length of 31 years.

In addition to the values used in the regression analyses, annual series of the 95-, 98-, and 99-percent duration discharges were computed for each streamflow-gaging station. The constructed annual series were used to test for trends in the streamflow data and to compute matrices of cross-correlations between the streamflows at the gaged sites. The matrices were used in the GLS regression analyses to assign weights to the sites used in the analyses based on their cross-correlations and periods of record.

Mann-Kendall tests for trend were done on the annual series of low-flows for all gaged sites with 10 or more years of record (Helsel and Hirsch, 1992, p. 326). Upward trends were significant in the annual series of at least one of the three computed duration discharges for 11 sites with periods of record beginning during the early to middle 1960's. The most extreme drought of record for most areas of Massachusetts occurred during 1962-66. Annual duration discharges were at or near their minimums for most gaging stations during these years. Trends were found for the stations with records beginning during the drought because their minimum annual values were at the beginning of the time series. Subsequent annual values could only increase from their minimums. No adjustments were made for these trends because their origin probably is due to short-term climatic variation rather than any real long-term trend.

Duration discharges for the partial-record sites used in the regression analyses were determined by correlating streamflow measurements made at the partial-record sites to concurrent daily mean discharges at nearby gaging stations. Correlation of the concurrent discharges was done by graphical (Searcy, 1959, p. 14) or mathematical (Hirsch, 1982) techniques, depending on the availability of data and on whether plots indicated the presence of curvature in the relation between the flows at the sites. Correlation coefficients





**Figure 6.** Type and locations of sites used to develop the regression models.

**Table 2.** Description of streamflow-gaging stations and low-flow partial-record sites used to develop the regression models

USGS station No.	Latitude	Longitude	Station name	Period of record or number of measurements	Remarks
01094760	42°23'49"	71°46'48"	Waushacum Brook near West Boylston	11	Partial-record site
01095220	42°24'39"	71°47'30"	Stillwater Brook near Sterling	14	Partial-record site
01095915	42°34'26"	71°37'43"	Mulpus Brook near Shirley	13	Partial-record site
01096000	42°38'03"	71°39'30"	Squannacook River near West Groton	1950-present	Occasional regulation by mill upstream
01096910	42°27'04"	71°13'43"	Boulder Brook at East Bolton	1972-83	--
01097300	42°30'39"	71°24'25"	Nashoba Brook near Acton	1964-present	--
01100700	42°48'41"	71°01'59"	East Meadow Brook near Haverhill	1962-74	--
01101000	42°45'10"	70°56'46"	Parker River at Byfield	1946-present	Occasional regulation by mill and ponds
01103253	42°08'27"	71°25'26"	Chicken Brook near West Medway	19	Partial-record site
01103440	42°17'45"	71°17'18"	Fuller Brook at Wellesley	17	Partial-record site
01104960	42°11'04"	71°13'29"	Germany Brook near Norwood	15	Partial-record site
01105100	42°09'36"	71°11'47"	Traphole Brook near Norwood	14	Partial-record site
01105575	42°11'02"	71°00'42"	Cranberry Brook at Braintree Highlands	16	Partial-record site
01105600	42°11'25"	70°56'43"	Old Swamp River near South Weymouth	1966-present	--
01105830	42°11'30"	70°46'49"	First Herring Brook near Scituate	14	Partial-record site
01106000	41°33'30"	71°07'47"	Adamsville Brook at Adamsville, RI	1941-78	--
01107000	42°03'41"	71°03'59"	Dorchester Brook near Brockton	1963-74	--
01107400	41°51'55"	70°54'32"	Fall Brook near Middleboro	36	Partial-record site. Continuous record during 1967 not used due to poor stage-discharge relation
01109087	41°55'23"	71°03'37"	Assonet River at Assonet	22	Partial-record site
01109090	41°46'36"	71°05'23"	Rattlesnake Brook near Assonet	11	Partial-record site
01109200	41°52'46"	71°15'18"	West Branch Palmer River near Rehoboth	1962-74	--
01109225	41°46'52"	71°15'03"	Rocky Run near Rehoboth	32	Partial-record site
01111200	42°06'17"	71°36'28"	West River at West Hill Dam near Uxbridge	1962-90	Flood-control dam upstream
01111300	41°58'52"	71°41'11"	Nipmuc River near Harrisville, RI	1964-91	--
01123200	42°03'41"	72°09'45"	Stevens Brook at Holland	23	Partial-record site
01162500	42°40'57"	72°06'56"	Priest Brook near Winchendon	1919-present	No daily record during climatic years 1936-37
01162900	42°33'52"	72°00'43"	Otter River at Gardner	16	Partial-record site
01165500	42°36'10"	72°21'36"	Moss Brook at Wendell Depot	1918-82	
01166105	42°35'39"	72°21'41"	Whetstone Brook, at Depot Road, Wendell Depot	1986-present	--
01168300	42°38'12"	72°56'10"	Cold River near Zoar	19	Partial-record site

**Table 2.** Description of streamflow-gaging stations and low-flow partial-record sites used to develop the regression models --*Continued*

USGS station No.	Latitude	Longitude	Station name	Period of record or number of measurements	Remarks
01169000	42°38'18"	72°43'32"	North River at Shattuckville	1939-present	Occasional small diurnal fluctuation
01169900	42°32'31"	72°41'39"	South River near Conway	1966-91	Small diurnal fluctuation since 1982
01170100	42°42'12"	72°40'16"	Green River near Colrain	1968-present	--
01171500	42°19'05"	72°39'21"	Mill River at Northampton	1940-present	--
01171800	42°18'09"	72°41'16"	Bassett Brook near Northampton	1963-74	--
01173260	42°23'52"	72°08'51"	Moose Brook near Barre	1962-74	--
01173450	42°14'56"	72°15'53"	Flat Brook near Ware	16	Partial-record site
01174000	42°28'42"	72°20'05"	Hop Brook near New Salem	1947-82	--
01174050	42°28'49"	72°13'27"	East Branch Fever River near Petersham	1984-85	--
01174565	42°27'18"	72°22'56"	West Branch Swift River at Shutesbury	1984-85	--
01174900	42°20'08"	72°22'12"	Cadwell Creek near Belchertown	1962-present	--
01175670	42°15'54"	72°00'19"	Sevenmile River near Spencer	1961-present	Occasional regulation by ponds upstream
01175850	42°15'50"	72°09'33"	Mill Brook near West Brookfield	12	Partial-record site
01176000	42°10'56"	72°15'51"	Quaboag River at West Brimfield	1913-present	Flood-retarding reservoirs upstream
01176100	42°10'13"	72°15'41"	Blodgett Mill Brook at West Brimfield	15	Partial-record site
01176200	42°09'41"	72°16'08"	Kings Brook at West Brimfield	14	Partial-record site
01176300	42°07'43"	72°15'31"	Foskett Mill Brook at Fentonville	10	Partial-record site
01180000	42°17'27"	72°52'15"	Sykes Brook at Knightville	1946-73	--
01180500	42°15'31"	72°52'23"	Middle Branch Westfield River at Goss Heights	1910-present	Data for climatic years 1965-66 not used due to construction of flood-control reservoir upstream
01180800	42°15'49"	73°02'48"	Walker Brook near Becket Center	1963-77	--
01181000	42°14'14"	72°53'46"	West Branch Westfield River at Huntington	1937-present	--
01184200	42°02'31"	72°41'00"	Still Brook near West Agawam	15	Partial-record site
01187400	42°02'03"	72°55'49"	Valley Brook near West Hartland, Ct.	1941-72	--
01197015	42°31'12"	73°13'48"	Town Brook at Bridge Street, Lanesborough	1980-83	--
01197180	42°17'59"	73°12'53"	Greenwater River at East Lee	8	Partial-record site
01197300	42°20'59"	73°17'56"	Marsh Brook at Lenox	1963-74	--
01198000	42°11'31"	73°23'28"	Green River near Great Barrington	1952-71	--
01331400	42°35'20"	73°06'48"	Dry Brook near Adams	1963-74	--
01332000	42°42'08"	73°05'37"	North Branch Hoosic River at North Adams	1933-90	Infrequent small diurnal fluctuation
01333000	42°42'32"	73°11'50"	Green River at Williamstown	1951-present	Infrequent small diurnal fluctuation
01333100	42°41'16"	73°13'50"	Hemlock Brook near Williamstown	14	Partial-record site

were computed between the logarithms of the concurrent discharges at the partial-record sites and those at the index gaging stations. Only those partial-record sites with correlation coefficients greater than 80 percent were included in the regression analyses.

The correlation procedures used to estimate duration discharges for partial-record sites for this study are identical to those used in the initial Basin Yield study; they are discussed in much greater detail in the report for that study (Ries, 1993, p. 22-30). For convenience, and to have the estimated duration discharges for the partial-record sites used in the regression analyses for this study represent long-term conditions, the estimates were computed for the same 25-year base period (October 1, 1962 through September 30, 1987) as that used for the initial Basin Yield study.

### **Selection and Measurement of Basin Characteristics**

Basin characteristics used in the regression analyses for this study include those used in the regression equations developed for the first Basin Yield study, plus additional characteristics measured for this study. All basin characteristics were measured using geographic information systems (GIS) computer software. Characteristics common to both studies were measured in the same manner as that described in the report for the first study (Ries, 1993, p. 33-36). The basin characteristics measured include (1) drainage-basin area; (2) total length of all streams in the basin; (3) total area of coarse-grained stratified-drift deposits in the basin; (4) total area of all water bodies in the basin; (5) mean, minimum, and maximum basin elevations; (6) mean, minimum, and maximum elevations of stratified-drift deposits; and (7) mean basin slope. The elevations of stratified-drift deposits and mean basin slope were not measured in the initial study.

Mean basin slopes, in percent, were calculated by the GRID option of ARC/INFO GIS software, from elevation data contained in the Survey's nationwide 1:250,000-scale digital elevation model (DEM) data layer (Elassal and Caruso, 1983; Environmental Systems Research Institute, Inc., 1991). These data are the same as those used to obtain mean, minimum, and maximum basin and stratified-drift elevations for both Basin Yield studies.

Because of enhancements to the data layer from which stream lengths were measured for the first Basin Yield study, stream lengths measured for some sites that were used in the initial study are slightly different than those for the same sites in this study. The new stream lengths supersede those published in the earlier report.

As with the first study, several additional basin characteristics were computed from the 11 characteristics listed above for this study: RELIEF, the difference, in feet, between the highest and lowest elevations in the basin; DENSITY, the total length of streams in the basin, in miles, divided by the basin area, in square miles; DRT/TST, the area of stratified-drift deposits, in square miles, divided by the total length of streams in the basin, in miles; and GWHEAD, the difference, in feet, between the mean basin elevation and the lowest basin elevation. All of the above-named basin characteristics were tested in initial regression analyses for possible use in the final regression equations.

The drainage areas for some sites did not contain stratified-drift deposits. Because the regression analyses were done on the logarithms of the streamflow statistics and the basin characteristics, it was necessary to add a constant to the areas of stratified-drift deposits and to the values of DRT/TST measured for all sites included in the regression analyses. A constant of 0.1 was selected on the basis of sensitivity tests done during the initial Basin Yield study.

GWHEAD was used in the first study as a surrogate for the effective head in the stratified-drift aquifers (the vertical difference between the highest point in the aquifer and the discharge point). This surrogate was felt to be a reasonable approximation to the actual head in the aquifers because stratified-drift aquifers generally are located in the lower elevations of most basins. GWHEAD was one of the three variables that, in combination, provided the best equations to predict duration discharges for the first Basin Yield study. Because of this, it was felt that use of a more direct measurement of aquifer head might improve the newer regression equations. A new measurement of aquifer head was obtained by determining the minimum and maximum elevations in the stratified-drift deposits of the basins from the overlain GIS drainage area, stratified drift, and DEM data layers. The minimum stratified-drift elevation for each basin was then subtracted from the maximum stratified-drift elevation to determine relief in the

stratified-drift deposits. These measurements were used to create a new variable for the regression analyses called GWRELIEF.

GWRELIEF was expected to be a more direct measure of the effective head in the stratified-drift aquifers than is GWHEAD. GWRELIEF and GWHEAD were used in the initial regression analyses, and are discussed in the following section. Interestingly, GWRELIEF had the highest individual correlation with discharges of 98 and 99 percent duration of any of the independent variables. Its correlation was just slightly less than that of DAREA, the independent variable with the largest individual correlation with the 95-percent duration discharges. However, when used in combination with other independent variables, GWHEAD was always more significant than GWRELIEF; therefore, GWRELIEF did not appear in the final regression equations.

## DEVELOPMENT OF THE GENERALIZED-LEAST-SQUARES REGRESSION MODELS

Multiple-regression analyses were used to relate the 95-, 98-, and 99-percent duration discharges (dependent variables) to measured basin characteristics (independent variables). Data for the 61 sites used in the analyses are listed in table 6 (at back of report). Logarithmic (base 10) transformations were done on all of the data to linearize the relations between the dependent and the independent variables and to normalize the distributions of the residual errors.

The models obtained from the multiple-regression analyses done for this study are similar to those developed for the initial Basin Yield study (Ries, 1993). The models are log-linear equations of the form

$$\log_{10} Y = b_0 + b_1 \log_{10} X_1 + b_2 \log_{10} X_2 + \dots + b_n \log_{10} X_n + \varepsilon, \quad (1)$$

or, after retransforming by taking antilogs to obtain the algebraically equivalent form

$$Y = 10^{b_0} (X_1^{b_1}) (X_2^{b_2}) \dots (X_n^{b_n}) 10^{\varepsilon}, \quad (2)$$

where

$Y_i$  is the dependent variable (the 99-, 98-, or 95-percent duration discharge) for site  $i$ ,

$X_1$  to  $X_n$  are the  $n$  independent variables (basin characteristics),

$b_0$  to  $b_n$  are the  $n+1$  regression model coefficients, and

$\varepsilon$  is the residual error for site  $i$ .

The initial Basin Yield study used a weighted-least-squares (WLS) regression analysis procedure to obtain the models. The WLS procedure gave each site in the analysis a different weight, assigned in proportion to the record lengths and inversely proportional to the variance of the regression residuals. The weights used for the WLS analyses did not compensate for cross (spatial) correlation between the concurrent streamflows of the sites used in the regression analyses.

The all-possible-regressions selection procedure BREG of the MINITAB statistical computer software was used to select subsets of independent variables for potential inclusion in the final regression equations (Ryan and others, 1985, p. 360). Data for each of the sites included in these analyses were weighted by the number of years of record at the site for gaging stations, and by the number of years during which streamflow measurements were obtained for the partial-record sites. These values are listed in the "years of record" column of table 6. The five best combinations of 2, 3, and 4 variables were selected from the weighted all-possible-subsets analysis for further testing by use of generalized-least-squares (GLS) regression analysis. Selection was based on minimizing Mallows's  $C_p$  and the regression standard error (Neter and others, 1985, p. 421-429).

## Generalized-Least-Squares Regression Analyses

The final regression equations described later in this report were developed by use of GLS regression analyses. The GLS procedure allows the weight given to each site in the analyses to be adjusted for cross correlation among the concurrent streamflows of the sites and for differences in their record lengths. Adjustments to the weights for cross correlation were based on an empirical relation between cross-correlations of constructed annual series of duration discharges for sites with 30 or more years of record and the distances between the sites. Tasker and Stedinger (1989) provide additional details on the theory and application of the GLS procedure.

GLS regression analysis has been shown to be more efficient and more statistically appropriate than ordinary-least-squares (OLS) or WLS regression analysis when used to regionalize streamflow statistics (Stedinger and Tasker, 1985). GLS also allows the prediction error for ungaged sites to be partitioned into model error and sampling error. This partitioning provides information about how future models can be improved. The model error is the error inherent in the imperfect model. The model error cannot be reduced by collecting more streamflow data at the sites used in the analysis, but it may be reduced by adding new sites to the analysis, or by use of physical or climatological characteristics that better describe the variation in the dependent variable. The sampling error is caused by the use of estimates for the actual values of the parameters in the regression equations. Sampling errors may be reduced by collecting additional streamflow data at the sites used in the regression analyses and by including additional sites in the analyses.

Subsets of independent variables selected for additional analysis from the weighted all-possible-subsets analyses were tested by use of the GLS procedure. Final model selection was based on (1) minimizing the sum of the model and sampling errors; (2) minimizing PRESS/n, an estimate of the mean prediction error sum of squares that indicates model performance when estimating duration discharges for ungaged sites (Thomas, W.O., Jr.; Lumb, A.M.; Flynn, K.M.; and Tasker, G.D., U.S. Geological Survey, written commun., 1987); (3) passing of diagnostic checks to test for model adequacy, violations of assumptions for regression analysis, and outliers; and (4) hydrologic reasonableness of the selected independent variables and the signs and magnitudes of their coefficients.

## Bias Correction

Because logarithmic (base 10) transformations were done to all data used in the regression analyses, the resulting equations are of the form of equation 1. Estimates of the dependent variables obtained from these equations are in logarithmic units. What is really desired are estimates in their original units of measure, cubic feet per second. Retransformation of the equations by taking the antilogs yields equations of the form of equation 2. These equations provide unbiased estimates of the median response of the dependent variable in their desired units of measure. As required in least squares regression, the estimates are unbiased because

the probability that an individual estimate is too high is the same as the probability that it is too low. However, estimates obtained from these retransformed equations, on average, do not equal their actual values because the retransformed residual errors are no longer normally distributed. When unbiased estimates of the mean response are actually desired, a bias correction factor (BCF) can be applied. Investigators have suggested ways to determine BCF's to reduce or eliminate this retransformation bias in the estimates (Duan, 1983; Miller, 1984; Koch and Smillie, 1986; Cohn and others, 1989; Gilroy and others, 1990).

Duan's (1983) "smearing estimate" of the mean residual error was used to reduce the bias inherent in the equations developed for the study. The smearing estimate BCF is applied by replacing the error term of equation 2 with the mean error of the retransformed residuals, yielding

$$Y = 10^{b_0} (X_1^{b_1}) (X_2^{b_2}) \dots (X_n^{b_n}) \left( \sum_{i=1}^N \frac{10^{\epsilon_i}}{N} \right), \quad (3)$$

where

$N$  is the number of sites used in the regression analysis.

The smearing estimate is the last factor in parentheses in equation 3; it is determined by summing the antilogs of the residual errors from the regression analyses and then dividing the sum by the number of sites used in the regression analyses. This value is then multiplied by the estimate obtained from the retransformed regression equation to provide a reasonably unbiased estimate of the mean response of the dependent variable. BCF's determined for this study were 1.1349, 1.2486, and 1.2970 for the equations used to predict the 95-, 98-, and 99-percent duration discharges, respectively. These factors were combined with the regression constants of the final equations.

## Final Equations for Predicting Long-Term Duration Discharges

The same combination of independent variables, DAREA, GWHEAD, and DRT/TST, provided the best estimates for each of the dependent variables. Regression equations using these variables to predict the long-term 95-, 98-, and 99-percent duration discharges are listed in table 3. Several measures of

**Table 3.** Summary of generalized-least-squares regression equations used to estimate duration discharges for selected sites in Massachusetts, bias correction factors, and measures of model adequacy

**REGRESSION EQUATIONS AND BIAS CORRECTION FACTORS (BCF)**

Regression equation <sup>1</sup>	BCF
95-percent duration discharge	
$\hat{P}_{95} = 0.01618 (\text{DAREA})^{0.9605} (\text{GWHEAD})^{0.7602} (\text{DRT/TST})^{1.3283}$	1.1349
98-percent duration discharge	
$\hat{P}_{98} = 0.00568 (\text{DAREA})^{0.9474} (\text{GWHEAD})^{0.9751} (\text{DRT/TST})^{1.6181}$	1.2486
99-percent duration discharge	
$\hat{P}_{99} = 0.00481 (\text{DAREA})^{0.9226} (\text{GWHEAD})^{0.9461} (\text{DRT/TST})^{1.4744}$	1.2970

**MEASURES OF MODEL ADEQUACY**

Dependent variable	$\gamma^2$	$\hat{\Sigma}_{ave}$	$s_r$	$s_p$	$ e_i _{0.5}$
$\hat{P}_{95}$	0.0490	0.0052	54.5	57.7	35.5
$\hat{P}_{98}$	.0954	.0082	81.1	85.6	47.5
$\hat{P}_{99}$	.1183	.0097	93.4	98.5	52.4

<sup>1</sup>Variables in the above equations, BCF, and measures of model adequacy are defined as follows:

$\hat{P}_{xx}$	the predicted <i>xx</i> -percent duration discharge, in cubic feet per second, corrected for bias;
DAREA	drainage area of the basin, in square miles;
GWHEAD	a surrogate for the effective head on the aquifer discharging to the stream, calculated by subtracting the lowest basin elevation from the mean basin elevation, in feet;
DRT/TST	the area of coarse-grained stratified drift in the basin, in square miles, divided by the total length of all streams in the basin, in miles, plus 0.1;
$\gamma^2$	average model error variance, in $\log_{10}$ units;
$\hat{\Sigma}_{ave}$	average sampling error variance, in $\log_{10}$ units;
BCF	bias correction factor multiplied into the regression constant (Duan, 1983);
$s_r$	standard error of the regression, in percent = $100[\exp(5.3018\gamma^2)-1]^{0.5}$ ;
$s_p$	root mean square prediction error, in percent = $100[\exp(5.3018(\gamma^2 + \hat{\Sigma}_{ave}))-1]^{0.5}$ ; and
$ e_i _{0.5}$	median absolute percentage error of the estimates, from units of cubic feet per second.

model adequacy also are in table 3, including average model and sampling error variances in logarithmic units, standard errors of estimate, root mean square prediction errors, and mean absolute errors of the estimates for the sites used in the regressions, in percent. The root-mean-square prediction error is the square root of the average variance of prediction in which residual errors of a set of sites with basin characteristics identical with the stations used in the regressions are averaged. It is a measure of the average predictive ability of the regression equations over the region (G.D. Tasker, U.S. Geological Survey, written commun., 1994).

The standard errors and the mean absolute errors in table 3 are larger than those obtained for the equations developed for the first Basin Yield study (Ries, 1993, p. 45). However, a direct comparison between the percentage errors obtained for the first study and those for the second study is inappropriate because 20 sites were added to the analyses for the second study. The drainage areas and the observed discharges at the newer sites are smaller on average than those at the sites included in the first study. For example, the median 95-percent duration discharge at the sites included in the first study is more than twice that at the 20 sites added for the second study. For given differences between their observed and predicted discharges, percentage errors in discharge at the new sites (with smaller drainage areas) are larger on average than those for the sites used in the first study (with larger drainage areas). The newer sites were added to the regression analyses to increase sample size and to provide a wider range of basin characteristics, thereby improving definition of the parameters of the models and increasing their ranges of applicability.

Some of the differences between the errors associated with the equations from the two studies also may be attributable to the fact that a base period was used to compute discharges for the initial study. In the first study, records for sites with less than complete record during the base period were extended so that the duration discharges computed for all sites represented the same 25-year period. Discharges for the second study were computed from the entire periods of record for the sites, possibly causing larger time-sampling errors for the equations from the second study than for those from the first study. These increased time-sampling errors could increase model errors.

The errors in the GLS regression equations generated for the second study probably reflect the actual errors associated with use of the equations better than do the errors in the equations for the previous study. The equations from the GLS procedure more closely satisfy the assumptions required for proper use of regression analysis than do the equations obtained from the WLS procedure used for the first study. However, previous studies have found that differences in errors between WLS and GLS regression equations generally are small (Stedinger and Tasker, 1985; Tasker and Stedinger, 1986). Differences between the results of OLS, GLS, and WLS regression analyses and usefulness of a base period will be examined more closely in a third Basin Yield study currently in progress.

### Prediction Intervals

Prediction intervals at the 90-percent confidence level were calculated for most estimated duration discharges that were obtained from the regression equations. The prediction intervals were calculated to provide estimates of the uncertainty inherent in use of the equations. Tasker and Driver (1988) have shown that a  $100(1-\alpha)$  prediction interval for the true value of a dependent variable obtained by use of GLS regression equations corrected for bias can be computed. Substituting terms used in this report, a prediction interval is computed by

$$\frac{1}{T} \frac{P_{xx}}{BCF_{xx}} < P_{xx} < \frac{P_{xx}}{BCF_{xx}}, \quad (4)$$

where

$P_{xx}$  is the predicted discharge at the  $xx$  percent duration,

$BCF_{xx}$  is the bias correction factor shown in table 3 for the equation used to obtain  $P_{xx}$ , and

$T$  is computed as explained below.

A row vector,  $X_i$ , of logarithms of the basin characteristics for site  $i$  is augmented by a 1 as the first element to obtain,  $X_i = [1, \log_{10}(\text{DAREA}), \log_{10}(\text{GWHEAD}), \log_{10}(\text{DRT/TST})]$  for the site. The standard error of prediction for site  $i$  is then estimated as

$$S_i = [\gamma^2 + X_i U X_i']^{0.5}, \quad (5)$$



where

- $\gamma^2$  is the model error variance,
- $U$  is the covariance matrix for the regression coefficients for the equation, both provided in table 4, and
- $X_i'$  is the transpose of  $X_i$  (Ludwig and Tasker, 1993).

The value for  $T$  is then computed as

$$T = 10^{\left[ t_{(\alpha/2, n-p)} S_i \right]}, \quad (6)$$

where

- $t_{(\alpha/2, n-p)}$  is the critical value from the  $t$ -distribution for  $n-p$  degrees of freedom,
- $n$  is the number of sites used in the regression equations, and
- $p$  is the number of basin characteristics used in the regression equations plus one.

Critical values from the  $t$ -distribution are available from many statistical textbooks. For the 90-percent prediction intervals computed for this study,  $\alpha/2 = 0.05$ ,  $n = 61$ , and  $p = 4$ ; therefore,  $t_{(\alpha/2, n-p)} = 1.672$  for all of the equations in table 3, simplifying calculation of  $T$  to

$$T = 10^{[1.672 S_i]}. \quad (7)$$

Values of  $\gamma^2$  and  $U$  for each of the regression equations in table 3 are provided in table 4.

## APPLICATION OF THE GENERALIZED-LEAST-SQUARES REGRESSION MODELS

The equations in table 3 were used to estimate long-term 95-, 98-, and 99-percent duration discharges for 35 selected sites in the five study basins. Basin characteristics for the sites, measured by use of the GIS procedures described previously, are listed in table 7 (at back of report). Estimated discharges and prediction intervals for the sites are listed in table 5.

The procedure necessary to obtain the estimates is explained by an example computation of the 95-percent duration discharge for the selected site on the Assabet River at Northborough (station number 01096640). First, the necessary basin characteristics for the site are measured from the various GIS data layers. Values for drainage area (DAREA), area of stratified drift, total length of streams, mean basin elevation, and minimum basin elevation, are 19.7 mi<sup>2</sup>, 6.99 mi<sup>2</sup>, 40.6 mi, 404 ft, and 259 ft, respectively. GWHEAD is computed by subtracting the minimum basin elevation from the mean basin elevation, to obtain a value of 145 ft. DRT/TST is

**Table 4.** Values needed to calculate 90-percent prediction intervals

[Dependent variables are from table 3]

Dependent variable	Model error variance, in log <sub>10</sub> units	Covariance matrix of the regression coefficients			
$\hat{P}_{95}$	0.04899	0.041016	0.002098	0.001470	- 0.016861
		.002098	.007287	- .006372	- .006249
		.001470	- .006372	.039259	.014216
		- .016861	- .006249	.014216	.014257
$\hat{P}_{98}$	.06347	.065744	.004040	.003050	- .027597
		.004040	.012946	- .011023	- .010902
		.003050	- .011023	.069338	.024726
		- .027597	- .010902	.024726	.024083
$\hat{P}_{99}$	.09543	.078499	.004926	.003827	- .033029
		.004926	.015722	- .013260	- .013154
		.003827	- .013260	.084119	.029857
		- .033029	- .013154	.029857	.028940

**Table 5.** Predicted 95-, 98-, and 99-percent duration discharges and prediction intervals for the selected sites in the study basins

[Duration discharges are in cubic feet per second; prediction intervals not provided for sites with basin characteristics outside the ranges of those in the regression analyses]

USGS station No.	95-percent duration discharge			98-percent duration discharge			99-percent duration discharge		
	Point estimate	Minimum	Prediction interval Maximum	Point estimate	Minimum	Prediction interval Maximum	Point estimate	Minimum	Prediction interval Maximum
<b>CONCORD RIVER BASIN</b>									
01096640	2.21	0.81	4.68	1.49	0.35	4.04	1.21	0.23	3.65
01096805	1.40	.52	2.95	.92	.22	2.45	.77	.15	2.31
01096935	1.62	.60	3.42	1.06	.25	2.86	.89	.17	2.66
01097000	16.0	5.62	35.2	11.4	2.55	32.7	8.83	1.64	28.2
01097050	18.2	6.40	40.1	13.2	2.96	38.0	10.2	1.89	32.5
01097280	1.94	.70	4.14	1.20	.28	3.28	1.00	.19	3.04
01097300	1.14	.41	2.45	.72	.16	1.97	.58	.11	1.78
01097370	2.03	.74	4.35	1.30	.30	3.56	1.05	.20	3.20
01097480	4.18	1.53	8.91	2.88	.68	7.86	2.33	.45	7.05
01098530	13.8	4.81	30.6	9.52	2.11	27.5	7.35	1.35	23.7
01098795	2.85	1.02	6.19	1.95	.45	5.42	1.51	.28	4.68
01098860	19.9	--	--	13.3	--	--	10.0	--	--
01099400	3.95	1.43	8.51	2.81	.65	7.77	2.19	.42	6.71
01099500	64.8	--	--	47.4	--	--	34.8	--	--
<b>NORTH COASTAL BASIN</b>									
01102031	0.035	--	--	0.017	--	--	0.004	--	--
01102033	.011	--	--	.004	--	--	.0	--	--
01102330	.80	0.28	1.76	.46	0.11	1.31	.37	0.063	1.17
01102344	2.21	.80	4.77	1.40	.32	3.88	1.12	.21	3.46
<b>SOUTH COASTAL BASIN</b>									
01105660	0.36	0.13	0.77	0.22	0.052	0.59	0.18	0.027	0.57
01105670	.021	--	--	.008	--	--	.0	--	--
01105730	2.35	.82	5.20	1.39	.31	3.98	1.10	.20	3.53

**Table 5.** Predicted 95-, 98-, and 99-percent duration discharges and prediction intervals for the selected sites in the study basins--  
Continued

USGS station No.	95-percent duration discharge			98-percent duration discharge			99-percent duration discharge		
	Point estimate	Minimum	Prediction interval Maximum	Point estimate	Minimum	Prediction interval Maximum	Point estimate	Minimum	Prediction interval Maximum
<b>SOUTH COASTAL BASIN--Continued</b>									
011058057	0.70	0.24	1.59	0.35	0.076	1.05	0.28	0.043	0.95
01105810	.84	.30	1.81	.51	.12	1.41	.41	.073	1.28
01105820	.14	.051	.30	.075	.017	.21	.056	.003	.20
01105845	.52	.18	1.16	.29	.065	.84	.23	.035	.75
01105870	3.00	1.06	6.59	2.08	.47	5.88	1.58	.29	4.98
<b>NARRAGANSETT BAY BASIN</b>									
01109135	0.49	0.18	1.06	0.28	0.067	0.78	0.23	0.038	0.73
01109215	2.05	.73	4.46	1.23	.28	3.44	.99	.18	3.11
01109225	.45	.16	.99	.25	.058	.71	.20	.031	.65
01109280	.23	.081	.50	.12	.026	.32	.090	.010	.31
<b>TENMILE RIVER BASIN</b>									
01109375	0.70	0.25	1.54	0.44	0.10	1.22	0.34	0.057	1.08
01109391	4.23	1.52	9.14	2.97	.68	8.25	2.29	.44	7.09
01109394	1.02	.37	2.16	.63	.15	1.70	.52	.097	1.59
01109397	.19	.067	.43	.11	.025	.32	.081	.008	.29
01109398	.088	.031	.19	.042	.010	.12	.028	.0	.12

computed by dividing the stratified-drift area by the total stream length and adding a constant of 0.1 to obtain a value of 0.2722 mi. Substituting these values into the equation to predict the 95-percent duration discharge from table 3 yields

$$P_{95} = 0.01618(19.7)^{0.9605}(145)^{0.7602}(0.2722)^{1.3283} \\ = 2.21 \text{ ft}^3/\text{s}.$$

To determine a 90-percent prediction interval for this estimate, the  $X_i$  vector is

$$X_i = \{1, \log_{10}(19.7), \log_{10}(145), \log_{10}(0.2722)\},$$

the model error variance from table 4 is  $\gamma^2 = 0.04899$ , and the covariance matrix,  $U$ , for the 95-percent duration is

$$U = \begin{bmatrix} 0.041016 & 0.002098 & 0.001470 & -0.016861 \\ 0.002098 & 0.007287 & -0.006372 & -0.006249 \\ 0.001470 & -0.006372 & 0.039259 & 0.014216 \\ -0.016861 & -0.006249 & 0.014216 & 0.014257 \end{bmatrix}.$$

The standard error of prediction computed from equation 5 is  $S_i = [0.048990 + 0.002876]^{0.5} = 0.227742$ , and  $T$  computed from equation 6 is  $T = 10^{1.672(0.227742)} = 2.4032$ . The 90-percent prediction interval is estimated from equation 4 as

$$\frac{1}{2.4032} \left( \frac{2.2118}{1.1349} \right) < P_{95} < \left( \frac{2.2118}{1.1349} \right) 2.4032,$$

$$\text{or } 0.81 < P_{95} < 4.68.$$

The 90-percent prediction interval is interpreted as follows: If 10 sites had the same basin characteristics as those for the Assabet River at Northborough, the true 95-percent duration discharge for nine of the sites (90 percent) would be from 0.81 to 4.68  $\text{ft}^3/\text{s}$ ; thus, assurance is 90-percent that the true value of  $P_{95}$  is in the given interval.

## LIMITATIONS OF THE EQUATIONS

The equations can be used to estimate natural long-term duration discharges for sites in most of Massachusetts. The measures of model adequacy listed in table 3, and the prediction intervals calculated by use of equations 4 to 7, indicate potential errors that can be expected when basin characteristics for the selected sites are within the ranges of those for the sites used in the regression analyses. Drainage areas (DAREA) for sites used in the regression analyses ranged from 1.61 to 149  $\text{mi}^2$ . Values of GWHEAD ranged from 18 to

1,035 ft, and values of DRT/TST ranged from 0.1 to 0.65 mi. Potential errors in the estimates can be substantially larger for sites where basin characteristics are outside the given ranges.

Drainage areas at five of the selected sites in the study basins were outside the ranges of the sites included in the regression analyses. Drainage areas at two of these sites (stations 01098860 and 01099500) were greater than 149  $\text{mi}^2$ , and drainage areas at three of the sites (stations 01102031, 01102033, and 01105670) were less than 1.61  $\text{mi}^2$  (see table 7). Duration discharges for the five sites were estimated by use of the equations. The estimates are provided in table 5, but prediction intervals were not provided because their true intervals are larger than those calculated by use of equations 4 to 7. Alternative estimates of duration discharges for the sites could be obtained by multiplying the duration discharges per unit area for a nearby unregulated streamflow-gaging station by the drainage area for the site where the estimate is desired; however, this method was not used because estimates from the regression equations are likely more accurate than estimates from the alternative method.

The equations generally are not applicable in almost all of the South Coastal Shore Subbasin of the South Coastal Basin, the eastern part of the Buzzards Bay Basin, Cape Cod, and the islands of Martha's Vineyard and Nantucket. These areas, which are almost entirely underlain by coarse-grained stratified-drift deposits, are not adequately represented by sites in the regression analyses. Streams in these areas commonly have ground-water drainage divides that are not coincident with topographic drainage divides. Estimates obtained by use of the regression equations for selected sites in these areas could have substantial errors.

## SUMMARY AND CONCLUSIONS

Physically based models in the form of mathematical equations were developed to estimate natural long-term discharges at the 95-, 98-, and 99-percent durations for ungaged sites in Massachusetts. The equations were developed by use of GLS regression analyses to relate computed duration discharges at 61 sites to physical characteristics of the areas drained by the sites. Twenty-four of the 61 sites used in the regression analyses were partial-record sites; 8 to 32 low-flow measurements were used to estimate the duration discharges at these sites. The 37 streamflow-gaging stations had periods of record of 3 to 88 years.

The weights assigned to each site used in the GLS regression analyses, were varied according to differences in length of record and cross-correlation among the concurrent discharges at the sites. The generalized-least-squares procedure was used for this study because other investigators have shown it to be more efficient and more statistically appropriate for use in regionalizing streamflow statistics than other regression analysis procedures.

All physical characteristics of the basins for the sites used in the regression analyses were determined from digital data bases by use of GIS computer software. The physical characteristics that provided the best estimates for each of the duration discharges were basin drainage area, area of stratified-drift deposits per unit of stream length in the basin, and a surrogate for the effective head of the aquifer in the stratified-drift deposits. The physical characteristics were each positively correlated to the duration discharges used as the dependent variables in the equations.

The equations to predict the 95-, 98-, and 99-percent duration discharges had standard errors of prediction of 57.7, 85.6, and 98.5 percent, respectively. These percentage errors are larger than those reported by the first Basin Yield study. Some of the differences may be explained by the fact that the drainage areas at the 20 sites added to the regression analyses for this study are smaller, on average, than those at the sites used in the first study. The medians of the observed natural duration discharges at the new sites are less than one-half those at the sites used in analyses for the first Basin Yield study. For given differences between observed and predicted discharges, percentage errors are larger for sites where observed discharges are small than for sites where observed discharges are large.

The average model error variances for the equations are about 10 times the sampling error variances; therefore, much larger gains in the precision of future models can be made by reducing the model error variances than by reducing the sampling error variances. Model error variances could be reduced by improving the methods used to measure the basin characteristics used in the present models, by measuring different basin characteristics that are more closely related to the streamflow statistics, or by adding new sites to the regression analyses.

Model error variances for future models could be improved when digital elevation data are available at a scale of 1:24,000 for all of Massachusetts. These data are currently being prepared for the entire continental United States by the USGS, National Mapping Division. The 1:24,000-scale digital elevation data could replace the 1:250,000-scale data used for this study. Improved basin characteristics also could be obtained by scanning hydrography from USGS 1:25,000-scale topographic maps. The scanned hydrography could provide improved data layers for streams, water bodies, and wetlands. An accurate wetlands data layer was not available for this study.

Adding new sites for future analyses is unlikely to lower the sampling errors for the model substantially, but adding new sites in areas that are under-represented by sites in the current models could decrease model errors and increase the range of applicability of the future models. Addition of new sites to the regression analyses also could increase confidence in the selection of the best independent variables for use in future equations.

The models were used to predict duration discharges for 35 sites in the Concord River, North Coastal, South Coastal, Narragansett Bay, and Tenmile River Basins. Ninety-percent prediction intervals were computed for the estimates at each of the sites except those for five sites that had basin characteristics outside the ranges of those used in the regression analyses. Alternative methods were not used to obtain the estimates for these sites because available alternative methods would not provide estimates more accurate than those calculated by use of the regression equations.

## REFERENCES CITED

- Cervione, M.A., Jr., 1982, Streamflow information for Connecticut with application to land-use planning: Connecticut Department of Environmental Protection Bulletin 35, p. 16.
- Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., and Wells, D.K., 1989, Estimating constituent loads: Water Resources Research, v. 25, no. 5, p. 937-942.
- de Lima, Virginia, 1991, Stream-aquifer relations and yield of stratified-drift aquifers in the Nashua River Basin, Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 88-4147, 47 p.
- Dingman, S.L., 1978, Synthesis of flow-duration curves for unregulated streams in New Hampshire: Water Resources Bulletin, v. 14, no. 6, p. 1481-1502.

- Duan, Naihua, 1983, Smearing estimate: a non-parametric retransformation method: *Journal of the American Statistical Association*, v. 78, no. 383, p. 605-610.
- Elasall, A.A., and Caruso, V.M., 1983, USGS digital cartographic data standards: digital elevation models: U.S. Geological Survey Circular 895-B, 40 p.
- Environmental Systems Research Institute, Inc., 1991, *Cell-based modeling with GRID*: Redlands, Calif., 7 chaps., various pagination.
- Fennessey, Neil, and Vogel, R.M., 1990, Regional flow-duration curves for ungaged sites in Massachusetts: *Journal of Water Resources Planning and Management*, v. 116, no. 4, p. 530-549.
- Gilroy, E.J., Hirsch, R.M., and Cohn, T.A., 1990, Mean square error of regression-based constituent transport estimates: *Water Resources Research*, v. 26, no. 9, p. 2069-2077.
- Helsel, D.R., and Hirsch, R.M., 1992, *Statistical methods in water resources*: Amsterdam, the Netherlands, Elsevier Science Publishers, 522 p.
- Hirsch, R.M., 1982, A comparison of four streamflow record extension techniques: *Water Resources Research*, v. 18, no. 4, p. 1081-1088.
- Iman, R.L., and Conover, W.J., 1983, *A modern approach to statistics*: New York, John Wiley, 497 p.
- Johnson, C.G., 1970, A proposed streamflow data program for central New England: U.S. Geological Survey Open-File Report, 38 p.
- Koch, R.W., and Smillie, G.M., 1986, Bias in hydrologic prediction using log-transformed regression models: *Water Resources Bulletin*, v. 22, no. 5, p. 717-723.
- Ku, H.F., Randall, A.D., and MacNish, R.D., 1975, Streamflow in the New York part of the Susquehanna River basin: New York State Department of Environmental Conservation Bulletin 71, 130 p.
- Lapham, W.W., 1988, Yield and quality of ground water from stratified drift aquifers, Taunton River basin, Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 86-4053, 69 p.
- Ludwig, A.H., and Tasker, G.D., 1993, Regionalization of low-flow characteristics of Arkansas streams: U.S. Geological Survey Water-Resources Investigations Report 93-4013, 19 p.
- Male, J.W., and Ogawa, Hisashi, 1982, Low flow of Massachusetts streams: Amherst, Massachusetts, University of Massachusetts, Water Resources Research Center Publication 125, 152 p.
- Miller, D.M., 1984, Reducing transformation bias in curve fitting: *The American Statistician*, v. 38, no. 2, p. 124-126.
- Myette, C.F., and Simcox, A.C., 1989 (revised 1991), Water resources and aquifer yields in the Charles River Basin, Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 88-4173, 53 p.
- National Oceanic and Atmospheric Administration, 1989, *Climatological data, annual summary, New England*, v. 101, no. 13, 51 p.
- Neter, John, Wasserman, William, and Kutner, M.H., 1985, *Applied linear statistical models*: Homewood, Illinois, Irwin, 1127 p.
- Olimpio, J.C., and de Lima, Virginia, 1984, Ground-water resources of the Mattapoissett River Valley, Plymouth County, Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 84-4043, 83 p.
- Parker, 1977, Methods for determining selected flow characteristics for streams in Maine: U.S. Geological Survey Open-File Report 78-871, 31 p.
- Persky, S.H., 1993, Yields and water quality of stratified-drift aquifers in the Southeast Coastal Basin, Cohasset to Kingston, Massachusetts: U.S. Geological Survey Water-Resources Investigations Report 91-4112, 47 p., 2 pl.
- Ries, K.G., III, 1993, Estimation of low-flow duration discharges in Massachusetts: U.S. Geological Survey Open-File Report 93-38, 59 p.
- Ryan, B.F., Joiner, B.L., and Ryan, T.A., Jr., 1985, *MINITAB handbook* (2d ed.): Boston, Mass., PWS-Kent, 385 p.
- Searcy, J.K., 1959, Flow-duration curves, manual of hydrology, part 2. low-flow techniques: U.S. Geological Survey Water-Supply Paper 1542-A, p. 1-33.
- Stedinger, J.R., and Tasker, G.D., 1985, Regional hydrologic analysis 1. Ordinary, weighted, and generalized-least-squares compared: *Water Resources Research*, v. 21, no. 9, p. 1421-1432.
- Tasker, G.D., 1972, Estimating low-flow characteristics of streams in southeastern Massachusetts from maps of ground water availability, in *Geological Survey Research, 1972*: U.S. Geological Survey Professional Paper 800-D, p. D217-D220.
- Tasker, G.D., and Driver, N.E., 1988, Nationwide regression models for predicting urban runoff water quality at unmonitored sites: *Water Resources Bulletin*, v. 24, no. 5, p. 1091-1101.
- Tasker, G.D., and Stedinger, J.R., 1986, Regional hydrologic analysis, 2, Model-error estimators, estimation of sigma and log-Pearson type 3 distributions: *Water Resources Research*, v. 22, no. 10, p. 1487-1499.
- \_\_\_\_\_, 1989, An operational GLS model for hydrologic regression: *Journal of Hydrology*, v. 3, p. 361-375.
- Thomas, M.P., 1966, Effect of glacial geology upon the time distribution of streamflow in eastern and southeastern Connecticut: U.S. Geological Survey Professional Paper 550-B, p. B209-B212.
- U.S. Geological Survey, 1987, List of U.S. Geological Survey water-supply reports and maps for Massachusetts, Rhode Island, and Connecticut, 98 p.
- \_\_\_\_\_, 1988, National Mapping Division digital data layer of population of the United States, from U.S. Census Bureau census blocks, scale 1:2,000,000.
- Vogel, R.M., and Kroll, C.N., 1990, Generalized low-flow frequency relationships for ungaged sites in Massachusetts: *Water Resources Bulletin*, v. 26, no. 2, p. 241-253.

---

## **Tables 6 and 7**

---

**Table 6.** Duration discharges and basin characteristics for sites used in the regression analyses

[Discharge is in cubic feet per second, area is in square miles, length is in miles, slope is in percent, and elevation is in feet]

Station No. (see fig. 6)	95-percent duration discharge	98-percent duration discharge	99-percent duration discharge	Years of record	Drainage area	Total stream length	Mean basin slope	Stratified drift area	Mean basin elevation
01094760	0.18	0.09	0.07	3	7.41	15.0	3.42	1.63	526
01095220	2.36	1.34	1.08	4	30.4	62.0	5.72	5.21	775
01095915	.90	.45	.35	4	15.7	24.0	2.92	4.52	440
01096000	11.0	7.80	6.40	42	63.6	144	4.99	16.8	615
01096910	.10	.06	.04	12	1.61	3.27	3.70	.18	433
01097300	.75	.33	.22	28	12.2	32.5	2.25	7.45	238
01100700	.41	.22	.16	13	5.48	7.92	2.79	2.75	133
01101000	.81	.36	.28	46	21.4	34.6	2.12	9.46	121
01103253	.53	.39	.31	6	7.23	17.6	2.56	1.08	267
01103440	.30	.21	.17	6	3.90	7.69	1.64	2.35	159
01104960	.22	.16	.14	3	2.37	3.65	1.30	.67	198
01105100	1.22	1.01	.94	3	3.40	3.54	2.64	1.96	225
01105575	.03	.01	.01	3	1.72	2.88	1.11	.0	185
01105600	.56	.33	.23	26	4.47	9.15	.50	1.19	147
01105830	.04	.02	.01	3	1.72	1.79	.57	.10	103
01106000	.12	.08	.07	38	7.99	17.6	1.51	.08	138
01107000	.07	.03	.02	12	4.71	10.6	1.06	.77	189
01107400	2.47	1.71	1.48	8	9.30	14.4	1.02	7.18	105
01109087	1.83	1.19	.97	4	20.7	28.3	1.35	8.88	113
01109090	.31	.16	.12	4	4.22	5.38	1.80	1.29	152
01109200	.09	.04	.02	12	4.33	12.7	.29	2.78	120
01109225	.17	.04	.01	9	7.39	11.3	.99	2.78	98
01111200	3.20	2.70	2.10	29	27.8	44.0	4.43	8.42	405
01111300	1.20	.81	.58	27	15.9	30.0	3.14	4.50	532
01123200	.23	.13	.11	10	4.39	4.62	4.36	.18	969
01162500	1.50	.80	.50	71	19.3	33.6	3.56	1.26	1,100
01162900	3.84	2.80	2.49	6	11.5	17.7	3.64	4.14	1,060
01165500	1.20	.87	.70	65	12.1	23.2	6.74	1.86	862
01166105	1.40	1.10	.99	5	5.24	14.2	7.39	1.24	934
01168300	2.60	1.60	1.20	9	29.6	50.5	10.9	.20	1,820
01169000	15.0	11.0	9.00	53	88.9	188	9.75	5.73	1,410
01169900	5.90	4.60	3.90	26	24.1	53.5	9.50	3.18	1,120
01170100	8.96	7.25	6.20	24	41.2	84.2	9.52	1.48	1,370
01171500	10.0	7.70	6.60	52	54.0	111	6.86	9.36	847
01171800	0.79	0.63	0.56	13	5.55	9.10	5.17	2.04	424
01173260	.05	.01	.0	12	4.62	4.24	1.62	.0	1,030
01173450	.36	.22	.18	8	6.60	11.0	4.24	1.00	667
01174000	.14	.05	.02	36	3.39	7.96	6.65	.07	1,030
01174050	.21	.17	.13	2	5.12	6.13	4.63	.72	872
01174565	1.30	.90	.90	2	12.6	28.8	7.33	2.05	947



Station No. (see fig. 6)	Minimum basin elevation	Maximum basin elevation	Mean elevation of drift	Minimum elevation of drift	Maximum elevation of drift	Basin relief	GWHEAD	GWRELIEF
01094760	443	748	459	443	604	305	83	171
01095220	443	2,000	569	443	1,020	1,560	332	587
01095915	285	748	365	285	554	463	155	279
01096000	249	1,450	401	249	1,030	1,200	366	791
01096910	299	548	338	299	397	249	134	108
01097300	158	463	216	158	338	305	80	190
01100700	59	266	97	59	184	207	74	135
01101000	49	351	106	49	236	302	72	197
01103253	174	397	226	174	348	223	93	184
01103440	128	289	149	128	190	161	31	72
01104960	141	282	192	148	201	141	57	63
01105100	98	530	219	98	361	432	127	272
01105575	131	249	131	131	131	118	54	10
01105600	98	197	139	98	148	99	49	60
01105830	98	190	98	98	98	92	5	10
01106000	16	203	136	72	180	187	122	118
01107000	98	299	183	98	256	201	91	168
01107400	56	190	101	56	190	134	49	144
01109087	23	200	88	23	197	177	90	184
01109090	52	299	133	52	295	247	90	253
01109200	102	266	120	102	134	164	18	42
01109225	49	200	81	49	98	151	49	59
01111200	240	630	314	240	453	390	165	223
01111300	361	758	447	361	548	397	171	197
01123200	699	1,200	810	699	997	501	270	308
01162500	899	1,800	951	899	1,100	901	201	211
01162900	902	1,300	1,010	902	1,200	398	158	308
01165500	594	1,520	706	594	902	926	268	318
01166105	499	1,300	939	499	1,080	801	435	591
01168300	787	2,800	998	791	1,330	2,010	1,030	549
01169000	499	2,200	1,020	499	1,830	1,710	911	1,340
01169900	499	1,800	947	499	1,510	1,300	621	1,020
01170100	499	2,400	1,090	499	1,970	1,900	871	1,480
01171500	197	1,600	560	197	1,360	1,400	650	1,170
01171800	272	797	319	276	512	525	152	246
01173260	932	1,180	932	932	932	248	98	10
01173450	472	899	583	472	676	427	195	214
01174000	794	1,200	806	797	833	406	236	46
01174050	699	1,200	777	728	797	501	173	79
01174565	568	1,300	810	568	1,030	732	379	472

Table 6 29

**Table 6.** Duration discharges and basin characteristics for sites used in the regression analyses--  
*Continued*

Station No. (see fig. 6)	95-percent duration discharge	98-percent duration discharge	99-percent duration discharge	Years of record	Drainage area	Total stream length	Mean basin slope	Stratified drift area	Mean basin elevation
01174900	.21	.13	.12	30	2.59	5.15	6.18	.017	934
01175670	.61	.36	.29	31	8.69	17.2	5.36	1.10	870
01175850	1.73	1.21	1.05	7	11.5	26.5	4.65	1.86	860
01176000	30.0	20.0	16.0	79	149.0	325	4.40	31.6	810
01176100	.61	.22	.08	8	7.70	17.9	4.34	2.70	839
01176200	1.13	.82	.73	8	3.96	4.95	6.99	1.08	669
01176300	2.79	2.37	2.20	5	6.57	9.07	8.47	1.39	832
01180000	.11	.08	.07	28	1.74	2.55	9.80	.0	1,090
01180500	4.70	3.60	2.90	80	52.8	101	8.40	1.50	1,380
01180800	.40	.30	.26	15	2.95	7.46	4.78	.12	1,560
01181000	12.0	9.10	7.40	55	94.0	166	8.74	3.90	1,420
01184200	.74	.55	.48	8	5.27	12.3	2.77	2.68	278
01187400	.50	.40	.30	32	7.37	12.2	11.0	.60	1,070
01197015	1.40	1.30	1.20	4	10.6	18.8	11.2	.52	1,560
01197180	2.46	1.91	1.50	5	7.62	8.60	12.2	.78	1,590
01197300	.06	.02	.01	12	2.18	2.71	9.64	.012	1,220
01198000	4.80	3.90	3.40	20	51.0	79.3	9.62	5.32	1,170
01331400	.57	.27	.17	12	7.68	9.32	8.12	.21	1,760
01332000	8.50	6.50	5.70	58	41.0	58.5	13.4	3.00	1,850
01333000	7.80	5.50	4.80	41	42.6	69.8	18.5	4.84	1,560
01333100	.22	.13	.11	7	5.25	9.41	20.0	.44	1,660

Station No. (see fig. 6)	Minimum basin elevation	Maximum basin elevation	Mean elevation of drift	Minimum elevation of drift	Maximum elevation of drift	Basin relief	GWHEAD	GWRELIEF
01174900	594	1,100	901	892	909	506	340	27
01175670	636	1,050	745	640	863	414	234	233
01175850	614	1,100	716	614	810	486	246	206
01176000	397	1,200	667	397	974	803	413	587
01176100	397	1,100	779	397	968	703	442	581
01176200	397	997	570	397	764	600	272	377
01176300	495	1,150	636	495	997	655	337	512
01180000	699	1,300	699	699	699	601	391	10
01180500	495	2,200	980	495	1,620	1,700	885	1,140
01180800	1,300	1,810	1,320	1,300	1,390	510	260	100
01181000	397	2,160	1,020	397	1,610	1,760	1,020	1,220
01184200	197	597	245	197	295	400	81	108
01187400	590	1,400	798	597	997	810	480	410
01197015	1,120	2,600	1,240	1,120	1,650	1,480	440	540
01197180	1,100	2,130	1,360	1,200	1,500	1,030	490	310
01197300	997	1,800	1,100	1,100	1,100	803	223	10
01198000	699	2,000	812	699	1,300	1,300	471	610
01331400	1,200	2,150	1,260	1,200	1,380	950	560	190
01332000	899	3,050	1,200	1,030	2,070	2,150	951	1,050
01333000	660	3,400	976	660	1,480	2,740	900	830
01333100	899	2,750	1,050	899	1,270	1,850	761	381

Table 6 31

**Table 7.** Basin characteristics for selected sites for which low-flow estimates are provided

[Area is in square miles, length is in miles, slope is in percent, and elevation is in feet]

USGS station No.	Drainage area	Total stream length	Mean basin slope	Stratified-drift area	Mean basin elevation	Minimum basin elevation	Maximum basin elevation	Mean elevation of drift	Minimum elevation of drift	Maximum elevation of drift	Basin relief	GWHEAD	GWRELIEF
<b>CONCORD RIVER BASIN</b>													
01096640	19.7	40.6	3.24	6.99	404	259	699	346	259	548	440	145	299
01096805	15.4	32.4	4.51	3.35	411	230	699	315	230	535	469	181	315
01096935	16.9	43.1	3.85	5.51	357	197	597	274	197	499	400	160	312
01097000	115	256	3.48	45.3	352	148	748	276	148	650	600	204	512
01097050	120	267	3.43	49.8	344	128	748	267	128	650	620	216	532
01097280	24.9	61.4	2.21	7.70	254	128	450	220	128	397	322	126	279
01097300	12.2	32.5	2.25	7.45	238	158	463	216	158	338	305	80	190
01097370	21.1	47.0	2.44	9.58	226	128	463	205	128	325	335	98	207
01097480	35.2	75.1	3.38	11.8	355	177	699	315	177	502	522	178	335
01098530	107	218	2.81	45.0	292	138	699	234	138	515	561	154	387
01098795	21.4	41.9	2.48	13.6	222	138	600	189	138	325	462	84	197
01098860	163	329	2.54	87.2	226	118	699	202	118	515	581	108	407
01099400	26.3	56.5	1.91	16.0	188	69	348	171	69	322	279	119	263
01099500	400	860	2.67	197	266	69	748	208	69	650	679	197	591
<b>NORTH COASTAL BASIN</b>													
01102031	1.09	3.48	1.29	0.25	99	46	148	95	72	105	102	53	43
01102033	.82	1.49	.39	.022	132	98	148	98	98	98	50	34	10
01102330	11.0	18.6	1.11	5.55	107	66	197	98	66	161	131	41	105
01102344	23.2	41.1	1.74	9.57	103	20	197	93	20	161	177	83	151
<b>SOUTH COASTAL BASIN</b>													
01105660	4.85	7.58	1.37	1.32	127	49	197	136	49	148	148	78	910
01105670	1.57	2.51	.79	.14	54	33	98	46	46	46	65	21	10

**Table 7.** Basin characteristics for selected sites for which low-flow estimates are provided--Continued

USGS station No.	Drainage area	Total stream length	Mean basin slope	Stratified-drift area	Mean basin elevation	Minimum basin elevation	Maximum basin elevation	Mean elevation of drift	Minimum elevation of drift	Maximum elevation of drift	Basin relief	GWHEA D	GWRELIEF
<b>SOUTH COASTAL BASIN-- Continued</b>													
01105730	30.3	77.6	20.7	20.7	103	49	180	96	49	177	131	54	138
011058057	14.9	31.5	.87	9.78	71	49	148	77	49	128	99	22	89
01105810	9.80	20.3	1.06	5.21	110	49	197	113	49	148	148	61	109
01105820	3.17	7.52	1.13	.77	114	49	197	96	49	148	148	65	109
01105845	7.56	16.0	.82	6.08	76	49	148	74	49	118	99	27	79
01105870	20.1	46.8	.95	19.3	80	10	200	81	16	200	190	70	194
<b>NARRAGANSETT BAY BASIN</b>													
01109135	7.77	10.5	1.15	1.77	103	36	184	82	39	125	148	67	96
01109215	26.8	55.5	.85	11.1	121	46	249	101	46	174	203	75	138
01109225	7.38	11.3	.99	2.78	93	52	200	81	52	98	148	41	56
01109280	5.94	12.4	.84	1.76	87	46	200	67	46	85	154	41	49
<b>TENMILE RIVER BASIN</b>													
01109375	7.29	14.9	.87	5.37	160	115	318	152	115	207	203	45	102
01109391	29.1	61.1	1.25	18.2	176	69	397	146	69	299	328	107	240
01109394	12.8	39.7	1.38	5.63	173	62	387	122	62	299	325	111	247
01109397	2.42	4.35	.66	1.56	99	66	138	90	66	121	72	33	65
01109398	3.10	6.69	.98	.40	124	69	200	85	69	98	131	55	39